Quantum Correlations and Communications

Tavakoli, Armin

Université de Genève
Departement de Physique Appliquée

Faculté des Sciences
Professeur N. Brunner
Professeur N. Gisin

Quantum Correlations and Communications

Thèse

présentée à la Faculté des Sciences de l'Université de Genève
pour obtenir le grade de Docteur ès sciences, mention Physique

par

Armin Tavakoli

Thèse N° 5489

Genève
Atelier de reproduction ReproMail
2020

UNIVERSITÉ DE GENÈVE
FACULTÉ DES SCIENCES

DOCTORAT ÈS SCIENCES, MENTION PHYSIQUE

Thèse de Monsieur Armin TAVAKOLI

intitulée :

«Quantum Correlations and Communications»

La Faculté des sciences, sur le préavis de Monsieur N. BRUNNER, professeur associé et directeur de thèse (Département de physique appliquée), Monsieur N. GISIN, professeur ordinaire et codirecteur de thèse (Groupe de physique appliquée), Monsieur J. D. BANCAL, docteur (Département de physique appliquée), Monsieur M. BOURENNANE, professeur (Department of Physics, Stockholm University, Sweden) et Monsieur A. ACIN, professeur (Institute of Photonic Sciences, Barcelona, Spain), autorise l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 6 juillet 2020

Thèse - 5489 -

Le Décanat

N.B. - La thèse doit porter la déclaration précédente et remplir les conditions énumérées dans les "Informations relatives aux thèses de doctorat à l'Université de Genève".

Abstract

The ontology of quantum theory, with features such as superpositions, entanglement and quantum measurements, is fascinating and radically different to that of many established physical models. It gives rise to predictions of correlations in laboratory data that distinguish quantum theory in the broader landscape of physical models. During the last three decades, the investigation of quantum correlations has received intense research attention. In foundational research, they allow for precise boundaries between classical, quantum and post-quantum phenomena. In quantum information science, correlations serve as resources for enhanced information processing, as tools physical inference and as signatures of quantum communications. In quantum technologies, correlations are a powerhouse for quantum advantages in practical tasks.

In this thesis, we present a broad investigation of quantum correlations ranging from foundations to applications. We begin with an introduction to three different forms of quantum correlations, each corresponding to a different setting in which quantum theory eludes classical models. Firstly, we focus on quantum correlations arising in communication experiments that cannot be modelled with classical communication resources. We establish general connections between quantum correlations and quantum communications, investigate different approaches within quantum theory towards communication advantages, develop efficient computational methods for bounding the advantages enabled by quantum theory and present a framework for studying the relation between correlations and information as manifested in classical and quantum models.

Secondly, we apply quantum correlations and communications towards physical inference. We introduce a framework for certification of different quantum devices in simple experiments within the state-of-the-art implementations. We show that key physical properties can be determined directly from the quantum correlations subject only to weak assumptions that require no precise prior characterisation of any part of the experiment. Qualitative and quantitative inference methods are developed for a variety of quantum devices implementing i.a. qubit ensembles, standard qubit measurements, non-projective measurements, quantum instruments, entangled states and entangled measurements.

Thirdly, we shift our attention to a different form of quantum correlations,

namely those that manifest contextuality. We show how quantum communication can be systematically employed to construct tests of contextuality and how such tests can be used to repeatedly harvest quantum correlations from a single system in many independent experiments. Moreover, we show remarkably strong connections to the ontology of quantum theory by proving general one-to-one relations between quantum contextuality, the failure of joint measurability and Einstein-Podolsky-Rosen steering.

Fourthly, we investigate quantum correlations that violate Bell inequalities. We introduce Bell inequalities that are tailored to a key resource for quantum information processing, namely the so-called mutually unbiased bases. We propose an operational formulation for such bases and prove that they can be certified through the quantum correlations. Departing from standard Bell experiments, we focus on quantum correlations in network settings that feature many sources and observers. For classes of such networks, we systematically derive Bell-type inequalities and demonstrate their quantum violation. Lastly, we address the standing suspicion that many known examples of quantum correlations in networks can be traced back to standard Bell inequalities rather than to the network structure; leading us to propose quantum correlations that appear to be genuine to the network structure.

The thesis serves as an overview of the selected scientific articles on which it is based. Additional discussions, related results and detailed proofs are provided in the original works.

Résumé

L'ontologie de la théorie quantique — états superposés, intrication, mesures probabilistes — est fascinante et radicalement différente de beaucoup d'autres modèles physiques, et les prédictions qu'elle établit quant aux corrélations observables expérimentalement l'en distinguent de façon frappante. Au cours des trois dernières décennies, l'exploration de ces corrélations quantiques a été l'objet de nombreuses investigations. En recherche fondamentale, elles permettent de séparer clairement les phénomènes classiques, quantiques et post-quantiques. En information quantique, elles tiennent lieu de ressources pour améliorer le traitement de l'information, d'outils d'inférence physique et de signature des communications quantiques. En technologie quantique, elles sont l'une des figures de proue en vue de la réalisation d'un avantage quantique pratique.

Dans cette thèse, nous réalisons une étude poussée des corrélations quantiques, tant d'un point de vue fondamental qu'appliqué. Nous commençons par une introduction à trois formes différentes qu'elles peuvent prendre, chacune correspondant à une situation dans laquelle la théorie quantique se soustrait à un modèle classique. Premièrement, nous nous intéressons aux corrélations quantiques issues d'expériences de communication et qui ne peuvent être reproduites par des ressources classiques de communication. Nous établissons des connexions générales entre corrélations et communications quantiques, explorons diverses approches au sein de la théorie quantique en vue d'obtenir un avantage de communication, développons des méthodes de calcul efficaces pour borner les avantages permis par la théorie quantique et présentons un cadre d'étude pour l'inspection des relations entre corrélation et information manifestées par les modèles classiques et quantiques.

Deuxièmement, nous utilisons ces corrélations et communications quantiques dans un but d'inférence physique. Nous introduisons un procédé pour certifier différents dispositifs quantiques utilisés dans des expériences simples de la recherche actuelle. Nous démontrons que des propriétés physiques essentielles peuvent être déterminées directement à partir des corrélations quantiques, et ce même avec de faibles hypothèses qui ne nécessitent aucune caractérisation préalable des éléments de l'expérience. Des méthodes qualitatives et quantitatives d'inférence sont élaborées pour des dispositifs quantiques variés parmi lesquels des ensembles de

qubits, des mesures ordinaires sur des qubits, des instruments quantiques, des mesures non projectives, des états et des mesures intriquées.

Troisièmement, nous portons notre attention sur une forme différente de corrélations quantiques: celles qui révèlent de la contextualité. Nous montrons comment la communication quantique peut être systématiquement utilisée pour construire des tests de contextualité et comment de tels tests peuvent être répétés afin de collecter des corrélations quantiques à partir d'un seul système, et ce pour de nombreuses expériences différentes. De plus, nous tissons des liens remarquablement étroits avec l'ontologie de la théorie quantique en prouvant une équivalence générale entre contextualité quantique, non mesurabilité conjointe et pilotage d'Einstein-Podolsky-Rosen.

Quatrièmement, nous nous intéressons aux corrélations quantiques qui violent des inégalités de Bell. Nous présentons de telles inégalités adaptées à une ressource phare du traitement de l'information quantique: les bases communément appelées mutuellement impartiales. Nous proposons une formulation opérationnelle pour ces bases et prouvons qu'elles peuvent être certifiées grâce aux corrélations quantiques. Nous éloignant ensuite des expériences de Bell habituelles, nous nous concentrons sur les corrélations quantiques dans des réseaux possédant plusieurs sources et observateurs. Pour de tels réseaux, nous dérivons systématiquement des inégalités à la Bell et démontrons leur violation quantique. Enfin, nous confirmons le soupçon de longue date selon lequel de nombreux exemples connus de corrélations quantiques dans des réseaux doivent en réalité leur non localité à des inégalités de Bell ordinaires plutôt qu'à leur structure en réseau, ce qui nous conduit à suggérer des corrélations quantiques tirant véritablement profit de cette structure.

Cette thèse donne un aperçu des quelques articles scientifiques sur laquelle elle s'appuie. Des discussions complémentaires, des résultats associés et des preuves détaillées sont fournis dans les travaux originaux.

List of papers and pre-prints during doctoral studies

Below follows a list of the author's scientific publications and pre-prints completed during the course of doctoral studies (specifically October 2016 to July 2020). The works that are not discussed in this thesis are indicated.

1. A. Tavakoli, and M. Żukowski, *Higher dimensional communication complexity problems: classical protocols vs quantum ones based on Bell's Theorem or prepare-transmit-measure schemes*, Phys. Rev. A **95**, 042305 (2017).

2. N. Gisin, Q. Mei, A. Tavakoli, M-O. Renou, and N. Brunner, *All entangled pure quantum states violate the bilocality inequality*, Phys. Rev. A **96**, 020304(R) (2017). **[Not discussed in the thesis]**

3. A. Tavakoli, M-O. Renou, N. Gisin, and N. Brunner, *Correlations in star networks: from Bell inequalities to network inequalities*, New J. Phys. **19**, 073003 (2017).

4. A. Hameedi*, A. Tavakoli*[1], B. Marques, and M. Bourennane, *Communication games reveal preparation contextuality*, Phys. Rev. Lett. **119**, 220402 (2017).

5. A. Tavakoli, and A. Cabello, *Quantum predictions for an unmeasured system cannot be simulated with a finite-memory classical system*, Rev. A **97**, 032131 (2018). **[Not discussed in the thesis]**

6. A. Tavakoli, G. Haack, M. Huber, N. Brunner, and J. B. Brask, *Heralded generation of maximal entanglement in any dimension via incoherent coupling to thermal baths*, Quantum **2**, 73 (2018). **[Not discussed in the thesis]**

7. A. Tavakoli, J. Kaniewski, T. Vértesi, D. Rosset, and N. Brunner, *Self-testing quantum states and measurements in the prepare-and-measure scenario*, Phys. Rev. A **98**, 062307 (2018).

[1] Stars indicate equal contribution.

8. M. Czechlewski, D. Saha, A. Tavakoli, and M. Pawłowski, *Device independent witness of arbitrary dimensional quantum systems employing binary outcome measurements*, Phys. Rev. A **98**, 062305 (2018). [**Not discussed in the thesis**]

9. A. Tavakoli, A. A. Abbott, M-O. Renou, N. Gisin, and N. Brunner, *Semi-device-independent characterisation of multipartite entangled states and measurements*, Phys. Rev. A **98**, 052333 (2018).

10. D. Martínez, A. Tavakoli, M. Casanova, G. Cañas, B. Marques, and G. Lima, *High-dimensional quantum communication complexity beyond strategies based on Bell's theorem*, Phys. Rev. Lett. **121**, 150504 (2018).

11. Z-X. Man, A. Tavakoli, J. B. Brask, L-Z. Hu, and Y-J. Xia, *Improving autonomous thermal entanglement generation using a common reservoir*, Phys. Scr. **94**, 075101 (2019). [**Not discussed in the thesis**]

12. A. Tavakoli*, D. Rosset*, and M-O. Renou, *Enabling computation of correlation bounds for finite-dimensional quantum systems via symmetrisation*, Phys. Rev. Lett. **122**, 070501 (2019).

13. K. Mohan, A. Tavakoli, and N. Brunner, *Sequential random access codes and self-testing of quantum measurement instruments*, New J. Phys. **21**, 083034 (2019).

14. A. Tavakoli* and R. Uola*, *Measurement incompatibility and steering are necessary and sufficient for operational contextuality*, Phys. Rev. Research **2**, 013011 (2020).

15. A. Tavakoli, G. Haack, N. Brunner, and J. B. Brask, *Autonomous multipartite entanglement engines*, Phys. Rev. A **101**, 012315 (2020).[**Not discussed in the thesis**]

16. A. Tavakoli, M. Smania, T. Vértesi, N. Brunner, and M. Bourennane, *Self-testing non-projective quantum measurements in prepare-and-measure experiments*, Science Advances **6**, 16 (2020).

17. H. Anwer*, N. Wilson*, R. Silva, S. Muhammad, A. Tavakoli, and M. Bourennane, *Noise-robust preparation contextuality shared between any number of observers via unsharp measurements*, arXiv:1904.09766.

18. N. Gisin, J-D. Bancal, Y. Cai, A. Tavakoli, E. Z. Cruzeiro, S. Popescu, and N. Brunner, *Constraints on nonlocality in networks from no-signaling and independence*, Nature Communications **11**, 2378 (2020). [**Not discussed in the thesis**]

19. G. Foletto, L. Calderaro, A. Tavakoli, M. Schiavon, F. Picciariello, A. Cabello, P. Villoresi, and G. Vallone, *Experimental certification of sustained entanglement and nonlocality after sequential measurements*, Phys. Rev. Applied **13**, 044008 (2020). [**Not discussed in the thesis**]

20. A. Tavakoli, Marek Żukowski, and Č. Brukner, *Does violation of a Bell inequality always imply quantum advantage in a communication complexity problem?*, Quantum **4**, 316 (2020).

21. A. Tavakoli, E. Z. Cruzeiro, J. B. Brask, N. Gisin, and N. Brunner, *Informationally restricted quantum correlations*, Accepted for publication in Quantum.

22. A. Tavakoli, M. Farkas, D. Rosset, J-D. Bancal, and J. Kaniewski, *Mutually unbiased bases and symmetric informationally complete measurements in Bell experiments: Bell inequalities, device-independent certification and applications*, arXiv:1912.03225.

23. A. Tavakoli and N. Gisin, *The Platonic solids and fundamental tests of quantum mechanics*, Quantum **4**, 293 (2020). [**Not discussed in the thesis**]

24. H. Anwer, S. Muhammad, W. Cherifi, N. Miklin, A. Tavakoli, and M. Bourennane, *Experimental characterisation of unsharp qubit observables and sequential measurement incompatibility via quantum random access codes*, Phys. Rev. Lett. **125**, 080403 (2020). [**Not discussed in the thesis**]

25. A. Tavakoli, *Semi-device-independent certification of independent state and measurement devices*, Accepted for publication in Physical Review Letters (arXiv:2003.03859). [**Not discussed in the thesis**]

26. A. Tavakoli, N. Gisin, and C. Branciard, *Bilocal Bell inequalities violated by the quantum Elegant Joint Measurement*, arXiv:2003.03859.

27. A. Tavakoli, I. Bengtsson, N. Gisin, and J. M. Renes, *Compounds of symmetric informationally complete measurements and their application in quantum key distribution*, Accepted for publication in Physical Review Research (arXiv:2007.01007). [**Not discussed in the thesis**]

Acknowledgements

The work that I accomplished during my doctoral studies was made possible thanks to my collaborators, doctoral advisors, family and friends. I am grateful to all of them: to some for the scientific discussions, to some for the guidance, to some for the opportunities they gave me and to some for the personal friendships. It was a pleasure to complete a PhD in the environment that they collectively created. However, as they are far too numerous to all be named here, I would like to extend gratitude towards a few people that have played a particularly important role.

I would like to extend my sincerest gratitude to my doctoral advisor Nicolas Brunner. The experience of completing a PhD in his group was very rewarding. The day-to-day discussions I had with him over the years did not only propel a substantial part of my PhD but also enabled me to mature as a researcher. I much appreciate the patience that the latter must have required of him. I am also grateful for the considerable number of concrete skills that I picked up either directly or indirectly from him.

Next, I would like to wholeheartedly thank my second doctoral advisor, Nicolas Gisin. While finishing my master degree in Germany in 2016, it was a bolt from the blue when Nicolas contacted me and extended an invitation to Geneva where he, together with Nicolas Brunner, offered me a PhD position that I had never applied for. During my PhD, he diluted many of my grievances with academia and academic work by setting an outstanding example. He was a source of inspiration and provided a valuable input on what good science is and how to be a good scientist.

I would also like to thank some of my closer scientific collaborators, including Mohamed Bourennane, Jonatan Brask, Géraldine Haack, Jędrzej Kaniewski, Gustavo Lima, Sadiq Muhammad, Marc-Olivier Renou, Denis Rosset, Tamás Vértesi and Marek Żukowski. May we have many more fruitful collaborations in the future.

Finally, I would like to thank my family and Elsa Sellin for their support and for always reminding me how much more there is to life than physics. I extend particular gratitude to my father, for his unparalleled enthusiasm for my work.

Introduction

Quantum theory is perhaps the most successful physical theory ever established. It is also the physical theory that most radically departs from the everyday human experience of nature. The apparent collision is avoided by the fact that quantum theory only concerns nature on small scales, i.e. atoms, electrons, photons etc. The picture of nature on these scales, as painted by quantum theory, is remarkable. Where pre-quantum physics upholds that physical systems always have well-defined observable properties, quantum theory withstands that they can be in so-called *superposition*, i.e. that they can simultaneously be in two mutually exclusive states prior to observation (following Schrödinger's famous cat, both dead and alive). While pre-quantum physics represents a measurement as an act of simply revealing an already existing property, a *quantum measurement* is a dramatic process in which the outcome is created through interaction with the state, leaving its future course altered. Whereas two pre-quantum particles can be fully understood by studying them separately and adding up the knowledge, two quantum particles can be *entangled* and therefore influence each other through "a spooky action at a distance" which makes their joint state more than the knowledge of the parts. It is therefore unsurprising that quantum theory provides novel tools for explaining and predicting physical phenomena. After roughly a hundred years of quantum theory, there is no shortage of examples: quantum theory is relevant for everything ranging from particle physics to spectroscopy, and quantum technologies such as microelectronics, lasers and magnetic resonance imaging are widely established. Indeed, for a multitude of problems in physics, quantum theory offers a physical model that has hitherto been remarkably consistent with experimental tests.

However, the success of quantum theory notwithstanding, it appears imperative to ask how quantum theory, being as radical and philosophically consequential as it is, distinguishes itself from other physical models that may be conceptually less dramatic and closer to home? How can we characterise physical phenomena as being genuinely quantum? What would such phenomena teach us about quantum theory? And what are the ultimate limitations of a reality governed by the laws of quantum theory? These foundational questions are key to our understanding of *what* quantum

theory is and *why* quantum theory conceptually distinguishes itself from other theories in physics.

Frequently, in fact, the predictions of quantum theory do not distinguish themselves from other reasonable models in the sense that the latters can be used to simulate the predictions. For instance, say that we build a laser which we assume emits a single photon that passes through an unbiased beam splitter. Quantum theory offers the explanation that the beam splitter puts the path taken by the photon in a superposition of the two alternatives (reflection and transmission). By placing a detector at the end of each path, we perform a measurement that interferes with the superposition, thus making the path well-defined and yielding a random click in one of the detectors. However, the uniformly random clicks that we observe could easily be explained with a classical model. For example, imagine that the beam splitter is subject to some internal process causing it to randomly direct the incoming photon either in the transmission path or in the reflection path. Then, the system acts like a coin-flip rather than as a superposition; its path is always well-defined. Nevertheless, it still produces the same outcome statistics as predicted by quantum theory. Clearly, we cannot conclude that we are observing a genuine quantum phenomenon. Conspicuously, does there exist situations in which the predictions of quantum theory can defy those of other model?

The matter of distinguishing quantum predictions from those of other models is naturally subject to some assumptions. Therefore, the more precise and relevant question is *under which assumptions can quantum theory distinguish itself from other physical models?* A milestone answer was given by John Bell in 1964 with a result known as *Bell's theorem* [1]. Bell considered an experiment involving two very distant parties who each are asked independent questions by a referee. In order to correlate their answers, the parties may share a pair of particles emitted from some source. Once they receive a question, they measure their particle and return the outcome as their answer. Importantly, these measurement events are space-like separated which justifies the assumption that the question that is asked to one party cannot influence the answer of the other party. This principle derives from the no-faster-than-light communication at the heart of relativity. Bell's theorem is the fact that using only this assumption, the data predicted by quantum theory cannot be reproduced in any model which assumes that the two parties cannot influence each other due to the space-like separation (locality)[2]. In a classical approach to physics, such local realist models are arguably natural. Nevertheless, Bell's theorem reduces the matter of distinguishing quantum theory from all local realist models to a question of experiment. Over the years many experiments have been performed, which to varying degree capture the rather extreme conditions required for a stringent test of Bell's theorem. The most rigorous tests were performed only recently and they arguably

[2]Notably, it is sometimes argued that Bell's theorem also requires an assumption of realism. Whereas this concept perhaps could be meaningful, it remains elusive to the author. Therefore, in this thesis, Bell's theorem is presented as the failure of a local variable model to account for quantum theory.

settled the matter decisively in the favour of quantum theory [2, 3, 4].

In fact, the predictions of quantum theory do not only distinguish themselves in the niche of Bell experiments. In 1967, the work of Kochen and Specker [5] showed that quantum theory is incompatible with any model that is both deterministic and noncontextual. A deterministic model upholds that all observable quantites are well-defined at all times and a noncontextual model upholds that observable properties do not depend on the precise manner in which they are measured. For instance, a measurement of the total angular momentum of an electron can be performed jointly with a measurement of its x-component. Alternatively, we could instead measure it jointly with its z-component. These are two different ways (contexts) of measuring the total angular momentum. The *Kochen-Specker theorem* is the fact that no deterministic noncontextual model can account for all the predictions of quantum theory. In principle, and in sharp contrast to Bell's theorem, such predictions can arise in experiments that involve communication. However, for a long time it was unclear precisely how the discrepancy between quantum theory and deterministic noncontextual models could be detected - leaving the result rather abstract. Nevertheless, in more recent years, that matter has solved by showing that quantum predictions, known as quantum contextuality, can be detected in the correlations seen between the outcomes of sequential measurements performed on a single quantum system [6]. Following this breakthrough, quantum contextuality has been experimentally demonstrated in several tests, see e.g. [7].

The predictions of quantum theory encountered in Bell's theorem and in the Kochen-Specker theorem are examples of *quantum correlations* in laboratory data. However they are only two, albeit important, examples of quantum correlations selected from a diverse fauna that arises in different physical scenarios and under different types of assumptions. While the foundational interest in quantum correlations is evident and has led to many foundational insights in the last few decades, their relevance extends well beyond the foundations of quantum theory. Quantum correlations are conceptually and practically crucial in the rise of quantum information science seen over the past three decades. It is the fact that they are signatures of post-classical data that make them a powerhouse for quantum information science. Landmark examples include the development of quantum technologies such as quantum computers, quantum cryptography and quantum communications that outperform their counterparts in conventional technologies.

The connection between quantum correlations and quantum advantages in information processing is not obvious (in fact it took decades before it was noticed by researchers). Let us illustrate its general spirit through an example of how quantum correlations can improve classical communications. Imagine that a number of distributed parties each hold a piece of data. Their aim is to collectively perform a pre-determined computation that depends on all their data pieces. Naturally,

in order to succeed, they must communicate. What is the smallest amount of communication necessary to perform the computation? It turns out there are different answers; one given by standard information processing techniques, and one given by information processing techniques that exploit quantum correlations. Say, for instance, that the distributed parties first perform a Bell experiment and establish quantum correlations in the spirit of Bell's theorem. In quantum theory, these correlations themselves cannot be used for transmitting information due to the space-like separation entering Bell's theorem. Therefore, we cannot consider these correlations as constituting a source for additional communication. However, they can be used to improve the efficiency of classical communication [8]. By using the quantum correlations to select the communication strategy of the parties, the amount of communication needed to perform the computation can be reduced beyond anything achievable by standard techniques. The advantage thus stems from a post-classical ability of coordinating classical resources.

In foundational science quantum correlations serve to distinguish quantum theory from other models and in quantum information science quantum correlations are employed to enhance information processing. However there is a third scene on which quantum correlations play a key role, namely in physical inference. If we assume quantum theory, what conclusions can we draw about the physical features of an experiment by inspecting its data? Such inferences are important as they create a bridge between the ontology of quantum theory and its laboratory predictions. Also, from a more practical point of view, they enable methods for the characterisation of quantum devices which is crucial in quantum information science and in exploratory experimental tests of quantum theory. A simple example of the type of inferences that quantum correlations enable is given by Bell's theorem. If we make no further assumption than space-like separation[3], a proof of Bell's theorem implies that the two distant particles appearing in a Bell experiment must be entangled. Thus, we can certify a quantum resource under minimal assumptions. However, this does not tell us what the entangled state actually is (indeed most quantum states are entangled). Remarkably, it turns out that the strongest forms of quantum correlations can be used to precisely infer the underlying quantum ontology in an experiment [9], i.e. to precisely determine both the state and the measurements that were used to obtain the data. Concretely, this means that correlations that to a larger extent defy classical models, and therefore are close to the correlation limits allowed in quantum theory, can be used to deduce more information about the physics of the experiment that created them. In general, quantum correlations enable the investigation of physical inference under weak assumption on the experiments in which the correlations are obtained.

In view of the above, the general investigation of quantum correlations is relevant to many lines

[3]Strictly speaking, one must also assume that there is no superdeterminism. Superdeterminism implies that the inputs of a Bell experiment always are correlated with the system that is measured.

of research that span from foundational considerations to applied matters. In this thesis, we explore quantum correlations in many physical scenarios, investigate their usefulness for physical inference and apply them as resources for quantum information processing.

Outline of the thesis

The content of this manuscript is organised as follows.

Chapter 1 is a brief review of quantum correlations. Its focus is on the three forms of quantum correlations that will permeate this thesis, namely quantum nonlocality, quantum contextuality and quantum communication complexity. Familiarity with basic quantum theory is assumed.

Chapter 2 focuses on quantum communication complexity i.e. how quantum resources can be employed to better perform communication tasks as compared to classical models. It first focuses on exploring the connection between quantum nonlocality and quantum communications. Then, it considers quantum nonlocality as opposed to quantum communication as resources for communication complexity. Their relationship is investigated. Subsequently, the focus shifts to quantum communication complexity powered by quantum communications. Efficient computer methods are presented for bounding the set of quantum correlations. Finally, a new route to quantum communication complexity is introduced that is based on the information content of quantum communication rather than the standard approach based on dimensionality.

Chapter 3 concerns certification of quantum devices. Its focus is on the so-called semi-device-independent setting in which only the dimension of physical systems is used to deduce interesting properties from various quantum devices. Firstly, precise and robust certification methods are developed for simple qubit states and measurements in prepare-and-measure experiments. Then, general certification methods are discussed for generalised qubit measurements (so-called non-projective measurements). This is followed by an extension of the experiment to a prepare-transform-measure scenario in which one can certify more sophisticated quantum operations that produce both classical and quantum outputs. Finally, a hybrid scheme is presented that involves both quantum communication and entanglement. This scheme is used for certification and characterisation of high-dimensional and multipartite entangled states.

Chapter 4 is an exploration of quantum contextuality. It exclusively concerns an operational approach to contextuality that is not limited to quantum theory. It is shown that tests of operational contextuality can be phrased as quantum communication games in which data is actively being hidden in the communication. The framework is used to derive families of noncontextuality inequalities. Next, general and powerful one-to-one relations are derived that link quantum contextuality to the failure of joint measurability and quantum steering. Finally, we consider the

possibility of sharing the contextuality enabled by a quantum ensemble between an indefinite number of independent observers.

Chapter 5 is devoted to Bell nonlocality. It begins with the construction of Bell inequalities tailored to so-called mutually unbiased bases of any Hilbert space dimension. We introduce an operational definition of mutually unbiased bases and prove that the quantum correlations appearing in the Bell experiments can be used for certification of these operational mutually unbiased bases. Subsequently, the application of these Bell inequalities to quantum key distribution is considered. This is followed by a discussion of Bell nonlocality in networks. It is shown how standard Bell inequalities can be systematically mapped to Bell inequalities valid on networks. Finally, we address the long-standing suspicion that known quantum correlations in the simplest network can be somehow traced back to standard Bell nonlocality: genuine quantum correlations in the so-called bilocality scenario are presented that bear no resemblance to standard Bell nonlocality.

Contents

1

Background in quantum correlations

1.1 Quantum nonlocality

Imagine an experiment in which two observers, for simplicity named Alice and Bob, are separated by a very large distance. A referee holds a source that emits a physical system such that one part of it is given to Alice and one part is given to Bob. In addition, the referee supplies each of them with independent inputs labelled x and y. Each of them are drawn randomly from the set $\{1, \ldots, n\} \equiv [n]$. Upon receiving their inputs, Alice and Bob are asked to produce outputs, belonging to the set $[m]$, that we respectively label a and b, and return them to the referee. The experiment is illustrated in Figure 1.1. This input/output process is repeated many times. In every round Alice and Bob receive a shared state and random and independent inputs and return their outputs. After a large number of rounds the process is stopped and the relative frequencies are used to determine a conditional probability distribution $p(a, b|x, y)$. This distribution is common referred to as *the correlations*.

A priori there is only one thing that we take for granted about the correlations, namely that the space-like separation between Alice and Bob guarantees that the input of one party cannot influence the output of the other party. That is, if we forget about the outcome of Alice (Bob), the output of Bob (Alice) depends only on his (her) input. This *no-signaling principle* is dictated by relativity, and we formalise it as follows:

$$\sum_a p(a, b|x, y) = p(b|x, y) \overset{\text{NS}}{=} p(b|y), \qquad \sum_b p(a, b|x, y) = p(a|x, y) \overset{\text{NS}}{=} p(a|x).$$

Having assumed away a clash with relativity, we now ask the following question: how do the realisable correlations depend on the physical model used to predict them?

If we are presented with the correlations $p(a, b|x, y)$, a broad class of reasonable physical models

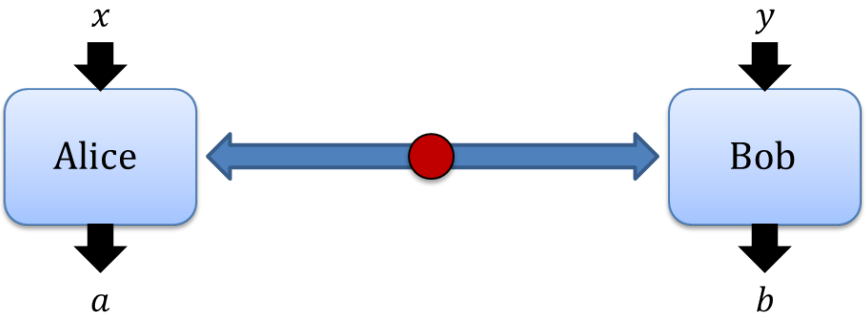

Figure 1.1: Bell experiment. Two space-like separated observers receive independent inputs x and y respectively and are asked to produce outputs, denoted a and b, by performing local measurements on parts of a shared physical system.

that potentially could explain the correlations are known as *local hidden variable* models - or for short *local models*. Local models are based on the following reasoning. Since the particles created by the source have a common origin, we may imagine that they are correlated in some way that perhaps is unknown to us. For example, when one particle is up the other particle is down. These correlations could also be stochastic; for instance sometimes one particle is up and the other is down and sometimes vice versa. We denote this shared influence by λ and out lack of knowledge about its precise nature is represented by a probability distribution $p(\lambda)$. In a local model, the influence is carried with the particles as they undergo separation and is then used by each observer to determine the output (for every given input). This means that a local model for explaining the correlations is written

$$p(a,b|x,y) = \sum_{\lambda} p(\lambda)p(a|x,\lambda)p(b|y,\lambda). \tag{1.1}$$

Notice that it is sufficient to let $p(\lambda)$ be a probability distribution rather than a probability density due to the fact that λ, together with the respective inputs, determines the outputs. We can think of it as Alice and Bob employing a different deterministic function to map their input into their output depending on λ. There are only finitely many such functions and thus only finitely many values of λ are required. Elaborating further on this picture, let us enumerate all deterministic functions from $[n]$ to $[m]$ in the list $\{f_1, \ldots, f_N\}$ where $N = m^n$ is the total number of such functions. We can then write the local model as a convex combination of deterministic distributions

$$p(a,b|x,y) = \sum_{\lambda_1=1}^{N} \sum_{\lambda_2=1}^{N} p(\lambda)\delta_{a,f_{\lambda_1}(x)}\delta_{b,f_{\lambda_2}(y)}, \tag{1.2}$$

where $\lambda = (\lambda_1, \lambda_2)$. In this form, it is evident that we can determine whether $p(a,b|x,y)$ admits a

local model through a linear program that searches for an appropriate probability distribution $p(\lambda)$. Since linear programs are easy to evaluate (at least for a small number of inputs and outputs), we can efficiently decide whether correlations admit a local model. It is worth noticing that this also means that the set of local correlations can geometrically be represented by a polytope where the vertices are obtained from the deterministic distributions [10].

The outstanding question is whether the correlations explainable with local models are any different from those predicted in quantum theory. The most insightful answer requires us to introduce the concept of a *Bell inequality*. A Bell inequality is a criterion that is respected by all local models. The most famous and well-studied Bell inequality is known as the Clauser-Horne-Shimony-Holt (CHSH) inequality [11]. It applies to the simplest possible Bell experiment, corresponding the scenario in which inputs and outputs are binary ($n = m = 2$). Consider the following functional that maps the correlations $p(a, b|x, y)$ onto a real number,

$$\mathcal{S}_{\text{chsh}} \equiv \langle A_1 B_1 \rangle + \langle A_1 B_2 \rangle + \langle A_2 B_1 \rangle - \langle A_2 B_2 \rangle, \tag{1.3}$$

where $\langle \cdot \rangle$ denotes the expectation value defined as

$$\langle A_x B_y \rangle \equiv \sum_{a,b} (-1)^{a+b} p(a, b|x, y). \tag{1.4}$$

In a local model, an expectation value can be written as

$$\langle A_x B_y \rangle \equiv \sum_{\lambda} p(\lambda) \sum_{a} (-1)^a p(a|x, \lambda) \sum_{b} (-1)^b p(b|y, \lambda) \equiv \sum_{\lambda} p(\lambda) \langle \tilde{A}_{x,\lambda} \rangle \langle \tilde{B}_{y,\lambda} \rangle, \tag{1.5}$$

where we have denoted the sum over a and b by $\langle \tilde{A}_{x,\lambda} \rangle$ and $\langle \tilde{B}_{y,\lambda} \rangle$ respectively. Evaluating the CHSH expression, we obtain that

$$\mathcal{S}_{\text{chsh}} = \sum_{\lambda} p(\lambda) \left[\left(\langle \tilde{A}_{1,\lambda} \rangle + \langle \tilde{A}_{2,\lambda} \rangle \right) \langle \tilde{B}_{1,\lambda} \rangle + \left(\langle \tilde{A}_{1,\lambda} \rangle - \langle \tilde{A}_{2,\lambda} \rangle \right) \langle \tilde{B}_{2,\lambda} \rangle \right]. \tag{1.6}$$

We want to place an upper bound on $\mathcal{S}_{\text{chsh}}$. Due to this being a convex combination over λ, we need to find the largest possible value of the square bracket and place all the weight of $p(\lambda)$ in front of that term. Since both $\langle \tilde{A}_{x,\lambda} \rangle$ and $\langle \tilde{B}_{y,\lambda} \rangle$ are numbers of magnitude at most one, it is clear that we should choose $\langle \tilde{B}_{1,\lambda} \rangle$ and $\langle \tilde{B}_{2,\lambda} \rangle$ to be of maximal magnitude (one) with the same sign as $\left(\langle \tilde{A}_{1,\lambda} \rangle + \langle \tilde{A}_{2,\lambda} \rangle \right)$ and $\left(\langle \tilde{A}_{1,\lambda} \rangle - \langle \tilde{A}_{2,\lambda} \rangle \right)$ respectively. Thus, we conclude that

$$\mathcal{S}_{\text{chsh}} \leq \left| \langle \tilde{A}_{1,\lambda} \rangle + \langle \tilde{A}_{2,\lambda} \rangle \right| + \left| \langle \tilde{A}_{1,\lambda} \rangle - \langle \tilde{A}_{2,\lambda} \rangle \right| \leq 2, \tag{1.7}$$

where the last step follows from the fact that $|s + r| + |s - r| \leq 2$ when $s, r \in [-1, 1]$. Thus, we have derived the CHSH inequality. The inequality itself is nothing more than an unremarkable statement

about set theory. What is remarkable about the CHSH inequality (and with Bell inequalities in general) is that they can be violated in quantum theory.

A violation of the CHSH inequality implies that $p(a,b|x,y)$ cannot be explained by *any* local model. In a quantum model of the correlations, the source emits some quantum state $|\psi\rangle$ that is a ray in Hilbert space. Holding a subsystem each, Alice and Bob implement different quantum measurements depending on their inputs. The measurements are labelled $\{A_{a|x}\}$ and $\{B_{b|y}\}$ respectively and they satisfy

$$A_{a|x} \geq 0, \qquad \sum_{a=1}^{m} A_{a|x} = \mathbb{I}, \qquad B_{b|y} \geq 0, \qquad \sum_{b=1}^{m} B_{b|y} = \mathbb{I}, \tag{1.8}$$

where $\mathbb{I}$ denotes the identity operator. Any complete set of positive operators constitutes a valid quantum measurement and is formally known as a *positive operator-valued measure* (POVM)[1]. The probabilities are then given by the Born rule

$$p(a,b|x,y) = \mathrm{tr}\left(A_{a|x} \otimes B_{b|y} |\psi\rangle\langle\psi|\right). \tag{1.9}$$

Let us consider that the source emits a pair of qubits in a maximally entangled state,

$$|\psi\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}. \tag{1.10}$$

When Alice receives $x = 1$ she measures the observable σ_x and when she receives $x = 2$ she measures the complementary observable σ_z. Throughout this thesis, we denote the standard Pauli matrices by $(\sigma_x, \sigma_y, \sigma_z)$. Bob measures the observables $(\sigma_x + \sigma_z)/\sqrt{2}$ and $(\sigma_x - \sigma_z)/\sqrt{2}$ corresponding to $y = 1$ and $y = 2$ respectively. From the Born rule, we compute the expectation values to be

$$\langle A_1 B_1 \rangle = \langle A_1 B_2 \rangle = \langle A_2 B_1 \rangle = -\langle A_2 B_2 \rangle = \frac{1}{\sqrt{2}}, \tag{1.11}$$

which leads to

$$\mathcal{S}_{\mathrm{chsh}} = 2\sqrt{2}. \tag{1.12}$$

This is a violation of the CHSH inequality and it means that the correlations predicted by quantum theory cannot be reproduced in any local model. Hence, we conclude that the set of quantum correlations is larger than the set of local variable correlations. This fact is known as *Bell's theorem*.

Let us emphasise the role of entanglement in quantum violations of Bell inequalities. Although the above example is based on the maximally entangled two-qubit state, it appears reasonable to ask if entanglement really is necessary for quantum correlations? The answer is yes and the reason

[1] It is the author's opinion that whoever introduced this terminology to physicists should consider undertaking an elementary course on public communication.

is simple. By definition, a state is separable (i.e. not entangled) if it can be written as a convex combination of local states,

$$\rho = \sum_\lambda q_\lambda \rho_\lambda^{\mathrm{A}} \otimes \sigma_\lambda^{\mathrm{B}}, \tag{1.13}$$

where $\rho_\lambda^{\mathrm{A}}$ and $\sigma_\lambda^{\mathrm{B}}$ are quantum states local to Alice's and Bob's respective laboratories and $\{q_i\}_i$ is a probability distribution. When such a state is measured locally, the Born rule gives the probabilities

$$p(a,b|x,y) = \sum_\lambda q_\lambda \operatorname{tr}\left(A_{a|x}\rho_\lambda^{\mathrm{A}}\right) \operatorname{tr}\left(B_{b|y}\sigma_\lambda^{\mathrm{B}}\right). \tag{1.14}$$

By direct comparison to Eq (1.2), we see that the correlations admit a local model by choosing $p(\lambda) = q_\lambda$, $p(a|x,\lambda) = \operatorname{tr}\left(A_{a|x}\rho_\lambda^{\mathrm{A}}\right)$ and $p(b|y,\lambda) = \operatorname{tr}\left(B_{b|y}\sigma_\lambda^{\mathrm{B}}\right)$. Hence, as intuition may suggest, entanglement is a necessary condition for quantum violations of Bell inequalities.

An interesting question is the largest violation of the CHSH inequality that can be achieved in quantum theory. This roughly corresponds to asking: *to what extent does quantum correlations defy local models?* As it turns out, the obtained violation is also the largest possible. It is instructive to see how one arrives at this conclusion. Let us define

$$|\alpha_x\rangle = A_x \otimes \mathbb{I}|\psi\rangle \qquad\qquad |\beta_y\rangle = \mathbb{I} \otimes B_y|\psi\rangle. \tag{1.15}$$

We can now write the CHSH expression in a quantum model as

$$\mathcal{S}_{\mathrm{chsh}} = \left(\langle\alpha_1| + \langle\alpha_2|\right)|\beta_1\rangle + \left(\langle\alpha_1| - \langle\alpha_2|\right)|\beta_2\rangle. \tag{1.16}$$

Clearly, to make the value as large as possible, we would like that $|\beta_1\rangle$ is aligned with $(\langle\alpha_1| + \langle\alpha_2|)$ and $|\beta_2\rangle$ is aligned with $(\langle\alpha_1| - \langle\alpha_2|)$. Note that the magnitude of both $|\alpha_x\rangle$ and $|\beta_y\rangle$ is one since $\langle\alpha_x|\alpha_x\rangle = \langle\psi|A_x^2 \otimes \mathbb{I}|\psi\rangle = 1$ where the last steps follows from the fact that every binary observable O obeys $O^2 = \mathbb{I}$. Hence, we have arrived at

$$\mathcal{S}_{\mathrm{chsh}} \leq \||\alpha_1\rangle + |\alpha_2\rangle\| + \||\alpha_1\rangle - |\alpha_2\rangle\| = \sqrt{2}\left(\sqrt{1 + \Re\langle\alpha_1|\alpha_2\rangle} + \sqrt{1 - \Re\langle\alpha_1|\alpha_2\rangle}\right). \tag{1.17}$$

Consider the function $f(x) = \sqrt{1+x} + \sqrt{1-x}$ for $x \in [-1,1]$. By writing

$$f(x) = \sqrt{f(x)^2} = \sqrt{2 + 2\sqrt{1-x^2}}, \tag{1.18}$$

it becomes clear that the maximum of this function is attained at $x = 0$. Since $f(0) = 2$, we have proven that quantum theory cannot exceed $\mathcal{S}_{\mathrm{chsh}} = 2\sqrt{2}$.

The CHSH inequality is the simplest of Bell inequality (at least in the modern sense of the word). It turns out that in the binary input/output scenario we considered, it constitutes a face

of the polytope of local variable correlations. Moreover, it is known that it in fact is the only non-trivial face of that polytope. However, when more inputs and/or outputs are introduced for Alice and Bob, the fauna of Bell inequalities rapidly grows larger and determining them all becomes an increasingly challenging problem. For a thorough guide to the study of quantum correlations in Bell experiments, we refer the reader to the review article [12]. In what follows, we will often substitute the cumbersome expression "quantum violation of a Bell inequality" with the easier (but not uncontroversial [13]) expression "quantum nonlocality" or "Bell nonlocality".

1.2 Quantum contextuality

Quantum nonlocality is relevant to experiments that do not involve communication. In contrast, *quantum contextuality* offers a generalisation of quantum nonlocality in such a way that it applies not only to Bell experiments but also to more general experiments that may feature explicit communication between the involved parties. Quantum contextuality is an approach to quantum correlations that does not privilege entangled states or space-like separation.

Quantum contextuality means that correlations in quantum theory cannot be explained by a hidden variable model that is both deterministic and noncontextual. Such a model ascribes pre-existing outcomes to every measurement without regard to the context in which the measurements are performed. Here, *context* refers to the other possible measurements that are performed jointly with the measurement of interest. For instance, if A and B are commuting observables and A and C also are commuting observables, we could either measure both A and B jointly or both A and C jointly. In the former, we say that we measured A in the context of B and in the latter we measured A in the context of C. Noncontextual models aim to explain the outcome statistics of A without regard to its context i.e. by ignoring which other compatible measurements that are jointly implemented. The fact that quantum theory does not admit a noncontextual description is known as the *Kochen-Specker theorem* [5]. It fundamentally traces back to quantum observables in general being non-commutative. For instance, in our example there is no need for B and C to commute.

However, this makes contextuality inherently quantum, in the sense that it is a property that is native to the Hilbert space formalism of quantum theory for closed systems. Depending on one's personal bent, this may be considered an unappealing feature. Should we not be able to talk about contextuality in nature without assuming that quantum theory governs its laws? For such an audience (which includes the author), it is therefore interesting that it is possible to generalise such "Kochen-Specker" contextuality to instead be formulated on operational grounds. In an operational approach, contextuality is no longer a property of operators on Hilbert spaces (to which quantum theory gives physical meaning) but instead a property of probability distributions measured in

laboratories. Such *operational contextuality*, as first introduced in Ref [14], is our focus in this thesis.

Consider a prepare-and-measure experiment, i.e. an experiment in which a sender (Alice) prepares states and a receiver (Bob) measures them. In her lab, Alice implements a preparation procedure that we denote $\mathbf{P}$. This can be thought of as a set of instructions to be implemented in a laboratory for creating a physical system. The preparation is communicated to Bob who implements a measurement procedure $\mathbf{M}$ which returns the outcome b. After many repetitions, the experiment generates the probability distribution $p(b|\mathbf{P},\mathbf{M})$. We can offer to explain the observed probabilities by employing a ontological model (hidden variable model). Such a model imagines that the preparation corresponds to a set of ontic states whose elements we denote λ. The ontic states represent the ontology of the system, i.e. they are not states of knowledge but in fact the true state of nature's affairs. A preparation may, however, correspond to a distribution over such states, which we denote $p(\lambda|\mathbf{P})$. Then, given only the preparation, we cannot determine the precise underlying ontology. Nevertheless, once a specific λ has been relayed to Bob, he can stochastically determine his measurement output. Thus, a general ontological model reads

$$p(b|\mathbf{P},\mathbf{M}) = \sum_{\lambda} p(\lambda|\mathbf{P})p(b|\mathbf{M},\lambda). \tag{1.19}$$

An important property of this model is that a convex combination of preparation procedures implies a convex combination of the corresponding distributions over the ontic states, i.e. if we implement $\mathbf{P}_1$ with probability q and $\mathbf{P}_2$ with probability $1-q$, then the resulting procedure $\mathbf{P}' = q\mathbf{P}_1 + (1-q)\mathbf{P}_2$ corresponds to the ontic state distribution

$$p(\lambda|\mathbf{P}') = qp(\lambda|\mathbf{P}_1) + (1-q)\,p(\lambda|\mathbf{P}_2). \tag{1.20}$$

Notice that every distribution can be explained through a ontological model, i.e. these models are always successful. For example, in quantum theory the preparation procedure is simply a density matrix $\mathbf{P} \simeq \rho$ living in Hilbert space. Ontic states in quantum theory are pure states while a general quantum state is mixed and represented by a probability distribution over some pure states $|\psi_\lambda\rangle$. Similarly, the measurement procedure is a POVM $\mathbf{M} \simeq \{M_b\}_b$ and the response of Bob is given by the Born rule, i.e. $p(b|\mathbf{M},\lambda) = \langle\psi_\lambda|M_b|\psi_\lambda\rangle$. In this manner, we can view quantum theory as a ontological model. This also highlights the fact that this operational approach to contextuality does not assume outcome determinism.

Matters become more interesting once we impose a non-trivial assumption on our ontological models. In order to do that, we need to introduce *operationally equivalent* procedures. Consider Alice having two preparation procedures $\mathbf{P}_1$ and $\mathbf{P}_2$. To these, we could apply any measurement

procedure $\mathbf{M}$. We say that if there exists no measurement procedure that allows us to distinguish between $\mathbf{P}_1$ and $\mathbf{P}_2$, then the two preparation procedures are operationally equivalent. Formally,

$$\forall \mathbf{M}: \quad p(b|\mathbf{P}_1, \mathbf{M}) = p(b|\mathbf{P}_2, \mathbf{M}) \quad \Leftrightarrow \quad \mathbf{P}_1 \sim \mathbf{P}_2. \tag{1.21}$$

In analogy, we say that two measurement procedures $\mathbf{M}_1$ and $\mathbf{M}_2$ are operationally equivalent if there exists no preparation procedure that can distinguish between them:

$$\forall \mathbf{P}: \quad p(b|\mathbf{P}, \mathbf{M}_1) = p(b|\mathbf{P}, \mathbf{M}_2) \quad \Leftrightarrow \quad \mathbf{M}_1 \sim \mathbf{M}_2. \tag{1.22}$$

We can now formulate the notion of a context in the operational framework. All operationally equivalent ways of realising a preparation (measurement) are said to be contexts of that preparation (measurement). We therefore say that a ontological model is *noncontextual* if it can model the distribution $p(b|\mathbf{P}, \mathbf{M})$ independently of the context of the preparation and/or measurement procedures. Specifically, we say that the model is *preparation noncontextual* if operationally equivalent preparation procedures imply identical distributions over the ontic states, i.e.

$$\mathbf{P}_1 \sim \mathbf{P}_2 \quad \Rightarrow \quad p(\lambda|\mathbf{P}_1) = p(\lambda|\mathbf{P}_2). \tag{1.23}$$

Similarly, we say that the model is *measurement noncontextual* if operationally equivalent measurement procedures imply identical response functions, i.e.

$$\mathbf{M}_1 \sim \mathbf{M}_2 \quad \Rightarrow \quad p(b|\mathbf{M}_1, \lambda) = p(b|\mathbf{M}_2, \lambda). \tag{1.24}$$

In summary, noncontextuality is the principle that if we in principle cannot distinguish two procedures then they are ontologically identical.

Let us exemplify all this in quantum theory. Assume that the preparation procedure corresponds to $\rho = \mathbb{I}/2$. We can realise it in two (or more) different contexts. In a first procedure, we flip an unbiased coin and prepare the corresponding eigenstate of σ_x. In a second procedure, we flip an unbiased coin and prepare the eigenstates of σ_z (see Figure 1.2). A preparation noncontextual model takes no consideration of the choice of context and therefore has that $\frac{1}{2}p(\lambda|+x) + \frac{1}{2}p(\lambda|-x) = \frac{1}{2}p(\lambda|+z) + \frac{1}{2}p(\lambda|-z)$. Similarly, consider three binary measurements $\{M_{b|y}\}$ for $b \in [2]$ and $y \in [3]$ as represented in Figure 1.2. We find two contexts by noticing that uniformly mixing $M_{1|1}$, $M_{1|2}$ and $M_{1|3}$ returns the same as uniformly mixing $M_{2|1}$, $M_{2|2}$ and $M_{2|3}$. Therefore, a measurement noncontextual model requires that $\frac{1}{3}p(b|M_{1|1}, \lambda) + \frac{1}{3}p(b|M_{1|2}, \lambda) + \frac{1}{3}p(b|M_{1|3}, \lambda) = \frac{1}{3}p(b|M_{2|1}, \lambda) + \frac{1}{3}p(b|M_{2|2}, \lambda) + \frac{1}{3}p(b|M_{2|3}, \lambda)$.

With the basic concepts in place, the key question is whether quantum correlations admit a noncontextual ontological model. As we have seen, this question is not entirely unambiguous

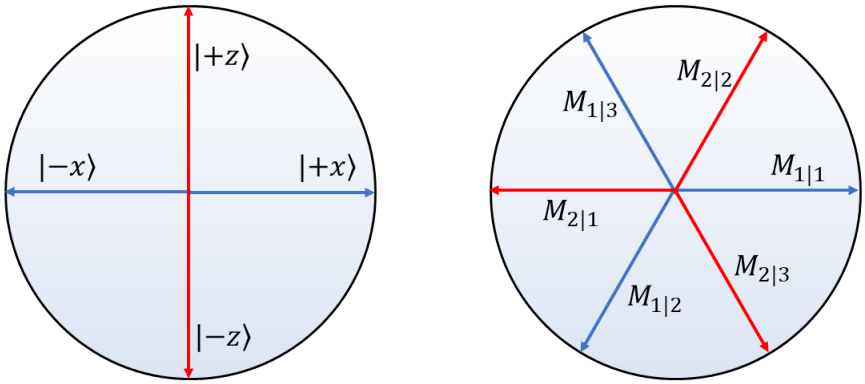

Figure 1.2: Illustratation of different contexts for preparations and measurements in a disk of the Bloch sphere. Left: realisation of the maximally mixed qubit state through the mixture of the eigenstates of σ_x and σ_z respectively. Right: realisation of the maximally mixed measurement through the mixture of the outcome-one and outcome-two operators respectively of three separate projective measurements.

since we could consider noncontextuality as it applies to states and measurements (or both of simultaneously). Let us begin with addressing measurement noncontextuality. If we assume that measurement outcomes are deterministic, i.e. that the response function $p(b|\mathbf{M}, \lambda)$ always is either zero or one, then there is no measurement noncontextual model that accounts for all predictions of quantum theory. This statement is in fact just the Kochen-Specker theorem; see Ref [15] for a considerably simpler proof than that originally given by Kochen and Specker. It serves to highlight that by additionally imposing outcome determinism on measurement noncontextual models, we recover standard Kochen-Specker contextuality as a limiting case of operational contextuality. However, in the operational approach we do not assume outcome determinism. Then, the question of noncontextuality in quantum theory becomes: does the Born rule only depends on the POVM elements or also on their contexts? Clearly, it does only depends on the POVM elements. Therefore, quantum theory admits a measurement noncontextual model.

However, while quantum theory is measurement noncontextual, it is in fact preparation contextual. The simplest proof known to the author follows the example of Ref [14]. We define the six qubit states $\{|\psi_{ij}\rangle\}$ where $i \in \{a, b, c\}$ and $j \in \{0, 1\}$ as follows

$$|\psi_{a0}\rangle = |0\rangle \qquad |\psi_{b0}\rangle = \frac{1}{2}|0\rangle + \frac{\sqrt{3}}{2}|1\rangle \qquad |\psi_{c0}\rangle = \frac{1}{2}|0\rangle - \frac{\sqrt{3}}{2}|1\rangle, \tag{1.25}$$

with $|\psi_{i1}\rangle$ defined as the orthogonal complement to $|\psi_{i0}\rangle$. These six states effectively correspond to the six Bloch vectors illustrated in Figure 1.2 (right). We write $\rho_{ij} = |\psi_{ij}\rangle\langle\psi_{ij}|$. Consider now

the three procedures corresponding to uniformly mixing ρ_{i0} and ρ_{i1} for each i and denote them $\mathbf{P}_i$. Define also the two procedures corresponding to uniformly mixing ρ_{1j}, ρ_{2j} and ρ_{3j} for each j and denote them $\mathbf{P}'_j$. Evidently, all five procedures are operationally equivalent:

$$\frac{1}{2}\mathbb{I} = \frac{1}{2}\rho_{10} + \frac{1}{2}\rho_{11} = \frac{1}{2}\rho_{20} + \frac{1}{2}\rho_{21} = \frac{1}{2}\rho_{30} + \frac{1}{2}\rho_{31} \tag{1.26}$$

$$= \frac{1}{3}\rho_{10} + \frac{1}{3}\rho_{20} + \frac{1}{3}\rho_{30} = \frac{1}{3}\rho_{11} + \frac{1}{3}\rho_{21} + \frac{1}{3}\rho_{31}. \tag{1.27}$$

The assumption of preparation noncontextuality implies that the ontic state distribution associated to each of these five procedures must be the same:

$$p(\lambda|\mathbf{P}_1) = p(\lambda|\mathbf{P}_2) = p(\lambda|\mathbf{P}_3) = p(\lambda|\mathbf{P}'_1) = p(\lambda|\mathbf{P}'_2). \tag{1.28}$$

When combining this with the convexity of ontic state distributions, we obtain that

$$\mu(\lambda) \equiv \frac{1}{2}p(\lambda|\rho_{10}) + \frac{1}{2}p(\lambda|\rho_{11}) = \frac{1}{2}p(\lambda|\rho_{20}) + \frac{1}{2}p(\lambda|\rho_{21}) = \frac{1}{2}p(\lambda|\rho_{30}) + \frac{1}{2}p(\lambda|\rho_{31})$$
$$= \frac{1}{3}p(\lambda|\rho_{10}) + \frac{1}{3}p(\lambda|\rho_{20}) + \frac{1}{3}p(\lambda|\rho_{30}) = \frac{1}{3}p(\lambda|\rho_{11}) + \frac{1}{3}p(\lambda|\rho_{21}) + \frac{1}{3}p(\lambda|\rho_{31}). \tag{1.29}$$

Now, notice that ρ_{i0} and ρ_{i1} are orthogonal for every i. This means that each such pair of states is fully distinguishable. Hence, they must belong to non-overlapping ontic state distributions. The reason is that if they had support for some common ontic variables then upon receiving such a variable one cannot decide whether it came from ρ_{i0} or from ρ_{i1} and hence one could not distinguish them. Thus, it must hold that

$$p(\lambda|\rho_{i0})p(\lambda|\rho_{i1}) = 0. \tag{1.30}$$

Thus, the existence of a preparation noncontextual model requires equations (1.29) and (1.30) to be compatible. However, a straightforward inspection of these equations shows that the only solution possible is the all-zero solution implied by $\mu(\lambda) = 0$, which evidently is not a probability distribution. This contradiction proves that there is no preparation noncontextual model for our ensemble of quantum states.

Whereas we phrased ontological models as appearing in prepare-and-measure experiments, the above formalism is in fact not limited to physical scenarios that involve communication. An illuminating example is that Bell nonlocality can emerge as a special instance of preparation contextuality. This fact has been noted in a number of works and proved in (perhaps) as many different ways (see e.g. [14, 16, 17, 18, 19]). The simplest derivation is (in the author's opinion) as follows. Consider a Bell scenario where Alice's and Bob's inputs are x and y respectively and their respective outputs are a and b. When Alice performs her measurement and registers her outcome, she remotely prepares

Bob in a post-measurement state. In an ontological model, we associate this remote preparation to a distribution over ontic states $p(\lambda|a,x)$. The conditional probability that Alice remotely prepares Bob in a particular state is given by $p(a|x,y)$. The response function of Bob reads $p(b|y,\lambda)$. Hence, the ontological model becomes

$$p(a,b|x,y) = \sum_{\lambda} p(a|x,y)p(\lambda|a,x)p(b|y,\lambda). \tag{1.31}$$

The no-signaling principle implies that the outcome of Alice is not influenced by Bob's setting; $p(a|x,y) = p(a|x)$. Then, by applying Bayes' rule, we have that

$$p(a|x,y)p(\lambda|a,x) \overset{\text{NS}}{=} p(a|x)p(\lambda|a,x) = p(\lambda|x)p(a|x,\lambda). \tag{1.32}$$

However, the no-signaling principle asserts that the convex combination of preparations associated to $\{(a,x)\}_a$ for a given x are operationally equivalent (since they cannot be distinguished by Bob). A preparation noncontextual model therefore has that $\sum_a p(a|x)p(\lambda|a,x)$ is independent of x. This necessitates that $p(\lambda|x) = p(\lambda)$. Thus, our preparation noncontextual ontological model takes the form

$$p(a,b|x,y) = \sum_{\lambda} p(\lambda)p(a|x,\lambda)p(b|y,\lambda). \tag{1.33}$$

This is precisely the form of a local hidden variable model. We conclude that every local model in a Bell scenario also is preparation noncontextual and that Bell nonlocality implies preparation contextuality.

1.3 Quantum communication complexity

Let us now focus entirely on quantum correlations arising in experiments that involve communication. Imagine that we have two separated parties who do not share any entanglement. Clearly, they cannot produce any correlations. However, if we allow them to communicate, correlations can be established. It appears natural that the more the parties are allowed to communicate, the stronger correlations should they be able to create. Therefore, a natural question is: *how strong correlations can be created from a given amount of communication?* This falls under the broad umbrella of *communication complexity problems* (CCPs). CCPs can be of two kinds, 1) find the smallest amount of communication needed to complete a task, and 2) with a given amount of communication, perform a task as well as possible. Our focus throughout this thesis is on the latter.

The simplest form of a CCP considers two separate parties, for simplicity named Alice and Bob, who each hold some data x and y respectively. Each party is unaware of the data held by the other party. Their aim is to collaborate in such a way that one of them (let us say Bob) can

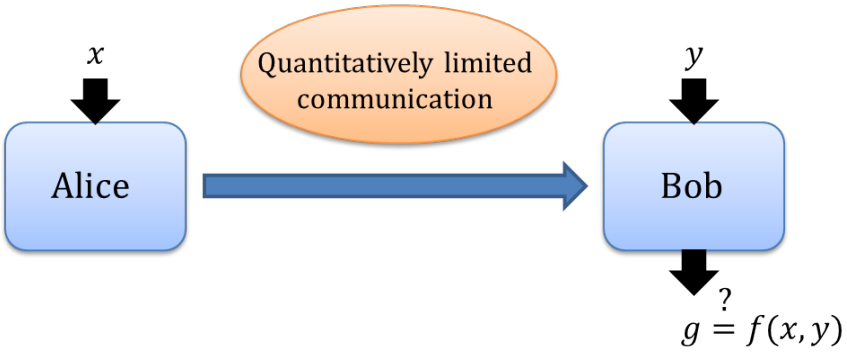

Figure 1.3: Communication complexity problem. Alice (Bob) receives input x (y). She is allowed to send Bob a restricted amount communication in order to aid his evaluation of a task function $f(x, y)$. Alice and Bob collaborate towards Bob outputting a guess g that as often as possible coincides with the value of $f(x, y)$.

evaluate some task function $f(x, y)$ that depends non-trivially on both their data. Of course, the only way that this could possibly happen is if Alice and Bob communicate. For example, Alice could simply send x to Bob. Now holding both x and y, he could compute $f(x, y)$ and the task would be successfully completed. The more interesting setting for a CCP is when the allowed amount of communication is quantitatively limited in such a way that Alice no longer can send all her data to Bob. Instead, the partners must adopt a more sophisticated strategy that allows Bob to accurately evaluate f with as high a probability as possible while respecting the allowed amount of communication (see Figure 1.3). The essence of the task is to optimally manage communication resources towards creating the strong correlations needed to evaluate $f(x, y)$.

Let us illustrate this in the simple example of Ref [20]. Imagine that a referee supplies Alice and Bob with random and independent inputs. Alice receives two bits $x_0, x_1 \in \{0, 1\}$ whereas Bob receives a single bit $y \in \{0, 1\}$. The referee then presents Alice and Bob with the following CCP: Bob must compute the binary function $f(x_0, x_1, y) = x_y$ but Alice may communicate no more than one bit to Bob. That is, if $y = 0$ Bob wants to know x_0 and if $y = 1$ he wants to know x_1. Since Alice does not know the value of y, she must try to compress her two bits into a single bit message from which Bob can extract as much useful information as possible. The average probability of succeeding with the task is then given by

$$p_{\text{suc}} = \frac{1}{8} \sum_{x_0, x_1, y} p(g = f | x_0, x_1, y), \tag{1.34}$$

where g denotes Bob's guess for the value of $f(x_0, x_1, y)$. In a classical model, Alice must use some encoding function $E : \{0, 1\}^2 \to \{0, 1\}$ to encode her two bits into one bit. The encoding creates

a message $m = E(x_0, x_1)$ that is relayed to Bob. Bob now holds two bits; his input y and the message m. He must apply some decoding function $D : \{0,1\}^2 \to \{0,1\}$ in order to construct his guess $g = D(y, m)$ for the task function $f(x_0, x_1, y)$. What is the best possible pair of encoding and decoding functions? Since the size of our problem is rather small, we could easily write out all possible encoding and decoding functions. In fact, there are only 16 different functions that map two bits into one bit. For clarity, let us write them all out:

$$E_1 = 0, \qquad E_2 = x_0, \qquad E_3 = x_1, \qquad E_4 = x_0 \oplus x_1, \qquad E_5 = x_0 x_1,$$
$$E_6 = x_0 \oplus x_0 x_1, \qquad E_7 = x_1 \oplus x_0 x_1, \qquad E_8 = x_0 \oplus x_1 \oplus x_0 x_1, \tag{1.35}$$

where $\oplus$ denotes addition modulo 2. Note that we obtain eight more encoding functions (bringing us to the total of 16) by simply adding 1 to each of the above eight encoding functions - but these can be neglected since we could simply have Bob adding 1 to the message he receives before proceeding with his decoding. Moreover, E_1 and E_4 are poor choices since the former is independent of x and the latter scrambles both useful pieces of information (x_0 and x_1). However, E_5 is a good choice: the values $x_0 x_1 = \{00, 01, 10\}$ are all mapped to 0 whereas $x_0 x_1 = 11$ is mapped to 1. Thus, if Bob receives 1 he knows both x_0 and x_1 and thus outputs $g = f$, but if he receives 0 the best he can do is to guess on $g = 0$ which is correct in 2/3 of the cases. Thus, he would find $p_{\text{suc}} = 1/4 \times 1 + 3/4 \times 2/3 = 3/4$. A little inspection shows that this value cannot be improved with any of the remaining encoding functions. Therefore, the best possible success rate in a classical implementation of the CCP is

$$p_{\text{suc}} \overset{\text{classical}}{=} \frac{3}{4}. \tag{1.36}$$

It turns out that Alice and Bob can do better if they let their one-bit communication be assisted by entanglement [20]. That is, we complement the picture by letting Alice and Bob share a bipartite state $|\psi\rangle$ which Alice and Bob can locally measure. This *entanglement-assisted classical communication* enables Alice and Bob to use their local outcomes as advice for how to encode and decode the classical communication. Importantly, this does not violate the one-bit communication limit due to the fact that entanglement alone cannot be used for communication. There is no way for Alice to send any information to Bob by means of $|\psi\rangle$ if the classical communication is removed. This is closely related to the no-signaling principle. However, the strong correlations that entanglement can give rise to, in particular through the violation of a Bell inequality, constitutes a useful aid for communications.

In what precise way can Alice and Bob exploit their shared entanglement to perform the task? Alice and Bob will conduct a test of the CHSH inequality. First, we remind ourselves that

$$\langle A_x B_y \rangle = p(a \oplus b = 0|x, y) - p(a \oplus b = 1|x, y) = (-1)^{xy} \left(2p(a \oplus b = xy|x, y) - 1\right), \tag{1.37}$$

where we have used the normalisation of $p(a,b|x,y)$ and for simplicity taken $x,y,a,b \in \{0,1\}$. This allows us to write the CHSH inequality (1.3) on the form

$$\mathcal{S}_{\text{chsh}} \equiv \frac{1}{4}\sum_{x,y} p(a \oplus b = xy|x,y) \leq \frac{3}{4}. \tag{1.38}$$

The best quantum implementation that we previously saw leading to $\langle A_x B_y \rangle = (-1)^{xy}/\sqrt{2}$ translates into

$$\mathcal{S}_{\text{chsh}} \overset{\text{Q}}{\leq} \frac{1}{2}\left(1 + \frac{1}{\sqrt{2}}\right) \approx 85.4\%. \tag{1.39}$$

In the quantum implementation of the CCP, Bob uses his bit y as his measurement setting. Alice maps her two bits into her binary measurement setting that enters the CHSH inequality test. She chooses her setting as $x \equiv x_0 \oplus x_1$. Now, Alice uses her outcome together with her input x_0 to construct the binary message $m = a \oplus x_0$ which she sends to Bob. On his side, Bob uses his outcome b to decode the received message into his final guess for the task function, $g = m \oplus b$. Thus, we have that

$$g = a \oplus b \oplus x_0. \tag{1.40}$$

The average success probability in the task then reads

$$p_{\text{suc}} = \frac{1}{8}\sum_{x_0,x_1,y} p(a \oplus b \oplus x_0 = x_y|x_0,x_1,y). \tag{1.41}$$

If we have $y = 0$ the winning condition is $a \oplus b = 0$ whereas if $y = 1$ the winning condition is $a \oplus b = x_0 \oplus x_1$. However, we know that a quantum implementation of the CHSH inequality can achieve $p(a \oplus b = xy|x,y) = 85.4\%$ for every x and y. Since we have $x = x_0 \oplus x_1$, this is precisely the same as the winning condition in the CCP. Hence, if Alice and Bob share a maximally entangled state and perform measurements that maximally violate the CHSH inequality, their entanglement-assisted classical communication can achieve

$$p_{\text{suc}} = \frac{1}{2}\left(1 + \frac{1}{\sqrt{2}}\right) \approx 85.4\%. \tag{1.42}$$

Importantly, this outperforms the best possible classical implementation and thus constitutes an example of quantum correlations $p(b|x_0,x_1,y)$ in the communication task.

The specific CCP that we have considered is known as a *random access code* [21]. It serves to illustrate that quantum resources can improve communication complexity beyond anything achievable by purely classical means. For a more general introduction to the use of quantum nonlocality to enhance communication complexity, we refer the reader to Ref [22].

However, quantum nonlocality is not the only approach to enhancing communication complexity beyond classical constraints. An interesting alternative is to instead of supplying Alice and Bob

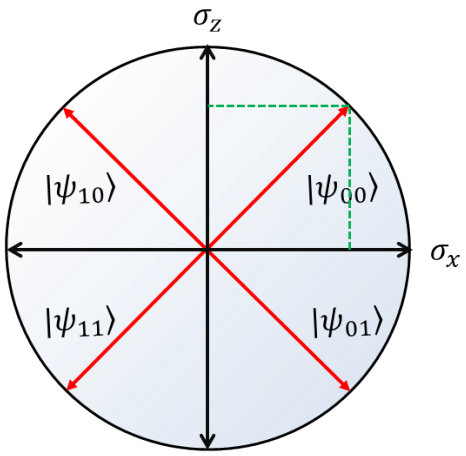

Figure 1.4: Random access code with quantum communication. Alice's four qubit states $|\psi_{x_0x_1}\rangle$ form a square in the xz-disk of the Bloch sphere (red arrows). Bob's two measurements form a rotated square in the same disk (black arrows).

with shared entanglement, one substitutes the classical communication for *quantum communication*. This means that instead of communicating a classical bit, Alice can communicate a qubit to Bob. The substitution is justified because of Holevo's theorem which implies that a quantum system of dimension d cannot carry more than $\log d$ bits of information [23]. Let us see how well Alice and Bob can perform the random access code when Alice sends qubits to Bob. We let Alice associate her four possible inputs to the four qubit states $\{|\psi_{x_0x_1}\rangle\}$. We choose them as

$$|\psi_{00}\rangle = \cos\frac{\pi}{8}|0\rangle + \sin\frac{\pi}{8}|1\rangle \qquad |\psi_{01}\rangle = \sin\frac{\pi}{8}|0\rangle + \cos\frac{\pi}{8}|1\rangle \tag{1.43}$$

$$|\psi_{10}\rangle = \cos\frac{\pi}{8}|0\rangle - \sin\frac{\pi}{8}|1\rangle \qquad |\psi_{11}\rangle = \sin\frac{\pi}{8}|0\rangle - \cos\frac{\pi}{8}|1\rangle. \tag{1.44}$$

It is more illuminating to geometrically display these states in the xz-disk of the Bloch sphere (see Figure 1.4) where we see that Alice's four states form a square. Bob's decoding procedure is a quantum measurement. We choose his observables as σ_x and σ_z for $y = 0$ and $y = 1$ respectively. In Figure 1.4, this corresponds to the black diagonals. The probability of Bob's output satisfying $g = x_y$ corresponds to the illustrated overlap in the figure. Notice that this overlap is the same for every state and measurement for the successful outcome. Therefore, we need only to consider one of them to evaluate the success probability in the CCP. We have that

$$p_{\text{suc}} = \langle\psi_{00}|\frac{\mathbb{I}+\sigma_z}{2}|\psi_{00}\rangle = \cos^2\frac{\pi}{8} = \frac{1}{2}\left(1+\frac{1}{\sqrt{2}}\right). \tag{1.45}$$

By means of quantum communication, we have again outperformed the limitations of classical communication. Moreover, we have found precisely the same advantage as we earlier found by means

of entanglement-assisted classical communication. However, this does not mean that entanglement-assisted classical communication in general is equivalent to quantum communication in terms of its ability of enhancing CCPs beyond their classical limitations.

2

Quantum communication complexity

The use of quantum resources allows distributed parties to establish stronger-than-classical correlations while not communicating more information. Here, we present a broad exploration of quantum communication complexity, i.e. the correlations that can be established through quantitatively limited communication. The first section investigates the relation between Bell inequalities and CCPs: we show how violations of the formers enable advantages in the latters but also that not all forms of quantum nonlocality are useful for solving CCPs [24, 25]. The second section explores the relation between entanglement-assisted classical communication and quantum communication as resources for quantum communication complexity: we prove that the two resources give rise to different correlations that can be strongly dependent on the dimension of Hilbert space [26]. The third section focuses on quantum correlations established through the communication of d-dimensional systems and enhances previous methods for bounding their advantages in communication complexity problems by introducing highly efficient computer tools [27]. In the final section, we depart from dimension-bounded systems and instead investigate the limits of classical and quantum correlations constrained only by communication whose information content is limited by entropic quantities. This approach allows us to go beyond standard dimension-bounded systems and make more explicit the connection between correlations and information [28].

2.1 Bell nonlocality as a resource for communication complexity

In section 1.3, we exemplified how violations of the CHSH inequality can be used to power better-than-classical communication complexity in the random access code. Moreover, we even found that every violation of the CHSH inequality can enable such improvements. This link between a Bell inequality and a CCP is by no means an isolated incident. A multitude of works have shown

than different types of Bell inequalities can be directly related to CCPs in such a way that their violation implies communication advantages. For instance, this has been shown for the multipartite Mermin Bell inequalities [8], the three-outcome Collins-Gisin-Linden-Massar-Popescu (CGLMP) Bell inequalities [29] as well as its generalisation to any number of outcomes [30, 31], the biased CHSH inequalities [32, 33], the Elegant Bell inequality [34] and Bell inequalities for generalised random access codes [35, 36]. More generally, and encompassing some of the above mentioned cases, the link between Bell inequalities and CCPs has been shown for any bipartite correlation Bell inequality with binary outcomes[1] [37]. In view of this diverse fauna of links between Bell inequalities and CCPs, it appears relevant to ask how general the connection actually is. Our first step towards answering this question is by showing how to map a broad class of Bell inequalities to CCPs [25]. An appealing feature of this map is that it reproduces all the mentioned examples as special cases.

Let us begin by outlining a model for how Bell nonlocality is used to power CCPs. For simplicity, we restrict it to two parties but afterwards we will consider the straightforward extension to more parties. Consider a bipartite CCP in which Alice receives an input $X \in [N_{\mathrm{A}}]_0$ and Bob receives an input $Y \in [N_{\mathrm{B}}]_0$. We use the notation $[s]_0$ to denote the set $\{0, \ldots, s-1\}$[2]. Alice may send a message $m \in [M]_0$ to Bob from which he constructs an output $g \in [G]_0$ which is rewarded (or penalised) with $t^g_{X,Y}$ points. To ensure that the game is non-trivial, we should have $M < N_{\mathrm{A}}$ since otherwise Alice can simply send her input to Bob. We say that the tuple $(N_{\mathrm{A}}, N_{\mathrm{B}}, M, G)$ corresponds to a scenario, i.e. it tells us the size of the inputs and outputs of all parties. Within the scenario, Alice and Bob can consider a general (linear) CCP. It corresponds to a task function of the form

$$\mathcal{S} = \sum_{g,X,Y} t^g_{X,Y} p(g|X,Y), \tag{2.1}$$

where $p(g|X,Y)$ is the probability distribution observed in the game.

Classical models

A classical model has Alice encoding her input into a message via an encoding function $E : [N_{\mathrm{A}}]_0 \to [M]_0$ and Bob uses a decoding function $D : [N_{\mathrm{B}}]_0 \times [M]_0 \to [G]_0$ to construct g. In addition, we allow classical models the additional resource of *shared randomness*. Shared randomness is a string of pre-established classical data shared between the two parties. We denote the value of the shared

[1] These Bell inequalities can be written as linear combinations of expectation values. They do not involve marginal probabilities.

[2] Since we frequently want to consider modular sums, it is more convenient to count from zero than from one.

randomness by λ and its distribution by $p(\lambda)$[3]. Alice and Bob may use the shared randomness to coordinate their encoding and decoding functions. Hence, for every given value of λ, Alice and Bob will choose a specific encoding and decoding function (E_λ, D_λ) and obtain the distribution $p_\lambda(g|X,Y)$. Specifically, this becomes

$$p^{\mathrm{C}}_\lambda(g|X,Y) = \sum_m p(m|X,\lambda)p(g|m,Y,\lambda), \tag{2.2}$$

where $p(m|X,\lambda) = \delta_{E_\lambda(X),m}$ and $p(g|m,Y,\lambda) = \delta_{D_\lambda(m,Y),g}$. These are deterministic distributions that can be used to create a general classical distribution by varying the distribution of the shared randomness. The total distribution becomes

$$p^{\mathrm{C}}(g|X,Y) = \sum_\lambda p(\lambda)p^{\mathrm{C}}_\lambda(g|X,Y). \tag{2.3}$$

Therefore, it also follows that the classical set of correlations arising in a given scenario can be geometrically represented as a polytope whose vertices correspond to the deterministic distributions $p^{\mathrm{C}}_\lambda(g|X,Y)$. The fact that a polytope is characterised by a set of linear inequalities is the reason behind us focusing on linear CCPs. A quantum advantage in a nonlinear CCPs necessitates a quantum advantage in some linear CCP.

We can now define what is meant by the best classical score in a CCP as follows: it is the largest value of the task function obtainable by means of Alice communicating integer messages coordinated via shared randomness, i.e.

$$\mathcal{S}^{\mathrm{C}} = \max_{p(\lambda)} \mathcal{S}[p^{\mathrm{C}}(g|X,Y)]. \tag{2.4}$$

Importantly, since the task function is linear, the best classical score is obtained by a deterministic strategy. This follows from the fact that if we use a mixed strategy, it can never perform better than the best deterministic strategy originally used to create the mixed strategy. We can therefore w.l.g. write

$$\mathcal{S}^{\mathrm{C}} = \max_{\lambda} \mathcal{S}[p^{\mathrm{C}}_\lambda(g|X,Y)]. \tag{2.5}$$

In summary, the best classical score in a given CCP is obtained from finding the best pair of deterministic encoding and decoding functions compatible with the scenario. There are only finitely many such functions (as long as the inputs/outputs are finite) and therefore one only needs to evaluate the list $\{\mathcal{S}[p_\lambda(g|X,Y)]\}_\lambda$ and choose the largest value.

[3]Since we will only consider scenarios with finitely many inputs and outputs, it is sufficient to only consider finite values of λ. Therefore we can safely restrict to a probability distribution $p(\lambda)$ rather than a probability density.

Entanglement-assisted models

Let us now consider entanglement-assisted models. Here, Alice and Bob still communicate classically, but they additionally share a quantum state ρ. On this state, they may perform measurements with respective outcomes a and b. To choose their settings, they may use their inputs for the CCP. This process allows them to share correlations that take the form

$$p(a,b|X,Y) = \mathrm{tr}\left(A_{a|X} \otimes B_{b|Y}\rho\right), \tag{2.6}$$

where $\{A_{a|X}\}_a$ and $\{B_{b|Y}\}_b$ are the POVMs of Alice and Bob respectively for each respective input X and Y. Notice that the outcomes of Alice and Bob could in principle be of any alphabet and the shared state could be of any dimension. With these quantum correlations in hand, Alice and Bob can employ a classical encoding procedure. Alice uses an encoding function $E : [|a|]_0 \times [N_\mathrm{A}]_0 \to [M]_0$ and Bob uses a decoding function $D : [M]_0 \times [N_\mathrm{B}]_0 \times [|b|]_0 \to [G]_0$. While we could simply absorb the shared randomness into the entangled state (and indeed, we could also have had absorbed the inputs of Alice and Bob into a and b respectively when defining E and D) we choose to treat it separately in order to distinguish it as a classical resource. For a given value of the shared randomness, the quantum model reads

$$p^\mathrm{Q}_\lambda(g|X,Y) = \sum_{a,b,m} p(a,b|X,Y)p(m|a,X,\lambda)p(g|m,b,Y,\lambda).$$

and the total distribution becomes

$$p^\mathrm{Q}(g|X,Y) = \sum_\lambda p(\lambda)p^\mathrm{Q}_\lambda(g|X,Y). \tag{2.7}$$

Notice that we first perform the Bell inequality test and then communicate, i.e. once Bob receives the message, he has already measured the entangled state. We choose this order of events since it is more in line with the concept of space-like separation in Bell inequality tests. It is, however, also interesting to consider models in which the message can be used to influence Bob's setting.

From quantum nonlocality to quantum communication complexity

Let us now show how one can map any N-observer correlation Bell inequality in which each observer has d outcomes to a CCP. To this end, we name our N observers $O_1, \ldots, O_N$, label their respective measurements $x_1, \ldots, x_N$ and label their respective outputs $a_1, \ldots, a_N$. The class of Bell inequalities that we consider takes the form

$$\mathcal{B} = \sum_{\vec{x}} \sum_r c^r_{\vec{x}} p_{\vec{x}} \left(\sum_{i=1}^N o_i = f^r_{\vec{x}} \right) \overset{\text{local}}{\leq} C, \tag{2.8}$$

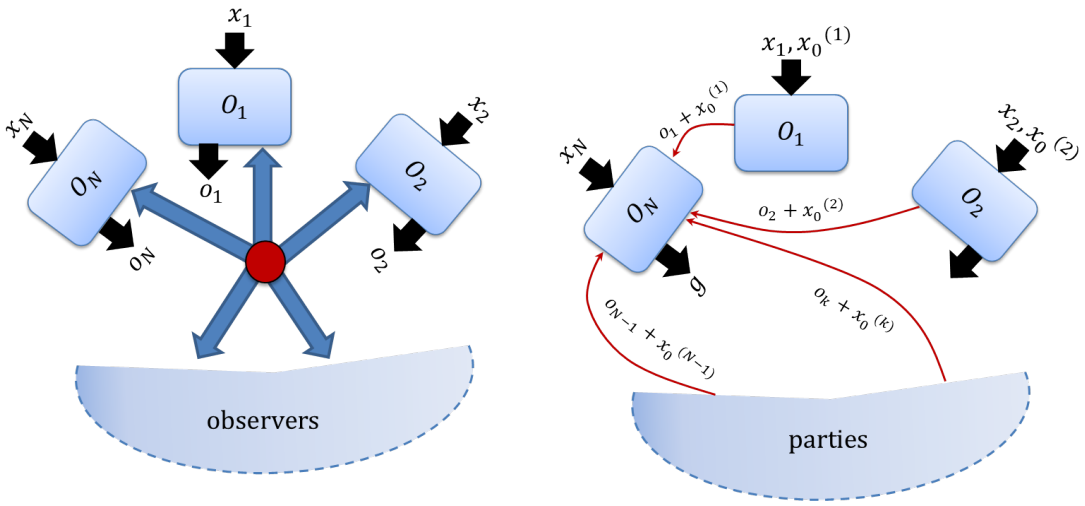

Figure 2.1: Map from Bell inequalities to CCPs. Left: multipartite Bell experiment with inputs $(x_1, \ldots, x_N)$ and outputs $(o_1, \ldots, o_N)$. Right: multipartite CCP with inputs $(x_k, x_0^{(k)})$ for party $k = 1, \ldots, N-1$ and input x_N for party O_N. The settings $\{x_i\}$ are used to perform the Bell experiment. Then, the first $N-1$ parties send the respective messages $m_k = o_k + x_0^{(k)}$ to O_N who outputs g.

where $\vec{x} = (x_1, \ldots, x_N)$, C is the local bound, $f_{\vec{x}}^r \in [d]_0$ and $c_{\vec{x}}^r$ are real coefficients. The relation $\sum_i o_i = f_{\vec{x}}^r$ is evaluated modulo d. These can be understood as the d-outcome multipartite generalisations of standard correlation Bell inequalities in the binary-outcome and bipartite setting.

In order to map such Bell inequalities to CCPs, consider the following procedure (see Figure 2.1). For $i = 1, \ldots, N-1$, let O_i have an input defined as the pair $X_i = (x_i, x_0^{(i)})$ where $x_0^{(i)} \in [d]_0$. The parties $O_1, \ldots, O_{N-1}$ may send a d-valued message $m_i \in [d]_0$ to O_N. O_N acts as the party who attempts to perform the computation. Therefore, O_N has a smaller input defined as $X_N = x_N$ and once the messages are received he produces a guess $g \in [d]_0$ and earns a payoff $c_{\vec{x}}^r / d^{N-1}$ whenever $g = f_{\vec{x}}^r + \sum_{i=1}^{N-1} x_0^{(i)}$. We choose the score in the CCP to emulate the witness in the Bell inequality. Specifically,

$$\mathcal{S} = \frac{1}{d^{N-1}} \sum_{\vec{x}, \vec{x}_0} \sum_r c_{\vec{x}}^r p_{\vec{x}} \left(g = f_{\vec{x}}^r + \sum_{i=1}^{N-1} x_0^{(i)} \right), \tag{2.9}$$

where $\vec{x}_0 = (x_0^{(1)}, \ldots, x_0^{(N-1)})$. In order to relate the Bell inequality to the CCP, we proceed as follows. Let the N parties share an entangled state and use their inputs $\vec{x}$ to perform a measurement with outcome $o_i \in [d]$. This corresponds to the Bell inequality test. Then, the parties O_i for $i = 1, \ldots, N-1$ send the respective message $m_i = o_i + x_0^{(i)} \mod d$ to O_N who outputs the guess $g = o_N + \sum_i m_i \mod d$. The intuition is that the input $x_0^{(i)}$ for each of the communicating parties acts as a scrambler that keeps O_N unaware of the input x_i. It is worth noticing that the parties

$O_1, \ldots, O_{N-1}$ only use part of their inputs for choosing a measurement setting. In this way, the scramblers cancel in the winning condition of the CCP and we find

$$\mathcal{S} = \mathcal{B}. \tag{2.10}$$

Therefore, for such additive communication strategies we classically have $\mathcal{S} \leq C$. Hence, a violation of the Bell inequality implies $\mathcal{S} > C$. This construction encompasses the many maps from specific Bell inequalities to CCPs encountered in the literature, see e.g. Refs [8, 29, 30, 31, 33, 35, 36, 37]. In what follows we refer to it as "the map from Bell inequalities to CCPs".

The instructive case of the CGLMP Bell inequality in CCPs

Our above presented construction rests on communication strategies that are additive. We saw that that violation of a Bell inequality of the form (2.8) is necessary and sufficient for a better-than-classical score under additive messages. What happens when classical messages are allowed to be more general? A partial answer is given in Ref [37]. By choosing binary outcomes in our Bell inequalities ($d = 2$), our map reduces to that derived in Ref [37]. Interestingly, it was shown that for these binary outcomes, additive messages yield the best classical score. In other words, one cannot hope to improve the classical score by adopting another communication strategy. This means that Bell inequality violation is necessary and sufficient for quantum advantages over general classical strategies. Several subsequent works have drawn on this result to establish the analogous relation between the violation of specific many-outcome Bell inequalities and CCPs. A prominent example is the mapping of the CGLMP inequalities to CCPs [29, 30, 31]. The CGLMP Bell inequalities [38] are facet Bell inequalities that generalise the CHSH inequality to any number of outputs per party. For instance, Ref [29] concluded that violation of the three-outcome CGLMP Bell inequality is necessary and sufficient for quantum advantages in the corresponding CCP. Formally, this was shown for additive classical communication strategies. However, as we now show, by going beyond binary-outcome Bell inequalities, one can no longer safely consider additive communication strategies to be optimal. By abandoning additive strategies, we disprove the main result of i.a. Ref [29] and thereby, find an interesting lead in the exploration of the relation between Bell nonlocality and CCPs.

Let us briefly present the CGLMP inequality. Alice and Bob receive binary inputs $x, y \in [2]_0$ and are asked to return ternary outputs $a, b \in [3]_0$. The Bell inequality reads

$$\mathcal{B}_{\text{cglmp}} = \frac{1}{4} \sum_{x,y} \left[P_{xy}(a + b = f_1) - P_{xy}(a + b = f_2) \right] \overset{\text{local}}{\leq} \frac{1}{2}, \tag{2.11}$$

where $f_1 = -xy$, $f_2 = -xy + (-1)^{x+y}$ and addition is modulo 3. It is well-known that quantum theory can violate this inequality. The maximal violation is $\mathcal{B}^{\text{Q}}_{\text{cglmp}} \approx 0.729$ and is obtained by Alice

and Bob performing suitable measurements on a partially entangled state [39]. Applying the map from Bell inequalities to CCPs, we supply Alice with a total of six inputs $x \in [2]_0$ and $x_0 \in [3]_0$ and Bob with a binary input $y \in [2]_0$. Alice may send a ternary message $m \in [3]_0$ to Bob whose guess $g \in [3]_0$ aims to maximise the task function

$$\mathcal{S}_{\text{cglmp}} = \frac{1}{12} \sum_{x_0,x,y} \left[p(g = x_0 + f_1 | x, y) - p(g = x_0 + f_2 | x, y) \right]. \tag{2.12}$$

Notice that this is a direct application of the presented map. As we have seen, additive communication strategies necessarily lead to $\mathcal{B}_{\text{cglmp}} = \mathcal{S}_{\text{cglmp}}$ and hence a quantum violation of the Bell inequality implies an advantage in the CCP.

However, let us now explore the full power of classical communication strategies. Alice's encoding function maps her six inputs into a ternary message - there are 3^6 such functions. Bob's decoding maps the ternary message and the binary input into a ternary guess - there are also 3^6 such functions. Hence, since we know that the best classical score is achieved for a deterministic strategy, we must consider a total of 3^{12} communication strategies and select the best one. This can be done by brute force. Interestingly, one finds that the additive strategy is not optimal. Instead, we have

$$\mathcal{S}^{\text{C}}_{\text{cglmp}} = \frac{2}{3}. \tag{2.13}$$

A strategy that achieves this optimal score is as follows. Choose

$$m(x_0, x) = \delta_{x,0}\delta_{x_0,2} + 2\delta_{x,1}\delta_{x_0,1} \mod 3 \tag{2.14}$$

$$g(m, y) = 2\delta_{y,0} m + \delta_{y,1} (m + 1) \mod 3. \tag{2.15}$$

Notice that by sending $m = 1$ or $m = 2$, Alice informs Bob of the precise value of x. This is conceptually very different from additive strategies since these are essentially emulating a Bell inequality test: by sending an additive message, one effectively preserves the no-signaling feature by scrambling any information one could have sent about the input in the Bell inequality test so that it cannot be extracted from the received message. Notably, entanglement-assisted strategies can still outperform the classical limitation since the maximal quantum violation of the CGLMP inequality is still larger than the classical limit of 2/3. But weak violations seem to no longer be sufficient to create a quantum advantage.

One natural question is whether also the quantum case can be improved by Alice and Bob maximally violating the CGLMP inequality but using another communication strategy. Numerical searches indicated no such improvement and the author conjectures that no such improvement exists. Moreover, if the distribution $p(g|x_0, x, y)$ obtained from violating the CGLMP inequality and using additive communication is not always useful for improving the CCP considered here,

could there be some other CCP for which it does offer a quantum advantage for as soon as the Bell inequality is violated? In other words, can we find another task function that is better suited to this entanglement-assisted strategy? To this end, we write

$$p(a,b|x,y) = vp^{\text{cglmp}}(a,b|x,y) + \frac{1-v}{9}, \tag{2.16}$$

where $p^{\text{cglmp}}(a,b|x,y)$ is the distribution that maximally violates the CGLMP inequality and $v \in [0,1]$ is the protocol visibility parameter. The distribution violates the CGLMP inequality when $v > 0.6861$. The probability distribution $p_v^{\text{Q}}(g|x_0,x,y)$, obtained from using the nonlocal correlations $p(a,b|x,y)$ in a CCP together with an additive communication strategy, beats the classical bound only when $v > 0.9149$. Is an intermediate value $v \in [0.6861, 0.9149]$ useful for a quantum advantage? We seek the largest v for which p_v^{Q} can be simulated by a classical model. This can be solved by means of the linear program

$$\begin{aligned} &\max_{p(\lambda)} v \qquad \text{s.t.} \qquad p(\lambda) \geq 0, \qquad \textstyle\sum_\lambda p(\lambda) = 1, \\ &\quad \text{and} \qquad p_v^{\text{Q}}(g|x_0,x,y) = \textstyle\sum_\lambda p(\lambda) p_\lambda^{\text{C}}(g|x_0,x,y). \end{aligned} \tag{2.17}$$

By considering $p_\lambda^{\text{C}}(g|x_0,x,y)$ for all possible deterministic strategies, we have found that the corresponding polytope of classical probability distributions has 47601 vertices. We have evaluated the linear program and found $v \approx 0.7943$. Hence, probability distributions $p_v^{\text{Q}}(g|x_0,x,y)$ for $0.7943 < v \leq 0.9149$ indeed imply an advantage over classical protocols in some CCP despite our particular CCP failing to detect it. However, when $0.6861 < v \leq 0.7942$ the CGLMP inequality is violated, but the probability distribution $p_v^{\text{Q}}(g|x_0,x,y)$ can be classically modeled. Hence, a violation of the CGLMP inequality combined with additive classical communication in our scenario does not always have the ability to outperform classical models.

Does Bell nonlocality imply communication advantages?

Although it is often believed that Bell nonlocality implies quantum advantages in CCPs, our example of the CGLMP inequality introduces a source of doubt. Is it really the case that Bell nonlocality always implies communication advantages? We now present evidence in favour of a negative answer. This evidence is based on proving a stronger statement than that already shown for the CGLMP example. We consider a specific facet Bell inequality and find a well-chosen set of correlations that violate it. Importantly, these correlations do not violate any other facet of the local polytope. We then fix the scenario for the CCP to be that commonly considered in the literature (and indeed also employed in our map). This is indeed an assumption, but arguably an intuitive one. Then, we allow Alice and Bob to employ any classical communication strategy that exploits the given quantum

nonlocality. Under these circumstances, we show that the nonlocal probability distribution does not have the ability of enhancing any CCP beyond its classical constraints in the scenario regardless of the communication strategy employed.

The simplest Bell scenario is that in which Alice and Bob have binary inputs and outputs. However, this is not useful for our purposes since the CHSH inequality is the only facet inequality and it is known that every violation of the CHSH inequality is useful for advantages in a CCP [20]. Therefore, we consider the second simplest setting; that in which Alice and Bob have ternary inputs and binary outputs. The facets of the local polytope are now two-fold; the lifted CHSH inequality (which is again not useful for us) and the so-called I_{3322} inequality (or the Froissart inequality). This inequality reads [40, 41]

$$I = -P_{\mathrm{A}}(0) - 2P_{\mathrm{B}}(0) - P_{\mathrm{B}}(1) + \sum_{x,y} T_{x,y} P(x,y) \leq 0, \tag{2.18}$$

where $P(x,y)$ is the probability of outputting $a = b = 0$, $P_{\mathrm{A}}(i) = p(a = 0|x = i)$, $P_{\mathrm{B}}(i) = p(b = 0|y = i)$ and $T = \{[1,1,1],[1,1,-1],[1,-1,0]\}$. Importantly, the I_{3322} inequality is not a correlation Bell inequality (it has marginals) and therefore is not within the scope of the map from Bell inequalities to CCPs. Therefore, even though it has binary outcomes, it does not fall into the broad class of binary outcome Bell inequalities whose violations are known to imply quantum advantages in CCPs.

Following previous literature and our map, a natural scenario in which to look for quantum advantages via violations of the I_{3322} inequality is that in which Alice receives a trit $x \in [3]_0$ and a bit $x \in [2]_0$ while Bob receives a trit $y \in [3]_0$. Alice sends a binary message $m \in [2]_0$ to Bob who outputs a binary guess $g \in [2]_0$. To confirm that the scenario indeed is a natural choice, let us consider a simple example. The maximal violation of the I_{3322} inequality using qubits is known to be $I^{\mathrm{Q}} = 1/4$ [40]. It can be achieved by Alice and Bob sharing a singlet state $|\psi^-\rangle$ and choosing their measurement Bloch vectors as

$$\begin{aligned} \vec{a}_1 &= [0,0,1] & \vec{a}_2 &= [\sqrt{3},0,1]/2 & \vec{a}_3 &= [\sqrt{3},0,-1]/2 \\ \vec{b}_1 &= -[\sqrt{3},0,1]/2 & \vec{b}_2 &= -[0,0,1] & \vec{b}_3 &= [\sqrt{3},0,-1]/2. \end{aligned} \tag{2.19}$$

The correlations are obtained from the Born rule

$$p^{3322}(a,b|x,y) = \langle\psi^-|\frac{\mathbb{I} + (-1)^a \vec{a}_x \cdot \vec{\sigma}}{2} \otimes \frac{\mathbb{I} + (-1)^b \vec{b}_y \cdot \vec{\sigma}}{2}|\psi^-\rangle \tag{2.20}$$

which can be more conveniently be written as

$$p^{3322}(a,b|x,y) = \frac{1}{4}\left[1 - (-1)^{a+b}\vec{a}_x \cdot \vec{b}_y\right]. \tag{2.21}$$

Let us now consider the one-parameter family of correlations obtained from mixing the optimal distribution $p^{3322}(a,b|x,y)$ with white noise,

$$p^v(a,b|x,y) = vp^{3322}(a,b|x,y) + \frac{1-v}{4}. \tag{2.22}$$

Inserting this into the I_{3322} inequality, we find that it gives a violation whenever $v > 4/5$.

Is it true that whenever $v > 4/5$, the correlations are also useful for creating a quantum advantage in some CCP in our considered scenario? It turns out that it is sufficient to employ the standard additive communication strategy in which Alice sends $m = a \oplus x_0$ and Bob outputs $g = m \oplus b$. As in our example of the CCP based on the CGLMP inequality, we can decide the critical v for a classical simulation of the full distribution $p(g|x_0,x,y)$ in the CCP by means of linear programming. Evaluating the relevant linear program, one finds that the critical value indeed is $v = 4/5$. All distributions of the form (2.22) that violate the I_{3322} inequality are useful for quantum communication complexity.

In spite of this, we now present a candidate for nonlocal correlations that cannot be used to enhance any CCP in the scenario with any communication strategy. This distribution was first presented in Ref [40] as an example of a distribution that can violate the I_{3322} inequality using a state that can never be used to violate the CHSH inequality. The distribution is defined as follows. Alice and Bob share the noisy state

$$\rho = \frac{17}{20}|\phi\rangle\langle\phi| + \frac{3}{20}|0,1\rangle\langle 0,1| \tag{2.23}$$

where $|\phi\rangle = (2|0,0\rangle + |1,1\rangle)/\sqrt{5}$. Let us choose Alice's and Bob's Bloch vectors in the xz-plane as $\vec{a}_x = [\sin\theta_x, \cos\theta_x]$ and $\vec{b}_y = [\sin\phi_y, \cos\phi_y]$ with

$$\begin{aligned} \theta_1 &= \eta & \theta_2 &= -\eta & \theta_3 &= -\frac{\pi}{2} \\ \phi_1 &= -\chi & \phi_2 &= \chi & \phi_3 &= \frac{\pi}{2}, \end{aligned}$$

and $\eta = \arccos\left(\sqrt{7/8}\right)$ and $\chi = \arccos\left(\sqrt{2/3}\right)$. This defines the candidate probability distribution

$$p^{\text{cand}}(a,b|x,y) = \frac{1}{4}\,\text{tr}\left[(\mathbb{I} + \vec{a}_x\cdot\vec{\sigma}) \otimes (\mathbb{I} + \vec{b}_y\cdot\vec{\sigma})\rho\right] \tag{2.24}$$

which achieves the small violation $I \approx 0.0129$.

Having fixed the nonlocal distribution of Alice and Bob, the set of quantum correlations in the CCP reads

$$p^{\text{Q}}(g|x,x_0,y) = \sum_\lambda p(\lambda)p^{\text{Q}}_\lambda(g|x,x_0,y) = \sum_\lambda p(\lambda)\sum_{a,b,m} p^{\text{cand}}(a,b|x,y)\delta_{E_\lambda(a,x_0,x),m}\delta_{D_\lambda(m,b,y),g}. \tag{2.25}$$

This is a polytope. The number of encoding/decoding strategies is 2^{24}. We have evaluated the list of all deterministic distributions $\{p^{\mathrm{Q}}_{\lambda}(g|x,x_0,y)\}_{\lambda=1}^{2^{24}}$. Fortunately, we find that many pairs of encoding/decoding functions give rise to the same distribution. Removing the duplicates, we are left with 8192992 unique distributions which is roughly half the original number. We must show that all these distributions can be simulated in a classical model. This is a rather painful task, since it involves the evaluation of just over eight million linear programs[4]. Having evaluated all linear programs, we found that all the distributions admit a classical simulation. Then, it immediately follows that any convex combination of them also can be classically simulated and therefore that there exists no CCP in the considered scenario for which p^{cand} offers a quantum advantage.

In conclusion, Bell nonlocality does not imply advantages in communication complexity using the standard scenario. If there nevertheless exists a way of harvesting an advantage from p^{cand}, it must occur within a more complicated scenario. This in itself would be interesting since it is likely to require a construction that significantly departs from the established ones. However, it is far from clear whether this is possible or not. In view of our negative evidence with regard to whether Bell nonlocality implies quantum advantages in CCPs, it is and outstanding open problem to settle the matter in full generality. One route to doing that is to attempt to generalise our argument based on p^{cand} to general scenarios.

2.2 Quantum communication versus entanglement-assisted classical communication

Hitherto, our approach to quantum communication complexity has been based quantum nonlocality. However, as we have seen in section 1.3, quantum advantages in CCPs can arise from entirely different quantum resources. We consider the setting in which parties share no entanglement and communicate quantum systems. How does such quantum communication compare to entanglement-assisted classical communication as a resource for CCPs? Are the resources equivalent? If not, when is one better than the other? How do their differences manifest themselves?

Quantum communication models

Let us begin by describing quantum communication models for CCPs. We focus on bipartite scenarios in which Alice receives the input X, Bob receives the input Y and the communication is from Alice to Bob. In a quantum communication model, we substitute the classical message for a

[4]The computation was distributed on several desktop computers, a two workstations and a high-performance cluster. It was completed in three weeks.

quantum message. Importantly, in order to not violate the quantitative communication constraint of the CCP, the dimension d of the quantum message must be no greater than the alphabet size of the classical message. While it is perhaps intuitive that a quantum d-level system cannot carry more information than a classical d-level system, this substitution is formally justified by the Holevo theorem [23] which shows that the accessible information in both the quantum and classical systems is no more than $\log d$ bits. Hence, Alice associates her input to a quantum state ρ_X living in d-dimensional Hilbert space. This state is relayed to Bob who's decoding procedure now corresponds to a quantum measurement. Given his input Y, he selects a POVM $\{M_{g|Y}\}_g$ and applies it to the incoming state. The resulting probability distribution is

$$p(g|X,Y) = \mathrm{tr}\left(\rho_X M_{g|Y}\right). \tag{2.26}$$

Bob's outcome g is his guess in the CCP. Notice that classical models also can be obtained from the quantum models under the additional constraint that all states of Alice are diagonal in the same basis.

From entanglement-assisted models to quantum communication models

It is many times possible to transform an entanglement-assisted model into a quantum communication model in such a way that success rate in the CCP stays the same [24]. This transformation is based on the map from Bell inequalities to CCPs introduced in the previous section where we found that additive message strategies in the classical communication allow a one-to-one link between (many) Bell inequalities and CCPs. The success rate in such entanglement-assisted CCP can be reproduced in a quantum communication model provided that two additional conditions are satisfied. These constraints are

1. The quantum state used to obtain the maximal violation of the relevant Bell inequality is of the same local dimension on Bob's side as the alphabet size of the classical communication.

2. The marginal distribution of Alice in the Bell experiment is uniform.

To see why these two conditions are needed, notice that when Alice applies the measurement operator $A_{a|x}$ in the Bell experiment, she remotely prepares Bob in the state

$$\rho_{a|x} = \frac{\mathrm{tr_A}\left(A_{a|x} \otimes \mathbb{I}\rho_{\mathrm{AB}}\right)}{\mathrm{tr}\left(A_{a|x} \otimes \mathbb{I}\rho_{\mathrm{AB}}\right)}. \tag{2.27}$$

Since we assume that Alice's marginals are uniform, we can instead write

$$\rho_{a|x} = d\,\mathrm{tr_A}\left(A_{a|x} \otimes \mathbb{I}\rho_{\mathrm{AB}}\right). \tag{2.28}$$

In the quantum communication model, due to the fact that Alice's outcome is uniformly random, we can treat it as the "scrambler" input appearing in the CCP[5]. Therefore, we can associate the set of states remotely prepared by Alice to the set of states explicitly communicated by Alice in a quantum communication model. However, for this communication to be compatible with the dimensional limitation, we must require that the optimal state for violating the Bell inequality, ρ_{AB}, lives on $\mathbb{C}^D \otimes \mathbb{C}^d$ for some D. Then, it must be that also $\rho_{a|x}$ is d-dimensional. Hence, if Alice has access to quantum communication, she can explicitly prepare this state and relay it to Bob. Because in both the quantum communication model and the entanglement-assisted model, Bob receives the same ensemble, his optimal action is the same in both cases. Therefore, the quantum communication model can reproduce the entanglement-assisted CCP score.

Note that a concrete example of applying this transformation of entanglement-assisted models to quantum communication models is implicitly presented in our example of the random access code in section 1.3. Recall that the random access code was obtained as a mapping from the CHSH inequality to a CCP. The CHSH inequality is compatible with both our additional assumptions; it reaches its maximal violation with two-qubit entanglement and the corresponding correlations have uniform marginals. Indeed, the states Alice sent in the quantum communication variant of the random access code were precisely those that she remotely prepared for Bob in the entanglement-assisted variant. With our present knowledge, it was therefore unsurprising that we found the same success rate in both cases.

What happens when we start off with a correlation Bell inequality that does not respect the two assumptions and map it to its corresponding CCP? Then matters are less clear. On the one hand, entanglement-assisted strategies are not limited in the dimension of the entangled state. Therefore, Alice could prepare remote states on Bob's side that are of a dimension higher than d. On the other hand, quantum communication allows Alice to prepare arbitrary states that are not constrained by the no-signaling principle. It is therefore interesting to know that there exists examples of such Bell inequalities in which entanglement-assisted classical communication outperforms quantum communication [35]. The simplest example (known to the author) is the so-called four-bit random access code. It is a generalisation of the random access code introduced in section 1.3. In this variant, Alice holds a four-bit input $X = x_1 \dots x_4 \in \{0,1\}^4$ and Bob holds $Y \in \{1,2,3,4\}$. Alice communicates no more than one bit to Bob, who attempts to maximise the task function

$$\mathcal{S}_{\mathrm{rac}} = \frac{1}{64} \sum_{X,Y} p(g = x_y | X, Y). \tag{2.29}$$

[5]Recall that in the considered construction, only one part of the input of each party is used as a measurement setting in the CCP. The other part is the scrambler.

An optimal entanglement-assisted strategy can achieve $\mathcal{S}_{\text{rac}} = 3/4$ [42] whereas a quantum communication strategy is broadly believed to be limited by $\mathcal{S}_{\text{rac}} \approx 74.1\%$ [43]. Indeed, the Bell inequality that Alice and Bob can use [36] to perform the entanglement-assisted protocol is optimally implemented with entangled states of local dimension larger than two.

Quantum communcation models can outperform entanglement-assisted models

Let us now turn to our main issue, namely that of determining whether there exists scenarios in which quantum communication models can outperform entanglement-assisted models. We answer this in the positive. More surprisingly, we find that there even exists scenarios in which the relation is activated once one steps over a critical dimensional threshold [31, 26].

Let us consider CCPs tailored to the d-outcome CGLMP Bell inequalities [38] - as obtained by applying our previously presented map from Bell inequalities to CCPs. The choice to focus on these Bell inequalities stems from the fact that these are the only known (at least to the author) family of d-outcome Bell inequalities that are also facets of the local polytope. In the CCP, Alice receives an input labelled by the pair $x \in [2]_0$ and $x_0 \in [d]_0$. Bob receives an input $y \in [2]_0$. Alice may communicate no more than a d-valued message to Bob. Having received the message, Bob constructs his guess $g \in [d]_0$. Depending on his guess, he either earns a certain amount of points or loses a certain amount of points. Specifically, consider the functions

$$f_k = x_0 - xy - (-1)^{x+y}k \mod d \qquad h_k = x_0 - xy + (-1)^{x+y}(k+1) \mod d, \tag{2.30}$$

for $k = 0, \ldots, \lfloor d/2 \rfloor - 1$. If Bob guesses f_k he will earn c_k points while if he instead guesses h_k he will lose c_k points. In all other cases, no points are won or lost. We then choose

$$c_k = 1 - \frac{2k}{d-1}. \tag{2.31}$$

The average score becomes

$$\Delta_d = \frac{1}{4d} \sum_{\substack{x_0,x \\ y,k}} c_k \left[p(g = f_k | x_0, x, y) - p(g = h_k | x_0, x, y) \right] \tag{2.32}$$

Clearly, the best Bob could possibly hope for is to always guess $g = f_1$ which would earn him 1 point in every round and therefore $\Delta_d = 1$. Following our previous discussion of entanglement-assisted models, Alice and Bob can share the state that enables the maximal violation of the CGLMP inequalities and use their inputs x and y to obtain outcomes a and b which have the strong correlations associated to the maximal quantum correlations. Then, Alice sends the message $m = x_0 + a \mod d$ to Bob who guesses $g = m - b \mod d$. The entanglement-assisted strategy is

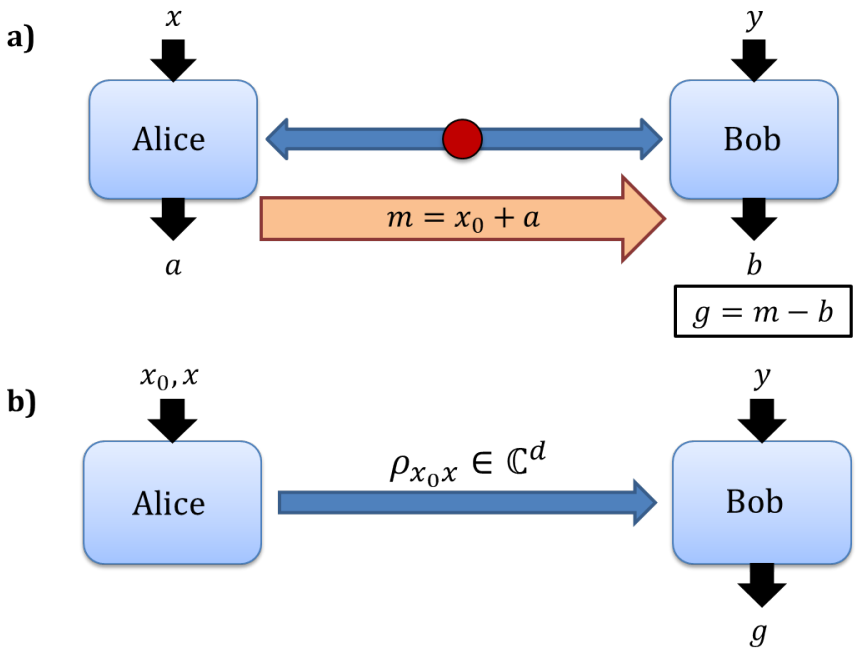

Figure 2.2: Quantum implementations of CCPs. a) Entanglement-assisted strategy for CCP based on the CGLMP inequality. b) Quantum communication strategy for CCP based on the CGLMP inequality.

illustrated in Figure 2.2. Due to our previous discussion, this gives a one-to-one relation between Δ_d and the violation of the CGLMP inequality. The maximal violation of the CGLMP inequalities does not have a known analytical form as a function of d, but it is known up to large values of d [44].

In a quantum communication model, we write Alice's d-dimensional quantum states as $\rho_{x_0 x} \in \mathbb{C}^d$ and Bob's measurements as $\{M_{g|y}\}_g$ (see Figure 2.2). We then have that

$$\Delta_d = \frac{1}{4d} \sum_{x_0,x,y,k} c_k \operatorname{tr}\left(\rho_{x_0 x}\left(M_{f_k|y} - M_{h_k|y}\right)\right). \tag{2.33}$$

How does the optimal quantum communication score compare to the optimal entanglement-assisted score? To answer this question, we have conducted extensive numerics which is summarised in Table 2.1. We have used semidefinite programs (SDPs) [45] in see-saw to optimise the quantum communication model. This gives a lower bound on the best performance. We have evaluated a lower bound under the additional constraint of Bob using rank-one projective measurements. In addition, we present the known maximal violations of the CGLMP inequality up to $d = 10$. Moreover, since the link between the CGLMP inequality and the score in the entanglement-assisted CCP is not restricted to quantum violations of the CGLMP inequality, we have also considered the

d	Lower bound QC	Optimal EACC	Optimal ML	Lower bound QC rank-one projective
2	0.7071	0.7071	0.7071	0.7071
3	0.7287	0.7287	0.7887	0.7287
4	0.7432	0.7432	0.8032	0.7432
5	0.7539	0.7539	0.8249	0.7539
6	0.8000	0.7624	0.8345	0.7624
7	0.8175	0.7694	0.8461	0.7814
8	0.8571	0.7753	0.8529	0.8006
9	0.8622	0.7804	0.8605	0.8188
10	0.8889	0.7849	0.8657	0.8396

Table 2.1: Lower bounds on the score in the CCP based on the CGLMP inequality using quantum communication (QC). Optimal scores via entanglement-assisted classical communication (EACC) using the maximal quantum violation of the CGLMP inequality as well as its relaxation to macroscopic locality (ML).

CCP score attainable in the post-quantum model based on Bell nonlocality that only is required to respect the principle of macroscopic locality [46]. The results reveal a peculiar pattern. The lower bound for quantum communication coincides accurately (up to several more decimal places than displayed) with the maximal violation of the CGLMP inequality for $d = 2, 3, 4, 5$. Since the CGLMP inequality is a correlation Bell inequality that is maximally violated with two entangled d-dimensional systems and the maximal violation has uniform marginals, we know due to our previous discussion that the quantum communication must be at least as good as the maximal violation of the CGLMP inequality. Indeed, we find that our numerically obtained quantum communication ensembles coincide with those remotely prepared in the corresponding Bell experiment. Thus, our results indicate that, in fact, that no better quantum communication strategy is possible for these low dimensions. However, surprisingly, from $d = 6$ and beyond, this pattern is broken and quantum communication outperforms the entanglement-assisted strategy. Dimension six appears to act as a threshold for a qualitatively different behaviour. It is interesting to note that the lower bound for quantum communication when Bob uses rank-one projective measurements for $d = 6$ still coincides with the entanglement-assisted bound. This indicates that starting from $d = 6$, Bob should use degenerate quantum measurements. Moreover, starting from dimension $d = 8$, quantum communication obtains an advantage large enough to even outperform the Bell nonlocality based strategy limited only by macroscopic locality.

These results motivate the following question: does dimension six really act as a threshold, or is quantum communication already advantageous for $d = 2, 3, 4, 5$ but we simply failed to detect it

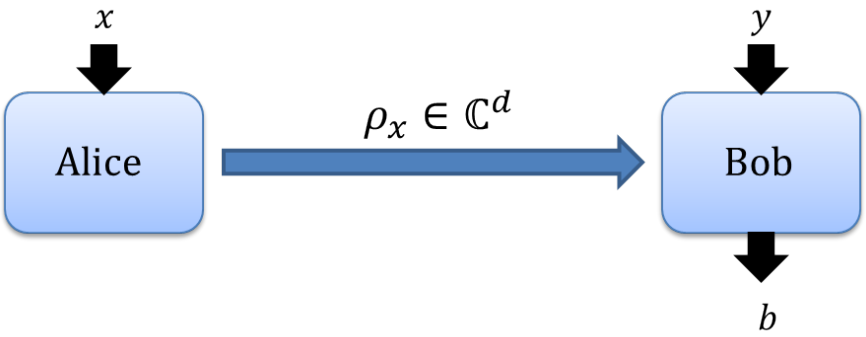

Figure 2.3: Alice and Bob receive inputs x and y of a given cardinality. Alice communicates a d-dimensional quantum state to Bob who applies a POVM to obtain an outcome b.

with our lower bound? We resolve this matter by evaluating symmetrised semidefinite relaxations of the set of quantum correlations [47, 27] (which is the subject of the next section). Such methods give us upper bounds on the best possible value of Δ_d achievable with quantum communication. In this manner, we can confirm that our lower bounds for $d = 2, 3, 4, 5$ in fact are optimal, and hence that the threshold at $d = 6$ is a genuine feature. In Ref [26] the quantum communication advantages over both the strategy using quantum nonlocality and the strategy using macroscopic locality were experimentally demonstrated.

This leaves a conspicuous open question: how can we understand and predict the emergence of a dimensional discontinuity in quantum communication complexity?

2.3 Bounding finite-dimensional quantum correlations

The set of classical correlations for a given communication scenario can be characterised by a polytope. How can we characterise the set of quantum correlations that arises from communication of quantum d-level systems (see Figure 2.3)? We will consider this question when Alice and Bob also have access to shared randomness. On the one hand, this is an admittedly fair comparison to classical models. On the other hand, shared randomness introduces convexity in the set of quantum correlations. This facilitates the analysis and allows us to employ tools for convex optimisation.

In general, it is a difficult task to evaluate the optimal score in a CCP under quantum communication of dimension d. Analytical solutions are rare. Therefore, it is relevant to develop general methods for establishing upper bounds on any given linear task function in a quantum communication model. In general, we can write the task as follows. If Alice's input is denoted x, Bob's input is denoted y and his output is denoted b, an arbitrary linear functional in a quantum model

is written

$$\mathcal{S} = \sum_{x,y,b} c_{xyb} \operatorname{tr}\left(\rho_x M_{b|y}\right), \tag{2.34}$$

where c_{xyb} can be arbitrary real coefficients, ρ_x are positive semi-definite trace-one operators of size d and $\{M_{b|y}\}$ are POVMs of size d. The maximal quantum correlations correspond to evaluating

$$\mathcal{S}^{\mathrm{Q}} = \max_{\{\rho_x\}} \max_{\{M_y\}} \mathcal{S}. \tag{2.35}$$

The problem of deriving upper bounds on $\mathcal{S}^{\mathrm{Q}}$ was tackled in Ref [47]. There, a hierarchy of SDPs was the developed for establishing a series of improving bounds on $\mathcal{S}^{\mathrm{Q}}$. Let us outline this Navascués-Vértesi (NV) hierarchy in a simple and brief manner. The reader is referred to Refs [47, 48] for a detailed description.

The Navascués-Vértesi hierarchy

To use the NV hierarchy, we must first choose the number of inputs and outputs present in our problem, the dimension of Alice's quantum communication as well as the specific objective (2.34) that we wish to evaluate in a quantum model. Then, we can make a list $X = \{\mathbb{I}, \{\rho_x\}, \{M_{b|y}\}\}$ of all operators that appear in the problem, namely the identity, all Alice's states and all Bob's POVM elements. We are now free to choose the *relaxation degree*. A higher relaxation degree means a better bound on $\mathcal{S}^{\mathrm{Q}}$ but also a more demanding computation. The relaxation degree corresponds to a list of monomials $\mathcal{O}$ which contains products of the operators that appear in X. Relaxation degree k means that all products of at most k operators from X must appear in $\mathcal{O}$. For instance, choosing $k = 1$ we just have $\mathcal{O} = X$. Choosing $k = 2$ we have $\mathcal{O} = \{\mathbb{I}, \{\rho\}, \{M\}, \{\rho\}\{\rho\}, \{\rho\}\{M\}, \{M\}\{M\}\}$ where by $\{A\}\{B\}$ we mean all products between all elements in set A and set B. Having chosen the relaxation degree, we evaluate the NV hierarchy to obtain an upper bound on the best possible quantum implementation, i.e. $\mathcal{S}_k \geq \mathcal{S}^{\mathrm{Q}}$. Below, we outline a step-by-step procedure for its implementation.

1. Sample a set of random pure states to be communicated by Alice; $\{|\psi_x\rangle\}_x \in \mathbb{C}^d$.

2. Sample a set of random projective measurements to be applied by Bob; $\{M_{b|y}\}_{b,y}$. The composition of the ranks in each measurement needs to be decided in advance.

3. The sampled states and measurements allow us to compute the corresponding operator list X and subsequently also the list of monomials $\mathcal{O}$. From the monomials, we can evaluate the sampled *moment matrix* which is defined as

$$\Gamma_{ij} = \operatorname{tr}\left(\mathcal{O}_i \mathcal{O}_j^\dagger\right). \tag{2.36}$$

4. Store the sampled moment matrix. Repeat the above three steps, each time storing the sampled moment matrix. Terminate this loop when the sampled moment matrix is linearly dependent on the previously sampled moment matrices. Then, throw away that final sample: left is a basis for the moment matrices that we write as $\{\Gamma^{(1)}, \Gamma^{(2)}, \ldots, \Gamma^{(m)}\}$.

5. Construct an affine combination of the moment matrix basis;

$$\Gamma = \sum_{i=1}^{m} c_i \Gamma^{(i)}, \tag{2.37}$$

for some arbitrary real coefficients c_i that are only required to satisfy $\sum_i c_i = 1$. These coefficients serve as our SDP variables.

6. As long as $k \geq 1$, the moment matrix contains elements that correspond to the quantum probabilities appearing in the objective (2.34). This means that we can write the objective as a linear combination $\mathcal{L}$ of the moment matrix elements,

$$\mathcal{S} = \mathcal{L}[\Gamma]. \tag{2.38}$$

We obtain an upper bound on $\mathcal{S}^{\mathrm{Q}}$ (for the given rank chosen for the measurement operators) by evaluating the SDP

$$\mathcal{S}_k \equiv \max_{\{c_i\}} \mathcal{L}[\Gamma], \qquad \Gamma \geq 0, \qquad \sum_i c_i = 1. \tag{2.39}$$

One should repeat this process for all rank combinations of the POVMs of Bob and choose the best result obtained. By again repeating this procedure for increasing relaxation degrees, one finds a hierarchy of bounds which obey

$$\mathcal{S}_1 \geq \mathcal{S}_2 \geq \ldots \geq \mathcal{S}^{\mathrm{Q}}. \tag{2.40}$$

While it is not known whether in the limit of large k one always recovers $\mathcal{S}_\infty = \mathcal{S}^{\mathrm{Q}}$, it is sometimes the case that already a relatively low relaxation degree is sufficient to obtain a tight bound on the quantum correlations. Also, for practical purposes, that is often what one must hope for since the computational requirements of evaluating this procedure grow rapidly with the number of inputs/outputs, the dimension of the communication and the relaxation degree. In practice, this restricts us to considering relatively small-sized problems. Since many interesting problems go beyond the few simplest ones, it is relevant to consider the development of tools for more efficiently bounding quantum correlations.

Symmetrised semidefinite relaxations

There are two tracks to reduce the computational requirements of the NV hierarchy. The first is to reduce the number of samples needed to form the moment matrix basis. Such a reduction means both less time and memory spent in the sampling stage and fewer variables in the final SDP. The second is to reduce the size of the positivity constraint in the final moment matrix Γ when evaluating the SDP. This means that instead of imposing the positivity of an $|\mathcal{O}| \times |\mathcal{O}|$ matrix, we can break it up into smaller blocks and impose the positivity of each of them. The key to achieving both these reductions is to exploit *symmetries* present in the problem [27].

In essence, symmetries are permutations of the indices of the monomials that leave the objective function invariant. Consider a permutation π of the indices of the operator list X. The permutation acts as $\pi(X_i) = X_{\pi(i)}$. In the monomial list, whose elements are products of the elements in the operator list, we define the action of the permutation as $\pi(X_i X_j \ldots) = X_{\pi(i)} X_{\pi(j)} \ldots$. If the operators are subject to additional constraints, we call the set of permutations that respect these constraints the *ambient group*. Ambient transformations preserve the structure of the scenario, but make no reference to the actual objective that we are considering. For example, a permutation belonging to the ambient group cannot permute a state with a measurement since these are subject to different constraints. Thus, if we apply an ambient permutation to the moment matrix, we preserve the structure of the scenario but necessarily the structure of the specific objective. Having established the ambient group, we now look for a subgroup that we call the *symmetry group* $\mathcal{G}$. This group contains all ambient permutations that also have the property that of preserving the objective function when written as a linear combination of the moment matrix elements, i.e. π is a symmetry if it is both ambient and respects

$$\mathcal{L}(\Gamma) = \mathcal{L}(\pi(\Gamma)). \tag{2.41}$$

To the author's knowledge, there is no simple and efficient method to construct the symmetry group for general problems. There are essentially two alternative approaches: either construct the ambient group and check all the elements and discard those that do not keep the objective invariant or carefully inspect the objective to spot simple symmetries[6]. Importantly, if there are subtle symmetries that manage go undetected, it is not a major issue. In a nutshell, the more symmetries found in a problem, the larger the computational advantages. Sometimes, finding only small symmetries in a problem already brings a large reduction of computational requirements.

With the symmetry group in hand, one computes the so-called Reynold's operator. It amounts

[6] Often a direct inspection is both simpler and more efficient.

to taking each sample of a moment matrix and mapping it into

$$\Gamma' = \frac{1}{|\mathcal{G}|} \sum_{\pi \in \mathcal{G}} \pi(\Gamma), \tag{2.42}$$

and then adding the symmetrised moment matrix Γ' to the moment matrix basis. Note that there are ways of rapidly evaluating this sum so that one does not need to consider all the group elements, see Ref [27]. Also, notice that averaging over the Reynold's operator is a standard procedure for symmetrisation techniques, see e.g. [49]. This procedure effectively confines the sampled moment matrices to the symmetric subspace of the space of feasible moment matrices. Often, this is a dramatically smaller space than that originally considered in the NV hierarchy. Therefore, we expect to require much fewer samples in order to construct a moment matrix basis.

A proper symmetrisation should combine the basis reduction with a reduction of the positivity constraint in the final SDP. This can be achieved due to the fact that the symmetrised moment matrix entering the SDP commutes with a representation of the symmetry group. Therefore, there must exist a unitary transformation that block-diagonalises the moment matrix. By implementing such a block-diagonalisation, we break up the original positivity constraint into the positivity of many small matrices. This is typically a major reduction of computational complexity. A complete block-diagonalisation requires one to find the irreducible components (with multiplicities) for the symmetry group. This is achievable (see [27]), but here it is omitted in favour of a simpler procedure that achieves a partial block-diagonalisation (it does not identify multiplicities)[7]. Pick a random set of real numbers $\{c_i\}_i$ satisfying $\sum_i c_i = 1$. The number of elements in the sum should equal the size of the symmetrised moment matrix basis. Evaluate the operator $O = \sum_i c_i \Gamma'^{,(i)}$ (where $\Gamma'^{,(i)}$) are the symmetrised samples. Do a spectral decomposition $O = UDU^\dagger$ where D is diagonal and U is unitary. The unitary U can be applied for block-diagonalisation. This is admittedly a somewhat dirty numerical trick, but it works rather well.

In view of the above, we can reduce the number of variables in the SDP, the size of the positivity constraint and the time and memory spent in sampling. We note that Ref [27] presents a matlab package for implementing these techniques. The conspicuous question now is how much reduction all this symmetrisation actually gives us and whether it is useful in practical problems. Let us also remark that whereas here we discussed symmetry techniques for the set of quantum correlations obtained from dimension-bounded quantum communication, these techniques in fact apply to a broad variety of SDPs (which need not concern physics at all). For this, we refer the reader to Ref [49].

[7] The author finds that this is typically a decent first attempt that often is sufficient to evaluate the SDP. However, if it proves insufficient, it is advisable to consider a full block-diagonalisation.

	# Basis elements		SDP (+ bkldiag) time (sec)	
d	standard	sym	standard	sym
2	28	6	2	2
3	545	13	15	2
4	3250	13	1900	2
5	12917	13	$> 5.5 \times 10^4$	3
6	-	13	-	3
10	-	13	-	15

Table 2.2: The number of basis elements the SDP evaluation time for a standard implementation of the NV hierarchy versus a symmetrised implementation of the random access code. The symbol "−" indicates that the computational requirements were too demanding to complete the procedure.

Exemplifying the power of symmetrised semidefinite relaxations

We give an example serving to illustrate the power of the symmetrisation of the NV hierarchy. For a more complete set of examples and applications, see Ref [27].

We first choose a problem in which the symmetries are easy to spot, namely the d-dimensional random access code problem [50]. This is a natural generalisation of the already discussed standard random access code. In this problem, Alice receives one of d^2 possible inputs represented by two d-valued integers $x_0, x_1 \in [d]_0$ and Bob receives a binary input $y \in [2]_0$. Alice communicates at most a d-dimensional system to Bob who attempts to output $b = x_y$. In a quantum model, the average success probability is therefore

$$\mathcal{S}_d = \frac{1}{2d^2} \sum_{x_0,x_1,y} \mathrm{tr}\left(\rho_{x_0 x_1} M_{x_y|y}\right). \tag{2.43}$$

The random access code is one of the rare cases in which the quantum maximum is known analytically. It was first conjectured in [50] and then proven in [51]:

$$\mathcal{S}_d^{\mathrm{Q}} = \frac{1}{2}\left(1 + \frac{1}{\sqrt{d}}\right). \tag{2.44}$$

Naturally, we do not need the NV hierarchy to bound S_d^{Q} since we already know it analyically. Nevertheless, since the random access code is simple and has plenty of symmetries, it serves as a good platform for illustrating the computational advantages of symmetrisation - which then readily extend well beyond this example.

We can easily spot several symmetries in the random access code. Firstly, we may permute x_0 with $\pi : [d]_0 \to [d]_0$ and simulatenously permute $b \to \pi(b)$ when $y = 0$. An inspection of the

objective shows that it remains invariant under this operation. In fact, the analogous symmetry can also be applied when $y = 1$. We can also swap x_0 and x_1 and simultaneously flip the bit-value of y. Each of these actions permutes states into states and measurements into measurements and also keep the objective function invariant. For several different values of d, we have evaluated both the standard NV hierarchy and the symmetrised variant on a standard desktop computer. The results are displayed in Table 2.2 for the intermediate hierarchy level known as $1 + AB$[8]. We see that the number of basis elements rapidly increases in the standard implementation as we increase d. At $d = 6$, the computer ran out of memory before finishing the sampling procedure. In contrast, the symmetries of the random access code allow us to complete the sampling part very quickly, using only 13 samples regardless of the considered value of d. Already at $d = 5$, the advantage in the sampling reduction is a factor of a thousand. Using symmetrisation, the evaluation of the SDP is completed in a matter of seconds. Again, for $d = 5$ the SDP evaluation for the standard case was not finalised as it was terminated after roughly 15 hours.

2.4 Informationally restricted correlations

Quantum communication complexity investigates the relationship between information and correlations. Hitherto, we have interpreted information as a limitation on the alphabet size of the quantum or classical communication. However, there are reasons to doubt that such dimensional limitations properly capture the concept of information as there are at least two conceptual inconveniences. Firstly, imagine a very high-dimensional ensemble of classical or quantum states. Due to its large dimension, it could carry a large amount of information. Consider now that each element of the ensemble is mixed with a large degree of white noise. The dimension of the communication remains unchanged, but it is evident that the information carried by the ensemble should decrease with the amount of white noise, and in the limit approach zero. The dimension alone does not necessarily capture the information carried in a classical or quantum ensemble. Indeed, there must exist noisy ensembles of dimension $d' > d$ that carry less than $\log d$ bits of information. Secondly, since dimensions are discrete, so is the associated information. However, it stands to reason that information ought to be represented as a continuous quantity in the spirit of Shannon theory. In view of these observations, a reasonable interpretation is that a dimensional limitation only is a sufficient, but not necessary, condition for a classical or quantum ensemble carrying at most $\log d$ bits. A more general concept that dissolves the raised concerns should not reference the Hilbert space dimension but nevertheless recover the set of dimensionally dimensionally restricted ensembles as a special

[8]This level corresponds to level $k = 1$ as well as all products of the form $\{\rho\}\{M\}$

case of ensembles of at most $\log d$ bits.

The correlation experiments we have considered, featuring classical or quantum communication, are of a single-shot nature. That is, Alice sends a single system to Bob who measures it and reports an outcome. A natural way of quantifying the information carried in Alice's communication is through the *min-entropy*. This is a conservative way of quantifying information. It rests on the following hypothetical game played between Alice and Bob. Alice's ensemble is denoted $\mathcal{E} = \{p_x, \rho_x\}_{x=1}^n$ where p_x is the probability distribution of her classical input represented by the random variable X. Alice prepares the state ρ_x and sends it to Bob. Bob applies a single measurement $\{N_z\}_{z=1}^n$ with the aim of guessing Alice's classical input. Thus, he is successful if and only if $b = x$. The average success probability of Bob when he uses the best possible measurement is called the *guessing probability*,

$$P_g(X|\mathcal{E}) = \max_{\{N_z\}} \sum_{x=1}^{n} p_x \operatorname{tr}(\rho_x N_x). \tag{2.45}$$

The better Bob could (in principle, via some POVM) guess Alice's input, the larger is the information content of Alice's ensemble. Importantly, this approach allows us to ask the key question: *information about what?* Here, we have chosen it to be information about Alice's classical input. The guessing probability can be used to evaluate the conditional min-entropy,

$$H_{\min}(X|\mathcal{E}) = -\log(P_g), \tag{2.46}$$

which can be interpreted as the maximal uncertainty of Bob about X once he has received Alice's communication. The information is then defined as the difference in uncertainty about X when Bob receives $\mathcal{E}$ and when he does not,

$$\mathcal{I}_X(\mathcal{E}) = H_{\min}(X) - H_{\min}(X|\mathcal{E}), \tag{2.47}$$

where the min-entropy $H_{\min}(X) = -\log \max_x p_x$ is the uncertaintly of Bob when no communication is received (naturally, Bob would just guess on the most likely value of X). We will use $\mathcal{I}_X$ as a our measure of the information content of a classical or quantum ensemble. Note that the classical case simply corresponds to choosing all ρ_x diagonal in the same basis. We remark that it is in general difficult to evaluate the value of $\mathcal{I}_X$ by hand. However, such evaluation can always be achieved by means of an SDP over Bob's extraction measurement.

Let us now define informationally restricted correlations. This takes into account the possibility of Alice and Bob sharing classical randomness. Therefore, in a general prepare-and-measure scenario (see Figure 2.3) the quantum correlations read

$$p(b|x,y) = \sum_{\lambda} p(\lambda) \operatorname{tr}\left(\rho_x^{(\lambda)} M_{b|y}^{(\lambda)}\right). \tag{2.48}$$

We are interested in the correlations that can arise when Alice's communication has bounded information, i.e. $\mathcal{I}_X \leq \alpha$, for some α to be chosen. In the presence of shared randomness, the information is interpreted as follows. For each strategy λ, Alice prepares the ensemble $\mathcal{E}_\lambda$ and Bob may apply an extraction measurement depending on λ. Hence, the guessing probability is the average of the guessing probabilities for the individual λ,

$$P_g(X|\mathcal{E}) = \sum_\lambda p(\lambda) P_g(X|\mathcal{E}_\lambda). \tag{2.49}$$

This averaged guessing probability is then used to compute the information. In summary, shared randomness applies on the level of the guessing probability, not directly on the level of the information.

Classical correlations

In a classical model, Alice sends integer messages, m, to Bob. The messages can be of any dimension d. The correlations therefore read

$$p(b|x,y) = \sum_\lambda p(\lambda) \sum_{m=1}^{d} p(m|x,\lambda) p(b|m,y,\lambda). \tag{2.50}$$

In analogy with what we have already seen, for a given d the set of classical correlations is a polytope. If we fix d, we can enumerate all encoding and decoding strategies of Alice and Bob. For each encoding strategy of Alice, we can compute its guessing probability and denote it $P_g^{(\lambda)}$. The constraint $\mathcal{I}_X \leq \alpha$ can then be stated in terms of the guessing probability:

$$\sum_\lambda p(\lambda) P_g^{(\lambda)} \leq 2^{\alpha - H_{\min}(X)}. \tag{2.51}$$

When phrased like this, it is clear that the information constraint for classical models is linear. Therefore, for a given d, the set of informationally restricted classical correlations is also a polytope; obtained from suitably cutting the unconstrained polytope with hyperplanes. However, we would like to assume nothing about d. How can we eliminate this variable? An intuitive solution is that in a classical model, Alice can never benefit from using a message that has a larger alphabet than the alphabet size of her input. Otherwose, her encoding function would simply ignore some elements in the image. It stands to reason that any correlation created with a larger message alphabet should therefore be reproducible with a message of at most dimension $d = n$. In Ref [28] this simple intuition is proven. Therefore, the full set of informationally restricted classical correlations is identical to the polytope obtained the message dimension equals the size of Alice's input. The faces of this polytope are linear inequalities that constitute tight tests of informationally restricted classical correlations.

The simplest scenario

Let us denote the scenario by (n, l, k), where n is the number of inputs for Alice, l is the number of inputs for Bob and k is his number of outputs. We seek the smallest scenario in which quantum correlations can outperform classical constraints. We focus on the case of a uniform input distribution $p_x = 1/n$. For instance, the scenarios $(2, 1, 2)$ and $(2, 1, 3)$ were found not to yield a non-trivial facet of the polytope of classical correlations. Therefore, we consider two measurements for Bob. Then, the simplest scenario is $(2, 2, 2)$. In this case, solving the classical polytope one finds that its faces are either trivial (positivity of probabilities) or correspond directly to the information constraint, which by assumption cannot be violated in a quantum model. Notably, the analogous is encountered for the $(2, 2, 3)$ scenario.

The first scenario in which we find a quantum advantage is $(3, 2, 2)$. Solving the classical polytope, one finds the facet

$$F_1 \equiv -E_{11} - E_{12} - E_{21} + E_{22} + E_{31} \leq 2^{\alpha+1} - 1 \tag{2.52}$$

where $E_{xy} = p(0|x, y) - p(1|x, y)$ and $\mathcal{I}_X \leq \alpha \in [0, \log 3]$. An interesting observation is that if we choose $\alpha = 1$ (i.e. one bit of information), then the inequality becomes identical to that of the simplest classical dimension witness for bits [52]. Since the set of quantum ensembles restricted by one bit of information is strictly larger than the set of qubit ensembles, it means that the inequality is in fact valid for the more general ensembles given by our definition of information.

Let us see how the quantum violation emerges. To this end, we consider an explicit quantum strategy (which probably is not optimal). Alice and Bob share a bit of randomness $\lambda \in \{0, 1\}$ with distribution $q = p(\lambda = 0)$. The value of λ corresponds to two different strategies. When $\lambda = 0$, Alice prepares three pure qubit states

$$|\psi_1\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right), \qquad |\psi_2\rangle = |0\rangle \qquad |\psi_3\rangle = \sin\frac{\pi}{8}|0\rangle - \cos\frac{\pi}{8}|1\rangle, \tag{2.53}$$

while Bob measures $-\frac{\sigma_x+\sigma_z}{\sqrt{2}}$ and $\frac{\sigma_z-\sigma_x}{\sqrt{2}}$. A simple calculation gives $F_1 = 1+2\sqrt{2}$. In contrast, when $\lambda = 1$ Alice sends white noise to Bob who always outputs $b = 1$. This leads to $F_1 = 1$. Averaging out the shared randomness, we have $F_1 = 1+2\sqrt{2}q$. A simple calculation gives the information content $\mathcal{I}_X = \log(1 + q)$. This strategy is valid for up to a bit of information but can be straightforwardly extended to $\mathcal{I} \in [1, \log 3]$ in a similar manner by mixing between the $\lambda = 0$ strategy and the strategy in which Alice sends her input to Bob (which costs $\log 3$ bits of information). The results are displayed in Figure 2.4. We see that the quantum strategy outperforms the classical bound at all times except at the end points. The latter is expected since the end points correspond to the trivial cases of sending no information and and sending x respectively.

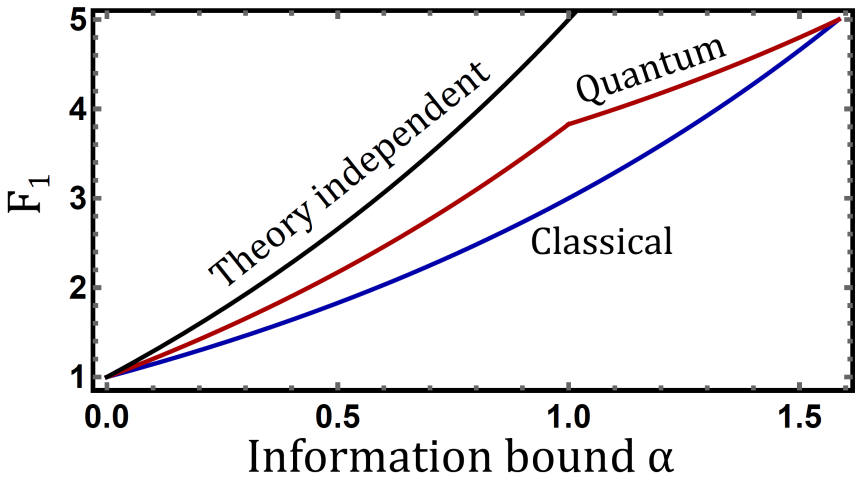

Figure 2.4: Correlations F_1 versus information content in classical, quantum and theory-independent models.

In addition, Figure 2.4 displayes a theory-independent bound on the correlations. This bound is valid for every physical theory used to describe the communication scenario. In order to derive it, we allow for arbitrary $p(b|x,y)$ subject to the constraint that there exists no post-processing of the distribution that allows one to extract more than the allowed amount of information. This amounts to imposing

$$\forall y: \quad \sum_{x,b} p_x p(b|x,y) p(b'=x|y,b) \leq 2^{\alpha - H_{\min}(X)}. \tag{2.54}$$

where $b' \in [n]$ is the post-processed guess of Bob for the value of x. Since the most general post-processing $p(b'|y,b)$ can be written as a convex combination of deterministic post-processings, it is sufficient to impose the above relation for every deterministic post-processing. Each of them is a linear constraint. Since there are only finitely many deterministic post-processings, we can evaluate the theory-independent bound on F_1 as a linear program over $p(b|x,y)$. It is emphasised that the theory-independent bounds obtained in this manner are not necessarily optimal.

Information versus dimension

Informationally restricted quantum correlations can outperform informationally restricted classical correlations. A natural next question is how informationally restricted quantum correlations relate to standard quantum correlations obtained from the communication of d-dimensional systems? Naturally, the comparison is only meaningful when $\mathcal{I}_X = \log d$ bits due to the discrete nature of

Hilbert space dimensions. It turns out that simple and general relations can be established between the two sets of quantum correlations.

Firstly, let us give a simple proof of the fact that every quantum ensemble of d-dimensional states, i.e. $\mathcal{E} = \{p_x, \rho_x\}$ where ρ_x is of size d, can carry no more than $\log d$ bits of information. To this end, we establish a bound on the guessing probability valid for all such ensembles:

$$P_g = \max_{\{N_z\}} \sum_x p_x \operatorname{tr}(\rho_x N_x) \leq \max_{\{N_z\}} \sum_x p_x \lambda_{\max}(N_x) \leq \max_{\{N_z\}} \sum_x p_x \operatorname{tr}(N_x) \leq d \max_x p_x, \tag{2.55}$$

where we have used that the best ρ_x is pure and aligned with the eigenvector of N_x with largest eigenvalue, and that for positive operators it holds that $\lambda_{\max}(A) \leq \operatorname{tr}(A)$. Inserting this bound in the definition of information, we obtain

$$\mathcal{I}_X = -\log\left(\max_x p_x\right) + \log(P_g) \leq \log d, \tag{2.56}$$

which is the desired relation. Hence, we conclude that the set of quantum ensembles of information at most $\log d$ bits is a strict superset of the set of quantum ensembles of dimension at most d. From that, it follows that the set of quantum correlations $p(b|x,y)$ in the former case contains that of the latter. However, a key question is whether $\log d$ bits of information invested in d-dimensional quantum systems is always just as good for creating correlations as investing the same amount of information in a higher-dimensional quantum ensemble? We now prove that the answer is negative through an explicit example. Recall the previous discussion of the four-bit random access code. This is a communication task in which Alice receives four input bits $x = x_1x_2x_3x_4 \in [2]^4$ and Bob receives $y \in [4]$. By sending no more than one bit of information to Bob, they aim to achieve $b = x_y$. The average success probability is

$$\mathcal{S}_{\text{rac}} = \frac{1}{64} \sum_{x,y} p(b = x_y|x,y). \tag{2.57}$$

It is known that qubit communication must satisfy $\mathcal{S}_{\text{rac}} < 3/4$ [43, 53]. A more precise bound that is supported by much numerics is $\mathcal{S}_{\text{rac}} \leq 74.1\%$. We show that if Alice communicates an ensemble of four-dimensional quantum systems carrying no more than one bit of information, she can outperform the qubit bound. To this end, we let Alice prepare states that each are uniform mixtures of two orthogonal states

$$\rho_x = \frac{1}{8}\Big(2\mathbb{I}\otimes\mathbb{I} - (-1)^{x_4}\mathbb{I}\otimes\sigma_y - (-1)^{x_1}\sigma_x\otimes\sigma_x - (-1)^{x_2}\sigma_y\otimes\sigma_x - (-1)^{x_3}\sigma_z\otimes\sigma_x\Big). \tag{2.58}$$

Then, we let Bob measure the observables $B_1 = \sigma_x\otimes\sigma_x$, $B_2 = \sigma_y\otimes\sigma_x$, $B_3 = \sigma_z\otimes\sigma_x$ and $B_4 = \mathbb{I}\otimes\sigma_y$. A simple calculation gives $\mathcal{S}_{\text{rac}} = 3/4$. This value is greater than that achievable with qubits. In

order to finish the proof, we need only to show that the ensemble carries no more than one bit of information. To show this, one could use an SDP to directly evaluate the guessing probability of the ensemble. An alternative method is to exploit the useful bound on the information derived in Ref [28]:

$$\mathcal{I}_X \leq \log(d) + \log\left(\frac{\max_x p_x \lambda_{\max}(\rho_x)}{\max_x p_x}\right), \tag{2.59}$$

which is valid for general ensembles. Since we have $p_x = 1/16$ and the spectra of all 16 states is $(1/2, 1/2, 0, 0)$, it immediately follows that $\mathcal{I}_X \leq 1$. We conclude that informationally restricted quantum correlations is a strict superset of the quantum correlations obtainable from d-dimensional quantum systems.

Quantum communication versus entanglement-assisted classical communication

Finally, informationally restricted quantum correlations motivate a reconsideration of the previous discussion of entanglement-assisted classical communication versus quantum communication as resources for creating correlations. Previously, we found that no simple relation exists between the two when quantum communication corresponds to a dimensional limitation. Interestingly, the relation simplifies signficantly if we instead involve informationally restricted correlations. Specifically, we show that all correlations possible by means of classical communication of a d-valued (potentially stochastic) message assisted by any amount o shared entanglement can be reproduced by means of the communication of a quantum system that carries no more than $\log d$ bits of information.

In order to prove this claim, let Alice and Bob share the entangled state ρ_{AB}. This state may be of any dimension. Alice's measurement of her subsystem leaves Bob in the state

$$\sigma_{a|x} = \frac{\mathrm{tr}_{\mathrm{A}}\left(A_{a|x} \otimes \mathbb{I}\rho_{\mathrm{AB}}\right)}{p(a|x)} \tag{2.60}$$

where $p(a|x) = \mathrm{tr}\left(A_{a|x} \otimes \mathbb{I}\rho_{\mathrm{AB}}\right)$ is the probability of outcome a when applying measurement x. Alice also sends a classical message to Bob, which we for simplicity write as a d-dimensional state $\mu_{a|x}$. Since the messages are classical, all states $\{\mu_{a|x}\}$ are diagonal in the same basis. Importantly, Alice can use both her input and her outcome to construct the message. Thus, Alice has supplied Bob with the net state $\mu_{a|x} \otimes \sigma_{a|x}$. Bob measures this state in order to create the correlations $p(b|x, y)$. Here, the cost of the communication lies only in the classical message.

Now, we show that there exists a quantum communication model in which the correlations $p(b|x, y)$ can be reproduced while the information cost never exceeds $\log d$ bits. In this model, Alice samples randomly from $p(a|x)$ and locally creates the state $\mu_{a|x} \otimes \sigma_{a|x}$ and sends it to Bob.

Evidently, the dimension of this state is high but the information content turns out to be much more restricted. The net state (averaged over a) seen by Bob is

$$\tau_x = \sum_a p(a|x)\mu_{a|x} \otimes \sigma_{a|x}. \tag{2.61}$$

The ensemble of Alice therefore corresponds to $\mathcal{E} = \{p_x, \tau_x\}$. Since Bob is supplied with the same ensemble as in the entanglement-assisted case, the correlations he can create are the same. The less trivial question is whether Alice's ensemble respects the information constraint. To this end, we consider the guessing probability

$$P_g^{\mathrm{QC}} = \max_{\{N_z\}} \sum_{a,x} p_x p(a|x) \operatorname{tr}(\mu_{a|x} \otimes \sigma_{a|x} N_x). \tag{2.62}$$

We can place an upper bound on the guessing probability by using that $\operatorname{tr}(\mu_{a|x} \otimes \sigma_{a|x} N_x) \leq \operatorname{tr}(\sigma_{a|x} N_x^{\mathrm{B}})$, where N_x^{B} is the partial trace of N_x over the first system. This leads to the following upper bound on the guessing probability

$$P_g^{\mathrm{QC}} \leq \max_{\{N_z\}} \sum_x p_x \operatorname{tr}\left(\sum_a p(a|x)\sigma_{a|x} N_x^{\mathrm{B}}\right). \tag{2.63}$$

However, the no-signaling nature of the remotely prepared ensemble implies that $\sum_a p(a|x)\sigma_{a|x} = \rho^{\mathrm{B}}$, where ρ^{B} is Bob's share of the entangled state. Consequently,

$$P_g^{\mathrm{QC}} \leq \max_{\{N_z\}} \sum_x p_x \operatorname{tr}\left(\rho^{\mathrm{B}} N_x^{\mathrm{B}}\right) \leq \left(\max_x p_x\right) \max_{\{N_z\}} \operatorname{tr}\left(\rho^{\mathrm{B}} \sum_x N_x^{\mathrm{B}}\right). \tag{2.64}$$

Since the POVM $\{N_z\}$ acts on the tensor product of a d-dimensional Hilbert space and an arbitrary-dimensional Hilbert space, it must hold that

$$\sum_x N_x^{\mathrm{B}} = \sum_x \operatorname{tr}_1(N_x) = \operatorname{tr}_1(\mathbb{I}_d \otimes \mathbb{I}) = d\mathbb{I}. \tag{2.65}$$

Thus, it follows that

$$P_g^{\mathrm{QC}} \leq d \max_x p_x \tag{2.66}$$

and that the information obeys the bound

$$\mathcal{I}_X = -\log\left(\max_x p_x\right) + \log(P_g) \leq \log d. \tag{2.67}$$

This concludes the proof.

Informationally restricted quantum correlations give a new take on the relation between quantum correlations and communication. An outstanding open problem is to develop methods for bounding these correlations in general quantum models. Another interesting question is to develop their applications in quantum information processing.

3

Certification of quantum devices

Quantum correlations that violate classical constraints can be used for physical inference. This opens up a path to certification of quantum devices based on the correlations that they produce in experiments on which only to weak assumptions are imposed. Here, we present a framework for certification and characterisation of many different types of quantum devices when the only assumption considered is a limitation on their Hilbert space dimension. This assumption is both reasonably weak, as it does not require precise control of any part of an experiment, and it is compatible with experiments that can be implemented with the current state-of-the-art. In the first section, we focus on prepare-and-measure experiments and consider the certification of the states and measurements appearing in the BB84 protocol for quantum key distribution [53]. In the second section, we present certification methods for generalised qubit measurements that correspond to POVMs that are non-projective [54]. The third section departs from prepare-and-measure scenarios and instead focuses on a three-party sequential experiment in which one can certify quantum instruments [55]. In the final section, we consider more sophisticated experiments that involve both quantum communication and entangled states. For these scenarios, we construct a scheme for certifying and characterising entangled states and measurements of arbitrary many subsystems and arbitrary dimension [56].

3.1 Certification of the BB84 states and measurements

In the previous chapter, we have seen many examples of quantum correlations established in prepare-and-measure experiments featuring d-dimensional systems. Let us change our perspective on them. Instead of focusing on how these correlations outperform classical constraints, let us instead consider what these correlations tell us about the states and measurements that give rise

to them. This entails that we *assume* the validity of quantum theory, but do not assume that we perfectly control the experiment that generates the correlations. We consider prepare-and-measure experiments (as in Figure 2.3) in which both Alice's and Bob's devices may perform general operations that are only constrained by their dimension. In practice, this roughly corresponds to experiments in which the degrees of freedom are known but their precise control is not assumed. If we observe quantum correlations $p(b|x,y)$, is it possible to deduce which ensemble of states $\{\rho_x\}$ and which set of measurements $\{M_{b|y}\}$ that was implemented in the experiment?

To answer this question, let us again focus on the simple case of the random access code. Recall that Alice receives two bits $x \equiv x_0, x_1 \in \{0,1\}$ and Bob receives a single bit $y \in \{0,1\}$. Alice is assumed to send qubit states to Bob who performs qubit measurements to construct his output $b \in \{0,1\}$. The score in the random access code in a quantum model reads

$$\mathcal{S}_{\text{rac}} = \frac{1}{8}\sum_{x,y} \text{tr}\left(\rho_x M_{x_y|y}\right). \tag{3.1}$$

Any value of $\mathcal{S}_{\text{rac}} > 3/4$ is a proof of quantum correlations. For simplicity, let us first consider the extremal case of the maximal value $\mathcal{S}_{\text{rac}} = 1/2\left(1+1/\sqrt{2}\right)$. What does this tell us about Alice's states and Bob's measurements?

Let us begin by considering Alice's states. To this end, let us re-write the random access code as follows

$$\mathcal{S}_{\text{rac}} = \frac{1}{2} + \frac{1}{8}\sum_{y} \text{tr}\left(M_{0|y}V_y\right) \leq \frac{1}{2} + \frac{1}{8}\sum_{y}\sqrt{\text{tr}\left(M_{0|y}V_y^2\right)\text{tr}\left(M_{0|y}\right)}, \tag{3.2}$$

where, we have used the fact that $M_{0|y} + M_{1|y} = \mathbb{I}$ to eliminate $M_{1|y}$ and then defined the effective preparation operator $V_y = \sum_{x_0,x_1}(-1)^{x_y}\rho_{x_0x_1}$. In order to obtain the right-hand-side, we have used the fact that $|\text{tr}\,(OR)|^2 \leq \text{tr}\,(OR^2)\,\text{tr}\,(O)$ for a positive semidefinite O and a Hermitian operator R. Let us now exploit the following useful fact: all binary-outcome measurements can be simulated by stochastically implementing projective binary-outcome measurements and post-processing their outcomes. Therefore, it is sufficient to restrict Bob's measurements to being projective. Since they are qubits, this means that they must be rank-one projective (rank-two projective measurements are simply the identity operator). Hence, we must have that $\text{tr}\left(M_{b|y}^0\right) = 1$. Now, we must evaluate the operator V_y^2. To this end, it is favourable to employ the Bloch sphere formalism. We write Alice's states as

$$\rho_{x_0x_1} = \frac{\mathbb{I} + \vec{m}_{x_0x_1}\cdot\vec{\sigma}}{2}, \tag{3.3}$$

for some Bloch vectors $\{\vec{m}_{x_0x_1}\}$. Hence, V_y^2 can also be written in terms of the Bloch vectors: finding the precise form is straightforward but somewhat tedious. The calculation gives

$$V_y^2 = \frac{1}{2}\left(\beta + (-1)^y\alpha\right)\mathbb{I} \tag{3.4}$$

where

$$\alpha = (\vec{m}_{00} - \vec{m}_{11}) \cdot (\vec{m}_{01} - \vec{m}_{10}), \tag{3.5}$$

$$\beta = \frac{1}{2} \sum_{x_0,x_1} |\vec{m}_{x_0x_1}|^2 - \vec{m}_{00} \cdot \vec{m}_{11} - \vec{m}_{01} \cdot \vec{m}_{10}. \tag{3.6}$$

A pivotal property is that V_y^2 is proportional to the identity. Therefore, we can use the rank-one projective property of Bob's measurements to evaluate $\mathrm{tr}\left(M_{0|y}V_y^2\right)$. Putting this together, we obtain the bound

$$\mathcal{S}_{\mathrm{rac}} \leq \frac{1}{2} + \frac{1}{8\sqrt{2}} \left[\sqrt{\beta + \alpha} + \sqrt{\beta - \alpha}\right]. \tag{3.7}$$

First, we determine the largest value of the right-hand-side. To do this, write $f(r,s) = \sqrt{r+s} + \sqrt{r-s}$ for $r \geq s$. The largest value of f is easily seen from writing

$$f(r,s) = \sqrt{f(r,s)^2} = \sqrt{2r + 2\sqrt{r^2 - s^2}}. \tag{3.8}$$

Evidently, the maximum is $2\sqrt{r}$ and is found at $s = 0$. Therefore, by choosing $\alpha = 0$ and β maximal, we obtain a bound on the quantum random access code. The algebraically maximal value is $\beta = 4$. Hence, we have

$$\mathcal{S} \leq \frac{1}{2}\left(1 + \frac{1}{\sqrt{2}}\right). \tag{3.9}$$

This is a proof of the already stated maximal value of $\mathcal{S}_{\mathrm{rac}}$. In order to have β maximal, we require that i) all states are pure (meaning that the Bloch vectors are of unit length) and that ii) Alice's states are pairwise antipodal on the Bloch sphere, i.e. $\vec{m}_{00} \cdot \vec{m}_{11} = -1$ and $\vec{m}_{01} \cdot \vec{m}_{10} = -1$. Now we need only to determine their relative angle. This is made clear from the fact that we need to have $\alpha = 0$. We have

$$\alpha = (\vec{m}_{00} - \vec{m}_{11}) \cdot (\vec{m}_{01} - \vec{m}_{10}) = 4\vec{m}_{00} \cdot \vec{m}_{01} \stackrel{!}{=} 0. \tag{3.10}$$

Thus, the two Bloch sphere diagonals must be unbiased. Hence, we have deduced that the only states compatible with a maximal score in the quantum random access code must necessarily form a square in some disk of the Bloch sphere (see Figure 1.4).

Let us shift focus to the measurements. In order to derive the implications of a maximal $\mathcal{S}_{\mathrm{rac}}$ on the measurements, we write the quantum random access code as follows.

$$\mathcal{S}_{\mathrm{rac}} = \frac{1}{2} + \frac{1}{16} \sum_{x_0,x_1} \mathrm{tr}\left(\rho_{x_0x_1}[(-1)^{x_0}M_0 + (-1)^{x_1}M_1]\right) \leq \frac{1}{2} + \frac{1}{16} \sum_{x_0,x_1} \lambda_{\max}\left[(-1)^{x_0}M_0 + (-1)^{x_1}M_1\right], \tag{3.11}$$

where we have written $M_y = M_{0|y} - M_{1|y}$ for the observable of Bob. To obtain the right-hand-side, we have used that the best state of Alice is that aligned with eigenvector corresponding to the largest eigenvalue of the operator $(-1)^{x_0} M_0 + (-1)^{x_1} M_1$. For qubit observables, we can precisely evaluate the relevant eigenvalues directly in terms of the observables. The right-hand-side becomes

$$\mathcal{S}_{\text{rac}} \leq \frac{1}{2} + \frac{1}{16}\left(\sqrt{2\mu + 2\nu - \eta_+^2} + \sqrt{2\mu - 2\nu - \eta_-^2}\right), \tag{3.12}$$

where $\mu = \text{tr}\,(M_0^2 + M_1^2)$, $\nu = \text{tr}\{M_0, M_1\}$ and $\eta_\pm = \text{tr}(M_0 \pm M_1)$. A simple calculation shows that the optimal value of the right-hand-side corresponds to $\mu = 4$ and $\nu = \eta_\pm = 0$. This means that the observables must be anticommuting and correspond to projective measurements. For qubits, this is equivalent to choosing the observables as σ_x and σ_z up to a global rotation on the Bloch sphere.

Robust certification

The above shows a precise certification of the states and measurements only when the maximal correlations in the quantum random access code are observed. However, no realistic implementation of the quantum random access code perfectly achieves the maximal quantum correlations. What can we say about the states and measurements when we observe sub-optimal quantum correlations?

Let us first focus on the states. As soon as the correlations are sub-optimal, more than a single ensemble (up to a global unitary) becomes compatible with the observed correlations. Therefore, it is important to specify precisely what it is that we aim to certify. One natural answer is to consider the distance between the ensemble that would have been certified had the correlations been optimal and the most distant ensemble still compatible with the observed correlations. In other words, how distant (in terms of distance to the square-like ensemble) is the most distant ensemble that could explain the quantum correlations? We employ the *fidelity* as our measure of distance[1] The fidelity between two quantum states is defined as

$$F(\rho, \sigma) = \text{tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}. \tag{3.13}$$

Then, denoting the ideal (square-like) ensemble by $\{\rho_x^{\text{ideal}}\}$, we write

$$S(\{\rho_x\}) = \max_{\Lambda} \frac{1}{4} \sum_{x_0 x_1} F(\rho_x^{\text{ideal}}, \Lambda[\rho_x]), \tag{3.14}$$

for the largest fidelity between the ensemble $\{\rho_x\}$ and the ideal ensemble obtainable by any extraction map Λ (formally a CPTP map). The extraction map plays an important role: it allows

[1] We note that fidelity is not a proper distance measure.

one to perform a global operation on the ensemble in an attempt to bring it closer to the ideal ensemble. This is non-trivial, as the following example illustrates. If our ensemble is $\{|0\rangle, |1\rangle\}$ and our ideal ensemble is $\{|1\rangle, |0\rangle\}$, we would like to enable the global bit-flip operation that makes these ensembles identical instead of maximally distant. This is made possible by choosing Λ as a bit-flip unitary.

Since the ideal states are pure, the fidelity simplifies to $F(\rho_x^{\text{ideal}}, \Lambda[\rho_x]) = \text{tr}\left(\rho_x^{\text{ideal}}\Lambda[\rho_x]\right)$. Using this, we can now write the fidelity of the most distant ensemble compatible with the quantum correlations as

$$\mathcal{F}\left(\mathcal{S}_{\text{rac}}\right) = \min_{\{\rho_x\}\in R(\mathcal{S}_{\text{rac}})} S\left[\{\rho_{x_0x_1}\}\right], \tag{3.15}$$

where $R(\mathcal{S}_{\text{rac}})$ denotes all ensembles compatible with the value $\mathcal{S}_{\text{rac}}$, i.e. all ensembles for which there exists measurements that lead to a $p(b|x,y)$ whose value in the random access code is $\mathcal{S}_{\text{rac}}$.

The aim is to place a lower bound on $\mathcal{F}(\mathcal{S})$ which serves as a robust certification. This is achieved with a technique based on operator inequalities inspired by Ref [57]. First, write the quantum random access code as

$$\mathcal{S}_{\text{rac}} = \frac{1}{2} + \sum_{x_0,x_1} \text{tr}\left(W_{x_0x_1}\rho_{x_0x_1}\right) \tag{3.16}$$

where we have exploited the observables of Bob to define the effective measurement operator $W_{x_0x_1} = \frac{1}{16}\sum_y (-1)^{x_y} M_y$. Then, we seek to find operator inequalities on the form

$$K_{x_0x_1}(M_0, M_1) \geq sW_{x_0x_1} + t_{x_0x_1}(M_0, M_1)\mathbb{I}, \tag{3.17}$$

where the operators $K_{x_0x_1}$ are defined as the action of the (dual) extraction map on the ideal states, i.e. $K_{x_0x_1}(M_0, M_1) = \Lambda^\dagger(M_0, M_1)[\rho_{x_0x_1}^{\text{ideal}}]$. Importantly, these operator inequalities must hold for all measurements. To make this happen, we must suitably choose the coefficients s and $t_{x_0x_1}$. Let us momentarily assume that we have constructed such inequalities. It is straightforward to see how they help towards achieving the robust certification. We have

$$S \geq \frac{1}{4}\sum_{x_0,x_1} \text{tr}\left(K_{x_0x_1}\rho_{x_0x_1}\right) \geq \frac{s}{4}\sum_{x_0,x_1} \text{tr}\left(W_{x_0x_1}\rho_{x_0x_1}\right) + \frac{1}{4}\sum_{x_0,x_1} t_{x_0x_1} = \frac{s}{4}(\mathcal{S}_{\text{rac}} - 1/2) + \frac{1}{4}\sum_{x_0,x_1} t_{x_0x_1}. \tag{3.18}$$

If we minimise the right-hand-side over the measurements, the inequality becomes valid for all preparations. Therefore, it would imply

$$\mathcal{F}(\mathcal{S}_{\text{rac}}) \geq \frac{s}{4}\left(\mathcal{S}_{\text{rac}} - 1/2\right) + t \equiv L\left(\mathcal{S}_{\text{rac}}\right), \tag{3.19}$$

where we have defined

$$t = \frac{1}{4}\min_{M_0,M_1} \sum_{x_0,x_1} t_{x_0x_1}. \tag{3.20}$$

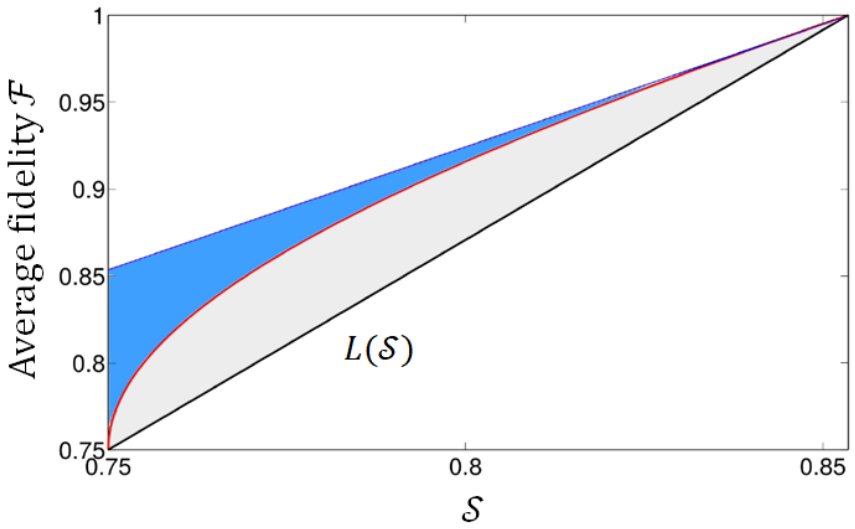

Figure 3.1: Robust certification of the fidelity of the state ensemble in the quantum random access code with the ideal ensemble. The black line is the lower bound $L(\mathcal{S}_{\text{rac}})$. The blue region, delimited by the red line, is reachable by single qubit strategies that do not exploit shared randomness. This region was obtained via numerics.

Thus, one only needs to find the operator inequalities (3.17) and perform the minimisation associated with t to derive the final result $L(\mathcal{S}_{\text{rac}})$. It remains only to construct the operator inequalities. This procedure is technical and it is detailed in Ref [53]. The main idea is pick a well-chosen CPTP map Λ and then fit the coefficients s and $t_{x_0x_1}$ so that the inequality is valid. How to choose Λ is a priori not clear. For the quantum random access code, it turns out that choosing Λ as a dephasing map (sometimes dephasing w.r.t. σ_x and sometimes w.r.t σ_z) is an optimal choice. The procedure eventually leads to the coefficients $s = 4\left(1+\sqrt{2}\right)$ and $t = \left(2-\sqrt{2}\right)/4$ which specifies the robust certification of Alice's states. In Figure 3.1 we illustrate the bound on the fidelity $\mathcal{F}$. We see that whenever the quantum random access code outperforms its classical counterpart, we can certify a non-trivial fidelity between Alice's ensemble and the ideal ensemble. This shows that our bound $L(\mathcal{S}_{\text{rac}})$ is optimal. Moreover, we have also illustrated the optimal $\mathcal{F}(\mathcal{S}_{\text{rac}})$ when Alice and Bob do not have access to shared randomness.

3.2 Certification of non-projective measurements

In the previous section, we have seen that quantum states and measurements can be certified in prepare-and-measure experiments in which only the dimension is known. The example of the BB84

states and measurements, certified via the quantum random access code, can be regarded as a proof-of-principle. However, the general technique readily extends to many other qubit state ensembles and sets of projective qubit measurements (see Ref [53] for examples). In this section, we go beyond these systems and instead focus on the certification of *generalised quantum measurements*. A generalised measurement corresponds to a POVM that is *non-projective*. Such non-projective measurements substantially enrich the standard (textbook) notion of a quantum measurement as a complete set of projectors. We develop methods for certifying any extremal non-projective qubit measurement and extend it also to higher dimension for interesting target POVMs. Moreover, we use both analytical and SDP techniques to investigate the robustness of the certification.

Extremal qubit non-projective measurements

The most general qubit POVM is a set of operators $\{E_i\}_{i=1}^{O}$ with the properties that $E_i \geq 0$ and $\sum_i E_i = \mathbb{I}$. Since all positive operators also are Hermitian, and all Hermitian qubit operators can be written as a linear combination of the Pauli matrices, we can w.l.g. write

$$E_i = \lambda_i \left(\mathbb{I} + \vec{n}_i \cdot \vec{\sigma}\right), \tag{3.21}$$

where $\vec{n}_i$ is the Bloch vector of the i'th measurement operator and $\lambda_i \geq 0$. In order to ensure that operators form a POVM, we require that

$$\sum_{i=1}^{O} \lambda_i = 1 \qquad \text{and} \qquad \sum_{i=1}^{O} \lambda_i \vec{n}_i = 0. \tag{3.22}$$

In principle, a qubit POVM can have any number of outcomes. However, every measurement with more than four outcomes can be simulated by stochastically implementing measurements that have no more than four outcomes [58]. When a measurement can be simulated in such a manner, it is said to be non-extremal. Non-extremal measurements cannot be certified in our prepare-and-measure experiments since one can never exclude the possibility of a stochastic simulation with other measurements solely from inspecting the quantum correlations. Instead, our interest is in *extremal* POVMs, i.e. the subset of POVMs with $O = 2, 3, 4$ outcomes that cannot be simulated with other POVMs. Among these POVMs, it is known that all extremal two-outcome POVMs are rank-one projective, i.e. standard measurements. The extremal non-projective qubit measurements are those that have $O = 3, 4$ outcomes [58]. They correspond to POVM elements whose Bloch vector is of unit length. Certifying such measurements is our goal.

From projective measurements to non-projective measurements

We present a general method for certifying non-projective qubit measurements. It is based on extending standard schemes focused on the certification of Alice's preparations. Let us denote the target non-projective measurement by $\mathcal{M}^{\text{target}}$ and the Bloch vector corresponding to outcome b is denoted $\vec{v}_b$.

Firstly, we must construct a standard correlation witness for certifying Alice's states. We choose a scenario in which Alice has O inputs and Bob receives an input of some cardinality $y \in [Y]$ and produces binary outputs $b \in \{0, 1\}$. The task is to construct a linear correlation witness $\mathcal{A}'$ for this scenario,

$$\mathcal{A}' = \sum_{x,y,b} c_{xyb} p(b|x, y). \tag{3.23}$$

The witness should have the property that its maximal value certifies that Alice's states $|\psi_x\rangle$ have the following relation: their Bloch vectors $\vec{u}_x$ form a mirror imagine of the Bloch vectors of the target POVM, i.e. $\vec{u}_x = -\vec{v}_x$. To find such a certificate, one needs only to employ projective measurements which can be achieved using techniques analogous to those discussed for the quantum random access code and the BB84 states. Notice that the witness $\mathcal{A}'$ that achieves the desired certification is not unique.

Secondly, with the standard certification achieved via the correlation witness $\mathcal{A}'$, we modify the inputs/outputs of Bob in order to accommodate the target measurement $\mathcal{M}^{\text{target}}$. To this end, we supply Bob with yet another measurement setting which we denote **povm**. It corresponds to a measurement that has $b \in [O]$ outcomes. The modified scenario is captured by the modified correlation witness

$$\mathcal{A} = \mathcal{A}' - k \sum_{x=1}^{O} p(b = x|x, \mathbf{povm}), \tag{3.24}$$

for some arbitrary constant $k > 0$. It is clear that the maximal value of $\mathcal{A}$ can be no greater than that of $\mathcal{A}'$. Furthermore, the only way in which $\mathcal{A}$ can precisely attain the maximal value of $\mathcal{A}'$ is if all terms $p(b = x|x, \mathbf{povm})$ vanish. In order for this to happen, we need the state $|\psi_x\rangle$ to be anti-aligned with the measurement operator corresponding to outcome $b = x$ for the setting **povm** and the Bloch vectors must be of unit length. This means that the Bloch vectors of the measurement must be identical to those of $\mathcal{M}^{\text{target}}$. Since the POVM is fully characterised by its Bloch vectors (the coefficients λ_i are fixed by normalisation and positivity), it follows that we have certified the target measurement $\mathcal{M}^{\text{target}}$. The procedure is illustrated in Figure 3.2.

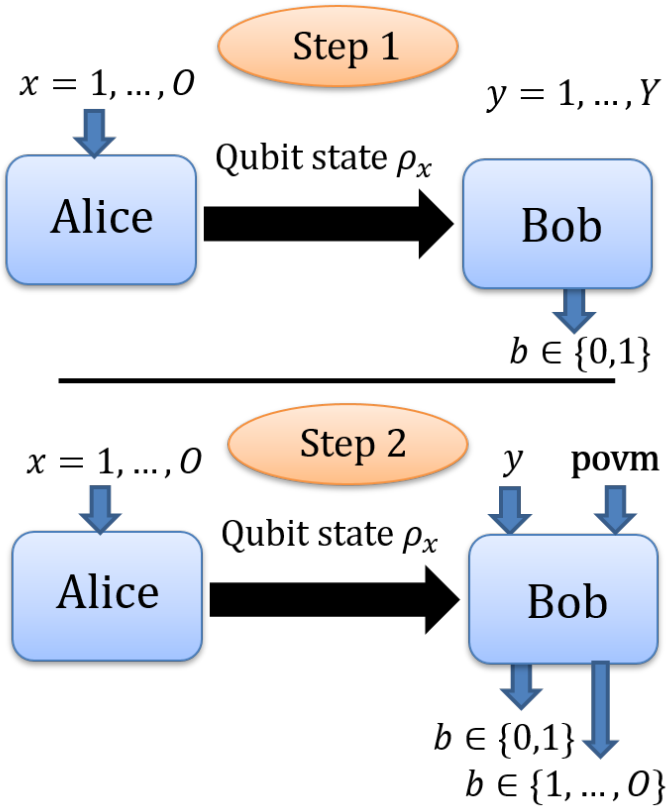

Figure 3.2: Method for certification of extremal non-projective measurements. In the first step, we find a standard prepare-and-measure scenario in which projective measurements are used to certify a specific relation on Alice's states. In the second step, we extend Bob's settings with one more input and exploit the property tailored in the first step to certify the additional setting as the target measurement.

Falsifying projective implementations

A maximal value of $\mathcal{A}$ certifies precisely the target non-projective measurement. However, if the observed correlations are sub-optimal, is it possible to simulate the measurement setting **povm** with stochastic projective measurements? Notice that the other, binary, settings of Bob always are optimally chosen as projective since such measurements are extremal. We show how to derive bounds on $\mathcal{A}$ valid for all projective measurements such that a violation therefore certifies the necessity of a non-projective measurement even for sub-optimal quantum correlations.

For all y, we label Bob's observable by M_y and denote the measurement corresponding to **povm** by $\{M^b_{\mathbf{povm}}\}$. If this measurement is a convex combination of projective measurements, the largest value of $\mathcal{A}$ occurs for a deterministic choice of $\{M^b_{\mathbf{povm}}\}$. Taken as a projective measurement, we may assign two of the O outcomes to rank-one projectors and the rest of the measurement

operators to the zero-operator. When $O = 3$ this can be done in three different ways. When $O = 4$, there are six choices. In general, all combinations of rank assignments must be considered. For the optimal implementation, we can associate the observable M_{Y+1} for Bob based on the two non-trivial outcomes of the **povm** setting. Now, it is possible to write the correlation witness on the form

$$\mathcal{A} = C(k) + \sum_x \mathrm{tr}\left[\rho_x \mathcal{L}_x^{(k)}(\{M_y\})\right], \tag{3.25}$$

where $C(k)$ is a constant, and $\mathcal{L}_x^{(k)}(\{M_y\})$ is a linear combination of the observables $\{M_1, \ldots, M_{Y+1}\}$. If we now apply the Cauchy-Schwarz inequality, we have that

$$\mathcal{A} \leq C(k) + \sum_x \sqrt{\mathrm{tr}\left[\rho_x \mathcal{L}_x^{(k)}(\{M_y\})^2\right]}. \tag{3.26}$$

Consider the effective operator $\mathcal{L}_x^{(k)}(\{M_y\})^2$. We can write each observable on the Bloch vector form $M_y = \vec{n}_y \cdot \vec{\sigma}$. Then, $\mathcal{L}^2$ is a linear combination of anti-commutators of the observables. However, for qubits we have that $\{M_k, M_l\} = 2\vec{n}_k \cdot \vec{n}_l \mathbb{I}$. This means that $\mathcal{L}^2$ is proportional to the identity, i.e. we can write

$$\mathcal{L}_x^{(k)}(\{M_y\})^2 = t_x^{(k)}(\{\vec{n}_y\})\,\mathbb{I}, \tag{3.27}$$

for some scalar function $t_x^{(k)}$ of the measurement Bloch vectors. Since $\mathrm{tr}\,\rho_x = 1$, we obtain the expression

$$\mathcal{A} \overset{\text{Proj}}{\leq} C(k) + \max_{\{\vec{n}_y\}} \sum_x \sqrt{t_x^{(k)}(\{\vec{n}_y\})} \equiv \mathcal{B}(k). \tag{3.28}$$

Specifically, we have eliminated the states from the evaluation of the bound on projective measurements. Hence, in order to compute the projective bound $\mathcal{B}(k)$, we need only to compute the maximisation of the right-hand-side over the observable Bloch vectors. Whereas this evaluation is hard in the general form presented here, it can be analytically evaluated for many specific witnesses of interest.

Case study: the qubit SIC-POVM

Let us apply the above to the case of the four-outcome qubit SIC-POVM (symmetric informationally complete [59]). This is an extremal POVM corresponding to $\lambda_b = 1/4$ and Bloch vectors that form a regular tetrahedron on the Bloch sphere. For instance, we can write the Bloch vectors as

$$\vec{v}_1 = \frac{1}{\sqrt{3}}(1,1,1) \qquad \vec{v}_2 = \frac{1}{\sqrt{3}}(1,-1,-1) \tag{3.29}$$

$$\vec{v}_3 = \frac{1}{\sqrt{3}}(-1,1,-1) \qquad \vec{v}_4 = \frac{1}{\sqrt{3}}(-1,-1,1)\,. \tag{3.30}$$

We begin by constructing a prepare-and-measure scenario and a correlation witness in which the states of Alice can be certified as pointing to the mirrored tetrahedron $\{-\vec{v}_x\}$. To this end, supply Alice with inputs $x \in [4]$ and Bob with three inputs $y \in [3]$ and binary outcomes b. We choose the following correlation witness

$$\mathcal{A}' = \frac{1}{12}\sum_{x,y} p(b = S_{x,y}|x, y), \tag{3.31}$$

where $S_{x,y} = 0$ if element number y in $\vec{v}_x$ is equal to one, and otherwise $S_{x,y} = 1$. Using techniques analogous to those employed in our investigation of the quantum random access code, one can prove the tight bound

$$\mathcal{A}' \leq \frac{1}{2}\left(1 + \frac{1}{\sqrt{3}}\right) \tag{3.32}$$

and that the maximum is saturated if and only if Alice's states are pure and form a tetrahedron on the Bloch sphere. Following the general recipe, we supply Bob with one more input **povm** which has four possible outcomes $b \in [4]$. The modified witness reads

$$\mathcal{A} = \frac{1}{12}\sum_{x,y} p(b = S_{x,y}|x, y) - k\sum_{x=1}^{4} p(b = x|x, \mathbf{povm}). \tag{3.33}$$

By observing $\mathcal{A} = 1/2(1 + 1/\sqrt{3})$ we therefore certify that the setting **povm** corresponds to a SIC-POVM.

Let us now employ $\mathcal{A}$ to derive a certificate for Bob's setting **povm** corresponding to a non-projective measurement. To this end, we follow the above procedure. The chosen witness simplifies the calculation of $\mathcal{B}(k)$ since it is symmetric with respect to Alice's inputs. That is, we may assign the projective operators to any two of the four measurement operators of Alice for the setting **povm**. Choosing the first two, we define $M_{\mathbf{povm}} \equiv M_4 = M_{\mathbf{1}|\mathbf{povm}} - M_{\mathbf{2}|\mathbf{povm}}$. A simple calculation then gives

$$\mathcal{L}^{(k)}_{x=0,1}(\{M_y\}) = \frac{1}{24}\left[1, (-1)^x, (-1)^x, (-1)^{x+1}12k\right] \cdot \vec{M} \tag{3.34}$$

$$\mathcal{L}^{(k)}_{x=2,3}(\{M_y\}) = \frac{1}{24}\left[-1, (-1)^x, (-1)^{x+1}, 0\right] \cdot \vec{M}, \tag{3.35}$$

where $\vec{M} = [M_1, M_2, M_3, M_4]$, with $M_y = \vec{n}_y \cdot \vec{\sigma}$. If we now use the fact that $\sqrt{z_1} + \sqrt{z_2} \leq \sqrt{2(z_1 + z_2)}$ for $z_1, z_2 \geq 0$, a fair share of simplification and optimal alignment choices for the Bloch vectors yields the bound

$$\mathcal{A} \leq \frac{1 - 2k}{2} + \frac{\sqrt{2}}{24}\sqrt{6 - 4q} + \frac{\sqrt{2}}{24}\sqrt{2r_k + 4q + 48k\sqrt{2}\sqrt{1 + q}} \equiv f_k(q), \tag{3.36}$$

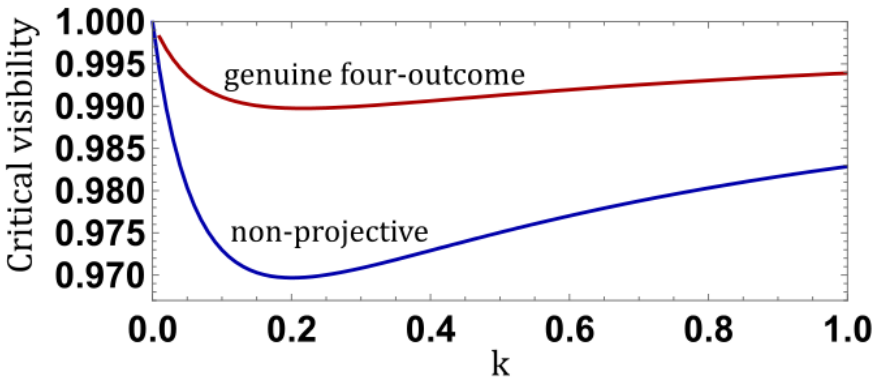

Figure 3.3: Critical visibility v for Alice's states in order to certify Bob's measurement as non-projective and genuine four-outcome respectively, as a function of k. The best choice of k is found at $k = 1/5$.

where $r_k = 3 + 144k^2$ and $q = \vec{n}_1 \cdot \vec{n}_2$. Hence, we find a bound on all projective measurements by evaluating

$$\mathcal{B}(k) = \max_{q\in[-1,1]} f_k(q). \tag{3.37}$$

This is an optimisation in a single variable and therefore straightforwardly solved. However, the final analytical expression is unwieldy. In this manner, we can choose any k and obtain a certificate of non-projectiveness for Bob's setting **povm**. It is interesting to note that this final bound appears to be tight for all values of k.

An even finer form of certification is relevant to the considered SIC-POVM example. A maximal value of $\mathcal{A}$ certifies a qubit SIC-POVM and correlations violating the bound $\mathcal{B}(k)$ certify a non-projective measurement which implies that the measurement must have had more than two outcomes. The natural question is, can we also derive a bound on $\mathcal{A}$ that is respected by all three-outcome measurements? Then, a violation would certify a genuine four-outcome measurement. To this end, the analytical method presented above no longer applies (we cannot work with observables). However, we can employ the symmetrised NV hierarchy discussed in the previous chapter to evaluate upper bounds on the maximal witness value attainable under ternary measurement. Recall, however, that the NV hierarchy only applies to projective measurements. In order to overcome this obstacle, one can embed the qubit states in a three-dimensional Hilbert space in which the ternary outcome non-projective qubit measurements can be represented as projective measurements on a three-level system. Hence, we obtain a bound valid for all ternary qubit measurements through an SDP.

Finally, the discussion is completed by a word about how to choose the coefficient k. The answer depends on circumstance. One simple example is the following. Consider that Alice's states are

noisy so that with probability v she prepares the optimal ensemble and with probability $1-v$ she communicates white noise. We can then calculate the critical v for certifying Bob's setting **povm** as a non-projective measurement and genuine four-outcome measurement respectively. However, this bound will depend on k. In Figure 3.3 we plot the critical v for both certifications as a function of k. We see that in general, a high value of v is needed to achieve the certification. Nevertheless, the best choice is found at $k = 1/5$.

Certification of d-dimensional SIC-POVMs

Let us now go beyond qubit systems and show how one can robustly certify higher-dimensional non-projective measurements. A prominent example of such measurements is the d-dimensional symmetric informationally complete POVM [59]. This is a measurement with d^2 outcomes whose POVM elements correspond to sub-normalised projectors that are equiangular lines in Hilbert space. In other words, the POVM elements read $E_i = \frac{1}{d}|\psi_i\rangle\langle\psi_i|$ where

$$|\langle\psi_i|\psi_j\rangle|^2 = \frac{1}{d+1}, \tag{3.38}$$

for $i \neq j$ where the constant on the right-hand-side is fixed by normalisation. SIC-POVMs are useful in many tasks in quantum information processing.

Let us show how they can be certified in prepare-and-measure experiments. The spirit of the certification is analogous to that presented for qubits. We begin with a prepare-and-measure scenario in which Bob only performs binary-outcome measurements. We prove that in this scenario, there exists a suitable witness whose maximal value certifies Alice's states to form a SIC. Then, we give an additional setting to Bob with d^2 possible outcomes and show that the modified scenario can be used to certify this setting as a SIC-POVM.

The scenario is as follows. Alice receives an input $x \in [N]$, where $N = d^2$, and Bob receives an input written as (y, y') that represents all ordered pairs of integers in the set $[N]$. Hence, Bob has $\binom{N}{2}$ settings. The outcome is denoted $b \in \{0, 1\}$ and the corresponding distribution becomes $p(b|x, (y, y'))$. Our correlation witness will only consider the events in which we either have $x = y$ or $x = y'$. Specifically, we define

$$\mathcal{A}'_d = \sum_{x<x'} \left[p(b = 0|x, (x, x')) + p(b = 1|x', (x, x'))\right]. \tag{3.39}$$

In order to evaluate its maximal value under d-dimensional systems and evaluate its implications

on Alice's states, we have

$$\mathcal{A}'_d = \max_{\{\rho\},\{M\}} \sum_{x<x'} \left[p(b=0|x,(x,x')) + p(b=1|x',(x,x'))\right] \tag{3.40}$$

$$= \max_{\{\rho\},\{M\}} \sum_{x<x'} \mathrm{tr}\left[(\rho_x - \rho_{x'})M_{0|(x,x')}\right] + \binom{N}{2} = \max_{\{\rho_x\}} \sum_{x<x'} \lambda_+ \left[\rho_x - \rho_{x'}\right] + \binom{N}{2}. \tag{3.41}$$

We have used that $M_{0|(x,x')} + M_{1|(x,x')} = \mathbb{I}$ to eliminate $M_{1|(x,x')}$ and then optimally chosen $M_{0|(x,x')}$ to be the projector onto the positive eigenspace of $\rho_x - \rho_{x'}$. By λ_+ we denote the sum of positive eigenvalues. From linearity, we know that the optimal preparations must be pure. Therefore, we can write $\rho_x \sim |\psi_x\rangle$. Since $\rho_x - \rho_{x'}$ therefore becomes a rank-2 operator with only one positive eigenvalue, we can replace λ_+ with the largest eigenvalue $\lambda_{\max}$. This gives

$$\mathcal{A}'_d = \max_{\{\psi_x\}} \sum_{x<x'} \lambda_{\max}\left[|\psi_x\rangle\langle\psi_x| - |\psi_{x'}\rangle\langle\psi_{x'}|\right] + \binom{N}{2}. \tag{3.42}$$

Every pair of two pure d-dimensional states can be viewed as pair of qubits embedded in a larger Hilbert space (they only span a qubit subspace). For two pure qubit states $|\phi_1\rangle$ and $|\phi_2\rangle$, it is straightforwardly shown that

$$\lambda_{\max}\left[|\phi_1\rangle\langle\phi_1| - |\phi_2\rangle\langle\phi_2|\right] = \sqrt{1 - |\langle\phi_1|\phi_2\rangle|^2}. \tag{3.43}$$

Consequently, it holds that

$$\mathcal{A}'_d = \max_{\{\psi_x\}} \sum_{x<x'} \sqrt{1 - |\langle\psi_x|\psi_{x'}\rangle|^2} + \binom{N}{2}. \tag{3.44}$$

Now, we exploit a simple inequality that follows from the concavity of the square-root function: for a real numbers $z_i \geq 0$ and $i = 1, \ldots, N$ it holds that

$$\sum_{i=1}^{N} \sqrt{z_i} \leq \sqrt{N \sum_{i=1}^{N} z_i}, \tag{3.45}$$

with equality if and only if all the z_i are equal. Applying this concavity inequality, we end up with

$$\mathcal{A}'_d \leq \max_{\{\psi_x\}} \sqrt{\binom{N}{2}^2 - \binom{N}{2} \sum_{x<x'} |\langle\psi_x|\psi_{x'}\rangle|^2} + \binom{N}{2}. \tag{3.46}$$

From here, we see that the maximal witness is obtained from evaluating the minimum of the sum under the square-root. Let us trivially re-write it by changing the domain of the summation

$$\sum_{x<x'} |\langle\psi_x|\psi_{x'}\rangle|^2 = \frac{1}{2}\left[\sum_{x,x'} |\langle\psi_x|\psi_{x'}\rangle|^2 - N\right]. \tag{3.47}$$

The sum on the right-hand-side is known as the frame-potential and its lowest value is known to be N^2/d [60]. We have therefore arrived at an upper bound on the witness,

$$\mathcal{A}'_d \leq \sqrt{\frac{N^3\,(N-1)\,(d-1)}{4d}} + \binom{N}{2}. \tag{3.48}$$

Let us now examine the conditions under which this bound is tight. In fact, the only condition for tightness is saturating of the inequality (3.45). This happens if and only if all overlaps $|\langle\psi_x|\psi_{x'}\rangle|^2$ are equal (for $x < x'$). This is precisely the defining characterisitic of the SIC-POVM. Hence, we conclude that our bound on $\mathcal{A}'_d$ can be saturated if and only if a SIC-POVM exists. In summary, observing a saturation of the bound implies that Alice's states form a SIC.

Now, we can immediately extend this to certify also the SIC-POVM. We supply Bob with the additional setting **povm** whose outcome is $b \in [N]$. The modified correlation witness becomes[2]

$$\mathcal{A}_d = \mathcal{A}'_d + \sum_{x=1}^{N} p(b=x|x, \mathbf{povm}). \tag{3.49}$$

Now, let us denote the latter sum by R. Evaluating its maximum, we find

$$R = \max_{\{\rho\},\{M_{\mathbf{povm}}\}} \sum_{x=1}^{N} p(b=x|x,\mathbf{povm}) = \max_{\{\rho_x\},\{M_{\mathbf{povm}}\}} \sum_{x=1}^{N} \mathrm{tr}\left(\rho_x M_{x|\mathbf{povm}}\right) \tag{3.50}$$

$$\leq \max_{\{M_{\mathbf{povm}}\}} \sum_{x=1}^{N} \lambda_{\max}\left[M_{x|\mathbf{povm}}\right] \leq \max_{\{M_{x|\mathbf{povm}}\}} \mathrm{tr}\left[\sum_{x=1}^{N} M_{x|\mathbf{povm}}\right] = d. \tag{3.51}$$

Equality in the first inequality requires that ρ_x is pure and aligned with $M_{x|\mathbf{povm}}$. Equality in the second inequality requires that $M_{x|\mathbf{povm}}$ is rank-one. Combine this fact with the maximal value of $\mathcal{A}'_d$ implying that $\{\rho_x\}$ forms a SIC, it follows that $\{M_{x|\mathbf{povm}}\}$ must be a SIC-POVM. We conclude that

$$\mathcal{A}_d \leq \sqrt{\frac{N^3\,(N-1)\,(d-1)}{4d}} + \binom{N}{2} + d \equiv \mathcal{A}_d^{\mathrm{Q}}, \tag{3.52}$$

is a tight bound on $\mathcal{A}_d$ (provided a SIC exists in dimension d) and that saturating it implies both that Alice prepares an ensemble of states forming a SIC and that Bob's setting **povm** corresponds to an aligned SIC-POVM.

Finally, we can ask the same question as we did for qubits: how much sub-optimality can we tolerate in the quantum correlations before we a simulation with stochastic projective measurements becomes possible? In other words, can we derive a non-trivial bound on $\mathcal{A}_d$ valid for

[2]Notice that in contrast to the qubit case we use a plus sign instead of a minus sign here. In fact, the sign does not matter too much: one can modify the expression $p(b=x|x,\mathbf{povm})$ somewhat and put a minus sign instead.

d	2	3	4	5	6
LB: $\mathcal{A}_d^P$	12.8484	70.0961	231.2685	578.7002	1219.0129
UB: $\mathcal{A}_d^P$	12.8484	70.1133	231.2685	578.7987	1219.2041
$\mathcal{A}_d^{\mathrm{Q}}$	12.8990	70.1769	231.3313	578.8613	1219.2667

Table 3.1: Upper bounds (UB) and lower bounds (LB) on $\mathcal{A}_d$ when measurements are constrained to be projective and d-dimensional. The lower bounds are obtained via SDPs in see-saw and the upper bounds are obtained via symmetrised semidefinite relaxations.

all projective measurements? The technique previously described for qubits is not useful in this higher-dimensional scenario. However, we can use the NV hierarchy (which per default uses projective measurements) to evaluate such a bound. The challenge is that our problem has many settings, many outputs and high-dimension. Therefore, the NV hierarchy in its standard form is unlikely to do the job. It is therefore imperative to make use of the symmetrisation techniques described in the previous chapter. In this manner, we have efficiently obtained bounds. The results (both upper bounds via the hierarchy and lower bounds via SDPs in see-saw) are presented in Table 3.1.

Moreover, this also serves as an illuminating example of the power of the symmetrised NV hierarchy in a practically motivated problem. We present some important computational parameters in Table 3.2. We compare the evaluation using i) no symmetries (standard implementation), ii) only reduction of the number of SDP variables (symmetry in the sampling stage), and iii) full symmetrisation (both reduction of SDP variables and block-diagonalisation). We see that the case i) is too demanding (on our standard desktop) already at $d = 3$. However, the advantages in the number of variables are very large already at that stage. Using ii), we can successfully evaluate the SDP up to $d = 4$ within a few minutes. However, using iii) we keep all the advantages in variable reduction but also obtain very substantial advantages via block-diagonalisation. It allows us to evaluate the case of $d = 6$ in just 1.2 seconds.

3.3 Certification of quantum instruments

A quantum measurement is a process that transforms a quantum system into a set of classical outputs. Consequently, there are many ways of implementing the same POVM. For instance, we could implement a binary-outcome identity measurement via the POVM $\{\frac{1}{2}\mathbb{I}, \frac{1}{2}\mathbb{I}\}$ or equally well by flipping an unbiased coin and measure σ_z for heads and $-\sigma_z$ for tails. We will see the same outcome statistics, but the procedures are inherently different; the former is a deterministic non-interacting measurement and the latter is a stochastic projective measurement. How could we tell

	d	2	3	4	5	6
Non-sym	#samples	221	>12000	-	-	-
	bl. sizes	1[43]	1[229]	1[741]	1[1831]	1[3823]
	SDP [s]	2.0	-	-	-	-
Sym no BD	#samples	65	134		137	
	bl. sizes	1[43]	1[229]	1[741]	1[1831]	1[3823]
	SDP [s]	0.5	19	500	-	-
Sym +BD	#samples	65	134		137	
	bl. sizes	4[6,16]	7[3,16]		8[3,16]	
	SDP [s]	0.3	0.6		1.2	

Table 3.2: Comparison between computational parameters for the task of bounding $\mathcal{A}_d$ under projective measurements using a standard implementation, symmetrisation to reduce the number of SDP variables, and symmetrisation to also perform block-diagonalisation (BD). The notation $D[a, b]$ means that there are D blocks with the smallest being of size a and the largest of size b. The symbol $-$ means that the evaluation was too demanding.

these operations apart? The answer is not to look at the POVM itself (which is indeed the same) but at the *quantum instrument*.

A quantum instrument is a process that maps a quantum system to a pair of outputs; one classical output and one quantum output. The classical output is a measurement outcome while the quantum output can be viewed as a post-measurement state (see Figure 3.4). In other words, a quantum instrument gives a broader picture than a POVM; we now care about what quantum state is left once the measurement is performed. Indeed, our above example corresponds to two very different quantum instruments. The former is the identity instrument. We can think of it as the state $|\psi\rangle$ being subjected to this process: the instrument flips a coin and outputs the classical result while $|\psi\rangle$ remains untouched and exits the instrument as its quantum output. Our stochastic measurements of $\pm\sigma_z$ are different. In this case, the state that exits the instrument is either $|0\rangle$ or $|1\rangle$ in every instance. Averaging out the stochastic element, the average state outputted by the instrument becomes $\frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \mathbb{I}/2$. Clearly, the two instruments are very different, but produce the same classical outcome statistics.

We consider the task of certifying quantum instruments operating on qubits solely from the statistics they produce in experiment. However, it is clear that our previously considered prepare-and-measure experiments are insufficient for the task since they make no regard to Alice's states after they have been measured by Bob. In order to enable the certification of instruments, we

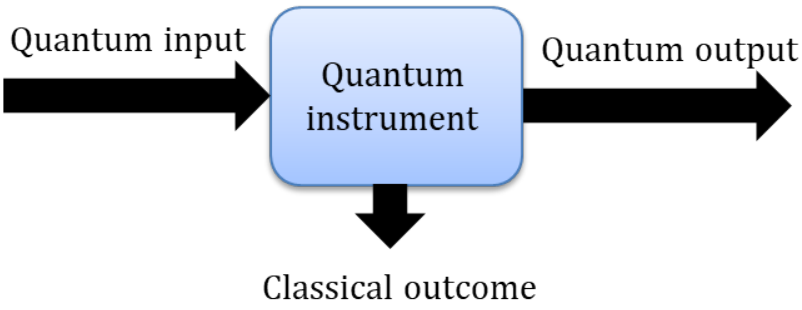

Figure 3.4: A quantum instrument transforms a quantum input state into a classical outcome and a quantum output state.

must add yet another party. We consider a prepare-transform-measure scenario as illustrated in Figure 3.5. Alice receives four inputs in the shape of two bits $x \equiv x_0, x_1 \in \{0,1\}$. She prepares a qubit state ρ_x that is sent to Bob. Bob receives an input $y \in \{0,1\}$ and implements a quantum instrument with binary classical outcomes $b \in \{0,1\}$ and a qubit output $\rho_x^{y,b}$. The latter is sent to Charlie who receives an input $z \in \{0,1\}$ and implements a POVM $\{C_{c|z}\}$ with binary outcomes $c \in \{0,1\}$.

Bob's instrument is represented by a set of Kraus operators $\{K_{b|y}\}$. These operators correspond to the POVM elements $M_{b|y} = K_{b|y}^\dagger K_{b|y}$. Since the POVM is complete, we require $\sum_b K_{b|y}^\dagger K_{b|y} = \mathbb{I}$. The post-measurement state sent from Bob to Charlie reads

$$\rho_x^{y,b} = \frac{K_{b|y}\rho_x K_{b|y}^\dagger}{\mathrm{tr}\left(\rho_x K_{b|y}^\dagger K_{b|y}\right)}. \tag{3.53}$$

Therefore, we can write the probability distribution for the prepare-transform-measure experiment as

$$p(b,c|x,y,z) = \mathrm{tr}\left[K_{b|y}\rho_x K_{b|y}^\dagger C_{c|z}\right]. \tag{3.54}$$

We use this distribution towards a specific correlation witness. We will once again employ the quantum random access code. This time, however, we will implement it twice: once between Alice and Bob and once between Alice and Charlie. Thus, we have a pair of witnesses that are defined as

$$W_{\mathrm{AB}} = \frac{1}{8}\sum_{x,y} p(b = x_y|x,y) \tag{3.55}$$

$$W_{\mathrm{AC}} = \frac{1}{8}\sum_{x,z} p(c = x_z|x,z). \tag{3.56}$$

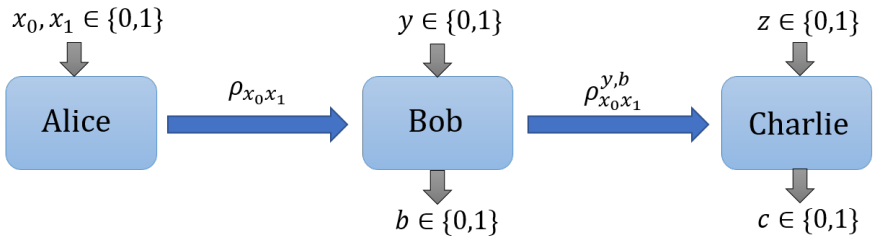

Figure 3.5: Prepare-transform-measure scenario. Alice receives inputs x_0, x_1 and prepares qubit states $\rho_{x_0x_1}$ that are transformed by Bob, using input y, into $\rho^{y,b}_{x_0x_1}$ with classical output b and finally measured by Charlie, using input z, returning the outcome c.

In the quantum model, the two quantum random access codes read

$$W_{\mathrm{AB}} = \frac{1}{8}\sum_{x,y} \mathrm{tr}\left[\rho_x M_{x_y|y}\right], \tag{3.57}$$

$$W_{\mathrm{AC}} = \frac{1}{16}\sum_{x,y,b,z} \mathrm{tr}\left[K_{b|y}\rho_x K^{\dagger}_{b|y} C_{x_z|z}\right]. \tag{3.58}$$

The intuiton behind these witnesses is the following. If Bob performs measurements that strongly interact with Alice's states, then he can have strong correlations meaning a large value of W_{AB}. However, the strong interaction means that he substantially disturbs Alice's state, hence washing out much of her encoded information. This will ensure that Charlie does not manage to have strong correlations with Alice, i.e. he finds W_{AC} small. Conversely, Bob could weakly interact with Alice's states such that much of her information remains to be harvested by Charlie. This would mean a large value of W_{AC}. However the weak interaction means that Bob chooses to not extract much information from the incoming states, thus suggesting a small W_{AB}. In summary, we expect there to be a non-trivial trade-off between the two quantum random access codes.

Therefore, it is interesting to characterise the region in the $(W_{\mathrm{AB}}, W_{\mathrm{AC}})$-plane attainable by quantum models. To this end, we consider the largest value of W_{AC} possible for a given value of W_{AB}. Formally, we can write this problem as

$$\begin{aligned} W^{\alpha}_{\mathrm{AC}} = \max_{\rho,U,M,C} & W_{\mathrm{AC}} \\ \text{such that } & \forall x : \rho_x \in \mathbb{C}^2,\ \rho_x \geq 0,\ \mathrm{tr}\,\rho_x = 1, \\ & \forall z, c : C_{c|z} \geq 0,\ C_{0|z} + C_{1|z} = \mathbb{I} \\ & \forall y, b : U_{yb} \in \mathrm{SU}(2),\ M_{b|y} \geq 0,\ M_{0|y} + M_{1|y} = \mathbb{I}, \\ & \text{and } \ W_{\mathrm{AB}} = \alpha, \end{aligned} \tag{3.59}$$

where we have used that the Kraus operators can be decomposed as $K_{b|y} = U_{yb}\sqrt{M_{b|y}}$ for some unitary U_{yb}. Evaluating this optimisation is not straightforward and quite lengthy (the full derivation appears in Ref [55]). We directly present the solution, which is

$$W_{\mathrm{AC}}^{\alpha} = \frac{1}{8}\left(4 + \sqrt{2} + \sqrt{16\alpha - 16\alpha^2 - 2}\right). \tag{3.60}$$

It turns out that up to global unitaries (or collective unitaries for Bob and Charlie), this optimal value is achieved if and only if Alice prepares the four states optimal in the quantum random access code (1.43). Bob performs weak measurements $M_0 = \eta\sigma_x$ and $M_1 = \eta\sigma_z$ where

$$\eta = \sqrt{2}\,(2W_{\mathrm{AB}} - 1) \tag{3.61}$$

and Charlie performs projective measurements of σ_x and σ_z. Here η is the *sharpness* of Bob's POVM. Importantly, whereas Bob's POVMs are non-extremal and can therefore be simulated with stochastic projective measurements, this result shows that Bob's instruments are extremal and in fact implied by the pair of quantum random access codes. On a simple form, we can write the boundary of the quantum region as

$$W_{\mathrm{AB}} = \frac{1}{4}\left(2 + \eta\sqrt{2}\right) \tag{3.62}$$

$$W_{\mathrm{AC}} = \frac{1}{8}\left(4 + \sqrt{2} + \sqrt{2 - 2\eta^2}\right). \tag{3.63}$$

We illustrate the quantum region in Figure 3.6. All along the boundary of the quantum region (its non-trivial part is marked by the solid red line), we can precisely certify the quantum instruments whose POVM component is a weak measurement of σ_x and σ_z respectively. The classically attainable correlations is much simpler: we already know that the best classical implementation of the random access code has a success rate of $3/4$. Since classical measurements simply reveal pre-existing properties, they do not disturb the system. Since the random access code between Alice and Charlie does not benefit from information held by Bob, it is clear that the best strategy is for Alice and Bob to perform the standard random access code and for Bob to relay his unperturbed state to Charlie who again performs the standard random access code with Alice. Then, we have $W_{\mathrm{AB}} = W_{\mathrm{AC}} = 3/4$. The classical region of correlations therefore becomes a rectangle in the $(W_{\mathrm{AB}}, W_{\mathrm{AC}})$-plane.

What can be said about the instruments when the quantum correlations are sub-optimal, i.e. when they are not on the boundary? It turns out that one may establish bounds on the sharpness η. For simplicity, assuming that both Bob's settings correspond to the same sharpness, we can confine η to an interval whose upper and lower delimitation is determined by $(W_{\mathrm{AB}}, W_{\mathrm{AC}})$.

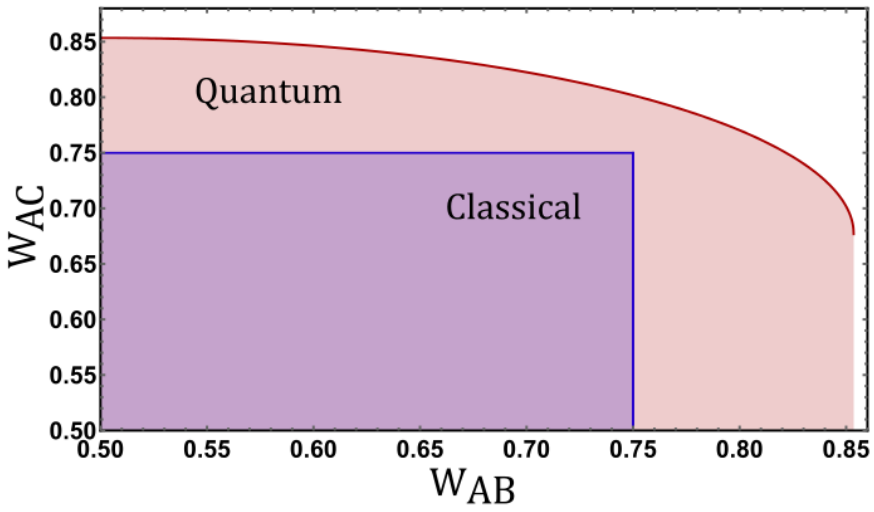

Figure 3.6: Quantum region of correlations attainble in the prepare-transform-measure experiment of two parallel quantum random access codes. The boundary (solid red) certifies Alice's preparations, Bob's instruments and Charlie's POVMs.

By the same type of procedure that initially led to (3.60), one finds that

$$\eta \geq \sqrt{2}\,(2W_{\mathrm{AB}} - 1)\,, \qquad \eta \leq 2\sqrt{\left(2 + \sqrt{2} - 4W_{\mathrm{AC}}\right)(2W_{\mathrm{AC}} - 1)}. \tag{3.64}$$

The latter relation is valid when $\frac{4+\sqrt{2}}{8} \leq W_{\mathrm{AC}} \leq \frac{2+\sqrt{2}}{4}$, otherwise the bound is trivialised. Along the boundary of the quantum region, the upper and lower bounds coincide and thus η is precisely determined. The closer the quantum correlations are to being on the boundary, the smaller is the width of the interval to which one can confine η. Hence, we see that the quality of the certification improves as the quantum correlations approach the boundary of the quantum region. These results have been experimentally demonstrated in Refs [61, 62].

3.4 Certification of entanglement

The experiments we have hitherto considered in this chapter have been based on single quantum systems. However, quantum correlations can also be used to certify composite quantum states. Among such states, the prominent resource of interest is entanglement. The strongest form of certification of entangled states is based on quantum nonlocality since it achieves the certification under what is arguably the smallest possible assumptions (quantum theory and the no-signaling principle). However, many entangled quantum states are not known to violate any Bell inequality. Even more problematically, many entangled quantum states are known to never violate any Bell

inequality [63]. This is a conceptual obstacle for entanglement certification based on Bell inequality violations. Also, for practical purposes, it is many times not motivated to make such weak assumptions in a certification protocol. Here, we continue on the theme of certifying quantum devices when only assuming their Hilbert space dimension. In order to extend these ideas to the certification of entanglement, we consider hybrid experiments that both involve entanglement and quantum communication. Importantly, it is many times the case that certification of entanglement is restricted to the entanglement of states. However, the certification of entangled measurements is conceptually at least as compelling. We present a scheme which works for both the certification of entangled states and entangled measurements subject to a dimension assumption. The scheme has the advantage of being versatile (applying to systems of any dimension d and any number of subsystems N) and it detects many entangled states that cannot be detected via Bell inequalities. For important families of entangled states and measurements, it can even detect every entangled system.

Scenario for entanglement certification

We outline the scenario in which we perform the entanglement certification. It is parameterised by the local dimension d and the number of subsystems in the state n. The state ρ is emitted by a source such that its subsystems are distributed to n separate parties named $A_1, \ldots, A_n$. Each party receives a random input denoted $x_k, y_k \in [d]_0^2$ and implements a corresponding transformation $\mathcal{T}_{x_k y_k}^{(k)}$ that maps the incoming d-dimensional system into an outgoing d-dimensional system. The quantum state outputted by each party is then sent to a final party B who performs a measurement $\{M_b\}$ and records the outcome $b \equiv b_1 \ldots b_n \in [d]_0^n$. The scenario is illustrated in Figure 3.7.

The probability distribution is denoted $p(b|x,y)$ where $x = x_1, \ldots, x_n$ and $y = y_1, \ldots, y_n$. It is given by

$$P(b|x,y) = \mathrm{tr}\left[\left(\bigotimes_{k=1}^{n} \mathcal{T}_{x_k y_k}^{(k)}\right)[\rho] \cdot M_b\right]. \tag{3.65}$$

It is used towards evaluating a specific correlation witness that can be phrased as a game. Whenever the following conditions are satisfied, the game is won

$$b_1 = \sum_{i=1}^{n} x_i \equiv C_1(x) \qquad \text{and} \qquad b_k = y_k - y_1 \equiv C_k(y), \tag{3.66}$$

for $k = 2, \ldots, N$ where computations are modulo d. We compactly write the winning condition as $b = C(x,y)$. The average score in the game becomes

$$\mathcal{A}_{n,d} = \frac{1}{d^{2n}} \sum_{x,y} P\left(b = C(x,y) \,|\, x, y\right). \tag{3.67}$$

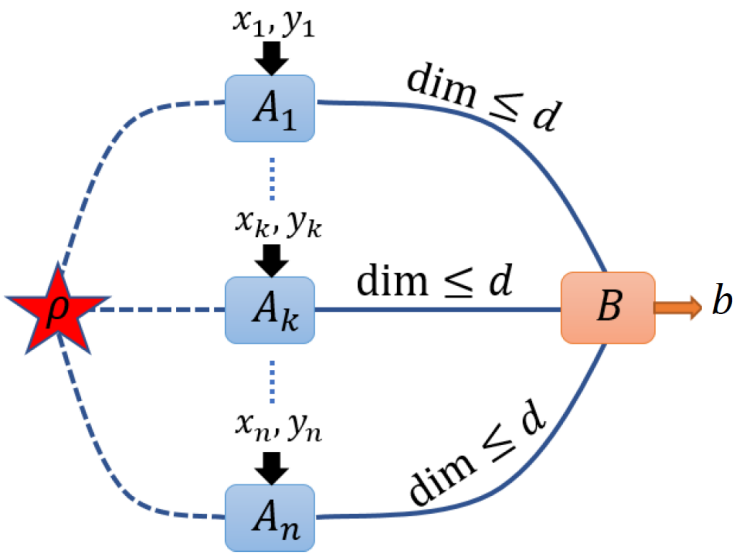

Figure 3.7: Scenario for entanglement certification. A state of n subsystems of local dimension d is shared between parties $A_1, \ldots, A_n$ who use their respective inputs (x_k, y_k) to locally output a d-dimensional system. The subsystems are relayed to B who performs a measurement with outcome b.

We use the value of $\mathcal{A}_{n,d}$ to certify and characterise the entanglement in the state ρ and the measurement $\{M_b\}$.

Entanglement of the state

Entanglement comes in many forms. Our interest is *genuine multipartite entanglement* (GME). A state is said to be GME when there exists no partition of its subsystems $\{1, \ldots, n\}$ into two sets $\{S, \bar{S}\}$ for which the state can be written on the form

$$\rho = \sum_S \sum_i p_{S,i} \rho_i^S \otimes \rho_i^{\bar{S}}, \tag{3.68}$$

where $p_{S,i}$ is a probability distribution. A state that can be written on this form is said to be biseparable. In order to witness GME through our entanglement certification scheme, we must derive the largest possible value of $\mathcal{A}_{n,d}$ attainable for biseparable states. Then, any value of $\mathcal{A}_{n,d}$ larger than that limit certifies that the state is GME.

Let us first notice the following simple classical strategy for the game. Each party $A_1, \ldots, A_n$ simply discards the received state and instead sends y_k to B over the channel. Hence, B has access to $\{y_1, \ldots, y_n\}$. This allows B to output $b_k = C_k(y)$ for $k = 2, \ldots, n$. The only remaining winning condition to be satisfied is the first one, i.e. $\sum_i x_i = C_1(x)$. However, since no information is held

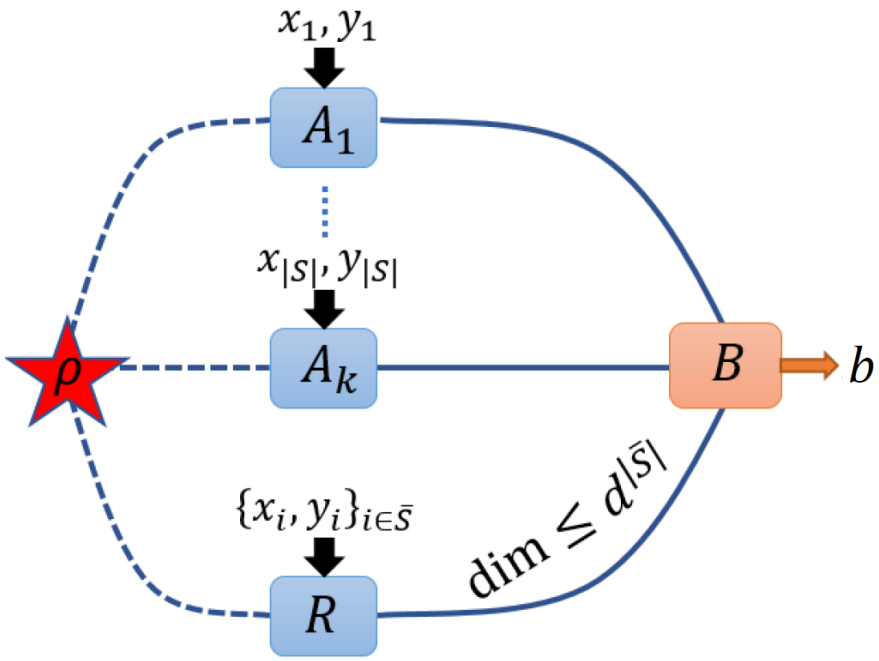

Figure 3.8: Relaxation of the entanglement certification scheme. The parties are grouped into two sets. The parties in the first set S are allowed to send their entire input to B. Parties in the complementary set $\bar{S}$ are grouped together into a single party R who is allowed to communicate with knowledge of the inputs of all these parties.

about x, the best option is to guess the sum $\sum_i x_i$. That guess is correct with probability $1/d$. Therefore, we have found that $\mathcal{A}_{n,d} = 1/d$. Since classical strategies also are biseparable, we know that the biseparable bound on $\mathcal{A}_{n,d}$ must be at least $1/d$.

As it turns out, biseparable models cannot do better than this simple classical strategy. Let us prove this statement. Firstly, any value of $\mathcal{A}_{n,d}$ attainable in a stochastic biseparable model (i.e. a non-trivial $p_{S,i}$) is also attainable in a deterministic biseparable model. This follows from linearity. Hence, we may w.l.g. consider an arbitrary partition of the subsystems $\{S, \bar{S}\}$ and ascribe the state $|\chi\rangle = |\psi\rangle_S \otimes |\phi\rangle_{\bar{S}}$. We must evaluate

$$\mathcal{A}_{n,d}^{\text{bisep}} = \max_{|\chi\rangle,\{\mathcal{T}\},\{M\}} \mathcal{A}_{n,d}. \tag{3.69}$$

In order to place an upper bound on this quantity, let us relax some constraints in the game. We allow the parties $\{A_k\}_{k\in S}$ to communicate all their information to B. The remaining parties $\{A_k\}_{k\in\bar{S}}$ are grouped together, forming an effective party R who holds the collection of all their inputs. This party is allowed to send the same amount of communication as the sum of the communication allowed for the original parties $\{A_k\}_{k\in\bar{S}}$, i.e. $|\bar{S}|$ d-dimensional systems. This is

equivalent to a quantum system of dimension $d^{|\bar{S}|}$. We illustrate this relaxed scenario in Figure 3.8

In this relaxed scenario, we can easily calculate the biseparable bound on the correlation witness. The relaxed scenario has effectively turned matters into a prepare-and-measure scenario between R and B. In order to win, R needs to relay $\sum_{i\in\bar{S}} x_i$ as well as $\{y_i\}_{i\in\bar{S}}$. However, this corresponds to $|\bar{S}|+1$ d-valued messages. That is one d-valued message too many to be sent over the channel to B. There is no strategy in which the average probability of B to recover all relevant information is more than $1/d$. This statement follows from the fact that the guessing probability of N quantum signals of dimension D respects

$$p^{\text{success}} \equiv \frac{1}{N}\sum_{x=1}^{N} p(b=x|x) = \frac{1}{N}\sum_{x=1}^{N} \text{tr}\left(\rho_x M_x\right) \leq \frac{1}{N}\sum_{x=1}^{N} \lambda_{\max}\left(M_x\right) \tag{3.70}$$

$$\leq \frac{1}{N}\sum_{x=1}^{N} \text{tr}\left(M_x\right) = \frac{1}{N}\,\text{tr}\left(\sum_{x=1}^{N} M_x\right) = \frac{D}{N}. \tag{3.71}$$

In our case, we have $D = d^{|\bar{S}|}$ and $N = d^{|\bar{S}|+1}$. This gives $p^{\text{success}} = 1/d$. We conclude that

$$\mathcal{A}_{n,d}^{\text{bisep}} = \frac{1}{d}. \tag{3.72}$$

How well can the game be performed with states that are GME? Consider the following quantum strategy. Let the state be the Greenberger-Horne-Zeilinger state

$$|\text{GHZ}_{n,d}\rangle = \frac{1}{\sqrt{d}}\sum_{i=0}^{d-1} |i\rangle^{\otimes n}. \tag{3.73}$$

We let each party $A_1, \ldots, A_n$ perform a unitary transformation

$$U_{x_k y_k}^{A_k} = Z^{x_k} X^{y_k}, \tag{3.74}$$

for $k = 1, \ldots, n$ where we have defined the so-called clock and shift operators

$$Z = \sum_{j=0}^{d-1} e^{2i\pi j/d} |j\rangle\langle j| \qquad\qquad X = \sum_{j=0}^{d-1} |j+1\rangle\langle j|. \tag{3.75}$$

The measurement of B is defined in terms of how the clock and shift operators act on the GHZ state. A basis of "GHZ-like states" for $(\mathbb{C}^d)^{\otimes n}$ is obtained by defining

$$|M_b\rangle = Z^{b_1} \otimes X^{b_2} \otimes \cdots \otimes X^{b_n} |\text{GHZ}_{n,d}\rangle. \tag{3.76}$$

This can be thought of as a generalised (multipartite and high-dimensional) Bell State Measurement [64]. This strategy is tailored to give

$$\mathcal{A}_{n,d}^{\text{Q}} = 1, \tag{3.77}$$

which coincides with the algebraically largest value of $\mathcal{A}$. Hence, this represents the quantum maximum. We see that by employing GME, we can outperform the biseparable bound. We remark that this protocol can be thought of a semi-device-independent scheme for superdense coding [65].

Let us illustrate the power of this scheme in an example. Consider the depolarisation of the GHZ state, defined as

$$\rho_{n,d}^{\mathrm{GHZ}}(v) = v\,|\mathrm{GHZ}_{n,d}\rangle\langle\mathrm{GHZ}_{n,d}| + \frac{1-v}{d^n}\mathbb{I} \tag{3.78}$$

for some $v \in [0,1]$. What is the critical v for certifying GME? Using the same unitaries and the same measurement as above, we arrive at a certification of GME whenever

$$v > \frac{d^{n-1}-1}{d^n-1}. \tag{3.79}$$

This result features some important special cases.

- If we have two subsystems ($n = 2$), then we find

$$v > \frac{1}{d+1}. \tag{3.80}$$

  This condition is identical to the condition for the state $\rho_{2,d}$ being entangled [66]. These states, known as Werner states, are broadly studied and applied in quantum information processing. For these states, the entanglement certification is optimal. Notably, many of these states cannot be certified in a Bell experiment [67].

- If we have many qubits ($d = 2$), we find

$$v > \frac{2^{n-1}-1}{2^n-1}. \tag{3.81}$$

  This condition is again identical to the condition for the state $\rho_{n,2}$ being GME [68]. Hence the entanglement certification is again optimal.

For systems of many higher-dimensional systems, the certification is not optimal. An interesting feature to note here is that states that cannot even be used for Einstein-Podolsky-Rosen steering[3] can still be detected in this scheme subject only to a dimension bound. A simple example is that the state $\rho_{2,2}$ is steerable only when $v > 1/2$ [69] but certified in our scheme when $v > 1/3$.

[3] A test of steering takes place in a Bell experiment in which the measurements one party are assumed to be perfectly controlled and known a priori.

Certifying the extractable GHZ fraction

The scheme can reveal more about the state than solely whether it is GME: a more precise characterisation of ρ is possible. Consider the following quantum strategy. The transformations of parties $A_1, \ldots, A_n$ correspond to first implementing a CPTP map $\Lambda_k[\rho]$ and then implementing the above defined unitaries $U^{A_k}_{x_k y_k}$. Hence, the transformation becomes

$$\mathcal{T}^{(k)}_{x_k y_k} = U^{A_k}_{x_k y_k} \Lambda_k[\rho] \left(U^{A_k}_{x_k y_k}\right)^\dagger . \tag{3.82}$$

Using the same measurement as outlined above and optimising over the extraction channels Λ_k leads to

$$\mathcal{A}_{n,d}(\rho) = \mathrm{EGF}_{n,d}(\rho), \tag{3.83}$$

where we have defined the *extractable GHZ fraction* as

$$\mathrm{EGF}_{n,d}(\rho) = \max_{\Lambda_1,\ldots,\Lambda_n} \mathrm{tr}\left(\left(\bigotimes_{k=1}^{n} \Lambda_k \right) [\rho] \cdot |\mathrm{GHZ}_{n,d}\rangle\langle \mathrm{GHZ}_{n,d}| \right). \tag{3.84}$$

This can be interpreted as "the amount of GHZ state" that can be extracted from ρ via local operations. It is a straightforward generalisation of the concept of a singlet fraction encountered in quantum teleportation [66]. This merely shows that every quantum state has the ability of creating correlations in the game that are equal to the extractable GHZ fraction. However, most interestingly, substantial numerical evidence obtained from many samples of states with reasonably small n and d strongly indicates that the extractable GHZ fraction in fact is the largest achievable score in the game for any given state. As it presently stands, it is a conjecture that allows us to certify a lower bound on the extractable GHZ fraction directly from the observed value of $\mathcal{A}_{n,d}$. Proving it is an open problem.

Certification of entangled measurements

The arguably most interesting aspect about the entanglement certification scheme is that it does not only apply to states but also to measurements. Specifically we certify qualitative properties in the d^n-outcome measurement performed by B directly from the value of $\mathcal{A}_{n,d}$.

We say that a measurement is entangled if at least one of the operators $\{M_b\}$ are entangled. Naturally, knowing that a measurement is entangled is interesting since it does not have a classical counterpart. However, from a quantum perspective, it tells us quite little about the measurement. It does not tell us how strong the entanglement is, nor to what extent it is present in the measurement - it may very well be that out of thousands of measurement operators, only one is entangled. This

obstacle is remedied by the proposed scheme. Specifically, by inspecting $\mathcal{A}_{n,d}$, we can determine a bound on the number of measurement operators that must be entangled in the set $\{M_b\}$. The larger the value of $\mathcal{A}_{n,d}$ is, the larger is the number of measurement operators that can be certified as entangled. The specific bounds for this certification are given by

$$\text{At least } k \text{ separable measurement operators } \implies \mathcal{A}_{n,d} \leq \frac{1}{d^n}\left(d^n - k + \frac{k}{d}\right). \tag{3.85}$$

Thus, violating this inequality certifies $d^n - k + 1$ entangled measurement operators. The proof of this result is of technical nature and it is detailed in Ref [56].

As an example, consider a joint measurement of two d-dimensional systems. We implement the Bell State Measurement subject to noise, i.e. the measurement operators read

$$M_{b_1 b_2}(v) = v\,|M_{b_1 b_2}\rangle\langle M_{b_1 b_2}| + \frac{1-v}{d^2}\mathbb{I}. \tag{3.86}$$

What is the critical visibility v for certifying entanglement in the measurement? Using the state $|\mathrm{GHZ}_{2,d}\rangle$ and the optimal unitaries previously discussed, one finds that the condition for witnessing at least one entangled measurement operator is

$$v > \frac{1}{d+1}, \tag{3.87}$$

which is also the condition for the measurement operators being entangled. In the other end, in order to witness that all measurement operators must be entangled, we require

$$v > \frac{d^2 + d - 1}{d(d+1)}. \tag{3.88}$$

For qubit systems, this corresponds to a visibility of $v = 5/6$.

4

Operational contextuality

This chapter investigates correlations that reveal operational contextuality. It is known that contextuality can be tested through quantum communication games. In the first section, we make this connection systematic and apply it to derive families of many-outcome noncontextuality inequalities [17]. In the second section, we prove general one-to-one connections between both quantum contextuality and measurement incompatibility as well as between quantum contextuality and steering [19]. In the third section, we explore how the ability of state ensembles to generate quantum contextuality can be independently harvested by several independent observers [70].

4.1 Communication games reveal contextuality

In section 1.2, we have seen that quantum contextuality is the failure of explaining quantum correlations in an ontological model that assigns equal ontology to operationally distinguishable preparation and/or measurement procedures. We also exemplified a quantum ensemble that enables preparation contextuality. Here, we discuss the systematic link between tests of contextuality and quantum communication games of the prepare-and-measure type encountered in previous chapters. Then, we apply the framework to derive families of preparation noncontextuality inequalities that can be robustly tested in experiment.

Again, our quantum communication games feature two parties, Alice and Bob. Alice receives an input x sampled from a space I_{A} with probability $p_{\mathrm{A}}(x)$ and Bob receives an input y sampled from a space I_{B} with probability $p_{\mathrm{B}}(y)$. Alice encodes her input into a state and Bob uses his input to choose a measurement to apply to said state. The outcome b is rewarded with c_{xyb} points. Then, the average score reads

$$\mathcal{A} = \sum_{x,y,b} c_{xyb} p_{\mathrm{A}}(x) p_{\mathrm{B}}(y) p(b|x,y). \tag{4.1}$$

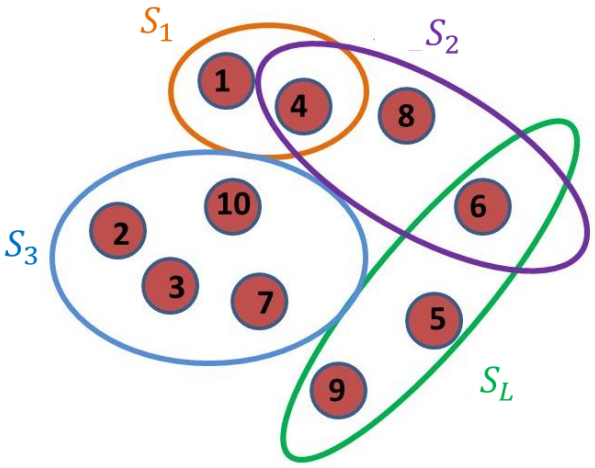

Figure 4.1: Alice's inputs (here represented by ten circles) are divided into L sets. Then an operational equivalence is imposed: Bob cannot distinguish to which set a any given preparation belongs.

In order to connect such a prepare-and-measure game to a test of contextuality, we must impose suitable communication constraints. These constraints must require Alice's preparations to constitute different contexts of the same preparation procedure. This can be achieved by imposing a data hiding constraint as follows. Take Alice's input space I_{A} and construct L non-empty subsets of it. We call them $S_k \subset I_{\mathrm{A}}$ for $k = 1, \ldots, L$ (see Figure 4.1). We would like to ensure that there exists no measurement that Bob could possibly perform that allows him to gain information about to which set his received preparation belongs. In other words, Alice must hide the set-membership of her preparations. We can write this data hiding constraint on the form

$$\forall y, b, k, k' : \quad \frac{1}{q_k} \sum_{x \in S_k} p(x|b,y) = \frac{1}{q_{k'}} \sum_{x \in S_{k'}} p(x|b,y), \tag{4.2}$$

where we have defined the normalisation $q_k = p(x \in S_k) = \sum_{x \in S_k} p_{\mathrm{A}}(x)$ as the prior probability that Alice's state is a member of S_k. Notice that since the sets $\{S_k\}_k$ do not need to be a partition of I_{A}, it is in general not the case that the q_k sum to one. Importantly, the way to interpret the above constraint is that it must hold for *every measurement that Bob could make* - not just the measurements he happens to make in the game. Presently, this condition is intuitive since Bob uses his data (b, y) to try to guess the membership of x. However, it is more handy to phrase the data hiding constraint in terms of the conditional probabilities $p(b|x, y)$ that appear in the communication game. Therefore, we apply Bayes' rule to write

$$p(x|b,y) = \frac{p(b|x,y)p(x|y)}{p(b|y)} = \frac{p(b|x,y)p_{\mathrm{A}}(x)}{p(b|y)}, \tag{4.3}$$

where we used that x and y are independent. Now, we can write the data hiding constraint on the form

$$\forall y, b, k, k' : \quad \sum_{x \in S_k} p(b|x,y) \frac{p_{\mathrm{A}}(x)}{q_k} = \sum_{x \in S_{k'}} p(b|x,y) \frac{p_{\mathrm{A}}(x)}{q_{k'}}. \tag{4.4}$$

Notice that $\{\frac{p_{\mathrm{A}}(x)}{q_k}\}_{x \in S_k}$ is a probability distribution over the inputs that belong to S_k. From convexity, we have that

$$p(b|x \in S_k, y) = \sum_{x \in S_k} p(b|x,y) \frac{p_{\mathrm{A}}(x)}{q_k}. \tag{4.5}$$

Our data hiding constraint therefore amounts to

$$\forall y, b, k, k' : \quad p(b|x \in S_k, y) = p(b|x \in S_{k'}, y). \tag{4.6}$$

We have recovered the definition of to preparation procedures (associated to sets S_k and $S_{k'}$) being operationally equivalent. Thereby, we confirm that our imposed data hiding condition link our communication games to tests of contextuality. However, in order to address noncontextual models, we must first recall that the hidden variable models that we are interested in have the property that the hidden variable distribution of a convex combination of preparation procedures is the convex combination of the hidden variable distributions associated to those preparation procedures. For our purposes, this means that

$$p(\lambda|x \in S_k) = \sum_{x \in S_k} p(b|x,y) \frac{p_{\mathrm{A}}(x)}{q_k}. \tag{4.7}$$

By the assumption of preparation noncontextuality, we impose that indistinguishable preparation procedures are ontologically equivalent, meaning that

$$\forall k, k' : \quad p(\lambda|x \in S_k) = p(\lambda|x \in S_{k'}). \tag{4.8}$$

Again, we use Bayes' rule to write this as

$$\frac{p(x \in S_k|\lambda)}{q_k} = \frac{p(x \in S_{k'}|\lambda)}{q_{k'}}, \tag{4.9}$$

which gives a simple interpretation: even if we know the hidden variable, the preparation procedures still remain indistinguishable to Bob. This is precisely the notion of preparation noncontextuality. With the noncontextual constraints in hand, we now know that there must exist a bound on the score in the game,

$$\mathcal{A} \overset{\mathrm{PNC}}{\leq} \mathcal{A}^{\mathrm{PNC}}, \tag{4.10}$$

that is respected by all preparation noncontextual models but can in principle be violated by contextual theories. Furthermore, the evaluation of this bound is a linear program and therefore also viable for practical purposes.

Preparation noncontextuality inequalities for high-dimensional systems

Let us put the above to use by deriving a family of high-dimensional preparation noncontextuality inequalities based on a simple communication game. We will employ the generalised random access code introduced in Ref [50]. In this game, Alice receives an input string $x = x_1 \dots x_n \in [d]_0^n \equiv I_{\mathrm{A}}$. Bob receives an input $y \in [n] \equiv I_{\mathrm{B}}$. The sampling of the inputs is uniform, i.e. $p_{\mathrm{A}} = 1/d^n$ and $p_{\mathrm{B}} = 1/n$. The aim of Alice and Bob is to maximise the score function

$$\mathcal{A} = \frac{1}{nd^n} \sum_{x,y} p(b = x_y|x, y). \tag{4.11}$$

Thus, the score is the same as in a standard random access code [50]. The key difference between our game and a standard random access code is the communication conditions under which the score is evaluated. The standard random access code is limited by the dimension of the physical system (as discussed in earlier chapters) whereas in our case, we will impose data hiding constraints while allowing for arbitrary high-dimensional systems. The data hiding constraints that we choose are as follows. Let r be a string $r \in [d]_0^n$ and let $Z(r)$ count the number of zeros that appear in the string. For every r that corresponds to $Z(r) \leq d - 2$, define the following subsets of Alice's input space,

$$S_k^r = \{x|x \cdot r \equiv \sum_{i=1}^{n} x_i r_i = k \mod d\}, \tag{4.12}$$

i.e. S_k^r contains all x such that the modular sum with r equals k. Since all these sets are of the same size and all priors are uniform, we can simply state the data hiding constraint as

$$\forall k, k', r, r' : \quad \sum_{x \in S_k^r} p(b|x, y) = \sum_{x \in S_{k'}^{r'}} p(b|x, y). \tag{4.13}$$

From our previous discussion, we know that when subjected to these data hiding constraints, there exists a bound on the score $\mathcal{A}$ that is respected by all preparation noncontextual models. However, since we are attempting to solve the problem for general d and n, we cannot simply run a linear program. The preparation noncontextual bound is obtained through a technical analysis based on Fourier expansions that is presented in the supplementary material of Ref [17]. It leads to the following family of preparation noncontextuality inequalities

$$\mathcal{A} = \frac{1}{nd^n} \sum_{y=1}^{n} \sum_{x \in \{0,\dots,d-1\}^n} p(b = x_y|x, y) \leq \frac{n + d - 1}{nd}. \tag{4.14}$$

These inequalities hold for general operational theories and can be violated in quantum theory. We note that the special case of $d = 2$ reduces to the preparation noncontextuality inequalities originally derived in Ref [71]. In Ref [17] some numerical examples were given for quantum violations

for some specific low values of d and n. However, no systematic investigation of quantum violations was conducted. In particular, an interesting open problem is to investigate the maximal quantum violations and the Hilbert space dimensions required for achieving them. It is the author's speculation that since the data hiding constraints in this task impose strong collective constraints on the preparation ensemble of Alice, a large quantum violation will typically require a high-dimensional Hilbert space (i.e. significantly larger than d) that can accommodate the large symmetries required to fulfill the constraint.

4.2 Contextuality, steering and measurement incompatibility

A fundamentally important question is how quantum contextuality relates to the ontology of quantum theory. For instance, incompatible measurements are necessary for Kochen-Specker contextuality, but not necessarily all incompatible measurements can give rise to such contextuality. In operational contextuality, does there exist one-to-one connections between quantum correlations and a nonclassical ontology? Here, we present two different links between contextuality and fundamental features of nonclassicality appearing in quantum theory, namely measurement incompatibility and steering. We show that measurement incompatibility is necessary and sufficient for preparation contextuality and that steering is necessary and sufficient for full contextuality in a Bell experiment. However, let us begin with briefly introducing measurement incompatibility and steering.

Briefly: measurement incompatibility

Measurements in quantum theory can be incompatible. Famously, one cannot precisely measure both position and momentum. For standard quantum measurements, corresponding to complete sets of projectors on Hilbert space, the definition of incompatibility is simple. Standard quantum measurements are incompatible if and only if they do not commute. However, as we have previously seen in our discussion of non-projective POVMs, general quantum measurements do not need to be projective. How can we define the compatibility of general sets of quantum measurements [72, 73]? The standard answer is that if for every x we have a POVM $\{A_{a|x}\}_a$, the full set of measurements is called compatible (or *jointly measurable*) if it can equally well be realised with just a single POVM $\{G_\lambda\}$ whose outcome λ is post-processed into the outcome a of our original POVMs. In other words, a set of POVMs is jointly measurable if and only if it can be simulated with a single POVM. Formally, we write this as

$$A_{a|x} = \sum_\lambda p(a|x,\lambda) G_\lambda, \tag{4.15}$$

for some probability distribution $p(a|x,\lambda)$. Consequently, if no such simulation is possible, the POVMs $\{A_{a,x}\}_{a,x}$ are said to be *incompatible*. Notice that deciding whether a given set of POVMs is jointly measurable can be achieved with an SDP.

Briefly: quantum steering

Steering in quantum theory [69] is the ability of one observer to remotely prepare states for another observer that cannot be simulated classically. This formalises the famous "spooky action at a distance" concept. Steering takes place in Bell experiments in which a bipartite state ρ is shared between two observers Alice and Bob. If Alice is going to steer Bob, she will randomly choose a measurement and apply it to her part of the state. Denoting her measurements by $\{A_{a|x}\}$, for each choice of setting x and each observed outcome a, she will remotely prepare Bob's system in the state

$$\rho_{a|x} = \frac{\mathrm{tr_A}\left(A_{a|x}\otimes \mathbb{I}\rho\right)}{\mathrm{tr}\left(A_{a|x}\otimes \mathbb{I}\rho\right)}. \tag{4.16}$$

For a given setting x, the probability of rendering Bob in $\rho_{a|x}$ is given by $p(a|x) = \mathrm{tr}\left(A_{a|x}\otimes \mathbb{I}\rho\right)$. For simplicity, one can describe the local states of Bob through a so-called assemblage, which is simply the collection of unnormalised states on Bob's side, $\{\sigma_{a|x}\}$ where

$$\sigma_{a|x} = \mathrm{tr_A}\left(A_{a|x}\otimes \mathbb{I}\rho\right). \tag{4.17}$$

The assemblage is said to be unsteerable if it can be simulated in a so-called *local hidden state model*. This model attempts to reproduce the assemblage through a source that emits a quantum state ρ_λ with probability $p(\lambda)$. When the state is given to Bob, he can post-process it with some distribution $p(a|x,\lambda)$ in order to simulate $\{\sigma_{a|x}\}$. If this is successful, we have that

$$\sigma_{a|x} = \sum_\lambda p(\lambda)p(a|x,\lambda)\rho_\lambda. \tag{4.18}$$

If a local hidden state model is not possible, the assemblage is said to be *steerable*.

Preparation contextuality is necessary and sufficient for measurement incompatibility

It is remarkably straightforward to show a one-to-one connection between preparation contextuality and measurement incompatibility. Let us assume that the set of POVMs $\{A_{a|x}\}_{a,x}$ when applied to a quantum state ρ give outcome statistics that is preparation noncontextual. This is written as

$$p(a|x,\mathbf{P}) = \sum_\lambda p(\lambda|\rho)p(a|x,\lambda), \tag{4.19}$$

where the preparation noncontextuality assumption is embodied in the fact that the hidden variable distribution only depends on the density matrix and not on its context. Thus, the above equality must hold for every context of the density matrix. That is, it holds true for every way of decomposing the density matrix into a convex combination of other states.

We must now characterise the hidden variable distribution $p(\lambda|\rho)$ as a convexity-preserving map from the full quantum state space to the real interval $[0, 1]$. The Riesz representation theorem provides us with a simple characterisation [74]. Every such map can be written as an inner product between the state ρ and some positive semidefinite operator G_λ,

$$p(\lambda|\rho) = \mathrm{tr}\left(\rho G_\lambda\right). \tag{4.20}$$

In order to obey normalisation, we must require that $0 \leq G_\lambda \leq \mathbb{I}$. Moreover, we need $p(\lambda|\rho)$ to be a probability distribution for every ρ. This means that

$$\forall \rho: \quad \sum_\lambda p(\lambda|\rho) = 1 \implies \sum_\lambda G_\lambda = \mathbb{I}. \tag{4.21}$$

Thus, we see that the operators $\{G_\lambda\}$ must form a POVM. Consequently, our preparation noncontextual probability distribution takes the form

$$p(a|x, \mathbf{P}) = \sum_\lambda p(a|x, \lambda)\, \mathrm{tr}\left(\rho G_\lambda\right). \tag{4.22}$$

We have recovered the outcome statistics corresponding to applying a set of jointly measurable observables to the state ρ. Conversely, if we assume that $\{A_{a|x}\}_{a,x}$ is jointly measurable, then the outcome statistics of measuring ρ is again given by (4.22) which is a preparation noncontextual model. We conclude the following

> *Preparation contextual outcome statistics implies measurement incompatibility. Conversely, every incompatible measurement can reveal preparation contextuality.*

In other words, preparation contextuality is necessary and sufficient for measurement incompatibility. An immediate corollary of this result is that every incompatible measurement is useful to obtain a quantum advantage in a communication task.

Contextuality is necessary and sufficient for steering

An essentially equally straightforward one-to-one connection can be established between quantum contextuality and quantum steering. Here, it is important to note that by quantum contextuality we mean the failure of an ontological model that is *both* preparation noncontextual and measurement

noncontextual. Thus, this "full" contextuality is a stronger constraint than the above considered preparation noncontextuality.

Consider a Bell experiment in which Alice remotely creates the assemblage $\{\sigma_{a|x}\}$ on Bob's side. We assume that if Bob applies any measurement M to the assemblage, the outcome statistics is measurement noncontextual. This means that

$$p(b|a,x,\mathbf{M})p(a|x) = p(a|x)\sum_{\lambda} p(\lambda|a,x)p(b|M,\lambda). \tag{4.23}$$

Here, the noncontextuality is embodied in the fact that Bob's response function does not depend on the measurement procedure $\mathbf{M}$ but only on the POVM elements M. Hence, this relation must hold for all measurement procedures compatible with M. Notice that we have not yet imposed preparation noncontextuality. We shall do so shortly.

Consider now the response of Bob. For every given λ, it is a map from the space of all POVMs to a probability distribution. All such maps were characterised by the works of Gleason [75] and Busch [76]. The Gleason-Busch theorem upholds that we can write

$$p(b|M,\lambda) = \mathrm{tr}\left(\rho_\lambda M_b\right), \tag{4.24}$$

for some quantum state ρ_λ. Notably, the quantum state is unique for each map. Equiped with this, the noncontextual distribution becomes

$$p(b|a,x,\mathbf{M})p(a|x) = p(a|x)\sum_{\lambda} p(\lambda|a,x)\,\mathrm{tr}\left(\rho_\lambda M_b\right). \tag{4.25}$$

From Bayes' rule, we have that

$$p(a|x)p(\lambda|a,x) = p(a|x,\lambda)p(\lambda|x). \tag{4.26}$$

Now, let us invoke preparation noncontextuality. Since we work in a Bell experiment, Alice's preparations (a,x) are already obeying the no-signaling principle. This means that her average preparations (for each x) are operationally equivalent. Therefore a preparation noncontextual model would impose their ontological equivalence, i.e. $p(\lambda|x) = p(\lambda)$. Putting it togther, we have that the outcome statistics in a fully noncontextual model reads

$$p(b|a,x,\mathbf{M})p(a|x) = \sum_{\lambda} p(\lambda)p(a|x,\lambda)\,\mathrm{tr}\left(\rho_\lambda M_b\right). \tag{4.27}$$

If we view $p(\lambda)\rho_\lambda$ as a subnormalised state, then this is precisely the outcome statistics obtained from applying a measurement to an unsteerable assemblage. Conversely, let us assume that the assemblage prepared by Alice for Bob is unsteerable. Then, it is immediately clear that the outcome statistics obeys a noncontextual model. Therefore, we conclude that

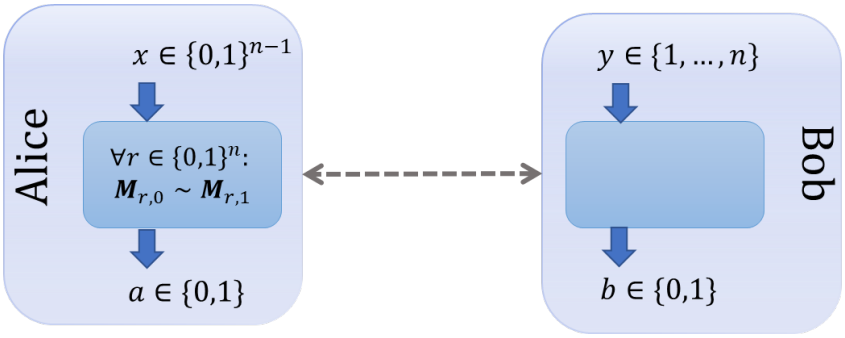

Figure 4.2: Contextuality test in a Bell scenario. Alice and Bob receive 2^{n-1} and n inputs respectively and produce binary outputs. Alice's measurements are constrained by operational equivalences; for every $r \in \{0,1\}^n$ with at least two instances of "1" the mixture $\mathbf{M}_{r,i} = \frac{1}{2^{n-1}} \sum_{a,x|r\cdot\bar{x}=i} M_{a|x}$ is independent of i.

> *An assemblage is unsteerable if and only if its statistics admits a preparation and measurement noncontextual model for all measurements.*

Thus, every steerable state is useful for a test of contextuality while every unsteerable state can never be used for the same purpose.

Application: tight steering inequalities for n projective qubit measurements

We apply the one-to-one link between contextuality and steering to tackle a concrete question in quantum steering. Imagine that Alice and Bob are given the two-qubit Werner state

$$\rho_v = v|\psi^-\rangle\langle\psi^-| + \frac{1-v}{4}\mathbb{I}, \tag{4.28}$$

where $|\psi^-\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}$ is the singlet state. Alice applies n projective qubit measurements on her system and thereby prepares an assemblage on Bob's side consisting of 2^n states. What is the smallest value of v for which there exists such measurements for Alice that she can steer Bob's state? We take the route through contextuality to answer this question.

In the previous section, we showed a family of preparation noncontextuality inequalities based on the quantum random access code. If we focus on the case of $d = 2$, we can re-cast them as noncontextuality inequalities in a Bell experiment. The scenario is illustrated in Figure 4.2. Instead of Alice preparing 2^n possible states, she receives one of 2^{n-1} possible inputs (x) and measures her part of the shared state, obtaining a binary outcome a. Bob acts in the same way as in the original scenario: he receives $y \in [n]$ and outputs a binary b. Preparation noncontextuality is imposed on the operational equivalences associated to the no-signaling condition. The operational equivalences on which we imposed measurement noncontextuality stem from those used in the preparation

noncontextuality inequality (4.14). Specifically, define the bit string $r \in \{0,1\}^{n-1}$. For every r with at least two instances of "1", we impose the operational equivalence

$$\sum_{a,x|r\cdot\bar{x}=0} M_{a|x} = \sum_{a,x|r\cdot\bar{x}=1} M_{a|x}, \tag{4.29}$$

where we have defined $\bar{x} = (a, x \oplus a)$. Whenever r has an even number of "1s" then this constraint is trivially satisfied due to $M_{0|x} + M_{1|x} = \mathbb{I}$. Imposing preparation and measurement noncontextuality in this manner, a straightforward adaption of our inequality (4.14) to the Bell scenario[1] gives

$$\mathcal{A}_n = \frac{1}{n2^{n-1}} \sum_{x,y} p(a \oplus b = \bar{x}_y | x, y) \overset{\text{NC}}{\leq} \frac{n+1}{2n}. \tag{4.30}$$

Our previous results guarantee that every violation of this noncontextuality inequality certifies that the shared state is steerable. However, it is not necessarily the case that every steerable state violates this inequality. The reason is that although we have shown that every steerable state does violate some noncontextuality inequality, it could very well not be our specific noncontextuality inequality. However, as it turns out, our inequality is useful towards finding the critical v for revealing the steerability of the state ρ_v using n projective measurements. We have used SDPs in see-saw to optimise over the measurements of Alice and Bob respectively in order to find the critical v below which a violation no longer is possible. We find the following results

$$\begin{aligned} v_2 &= 0.7071 & v_3 &= 0.5774 & v_4 &= 0.5547 \\ v_5 &= 0.5422 & v_6 &= 0.5270 & v_7 &= 0.5234. \end{aligned} \tag{4.31}$$

Interestingly, Ref [77] considered the steerability of ρ_v under n projective measurements using methods not based on contextuality. The numbers we have obtained precisely coincide with those presented in Ref [77]. This is a strong indication that our noncontextuality inequalities also serve as tight steering inequalities.

4.3 Harvesting contextuality in multiple sequential experiments

The typical tests of quantum correlations, be it quantum communication complexity, quantum nonlocality or quantum contextuality, aim to generate strong correlations in scenarios in which a state undergoes a single measurement. By this, we mean that parties measure a state (or a part of it) and have no regard for what happens to the state after the measurement. Practically speaking,

[1] Recall that the Bell scenario and the Bell correlation witness can be found by reversing the map from Bell inequalities to CCPs discussed in chapter 2.

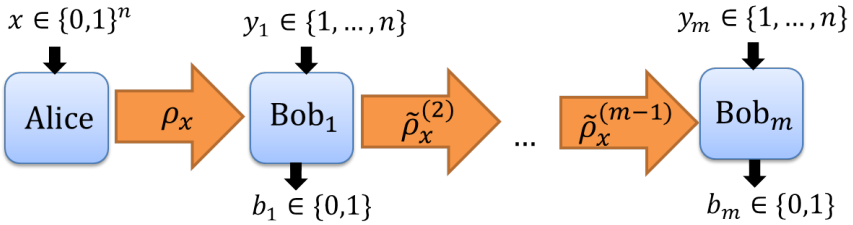

Figure 4.3: Sequential tests of preparation contextuality from a single ensemble. Alice implements a quantum random access code with each Bob in a a way that the quantum output of one Bob is the quantum input of the next Bob.

they typically even demolish the physical carrier during the measurement and the state ceases to exist. Here, we ask a very different question: *is it possible for many independent observers to each harvest the contextuality enabled by a single ensemble of quantum states?* That is, can we use the same quantum system to create many sequential proofs of contextuality separately obtained by observers conducting independent experiments? We will focus on the preparation noncontextuality inequalities based on the quantum random access code discussed in the first section of this chapter (specifically the inequality (4.14)). We choose to focus on the case of $d = 2$ in which the preparation noncontextuality inequality reduces to those derived in Ref [71]:

$$\mathcal{A}_n = \frac{1}{n2^n} \sum_{x,y} p(b = x_y|x, y) \overset{\text{PNC}}{\leq} \frac{n+1}{2n}, \tag{4.32}$$

subject to the data hiding constraint (in a quantum model)

$$\forall r : \quad \sum_{x \cdot r = 0} \rho_x = \sum_{x \cdot r = 1} \rho_x, \tag{4.33}$$

for each string $r \in \{0,1\}^n$ with at least two instances of "1". Using these contextuality tests, we investigate the possibility of *sharing contextuality* between many independent observers.

Consider the following scenario. Alice receives a random bit string $x \in \{0,1\}^n$ and prepares the state ρ_x which she communicates to a receiver Bob_1. Bob_1 receives a random input $y \in [n]$ and produces a binary output $b_1 \in \{0,1\}$. The post-measurement state is relayed to another observer Bob_2 who receives a random input $y_2 \in [n]$ and outputs $b_2 \in \{0,1\}$. Again, the post-measurement state is relayed to an analogous observer Bob_3 etc. This continues until the state is received by the final observer Bob_m who receives a random input $y_m \in [n]$ and produces the output $b_m \in \{0,1\}$, see Figure 4.3. We are interested in whether Alice can demonstrate contextuality *independently* with each of the receivers $\text{Bob}_1, \ldots, \text{Bob}_m$. Clearly, if she demonstrates a strong violation of the noncontextuality inequality with Bob_1, this will lead to weaker correlations with Bob_2 etc. Each

time a party measures the state to create correlations, the coherence of Alice's initial ensemble decreases - at some point it must not longer enable contextuality. Is there a critical number of observers that can harvest the contextuality of Alice's state ensemble? We prove that for any given number of sequential observers, there exists an ensemble of preparations for Alice such that they all can share its contextuality. In order to arrive to this result, we benefit from first deriving the maximal quantum value of $\mathcal{A}_n$ in a standard (non-sequential) scenario.

Maximal quantum contextuality

The problem of deriving the largest quantum violation of (4.32) was solved in Ref [78]. Here, we summarise the optimal measurements of Bob and the optimal preparations of Alice.

We can characterise Bob's measurement through an observable $G^T_{n,y}$. The easiest way to present it is through a recursive definition. The recursion starts from $G_{2,1} = \sigma_x$, $G_{2,2} = \sigma_y$, and $G_{3,1} = \sigma_x$, $G_{3,2} = \sigma_y$ and $G_{3,3} = \sigma_z$. All the subsequent observables are obtained from

$$\begin{aligned} &n \text{ even:} && G_{n,k} = G_{n-1,k} \otimes \sigma_x \quad \forall k \in \{1, \ldots, n-1\}, \\ &n \text{ odd:} && G_{n,k} = G_{n-2,k} \otimes \sigma_x \quad \forall k \in \{1, \ldots, n-2\} \end{aligned} \tag{4.34}$$

with $G_{n,n} = \mathbb{I} \otimes \sigma_y$ if $n > 3$ is even, and $G_{n,n} = \mathbb{I} \otimes \sigma_z$ and $G_{n,n-1} = \mathbb{I} \otimes \sigma_y$ if $n > 3$ is odd.

Alice has 2^n states. We write each of them as a state of $\lfloor n/2 \rfloor$ qubits. For simplicitly, we can phrase it as the subsystems of a collection of $\lfloor n/2 \rfloor$ entangled state as obtained after a local operation:

$$\rho_x = \mathrm{tr}_{\mathrm{A}} \left[(\mathbb{I} + A_x) \otimes \mathbb{I} \left(|\phi_{\max}\rangle\langle\phi_{\max}| \right)^{\otimes \lfloor n/2 \rfloor} \right], \tag{4.35}$$

where

$$A_x = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (-1)^{x_i} G_{n,i} \tag{4.36}$$

$$|\phi_{\max}\rangle = \frac{1}{\sqrt{2}} \left(|0,0\rangle + |1,1\rangle \right) \tag{4.37}$$

and the trace is taken over the first system in every entangled pair. Interestingly, these states are mixed and their purity is only $\mathrm{tr}\left(\rho_x^2\right) = 1/2$. Notice, however, that no entanglement is present in the actual problem - this is merely a convenient way of writing Alice's states. It is not straightforward to compute the probabilities $p(b = x_y|x, y)$; one finds that they are all equal. The maximal quantum violation becomes

$$\mathcal{A}_n = \frac{1}{2} \left(1 + \frac{1}{\sqrt{n}} \right), \tag{4.38}$$

for all $n \geq 2$.

Sequential decoherence of the ensemble

In the sequential scenario, the actions of each Bob corresponds to a quantum instrument that is represented by Kraus operators $\{K_{y_k}^{b_k}\}$. The state received by Bob_k is then determined by Alice's input x and the list of all previous inputs and outputs, $(y_1, \ldots, y_{k-1})$ and $(b_1, \ldots, b_{k-1})$. Importantly, Bob_k does not know the previous inputs and outputs because the observers are assumed to be independent. Therefore, although in each instance Bob_k receives a state that depends on all these parameters, the effective state that is visibile to him is their average. This effective state $\tilde{\rho_x}^{(k)}$ determines his correlations and it is written recursively as

$$\tilde{\rho}_x^{(k)} = \frac{1}{n} \sum_{y_{k-1}, b_{k-1}} K_{y_{k-1}}^{b_{k-1}} \tilde{\rho}_x^{(k-1)} (K_{y_{k-1}}^{b_{k-1}})^\dagger, \tag{4.39}$$

where we define $\tilde{\rho}_x^{(1)} = \rho_x$. Then, the probability distribution between Alice and Bob_k, that is used to evaluate the k'th independent contextuality test, is given by

$$p(b_k|x, y_k) = \text{tr}\left(\tilde{\rho}_x^{(k)} (K_{y_k}^{b_k})^\dagger K_{y_k}^{b_k}\right). \tag{4.40}$$

Let us now specify a simple and useful strategy for the encoding of Alice's states and the instruments of all the Bobs. It will draw substantially on the optimal quantum strategy for the previously discussed standard (non-sequential) scenario. We let Alice prepare the same states as those that were optimal in the non-sequential scenario. The Bobs, however, cannot perform the same measurements as in that scenario since the first Bob would end up harvesting all the contextuality for himself, leaving only noncontextual correlations for all the subsequent Bobs. Instead, we define their observables G_{n,y_k}^T as corresponding to the projectors $\Pi_{n,y}^b = \frac{\mathbb{I}+(-1)^b G_{n,y}^T}{2}$. We will let each Bob perform a weaker variant of this measurement, in which the projectors are modified to

$$\frac{\mathbb{I} + (-1)^b \eta_k G_{n,y}^T}{2} \tag{4.41}$$

for some sharpness parameter $\eta_k \in [0, 1]$ that can be taken differently for each Bob. These POVM elements are realised with the following choice of Kraus operators:

$$K_{y_k}^{b_k} = \sqrt{\frac{1+\eta_k}{2}} \Pi_{n,y_k}^{b_k} + \sqrt{\frac{1-\eta_k}{2}} \Pi_{n,y_k}^{\bar{b}_k}, \tag{4.42}$$

where the bar-sign denotes a bit-flip. Hence, this family of quantum strategies is parameterised by the collection of sharpness parameters $\{\eta_k\}_{k=1}^m$.

What is the effective state $\tilde{\rho}_x^{(k)}$ as seen by Bob_k for a given choice of sharpness parameters $\{\eta_k\}$? The answer can be recursively phrased as follows. The effective state received by Bob_k takes the form

$$\tilde{\rho}_x^{(k)} = v_k \rho_x + (1 - v_k)\, \rho_{\text{mix}}, \tag{4.43}$$

where $v_k \in [0,1]$ is a visibility parameter and ρ_{mix} is the maximally mixed state in the relevant dimension. The larger this visibility parameter is, the lesser has the state been decohered from the original preparation ρ_x of Alice. Hence, a larger v_k enables stronger correlations. It turns out that the precise value of v_k can be recursively written as

$$v_k = v_{k-1} f_{k-1} = \prod_{j=1}^{k-1} f_j, \tag{4.44}$$

where $v_1 = 1$ by definition, and the "quality factor" f_k of the measurement of Bob_k is defined from the sharpness η_k as

$$f_k = \frac{1 + (n-1)\sqrt{1-\eta_k^2}}{n}. \tag{4.45}$$

Before we proceed any further, let us give a proof for these claims.

Proof of state decoherence lemma

This section reproduces the proof originally given in Ref [70]. The state prepared by Alice is defined as

$$\rho_x = \text{tr}_{\text{A}} \left[(\mathbb{I} + A_x) \otimes \mathbb{I} \left(|\phi_{\text{max}}\rangle\langle\phi_{\text{max}}| \right)^{\otimes \lfloor n/2 \rfloor} \right], \tag{4.46}$$

where $(|\phi_{\text{max}}\rangle\langle\phi_{\text{max}}|)^{\otimes \lfloor n/2 \rfloor}$ is $\lfloor n/2 \rfloor$ copies of the two-qubit maximally entangled state, and the partial trace is taken over all the first qubits in each pair. Consider that the sequence of Bobs, labelled by $\{1, 2, ..., m-1\}$, apply measurements of intermediate sharpness to the state above, each denoted by $\eta_k = \sin\theta_k$. We proceed to prove that the average state $\tilde{\rho}_x^{(m)}$ received by Bob_m will be of the form

$$\tilde{\rho}_x^{(m)} = \text{tr}_{\text{A}} \left[(\mathbb{I} + v_m A_x) \otimes \mathbb{I} \left(|\phi_{\text{max}}\rangle\langle\phi_{\text{max}}| \right)^{\otimes \lfloor n/2 \rfloor} \right], \tag{4.47}$$

where v_m (the "visibility" of the state) is given by

$$v_m = v_{m-1} f_{m-1} = \prod_{j=1}^{m-1} f_j, \tag{4.48}$$

$$\text{where} \quad f_j = \frac{1 + (n-1)\cos\theta_j}{n}. \tag{4.49}$$

We call f_j the "quality factor" of the measurement of the j^{th} Bob. The visibility of the first Bob is $v_1 = 1$, since he possesses the undisturbed state received directly from Alice.

The proof is inductive. For the first Bob, the statement holds trivially. Consider that it holds true for $m-1$ Bobs, so that the average state $\tilde{\rho}_x^{(m)}$ received by Bob_m is given by (4.47). Then using

the Kraus operators, the average state $\tilde{\rho}_x^{(m+1)}$ (averaging over all Bob$_m$'s possible and equiprobable inputs, and with no knowledge of his outcome), is given by

$$\tilde{\rho}_x^{(m+1)} = \frac{1}{n}\sum_{y,b} K_y^b \tilde{\rho}_x^{(m)} (K_y^b)^\dagger = \frac{1}{n}\sum_{y,b} \mathrm{tr_A}\left[(\mathbb{I} + v_m A_x)\otimes K_y^b \ (|\phi_{\max}\rangle\langle\phi_{\max}|)^{\otimes\lfloor n/2\rfloor} \ \mathbb{I}\otimes (K_y^b)^\dagger\right], \tag{4.50}$$

where the Kraus operators are acting on the part of the Hilbert space complementary to that being traced out. First, using the property of the maximally entangled state that $(\mathbb{I}\otimes O)\,|\phi_{\max}\rangle\langle\phi_{\max}|\left(\mathbb{I}\otimes O^\dagger\right) = \left(O^T\otimes\mathbb{I}\right)|\phi_{\max}\rangle\langle\phi_{\max}|\,(O^*\otimes\mathbb{I})$, and then using the cyclicity of the trace, we obtain

$$\tilde{\rho}_x^{(m+1)} = \frac{1}{n}\sum_{y,b} \mathrm{tr_A}\left[(\mathbb{I} + v_m A_x)\,(K_y^b)^T\otimes\mathbb{I} \ (|\phi_{\max}\rangle\langle\phi_{\max}|)^{\otimes\lfloor n/2\rfloor} \ (K_y^b)^{\dagger T}\otimes\mathbb{I}\right] \tag{4.51}$$

$$= \frac{1}{n}\sum_{y,b} \mathrm{tr_A}\left[(K_y^b)^{\dagger T}(\mathbb{I} + v_m A_x)\,(K_y^b)^T\otimes\mathbb{I} \ (|\phi_{\max}\rangle\langle\phi_{\max}|)^{\otimes\lfloor n/2\rfloor}\right]. \tag{4.52}$$

Splitting the above into the sum of the two terms from the $(\mathbb{I} + v_m A_x)$, the contribution of the identity part is

$$\frac{1}{n}\sum_{y,b} \mathrm{tr_A}\left[(K_y^b)^{\dagger T}(K_y^b)^T\otimes\mathbb{I} \ (|\phi_{\max}\rangle\langle\phi_{\max}|)^{\otimes\lfloor n/2\rfloor}\right] = \frac{1}{n}\sum_{y} \mathrm{tr_A}\left[\mathbb{I}\otimes\mathbb{I} \ (|\phi_{\max}\rangle\langle\phi_{\max}|)^{\otimes\lfloor n/2\rfloor}\right]$$
$$= \mathrm{tr_A}\left[(|\phi_{\max}\rangle\langle\phi_{\max}|)^{\otimes\lfloor n/2\rfloor}\right], \tag{4.53}$$

where we have used that the K_y^b are Hermitian and that measurements are complete i.e., $\sum_b (K_y^b)^{\dagger T}(K_y^b)^T = \mathbb{I}$. For the term involving A_x, we calculate the sum using the Kraus operators, denoting by $\eta_m = \sin\theta_m$ the sharpness of the measurement of Bob m,

$$K_y^b = \sqrt{\frac{1+\eta_m}{2}}\Pi_{n,y}^b + \sqrt{\frac{1-\eta_m}{2}}\Pi_{n,y}^{\bar{b}} = \left(\frac{\cos\frac{\theta_m}{2}\mathbb{I} + (-1)^b \sin\frac{\theta_m}{2} G_{n,y}}{\sqrt{2}}\right), \tag{4.54}$$

which results in

$$\frac{1}{n}\sum_{y,b}(K_y^b)^{\dagger T} A_x (K_y^b)^T = \frac{1}{n}\sum_{y,b}\left(\frac{\cos\frac{\theta_m}{2}\mathbb{I} + (-1)^b \sin\frac{\theta_m}{2} G_{n,y}}{\sqrt{2}}\right) A_x \left(\frac{\cos\frac{\theta_m}{2}\mathbb{I} + (-1)^b \sin\frac{\theta_m}{2} G_{n,y}}{\sqrt{2}}\right)$$
$$= \frac{1}{2n}\sum_{y,b}\cos^2\left(\frac{\theta_m}{2}\right) A_x + (-1)^b \cos\left(\frac{\theta_m}{2}\right)\sin\left(\frac{\theta_m}{2}\right)\{G_{n,y}, A_x\}$$
$$+ \sin^2\left(\frac{\theta_m}{2}\right) G_{n,y} A_x G_{n,y} \tag{4.55}$$
$$= \frac{1}{n}\sum_{y}\left(\frac{1+\cos\theta_m}{2}\right) A_x + \left(\frac{1-\cos\theta_m}{2}\right) G_{n,y} A_x G_{n,y}$$
$$= \left(\frac{1+\cos\theta_m}{2}\right) A_x + \left(\frac{1-\cos\theta_m}{2}\right)\frac{1}{n}\sum_{y} G_{n,y} A_x G_{n,y}. \tag{4.56}$$

We may now use the expansion $A_x = \frac{1}{\sqrt{n}} \sum_i (-1)^{x_i} G_{n,i}$, together with the anti-commutation relation

$$\{G_{n,j}, G_{n,k}\} = 2\delta_{j,k}\mathbb{I} \tag{4.57}$$

to simplify the calculation into

$$\begin{aligned}
\frac{1}{n}\sum_y G_{n,y} A_x G_{n,y} &= \frac{1}{\sqrt{n}}\sum_i (-1)^{x_i} \frac{1}{n}\sum_y G_{n,y} G_{n,i} G_{n,y} \\
&= \frac{1}{\sqrt{n}}\sum_i (-1)^{x_i} \frac{1}{n}\sum_y (2\delta_{i,y} G_{n,y} - G_{n,i}) \\
&= \frac{1}{\sqrt{n}}\sum_i (-1)^{x_i} \frac{1}{n}(2-n) G_{n,i} \\
&= \frac{2-n}{n} A_x.
\end{aligned} \tag{4.58}$$

Inserting this into Eq. (4.56), we obtain

$$\frac{1}{n}\sum_{y,b} (K_y^b)^{\dagger T} A_x (K_y^b)^T = f_m A_x, \tag{4.59}$$

$$\text{where} \quad f_m = \left(\frac{1 + (n-1)\cos\theta_m}{n}\right) = \left(\frac{1 + (n-1)\sqrt{1-\eta^2}}{n}\right), \tag{4.60}$$

is the quality factor of the measurement of Bob_m. Putting it all together, we find the final expression for the average state after Bob_m's measurement:

$$\tilde{\rho}_x^{(m+1)} = \text{tr}_\text{A}\left[(|\phi_\text{max}\rangle\langle\phi_\text{max}|)^{\otimes \lfloor n/2 \rfloor}\right] + \text{tr}_\text{A}\left[v_m f_m A_x \otimes \mathbb{I}\ (|\phi_\text{max}\rangle\langle\phi_\text{max}|)^{\otimes \lfloor n/2 \rfloor}\right] \tag{4.61}$$

$$= \text{tr}_\text{A}\left[(\mathbb{I} + v_m f_m A_x) \otimes \mathbb{I}\ (|\phi_\text{max}\rangle\langle\phi_\text{max}|)^{\otimes \lfloor n/2 \rfloor}\right]. \tag{4.62}$$

This is the desired relation stated in the previous section.

Sharing contextuality between any number of observers

We are now ready to show that for any given number of observers m, there exists a choice of n so that the outlined quantum strategy allows all observers to independently violate the preparation noncontextuality inequality. It is now a straightforward calculation to show that the correlations seen by Bob_k correspond to the witness value

$$\mathcal{A}_n^{(k)} = \frac{1}{2}\left(1 + \frac{v_k \eta_k}{\sqrt{n}}\right). \tag{4.63}$$

Comparing this to the preparation noncontextual bound, we have a proof of contextuality whenever

$$\eta_k > \frac{1}{v_k\sqrt{n}}. \tag{4.64}$$

Starting from $v_1 = 1$, we can recursively calculate the critical values of η_k and v_k. Then, it is sufficient for each Bob to given an infinitesimal perturbation to the positive of η_k in order to witness preparation contextuality. Let us write the sharpness parameters as $\eta_k = \sin\theta_k$ for some angle $\theta_k \in [0, \frac{\pi}{2}]$. The condition for contextuality therefore reads

$$\sin\theta_k > \frac{1}{v_k\sqrt{n}}. \tag{4.65}$$

Keeping this in mind, let us also put the following (entirely trivial) lower bound on the quality factor

$$f_k = \frac{1 + (n-1)\cos\theta_k}{n} \geq \cos\theta_k. \tag{4.66}$$

Putting these two together, we can square both sides to obtain

$$f_k^2 \geq \cos^2\theta_k = 1 - \sin^2\theta_k \geq 1 - \frac{1}{v_k^2 n}. \tag{4.67}$$

Similarly, we can bound the visibility of the subsequent Bob by using that $v_{k+1} = v_k f_k$ and the analogous squaring gives

$$v_{k+1}^2 = v_k^2 f_k^2 \geq v_k^2\left(1 - \frac{1}{v_k^2 n}\right). \tag{4.68}$$

In this manner, we can bound the difference in visibility in any two sequential steps as

$$v_k^2 - v_{k+1}^2 \leq \frac{1}{n}. \tag{4.69}$$

Since we start from $v_1 = 1$, it implies that

$$v_{k+1}^2 \geq 1 - \frac{k}{n}. \tag{4.70}$$

Choosing $k + 1 = n$, we therefore obtain that

$$v_n \geq \frac{1}{\sqrt{n}}. \tag{4.71}$$

In conclusion, the visibility of the n'th Bob must obey precisely the condition required for violating the preparation noncontextuality inequality. Thus, in order to have m sequential violations, we must choose at least $n = m$ in our protocol. Furthermore, we can obtain proofs of contextuality that are robust to noise (for each and every Bob) simply by increasing n to be suitably larger than m.

Finally, let us mention that the analogous sharing of nonlocality has been investigated in Ref [79]. It was found that only two sequential observers can share the nonlocality of one part of a singlet state in a CHSH Bell experiment subject to uniformly random inputs. This stands in sharp contrast to the unbounded sequence of contextuality shown here.

5

Quantum nonlocality

In this chapter, we depart entirely from correlation experiments featuring communication and focus on quantum correlations arising in Bell-type experiments. In the first section, we present Bell inequalities that are tailored to one of the most celebrated, elegant and useful discrete Hilbert space structures encountered in quantum theory, namely mutually unbiased bases (MUBs) [80]. We then introduce an operational definition of mutual unbiasedness and prove that a maximal quantum violation certifies such measurements [81]. In the third section, we depart from standard Bell experiments and instead consider quantum nonlocality in networks. We discuss a method in which standard Bell inequalities can be systematically mapped to network Bell inequalities in such a way that their local bounds and quantum violations are preserved [82]. In the final section, we focus on the simplest network Bell scenario known as the *bilocality scenario*. A commonly held suspicion is that the known examples of Bell inequalities for the bilocality scenario can be traced back to quantum violations of standard Bell inequalities. We investigate quantum correlations in the bilocality scenario that have no resemblance to standard Bell nonlocality and present a network Bell inequality that is inspired by quantum correlations rather than standard Bell inequalities [83].

5.1 Bell inequalities and mutual unbiasedness

Mutually unbiased bases (MUBs) are some of the most intriguing and well-researched discrete structures in quantum theory. They also appear in many quantum information protocols such as for random number generation and quantum key distribution. Here, we place these structures in the context of Bell experiments where we investigate both their ability to maximally violate Bell inequalities and the possibility of certifying them in a device-independent framework. Towards this, we first define the concept. Let $\{|e_j\rangle\}_{j=1}^{d}$ and $\{|f_k\rangle\}_{k=1}^{d}$ be two orthonormal bases of the

d-dimensional Hilbert space $\mathbb{C}^d$. The two bases are mutually unbiased if

$$|\langle e_j|f_k\rangle|^2 = \frac{1}{d} \tag{5.1}$$

for all j and k, where the constant on the right-hand-side is fixed by normalisation. A simple interpretation for MUBs is that if a state is prepared in an eigenstate of one basis and measured in the other basis, all outcomes are equally probable. Indeed, such bases always exist in any Hilbert space dimension; for instance one can consider the computational basis and its Fourier transform. Moreover, in dimension two that would amount to the Pauli matrices σ_z and σ_x.

In what follows, we first derive a family of Bell inequalities whose maximal quantum violation is obtained with pairs of MUBs. Then, we show that the maximal quantum violation also can be used for device-independent certification of a operational MUBs (which we introduce and define). Finally, we apply the Bell inequalities for device-independent quantum key distribution. Notably, the original work on which this chapter is based also develops Bell inequalities, device-independent certification and random number generation based on symmetric informationally complete measurements [81]. However, these results are omitted here in the interest of space and accessibility.

Bell inequalities for MUBs

Consider a bipartite Bell experiment involving Alice and Bob. Alice randomly receives one of d^2 possible inputs denoted by the two-dit stringt $x \equiv x_1x_2 \in [d]^2$ and is asked to return a ternary output $a \in \{1, 2, \perp\}$. Bob receives a binary input $y \in [2]$ and returns a d-valued output $b \in [d]$. The Bell scenario is illustrated in Figure 5.1.

We let Alice and Bob play a game in which they collectively win or lose points depending on their collective inputs and outputs. Their aim is to cooperate to play the game as well as possible. Alice can singlehandedly decide that no points will won or lost in a round by deciding to output $a = \perp$. Instead, if she wants to play for points, she outputs $a \in \{1, 2\}$. In these cases, points are only won or lost if Bob's output satisfies the relation $b = x_y$. Specifically, a point is won if $a = y$ but lost if $a = \bar{y}$ (the bar-sign denotes a bit flip). The total score in the game can therefore be written as

$$\mathcal{R}_d^{\text{MUB}} \equiv \sum_{x,y} p(a = y, b = x_y|x, y) - p(a = \bar{y}, b = x_y|x, y). \tag{5.2}$$

Now, it may seem as if the outcome $a = \perp$ is artificial. Why would Alice ever decide to output $a = \perp$? Presently, there is no good answer. In order to make the $\perp$ outcome relevant, we must modify the game. We explain this modification through a simple intuition.

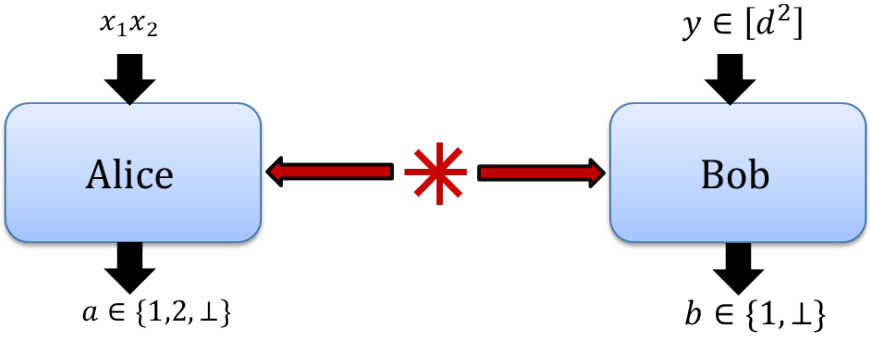

Figure 5.1: Bell experiment tailored to pairs of d-dimensional MUBs. Alice receives a d^2-valued input and returns a ternary output. Bob receives a binary input and returns a d-valued output.

Let us imagine that Alice and Bob share the maximally entangled state of local dimension d,

$$|\psi_d^{\max}\rangle = \frac{1}{\sqrt{d}} \sum_{k=1}^{d} |k,k\rangle. \tag{5.3}$$

We want that Alice remotely prepares Bob in a pure state whenever she decides that a round should play for points (i.e. when $a \in \{1,2\}$). The reason for this is that pure states allow for the strong correlations that will be needed for Alice and Bob to obtain high scores. In order to prepare pure states for Bob, we need that Alice's measurement operators correspond to rank-one projectors. When $d > 2$ this cannot fill the full space of $\mathbb{C}^d$. Therefore, we consider that the outcome $\perp$ corresponds to a projection on the complementary $d-2$ dimensional Hilbert space. Moreover, since the local states of Alice and Bob are both the maximally mixed state, it would then follow that $p(a=1|x) = p(a=2|x) = 1/d$. Therefore, we wish to motivate Alice to employ a strategy in which she only outputs $a=1$ and $a=2$ with probability $1/d$ respectively. Consequently, we want her to render the round moot by outputing $a = \perp$ with probability $p(a = \perp |x) = 1 - 2/d$. This is clearly different from what is expected to be a good strategy to earn a large value of $\mathcal{R}_d^{\text{MUB}}$: evidently it appears better for Alice to always output $a \in \{1,2\}$. In order to tailor the game to our goal, we introduce a tunable penalty. Specifically, Alice will lose γ_d points whenever she decides to output $a \in \{1,2\}$. It stands to reason that by tuning the penalty, the optimal rate of outputing $a \in \{1,2\}$ should change accordingly. Outputing $a \in \{1,2\}$ contributes to $\mathcal{R}_d^{\text{MUB}}$ but simultaneously costs γ_d points. Our total score now reads

$$\mathcal{S}_d^{\text{MUB}} \equiv \mathcal{R}_d^{\text{MUB}} - \gamma_d \sum_x \left(p(a=1|x) + p(a=2|x)\right). \tag{5.4}$$

This serves as our Bell functional.

We are now faced with some natural questions. Firstly, how do we choose the penalty γ_d to establish the connection to MUBs? Secondly, what is the local bound and the quantum bound of the resulting Bell functional? We can answer these questions in one swoop. For this purpose, we define an observable for Alice as $A_x = A_{1|x} - A_{2|x}$. We also define $\{P_b\}$ to be Bob's POVM for $y = 1$ and $\{Q_b\}$ to be his POVM for $y = 2$. The shared state can without loss of generality be assumed pure. Then, we can write

$$\mathcal{R}_d^{\mathrm{MUB}} = \sum_x \langle\psi| A_x \otimes (P_{x_1} - Q_{x_2}) |\psi\rangle. \tag{5.5}$$

If we apply the Cauchy-Schwarz inequality to each term in the sum, we obtain an upper bound on the form

$$\mathcal{R}_d^{\mathrm{MUB}} \leq \sum_x \sqrt{\langle\psi| A_{1|x} + A_{2|x} |\psi\rangle} \sqrt{\langle\psi| (P_{x_1} - Q_{x_2})^2 |\psi\rangle}, \tag{5.6}$$

where we used that Alice's measurements w.l.g. can be assumed as projective, for which it holds that $(A_x)^2 = A_{1|x} + A_{2|x}$. Next, we use the elementary inequality

$$\sum_i \sqrt{s_i}\sqrt{r_i} \leq \sqrt{\sum_i s_i} \sqrt{\sum_i r_i} \tag{5.7}$$

which is valid for non-negative s_i and r_i and which gives equality if and only if $s_i = k r_i$ for some proportionality constant k. Our new upper bound then reads

$$\mathcal{R}_d^{\mathrm{MUB}} \leq \sqrt{\sum_x \langle\psi| A_{1|x} + A_{2|x} |\psi\rangle} \sqrt{\sum_x \langle\psi| (P_{x_1} - Q_{x_2})^2 |\psi\rangle}. \tag{5.8}$$

To proceed further, we note the following useful property of projective measurements:

$$\sum_x (P_{x_1} - Q_{x_2})^2 = \sum_x P_{x_1} + Q_{x_2} - \{P_{x_1}, Q_{x_2}\} = 2\,(d-1)\,\mathbb{I}. \tag{5.9}$$

This allows us to write

$$\mathcal{S}_d^{\mathrm{MUB}} \leq \sqrt{2(d-1)t} - \gamma_d t, \tag{5.10}$$

where we have defined $t = \sum_x \langle\psi| A_{1|x} + A_{2|x} |\psi\rangle$. We can now find the maximal value of the right-hand-side by differentiating w.r.t. t and setting the derivative equal to zero. That equation gives

$$t = \frac{d-1}{2\gamma_d^2}. \tag{5.11}$$

Now, we recall that our goal is to have $p(a = 1|x) = p(a = 2|x) = 1/d$. This would mean that we aim for $t = 2d$. Imposing this constraint determines the penalty to be

$$\gamma_d = \frac{1}{2}\sqrt{\frac{d-1}{d}}. \tag{5.12}$$

With this fixed penalty, our Bell functional is fully defined. Then, we can turn to the local bound and the quantum bound. We have essentially already done the main work for both of them. In the quantum case, we just insert our value of t into (5.10) which immediately gives the bound

$$\mathcal{S}_d^{\mathrm{MUB}} \overset{\mathrm{Q}}{\leq} \sqrt{d(d-1)}. \tag{5.13}$$

As we will prove soon, this bound is tight.

The local bound requires a only a simple calculation which we include for completeness. Recall that it suffices to optimise over deterministic strategies to determine the bound. Moreover, once the strategy of Bob is fixed, finding the optimal strategy of Alice is easy. If Bob outputs $b = u_1$ for $y = 1$ and $b = u_2$ for $y = 2$, the Bell functional becomes

$$\mathcal{S}_d^{\mathrm{MUB}} = \sum_x \left(\delta_{x_1,u_1} - \delta_{x_2,u_2}\right) \langle\psi|A_{1|x} - A_{2|x}|\psi\rangle - \frac{1}{2}\sqrt{\frac{d-1}{d}} \sum_x \langle\psi|A_{1|x} + A_{2|x}|\psi\rangle. \tag{5.14}$$

We define $R_\pm = \{x \in [d]^2 | \delta_{x_1,u_1} - \delta_{x_2,u_2} = \pm 1\}$ and $R_0 = [d]^2 \setminus (R_+ \cup R_-)$. By expanding the above expression for $\mathcal{S}_d^{\mathrm{MUB}}$ into the separate sums over R_+, R_- and R_0 it becomes clear that the optimal choice of Alice is to choose $A_{1|x} = \mathbb{I}$ and $A_{2|x} = A_{\perp|x} = 0$ (always output $a = 1$) when $x \in R_+$, choose $A_{2|x} = \mathbb{I}$ and $A_{1|x} = A_{\perp|x} = 0$ (always output $a = 2$) when $x \in R_-$ and choose $A_{\perp|x} = \mathbb{I}$ and $A_{1|x} = A_{2|x} = 0$ (always output $a = \perp$) when $x \in R_0$. Since $|R_\pm| = d - 1$, this leads to the local bound

$$\mathcal{S}_d^{\mathrm{MUB}} \overset{\mathrm{local}}{\leq} 2\,(d-1)\left(1 - \frac{1}{2}\sqrt{\frac{d-1}{d}}\right). \tag{5.15}$$

Let us now return to the MUBs. We have attempted to tailor the Bell inequality in such a way that the maximal quantum value is achieved with any pair of MUBs of dimension d. It can be confirmed that this is indeed the case. To this end, we let Alice and Bob share the maximally entangled state $|\psi_d^{\max}\rangle$ and let Bob's measurements $\{P_b\}$ and $\{Q_b\}$ be a pair of MUBs, i.e. we define $P_b = |\phi_b\rangle\langle\phi_b|$ and $Q_b = |\varphi_b\rangle\langle\varphi_b|$ such that $|\langle\phi_b|\varphi_b\rangle|^2 = 1/d$. Finally, we choose Alice's observables as

$$A_x = \sqrt{\frac{d}{d-1}}(P_{x_1} - Q_{x_2})^{\mathrm{T}}, \tag{5.16}$$

where the pre-factor is fixed by normalisation. Then, we have that

$$\mathcal{S}_d^{\mathrm{MUB}} = \sum_x \langle\psi|A_x \otimes (P_{x_1} - Q_{x_2}) - \gamma_d \left(A_{1|x} + A_{2|x}\right) \otimes \mathbb{I}|\psi\rangle \tag{5.17}$$

$$= \sqrt{\frac{d}{d-1}} \sum_x \langle\psi|\mathbb{I} \otimes (P_{x_1} - Q_{x_2})^2 |\psi\rangle - \sum_x \langle\psi|\gamma_d \left(A_{1|x} + A_{2|x}\right) \otimes \mathbb{I}|\psi\rangle \tag{5.18}$$

$$= \sqrt{\frac{d}{d-1}} 2(d-1) - 2d\gamma_d = \sqrt{d(d-1)}. \tag{5.19}$$

Indeed, our quantum strategy based on MUBs can saturate the maximal quantum violation of the Bell inequality.

Certification of mutually unbiased measurements

Any pair of MUBs can saturate the maximal quantum violation of the Bell inequalities. However, what can be deduced from the observation of correlations that achieve the maximal violation? The answer to this question will lead us to introduce the concept of *mutually unbiased measurements* and prove that such structures can be certified by the quantum correlations.

Previously, we used the Cauchy-Schwarz inequality to obtain a tight upper bound on the quantum violation of the Bell inequality. Saturating the Cauchy-Schwarz inequalities implies the relation

$$A_x|\psi\rangle = \mu_x \left(P_{x_1} - Q_{x_2}\right)|\psi\rangle, \tag{5.20}$$

where μ_x is some constant. We need to determine the value of μ_x. To do this, we left-multiply by $\langle\psi|(P_{x_1} - Q_{x_2})$ which tells us that μ_x is real and non-negative. The above relation implies that

$$\langle\psi|(A_x)^2|\psi\rangle = \mu_x^2\langle\psi|\left(P_{x_1} - Q_{x_2}\right)^2|\psi\rangle. \tag{5.21}$$

We have already seen that the left-hand-side is constant in x for the maximal violation, and therefore we must have $\mu_1 = \mu_2 = \ldots = \mu_{d^2} \equiv \mu$. Then, it follows from our previous analysis that

$$\mu = \sqrt{\frac{d}{d-1}}. \tag{5.22}$$

This is indeed the normalisation constant used in our specific quantum strategy based on pairs of MUBs. We conclude that the following cross-relation must hold

$$A_x|\psi\rangle = \sqrt{\frac{d}{d-1}}\left(P_{x_1} - Q_{x_2}\right)|\psi\rangle. \tag{5.23}$$

This relation can be manipulated into the certification statement that we aim to derive concerning only Bob's pair of measurements.

The spectrum of Alice's observables is $\{+1, -1, 0\}$. Therefore, it must be that $(A_x)^3 = A_x$. Using this, we can eliminate Alice's operators from the cross-relation and write it as

$$\sqrt{\frac{d}{d-1}}\left(P_{x_1} - Q_{x_2}\right)|\psi\rangle = \left(\frac{d}{d-1}\right)^{3/2}\left(P_{x_1} - Q_{x_2}\right)^3|\psi\rangle. \tag{5.24}$$

Now, we trace-out Alice's system. If we assume that Bob's local state is full-rank, then we can right-multiply by its inverse to obtain a relation that only concerns Bob's operators:

$$P_{x_1} - Q_{x_2} = \frac{d}{d-1}\left(P_{x_1} - Q_{x_2}\right)^3, \tag{5.25}$$

Recalling that Bob's POVMs can be assumed projective, expanding the right-hand-side leads to the simple relation

$$P_{x_1} - Q_{x_2} = d\left(P_{x_1}Q_{x_2}P_{x_1} - Q_{x_2}P_{x_1}Q_{x_2}\right). \tag{5.26}$$

If we sum over either x_1 or x_2, we obtain the two final relations

$$P_{x_1} = dP_{x_1}Q_{x_2}P_{x_1} \qquad \text{and} \qquad Q_{x_2} = dQ_{x_2}P_{x_1}Q_{x_2}. \tag{5.27}$$

We conclude that a maximal quantum violation of the Bell inequalities implies that Bob's two measurements must obey these relations. However, what do these relations actually mean? It is clear that if Bob's two POVMs are MUBs, then relations (5.27) are satisfied, as expected. What is the physics of this certification?

As it turns out, the relation between Bob's measurements that we have certified corresponds to a natural development of the concept of mutual unbiasedness. In the standard definition, which we gave in the beginning of this chapter, MUBs are native to a specific Hilbert space dimension. If we assume the Hilbert space dimension, we could describe them entirely in terms of operational quantities (probabilities): if every state for which the outcome a d-outcome measurement is deterministic yields a uniform distribution when measured in another basis, then the two measurementes are MUBs. However, Hilbert space dimension is not an observable quantity. Can one determine the notion of mutual unbiasedness on fully operational grounds? This leads us to introduce the concept of *mutually unbiased measurements* (MUMs). We say that two n-outcome measurements $\{P_a\}_{a=1}^n$ and $\{Q_b\}_{b=1}^n$ are mutually unbiased if they are projective and the following implications hold:

$$\begin{aligned} \langle\psi|P_a|\psi\rangle = 1 \Rightarrow \langle\psi|Q_b|\psi\rangle = \frac{1}{n} \\ \langle\psi|Q_b|\psi\rangle = 1 \Rightarrow \langle\psi|P_a|\psi\rangle = \frac{1}{n}, \end{aligned} \tag{5.28}$$

for all a and b. That is, two projective measurements are mutually unbiased if the eigenvectors of one measurement give rise to a uniform outcome distribution for the other measurement. Importantly, this concept is strictly different from the concept of MUBs since all MUBs indeed are MUMs but there also exists MUMs that are not MUBs (see Ref [81] for an exploration of MUBs versus MUMs).

It is relevant to address the question of how one can characterise pairs of MUMs. Interestingly, one can show that a necessary and sufficient condition for two measurements to be MUMs is precisely the condition that we have derived as an implication of the maximal quantum violation of the Bell inequalities, namely Eq (5.27). This means that the obtained certification in fact has a natural physical interpretation in terms of MUMs. For the sake of completeness, we reproduce the proof of this statement originally given in Ref [81].

Let us first assume that the relations (5.27) hold. By summing over the middle index, one finds that both measurements are projective. Moreover, if $|\psi\rangle$ is an eigenvector of P_a, then $\langle\psi|Q_b|\psi\rangle = \langle\psi|P_aQ_bP_a|\psi\rangle = \frac{1}{n}\langle\psi|P_a|\psi\rangle = \frac{1}{n}$. By symmetry, the analogous property holds if $|\psi\rangle$ is an eigenvector of Q_b. Conversely, let us show that MUMs must satisfy the relation (5.27). Since $\sum_a P_a = \mathbb{I}$ we can choose an orthonormal basis of the Hilbert space composed only of the eigenvectors of the measurement operators. Let $\{|e_j^a\rangle\}_{a,j}$ be an orthonormal basis, where $a \in [n]$ tells us which projector the eigenvector corresponds to and j labels the eigenvectors within a fixed projector (if P_a has finite rank, then $j \in [\operatorname{tr} P_a]$, otherwise $j \in \mathbb{N}$). By construction for such a basis we have $P_a|e_j^{a'}\rangle = \delta_{aa'}|e_j^a\rangle$. To show that $P_a = nP_aQ_bP_a$ it suffices to show that the two operators have the same coefficients in this basis. Since

$$\langle e_j^{a'}|nP_aQ_bP_a|e_k^{a''}\rangle = n\delta_{aa'}\delta_{aa''}\langle e_j^a|Q_b|e_k^a\rangle, \tag{5.29}$$

$$\langle e_j^{a'}|P_a|e_k^{a''}\rangle = \delta_{aa'}\delta_{aa''}\delta_{jk} \tag{5.30}$$

it suffices to show that $n\langle e_j^a|Q_b|e_k^a\rangle = \delta_{jk}$. For $j = k$ this is a direct consequence of the definition in Eq. (5.28). To prove the other case, define $|\phi_\theta\rangle = \left(|e_j^a\rangle + \mathrm{e}^{\mathrm{i}\theta}|e_k^a\rangle\right)/\sqrt{2}$, for $\theta \in [0, 2\pi)$. Since $P_a|\phi_\theta\rangle = |\phi_\theta\rangle$, we have $\langle\phi_\theta|Q_b|\phi_\theta\rangle = 1/n$. Writing this equality out gives

$$\frac{1}{n} = \frac{1}{2}\left(\frac{2}{n} + \mathrm{e}^{\mathrm{i}\theta}\langle e_j^a|Q_b|e_k^a\rangle + \mathrm{e}^{-\mathrm{i}\theta}\langle e_k^a|Q_b|e_j^a\rangle\right). \tag{5.31}$$

Choosing $\theta = 0$ implies that the real part of $\langle e_j^a|Q_b|e_k^a\rangle$ vanishes, while $\theta = \pi/2$ implies that the imaginary part vanishes. Proving the relation $Q_b = nQ_bP_aQ_b$ proceeds in an analogous fashion.

Application in quantum key distribution

To conclude our discussion of Bell inequalities tailored to MUBs, we outline their application towards the task of quantum key distribution (QKD). QKD is the task of distributing shared data between two separate parties in such a way that an eavesdropper cannot gain knowledge of the data. Many of the most well-known protocols for QKD rely on the implementation of MUBs; examples include the BB84 protocol and the six state protocol. However, these are protocols in which both Alice and Bob are assumed to operate characterised quantum devices. In contrast, quantum nonlocality enables device-independent protocols for QKD in which neither Alice nor Bob are assumed to operate a priori characterised devices. In these low-trust schemes, both the key rate and its security is derived directly from the Bell inequality violation. Since our Bell inequalities offer a device-independent approach to MUBs, it appears natural to consider their application for device-independent QKD.

In our QKD protocol, we give Alice an additional input x^* that is perfectly correlated with the first setting ($y = 1$) of Bob. From these settings, the parties can extract a shared key. Note that the correlations will be perfect if the shared state is maximally entangled. Alice and Bob run the Bell experiment. In most rounds, they both choose the settings from which they extract the key while in a few rounds they implement the actual Bell experiment (i.e. measurement settings x and y). From the fewer rounds they can estimate the Bell functional $\mathcal{S}_d^{\text{MUB}}$ which serves as their security parameter from which they also decide the key rate K. It is well-known that the key rate can be lower bounded via the following formula [84]:

$$K \geq -\log\left(P_g^\beta\right) - H(B_{y=1}|A_{x^*}), \tag{5.32}$$

where $H(B_{y=1}|A_{x^*})$ is the conditional Shannon entropy and P_g^β is the guessing probability of an eavesdropper. Specifically, the guessing probability is the largest probability of guessing the outcome of, say, Bob that is still compatible with the observed value of the Bell functional $\beta = \mathcal{S}_d^{\text{MUB}}$. We write it as

$$P_g^\beta \equiv \sup\left\{\sum_{c=1}^{d}\langle\psi_{\text{ABE}}|\mathbb{I}\otimes P_c\otimes E_c|\psi_{\text{ABE}}\rangle\right\}, \tag{5.33}$$

where we have denoted the POVM of the eavesdropper by $\{E_c\}$ and allowed Alice, Bob and the eavesdropper to share some arbitrary tripartite state $|\psi_{\text{ABE}}\rangle$. Thus, we need to evaluate this guessing probability as a function of the Bell functional. This is in general difficult. However, interestingly, the case of a maximal quantum violation of the Bell inequality admits a simple solution.

In order to present the analysis, we must first add two important pieces of knowledge. We already know that if we observe the maximal quantum value $\mathcal{S}_d^{\text{MUB}} = \sqrt{d(d-1)}$, then the measurements of Bob must be a pair of MUMs. Importantly, the maximal violation actually contains even more information about the experiment. The following two additional pieces of information were derived in Ref [81]

- From the shared state, it is possible to extract a maximally entangled d-dimensional state by means of local extraction maps.

- There is only a single probability distribution $p(a, b|x, y)$ that is compatible with the maximal quantum violation.

Equipped with these results, we first conclude that a maximal $\mathcal{S}_d^{\text{MUB}}$ implies perfect correlations between Alice and Bob (for the settings x^* and $y = 1$) which means $H(B_{y=1}|A_{x^*}) = 0$. Moreover,

it must be that $P_g = 1/d$. To prove the latter, let us denote the state after the eavesdropper's measurement by

$$\rho_{\mathrm{AB}}^{(c)} = \frac{1}{p(c)} \mathrm{tr}_C \left[(\mathbb{I} \otimes \mathbb{I} \otimes E_c) |\psi_{\mathrm{ABE}}\rangle\langle\psi_{\mathrm{ABE}}| \right], \tag{5.34}$$

where $p(c)$ is the probability of the outcome c. Alice and Bob are performing their Bell inequality test on the state $\rho_{\mathrm{AB}}^{(c)}$. Therefore, since they are unaware of c, they will obtain different distributions $p_c(a, b|x, y)$. Since we have assumed the maximal value of $\mathcal{S}_d^{\mathrm{MUB}}$ and we know that only a single distribution is compatible with it, it must follow that $p_c(a, b|x, y)$ is independent of c. Then, recalling that the marginals of Bob are uniformly random, it follows that

$$P_g = \sum_{c=1}^{d} p(c) p_c(b = c|y = 1) = \frac{1}{d}. \tag{5.35}$$

We conclude that the maximal value of $\mathcal{S}_d^{\mathrm{MUB}}$ implies the largest possible key rate in any experiment in which the key is extracted from a d-outcome measurement, namely

$$K = \log d. \tag{5.36}$$

Matters become more complicated when we deal with values of $\mathcal{S}_d^{\mathrm{MUB}}$ that are not maximal. In these cases, one must first choose a noise model in which to evaluate $H(B_{y=1}|A_{x^*})$ and then bound P_g^{β}. Regarding the former, we consider the simple case of depolarising noise, namely

$$\rho_v = v|\psi_d^{\max}\rangle\langle\psi_d^{\max}| + \frac{1 - v}{d^2}\mathbb{I}. \tag{5.37}$$

The visibility v relates to the observed Bell functional as

$$v = \frac{1}{2}\left(1 + \frac{\mathcal{S}_d^{\mathrm{MUB}}}{\sqrt{d(d-1)}}\right). \tag{5.38}$$

This enables a straightforward calculation of the conditional entropy $H(B_{y=1}|A_{x^*})$ as a function of v.

An exact computation of P_g^{β} is challenging, we resort to placing an upper bound on this quantity which then translates into a lower bound on the key rate. One way of achieving this is through a semidefinite relaxation. Importantly, the precise computation of the guessing probability is not an SDP since it involves three unknown sets of operators that multiply each other; the state, Bob's measurement and the eavesdropper's measurement. Nevertheless, an upper bound on P_g can be obtained via SDP since it is known that the set of quantum correlations can be approximated from above by such means. We implement the (tripartite) hierarchy of quantum correlations [85]. We focus on the case of $d = 3$ for which we implement the associated SDPs using an intermediate

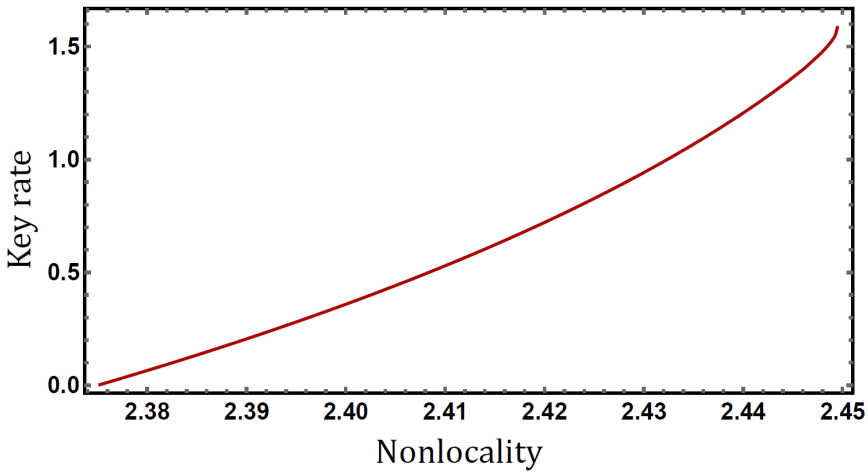

Figure 5.2: Certified key rate versus Bell functional $\mathcal{S}_d^{\mathrm{MUB}}$ for the case of $d = 3$. The result is obtained via symmetrised semidefinite relaxations of the quantum set of correlations.

level corresponding to 532 monomials. The corresponding SDP has 15617 variables which makes the evaluation computationally demanding. Fortunately, we can employ the SDP symmetrisation techniques discussed in earlier chapters for the NV hierarchy. It is worth pointing out that those methods readily also apply to Bell scenarios (they are arguably easier to apply in Bell scenarios). To perform a symmetrisation of the SDP, we need to find some symmetries. Importantly, these symmetries must preserve both the objective (the guessing probability) and its constraints (the Bell functional). A simple symmetry that does the job is the following: let π be any permutation of three elements, then a symmetry is to permute Bob's output $b \to \pi(b)$ conditioned on $y = 1$ while simultaneously permuting the outcome of the eavesdropper $c \to \pi(c)$ and the first input trit of Alice $x_1 \to \pi(x_1)$. A straightforward inspection of the guessing probability and the Bell functional shows that the transformations indeed keep both of them invariant. Exploiting these symmetries, the number of variables in the moment matrix reduces to 2823 which allows for a fast SDP evaluation[1]. The resulting lower bound on the key rate as a function of the Bell functional is displayed in Figure 5.2. We see that for a maximal value, we indeed have $K = \log 3$. Then, the key rate falls of and reaches one bit around $\mathcal{S}_d^{\mathrm{MUB}} \approx 2.432$ and eventually reaches zero around $\mathcal{S}_d^{\mathrm{MUB}} \approx 2.375$. Notice that the local bound $\mathcal{S}_d^{\mathrm{MUB}} \approx 2.367$ is not far from the zero key rate threshold.

[1] In this specific case, the evaluation time was reduced by a factor 60 and the memory use was reduced by a factor 55.

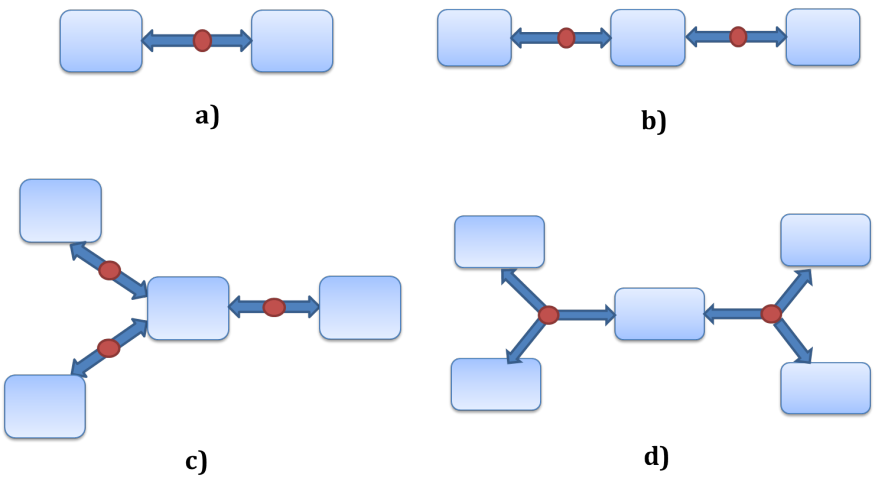

Figure 5.3: Examples of network Bell experiments. a) is a standard Bell experiment. b) is the simplest network Bell scenario and is commonly called the bilocality scenario. c) is an example of a star-network (and so is b). d) is a more elaborate network.

5.2 Bell inequalities for star-networks

Recent years have seen an increased interest in generalisations of the standard Bell experiment to network Bell experiments [86]. Whereas the former features a number of observers that share a physical system emitted by a source, the latter features many independent sources that emit states whose subsystems are distributed between the observers. A few examples are illustrated in Figure 5.3. As a consequence, two observers may receive subsystems of different states, and thereby a priori lack the ability of establishing correlations. Such *network Bell scenarios* are interesting from the point of view of quantum theory since they enable processes such as entanglement swapping to be used by some observers to distribute entanglement to initially independent observers throughout the network. Naturally, such procedures do not have a classical analogy. However, network Bell experiments also pose technical challenges since the assumption of independent sources means that the notion of local correlations now features several independent local hidden variables. Consequently, the set of local correlations is no longer convex. This makes it more difficult to derive network Bell inequalities for witnessing quantum correlations.

Here, we focus on a class of networks that are shaped like a star. A star-network features N independent sources and $N+1$ observers. We call one observer "the node". The node receives an input y and produces an outcome b. This party has a special status since he independently shares a bipartite state with each of the other N "edge observers". Thus, the edge observers are only connected to the node and therefore independent of each other. Edge observer number k receives

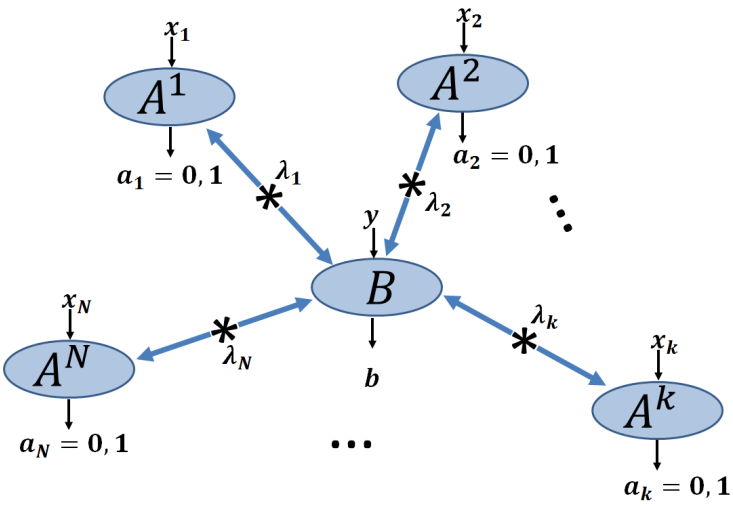

Figure 5.4: Bell experiment in a star-network. The node observer B independently shares a bipartite state with each of the N edge observers A^k. In a local model, each source is associated to an independent local hidden variable.

an input x_k and produces an output a_k. We illustrate the star-network in Figure 5.4.

In a local model, each of the N sources is associated to an independent local hidden variable λ_k. The response function of each of the k'th edge observer is therefore determined by the input x_k and λ_k. For the node, having access to all the hidden variables, the output is determined by y and $\vec{\lambda} = (\lambda_1, \ldots, \lambda_N)$. A local model therefore reads

$$p(a_1 \ldots a_N b|x_1 \ldots x_N y) = \int \left(\prod_{k=1}^{N} d\lambda_k q_k(\lambda_k) p(a_k|x_k, \lambda_k) \right) p(b|y, \vec{\lambda}). \tag{5.39}$$

Notice that if we choose $N = 1$, this reduces to a standard Bell experiment featuring only one source. However, whenever $N \geq 2$ the geometry of the set of local correlations becomes non-convex due to the products $q_1(\lambda_1) \ldots q_N(\lambda_N)$. Therefore, Bell inequalities in the network scenario no longer corresponds linear inequalities constituting faces of some local polytope. Instead, they must be nonlinear in order to capture features of this more complicated set of local correlations.

From Bell inequalities to star inequalities

We present a method for systematically constructing Bell inequalities for the star-network (star inequalities). The main idea is to start from the familiar Bell inequalities for standard Bell experiments and then transform them into star inequalities. Therefore, let us begin with considering a

general correlation Bell inequality of the form

$$\mathcal{S}_M^{bs} \equiv \sum_{x=1}^{n_A} \sum_{y=1}^{n_B} M_{yx} \langle A_x^{bs} B_y^{bs} \rangle \leq C, \tag{5.40}$$

where M_{yx} is a real number, n_A is the number of settings for Alice, n_B is the number of settings for Bob and (A_x^{bs}, B_y^{bs}) are the observables of Alice and Bob that take values ± 1. By specifying the matrix M, one can determine the local bound C. It is favourable to re-write the Bell inequality on the following simple form

$$\mathcal{S}_M^{bs} = \sum_{y=1}^{n_B} \left(\sum_{x=1}^{n_A} M_{yx} A_x^{bs} \right) B_y^{bs} = \sum_{y=1}^{n_B} \hat{A}_y^{bs} B_y^{bs}, \tag{5.41}$$

where we have defined $\hat{A}_y^{bs} = \sum_{x=1}^{n_A} M_{yx} A_x^{bs}$. The local bound can now be calculated as follows;

$$C = \max_{A_1 \ldots A_{n_A} \in \{\pm 1\}} \sum_{y=1}^{n_B} \left| \sum_{x=1}^{n_A} M_{yx} A_x^{bs} \right| = \max_{A_1 \ldots A_{n_A} \in \{\pm 1\}} \sum_{y=1}^{n_B} \left| \hat{A}_y^{bs} \right|. \tag{5.42}$$

In order to transform these Bell inequalities into star inequalities, consider the following ansatz. Consider binary outputs for all edge parties but not necessarily for the node. Then, we define the global expectation value for all parties involved in the network by

$$\langle A_{x_1}^1 \ldots A_{x_N}^N B_i \rangle = \sum_{a_1 \ldots a_N = 0,1} \sum_b (-1)^{a_1 + \ldots + a_N + f_i(b)} \times p(a_1 \ldots a_N b | x_1 \ldots x_N y_i), \tag{5.43}$$

where $i = 1, \ldots, n_B$ and $f_i(b)$ is some binary-valued function for each i. These expectation values are applied to define the correlation quantities

$$I_i = \sum_{x_1 \ldots x_N = 1}^{n_A} M_{i x_1} \ldots M_{i x_N} \langle A_{x_1}^1 \ldots A_{x_N}^N B_i \rangle = \langle \hat{A}_i^1 \ldots \hat{A}_i^N B_i \rangle.$$

Using the quantities $\{I_i\}_{i=1}^{n_B}$, we define the following Bell functional for the network,

$$\mathcal{S}_{M,\{f_i\}}^{net} \equiv \sum_{i=1}^{n_B} |I_i|^{1/N}. \tag{5.44}$$

Notice that for each choice of coefficient matrix M and each choice of functions $\{f_i\}_i$, we have a different Bell functional. Our aim is to prove a bound on $\mathcal{S}_{M,\{f_i\}}^{net}$ valid for all local correlations.

Assuming a local model, the correlation quantities can be written on the form

$$I_i = \int \left[\prod_{k=1}^{N} d\lambda_k q_k(\lambda_k) \hat{A}_i^k(\lambda_k) \right] B_i(\vec{\lambda}). \tag{5.45}$$

Taking the absolute value of both sides and using that $|B_i| \leq 1$, we obtain the bound

$$|I_i| \leq \prod_{k=1}^{N} \int d\lambda_k q_k(\lambda_k) \left| \hat{A}_i^k(\lambda_k) \right|. \tag{5.46}$$

Inserting this into the chosen Bell functional and applying the Hölder inequality, we find that

$$\mathcal{S}^{net}_{\{M\},\{f_i\}} = \sum_{i=1}^{n_B} |I_i|^{1/N} \leq \left[\prod_{k=1}^{N} \int d\lambda_k q_k(\lambda_k) \sum_{i=1}^{n_B} \left| \hat{A}_i^k(\lambda_k) \right| \right]^{1/N}. \tag{5.47}$$

The sum inside the square bracket is precisely the expression that previously appeared when we established the local bound of the standard Bell inequality. We can therefore use that result to conclude that

$$\mathcal{S}^{net}_{M,\{f_i\}} \equiv \sum_{i=1}^{n_B} |I_i|^{1/N} \overset{\text{local}}{\leq} C, \tag{5.48}$$

is a star inequality.

In this manner, the original Bell inequality can be mapped to a star inequality in which the local bound is the same as in the original Bell inequality. It is interesting to point out that some previous examples of star inequalities [87, 88] can be reproduced by the above by choosing the standard Bell inequality to be the CHSH inequality.

In fact, it is not only the local bound that carries over from the Bell inequality to the star inequality. Quantum strategies that yield a violation of the Bell inequality can also be used to obtain violations of the star inequality. Give the node the same observables as Bob has in the Bell inequality test, but not presented as an N-fold tesor; $\mathcal{B} = (\mathcal{B}_y^{bs})^{\otimes n}$. Similarly, give each of the edge observers the same measurements as Alice has in the Bell inequality test; $\mathcal{A}_x^1 = \ldots, = \mathcal{A}_x^N$. Finally, the state used in the Bell inequality test is distributed in each of the N sources in the star-network. This somewhat trivial quantum strategy now gives a factorisation of the expectation values:

$$\langle A_{x_1}^1 \ldots A_{x_N}^N B_y \rangle_{\rho^{\otimes N}} = \langle \mathcal{A}_{x_1}^1 \ldots \mathcal{A}_{x_N}^N \mathcal{B}_y \rangle_{\rho^{\otimes N}} = \langle \mathcal{A}_{x_1}^{bs} \mathcal{B}_y^{bs} \rangle_\rho \ldots \langle \mathcal{A}_{x_N}^{bs} \mathcal{B}_y^{bs} \rangle_\rho. \tag{5.49}$$

Consequently, the correlation quantities take on the simple form

$$I_i = \left(\sum_{x=1}^{n_B} M_{ix} \langle \mathcal{A}_x^{bs} \mathcal{B}_i^{bs} \rangle_\rho \right)^N = \left(\langle \hat{\mathcal{A}}_i^{bs} \mathcal{B}_i^{bs} \rangle_\rho \right)^N, \tag{5.50}$$

which when inserted into the Bell functional evidently leads to

$$\mathcal{S}^{net} = \mathcal{S}^{bs}. \tag{5.51}$$

Therefore, every state that can violate the standard Bell inequality can also violate the star inequality in a quantum model.

5.3 Quantum violations of bilocality

It is interesting that Bell inequalities systematically can be mapped to star-inequalities and that the most well-known network Bell inequalities are instances of this map. Nevertheless, the fact that such a map is possible strongly suggests that the star-inequalities nor their quantum violations will conceptually differ from that encountered in standard Bell experiments. Indeed, we saw that the quantum violations simply amounted to conducting several coordinated violations of standard Bell inequalities. Notably, there are star-inequalities that employ an entangled collective measurement in the node [87, 88]. Commonly, this measurement is a Bell State Measurement, projecting the qubits into a basis of GHZ-like states. However, such Bell State Measurements turn out to effectively correspond to simultaneous measurements of $\sigma_x \otimes \sigma_x$ and $\sigma_z \otimes \sigma_z$, which therefore allows a simulation of Bell inequality violations by coordinated separate violations of the CHSH inequality (again, via the map from the previous section) [89]. However, one of the main motivations for studying network Bell nonlocality is quantum correlations that in some sense are genuine to a network and do not bear any resemblance to standard Bell nonlocality.

Here, we consider the simplest star-network which corresponds to $N = 2$. This is known as *the bilocality scenario* and correlations attainable in local models are said to be *bilocal*. We investigate quantum correlations arising in this network from the node implementing a measurement that is different from the Bell State Measurement. This leads us to quantum correlations and bilocality inequalities that admit no apparent resemblance to standard Bell nonlocality.

The bilocality scenario

In a bilocality scenario, we have three observers named Alice, Bob and Charlie. Alice and Bob share a state while Bob and Charlie share an independent state. Hence, this can be viewed as the simplest star-network featuring only two edge observers (Alice and Charlie). Since we eventually will choose the states to be entangled qubits, it appears natural to consider a setting in which Alice and Charlie have binary outcomes $a = \pm 1$ and $c = \pm 1$ respectively and Bob has four possible outcomes $b \in [4]$. The choice for Bob is based on him jointly measuring his two independent particles, which will correspond to a rank-one projective measurement on $\mathbb{C}^2 \otimes \mathbb{C}^2$. For simplicity, we let Bob only perform a single measurement (he has no input). Then, the bilocal model for the correlations $p(a, b, c|x, z)$ reads

$$p_{\text{biloc}}(a, b, c|x, z) = \sum_{\alpha,\gamma} q_1(\lambda_1) q_2(\lambda_2) p(a|x, \lambda_1) p(c|z, \lambda_2) p(b|\lambda_1, \lambda_2). \tag{5.52}$$

Any distribution that does not admit this form is said to be non-bilocal.

The elegant joint measurement in the bilocality scenario

How should we choose Bob's four-outcome measurement on his two qubits? Our intention is to depart from the well-studied Bell State Measurement and instead introduce a different measurement that gives rise to qualitatively different quantum correlations. To this end, we consider to so-called elegant joint measurement (EJM) [90]. This measurement features four rank-one projectors that are equally entangled $\{|\Phi_j\rangle\}_{j=1}^4$. They can be written as

$$|\Phi_j\rangle = \frac{\sqrt{3}+1}{2\sqrt{2}}|\vec{m}_j, -\vec{m}_j\rangle + \frac{\sqrt{3}-1}{2\sqrt{2}}|-\vec{m}_j, \vec{m}_j\rangle, \tag{5.53}$$

where the qubit states $|\vec{m}_j\rangle$ correspond to Bloch vectors that form a tetrahedron on the Bloch sphere. It is easier to write the Bloch vectors in cylindrical coordinates

$$\vec{m}_j = \left(\sqrt{1-\eta_j^2}\cos\phi_j, \sqrt{1-\eta_j^2}\sin\phi_j, \eta_j\right). \tag{5.54}$$

We choose them to correspond to the four Bloch vectors that form a tetrahedron. Namely,

$$\begin{aligned} \vec{m}_1 &= \frac{1}{\sqrt{3}}(1,1,1) & \vec{m}_2 &= \frac{1}{\sqrt{3}}(1,-1,-1) \\ \vec{m}_3 &= \frac{1}{\sqrt{3}}(-1,1,-1) & \vec{m}_4 &= \frac{1}{\sqrt{3}}(-1,-1,1), \end{aligned} \tag{5.55}$$

from which we obtain the state vector

$$|\vec{m}_j\rangle = \sqrt{\frac{1-\eta_j}{2}}e^{i\phi_j/2}|0\rangle + \sqrt{\frac{1+\eta_j}{2}}e^{-i\phi_j/2}|1\rangle. \tag{5.56}$$

Then, $|-\vec{m}_j\rangle$ denotes the unique state orthogonal to $|\vec{m}_j\rangle$. The EJM has the following nice property. If we trace-out the first system, then the four local states of the other system correspond to a tetrahedron on the Bloch sphere (now shrunk by a factor 1/2). Similarly, if we trace-out the second system, then the four local states of the first system correspond to a mirror image of that tetrahedron.

In view of this property, it appears most reasonable to allow Alice and Charlie to perform three measurements (see Figure 5.5). That is, because the tetrahedron spans all three dimensions of the Bloch sphere, Alice and Charlie ought to benefit from not ignoring a subspace of the Bloch sphere. Hence, we choose $x, z \in [3]$. The most natural choices for the measurements of Alice and Charlie are to let them be σ_x, σ_y and σ_z, because these are symmetric with respect to the two local tetrahedra. Choosing the shared states to be maximally entangled, i.e. $|\psi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$, we can calculate that the joint expectation value of Alice and Charlie conditioned on the outcome of Bob is

$$E_b^{\mathrm{AC}}(x,z) \equiv \sum_{a,c} ac \cdot p(a,c|b,x,z) = -\frac{\tilde{m}_{b,x}\tilde{m}_{b,z}}{2}(1-\delta_{x,z}), \tag{5.57}$$

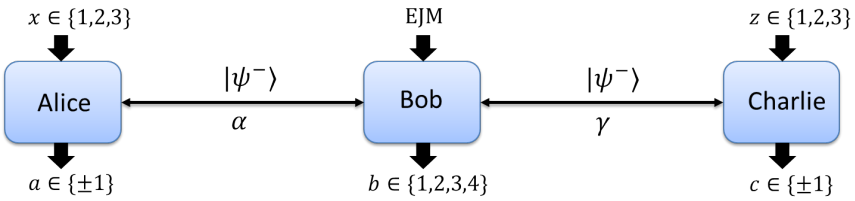

Figure 5.5: Bilocality scenario based on the elegant joint measurement. Bob performs the EJM while Alice and Charlie perform one of three measurements with binary outcomes. In quantum (classical) implementation, the network features two independent singlet states (hidden variables).

where $\tilde{m}_{j,k}$ is the sign of the k'th component in $\vec{m}_j$. Similarly, we can compute the conditional single party expectation values to be

$$E_b^{\mathrm{A}}(x) = \frac{1}{2}\tilde{m}_{b,x} \qquad\qquad E_b^{\mathrm{C}}(z) = -\frac{1}{2}\tilde{m}_{b,z}. \tag{5.58}$$

Using these expectation values, we can write the full distribution as

$$p_{\mathrm{Q}}(a,b,c|x,z) = \frac{1}{16}\left(1 + aE_b^{\mathrm{A}}(x) + cE_b^{\mathrm{C}}(z) + acE_b^{\mathrm{AC}}(x,z)\right). \tag{5.59}$$

A relevant generalisation of this strategy is that in which each of the states are subjected to depolarising noise. Then the state becomes

$$\rho_i = V_i|\psi^+\rangle\langle\psi^+| + \frac{1-V_i}{4}\mathbb{I}, \tag{5.60}$$

for $i \in \{1,2\}$ where $V_i \in [0,1]$ denotes the visibility. In similar spirit, the probability distribution takes the same form as above now the expectation values are now modified to

$$\begin{aligned} E_b^{\mathrm{A}}(x) &= \frac{V_1}{2}\tilde{m}_{b,x}, \qquad E_b^{\mathrm{C}}(z) = -\frac{V_2}{2}\tilde{m}_{b,z}, \\ E_b^{\mathrm{AC}}(x,z) &= -\frac{V_1V_2}{2}\tilde{m}_{b,x}\tilde{m}_{b,z}\delta_{x\neq z}, \end{aligned} \tag{5.61}$$

We proceed to investigate whether the distribution p_{Q} is bilocal.

Bilocal simulation

To what extent can bilocal models simulate the quantum correlations? The answer is much less trivial than in standard Bell nonlocality. To tackle the problem, we can w.l.g. restrict the cardinality of the hidden variables to be eight since the number of ways of mapping Alice's (Charlie's) three inputs to her (his) binary output is eight. For simplicity, we rename the hidden variables as $\lambda_1 \equiv \alpha$ and $\lambda_2 \equiv \gamma$ (as in Figure 5.5). We define $\alpha, \gamma \in \{-4,\ldots,-1,1,\ldots,4\}$ and write their respective

distributions as $q_\alpha^{(\mathrm{A})}$ and $q_\gamma^{(\mathrm{C})}$. The list of binary answers to the three questions of Alice and Charlie can be written as a three-element tuple. We associate these tuples to the hidden variable by defining $w_j = \sqrt{3}\vec{m}_j$ and $w_{-j} = -\sqrt{3}\vec{m}_j$, where $\vec{m}_j$ are point to the tetrahedron as given in Eq (5.55). Now, we can write the bilocal model on the form

$$p_{\mathrm{biloc}}(a,b,c|x,z) = \sum_{\alpha,\gamma} q_\alpha^{(\mathrm{A})} q_\gamma^{(\mathrm{C})} \delta_{a,w_{\alpha x}} \delta_{c,w_{\gamma z}} p(b|\alpha,\gamma). \tag{5.62}$$

Note that the response of Bob, $p(b|\alpha,\gamma)$, is unconstrained.

We have employed a numeber of methods to decide whether the correlations generated via the EJM admit a bilocal model. Firstly, we define the distance between the quantum distribution and the bilocal distribution by

$$D \equiv \sqrt{\sum_{a,b,c,x,z} \left(p_{\mathrm{Q}}(a,b,c|x,z) - p_{\mathrm{biloc}}(a,b,c|x,z)\right)^2}. \tag{5.63}$$

Considering the case of $v \equiv v_1 = v_2$, one can use matlab's fmincon function to find the largest v for which one can up to numerical precision satisfy $D = 0$. Hence, this minimisation takes places over the response of Bob and the two hidden variable distributions. Repeating this optimisation many times reveals an interesting pattern. We find that the optimum occurs for hidden variable distributions that are uniform over each of the tetrahedral configurations present in the eight tuples $(\pm 1, \pm 1, \pm 1)$. In other words, all entries associated to $\alpha, \gamma = 1, \ldots, 4$ and $\alpha, \gamma = -1, \ldots, -4$ are respectively equal. This leads us to an ansatz of tetrahedral symmetry.

$$q^{(\mathrm{A})} = \frac{1}{4}\left[u, u, u, u, 1-u, 1-u, 1-u, 1-u\right] \tag{5.64}$$

$$q^{(\mathrm{C})} = \frac{1}{4}\left[r, r, r, r, 1-r, 1-r, 1-r, 1-r\right], \tag{5.65}$$

for some $u, r \in [0,1]$. Let us extend this tetrahedral symmetry also to Bob's response function. Specifically, let π be a permutation of $\{1,2,3,4\}$. We impose that

$$\forall \pi: \quad p\left(\pi(b)|\pi(\alpha),\pi(\gamma)\right) = p(b|\alpha,\gamma), \tag{5.66}$$

where $\pi(-\alpha) = -\pi(\alpha)$. This assumption appears natural given the tetrahedral symmetries present in the quantum probability distribution and in the hidden variable distribution. Exploiting it, one can reduce the existence of a bilocal model to only eight equations involving 24 variables. These variables are (u, r, V_1, V_2) and an additional 20 variables associated to the response of Bob. Considering the case of $V \equiv V_1 = V_2$, solving these equations leads us to

$$18V^2 + 6V + \sqrt{3}\sqrt{4V^2 + 12V - 9} = 19. \tag{5.67}$$

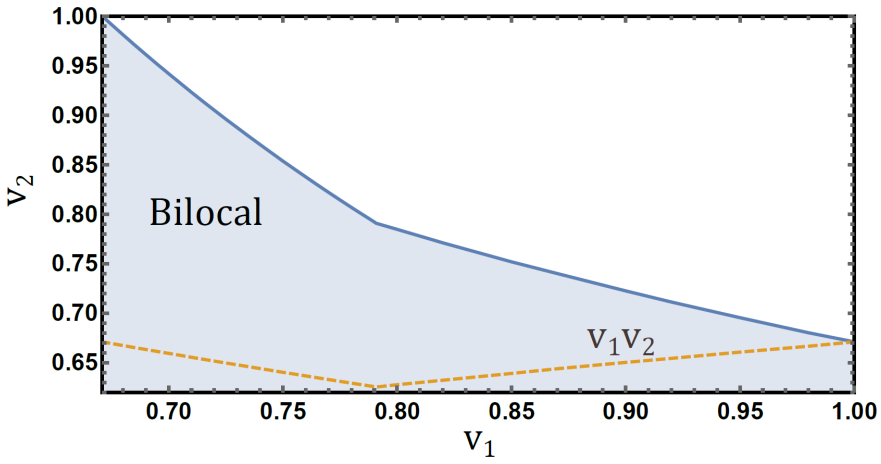

Figure 5.6: The set of bilocal correlations in the plane of visibilities (v_1, v_2). The dashed line represents the product of the visibilites on the boundary of the bilocal set (solid blue curve).

The critical visibility is therefore the solution to this equation. It has an analytical form, but it is unwieldy. Approximately, it reads

$$V \approx 79.09\%. \tag{5.68}$$

Thus, under the tetrahedral assumption, this is the best bilocal simulation possible. Furthermore, we can characterise general pairs (V_1, V_2) in this manner. Solving the symmetrised system of equations, one finds that

$$V_2 = \frac{58 + 9V_1 - 4\sqrt{18V_1 - 8}}{27 + 54V_1}, \tag{5.69}$$

which is valid when $V_1 \geq V_2$. Due to symmetry, for the region $V_1 \leq V_2$, one may simply interchange V_1 and V_2 in the above formula. Notice that choosing $V_1 = V_2$ returns the equation (5.67).

One can reproduce the critical visibilities (up to numerical precision) without the tetrahedra assumption on Bob's response. To this end, notice that for a fixed value of (u, r), the existence of a bilocal model is a linear program over Bob's response function. We have made a fine-grained grid of the values of (u, r) and evaluated the corresponding linear program for each case. This allows us to reliably find the largest value of V_2 (for a given V_1) for which there is a bilocal model. In Figure 5.6 we present the critical pairs (V_1, V_2) for the existance of a bilocal model. Importantly, notice that the product of the critical pairs is not constant (as displayed in the Figure). If the quantum correlations were based on wiring two separate standard Bell experiments, one would expect the critical pairs to factor (this has been the case in previous bilocality inequalities). This is yet another indication that the quantum correlations are not based on wiring of standard Bell experiments. This leaves a natural open problem to prove the bilocal set in the (V_1, V_2) plane in

full generality (i.e. to prove the tetrahedral assumption on the hidden variable distributions).

Bilocality inequality

The bilocal simulation applies only to our target distribution p_{Q}. It is, however, natural to consider the construction of Bell inequalities for the bilocality scenario that both bear no resemblance to standard Bell inequalities and apply to general probability distributions. Our results for the simulation of p_{Q} suggests a natural candidate for such a bilocality inequality. Consider the Bell expression

$$\mathcal{A} \equiv \sum_{x\neq z,b} \sqrt{p(b)\left(1-\tilde{m}_{b,x}\tilde{m}_{b,z}E_b^{\mathrm{AC}}(x,z)\right)} + \sum_{j,b}\left(\sqrt{p(b)\left(1+\tilde{m}_{b,j}E_b^{\mathrm{A}}(j)\right)}+\sqrt{p(b)\left(1-\tilde{m}_{b,j}E_b^{\mathrm{C}}(j)\right)}\right).$$

This expression is tailored to p_{Q}. When $V_1 = V_2 = 1$, the quantum value becomes $\mathcal{A} = 12\sqrt{6} \approx 29.39$. Under the assumption of qubit states and rank-one projective measurements for Bob, numerical searches found no quantum strategy that improves on this result.

What is the largest value of $\mathcal{A}$ achievable in a bilocal model? We have again employed the tetrahedral symmetry to write the problem as a nonlinear maximisation (of $\mathcal{A}$) which depends on 14 variables. Of these variables, 12 characterise the symmetric response of Bob and the remaining two are (u, r). Then one finds the bilcal maximum

$$\mathcal{A} = 2\sqrt{3}\left(6+\sqrt{5}\right) \approx 28.53. \tag{5.70}$$

Thus, under the tetrahedral assumption, the bilocality inequality admits a quantum violation. Notably, we have also searched for the bilocal maximum by numerical means without imposing tetrahedra symmetry and in that manner independently recovered the above bound. We conjecture that the obtained bilocal result is optimal.

Finally, we remark that the bilocality inequality does not successfully detect all non-bilocal p_{Q}. Nevertheless, it does manage to detect quantum correlations at a reasonable visibility. Choosing $V \equiv V_1 = V_2$, one finds the critical visibility

$$V = 88.0\% \tag{5.71}$$

This should make the bilocality inequality applicable to experimental demonstrations.

Conclusions

Motivation

Quantum theory describes a reality that is radically different from that encountered in established, pre-quantum, physical theories such as classical mechanics, electrodynamics, thermodynamics and even the theory of relativity. The essence of the quantum world is its remarkable ontology in which physical states can lack definite properties (Schrödinger's cat is *both* dead and alive), the act of observation changes the future of a physical system (nature changes when observed), the outcomes of measurements can be intrinsically random[2] and knowledge is subject to fundamental uncertainty[3]. It appears that quantum theory is describing a world that is very different from everyday experience. Nevertheless, most of us firmly believe that there is only one reality. How do we reconcile this with the remarkable (and arguably unparalleled) success of quantum theory, in terms of explaining and predicting nature, witnessed over the last hundred years of physical science? In order to find an answer, we must truly understand quantum theory. We must understand what makes it so different from the more familiar physics of everyday experience. We must understand why nature seems to follow the predictions of quantum theory, rather than some other (potentially even more conceptually radical) theory. We must understand which concepts of a quantum theory are essential in nature and which, perhaps, could be different in an even more sophisticated post-quantum theory of physics. And we must understand what are the ultimate limitations of a reality governed by quantum theory so that we can probe and explore its most extreme predictions in state-of-the-art experiments [91, 92, 93].

However, it is an astonishing fact that such research in the conceptual foundations of the quantum world constitutes only one side of a coin. The era in which we can individually control and

[2]Intrinsic randomness is to be distinguished from randomness that arises simply from lack of knowledge. Many processes in everyday life feature the latter form of randomness. For instance, a coin toss appears random but it could be predicted if the appropriate knowledge about the toss, the coin and its environment is provided.

[3]That uncertainty is fundamental means that nature, rather than technological limitations, hinder us from eliminating the uncertainty.

manipulate atoms and sub-atomic particles is presently unfolding. Aptly capturing its magnitude of impact, it is often referred to as the second quantum revolution [94] which makes research in quantum theory not only a matter of conceptual development but also a matter of technological applications. These quantum technologies either promise to dramatically improve on conventional technologies or to offer solutions to tasks that are not known to have any solution with technologies based on pre-quantum physics. This is possible due to the close interplay with advances in our foundational understanding of quantum theory. Let us give three examples of this pivotal connection. 1) In order to encode, say, a million different objects in a computer, we need to construct all bit-strings of length roughly 20 (since 2^{20} is approximately a million). However, since quantum systems can occupy many different states at the same time, we could probabilistically find all one million different cases by measuring just a single string of 20 quantum bits. This possibility is harvested towards the construction of quantum computers. Very recently, the first demonstration of quantum supremacy over present super computers has been achieved [95]. 2) The field of quantum cryptography rests on the fact that quantum systems change when they are observed. Practically, this means that the presence of adversaries attempting to read a secret message can be detected. Therefore, powered by the laws of the quantum world, the security of cryptography is no longer based on the assumption that adversaries do not have strong enough computers to break the encryption but, instead, directly on the laws of nature. In recent years, the quantum distribution of cryptographic keys has been achieved over very long distances [96, 97]. 3) One of the most remarkable fundamental features of quantum theory is that it predicts laboratory data that cannot be accounted for in any (perhaps unknown) physical theory based on local hidden variables. If such data is observed, we can exclude the possibility that some unknown influence is governing its output, even without knowing how the device precisely operates. Therefore, quantum theory opens an avenue to random number generators whose security is certified independently by the inner workings of the device. Such quantum random number generators, operating at the highest security levels, have been achieved in recent experiments [98, 99, 100].

Summary

In view of the above, it is evident that the many crucial questions surrounding the foundational and applied aspects of quantum theory demand research attention. This thesis is a contribution to that ongoing research effort. The work that has been presented here places particular emphasis on the study of quantum correlations, i.e. the study of predictions made by quantum theory that cannot be explained by classical physical models. Topics that have been investigated in the above chapters are

- Understanding the creation and limitations of quantum correlations in communication experiments.
- Exploring how quantum correlations arise in network of distant observers.
- Investigating and developing the relationship between quantum correlations and information.
- Studying the relationship between different quantum resources and their applications in distributed computation.
- Finding methods for deducing the precise properties of a quantum resource based on the correlations it produces in simple experiments.

While the precise character of the contributions of this thesis is to be found in the above chapters, they span a number of qualitatively different types including the introduction of novel concepts (see e.g. section 2.4), exploring already established but poorly understood concepts (see e.g. sections 2.1 and 5.3), developing connections between different fundamental features of quantum theory (see e.g. sections 4.2 and 5.1), solving practically motivated tasks (see e.g. section 3) and developing computational tools for quantum information processing (see section 2.3). Notably, while these contributions are theoretical, many of them have motivated recent experimental works focused on the realisation of different forms of quantum correlations (see e.g. Refs [26, 61, 101, 70]).

Outlook

Naturally, on the topic of quantum correlations, many questions remain open and many new questions are made relevant due to more recent theoretical or experimental advances. I shall highlight three avenues of research that I find particularly interesting and promising.

Quantum correlations in large systems.— Naively, it may appear as if modern physics does not describe a single reality, but rather a quantum world of the microscopic and a classical world of the macroscopic. However, nature does (probably) not draw a sharp border between what we call quantum and what we call classical. This suggests that the borderland between the microscopic and the macroscopic should witness a natural merge of the quantum and classical phenomena. In other words, as systems grow larger, we would expect quantum properties to degrade and finally vanish in favour of a classical description. Theoretically [102] and experimentally [103, 104, 105] exploring the emergence of classicality in a quantum world is a fundamental and intriguing question.

Semi-device-independent quantum information processing.— Quantum theory can enhance many practical task beyond the limitations of conventional technologies. However, within the realm of quantum technologies, there is a multitude of qualitatively different paths to developing

protocols for the same task. Often, a desirable feature of a quantum technology is that it can achieve its purpose while making small and reasonable assumptions. The most rigorous path is known as device-independent quantum information processing in which tasks are completed with minimal assumptions on the quantum devices. However, the price to pay for this is that implementations become very demanding. Therefore, a compromise becomes important: how can quantum technologies run efficiently while operating only under mild and well-chosen assumptions? A breadth of different answers are possible and they are collectively referred to as semi-device-independent quantum information processing. In recent years, a number of different paradigms for such an approach to quantum technologies has been proposed. Important examples are based on physical dimension [106], energy [107] and overlap [108]. Notably, this thesis laid out the foundations for the development of another approach to semi-device-independent quantum information processing based on the information carried in physical ensembles (see section 2.4). This diversity in semi-device-independent quantum information processing is both interesting and necessary due to the many different types of applications, and physical platforms for their realisation, that are relevant to the field. The theoretical and experimental development of semi-device-independent protocols for practically useful tasks is a promising route towards high-performing and efficient quantum technologies.

Entanglement in measurements.— Entanglement is perhaps the most puzzling feature of quantum theory: it is the impossibility of understanding a quantum system by examining only its parts. This remains true even if parts of the system are separated by a huge distance. Entanglement has been studied intensively over many decades and the state-of-the-art knowledge is rather advanced. The research focus has largely been on classifying, quantifying and detecting entanglement in physical states. However, entanglement in quantum theory is as relevant to measurements as it is to states. Entanglement in measurements is arguably a more complex phenomenon since quantum measurements are composed of collections of operators and entanglement may arise both individually and collectively among them. In comparison, the entanglement of measurements is much less explored. Nevertheless, it appears reasonable to believe that it should carry an equally crucial signficance for the understanding of quantum theory as do entangled states. Therefore, a research effort devoted to the exploration of entangled measurements is interesting and relevant. Furthermore, studying more sophisticated forms of entanglement is also timely due to experimental advances seen over the last few years. Entangled states of both many particles [109, 110] and of high dimension [111] has been realised, as well as entangled measurements beyond qubit systems [112]. In view of these experimental advances, it is also intriguing to investigate quantum correlations phenomena that is native to more sophisitcated forms of entangled states and/or measurements.

Bibliography

[1] J. S. Bell. On the Einstein Podolsky Rosen Paradox. *Physics*, 1964(3):195–290.

[2] Marissa Giustina, Marijn A. M. Versteegh, Sören Wengerowsky, Johannes Handsteiner, Armin Hochrainer, Kevin Phelan, Fabian Steinlechner, Johannes Kofler, Jan-Å ke Larsson, Carlos Abellán, Waldimar Amaya, Valerio Pruneri, Morgan W. Mitchell, Jörn Beyer, Thomas Gerrits, Adriana E. Lita, Lynden K. Shalm, Sae Woo Nam, Thomas Scheidl, Rupert Ursin, Bernhard Wittmann, and Anton Zeilinger. Significant-loophole-free test of Bell's theorem with entangled photons. *Physical Review Letters*, 115(25):250401, December 2015.

[3] Lynden K. Shalm, Evan Meyer-Scott, Bradley G. Christensen, Peter Bierhorst, Michael A. Wayne, Martin J. Stevens, Thomas Gerrits, Scott Glancy, Deny R. Hamel, Michael S. Allman, Kevin J. Coakley, Shellee D. Dyer, Carson Hodge, Adriana E. Lita, Varun B. Verma, Camilla Lambrocco, Edward Tortorici, Alan L. Migdall, Yanbao Zhang, Daniel R. Kumor, William H. Farr, Francesco Marsili, Matthew D. Shaw, Jeffrey A. Stern, Carlos Abellán, Waldimar Amaya, Valerio Pruneri, Thomas Jennewein, Morgan W. Mitchell, Paul G. Kwiat, Joshua C. Bienfang, Richard P. Mirin, Emanuel Knill, and Sae Woo Nam. Strong Loophole-Free Test of Local Realism. *Physical Review Letters*, 115(25):250402, December 2015.

[4] Wenjamin Rosenfeld, Daniel Burchardt, Robert Garthoff, Kai Redeker, Norbert Ortegel, Markus Rau, and Harald Weinfurter. Event-Ready Bell Test Using Entangled Atoms Simultaneously Closing Detection and Locality Loopholes. *Physical Review Letters*, 119(1):010402, July 2017.

[5] S. Kochen and E. P. Specker. On the problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics*, 1967(1):59–87.

[6] Alexander A. Klyachko, M. Ali Can, Sinem Binicioğlu, and Alexander S. Shumovsky. Simple Test for Hidden Variables in Spin-1 Systems. *Physical Review Letters*, 101(2):020403, July 2008.

[7] G. Kirchmair, F. Zähringer, R. Gerritsma, M. Kleinmann, O. Gühne, A. Cabello, R. Blatt, and C. F. Roos. State-independent experimental test of quantum contextuality. *Nature*, 460(7254):494–497, July 2009.

[8] Richard Cleve and Harry Buhrman. Substituting quantum entanglement for communication. *Physical Review A*, 56(2):1201–1204, August 1997.

[9] Dominic Mayers and Andrew Yao. Self testing quantum apparatus. *QIC*, 4(4):273–286.

[10] Arthur Fine. Hidden Variables, Joint Probability, and the Bell Inequalities. *Physical Review Letters*, 48(5):291–295, February 1982.

[11] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt. Proposed Experiment to Test Local Hidden-Variable Theories. *Physical Review Letters*, 23(15):880–884, October 1969.

[12] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner. Bell nonlocality. *Reviews of Modern Physics*, 86(2):419–478, April 2014.

[13] Marek Zukowski and Časlav Brukner. Quantum non-locality—it ain't necessarily so... *Journal of Physics A: Mathematical and Theoretical*, 47(42):424009, October 2014.

[14] R. W. Spekkens. Contextuality for preparations, transformations, and unsharp measurements. *Physical Review A*, 71(5):052108, May 2005.

[15] Adán Cabello, JoséM. Estebaranz, and Guillermo García-Alcaine. Bell-Kochen-Specker theorem: A proof with 18 vectors. *Physics Letters A*, 212(4):183–187, March 1996.

[16] Matthew F. Pusey. Robust preparation noncontextuality inequalities in the simplest scenario. *Physical Review A*, 98(2):022112, August 2018.

[17] Alley Hameedi, Armin Tavakoli, Breno Marques, and Mohamed Bourennane. Communication Games Reveal Preparation Contextuality. *Physical Review Letters*, 119(22):220402, November 2017.

[18] Debashis Saha and Anubhav Chaturvedi. Preparation contextuality as an essential feature underlying quantum communication advantage. *Physical Review A*, 100(2):022108, August 2019.

[19] Armin Tavakoli and Roope Uola. Measurement incompatibility and steering are necessary and sufficient for operational contextuality. *Physical Review Research*, 2(1):013011, January 2020.

[20] Harry. Buhrman, Richard. Cleve, and Wim. van Dam. Quantum Entanglement and Communication Complexity. *SIAM Journal on Computing*, 30(6):1829–1841, January 2001.

[21] Andris Ambainis, Ashwin Nayak, Amnon Ta-Shma, and Umesh Vazirani. Dense quantum coding and quantum finite automata. *Journal of the ACM*, 49(4):496–511, July 2002.

[22] Harry Buhrman, Richard Cleve, Serge Massar, and Ronald de Wolf. Nonlocality and communication complexity. *Reviews of Modern Physics*, 82(1):665–698, March 2010.

[23] A.S. Holevo. The capacity of the quantum channel with general signal states. *IEEE Transactions on Information Theory*, 44(1):269–273, January 1998.

[24] Armin Tavakoli and Marek Żukowski. Higher-dimensional communication complexity problems: Classical protocols versus quantum ones based on Bell's theorem or prepare-transmit-measure schemes. *Physical Review A*, 95(4):042305, April 2017.

[25] Armin Tavakoli, Marek Żukowski, and Časlav Brukner. Does violation of a Bell inequality always imply quantum advantage in a communication complexity problem? *arXiv:1907.01322 [quant-ph]*, October 2019.

[26] Daniel Martínez, Armin Tavakoli, Mauricio Casanova, Gustavo Cañas, Breno Marques, and Gustavo Lima. High-Dimensional Quantum Communication Complexity beyond Strategies Based on Bell's Theorem. *Physical Review Letters*, 121(15):150504, October 2018.

[27] Armin Tavakoli, Denis Rosset, and Marc-Olivier Renou. Enabling Computation of Correlation Bounds for Finite-Dimensional Quantum Systems via Symmetrization. *Physical Review Letters*, 122(7):070501, February 2019.

[28] Armin Tavakoli, Emmanuel Zambrini Cruzeiro, Jonatan Bohr Brask, Nicolas Gisin, and Nicolas Brunner. Informationally restricted quantum correlations. *arXiv:1909.05656 [quant-ph]*, September 2019.

[29] Časlav Brukner, Marek Żukowski, and Anton Zeilinger. Quantum Communication Complexity Protocol with Two Entangled Qutrits. *Physical Review Letters*, 89(19):197901, October 2002.

[30] Časlav Brukner, Tomasz Paterek, and Marek Żukowski. Quantum communication complexity protocols based on higher-dimensional entangled systems. *International Journal of Quantum Information*, 01(04):519–525, December 2003.

[31] Armin Tavakoli, Marcin Pawł owski, Marek Żukowski, and Mohamed Bourennane. Dimensional discontinuity in quantum communication complexity at dimension seven. *Physical Review A*, 95(2):020302, February 2017.

[32] Thomas Lawson, Noah Linden, and Sandu Popescu. Biased nonlocal quantum games. *arXiv:1011.6245 [quant-ph]*, November 2010.

[33] Sadiq Muhammad, Armin Tavakoli, Maciej Kurant, Marcin Pawł owski, Marek Żukowski, and Mohamed Bourennane. Quantum Bidding in Bridge. *Physical Review X*, 4(2):021047, June 2014.

[34] Nicolas Gisin. Bell inequalities: Many questions, a few answers. *arXiv:quant-ph/0702021*, May 2007.

[35] Marcin Pawł owski and Marek Żukowski. Entanglement-assisted random access codes. *Physical Review A*, 81(4):042326, April 2010.

[36] Armin Tavakoli, Breno Marques, Marcin Pawł owski, and Mohamed Bourennane. Spatial versus sequential correlations for random access coding. *Physical Review A*, 93(3):032336, March 2016.

[37] Časlav Brukner, Marek Żukowski, Jian-Wei Pan, and Anton Zeilinger. Bell's Inequalities and Quantum Communication Complexity. *Physical Review Letters*, 92(12):127901, March 2004.

[38] Daniel Collins, Nicolas Gisin, Noah Linden, Serge Massar, and Sandu Popescu. Bell Inequalities for Arbitrarily High-Dimensional Systems. *Physical Review Letters*, 88(4):040404, January 2002.

[39] A. Acín, T. Durt, N. Gisin, and J. I. Latorre. Quantum nonlocality in two three-level systems. *Physical Review A*, 65(5):052325, May 2002.

[40] Daniel Collins and Nicolas Gisin. A relevant two qubit Bell inequality inequivalent to the CHSH inequality. *Journal of Physics A: Mathematical and General*, 37(5):1775–1787, January 2004.

[41] M. Froissart. Constructive generalization of Bell's inequalities. *Il Nuovo Cimento B (1971-1996)*, 64(2):241–251, August 1981.

[42] Jonathan Oppenheim and Stephanie Wehner. The Uncertainty Principle Determines the Nonlocality of Quantum Mechanics. *Science*, 330(6007):1072–1074, November 2010.

[43] Andris Ambainis, Debbie Leung, Laura Mancinska, and Maris Ozols. Quantum Random Access Codes with Shared Randomness. *arXiv:0810.2937 [quant-ph]*, June 2009.

[44] Stefan Zohren and Richard D. Gill. Maximal Violation of the Collins-Gisin-Linden-Massar-Popescu Inequality for Infinite Dimensional States. *Physical Review Letters*, 100(12):120406, March 2008.

[45] Lieven Vandenberghe and Stephen Boyd. Semidefinite Programming. *SIAM Review*, 38(1):49–95, March 1996.

[46] Miguel Navascués and Harald Wunderlich. A glance beyond the quantum model. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 466(2115):881–890, March 2010.

[47] Miguel Navascués and Tamás Vértesi. Bounding the Set of Finite Dimensional Quantum Correlations. *Physical Review Letters*, 115(2):020501, July 2015.

[48] Miguel Navascués, Adrien Feix, Mateus Araújo, and Tamás Vértesi. Characterizing finite-dimensional quantum behavior. *Physical Review A*, 92(4):042117, October 2015.

[49] Karin Gatermann and Pablo A. Parrilo. Symmetry groups, semidefinite programs, and sums of squares. *Journal of Pure and Applied Algebra*, 192(1):95–128, September 2004.

[50] Armin Tavakoli, Alley Hameedi, Breno Marques, and Mohamed Bourennane. Quantum Random Access Codes Using Single d-Level Systems. *Physical Review Letters*, 114(17):170502, April 2015.

[51] Máté Farkas and Jędrzej Kaniewski. Self-testing mutually unbiased bases in the prepare-and-measure scenario. *Physical Review A*, 99(3):032316, March 2019.

[52] Rodrigo Gallego, Nicolas Brunner, Christopher Hadley, and Antonio Acín. Device-Independent Tests of Classical and Quantum Dimensions. *Physical Review Letters*, 105(23):230501, November 2010.

[53] Armin Tavakoli, Jędrzej Kaniewski, Tamás Vértesi, Denis Rosset, and Nicolas Brunner. Self-testing quantum states and measurements in the prepare-and-measure scenario. *Physical Review A*, 98(6):062307, December 2018.

[54] Armin Tavakoli, Massimiliano Smania, Tamás Vértesi, Nicolas Brunner, and Mohamed Bourennane. Self-testing non-projective quantum measurements in prepare-and-measure experiments. *arXiv:1811.12712 [quant-ph]*, December 2018.

[55] Karthik Mohan, Armin Tavakoli, and Nicolas Brunner. Sequential random access codes and self-testing of quantum measurement instruments. *New Journal of Physics*, 21(8):083034, August 2019.

[56] Armin Tavakoli, Alastair A. Abbott, Marc-Olivier Renou, Nicolas Gisin, and Nicolas Brunner. Semi-device-independent characterization of multipartite entanglement of states and measurements. *Physical Review A*, 98(5):052333, November 2018.

[57] Jędrzej Kaniewski. Analytic and Nearly Optimal Self-Testing Bounds for the Clauser-Horne-Shimony-Holt and Mermin Inequalities. *Physical Review Letters*, 117(7):070402, August 2016.

[58] Giacomo Mauro D'Ariano, Paoloplacido Lo Presti, and Paolo Perinotti. Classical randomness in quantum measurements. *Journal of Physics A: Mathematical and General*, 38(26):5979–5991, June 2005.

[59] Joseph M. Renes, Robin Blume-Kohout, A. J. Scott, and Carlton M. Caves. Symmetric informationally complete quantum measurements. *Journal of Mathematical Physics*, 45(6):2171–2180, May 2004.

[60] John J. Benedetto and Matthew Fickus. Finite Normalized Tight Frames. *Advances in Computational Mathematics*, 18(2):357–385, February 2003.

[61] Hammad Anwer, Sadiq Muhammad, Walid Cherifi, Nikolai Miklin, Armin Tavakoli, and Mohamed Bourennane. Experimental characterisation of unsharp qubit measurements in a semi-device-independent setting. *arXiv:2001.04768 [quant-ph]*, January 2020.

[62] Giulio Foletto, Luca Calderaro, Giuseppe Vallone, and Paolo Villoresi. Experimental Demonstration of Sequential Quantum Random Access Codes. *arXiv:2001.04885 [quant-ph]*, January 2020.

[63] R. Augusiak, M. Demianowicz, and A. Acín. Local hidden–variable models for entangled quantum states. *Journal of Physics A: Mathematical and Theoretical*, 47(42):424002, October 2014.

[64] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Physical Review Letters*, 70(13):1895–1899, March 1993.

[65] Charles H. Bennett and Stephen J. Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Physical Review Letters*, 69(20):2881–2884, November 1992.

[66] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. General teleportation channel, singlet fraction, and quasidistillation. *Physical Review A*, 60(3):1888–1898, September 1999.

[67] Reinhard F. Werner. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Physical Review A*, 40(8):4277–4281, October 1989.

[68] S. M. Hashemi Rafsanjani, M. Huber, C. J. Broadbent, and J. H. Eberly. Genuinely multipartite concurrence of N-qubit X matrices. *Physical Review A*, 86(6):062303, December 2012.

[69] H. M. Wiseman, S. J. Jones, and A. C. Doherty. Steering, Entanglement, Nonlocality, and the Einstein-Podolsky-Rosen Paradox. *Physical Review Letters*, 98(14):140402, April 2007.

[70] Hammad Anwer, Natalie Wilson, Ralph Silva, Sadiq Muhammad, Armin Tavakoli, and Mohamed Bourennane. Noise-robust preparation contextuality shared between any number of observers via unsharp measurements. *arXiv:1904.09766 [quant-ph]*, April 2019.

[71] Robert W. Spekkens, D. H. Buzacott, A. J. Keehn, Ben Toner, and G. J. Pryde. Preparation Contextuality Powers Parity-Oblivious Multiplexing. *Physical Review Letters*, 102(1):010401, January 2009.

[72] Paul Busch, Pekka J. Lahti, and Peter Mittelstaedt. *The Quantum Theory of Measurement*. Lecture Notes in Physics Monographs. Springer-Verlag, Berlin Heidelberg, second edition, 1996.

[73] Teiko Heinosaari, Takayuki Miyadera, and Mário Ziman. An invitation to quantum incompatibility. *Journal of Physics A: Mathematical and Theoretical*, 49(12):123001, February 2016.

[74] Paul Busch, Pekka J. Lahti, Juha-Pekka Pellonpää, and Kari Ylinen. *Quantum Measurement*. Theoretical and Mathematical Physics. Springer International Publishing, 2016.

[75] Andrew Gleason. Measures on the Closed Subspaces of a Hilbert Space. *Journal of Mathematics and Mechanics*, 6(6):885–893, 1957.

[76] P. Busch. Quantum States and Generalized Observables: A Simple Proof of Gleason's Theorem. *Physical Review Letters*, 91(12):120403, September 2003.

[77] Jessica Bavaresco, Marco Túlio Quintino, Leonardo Guerini, Thiago O. Maciel, Daniel Cavalcanti, and Marcelo Terra Cunha. Most incompatible measurements for robust steering tests. *Physical Review A*, 96(2):022110, August 2017.

[78] André Chailloux, Iordanis Kerenidis, Srijita Kundu, and Jamie Sikora. Optimal bounds for parity-oblivious random access codes. *New Journal of Physics*, 18(4):045003, April 2016.

[79] Ralph Silva, Nicolas Gisin, Yelena Guryanova, and Sandu Popescu. Multiple Observers Can Share the Nonlocality of Half of an Entangled Pair by Using Optimal Weak Measurements. *Physical Review Letters*, 114(25):250401, June 2015.

[80] Julian Schwinger. Unitary Operator Bases. *Proceedings of the National Academy of Sciences*, 46(4):570–579, April 1960.

[81] Armin Tavakoli, Máté Farkas, Denis Rosset, Jean-Daniel Bancal, and Jędrzej Kaniewski. Mutually unbiased bases and symmetric informationally complete measurements in Bell experiments: Bell inequalities, device-independent certification and applications. *arXiv:1912.03225 [quant-ph]*, December 2019.

[82] Armin Tavakoli, Marc Olivier Renou, Nicolas Gisin, and Nicolas Brunner. Correlations in star networks: From Bell inequalities to network inequalities. *New Journal of Physics*, 19(7):073003, July 2017.

[83] Nicolas Gisin and Armin Tavakoli. Geninue quantum violations of bilocality. *In preparation*, 2020.

[84] Lluís Masanes, Stefano Pironio, and Antonio Acín. Secure device-independent quantum key distribution with causally independent measurement devices. *Nature Communications*, 2(1):1–7, March 2011.

[85] Miguel Navascués, Stefano Pironio, and Antonio Acín. Bounding the Set of Quantum Correlations. *Physical Review Letters*, 98(1):010401, January 2007.

[86] Tobias Fritz. Beyond Bell's theorem: Correlation scenarios. *New Journal of Physics*, 14(10):103001, October 2012.

[87] Cyril Branciard, Denis Rosset, Nicolas Gisin, and Stefano Pironio. Bilocal versus nonbilocal correlations in entanglement-swapping experiments. *Physical Review A*, 85(3):032119, March 2012.

[88] Armin Tavakoli, Paul Skrzypczyk, Daniel Cavalcanti, and Antonio Acín. Nonlocal correlations in the star-network configuration. *Physical Review A*, 90(6):062109, December 2014.

[89] Nicolas Gisin, Quanxin Mei, Armin Tavakoli, Marc Olivier Renou, and Nicolas Brunner. All entangled pure quantum states violate the bilocality inequality. *Physical Review A*, 96(2):020304, August 2017.

[90] Nicolas Gisin. Entanglement 25 Years after Quantum Teleportation: Testing Joint Measurements in Quantum Networks. *Entropy*, 21(3):325, March 2019.

[91] C. Abellán, A. Acín, A. Alarcón, O. Alibart, C. K. Andersen, F. Andreoli, A. Beckert, F. A. Beduini, A. Bendersky, M. Bentivegna, P. Bierhorst, D. Burchardt, A. Cabello, J. Cariñe, S. Carrasco, G. Carvacho, D. Cavalcanti, R. Chaves, J. Cortés-Vega, A. Cuevas, A. Delgado, H. de Riedmatten, C. Eichler, P. Farrera, J. Fuenzalida, M. García-Matos, R. Garthoff, S. Gasparinetti, T. Gerrits, F. Ghafari Jouneghani, S. Glancy, E. S. Gómez, P. González, J.-Y. Guan, J. Handsteiner, J. Heinsoo, G. Heinze, A. Hirschmann, O. Jiménez, F. Kaiser, E. Knill, L. T. Knoll, S. Krinner, P. Kurpiers, M. A. Larotonda, J.-Å. Larsson, A. Lenhard, H. Li, M.-H. Li, G. Lima, B. Liu, Y. Liu, I. H. López Grande, T. Lunghi, X. Ma, O. S. Magaña-Loaiza, P. Magnard, A. Magnoni, M. Martí-Prieto, D. Martínez, P. Mataloni, A. Mattar, M. Mazzera, R. P. Mirin, M. W. Mitchell, S. Nam, M. Oppliger, J.-W. Pan, R. B. Patel, G. J. Pryde, D. Rauch, K. Redeker, D. Rieländer, M. Ringbauer, T. Roberson, W. Rosenfeld, Y. Salathé, L. Santodonato, G. Sauder, T. Scheidl, C. T. Schmiegelow, F. Sciarrino, A. Seri, L. K. Shalm, S.-C. Shi, S. Slussarenko, M. J. Stevens, S. Tanzilli, F. Toledo, J. Tura, R. Ursin, P. Vergyris, V. B. Verma, T. Walter, A. Wallraff, Z. Wang, H. Weinfurter, M. M. Weston, A. G. White, C. Wu, G. B. Xavier, L. You, X. Yuan, A. Zeilinger, Q. Zhang, W. Zhang, J. Zhong, and The BIG Bell Test Collaboration. Challenging local realism with human choices. *Nature*, 557(7704):212–216, May 2018.

[92] Ming-Han Li, Cheng Wu, Yanbao Zhang, Wen-Zhao Liu, Bing Bai, Yang Liu, Weijun Zhang, Qi Zhao, Hao Li, Zhen Wang, Lixing You, W. J. Munro, Juan Yin, Jun Zhang, Cheng-Zhi Peng, Xiongfeng Ma, Qiang Zhang, Jingyun Fan, and Jian-Wei Pan. Test of Local Realism into the Past without Detection and Locality Loopholes. *Physical Review Letters*, 121(8):080404, August 2018.

[93] Johannes Handsteiner, Andrew S. Friedman, Dominik Rauch, Jason Gallicchio, Bo Liu, Hannes Hosp, Johannes Kofler, David Bricher, Matthias Fink, Calvin Leung, Anthony Mark, Hien T. Nguyen, Isabella Sanders, Fabian Steinlechner, Rupert Ursin, Sören Wengerowsky, Alan H. Guth, David I. Kaiser, Thomas Scheidl, and Anton Zeilinger. Cosmic Bell Test: Measurement Settings from Milky Way Stars. *Physical Review Letters*, 118(6):060401, February 2017.

[94] Jonathan P. Dowling and Gerard J. Milburn. Quantum technology: The second quantum revolution. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 361(1809):1655–1674, August 2003.

[95] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando G. S. L. Brandao, David A. Buell, Brian Burkett, Yu Chen, Zijun Chen, Ben Chiaro, Roberto Collins, William Courtney, Andrew Dunsworth, Edward Farhi, Brooks Foxen, Austin Fowler, Craig Gidney, Marissa Giustina, Rob Graff, Keith Guerin, Steve Habegger, Matthew P. Harrigan, Michael J. Hartmann, Alan Ho, Markus Hoffmann, Trent Huang, Travis S. Humble, Sergei V. Isakov, Evan Jeffrey, Zhang Jiang, Dvir Kafri, Kostyantyn Kechedzhi, Julian Kelly, Paul V. Klimov, Sergey Knysh, Alexander Korotkov, Fedor Kostritsa, David Landhuis, Mike Lindmark, Erik Lucero, Dmitry Lyakh, Salvatore Mandrà, Jarrod R. McClean, Matthew McEwen, Anthony Megrant, Xiao Mi, Kristel Michielsen, Masoud Mohseni, Josh Mutus, Ofer Naaman, Matthew Neeley, Charles Neill, Murphy Yuezhen Niu, Eric Ostby, Andre Petukhov, John C. Platt, Chris Quintana, Eleanor G. Rieffel, Pedram Roushan, Nicholas C. Rubin, Daniel Sank, Kevin J. Satzinger, Vadim Smelyanskiy, Kevin J. Sung, Matthew D. Trevithick, Amit Vainsencher, Benjamin Villalonga, Theodore White, Z. Jamie Yao, Ping Yeh, Adam Zalcman, Hartmut Neven, and John M. Martinis. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510, October 2019.

[96] Sheng-Kai Liao, Wen-Qi Cai, Wei-Yue Liu, Liang Zhang, Yang Li, Ji-Gang Ren, Juan Yin, Qi Shen, Yuan Cao, Zheng-Ping Li, Feng-Zhi Li, Xia-Wei Chen, Li-Hua Sun, Jian-Jun Jia, Jin-Cai Wu, Xiao-Jun Jiang, Jian-Feng Wang, Yong-Mei Huang, Qiang Wang, Yi-Lin Zhou, Lei Deng, Tao Xi, Lu Ma, Tai Hu, Qiang Zhang, Yu-Ao Chen, Nai-Le Liu, Xiang-Bin Wang, Zhen-Cai Zhu, Chao-Yang Lu, Rong Shu, Cheng-Zhi Peng, Jian-Yu Wang, and Jian-Wei Pan. Satellite-to-ground quantum key distribution. *Nature*, 549(7670):43–47, September 2017.

[97] Alberto Boaron, Gianluca Boso, Davide Rusca, Cédric Vulliez, Claire Autebert, Misael Caloz, Matthieu Perrenoud, Gaëtan Gras, Félix Bussières, Ming-Jun Li, Daniel Nolan, Anthony

Martin, and Hugo Zbinden. Secure Quantum Key Distribution over 421 km of Optical Fiber. *Physical Review Letters*, 121(19):190502, November 2018.

[98] S. Pironio, A. Acín, S. Massar, A. Boyer de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe. Random numbers certified by Bell's theorem. *Nature*, 464(7291):1021–1024, April 2010.

[99] Carlos Abellán, Waldimar Amaya, Daniel Mitrani, Valerio Pruneri, and Morgan W. Mitchell. Generation of Fresh and Pure Random Numbers for Loophole-Free Bell Tests. *Physical Review Letters*, 115(25):250403, December 2015.

[100] Yanbao Zhang, Lynden K. Shalm, Joshua C. Bienfang, Martin J. Stevens, Michael D. Mazurek, Sae Woo Nam, Carlos Abellán, Waldimar Amaya, Morgan W. Mitchell, Honghao Fu, Carl A. Miller, Alan Mink, and Emanuel Knill. Experimental Low-Latency Device-Independent Quantum Randomness. *Physical Review Letters*, 124(1):010505, January 2020.

[101] Giulio Foletto, Luca Calderaro, Giuseppe Vallone, and Paolo Villoresi. Experimental Demonstration of Sequential Quantum Random Access Codes. *arXiv:2001.04885 [quant-ph]*, January 2020.

[102] Florian Fröwis, Pavel Sekatski, Wolfgang Dür, Nicolas Gisin, and Nicolas Sangouard. Macroscopic quantum states: Measures, fragility, and implementations. *Reviews of Modern Physics*, 90(2):025004, May 2018.

[103] Florian Fröwis, Peter C. Strassmann, Alexey Tiranov, Corentin Gut, Jonathan Lavoie, Nicolas Brunner, Félix Bussières, Mikael Afzelius, and Nicolas Gisin. Experimental certification of millions of genuinely entangled atoms in a solid. *Nature Communications*, 8(1):1–6, October 2017.

[104] Igor Marinković, Andreas Wallucks, Ralf Riedinger, Sungkun Hong, Markus Aspelmeyer, and Simon Gröblacher. Optomechanical Bell Test. *Physical Review Letters*, 121(22):220404, November 2018.

[105] Rainer Kaltenbaek, Markus Aspelmeyer, Peter F. Barker, Angelo Bassi, James Bateman, Kai Bongs, Sougato Bose, Claus Braxmaier, Časlav Brukner, Bruno Christophe, Michael Chwalla, Pierre-François Cohadon, Adrian Michael Cruise, Catalina Curceanu, Kishan Dholakia, Lajos Diósi, Klaus Döringshoff, Wolfgang Ertmer, Jan Gieseler, Norman Gürlebeck, Gerald Hechenblaikner, Antoine Heidmann, Sven Herrmann, Sabine Hossenfelder, Ulrich Johann, Nikolai Kiesel, Myungshik Kim, Claus Lämmerzahl, Astrid Lambrecht, Michael Mazilu,

Gerard J. Milburn, Holger Müller, Lukas Novotny, Mauro Paternostro, Achim Peters, Igor Pikovski, André Pilan Zanoni, Ernst M. Rasel, Serge Reynaud, Charles Jess Riedel, Manuel Rodrigues, Loïc Rondin, Albert Roura, Wolfgang P. Schleich, Jörg Schmiedmayer, Thilo Schuldt, Keith C. Schwab, Martin Tajmar, Guglielmo M. Tino, Hendrik Ulbricht, Rupert Ursin, and Vlatko Vedral. Macroscopic Quantum Resonators (MAQRO): 2015 update. *EPJ Quantum Technology*, 3(1):1–47, December 2016.

[106] Marcin Pawł owski and Nicolas Brunner. Semi-device-independent security of one-way quantum key distribution. *Physical Review A*, 84(1):010302, July 2011.

[107] Thomas Van Himbeeck, Erik Woodhead, Nicolas J. Cerf, Raúl García-Patrón, and Stefano Pironio. Semi-device-independent framework based on natural physical assumptions. *Quantum*, 1:33, November 2017.

[108] Jonatan Bohr Brask, Anthony Martin, William Esposito, Raphael Houlmann, Joseph Bowles, Hugo Zbinden, and Nicolas Brunner. Megahertz-Rate Semi-Device-Independent Quantum Random Number Generators Based on Unambiguous State Discrimination. *Physical Review Applied*, 7(5):054018, May 2017.

[109] Nicolai Friis, Oliver Marty, Christine Maier, Cornelius Hempel, Milan Holzäpfel, Petar Jurcevic, Martin B. Plenio, Marcus Huber, Christian Roos, Rainer Blatt, and Ben Lanyon. Observation of Entangled States of a Fully Controlled 20-Qubit System. *Physical Review X*, 8(2):021012, April 2018.

[110] Karsten Lange, Jan Peise, Bernd Lücke, Ilka Kruse, Giuseppe Vitagliano, Iagoba Apellaniz, Matthias Kleinmann, Géza Tóth, and Carsten Klempt. Entanglement between two spatially separated atomic modes. *Science*, 360(6387):416–418, April 2018.

[111] Jessica Bavaresco, Natalia Herrera Valencia, Claude Klöckl, Matej Pivoluska, Paul Erker, Nicolai Friis, Mehul Malik, and Marcus Huber. Measurements in two bases are sufficient for certifying high-dimensional entanglement. *Nature Physics*, 14(10):1032–1037, October 2018.

[112] Yi-Han Luo, Han-Sen Zhong, Manuel Erhard, Xi-Lin Wang, Li-Chao Peng, Mario Krenn, Xiao Jiang, Li Li, Nai-Le Liu, Chao-Yang Lu, Anton Zeilinger, and Jian-Wei Pan. Quantum Teleportation in High Dimensions. *Physical Review Letters*, 123(7):070505, August 2019.

PHYSICAL REVIEW A **95**, 042305 (2017)

Higher-dimensional communication complexity problems: Classical protocols versus quantum ones based on Bell's theorem or prepare-transmit-measure schemes

Armin Tavakoli[1,2] and Marek Żukowski[3]
[1]*Department of Physics, Stockholm University, S-10691 Stockholm, Sweden.*
[2]*Groupe de Physique Appliquée, Université de Genéve, CH-1211 Genéve, Switzerland*
[3]*Institute of Theoretical Physics and Astrophysics, Uniwersytet Gdański, PL-80-308 Gdańsk, Poland*
(Received 15 November 2016; published 4 April 2017)

Communication complexity problems (CCPs) are tasks in which separated parties attempt to compute a function whose inputs are distributed among the parties. Their communication is limited so that not all inputs can be sent. We show that broad classes of Bell inequalities can be mapped to CCPs and that a quantum violation of a Bell inequality is a necessary and sufficient condition for an enhancement of the related CCP beyond its classical limitation. However, one can implement CCPs by transmitting a quantum system, encoding no more information than is allowed in the CCP, and extracting information by performing measurements. We show that for a large class of Bell inequalities, the improvement of the CCP associated with a quantum violation of a Bell inequality can be no greater than the improvement obtained from quantum prepare-transmit-measure strategies.

DOI: 10.1103/PhysRevA.95.042305

I. INTRODUCTION

Bell's theorem asserts that measurements on separated entangled quantum states can give rise to outcome correlations that have no local realistic model [1]. This fact can be used to break classical limits in communication complexity problems (CCPs) [2]. However, quantum protocols for CCPs violating classical bounds, that are based on prepare-transmit-measure schemes involving just a single quantum system, are also possible [3]. This can be certified by a violation of an inequality bounding the strength of the classical counterpart of such a protocol.

Nonclassical features of various quantum predictions are an essential tool in many quantum information tasks such as (semi) device-independent cryptography [4,5], randomness generation [6–8], and dimension witnesses [9,10]. However, in terms of studying the fundamental physical phenomena, correlations due to the entanglement of two or more systems have been given significantly more attention than those obtained from preparing and measuring a single quantum system. Indeed, little is known about the relation between the strength of the two general types of nonclassical correlations enabled by quantum theory, and their comparative applicability in quantum protocols violating classical bounds in information processing tasks. For instance, it was shown in Ref. [11] that in some tasks entanglement is as least as good a resource as transmission of a single qubit. Examples in which entanglement is strictly better were given in [12]. In contrast, Ref. [13] showed that there are tasks in which a single-quantum system is a better resource than entanglement. Also, Ref. [14] showed that single-quantum systems can exhibit a discontinuous jump at dimension seven when performing an information processing task, while no such jump occurs for entanglement-based strategies.

Here, we aim to construct a framework based on games in which one can, on equal footing, compare the communication complexity reduction power of entanglement-based protocols and single-quantum system approaches. For this purpose, we will use a class of information-theoretic games related in fact to CCPs.

In CCPs, a number of parties, say N, attempt to jointly compute a task function $f(X_1,\ldots,X_N)$. However, the input X_i is only known to party i. The task the N parties attempt to solve is either (i) to minimize the communication required for one of them to compute f, or (ii) to maximize the probability of one party to correctly compute f when the communication between the parties is limited by some rule, which does not allow one to transmit all the data contained in any X_i. In this work we consider the latter. On the one hand, since single-system protocols are based on measurements on a transmitted quantum system of a specific dimension d, which constrains its information-carrying capacity to $\log d$ bits, appropriate CCPs are a natural habitat in which the quantum strength of such CCP protocols can be studied. On the other hand, Bell inequalities are known to exhibit links to games [15]. The relation between CCPs and correlations due to entanglement has been extensively studied [2,16,17].

We will show that for every bipartite Bell inequality, we can formulate a CCP such that the reduction of communication complexity obtained from using classical communication assisted by correlations due to shared entanglement directly corresponds to the ability of quantum theory to violate the original Bell inequality. However, the CCP can also be implemented in quantum theory by the preparation, transmission, and measurement of a single-quantum system. Using such CCPs as a framework for both types of quantum resources, we will show that for large classes of Bell inequalities, correlations due to measurements on entangled states cannot beat the performance of quantum prepare-transmit-measure protocols.

II. THE STUDIED CLASS OF BELL INEQUALITIES

In a bipartite Bell inequality, observers Alice and Bob perform measurements $x \in \{0,\ldots,m_A-1\}$ and $y \in \{0,\ldots,m_B-1\}$, respectively, with a distribution $p(x,y)$. Each measurement has an outcome $a,b \in \{0,\ldots,d-1\}$, respectively. Such Bell inequalities can in a general way be written

as

$$\sum_{x,y} p(x,y) \sum_{a,b=0}^{d-1} \sum_{k=0}^{K} c^k_{ab|xy} P(a,b|x,y) \leqslant B. \qquad (1)$$

B is the classical bound, $c^k_{ab|xy}$ are real numbers, and k is an index with some range $k \in \{0,\ldots,K\}$ for some natural number K. This index will allow us to put the inequalities in a form which generalizes the form of the CGLMP inequalities [18]. Note that we can without loss of generality assume that $\forall\, a,b,x,y$, there exists at most one $k \in \{0,\ldots,K\}$ such that $c^k_{ab|xy} \neq 0$. To see this, simply note that if $c^k_{ab|xy}$ and $c^{k'}_{ab|xy}$ with $k \neq k'$ were both nonzero, then in Eq. (1) we would encounter the terms $c^k_{ab|xy} P(a,b|x,y) + c^{k'}_{ab|xy} P(a,b|x,y) = (c^k_{ab|xy} + c^{k'}_{ab|xy}) P(a,b|x,y)$ where the prefactor is again just some real number.

The Bell inequalities of our interest have the following structure. First, we draw inspiration from a variety of known Bell inequalities in which correlations between Alice's and Bob's local outcomes are quantified using their sum $a+b$ mod d. With this in mind, for any given pair of measurements (x,y), we construct the set $S_{xy} = \{\forall\ (a,b,k)$ such that $c^k_{ab|xy} \neq 0\}$. We we require that S_{xy} admits a partition of the form $\{S^{i,k}_{xy}\}_{i,k}$ for $i = 1,\ldots,N$ and $k = 0,\ldots,K$, for some integers N,K, with $S^{i,k}_{xy} = \{(a,b)|a+b = F^i_{xy}(k) \mod d\}$ for some functions $F^i_{xy}(k)$, i.e.,

$$\forall\, x\ \forall\, y : S_{xy} = \bigcup_{i=1}^{N} \bigcup_{k=0}^{K} S^{i,k}_{xy},$$

$$\text{and} \quad S^{i,k}_{xy} \cap S^{i',k'}_{xy} = \emptyset \text{ for } (i,k) \neq (i',k'). \qquad (2)$$

Remark: since the sets $S^{i,k}_{xy}$ are disjoint for $i = 1,\ldots,N$ and $k = 0,\ldots,K$ it follows that there can be no set (a,b) that simultaneously satisfies both $a+b = F^j_{xy}(k)$ and $a+b = F^{j'}_{xy}(k')$ for $(j,k) \neq (j',k')$. This implies that the range of F^j_{xy} is disjoint with that of $F^{j'}_{xy}$ for $j \neq j'$. Also, since $a+b$ mod d can have at most d different values it follows that $(K+1)N \leqslant d$.

Second, we restrict the structure of $c^k_{ab|xy}$ such that we later can make the connection to a related family of CCPs. To see why this restriction is necessary, we remind ourselves that in CCPs, Alice and Bob attempt to compute the value of some function from which they earn some payoff. Although the local outcomes produced from measurements on a (perhaps) entangled state may assist Alice and Bob in performing the computation, the values of these local outcomes are per se of no interest in the CCP. Therefore, we require that the Bell inequality is such that the same coefficient $c^k_{ab|xy}$ is assigned to any pair $(a,b) \in S^{i,k}_{xy}$, i.e., we may write $c^k_{ab|xy} = c^{i,k}_{xy}$.

Thus, the Bell inequalities we will consider are written

$$I^{\text{Bell}} \equiv \sum_{x,y} p(x,y) \sum_{k=0}^{K} \sum_{i=1}^{N} c^{i,k}_{xy} P_{xy}[a+b = F^i_{xy}(k)] \leqslant B. \qquad (3)$$

By $\mathscr{B}$, we will denote some arbitrary Bell inequality of this form. This class can be viewed as a generalization of the inequalities considered in Ref. [16], from two-outcome to many-outcome Bell scenarios. A simple example of a well-known Bell inequality that can be written in the form (3) is the CHSH inequality [19]. In fact, every Bell inequality using correlation functions admits this form [16]. Several known Bell inequalities with arbitrary many outcomes and/or arbitrary many settings admit the form (3), see Refs. [13,18,20–22]. Also, known families of full-correlation two-outcome Bell inequalities with nonuniform input distributions [23] fall into the class (3). However, Bell inequalities with marginal probabilities are not taken into account. For example, the Clauser-Horne inequality [24] and the noncorrelation parts of the inequalities presented in [25] do not admit the form (3). To include also those inequalities, a further generalization would be necessary.

III. REPRESENTING BELL INEQUALITIES $\mathscr{B}$ AS PAYOFF BOUNDS FOR A CLASSICAL CCP

Consider the following family of CCPs, which we label $\mathbb{G}_{\mathscr{B}}$. Alice is given one input $x_0 \in \{0,\ldots,d-1\}$ with $p(x_0) = 1/d$, and one input $x \in \{0,\ldots,m_A - 1\}$, while Bob receives one input $y \in \{0,\ldots,m_B - 1\}$. The inputs x of Alice and y of Bob are distributed according to a joint probability distribution $p(x,y)$. There is a communication channel from Alice to Bob over which Alice may send at most $\log d$ bits of information in the form of a message $m(x_0,x) \in \{0,\ldots,d-1\}$. Having received the message, Bob outputs his guess $G(y,m) \in \{0,\ldots,d-1\}$. If G coincides with the value of one of the functions $f_{i,k}(x_0,x,y) = x_0 + F^i_{xy}(k) \mod d$, then Alice and Bob jointly earn a payoff $c^{i,k}_{xy}$. The average earned payoff in $\mathbb{G}_{\mathscr{B}}$ is

$$I^{\text{cc}}_{\mathbb{G}_{\mathscr{B}}} = \frac{1}{d} \sum_{x_0,x,y} p(x,y) \sum_{k=0}^{K} \sum_{i=1}^{N} c^{i,k}_{xy} P_{xy}[G = f_{i,k}(x_0,x,y)]. \qquad (4)$$

In a quantum version of such a CCP, to assist Alice and Bob's attempts to perform optimally, they may perform measurements on their subsystems in an entangled state. Alice performs a local measurement of a setting labeled by x and obtains the outcome $a \in \{0,\ldots,d-1\}$. Similarly, Bob performs a local measurement labeled by y and obtains the outcome $b \in \{0,\ldots,d-1\}$. Alice sends a message which depends on x_0 and a. The other method is that Alice sends to Bob a quantum system of dimension d in a state which depends on x_0 and x, upon which Bob performs a measurement of his choice, and somehow produces a guess.

There are many possible ways of implementing $\mathbb{G}_{\mathscr{B}}$ by choosing different ways of coding the message m and outputting the guess G. However, we shall limit the strategies under consideration to only those in which Bob's guess is of the form $m + b(y) \mod d$. In particular, we call any strategy linear, both in the case of classical and entanglement assisted CCPs, if $m = x_0 + a(x) \mod d$. Any other strategy of Alice we call nonlinear. We shall state and prove a theorem about the optimality of such linear strategies in $\mathbb{G}_{\mathscr{B}}$. Before that, we provide two useful lemmas.

Lemma 1. Take a function $R(x_0)$, where $x_0 = 0,1,\ldots,d-1$, such that its values are only in the set $\omega^0,\omega,\ldots,\omega^{d-1}$. The discrete Fourier transform of R, defined as $K(l,R) =$

$\frac{1}{d}\sum_{x_0=0}^{d-1}\omega^{-lx_0}R(x_0)$, has the following property: either it is such that (a) only for one value of l, say $l=s$, one has $K(s,R)\neq 0$, and then $K(s,R)$ is a power of ω; or (b) for every l the value $K(l,R)$ is some *convex* combination of some subsets of numbers $\omega^0,\omega,\ldots,\omega^{d-1}$.

Proof. The values of $R(x_0)$, are in the form $\omega^{n(x_0)}$, where n is an integer function of x_0. Its discrete Fourier transform is $K(l,R)=\frac{1}{d}\sum_{x_0=0}^{d-1}\omega^{n(x_0)}\omega^{-lx_0}$, which for every l is exactly such a convex combination. In particular, if $K(l,R)$ are not proper convex combinations then only one of them is nonzero. If this is the case for $K(l=s,R)$, then this is if, and only if, $R(x_0)=\omega^t\omega^{sx_0}$, where s,t are integers, and $K(l=s,R)=\omega^t$. ■

Lemma 2. The Fourier expansion coefficients of powers of function $R(x_0)$, that is, $R(x_0)^r$ where r is an integer, have the following form. Assume that for the function R the Fourier transform $K(l=1,R)$ is in the form of the following convex combination: $K(1,R)=\sum_{v=0}^{d-1}\lambda_v\omega^v$. Then the $l=r$ value of the Fourier transform of R^r is given by $K(l=r,R^r)=\sum_{v=0}^{d-1}\lambda_v\omega^{rv}$. That is, it is a convex combination of rth powers of ω^v, with the same coefficients λ_v, as $K(1,R)$.

Proof. The convex combination coefficients of $K(1,R)$, that is, λ_v's, in fact, are equal to $\frac{k_v}{d}$, where each $k_v=0,1,\ldots,d$ tells us how many times the number ω^v appears in $K(1,R)=\frac{1}{d}\sum_{x_0=0}^{d-1}\omega^{n(x_0)}\omega^{-x_0}$. Because $\omega^{ln(x_0)}\omega^{-lx_0}=(\omega^{n(x_0)}\omega^{-x_0})^l$ and $K(r,R^r)=\frac{1}{d}\sum_{x_0=0}^{d-1}\omega^{rn(x_0)}\omega^{-rx_0}$, we see that if ω^v appears k_v times in $K(1,R)$, so does ω^{rv} in $K(r,R^r)$. ■

Theorem 1. The optimal performance in classical $\mathbb{G}_{\mathscr{B}}$ is achieved with a linear strategy. Moreover, the performance of any nonlinear strategy is a probabilistic mixing of the performances of linear strategies.

Proof. We first rewrite our Bell inequalities in Eq. (3). The discrete Fourier transform of $P(a+b|x,y)$ can be defined as $E(l|x,y)=\sum_{z=0}^{d-1}\omega^{lz}P(a+b=z|x,y)$ where $\omega=e^{\frac{2\pi i}{d}}$. Its inverse reads

$$P(a+b=z|x,y)=\frac{1}{d}\sum_{l=0}^{d-1}\omega^{-lz}E(l|x,y). \tag{5}$$

Therefore we get

$$P\big(a+b=F^i_{xy}(k)|x,y\big)=\frac{1}{d}\sum_{l=0}^{d-1}\omega^{-lF^i_{xy}(k)}E(l|x,y). \tag{6}$$

By direct insertion into Eq. (3), we may write any Bell inequality $\mathscr{B}$ in the form

$$I^{\text{Bell}}\equiv\sum_{x,y}\frac{p(x,y)}{d}\sum_{l=0}^{d-1}\sum_{i=0}^{N}\sum_{k=0}^{K}c^{i,k}_{xy}\omega^{-lF^i_{xy}(k)}E(l|x,y)\leqslant B. \tag{7}$$

Next, notice that $E(l|x,y)$ is the average value of the products of the local results, each represented by specific powers of ω, for local settings x, y. Namely, $E(l|x,y)=\langle\omega^{la}\omega^{lb}\rangle_{x,y}$. Thus, for each l we have a different form of correlation function.

Having written $\mathscr{B}$ in terms of correlators, it is now straightforward to write down the performance (4) in $\mathbb{G}_{\mathscr{B}}$ in this terminology. The property of the dth roots of unity $\sum_{l=0}^{d-1}\omega^{l(z-q)}=d\delta_{z,q}$, where z,q are integers, allows one to put the logical value of the question of whether a guess G equals $f_{i,k}$ in the form of $\frac{1}{d}\sum_{l=0}^{d-1}\omega^{l(G-f_{i,k})}$. Thus the payoff of a $\mathbb{G}_{\mathscr{B}}$ class game, if the answer is G, is given by

$$I^{\text{cc}}_{\mathscr{B}}=\frac{1}{d}\sum_{x,y}p(x,y)\sum_{x_0=0}^{d-1}\sum_{l=0}^{d-1}\sum_{i=1}^{N}\sum_{k=0}^{K}\times c^{i,k}_{xy}\omega^{-lF^i_{xy}(k)}\omega^{-lx_0}\tilde{G}^*(x_0,x,y)^l, \tag{8}$$

where $\tilde{G}=\omega^G$ represents the guess output by Bob, transformed into a power of ω. Representation of the guess and the message in the form of powers of ω will play an important technical role in what follows.

Assume now that Alice and Bob apply some general strategy in $\mathbb{G}_{\mathscr{B}}$, i.e., $G=G(m(x_0,x),y)=m(x_0,x)+b(y)$. The guess $\tilde{G}_{xy}(x_0)=\omega^{G_{xy}(x_0)}=\tilde{M}_x(x_0)\omega^{b(y)}$ where $\tilde{M}_x(x_0)=\omega^{m(x_0,x)}$ is always equal to some integer power of ω. We shall analyze the function $\tilde{M}$ by treating x as an index for fixed values of which $\tilde{M}_x(x_0)$ is a function of x_0 only.

Notice that the part of the expression in Eq. (8) which depends on x_0 is $\frac{1}{d}\sum_{x_0=0}^{d-1}\omega^{-lx_0}\tilde{M}^*_x(x_0)^l$. We see that we have here the lth value of a discrete Fourier transform of $(\tilde{M}_x)^l$. In Lemmas 1 and 2 we saw that the Fourier transforms of powers of functions, which can have values only in the form of powers of ω, have very specific properties.

Using the lemmas, one can replace in Eq. (8) the expression $\sum_{x_0=0}^{d-1}\omega^{-lx_0}\tilde{M}^*_x(x_0)^l$ by $K(l,M^l_x)=\sum_{v=0}^{d-1}\lambda_v(x)\omega^{lv}$. Thus any strategy which is different from the linear one is effectively in terms of payoffs equivalent to a probabilistic strategy in which Alice with probabilities $\lambda_v(x)$ chooses the value of the message to be sent to Bob. Such probabilistic strategies are never better than the optimal deterministic one. In the case of a linear strategy we have situation (a) of Lemma 1, and thus it is deterministic. Obviously the bound for such a strategy is given by B. ■

Let us now move to the quantum strategies which use classical communication and correlations due to entanglement as a source for information processing, which supplies the partners with partially correlated random noise. The following theorem holds.

Theorem 2. The optimal quantum strategy based on classical communication assisted by entanglement for $\mathbb{G}_{\mathscr{B}}$ employs the linear strategy of messaging. It achieves its best performance identical to the maximal quantum violation for the associated Bell inequality $\mathscr{B}$.

Proof. Using a linear strategy, Bob effectively outputs $G=x_0+a+b \mod d$. In order to compute $f_{i,k}$, note that $G=f_{i,k}\Leftrightarrow a+b=F^i_{xy}(k)$. In particular, this strategy eliminates the dependence in Eq. (4) on x_0. Therefore, the average payoff becomes

$$I^{\text{cc}}_{\mathbb{G}_{\mathscr{B}}}=\sum_{x,y}p(x,y)\sum_{k=0}^{K}\sum_{i=1}^{N}c^{i,k}_{xy}P_{xy}\big[a+b=F^i_{xy}(k)\big]. \tag{9}$$

However, this is precisely the same as the left-hand side of Eq. (3). Since Theorem 1 asserts that linear strategies are optimal for implementing $\mathbb{G}_{\mathscr{B}}$, it follows from Eq. (3) that

$$I^{\text{cc}}_{\mathbb{G}_{\mathscr{B}}}\leqslant B, \tag{10}$$

and that the performance in $\mathbb{G}_{\mathscr{B}}$ with a strategy based on classical communication assisted by entanglement can achieve the maximal quantum violation of $\mathscr{B}$.

But is the linear classical messaging strategy also optimal in the entanglement-assisted protocol? Notice that in the case of a quantum protocol, we have to introduce an "answer" observable $\hat{Q}(x_0,x,y)$ of eigenvalues which are powers of ω. This is because Alice, if she follows a deterministic messaging strategy based on her measurement results, the setting of which are determined by her local data x, as well as directly on her data, must act as follows. She measures an observable $\hat{A}(x)$, and if her ith detector fires, she gets an eigenvalue $\xi_i(x)$, whatever it is. Therefore her message will be a function of ξ_i and x_0,x, in the form of $m(x_0,x,\xi_i(x))$. But this can be treated as a direct measurement of an observable $\hat{m} = m(x_0,x,\hat{A}(x))$, which as we know always commutes with $\hat{A}(x)$. Nondegenerate commuting observables differ only by their eigenvalues, but share projectors onto eigenstates. Any degenerate observable can always be put in a form which also has the above features.

The performance of the entanglement-assisted protocol is therefore measured by

$$I^{\mathrm{cc}}_{\mathscr{B}} = \frac{1}{d}\sum_{x,y} p(x,y) \sum_{x_0=0}^{d-1}\sum_{l=0}^{d-1}\sum_{i=1}^{N}\sum_{k=0}^{K} \times c^{i,k}_{xy}\omega^{-lF^{i}_{xy}(k)}\omega^{-lx_0}\,\mathrm{Tr}[\hat{Q}(x_0,x,y)^{l}\rho^{AB}], \quad (11)$$

where ρ^{AB} is the state. With the assumption that the guess of Bob has the structure $m + b(y)$, the structure of $\hat{Q}_{xy}(x_0)$ must be as follows:

$$\hat{Q}_{xy}(x_0) = \omega^{\hat{m}(x_0,x)} \otimes \omega^{\hat{b}(y)}, \quad (12)$$

where we have dropped $\hat{A}(x)$ in the argument of $\hat{m}$. The hats denote here local observables of integer eigenvalues. Just as in the classical case, the crucial point is the analysis of the operators given by the expression $\hat{M}(l,x) = \frac{1}{d}\sum_{x_0=0}^{d-1}\omega^{-lx_0}\omega^{\hat{m}(x_0,x)l}$. The message observable $\omega^{\hat{m}(x_0,x)}$ can be split into sum of projectors $\Pi^i(x)$ multiplied by the associated eigenvalues $\omega^{\eta_i(x_0,x_1,\xi_i)}$. Each η_i can be a different function of integer values. This represents the possible strategies of Alice, of how to form the message, once the result of her measurement of $\hat{m}$ is a collapse of the state given by the projector $\Pi^i(x)$. This reflects all possible value assignments to the obtained measurement results, represented by the projectors. Of course $\sum_{i=1}^{d}\Pi^i(x) = \hat{I}$, where $\hat{I}$ is the local identity operator. With all that, one has

$$\hat{M}(l,x) = \frac{1}{d}\sum_{x_0=0}^{d-1}\omega^{-lx_0}\sum_{i=1}^{d}\omega^{l\eta_i(x_0,x,\xi_i)}\Pi^i(x). \quad (13)$$

Therefore our analysis now moves to the properties of the "effective eigenvalue" $\frac{1}{d}\sum_{x_0=0}^{d-1}\omega^{-lx_0}\omega^{l\eta_i(x_0,x,\xi_i)}$. The messaging protocol strategies are defined the by the structure of the functions η_i. If one has $\omega^{\eta_i(x_0,x,\xi_i)} \neq \omega^{x_0}\omega^{\xi_i(x)}$, then just as in the classical case the effective eigenvalue which survives the summation over x_0 is a convex combination $\sum_{v=0}^{d-1}\lambda_v(x,i)\omega^{lv}$, where as before $\lambda_v = k_v/d$, and k_v tell us how many times in the sums ω^v is repeated in the sum $\sum_{x_0=0}^{d-1}\omega^{-lx_0}\omega^{l\eta_i(x_0,x,\xi_i)}$. The convex combination can be interpreted as a probabilistic mixture of eigenvalues which are powers of ω. Thus, it represents a probabilistic mixture of eigenvalue strategies. However, a mixture of strategies is never better than some deterministic strategy, which thus can give the upper bound of Eq. (11). Thus, the eigenvalues should read $\omega^{x_0}\omega^{\xi_i(x)}$. In such a case, our message observable $\omega^{\hat{m}(x_0,x)}$ factorizes to $\omega^{x_0}\sum_{i=1}^{d}\omega^{\xi_i(x)}\Pi^i(x)$. We get a linear strategy and the message, if detector i fires, is $x_0 + \xi_i(x)$. ■

Essentially, the linear strategy allows us to interpret x_0 as a scrambler that Alice uses to randomize her message, as Bob has no information whatsoever on $a(x)$ for the classical case or $\xi_i(x)$ for the quantum one. It is never unscrambled; however, the linear strategy allows Bob to guess effectively the functions $f_{i,k}$. This places the original Bell inequality $\mathscr{B}$ and the performance of the linear strategy in $\mathbb{G}_{\mathscr{B}}$ on equal footing: whenever quantum correlations can be used to achieve some value of I^{Bell} [Eq. (3)], they can be used to assist classical communication in $\mathbb{G}_{\mathscr{B}}$ such that $I^{\mathrm{cc}}_{\mathbb{G}_{\mathscr{B}}} = I^{\mathrm{Bell}}$, and vice versa.

IV. IMPLEMENTING COMMUNICATION COMPLEXITY REDUCTION PROTOCOLS WITH QUANTUM PREPARE-TRANSMIT-MEASURE STRATEGY

We now turn our attention to quantum implementations of $\mathbb{G}_{\mathscr{B}}$ with prepare-transmit-measure protocols. In such a scenario, Alice uses her input data (x_0,x) to prepare a physical state of information content at most $\log d$ bits, i.e., a density matrix ρ_{x_0x} of a d-dimensional system. She sends the system to Bob who performs a measurement on it using an observable, the choice of which is dictated by y, and obtains an outcome b_y. We can easily transform the performance metric of $\mathbb{G}_{\mathscr{B}}$ in Eq. (4) to this alternative implementation in a prepare-transmit-measure scenario. Whenever the output of Bob satisfies $b_y = f_{i,k}(x_0,x,y)$ the partnership earns a payoff $c^{i,k}_{xy}$. The average earned payoff is

$$I^{qc}_{\mathbb{G}_{\mathscr{B}}} = \frac{1}{d}\sum_{x_0=0}^{d-1}\sum_{x,y} p(x,y)\sum_{i=1}^{N}\sum_{k=0}^{K} c^{i,k}_{xy}P(b_y = f_{i,k}|\rho_{x_0x},y). \quad (14)$$

Thus, since $\mathbb{G}_{\mathscr{B}}$ implemented with entanglement and classical communication always can be implemented also with a prepare-transmit-measure quantum scenario, we can use the considered CCPs as a framework in which we can speak about the two types of quantum protocols on equal footing.

V. ENTANGLEMENT VS TRANSMISSION OF A QUANTUM SYSTEM

We make two limiting assumptions: **AI**, for any Bell inequality $\mathscr{B}$ we consider situations in which it is violated by quantum predictions, which are achievable with some sets of d-outcome measurements of Alice and Bob on entangled systems in a state $\rho^{AB} \in \mathbb{C}^D \otimes \mathbb{C}^d$ for some integer $D \geqslant d$; and **AII**, we consider only such measurements and states used to achieve the maximal violation of the classical bound of the inequality $\mathscr{B}$ for which the following holds: local measurements of whichever observable in Alice's set give uniformly random local results.

Theorem 3. Assume **AI** and **AII**. With any given correlation due to entanglement we can associate a prepare-transmit-measure protocol which achieves $I^{cc}_{\mathbb{G}_{\mathscr{B}}} = I^{qc}_{\mathbb{G}_{\mathscr{B}}}$.

Proof. We have already shown that in quantum theory, for given sets of measurements and a given state, the maximal value of I^{Bell} is the same as that of $I^{cc}_{\mathbb{G}_{\mathscr{B}}}$. Thus, let us study quantum violations of the Bell inequality $\mathscr{B}$. Consider the state, ρ^{AB}, and the measurements used to achieve a violation of $\mathscr{B}$. The projector $\Pi^i(x)$ of Alice is associated with her measurement setting x and her outcome i. Now, in the prepare-transmit-measure protocol of $\mathbb{G}_{\mathscr{B}}$, we define the preparations of Alice as the local states of Bob in the Bell scenario after Alice's local measurement, i.e.,

$$\varrho_{i,x} = d\,\mathrm{Tr}_A(\Pi^i(x) \otimes \mathbf{1}\rho^{AB}). \tag{15}$$

Note that because of assumption **AI**, communication of the states in Eq. (15) is always possible. Because of assumption **AII**, we have $p(i|x) = 1/d$. Remember that $p(x_0|x) = 1/d$ was a premise when we defined $\mathbb{G}_{\mathscr{B}}$. Therefore in the prepare-transmit-measure scenario we define the set of Alice's states as $\rho_{x_0 x} = \varrho_{i=x_0,x}$. If Bob performs the same measurements as those used to achieve the violation of $\mathscr{B}$ it follows by construction that there is an analogous violation of Eq. (14) yielding $I^{cc}_{\mathbb{G}_{\mathscr{B}}} = I^{qc}_{\mathbb{G}_{\mathscr{B}}}$. ■

However, the opposite of Theorem 3 need not be true. Simply by giving some suitable alterations to some particular states in the set of preparations $\{\rho_{x_0 x}\}$, we would not be able to reproduce the communicated states by local measurements on an entangled state. This leads to a qualitative relation between the two types of quantum resources:

$$\textbf{AI and AII} \Rightarrow I^{qc}_{\mathbb{G}_{\mathscr{B}}} \geqslant I^{cc}_{\mathbb{G}_{\mathscr{B}}}. \tag{16}$$

For any Bell inequality satisfying the given assumptions, prepare-transmit-measure methods are at least as powerful as correlations due to entanglement. Of course, from our discussion so far, it is not necessarily the case that a strict inequality can be observed. However, case studies [13,14] based on two different families of Bell inequalities satisfying assumptions **AI** and **AII** have revealed multiple such examples.

However, if we are given a Bell inequality that does not fulfill both **AI** and **AII**, we may find that quantum correlations due to entangled states are more powerful than prepare-transmit-measure protocols. In fact, for any Bell inequality $\mathscr{B}$ with binary outcomes that achieves its maximal quantum violation for an entangled state of two D-level quantum systems with $D > 2$, entanglement is a strictly stronger resource than the preparation-transmission-measurement method with a qubit in $\mathbb{G}_{\mathscr{B}}$. To show this, note that it was shown in Ref. [11] that for any CCP with binary answers, entanglement is as least as good a resource as transmission of a single qubit. Note also that $\mathbb{G}_{\mathscr{B}}$ are such CCPs when $\mathscr{B}$ has binary outcomes, i.e., when $d = 2$. When the maximal quantum violation of $\mathscr{B}$ is obtained from an entangled state with $D > 2$, a strict inequality follows immediately from the fact that the state of Bob after Alice's measurement cannot be reproduced by sending a qubit. Explicit examples of such Bell inequalities in which prepare-transmit-measure protocols are weaker than their entanglement-assisted counterparts have been given in Refs. [12,26].

VI. DISCUSSION

We have introduced a framework based on games for studying the ability of quantum correlations obtained from entangled states to assist information processing tasks, as compared to that of prepare-transmit-measure protocols involving only a single-quantum system. Importantly, concerning the former resource, we showed that the performance in our CCPs is analogous to the ability of quantum theory to violate a Bell inequality. This opens the door for systematic studies of the comparative nonclassical abilities of the two quantum resources. In particular, we show that for CCPs corresponding to a large class of Bell inequalities, the degree of achievable nonclassicality using a prepare-transmit-measure protocol is as least as much as an entanglement-assisted strategy. Previous case studies [12–14] further support the potential richness of the relation between the two types of quantum protocols. Furthermore, the part of our work concerning correlations due to entanglement can be understood as a generalization of the results of Ref. [16] from two-outcome to many-outcome Bell inequalities. Additionally, we presented a proof of the optimality of linear messaging strategies, which was missing in Ref. [16].

From the point of view of possible applications, we note that using our mapping between Bell inequalities and CCPs one can systematically transform many certificates of genuine nonclassical behavior in device-independent entanglement-assisted protocols to analogous semi-device-independent prepare-transmit-measure protocols. Typically, such semi-device-independent protocols are somewhat less secure but more efficient than their device-independent counterparts. However, due to our relation in Eq. (16), one may obtain further advantages in the efficiency of semi-device-independent information processing tasks from the fact that CCPs in a prepare-transmit-measure scheme can to a further extent outperform the classical bound as compared to Bell inequality violations.

Our work leaves multiple open questions of which we mention some of the more challenging ones: (1) Further qualitative and quantitative characterization of the relation between correlations due to entanglement and protocols based on preparations and measurements of single-uantum systems is a key open problem in understanding the extent of nonclassicality enabled by quantum theory. (2) We have only considered bipartite Bell inequalities. Can the mapping between Bell inequalities and CCPs be extended to multipartite scenarios? How will prepare-transmit-measure protocols behave in such scenarios when intermediate partners appear in the chain of communication? (3) In recent years, much effort has been directed at characterizing Bell-type quantum correlations from information-theoretic principles. Our results suggest that similar attempts to understand the correlations due to single-quantum systems may be of interest.

ACKNOWLEDGMENTS

A.T. acknowledges financial support from the Swiss National Science Foundation (Starting grant DIAQ). M.Z. is supported by the ERC advanced grant QOLAPS, and COPERNICUS grant-award of Deutsche Forschungsgemeinschaft/FNP.

[1] J. S. Bell, Physics **1**, 195 (1964).
[2] H. Buhrman, W. van Dam, P. Høyer, and A. Tapp, Phys. Rev. A **60**, 2737 (1999); H. Buhrman, R. Cleve, S. Massar, and R. de Wolf, Rev. Mod. Phys. **82**, 665 (2010)
[3] E. F. Galvão, Phys. Rev. A **65**, 012318 (2001); P. Trojek, C. Schmid, M. Bourennane, C. Brukner, M. Zukowski, and H. Weinfurter, *ibid.* **72**, 050305(R) (2005).
[4] A. Acín, N. Gisin, and L. Masanes, Phys. Rev. Lett. **97**, 120405 (2006).
[5] M. Pawłowski and N. Brunner, Phys. Rev. A **84**, 010302(R) (2011).
[6] S. Pironio, A. Acín, S. Massar, A. Boyer de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe, Nature **464**, 1021 (2010).
[7] H.-W. Li, M. Pawłowski, Z.-Q. Yin, G.-C. Guo, and Z.-F. Han, Phys. Rev. A **85**, 052308 (2012).
[8] P. Mironowicz, A. Tavakoli, A. Hameedi, B. Marques, M. Pawłowski, and M. Bourennane, New J. Phys. **18**, 065004 (2016).
[9] N. Brunner, S. Pironio, A. Acín, N. Gisin, A. A. Methot, and V. Scarani, Phys. Rev. Lett. **100**, 210503 (2008).
[10] R. Gallego, N. Brunner, C. Hadley, and A. Acín, Phys. Rev. Lett. **105**, 230501 (2010).
[11] M. Pawłowski and A. Winter, Phys. Rev. A **85**, 022331 (2012).
[12] M. Pawłowski and M. Żukowski, Phys. Rev. A **81**, 042326 (2010).
[13] A. Tavakoli, B. Marques, M. Pawłowski, and M. Bourennane, Phys. Rev. A **93**, 032336 (2016).
[14] A. Tavakoli, M. Pawłowski, M. Żukowski, and M. Bourennane, Phys. Rev. A **95**, 020302(R) (2017).
[15] N. Brunner and N. Linden, Nat. Commun. **4**, 2057 (2013).
[16] C. Brukner, M. Żukowski, J.-W. Pan, and A. Zeilinger, Phys. Rev. Lett. **92**, 127901 (2004).
[17] C. Brukner, M. Żukowski, and A. Zeilinger, Phys. Rev. Lett. **89**, 197901 (2002); H. Buhrman, L. Czekaj, A. Grudka, M. Horodecki, P. Horodecki, M. Markiewicz, Fl. Speelman, and S. Strelchuk, Proc. Natl. Acad. Sci. USA **113**, 3191 (2016).
[18] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu, Phys. Rev. Lett. **88**, 040404 (2002).
[19] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).
[20] A. Salavrakos, R. Augusiak, J. Tura, P. Wittek, A. Acín, and S. Pironio, arXiv:1607.04578.
[21] J. Barrett, A. Kent, and S. Pironio, Phys. Rev. Lett. **97**, 170409 (2006).
[22] A. Tavakoli, S. Zohren, and M. Pawłowski, J. Phys. A: Math. Theor. **49**, 145301 (2016).
[23] T. Lawson, N. Linden, and S. Popescu, arXiv:1011.6245.
[24] J. F. Clauser and M. A. Horne, Phys. Rev. D **10**, 526 (1974).
[25] D. Collins and N. Gisin, J. Phys. A: Math. Gen. **37**, 1775 (2004).
[26] J. Oppenheim and S. Wehner, Science **19**, 330 (2010).

PHYSICAL REVIEW A **96**, 020304(R) (2017)

All entangled pure quantum states violate the bilocality inequality

Nicolas Gisin, Quanxin Mei, Armin Tavakoli, Marc Olivier Renou, and Nicolas Brunner
Groupe de Physique Appliquée, Université de Genève, CH-1211 Genève, Switzerland
(Received 1 February 2017; published 22 August 2017)

The nature of quantum correlations in networks featuring independent sources of entanglement remains poorly understood. Here, focusing on the simplest network of entanglement swapping, we start a systematic characterization of the set of quantum states leading to violation of the so-called "bilocality" inequality. First, we show that all possible pairs of entangled pure states can violate the inequality. Next, we derive a general criterion for violation for arbitrary pairs of mixed two-qubit states. Notably, this reveals a strong connection between the Clauser-Horne-Shimony-Holt (CHSH) Bell inequality and the bilocality inequality, namely, that any entangled state violating CHSH also violates the bilocality inequality. We conclude with a list of open questions.

DOI: 10.1103/PhysRevA.96.020304

Introduction. Quantum nonlocality, in the sense of violation of a Bell inequality, was considered as a mere curiosity—when not entirely ignored—during several decades after Bell's seminal work [1]. Things changed dramatically in the early 1990s when Ekert showed that nonlocality can be exploited to establish cryptographic keys between two remote observers [2]. How could one ignore something useful for cryptography, especially in our information-based society? Moreover, also in the early 1990s, experiments showed that the violation of Bell inequalities can be demonstrated over several kilometers using special optical fibers [3] and even outside the controlled laboratory environment using standard telecom fibers [4]. This led to rapid developments, both conceptually and for applications. Today, Bell inequality violation is routinely used in order to demonstrate the presence of entanglement in some physical systems. This demonstrates quantumness beyond any doubt.

In the context of applications, quantum nonlocality led to the development of the field of device-independent quantum information processing [5], a way of processing information requiring no assumption about the details of the physical implementation; not even the dimension of the Hilbert space in which the quantum systems are represented. The measurement statistics suffice to guarantee security for generating, e.g., cryptographic keys [6], or random numbers [7,8]. It is impressive that NIST has already made available online a beta version of a randomness beacon that will soon be offered to the public [9].

In the conceptual context, novel developments in quantum nonlocality have been inspired by experimental work on quantum networks. In such networks, there is not just one source of entanglement (the resource exploited for Bell inequality violation), but several sources distributing entanglement between different nodes, which can perform joint quantum measurements [10]. This leads to strong correlations across the entire network. The understanding of such correlations is highly desirable, although still very limited at the moment.

The simplest example of a joint quantum measurement is the so-called Bell state measurement (BSM), a central ingredient in quantum teleportation [11] and in entanglement swapping [12]. Formally, the BSM is represented by its four eigenvectors, namely, the Bell states:

$$|\phi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|0,0\rangle \pm |1,1\rangle), \tag{1}$$

$$|\psi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|0,1\rangle \pm |1,0\rangle), \tag{2}$$

hence referred to as a joint (or entangled) measurement. Since all Bell states are maximally entangled, their marginals are given by the maximally mixed state. Consequently, when one performs a BSM on independent qubits, all four results are equally likely, i.e., 25% probability for each.

Figure 1 illustrates the simplest quantum network, with only three observers and two sources. This is the scenario we consider in this Rapid Communication. In the standard analysis of this scenario, i.e., following Bell locality, one would contrast the correlations achievable with quantum resources, e.g., two sources of entangled pairs and the BSM in the middle, with classical resources, i.e., all three parties share some common local hidden variable (LHV). Note that "local hidden variable" is the old terminology, going back to Einstein *et al.* [13] and Bell [1]. Nowadays one refers to shared randomness, a terminology closer to cryptography, although technically synonymous. Hence, all three parties—named Alice, Bob, and Charlie—would share a common classical random variable.

However, looking at Fig. 1, it is arguably much more natural to contrast quantum correlations with classical correlations achievable via two independent sources of shared randomness. More precisely, Alice and Bob would share some variable λ (originating from the source between them), while Bob and Charlie would share another variable μ (originating from the second source). Importantly the variables λ and μ should be uncorrelated, as the two sources are independent. This independence assumption is very natural, given that the quantum network of Fig. 1 features two fully independent sources of entanglement. There is thus no reason to assume that λ and μ are correlated. This very natural assumption changes everything.

This new scenario has been studied under the name of 2-locality (2- because of the two sources) or merely bilocality. More formally, 2-local correlations are characterized as follows. Consider that Alice receives measurement setting (or input) x, while Bob gets input y, and Charlie z. Upon receiving their inputs, each party should provide a measurement result (an output), denoted A for Alice, B for Bob, and C for Charlie. In this context, the observed statistics is said to be 2-local when

$$p(ABC|xyz) = \int d\lambda\, d\mu\, q_1(\lambda) q_2(\mu)\, p(A|x\lambda)\, p(B|y\lambda\mu)\, p(C|z\mu),$$

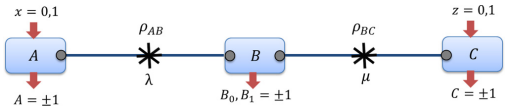

FIG. 1. Scenario of bilocality, the network we consider in this work. In the quantum setting, two independent sources distribute entangled states, ρ_{AB} and ρ_{BC}, between three distant observers—Alice, Bob, and Charlie. In order to compare the resulting quantum correlations to classical ones, we discuss 2-local correlations obtained by two independent sources of shared classical random variables, λ and μ. For the bilocality inequality we consider, Alice and Charlie perform two dichotomic measurements, while Bob performs a fixed measurement with four possible outcomes. In the quantum setting, Bob's measurement is taken to be the Bell state measurement.

where λ and μ are the independent shared random variables distributed according to the densities $q_1(\lambda)$ and $q_2(\mu)$, respectively. The set of 2-local correlations (i.e., the set of all correlations of the above form) is nonconvex [14], rendering its analysis challenging. In particular, in order to efficiently characterize the 2-local set, nonlinear Bell inequalities are required. Note that this is in stark contrast to the set of Bell-local (or 1-local) correlations which is convex and can thus be fully characterized by linear Bell inequalities [5].

In Refs. [14,15], first nonlinear inequalities that allow one to efficiently capture 2-local correlations (better than any linear inequality) were derived. Here we focus on an inequality presented in [15], which we will refer to as the bilocality inequality (for simplicity). Consider that Alice and Charlie receive binary inputs, $x = 0,1$ and $z = 0,1$, and must give binary outputs, denoted $A_x = \pm 1$ and $C_z = \pm 1$, respectively. The middle party Bob always performs the same measurement (hence receives no input y) with four possible outcomes, as, e.g., the BSM. Denote Bob's outcome by two bits $B_0 = \pm 1$ and $B_1 = \pm 1$. The bilocality inequality reads

$$S_{\text{biloc}} \equiv \sqrt{|I|} + \sqrt{|J|} \leqslant 2, \tag{3}$$

where

$$I \equiv \langle (A_0 + A_1) B_0 (C_0 + C_1) \rangle, \tag{4}$$

$$J \equiv \langle (A_0 - A_1) B_1 (C_0 - C_1) \rangle. \tag{5}$$

The bracket $\langle \cdot \rangle$ denotes the expectation value of many experimental runs.

Interestingly, this inequality can be violated by certain quantum correlations [15], which would have to be considered local in the usual Bell approach (i.e., when all three parties could have common shared randomness). In particular, consider the case where Alice-Bob, as well as Bob-Charlie, share a noisy Bell state (with visibility V), a so-called Werner state, of the form $\rho = V|\phi^+\rangle\langle\phi^+| + (1-V)\frac{\mathbb{I}}{4}$. Conditioned on one outcome of Bob's BSM, the state shared by Alice and Charlie is again a Werner state, but with lower visibility V^2. The bilocality inequality can be violated whenever $V^2 > 1/2$. This is in strong contrast with the usual Bell approach, where in order to detect quantum nonlocality, one would require a visibility $V > 1/\sqrt{2}$ using the Clauser-Horne-Shimony-Holt (CHSH) [16] Bell inequality,[1] while for visibilities up to $V \simeq 0.682$ the Werner state admits a LHV model [18] and can thus not violate any Bell inequality.[2]

The above results demonstrated the relevance of the 2-locality approach for detecting quantum correlations in networks. This triggered further research. On the theory side, novel nonlinear inequalities were derived and more sophisticated networks were considered (see, e.g., [19–27]). On the experimental side, violations of the bilocality inequality were demonstrated [28,29]. However, the extent of quantum correlations in networks remains poorly understood. This is precisely the goal of the present work, where we start a systematic characterization of the class of quantum states leading to violation of the bilocality inequality (3).

All pairs of pure entangled states. We start our analysis by considering that both sources emit pure entangled states. Denote $|\psi_{AB}\rangle = c_0|00\rangle + c_1|11\rangle$ and $|\phi_{BC}\rangle = q_0|00\rangle + q_1|11\rangle$ the two normalized (two-qubit) pure states shared by Alice and Bob and by Bob and Charlie, respectively, written in the Schmidt basis, with real and positive coefficients c_j and q_j. Note that if these Schmidt bases would differ from the computational basis in which the BSM (1) is written, then it would suffice to add local unitary rotations on each qubit to recover the case we discuss here. Define $c = 2c_0c_1$ and $q = 2q_0q_1$; $|\psi_{AB}\rangle$ ($|\phi_{BC}\rangle$) are entangled whenever $c > 0$ ($q > 0$). Note that we can restrict to two-qubit entangled states here. If the states are of larger dimension, Alice, Bob, and Charlie can first project them onto qubit subspaces, hence our setting is fully general for the case of two pure states [30].

Let Alice's inputs correspond to projective measurements in the Z-X plane of the Bloch sphere. Thus each measurement can be characterized by one angle; in fact, it is straightforward to see that optimal settings are given by angles $\pm\alpha$ symmetric with respect to the Z axis. The observable corresponding to the first input reads $\vec{a} \cdot \vec{\sigma}$, where $\vec{a} = [\sin(\alpha), 0, \cos(\alpha)]$ and $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ denotes the vector of Pauli matrices. Similarly for Charlie, we have angles $\pm\gamma$. Bob performs the usual BSM. For all $x,z = 0,1$ one gets

$$\begin{aligned}\langle A_x B_0 C_z \rangle &= \langle [\cos(\alpha)\sigma_z + (-1)^x \sin(\alpha)\sigma_x] \otimes (\sigma_z \otimes \sigma_z) \\ &\quad \otimes [\cos(\gamma)\sigma_z + (-1)^z \sin(\gamma)\sigma_x] \rangle_{\psi_{AB} \otimes \phi_{BC}} \\ &= \cos(\alpha)\cos(\gamma). \end{aligned} \tag{6}$$

Hence $I = 4\cos(\alpha)\cos(\gamma)$. A similar calculation gives $J = 4\sin(\alpha)\sin(\gamma)cq$.

Maximizing expression (3) with respect to α and γ leads to

$$\cos(\alpha) = \cos(\gamma) = \frac{1}{\sqrt{1+cq}} \tag{7}$$

[1] Note that one could do marginally better ($V \simeq 0.705$) by using an inequality introduced by Vértesi [17].

[2] Notice that this does not allow one to reveal the nonlocality of a Werner state ρ with $V \leqslant 1/\sqrt{2}$ by distributing two copies of ρ in the considered network and violate the bilocality inequality (see discussion in the *Conclusions*) However, it does constitute a significant advantage as compared to entanglement swapping experiments based on the CHSH Bell inequality.

and the maximum takes the value

$$S_{\text{biloc}}^{\max} = \sqrt{4\cos(\alpha)\cos(\gamma)} + \sqrt{4\sin(\alpha)\sin(\gamma)cq} = 2\sqrt{1+cq}\,. \tag{8}$$

Accordingly, for all possible pairs of entangled pure states, i.e., when $c > 0$ and $q > 0$, we get violation the standard bilocality inequality (3) and thus nonbilocal correlations.

Note that if $|\psi_{AB}\rangle = |\phi_{BC}\rangle$, then the optimal settings α and γ for bilocality are the same as the optimal settings for the CHSH inequality. Furthermore, $S_{\text{biloc}}^{\max}$ takes the same value as the maximum CHSH value for $|\psi_{AB}\rangle$ [30].

Interestingly, if the states differ, then Alice's optimal settings depend on the state $|\phi_{BC}\rangle$ shared by Bob and Charlie, and similarly Charlie's optimal settings depend on $|\psi_{AB}\rangle$, as can be seen from Eq. (7).

Note that if one would now consider noisy states of the form $V_{AB}|\psi_{AB}\rangle\langle\psi_{AB}| + (1 - V_{AB})\mathbb{1}/4$ and similarly for $V_{BC}|\phi_{BC}\rangle\langle\phi_{BC}| + (1 - V_{BC})\mathbb{1}/4$, then one can characterize the critical visibilities (V_{AB}^{biloc} and V_{BC}^{biloc}), i.e., the minimum visibilities for which violation of the bilocality inequality is still possible, which are in general related. More precisely, one finds that the product of the critical visibilities (for bilocality) is larger than the product of the visibilities for Bell locality (i.e., 1-locality): $V_{AB}^{\text{biloc}} V_{BC}^{\text{biloc}} = \frac{1}{1+cq} \geqslant V_{AB}^{\text{loc}} V_{BC}^{\text{loc}} = \sqrt{\frac{1}{1+c^2}}\sqrt{\frac{1}{1+q^2}}$, with equality holding only when $c = q$, i.e., when the two states are equal, $|\psi_{AB}\rangle = |\phi_{BC}\rangle$.

Criterion for arbitrary pairs of mixed states. We now move to mixed states, and start our analysis with the case of two-qubit density matrices. Let

$$\rho_{AB} = \frac{1}{4}\left(\mathbb{1} + \vec{m}_A \cdot \vec{\sigma} \otimes \mathbb{1} + \mathbb{1} \otimes \vec{m}_B \cdot \vec{\sigma} + \sum_{ij} t_{ij}^{AB} \sigma_i \otimes \sigma_j\right)$$

be the state shared by Alice and Bob, expressed in the Pauli basis; here the vector $\vec{m}_A$ ($\vec{m}_B$) represents the Bloch vector of Alice's (Bob's) reduced state, while t_{ij}^{AB} (with $i, j \in \{x, y, z\}$) is the correlation matrix. Similarly we express ρ_{BC}, the state shared by Bob and Charlie, in the Pauli basis.

Alice's settings are represented by Bloch vectors $\vec{a}$ and $\vec{a}'$, and similarly for Charlie $\vec{c}$ and $\vec{c}'$. Assume Bob performs a BSM in a well chosen basis to be defined below. The quantity I in Eq. (4) can be expressed as follows:

$$\begin{aligned} I &= \text{Tr}[(\vec{a} + \vec{a}') \cdot \vec{\sigma} \otimes \sigma_z \otimes \sigma_z \otimes (\vec{c} + \vec{c}') \cdot \vec{\sigma}\ \rho_{AB} \otimes \rho_{BC}] \\ &= \text{Tr}[(\vec{a} + \vec{a}') \cdot \vec{\sigma} \otimes \sigma_z\ \rho_{AB}]\,\text{Tr}[\sigma_z \otimes (\vec{c} + \vec{c}') \cdot \vec{\sigma}\ \rho_{BC}] \\ &= \sum_i (a_i + a_i') t_{iz}^{AB} \sum_k t_{3k}^{BC} (c_k + c_k'). \end{aligned} \tag{9}$$

Using the polar decomposition, the correlation matrix can be written as $t^{AB} = U^{AB} R^{AB}$, where U^{AB} is a unitary matrix and $R^{AB} = \sqrt{t^{AB\dagger} t^{AB}} \geqslant 0$. Denote $\xi_1 \geqslant \xi_2 \geqslant \xi_3 \geqslant 0$ the three non-negative eigenvalues of R^{AB}. Similarly denote $\zeta_1 \geqslant \zeta_2 \geqslant \zeta_3 \geqslant 0$ the non-negative eigenvalues of the corresponding matrix R^{BC}.

This allows us to characterize Bob's BSM. Specifically, the Bell states [as given in Eq. (1)] has been defined such that the Z and X Bloch directions (on the first subsystem, connected to Alice) are given by the eigenvectors of the matrix R^{AB} corresponding to the two largest eigenvalues, ξ_1 and ξ_2, respectively. Similarly we use R^{BC} for aligning the second subsystem of Bob, connected to Charlie. Note that the Z and X axes Bob uses with Alice may differ from those he uses with Charlie, i.e., Bob may have to apply different unitaries to the two qubits he shares with Alice and with Charlie before performing a standard BSM.

Next our goal is to maximize S_{biloc} with respect to the Bloch vectors $\vec{a}$, $\vec{a}'$, $\vec{c}$, and $\vec{c}'$. It is clear that they should lie within the two-dimensional subspace spanned by the two eigenvectors with largest eigenvalues: $\vec{a} = (\sin\alpha, 0, \cos\alpha)$, $\vec{a}' = (\sin\alpha', 0, \cos\alpha')$, $\vec{c} = (\sin\gamma, 0, \cos\gamma)$, and $\vec{c}' = (\sin\gamma', 0, \cos\gamma')$. The maximum is easily found by imposing $\partial_\alpha S = 0$, $\partial_{\alpha'} S = 0$, $\partial_\gamma S = 0$, and $\partial_{\gamma'} S = 0$. One finds $\alpha' = -\alpha$, $\gamma' = -\gamma$ and

$$\cos\alpha = \cos\gamma = \sqrt{\frac{\xi_1\zeta_1}{\xi_1\zeta_1 + \xi_2\zeta_2}}, \tag{10}$$

and the maximal value of the left-hand side of the bilocality inequality

$$S_{\text{biloc}}^{\max} = 2\sqrt{\xi_1\zeta_1 + \xi_2\zeta_2}. \tag{11}$$

Consequently, a pair of states ρ_{AB} and ρ_{BC} can violate the bilocality inequality (3) if and only if $\xi_1\zeta_1 + \xi_2\zeta_2 > 1$. Note that for the case of two pure states considered previously, $\xi_1 = \zeta_1 = 1$, $\xi_2 = 2c_0c_1 = c$, and $\zeta_2 = 2q_0q_1 = q$; hence (11) reduces to (8), as it should.

The above criterion is analogous to the Horodecki criterion for violation of the CHSH Bell inequality [31]. In fact, there is a direct connection between the two criteria. According to the Horodecki criterion the maximal CHSH value for ρ_{AB} is given by $S_{AB}^{\max} = 2\sqrt{\xi_1^2 + \xi_2^2} = 2\|\vec{\xi}\|$ where $\vec{\xi} = (\xi_1, \xi_2)$. Similarly, for ρ_{BC} we have $S_{BC}^{\max} = 2\sqrt{\zeta_1^2 + \zeta_2^2} = 2\|\vec{\zeta}\|$ with $\vec{\zeta} = (\zeta_1, \zeta_2)$. From Eq. (11) it follows that

$$S_{\text{biloc}}^{\max} = 2\sqrt{\vec{\xi} \cdot \vec{\zeta}} \leqslant 2\sqrt{\|\vec{\xi}\|\,\|\vec{\zeta}\|} = \sqrt{S_{AB}^{\max}\, S_{BC}^{\max}}. \tag{12}$$

Hence, violation of the bilocality inequality implies that either ρ_{AB} or ρ_{BC} (or both) must violate CHSH. Moreover, when the two states are the same, i.e., $\rho_{AB} = \rho_{BC} = \rho$, the criterion of Eq. (11) reduces to the Horodecki criterion. This is easily seen from Eq. (12), where the inequality becomes an equality when the vectors $\vec{\xi}$ and $\vec{\zeta}$ are the same. Therefore, CHSH violation implies violation of the bilocality inequality in the sense that

$$\rho \text{ violates CHSH} \;\rightarrow\; \rho \otimes \rho \text{ violates } S_{\text{biloc}}. \tag{13}$$

Note that under the assumption that Bob performs the BSM, the reverse link also holds. In this case activation of nonlocality is thus impossible for two-qubit entangled states using the bilocality inequality (see discussion below). Note also that the connection (13) holds true for arbitrary bipartite mixed states ρ, not only for two-qubit states [33].

Conclusion. In quantum networks involving several independent sources of entangled states, it is natural to contrast the obtained quantum correlations with "classical" correlations that can be realized using independent sources of shared

randomness between the observers. Indeed, this picture is arguably a natural generalization to networks of John Bell's original intuition [1,32]. In the simplest case, i.e., with two independent sources as in entanglement swapping, we analyzed the standard bilocality inequality and proved that all pairs of entangled pure states can violate it, in analogy to the case of the CHSH-Bell inequality which can be violated by any pure entangled state. Moving to mixed entangled states, we then derived a general criterion for violation of the bilocality inequality, providing a natural extension of the Horodecki criterion for violation of CHSH. In particular, this reveals a strong connection between CHSH and the bilocality inequality, namely, that any entangled state violating CHSH can also be used to demonstrate violation of the bilocality inequality.

While the results presented in this Rapid Communication were obtained analytically, we conclude with a list of open questions that we could so far tackle only numerically:

(1) Here we assumed that Bob performs a BSM, defined in local basis depending on the shared entangled states. One may expect that this is always optimal, which is confirmed numerically for any pair of pure states. However, numerical evidence suggests that there are cases, in which one or both states are mixed, for which no BSM is optimal. So far, we could not find any structure in the optimal joint measurements and leave it for future work.

(2) The bilocality inequality (3) used here assumes a scenario in which Bob has no choice of input and four possible outcomes. However, an inequality formally identical to (3) is also valid for the scenario in which Bob has a choice between two inputs with binary outcomes: it suffices to label B_0 and B_1 the outcomes corresponding to the two inputs, respectively. The bilocal bound of the inequality remains the same (because classically Bob could always compute and output both the value of B_0 and of B_1). However, quantum mechanically, Bob's two joint measurements may be incompatible, leading possibly to larger violations. We could confirm this possibility, although only numerically so far.

(3) It would be interesting to generalize the present results to the case of more sophisticated networks, such as star networks [21] with an arbitrary number of branches.

(4) A central open question is the possibility to activate the nonlocality of certain entangled quantum states—admitting a LHV model in the usual Bell scenario—by placing several copies of them in a network. While such effect is possible even when considering the standard Bell approach [34] (see also Ref. [35]), intuition suggests that the notion of N locality should be very useful in this context. However, no examples have been reported so far. Here, we have proven that activation is impossible for the bilocality inequality when Bob performs the BSM. We also performed intensive numerical search considering more general measurements for Bob. The results suggest that activation is impossible for the bilocality inequality. A formal proof of this statement would be desirable. A counterexample would be even more interesting.

Acknowledgments. This work was supported by the Swiss National Science Foundation (SNSF 200021-149109, Starting grant DIAQ, and QSIT), and the European Research Council (ERC-AG MEC).

[1] J. S. Bell, Physics **1**, 195 (1964).
[2] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).
[3] P. R. Tapster, J. G. Rarity, and P. C. M. Owens, Phys. Rev. Lett. **73**, 1923 (1994).
[4] W. Tittel, J. Brendel, H. Zbinden, and N. Gisin, Phys. Rev. Lett. **81**, 3563 (1998).
[5] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Rev. Mod. Phys. **86**, 419 (2014).
[6] A. Acin, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Phys. Rev. Lett. **98**, 230501 (2007).
[7] R. Colbeck, Ph.D. thesis, University of Cambridge, 2007.
[8] S. Pironio *et al.*, Nature (London) **464**, 1021 (2010).
[9] See https://www.nist.gov/programs-projects/nist-randomness-beacon.
[10] H. J. Kimble, Nature (London) **453**, 1023 (2008).
[11] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Phys. Rev. Lett. **70**, 1895 (1993).
[12] M. Żukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert, Phys. Rev. Lett. **71**, 4287 (1993).
[13] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
[14] C. Branciard, N. Gisin, and S. Pironio, Phys. Rev. Lett. **104**, 170401 (2010).
[15] C. Branciard, D. Rosset, N. Gisin, and S. Pironio, Phys. Rev. A **85**, 032119 (2012).
[16] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).
[17] T. Vértesi, Phys. Rev. A **78**, 032112 (2008).
[18] F. Hirsch, M. T. Quintino, T. Vértesi, M. Navascués, and N. Brunner, Quantum **1**, 3 (2017).
[19] T. Fritz, New J. Phys. **14**, 103001 (2012).
[20] C. Branciard, N. Brunner, H. Buhrman, R. Cleve, N. Gisin, S. Portmann, D. Rosset, and M. Szegedy, Phys. Rev. Lett. **109**, 100401 (2012).
[21] A. Tavakoli, P. Skrzypczyk, D. Cavalcanti, and A. Acin, Phys. Rev. A **90**, 062109 (2014).
[22] J. Henson, R. Lal, and M. F. Pusey, New J. Phys. **16**, 113043 (2014).
[23] K. Mukherjee, B. Paul, and D. Sarkar, Quantum Inf. Process. **14**, 2025 (2015).
[24] D. Rosset, C. Branciard, T. J. Barnea, G. Pütz, N. Brunner, and N. Gisin, Phys. Rev. Lett. **116**, 010403 (2016).
[25] R. Chaves, Phys. Rev. Lett. **116**, 010402 (2016).
[26] A. Tavakoli, Phys. Rev. A **93**, 030101 (2016).
[27] A. Tavakoli, J. Phys. A: Math. Theor. **49**, 145304 (2016).
[28] D. J. Saunders, A. J. Bennet, C. Branciard, and G. J. Pryde, Sci. Adv. **3**, 28 (2017).
[29] G. Carvacho, F. Andreoli, L. Santodonato, M. Bentivegna, R. Chaves, and F. Sciarrino, Nat. Commun. **8**, 14775 (2017).
[30] N. Gisin, Phys. Lett. A **154**, 201 (1991).

[31] R. Horodecki, P. Horodecki, and M. Horodecki, Phys. Lett. A **200**, 340 (1995).

[32] J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy*, revised edition 2004 (Cambridge University Press, Cambridge, 1987).

[33] A. Tavakoli, M. O. Renou, N. Gisin, and N. Brunner, New J. Phys. **19**, 073003 (2017).

[34] D. Cavalcanti, M. L. Almeida, V. Scarani, and A. Acin, Nat. Commun. **2**, 184 (2011).

[35] W. Klobus, W. Laskowski, M. Markiewicz, and A. Grudka, Phys. Rev. A **86**, 020302(R) (2012).

Correlations in star networks: from Bell inequalities to network inequalities

IOP Publishing New J. Phys. 19 (2017) 073003 https://doi.org/10.1088/1367-2630/aa7673

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

OPEN ACCESS

RECEIVED
15 February 2017

REVISED
9 May 2017

ACCEPTED FOR PUBLICATION
1 June 2017

PUBLISHED
7 July 2017

PAPER

Correlations in star networks: from Bell inequalities to network inequalities

Armin Tavakoli, Marc Olivier Renou, Nicolas Gisin and Nicolas Brunner

Groupe de Physique Appliquée, Université de Genève, CH-1211 Genève, Switzerland

E-mail: armin.tavakoli@unige.ch

Keywords: quantum correlations, quantum networks, Bell inequalities

Abstract

The problem of characterizing classical and quantum correlations in networks is considered. Contrary to the usual Bell scenario, where distant observers share a physical system emitted by one common source, a network features several independent sources, each distributing a physical system to a subset of observers. In the quantum setting, the observers can perform joint measurements on initially independent systems, which may lead to strong correlations across the whole network. In this work, we introduce a technique to systematically map a Bell inequality to a family of Bell-type inequalities bounding classical correlations on networks in a star-configuration. Also, we show that whenever a given Bell inequality can be violated by some entangled state ρ, then all the corresponding network inequalities can be violated by considering many copies of ρ distributed in the star network. The relevance of these ideas is illustrated by applying our method to a specific multi-setting Bell inequality. We derive the corresponding network inequalities, and study their quantum violations.

1. Introduction

Bell inequalities bound the strength of correlations between the outcomes of measurements performed by distant observers who share a physical system under the assumption of Bell-like locality. Famously, quantum theory predicts correlations, mediated by entangled states, that violate Bell inequalities [1]. Such nonlocal quantum correlations are central for many quantum information tasks as well as foundational challenges [2]. Classical and quantum correlations in the standard Bell scenario, i.e., where distant observers share a physical system produced by a single common source, have been intensively studied and are by now relatively well understood.

In comparison, only very little is known about classical and quantum correlations in networks. The latter are generalizations of the Bell scenario to more sophisticated configurations featuring several independent sources. Each source distributes a physical system to a subset of the distant observers. In the classical setting, each physical system is represented by a classical random variable. Importantly, random variables from different source are assumed to be independent. In the quantum setting, each source can produce an entangled quantum state. Moreover, each observer can perform joint (or entangled) measurements on systems coming from different sources—e.g., as in entanglement swapping [3]—thus potentially creating strong correlations across the entire network. Understanding the strength of quantum correlations in networks is a challenging problem, but of clear foundational interest. In addition, practical developments of quantum networks make these questions timely, see e.g. [4, 5].

One of the main hurdles for solving the above problem, is to first characterize classical correlations in networks. This turns out to be a challenging problem. Due to the assumption that the sources are independent, the set of classical correlations does not form a convex set anymore, as it is the case in the usual Bell scenario. Therefore, in order to efficiently characterize classical correlations, one should now derive nonlinear Bell-type inequalities. Only a handful of these inequalities have been derived so far. First works derived inequalities for the simplest network of entanglement swapping [6, 7], for which experimental violations were recently reported

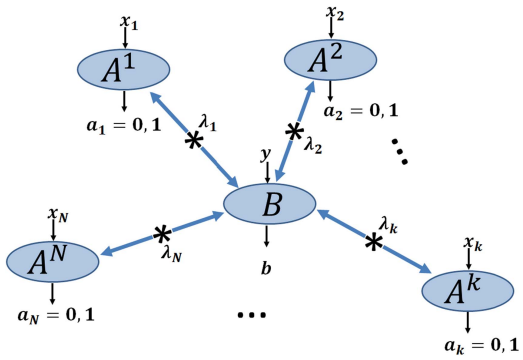

Figure 1. Star-network with a node observer B and N edge observers $A^1 \ldots A^N$ each independently sharing a bipartite physical system with the node observer.

[8, 9]. Then inequalities for networks in the star-configuration were presented [10]. There exists also methods for deriving nontrivial Bell-type inequalities for other classes of networks [11–13]. Entropic Bell inequalities has also been derived for several networks [14], but are usually not very efficient at capturing classical correlations. Furthermore, another approach to study correlations in networks is from the point of view of Bayesian inference [15–21].

In this work we aim to establish a direct link between the well-developed machinery of Bell inequalities, and the much less developed study of Bell-type inequalities for networks. Here, we focus on star-networks. We introduce a technique that allows one to map any full-correlation two-outcome bipartite Bell inequality into a family of Bell-type inequalities for star-networks (henceforth referred to as star inequalities). Specifically, starting from any such Bell inequality, we construct star inequalities for any possible star-network, which efficiently capture classical correlations. As a special case, this allows us to recover previously derived star inequalities [7, 10] by starting from the CHSH Bell inequality [22]. In general, our approach has three appealing features. First, the star inequalities we derive can have any number of settings for all observers. Second, our star inequalities efficiently capture the non-convex set of classical correlations, as there exist probability distributions violating our inequalities that would not violate any standard Bell inequality. Third, their quantum violations can be directly related to the quantum violation of the initial Bell inequality. More precisely, we show that whenever an entangled state ρ violates a Bell inequality, then all the corresponding star inequalities can be violated by placing many copies of ρ in the star-network. Conversely, we show that certain quantum correlations in star-networks can be used to infer bounds on independent Bell tests. Finally, we illustrate the relevance of this method by an explicit example in which we start from a Bell inequality with more than two settings and construct the mapping to a particular star inequality and study its violation in the simplest network of entanglement swapping.

2. Star networks and N-locality

Star-networks are a class of networks parametrized by the number of independent sources N. The network thus involves $N + 1$ observers: N so-called edge observers each of whom independently shares a state with one common central observer called the node observer. See figure 1 for an illustration.

The k'th edge observer performs a measurement labeled by x_k (chosen among a finite set) returning a binary outcome a_k. Depending on the context, to be consistent with previous work, we label it either $\{0, 1\}$ (sections 4, 6 and 7) or $\{-1, +1\}$ (as soon as correlators are involved). The node observer performs a measurement labeled by y returning an outcome b. The resulting statistics is given by a conditional probability distribution of the outcomes of all observers given their inputs. This probability distribution is called N-local if it admits the following form:

$$P(a_1 \ldots a_N b|x_1 \ldots x_N y) = \int \left(\prod_{k=1}^{N} \mathrm{d}\lambda_k q_k(\lambda_k) P(a_k|x_k \lambda_k) \right) P(b|y\vec{\lambda}), \tag{1}$$

where we have used the short-hand notation $\vec{\lambda} \equiv \lambda_1 \ldots \lambda_N$. In an N-local model (which is the analogue of a local model in the Bell scenario), each independent source emits a random variable λ_k which is shared between a

subset of the observers. In particular, for the star network, each edge observer shares a λ_k (possibly encoding an unlimited amount of shared randomness) with the node observer. Importantly, the sources are assumed to be independent from each other, and thus the variables λ_k are uncorrelated. Since the node observer holds $\vec{\lambda}$, he can create correlations among all observers. Notice that if $N = 1$ we recover the definition of classical correlations in the Bell scenario. If the probability distribution cannot be written on the above form, it is said to be non N-local. Inequalities bounding the strength of N-local correlations arising in a star network are called star inequalities.

3. Mapping Bell inequalities to star inequalities

Consider a bipartite Bell scenario, where two observers Alice and Bob each perform one of n_{A} respectively n_{B} measurements on a shared physical system. Each measurement returns a binary outcome, now denoted A_x, $B_y = \pm 1$ for convenience, where x and y indicate the choice of measurement of Alice and Bob respectively. Any full-correlation Bell inequality can be written

$$S_M^{\mathrm{bs}} \equiv \sum_{x=1}^{n_{\mathrm{A}}} \sum_{y=1}^{n_{\mathrm{B}}} M_{yx} \langle \mathrm{A}_x^{\mathrm{bs}} \mathrm{B}_y^{\mathrm{bs}} \rangle \leqslant C, \tag{2}$$

where M_{yx} are real numbers, and C is the local bound. Note that $\langle A_x^{\mathrm{bs}} B_y^{\mathrm{bs}} \rangle$ denotes the expectation value of the product of the outcomes of Alice and Bob. In equation (2) the superscript bs only serves as a label for the Bell scenario. Importantly, one can fully characterize the Bell inequality by specifying the matrix $M \in \mathbb{R}^{n_{\mathrm{B}} \times n_{\mathrm{A}}}$, from which the local bound C can be computed as follows. It is sufficient to consider deterministic strategies of Alice and Bob, due to the fact that the set of local correlations in the Bell scenario is a polytope [2]. Hence, we can write

$$S_M^{\mathrm{bs}} = \sum_{y=1}^{n_{\mathrm{B}}} \left(\sum_{x=1}^{n_{\mathrm{A}}} M_{yx} \mathrm{A}_x^{\mathrm{bs}} \right) \mathrm{B}_y^{\mathrm{bs}} = \sum_{y=1}^{n_{\mathrm{B}}} \hat{\mathrm{A}}_y^{\mathrm{bs}} \mathrm{B}_y^{\mathrm{bs}}, \tag{3}$$

where $\hat{\mathrm{A}}_y^{\mathrm{bs}} = \sum_{x=1}^{n_{\mathrm{A}}} M_{yx} \mathrm{A}_x^{\mathrm{bs}}$. From now on, we use this notation for a linear transformation M of Alice's set of correlators. To maximize $\mathrm{S}_M^{\mathrm{bs}}$, we choose $B_y^{\mathrm{bs}} = \mathrm{sign}(\hat{A}_y^{\mathrm{bs}})$, which allows us to write the classical bound as

$$C = \max_{\mathrm{A}_1 \ldots \mathrm{A}_{n_{\mathrm{A}}} \in \{\pm 1\}} \sum_{y=1}^{n_{\mathrm{B}}} \left| \sum_{x=1}^{n_{\mathrm{A}}} M_{yx} \mathrm{A}_x^{\mathrm{bs}} \right| \tag{4}$$

$$= \max_{\mathrm{A}_1 \ldots \mathrm{A}_{n_{\mathrm{A}}} \in \{\pm 1\}} \sum_{y=1}^{n_{\mathrm{B}}} |\hat{\mathrm{A}}_y^{\mathrm{bs}}|. \tag{5}$$

We now show how any Bell inequality of the form (2) can be mapped into a family of star inequalities for star-networks with N sources.

Theorem 3.1. *For any full-correlation Bell inequality represented by the matrix $M \in \mathbb{R}^{n_{\mathrm{B}} \times n_{\mathrm{A}}}$ with corresponding classical bound C, we can associate star inequalities as follows:*

$$S_{M,\{f_i\}}^{\mathrm{net}} \equiv \sum_{i=1}^{n_{\mathrm{B}}} |I_i|^{1/N} \leqslant C, \tag{6}$$

where

$$I_i = \sum_{x_1 \ldots x_N = 1}^{n_{\mathrm{A}}} M_{ix_1} \ldots M_{ix_N} \langle \mathrm{A}_{x_1}^1 \ldots \mathrm{A}_{x_N}^N B_i \rangle = \langle \hat{\mathrm{A}}_i^1 \ldots \hat{\mathrm{A}}_i^N B_i \rangle, \tag{7}$$

and

$$\langle A_{x_1}^1 \ldots A_{x_N}^N B_i \rangle = \sum_{a_1 \ldots a_N = 0,1} \sum_b (-1)^{a_1 + \ldots + a_N + f_i(b)} \times P(a_1 \ldots a_N b | x_1 \ldots x_N y), \tag{8}$$

for some boolean functions $\{f_i\}_{i=1}^{n_{\mathrm{B}}}$. Thus, specifying the real-valued matrix M and the functions $\{f_i\}_{i=1}^{n_{\mathrm{B}}}$ returns a specific star inequality for a star-network with N sources.

The proof is rather technical hence we defer it to appendix A, where we prove a generalized version of the above theorem in which the star inequality is obtained as a mapping of up to N different full-correlation Bell inequalities, each characterized by a real-valued matrix $M^{(k)}$ for $k = 1, \ldots, N$. The only restriction on the N Bell inequalities is that one observer (the one that by theorem 3.1 is mapped to the node observer) in each inequality chooses between the same number of measurements. For sake of simplicity, we have in the above taken all these

N Bell inequalities to be represented by the same matrix, namely M. Furthermore, we note that generalizations of our theorem to networks of the type studied in [23], in which each source emits a multiparty physical system, are possible[1].

An important feature of the star inequalities generated via the above construction is that they give a better characterization of the N-local set compared to standard Bell inequalities. That is, there exist certain probability distributions that are not N-local (as witnessed by the violation of some of our inequalities) that are nevertheless local in the usual Bell scenario (i.e. 1-local), and thus cannot violate any standard Bell inequality. This is made explicit in the next section.

4. Recovering the inequalities of [7, 10]

As an example of our technique, consider the CHSH inquality [22] which corresponds to the 2×2 matrix $M_{xy}^{\text{chsh}} = \frac{1}{2}(-1)^{xy}$ for $x, y = 0, 1$. The local bound (4) is straightforwardly evaluated to $C = 1$. Choosing a star-network with two sources ($N = 2$), and letting the node observer perform one complete two-qubit measurement with outcomes $b = b_1 b_2 \in \{0, 1\}^{\otimes 2}$, we can define $f_i(b_1 b_2) = b_i$ and immediately recover the inequality of [7]:

$$\sqrt{|I_1|} + \sqrt{|I_2|} \leqslant 1, \tag{9}$$

where I_1 and I_2 are defined via equation (7). Also, by letting the node observer have two measurement settings ($y \in \{0, 1\}$), one associated to I_1 and one associated to I_2, returning an output bit $b \in \{0, 1\}$, we recover the other inequality of [7] with $f_i(b) = b$ (this example will be studied in more detail in appendix B3). Similar mappings of the CHSH inequality also return the star inequalities of [10] valid for an arbitrary number of sources:

$$|I_1|^{1/N} + |I_2|^{1/N} \leqslant 1. \tag{10}$$

In this scenario, all observers have two settings and two outcomes, and $f_i(b_1 b_2) = b_i$. For $N = 1$ this reduces to the CHSH inequality.

Importantly, the above star inequalities are nonlinear, and thus give a better characterization of the N-local set compared to standard (linear) Bell inequalities. In particular, there exist quantum correlations admitting a local description (hence not violating any standard Bell inequality) that nevertheless violate these star inequalities [7, 10].

5. Optimal classical strategies and tightness

We now demonstrate a property of optimal N-local strategies regarding our star inequalities. We show that for any star inequality obtained from theorem 3.1, any N-local strategy achieving $S^{\text{net}} = C$ with given values $\{I_i\}$ can be replaced with another N-local strategy achieving the same $\{I_i\}$ in which the node observer B acts trivially i.e. gives a deterministic output depending on the input. Moreover, this is achieved with the same strategy for each edge observer A^k. More precisely, we have the following:

Proposition 5.1. *For any N-local strategy $\mathcal{S}: A_{x_k}^k(\lambda_k),\ B_y(\vec{\lambda})$ reaching the the N-local bound in equation (6) with $0 \leqslant I_i$, there is a reduced strategy $\mathcal{S}': A_{x_k}^{\prime k}(\lambda_k),\ B_y'(\vec{\lambda})$ such as:*

1. *The node observer B has a deterministic output: $B_i'(\vec{\lambda})$ is independent from $\vec{\lambda}$ and only depends on i. We note it b_i: this is the deterministic output of B for input $y = i$. Thus each source of randomness λ_k can be considered as local and held by the edge observer A_k.*
2. *Each edge observer A^k chooses her output according to the same strategy: the functions $(\lambda_k, x_k) \mapsto A_{x_k}^{\prime k}(\lambda_k)$ are independent from k (then we write $A_{x_k}'(\lambda_k)$).*
3. *The quantities I_i remain unchanged:*
   $\langle \hat{A}_i^{\ 1} \ldots \hat{A}_i^{\ N} B_i \rangle = \langle \hat{A'}_i^{\ 1} \ldots \hat{A'}_i^{\ N} B_i' \rangle = b_i \langle \hat{A'}_i \rangle^N$.

This proposition is proven in appendix B, in which we also illustrate it by applying it to a particular example.

Another question is whether any set $\{I_i\}$ saturating the inequality (6) can be obtained by an N-local strategy. We see in appendix C that this is not the case and give a way to find and enumerate all the sets $\{I_i\}$ satisfying this property.

[1] Also, one may consider variations of theorem 3.1 in which one constructs more than n_B quantities $\{I_i\}_i$.

So far, we have shown how the limitations of classical correlations in the Bell scenario can be mapped to analog limitations in networks. Next, we explore if an analogous statement can be made for quantum correlations.

6. Quantum violations

We shall relate the quantum violation of the initial Bell inequality to the quantum violation of the corresponding star inequalities. Specifically, we will see that for any state ρ violating the initial Bell inequality, taking a sufficient number of copies of ρ distributed in the network will lead to violation of the corresponding star inequality. Also, the robustness to white noise of every quantum state violating a Bell inequality (2), is the same as that of N copies of the same state violating a star inequality.

Consider a Bell scenario where Alice and Bob share an entangled state ρ and perform n_A and respectively n_B binary local measurements represented by observables $\mathcal{A}_x^{\text{bs}}$ and $\mathcal{B}_y^{\text{bs}}$. This leads to violation of some full-correlation Bell inequality, i.e. achieving $S^{\text{bs}} > C$. Then we obtain a quantum strategy for violating the corresponding star inequalities as follows.

Let the node observer in the star-network perform n_{B} different measurements. Each one is represented by an observable which is simply the N-fold tensor product of the measurements performed by Bob in the Bell scenario: $\forall\, y : \mathcal{B}_y \equiv \mathcal{B}_y^{\text{bs}} \otimes \ldots \otimes \mathcal{B}_y^{\text{bs}}$, and let all the edge observers perform the same n_{A} measurements as Alice in the Bell scenario: $\forall\, x$: $\mathcal{A}_x^1 = \ldots = \mathcal{A}_x^N \equiv \mathcal{A}_x^{\text{bs}}$. Finally, let all N sources emit the same bipartite state ρ as in the Bell scenario. This causes the factorization $\langle \mathcal{A}_{x_1}^1 \ldots \mathcal{A}_{x_N}^N \mathcal{B}_y \rangle_{\rho^{\otimes N}} = \langle \mathcal{A}_{x_1}^{\text{bs}} \mathcal{B}_y^{\text{bs}} \rangle_\rho \ldots \langle \mathcal{A}_{x_N}^{\text{bs}} \mathcal{B}_y^{\text{bs}} \rangle_\rho$ which implies

$$I_i = \left(\sum_{x=1}^{n_{\text{B}}} M_{ix} \langle \mathcal{A}_x^{\text{bs}} \mathcal{B}_i^{\text{bs}} \rangle_\rho \right)^N = (\langle \hat{\mathcal{A}}_i^{\text{bs}} \mathcal{B}_i^{\text{bs}} \rangle_\rho)^N. \tag{11}$$

Inserting this into equation (6) we recover $S^{\text{net}} = S^{\text{bs}} > C$. We conclude that

$$\rho \text{ violates Bell inequality } \Rightarrow \rho^{\otimes N} \text{ violates star inequality.} \tag{12}$$

Note the generality of the above statement, which holds true for any full-correlation Bell inequality and all its corresponding star inequalities (in particular for all possible choices of functions $f_i(b)$). Moreover, the statement holds for an entangled state ρ of arbitrary Hilbert space dimension.

The case of CHSH Bell inequality deserves to be discussed. The above statement implies that any entangled state violating CHSH will violate all its corresponding star inequalities when enough copies are distributed in the network. In particular, this is case for any pure entangled bipartite state [24], and more generally for any two-qubit state detected by the Horodecki criterion [25] (necessary for CHSH violation). Note that the latter statement was recently derived in [26] for the case $N = 2$, however, with the important difference that there the node observer performed a Bell state measurement whereas in our case we consider product measurements.

Conversely, if the node observer performs some product measurement, i.e., a measurement of the form $\forall\, y : \mathcal{B}_y = \mathcal{B}_y^1 \otimes \ldots \otimes \mathcal{B}_y^N$, with otherwise arbitrary choices of measurements for all edge observers and N arbitrary states distributed in the network, then the achieved value of S^{net} is upper bounded by the geometric average of S^{bs} as obtained in N independent Bell tests. Due to the separability of $\mathcal{B}_y$, we have $I_i = \prod_{k=1}^N \langle \hat{\mathcal{A}}_i^{\text{bs},k} \mathcal{B}_i^{\text{bs}} \rangle_{\rho_k}$. Inserting this into equation (6) we find

$$S^{\text{net}} = \sum_{i=1}^{n_{\text{B}}} \prod_{k=1}^{N} |\langle \hat{\mathcal{A}}_i^{\text{bs},k} \mathcal{B}_i^{\text{bs}} \rangle_{\rho_k}|^{1/N} \leqslant \prod_{k=1}^{N} \left[\sum_{i=1}^{n_{\text{B}}} |\langle \hat{\mathcal{A}}_i^{\text{bs},k} \mathcal{B}_i^{\text{bs}} \rangle_{\rho_k}| \right]^{1/N}. \tag{13}$$

To obtain the upper bound, we have used lemma A.1 stated in appendix A, which may be regarded as a generalization of the Cauchy–Schwarz inequality. The expression on the right-hand-side of equation (13) is the geometric average of $\{S^{\text{bs}}(i)\}_{i=1}^N$ as obtained in N independent Bell tests M, each performed on the state ρ_k with settings of Alice and Bob determined by the settings used to achieve S^{net} in the star-network. This upper bound coincides with S^{net} only when all N Bell tests yield the same value $S^{\text{net}} = S^{\text{bs}}(i) \forall\, i$.

So far, we have only considered product measurements of the node observer, which were sufficient to map quantum strategies in Bell inequalities to analog strategies in networks. Next, we consider an explicit example of a multisetting Bell inequality from which we construct a star inequality for $N = 2$ and study the quantum violations using product and joint measurements.

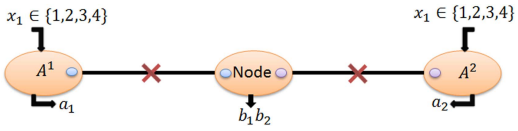

Figure 2. The node observer performs a single measurement with four possible outcomes and the two edge observers each perform one of four measurements with binary outcomes.

7. Example: quantum correlations from entanglement swapping

We consider the full-correlation Bell inequality presented in [27] [2]. It is represented by the following matrix;

$$M^{3\times 4} = \begin{pmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1. \end{pmatrix}. \tag{14}$$

We can calculate the classical bound using equation (4), and write the associated Bell inequality, in which Alice has four settings and Bob has three settings, as follows:

$$\sum_{x=1}^{4}\sum_{y=1}^{3} M_{yx}\langle A_x B_y\rangle \leqslant 6. \tag{15}$$

The maximal quantum violation of this inequality is given by $4\sqrt{3} > 6$, obtained with a maximally entangled two-qubit state $|\psi_{00}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. Alice's measurements are characterized by Bloch vectors forming the vertices of a tetrahedron on the Bloch sphere:

$$\vec{m}_1 = \frac{1}{\sqrt{3}}(1,\,1,\,1)\ \vec{m}_2 = \frac{1}{\sqrt{3}}(1,\,-1,\,-1) \qquad \vec{m}_3 = \frac{1}{\sqrt{3}}(-1,\,1,\,-1)\ \vec{m}_4 = \frac{1}{\sqrt{3}}(-1,\,-1,\,1). \tag{16}$$

Bob's measurements are simply given by the three Pauli matrices σ_1, σ_2, and σ_3. If we consider the mixture of $|\psi_{00}$ with white noise, i.e. a Werner state of the form

$$\rho_v = v|\psi_{00}\rangle\langle\psi_{00}| + \frac{1-v}{4}\mathbf{1}, \tag{17}$$

the inequality can be violated whenever $v > \sqrt{3}\,2^{-1/2}$. Note that a sufficiently high violation of this inequality implies that the measurements settings do not lie in a plane of the Bloch sphere, i.e. they feature complex phases [28].

Next, we obtain a particular star inequality for $N = 2$ in which we let each edge observer perform one of four measurements $x_1, x_2 \in \{1,\,2,\,3,\,4\}$, whereas the node observer performs a single measurement (i.e. no input y) with four possible outcomes $b = b_1 b_2 \in \{0,\,1\}^{\otimes 2}$. This is illustrated in figure 2. To this end, we apply our theorem 1. We define three quantities

$$I_i = \sum_{x_1,x_2=1}^{4} M_{ix_1} M_{ix_2} \langle A^1_{x_1} A^2_{x_2} B_i\rangle, \tag{18}$$

where

$$\langle A^1_{x_1} A^2_{x_2} B_i\rangle = \sum_{a_1,a_2,b} (-1)^{a_1+a_2+f_i(b)} P(a_1 a_2 b|x_1 x_2). \tag{19}$$

We choose the functions $f_i(b)$ for $i = 1,\,2,\,3$ as: $(f_1, f_2, f_3) = (b_1,\, b_1 + b_2 + 1,\, b_2)$. Hence, our star inequality reads

$$S^{\text{net}} \equiv \sqrt{|I_1|} + \sqrt{|I_2|} + \sqrt{|I_3|} \leqslant 6. \tag{20}$$

Next we discuss quantum violations. Both sources in the network emit the Bell state $|\psi_{00}\rangle$. The two edge observers perform the four tetrahedron measurements given in equation (16). The node observer performs the Bell state measurement projecting her two systems onto the basis of maximally entangled two-qubit states: $|\psi_{b_1 b_2}\rangle = \sigma_3^{b_1} \otimes \sigma_1^{b_2}|\psi_{00}\rangle$. Such a Bell state measurement typically causes the joint state of the subsystems of the two edge observers to become entangled, with its exact form depending on the outcome of the node observer. The resulting expectation values are

[2] This inequality is referred to in [27] as the 'elegant Bell inequality' due to the high symmetry of the observables leading to its maximal quantum violation.

$$\langle A^1_{x_1} A^2_{x_2} B_i\rangle = \mathrm{tr}[(\rho_{00}\otimes\rho_{00})\,\vec{m}_{x_1}\cdot\vec{\sigma}\otimes(\sigma_i\otimes\sigma_i)\otimes\vec{m}_{x_2}\cdot\vec{\sigma}] = \frac{M_{ix_1}M_{ix_2}}{3}\,, \tag{21}$$

where $\rho_{00} = |\psi_{00}\rangle\langle\psi_{00}|$. This leads to $I_1 = I_2 = I_3 = 16/3$ which inserted into equation (20) returns $S^{\mathrm{net}} = 4\sqrt{3} > 6$. Hence, quantum correlations generated in an entanglement swapping scenario violate the considered star inequality. If both sources are noisy and each emits a Werner state ρ_v, then one can violate the inequality (20) whenever $v > \sqrt{3}/2$. This coincides with the critical noise level of the Bell inequality in equation (14).

Furthermore, note that we can with minor modification re-cast our inequality (20) so that the node observer performs three different measurements, each with a binary outcome b. In this scenario, one can again obtain the quantum violation $S^{\mathrm{net}} = 4\sqrt{3}$. Note in this case the node observer uses three product measurements of the form $\sigma_i\otimes\sigma_i$, i.e. a product of Pauli matrices. It turns out that these three measurements are compatible (they commute). They can thus be measured jointly, which is done via the Bell state measurement.

Finally, we point out that we can swap the roles of Alice and Bob in the Bell inequality equation (15) so that when mapped to the star inequality, the node observer has four settings and the edge observers each have three settings. That inequality reads

$$\sqrt{|I'_1|} + \sqrt{|I'_2|} + \sqrt{|I'_3|} + \sqrt{|I'_4|} \leqslant 6, \tag{22}$$

where $I'_y = \sum_{x_1,x_2=1}^{3} M^T_{yx_1} M^T_{yx_2} \langle A^1_{x_1} A^2_{x_2} B_y\rangle$, where $\langle A^1_{x_1} A^2_{x_2} B_i\rangle = \sum_{a_i,a_2,b}(-1)^{a_1+a_2+b} P(a_1 a_2 b|x_1 x_2 y)$. By letting the node observer perform products of the measurements in equation (16) and the edge observers perform the Pauli measurements σ_i for $i = 1, 2, 3$, we again find a violation $S^{net} = 4\sqrt{3}$, for which the critical visibility again is $v = \sqrt{3}/2$.

8. Conclusions

Our main result is a method for systematically mapping any multi-setting full-correlation Bell inequality into a family of inequalities bounding the strength of classical correlations in star networks. This construction also allows us to show that quantum strategies for Bell inequalities can be mapped into analogous quantum strategies on star-networks. Specifically, for any entangled state ρ violating the initial Bell inequality, it follows that by taking enough copies of ρ in the star network one obtains a quantum violation of the corresponding star inequalities. Finally, we considered an explicit scenario involving more than two settings and show that quantum correlations in an entanglement swapping experiment can violate our inequalities.

To conclude, we mention some open problems: (1) Can our technique be extended to also include mappings of Bell inequalities with marginals, i.e. not only full-correlation terms as in equation (2). Whether the technique can be adapted to full-correlation Bell inequalities with more than two outputs (see e.g. [29]) is also relevant. (2) In particular, our technique allows us to explore quantum correlations in entanglement swapping experiments with many settings. Exploring the ability of these correlations to violate the inequalities would be of interest. (3) How can one extend our technique to involve networks that are not of the star configuration? (4) Can one construct star inequalities analogous to the one in equation (22) in which the node observer performs a single joint measurement with four outcomes? To what extent can quantum theory violate these inequalities? (5) It appears, after considering several particular examples, that all star inequalities derived by the presented technique cannot outperform the original Bell inequality in terms of noise tolerance when mixed with the maximally mixed state. Is this the case for any joint measurement? Or on the contrary, can one find an example where the use of an adequate joint measurements allows for activation of nonlocality. That is, while the entangled state ρ would not violate the initial Bell inequality, many copies of ρ distributed in the network would lead to violation of the star inequality. While such activation phenomena are proven to exist even when considering the standard definition of Bell locality [30, 31], we expect that the effect of activation should become much stronger when considering N-locality.

Acknowledgments

This work was supported by the Swiss national science foundation (SNSF 200021-149109 and Starting Grant DIAQ and QSIT), and the European Research Council (ERC-AG MEC).

Appendix A. Proof of main theorem

In this appendix we prove a generalized version of theorem 3.1, in which the star inequality is obtained as mapping of up to N different full-correlation Bell inequalities in all of which at least one observer has the same number of settings. However, we first state a useful lemma that was presented and proven in [10]:

Lemma A.1. *Let x_i^k be non-negative real numbers and let n_{B}, $N \geqslant 1$ be integers. Then, the following relation holds:*

$$\sum_{k=1}^{n_B}\left(\prod_{i=1}^{N} x_i^k\right)^{1/N} \leqslant \prod_{i=1}^{N}\left(\sum_{k=1}^{n_B} x_i^k\right)^{1/N}, \tag{A1}$$

with equality if and only if $\forall\, k$: $x_1^k = \ldots = x_N^k$.

Equipped with this lemma, we state and prove our main theorem.

Theorem A.2. *Consider any set of N full-correlation Bell inequalities such that in every Bell scenario Bob has n_{B} measurement settings, whereas in the* k*'th Bell scenario Alice has $n_A^{(k)}$ measurement settings. The* k*'th Bell inequality is represent by the matrix $M^{(k)} \in \mathbb{R}^{n_B} \times \mathbb{R}^{n_A^{(k)}}$ with associated classical bound C_k. To every set of such matrices, $\{M^{(k)}\}_{k=1}^{N}$, we can associate a family of star inequalities as follows:*

$$S^{\mathrm{net}}_{\{M\},\{f_i\}} \equiv \sum_{i=1}^{n_{\mathrm{B}}} |I_i|^{1/N} \leqslant (C_1 \ldots C_N)^{1/N}, \tag{A2}$$

where

$$I_i = \sum_{x_1=1}^{n_A^{(1)}} \ldots \sum_{x_N=1}^{n_A^{(N)}} M^1_{ix_1} \ldots M^N_{ix_N} \langle \mathrm{A}^1_{x_1} \ldots \mathrm{A}^N_{x_N} B_i \rangle \tag{A3}$$

$$= \langle \hat{A}_i^{\,1} \ldots \hat{A}_i^{\,N} B_i \rangle, \tag{A4}$$

and

$$\langle A^1_{x_1} \ldots A^N_{x_N} B_i \rangle = \sum_{a_1 \ldots a_N = 0,1} \sum_{b} (-1)^{a_1 + \ldots + a_N + f_i(b)} \times P(a_1 \ldots a_N b | x_1 \ldots x_N y), \tag{A5}$$

for some boolean functions $\{f_i\}_{i=1}^{n_{\mathrm{B}}}$. Thus, specifying the real-valued matrices $\{M^{(k)}\}_k$ and the functions $\{f_i\}_{i=1}^{n_{\mathrm{B}}}$ returns specific star inequality for the star-network with N sources.

Proof. Impose a classical model (1) on the probabilities in the quantities I_i defined in equation (7). This gives

$$I_i = \int \left[\prod_{k=1}^{N} \mathrm{d}\lambda_k q_k(\lambda_k) \hat{A}_i^{\,k}(\lambda_k) B_i(\vec{\lambda}) \right]. \tag{A6}$$

Applying an absolute value to both sides allows for the following upper bound;

$$|I_i| \leqslant \prod_{k=1}^{N} \int \mathrm{d}\lambda_k q_k(\lambda_k) |\hat{A}_i^{\,k}(\lambda_k)|. \tag{A7}$$

Each integral in the product series is a non-negative number. Hence, the quantity $|I_i|^{1/N}$ can be upper bounded by a geometric average of such integrals. Applying the lemma A.1 to put an upper bound S^{net}, which is a sum of such quantities, we obtain the following:

$$S^{\mathrm{net}}_{\{M\},\{f_i\}} = \sum_{i=1}^{n_{\mathrm{B}}} |I_i|^{1/N} \leqslant \left[\prod_{k=1}^{N} \int \mathrm{d}\lambda_k q_k(\lambda_k) \sum_{i=1}^{n_{\mathrm{B}}} |\hat{A}_i^{\,k}(\lambda_k)| \right]^{1/N}. \tag{A8}$$

Remember that each correlator of Alice obeys $-1 \leqslant A^k_{x_k}(\lambda_k) \leqslant 1$ and hence, using the classical bound (4) of the Bell inequality associated to $M^{(k)}$ to substitute in the integrand, we find

$$S^{\mathrm{net}}_{\{M\},\{f_i\}} \leqslant \left[\prod_{k=1}^{N} \int \mathrm{d}\lambda_k q_k(\lambda_k) C_k \right]^{1/N}. \tag{A9}$$

Using that $\forall\, k : \int \mathrm{d}\lambda_k q_k(\lambda_k) = 1$, we obtain the final result

$$S^{\mathrm{net}}_{\{M\},\{f_i\}} \leqslant (C_1 \ldots C_N)^{1/N}. \tag{A10}$$

■

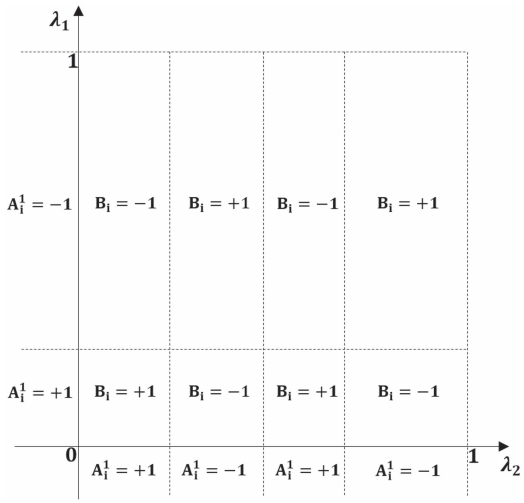

Figure 3. Illustration of the argument for two edge observers. As $A_i^1(\lambda_1)A_i^2(\lambda_2)B_i(\vec{\lambda})$ is of constant sign (here positive) and $B_i(\vec{\lambda}) = \pm 1$, the sign of $A_i^1(\lambda_1)$ and $A_i^2(\lambda_2)$ totally determine $B_i(\vec{\lambda})$, which is of the form given by (B3).

Remark. By choosing all the N Bell inequalities to be the same, i.e. setting $M \equiv M^{(1)} = \ldots = M^{(N)}$, we obtain the special case of this theorem considered in the main text.

Appendix B. Redundancy of node observer in classical strategies

Here we prove proposition 5.1 and illustrate it on a simple example.

Proof. Let us consider that we already have a strategy $\mathcal{S}$ reaching the bound in equation (6) with given I_i. $\mathcal{S}$ is defined by the correlators of each edge observer A^k (resp. node observer B) given (λ_k, x_k) (resp. $(\vec{\lambda}, y)$) i.e. $A^k_{x_k}(\lambda_k)$ (resp. $B^k_y(\vec{\lambda})$). These correlators are such as:

$$I_i = \langle \hat{A}_i^{\,1} \ldots \hat{A}_i^{\,N} B_i \rangle \tag{B1}$$

$$= \int \left[\prod_{k=1}^{N} \mathrm{d}\lambda_k q_k(\lambda_k) \hat{A}_i^{\,k}(\lambda_k) B_i(\vec{\lambda}) \right]. \tag{B2}$$

As we have equality in equation (6), going back in the proof of theorem 3.1, $\mathcal{S}$ must be such as equation (A7) and equation (A8) are equalities. From the equality condition of equation (A7), we will deduce a $\mathcal{S}'$ satisfying condition 1. and 2. of the proposition. We then improve it in a strategy $\mathcal{S}''$ satisfying *3.*, using the equality condition of equation (A8).

Equation (A7) is the continuous triangle inequality. As we have equality, for any i, the integrand $\vec{\lambda} \mapsto \prod_{k=1}^{N} \hat{A}_i^{\,k}(\lambda_k) B_i(\vec{\lambda})$ must be of constant sign (the weights q_k are positive). Then, any change in the sign of some $\hat{A}_i^{\,k}(\lambda_k)$ at a specific λ_j^0 must be compensated by a change of the sign of $B_i(\vec{\lambda})$ at the same λ_j^0, whatever are the other λ_j's (see figure 3). As $B_i(\vec{\lambda}) = \pm 1$, we have that:

$$B_i(\vec{\lambda}) = \prod_k B_i^k(\lambda_k), \tag{B3}$$

where $B_i^k(\lambda_k)$ depends on the sign of $\hat{A}_i^{\,k}(\lambda_k)$.

We now can define the new strategy $\mathcal{S}'$:

- $A'^k_{x_k}(\lambda_k) \equiv A^k_{x_k}(\lambda_k) B_k(\lambda_k)$
- $B'_i(\vec{\lambda}) \equiv 1$.

Through the transformation induced by M, this corresponds to corresponding $\hat{A}_i^{\,\prime k}(\lambda_k) = \hat{A}_i^{\,k}(\lambda_k) B_k(\lambda_k)$.

As $\prod_{k=1}^{N} \hat{A}_i^{\,k}(\lambda_k) B_i(\vec{\lambda}) = \prod_{k=1}^{N} \hat{A'}_i^{\,k}(\lambda_k) B'_i(\vec{\lambda})$, the new I'_i corresponding to strategy $\mathcal{S}'$ are equal to the I_i corresponding to strategy $\mathcal{S}$. Then *1.* and *2.* of 5.1 are satisfied by $\mathcal{S}'$. Moreover, we have:

$$I_i = \langle \hat{A'}_i^{\,1} \rangle \ldots \langle \hat{A'}_i^{\,N} \rangle. \tag{B4}$$

We now use the equality condition of lemma A.1 and equation (A8) to prove *3.* It is a convexity inequality which now reads:

$$\sum_{i=1}^{n_B} |I_i|^{1/N} = \sum_{i=1}^{n_B} \prod_{k=1}^{N} |\langle \hat{A'}_i^{\,k} \rangle|^{1/N} \leqslant \prod_{k=1}^{N} \sum_{i=1}^{n_B} |\langle \hat{A'}_i^{\,k} \rangle|^{1/N}, \tag{B5}$$

were here the inequality is an equality. The condition for the convexity inequality in lemma A.1 to be an equality is that $|\langle \hat{A'}_i^{\,k} \rangle|$ is independent of k: we have here that for each k, $|\langle \hat{A'}_k^{\,1} \rangle| = |\langle \hat{A'}_i^{\,1} \rangle|$. Replacing each the strategy and local random source of each edge observer A^k by a copy of the first edge observer A^1 strategy and random source in $\mathcal{S}'$ (we then leave the exponent k), we may only change the sign of I_i. This can be compensate by an appropriate choice of $B''_i(\vec{\lambda}) = b_i = \pm 1$. Then, we do not change *1.* and *2.* and obtain *3.*, with:

$$I_i = b_i \langle \hat{A''}_i \rangle^N. \tag{B6}$$

■

To illustrate the proposition, let us recall the proof of the tightness of an inequality (already introduced in section 4) presented in [7], for a star network with N edges, two inputs and two outputs for each observer. As illustrated in the section *Recovering the inequalities of* ([7, 10]), the inequality can be seen as a direct application of theorem 3.1, taking a matrix M corresponding to a renormalized CHSH inequalitie, $M_{xy} = \frac{1}{2}(-1)^{xy}$. Then $\hat{A}_1^{\,k} = \frac{1}{2}(A_1^{\,k} + A_2^{\,k})$ and $\hat{A}_2^{\,k} = \frac{1}{2}(A_1^{\,k} - A_2^{\,k})$ and (I_1, I_2) in equation (B2) is (I, J) in [7], with an inequality which writes:

$$|I_1|^{1/N} + |I_2|^{1/N} \leqslant 1. \tag{B7}$$

The authors obtained the classical bound (restricting to $0 \leqslant I_1, I_2$) for each possible $I_1 = r^N$, $I_2 = (1 - r)^N$ with the following strategy:

$$A_{x_k}^k(x_k, \lambda_k) = (-1)^{\lambda_k}(-1)^{\mu_k x_k} \tag{B8}$$

$$B_y(\vec{\lambda}) = \prod_k (-1)^{\lambda_k}, \tag{B9}$$

where the $\lambda_k \in \{0, 1\}$ are uniform shared variables between each of the edge observer and the node observer and the $\mu_k \in \{0, 1\}$ ($\mu_k = 0$ with probability r) are sources of local randomness for each edge observer. We see here, as shown by the proposition, that all edge observer have the same strategy and that the node observer's strategy factorizes in $B_y(\vec{\lambda}) = \prod_k B_y^k(\lambda)$ with $B_y^k(\vec{\lambda}) = \prod_k (-1)^{\lambda_k}$. Then, as suggested by the proposition, defining:

$$A'^k_{x_k}(x_k, \lambda_k) = A_{x_k}^k(x_k, \lambda_k) B_y^k(\lambda_k) = (-1)^{\mu_k x_k} \tag{B10}$$

$$B'_y(\vec{\lambda}) = 1, \tag{B11}$$

we see that the I_i are unchanged by the transformation, and obtain a reduced strategy in which all the conditions of the proposition are satisfied. The proposition states that such a transformation is always possible.

Appendix C. Partial tightness of star inequalities

We now study the tightness of the bound in equation (4):

$$\sum_{i=1}^{n_B} |I_i|^{1/N} \leqslant C. \tag{C1}$$

In the following, using proposition 5.1, we find all the sets of $\{I_i\}$ which are reachable by N-local strategies and saturate (C1).

We start by enumerating all the possible deterministic strategies for each edge observer: ${}^rX = ({}^r\mathrm{A}_1 \ldots {}^r\mathrm{A}_{n_A}) \in \{\pm 1\}^{n_A}$ for $r = 1 \ldots 2^{n_A}$. For each one, we note ${}^rY = ({}^r\hat{A}_1 \ldots {}^r\hat{A}_{n_B})$ the vector obtained after transformation of X by M:

$${}^rY = M\,{}^rX. \tag{C2}$$

Suppose that a given set $\{I_i\}$ satisfying condition (C1) can be obtained with an N-local strategy. Then, by proposition 5.1, we can suppose that it is obtained with a strategy in which the node observer B has deterministic strategy $B_i(\vec{\lambda}) = b_i = \pm 1$ and all edge observers A^k play the same strategy (we then leave the exponent k) based

on a shared random variable. Hence,

$$I_i = \langle \hat{A}_i^{\,1} \ldots \hat{A}_i^{\,N} B_i \rangle = b_i \langle \hat{A}_i \rangle^N. \tag{C3}$$

As A has only 2^{n_A} possible deterministic strategies, there exists probabilities $p_1 + \ldots + p_{n_A} = 1$ such that the strategy of each A is 'play deterministic strategy X_r with probability p_r'. Then:

$$\langle \hat{A}_i \rangle = \sum_{r=1}^{2^{n_A}} p_r \; {}^r\hat{A}_i. \tag{C4}$$

We then have:

$$\sum_i |I_i|^{1/N} = \sum_i \left| \sum_{r=1}^{2^{n_A}} p_r \; {}^r\hat{A}_i \right| \leqslant \sum_i \sum_{r=1}^{2^{n_A}} p_r |{}^r\hat{A}_i| \leqslant \max_s \sum_i |{}^s\hat{A}_i| = C, \tag{C5}$$

where the inequalities are equalities, which implies:

- $|\sum_{r=1}^{2^{n_A}} p_r \; {}^r\hat{A}_i| = \sum_{r=1}^{2^{n_A}} p_r |{}^r\hat{A}_i|$ i.e. for any i, the sign of all ${}^r\hat{A}_i$ such as $p_r \neq 0$ is the same (but may differ from one i to the other).
- $\forall\, r$ such as $p_r \neq 0$, $\sum_i |{}^r\hat{A}_i| = C$.

Then, this proves that any distribution of $\{I_i\}$ such as (C1) can be generated from the following method:

1. Enumerate all the possible $({}^rX, {}^rY)$
2. Keep the one such as $\sum_i |{}^r\hat{A}_i| = C$.
3. Sort them in different sets S_ν of size s_ν, each S_ν containing $({}^rX, {}^rY)$ where $sign({}^rY_i)$ is constant over r (but may differ depending on i).
4. The set of all $\{I_i\}$ such as the condition (C1) is fulfilled is :

$$\bigcup_{b_i=\pm 1} \bigcup_{\nu} \bigcup_{\substack{p_1+\cdots \\ +p_{s_\nu}=1}} \{I_i^{p_1,\ldots,p_{s_\nu},b_i}\}, \tag{C6}$$

where $\{I_i^{p_1,\ldots,p_{s_\nu},b_i}\}$ are obtained when each A 'play deterministic strategy $X_r \in S_\nu$ with probability p_r' and B deterministically answer b_i : $I_i^{p_1,\ldots,p_{s_\nu},b_i} = b_i(\sum_r p_r \; {}^r\hat{A}_i)^N$. Conversely, this gives a strategy proving that any distribution of $\{I_i\}$ given by (C6) can be obtain by an N-local strategy.

References

[1] Bell J S 1964 *Physics* **1** 195–200
[2] Brunner N, Cavalcanti D, Pironio S, Scarani V and Wehner S 2014 *Rev. Mod. Phys.* **86** 419
[3] Żukowski M, Zeilinger A, Horne M A and Ekert A K 1993 *Phys. Rev. Lett.* **71** 4287
[4] Kimble H J 2008 *Nature* **453** 1023
[5] Sangouard N, Simon C, de Riedmatten H and Gisin N 2011 *Rev. Mod. Phys.* **83** 33
[6] Branciard C, Gisin N and Pironio S 2010 *Phys. Rev. Lett.* **104** 170401
[7] Branciard C, Rosset D, Gisin N and Pironio S 2012 *Phys. Rev.* A **85** 032119
[8] Saunders D J, Bennet A J, Branciard C and Pryde G J 2017 *Sci. Adv.* **3** 4
[9] Carvacho G, Andreoli F, Santodonato L, Bentivegna M, Chaves R and Sciarrino F 2017 *Nat. Commun.* **8** 14775
[10] Tavakoli A, Skrzypczyk P, Cavalcanti D and Acín A 2014 *Phys. Rev.* A **90** 062109
[11] Chaves R 2016 *Phys. Rev. Lett.* **116** 010402
[12] Rosset D, Branciard C, Barnea T J, Pütz G, Brunner N and Gisin N 2016 *Phys. Rev. Lett.* **116** 010403
[13] Tavakoli A 2016 *Phys. Rev.* A **93** 030101
[14] Chaves R and Fritz T 2012 *Phys. Rev.* A **85** 032113
[15] Fritz T 2012 *New J. Phys.* **14** 103001
[16] Henson J, Lal R and Pusey M F 2014 *New J. Phys.* **16** 113043
[17] Chaves R, Majenz C and Gross D 2015 *Nat. Commun.* **6** 5766
[18] Wood C J and Spekkens R W 2015 *New J. Phys.* **17** 033002
[19] Chaves R, Kueng R, Brask J B and Gross D 2015 *Phys. Rev. Lett.* **114** 140403
[20] Wolfe E, Spekkens R W and Fritz T 2016 arXiv:1609.00672
[21] Kela A, von Prillwitz K, Aberg J, Chaves R and Gross D 2017 arXiv:1701.00652
[22] Clauser J F, Horne M A, Shimony A and Holt R A 1969 *Phys. Rev. Lett.* **23** 880
[23] Tavakoli A 2016 *J. Phys. A: Math. Theor.* **49** 145304
[24] Gisin N 1991 *Phys. Lett.* A **154** 201
[25] Horodecki R, Horodecki P and Horodecki M 1995 *Phys. Lett.* A **200** 340
[26] Gisin N, Mei Q, Tavakoli A, Renou M O and Brunner N 2017 arXiv:1702.00333

[27] Gisin N 2009 Bell inequalities: many questions, a few answers, quant-ph/0702021 *Quantum Reality, Relativistic Causality, and Closing the Epistemic Circle* essays in honor of A. Shimony (*The Western Ontario Series in Philosophy of Science*) ed W C Myrvold and J Christian (Berlin: Springer) pp 125–40
[28] Christensen B G, Liang Y-C, Brunner N, Gisin N and Kwiat P G 2015 *Phys. Rev.* X **5** 041052
[29] Bancal J-D, Branciard C, Brunner N, Gisin N and Liang Y-C 2012 *J. Phys. A: Math. Theor.* **45** 125301
[30] Cavalcanti D, Almeida M L, Scarani V and Acin A 2011 *Nat. Commun.* **2** 184
[31] Sen(De) A, Sen U, Brukner C, Buzek V and Żukowski M 2005 *Phys. Rev.* A **72** 042310

Communication Games Reveal Preparation Contextuality

Alley Hameedi,[1] Armin Tavakoli,[1,2] Breno Marques,[3,*] and Mohamed Bourennane[1]
[1]Department of Physics, Stockholm University, S-10691 Stockholm, Sweden
[2]Groupe de Physique Appliquée, Université de Genève, CH-1211 Genève, Switzerland
[3]Instituto de Física, Universidade de São Paulo, 05315-970 São Paulo, Brazil
(Received 18 May 2017; published 29 November 2017)

A communication game consists of distributed parties attempting to jointly complete a task with restricted communication. Such games are useful tools for studying limitations of physical theories. A theory exhibits preparation contextuality whenever its predictions cannot be explained by a preparation noncontextual model. Here, we show that communication games performed in operational theories reveal the preparation contextuality of that theory. For statistics obtained in a particular family of communication games, we show a direct correspondence with correlations in spacelike separated events obeying the no-signaling principle. Using this, we prove that all mixed quantum states of any finite dimension are preparation contextual. We report on an experimental realization of a communication game involving three-level quantum systems from which we observe a strong violation of the constraints of preparation noncontextuality.

DOI: 10.1103/PhysRevLett.119.220402

Introduction.—Communication games are tools by which one can study fundamental limiting features of physical theories in terms of their ability to process information [1–3]. In these games, a number of parties intend to jointly solve a task despite the amount and type of communication being constrained by some rules. Thus, the task can be solved only with some probability, which depends on the theory by which they are assumed to operate. Therefore, communication games are frequent tools for identifying and quantifying quantum advantages over classical theories [4–10].

Interestingly, there are known examples of communication games in which the better-than-classical performance constitutes a certificate of the system lacking a preparation noncontextual ontological model [11–13]. An ontological model is a way of explaining the physics of an operational theory, by assuming that there are independent and objective (ontic) states subject to experiment. However, specifying a preparation does not necessarily specify the ontic state. A preparation may be represented by a distribution μ over the ontic states. Let two preparations P_1 and P_2 associated to distributions μ_1 and μ_2 be indistinguishable, i.e., satisfy $p(b|P_1, M) = p(b|P_2, M)$ for any measurement M with outcome b. The assumption of preparation noncontextuality asserts that no additional features (called *contexts*) influence the physics of the preparations and, therefore, asserts that both preparations have equivalent representation in the ontological model: $\mu_1 = \mu_2$ [14]. If a theory does not satisfy this assumption, it is said to be preparation contextual. Preparation contextuality has been shown relevant for many foundational topics [3,15–18].

Here, we show that the performance of an operational theory in communication games constitutes a certificate of that theory exhibiting preparation contextuality. Specifically, we introduce communication constraints which keep the receiver oblivious about subsets of the information held by the sender. Preparation noncontextuality imposes a bound on the performance of any communication game executed under such an obliviousness constraint. This bound is violated by preparation contextual theories. Subsequently, we show how to understand no-signaling correlations from spacelike separated measurements (perhaps violating a Bell inequality) through a subclass of communication games. In particular, we find that quantum preparation contextuality manifested in communication games imposes a quantitative bound on quantum nonlocality (i.e., Bell inequality violations). Furthermore, we apply this result to resolve an open problem in this field: Which quantum states are preparation contextual? We show that all mixed quantum states in any finite dimension are preparation contextual. Finally, we present an experimental implementation of a quantum strategy in a specific communication game, inspired by the Collins-Gisin-Linden-Massar-Popescu (CGLMP) Bell inequality, in which three-level quantum systems are communicated. We certify a large violation of a preparation noncontextuality inequality.

Communication games.—In a two-player communication game, a party Alice (Bob) holds a set of data denoted $x \in I_A$ ($y \in I_B$) with associated probability distribution $p_A(x)$ [$p_B(y)$]. Alice encodes x by preparing a state which is sent to Bob, who attempts to decode it with a measurement labeled y. This returns an outcome b. Subsequently, a payoff $C^b_{x,y} \in \mathbb{R}$ is awarded. The average payoff earned by the partnership is

$$A \equiv \sum_{x\in I_A}\sum_{y\in I_B} C^b_{x,y} p_A(x) p_B(y) p(b|x,y). \tag{1}$$

Equation (1) quantifies the *performance* in the game. However, the content of Alice's communication to Bob is restricted by some communication constraints. These ensure that the game is nontrivial; i.e., Alice cannot simply send x to Bob. A suitable choice of these constraints enables the connection to tests of preparation contextuality.

Communication games as tests of preparation contextuality.—An operational theory is said to be preparation noncontextual [14] if operationally equivalent preparations imply equivalent distributions over the ontic states:

$$\forall y \quad \forall\, b\colon\ p(b|x,y) = p(b|x',y) \Rightarrow p(\lambda|x) = p(\lambda|x'), \tag{2}$$

where λ is a hidden variable, x and x' are two preparations, and y denotes a measurement.

We will now define a class of communication constraints which enables a connection to the premise of Eq. (2). The assumption of preparation noncontextuality then leads to a preparation noncontextuality inequality in which the performance in the communication game is the operator.

Construct L subsets of the space I_A; $S_k \subset I_A$ for $k = 1, \ldots, L$. Now, choose communication constraints as follows: Impose an *obliviousness* constraint

$$\forall\ y, b, k, k'\colon\ \frac{1}{q_k}\sum_{x\in S_k} p(x|b,y) = \frac{1}{q_{k'}}\sum_{x\in S_{k'}} p(x|b,y). \tag{3}$$

Here $q_k = p(x \in S_k) = \sum_{x\in S_k} p_A(x)$ serves as a normalization. In other words, Eq. (3) states that, no matter the performed measurement and observed outcome, Bob gains no information, as compared to what he knew before the communication, about to which set S_k the data x of Alice belong. Let us now apply Bayes' rule to the above summands: $p(x|b,y) = p(b|x,y)p(x|y)/p(b|y)$. Since x and y are independent, Eq. (3) becomes

$$\forall y \quad \forall\, b\colon \sum_{x\in S_k} p(b|x,y)\frac{p_A(x)}{q_k} = \sum_{x\in S_{k'}} p(b|x,y)\frac{p_A(x)}{q_{k'}}. \tag{4}$$

Note that each side is a convex combination, since $\{p_A(x)/q_k\}_{x\in S_k}$ is a probability distribution over the set S_k. Now, note that the probability that the outcome b was obtained from a measurement on a preparation associated to S_k is the convex mixing of its constitutes: $p(b|x \in S_k, y) = \sum_{x\in S_k} p(b|x,y)p_A(x)/q_k$. Similarly, the distribution of the hidden variable is $p(\lambda|x \in S_k) = \sum_{x\in S_k} p(\lambda|x)p_A(x)/q_k$. Putting it all together, we have $\forall y \quad \forall\, b\colon\ p(b|x \in S_k, y) = p(b|x \in S_{k'}, y)$, which takes the form of the premise of the preparation noncontextuality statement in Eq. (2). Thus, preparation noncontextuality imposes that $p(\lambda|x \in S_k) = p(\lambda|x \in S_{k'})$. Using Bayes' rule, we find that $p(x \in S_k|\lambda)/q_k = p(x \in S_{k'}|\lambda)/q_{k'}$. This means that, despite knowledge of the hidden variable, Eq. (3) remains satisfied.

Given any λ, Alice encodes x classically knowing that the obliviousness constraint is satisfied. Therefore, the preparation noncontextual bound p^{pnc} of Eq. (1) is obtained from maximizing Eq. (1) over all classical encodings respecting the obliviousness constraint. Hence, $A \le p^{\text{pnc}}$ is a preparation noncontextuality inequality. ∎

Clearly, for a given communication game, there are a plethora of ways in which one can choose the obliviousness constraint and construct the associated preparation noncontextuality inequality. In what follows, we will examine some interesting cases of the presented framework.

Communication games based on Bell inequalities.—Consider a general bipartite Bell experiment in which Alice and Bob share a two-particle state with each of them choosing measurements $X \in \{1, \ldots, m_A\}$, for some positive integer m_A, and $Y \in \{1, \ldots, m_B\}$, for some positive integer m_B, sampled from a distribution $p_A(X)$ and $p_B(Y)$, respectively. Each measurement returns an outcome $a, b \in \{1, \ldots, d\}$. From the resulting probability distribution $p(a,b|X,Y)$, one constructs a general Bell inequality

$$I_b \equiv \sum_{abXY} \mathcal{C}^{a,b}_{X,Y} p_A(X) p_B(Y) p(a,b|X,Y) \le C, \tag{5}$$

where C is the local realist bound and $\mathcal{C}^{a,b}_{X,Y}$ are real coefficients.

In the following, we construct a family of communication games and obliviousness constraints inspired by such Bell experiments. Alice is given inputs $(x, x_0) \in \{1, \ldots, m_A\} \times \{1, \ldots, d\}$ admitting the distribution $p(x_0, x) = p_g(x_0|x)p_A(x)$, with $p_A(x = i) = p_A(X = i)$ whereas $p_g(x_0|x)$ is yet to be specified. Bob has an input $y \in \{1, \ldots, m_B\}$ with distribution $p_B(y = i) = p_B(Y = i)$. The inputs (x_0, x, y) in the communication game, respectively, correspond to (a, X, Y) in the Bell experiment. Having received Alice's communication, Bob earns a payoff $\mathcal{C}^{x_0,b}_{x,y}$ if he outputs b given a measurement of y and that Alice held (x_0, x). The performance is written

$$I_g[\{p_g(x_0|x)\}_x] \equiv \sum_{x_0 x y b} \mathcal{C}^{x_0,b}_{x,y} p_A(x) p_B(y) p_g(x_0|x) p(b|x_0, x, y). \tag{6}$$

Notice that, for every choice of $\{p_g(x_0|x)\}_x$, we have a different communication game.

Alice's communication must satisfy the following obliviousness constraint. Partition Alice's $m_A d$ possible inputs into m_A sets each containing d elements; we define $S_k = \{x_0 x | x = k\}$ for $k = 1, \ldots, m_A$. The obliviousness constraint requires that Bob gains no information about to which S_k the data (x_0, x) belong. Inserting this into Eq. (4) with $q_k = p_A(x = k)$ and using Bayes' rule, we obtain

$$\forall\, b,y,k,k'\colon \sum_{x_0=1}^{d} p(x_0,b|x=k,y)=\sum_{x_0=1}^{d} p(x_0,b|x=k',y). \tag{7}$$

This constraint is an analogy of the directed no-signaling principle imposed by special relativity on correlations in spacelike separated measurement events: The probability of Bob's outcome marginalized over Alice's input x_0 is independent of Alice's other input x. One needs only to relabel x_0 by a and (x,y) by (X,Y) to recover the corresponding statement in Bell experiments.

On the one hand, imagine we run a Bell experiment and achieve some value of I_b. Using Bayes' theorem and the obliviousness constraint (7), it is straightforwardly shown that if we choose the communication game in which p_g coincides with the observed marginals of Alice, $p(a|X)$, one finds $I_g = I_b$. We explicitly consider the case of the quantum theory. In a Bell experiment, when Alice performs her measurement X, she renders Bob's local state in one of d possible states labeled ϱ_l^X for $l = 1,\ldots,d$. The probability of Bob's local state being ϱ_l^X is the probability of Alice obtaining outcome l, i.e., $p(a=l|X)$. No signaling implies that the average state of Bob is independent of the measurement choice X of Alice. We associate for every X the set $\{\varrho_l^X\}_{l=1}^d$ to the states in S_X prepared by Alice in our communication game. As shown, these will necessarily satisfy the obliviousness constraint (7) while by construction returning the same performance in the communication game (6) as in the Bell experiment, namely, $I_g = I_b$.

On the other hand, imagine we have not specified $p_g(x_0|x)$. Let λ index all functions $f_\lambda(x)\colon\{1,\ldots,m_A\}\to\{1,\ldots,d\}$. By choosing a suitable probability distribution $\mu(\lambda)$, we can write $p_g(x_0|x)=\sum_\lambda \mu(\lambda) D_A(x_0|x\lambda)$, where $D_A(x_0|x\lambda)=\delta_{f_\lambda(x),x_0}$. Alice then communicates λ, which contains no information about x, to Bob, who decodes the message using some strategy D_B. We find

$$I_g=\sum_{x_0xyb} \mathcal{C}_{x,y}^{x_0,b}\, p_A(x)p_B(y)\sum_\lambda \mu(\lambda) D_A(x_0|x\lambda)D_B(b|y\lambda). \tag{8}$$

This is precisely the notion of local realist models for the Bell experiment (5). Hence, if we choose $p_g(x_0|x)$ such that there is a local hidden variable strategy that both (i) has $p_g(x_0|x)=p(a|X)$ as a marginal of Alice and (ii) saturates the local realist bound C of (5), the preparation noncontextuality inequality $I_g \le C$ will be tight. Of particular interest is to choose $p_g(x_0|x)$ such that it coincides with Alice's marginals in a maximal violation of a Bell inequality given some operational no-signaling theory. Then, we assert that I_g can witness a violation of preparation noncontextuality corresponding to the maximal Bell inequality violation.

Note that only very particular obliviousness constraints and communication games retain the analogy to the no-signaling principle through our construction. In Ref. [19], we present a family of games that is not of the type presented in this section. The corresponding preparation noncontextuality inequalities are many-outcome generalizations of the those based on parity-oblivious multiplexing [11].

Quantum preparation contextuality limits maximal quantum nonlocality.—If Alice and Bob share entangled states, all mixed states can be prepared on Bob's side by considering the average of his local state computed over the outcomes of Alice obtained from some measurement. Thus, due to our previous discussion, it follows that the maximal quantum violation of a bipartite Bell inequality is a limitation imposed by the preparation contextuality allowed in the quantum theory. This generalizes the result of Ref. [3], showing this statement for the Clauser-Horne-Shimony-Holt inequality [20]. We exemplify this generalization by shining light on the numerical quantum violations of the preparation noncontextuality inequalities considered in Ref. [13]. These inequalities were based on communication games which happen to admit an obliviousness constraint of the form considered in the above section. The corresponding Bell inequalities were in fact studied in Ref. [21] in a different context. Comparing the numerics for quantum preparation contextuality [13] and the quantum nonlocality [21], one indeed finds that these agree very accurately.

All mixed states are preparation contextual.—The maximally mixed quantum state of dimension $d = 2, 3, 4, 5$ is known to be preparation contextual [13,14]. So is every mixed qubit state [22]. Our mapping between communication games and Bell inequalities allows us to straightforwardly show that all mixed quantum states of any dimension d are preparation contextual. For this purpose, consider the CGLMP Bell inequality [23], which is a bipartite facet Bell inequality with d outcomes for both observers. For any d, this inequality can be violated by all pure bipartite entangled states of dimension d [24]. Hence, all possible mixed quantum states of dimension d can appear as the average state of Bob after either of Alice's measurements. That average state is just the state of Bob's part of the entangled system. Since quantum strategies in the Bell scenario can be mapped to quantum strategies in a communication game (of the form previously discussed) testing preparation contextuality, it follows that all mixed quantum states of dimension d are preparation contextual.

A specific communication game.—Let us focus on the CGLMP Bell inequality with $d = 3$ and construct the preparation noncontextuality inequality based on the associated communication game. Following our previous discussion, we let Alice hold $x = x_0 x \in \{0,1,2\}\times\{0,1\}$ with $p(x_0,x)=1/6$ and Bob hold $y\in\{0,1\}$ with $p(y)=1/2$. In order to satisfy the obliviousness constraint, Alice's communication ρ_{x_0x} must in the quantum theory obey $\sum_{x_0=0}^{2} p(x_0|x=0)\rho_{x_0 0}=\sum_{x_0=0}^{2} p(x_0|x=1)\rho_{x_0 1}$. Since the preparation noncontextual bound coincides with the local bound of the CGLMP inequality (which achieves its maximal quantum violation with uniform marginals on Alice), our preparation noncontextuality inequality reads

$$A_3 \equiv \frac{1}{12}\sum_{x_0xyk}(-1)^k p(b = T_k|x_0, x, y) \leq 1/2, \quad (9)$$

where $T_k = x_0 - (-1)^{x+y+k}k - xy \bmod 3$ for $k = 0, 1$. The maximal quantum violation of the CGLMP Bell inequality is $A_3 = (3 + \sqrt{33})/12 \approx 0.7287$ [25], which immediately translates into an equal quantum violation of the inequality (9). In Ref. [19], we give the details of the corresponding quantum strategy in the communication game.

Experiment.—We experimentally confirm the above prediction of quantum preparation contextuality. The experimental implementation of the communication game uses three-path encoding for preparing qutrits. Single photons are initially prepared in the $|H\rangle$ polarization state by the use of polarization fiber controllers in a single-mode fiber (SMF). The qutrit state is prepared using the two spatial modes of three polarization beam splitters (PBSs) (see Fig. 1). The states required for the game, $|\psi_{\text{in}}\rangle = \cos(2\chi_1)|0\rangle + \sin(2\chi_1)\sin(2\chi_2)|1\rangle + \sin(2\chi_1)\cos(2\chi_2)|2\rangle$, are prepared by suitably orienting the half-wave plates (HWPs) χ_1 and χ_2. Details are given in Ref. [19].

We use a heralded single-photon source generating twin photons at 780 nm by spontaneous parametric down-conversion. In this process, a nonlinear crystal type II (β-barium borate) is pumped using a high-power femtosecond laser such that a pump photon probabilistically converts into two lower-energy photons, called the signal and idler. The twin photons pass through a 3 nm filter and are coupled into single-mode fibers to have well-defined spatial and spectral properties. A detection of the idler then heralds the signal photon.

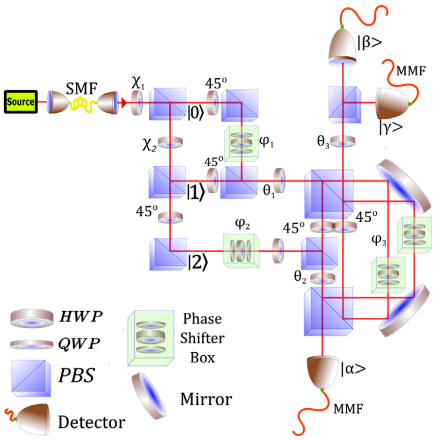

FIG. 1. Experimental setup. Suitable settings of χ_1 and χ_2 allow us to produce the desired qutrit states for the task. Measurement basis selection is implemented by appropriate settings of HWPs θ_1, θ_2, and θ_3 and by setting the total experimental phase ($\phi_i; i \in 1, 2, 3$) between path modes by employing a phase shifter box (QWP-HWP-QWP) inside the setup. Detection events in detectors $|\alpha\rangle$, $|\beta\rangle$, and $|\gamma\rangle$ are used to obtain the respective probabilities.

The corresponding experimental setup consists of three subsequent interferometers comprising of single-photon interferometers between all three paths followed by a stable and compact Sagnac interferometer, such that, while performing a measurement in a given measurement basis, the state is projected into basis vectors of the chosen basis. The protocol requires measurements in the computational basis and a second basis defined in Ref. [19]. Moreover, state tomography is performed using measurements in four mutually unbiased bases (MUBs), so that the total set of measurements is informationally complete [26]. For this purpose, the choice of a given measurement basis is enabled by suitable orientations of the HWPs θ_1, θ_2, and θ_3 (see Table I in Ref. [19]) and by the introduction of a phase ($\phi_i; i \in 1, 2, 3$) between the special modes by employing a set of three wave plates QWP-HWP-QWP (phase shifter box) geometries at different tilding positions [27].

A measurement projects the state onto the basis vectors. These are represented by the spatial modes of the two PBSs in the Sagnac interferometer (denoted by $|\alpha\rangle$, $|\beta\rangle$, and $|\gamma\rangle$). In our experiment, the photons arriving at $|\alpha\rangle$, $|\beta\rangle$, and $|\gamma\rangle$ are collected by multimode fibers that are in turn coupled to single-photon silicon avalanche photodiodes from Excelitas Technologies with an effective detection efficiency $\eta_d = 0.55$. A home-built field-programmable gate array-based timing system records the coincidence events between the arriving and trigger (idler) photons with a detection time window of 1.7 ns. The number of detection events at each detector is used to compute the respective probabilities. In each measurement round, approximately 60000 photons were detected per second. The measurement time was 10 s.

From the measured probabilities, we computed $A_3^{\text{pri}} \approx 0.7172 \pm 0.0365$, which is in good agreement with the theoretical prediction. We reconstructed the states using variational quantum tomography [26,28] and the experimental results from four MUBs. We found the following fidelities for the six states: $|\psi_{11}\rangle \sim 0.9826$, $|\psi_{12}\rangle \sim 0.9804$, $|\psi_{13}\rangle \sim 0.9893$, $|\psi_{21}\rangle \sim 0.9838$, $|\psi_{22}\rangle \sim 0.9876$, and $|\psi_{23}\rangle \sim 0.9840$. These small imperfections cause the obliviousness constraint not to be perfectly satisfied. Next, we shall see how to overcome this issue.

Data analysis.—Reference [29] constructed a method in which one maps measured outcome probabilities (primary data), which does not perfectly satisfy a strict equivalence constraint, into another set of probabilities (secondary data) that satisfies that equivalence constraint. Then, one uses the secondary data to calculate the parameter of interest in the experiment. We will use this method to strictly enforce the obliviousness constraint and then compute A_3.

The primary data in our experiment consist of six 2×3 matrices [one for each preparation (x_0, x)] with elements $\mathbf{P}_{i,j}^{x_0 x} \equiv P^{\text{lab}}(j|x_0, x, i)$ corresponding to performing measurement i in the laboratory and obtaining outcome j. We will

assume that the underlying physical theory governing the system is linear, allowing us to search for secondary data in the form of six other matrices $\{\mathbf{P}'^{x_0,x}\}_{x_0,x}$ that are in the convex hull of $\{\mathbf{P}^{x_0,x}\}_{x_0,x}$. That is, we let $\forall\, x_0,x\colon \mathbf{P}'^{x_0,x} = \sum_{x_0'=0}^{2}\sum_{x'=0}^{1} w_{x_0',x'}^{x_0,x}\mathbf{P}^{x_0',x'}$, where $\forall\ x_0,x\colon w_{x_0',x'}^{x_0,x}$ is a probability distribution. We seek secondary data which (i) satisfy the obliviousness constraint and (ii) on average are as close to the primary data as possible. This corresponds to a linear program:

$$S \equiv \max_{\{w\}} \frac{1}{6}\sum_{x_0=0}^{2}\sum_{x=0}^{1} w_{x_0,x}^{x_0,x},$$
$$\text{such that} \sum_{x_0}\mathbf{P}'^{x_0 0} = \sum_{x_0}\mathbf{P}'^{x_0 1}. \tag{10}$$

We find $S \approx 0.9938$, indicating that the secondary data are close to the primary data. Using the secondary data to compute A_3, we obtain $A_3^{\text{sec}} \approx 0.7118 \pm 0.0365$. This is only marginally smaller than A_3^{pri}. It is in good agreement with the theoretical prediction of the quantum theory and strictly satisfies the obliviousness constraint.

Conclusions.—We have established relations between operational statistics in a class of communication games and tested preparation contextuality. We showed close relations between quantum nonlocality and quantum correlations in such communication games and also shown all mixed quantum states of finite dimension to be preparation contextual. Furthermore, we provided an experimental demonstration of a quantum communication game showing a large violation of a preparation noncontextuality inequality.

We conclude with some open problems: (i) Do communication games without obliviousness constraints admit a connection to some operational physical assumption in the same spirit as presented here for games respecting an obliviousness constraint? (ii) Are generalizations of the presented framework to more than two players possible? (iii) Can the considered communication games be used in one-sided device-independent cryptography protocols?

The authors thank Adán Cabello, Nicolas Gisin, Nicolas Brunner, Thiago Maciel, and Artur Matoso for the useful discussions and comments. We extend particular gratitude to Debashis Saha and Anubhav Chaturvedi for enlightening comments and criticism. The project was financially supported by Knut and Alice Wallenberg foundation and the Swedish research council. A. T. acknowledges financial support from the Swiss National Science Foundation (starting grant DIAQ). B. M. is supported by FAPESP No. 2014/27223-2.

A. H. and A. T. contributed equally for this project.

*bmgt@if.usp.br

[1] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski, Nature (London) **461**, 1101 (2009).
[2] A. Grudka, K. Horodecki, M. Horodecki, W. Kłobus, and M. Pawłowski, Phys. Rev. Lett. **113**, 100401 (2014).
[3] M. Banik, S. S. Bhattacharya, A. Mukherjee, A. Roy, A. Ambainis, and A. Rai, Phys. Rev. A **92**, 030103(R) (2015).
[4] R. Gallego, N. Brunner, C. Hadley, and A. Acín, Phys. Rev. Lett. **105**, 230501 (2010).
[5] M. Hendrych, R. Gallego, M. Micuda, N. Brunner, A. Acín, and J. P. Torres, Nat. Phys. **8**, 588 (2012).
[6] J. Ahrens, P. Badziag, A. Cabello, and M. Bourennane, Nat. Phys. **8**, 592 (2012).
[7] M. Pawłowski and N. Brunner, Phys. Rev. A **84**, 010302(R) (2011).
[8] H.-W. Li, M. Pawłowski, Z.-Q. Yin, G.-C. Guo, and Z.-F. Han, Phys. Rev. A **85**, 052308 (2012).
[9] V. D'Ambrosio, F. Bisesto, F. Sciarrino, J. F. Barra, G. Lima, and A. Cabello, Phys. Rev. Lett. **112**, 140503 (2014).
[10] A. Tavakoli, A. Hameedi, B. Marques, and M. Bourennane, Phys. Rev. Lett. **114**, 170502 (2015).
[11] R. W. Spekkens, D. H. Buzacott, A. J. Keehn, B. Toner, and G. J. Pryde, Phys. Rev. Lett. **102**, 010401 (2009).
[12] A. Chailloux, I. Kerenidis, S. Kundu, and J. Sikora, New J. Phys. **18**, 045003 (2016).
[13] A. Ambainis, M. Banik, A. Chaturvedi, D. Kravchenko, and A. Rai, arXiv:1607.05490.
[14] R. W. Spekkens, Phys. Rev. A **71**, 052108 (2005).
[15] R. W. Spekkens, Phys. Rev. Lett. **101**, 020401 (2008).
[16] M. F. Pusey, Phys. Rev. Lett. **113**, 200401 (2014).
[17] M. S. Leifer and O. J. E. Maroney, Phys. Rev. Lett. **110**, 120401 (2013).
[18] R. Kunjwal and R. W. Spekkens, Phys. Rev. Lett. **115**, 110403 (2015).
[19] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.119.220402 for (i) the theory part of SM shows a family of preparation noncontextuality inequalities based on parity-obliviousness and the details of the quantum strategy realized by the experiment, (ii) the experimental part of SM shows details for the experimental settings and the results obtain by our implementation.
[20] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).
[21] A. Tavakoli, B. Marques, M. Pawłowski, and M. Bourennane, Phys. Rev. A **93**, 032336 (2016).
[22] M. Banik, S. S. Bhattacharya, S. K. Choudhary, A. Mukherjee, and A. Roy, Found. Phys. **44**, 1230 (2014).
[23] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu, Phys. Rev. Lett. **88**, 040404 (2002).
[24] J.-L. Chen, D.-L. Deng, and M.-G. Hu, Phys. Rev. A **77**, 060306(R) (2008).
[25] A. Acin, T. Durt, N. Gisin, and J. I. Latorre, Phys. Rev. A **65**, 052325 (2002).
[26] D. S. Gonçalves, C. Lavor, M. A. Gomes-Ruggiero, A. T. Cesário, R. O. Vianna, and T. O. Maciel, Phys. Rev. A **87**, 052140 (2013).
[27] B. G. Englert, C. Kurtsiefer, and H. Weinfurter, Phys. Rev. A **63**, 032303 (2001).
[28] T. O. Maciel, A. T. Cesário, and R. O. Vianna, Int. J. Mod. Phys. C **22**, 1361 (2011).
[29] M. D. Mazurek, M. F. Pusey, R. Kunjwal, K. J. Resch, and R. W. Spekkens, Nat. Commun. **7**, ncomms11780 (2016).

Quantum predictions for an unmeasured system cannot be simulated with a finite-memory classical system

Armin Tavakoli[1,2,*] and Adán Cabello[3,†]
[1]*Groupe de Physique Appliquée, Université de Genève, 1211 Genève, Switzerland*
[2]*Department of Physics, Stockholm University, 10691 Stockholm, Sweden*
[3]*Departamento de Física Aplicada II, Universidad de Sevilla, 41012 Sevilla, Spain*

(Received 31 May 2017; published 29 March 2018)

We consider an ideal experiment in which unlimited nonprojective quantum measurements are sequentially performed on a system that is initially entangled with a distant one. At each step of the sequence, the measurements are randomly chosen between two. However, regardless of which measurement is chosen or which outcome is obtained, the quantum state of the pair always remains entangled. We show that the classical simulation of the reduced state of the distant system requires not only unlimited rounds of communication, but also that the distant system has infinite memory. Otherwise, a thermodynamical argument predicts heating at a distance. Our proposal can be used for experimentally ruling out nonlocal finite-memory classical models of quantum theory.

DOI: 10.1103/PhysRevA.97.032131

I. INTRODUCTION

It has been shown recently that an experiment in which a single quantum system is subjected to many sequential measurements, each randomly chosen among n alternatives, cannot be simulated with a classical system that has access only to finite memory. Such a classical simulation has an additional cost, namely, that after sufficiently many measurements the system dissipates heat. Specifically, the amount of heat per measurement tends to infinity and the divergence is linear in n [1].

Such finite-memory classical simulations can be put in one-to-one correspondence with interpretations of quantum theory in which measurement outcomes are governed by intrinsic properties and, in addition, satisfy some assumptions [1]. Therefore, an experiment testing the presence or absence of such heat would rule out some interpretations of quantum theory. However, such an experiment is exposed to the practical problem of the implementation of the sequential measurements themselves producing heat, making it difficult to distinguish the hypothetical heat emitted by the finite-memory classical systems. Therefore, an interesting problem is whether a similar phenomenon can be demonstrated using sequential measurement that are *not* performed on the same physical system as from which the heat would originate. Such an experiment would require at least two systems, one that is being repeatedly measured and one in which the heat could appear.

In order for measurements on one system to influence the quantum state of the other, the joint state of the two systems must have some entanglement. A complete projective local measurement performed on an entangled state renders the postmeasurement state separable. Thus, a second local measurement can no longer change the quantum state of the distant system. Therefore, the local measurements must necessarily be nonprojective positive-operator-valued measures (POVMs) in order to both induce a change in the local state of the distant system and retain this ability in a subsequent measurement. Sequential nonprojective local measurements have previously been shown useful in Bell experiments [2] and random number generation [3]. Here we show that sequential measurements can also be used to distinguish the predictions of quantum theory from classical simulations with finite memory in experiments involving two distant entangled systems.

In Sec. II, we introduce a protocol in which entanglement is preserved indefinitely for all measurement choices. Then, in Sec. III, we compute the cost of classically simulating some possible predictions of quantum theory for this experiment. We show that, in addition to an always increasing (but finite) amount of communication required for simulating entanglement in standard Bell experiments [4,5], a thermodynamical analysis imposes an additional and qualitatively different cost: infinite local memory. Otherwise, an experiment would be able to detect the heat emitted by the system that is not measured.

II. PROTOCOL

A. Scenario

We consider two parties, Alice and Bob, who at time t_0 share two qubits in a maximally entangled state

$$|\psi_0\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle). \tag{1}$$

At later times $t_1 < t_2 < \cdots < t_N$, Alice randomly chooses between two measurements x_k and $\bar{x}_k$ and performs this measurement on her qubit. Each measurement has two possible outcomes denoted by 0 and 1. The two measurements between which Alice measures are not preestablished but depend on the previous measurements and outcomes. Bob does not perform any operation over the course of the protocol. However, at any time, the parties can stop the protocol and perform

*armin.tavakoli@unige.ch
†adan@us.es

measurements (including Bob) to test some predictions of quantum theory.

Any of Alice's measurement at time t_k, denoted by $j_k \in \{x_k, \bar{x}_k\}$, will be a two-outcome POVM which has, associated with outcome 0, the POVM element $E_k^{j_k} = K_{\hat{n}_{j_k}}(\mu_k) K_{\hat{n}_{j_k}}(\mu_k)^\dagger$, where $K_{\hat{n}_{j_k}}(\mu_k)$ is the Kraus operator [6]

$$K_{\hat{n}_{j_k}}(\mu_k) = \cos(\mu_k)|\hat{n}_{j_k}\rangle\langle\hat{n}_{j_k}| + \sin(\mu_k)|-\hat{n}_{j_k}\rangle\langle-\hat{n}_{j_k}|, \tag{2}$$

where $\hat{n}_{j_k}$ is a vector on the Bloch sphere that will be specified later. This POVM is a noisy version of the measurement represented by a Pauli matrix along $\hat{n}_{j_k}$. The amount of noise is controlled by the value of $\mu_k \in [0, \pi/2]$ and will be specified later. If $\mu_k \in \{0, \pi/2\}$, the measurement is projective. If $\mu_k = \pi/4$, then $K_{\hat{n}_{j_k}} = \mathbb{1}/2$, implying a noninteractive measurement. Other values of μ_k correspond to weak measurements.

In addition, we assume that the time evolution is trivial, that is, that the state of Alice's and Bob's qubits just after Alice's measurement at t_k is the state just before Alice's measurement at t_{k+1}, and is determined by Alice's sequence of measurements and outcomes at $\{t_1, \ldots, t_k\}$. The list of measurements and outcomes of Alice from t_1 to t_k will be denoted by l_k.

B. Choosing sequential measurements that always enable Bell inequality violation

One of the features of the protocol that we are about to introduce is that, at each t_k, for each pair of measurements of Alice, there exist two measurements that Bob *could* perform (if the parties agreed to stop the protocol at this particular t_k) such that the outcome statistics of Alice and Bob would violate the Clauser-Horne-Shimony-Holt (CHSH) inequality [7]. Recall that in a CHSH experiment, Alice and Bob perform measurements A_i and B_j, respectively, with $i, j \in \{0, 1\}$, on shared pairs of systems. The measurement on each system is chosen independently and randomly. Any local realistic model of the outcome statistics must satisfy the CHSH inequality

$$S_{\text{CHSH}} \equiv \langle A_0 B_0\rangle + \langle A_0 B_1\rangle + \langle A_1 B_0\rangle - \langle A_1 B_1\rangle \leqslant 2, \tag{3}$$

where $\langle\cdot\rangle$ denotes expectation value.

The following lemma (which is a corollary of the main result of Ref. [8]) explains why it is possible to achieve the feature described above.

Lemma. Consider any pure entangled state $|\Psi_\eta\rangle = \cos(\eta)|00\rangle + \sin(\eta)|11\rangle$, with $\eta \in (0, \pi/2)$. For every $|\Psi_\eta\rangle$, Alice can find measurements associated with Kraus operators (2) with $\hat{n}_{j_k}$ equal to (0,0,1) and (1,0,0), respectively (i.e., noisy measurements of σ_z and σ_x), for which she can choose a noise parameter $\mu \notin \{0, \pi/2\}$ such that there exist two projective measurements for Bob leading to outcome statistics violating the CHSH inequality (3).

Proof. The Bloch vectors associated with the measurements A_0 and A_1 of Alice are $[0, 0, \cos(2\mu)]$ and $[\cos(2\mu), 0, 0]$, respectively. These are unnormalized for $\mu \notin \{0, \pi/2\}$. Let us choose the Bloch vectors representing Bob's measurements B_0 and B_1 to be of the form $[\cos(\theta), 0, \sin(\theta)]$ and $[-\cos(\theta), 0, \sin(\theta)]$, respectively, for some θ. These Bloch vectors are normalized and hence correspond to projective measurements. A direct computation of S_{CHSH} in (3) gives

$$S_{\text{CHSH}} = 2\cos(2\mu)[\sin(\theta) + \sin(2\eta)\cos(\theta)]. \tag{4}$$

We choose θ so that S_{CHSH} is maximal, i.e., we solve the equation $\partial S_{\text{CHSH}}/\partial\theta = 0$. The solution of our interest is $\theta = \arctan[1/\sin(2\eta)]$, which is independent of μ. Inserting this in Eq. (4), we find

$$S_{\text{CHSH}} = \sqrt{6 - 2\cos(4\eta)}\cos(2\mu). \tag{5}$$

The minimal value of the square root is 2 and is achieved for product states. Therefore, for every entangled state corresponding to $\eta \notin \{0, \pi/2\}$, the square root is larger than 2. Hence, if Alice chooses her noise parameter μ such that

$$0 < \mu < \frac{1}{2}\arccos\left[\frac{2}{\sqrt{6 - 2\cos(4\eta)}}\right] \equiv F(\eta), \tag{6}$$

then the outcome statistics of Alice and Bob will violate the CHSH inequality (3).

C. Protocol

Let us now describe the protocol itself.

(0) At time t_0, Alice and Bob share the maximally entangled state $|\psi_0\rangle$ given in Eq. (1).

(1a) At $t_1 > t_0$, Alice chooses some nonzero $\mu_1 < F(\pi/4) = \pi/8$. Then she randomly chooses between x_1 and $\bar{x}_1$, each of them associated with a Bloch vector (0,0,1) and (1,0,0), respectively. Alice's choice is denoted by j_1. Then Alice performs the measurement $\{E_1^{j_1}, \mathbb{1} - E_1^{j_1}\}$. The Lemma ensures that, for the state before the measurement and Alice's two possible measurements, there are two possible measurements on Bob's system violating the CHSH inequality.

(1b) From her observed outcome, Alice calculates the postmeasurement state $|\psi_1^{l_1}\rangle$ of the two qubits. This state is necessarily pure and entangled and can be written in the form

$$|\psi_1^{l_1}\rangle = U_A^{l_1} \otimes U_B^{l_1}[\cos(\theta^{l_1})|00\rangle + \sin(\theta^{l_1})|11\rangle], \tag{7}$$

where $\theta^{l_1} \notin \{0, \pi/2\}$ and $U_A^{l_1}$ and $U_B^{l_1}$ are unitary operators. Here θ^{l_1} does not refer to an actual operation but is a hypothetical angle which would maximize Eq. (5). Then Alice applies on her qubit the unitary $(U_A^{l_1})^\dagger$, which cancels the unitary $U_A^{l_1}$ in Eq. (7). After Alice's actions at t_1, the reduced state of Bob's qubit is one of four possible states (see Fig. 1).

(2a) At t_2, Alice again chooses some positive $\mu_2 < F(\theta^{l_1})$. She makes a random choice of measurement $j_2 \in \{x_2, \bar{x}_2\}$ associated with Bloch vectors (0,0,1) and (1,0,0), respectively, and performs the measurement $\{E_2^{j_2}, \mathbb{1} - E_2^{j_2}\}$. Again, the Lemma ensures that, for the state before the measurement and Alice's two possible measurements, there are two possible measurements on Bob's system violating the CHSH inequality.

(2b) From her observed outcome, Alice calculates the new postmeasurement state $|\psi_2^{l_2}\rangle$ of the two qubits. Just as in (1b), Alice rotates her reduced state back to the computational basis by applying a suitable unitary. After Alice's actions at t_2, the reduced state of Bob's qubit is one of 16 possible states (see Fig. 1).

Alice continues this process of measuring, recording the outcome, and choosing the next measurement indefinitely. That is:

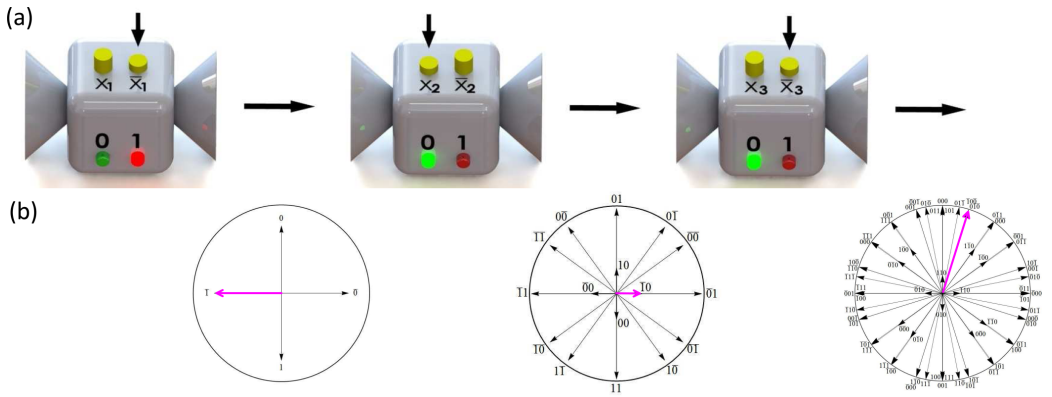

FIG. 1. (a) From left to right, Alice's sequential measurements at times $t_1 < t_2 < t_3$, respectively. At each t_k, Alice performs a measurement, either x_k or $\bar{x}_k$. Each measurement has two possible outcomes: 0 and 1. Alice's measurements are such that the state of the two qubits after her measurement is always entangled. (b) From left to right, possible reduced states of Bob's qubit after Alice's measurements at $t_1 < t_2 < t_3$, respectively. States are represented by nonunit arrows in the equatorial plane of the Bloch sphere. For example, $\bar{1}$ denotes the state when Alice measured $\bar{x}_1$ at t_1 and obtained outcome 1; $\bar{1}0$ denotes the state when Alice measured $\bar{x}_1$ at t_1 and x_2 at t_2 and obtained outcomes 1 and 0, respectively. Bob's states highlighted in purple are those produced in the particular sequence of Alice's measurements and outcomes shown in (a).

(ta) At t_k, she chooses some positive $\mu_k < F(\theta^{l_{k-1}})$, randomly chooses a measurement $j_k \in \{x_k, \bar{x}_k\}$ associated with Bloch vectors (0,0,1) and (1,0,0), respectively, and performs the measurement $\{E_k^{j_k}, \mathbb{1} - E_k^{j_k}\}$. The Lemma guarantees that, for the state before the measurement and Alice's two possible measurements, there are two possible measurements on Bob's system violating the CHSH inequality.

(tb) From her observed outcome, Alice calculates the postmeasurement state $|\psi_k^{l_k}\rangle$, which takes the form

$$\left|\psi_k^{l_k}\right\rangle = U_A^{l_k} \otimes U_B^{l_k}[\cos(\theta^{l_k})|00\rangle + \sin(\theta^{l_k})|11\rangle] \qquad (8)$$

for some angle $\theta^{l_k} \notin \{0, \pi/2\}$. Subsequently, she undoes the rotation of her local state by applying $(U_A^{l_k})^\dagger$. This renders the reduced state of Bob's qubit in one of 4^k possible states (see Fig. 1).

D. Properties of the protocol

At each time t_k, Alice's alternative measurements are *both* nonprojective and depend on Alice's previous choices of measurements and also on the outcomes of the previous measurements. This way, the initial entanglement is never consumed regardless of Alice's performed measurements and observed outcomes and Alice's two measurement options enable a violation of the CHSH inequality.

To illustrate the properties of the protocol, in Table I we display data from the first few steps of one possible execution of the protocol. There we can see that at each time step, the measurement of Alice becomes stronger without ever becoming projective. Furthermore, the entanglement, quantified by the negativity [9], remains nonzero. From t_2 onward, not all of the 4^k possible states just after t_k contain the same amount of entanglement and therefore we must consider the weakest possible entanglement. As displayed in Table I, the negativity of the weakest entangled state quickly decreases. However, some entanglement is always present. When choosing her noise parameter μ_k, Alice ensures that even the weakest entangled state violates the CHSH inequality (3). That this is indeed the case can be seen from the corresponding smallest values of S_{CHSH} in Table I. In contrast, from the largest possible

TABLE I. Data from the first four timesteps in one possible execution of the quantum protocol: choices of the noise parameter for Alice's measurement at t_k, the number of different local states of Bob just after t_k, the smallest and largest negativity of the 4^k possible global states just after t_k, and the smallest and largest values of S_{CHSH} achieved with the 4^{k-1} possible states. The choices of μ carry no special significance other than that they satisfy the relation $0 < \mu_k < F(\theta^{l_{k-1}})$ for all k.

Time	μ	Number of possible states of Bob's qubit	Smallest negativity	Largest negativity	Smallest value of S_{CHSH}	Largest value of S_{CHSH}
t_0		1	0.5	0.5		
t_1	$\pi/9$	4	0.3214	0.3214	2.1667	2.1667
t_2	$\pi/12$	16	0.0966	0.4774	2.0590	2.0590
t_3	$\pi/40$	64	0.0077	0.4887	2.0119	2.7313
t_4	$\pi/500$	256	0.00005	0.4902	2.00008	2.7965

negativity we see that the protocol sometimes, albeit with small probability, acts as a probabilistic entanglement amplification scheme [10,11].

III. CLASSICAL SIMULATION OF BOB'S LOCAL STATE

A. Cost 1: Unlimited rounds of communication

We now consider the cost of classically simulating the evolution of the quantum reduced state of Bob's system induced by Alice's sequential measurements on her distant system. A classical simulation of Bob's local state must be able to, at any time t_k, account for the statistical outcomes of every possible quantum measurement that Bob may apply. We will not consider this problem in its full generality, as it is not the emphasis of our work. For our purposes, it suffices to show that after each of Alice's sequential measurements, the classical simulation needs to be supplemented with some amount of communication.

Since the postmeasurement state just after t_k and the two measurement options Alice has at time t_{k+1} could always be used, in conjunction with suitable measurements of Bob, to violate the CHSH inequality, any local realist model aiming to simulate these quantum predictions has to be supplemented with some communication [4,5]. Therefore, there must be a round of communication between Alice and Bob after every measurement performed by Alice. In this round, depending on her measurement and the resulting local realistic state of her system, Alice communicates some information to Bob. Communication from Bob to Alice is of no use since Bob does not perform any operations on his qubit over the course of the protocol.

The critical observation is that the communication required to simulate the quantum predictions just after t_k will not be enough to reproduce the quantum predictions after t_{k+1}. The reason is that, at t_{k+1}, the new quantum predictions could again be use to violate the CHSH inequality. Thus, regardless of Alice's measurement choice and observed outcome, any simulation of the predictions of quantum theory for the experiment based on a local realistic model complemented with communication requires unlimited rounds of communication.

Note that the *total* amount of communication required to simulate the ability of the state to violate the CHSH inequality at each time step is finite. To show this, we use that the average amount of communication $\mathcal{C}$ required to simulate a nonsignaling probability distribution achieving the value S_{CHSH} is given by $\mathcal{C} = S_{\mathrm{CHSH}}/2 - 1$ [5]. Let us denote the average communication over all possible postmeasurement states at time t_k by $\mathcal{C}_k$. The total amount of communication is finite if $\bar{\mathcal{C}} \equiv \sum_{k=1}^{\infty} \mathcal{C}_k$ is finite. To show that $\bar{\mathcal{C}}$ is finite, we consider the states $|\psi_\eta\rangle$, which are unitarily equivalent to the states shared by Alice and Bob. Applying the Horodecki criterion [12] to $|\psi_\eta\rangle$, we find the maximal value of S_{CHSH} at time t_k. It is a straightforward calculation to show that this quantity is an upper bound on the sum of the CHSH value (5) of $|\psi_\eta\rangle$, as obtained when applying a noisy measurement in our protocol at time t_k, and the average maximal CHSH value, obtained from applying the Horodecki criterion to the four possible postmeasurement states at t_{k+1} weighted by the respective probability of obtaining each state. This argument can be repeated throughout the protocol and consequently $\bar{\mathcal{C}} < \sqrt{2} - 1$, which is the communication cost of simulating a maximal violation of the CHSH inequality achieved with $|\psi_0\rangle$.

B. Cost 2: Unbounded local memory

At each time step in the protocol, Alice chooses with uniform probability between two measurement options. After each measurement of Alice, the number of possible reduced states of Bob's qubit quadruples. Any classical simulation must account for this exponentially increasing number of possible states. Since each of Alice's measurement choices is random, any classical simulation requires having at least the same number of local realist states as the number of pure quantum states achieved during the experiment. The proof is as follows.

A stochastic process is a one-dimensional chain of discrete random variables that attains values in a finite or countably infinite alphabet. An input-output process [13] is a collection of stochastic processes in which each such process corresponds to all possible output sequences given a particular infinite input sequence. The experiment is an example of an input-output process. It has input alphabet $\{x_k, \bar{x}_k\}$ and output alphabet $\{0, 1\}$. As shown in Ref. [13], for any input-output process there is a unique finite-state machine, i.e., an abstract machine that can be in exactly one of a finite number of states at any given time, with the following property: It has minimal entropy over the state probability distribution and maximal mutual information with the future output of the process given the past choices of inputs and past observed outputs, and the future input of the process. This machine is called the ε transducer [13] of the input-output process. It consists of the input and output alphabets, a set of causal states, and the set of conditional transition probabilities between the causal states. Each causal state is associated with the set of input-output pasts producing the same probabilities for all possible input-output futures. Thus, the causal states constitute equivalence classes for the set of input-output pasts. A causal state stores all the information about the past needed to predict the future output but as little as possible of the remaining information overhead contained in the past. The Shannon entropy over the stationary distribution of the causal states represents the minimum internal entropy needed to be stored to optimally compute future outputs. It depends on how Alice's measurements are chosen; here we have assumed that they are selected from a uniform probability distribution with entropy one bit at each time step.

The number of causal states of the ε transducer corresponding to our experiment is infinite. This implies that the classical system that simulates the experiment has to store new information in its memory. This leads to two possibilities: Either the memory is infinite and additional information can always be stored without needing to erase previous information or the memory is finite and the system has to erase a part of it to allocate new information. However, due to Landauer's principle [14], the erasure of information has a thermodynamical cost. Landauer's principle states that the erasure of information in an information-carrying degree of freedom is accompanied by an associated increase of entropy in some non-information-carrying degree of freedom. There is strong evidence supporting the validity of Landauer's principle in both the classical and quantum domains [15–21]. Since

we are assuming a local realistic model supplemented by communication, the memory should be allocated in the local systems. Since Bob's quantum state and its classical counterpart (represented by a causal state of the ε transducer) are changing after each of Alice's measurements, this implies that there should be some information erasure in the local memory associated with Bob's system. Therefore, after sufficiently many measurements of Alice, Bob's system begins to emit heat. Such heating at a distance is a form of signaling.

IV. CONCLUSION

We have introduced a protocol in which sequential nonprojective measurements are performed on one of two entangled systems while no measurements are performed on the other distant system. Regardless of which local measurements are chosen and which outcomes are obtained, both entanglement and the possibility of violating a Bell inequality *never* vanish. We showed that, to simulate the predictions of quantum theory for the local state of the distant system, it is not sufficient to supplement finite-memory classical models with unlimited rounds of communication. In addition, the distant system must have infinite memory. Whenever the distant system fails to have infinite memory, a thermodynamical argument implies that it will be heated at a distance after sufficiently many local measurements on its companion.

Our protocol shows that (i) there is a way for experimentally ruling out nonlocal finite-memory classical models without measuring the system that will, hypothetically, emit heat and (ii) there are problems whose solution would require classical systems with infinite memory and communication but which can be solved combining sequential quantum measurements and entanglement.

ACKNOWLEDGMENTS

We thank Nicolas Brunner, Nicolas Gisin, Mile Gu, and Matthias Kleinmann for their comments on the manuscript, Gustavo Cañas for his help with Fig. 1, Jim Crutchfield for discussions, and Matthias Kleinmann for checking the calculations. This work was supported by the project "Photonic Quantum Information" (Knut and Alice Wallenberg Foundation, Sweden), Project No. FIS2014-60843-P, "Advanced Quantum Information" (MINECO, Spain), with FEDER funds, and the FQXi Large Grant "The Observer Observed: A Bayesian Route to the Reconstruction of Quantum Theory." A.T. acknowledges financial support from the Swiss National Science Foundation (Starting Grant DIAQ).

[1] A. Cabello, M. Gu, O. Gühne, J.-Å. Larsson, and K. Wiesner, Thermodynamical cost of some interpretations of quantum theory, Phys. Rev. A **94**, 052127 (2016).

[2] R. Silva, N. Gisin, Y. Guryanova, and S. Popescu, Multiple Observers Can Share the Nonlocality of Half of an Entangled Pair by Using Optimal Weak Measurements, Phys. Rev. Lett. **114**, 250401 (2015).

[3] F. J. Curchod, M. Johansson, R. Augusiak, M. J. Hoban, P. Wittek, and A. Acín, Unbounded randomness certification using sequences of measurements, Phys. Rev. A **95**, 020102(R) (2017).

[4] G. Brassard, R. Cleve, and A. Tapp, Cost of Exactly Simulating Quantum Entanglement with Classical Communication, Phys. Rev. Lett. **83**, 1874 (1999).

[5] S. Pironio, Violations of Bell inequalities as lower bounds on the communication cost of nonlocal correlations, Phys. Rev. A **68**, 062102 (2003).

[6] K. Kraus, *States, Effects and Operations: Fundamental Notions of Quantum Theory* (Springer, Berlin, 1983).

[7] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, Phys. Rev. Lett. **23**, 880 (1969).

[8] M. M. Wolf, D. Pérez-García, and C. Fernández, Measurements Incompatible in Quantum Theory Cannot Be Measured Jointly in Any Other No-Signaling Theory, Phys. Rev. Lett. **103**, 230402 (2009).

[9] K. Życzkowski, P. Horodecki, A. Sanpera, and M. Lewenstein, Volume of the set of separable states, Phys. Rev. A **58**, 883 (1998).

[10] Y. Ota, S. Ashhab, and F. Nori, Entanglement amplification via local weak measurements, J. Phys. A: Math. Theor. **45**, 415303 (2012).

[11] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Concentrating partial entanglement by local operations, Phys. Rev. A **53**, 2046 (1996).

[12] R. Horodecki, P. Horodecki, and M. Horodecki, Violating Bell inequality by mixed spin-1/2 states: Necessary and sufficient condition, Phys. Lett. A **200**, 340 (1995).

[13] N. Barnett and J. P. Crutchfield, Computational mechanics of input-output processes: Structured transformations and the ϵ-transducer, J. Stat. Phys. **161**, 404 (2015).

[14] R. Landauer, Irreversibility and heat generation in the computing process, IBM J. Res. Dev. **5**, 183 (1961).

[15] B. Piechocinska, Information erasure, Phys. Rev. A **61**, 062314 (2000).

[16] S. Hilt, S. Shabbir, J. Anders, and E. Lutz, Landauer's principle in the quantum regime, Phys. Rev. E **83**, 030102(R) (2011).

[17] A. Bérut, A. Arakelyan, A. Petrosyan, S. Ciliberto, R. Dillenschneider, and E. Lutz, Experimental verification of Landauer's principle linking information and thermodynamics, Nature (London) **483**, 187 (2012).

[18] Y. Jun, M. Gavrilov, and J. Bechhoefer, High-Precision Test of Landauer's Principle in a Feedback Trap, Phys. Rev. Lett. **113**, 190601 (2014).

[19] D. Reeb and M. M. Wolf, An improved Landauer principle with finite-size corrections, New J. Phys. **16**, 103011 (2014).

[20] S. Lorenzo, R. McCloskey, F. Ciccarello, M. Paternostro, and G. M. Palma, Landauer's Principle in Multipartite Open Quantum System Dynamics, Phys. Rev. Lett. **115**, 120403 (2015).

[21] J. P. S. Peterson, R. S. Sarthour, A. M. Souza, I. S. Oliveira, J. Goold, K. Modi, D. O. Soares-Pinto, and L. C. Céleri, Experimental demonstration of information to energy conversion in a quantum system at the Landauer limit, Proc. R. Soc. A **472**, 20150813 (2016).

Heralded generation of maximal entanglement in any dimension via incoherent coupling to thermal baths

Armin Tavakoli[1], Géraldine Haack[1], Marcus Huber[2], Nicolas Brunner[1], and Jonatan Bohr Brask[1]

[1]Department of Applied Physics, University of Geneva, 1211 Geneva, Switzerland

[2]Institute for Quantum Optics and Quantum Information (IQOQI), Austrian Academy of Sciences, Boltzmanngasse 3, A-1090 Vienna, Austria

June 12, 2018

We present a scheme for dissipatively generating maximal entanglement in a heralded manner. Our setup requires incoherent interactions with two thermal baths at different temperatures, but no source of work or control. A pair of $(d+1)$-dimensional quantum systems is first driven to an entangled steady state by the temperature gradient, and maximal entanglement in dimension d can then be heralded via local filters. We discuss experimental prospects considering an implementation in superconducting systems.

1 Introduction

Entanglement is a key phenomenon distinguishing quantum from classical physics, and is the paradigmatic resource enabling many applications of quantum information science. Generating and maintaining entanglement is therefore a central challenge. Decoherence caused by unavoidable interactions of a system with its environment generally degrades entanglement, and significant effort is invested in minimising the effect of such dissipation in experiments.

However, dissipation can also be advantageous, and may indeed be exploited for the generation of entangled quantum states under the right conditions [1–8]. In particular, it is possible for dissipative processes to drive the system into an entangled steady state [9–13]. This was studied in a variety of physical systems [14–20] and demonstrated experimentally for atomic ensembles [21], trapped ions [22, 23], and superconducting qubits [24]. The main ingredients are engineered decay processes and quantum bath engineering [25–27], and coherent external driving is employed, which, from a thermodynamic point of view, can be considered a source of work.

More generally, it is natural to look for the minimal setting in which dissipative entanglement generation is possible. In particular, one may ask if entanglement can be generated from purely thermal processes alone, without the need for work input or external control. This can in principle be achieved in equilibrium situations, as any entangled state can be obtained as the ground state of a specific Hamiltonian. However, this requires highly nonlocal Hamiltonians which may be extremely difficult to implement in practice.

On the other hand, it was shown that steady-state entanglement can be obtained in systems out of thermal equilibrium. This was first discussed for an atom coupled to two cavities driven by incoherent light [28], and later for many-body systems [29, 30], interacting spins [31, 32], atoms in a thermal environment [33, 34], and mechanical oscillators [35]. In this context, Ref. [36] discussed what is arguably the simplest setting, namely a two-qubit system, where one qubit is connected to a hot bath and the other to a cold bath. This setup is promising for implementations in superconducting systems and quantum dots. Overall, the out-of-equilibrium approach thus opens interesting perspectives for dissipative entanglement generation. However, its main drawback so far is the fact that the generated entanglement is typically very weak, and thus not directly useful for applications.

Here we offer a solution to this problem, presenting a scheme in which maximal entanglement can be generated in a heralded manner, through incoherent interactions with thermal baths alone. Specifically, a pair of $(d+1)$-dimensional systems is first driven to an entangled steady state, from which maximal entanglement in dimension d can

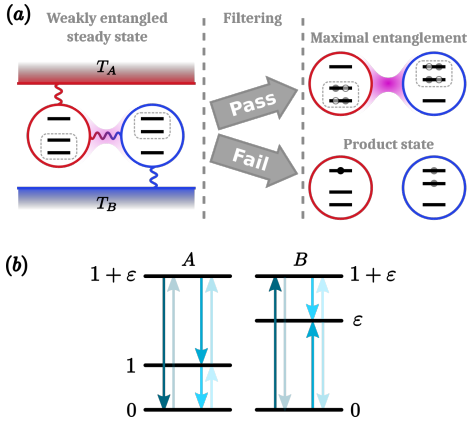

Figure 1: **(a)** Qutrit thermal machine. Two qutrits are coupled to each other and to hot and cold thermal baths. The interaction with the baths drives the qutrits into a steady state featuring weak entanglement. Local filters project onto qubit subspaces on each side (dashed boxes). Upon success, the system is projected into a strongly entangled two-qubit state. Failure leaves the qutrits in a separable state, and the process must be restarted. **(b)** Level structure for the two qutrits. Arrows indicate the transitions involved in the interaction Hamiltonian.

then be heralded via local filters. The procedure is implemented by a simple quantum thermal machine, operating out of equilibrium between two heat baths at different temperatures. Moreover, for $d = 2, 3$ we prove that any pure entangled state can be obtained without additional filtering, indicating that this holds for any d. Finally, we discuss experimental prospects considering an implementation in superconducting systems.

2 Two-qutrit thermal machine

The setup we consider is illustrated in Fig. 1(a). Two three-level systems (i.e. qutrits) interact with each other, and independently with two thermal baths at different temperatures T_A and T_B (in the following, $T_A > T_B$ will be the relevant setting for entanglement generation). This out-of-equilibrium situation drives the two-qutrit system into a steady state, which is weakly entangled. A local filter is then applied to each qutrit, projecting the system onto a two-qubit subspace (as indicated by the dashed boxes). If the filter succeeds, the final state is arbitrarily close to a target two-qubit state. This target state can be any pure, entangled state, and in particular may be maximally entangled. If the filter fails, the system is left in a product state with no entanglement, and the process is restarted.

Each qutrit is described by a Hamiltonian H_A, H_B, and their interaction by H_{int}. We take the energy level structure illustrated in Fig. 1(b)

$$H_A = (|1\rangle_A\langle 1| + (1+\varepsilon)|2\rangle_A\langle 2|) \otimes \mathbb{1}_B, \quad (1)$$
$$H_B = \mathbb{1}_A \otimes (\varepsilon|1\rangle_B\langle 1| + (1+\varepsilon)|2\rangle_B\langle 2|), \quad (2)$$

where, without loss of generality, we set the ground state energies to zero and the first gap of qutrit A to 1 (throughout the paper, we work in units where $\hbar = k_B = 1$). We are interested in autonomous processes, which require no external work input. This means that H_{int} must be time independent and preserve the total energy, i.e. $[H_{int}, H_A + H_B] = 0$. There are three possible energy-preserving transitions. Hence, writing $|ij\rangle = |i\rangle_A|j\rangle_B$, the most general form of the interaction is

$$H_{int} = g_1|02\rangle\langle 20| + g_2|11\rangle\langle 20| + g_3|11\rangle\langle 02| + h.c., \quad (3)$$

where g_1, g_2, and g_3 denote the interaction strengths.

To enable a fully analytical treatment, we first describe the evolution of the system in contact with the thermal baths by a simple reset model [37]. When considering potential implementations below, we confirm that our results hold also under a Lindblad-type description of the open system. The reset model leads to the following master equation [1]

$$\frac{\partial\rho}{\partial t} = i[\rho, H] + p_A\left(\tau_A \otimes \mathrm{Tr}_A\,\rho - \rho\right) + p_B\left(\mathrm{Tr}_B\,\rho \otimes \tau_B - \rho\right), \quad (4)$$

where $H = H_A + H_B + H_{int}$ is the total Hamiltonian, p_A, p_B are coupling constants, and τ_A, τ_B are thermal states, that is, $\tau_i = \exp(-H_i/T_i)/\,\mathrm{Tr}[\exp(-H_i/T_i)]$ for $i = A, B$.

[1] Note that we are using a local master equation, where the dissipation induced by each bath acts locally on the subsystem connected to that bath. Recent works, analysing thermal machines similar to those employed here, have shown such a local approach to provide very good agreement with the exact dynamics for weak coupling, which is the regime of interest here [38, 39].

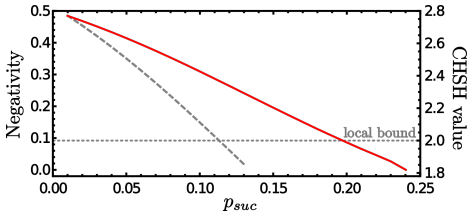

Figure 2: Optimal negativity (solid, left axis) and CHSH value (dashed, right axis) vs. postselection success probability. The dotted line shows the local bound above which the CHSH Bell inequality is violated.

One can interpret (4) as describing a process where, at each instance of time, each qutrit is either left unchanged or reset to a thermal state at the temperature of the bath, with resets happening at rates p_A, p_B. To ensure validity of our master equation, we always work in the perturbative regime where $g_1, g_2, g_3, p_A, p_B \ll 1, \varepsilon$.

To understand how the machine can generate entanglement, first note that the two-qubit subspace selected by the filters is spanned by the states $\{|01\rangle, |12\rangle, |11\rangle, |12\rangle\}$. Clearly, the transition g_3 generates coherence if the system is already in this subspace, between $|11\rangle$, and $|02\rangle$ thus creating entanglement. Interaction with the cold bath will tend to drive the cold qutrit towards the ground state, taking the system out of the filtered subspace. Transition g_3 cannot bring the system back, but transitions g_1 and g_2 do. In addition, the combination of these two transitions also generates entanglement because

$$[H_{g1}, H_{g2}] = g_1 g_2 \left(|02\rangle\langle 11| - |11\rangle\langle 02|\right) , \qquad (5)$$

where $H_{g1} = g_1|02\rangle\langle 20| + h.c.$ etc. If the cold bath temperature is low, the cold qutrit will tend to be in the ground state, and the system will only get excited into the filtered subspace whenever the joint state is $|20\rangle$. The interaction will then generate a pure, entangled state. Resets induced by the cold bath drive the system out of the filtered subspace and hence do not degrade the purity of the filtered state. Resets induced by the hot bath, on the other hand, do destroy coherence there, reducing the purity. Nevertheless, some hot resets are necessary to populate the state $|20\rangle$. We thus expect the best entanglement to be generated when T_A is large, T_B is close to zero, and $p_A \ll p_B$.

We have derived the steady-state solution $\bar{\rho}$ of (4) in the limit of a maximal temperature gradient, $T_A \to \infty$, $T_B \to 0$ (see App. A). To obtain the final state, a local filter is applied to each qutrit, defined by projectors $\Pi_A = |0\rangle_A\langle 0| + |1\rangle_A\langle 1|$ and $\Pi_B = |1\rangle_B\langle 1| + |2\rangle_B\langle 2|$. The normalised, postselected state is

$$\rho' = \frac{1}{p_{suc}} (\Pi_A \otimes \Pi_B)\bar{\rho}(\Pi_A \otimes \Pi_B), \qquad (6)$$

where $p_{suc} = \mathrm{Tr}[(\Pi_A \otimes \Pi_B)\bar{\rho}]$ is the probability for the filtering to succeed. We take all the interaction strengths equal, $g_1 = g_2 = g_3 = g$. In this case, the state after filtering becomes

$$\rho' = \begin{pmatrix} \frac{p_A}{4p_A+6p_B} & 0 & 0 & 0 \\ 0 & \frac{p_A+3p_B}{4p_A+6p_B} & \frac{3p_B}{4p_A+6p_B} & 0 \\ 0 & \frac{3p_B}{4p_A+6p_B} & \frac{p_A+3p_B}{4p_A+6p_B} & 0 \\ 0 & 0 & 0 & \frac{p_A}{4p_A+6p_B} \end{pmatrix} . \qquad (7)$$

As expected, the highest purity of ρ' is obtained when the ratio $\mu = p_A/p_B$ is small. For $\mu \to 0$, the state ρ' tends to a pure, maximally entangled state (relabelling the basis states of the qubit subspace to $|0\rangle$, $|1\rangle$)

$$|\psi_+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle). \qquad (8)$$

Thus our machine can generate entanglement arbitrarily close to maximal. In addition, it is interesting to note that different choices for the interaction strengths enable the generation of other entangled states. Specifically, as shown in App. A, taking $g_1 = g\cos(\theta)$, $g_2 = g\sin(\theta)$, $g_3 = 0$ generates any partially entangled state of the form $|\psi_\theta\rangle = \sin(\theta)|01\rangle + \cos(\theta)|10\rangle$. We note that (7) holds for any value of g. Hence the limit $\mu \to 0$ can be taken while keeping the ratio of g/p_A fixed, retaining the validity of the local master equation.

There is a trade-off between the probability for successful filtering and the quality of ρ'. The success probability tends to zero for both small μ (for fixed g) and small g (for fixed μ). In the two cases, respectively

$$\begin{aligned} p_{suc} &\approx \frac{1}{3}\frac{p_A}{p_B}, \quad \text{and} \\ p_{suc} &\approx \frac{2(2p_A + 3p_B)}{9p_B(p_A + p_B)^2} g^2. \end{aligned} \qquad (9)$$

Adjusting the coupling parameters to increase p_{suc} results in a final state ρ' with a smaller

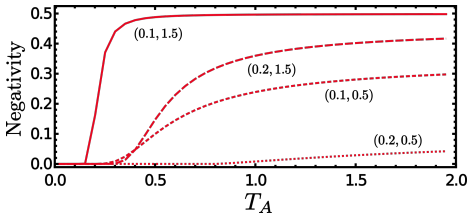

Figure 3: Entanglement generation for finite temperatures. The numbers given for each curve are (T_B, ϵ) (in units of the first energy gap of qutrit A equal to 1). To make optimisation over the coupling parameters tractable, we maximise the off-diagonal element of the output state rather than the negativity directly. The curves therefore represent lower bounds.

overlap with the target pure state (8). Nevertheless, states of high quality can be generated. In Fig. 2 we show the maximal negativity [40] as well as the value of the Clause-Horne-Shimony-Holt (CHSH) quantity [41] for varying p_{suc} (we optimise over g, p_A, and p_B, while imposing the perturbative regime). The negativity is an entanglement monotone ranging from 0 (separable) to 1/2 (maximally entangled) for qubits. Twice the negativity is a lower bound on the concurrence (which ranges from 0 to 1) [42]. We see that ρ' remains entangled up to $p_{suc} \approx 0.25$ and nonlocal up to $p_{suc} \approx 0.12$.

The machine thus provides a heralded source of entangled states: running the machine continuously, the system remains in the steady state until the entangled state is needed, at which point the filtering is performed. If filtering fails, the machine is allowed to return to the steady state, and another attempt can be made. A quasi-deterministic source can be constructed by running several machines in parallel. With n machines, the probability for obtaining a successful projection in at least one of them scales as $1-(1-p_{suc})^n$. Failure is exponentially suppressed in n.

In addition to the trade-off between success probability and quality of the postselected state, controlled by the coupling parameters, the temperatures also influence the generated entanglement. So far, we have taken a maximal temperature gradient, $T_A \to \infty$, $T_B \to 0$. In Fig. 3 we plot attainable negativity for finite temperatures. We see that, as might be expected, it is always better to take the hot bath temperature as large as possible, maximising the temperature gradient. As the cold bath temperature increases or the gap size ε decreases, the hot bath temperature required to generate entanglement increases, and the maximal amount of attainable entanglement decreases. So, to maximise the entanglement, it is desirable to make T_B small and ϵ large (note though, that p_{suc} decreases with increasing ε).

3 Two-qudit thermal machine

The scheme considered above can be generalised to create entangled states of two d-level systems, using a $(d+1)$-level thermal machine. The setup is the same as in Fig. 1(a), with the qutrits replaced by $(d+1)$-level systems, with level structures as illustrated in Fig. 4. Denoting the energy gaps by ε_k (with $\varepsilon_1 = 1$), and setting $E_k^A = \sum_{l=1}^k \varepsilon_l$ and $E_k^B = \sum_{l=1}^k \varepsilon_{d-l+1}$, the free Hamiltonians are

$$\begin{aligned} H_A &= \sum_{k=1}^{d} E_k^A |k\rangle_A\langle k| \otimes \mathbb{1}, \\ H_B &= \sum_{k=1}^{d} \mathbb{1} \otimes E_k^B |k\rangle_B\langle k|, \end{aligned} \tag{10}$$

and the interaction Hamiltonian is

$$H_{int} = \sum_{k=1}^{d} g_k |d, 0\rangle\langle k-1, d-k+1| + h.c., \tag{11}$$

corresponding to the transitions indicated on Fig. 4. The evolution is again described by the master equation (4).

We will focus on the generation of a maximally entangled qudit state

$$|S_d\rangle = \frac{1}{\sqrt{d}} \sum_{k=1}^{d} |k-1, d-k\rangle. \tag{12}$$

In that case, it suffices to set all the interaction strengths equal, $g_k = g/\sqrt{2}$ (the $\sqrt{2}$ ensures consistency with the qutrit case). In the limit $T_A \to \infty$, $T_B \to 0$, the steady state solution of (4) can then be derived analytically for any value of d. It is given in App. B. In analogy with the qutrit case, we consider local projections onto d-dimensional subsystems on each side, given by $\Pi_A = \mathbb{1} - |d\rangle_A\langle d|$ and $\Pi_B = \mathbb{1} - |0\rangle_B\langle 0|$, and the state after successful filtering is again computed as in (6). We find that, as before, high purity

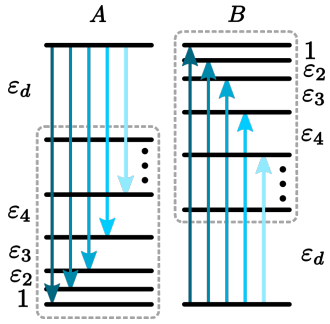

Figure 4: Level structure of the two $(d+1)$-level systems in the qudit thermal machine. Arrows indicate the transitions involved in the interaction Hamiltonian. The dashed boxes indicate the d-dimensional subspaces to which the steady state is filtered to obtain the final state.

is attained when $p_A \ll p_B$, and the state tends to $|S_d\rangle$, as desired. Thus, our scheme is able to generate maximally entangled states in any dimension. The success probability is given by

$$p_{suc} = \frac{(d-1)g^2 p_A((d-1)p_A + dp_B)}{d^2\left(g^2\xi + p_A p_B(p_A+p_B)^2\right)}, \qquad (13)$$

where $\xi = (2(d-1)p_A p_B + (d-1)p_B^2 + p_A^2)$. One can check that this agrees with (9) for $d=3$. Note that p_{suc} scales like $1/d$ for large d, unless $g \sim 1/\sqrt{d}$.

From $|S_d\rangle$, any pure two-qudit state can be obtained via biased filtering and local operations [43]. However, given that any pure, entangled state of two qubits can be generated directly using the qutrit machine by adjusting the coupling strengths, it is natural to ask whether the same holds for qudits. In App. C, we prove this for $d=3$, suggesting that it generalises to arbitrary d. Note that such direct generation can be advantageous in terms of success probability.

4 Implementation

A variety of physical platforms might be considered for implementation of our scheme, including trapped atoms, ions, or solid-state artificial atoms. A promising platform is superconducting, circuit QED systems, which are generally good candidates for realizing quantum thermal machines [36, 44–46]. Here, we discuss prospects for a circuit QED implementation of the qutrit machine in more detail, and provide numerical evidence that strong entanglement generation can be achieved with parameter settings corresponding to state-of-the-art experimental capabilities, see Fig. 5.

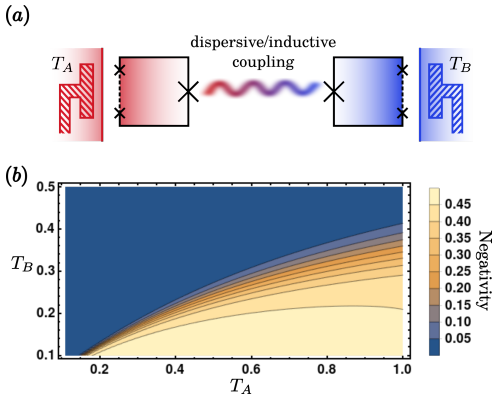

Figure 5: **(a)** Implementation of the qutrit machine in circuit QED. Each fluxonium qutrit, depicted by their quantum circuit made of Josephson junctions [47], is capacitively coupled to a transmission line that plays the role of a thermal reservoir, see main text and Ref. [36]. The flip-flop type interaction Hamiltonian between the two qutrits can be implemented either in the dispersive regime or by direct inductive coupling. **(b)** Negativity computed from the Lindblad model for $(\Gamma_A, \Gamma_B, \gamma, g, \epsilon) = (10^{-4}, 5\times10^{-3}, 3.5\times10^{-5}, 1.6\times10^{-3}, 3)$ and $\Gamma_{B,12} = \Gamma_B/50$, see App. D for the full Lindblad equation. All temperatures and energies are given in units of the first energy gap of qutrit A, taken to be 1 GHz. Near-maximal entanglement is generated in the bright region for experimentally relevant parameter values.

Considering that the interaction (3) requires the transition $|0\rangle \leftrightarrow |2\rangle$, fluxonium qutrits are good candidates for realizing the machine. In contrast to transmon qubits, for which selection rules forbid this transition, tuning of the magnetic quantum flux away of the sweet spot breaks quantum parity without inducing additional decoherence [48, 49]. Consequently, simple selection rules are absent and the transition $|0\rangle \leftrightarrow |2\rangle$ is allowed. Fluxomium artificial atoms have also recently shown outstanding performances in the context of quantum information processing thanks to their high tunability. In particular, their transition frequencies are in the range of hundreds of MHz to 30 GHz and the couplings

to the baths can also be tuned from several kHZ to a few MHz [50–52]. In [52], it was even shown that complete decoupling from the environment is achievable.

With respect to the implementation of the interaction Hamiltonian, several coupling mechanisms are already available with fluxonium systems. First, similarly to transmon qubits [53–55], fluxonium qutrits can be coupled capacitively or inductively via a cavity bus in the dispersive regime characterized by a strong frequency detuning of the qutrits and cavity with respect to their respective coupling strength to the cavity [50, 56]. Second, a possibly advantageous alternative is provided by a direct mutual inductive coupling as described in [57] and proposed for fluxonium qutrits in [50]. Technicalities will depend on the actual frequencies that can be achieved experimentally.

Regarding the description of coupling mechanisms of each qutrit to a thermal bath, as well as the nature of the thermal baths in this setup, we refer to [36]. It is also worth mentioning that fluxonium qutrits allow for flux-resolved spectroscopy, a technique to precisely determine all system frequencies [52].

Finally, the filtering procedure requires binary projective measurements onto a single energy level for each qutrit. That is, measurements which reveal whether or not the qutrit is in the corresponding state, but do not distinguish the remaining two states. This can be achieved by dispersive read-out in the regime where the dispersive shift is larger than the readout cavity line width (the photon-resolved regime) [58]. The shifts corresponding to each qutrit state will then be well separated and the transmittivity of the cavity at a frequency corresponding to, say, state $|0\rangle$ will be significant only when the qutrit is in this state, allowing for a binary projective measurement. A recent experiment operating in this regime was reported in [59]. Alternatively, two of the three shifts can be tuned to be identical. A binary projective measurement on qutrits using this technique was demonstrated in Ref. [60].

To model a circuit-QED implementation of the two-qutrit thermal machine and determine how much entanglement can be generated for reasonable parameter values, we use a master equation on standard Lindblad form. It describes dissipation due to coupling to bosonic baths, as well as pure dephasing, which is usually present in experiments. We note that it is possible to exactly map the reset model of Sec. 2 to a Lindblad master equation of the form described here. This is discussed in App. D. The equation (which replaces (4)) can be written

$$\frac{\partial\rho}{\partial t} = i[\rho, H] + \mathcal{L}_A(\rho) + \mathcal{L}_A^z(\rho) + \mathcal{L}_B(\rho) + \mathcal{L}_B^z(\rho). \tag{14}$$

Here, the dissipators $\mathcal{L}_A$ and $\mathcal{L}_B$ describe the effect of the thermal baths while $\mathcal{L}_A^z(\rho)$ and $\mathcal{L}_B^z(\rho)$ describe pure dephasing. We define

$$\mathcal{D}[O]\rho = O\rho O^\dagger - \frac{1}{2}\{O^\dagger O, \rho\} \tag{15}$$

to denote a standard Lindblad-type dissipator. Then

$$\mathcal{L}_A(\rho) = \sum_{l=\pm}\sum_{k\in\{01,12,02\}} \Gamma_{A,k}^l \mathcal{D}[\sigma_k^l \otimes \mathbb{1}]\rho\,, \tag{16}$$

$$\mathcal{L}_B(\rho) = \sum_{l=\pm}\sum_{k\in\{01,12,02\}} \Gamma_{B,k}^l \mathcal{D}[\mathbb{1} \otimes \sigma_k^l]\rho\,, \tag{17}$$

and

$$\mathcal{L}_A^z(\rho) = \sum_{k\in\{01,12,02\}} \gamma_{A,k}\, \mathcal{D}[\sigma_k^z \otimes \mathbb{1}]\rho\,, \tag{18}$$

$$\mathcal{L}_B^z(\rho) = \sum_{k\in\{01,12,02\}} \gamma_{B,k}\, \mathcal{D}[\mathbb{1} \otimes \sigma_k^z]\rho\,. \tag{19}$$

Here, $\sigma_{mn}^{\pm}$ describe jumps between states $|m\rangle$ and $|n\rangle$ while σ_{mn}^z describe phase flips between these states. Specifically,

$$\sigma_{mn}^+ = |n\rangle\langle m|\,, \qquad \sigma_{mn}^- = |m\rangle\langle n|\,, \tag{20}$$

and

$$\sigma_{mn}^z = |m\rangle\langle m| - |n\rangle\langle n|. \tag{21}$$

The jump rates follow bosonic statistics ($j = A, B$)

$$\Gamma_{j,mn}^+ = \Gamma_{j,mn}\, n_B(\Delta E_{mn}, T_j)\,, \tag{22}$$

$$\Gamma_{j,mn}^- = \Gamma_{j,mn}\,[1 + n_B(\Delta E_{mn}, T_j)]\,. \tag{23}$$

In principle, the bath coupling constants $\Gamma_{A,k}$, $\Gamma_{B,k}$, and the pure dephasing rates $\gamma_{A,k}$, $\gamma_{B,k}$ could be different for each possible transition. For simplicity, here we take $\gamma_{A,k} = \gamma_{B,k} = \gamma$ to be the same for all transitions for both qutrits, and we take the bath couplings to be the same for all transitions $\Gamma_{A,k} = \Gamma_A$, $\Gamma_{B,k} = \Gamma_B$ with one exception. Jumps beteween states $|1\rangle_B$ and

$|2\rangle_B$ of the cold qubit degrade coherence within the filtered subspace. Good entanglement generation therefore requires that $\Gamma_{B,12} < \Gamma_B$. This can be achieved by coupling through a bandpass filter centered away from the relevant transition frequency, reducing environmental damping for such transitions more strongly relative to jumps between the ground and excited states. The use of bandpass filtering to suppress environmental damping has been experimentally demonstrated [61, 62].

We numerically solve (14) in the steady state and compute the amount of entanglement generate by our scheme. Values for the different parameters (interaction strength g, qutrit energies ϵ, bath coupling rates Γ_A, Γ_B, and pure dephasing rate γ) are taken from recent experimental achievements in circuit-QED architectures using fluxonium qutrits [50, 59, 63]. The result is shown in Fig. 5(b). We see that near-maximal entanglement can be obtained. Thus, the scheme is a promising approach to demonstrating heralded entanglement using incoherent couplings to thermal baths. It is interesting to note in Fig. 5(b) that for fixed couplings, it is not optimal to maximise the temperature gradient. Maximal entanglement is obtained at a finite gradient.

5 Conclusion

We have demonstrated that combining incoherent couplings to thermal baths out of equilibrium with local filtering enables heralded generation of maximally entangled states in any dimension. The generated states can be made arbitrarily pure, at the price of lowering the filtering success probability. We have discussed an implementation of our scheme for qubit entanglement in superconducting systems, and found that prospects for a proof-of-principle experiment are good, with significant amounts of entanglement generated in the presence of decoherence and with limited temperature gradients. Interesting future perspectives include thermal generation of multipartite entanglement, and states useful for quantum computation or metrology.

We acknowledge helpful discussions with N. Cottet and B. Huard on implementations in superconducting systems. We acknowledge the Swiss National Science Foundation (Starting grant DIAQ, grant 200021_169002, and QSIT). GH acknowledges support from the Swiss National Science Foundation through the Marie-Heim Vögtlin grant no. 164466. MH acknowledges funding from the Swiss National Science Foundation (AMBIZIONE $PZ00P2\_161351$) and the Austrian Science Fund (FWF) through the START project Y879-N27.

References

[1] M. B. Plenio, S. F. Huelga, A. Beige, and P. L. Knight, "Cavity-loss-induced generation of entangled atoms," Phys. Rev. A **59**, 2468–2475 (1999).

[2] M. S. Kim, Jinhyoung Lee, D. Ahn, and P. L. Knight, "Entanglement induced by a single-mode heat environment," Phys. Rev. A **65**, 040101 (2002).

[3] L. Jakóbczyk, "Entangling two qubits by dissipation," J. Phys. A: Math. Gen. , 6383 (2002).

[4] D. Braun, "Creation of entanglement by interaction with a common heat bath," Phys. Rev. Lett. **89**, 277901 (2002).

[5] F. Benatti, R. Floreanini, and M. Piani, "Environment induced entanglement in markovian dissipative dynamics," Phys. Rev. Lett. **91**, 070402 (2003).

[6] D. Burgarth and V. Giovannetti, "Mediated homogenization," Phys. Rev. A **76**, 062307 (2007).

[7] B. Bellomo, R. Lo Franco, S. Maniscalco, and G. Compagno, "Entanglement trapping in structured environments," Phys. Rev. A **78**, 060302 (2008).

[8] D. Manzano, M. Tiersch, A. Asadian, and H. J. Briegel, "Quantum transport efficiency and fourier's law," Phys. Rev. E **86**, 061118 (2012).

[9] S. Diehl, A. Micheli, A. Kantian, B. Kraus, H. P. Buchler, and P. Zoller, "Quantum states and phases in driven open quantum systems with cold atoms," Nat Phys **4**, 878–883 (2008).

[10] F. Verstraete, M. M. Wolf, and I. J. Cirac, "Quantum computation and quantum-state engineering driven by dissipation," Nat Phys **5**, 633–636 (2009).

[11] B. Kraus, H. P. Büchler, S. Diehl, A. Kantian, A. Micheli, and P. Zoller, "Preparation

of entangled states by quantum markov processes," Phys. Rev. A **78**, 042307 (2008).

[12] F. Ticozzi and L. Viola, "Steady-state entanglement by engineered quasi-local markovian dissipation," Quant. Inf. and Comp. **14**, 0265 (2014).

[13] F. Tacchino, A. Auffèves, M. F. Santos, and D. Gerace, "Steady state entanglement beyond thermal limits," Phys. Rev. Lett. **120**, 063604 (2018).

[14] S. Schneider and G. J. Milburn, "Entanglement in the steady state of a collective-angular-momentum (dicke) model," Phys. Rev. A **65**, 042107 (2002).

[15] M. J. Kastoryano, F. Reiter, and A. S. Sørensen, "Dissipative preparation of entanglement in optical cavities," Phys. Rev. Lett. **106**, 090502 (2011).

[16] X. Wang and S. G. Schirmer, "Generating maximal entanglement between non-interacting atoms by collective decay and symmetry breaking," arXiv e-print , 1005.2114 (2010).

[17] F. Reiter, L. Tornberg, G. Johansson, and A. S. Sørensen, "Steady-state entanglement of two superconducting qubits engineered by dissipation," Phys. Rev. A **88**, 032317 (2013).

[18] M. J. A. Schuetz, E. M. Kessler, L. M. K. Vandersypen, J. I. Cirac, and G. Giedke, "Steady-state entanglement in the nuclear spin dynamics of a double quantum dot," Phys. Rev. Lett. **111**, 246802 (2013).

[19] J. Cai, S. Popescu, and H. J. Briegel, "Dynamic entanglement in oscillating molecules and potential biological implications," Phys. Rev. E **82**, 021921 (2010).

[20] S. Walter, J. C. Budich, J. Eisert, and B. Trauzettel, "Entanglement of nanoelectromechanical oscillators by cooper-pair tunneling," Phys. Rev. B **88**, 035441 (2013).

[21] H. Krauter, C. A. Muschik, K. Jensen, W. Wasilewski, J. M. Petersen, J. I. Cirac, and E. S. Polzik, "Entanglement generated by dissipation and steady state entanglement of two macroscopic objects," Phys. Rev. Lett. **107**, 080503 (2011).

[22] J. T. Barreiro, M. Muller, P. Schindler, D. Nigg, T. Monz, M. Chwalla, M. Hennrich, C. F. Roos, P. Zoller, and R. Blatt, "An open-system quantum simulator with trapped ions," Nature **470**, 486–491 (2011).

[23] Y. Lin, J. P. Gaebler, F. Reiter, T. R. Tan, R. Bowler, A. S. Sorensen, D. Leibfried, and D. J. Wineland, "Dissipative production of a maximally entangled steady state of two quantum bits," Nature **504**, 415–418 (2013).

[24] S. Shankar, M. Hatridge, Z. Leghtas, K. M. Sliwa, A. Narla, U. Vool, S. M. Girvin, L. Frunzio, M. Mirrahimi, and M. H. Devoret, "Autonomously stabilized entanglement between two superconducting quantum bits," Nature **504**, 419–422 (2013).

[25] G. Vacanti and A. Beige, "Cooling atoms into entangled states," New Journal of Physics **11**, 083008 (2009).

[26] F. Reiter, M. J. Kastoryano, and A. S. Sørensen, "Driving two atoms in an optical cavity into an entangled steady state using engineered decay," New Journal of Physics **14**, 053022 (2012).

[27] C. Aron, M. Kulkarni, and H. E. Türeci, "Steady-state entanglement of spatially separated qubits via quantum bath engineering," Phys. Rev. A **90**, 062305 (2014).

[28] M. B. Plenio and S. F. Huelga, "Entangled light from white noise," Phys. Rev. Lett. **88**, 197901 (2002).

[29] L. Hartmann, W. Dür, and H.-J. Briegel, "Steady-state entanglement in open and noisy quantum systems," Phys. Rev. A **74**, 052304 (2006).

[30] L. Hartmann, W. Dür, and H. J. Briegel, "Entanglement and its dynamics in open, dissipative systems," New Journal of Physics **9**, 230 (2007).

[31] L. Quiroga, F. J. Rodríguez, M. E. Ramírez, and R. París, "Nonequilibrium thermal entanglement," Phys. Rev. A **75**, 032308 (2007).

[32] M. Žnidarič, "Entanglement in stationary nonequilibrium states at high energies," Phys. Rev. A **85**, 012324 (2012).

[33] B. Bellomo and M. Antezza, "Steady entanglement out of thermal equilibrium," EPL (Europhysics Letters) **104**, 10006 (2013).

[34] B. Bellomo and M. Antezza, "Creation and protection of entanglement in systems out of thermal equilibrium," New Journal of Physics **15**, 113052 (2013).

[35] D. Boyanovsky and D. Jasnow, "Coherence of mechanical oscillators mediated by coupling to different baths," Phys. Rev. A **96**, 012103 (2017).

[36] J. B. Brask, G. Haack, N. Brunner, and M. Huber, "Autonomous quantum thermal machine for generating steady-state entanglement," New Journal of Physics **17**, 113029 (2015).

[37] N. Linden, S. Popescu, and P. Skrzypczyk, "How small can thermal machines be? the smallest possible refrigerator," Phys. Rev. Lett. **105**, 130401 (2010).

[38] P. P. Hofer, M. Perarnau-Llobet, L. D. M. Miranda, G. Haack, R.Silva, J. B. Brask, and N. Brunner, "Markovian master equations for quantum thermal machines: local versus global approach," New Journal of Physics **19**, 123037 (2017).

[39] J. O. González, L. A. Correa, G. Nocerino, J. P. Palao, D. Alonso, and G. Adesso, "Testing the Validity of the 'Local' and 'Global' GKLS Master Equations on an Exactly Solvable Model," Open Systems & Information Dynamics **24**, 1740010 (2017).

[40] G. Vidal and R. F. Werner, "Computable measure of entanglement," Phys. Rev. A **65**, 032314 (2002).

[41] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, "Proposed experiment to test local hidden-variable theories," Phys. Rev. Lett. **23**, 880–884 (1969).

[42] F. Verstraete, K. Audenaert, J. Dehaene, and B. De Moor, "A comparison of the entanglement measures negativity and concurrence," J. Phys. A: Math. Gen. **34**, 10327 (2001).

[43] M. A. Nielsen, "Conditions for a class of entanglement transformations," Phys. Rev. Lett. **83**, 436–439 (1999).

[44] Y.-X. Chen and S.-W. Li, "Quantum refrigerator driven by cur-rent noise," Europhys. Lett. **97**, 40003 (2012).

[45] P. P. Hofer, J.-R. Souquet, and A. A. Clerk, "Quantum heat engine based on photon-assisted cooper pair tunneling," Phys. Rev. B **93**, 041418 (2016).

[46] P. P. Hofer, M. Perarnau-Llobet, J. B. Brask, R. Silva, M. Huber, and N. Brunner, "Autonomous quantum refrigerator in a circuit qed architecture based on a josephson junction," Phys. Rev. B **94**, 235420 (2016).

[47] X. Gu, A. F. Kockum, A. Miranowicz, Y.-X. Liu, and F. Nori, "Microwave photonics with superconducting quantum circuits," Physics Reports **718-719**, 1 – 102 (2017).

[48] V. E. Manucharyan, J. Koch, L. I. Glazman, and M. H. Devoret, "Fluxonium: Single cooper-pair circuit free of charge offsets," Science **326**, 113–116 (2009).

[49] G. Zhu, D. G. Ferguson, V. E. Manucharyan, and J. Koch, "Circuit qed with fluxonium qubits: Theory of the dispersive regime," Phys. Rev. B **87**, 024510 (2013).

[50] V.E. Manucharyan, *Superinductance*, Ph.D. thesis (2012).

[51] I. M. Pop, K. Geerlings, G. Catelani, R. J. Schoelkopf, L. I. Glazman, and M. H. Devoret, "Coherent suppression of electromagnetic dissipation due to superconducting quasiparticles," Nature **508**, 369 (2014).

[52] Y.-H. Lin, L. B. Nguyen, N. Grabon, J. San Miguel, N. Pankratova, and V. E. Manucharyan, "Demonstration of protection of a superconducting qubit from energy decay," Phys. Rev. Lett. **120**, 150503 (2018).

[53] J. Majer, J. M. Chow, J. M. Gambetta, Jens Koch, B. R. Johnson, J. A. Schreier, L. Frunzio, D. I. Schuster, A. A. Houck, A. Wallraff, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, "Coupling superconducting qubits via a cavity bus," Nature **449**, 443 EP – (2007).

[54] Mika A. Sillanpää, Jae I. Park, and Raymond W. Simmonds, "Coherent quantum state storage and transfer between two phase qubits via a resonant cavity," Nature **449**, 438 EP – (2007).

[55] L. DiCarlo, J. M. Chow, J. M. Gambetta, Lev S. Bishop, B. R. Johnson, D. I. Schuster, J. Majer, A. Blais, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, "Demonstration of two-qubit algorithms with a superconducting quantum processor," Nature **460**, 240 EP – (2009).

[56] N. Cottet, "Private communication," .

[57] Y. Chen, C. Neill, P. Roushan, N. Leung, M. Fang, R. Barends, J. Kelly, B. Campbell, Z. Chen, B. Chiaro, A. Dunsworth, E. Jeffrey, A. Megrant, J. Y. Mutus, P. J. J. O'Malley, C. M. Quintana, D. Sank,

A. Vainsencher, J. Wenner, T. C. White, Michael R. Geller, A. N. Cleland, and J. M. Martinis, "Qubit architecture with high coherence and fast tunable coupling," Phys. Rev. Lett. **113**, 220502 (2014).

[58] D. I. Schuster, A. A. Houck, J. A. Schreier, A. Wallraff, J. M. Gambetta, A. Blais, L. Frunzio, J. Majer, B. Johnson, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, "Resolving photon number states in a superconducting circuit," Nature **445**, 515 EP – (2007).

[59] N. Cottet, S. Jezouin, L. Bretheau, P. Campagne-Ibarcq, Q. Ficheux, J. Anders, A. Auffèves, R. Azouit, P. Rouchon, and B. Huard, "Observing a quantum maxwell demon at work," Proc. Natl. Acad. Sci. U.S.A. **114**, 7561–7564 (2017).

[60] M. Jerger, P. Macha, A. R. Hamann, Y. Reshitnyk, K. Juliusson, and A. Fedorov, "Realization of a binary-outcome projection measurement of a three-level superconducting quantum system," Phys. Rev. Applied **6**, 014014 (2016).

[61] E. Jeffrey, D. Sank, J. Y. Mutus, T. C. White, J. Kelly, R. Barends, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. Megrant, P. J. J. O'Malley, C. Neill, P. Roushan, A. Vainsencher, J. Wenner, A. N. Cleland, and J. M. Martinis, "Fast accurate state measurement with superconducting qubits," Phys. Rev. Lett. **112**, 190504 (2014).

[62] N. T. Bronn, Y. Liu, J. B. Hertzberg, A. D. Córcoles, A. A. Houck, J. M. Gambetta, and J. M. Chow, "Broadband filters for abatement of spontaneous emission in circuit quantum electrodynamics," Applied Physics Letters **107**, 172601 (2015).

[63] A. Kou, W. C. Smith, U. Vool, I. M. Pop, K. M. Sliwa, M. H. Hatridge, L. Frunzio, and M. H. Devoret, "Simultaneous monitoring of fluxonium qubits in a waveguide," arXiv e-print , 1705.05712 (2017).

[64] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan, "Completely positive dynamical semigroups of nlevel systems," J. Math. Phys. **17**, 821–825 (1976).

[65] G. Lindblad, "On the generators of quantum dynamical semigroups," Commun. Math. Phys. **48**, 119–130 (1976).

[66] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, Oxford, 2002).

[67] C. Gardiner and P. Zoller, *Quantum Noise* (Springer-Verlag Berlin Heidelberg, 2004).

[68] G. Schaller, *Non-Equilibrium Master Equations* (Technische Universität Berlin, 2015).

Appendices

In Secs. A and B, we provide details of the derivations steady-state solutions of the two-qutrit and two-qudit master equations. In Sec. C we show how to generate any pure, entangled qutrit state. Finally, in Sec. D we provide details of the implementation of our scheme in circuit QED.

A Finding the steady state and filtered state for the two-qutrit machine

Here, we explain how to derive the steady-state solution of the reset model master equation, Eq. (4) of the main text, for two qutrits, and how to obtain the filtered two-qubit state Eq. (7).

The problem of finding an analytical solution is significantly simplified by the following observation: unless the interaction Hamiltonian induces transitions between $|k,j\rangle$ and $|k',j'\rangle$, there can be no coherence between these states in the steady state and the corresponding element in the density matrix vanishes, i.e., $\langle k,j|\rho|k',j'\rangle = 0$. This is because the dissipative processes locally reset each qutrit to a thermal state, which is diagonal, and hence do not generate any coherence. In the absence of the interaction Hamiltonian, dissipation would drive the system to a product of thermal states with no coherence. Note that, as seen in the previous section, in addition to the transitions directly present in H_{int}, it also induces second order transitions which need to be taken into account. Following such reasoning, one finds that there are only three non-zero off-diagonal elements in ρ. The density operator then takes the form:

$$\rho = \sum_{k,l=0}^{2} q_{kl}|k,l\rangle\langle k,l| + c_0|0,2\rangle\langle 2,0| + c_1|1,1\rangle\langle 2,0| + c_2|0,2\rangle\langle 1,1| + h.c \tag{24}$$

where q_{kl} are non-negative numbers that sum to one, and c_0, c_1 and c_2 are complex numbers which we can write as $c_k = v_k + iu_k$ with v_k, u_k real. Plugging this ansatz into the master equation and requiring $\partial\rho/\partial t = 0$, we obtain three independent equations for the off-diagonals terms. Solving the real and imaginary parts of this equation system returns v_k and u_k in terms of the q_{kl}. We are now faced with solving the system of equations corresponding to the diagonal of the right-hand-side of the master equation. This system of eight independent linear inhomogeneous equations can be written in the form $0 = AX + W$ where $X = (q_{00}, q_{01} \ldots, q_{21})^T$ and A is a 8×8 matrix depending on g_1, g_2, g_3, p_A and p_B, and W is a 8×1 row-matrix accounting for the inhomogeneous part of the equation system. The solution can then be written $X = -A^{-1}W$. We note that, as the dissipation induced by resets leaves no subspace invariant, A is always invertible when the rates p_A, p_B are non-zero, and there exists a unique steady state. For maximal temperature gradient, $T_A \to \infty$, $T_B = 0$, the solution can be computed analytically, although the expression is too unwieldy to display here.

To obtain the state given in the main text, one sets $g_1 = g_2 = g_3 = g$. Applying the local filters to the steady state, as explained in the main text, and renormalising, one directly obtains Eq. (7). Interestingly, the filtered state in this case is independent of g.

Interestingly, the scheme can also be adapted to generate any pure, entangled two-qubit state. This can be achieved by setting $g_3 = 0$ and taking $g_1 = g\cos(\theta)$ and $g_2 = g\sin(\theta)$. In this case, the filtered state becomes

$$\rho' = \begin{pmatrix} r_1 & 0 & 0 & 0 \\ 0 & r_2 & t & 0 \\ 0 & t^* & r_3 & 0 \\ 0 & 0 & 0 & 1-r_1-r_2-r_3 \end{pmatrix}, \tag{25}$$

with

$$r_1 = \frac{2p_A\cos^2(\theta)\left(-g^2p_A\cos(2\theta)+g^2(p_A+3p_B)+3p_B(p_A+p_B)^2\right)}{(2p_A+3p_B)\left(-g^2p_A\cos(4\theta)+g^2(p_A+6p_B)+6p_B(p_A+p_B)^2\right)} \tag{26}$$

$$r_2 = \frac{2\sin^2(\theta)(p_A+3p_B)\left(g^2p_A\cos(2\theta)+g^2(p_A+3p_B)+3p_B(p_A+p_B)^2\right)}{(2p_A+3p_B)\left(-g^2p_A\cos(4\theta)+g^2(p_A+6p_B)+6p_B(p_A+p_B)^2\right)} \tag{27}$$

$$r_3 = \frac{2\cos^2(\theta)(p_A+3p_B)\left(-g^2p_A\cos(2\theta)+g^2(p_A+3p_B)+3p_B(p_A+p_B)^2\right)}{(2p_A+3p_B)\left(-g^2p_A\cos(4\theta)+g^2(p_A+6p_B)+6p_B(p_A+p_B)^2\right)} \tag{28}$$

$$t = \frac{3p_B\sin(2\theta)\left(g^2(p_A+3p_B)+3p_B(p_A+p_B)^2\right)}{(2p_A+3p_B)\left(-g^2p_A\cos(4\theta)+g^2(p_A+6p_B)+6p_B(p_A+p_B)^2\right)}. \tag{29}$$

To first order in the ratio $\mu = p_A/p_B$, when additionally $g \ll p_B$, one finds the simple expression

$$\rho' = \begin{pmatrix} \frac{1}{3}\mu c_\theta^2 & 0 & 0 & 0 \\ 0 & (1-\frac{1}{3}\mu)s_\theta^2 & (1-\frac{2}{3}\mu)c_\theta s_\theta & 0 \\ 0 & (1-\frac{2}{3}\mu)c_\theta s_\theta & (1-\frac{1}{3}\mu)c_\theta^2 & 0 \\ 0 & 0 & 0 & \frac{1}{3}\mu s_\theta^2 \end{pmatrix}, \tag{30}$$

where $c_\theta = \cos(\theta)$, $s_\theta = \sin(\theta)$. For $\mu \to 0$, the state ρ' thus tends to the pure state (relabelling the basis states of both qubit subspaces to $|0\rangle$, $|1\rangle$)

$$|\psi_\theta\rangle = \sin\theta|0,1\rangle + \cos\theta|1,0\rangle. \tag{31}$$

Hence, any pure, entangled two-qubit state can be obtained from the qutrit thermal machine (up to local unitaries). In particular, for $\theta = \pi/4$ (i.e. $g_1 = g_2$), we again get a maximally entangled state.

The filtering success probability is given by

$$p_{suc} = \frac{2g^2p_A(2p_A+3p_B)\left(g^2p_A\cos(4\theta)-g^2(p_A+6p_B)-6p_B(p_A+p_B)^2\right)}{9A-9\left(B+C+D\right)}, \tag{32}$$

where

$$A = g^4p_A\cos(4\theta)(p_A+p_B)(p_A+2p_B), \tag{33}$$

$$B = g^4\left(p_A^3+11p_A^2p_B+26p_Ap_B^2+12p_B^3\right), \tag{34}$$

$$C = 2g^2p_B(p_A+p_B)^2\left(4p_A^2+15p_Ap_B+6p_B^2\right), \tag{35}$$

$$D = 6p_Ap_B^2(p_A+p_B)^4. \tag{36}$$

We note that for small p_A or g, the success probability depends only weakly on θ.

B Finding the steady state and filtered state for the qudit machine

In the following, we give the steady-state solution of the master equation, Eq. (4) in the main text, for any $d \geq 3$, with the interaction Hamiltonian Eq. (10). We work in the limit $T_A \to \infty$ and $T_B \to 0$. That is, we solve

$$0 = i[\rho, H_{int}] + p_A\left(\frac{\mathbb{1}}{d+1}\otimes \mathrm{Tr}_A\,\rho - \rho\right) + p_B\left(\mathrm{Tr}_B\,\rho\otimes|0\rangle\langle 0| - \rho\right). \tag{37}$$

Note that we have ignored the free Hamiltonian. We can do that since ρ commutes with the free Hamiltonian in the steady state. This is because H_{int} is energy preserving and can only generate coherence between states of the free Hamiltonians which are degenerate in energy (c.f. the previous section).

We show that the following state solves (37) for any $d \geq 3$.

$$\begin{aligned}\rho_{d+1} = \frac{1}{N}\Bigg[&\sum_{k,l=0}^{d} 2g^2 p_A^2 |k,l\rangle\langle k,l| + \sum_{k=0}^{d-1} c_1 |k,0\rangle\langle k,0| + c_2 |d,0\rangle\langle d,0| \\ &+ \sum_{k=0}^{d-1} 2(d+1)g^2 p_A p_B |k,d-k\rangle\langle k,d-k| + \sum_{k=0}^{d-1} c_3 |d,0\rangle\langle k,d-k| + h.c \\ &+ \sum_{k=1}^{d-1}\sum_{l=1}^{d-k} 2(d+1)g^2 p_A p_B |k+l-1,d-k-l+1\rangle\langle k-1,d-k+1| + h.c\Bigg],\end{aligned} \tag{38}$$

where we have defined coefficients

$$\begin{aligned} c_1 &= (d+1)p_A p_B \left(p_A + p_B\right)^2 + 2g^2\left((d+1)^2 p_B^2 + 2d(d+1)p_A p_B\right) \\ c_2 &= p_A\left((d+1)p_B\left(p_A+p_B\right)^2 + 2(d+1)g^2 d p_B\right) \\ c_3 &= i(d+1)g p_A p_B (p_A + p_B) \\ N &= (d+1)^2\left(p_A p_B (p_A+p_B)^2 + 2g^2\left(p_A^2 + 2d p_A p_B + d p_B^2\right)\right). \end{aligned} \tag{39}$$

First we compute the following partial traces

$$\mathrm{Tr}_A(\rho) = \frac{1}{N}\left[\sum_{l=0}^{d} 2(d+1)g^2 p_A^2 |l\rangle\langle l| + (dc_1 + c_2)|0\rangle\langle 0| + \sum_{k=0}^{d-1} 2(d+1)g^2 p_A p_B |d-k\rangle\langle d-k|\right] \tag{40}$$

$$\mathrm{Tr}_B(\rho) = \frac{1}{N}\left[\sum_{k=0}^{d} 2(d+1)g^2 p_A^2 |k\rangle\langle k| + \sum_{k=0}^{d-1} c_1 |k\rangle\langle k| + c_2 |d\rangle\langle d| + \sum_{k=0}^{d-1} 2(d+1)g^2 p_A p_B |k\rangle\langle k|\right]. \tag{41}$$

Subsequently, one can show that

$$\begin{aligned} &p_A\left(\frac{\mathbb{1}}{d+1}\otimes \mathrm{Tr}_A\,\rho - \rho\right) + p_B\left(\mathrm{Tr}_B\,\rho \otimes |0\rangle\langle 0| - \rho\right) = \\ &-\frac{1}{N}\Bigg[-2d(d+1)g^2 p_A p_B (p_A+p_B)|d,0\rangle\langle d,0| + 2(d+1)g^2 p_A p_B (p_A+p_B)\sum_{k=0}^{d-1}|k,d-k\rangle\langle k,d-k| \\ &\qquad + c_3(p_A+p_B)\sum_{k=0}^{d-1}|d,0\rangle\langle k,d-k| + c_3^*(p_A+p_B)\sum_{k=0}^{d-1}|k,d-k\rangle\langle d,0| \\ &+2(d+1)g^2 p_A p_B (p_A+p_B)\sum_{k=1}^{d-2}\sum_{l=1}^{d-1-k}\left(|k+l-1,d-k-l\rangle\langle k-1,d-k| + |k-1,d-k\rangle\langle k+l-1,d-k-l|\right). \end{aligned} \tag{42}$$

Similarly, extensive simplification of the commutator in (37) gives

$$\begin{aligned}[\rho, H_{int}] = &\sum_{k=0}^{d-1}\left(c_2 g - 2d(d+1)g^3 p_A p_B\right)|d,0\rangle\langle k,d-k| + 2idg\,\mathrm{Im}\left(c_3\right)|d,0\rangle\langle d,0| \\ &+\sum_{k=0}^{d-1}\left(-c_2 g + 2d(d+1)g^3 p_A p_B\right)|k,d-k\rangle\langle d,0| - 2ig\,\mathrm{Im}\left(c_3\right)\sum_{k,l=0}^{d-1}|k,d-k\rangle\langle l,d-l|. \end{aligned} \tag{43}$$

Inserting (42) and (43) back into (37), the verification reduces to two equations

$$i\left(g c_2 - 2d(d+1)g^3 p_A p_B\right) = c_3(p_A+p_B) \tag{44}$$

$$i\left(2dg i\mathrm{Im}(c_3)\right) = -2d(d+1)g^2 p_A p_B (p_A+p_B) \tag{45}$$

From the definition of c_2 and c_3, it is easily shown that both these equations are satisfied. Hence, the state (38) is the steady-state of the thermal machine.

Finally, we show that by applying suitable local filters to ρ, we obtain two maximally entangled d-level systems. The local projectors are

$$\Pi_A = \sum_{k=0}^{d-1} |k\rangle\langle k| \qquad \Pi_B = \sum_{l=1}^{d} |l\rangle\langle l|. \tag{46}$$

The filtered state becomes

$$\rho' = \frac{\Pi_A \otimes \Pi_B \rho \Pi_A \otimes \Pi_B}{\mathrm{Tr}\left[\Pi_A \otimes \Pi_B \rho\right]} = \frac{1}{(2g^2 p_A^2 d^2 + 2d(d+1)g^2 p_A p_B)}\Big[\sum_{k=0}^{d-1}\sum_{l=1}^{d} 2g^2 p_A^2 |k,l\rangle\langle k,l| + \sum_{k=0}^{d-1} 2(d+1)g^2 p_A p_B |k,d-k\rangle\langle k,d-k| \sum_{k=1}^{d-2}\sum_{l=1}^{d-k-1} 2dg^2 p_A p_B \big(|k+l-1,d-k-l\rangle\langle k-1,d-k| + |k-1,d-k\rangle\langle k+l-1,d-k-l|\big)\Big]. \tag{47}$$

In the limit $p_A \ll p_B$ this indeed reduces to the maximally entangled state of two d-level systems

$$\rho' = |S_d\rangle\langle S_d| + O(\frac{p_A}{p_B}). \tag{48}$$

C Generating all pure, entangled states of two qutrits

All pure entangled two-qutrit states can be written using the Schmidt-decomposition as

$$|\psi^3_{\lambda_1,\lambda_2,\lambda_3}\rangle = \sum_{i=0}^{2} \lambda_i |i,i\rangle, \tag{49}$$

with $\lambda \geq 0$ and $\lambda_0^2 + \lambda_1^2 + \lambda_2^2 = 1$. Here, we show that any such state can be generated using a two-ququart thermal machine and local filtering.

The machine consists of two ququarts (four-level systems) with an interaction Hamiltonian

$$H_{int} = g_0(|0,3\rangle\langle 3,0| + |3,0\rangle\langle 0,3|) + g_1(|1,2\rangle\langle 3,0| + |3,0\rangle\langle 1,2|) + g_2(|2,1\rangle\langle 3,0| + |3,0\rangle\langle 2,1|). \tag{50}$$

Where we choose $g_i = g\lambda_i$ for some small constant g. In the limit of maximal thermal gradient, $T_A \to \infty$ and $T_B = 0$, the steady-state solution of the master equation can be derived using the method outlined in Sec. A. The steady state ρ is then filtered to a space of two qutrits corresponding to the projectors $\Pi_A = \mathbb{1} - |3\rangle\langle 3|$ and $\Pi_B = \mathbb{1} - |0\rangle\langle 0|$. The filtered state ρ' depends on $\lambda_0, \lambda_1, \lambda_2, p_A, p_B$ and g. We consider the limit in which $p_A \ll p_B$. This eliminates the dependence on g and p_B. The resulting state is found to be

$$\rho' = |\psi^3_{\lambda_0,\lambda_1,\lambda_2}\rangle\langle\psi^3_{\lambda_0,\lambda_1,\lambda_2}| + O(\frac{p_A}{p_B}). \tag{51}$$

Thus, we can generate any pure entangled state of two qutrits.

Based on this result, and the corresponding case for qubits in the main text, we conjecture that any pure entangled state in any dimension can be generated by a generalisation of this thermal machine. Specifically

Conjecture

Let the autonomous thermal machine of two $d+1$-level systems coupled to baths of temperature $T_A \to \infty$ and $T_B = 0$ respectively, operate with an interaction Hamiltonian of the form

$$H_{int} = \sum_{k=0}^{d-1} g_k |d,0\rangle\langle k, d-k| + h.c. \tag{52}$$

where we take $g_i = g\lambda_i$ for some small constant g, and where $\{\lambda_i\}_i$ are the Schmidt coefficients of any pure entangled state of two systems of dimension d. Applying the projectors $\Pi_A = \mathbb{1} - |d\rangle\langle d|$ and $\Pi_B = \mathbb{1} - |0\rangle\langle 0|$ to the steady-state of the system, and considering the limit $p_A \ll p_B$, the filtered state becomes

$$|\psi^d_{\lambda_0,\ldots,\lambda_{d-1}}\rangle = \sum_{i=0}^{d-1} \lambda_i |i,i\rangle. \tag{53}$$

In this work, we have shown this conjecture to be true for $d = 2$ and $d = 3$. In addition, we have checked numerically that the conjecture holds for $d = 4$ and $d = 5$ for 100 randomly chosen pure entangled states. Note that the number of adjustable paramters (the g_k) exactly match the number of Schmidt coefficients required to describe a pure state of two systems of dimension d.

D Reset vs Lindblad master equation

The reset model considered in the main text is intuitive, amenable to analytical analysis, and captures the essential physics of a multipartite quantum system in contact with thermal baths. However, instantaneous thermal resets are a simplification with respect to realistic implementations. In this appendix, we first show that a reset master equation is exactly equivalent to a master equation on standard Lindblad form and derive an explicit mapping between the two. The corresponding Linblad master equation describes dissipation due to local coupling with bosonic thermal baths combined with additional pure dephasing. We then discuss how the optimal conditions for entanglement generation derived for the reset model translate to the Lindblad model.

D.1 Equivalence for single qutrits

Since a reset master equation generates Markovian (specifically semi-group) dynamics, there must exist a master equation of standard Gorini-Kossakowski-Sudarshan-Lindblad form which generates the same dynamics [64, 65]. Here, we give an explicit mapping between these two forms.

We first consider a single qutrit and show that any reset master equation of the form

$$\frac{\partial \rho}{\partial t} = i[\rho, H] + \mathcal{L}_{res}(\rho) = i[\rho, H] + p\,(\tau - \rho)\ , \tag{54}$$

where p is a positive rate and τ is a thermal state, is equivalent to a master equation on standard Lindblad form given by

$$\frac{\partial \rho}{\partial t} = i[\rho, H] + \mathcal{L}_{lin}(\rho) = i[\rho, H] + \sum_{k\in\{01,12,02\}} \left(\Gamma_k^+ \mathcal{D}[\sigma_k^+]\rho + \Gamma_k^- \mathcal{D}[\sigma_k^-]\rho + \gamma_k \mathcal{D}[\sigma_k^z]\rho\right)\ . \tag{55}$$

where the label k runs over the three possible qubit subspaces of the qutrit, $\Gamma_k^\pm$ and γ_k are positive rates, and $\sigma_k^\pm$ and σ_k^z are jump operators acting on the qubit subspace labeled by k. Specifically

$$\sigma^+_{mn} = |n\rangle\langle m|\,, \qquad \sigma^-_{mn} = |m\rangle\langle n|\,, \qquad \sigma^z_{mn} = |m\rangle\langle m| - |n\rangle\langle n|. \tag{56}$$

The dissipators take the standard Lindblad form

$$\mathcal{D}[A]\rho = A\rho A^\dagger - \frac{1}{2}\{A^\dagger A, \rho\}. \tag{57}$$

Having established a mapping between (54) and (55), we generalise it to two coupled qubits below.

By a mapping between (54) and (55) we mean a set of relations defining $\Gamma_k^{\pm}$ and γ_k in terms of p and the elements of τ such that the right-hand sides of the two equations become equal. Since the Hamiltonian parts of (54) and (55) are the same, we only need to match the dissipators

$$\mathcal{L}_{res}(\rho) = p\,(\tau - \rho)\,, \tag{58}$$

and

$$\mathcal{L}_{lin}(\rho) = \sum_{k\in\{01,12,02\}} \left(\Gamma_k^+ \mathcal{D}[\sigma_k^+]\rho + \Gamma_k^- \mathcal{D}[\sigma_k^-]\rho + \gamma_k \mathcal{D}[\sigma_k^z]\rho\right). \tag{59}$$

The space of 3×3 hermitian matrices is spanned by the projectors $|m\rangle\langle m|$, $m = 0,1,2$ and off-diagonals $|m\rangle\langle n| + |n\rangle\langle m|$ and $i|m\rangle\langle n| - i|n\rangle\langle m|$ with $m,n = 0,1,2$, $m < n$. The two dissipators will therefore act the same on any state ρ if they act the same on each of these basis elements. By demanding $\mathcal{L}_{res}(|m\rangle\langle m|) = \mathcal{L}_{lin}(|m\rangle\langle m|)$ we obtain a set of six equations (plus three redundant ones) which determine the $\Gamma_k^{\pm}$ in terms of p and τ. Similarly, by requiring $\mathcal{L}_{res}(|m\rangle\langle n| + |n\rangle\langle m|) = \mathcal{L}_{lin}(|m\rangle\langle n| + |n\rangle\langle m|)$ for the three off-diagonals we obtain three more equations which determine the γ_k. Specifically, the solution is

$$\begin{array}{lll} \Gamma_{01}^- = p\tau_0\,, & \Gamma_{01}^+ = p\tau_1\,, & \gamma_{01} = \frac{1}{9}p(2-3\tau_2)\,, \\ \Gamma_{02}^- = p\tau_0\,, & \Gamma_{02}^+ = p\tau_2\,, & \gamma_{02} = \frac{1}{9}p(2-3\tau_1)\,, \\ \Gamma_{12}^- = p\tau_1\,, & \Gamma_{12}^+ = p\tau_2\,, & \gamma_{12} = \frac{1}{9}p(2-3\tau_0)\,. \end{array} \tag{60}$$

where τ_0, τ_1, τ_2 are the populations of the states $|0\rangle$, $|1\rangle$, $|2\rangle$ in the thermal state (i.e. the diagonal elements of τ). One can check that indeed using (60) one has $\mathcal{L}_{lin}(\rho) = \mathcal{L}_{res}(\rho)$ for any arbitrary qutrit state ρ. Explicitly, at a given temperature T, the populations are given by

$$\tau_m = \frac{e^{-E_m/T}}{\sum_{n=0}^{2} e^{-E_n/T}}, \tag{61}$$

where E_m is the energy of state $|m\rangle$, $m = 0,1,2$. It follows that the jump rates in the Lindblad master equation satisfy detailed balance, as one would expect

$$\frac{\Gamma_{mn}^+}{\Gamma_{mn}^-} = e^{-(E_n - E_m)/T} \tag{62}$$

We can then understand these jumps as being induced by a bosonic bath [66–68]

$$\Gamma_{mn}^+ = \Gamma_{mn} n_B(E_n - E_m, T)\,, \tag{63}$$

$$\Gamma_{mn}^- = \Gamma_{mn}[1 + n_B(E_n - E_m, T)]\,, \tag{64}$$

where

$$n_B(E,T) = \frac{1}{e^{E/T} - 1} \tag{65}$$

is the Bose-Einstein distribution, and the coupling constant Γ_{mn} for transitions between states $|m\rangle$ and $|n\rangle$ is given by

$$\Gamma_{mn} = p\frac{\tau_n}{n_B(E_n - E_m, T)}. \tag{66}$$

D.2 Equivalence for two qutrits

The mapping derived between the single-qutrit master equations (54) and (55) can be applied directly to a system of two weakly coupled qutrits, as considered in the main text. Specifically, the reset master equation

$$\frac{\partial\rho}{\partial t} = i[\rho, H] + p_A(\tau_A \otimes \mathrm{Tr}_A(\rho) - \rho) + p_B(\mathrm{Tr}_B(\rho)\otimes\tau_B - \rho) \tag{67}$$

is equivalent to the following local Lindblad master equation

$$\frac{\partial\rho}{\partial t} = i[\rho, H] + \sum_{k\in\{01,12,02\}} \left(\Gamma^{+}_{A,k}\mathcal{D}[\sigma^{+}_{A,k}]\rho + \Gamma^{-}_{A,k}\mathcal{D}[\sigma^{-}_{A,k}]\rho + \gamma_{A,k}\mathcal{D}[\sigma^{z}_{A,k}]\rho\right) + \sum_{k\in\{01,12,02\}} \left(\Gamma^{+}_{B,k}\mathcal{D}[\sigma^{+}_{B,k}]\rho + \Gamma^{-}_{B,k}\mathcal{D}[\sigma^{-}_{B,k}]\rho + \gamma_{B,k}\mathcal{D}[\sigma^{z}_{B,k}]\rho\right), \tag{68}$$

where the jump operators are defined analogously to (56) above for each qutrit A and B locally. That is $\sigma^{+}_{A,k} = \sigma^{+}_{k} \otimes \mathbb{1}$ and $\sigma^{+}_{B,k} = \mathbb{1} \otimes \sigma^{+}_{k}$, and similarly for the other jump operators. The mapping which makes the two master equations equivalent is given by (60) applied to each system A and B individually, as one can check.

Just as in the single-qutrit case, the jump rates in the Lindblad master equation correspond to bosonic baths.

$$\begin{aligned} \Gamma^{+}_{A,mn} &= \Gamma_{A,mn} n_B(E^A_n - E^A_m, T_A)\,, & \Gamma^{+}_{B,mn} &= \Gamma_{A,mn} n_B(E^B_n - E^B_m, T_B)\,, \\ \Gamma^{-}_{A,mn} &= \Gamma_{A,mn}[1 + n_B(E^A_n - E^A_m, T_A)]\,, & \Gamma^{-}_{A,mn} &= \Gamma_{B,mn}[1 + n_B(E^B_n - E^B_m, T_B)]. \end{aligned} \tag{69}$$

When considering potential implementations of our scheme in the main text, we use a master equation of the form (68) for the numerical simulation, taking values for the bath coupling strengths $\Gamma_{A,mn}$, $\Gamma_{B,mn}$ and pure dephasing rates $\gamma_{A,mn}$, $\gamma_{B,mn}$ based on recent experimental works, as explained in the text.

D.3 Optimal settings for generating maximal entanglement

In the main text, we identified conditions under which our scheme generates a pure, maximally entangled state, using the reset model. Using the mapping above, we can translate these conditions to the Lindblad model.

The ideal temperatures for entanglement generation in the reset model are $T_A \to \infty$ and $T_B \to 0$. This means that the thermal populations become $\tau^A_0 = \tau^A_1 = \tau^A_2 = 1/3$ and $\tau^B_0 = 1$, $\tau^B_1 = \tau^B_2 = 0$. In turn, for the Lindblad jump rates, using (60) this implies that

$$\Gamma^{+}_{A,mn} = \Gamma^{-}_{A,mn}, \tag{70}$$

and

$$\Gamma^{-}_{B,01} = \Gamma^{-}_{B,02}\,, \qquad \Gamma^{-}_{B,12} = \Gamma^{+}_{B,mn} = 0. \tag{71}$$

The former condition is satisfied in the Lindblad model also in the limit $T_A \to \infty$ since then $n_B(E^A_n - E^A_m, T_A) \gg 1$. The latter condition can be satified in the limit $T_B \to 0$, where $n_B(E^B_n - E^B_m, T_B) \to 0$, if the coupling strength $\Gamma_{B,12}$ also vanishes.

Thus, we see that the Lindblad model is in principle compatible with the ideal limit for entanglement generation identified using the reset model, and one can thus expect entanglement generation to be possible also under such a more realistic model. We stress that it is not necessary to go to the ideal limit to achieve near-perfect entanglement generation. As shown in Fig. 5 in the main text, using parameter values which are reasonable in the context of the current experimental state of the art, entanglement close to maximal can be attained.

Self-testing quantum states and measurements in the prepare-and-measure scenario

Armin Tavakoli,[1] Jędrzej Kaniewski,[2] Tamás Vértesi,[3] Denis Rosset,[4,5] and Nicolas Brunner[1]
[1]*Département de Physique Appliquée, Université de Genève, CH-1211 Genève, Switzerland*
[2]*QMATH, Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark*
[3]*Institute for Nuclear Research, Hungarian Academy of Sciences, P.O. Box 51, 4001 Debrecen, Hungary*
[4]*Perimeter Institute for Theoretical Physics, 31 Caroline St. N, Waterloo, Ontario, Canada N2L 2Y5*
[5]*Institute for Quantum Optics and Quantum Information (IQOQI), Boltzmangasse 3, 1090 Vienna, Austria*

(Received 30 January 2018; revised manuscript received 11 September 2018; published 6 December 2018)

The goal of self-testing is to characterize an *a priori* unknown quantum system based solely on measurement statistics, i.e., using an uncharacterized measurement device. Here we develop self-testing methods for quantum prepare-and-measure experiments, thus not necessarily relying on entanglement and/or violation of a Bell inequality. We present noise-robust techniques for self-testing sets of quantum states and measurements, assuming an upper bound on the Hilbert space dimension. We discuss in detail the case of a $2 \to 1$ random access code with qubits, for which we provide analytically optimal self-tests. The simplicity and noise robustness of our methods should make them directly applicable to experiments.

DOI: 10.1103/PhysRevA.98.062307

I. INTRODUCTION

Predicting the results of measurements performed on a given physical system has traditionally been the main concern of physics. However, with the advent of device-independent quantum information processing [1–3], the opposite question has become relevant. More specifically, given an initially unknown system and an uncharacterized measurement device, what can be inferred about the physics of the experiment based solely on the observed measurement statistics? Despite the apparent generality of this question, certain cases do allow for a precise characterization of the system. This is referred to as self-testing [4,5].

The possibility to self-test quantum states and measurements usually relies on quantum nonlocality. Consider two distant observers performing local measurements on a shared quantum state. When the resulting statistics leads to violation of a Bell inequality [6], it is necessarily the case that the shared quantum state is entangled and, moreover, that the local quantum measurements are incompatible; see, e.g., Ref. [7]. Furthermore, for specific Bell inequalities, maximal violation (i.e., the largest possible value in quantum theory) implies that the quantum state and the measurements can be uniquely identified (up to local isometries). For instance, a maximal violation of the Clauser-Horne-Shimony-Holt (CHSH) Bell inequality [8] implies maximally incompatible measurements (two anticommuting Pauli observables) and a shared maximally entangled two-qubit state [9–12]. More recently, it has been demonstrated that all bipartite pure entangled states can be self-tested [13], as well as certain multipartite entangled states [14–16]. Another important progress is the development of self-testing methods robust to noise [17–23]. For instance, given a certain level of violation of a Bell inequality (but not necessarily maximal), the fidelity between the initially unknown state and a given target state can be lower bounded.

Self-testing thus offers promising perspectives for the certification of quantum systems in experiments (see, e.g., Ref. [24]), as well as for device-independent quantum information protocols [25]. It is therefore natural to ask whether the concept of self-testing can be applied to more general quantum experiments, beyond those based on entanglement and nonlocality.

In the present work, we develop self-testing methods tailored to the prepare-and-measure scenario. This covers a broad class of experiments, where quantum communication schemes [e.g., the BB84 quantum key distribution (QKD) protocol] are prominent examples. In this setting, a preparation device initially prepares a quantum system in different possible states. The system is then transmitted to a measurement device, which performs different possible measurements on it. While it is still possible in this case to characterize certain physical properties of the system based only on statistics, this requires in general an assumption on the devices. One possibility, which we will follow here, is to assume that the set of quantum states and measurements admit a full description in a Hilbert space of given dimension [26–28]. Intuitively this means that the amount of information communicated from the preparation device to the measurement device is assumed to be upper bounded. Such a scenario considering quantum systems of fixed dimension, but otherwise uncharacterized, is referred to as semi-device-independent, and opens interesting possibilities for quantum information processing [29–33].

Here we demonstrate techniques for robustly self-testing sets of prepared quantum states, as well as sets of quantum measurements. These methods allow one to (i) assess the compatibility of given sets of preparations and measurements with the observed statistics and (ii) lower bound the average fidelity between the unknown preparations (measurements) and a set of ideal quantum states (measurements). We discuss in detail a simple prepare-and-measure scenario, namely the $2 \to 1$ random access code (RAC). This allows us to provide analytically optimal self-tests for a pair of anticommuting Pauli observables, and for a set of four qubit states corresponding

to the eigenstates of two anticommuting Pauli observables. We then generalize these results to other prepare-and-measure scenarios. The simplicity and robustness of our methods should make them directly applicable to experiments. We conclude with a number of open questions.

II. SCENARIO

We consider a quantum prepare-and-measure experiment. Upon receiving input x, a preparation device emits a physical system in a quantum state ρ_x. The system is then transmitted to a measurement device, which, upon receiving an input y, performs a quantum measurement returning an outcome b. Formally, the measurement is described by a set of positive operators M_y^b, that equal identity when summed over b. Importantly both the specific states ρ_x and measurements M_y^b are *a priori* unknown to the observer. The statistics of the experiment is then given by $P(b|x, y) = \mathrm{tr}(\rho_x M_y^b)$. In this setting, any possible probability distribution can be obtained, given that the prepared states ρ_x can be taken in a sufficiently large Hilbert space. This is however no longer the case when we limit the Hilbert space dimension; specifically we impose that $\rho_x \in \mathcal{L}(\mathbb{C}^d)$ for some given $d < |x|$ (where $|x|$ denotes the number of possible inputs x). In this case, limits on the set of possible distributions can be captured via inequalities of the form

$$\mathcal{A} = \sum_{x,y,b} \alpha_{xyb} P(b|x, y) \leqslant Q_d, \tag{1}$$

where α_{xyb} are real coefficients. These "dimension witnesses" allow one to place device-independent lower bounds on the dimension of the quantum system [26].

Subsequently, one can ask what the limitations are on the set of distributions $P(b|x, y)$ given that the preparations admit a classical d-dimensional representation, i.e., there exists a d-dimensional basis such that all states ρ_x are diagonal in this basis. We denote by C_d the maximal value of the quantity $\mathcal{A}$ in this case. Interestingly, for well-chosen quantities $\mathcal{A}$, one finds that $C_d < Q_d$. Thus, for a given system dimension d, quantum systems outperform classical ones, in the sense that certain quantum distributions cannot be reproduced classically [26]. This quantum advantage can be viewed as the origin for the possibility of developing self-testing methods for the prepare-and-measure scenario, in analogy to Bell inequality violation being the root for self-testing entangled states.

In the following we present robust self-testing techniques based on specific dimension witnesses $\mathcal{A}$. Based only on the value of $\mathcal{A}$, which is directly accessible from the experiment statistics, we characterize the (initially unknown) prepared states and measurements. In particular, when the maximal value of the witness is obtained, i.e., $\mathcal{A} = Q_d$, then a specific set of pure states $\rho_x = |\psi_x\rangle\langle\psi_x|$ and a specific set of projective measurements M_y^b must have been used (up to a unitary). Moreover, when a nonmaximal value $\mathcal{A} < Q_d$ is obtained, the compatibility of given sets of preparations and measurements can be assessed. Finally, one can efficiently lower bound the fidelity between the prepared states and measurements and the ideal (or target) states and measurements leading to $\mathcal{A} = Q_d$.

Note that a recent series of works followed a related though conceptually different approach, based on hypothesis testing [34–36]. This method does however not allow for self-testing.

III. 2 → 1 RANDOM ACCESS CODE

We discuss in detail a simple prepare-and-measure experiment. This involves four possible preparations, denoted by $x = (x_0, x_1)$ (where $x_j \in \{0, 1\}$), and two possible binary measurements, $y \in \{0, 1\}$ and $b \in \{0, 1\}$. The score is given by

$$\mathcal{A}_2 = \frac{1}{8} \sum_{x_0,x_1,y} P(b = x_y|x_0, x_1, y). \tag{2}$$

This means that, upon receiving input y, the measurement device should return the output $b = x_y$, i.e., the yth bit of the input bit-string x received by the preparation device. Hence the name of a $2 \to 1$ RAC [37–39]. Note that all inputs are assumed to be chosen uniformly at random. Indeed, this task is nontrivial only when $d < 4$; here we will consider the case $d = 2$, i.e., qubits. In this case, one finds the tight bounds $C_2 = 3/4$ and $Q_2 = (1 + 1/\sqrt{2})/2 \approx 0.85$ [37]. The classical bound C_2 can be obtained by simply always sending the bit x_0. The quantum bound Q_2 is obtained via the following "ideal" strategy. The four qubit preparations correspond to the pure states

$$\rho_{jj}^{\text{ideal}} = \frac{\mathbb{1} + (-1)^j \sigma_x}{2}, \quad \rho_{j\bar{j}}^{\text{ideal}} = \frac{\mathbb{1} + (-1)^j \sigma_z}{2} \tag{3}$$

for $j \in \{0, 1\}$ and $\bar{j} = 1 - j$. These are simply the eigenstates of the Pauli observables σ_x and σ_z. Next, the measurements are projective and given by two anticommuting Pauli observables

$$M_y^{\text{ideal}} = \left(M_y^0\right)^{\text{ideal}} - \left(M_y^1\right)^{\text{ideal}} = \frac{\sigma_x + (-1)^y \sigma_z}{\sqrt{2}}. \tag{4}$$

These qubit preparations and measurements represent the ideal situation, where the maximal value $\mathcal{A}_2 = Q_2$ is achieved. In the following we will determine what restrictions apply to the possible preparations and measurements, given that a particular value of $\mathcal{A}_2$ is observed. In particular, when the maximal value $\mathcal{A}_2 = Q_2$ is attained, both the states and the measurements must be the ideal ones as given above (up to a unitary).

IV. SELF-TESTING PREPARATIONS

Here we find restrictions on the set of prepared states given an observed value of $\mathcal{A}_2$. For convenience, we write the qubit preparations as $\rho_{x_0x_1} = (\mathbb{1} + \vec{m}_{x_0x_1} \cdot \vec{\sigma})/2$, where $\vec{m}_{x_0x_1}$ denotes the Bloch vector (satisfying $|\vec{m}_{x_0x_1}| \leqslant 1$) and $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ denotes the vector of Pauli matrices.

The first step consists in reexpressing

$$\begin{aligned} \mathcal{A}_2 &= \frac{1}{2} + \frac{1}{8} \sum_y \mathrm{tr}\left(M_y^0 V_y\right) \\ &\leqslant \frac{1}{2} + \frac{1}{8} \sum_y \sqrt{\mathrm{tr}\left(M_y^0 V_y^2\right) \mathrm{tr}\left(M_y^0\right)}, \end{aligned} \tag{5}$$

where $V_y = \sum_{x_0,x_1}(-1)^{x_y}\rho_{x_0x_1}$. In the second step we used that for a positive semidefinite O and a Hermitian operator R, it holds that $|\mathrm{tr}(OR)|^2 \leqslant \mathrm{tr}(OR^2)\,\mathrm{tr}(O)$ [23]. Without loss of generality, we can restrict ourselves to extremal qubit measurements, which are here projective rank-one operators. Consequently, we have that $\mathrm{tr}(M_y^0) = 1$. Next, we obtain $V_y^2 = \frac{1}{2}[\beta + (-1)^y\alpha]\mathbb{1}$, where $\beta = \frac{1}{2}\sum_{x_0,x_1}|\vec{m}_{x_0x_1}|^2 - \vec{m}_{00}\cdot\vec{m}_{11} - \vec{m}_{01}\cdot\vec{m}_{10}$ and $\alpha = (\vec{m}_{00} - \vec{m}_{11})\cdot(\vec{m}_{01} - \vec{m}_{10})$. Finally, we find that Eq. (5) reduces to

$$\mathcal{A}_2 \leqslant \frac{1}{2} + \frac{1}{8\sqrt{2}}[\sqrt{\beta+\alpha} + \sqrt{\beta-\alpha}]. \tag{6}$$

This provides a tight self-test of the prepared states (in terms of their Bloch vectors), for any given value of $\mathcal{A}_2$. Let us start with the case $\mathcal{A}_2 = Q_2$. Since $\sqrt{\beta+\alpha} + \sqrt{\beta-\alpha} = \sqrt{2\beta + 2\sqrt{\beta^2-\alpha^2}}$, we see that Eq. (6) is maximal iff $\alpha = 0$ and β is maximal. This turns out to be achievable. In order to maximize β, we need (i) $\forall x_0x_1 : |\vec{m}_{x_0x_1}| = 1$, i.e., that all preparations are pure states, and (ii) that $\vec{m}_{00}\cdot\vec{m}_{11} = \vec{m}_{01}\cdot\vec{m}_{10} = -1$, i.e., the states correspond to (pairwise) antipodal Bloch vectors. We define $\vec{r}_0 = \vec{m}_{00} = -\vec{m}_{11}$ and $\vec{r}_1 = \vec{m}_{01} = -\vec{m}_{10}$. Consequently, we find $\alpha = 4\vec{r}_0\cdot\vec{r}_1$. Therefore, in order to have $\alpha = 0$, we must choose $\vec{r}_0\cdot\vec{r}_1 = 0$. This implies that the right-hand side of Eq. (6) is upper bounded by Q_2. Therefore, we conclude that when observing maximal value $\mathcal{A}_2 = Q_2$, the set of four prepared states must be equivalent (up to a unitary rotation) to the set of four ideal states; we note that this was also shown in Ref. [40] in the context of QKD.

More generally, for any value $\mathcal{A}_2$, one can find a set of preparations (and corresponding measurements) such that the inequality (6) is saturated; see Appendix A. For the case of classical preparations (i.e., diagonal in a given basis), the Bloch vectors can simply be replaced by numbers $m_{x_0x_1} \in [-1, 1]$, and we get $\mathcal{A}_2 \leqslant C_2$.

V. SELF-TESTING MEASUREMENTS

Let us now consider self-testing of measurements. Using that $M_y = M_y^0 - M_y^1$, we write

$$\mathcal{A}_2 \leqslant \frac{1}{2} + \frac{1}{16}\sum_{x_0,x_1}\lambda_{\max}[(-1)^{x_0}M_0 + (-1)^{x_1}M_1], \tag{7}$$

where $\lambda_{\max}[X]$ is the largest eigenvalue of the (Hermitian) operator X. Since the upper bound corresponds to choosing the optimal preparations for a fixed pair of observables, it simply quantifies the optimal performance achievable using these observables. If M_0 and M_1 are qubit observables the upper bound can be evaluated exactly (see Appendix A) to give

$$\mathcal{A}_2 \leqslant \tfrac{1}{2} + \tfrac{1}{16}(\sqrt{2\mu + 2\nu - \eta_+^2} + \sqrt{2\mu - 2\nu - \eta_-^2}), \tag{8}$$

where $\mu = \mathrm{tr}(M_0^2 + M_1^2)$, $\nu = \mathrm{tr}\{M_0, M_1\}$, and $\eta_\pm = \mathrm{tr}(M_0 \pm M_1)$. The right-hand side reaches the optimal value Q_2 iff $\mu = 4$, $\eta_\pm = 0$, and $\nu = 0$, which implies anticommuting projective observables (i.e., projective measurement operators). In other words, observing $\mathcal{A}_2 = Q_2$ implies that the measurements are unitarily equivalent to the ideal ones. Moreover, note that inequality (8) is tight; for any value of $\mathcal{A}_2$ one can find measurements (and corresponding states) such that inequality is saturated (see Appendix A). It follows that any pair of projective, rank-one observables that is incompatible ($|\nu| < 4$) can lead to $\mathcal{A}_2 > C_2$.

VI. ROBUST SELF-TESTING OF THE PREPARATIONS

We now discuss the problem of characterizing the fidelity between the realized preparations and the ideal ones. This will allow us to quantify the distance of the prepared states with respect to the ideal ones. Again, we want to develop self-testing methods which are based only on the value of $\mathcal{A}_2$.

More formally, given an arbitrary set of preparations, we define the average fidelity with the ideal preparations to be $S(\{\rho_{x_0x_1}\}) = \max_\Lambda \sum_{x_0,x_1} F(\rho_{x_0x_1}^{\mathrm{ideal}}, \Lambda[\rho_{x_0x_1}])/4$, where Λ is a quantum channel, i.e., a completely positive trace-preserving map. Here the fidelities $F(\rho,\sigma) = \mathrm{tr}(\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}})$ simplify to $F(\rho_{x_0x_1}^{\mathrm{ideal}}, \Lambda[\rho_{x_0x_1}]) = \mathrm{tr}(\Lambda[\rho_{x_0x_1}]\rho_{x_0x_1}^{\mathrm{ideal}})$, as the $\rho_{x_0x_1}^{\mathrm{ideal}}$ are pure states. We derive lower bounds on the smallest possible value of S given a value of $\mathcal{A}_2$, i.e.,

$$\mathcal{F}(\mathcal{A}_2) = \min_{\{\rho_{x_0x_1}\}\in R(\mathcal{A}_2)} S[\{\rho_{x_0x_1}\}]. \tag{9}$$

Note that this involves a minimization over all sets of four preparations $R(\mathcal{A}_2)$ that are compatible with an observed value $\mathcal{A}_2$.

In order to lower bound $\mathcal{F}$, we use an approach inspired by Ref. [22]. From Eq. (7), we have $\mathcal{A}_2 = \frac{1}{2} + \sum_{x_0,x_1}\mathrm{tr}(W_{x_0x_1}\rho_{x_0x_1})$, where $W_{x_0x_1} = \frac{1}{16}\sum_y(-1)^{x_y}M_y$. We define operators corresponding to some suitably chosen channel acting on the ideal preparations:

$$K_{x_0x_1}(M_0, M_1) = \Lambda^\dagger(M_0, M_1)\big[\rho_{x_0x_1}^{\mathrm{ideal}}\big], \tag{10}$$

where $\Lambda^\dagger$ is the channel dual to Λ. We aim to construct operator inequalities of the form

$$K_{x_0x_1}(M_0, M_1) \geqslant sW_{x_0x_1} + t_{x_0x_1}(M_0, M_1)\mathbb{1}, \tag{11}$$

for all inputs (x_0, x_1), for any given measurements, where s and $t_{x_0x_1}(M_0, M_1)$ are real coefficients. Finding such inequalities, as well as a suitable channel Λ, allows us to lower bound S as follows:

$$\begin{aligned} S &\geqslant \frac{1}{4}\sum_{x_0,x_1}\mathrm{tr}\left(K_{x_0x_1}\rho_{x_0x_1}\right) \geqslant \frac{s}{4}\sum_{x_0,x_1}\mathrm{tr}\left(W_{x_0x_1}\rho_{x_0x_1}\right) \\ &\quad + \frac{1}{4}\sum_{x_0,x_1}t_{x_0x_1} = \frac{s}{4}(\mathcal{A}_2 - 1/2) + \frac{1}{4}\sum_{x_0,x_1}t_{x_0x_1}. \end{aligned} \tag{12}$$

Applying a minimization over M_0 and M_1 to the right-hand side, the above inequality becomes valid for all preparations. Consequently,

$$\mathcal{F}(\mathcal{A}_2) \geqslant \frac{s}{4}(\mathcal{A}_2 - 1/2) + t \equiv L(\mathcal{A}_2), \tag{13}$$

where $t \equiv 1/4\,\min_{M_0,M_1}\sum_{x_0,x_1}t_{x_0x_1}(M_0, M_1)$. In Appendix B, we construct explicitly the channel and derive an operator inequality leading to a lower bound, given by $s = 4(1+\sqrt{2})$ and $t = (2-\sqrt{2})/4$.

This provides a robust self-testing for the preparations. A maximal value $\mathcal{A}_2 = Q_2$ implies $\mathcal{F} = 1$, i.e., the preparations must be the ideal ones (up to a unitary). For $\mathcal{A}_2 = C_2$, i.e.,

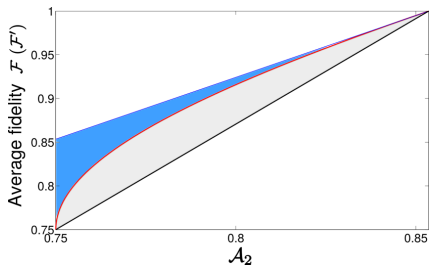

FIG. 1. Average fidelity $\mathcal{F}$ ($\mathcal{F}'$) for prepared states (measurements), as a function of the observed value of $\mathcal{A}_2$. The black line is our analytical lower bound of Eq. (13). The blue region is accessible via single qubit strategies without shared randomness, as confirmed by strong numerical evidence (see Appendixes). When allowing for shared randomness between the devices, the accessible region (obtained by taking the convex hull of the blue region) now also includes the gray area, and our analytic lower bound is tight in general.

a maximal value given a set of classical states, we get that $\mathcal{F} \geqslant 3/4$. This bound can be attained via the set of pure states $\rho_{x_0x_1} = [\mathbb{1} + (-1)^{x_0x_1}\sigma_z]/2$ (diagonal in the same basis, hence classical), combined with the measurements $M_0 = M_1 = \sigma_z$. Therefore, we see that our bound $\mathcal{F}(\mathcal{A}_2) \geqslant L(\mathcal{A}_2)$ is optimal, as far as linear inequalities are concerned (see Fig. 1). It is then interesting to consider the intermediate region $C_2 < \mathcal{A}_2 < Q_2$. First, focusing on strategies involving a single set of states and measurements, we observe numerically that the linear bound $\mathcal{F}(\mathcal{A}_2) \geqslant L(\mathcal{A}_2)$ cannot be saturated anymore, and conjecture the form of optimal states and measurements; see red curve in Fig. 1 and Appendix C for details. Second, allowing for shared randomness between the preparation and measurement device (such that convex combinations of qubit strategies are now possible), the linear bound becomes tight, a direct consequence of the linearity of $\mathcal{F}$ and $\mathcal{A}_2$ in terms of the states and measurements.

VII. ROBUST SELF-TESTING OF THE MEASUREMENTS

Similarly, we can quantify the average fidelity of the measurements with respect to the ideal ones: $S'(\{M_y^b\}) = \max_\Lambda \sum_{y,b} F((M_y^b)^{\text{ideal}}, \Lambda[M_y^b])/4$, where Λ must be a unital channel (i.e., mapping the identity to itself), in order to ensure that measurements are mapped to measurements. In analogy with the case of preparations, our goal is to lower bound the following quantity:

$$\mathcal{F}'(\mathcal{A}_2) = \min_{\{M_y^b\} \in R'(\mathcal{A}_2)} S'\left(\left\{M_y^b\right\}\right), \tag{14}$$

where $R'(\mathcal{A}_2)$ represents all sets of measurements compatible with a certain value of $\mathcal{A}_2$.

We first rewrite $\mathcal{A}_2 = \sum_{y,b} \text{tr}(M_y^b Z_{yb})$, where $Z_{yb} = \frac{1}{8}\sum_{x_0,x_1} \rho_{x_0x_1}\delta_{b,x_y}$. Next, we construct operator inequalities

$$K_{yb}(\{\rho_{x_0x_1}\}) \geqslant s Z_{yb} + t_y(\{\rho_{x_0x_1}\})\mathbb{1}, \tag{15}$$

given the unital channel $K_{yb} = \Lambda^\dagger[(M_y^b)^{\text{ideal}}]$. Similar to the case of preparations, strong operator inequalities can be derived by choosing carefully the channel; all details are given in Appendix D. Finally, this leads to a lower bound on the average fidelity

$$\mathcal{F}'(\mathcal{A}_2) \geqslant \min_{\{\rho_{x_0x_1}\}} \frac{1}{4}\sum_{y,b} \text{tr}\left(K_{yb}M_y^b\right) \geqslant L(\mathcal{A}_2). \tag{16}$$

That is, we find that $\mathcal{F}'$ can be lower bounded by a linear expression in terms of $\mathcal{A}_2$, which turns out to be the same as for the case of preparations.

This provides a robust self-test for the measurements. Observing $\mathcal{A}_2 = Q_2$ implies that $\mathcal{F}' = 1$; hence the measurements are equivalent to the ideal ones (up to a unitary). For $\mathcal{A}_2 = C_2$, we have that $\mathcal{F} \geqslant 3/4$. This lower bound can be attained by choosing $M_0 = \sigma_z$ and $M_1 = \mathbb{1}$, with the states $\rho_{00} = \rho_{01} = (\mathbb{1} + \sigma_z)/2$ and $\rho_{10} = \rho_{11} = (\mathbb{1} - \sigma_z)/2$. For $C_2 < \mathcal{A}_2 < Q_2$, we find numerically that the inequality (16) cannot be saturated using a single set of measurements and states (see Fig. 1). Details, in particular a conjecture for the form of the optimal measurements, are given in Appendix C. Similarly as for the case of states, when allowing for convex combinations of qubit strategies, our linear bound is tight.

VIII. GENERALIZATIONS

The above results can be generalized in several directions. First, a generalization of the $2 \to 1$ RAC enables self-testing of any pair of incompatible Pauli observables (see Appendix E). Secondly, we consider the $N \to 1$ RAC, where the preparation device receives as input an N-bit string $x = (x_1, \ldots, x_N)$ and the measurement device gets input $y \in \{1, \ldots, N\}$. The average score is then given by

$$\mathcal{A}_N = \frac{1}{N2^N}\sum_{x,y} P(b = x_y|x, y). \tag{17}$$

The methods discussed above (for $N = 2$) can be generalized and lead to self-testing conditions for states and measurements; details are given in Appendix F. The case of $N = 3$ is of particular interest. Here, the best possible score with qubits is $\mathcal{A}_3 = (1 + 1/\sqrt{3})/2$; see, e.g., Ref. [39]. In this case, our self-testing conditions can certify that (i) the eight prepared states correspond to Bloch vectors forming a cube on the Bloch sphere and (ii) the measurements correspond to three mutually unbiased bases (i.e., three pairwise anticommuting Pauli observables). Thirdly, we self-test qutrit preparations and projective measurements in the $2 \to 1$ RAC (see Appendix G).

Finally, we present a numerical method for robust self-testing of preparations applicable in scenarios beyond RACs. The method is based on semidefinite programing and combines (i) the swap method [21] used for self-testing in Bell scenarios with (ii) the hierarchy of finite-dimensional quantum correlations [41–43]. The idea is to first construct a swap operator, based on the measurement operators, which maps the state of the preparation onto an ancilla. The average fidelity between the ancilla and the ideal states can then be expressed in terms of strings of products of measurement operators and the extracted states. The last step is to miminize this average fidelity over all quantum realizations that are compatible with a given witness value, using the hierarchy of Refs. [41–43].

Although typically returning suboptimal bounds on $\mathcal{F}$, this method is widely applicable. In Appendix H, we describe in detail the methodology and apply to two examples, including the $2 \to 1$ RAC.

IX. OUTLOOK

We presented methods for self-testing quantum states and measurements in the prepare-and-measure scenario. These techniques demonstrate strong robustness to noise, and should therefore be directly amenable to experiments, providing useful certification techniques in a semi-device-independent setting. Moreover, these ideas should find applications in quantum communications. Our methods apply to the states and measurements used in QKD (e.g., in BB84), as well as in semi-device-independent QKD and randomness generation protocols [29–33].

It would be interesting to develop robust self-testing techniques for more general scenarios, e.g., for higher-dimensional quantum systems. Another direction would be to consider scenarios beyond prepare and measure, for instance, adding between the preparation and measurement devices a transformation device [44,45] and self-testing the latter.

Finally, while we have focused here on self-testing based on an assumption on the dimension, one could develop methods based on different assumptions, such as a bound on the mean energy [46], the overlap [47], or the entropy [48].

ACKNOWLEDGMENTS

This work was supported by the Swiss National Science Foundation (Starting grant DIAQ, QSIT, and Early Postdoc Mobility fellowship No. P2GEP2_162060), the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Action ROSETTA (Grant No. 749316), the European Research Council (Grant No. 337603), the Danish Council for Independent Research (Sapere Aude), VILLUM FONDEN via the QMATH Centre of Excellence (Grant No. 10059), and the National Research, Development and Innovation Office NKFIH (Grants No. K111734 and No. KH125096).

APPENDIX A: SELF-TESTING RELATIONS FOR PREPARATIONS AND MEASUREMENTS

In this section we provide a simple example of preparations that saturate the compatibility bound for $\mathcal{A}_2$ given in the main text. Moreover, we derive the upper bound for compatibility of measurements given in the main text.

First, we consider the case of preparations. Consider preparations such that ρ_{00} and ρ_{11}, and ρ_{01} and ρ_{10} correspond to antipodal Bloch vectors with a relative angle θ, the maximal quantum value of $\mathcal{A}_2$, is obtained from

$$\mathcal{A}_2 = \frac{1}{2} + \frac{1}{8}\sum_y \lambda_{\max}[V_y], \tag{A1}$$

where $V_y = \sum_{x_0,x_1}(-1)^{x_y}\rho_{x_0x_1}$. We represent the preparations on the Bloch sphere as $\rho_{x_0x_1} = 1/2(\mathbb{1} + \vec{m}_{x_0x_1}\cdot\vec{\sigma})$, where $\vec{m}_{00} = [\cos(\theta/2), 0, \sin(\theta/2)]$ and $\vec{m}_{01} = [\cos(\theta/2), 0, -\sin(\theta/2)]$, with $\vec{m}_{11} = -\vec{m}_{00}$ and $\vec{m}_{10} = -\vec{m}_{01}$. This gives $V_0 = 2\cos(\theta/2)\sigma_x$ and $V_1 = 2\sin(\theta/2)\sigma_z$. The respective largest eigenvalues are $\lambda_{\max}[V_0] = 2\cos(\theta/2)$ and $\lambda_{\max}[V_1] = 2\sin(\theta/2)$, leading to

$$\mathcal{A}_2 = \frac{1}{2} + \frac{1}{4\sqrt{2}}[\sqrt{1+\cos\theta} + \sqrt{1-\cos\theta}]. \tag{A2}$$

It is straightforward to see that this achieves the upper bound in the main text; indeed the above choice of preparations leads to $\beta = 4$ and $\alpha = 4\cos\theta$.

In order to derive the upper bound on $\mathcal{A}_2$ for compatibility of measurements in the main text we evaluate

$$\sum_{x_0,x_1}\lambda_{\max}[(-1)^{x_0}M_0 + (-1)^{x_1}M_1] \tag{A3}$$

for arbitrary qubit observables M_0, M_1. We take advantage of the fact that

$$\lambda_{\max}[T] + \lambda_{\max}[-T] = \lambda_{\max}[T] - \lambda_{\min}[T], \tag{A4}$$

which for a 2×2 matrix can be evaluated analytically. More specifically, if T is a 2×2 Hermitian matrix with eigenvalues $\lambda_0 \geqslant \lambda_1$, let

$$\chi := \operatorname{tr} T = \lambda_0 + \lambda_1,$$
$$\zeta := \operatorname{tr} T^2 = \lambda_0^2 + \lambda_1^2$$

and then

$$\lambda_0 - \lambda_1 = \sqrt{2\zeta - \chi^2}. \tag{A5}$$

Evaluating this expression for $T = M_0 \pm M_1$ gives the desired upper bound.

APPENDIX B: OPERATOR INEQUALITIES FOR ROBUST SELF-TESTING OF PREPARATIONS

In this section we provide a detailed derivation of the lower bound on the average fidelity $\mathcal{F}(\mathcal{A}_2)$. For a real constant $s > 0$, to be chosen later, consider for each pair (x_0, x_1) the operator $K_{x_0x_1} - sW_{x_0x_1}$, where $W_{x_0x_1} = \frac{1}{16}\sum_y(-1)^{x_y}M_y$ and $K_{x_0x_1} = \Lambda^\dagger[\rho^{\text{ideal}}_{x_0x_1}]$, for some channel Λ. Suppose now that $t_{x_0x_1} \in \mathbb{R}$ is a lower bound on its eigenvalues, or, equivalently, that the operator inequality

$$K_{x_0x_1} \geqslant s\, W_{x_0x_1} + t_{x_0x_1}\,\mathbb{1} \tag{B1}$$

holds. Then, computing the trace of this inequality with $\rho_{x_0x_1}$ and averaging over inputs leads to

$$S \geqslant \frac{1}{4}\sum_{x_0,x_1} F\big(\rho^{\text{ideal}}_{x_0x_1}, \Lambda[\rho_{x_0x_1}]\big) \geqslant \frac{s}{4}\left(\mathcal{A}_2 - \frac{1}{2}\right) + t,$$
$$t \equiv \frac{1}{4}\sum_{x_0,x_1} t_{x_0x_1}, \tag{B2}$$

where the first inequality holds because S is defined as maximization over all possible channels, and the Λ used there is one possible choice. In turn, if (B1) holds *as an operator inequality*, it is valid for any set of preparations $\{\rho_{x_0x_1}\}$, and thus $\mathcal{F}(\mathcal{A}_2) \geqslant \frac{s}{4}(\mathcal{A}_2 - \frac{1}{2}) + t$. Note that (B1) has a dependence on M_0, M_1 through $W_{x_0x_1}$. If (B1) holds for a particular choice of measurement operators M_0, M_1, then the bound on $\mathcal{F}(\mathcal{A}_2)$ holds for all preparations, for that particular choice of M_0, M_1. However, if (B1) holds for *all* possible

M_0, M_1, then the bound on $\mathcal{F}(\mathcal{A}_2)$ is valid for all quantum setups and is thus a robust self-testing inequality. To derive the appropriate constants s and $t_{x_0x_1}$, we first allow $t_{x_0x_1}$ and Λ to have a dependence on M_0 and M_1. We then minimize over M_0 and M_1 the constants $t_{x_0x_1}$, for a suitable choice of s, such that, at the end, Eq. (B1) holds regardless of the choice of measurement operators.

We choose a dephasing channel of the form

$$\Lambda_\theta(\rho) = \frac{1+c(\theta)}{2}\rho + \frac{1-c(\theta)}{2}\Gamma(\theta)\rho\Gamma(\theta), \tag{B3}$$

where for $0 \leqslant \theta \leqslant \pi/4$ we use $\Gamma = \sigma_x$, while for $\pi/4 < \theta \leqslant \pi/2$ we use $\Gamma = \sigma_z$. The function $c(\theta) \in [-1, 1]$ will be specified later.

In the interval $0 \leqslant \theta \leqslant \pi/4$, the action of the channel leads to

$$\begin{aligned} K_{00} &= \frac{\mathbb{1}+\sigma_x}{2}, \quad K_{01} = \frac{\mathbb{1}+c(\theta)\sigma_z}{2}, \\ K_{10} &= \frac{\mathbb{1}-c(\theta)\sigma_z}{2}, \quad K_{11} = \frac{\mathbb{1}-\sigma_x}{2}, \end{aligned} \tag{B4}$$

whereas in the interval $\pi/4 < \theta \leqslant \pi/2$, we have

$$\begin{aligned} K_{00} &= \frac{\mathbb{1}+c(\theta)\sigma_x}{2}, \quad K_{01} = \frac{\mathbb{1}+\sigma_z}{2}, \\ K_{10} &= \frac{\mathbb{1}-\sigma_z}{2}, \quad K_{11} = \frac{\mathbb{1}-c(\theta)\sigma_x}{2}. \end{aligned} \tag{B5}$$

As discussed in the main text, for any given set of preparations, the optimal measurements are projective and rank-one. Furthermore, any two such measurements can be represented on an equator of the Bloch sphere. Due to the freedom of setting the reference frame, we can without loss of generality represent the two measurements in the xz plane, i.e.,

$$\begin{aligned} M_0 &= \cos\theta\ \sigma_x + \sin\theta\ \sigma_z, \\ M_1 &= \cos\theta\ \sigma_x - \sin\theta\ \sigma_z. \end{aligned}$$

We can therefore write $W_{x_0x_1}$ as

$$\begin{aligned} W_{00} &= \tfrac{1}{8}\cos\theta\sigma_x, \quad W_{01} = \tfrac{1}{8}\sin\theta\sigma_z, \\ W_{10} &= -\tfrac{1}{8}\sin\theta\sigma_z, \quad W_{11} = -\tfrac{1}{8}\cos\theta\sigma_x. \end{aligned} \tag{B6}$$

We can reduce the number of operator inequalities (B1) by exploiting the apparent symmetries in the expressions for $W_{x_0x_1}$ and $K_{x_0x_1}$: we restrict ourselves so that $t_o \equiv t_{01} = t_{10}$ and $t_e \equiv t_{00} = t_{11}$. Thus we have to consider two operator inequalities in each interval $\theta \in [0, \pi/4]$ and $\theta \in (\pi/4, \pi/2]$. In the first interval, the two operator inequalities are

$$\begin{aligned} \frac{1+\sigma_x}{2} - \frac{s}{8}\cos\theta\sigma_x - t_e\mathbb{1} &\geqslant 0, \\ \frac{1+c(\theta)\sigma_z}{2} - \frac{s}{8}\sin\theta\sigma_z - t_o\mathbb{1} &\geqslant 0. \end{aligned} \tag{B7}$$

In the second interval, the two operator inequalities are

$$\begin{aligned} \frac{1+c(\theta)\sigma_x}{2} - \frac{s}{8}\cos\theta\sigma_x - t_e\mathbb{1} &\geqslant 0, \\ \frac{1+\sigma_z}{2} - \frac{s}{8}\sin\theta\sigma_z - t_o\mathbb{1} &\geqslant 0. \end{aligned} \tag{B8}$$

We now focus on the former interval. Solving the two inequalities for t_o and t_e we obtain

$$t_e \leqslant 1 - \frac{s}{8}\cos\theta, \quad t_o \leqslant \frac{1}{8}[4 + 4c(\theta) - s\ \sin\theta], \tag{B9}$$

$$t_e \leqslant \frac{s}{8}\cos\theta, \quad t_o \leqslant \frac{1}{8}[4 - 4c(\theta) + s\ \sin\theta]. \tag{B10}$$

Any choice of t_o and t_e satisfying these constraints gives rise to valid operator inequalities. In order to obtain the strongest bound, we choose the largest values of t_o and t_e consistent with their respective constraints, i.e.,

$$\begin{aligned} t_e &= \min\left\{1 - \frac{s}{8}\cos\theta, \frac{s}{8}\cos\theta\right\}, \\ t_o &= \min\left\{\frac{1}{8}[4 + 4c(\theta) - s\sin\theta], \frac{1}{8}[4 - 4c(\theta) + s\sin\theta]\right\}. \end{aligned} \tag{B11}$$

A similar procedure for the interval $\theta \in (\pi/4, \pi/2]$ leads to

$$\begin{aligned} t_e &= \min\left\{\frac{1}{8}[4 + 4c(\theta) - s\ \cos\theta], \frac{1}{8}[4 - 4c(\theta) + s\ \cos\theta]\right\}, \\ t_o &= \min\left\{1 - \frac{s}{8}\sin\theta, \frac{s}{8}\sin\theta\right\}. \end{aligned} \tag{B12}$$

It is worth pointing out that the two intervals only differ by exchanging $t_e \leftrightarrow t_o$ and $\sin\theta \leftrightarrow \cos\theta$. Hence, for any given θ, we have constructed operator inequalities of the form (B1).

As shown in the main text, we obtain our lower bound on the average fidelity from

$$\mathcal{F}(\mathcal{A}_2) \geqslant \frac{s}{4}(\mathcal{A}_2 - 1/2) + \min_{M_0, M_1} t(M_0, M_1) \equiv L(\mathcal{A}_2), \tag{B13}$$

where $t(M_0, M_1) = (t_e + t_o)/2$. To compute this quantity we fix the value of s to be

$$s = 4(1 + \sqrt{2}) \tag{B14}$$

and choose the dephasing function as $c(\theta) = \min\{1, \frac{s}{4}\sin\theta\}$ whenever $\theta \in [0, \pi/4]$ and $c(\theta) = \min\{1, \frac{s}{4}\cos\theta\}$ whenever $\theta \in (\pi/4, \pi/2]$. It is easy to see that $c(\theta) \in [0, 1]$, which ensures that Λ_θ is a valid quantum channel, and that $c(\theta)$ is continuous at $\theta = \pi/4$. A simple calculation shows that in this case

$$t = \frac{2 - \sqrt{2}}{4}, \tag{B15}$$

which gives the lower bound

$$\mathcal{F}(\mathcal{A}_2) \geqslant (1 + \sqrt{2})\mathcal{A}_2 - \frac{3}{2\sqrt{2}} \equiv L(\mathcal{A}_2). \tag{B16}$$

One can check that choosing distinct values of s will not lead to improved lower bounds.

APPENDIX C: TIGHTNESS OF FIDELITY BOUNDS

In the main text, we have derived fidelity bounds for both the preparations and the measurements, based on operator

inequalities. Specifically, we obtain a lower bound on the average fidelity $\mathcal{F}$ of the prepared states (with respect to the ideal ones) given by the linear expression

$$\mathcal{F}(\mathcal{A}_2) \geqslant (1+\sqrt{2})\mathcal{A}_2 - \frac{3}{2\sqrt{2}} \equiv L(\mathcal{A}_2). \tag{C1}$$

For measurements, a similar bound is obtained on the average fidelity $\mathcal{F}'$ with respect to the ideal ones. In the present appendix, we discuss the tightness of these bounds.

We start with our bound on the fidelity of the states. As discussed in the main text, obtaining $\mathcal{A}_2 = Q_2$ implies $\mathcal{F} = 1$, i.e., the states are the ideal ones (up to a unitary). Let us refer to the optimal strategy (with the ideal states) as strategy S_1. Then, for $\mathcal{A}_2 = C_2$, our bound gives $\mathcal{F} \geqslant 3/4$. This bound is tight and can be obtained via the set of pure states $\rho_{x_0x_1} = [\mathbb{1} + (-1)^{x_0,x_1}\sigma_z]/2$ (diagonal in the same basis, hence classical), combined with the measurements $M_0 = M_1 = \sigma_z$. Let us refer to this strategy as S_2.

The above shows that our bound (C1) is tight as far as linear inequalities are concerned. More generally, the bound is in fact tight in general, when shared randomness between the preparation and measurement devices is taken into account. In this case, taking a convex combination between strategies S_1 and S_2 allows us to get any point on the line (i.e., pair of values $\mathcal{F}$ and $\mathcal{A}_2$) between S_1 and S_2.

It is also interesting to understand what happens when shared randomness between the devices is not taken into account. In this case, the end points ($\mathcal{A}_2 = Q_2, \mathcal{F} = 1$) and ($\mathcal{A}_2 = C_2, \mathcal{F} = 3/4$) can still be obtained. To understand what happens in the intermediate region $C_2 < \mathcal{A}_2 < Q_2$, we first performed a numerical analysis. Specifically, we choose randomly four qubit states, and compute (i) the maximal value of $\mathcal{A}_2$ (optimizing over the measurements) and (ii) the average fidelity $\mathcal{F}$ (where the optimization over channels is restricted here to unitaries). The resulting points are shown on Fig. 2 (blue circles). This indicates that, for $C_2 < \mathcal{A}_2 < Q_2$, the bound (C1) cannot be saturated anymore. Moreover, we conjecture that an optimal class of strategies is given by the pure states

$$|\psi_{00}\rangle = |0\rangle, \quad |\psi_{11}\rangle = |1\rangle, \quad |\psi_{01}\rangle = \cos\theta|0\rangle + \sin\theta|1\rangle,$$
$$|\psi_{10}\rangle = \cos\theta|0\rangle - \sin\theta|1\rangle \tag{C2}$$

and the measurements $M_y = \cos(\varphi)\sigma_z + (-1)^y \sin(\varphi)\sigma_x$. Straightforward calculations show that taking $\tan\varphi = \sin 2\theta$ leads to

$$\mathcal{A}_2 = \tfrac{1}{2} + \tfrac{1}{4}\sqrt{1+\tan^2(\varphi)}, \quad \mathcal{F} = \tfrac{1}{4}(3+\tan\varphi). \tag{C3}$$

This gives a parametric curve, as a function of $\varphi \in [0, \pi/4]$, given by the red curve in Fig. 2. This curve is in excellent agreement with the numerical results obtained before. Note that this class of strategies interpolates between the strategies S_1 (setting $\varphi = 0$) and S_2 (setting $\varphi = \pi/4$).

Next we discuss the bound on the average fidelity of measurements. As discussed in the main text, the linear bound $\mathcal{F}'(\mathcal{A}_2) \geqslant L(\mathcal{A}_2)$ is optimal as far as linear inequalities are concerned. Moreover, when allowing for shared randomness the bound is tight in general for $C_2 \leqslant \mathcal{A}_2 \leqslant Q_2$. This is obtained by considering convex combinations of strategy S_1' (defined as the optimal strategy S_1, up to a rotation of $\pi/8$ around the y axis; see below), and the following strategy (referred to as S_3): take $M_0 = \sigma_z$ and $M_1 = \mathbb{1}$, with the states $\rho_{00} = \rho_{01} = (\mathbb{1} + \sigma_z)/2$ and $\rho_{10} = \rho_{11} = (\mathbb{1} - \sigma_z)/2$.

Similar to the case of states, we now consider the situation where shared randomness between the devices is not allowed. Performing a numerical analysis similar to the one described above (except that measurements are now generated randomly), we observe that the accessible region (in terms of $\mathcal{F}'$ vs $\mathcal{A}_2$) appears to be exactly the same as for the case of states (i.e., the blue region in Fig. 2). We conjecture that the lower bound is given by the following class of optimal strategies: take the measurements

$$M_0 = \sigma_z, \quad M_1 = \eta\sigma_x + (1-\eta)\mathbb{1}, \tag{C4}$$

with the states $|\psi_{00}\rangle = \cos\theta|0\rangle + \sin\theta|1\rangle$, $|\psi_{01}\rangle = \cos\theta|0\rangle - \sin\theta|1\rangle$, $|\psi_{10}\rangle = \cos\theta|1\rangle + \sin\theta|0\rangle$, and $|\psi_{11}\rangle = \cos\theta|1\rangle - \sin\theta|0\rangle$. Setting $\eta = \tan 2\theta$, we get

$$\mathcal{A}_2 = \frac{\cos^2(\theta)}{2} + \frac{1}{4} + \frac{\sin^2(2\theta)}{\cos(2\theta)}, \quad \mathcal{F} = \frac{1}{4}[3 + \tan(2\theta)]. \tag{C5}$$

This gives a parametric curve, as a function of $\theta \in [0, \pi/8]$, given by the red curve in Fig. 2. This curve is in excellent agreement with the numerical results obtained before. Also, this curve turns out to be exactly the same as the curve we obtained above for the case of states. Note that this class of strategies interpolates between the strategies S_1' (setting $\theta = \pi/8$) and S_3 (setting $\theta = 0$).

Finally, note that the numerics also suggests that there is a linear upper bound on the average fidelities $\mathcal{F}$ ($\mathcal{F}'$) as a function of $\mathcal{A}_2$ (see Fig. 2); specifically $\mathcal{F} \leqslant \frac{1-Q_2}{Q_2-3/4}\mathcal{A}_2 +$

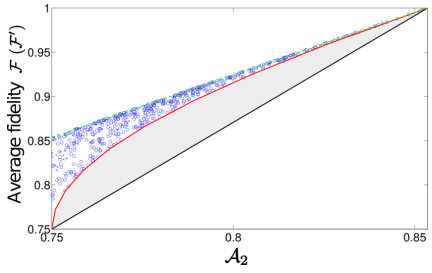

FIG. 2. Black line is the analytic lower bound on the average fidelity $\mathcal{F}$ ($\mathcal{F}'$) for prepared states (measurements), as a function of the observed value of $\mathcal{A}_2$. To characterize the region accessible via pure qubit strategies (i.e., without shared randomness), we perform numerics generating randomly sets of qubit preparations (blue circles and crosses); here we show the numerical results for the case of states, but similar results are obtained for the case of measurements. In the region $C_2 < \mathcal{A}_2 < Q_2$, we conjecture that the class of strategies given in the text (corresponding to the red curve) are optimal, both for $\mathcal{F}$ and $\mathcal{F}'$. Finally, the green dashed line is our conjectured upper bound on the average fidelity.

$\frac{Q_2^2-3/4}{Q_2-3/4}$ and similarly for $\mathcal{F}'$. It would be interesting to provide a proof of these upper bounds.

APPENDIX D: OPERATOR INEQUALITIES FOR ROBUST SELF-TESTING OF MEASUREMENTS

In this section, we account for the detailed derivation of the lower bound on the average fidelity of the measurements $\mathcal{F}'(\mathcal{A}_2)$. The approach bears significant resemblance to the case of robustly self-testing preparations, as outlined in Appendix B.

We aim to derive operator inequalities of the form

$$K_{yb}(\{\rho_{x_0x_1}\}) \geqslant s Z_{yb} + t_y(\{\rho_{x_0x_1}\})\mathbb{1}, \tag{D1}$$

where $Z_{yb} = \frac{1}{8}\sum_{x_0,x_1}\rho_{x_0x_1}\delta_{b,x_y}$ and $K_{yb}(\{\rho_{x_0x_1}\}) = \Lambda^\dagger[(M_y^b)^{\text{ideal}}]$. For the sake of simplicity, we first apply a unitary channel to $(M_y^b)^{\text{ideal}}$ to align these operators with the eigenstates of σ_x and σ_z. Then, we adopt the same (unital, trace-preserving) channel Λ as specified in the main text, with the same coefficients as used to robustly self-test the preparations: $c(\theta) = \min\{1, \frac{s}{4}\sin\theta\}$ when $\theta \in [0, \pi/4]$ and $c(\theta) = \min\{1, \frac{s}{4}\cos\theta\}$ when $\theta \in (\pi/4, \pi/2]$.

It is straightforward to see that, for any given pair of measurements, the optimal choice of preparations are four pure qubit states, such that ρ_{00} and ρ_{11}, and ρ_{01} and ρ_{10}, respectively, correspond to antipodal vectors on the Bloch sphere. Therefore, we can without loss of generality restrict to such preparations since these impose the weakest constraints on the measurements of our interest. We can therefore parametrize the preparations $\rho_{x_0x_1} = 1/2(\mathbb{1} + \vec{m}_{x_0x_1}\cdot\vec{\sigma})$ by Bloch vectors

$$\begin{aligned}\vec{m}_{00} &= [\cos\theta, 0, \sin\theta], \quad \vec{m}_{11} = -[\cos\theta, 0, \sin\theta],\\ \vec{m}_{01} &= [\cos\theta, 0, -\sin\theta], \quad \vec{m}_{10} = [-\cos\theta, 0, \sin\theta].\end{aligned} \tag{D2}$$

Expressing Z_{yb} in terms of these preparations gives

$$\begin{aligned}Z_{00} &= \tfrac{1}{8}(\mathbb{1} + \cos\theta\sigma_x), \quad Z_{01} = \tfrac{1}{8}(\mathbb{1} - \cos\theta\sigma_x),\\ Z_{10} &= \tfrac{1}{8}(\mathbb{1} + \sin\theta\sigma_z)\sigma_z, \quad Z_{11} = \tfrac{1}{8}(\mathbb{1} - \sin\theta\sigma_z).\end{aligned} \tag{D3}$$

Due to symmetries, we restrict ourselves so that $t_o \equiv t_{01} = t_{10}$ and $t_e \equiv t_{00} = t_{11}$. Thus we have to consider two operator inequalities in each interval $\theta \in [0, \pi/4]$ and $\theta \in (\pi/4, \pi/2]$. In the first interval, the two operator inequalities are

$$\begin{aligned}\frac{1+\sigma_x}{2} - \frac{s}{8}(\mathbb{1} + \cos\theta\sigma_x) - t_e\mathbb{1} &\geqslant 0,\\ \frac{1+c(\theta)\sigma_z}{2} - \frac{s}{8}(\mathbb{1} + \sin\theta\sigma_z) - t_o\mathbb{1} &\geqslant 0.\end{aligned} \tag{D4}$$

In the second interval, the two operator inequalities are

$$\begin{aligned}\frac{1+c(\theta)\sigma_x}{2} - \frac{s}{8}(\mathbb{1} + \cos\theta\sigma_x) - t_e\mathbb{1} &\geqslant 0,\\ \frac{1+\sigma_z}{2} - \frac{s}{8}(\mathbb{1} + \sin\theta\sigma_z) - t_o\mathbb{1} &\geqslant 0.\end{aligned} \tag{D5}$$

Just as in Appendix B, we solve these inequalities for t_e and t_o, and choose the largest value compatible with the solutions. In the first interval, this gives

$$\begin{aligned}t_e &= \min\left\{\frac{1}{8}(8 - s - s\cos\theta), \frac{s}{8}(\cos\theta - 1)\right\},\\ t_o &= \min\left\{\frac{1}{8}[4c(\theta) - s\sin\theta - s + 4],\right.\\ &\qquad \left.\frac{1}{8}[-4c(\theta) + s\sin\theta - s + 4]\right\}.\end{aligned} \tag{D6}$$

A similar procedure for the interval $\theta \in (\pi/4, \pi/2]$ leads to

$$\begin{aligned}t_e &= \min\left\{\frac{1}{8}[4c(\theta) - s\cos\theta - s + 4],\right.\\ &\qquad \left.\frac{1}{8}[-4c(\theta) + s\cos\theta - s + 4]\right\},\\ t_o &= \min\left\{\frac{s}{8}(\sin\theta - 1), \frac{1}{8}(8 - s - s\sin\theta)\right\}.\end{aligned} \tag{D7}$$

For any choice of θ, we have constructed operator inequalities of the form (D1).

In order to obtain our lower bound on $\mathcal{F}'$, we must minimise the quantity $t(\theta) = (t_e + t_o)/2$ for a specific choice of s. In analogy with the procedure in Appendix D, we choose $s = 4(1+\sqrt{2})$, which returns $\min_\theta t(\theta) = -3/(2\sqrt{2})$. Hence we have obtained the lower bound

$$\mathcal{F}'(\mathcal{A}_2) \geqslant (1+\sqrt{2})\mathcal{A}_2 - \frac{3}{2\sqrt{2}} = L(\mathcal{A}_2). \tag{D8}$$

APPENDIX E: SELF-TESTING ALL PAIRS OF INCOMPATIBLE PAULI OBSERVABLES

Consider a generalization of the $2 \to 1$ RAC, in which we introduce a bias on the score associated to certain inputs. Specifically, whenever the game is successful, i.e., $b = x_y$, the awarded score is $q/2$ if $x_0 \oplus x_1 = 0$, and $(1-q)/2$ if $x_0 \oplus x_1 = 1$, for some $q \in [0, 1]$. The average score reads

$$\mathcal{A}_2^q = \frac{1}{2}\sum_{x_0,x_1,y} r(x_0, x_1)P(b = x_y|x_0, x_1, y), \tag{E1}$$

where $r(x_0, x_1) = q/2$ if $x_0 \oplus x_1 = 0$ and $r(x_0, x_1) = (1-q)/2$ if $x_0 \oplus x_1 = 1$. Note that, for $q = 1/2$, we recover the standard $2 \to 1$ RAC. Based on the quantity $\mathcal{A}_2^q$, we will now see how to derive a self-testing condition for any pair of incompatible Pauli observables, i.e., any pair of noncommuting projective rank-one qubit measurements.

We start by expressing $\mathcal{A}_2^q$ for a quantum strategy:

$$\begin{aligned}\mathcal{A}_2^q &= \frac{1}{2} + \frac{1}{4}\sum_{x_0,x_1} r(x_0, x_1)\,\mathrm{tr}\{\rho_{x_0x_1}[(-1)^{x_0}M_0 + (-1)^{x_1}M_1]\}\\ &\leqslant \frac{1}{2} + \frac{1}{4}\sum_{x_0,x_1} r(x_0, x_1)\lambda_{\max}[(-1)^{x_0}M_0 + (-1)^{x_1}M_1].\end{aligned} \tag{E2}$$

Denoting $\mu_k = \lambda_{\min}[M_0 + (-1)^k M_1]$ and $\nu_k = \lambda_{\max}[M_0 + (-1)^k M_1]$, for $k = 0, 1$, we obtain

$$\mathcal{A}_2^q \leqslant \tfrac{1}{2} + \tfrac{1}{8}[q(\mu_0 - \nu_0) + (1-q)(\mu_1 - \nu_1)]. \tag{E3}$$

Following a derivation analogous to that appearing in Appendix A to obtain, we obtain

$$\mathcal{A}_2^q \leqslant \tfrac{1}{2} + \tfrac{1}{8}[q\sqrt{\beta+\alpha} + (1-q)\sqrt{\beta-\alpha}], \tag{E4}$$

where $\beta = 2\,\mathrm{tr}\,(M_0^2 + M_1^2) - \mathrm{tr}\,(M_0)^2 - \mathrm{tr}\,(M_1)^2$ and $\alpha = 2\,\mathrm{tr}\,(\{M_0, M_1\}) - 2\,\mathrm{tr}\,(M_0)\,\mathrm{tr}\,(M_1)$. Treating α and β as independent variables, we obtain the largest value of the right-hand side of Eq. (E4) by demanding that the derivative with respect to α equals zero, and checking that the second derivative is negative at this point. We obtain the optimality constraint

$$\alpha = \frac{2q-1}{1-2q+2q^2}\beta. \tag{E5}$$

Inserting this value back into Eq. (E4), we find an upper bound on $\mathcal{A}_2^q$ as obtained by independent variables α and β. It turns out that this bound can be saturated by the de facto coupled variables α and β. From Eq. (E4), it is clear that a necessary condition for optimality is to maximize β. This amounts to the observables M_0 and M_1 being traceless and such that $M_0^2 = M_1^2 = \mathbb{1}$, leading to $\beta = 8$. This implies that the observables represent projective rank-one measurements. Hence we can write $M_y = \vec{n}_y \cdot \vec{\sigma}$ where the Bloch vector satisfies $|\vec{n}_y| = 1$. Hence we have $\alpha = 8\vec{n}_0 \cdot \vec{n}_1$. Thus Eq. (E5) becomes

$$\vec{n}_0 \cdot \vec{n}_1 = \frac{2q-1}{1-2q+2q^2}, \tag{E6}$$

which has a solution for any choice of q. Note that setting $q = 1/2$ reduces the above to $\vec{n}_0 \cdot \vec{n}_1 = 0$, which we recognize as the optimality constraint for the standard $2 \to 1$ random access code. In conclusion, for any pair of incompatible Pauli observables (characterized by the scalar product $\vec{n}_0 \cdot \vec{n}_1$), we have a game $\mathcal{A}_2^q$ (where q is chosen in order to satisfy the above equation), such that the maximal score can only be attained by using that specific pair of Pauli observables. We thus obtain a general class of self-tests for any pair of Pauli observables, corresponding to saturating the maximal quantum value of $\mathcal{A}_2^q$ for a given value of q:

$$\mathcal{A}_2^q \leqslant \tfrac{1}{2}(1 + \sqrt{1-2q+2q^2}). \tag{E7}$$

APPENDIX F: SELF-TESTING FOR THE $N \to 1$ RANDOM ACCESS CODE

In this appendix, we extend the results presented in the main text to self-test the preparations and measurements in an $N \to 1$ RAC. The latter is a straightforward generalization of the $2 \to 1$ RAC considered in the main text. The input of the preparation device is a random N-bit string $x \equiv (x_1, \ldots, x_N)$, while the input of the measurement device is $y \in \{1, \ldots, N\}$. The average score is

$$\mathcal{A}_N = \frac{1}{N2^N} \sum_{x,y} P(b = x_y | x, y). \tag{F1}$$

Considering qubit states ρ_x, and measurement observables M_y, we get

$$\mathcal{A}_N = \frac{1}{2} + \frac{1}{N2^{N+1}} \sum_{x,y} (-1)^{x_y}\,\mathrm{tr}(\rho_x M_y). \tag{F2}$$

1. Compatibility of measurements

We determine whether a set of measurements can explain (i.e., are compatible with) a given value of $\mathcal{A}_N$. Since rank-one projective measurements are optimal for any set of preparations, we choose for simplicity to restrict our consideration to such measurements. However, it is straightforward to consider general measurements using the method outlined in the main text and Appendix A.

Specifically, we first write

$$\begin{aligned}\mathcal{A}_N &= \frac{1}{2} + \frac{1}{N2^{N+1}} \sum_x \mathrm{tr}\,(\rho_x W_x) \\ &\leqslant \frac{1}{2} + \frac{1}{N2^{N+1}} \sum_x \lambda_{\max}[W_x],\end{aligned} \tag{F3}$$

where $W_x = \sum_y (-1)^{x_y} M_y$.

Note $\lambda_{\max}[W_x] = \lambda_{\min}[W_{\bar{x}}]$, where $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_N)$ is the bit string obtained from x by flipping all bits. Thus it is sufficient to only calculate eigenvalues for the strings not obtainable from each other under a full bit-flip operation. To this end let $z = x_1 \ldots x_{N-1}, 0$ and $\lambda_{z,0}$ ($\lambda_{z,1}$) be the largest (smallest) eigenvalue of W_z. Thus we write

$$\mathcal{A}_N \leqslant \frac{1}{2} + \frac{1}{N2^{N+1}} \sum_z [\lambda_{z,0} - \lambda_{z,1}]. \tag{F4}$$

Since $\lambda_{z,0}^2$ and $\lambda_{z,1}^2$ are eigenvalues of W_z^2, we have $\lambda_{z,0}^2 + \lambda_{z,1}^2 = \mathrm{tr}\,(W_z^2)$, which is equivalent to

$$\lambda_{z,0}^2 + \lambda_{z,1}^2 = \sum_{y=1}^N \mathrm{tr}\left(M_y^2\right) + \sum_{k<l} (-1)^{z_k + z_l}\,\mathrm{tr}\,(\{M_k, M_l\}). \tag{F5}$$

This equation, together with the relation $(\lambda_{z,0} - \lambda_{z,1})^2 \leqslant 2(\lambda_{z,0}^2 + \lambda_{z,1}^2)$, imply that Eq. (F4) becomes

$$\begin{aligned}\mathcal{A}_N \leqslant \frac{1}{2} + \frac{\sqrt{2}}{N2^{N+1}} \sum_z \Bigg[&\sum_{y=1}^N \mathrm{tr}\left(M_y^2\right) \\ &+ \sum_{k<l} (-1)^{z_k + z_l}\,\mathrm{tr}\,(\{M_k, M_l\}) \Bigg]^{1/2}.\end{aligned} \tag{F6}$$

This provides a robust self-testing condition, allowing one to determine whether a given set of measurements is compatible with the observed value of $\mathcal{A}_N$. Furthermore, we can derive an upper bound on the maximal value of $\mathcal{A}_N$ by assuming (incorrectly for $N > 3$) that there exists N mutually unbiased bases in $\mathbb{C}^2$. This means that all measurements are maximally incompatible, i.e., that $\mathrm{tr}\,(\{M_k, M_l\}) = 0$ for $k \neq l$. Consequently, Eq. (F6) reduces to

$$\mathcal{A}_N \leqslant \frac{1}{2}\left(1 + \frac{1}{\sqrt{N}}\right). \tag{F7}$$

We emphasize that only three mutually unbiased bases exist in $\mathbb{C}^2$ and hence this bound is only tight for $N = 2, 3$. For $N = 2$, we recover the result presented in the main text. For $N = 3$, this implies that a maximal value of $\mathcal{A}_3$ (i.e., achieving the right-hand side of the above inequality) ensures that the

measurements are three mutually unbiased qubit observables, such as the three Pauli observables σ_x, σ_y, and σ_z.

Going one step further, we can then also self-test the preparations (still assuming maximal value of $\mathcal{A}_3$). Indeed, each preparation ρ_x must be pure, and correspond to the eigenvector of W_x associated to its largest eigenvalue. Such a set of preparations corresponds to a set of Bloch vectors forming a cube on the surface of the Bloch sphere.

2. Compatibility of preparations

We ask whether a given value of $\mathcal{A}_N$ can be explained by a particular set of preparations. We suitably express (F2) in a quantum model and subsequently apply the Cauchy-Schwarz inequality for operators to obtain

$$\mathcal{A}_N = \frac{1}{2} + \frac{1}{N2^N}\sum_{y=1}^{N} \mathrm{tr}\left[M_y^0 \sum_x (-1)^{x_y}\rho_x\right] \leqslant \frac{1}{2} + \frac{1}{N2^N}\sum_{y=1}^{N}\sqrt{\mathrm{tr}\left[M_y^0\left(\sum_x (-1)^{x_y}\rho_x\right)^2\right]}. \quad \text{(F8)}$$

In the last expression, the squared operator is evaluated to

$$\left(\sum_x (-1)^{x_y}\rho_x\right)^2 = \sum_x \rho_x^2 + \sum_{k<l}(-1)^{k_y+l_y}\{\rho_k, \rho_l\}. \quad \text{(F9)}$$

If necessary, the anticommutators can be evaluated using Bloch sphere representation with the relation $\{\rho_k, \rho_l\} = 1/2[(1 + \vec{m}_k \cdot \vec{m}_l)\mathbb{1} + (\vec{m}_k + \vec{m}_l)\cdot\vec{\sigma}]$. However, it is more convenient to consider a basis-independent representation. Importantly, note that since an equal number of positive and negative terms appear inside the square, the operator $\sum_x (-1)^{x_y}\rho_x$ is a linear combination of $\{\sigma_x, \sigma_y, \sigma_z\}$ and hence its square is proportional to the identity operator. Therefore, when reinserting Eq. (F9) into Eq. (F8), we find

$$\mathcal{A}_N \leqslant \frac{1}{2} + \frac{1}{N2^N}\sum_{y=1}^{N}\left[\sum_x \mathrm{tr}\left(\rho_x^2\right) + \sum_{k<l}(-1)^{k_y+l_y}\,\mathrm{tr}\left(\{\rho_k, \rho_l\}\right)\right]^{1/2}. \quad \text{(F10)}$$

This is a self-testing condition for preparations, assessing whether a given set of preparations is compatible with a given value of $\mathcal{A}_N$. In particular, a classical strategy in which the preparations are binary messages corresponds to $\forall x$: $\mathrm{tr}\,(\rho_x^2) = 1$ and $\mathrm{tr}\,(\{\rho_k, \rho_l\}) = 2\delta_{E(k),E(l)}$, where E is the specific classical encoding strategy, i.e., a function $E : \{0, 1\}^N \to \{0, 1\}$.

APPENDIX G: SELF-TESTING WITH THREE-LEVEL SYSTEMS

In the main text, we have considered self-testing in the $2 \to 1$ random access code when the physical system transmitted from Alice to Bob is a qubit. Clearly, if that system is allowed to carry two bits of information, the task is trivial since Alice can send both her inputs to Bob. Here, we consider the remaining nontrivial case of Alice communicating a three-level quantum system. To simplify the analysis we restrict ourselves to projective measurements for which all possible arrangements admit a compact characterization. We show that the optimal quantum value equals $\mathcal{A}_2 = (5 + \sqrt{5})/8 \approx 0.9045$ and find all the optimal arrangements of observables (we argue that the optimal value is achieved only if both measurements are projective). Our argument is robust in the sense that we are able to certify incompatibility of M_0 and M_1 whenever the success probability exceeds the classical bound for three-level systems, which turns out to be $\mathcal{A}_2 \leqslant 7/8$.

To obtain a statement which only depends on the observables we follow the main text and evaluate the sum

$$\sum_{x_0,x_1}\lambda_{\max}[(-1)^{x_0}M_0 + (-1)^{x_1}M_1]. \quad \text{(G1)}$$

Jordan's lemma states that any two projective observables can be simultaneously diagonalized such that the resulting blocks are 1×1 or 2×2. For observables acting on a qutrit, we only need to consider two cases: (a) three one-dimensional subspaces or (b) one subspace of each type. Case (a) corresponds to classical strategies and it is easy to check that these satisfy $\mathcal{A}_2 \leqslant 7/8$. In case (b) the observables (up to a unitary) can be written as

$$M_0 = \begin{pmatrix} \cos\alpha\,\sigma_x + \sin\alpha\,\sigma_z & \\ & r \end{pmatrix}, \quad M_1 = \begin{pmatrix} \cos\alpha\,\sigma_x - \sin\alpha\,\sigma_z & \\ & s \end{pmatrix} \quad \text{(G2)}$$

for some angle $\alpha \in [0, 2\pi]$ and $r, s \in \{\pm 1\}$. A simple calculation yields

$$\begin{aligned}\lambda_{\max}[M_0 + M_1] &= \max\{2|\cos\alpha|, r + s\},\\ \lambda_{\max}[M_0 - M_1] &= \max\{2|\sin\alpha|, r - s\},\\ \lambda_{\max}[-M_0 + M_1] &= \max\{2|\sin\alpha|, -r + s\},\\ \lambda_{\max}[-M_0 - M_1] &= \max\{2|\cos\alpha|, -r - s\}\end{aligned}$$

and, therefore,

$$\sum_{x_0,x_1}\lambda_{\max}[(-1)^{x_0}M_0 + (-1)^{x_1}M_1] = \begin{cases} 2 + 4|\sin\alpha| + 2|\cos\alpha| & \text{if } r = s,\\ 2 + 2|\sin\alpha| + 4|\cos\alpha| & \text{if } r \neq s.\end{cases} \quad \text{(G3)}$$

For $r = s$ the right-hand side is maximized for $\alpha \in \{c_1, c_1 + \pi, -c_1 + \pi, -c_1 + 2\pi\}$, where c_1 is the unique solution to $\tan c_1 = 2$ in the interval $[0, \pi/2]$. Similarly, for $r \neq s$ the right-hand side is maximized for $\alpha \in \{c_2, c_2 + \pi, -c_2 + \pi, -c_2 + 2\pi\}$, where c_2 is the unique solution to $\tan c_2 = 1/2$ in the interval $[0, \pi/2]$.

While the different optimal arrangements are not unitarily equivalent, they are of similar form. The optimal arrangement characterized by $r = s = 1$ and $\alpha = c_1$ yields the following

optimal preparations:

$$\rho_{00} = \begin{pmatrix} 0 & \\ & 1 \end{pmatrix}, \quad \rho_{01} = \begin{pmatrix} (\mathbb{1}+\sigma_z)/2 & \\ & 0 \end{pmatrix},$$
$$\rho_{10} = \begin{pmatrix} (\mathbb{1}-\sigma_z)/2 & \\ & 0 \end{pmatrix}, \quad \rho_{11} = \begin{pmatrix} (\mathbb{1}-\sigma_x)/2 & \\ & 0 \end{pmatrix}. \tag{G4}$$

Indeed, it is always the case that one preparation lives in the 1×1 subspace, whereas the other three occupy the 2×2 subspace (two of them form a basis to which the last one is unbiased). To see that the optimal winning probability requires projective measurements, note that for every set of preparations the optimal observables can be chosen projective. However, all sets of preparations optimal for projective observables are of the form given above and one can check that for these preparations the optimal measurements must be projective (a direct consequence of the fact that the operators $\rho_{00} + \rho_{01} - \rho_{10} - \rho_{11}$ and $\rho_{00} - \rho_{01} + \rho_{10} - \rho_{11}$ are full rank).

It is the presence of multiple inequivalent maximizers that prevents us from writing down a simple self-testing statement. However, Eq. (G3) allows us to deduce the range of α compatible with the observed value of $\mathcal{A}_2$ (note that the conclusion will be stronger if we know whether $r = s$ or $r \neq s$). In particular, any value exceeding the classical bound of $7/8$ implies a lower bound on the incompatibility between M_0 and M_1 on the 2×2 subspace.

APPENDIX H: NUMERICAL METHOD FOR ROBUST SELF-TESTING

In the main text, we focused on the RAC and derived an optimal robust self-test. However, robust self-testing is relevant also for many other tasks that are not RACs. Here, we outline a numerical method based on semidefinite programming for inferring lower bounds on the worst-case average fidelity of preparations $\mathcal{F}$ in more general tasks. Specifically, we adapt the so-called swap method of [21] (constructed for Bell scenarios) to prepare-and-measure scenarios by combining it with the hierarchy of dimensionally bounded quantum correlations [41]. For sake of instruction, we first present the method by applying it to the RAC, and then use it to robustly self-test preparations in another prepare-and-measure scenario.

The preparations in the random access code are self-tested up to a collective unitary transformation. A robust self-test must therefore be valid under this degree of freedom. However, one can only consider the fidelity of the unknown preparations with respect to the optimal states in some chosen basis. Therefore, in order to achieve a robust self-test, one needs to find a way to avoid the possibility of a collective unitary misaligning the bases. This can be done by supplying Bob's measurement device with an auxillary system, say it is initialized in the state $|0\rangle_{\mathrm{A}}$, into which the unknown received preparations can be swapped [21]. In the RAC, the optimal measurements are anticommuting Pauli measurements. Therefore, with inspiration from this ideal case, Bob's swap operator S can be composed as follows: $S = UVU$, where

$$\begin{aligned} U &= \mathbb{1} \otimes |0\rangle\langle 0| + B_1 \otimes |1\rangle\langle 1|, \\ V &= \frac{1+B_0}{2} \otimes \mathbb{1} + \frac{1-B_0}{2} \otimes \sigma_x, \end{aligned} \tag{H1}$$

where B_0 and B_1 denote the observables of Bob. If B_0 and B_1 correspond to σ_z and σ_x, respectively, the above returns the two-qubit swap operator. Bob applies S to the joint system of received preparation (labeled B) and ancilla (labeled A). The state swapped into Bob's ancilla reads

$$\rho^{\mathrm{SWAP}}_{x_0 x_1} = \mathrm{tr}_{\mathrm{B}}[S(\rho_{x_0 x_1} \otimes |0\rangle_{\mathrm{AA}}\langle 0|)S^\dagger]. \tag{H2}$$

Consequently, the worst-case average fidelity of Alice's preparations with her optimal preparations is

$$\begin{aligned} \mathcal{F}(\mathcal{A}_2^*) &= \min_{\rho \in R(\mathcal{A}_2^*)} \max_{\Lambda} \frac{1}{4} \sum_{x_0 x_1} \mathrm{tr}\left[\Lambda\left[\rho^{\mathrm{ideal}}_{x_0 x_1}\right] \rho^{\mathrm{SWAP}}_{x_0 x_1}\right] \\ &= \min_{\rho \in R(\mathcal{A}_2^*)} \max_{\Lambda} \frac{1}{4} \sum_{x_0 x_1} \mathrm{tr}\left[S(\Lambda[\rho_{x_0 x_1}] \otimes |0\rangle_{\mathrm{AA}}\langle 0|) \right. \\ &\qquad \left. S^\dagger\left(\mathbb{1} \otimes \rho^{\mathrm{ideal}}_{x_0 x_1}\right)\right], \end{aligned} \tag{H3}$$

where $R(\mathcal{A}_2^*)$ is the set of all preparations that are compatible with the value $\mathcal{A}_2^*$ and Λ is the extraction channel, the duality of which is used above.

We may write the operator S in terms Bob's observables as follows:

$$S = \frac{1}{2} \sum_{ij} s_{ij} \otimes |i\rangle_{\mathrm{AA}}\langle j|, \tag{H4}$$

where

$$\begin{aligned} s_{00} &= \mathbb{1} + B_0, \quad s_{01} = B_1 - B_0 B_1, \\ s_{10} &= B_1 - B_1 B_0, \quad s_{11} = \mathbb{1} + B_1 B_0 B_1. \end{aligned} \tag{H5}$$

Inserting this into (H3) we find

$$\begin{aligned} \mathcal{F}(\mathcal{A}_2^*) &= \min_{\rho \in R(\mathcal{A}^*)} \max_{\Lambda} \frac{1}{16} \sum_{x_0 x_1} \sum_{ijkl} \mathrm{tr}\left[(s_{ij} \otimes |i\rangle_{\mathrm{AA}}\langle j|)(\Lambda[\rho_{x_0 x_1}] \otimes |0\rangle_{\mathrm{AA}}\langle 0|)(s_{kl} \otimes |k\rangle_{\mathrm{AA}}\langle l|)^\dagger\left(\mathbb{1} \otimes \rho^{\mathrm{ideal}}_{x_0 x_1}\right)\right] \\ &= \min_{\rho \in R(\mathcal{A}_2^*)} \max_{\Lambda} \frac{1}{16} \sum_{x_0 x_1} \sum_{ijkl} \mathrm{tr}[s_{ij} \Lambda[\rho_{x_0 x_1}] s_{kl}^\dagger] \, \mathrm{tr}\left[|i\rangle\langle j|0\rangle\langle 0|l\rangle\langle k| \rho^{\mathrm{ideal}}_{x_0 x_1}\right] \\ &= \min_{\rho \in R(\mathcal{A}_2^*)} \max_{\Lambda} \frac{1}{16} \sum_{x_0 x_1} \sum_{ik} \mathrm{tr}[s_{k0}^\dagger s_{i0} \Lambda[\rho_{x_0 x_1}]] \langle k| \rho^{\mathrm{ideal}}_{x_0 x_1} |i\rangle \\ &= \min_{\rho \in R(\mathcal{A}_2^*)} \max_{\Lambda} \frac{1}{16} \sum_{x_0 x_1} \sum_{ik} \mathrm{tr}[T_{ik} \Lambda[\rho_{x_0 x_1}]] \langle k| \rho^{\mathrm{ideal}}_{x_0 x_1} |i\rangle, \end{aligned} \tag{H6}$$

where we defined $T_{ik} = s_{k0}^{\dagger} s_{i0}$. The four elements of T are straightforwardly computed to

$$T_{00} = 2(\mathbb{1} + B_0), \quad T_{01} = B_1(\mathbb{1} - B_0) + B_0 B_1(\mathbb{1} - B_0), \tag{H7}$$

$$T_{11} = 2(\mathbb{1} - B_0), \quad T_{10} = B_1(\mathbb{1} + B_0) - B_0 B_1(\mathbb{1} + B_0). \tag{H8}$$

In the calculation of the fidelity, the same channel is applied to all Alice's preparations. We may simply consider that as four other valid preparations $\bar{\rho}_{x_0 x_1} = \Lambda[\rho_{x_0 x_1}]$. The fidelity in (H6) is then a linear combination of variables $\{\mathrm{tr}\,(\bar{\rho}_{x_0 x_1}\mathbb{1}), \mathrm{tr}\,(\bar{\rho}_{x_0 x_1} B_0), \ldots, \mathrm{tr}\,(\bar{\rho}_{x_0 x_1} B_0 B_1 B_0)\}$. Therefore, we may establish a lower bound on (H6) using the dimensionally bounded hierarchy of quantum correlations [41]. The accuracy of this bound depends on the level of the hierarchy employed. We choose to consider the following level: we define a moment matrix

$$\begin{aligned} \chi_{ijkl} &= \mathrm{tr}[R_j^{\dagger} Q_i^{\dagger} Q_k R_l], \quad \text{where} \\ Q &= (\mathbb{1}, B_0, B_1, B_0 B_1, B_1 B_0), \\ R &= (\mathbb{1}, \bar{\rho}_{00}, \bar{\rho}_{01}, \bar{\rho}_{10}, \bar{\rho}_{11}), \end{aligned} \tag{H9}$$

for $i, j, k, l = 1, \ldots, 5$. From the moment matrix we calculate all terms needed to evaluate the average fidelity (H6), using the labels $x = 2x_0 + x_1 + 2$,

$$\begin{aligned} \mathrm{tr}\left[T_{00}\bar{\rho}_{x_0 x_1}\right] &= 2\chi_{111x} + 2\chi_{112x}, \\ \mathrm{tr}\left[T_{11}\bar{\rho}_{x_0 x_1}\right] &= 2\chi_{111x} - 2\chi_{112x}, \end{aligned} \tag{H10}$$

$$\begin{aligned} \mathrm{tr}\left[T_{01}\bar{\rho}_{x_0 x_1}\right] &= \chi_{113x} + \chi_{114x} - \chi_{115x} - \chi_{215x}, \\ \mathrm{tr}\left[T_{10}\bar{\rho}_{x_0 x_1}\right] &= \chi_{113x} - \chi_{114x} + \chi_{115x} - \chi_{215x}. \end{aligned} \tag{H11}$$

In order to enforce that the average fidelity is extremized for a particular value $\mathcal{A}_2^*$ of the random access code, we write the probability distribution of Bob's outcomes in terms of the moment matrix as

$$P(b|x_0, x_1, y) = \frac{1 + (-1)^b \chi_{1,1,y+2,x}}{2}. \tag{H12}$$

Thus we can evaluate $\mathcal{A}_2$ as a linear combination of moment matrix elements. Fixing the value of $\mathcal{A}_2$ corresponds to introducing an affine constraint on the moment matrix. Therefore, the following semidefinite program establishes a lower bound on $\mathcal{F}(\mathcal{A}_2)$:

$$\mathcal{F}(\mathcal{A}_2^*) \geqslant \min_{\chi} \frac{1}{16} \sum_{x_0 x_1} \sum_{i,k=0}^{1} \mathrm{tr}\left(T_{ik}\bar{\rho}_{x_0 x_1}\right) \langle k|\rho_{x_0 x_1}^{\mathrm{ideal}}|i\rangle \tag{H13}$$

$$\text{such that } \chi \geqslant 0, \quad \mathcal{A}_2 \geqslant \mathcal{A}_2^*.$$

We have implemented the semidefinite program and the results are presented in Fig. 3, together with the lower bound on $\mathcal{F}(\mathcal{A}_2)$ obtained from the analytical method presented in the main text. Evidently, the swap method returns a suboptimal but still nontrivial result. Using the swap method, we find a higher-than-classical value of $\mathcal{F}(\mathcal{A}_2)$, i.e., $\mathcal{F}(\mathcal{A}_2) > 3/4$, whenever $\mathcal{A}_2 > 0.802$.

The advantage of the swap method is that it applies also to other prepare-and-measure scenarios beyond RACs. The drawback of the method is that the self-tests are typically not optimal, and that the complexity of evaluating the dimensionally bounded hierarchy of quantum correlations increases exponentially with the number of preparations and measurements, thus making more complicated scenarios infeasible to study.

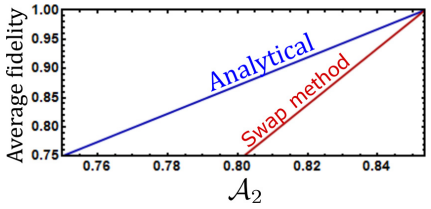

FIG. 3. Lower bound on $\mathcal{F}(\mathcal{A}_2)$ as obtained by the swap method and by analytical technique.

To exemplify the usefulness of this method also for other prepare-and-measure scenarios, we present a second example. Consider a prepare-and-measure scenario in which Alice has a random input $x \in \{0, 1, 2\}$ and Bob has a random input $y \in \{0, 1\}$. Alice may only communicate a qubit to Bob. The objective of the scenario reads

$$\mathcal{A} = \sum_{x,y} c_{x,y} E(x, y), \tag{H14}$$

where $E(x, y) = p(b = 0|x, y) - p(b = 1|x, y)$ and $c_{x,0} = [1, 1, -1]$ and $c_{x,1} = [\sqrt{3}, -\sqrt{3}, 0]$. One straightforwardly finds that the maximal classical value is $\mathcal{A} = 1 + 2\sqrt{3}$. We wish to robustly self-test Alice's preparations solely based on the value of $\mathcal{A}$. From numerical brute-force maximizations of $\mathcal{A}$, we find that its maximal value is $\mathcal{A} = 5$ and that this value is saturated using anticommuting Pauli measurements and preparations forming an equilateral triangle in a disk of the Bloch sphere. Such preparations can up to a unitary be written

$$\begin{aligned} \rho_0^{\mathrm{ideal}} &= \frac{1}{2}(\mathbb{1} + \sigma_x), \quad \rho_1^{\mathrm{ideal}} = \frac{1}{2}\left(\mathbb{1} + \frac{\sqrt{3}}{2}\sigma_z - \frac{1}{2}\sigma_x\right), \\ \rho_2^{\mathrm{ideal}} &= \frac{1}{2}\left(\mathbb{1} - \frac{\sqrt{3}}{2}\sigma_z - \frac{1}{2}\sigma_x\right). \end{aligned} \tag{H15}$$

We make the ansatz that this constitutes a self-test of the preparations. We supply Bob with an ancilla state and define the swap operator as done in the RAC. Performing calculations fully analogous to the case of the RAC, we obtain a semidefinite program that gives a lower bound on the

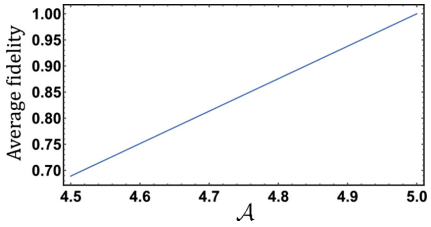

FIG. 4. Lower bound on $\mathcal{F}(\mathcal{A})$ as obtained by the swap method.

worst-case average fidelity

$$\mathcal{F}(\mathcal{A}) = \min_{\rho\in R(\mathcal{A})} \max_{\Lambda} \frac{1}{3}\sum_{x} \mathrm{tr}\left[\Lambda\left[\rho_x^{\mathrm{ideal}}\right]\rho_x\right], \tag{H16}$$

where $\mathcal{R}(\mathcal{A})$ is the set of preparations compatible with the value $\mathcal{A}$ and Λ is the extraction channel. We have used an intermediate level of the hierarchy of dimensionally bounded quantum correlations (sometimes referred to as 1+AB+BB+BBA) corresponding to an SDP matrix of size 20. The corresponding lower bound on $\mathcal{F}(\mathcal{A})$ is presented in Fig. 4. We first see that the maximal value $\mathcal{A} = 5$ indeed self-tests (up to numerical precision) the preparations of Alice to form an equilateral triangle on the Bloch sphere (the fidelity is one). For nonmaximal values of $\mathcal{A}$, we still obtain a nontrivial bound on the average fidelity of Alice's preparations with the optimal ones.

[1] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Phys. Rev. Lett. **98**, 230501 (2007).

[2] R. Colbeck, Ph.D. thesis, University of Cambridge, 2007; arXiv:0911.3814.

[3] S. Pironio, A. Acín, S. Massar, A. Boyer De La Giroday, N. D. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe, Nature (London) **464**, 1021 (2010).

[4] D. Mayers and A. Yao, *Proceedings of the 39th FOCS* (IEEE Computer Society, Washington, DC, 1998), p. 503.

[5] D. Mayers and A. Yao, Quantum Inf. Comput. **4**, 273 (2004).

[6] J. S. Bell, Physics **1**, 195 (1964).

[7] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Rev. Mod. Phys. **86**, 419 (2014).

[8] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).

[9] S. J. Summers and R. Werner, J. Math. Phys. **28**, 2440 (1987).

[10] S. Popescu and D. Rohrlich, Phys. Lett. A **169**, 411 (1992).

[11] B. S. Tsirelson, Hadron. J. Suppl. **8**, 329 (1993).

[12] B. W. Reichardt, F. Unger, and U. Vazirani, Nature (London) **496**, 456 (2013).

[13] A. Coladangelo, K. T. Goh, and V. Scarani, Nat. Commun. **8**, 15485 (2017).

[14] M. McKague, in *Theory of Quantum Computation, Communication, and Cryptography TQC 2011*, edited by D. Bacon, M. Martin-Delgado, and M. Roetteler, Lecture Notes in Computer Science Vol. 6745 (Springer, Berlin, Heidelberg, 2014), pp. 104–120.

[15] K. F. Pál, T. Vértesi, and M. Navascués, Phys. Rev. A **90**, 042340 (2014).

[16] X. Wu, Y. Cai, T. H. Yang, H. N. Le, J.-D. Bancal, and V. Scarani, Phys. Rev. A **90**, 042339 (2014).

[17] C.-E. Bardyn, T. C. H. Liew, S. Massar, M. McKague, and V. Scarani, Phys. Rev. A **80**, 062327 (2009).

[18] M. McKague, T. H. Yang, and V. Scarani, J. Phys. A: Math. Theor. **45**, 455304 (2012).

[19] T. H. Yang and M. Navascués, Phys. Rev. A **87**, 050102(R) (2013).

[20] C. Bamps and S. Pironio, Phys. Rev. A **91**, 052111 (2015).

[21] T. H. Yang, T. Vértesi, J.-D. Bancal, V. Scarani, and M. Navascués, Phys. Rev. Lett. **113**, 040401 (2014).

[22] J. Kaniewski, Phys. Rev. Lett. **117**, 070402 (2016).

[23] J. Kaniewski, Phys. Rev. A **95**, 062323 (2017).

[24] T. R. Tan, Y. Wan, S. Erickson, P. Bierhorst, D. Kienzler, S. Glancy, E. Knill, D. Leibfried, and D. J. Wineland, Phys. Rev. Lett. **118**, 130403 (2017).

[25] I. Supic, R. Augusiak, A. Salavrakos, and A. Acin, New J. Phys. **18**, 035013 (2016).

[26] R. Gallego, N. Brunner, C. Hadley, and A. Acin, Phys. Rev. Lett. **105**, 230501 (2010).

[27] M. Hendrych, R. Gallego, M. Micuda, N. Brunner, A. Acin, and J. P. Torres, Nat. Phys. **8**, 588 (2012).

[28] J. Ahrens, P. Badziag, A. Cabello, and M. Bourennane, Nat. Phys. **8**, 592 (2012).

[29] M. Pawłowski and N. Brunner, Phys. Rev. A **84**, 010302(R) (2011).

[30] H.-W. Li, Z.-Q. Yin, Y.-C. Wu, X.-B. Zou, S. Wang, W. Chen, G.-C. Guo,and Z.-F. Han, Phys. Rev. A **84**, 034301 (2011).

[31] E. Woodhead and S. Pironio, Phys. Rev. Lett. **115**, 150501 (2015).

[32] T. Lunghi, J. B. Brask, C. C. W. Lim, Q. Lavigne, J. Bowles, A. Martin, H. Zbinden, and N. Brunner, Phys. Rev. Lett. **114**, 150501 (2015).

[33] P. Mironowicz, A. Tavakoli, A. Hameedi, B. Marques, M. Pawłowski, and M. Bourennane, New J. Phys. **18**, 065004 (2016).

[34] M. Dall'Arno, S. Brandsen, F. Buscemi, and V. Vedral, Phys. Rev. Lett. **118**, 250501 (2017).

[35] M. Dall'Arno, S. Brandsen, and F. Buscemi, Proc. R. Soc. A **473**, 20160721 (2017).

[36] M. Dall'Arno, arXiv:1702.00575.

[37] A. Ambainis, A. Nayak, A. Ta-Shama, and U. Vazirani, *Proceedings of 31st ACM Symposium on Theory of Computing* (ACM, Atlanta, 1999), pp. 376–383.

[38] A. Nayak, *Proceedings of the 40th IEEE Symposium on Foundations of Computer Science (FOCS'99)* (IEEE, New York, 1999), pp. 369–376.

[39] A. Tavakoli, A. Hameedi, B. Marques, and M. Bourennane, Phys. Rev. Lett. **114**, 170502 (2015).

[40] E. Woodhead, C. C. W. Lim, and S. Pironio, in *Theory of Quantum Computation, Communication, and Cryptography, TQC 2012*, edited by K. Iwama, Y. Kawano, and M. Murao, Lecture Notes in Computer Science Vol. 7582 (Springer, Berlin, Heidelberg, 2013), pp. 107–115

[41] M. Navascués and T. Vértesi, Phys. Rev. Lett. **115**, 020501 (2015).

[42] M. Navascués, A. Feix, M. Araújo, and T. Vértesi, Phys. Rev. A **92**, 042117 (2015).

[43] A. Tavakoli, D. Rosset, and M. O. Renou, arXiv:1808.02412.

[44] P. Trojek, C. Schmid, M. Bourennane, C. Brukner, M. Żukowski, and H. Weinfurter, Phys. Rev. A **72**, 050305(R) (2005).

[45] J. Bowles, N. Brunner, and M. Pawłowski, Phys. Rev. A **92**, 022351 (2015).

[46] T. Van Himbeeck, E. Woodhead, N. J. Cerf, R. García-Patron, and S. Pironio, Quantum **1**, 33 (2017).

[47] J. B. Brask, A. Martin, W. Esposito, R. Houlmann, J. Bowles, H. Zbinden, and N. Brunner, Phys. Rev. Appl. **7**, 054018 (2017).

[48] R. Chaves, J. B. Brask, and N. Brunner, Phys. Rev. Lett. **115**, 110501 (2015).

Device-independent witness of arbitrary-dimensional quantum systems employing binary-outcome measurements

Mikołaj Czechlewski,[1,*] Debashis Saha,[2,3,†] Armin Tavakoli,[4,‡] and Marcin Pawłowski[2,§]

[1]*Institute of Informatics, National Quantum Information Centre, Faculty of Mathematics, Physics and Informatics, University of Gdańsk, Wita Stwosza 57, 80-308 Gdańsk, Poland*

[2]*Institute of Theoretical Physics and Astrophysics, National Quantum Information Centre, Faculty of Mathematics, Physics and Informatics, University of Gdańsk, Wita Stwosza 57, 80-308 Gdańsk, Poland*

[3]*Center for Theoretical Physics, Polish Academy of Sciences, Aleja Lotników 32/46, 02-668 Warsaw, Poland*

[4]*Département de Physique Appliquée, Université de Genève, 1211 Genève, Switzerland*

(Received 27 March 2018; published 4 December 2018)

Device-independent dimension witnesses (DWs) are a remarkable way to test the dimension of a quantum system in a prepare-and-measure scenario imposing minimal assumptions on the internal features of the devices. However, as the dimension increases, the major obstacle in the realization of DWs arises due to the requirement of many-outcome quantum measurements. In this article we propose a variant of a widely studied communication task (random access code) and take its average payoff as the DW. The presented DW applies to arbitrarily large quantum systems employing only binary-outcome measurements.

DOI: 10.1103/PhysRevA.98.062305

I. INTRODUCTION

Realizing higher-dimensional quantum systems with full control is one of the crucial barriers towards implementing many quantum information processing protocols and testing the foundations of physics. While the process of quantum tomography allows us to reconstruct a quantum system, it requires the assumption of fully characterized measurement devices. The device-independent framework [1,2] in a prepare-and-measure experiment provides a methodology to obtain a lower bound on the dimension without assuming the internal features of the devices. Moreover, quantum advantages in information processing, for example, quantum communication complexity [3,4], are linked to this approach. Despite its merits, implementing device-independent dimension witnesses (DWs) for higher-dimensional quantum systems [5–9] faces several complications.

One of the problems in many existing protocols is the requirement of d-outcome measurements. As the dimension increases, performing many-outcome measurements [10] becomes practically difficult due to the facts that (a) measurement outcomes turn coarse grained and (b) the system becomes more prone to decoherence. In some cases, one may impose additional assumptions, for instance, simulating d-outcome measurements by many binary-outcome measurements. However, this approach fails to fulfill the requirements of DWs in the strict sense.

Another difficulty arises from the fact that the number of different preparations and measurements (i.e., the total number of inputs in the devices) also increases as one seeks to certify a higher-dimensional system. As a result, the experimental errors grow large due to the finite number of trials and imperfections in the experiment.

Furthermore, the applicability of a desired figure of merit, used as a DW, should not be limited by a particular dimension. Rather, it should be applicable to test systems of an arbitrarily high dimension.

In this article we overcome these challenges by proposing a class of DWs based on random access codes (RACs) [11] for quantum systems of an arbitrary dimension. In the simplest scenario, a DW can be interpreted as a task carried out by two parties. In each run of the task, the sender Alice obtains an input in the form of a classical variable a and communicates a system to the receiver Bob. Apart from the communicated message, Bob also receives an input y and produces an output b. The figure of merit, denoted by $\overline{T}$, of the task could be an arbitrary linear function of the statistics $\overline{T} = \sum_{a,y,b} p(a,y)T(a,y,b)p(b|a,y)$, where $p(b|a,y)$ refers to the probability of obtaining the output b given the inputs a and y, and $T(a,y,b)$ denotes the payoff to that event.[1] Assuming that the dimension of the communicated system is d, one can obtain the optimal value of the figure of merit, denoted by $\overline{T}^c$, for a classical implementation. Obtaining a value greater than $\overline{T}^c$ from the observed statistics certifies the communicated quantum system to be of at least dimension d. Quantum random access codes (QRACs), a primitive quantum communication protocol [12–14], can be used for this purpose. The original study of QRACs was restricted to two-dimensional systems [11] and was later generalized to

*mczechlewski@inf.ug.edu.pl
†saha@cft.edu.pl
‡armin.tavakoli@unige.ch
§dokmpa@univ.gda.pl

[1]Note that $p(a,y)$ could be absorbed into $T(a,y,b)$. Nevertheless, the stated form provides a simple intuition.

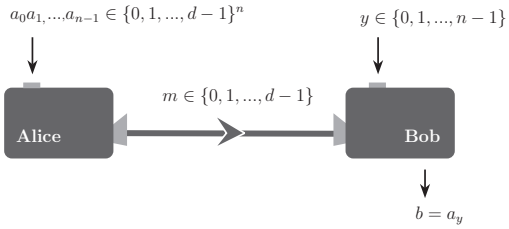

FIG. 1. Scheme of the dRAC. Alice gets the input $a_0, \ldots, a_{n-1}$ and sends a message m to Bob. Besides the message, Bob also receives the input $y \in \{0, \ldots, n-1\}$. His task is to give the output b, which obeys the relation $b = a_y$.

higher dimensions [15–17], yielding several interesting results in quantum communication [18–21].

There are advantages of using RAC as DWs. The upper bound on $\overline{T}^c$ can be obtained for any d. In addition, the number of inputs in the devices increases polynomially with d. Note that one can exploit the quantum communication complexity tasks [4], which involve binary-outcome measurement for dimension witnessing, but in that case, the input size grows exponentially with d. However, the generalized RAC requires d-outcome measurements. To tackle this issue we introduce a version of RAC, namely, binary RAC. This involves only binary-outcome measurements and provides a method to obtain the upper bound of $\overline{T}^c$ applicable to arbitrary d.

The paper is organized as follows. First, we describe the generalization of the d-dimensional RAC, along with the proof of optimal classical protocols and bounds. Next, we propose the binary version (i.e., the outcome b is binary) of the d-dimensional RAC, taking into account a wider class of payoff function. Then, we derive a condition on the payoff function such that the optimal classical protocol is the same as in a standard RAC. Further, we provide the classical bound and a quantum protocol that violates the proposed DW for arbitrary d.

II. STANDARD d-DIMENSIONAL RANDOM ACCESS CODE

Standard d-dimensional random access codes (dRACs) are a natural generalization of the two-dimensional random access code [11,17]. Alice receives n numbers $a_0, \ldots, a_{n-1}$, where $a_i \in \{0, \ldots, d-1\}$. Then she sends a d-valued (one dit[2]) message $m \in \{0, \ldots, d-1\}$ to Bob. Bob gets an input $y \in \{0, \ldots, n-1\}$. He needs to give the output b, which obeys the relation $b = a_y$ (Fig. 1). Specifically, we are interested in the average success probability in the case of the inputs a and y being uniformly distributed and $T(a, y, b) = \delta_{b,a_y}$,

$$\overline{T}_S = \frac{1}{nd^n} \sum_{a,y} p(b = a_y|a, y). \tag{1}$$

Since the communicated message m is constrained to be d valued, it is evident that achieving average success probability equal to 1 is impossible. The aim is to find an optimal strategy for the parties, which gives the largest average success probability.

Following the result in [11] for $d = 2$, it was mentioned in [17] and shown later in [22] that coding by majority and identity decoding is an optimal strategy for dRACs. In the next two sections, we demonstrate an alternative shorter proof of this fact and subsequently provide an expression of the optimal average success probability.

A. Optimal classical strategy

Due to the linearity of the figure of merit, it is sufficient to consider only deterministic encoding and decoding strategies to maximize the average success probability. Let us denote the dit-string $a_0, \ldots, a_{n-1}$ by a. Any encoding strategy can be described by a function $E : \{a\} \equiv \{0, \ldots, d-1\}^n \mapsto \{m\} \equiv \{0, 1, \ldots, d-1\}$ and the probability of sending m for input a is $\delta_{m,E(a)}$. In contrast, any decoding for Bob's input y is described as a function $D_y : \{m\} \equiv \{0, 1, \ldots, d-1\} \mapsto \{b\} \equiv \{0, 1, \ldots, d-1\}$ and $\delta_{b,D_y(m)}$ is the probability of outputting b when message m is received. Thus, the classical average success probability in the standard RAC is

$$\begin{aligned}
\overline{T}^c_S &= \frac{1}{nd^n} \sum_{a,y} p(b = a_y|a, y) \\
&= \frac{1}{nd^n} \sum_{m} \sum_{a,y} \delta_{m,E(a)} \delta_{a_y,D_y(m)} \\
&= \frac{1}{nd^n} \sum_{m,a} \delta_{m,E(a)} \left(\sum_{y} \delta_{a_y,D_y(m)} \right) \\
&\leqslant \frac{1}{nd^n} \sum_{a} \max_{m} \left(\sum_{y} \delta_{a_y,D_y(m)} \right).
\end{aligned} \tag{2}$$

From this expression, we can observe that for given decoding strategy $D_y(m)$, the optimal encoding will be

$$\begin{aligned}
&\delta_{m,E(a)} = 1 \quad \text{if, } \forall m' \in \{0, \ldots, d-1\}, \\
&\sum_{y} \delta_{a_y,D_y(m)} \geqslant \sum_{y} \delta_{a_y,D_y(m')}.
\end{aligned} \tag{3}$$

We can reduce the possibility of all decoding functions into two ways: (a) identity decoding, i.e., $\forall y, m$, $D_y(m) = m$, and (b) not identity decoding, $\exists y, m$ such that $D(m) \neq m$. Here the mapping D_y could be one to many in general.

Lemma. There exists an optimal classical strategy with identity decoding (a).

Proof. We will show that for the case described in (b), there exists a strategy obtaining the same average success probability as for the identity decoding (a). Let $D_y^{\leftarrow}(b)$ be the domain of b, i.e., the set of m such that $D_y(m) = b$. If b does not exist in the range of D_y, we define $D_y^{\leftarrow}(b) = b$. We denote by $D_y^{\leftarrow}(a)$ the set of a dit string $a' \equiv a'_0, \ldots, a'_{n-1}$ such that $D_y(a'_y) = a_y$. Thus, $D_y^{\leftarrow}(a)$ acts on the yth dit of the dit string. If there is a classical strategy having an encoding function E and decoding functions D_y [where $D_y(m) \neq m$ for some y, m], we can construct different encoding and decoding

[2]By dit we mean a d-dimensional classical system.

functions

$$E'(D_0^{\leftarrow} D_1^{\leftarrow} \cdots D_{n-1}^{\leftarrow}(a)) = E(a)$$
$$\forall y, m,\ D'_y(m) = m. \tag{4}$$

Now, if the strategy (E, D_y) gives the correct answer for the input (a, y), then the modified strategy (E', D'_y) gives the correct answer for at least one of the inputs $(D_0^{\leftarrow} \cdots D_{n-1}^{\leftarrow}(a), y)$. Thus, the average success probability for the modified strategy (E', D') is equal to or greater than the strategy (E, D). ■

From (3) we conclude that optimal encoding is

$$\delta_{m,E(a)} = 1 \quad \text{such that,}$$
$$\forall m' \in \{0, \ldots, d-1\}, \quad \sum_y \delta_{a_y,m} \geqslant \sum_y \delta_{a_y,m'}. \tag{5}$$

In other words, the optimal strategy for Alice is to communicate the majority dit of the input string and $b = m$.

B. Average success probability

Now we calculate the classical average success probability for an n-dit string. The total number of possible inputs is nd^n. In the n-dit string, which is given to Alice, the ith dit ($i \in \{0, 1, 2, \ldots, n-1\}$) appears n_i times in the string a. The number of ways it may occur is the same as the number of solutions in non-negative integers of the equation

$$n_0 + n_1 + n_2 + \cdots + n_{d-1} = n. \tag{6}$$

The Equation (6) is a special case of the equation

$$c_0 n_0 + c_1 n_1 + c_2 n_2 + \cdots + c_{d-1} n_{d-1} = n, \tag{7}$$

with all coefficients $\{c_0, c_1, c_2, \ldots, c_{d-1}\}$ equal 1. Equation (7) is known in number theory as the Diophantine equation of Frobenius and it is connected to the Frobenius coin problem and the Frobenius number [23,24]. The total number of possible solutions of (6) is $\binom{n+d-1}{d-1}$ [25]. For each solution Alice will communicate $\max\{n_0, n_1, \ldots, n_{d-1}\}$ to Bob. So the number of successful inputs is given by $\frac{n!}{n_0!n_1!\cdots n_{d-1}!} \max\{n_0, n_1, \ldots, n_{d-1}\}$, as $\frac{n!}{n_0!n_1!\cdots n_{d-1}!}$ is the number of possible combinations for an n-dit string with a given set of n_i's and $\max\{n_0, n_1, \ldots, n_{d-1}\}$ is the number of times where Bob will guess the correct dit. Therefore, the average success probability is given by

$$\overline{T}_S^c = \frac{1}{nd^n} \sum \frac{n!}{n_0!n_1! \cdots n_{d-1}!} \max\{n_0, n_1, \ldots, n_{d-1}\}, \tag{8}$$

where the summation is over all $\binom{n+d-1}{d-1}$ possible solutions of (6).

III. BINARY RANDOM ACCESS CODE

A binary random access code (Fig. 2) is a communication complexity problem based on the standard dRAC. Two parties, Alice and Bob, are given the following task. Alice receives n dits $a = a_0, \ldots, a_{n-1}$, the same as in the standard dRAC. She sends a d-valued message to Bob. However, Bob gets two inputs $y \in \{0, 1, \ldots, n-1\}$ and $k \in \{0, 1, \ldots, d-1\}$. He needs to answer the following question: Is $a_y = k$? Bob encodes his answer in a variable G, which is 0 when his guess is yes and 1 for no.

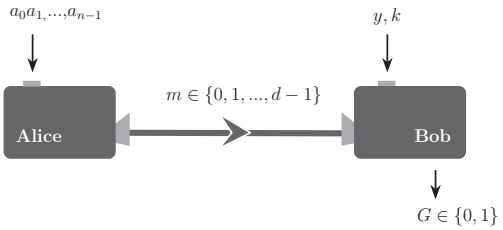

FIG. 2. Scheme of the binary RAC. Alice gets the input $a_0, \ldots, a_{n-1}$ and sends the message m to Bob. Besides the message Bob receives two inputs $y \in \{0, 1, \ldots, n-1\}$ and $k \in \{0, 1, \ldots, d-1\}$. His task is to guess whether $a_y = k$ or not. His answer is encoded in G, which is 0 when his guess is yes and 1 when it is no.

A. Defining the average payoff function

We are free to reward the parties with any number of points, specified by a payoff function $T(a, y, k, G)$. Therefore, for simplicity, we assume that this function does not depend on the values of numbers a_i in the input a with indices different from y. Hence, we assign T only two values

$$T(a_y, k, G) = \begin{cases} T_{\text{yes}} & \text{when } G = 0,\ a_y = k \\ 1 & \text{when } G = 1,\ a_y \neq k. \end{cases} \tag{9}$$

We are interested in the average payoff function, which is a linear combination of payoffs for all possible uniformly distributed inputs. Without loss of generality, we can normalize the average payoff such that it takes a value within [0,1]. Thus, for the binary RAC with payoffs defined in (9) we have

$$\overline{T}_B = \frac{1}{nd^n T_d} \Bigg[\sum_{a,y,k} [p(G = 0|a, y, k, a_y = k) T_{\text{yes}} + p(G = 1|a, y, k, a_y \neq k)] \Bigg], \tag{10}$$

where $T_d = T_{\text{yes}} + d - 1$ such that $\overline{T}_B$ is normalized.

B. Optimal classical strategy for Bob

To find the optimal classical strategy for Bob, first we split him into two parts: B_I (initial Bob) and B_F (final Bob). Here B_I gets the message m from Alice, receives input y, and forwards a d-long-bit string $b = b_0, \ldots, b_{d-1}$ to B_F. Each of the bits in the string represents the given answer of B_F for a different question ruled by k. Thus, when B_F gets k and the bit string b he returns $G = b_k$ (Fig. 3). This splitting in no way reduces the generality of Bob's behavior since the whole information processing part is done locally by B_I; B_F only returns one of the values from a table provided by B_I.

Notice that before receiving Alice's message Bob knows nothing about the string a, so his entropy $H(a) = n \log_2 d$ (we assume that Alice's inputs are uniformly distributed). After receiving the message, Bob's entropy for each a_i is reduced to $H_i^m = H(a_i|m)$. These two entropies are related by the

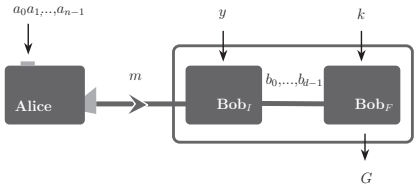

FIG. 3. Scheme of the binary RAC. Bob is split into two parts: B_I (initial Bob) and B_F (final Bob).

information causality principle [26]

$$H(a) - \sum_{i=0}^{n-1} H_i \leqslant C, \tag{11}$$

where $H_i = \sum_{m=0}^{d-1} p(m) H_i^m$ is the averaged conditional Shannon entropy and C is the capacity of a classical channel. Hence, from (11) we obtain the lower bound for H_i, which is determined by two established quantities: entropy $H(a)$ and the channel capacity C.

Besides the message m, B_I receives the input y, which makes him interested in the particular dit a_y from the string a. Let us introduce the probability distribution $p_j = p(a_y = j|m, y)$, where $j \in \{0, \ldots, d-1\}$, which represents B_I's knowledge about dit y. First, we see that the entropy $H_{i=y}^m$ can be presented in terms of this probability distribution

$$H_{i=y}^m = -\sum_{j=0}^{d-1} p_j \log_2 p_j. \tag{12}$$

Second, we notice that, depending on the payoff function, there exists a critical value of probability (p_{crit}) such that if $p_j > p_{\text{crit}}$, then sending $b_j = 0$ leads to larger average payoff than $b_j = 1$. We derive a formula for p_{crit} in the following way. We know that sending $b_j = 0$ leads to the answer $G = 0$ for $j = k$. This gives T_{yes} points with probability p_j. For $b_j = 1$ we get one point with $1 - p_j$. The first option is better if $T_{\text{yes}} p_j \geqslant 1 - p_j$, so

$$p_j \geqslant \frac{1}{T_{\text{yes}} + 1} = p_{\text{crit}}. \tag{13}$$

Furthermore, let us analyze the average payoff $T = T(m, y)$ for a message set m, given encoding strategy E, the input y and T_d defined in (10),

$$T = \frac{1}{T_d} \sum_{j=0}^{d-1} [T_{\text{yes}} p(b_j = 0|m, y) p(a_y = j|m, y) + p(b_j = 1|m, y) p(a_y \neq j|m, y)]. \tag{14}$$

We introduce a variable x as the number of bits in the string b for which the optimal strategy is set to 0 for the probability distribution $p(b_j|m, y)$. In other words, x is the number of p_j that are greater than p_{crit}. Using x we can rewrite the entropy $H_{i=y}^m$ [Eq. (12)] as

$$H_{i=y}^m = -\sum_{j=0}^{x-1} p_j \log_2 p_j - \sum_{j=x}^{d-1} p_j \log_2 p_j. \tag{15}$$

Additionally, without loss of generality, we may assume that p_j are ordered in such way that $p_j \geqslant p_{j+1}$. Then the average payoff becomes

$$T = \frac{1}{T_d} \left[\sum_{j=0}^{x-1} T_{\text{yes}} p_j + \sum_{j=x}^{d-1} (1 - p_j) \right]. \tag{16}$$

Because the value of T [Eq. (16)] depends only on the sums $\sum_{j=0}^{x-1} p_j$ and $\sum_{j=x}^{d-1} p_j$ and not on the individual elements of the sums, we can choose that all the elements in each sum are equal because this makes the entropy $H_{i=y}^m$ [Eq. (15)] the largest without changing T. In other words, the probability distribution p_j becomes a step function: The values of all p_j for $j = \{0, \ldots, x-1\}$ are uniform (denoted by p) and the values of the remaining p_j for $j = \{x, \ldots, d-1\}$ are uniform as well and, according to the normalization condition $\sum_j p_j = 1$, they must be equal to $\frac{1-xp}{d-x}$. Obviously, we assume that the encoding strategy E reaches $p > \frac{1}{d}$. Due to the above assumptions, we can express T as a function of x and p,

$$T = \frac{1}{T_d} \{x[T_{\text{yes}}\, p - (1 - p)] + d - 1\}. \tag{17}$$

The entropy (15) (henceforth denoted by H^x) can also be expressed by these parameters

$$H^x = -xp \log_2 p - (1 - xp) \log_2 \frac{1 - xp}{d - x}. \tag{18}$$

Imposing (13), we substitute T_{yes} in (17) and find

$$p = \frac{T + p_{\text{crit}}[d(T-1) - 2T + x + 1]}{x}. \tag{19}$$

One can further plug the above expression into (18) to get the entropy H^x as a function of d, T, x, and p_{crit}.

C. Optimal x for our case

It has been shown in Sec. II A that the majority of the encoding is optimal in the standard RAC scenario, where Alice is allowed to send only one dit of information to Bob. To employ this result in the binary RAC protocol (in this case B_I sends to B_F a bit string $b_0, \ldots, b_{d-1}$ with exactly one 0 in the established position and 1's in the others) we must set the restriction that for any T, the probability p for $x = 1$ is always greater than any p for $x \neq 1$ [Eq. (19)]. To do this we must find a lower bound of p_{crit} such that the entropy $H^{x=1}$ is always greater than any entropy $H^{x\neq 1}$ for any given value of T from the relevant range. Hence, in the beginning, we define a function Δ_i in the following way:

$$\forall i \neq 1, \quad \Delta_i = H^{x=1} - H^{x=i}. \tag{20}$$

Notice that the symmetry of the entropy $H^x = H^{d-x}$ for $x \in \{1, \ldots, d-1\}$ makes it sufficient to check the condition (20) only for Δ_i, $i \in \{2, 3, \ldots, \lceil \frac{d}{2} \rceil\}$.

Let us outline the methodology of obtaining the minimum value of p_{crit} for which $\Delta_i > 0$. Clearly, Δ_i is a function of d, T, and p_{crit}. We first find the range of T in terms of d and p_{crit} within which Δ_i is well defined. After that, we fix the value of d and p_{crit} and obtain the minimum value of Δ_i within the relevant range of T for all i. If the minimum value of Δ_i is

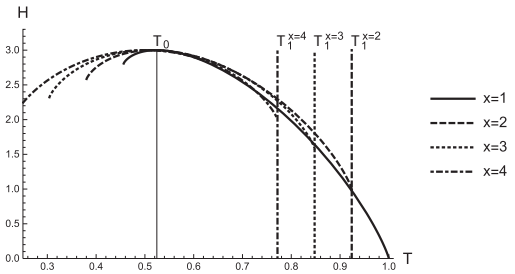

FIG. 4. Dependence of H on T for $d = 8$, $x = 1, 2, 3, 4$, and $p_{\text{crit}} = 0.14$. We note that the entropy for the strategy with $x = 1$ is not always the largest in the established ranges of T. According to the numerical procedure, this is an example in which, at step III C, $\Delta_i \leqslant 0$ and our algorithm skips from step III C to step III C. Vertical lines indicate the limits of the ranges $[T_0, T_1^{x=i}]$.

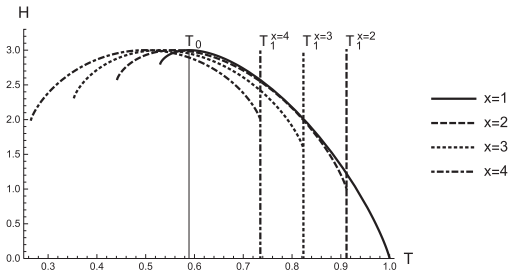

FIG. 5. Dependence of H on T for $d = 8$, $x = 1, 2, 3, 4$, and $p_{\text{crit}} = 0.184\,95$. The largest entropy is obtained with exactly one strategy for which $x = 1$. According to our procedure, this is an example in which, at step 6, $\Delta_i > 0$ for every $i \in \{2, 3, \ldots, \lceil \frac{d}{2} \rceil\}$ and our algorithm returns p_{crit}. Vertical lines indicate the limits of the ranges $[T_0, T_1^{x=i}]$.

nonpositive for some $i \in \{2, \ldots, \lceil \frac{d}{2} \rceil\}$, we know that such a value of p_{crit} is not suitable. We repeat the evaluation of Δ_i for another value of p_{crit} increased by a small interval than before. Once we find that Δ_i is positive for all $i \in \{2, \ldots, \lceil \frac{d}{2} \rceil\}$, we conclude that the taken value of p_{crit} is approximately the same as the desired value.

For every Δ_i we must determine the range of T. The lower limit of the range is the value of T for which $H^{x=1}$ is maximal. According to (18), $H^{x=1}$ takes a maximum for $p = \frac{1}{d}$. Putting it in (19) gives an analytical expression for the lower limit

$$T_0 = \frac{1 + (d-2)dp_{\text{crit}}}{d + (d-2)dp_{\text{crit}}}. \tag{21}$$

On the other hand, the upper limit of the range is the value of $T > T_0$ for which H^x takes the bound. The bound is established by setting $xp = 1$ in (18) so that it strictly depends on x. Hence, setting $p = \frac{1}{x}$ in (19) gives

$$T_1^{x=i} = \frac{1 + p_{\text{crit}}(d - i - 1)}{1 + (d-2)p_{\text{crit}}}. \tag{22}$$

Thus, for every Δ_i there is a different range $[T_0, T_1^{x=i}]$.

We have found p_{crit} numerically using a method described by the following algorithm.

1. For a chosen dimension d, set $p_{\text{crit}} = \frac{1}{d}$ and $\varepsilon_{p_{\text{crit}}}$, which is its numerical increase.
2. Substitute $p_{\text{crit}} := p_{\text{crit}} + \varepsilon_{p_{\text{crit}}}$.
3. Calculate T_0 from (21).
4. Set the variable $i := 2$.
5. Calculate $T_1^{x=i}$ from (22).
6. Calculate Δ_i for T_0 and $T_1^{x=i}$ and find the minimal value of Δ_i in the range $[T_0, T_1^{x=i}]$ (if the minimal value does not exist do not take it into account). If $\Delta_i \leqslant 0$ for at least one of these three (or two) points, then go to point III C. Otherwise, $i := i + 1$.
7. Check if $i \leqslant \lceil \frac{d}{2} \rceil$. If it is fulfilled then go to step 5. Otherwise return p_{crit}.

Obviously, the accuracy of our method depends strictly on $\varepsilon_{p_{\text{crit}}}$. The smaller it is, the more precise the result is. Additionally, it is noteworthy that the criterion for optimal encoding is derived from $H^m_{i=y}$ [Eq. (12)], which is valid for all $y \in \{0, \ldots, n-1\}$ and thus it is independent of n.

To illustrate the procedure described above we plot the dependence of H on T for some small p_{crit} and different values of x in Figs. 4 and 5. Obtained values of p_{crit} along with their corresponding T_{yes} are shown in Fig. 6 and the values for some particular dimensions are given in Table I.

D. Average classical and quantum payoff function for $n = 2$ and arbitrary dimension

Now we calculate the average classical and quantum payoff (10) for binary RAC. First, for a given dimension d, we must determine the value of T_{yes} corresponding to $x = 1$ as it was presented in the preceding section. It follows that the optimal encoding strategy is sending the majority dit, which is the same as for the standard dRAC [Eq. (5)]. Further, it can be readily seen that, given an encoding E, the optimal

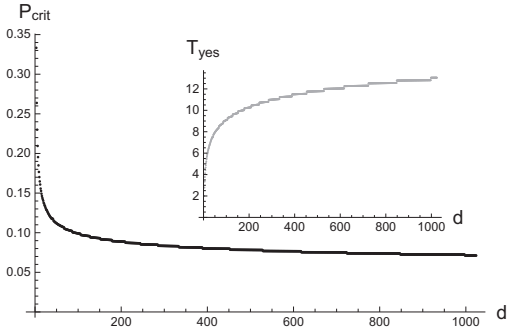

FIG. 6. Numerical calculation of values of the minimal p_{crit} and the corresponding maximal T_{yes} as a function of dimension with accuracy $\varepsilon_{p_{\text{crit}}} = 10^{-5}$.

TABLE I. Values of minimal p_{crit} and the corresponding maximal T_{yes} for chosen dimensions d.

d	p_{crit}	T_{yes}
3	0.33340	1.99940
8	0.18495	4.40687
10	0.17021	4.87510
50	0.11180	7.94454
200	0.08885	10.25490
700	0.07524	12.29080
1000	0.07121	13.04300

decoding is

$$G = \begin{cases} 0 & \text{if } \sum_{a|a_y=k} \delta_{m,E(a)} \geqslant \sum_{a|a_y\neq k} \delta_{m,E(a)} \\ 1 & \text{otherwise.} \end{cases} \tag{23}$$

Therefore, in the case of majority encoding, Bob returns $G = 0$ if the received message $m = k$; otherwise 1. Given an input a, the total payoff over all possible y and k is

$$T_d\tilde{n} + (d-2)(n-\tilde{n}), \tag{24}$$

where we define $\tilde{n} = \max\{n_0, n_1, \ldots, n_{d-1}\}$. This is due to the fact that if y is such that n_y is the maximum, i.e., a_y is the majority dit, then Bob gives the correct answer for all k, obtaining the maximum payoff T_d. Such an event occurs $\tilde{n}$ times. In the other $n-\tilde{n}$ cases Bob returns the correct answer only if $k \neq a_y$ and $k \neq E(a)$, obtaining the $d-2$ payoff. Subsequently, the average payoff is given by

$$\overline{T}^c_B = \frac{1}{nd^nT_d} \sum \frac{n!}{n_0!n_1!\cdots n_{d-1}!} \times [\tilde{n}(T_{\text{yes}}+1) + n(d-2)], \tag{25}$$

where the summation is over all $\binom{n+d-1}{d-1}$ possible solutions of (6). Imposing the expression of the average payoff of the dRAC [Eq. (8)], $\overline{T}^c_B$ simplifies to

$$\overline{T}^c_B = \frac{(T_{\text{yes}}+1)\overline{T}^c_S + d - 2}{T_{\text{yes}} + d - 1}. \tag{26}$$

For $n = 2$ we can find $\overline{T}^c_S = \frac{1}{2} + \frac{1}{2d}$, and substituting this in (26) leads to

$$\overline{T}^c_B = \frac{1}{T_d}\left[\frac{T_{\text{yes}} + 1 + d(2d + T_{\text{yes}} - 3)}{2d}\right]. \tag{27}$$

Let us consider a quantum strategy based on the quantum dRAC presented in [17]. Alice codes her input a_0a_1 in a d-dimensional quantum state as

$$|\psi_{a_0a_1}\rangle = \frac{1}{N_{2,d}}\left(|a_0\rangle + \frac{1}{\sqrt{d}}\sum_{j=0}^{d-1}\omega^{ja_1}|a_1+j\rangle\right), \tag{28}$$

where $N_{2,d} = \sqrt{2+\frac{2}{\sqrt{d}}}$ is the normalization factor and $\omega = e^{2\pi i}$ is the quantum Fourier transform factor. For the decoding Bob uses the projective measurements M^y_k, depending on inputs y and k,

$$M^0_k = \{P^0_k, \mathbb{I} - P^0_k\}, \quad M^1_k = \{P^1_k, \mathbb{I} - P^1_k\}. \tag{29}$$

Here $P^0_k = |k\rangle\langle k|$ and $P^1_k = |\bar{k}\rangle\langle\bar{k}|$, where $|\bar{k}\rangle = \frac{1}{\sqrt{d}}\sum_{k=0}^{d-1}\omega^{k\bar{k}}|k\rangle$ corresponds to the outcome $G = 0$. Simple calculations lead to the quantum average payoff

$$\overline{T}^q_B = \frac{1}{T_d}\left[\frac{T_{\text{yes}} + 1 + \sqrt{d}(2d + T_{\text{yes}} - 3)}{2\sqrt{d}}\right]. \tag{30}$$

The difference between (30) and (27) is given by

$$\overline{T}^q_B - \overline{T}^c_B = \frac{1}{T_d}\left[\frac{(T_{\text{yes}}+1)(\sqrt{d}-1)}{2d}\right], \tag{31}$$

which is always greater than zero for $d \geqslant 2$. Thus, the binary version of the RAC provides a device-independent way to test an arbitrary-dimensional quantum system employing only binary-outcome measurements.

IV. SUMMARY

The primary feature of this article was to present a DW applicable to test arbitrarily large quantum systems implementing only binary-outcome measurements. We proposed a variant of the RAC and took the average payoff of this communication task as the indicator of the dimension. We have provided the optimal classical bound for the binary version of the generalized RAC. In contrast to the other quantum communication complexity problems in which the number of prepared states grows exponentially with dimension, the proposed DW requires d^2 different preparations and 2d measurements. In the future, it would be interesting to prove the optimality of the quantum strategy for binary RAC and look for more robust DWs retaining the aforementioned significant features.

ACKNOWLEDGMENTS

We thank Máté Farkas and Edgar A. Aguilar for helpful discussions and comments. We are also grateful to Edgar A. Aguilar for a critical reading of the manuscript. This work was supported by FNP program First TEAM (Grants No. First TEAM/2016-1/5 and No. First TEAM/2017-4/31) and NCN Grants No. 2014/14/E/ST2/00020 and No. 2016/23/N/ST2/02817.

[1] S. Wehner, M. Christandl, and A. C. Doherty, Phys. Rev. A **78**, 062112 (2008).

[2] R. Gallego, N. Brunner, C. Hadley, and A. Acín, Phys. Rev. Lett. **105**, 230501 (2010).

[3] G. Brassard, Found. Phys. **33**, 1593 (2003).
[4] H. Buhrman, R. Cleve, S. Massar, and R. de Wolf, Rev. Mod. Phys. **82**, 665 (2010).
[5] J. Ahrens, P. Badziag, A. Cabello, and M. Bourennane, Nat. Phys. **8**, 592 (2012).
[6] M. Hendrych, R. Gallego, M. Mičuda, N. Brunner, A. Acín, and J. P. Torres, Nat. Phys. **8**, 588 (2012).
[7] V. D'Ambrosio, F. Bisesto, F. Sciarrino, J. F. Barra, G. Lima, and A. Cabello, Phys. Rev. Lett. **112**, 140503 (2014).
[8] J. Ahrens, P. Badziag, M. Pawłowski, M. Żukowski, and M. Bourennane, Phys. Rev. Lett. **112**, 140401 (2014).
[9] E. A. Aguilar, M. Farkas, D. Martínez, M. Alvarado, J. Cariñe, G. B. Xavier, J. F. Barra, G. Cañas, M. Pawłowski, and G. Lima, Phys. Rev. Lett. **120**, 230503 (2018).
[10] A. Tavakoli, M. Pawłowski, M. Żukowski, and M. Bourennane, Phys. Rev. A **95**, 020302(R) (2017).
[11] A. Ambainis, D. Leung, L. Mančinska, and M. Ozols, arXiv:0810.2937.
[12] S. Wiesner, SIGACT News **15**, 78 (1983).
[13] A. Ambainis, A. Nayak, A. Ta-Shma, and U. Vazirani, in *Proceedings of the 31st Annual ACM Symposium on Theory of Computing (STOC'99), Atlanta, 1999* (ACM, New York, 1999), pp. 376–383.
[14] A. Nayak, in *Proceedings of the 40th IEEE Symposium on Foundations of Computer Science (FOCS'99), New York, 1999* (IEEE, Piscataway, 1999), pp. 369–376.
[15] E. F. Galvão, Ph.D. thesis, University of Oxford, 2002.
[16] A. Casaccino, E. F. Galvão, and S. Severini, Phys. Rev. A **78**, 022310 (2008).
[17] A. Tavakoli, A. Hameedi, B. Marques, and M. Bourennane, Phys. Rev. Lett. **114**, 170502 (2015).
[18] A. Hameedi, D. Saha, P. Mironowicz, M. Pawłowski, and M. Bourennane, Phys. Rev. A **95**, 052345 (2017).
[19] A. Hameedi, B. Marques, P. Mironowicz, D. Saha, M. Pawłowski, and M. Bourennane, arXiv:1511.06179v2.
[20] E. A. Aguilar, J. J. Borkała, P. Mironowicz, and M. Pawłowski, Phys. Rev. Lett. **121**, 050501 (2018).
[21] M. Farkas and J. Kaniewski, arXiv:1803.00363v2.
[22] A. Ambainis, D. Kravchenko, and A. Rai, arXiv:1510.03045v1.
[23] P. Erdős and L. R. Graham, Acta Arithmetica **21**, 339 (1972).
[24] J. Dixmier, J. Number Theory **34**, 198 (1990).
[25] J. H. van Lint and R. M. Wilson, *A Course in Combinatorics*, 2nd ed. (Cambridge University Press, Cambridge, 2001).
[26] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski, Nature (London) **461**, 1101 (2009).

Semi-device-independent characterization of multipartite entanglement of states and measurements

Armin Tavakoli,[1] Alastair A. Abbott,[2] Marc-Olivier Renou,[1] Nicolas Gisin,[1] and Nicolas Brunner[1]
[1]*Département de Physique Appliquée, Université de Genève, CH-1211 Genève, Switzerland*
[2]*Université de Grenoble Alpes, CNRS, Grenoble INP, Institut Néel, 38000 Grenoble, France*

(Received 3 May 2018; published 26 November 2018)

The semi-device-independent framework allows one to draw conclusions about properties of an unknown quantum system under weak assumptions. Here we present a semi-device-independent scheme for the characterization of multipartite entanglement based around a game played by several isolated parties whose devices are uncharacterized beyond an assumption about the dimension of their Hilbert spaces. Our scheme can simultaneously certify that an n-partite high-dimensional quantum state features genuine multipartite entanglement, and that a joint measurement on n subsystems is entangled. Moreover, it provides a lower bound on the number of entangled measurement operators. These tests are strongly robust to noise, and even optimal for certain classes of states and measurements, as we demonstrate with illustrative examples. Notably, our scheme allows the certification of many entangled states admitting a local model, which therefore cannot violate any Bell inequality.

DOI: 10.1103/PhysRevA.98.052333

I. INTRODUCTION

Entanglement represents a central feature of quantum theory and a key resource for quantum information processing [1]. Therefore, the task of characterizing entanglement experimentally is of fundamental significance. In particular, the development of quantum networks, multiparty cryptography, quantum metrology, and quantum computing necessitates certification methods tailored to multipartite entangled states, as well as to entangled joint measurements.

Standard methods for the certification of multipartite entanglement rely on entanglement witnesses [2], as quantum tomography quickly becomes infeasible as the number of subsystems increases. Entanglement witnesses can also be used for the particularly important subcase of certifying genuine multipartite entanglement (GME), the strongest form of multipartite entanglement, where all subsystems are genuinely entangled together; see, e.g., Refs. [1–3]. In practice, the main drawback of entanglement witnesses is that they crucially rely on the correct calibration of the measurement devices, as a set of specific observables must be measured. Importantly, even small alignment errors can have undesirable consequences, e.g., leading to false positives [4], and it is generally cumbersome to estimate these errors and take them into account rigorously.

This motivates the development of certification methods that require minimal assumptions on the measurement devices and in particular do not rely on their detailed characterization. This is the spirit of the device-independent (DI) approach to entanglement characterization, which also leads to interesting possibilities for quantum information processing [5–7]. The idea is to use the violation of Bell inequalities, as these necessarily imply that the state is entangled even without any knowledge about the measuring devices. Moreover, GME can also be detected via Bell-like inequalities [8–15]. Experimentally, however, this approach is very demanding as high visibilities are typically required. More generally, a broad range of entangled states (including many GME states [16–18]) cannot violate any Bell inequality [19] as they admit local hidden variable models [20,21]. Finally, although Bell inequalities can in principle be used for the certification of entangled joint measurements [22], no practical scheme has been reported thus far.

This motivates the exploration of partially DI scenarios, in between the fully DI case of Bell inequalities and the device-dependent case of entanglement witnesses. Here, only weak assumptions about the devices are typically made. One possibility is to consider that a subset of parties perform well-characterized measurements, while the others are uncharacterized [23–25]. Another option is to consider Bell experiments with quantum inputs, leading to the so-called measurement-device-independent characterization of entanglement [26,27]. While experimental demonstrations have been reported, both of these approaches have the drawback of requiring certain parts of the experiment to be fully characterized.

In the present work, we follow a different approach for entanglement characterization. Specifically, we will assume only an upper bound on the Hilbert space dimension of the subsystems of interest, but require no detailed characterization of any of the devices. Roughly speaking, this assumption means that all the relevant degrees of freedom are described in a Hilbert space of given dimension [28–30], and that other potential side channels can be neglected. This scenario, usually referred to as the semi-DI (SDI) setting, has been considered for the characterization of entanglement in the simplest setting of two-qubit states [31,32], as well as the detection of two-qubit entangled measurements [33,34].

Here, we present a versatile scheme for characterizing multipartite entanglement in both states and measurements simultaneously in a semi-DI setting that significantly improves previous approaches. Our scheme allows one to simultaneously certify that (I) an n-partite quantum state (of arbitrary local

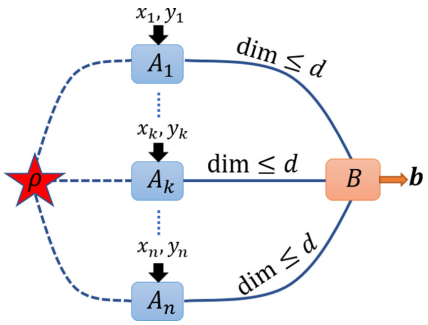

FIG. 1. An n-partite state is distributed between parties $A_1, \ldots, A_n$ who perform transformations on their local systems depending on their random inputs (x_k, y_k). Each party sends their transformed d-dimensional system to B who performs a measurement and obtains an outcome $\mathbf{b}$.

dimension) is GME, and that (II) a measurement performed on the n subsystems is entangled. Furthermore, we obtain a finer characterization for the measurement, namely, a lower bound on the number of entangled measurement operators. In general our scheme is strongly robust to noise, and even optimal in certain cases. It certifies all noisy qubit Greenberger-Horne-Zeilinger (GHZ) states that are GME, and, in the bipartite case, all entangled isotropic states of arbitrary dimension. Other classes of GME states, e.g., Dicke states, can also be certified, although not optimally. For the case of entangled measurements, we give two illustrative examples. In particular, we optimally certify the presence of entanglement in a noisy Bell-state measurement. Finally, we conclude with a list of open questions.

II. SCENARIO

The scenario we consider consists of a state being initially prepared, then transformed by several separated parties, and finally measured. Since we operate within the SDI framework no assumptions are made on the internal workings of any of these parties' devices, other than a bound on their local Hilbert space dimensions, see Fig. 1.

Let an uncharacterized source distribute a state ρ of arbitrary dimension between n parties $A_1, \ldots, A_n$, each of which receives a subsystem. Each party A_k receives uniformly random inputs $x_k, y_k \in \{0, \ldots, d-1\}$. Subsequently, they perform local transformations $\mathcal{T}^{(k)}_{x_k y_k}$ [some completely positive trace-preserving (CPTP) map] which map their local states into a d-dimensional state. The transformed state is sent to a final party denoted by B who performs a measurement $\{M_{\mathbf{b}}\}_{\mathbf{b}}$ (i.e., $M_{\mathbf{b}} \geqslant 0$ and $\sum_{\mathbf{b}} M_{\mathbf{b}} = \mathbb{1}$) which produces an outcome string $\mathbf{b} = b_1 \ldots b_n \in \{0, \ldots, d-1\}^n$ (see Fig. 1).The experiment generates a probability distribution $P(\mathbf{b}|\mathbf{x}, \mathbf{y})$, where $\mathbf{x} = x_1 \ldots x_n$ and $\mathbf{y} = y_1 \ldots y_n$, given by

$$P(\mathbf{b}|\mathbf{x}, \mathbf{y}) = \mathrm{tr}\left[\left(\bigotimes_{k=1}^{n} \mathcal{T}^{(k)}_{x_k y_k}\right)[\rho] M_{\mathbf{b}}\right]. \tag{1}$$

The goal of the task we consider is for the parties to cooperate so that B's output satisfies the conditions

$$b_1 = \sum_{i=1}^{n} x_i \equiv C_1(x) \quad \text{and} \quad b_k = y_k - y_1 \equiv C_k(y), \tag{2}$$

for $k = 2, \ldots, n$, where all quantities are computed modulo d. Compactly, we write $C(\mathbf{x}, \mathbf{y})$ for the (unique) string $\mathbf{b}$ satisfying all the above conditions. Note that these conditions are not totally symmetric (except when $n = 2$). Given a strategy leading to a probability distribution $P(\mathbf{b}|\mathbf{x}, \mathbf{y})$, the probability of winning the task (or if a win is rewarded with a point, the average score) is thus given by

$$\mathcal{A}_{n,d} = \frac{1}{d^{2n}} \sum_{\mathbf{x},\mathbf{y}} P(\mathbf{b} = C(\mathbf{x}, \mathbf{y}) \,|\, \mathbf{x}, \mathbf{y}). \tag{3}$$

We now show how, from the value of an observed average score $\mathcal{A}_{n,d}$, one can make inferences about the entanglement of the state ρ and the measurement $\{M_{\mathbf{b}}\}_{\mathbf{b}}$.

III. CHARACTERIZING ENTANGLEMENT OF STATES

We first consider certifying the GME of the shared state. A state is said to be GME if it is not biseparable, i.e., if it cannot be written in the form $\rho = \sum_S \sum_i p_{S,i} \rho_i^S \otimes \rho_i^{\bar{S}}$, for any possible bipartition $\{S, \bar{S}\}$ of the subsystems $\{1, \ldots, n\}$, where $\sum_{S,i} p_{S,i} = 1$ and $p_{S,i} \geqslant 0$.

We now show that the value of $\mathcal{A}_{n,d}$ can be nontrivially upper bounded for any n-partite biseparable state. This will allow us to certify GME of many states of interest.

Result 1. Let ρ be a state of n subsystems. For any measurement $\{M_{\mathbf{b}}\}_{\mathbf{b}}$ and any transformations $\{\mathcal{T}^{(k)}_{x_k y_k}\}_k$, it holds that

$$\rho \text{ is biseparable} \Rightarrow \mathcal{A}_{n,d} \leqslant 1/d. \tag{4}$$

Hence, whenever $\mathcal{A}_{n,d} > 1/d$, ρ is certified to be GME. Moreover, this inequality is tight and the bound can be saturated with fully separable states.

Proof. The full details of the proof are given in Appendix A. In order to prove the upper bound in Eq. (4), we consider a relaxed SDI task in which any distribution $P(\mathbf{b}|\mathbf{x}, \mathbf{y})$ obtainable in the original task is also possible in the relaxed setting, but not vice versa. The relaxation is chosen so that the average score $\mathcal{A}_{n,d}$ can easily be upper bounded using a result of Ref. [35]. By construction, the upper bound obtained in the relaxed scenario is also valid for the original task.

To see that the bound is tight, we give a strategy that utilizes only product states and saturates the bound (4). Let A_k (for $k = 1, \ldots, n$) send y_k to party B. B then outputs $b_i = y_i - y_1$, satisfying condition C_i, for $i = 2, \ldots, n$. However, this strategy forces B to guess b_1 in order to satisfy condition C_1. Any such guess succeeds, on average, with probability $1/d$, thus saturating the bound. ■

To show the relevance of the relation in Eq. (4), we show that it can be violated by GME states. In particular, we first consider the largest achievable value of $\mathcal{A}_{n,d}$. It turns out that the algebraically maximal value, i.e., $\mathcal{A}_{n,d} = 1$, can be achieved for all n and d via the following strategy. A GME state of n subsystems of local dimension d, namely,

the generalized GHZ state, $|\mathrm{GHZ}_{n,d}\rangle = 1/\sqrt{d}\sum_{i=0}^{d-1}|i\rangle^{\otimes n}$, is distributed among the parties $A_1,\ldots,A_n$. Each party then performs the unitary transformation $U_{x_k y_k}^{A_k} = Z^{x_k}X^{y_k}$ for $k = 1,\ldots,n$, where

$$Z = \sum_{j=0}^{d-1} e^{2i\pi j/d}|j\rangle\langle j|, \qquad X = \sum_{j=0}^{d-1}|j+1\rangle\langle j| \tag{5}$$

are the usual clock and shift operators. Finally, B performs a joint projective measurement in the basis of generalized GHZ states given by

$$|M_b\rangle = Z^{b_1}\otimes X^{b_2}\otimes\cdots\otimes X^{b_n}|\mathrm{GHZ}_{n,d}\rangle. \tag{6}$$

Note that in the simplest case of two qubits ($n = d = 2$), the four unitaries are simply the three Pauli matrices and the identity matrix, while the measurement is the Bell-state measurement [36].

Let us now consider noisy GHZ states: $\rho_{n,d}^{\mathrm{GHZ}}(v) = v\,|\mathrm{GHZ}_{n,d}\rangle\langle\mathrm{GHZ}_{n,d}| + (1-v)\mathbb{1}/d^n$, where $v\in[0,1]$ is the visibility of the state. For the above strategy, one has $\mathcal{A}_{n,d}(\mathbb{1}/d^n) = 1/d^n$. Hence, from the linearity of $\mathcal{A}_{n,d}$ in ρ, it follows that a violation of Eq. (4) is obtained whenever $v + (1-v)/d^n > 1/d$, that is, when $v > (d^{n-1}-1)/(d^n-1)$. We discuss the implications of this result in three separate cases of interest.

(I) For two d-dimensional systems ($n = 2$), the criterion is $v > 1/(d+1)$, which is precisely the condition for the entanglement of $\rho_{2,d}^{\mathrm{GHZ}}(v)$ [37]. Hence, every entangled isotropic bipartite state is certified by our protocol. Interestingly, such certification is impossible using Bell inequalities; for instance, we note that the state $\rho_{2,2}^{\mathrm{GHZ}}(v)$ has a local hidden variable model (for projective measurements) when $v < 0.6829$ [38], and it will therefore not violate any Bell inequality. For large d, such models are known for $v \lesssim \ln d/d$ [39], while Bell inequality violations are possible for $v\sim \ln d/\sqrt{d}$ [40]. In contrast, our protocol can certify entanglement for $v\sim 1/d$.

(II) In the case of a system of many qubits ($d = 2$), our visibility criterion coincides with the condition for $\rho_{n,2}^{\mathrm{GHZ}}(v)$ to be GME [41]. Hence our scheme can certify all noisy qubit GHZ states that are GME. Again, this would not be possible using Bell inequalities, as (for instance) the state $\rho_{3,2}^{\mathrm{GHZ}}(v = 1/2)$ is GME but admits a biseparable model reproducing all correlations from projective measurements [10]. Furthermore, the GME of $\rho_{3,2}^{\mathrm{GHZ}}(v)$ is known to be certifiable via a Bell inequality when $v > 0.64$ [11], whereas our criterion reads $v > 3/7$.

(III) When considering many high-dimensional systems ($n > 2$ and $d > 2$), our setup is no longer optimal since there exist generalized GHZ states that are GME below the critical visibility of our scheme [42]. Nevertheless, choosing, for instance, $n = d = 3$, the criterion is $v > 4/13$ which substantially outperforms the certification obtainable via known Bell inequalities, which is possible when $v > 0.81$ [10,14].

More generally, we derive a lower bound on the maximal value of $\mathcal{A}_{n,d}$ achievable for an arbitrary state ρ. To this end, consider the following strategy. Let B perform the measurement given by Eq. (6), and let each party A_k perform the transformation $\mathcal{T}_{x_k y_k}^{(k)}[\rho] = (U_{x_k y_k}^{A_k})\,\Lambda_k[\rho]\,(U_{x_k y_k}^{A_k})^\dagger$, where $U_{x_k y_k}^{A_k} = Z^{x_k}X^{y_k}$ and the Λ_k, for $k = 1,\ldots,n$, are CPTP maps. Evaluating the score with this strategy and optimizing over the CPTP maps $\Lambda_1,\ldots,\Lambda_n$ straightforwardly leads to

$$\mathcal{A}_{n,d}(\rho) = \mathrm{EGF}_{n,d}(\rho), \tag{7}$$

where we have defined the quantity $\mathrm{EGF}_{n,d}(\rho) = \max_{\Lambda_1,\ldots,\Lambda_n}\mathrm{tr}\,[(\bigotimes_{k=1}^n\Lambda_k)[\rho]|\mathrm{GHZ}_{n,d}\rangle\langle\mathrm{GHZ}_{n,d}|]$. If one instead maximizes only over unitary maps $\Lambda_k[\rho] = V_k\rho V_k^\dagger$ one obtains $A_{n,d} = \mathrm{GF}_{n,d}(\rho)$, where $\mathrm{GF}_{n,d}(\rho) = \max_{V_1,\ldots,V_n}\mathrm{tr}\,[(\bigotimes_{k=1}^n V_k)\rho(\bigotimes_{k=1}^n V_k)^\dagger|\mathrm{GHZ}_{n,d}\rangle\langle\mathrm{GHZ}_{n,d}|]$ is the *GHZ fraction* [43], a multipartite generalization of the singlet fraction [37,44]. $\mathrm{EGF}_{n,d}(\rho)$ can then be seen as the "extractable" GHZ fraction, an important generalization since, even in the bipartite case, local CPTP maps can increase the singlet fraction of an entangled state [45].

To determine whether one can obtain a better score than that given by Eq. (7) by considering arbitrary transformations and measurements, we conducted extensive numerical tests. Focusing on the cases $(n,d)\in\{(2,2),(2,3),(3,2)\}$ and optimizing numerically $\mathcal{A}_{n,d}$ starting from randomly chosen transformations and measurements, we were unable to obtain a better score than $\mathrm{EGF}_{n,d}(\rho)$. We note also that when restricted to unitary transformations, we were similarly unable to obtain a score larger than $\mathrm{GF}_{n,d}(\rho)$. Motivated by this numerical evidence, we make the following conjecture:

Conjecture. Let ρ be an n-partite state of local dimension d. Then the maximal value of $\mathcal{A}_{n,d}$ achievable for any measurement $\{M_{\mathbf{b}}\}_{\mathbf{b}}$ and transformations $\{\mathcal{T}_{x_k y_k}^{(k)}\}_k$ is $\mathrm{EGF}_{n,d}(\rho)$, i.e.,

$$\max_{\{\mathcal{T}_{x_k y_k}^{(k)}\}_k,\{M_{\mathbf{b}}\}_{\mathbf{b}}}\mathcal{A}_{n,d}(\rho) = \mathrm{EGF}_{n,d}(\rho). \tag{8}$$

Proving this conjecture would be particularly interesting in the bipartite case, as violation of our witness would then certify an extractable singlet fraction greater than $1/d$, which implies that maximal entanglement can be distilled from the state ρ [37].

Finally, we discuss other classes of GME states that are qualitatively inequivalent to GHZ states. First, consider noisy W states of three qubits, i.e., $W(v) = v\,|W\rangle\langle W| + (1-v)\mathbb{1}/8$, where $|W\rangle = \frac{1}{\sqrt{3}}(|001\rangle + |010\rangle + |100\rangle)$. Numerical optimization gives a (seemingly optimal) strategy obtaining $\mathcal{A}_{3,2} = \frac{1}{8}(1+5v)$. This implies that our scheme certifies the GME of $W(v)$ for $v > 3/5$. This is relatively close to the optimal visibility for GME of $v > 0.48$ [46]. In comparison, known DI schemes based on Bell inequalities would require $v > 0.72$ [10].

Second, consider a noisy four-qubit Dicke state $D(v) = v\,|D\rangle\langle D| + (1-v)\mathbb{1}/16$, where $|D\rangle = \frac{1}{\sqrt{6}}(|0011\rangle + |0101\rangle + |0110\rangle + |1001\rangle + |1010\rangle + |1100\rangle)$. Numerically, the best strategy we find certifies the GME of $D(v)$ for $v > 7/11\approx 0.64$, while it is known to be GME for $v > 0.46$ [46].

IV. CHARACTERIZING ENTANGLEMENT OF MEASUREMENTS

Next, we consider characterizing the entanglement of the joint measurement performed by B. A measurement $\{M_{\mathbf{b}}\}_{\mathbf{b}}$ is said to be entangled if at least one measurement operator $M_{\mathbf{b}}$

does not have a fully separable decomposition $M_{\mathbf{b}} = \sum_i M_{\mathbf{b},i}$, where $M_{\mathbf{b},i} \geqslant 0$ and each $M_{\mathbf{b},i}$ has the tensor product form $M_{\mathbf{b},i} = \bigotimes_{k=1}^{n} M_{\mathbf{b},i}^{(k)}$. We will see that the separability of $\{M_{\mathbf{b}}\}_{\mathbf{b}}$ imposes a nontrivial bound on $\mathcal{A}_{n,d}$ which will allow us to certify the entanglement of the measurement used. However, it is sufficient for a single measurement operator to be entangled in order for the measurement to be entangled. This qualitative property rules out a classical description, but reveals little about the extent to which entanglement is present in the measurement. Interestingly, we can go a step further and show that the value of $\mathcal{A}_{n,d}$ implies a bound on the minimum number of the d^n measurement operators that are entangled. This provides a much finer characterization of the joint entangled measurement performed by B.

Result 2. Let $\{M_{\mathbf{b}}\}_{\mathbf{b}}$ be a joint measurement of an n-partite system of local dimension d with at least $k \in \{0, \ldots, d^n\}$ fully separable measurement operators. For any n-partite state ρ and any transformations $\{\mathcal{T}_{x_k y_k}^{(k)}\}_k$ it holds that

$$\text{At least } k \text{ separable measurement operators} \Rightarrow \mathcal{A}_{n,d} \leqslant \frac{1}{d^n}\left(d^n - k + \frac{k}{d}\right). \tag{9}$$

Hence, a violation of this inequality for a particular k implies that at least $d^n - k + 1$ measurement operators are entangled.

The proof of this result is given in Appendix B.

In the extremal case of a fully separable measurement ($k = d^n$), the bound (9) reduces to $\mathcal{A}_{n,d} \leqslant 1/d$, and observing a violation of this bound certifies the entanglement of the measurement. In the other extreme of witnessing a measurement whose operators are all entangled [i.e., violating Eq. (9) for $k = 1$], the bound (9) is $\mathcal{A}_{n,d} \leqslant 1 - (d-1)/d^{n+1}$. In Appendix C we give some partial results on the tightness of Eq. (9).

To demonstrate the usefulness of Result 2, we now discuss two illustrative examples in the bipartite case. First, consider a noisy version of a generalized Bell-state measurement (for arbitrary d), which is a joint measurement of two d-level systems given by the projection onto a basis of d^2 maximally entangled states $\{|M_{b_1 b_2}\rangle\}_{b_1 b_2}$. To this end, we consider the measurement operators $M_{b_1 b_2}(v) = v\,|M_{b_1 b_2}\rangle\langle M_{b_1 b_2}| + (1 - v)\mathbb{1}/d^2$. Note that each of these measurement operators is equivalent (up to local unitaries) to an isotropic state $\rho_{2,d}^{\mathrm{GHZ}}(v)$, and is thus entangled precisely when $v > 1/(d+1)$. To certify the entanglement of this measurement, consider again the strategy used previously in which the parties share a maximally entangled state of two d-level systems, and perform the unitary transformations $U_{x_k y_k}^{A_k} = Z^{x_k} X^{y_k}$. Note that the score obtained here is the same as when considering a shared noisy isotropic state $\rho_{2,d}^{\mathrm{GHZ}}(v)$, combined with an ideal Bell-state measurement. It thus follows from the previous results that we can certify the entanglement of the measurement whenever $v > 1/(d+1)$. Hence, we certify entanglement whenever the level of noise is low enough to keep the measurement operators entangled. Nevertheless, note that in this case our witness certifies only that at least one measurement operator is entangled, while all the measurement operators are in fact entangled. To certify the entanglement of all d^2 measurement operators, a much higher visibility of $v > (d^2 + d - 1)/(d^2 + d)$ is required.

Second, we consider a two-qubit Bell-state measurement subjected to colored noise for the case $n = d = 2$. Defining the usual Bell basis $|\phi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle)$, $|\psi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle)$, and $\Phi^{\pm} = |\phi^{\pm}\rangle\langle\phi^{\pm}|$, $\Psi^{\pm} = |\psi^{\pm}\rangle\langle\psi^{\pm}|$, consider the measurement given by the operators

$$E_{00} = v\Phi^+ + \frac{1-v}{4}(\Phi^+ + 2\Phi^- + \Psi^+), \tag{10}$$

$$E_{01} = v\Psi^+ + \frac{1-v}{4}(\Phi^+ + \Phi^- + 2\Psi^+), \tag{11}$$

$$E_{10} = v\Phi^- + \frac{1-v}{4}(2\Phi^+ + \Phi^- + \Psi^+), \tag{12}$$

$$E_{11} = \Psi^-, \tag{13}$$

for some visibility $v \in [0, 1]$. When $1/3 < v \leqslant 1$, all four operators are entangled; when $0 < v \leqslant 1/3$, only E_{01} and E_{11} are entangled; and when $v = 0$, only E_{11} is entangled.

Considering the same strategy as in the previous example (i.e., sharing a maximally entangled state and applying the transformations $U_{x_k y_k}^{A_k}$), we find $\mathcal{A}_{2,2} = (1 + v)/2$. By virtue of Eq. (9), this certifies the presence of four entangled measurement operators when $v > 3/4$, at least three when $v > 1/2$, at least two when $v > 1/4$, and at least one when $v > 0$.

V. CONCLUSION

We presented a scheme for the semi-DI characterization of high-dimensional and multipartite entanglement. It can simultaneously certify that states are GME and provide a lower bound on the number of measurement operators that are entangled. It is highly robust to noise and even certifies the entanglement of some families of states and measurements optimally. Since many entangled states admitting a local model can be certified in our scheme, it overcomes the fundamental limitations of entanglement certification using Bell inequalities at the price of a bound on Hilbert space dimension.

The noise robustness of our setup makes it promising for experimental tests, in particular for atomic and solid-state systems where both entangled states and entangled measurements have been demonstrated [47,48]. An interesting question is whether our scheme could be adapted to optical setups, where only partial entangled measurements are typically available.

We conclude with some relevant open questions. (I) Does Eq. (9) hold also for biseparable measurement operators, thereby allowing the number of GME measurement operators to be bounded? (II) Can a more sophisticated semi-DI scheme certify the entanglement of all GME states (not just particular classes of states, as shown here)? (III) Can our scheme be used as a dimension witness for both states and measurements? (IV) Note that a score $\mathcal{A}_{n,d} = 1$ implies that all the relations (2) are satisfied. This makes the scheme a good candidate for multiparty cryptographic tasks (e.g., secret sharing of classical data with quantum resources). Exploring this possibility would be interesting. (V) Can the scheme be used for self-testing states, measurements, and transformations in prepare-and-measure experiments [49]?

ACKNOWLEDGMENTS

We thank Cyril Branciard and Jędrzej Kaniewski for discussions and comments. This work was supported by the Swiss National Science Foundation (Starting Grant DIAQ, NCCR QSIT and SwissMAP) and the French National Research Agency (Retour Post-Doctorants Programme Grant ANR-13-PDOC-0026).

APPENDIX A: PROOF OF RESULT 1

In this Appendix, we prove result 1. Specifically, we bound the maximal value of $\mathcal{A}_{n,d}$ that can be obtained by biseparable states, regardless of the choice of transformations and measurements. Since $\mathcal{A}_{n,d}$ depends linearly on ρ, no mixed biseparable state can be used to obtain a larger score than some pure biseparable state. Hence, we need only consider pure states of the form $|\chi\rangle_S \equiv |\psi\rangle_S \otimes |\phi\rangle_{\bar{S}}$, for any nontrivial bipartition $\{S, \bar{S}\}$ of the set of subsystems $\{1, \ldots, n\}$.

Consider $|\chi\rangle_S$ for a particular bipartition $\{S, \bar{S}\}$. To give an upper bound on $\mathcal{A}_{n,d}$, we will relax some of the constraints in the scenario considered by the scheme and then evaluate the maximal average score in this less restrictive setting. In particular, we consider the scenario in which the parties $\{A_k\}_{k\in S}$ are permitted to communicate unbounded information to B, while the remaining parties $\{A_k\}_{k\in\bar{S}}$ are grouped into a single "effective" party R, and which receives all their inputs. The party R is allowed to send $|\bar{S}|$ d-level quantum systems (equivalently, a system of dimension $d^{|\bar{S}|}$) to B. This scenario is illustrated in Fig. 2.

It is clear that any probability distribution $P(\mathbf{b}|\mathbf{x}, \mathbf{y})$ obtainable in the original scenario on a state $|\chi\rangle_S$ (i.e., with some choice of transformations and measurements) is also obtainable in the relaxed scenario, but not vice versa. Since, by assumption, there is no entanglement over the bipartition $\{S, \bar{S}\}$, the parties $\{A_k\}_{k\in S}$ cannot do better than to simply send all their inputs to B. To win the game, R therefore needs to communicate to B the values of $\sum_{i\in\bar{S}} x_i$ and all $\{y_i\}_{i\in\bar{S}}$ for the conditions $C_1, \ldots, C_n$ to be satisfied. However, this amounts to $(|\bar{S}| + 1)\log_2 d$ bits of information, while R can generally only send $|\bar{S}|\log_2 d$ bits using a $d^{|\bar{S}|}$-level system, and must therefore employ a nontrivial optimal communication strategy. As we demonstrate more formally below, there is no quantum strategy that allows B to know this information with a probability higher than $1/d$, and therefore cannot win the game with a probability better than this either. Consequently, the desired bound follows.

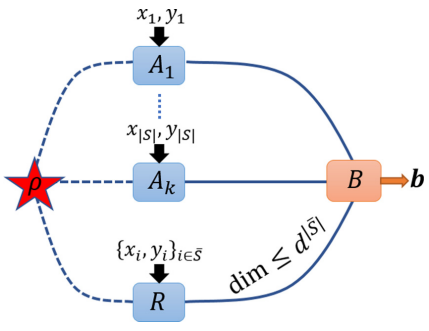

FIG. 2. Modified scenario (cf. Fig. 1). Parties $\{A_k\}_{k\in S}$ are allowed unlimited communication to B, whereas the remaining parties $\{A_k\}_{k\in\bar{S}}$ are grouped into a single party, R, allowed $\bar{S}\log_2 d$ bits of communication. The case shown is for $S = \{1, \ldots, |S|\}$.

To this end, let us consider the following general setting for an information compression task between two parties, Alice and Bob. Let Alice receive a uniformly distributed input $x \in \{1, \ldots, N\}$ which she must communicate to Bob. She encodes this input into a quantum state ρ_x of dimension at most d, with $N > d$ (if $N \leqslant d$ then Alice can simply send x). This amounts to Alice compressing her input into a smaller message to send. This message is then sent to Bob, who must attempt to retrieve the value of x from the state ρ_x by performing a suitable measurement $\{M_b\}_{b=1}^N$. What is the average probability of success for Bob to correctly obtain x following this measurement? As is shown, e.g., in Ref. [35], any quantum strategy must obey the following bound:

$$\begin{aligned} p^{\text{success}} &\equiv \frac{1}{N}\sum_{x=1}^{N} P^{\text{Q}}(b = x|x) = \frac{1}{N}\sum_{x=1}^{N} \text{tr}\,(\rho_x M_x) \\ &\leqslant \frac{1}{N}\sum_{x=1}^{N} \lambda_{\max}(M_x) \leqslant \frac{1}{N}\sum_{x=1}^{N} \text{tr}\,(M_x) \\ &= \frac{1}{N}\,\text{tr}\left(\sum_{x=1}^{N} M_x\right) = \frac{d}{N}, \end{aligned} \tag{A1}$$

where we have used the fact that the optimal state ρ_x maximizing $\text{tr}(\rho_x M_x)$ is the eigenvector of M_x corresponding to its largest eigenvalue, and that $\lambda_{\max}(M_x) \leqslant \text{tr}\,(M_x)$. In the last step we used that $\sum_x M_x = \mathbb{1}_d$ for any positive-operator valued measure (POVM) $\{M_x\}_x$.

Moreover, there is no quantum advantage over classical approaches to encode and decode this information. This is straightforwardly seen by noting that the quantum bound (A1) can be saturated with a classical strategy as follows. Let Alice send the classical message $m(x) = x$ whenever $x = 1, \ldots, d$, and $m(x) = d$ if $x \in \{d + 1, \ldots, N\}$. Bob always outputs the message he receives, i.e., $b = m(x)$. Whenever $x \in \{1, \ldots, d\}$, it is indeed true that $b = x$ and Bob correctly obtains x; when $x \in \{d + 1, \ldots, N\}$, Bob never succeeds. The average success probability of this simple classical strategy reads

$$p^{\text{success}} \equiv \frac{1}{N}\sum_{x=1}^{N} P^{\text{C}}(b = x|x) = \frac{d}{N}. \tag{A2}$$

Hence, quantum theory provides no advantage over classical coding for the stated task. Note that the above bears resemblance to the Holevo bound, with the difference that we are probabilistically accessing the information.

We can now apply this result to the relaxed scenario under consideration in the proof of Result 1. There, the effective party R plays the role of Alice and has to transmit to Bob which of the $d^{|\bar{S}|+1}$ possible inputs they received by encoding it in a $d^{|\bar{S}|}$-dimensional quantum system. From the above

result, the average success probability of B correctly guessing the inputs of R—and therefore winning the game—cannot be better than $1/d$, which is indeed the desired result.

APPENDIX B: PROOF OF RESULT 2

Here we present the proof of result 2. We begin with a useful lemma, which is straightforward:

Lemma 1. Let σ_{AB} be a bipartite density matrix in $\mathcal{H}_A \otimes \mathcal{H}_B$ and $0 \leqslant M \leqslant \mathbb{1}$ (N) be a symmetric operator on $\mathcal{H}_A$ ($\mathcal{H}_B$). Let $\sigma_A = \mathrm{tr}_B[\sigma_{AB}]$. Then the following inequality holds:

$$\mathrm{tr}_{AB}[\sigma_{AB}(M \otimes N)] \leqslant \mathrm{tr}_A[\sigma_A M]\lambda_{\max}[N]. \tag{B1}$$

Proof. Let $N = \sum_i n_i |i\rangle\langle i|$ be the spectral decomposition of N. Remark that $\sigma_A^{(i)} = \langle i|\sigma_{AB}|i\rangle$ is an (unnormalized) positive semidefinite operator. We then have $\mathrm{tr}_{AB}[\sigma_{AB}(M \otimes N)] = \sum_i n_i \mathrm{tr}_A[\sigma_A^{(i)} M] \leqslant \lambda_{\max}[N]\,\mathrm{tr}_A[\sum_i \sigma_A^{(i)} M] = \lambda_{\max}[N]\,\mathrm{tr}_A[\sigma_A M]$. ∎

Equipped with this lemma, we can now prove the statement of result 2. Party B performs a measurement with d^n different possible outcomes, described by measurement operators $\{M_{\mathbf{b}}\}_{\mathbf{b}}$ with $M_{\mathbf{b}} \geqslant 0$ and $\sum_{\mathbf{b}} M_{\mathbf{b}} = \mathbb{1}$. Let SEP be the set of strings $\mathbf{b}$ for which $M_{\mathbf{b}}$ is a separable measurement operator, with $|\mathrm{SEP}| \geqslant k$; for $\mathbf{b} \in \mathrm{SEP}$ we thus have $M_{\mathbf{b}} = \sum_i \bigotimes_{k=1}^n M_{\mathbf{b},i}^{(k)}$. We will initially assume for simplicity that these separable POVM elements have the simpler tensor product form $M_{\mathbf{b}} = \bigotimes_{k=1}^n M_{\mathbf{b}}^{(k)}$; we will see later that because of the linearity of $\mathcal{A}_{n,d}$, the proof nonetheless holds for arbitrary separable operators.

The parties $A_1, \ldots, A_n$ may perform arbitrary transformations $\mathcal{T}_{x_k y_k}^{(k)}$, which are formally represented by completely positive trace-preserving (CPTP) maps. Such transformations can always be written as a unitary $U_{x_k y_k}^{(k)}$ applied jointly to the local system and some environment state $\xi_{\mathcal{E}_k}$, with the environment being subsequently traced out. Formally, we have

$$\mathcal{T}_{x_k y_k}^{(k)} : \sigma \mapsto \mathrm{tr}_{\mathcal{E}_k}\left[U_{x_k y_k}^{(k)}(\sigma \otimes \xi_{\mathcal{E}_k})U_{x_k y_k}^{(k)\dagger}\right]. \tag{B2}$$

We denote the Hilbert space for each party by $\mathcal{S}_k$ and the total Hilbert space of the parties by $\mathcal{S} = \bigotimes_{k=1}^n \mathcal{S}_k$. Similarly, we denote the local and total Hilbert spaces of the environment by $\mathcal{E}_k$ and $\mathcal{E} = \bigotimes_{k=1}^n \mathcal{E}_k$, respectively. The total initial environment state is thus $\xi_{\mathcal{E}} = \bigotimes_{k=1}^n \xi_{\mathcal{E}_k}$. (Note that in our SDI scheme we only assume a bound on the dimension of the output space of each party so, *a priori*, this may be different from that of the input spaces so that the $U_{x_k y_k}^{(k)}$ instead map $\mathcal{S}_k^i \otimes \mathcal{E}_k^i \to \mathcal{S}_k^f \otimes \mathcal{E}_k^f$. For simplicity we assume the input and output spaces have the same dimensions in the proof below; the argument generalizes trivially to the more general case.)

Evaluating explicitly $\mathcal{A}_{n,d}$, we have

$$\begin{aligned}
\mathcal{A}_{n,d} &= \frac{1}{d^{2n}} \sum_{\substack{\mathbf{x},\mathbf{y},\mathbf{b}:\\ \mathbf{b}=C(\mathbf{x},\mathbf{y})}} \mathrm{tr}_{\mathcal{S}}\left[\left(\bigotimes_{k=1}^n \mathcal{T}_{x_k y_k}^{(k)}\right)[\rho]M_{\mathbf{b}}\right] \\
&= \frac{1}{d^{2n}} \sum_{\substack{\mathbf{x},\mathbf{y},\mathbf{b}:\\ \mathbf{b}=C(\mathbf{x},\mathbf{y})}} \mathrm{tr}_{\mathcal{S}\mathcal{E}}\left[\left(\bigotimes_{k=1}^n U_{x_k y_k}^{(k)}\right)(\rho \otimes \xi_{\mathcal{E}})\left(\bigotimes_{k=1}^n U_{x_k y_k}^{(k)\dagger}\right)(M_{\mathbf{b}} \otimes \mathbb{1}_{\mathcal{E}})\right] \\
&= \frac{1}{d^{2n}} \sum_{\substack{\mathbf{x},\mathbf{y},\mathbf{b}\in\mathrm{SEP}:\\ \mathbf{b}=C(\mathbf{x},\mathbf{y})}} \mathrm{tr}_{\mathcal{S}\mathcal{E}}\left[(\rho \otimes \xi_{\mathcal{E}})\bigotimes_{k=1}^n \left\{U_{x_k y_k}^{(k)\dagger}\left(M_{\mathbf{b}}^{(k)} \otimes \mathbb{1}_{\mathcal{E}_k}\right)U_{x_k y_k}^{(k)}\right\}\right] \\
&\quad + \frac{1}{d^{2n}} \sum_{\substack{\mathbf{x},\mathbf{y},\mathbf{b}\notin\mathrm{SEP}:\\ \mathbf{b}=C(\mathbf{x},\mathbf{y})}} \mathrm{tr}_{\mathcal{S}\mathcal{E}}\left[(\rho \otimes \xi_{\mathcal{E}})\left(\bigotimes_{k=1}^n U_{x_k y_k}^{(k)\dagger}\right)(M_{\mathbf{b}} \otimes \mathbb{1}_{\mathcal{E}})\left(\bigotimes_{k=1}^n U_{x_k y_k}^{(k)}\right)\right].
\end{aligned} \tag{B3}$$

Restricting ourselves to the first term above, we have

$$\begin{aligned}
T &\equiv \frac{1}{d^{2n}} \sum_{\substack{\mathbf{x},\mathbf{y},\mathbf{b}\in\mathrm{SEP}:\\ \mathbf{b}=C(\mathbf{x},\mathbf{y})}} \mathrm{tr}_{\mathcal{S}\mathcal{E}}\left[(\rho \otimes \xi_{\mathcal{E}})\bigotimes_{k=1}^n \left\{U_{x_k y_k}^{(k)\dagger}\left(M_{\mathbf{b}}^{(k)} \otimes \mathbb{1}_{\mathcal{E}_k}\right)U_{x_k y_k}^{(k)}\right\}\right] \\
&\leqslant \frac{1}{d^{2n}} \sum_{\substack{\mathbf{x},\mathbf{y},\mathbf{b}\in\mathrm{SEP}:\\ \mathbf{b}=C(\mathbf{x},\mathbf{y})}} \mathrm{tr}_{\mathcal{S}_1\mathcal{E}_1}\left[(\rho_{\mathcal{S}_1} \otimes \xi_{\mathcal{E}_1})\left\{U_{x_1 y_1}^{(1)\dagger}\left(M_b^{(1)} \otimes \mathbb{1}_{\mathcal{E}_1}\right)U_{x_1 y_1}^{(1)}\right\}\right]\prod_{k=2}^n \lambda_{\max}\left[M_{\mathbf{b}}^{(k)}\right],
\end{aligned} \tag{B4}$$

where $\rho_{\mathcal{S}_1} = \mathrm{tr}_{\mathcal{S}_2\ldots\mathcal{S}_n}[\rho_{\mathcal{S}}]$ and we have used Lemma B $n-1$ times, together with the identity $\lambda_{\max}[U_{x_k y_k}^{(k)\dagger}(M_{\mathbf{b}}^{(k)} \otimes \mathbb{1}_{\mathcal{E}_k})U_{x_k y_k}^{(k)}] = \lambda_{\max}[M_{\mathbf{b}}^{(k)}]$.

Since $\{M_{\mathbf{b}}\}_{\mathbf{b}}$ is a valid POVM it must satisfy $\sum_{\mathbf{b}} M_{\mathbf{b}} = \sum_{\mathbf{b}\in\mathrm{SEP}} \bigotimes_{k=1}^{n} M_{\mathbf{b}}^{(k)} + \sum_{\mathbf{b}\notin\mathrm{SEP}} M_{\mathbf{b}} = \mathbb{1}_{\mathcal{S}}$. Tracing out subsystems $\{2,\ldots,n\}$, we see that

$$\sum_{\mathbf{b}\in\mathrm{SEP}} M_{\mathbf{b}}^{(1)} \prod_{k=2}^{n} \mathrm{tr}\left[M_{\mathbf{b}}^{(k)}\right] = d^{n-1}\mathbb{1}_{\mathcal{S}_1} - \sum_{\mathbf{b}\notin\mathrm{SEP}} \mathrm{tr}_{\mathcal{S}_2\cdots\mathcal{S}_n}[M_{\mathbf{b}}]. \tag{B5}$$

By noting that, given the values of $y_1, x_1, \ldots, x_{n-1}, \mathbf{b}$, the condition $\mathbf{b} = C(\mathbf{x}, \mathbf{y})$ fixes the values of the remaining variables and that these remaining variables do not appear in the summand in Eq. (B4), we can rewrite the summation simply over $y_1, x_1, \ldots, x_{n-1}, \mathbf{b}$. Noting also that $0 \leqslant \lambda_{\max}[P] \leqslant \mathrm{tr}\,[P]$ for any positive semidefinite operator P and using Eq. (B5) we have

$$\begin{aligned}
T &\leqslant \frac{1}{d^{2n}} \sum_{\substack{y_1, x_1, \ldots, x_{n-1},\\ \mathbf{b}\,\in\,\mathrm{SEP}}} \mathrm{tr}_{\mathcal{S}_1\mathcal{E}_1}\left[\left(\rho_{\mathcal{S}_1}\otimes\xi_{\mathcal{E}_1}\right)\left\{U_{x_1y_1}^{(1)\dagger}\left(M_{\mathbf{b}}^{(1)}\prod_{k=2}^{n}\lambda_{\max}\left[M_{\mathbf{b}}^{(k)}\right]\otimes\mathbb{1}_{\mathcal{E}_1}\right)U_{x_1y_1}^{(1)}\right\}\right] \\
&\leqslant \frac{1}{d^{2n}} \sum_{y_1,x_1,\ldots,x_{n-1}} \mathrm{tr}_{\mathcal{S}_1\mathcal{E}_1}\left[\left(\rho_{\mathcal{S}_1}\otimes\xi_{\mathcal{E}_1}\right)\left\{U_{x_1y_1}^{(1)\dagger}\left(\sum_{\mathbf{b}\in\mathrm{SEP}} M_{\mathbf{b}}^{(1)}\prod_{k=2}^{n}\mathrm{tr}\left[M_{\mathbf{b}}^{(k)}\right]\otimes\mathbb{1}_{\mathcal{E}_1}\right)U_{x_1y_1}^{(1)}\right\}\right] \\
&= \frac{1}{d^{2n}} \sum_{y_1,x_1,\ldots,x_{n-1}} \mathrm{tr}_{\mathcal{S}_1\mathcal{E}_1}\left[\left(\rho_{\mathcal{S}_1}\otimes\xi_{\mathcal{E}_1}\right)\left\{U_{x_1y_1}^{(1)\dagger}\left(d^{n-1}\mathbb{1}_{\mathcal{S}_1\mathcal{E}_1} - \sum_{\mathbf{b}\notin\mathrm{SEP}}\mathrm{tr}_{\mathcal{S}_2\cdots\mathcal{S}_n}[M_{\mathbf{b}}]\otimes\mathbb{1}_{\mathcal{E}_1}\right)U_{x_1y_1}^{(1)}\right\}\right] \\
&\leqslant \frac{1}{d} - \frac{1}{d^{2n}}\lambda_{\max}\left[\sum_{y_1,x_1,\ldots,x_{n-1}} U_{x_1y_1}^{(1)\dagger}\left(\sum_{\mathbf{b}\notin\mathrm{SEP}}\mathrm{tr}_{\mathcal{S}_2\cdots\mathcal{S}_n}[M_{\mathbf{b}}]\otimes\mathbb{1}_{\mathcal{E}_1}\right)U_{x_1y_1}^{(1)}\right].
\end{aligned} \tag{B6}$$

We note briefly that one also obtains Eq. (B6) if one considers general separable operators of the form $M_{\mathbf{b}} = \sum_i \bigotimes_{k=1}^{n} M_{\mathbf{b},i}^{(k)}$ (rather than simple tensor products). This follows readily from the linearity of both Eqs. (B4) and (B5), so that the sum over separable measurement operators can be eliminated in the same way as above.

Continuing with the proof, note that for a positive semidefinite operator P in a d-dimensional Hilbert space, $\lambda_{\max}[P] \geqslant \frac{1}{d}\,\mathrm{tr}\,[P]$. Hence, letting D be the dimension of $\mathcal{E}_1$, we have

$$\begin{aligned}
T &\leqslant \frac{1}{d} - \frac{1}{Dd^{2n+1}}\,\mathrm{tr}_{\mathcal{S}_1\mathcal{E}_1}\left[\sum_{y_1,x_1,\ldots,x_{n-1}} U_{x_1y_1}^{(1)\dagger}\left(\sum_{\mathbf{b}\notin\mathrm{SEP}}\mathrm{tr}_{\mathcal{S}_2\cdots\mathcal{S}_n}[M_{\mathbf{b}}]\otimes\mathbb{1}_{\mathcal{E}_1}\right)U_{x_1y_1}^{(1)}\right] \\
&= \frac{1}{d} - \frac{1}{Dd^{n+1}}\sum_{\mathbf{b}\notin\mathrm{SEP}}\mathrm{tr}_{\mathcal{S}\mathcal{E}_1}[M_{\mathbf{b}}\otimes\mathbb{1}_{\mathcal{E}_1}] \\
&\leqslant \frac{1}{d} - \frac{1}{d^{n+1}}\sum_{\mathbf{b}\notin\mathrm{SEP}}\lambda_{\max}[M_{\mathbf{b}}],
\end{aligned} \tag{B7}$$

where, in the last step, we used the relation $\mathrm{tr}_{\mathcal{S}\mathcal{E}_1}[M_{\mathbf{b}}\otimes\mathbb{1}_{\mathcal{E}_A}] = D\,\mathrm{tr}_{\mathcal{S}}[M_{\mathbf{b}}] \geqslant D\lambda_{\max}[M_{\mathbf{b}}]$.

Substituting this back into the expression (B3) for $\mathcal{A}_{n,d}$ and noting that

$$\mathrm{tr}_{\mathcal{S}\mathcal{E}}\left[(\rho\otimes\xi_{\mathcal{E}})\left(\bigotimes_{k=1}^{n}U_{x_ky_k}^{(k)\dagger}\right)(M_{\mathbf{b}}\otimes\mathbb{1}_{\mathcal{E}_k})\left(\bigotimes_{k=1}^{n}U_{x_ky_k}^{(k)}\right)\right] \leqslant \lambda_{\max}[M_{\mathbf{b}}] \leqslant 1, \tag{B8}$$

we have the bound

$$\mathcal{A}_{n,d} \leqslant \frac{1}{d} + \left(\frac{1}{d^n} - \frac{1}{d^{n+1}}\right)\sum_{\mathbf{b}\notin\mathrm{SEP}}\lambda_{\max}[M_{\mathbf{b}}] \leqslant \frac{1}{d} + \left(\frac{1}{d^n} - \frac{1}{d^{n+1}}\right)(d^n - k) = \frac{1}{d^n}\left(d^n - k + \frac{k}{d}\right), \tag{B9}$$

as desired.

APPENDIX C: PARTIAL TIGHTNESS OF RESULT 2 FOR TWO PARTIES

In this Appendix we consider the tightness of the inequality derived in result 2 for the bipartite case, which gives a lower bound on the number of entangled measurement operators compatible with a given average score $\mathcal{A}_{2,d}$. Specifically, we present a simple strategy that saturates the bound using a measurement for which there are $k = md$ separable measurement operators, for $m = 0, \ldots, d$.

Consider the following projective joint measurement of two d-dimensional quantum systems, in which the measurement operators $M_{\mathbf{b}} = |M_{\mathbf{b}}\rangle\langle M_{\mathbf{b}}|$ are separable for all (b_1, b_2) satisfying $b_2 - b_1 \in \{0, \ldots, m-1\}$, and entangled otherwise (where, as

always, $b_2 - b_1$ is computed modulo d). Hence, there are md separable operators and $(d-m)d$ entangled operators. Specifically,

$$|M_{b_1 b_2}\rangle = \begin{cases} |b_1, b_2\rangle & \text{for } b_2 - b_1 \in \{0, \ldots, m-1\}, \\ Z^{b_1} \otimes X^{b_2 - b_1} |\phi_{\max}\rangle & \text{for } b_2 - b_1 \notin \{0, \ldots, m-1\}. \end{cases} \tag{C1}$$

To see that this is a valid measurement, note that all the inner products of two different separable basis elements is zero. Similarly, the entangled basis elements constitute a subset of the Bell basis and are thus orthonormal. To show that the inner products between the separable basis elements and entangled basis elements are all zero, consider the straightforward calculation

$$\langle b_1', b_2' | Z^{b_1} \otimes X^{b_2 - b_1} |\phi_{\max}\rangle = \frac{1}{\sqrt{d}} \sum_{\ell=0}^{d-1} \omega^{\ell b_1} \langle b_1', b_2' | \ell, \ell + b_2 - b_1 \rangle = \frac{\omega^{b_1' b_1}}{\sqrt{d}} \delta_{b_2' - b_1', b_2 - b_1}. \tag{C2}$$

Since all the separable basis elements have $b_2' - b_1' \in \{0, \ldots, m-1\}$ while all entangled basis elements have $b_2 - b_1 \notin \{0, \ldots, m-1\}$, the final Kronecker delta function is zero. Hence, Eq. (C1) defines an orthonormal basis.

Consider thus the following strategy. Let A_1 and A_2 share a maximally entangled state, and apply (a relabelled variant of) the transformation strategy exploited several times in the main text, namely, take the unitary transformations $U^{A_1}_{x_1 y_1} = Z^{x_1} X^{y_1 + x_1}$ and $U^{A_2}_{x_2 y_2} = Z^{x_2} X^{y_2 - x_2}$. It follows straightforwardly that

$$\mathcal{A}_{2,d} = \frac{d-m}{d} + \frac{m}{d^2}, \tag{C3}$$

which saturates the upper bound described in result 2 for any m. We leave it as an open question whether a similar partial tightness result holds in the more general n partite case.

[1] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, Rev. Mod. Phys. **81**, 865 (2009).

[2] O. Gühne and G. Toth, Entanglement detection, Phys. Rep. **474**, 1 (2009).

[3] C. Eltschka and J. Siewert, Quantifying entanglement resources, J. Phys. A **47**, 424005 (2014).

[4] D. Rosset, R. Ferretti-Schöbitz, J.-D. Bancal, N. Gisin, and Y.-C. Liang, Imperfect measurement settings: Implications for quantum state tomography and entanglement witnesses, Phys. Rev. A **86**, 062325 (2012).

[5] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Device-Independent Security of Quantum Cryptography Against Collective Attacks, Phys. Rev. Lett. **98**, 230501 (2007).

[6] R. Colbeck, Quantum and relativistic protocols for secure multiparty computation, Ph.D. thesis, University of Cambridge, 2006, arXiv:0911.3814.

[7] S. Pironio, A. Acín, S. Massar, A. Boyer de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe, Random numbers certified by Bell's theorem, Nature **464**, 1021 (2010).

[8] D. Collins, N. Gisin, S. Popescu, D. Roberts, and V. Scarani, Bell-Type Inequalities to Detect True n-Body Nonseparability, Phys. Rev. Lett. **88**, 170405 (2002).

[9] M. Seevinck and G. Svetlichny, Bell-Type Inequalities for Partial Separability in n-Particle Systems and Quantum Mechanical Violations, Phys. Rev. Lett. **89**, 060401 (2002).

[10] J.-D. Bancal, N. Gisin, Y.-C. Liang, and S. Pironio, Device-Independent Witnesses of Genuine Multipartite Entanglement, Phys. Rev. Lett. **106**, 250404 (2011).

[11] K. F. Pál and T. Vértesi, Multisetting Bell-type inequalities for detecting genuine multipartite entanglement, Phys. Rev. A **83**, 062123 (2011).

[12] J.-D. Bancal, C. Branciard, N. Brunner, N. Gisin, and Y.-C. Liang, A framework for the study of symmetric full-correlation Bell-like inequalities, J. Phys. A **45**, 125301 (2012).

[13] T. Moroder, J.-D. Bancal, Y.-C. Liang, M. Hofmann, and O. Gühne, Device-Independent Entanglement Quantification and Related Applications, Phys. Rev. Lett. **111**, 030501 (2013).

[14] G. Murta, R. Ramanathan, N. Móller, and M. Terra Cunha, Quantum bounds on multiplayer linear games and device-independent witness of genuine tripartite entanglement, Phys. Rev. A **93**, 022305 (2016).

[15] F. Baccari, D. Cavalcanti, P. Wittek, and A. Acín, Efficient Device-Independent Entanglement Detection for Multipartite Systems, Phys. Rev. X **7**, 021042 (2017).

[16] G. Tóth and A. Acín, Genuine tripartite entangled states with a local hidden-variable model, Phys. Rev. A **74**, 030306 (2006).

[17] R. Augusiak, M. Demianowicz, J. Tura, and A. Acín, Entanglement and Nonlocality are Inequivalent for Any Number of Parties, Phys. Rev. Lett. **115**, 030404 (2015).

[18] J. Bowles, J. Francfort, M. Fillettaz, F. Hirsch, and N. Brunner, Genuinely Multipartite Entangled Quantum States with Fully Local Hidden Variable Models and Hidden Multipartite Nonlocality, Phys. Rev. Lett. **116**, 130401 (2016).

[19] Note that nonlocality can nevertheless be activated in some more sophisticated Bell scenarios involving, e.g., sequential measurements [50] or processing of multiple copies [51]. Furthermore, with the help of additional sources of perfect singlets, any entangled bipartite state can be certified [52].

[20] R. F. Werner, Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model, Phys. Rev. A **40**, 4277 (1989).

[21] R. Augusiak, M. Demianowicz, and A. Acín, Local hidden variable models for entangled quantum states, J. Phys. A **42**, 424002 (2014).

[22] R. Rabelo, M. Ho, D. Cavalcanti, N. Brunner, and V. Scarani, Device-Independent Certification of Entangled Measurements, Phys. Rev. Lett. **107**, 050502 (2011).

[23] Q. Y. He and M. D. Reid, Genuine Multipartite Einstein-Podolsky-Rosen Steering, Phys. Rev. Lett. **111**, 250403 (2013).

[24] D. Cavalcanti, P. Skrzypczyk, G. H. Aguilar, R. V. Nery, P. H. Souto Ribeiro, and S. P. Walborn, Detection of entanglement in asymmetric quantum networks and multipartite quantum steering, Nat. Commun. **6**, 7941 (2015).

[25] W. McCutcheon, A. Pappa, B. A. Bell, A. McMillan, A. Chailloux, T. Lawson, M. Mafu, D. Markham, E. Diamanti, I. Kerenidis, J. G. Rarity, and M. S. Tame, Experimental verification of multipartite entanglement in quantum networks, Nat. Commun. **7**, 13251 (2016).

[26] F. Buscemi, All Entangled Quantum States are Nonlocal, Phys. Rev. Lett. **108**, 200401 (2012).

[27] C. Branciard, D. Rosset, Y.-C. Liang, and N. Gisin, Measurement-Device-Independent Entanglement Witnesses for All Entangled Quantum States, Phys. Rev. Lett. **110**, 060405 (2013).

[28] M. Pawłowski and N. Brunner, Semi-device-independent security of one-way quantum key distribution, Phys. Rev. A **84**, 010302 (2011).

[29] H.-W. Li, M. Pawłowski, Z.-Q. Yin, G.-C. Guo, and Z.-F. Han, Semi-device-independent randomness certification using $n \rightarrow 1$ quantum random access codes, Phys. Rev. A **85**, 052308 (2012).

[30] A. Tavakoli, A. Hameedi, B. Marques, and M. Bourennane, Quantum Random Access Codes using Single d-Level Systems, Phys. Rev. Lett. **114**, 170502 (2015).

[31] Y.-C. Liang, T. Vértesi, and N. Brunner, Semi-device-independent bounds on entanglement, Phys. Rev. A **83**, 022108 (2011).

[32] K. T. Goh, J.-D. Bancal, and V. Scarani, Measurement-device-independent quantification of entanglement for given Hilbert space dimension, New J. Phys. **18**, 045022 (2016).

[33] T. Vértesi and M. Navascués, Certifying entangled measurements in known Hilbert spaces, Phys. Rev. A **83**, 062112 (2011).

[34] A. Bennet, T. Vértesi, D. J. Saunders, N. Brunner, and G. J. Pryde, Experimental Semi-Device-Independent Certification of Entangled Measurements, Phys. Rev. Lett. **113**, 080405 (2014).

[35] N. Elron and Y. C. Eldar, Optimal encoding of classical information in a quantum medium, IEEE Trans. Inf. Theory **53**, 1900 (2007).

[36] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Teleporting an Unknown Quantum State via Dual Classical and Einstein-Podolsky-Rosen Channels, Phys. Rev. Lett. **70**, 1895 (1993).

[37] M. Horodecki, P. Horodecki, and R. Horodecki, General teleportation channel, singlet fraction, and quasidistillation, Phys. Rev. A **60**, 1888 (1999).

[38] F. Hirsch, M. T. Quintino, T. Vértesi, M. Navascués, and N. Brunner, Better local hidden variable models for two-qubit Werner states and an upper bound on the Grothendieck constant $K_G(3)$, Quantum **1**, 3 (2017).

[39] M. L. Almeida, S. Pironio, J. Barrett, G. Tóth, and A. Acín, Noise Robustness of the Nonlocality of Entangled Quantum States, Phys. Rev. Lett. **99**, 040403 (2007).

[40] H. Buhrman, O. Regev, G. Scarpa, and R. de Wolf, Near-optimal and explicit Bell inequality violations, *2011 IEEE 26th Annual Conference on Computational Complexity* (IEEE, New York, 2011), p. 157.

[41] S. M. Hashemi Rafsanjani, M. Huber, C. J. Broadbent, and J. H. Eberly, Genuinely multipartite concurrence of N-qubit X-matrices, Phys. Rev. A **86**, 062303 (2012).

[42] F. Clivaz, M. Huber, L. Lami, and G. Murta, Genuine-multipartite entanglement criteria based on positive maps, J. Math. Phys. **58**, 082201 (2017).

[43] J. Xu, Multipartite fully entangled fraction, Int. J. Theor. Phys. **55**, 2904 (2016).

[44] Note that EGF is defined for arbitrary states whereas GF is only defined for d^n-dimensional states.

[45] P. Badziag, M. Horodecki, P. Horodecki, and R. Horodecki, Local environment can enhance fidelity of quantum teleportation, Phys. Rev. A **62**, 012311 (2000).

[46] B. Jungnitsch, T. Moroder, and O. Gühne, Taming Multiparticle Entanglement, Phys. Rev. Lett. **106**, 190502 (2011).

[47] M. Riebe, T. Monz, K. Kim, A. S. Villar, P. Schindler, M. Chwalla, M. Hennrich, and R. Blatt, Deterministic entanglement swapping with an ion-trap quantum computer, Nat. Phys. **4**, 839 (2008).

[48] W. Pfaff, B. J. Hensen, H. Bernien, S. B. van Dam, M. S. Blok, T. H. Taminiau, M. J. Tiggelman, R. N. Schouten, M. Markham, D. J. Twitchen, and R. Hanson, Unconditional quantum teleportation between distant solid-state quantum bits, Science **345**, 532 (2014).

[49] A. Tavakoli, J. Kaniewski, T. Vértesi, D. Rosset, and N. Brunner, Self-testing quantum states and measurements in the prepare-and-measure scenario, arXiv:1801.08520.

[50] S. Popescu, Bell's Inequalities and Density Matrices: Revealing "Hidden" Nonlocality, Phys. Rev. Lett. **74**, 2619 (1995).

[51] C. Palazuelos, Superactivation of Quantum Nonlocality, Phys. Rev. Lett. **109**, 190401 (2012).

[52] J. Bowles, I. Šupić, D. Cavalcanti, and A. Acín, Device-independent entanglement certification of all entangled states, arXiv:1801.10444.

High-Dimensional Quantum Communication Complexity beyond Strategies Based on Bell's Theorem

Daniel Martínez,[1,2] Armin Tavakoli,[3] Mauricio Casanova,[1,2] Gustavo Cañas,[4] Breno Marques,[5,6] and Gustavo Lima[1,2]
[1]*Departamento de Física, Universidad de Concepción, 160-C Concepción, Chile*
[2]*Millennium Institute for Research in Optics, Universidad de Concepción, 160-C Concepción, Chile*
[3]*Groupe de Physique Appliquée, Université de Genève, CH-1211 Genève, Switzerland*
[4]*Departamento de Física, Universidad del Bio-Bio, Avenida Collao 1202, Concepción, Chile*
[5]*Instituto de Física, Universidade de São Paulo, 05315-970 São Paulo, Brazil*
[6]*Centro de Ciências Naturais e Humanas, Universidade Federal do ABC, Avenida dos Estados 5001, 09210-580 Santo André, São Paulo, Brazil*

(Received 16 July 2018; published 12 October 2018)

Quantum resources can improve communication complexity problems (CCPs) beyond their classical constraints. One quantum approach is to share entanglement and create correlations violating a Bell inequality, which can then assist classical communication. A second approach is to resort solely to the preparation, transmission, and measurement of a single quantum system, in other words, quantum communication. Here, we show the advantages of the latter over the former in high-dimensional Hilbert space. We focus on a family of CCPs, based on facet Bell inequalities, study the advantage of high-dimensional quantum communication, and realize such quantum communication strategies using up to ten-dimensional systems. The experiment demonstrates, for growing dimension, an increasing advantage over quantum strategies based on Bell inequality violation. For sufficiently high dimensions, quantum communication also surpasses the limitations of the postquantum Bell correlations obeying only locality in the macroscopic limit. We find that the advantages are tied to the use of measurements that are not rank-one projective, and provide an experimental semi-device-independent falsification of such measurements in Hilbert space dimension six.

DOI: 10.1103/PhysRevLett.121.150504

Introduction.—Communication complexity problems (CCPs) are tasks in which distant parties hold local data, the collection of which is needed for a computation of their interest. To make the computation possible, the parties communicate with each other. However, the amount of communication is limited, and therefore, not all data can be sent. The CCPs consist in parties adopting an efficient communication strategy which allows them to perform the desired computation with a probability as high as possible. Efficient use of quantitatively limited communication is a broadly relevant matter [1], which provides fundamental insights on physical limitations [2,3].

The ability to process information depends on the choice of the physical system into which the information is encoded [4]. Consequently, quantum entities without a classical counterpart can be regarded as tools for quantum information processing. The most famous example is entanglement. In a quantum CCP, parties may share an entangled state on which they perform local measurements, generating strongly correlated data which violate a Bell inequality. That data can then be used to assist a classical communication strategy [5]. In fact, Bell inequalities have been systematically linked to CCPs [6–8], and their violation enables better-than-classical communication efficiencies [7–15].

Nevertheless, quantum theory also presents a second approach to CCPs: substituting classical communication with quantum communication. Such a substitution must ensure that no more than the allowed amount of classical information can be extracted from the quantum communication, i.e., that the constraints of the CCP are respected. Since the Holevo theorem [16] implies that no more information can be extracted from a quantum d-level system than from a classical d-level system, a valid quantum communication strategy may encode information in quantum states of a specified limited Hilbert space dimension, and subsequently extract it by a measurement. The ability of quantum communication to outperform classical constraints in CCPs is well established [17–25].

Many quantum communication tasks can be successfully completed both by means of local measurements on an entangled state followed by classical communication, or by the communication of a single quantum system [26–28]. For two-party CCPs with both binary communication and outcomes, classical communication assisted by correlations violating a Bell inequality is always at least as good as an implementation based on quantum communication [29]. Explicit examples in which the advantage is strict are known [30,31]. However, examples also exist of two-party CCPs with more than two outcomes in which quantum

communication holds an advantage over the Bell inequality based approach [32,33].

In this Letter, we theoretically explore and experimentally demonstrate advantages of performing CCPs with quantum communication in high-dimensional Hilbert space, as compared to exploiting the violation of a Bell inequality. To this end, we focus on a family of CCPs [12,13] based on the (to the best of our knowledge) only known family of bipartite facet Bell inequalities. Facet inequalities optimally bound correlations with a local hidden variable model [34]. We consider the Collins-Gisin-Linden-Massar-Popescu (CGLMP) inequalities, involving any d number of outcomes [35,36]. We demonstrate the advantage of quantum communication over strategies based on violations of the CGLMP inequalities, which we show to be even larger than previously thought [33]. In particular, while resolving two conjectures of [33], we show that, below dimension six, both quantum CCP implementations are equally efficient, whereas above (and including) dimension six, they are not. In this sense, dimension six acts as a threshold for revealing the advantages of quantum communication. To understand the suddenly emerging discrepancy between the two quantum CCP implementations, we evidence that optimal quantum communication strategies in high-dimensional Hilbert space require projective measurements that are not rank one. Subsequently, we present an experimental realization. Using high-dimensional photonic systems, specifically up to dimension ten, we outperform strategies based on violating the CGLMP inequalities, emerging from dimension six, by means of quantum nonlocal correlations. Furthermore, we also outperform strategies based on superquantum violations of said inequalities respecting only no-signaling and macroscopic locality [37]. Finally, we prove that the experimental data cannot be simulated with any rank-one projective measurement without additional postprocessing of the data. Since only a dimensional bound on the relevant Hilbert space is assumed, this constitutes a semi-device-independent [38] falsification of said property.

The communication complexity problems.—Every quantum implementation of a Bell experiment that leads to a violation of a Bell inequality can be mapped to a particular strategy in an associated CCP. The corresponding CCP strategy leads to an advantage over classical methods analogous to the violation of the Bell inequality [7,8]. A natural candidate for such mappings are optimal (facet) Bell inequalities. The CGLMP inequalities [35] constitute a family of facet Bell inequalities for two parties, each with two choices of measurements and with $d \geq 2$ possible outcomes.

The construction of CCPs based on the CGLMP inequalities has been developed in [12,13]. In this family of CCPs (parametrized by d), a party Alice is given random inputs $x_0 \in \{0, \ldots, d-1\}$ and $x \in \{0, 1\}$, and another party, Bob, is given a random input $y \in \{0, 1\}$. Alice may communicate no more than $\log d$ bits to Bob, after which he outputs a guess $g \in \{0, \ldots, d-1\}$. If g coincides with the value of a function $f_k(x_0, x, y) = x_0 - xy - (-1)^{x+y} k \bmod d$, for some $k = 0, \ldots, \lfloor d/2 \rfloor - 1$, the partnership is awarded $c_k = 1 - 2k/(d-1)$ points. However, if g coincides with $h_k = x_0 - xy + (-1)^{x+y}(k+1) \bmod d$, the partnership loses c_k points. The task is to efficiently communicate such that the average number of points earned is large. The payoff function is given by

$$\Delta_d^{\mathrm{Bell}} = \frac{1}{4d} \sum_{\substack{x_0, x \\ y, k}} c_k [P(g = f_k | x_0, x, y) - P(g = h_k | x_0, x, y)].$$

On the one hand, in an approach based on Bell inequalities, Alice and Bob share an entangled state and perform local measurements x and y with d-valued outcomes a and b, respectively. In order to exploit the fact that the CCP is tailored to the CGLMP inequalities, Alice encodes the classical communication $m(a, x_0, x) \in \{0, \ldots, d-1\}$ using $m = x_0 + a \bmod d$, and Bob subsequently decodes it using $g = m - b \bmod d$ (see Fig. 1). It was shown [12,13,33] that the resulting value of Δ_d^{Bell} is in one-to-one correspondence with the quantity evaluated from the statistics $p(a, b|x, y)$ in a test of the CGLMP inequalities. In this sense, the efficiency in the CCP is determined by the amount of nonlocality present in the distribution $p(a, b|x, y)$. In particular, if $p(a, b|x, y)$ generates a maximal violation of the (suitably normalized) CGLMP inequalities, then by the outlined communication strategy, it can be used to achieve an equally large value of Δ_d^{Bell}. The maximal quantum value achievable in a test of the CGLMP inequalities lacks a simple analytical form but is known up to large d and achieved with nonmaximally entangled states [39].

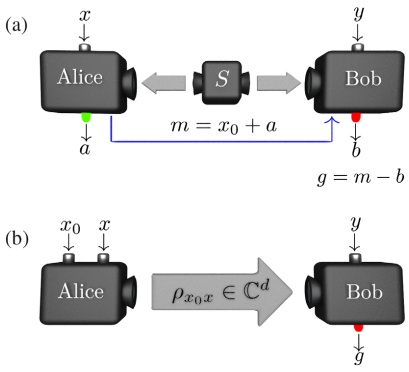

FIG. 1. (a) Quantum CCP implementation based on the violation of the CGLMP inequalities. (b) Quantum CCP implementation based on communicating a single d-level quantum system.

Violations have been experimentally observed for high-dimensional systems [40–42].

On the other hand, these CCPs can also be implemented without exploiting entanglement and Bell inequality violations [33]. Instead, Alice and Bob can use single quantum systems for direct quantum communication. In such an implementation, Alice associates her random inputs (x, x_0) to a d-dimensional quantum state, $\rho_{x_0 x} \in \mathbb{C}^d$, which is sent to Bob who performs a measurement $\{M_y^g\}_{g=0}^{d-1}$, the outcome g of which determines his output guess (see Fig. 1). In a quantum model, the performance of the CCP reads

$$\Delta_d^{\mathrm{QS}} = \frac{1}{4d} \sum_{x_0,x,y,k} c_k \mathrm{tr}[\rho_{x_0 x}(M_y^{f_k} - M_y^{h_k})]. \tag{1}$$

An efficient quantum communication strategy, i.e., a suitable choice of state preparations and measurements, aims to find the largest value of Δ_d^{QS}. In the Supplemental Material [43], we discuss the advantages and limitations of the two quantum CCP implementations, partly based on the results of Ref. [44]. In Ref. [33], it was shown that the optimal performance of the two different quantum approaches is equal, i.e., $\Delta_d^{\mathrm{QS}} = \Delta_d^{\mathrm{Bell}}$, when $d = 2, 3, 4$. Numerical results also suggested the same relation for $d = 5, 6$. However, for $d \geq 7$, lower bounds on Δ_d^{QS} were shown to outperform the maximal value of Δ_d^{Bell}. Next, we revisit this analysis, show improved quantitative results, establish the precise dimension revealing the inequivalence, and provide insight to the qualitative differences between the two quantum implementations of the CCPs.

The efficiency of quantum communication.—We begin by quantifying the advantage of quantum communication over strategies based on the violation of the CGLMP inequalities. Specifically, we numerically infer lower bounds on the maximal value of Δ_d^{QS} for $d \leq 10$. This has been done by running two optimizations in seesaw [45,46]: first optimizing over the states of Alice, and then over the measurements of Bob, repeatedly. Each such optimization constitutes a semidefinite program [47]. The best found states and measurements are listed in Supplemental Material [43]. The results are presented in Table I together with the known [33,39] optimal CGLMP-based values of Δ_d^{Bell} as obtained both in quantum theory, and by the superquantum principle of macroscopic locality [37]. The latter correlations are only constrained by the inability of violating a Bell inequality when the measurements are macroscopic, i.e., that a large number of particles are collectively measured instead of microscopic measurements on single particles. The results substantially improve on the lower bounds for Δ_d^{QS} obtained in [33] and, thus, establish an increased quantitative advantage of high-dimensional quantum communication over strategies based on Bell inequality violation. In particular, note that, for $d = 8, 9, 10$, quantum communication can even outperform the Bell inequality based approach when the correlations established are only required to be macroscopically local; i.e., the violation of the CGLMP inequalities is larger-than quantum.

TABLE I. Lower bounds for the maximal value of Δ_d^{QS} as compared to the maximal value of Δ_d^{Bell} obtained via the maximal quantum (and macroscopically local, i.e., Δ_d^{ML}) violation of the CGLMP inequalities. The final column was obtained through optimization over unit-trace measurement operators and optimal measurements were always found to be rank-one projective.

d	Lower bound Δ_d^{QS}	Lower bound Δ_d^{QS} from [33]	Δ_d^{Bell}	Δ_d^{ML}	Lower bound Δ_d^{QS} rank-one projective
2	⋯	0.7071	0.7071	0.7071	0.7071
3	⋯	0.7287	0.7287	0.7887	0.7287
4	⋯	0.7432	0.7432	0.8032	0.7432
5	⋯	0.7539	0.7539	0.8249	0.7539
6	0.8000	0.7624	0.7624	0.8345	0.7624
7	0.8175	0.7815	0.7694	0.8461	0.7814
8	0.8571	0.8006	0.7753	0.8529	0.8006
9	0.8622	0.8622	0.7804	0.8605	0.8188
10	0.8889	0.8778	0.7849	0.8657	0.8396

Furthermore, we rectify the main result of [33] by resolving two of its conjectures: that the optimal quantum communication strategy performs equally well as that based on the quantum violation of the CGLMP inequalities when $d = 5$ and when $d = 6$. For $d = 5$, we have used the second hierarchy level of dimensionally bound quantum correlations [48] combined with symmetrization techniques [49]. We obtain a tight bound on the efficiency of quantum communication matching that obtained through a maximal violation of the CGLMP inequalities. This proves the conjecture. For $d = 6$, the presented lower bound on Δ_d^{QS} shows that quantum communication outperforms the analogous Bell inequality based result. Thus, the improved lower bound falsifies the conjecture. This establishes dimension six as the dimension revealing the quantitative inequivalence between the two quantum implementations of the CCPs.

A relevant question is whether the breaking of the equivalence of the two quantum implementations, emerging when $d \geq 6$, is linked to qualitatively different properties in the optimal use of the respective quantum systems. Below the critical dimension six, the optimal found preparations of Alice can be effectively prepared by Alice locally measuring an entangled state, and then considering the post-measurement state of Bob for her given outcome. The collection of Bob's post-measurement states is then identical to the collection of states communicated over a quantum channel in an optimal strategy. Consequently, there is no advantage over Bell inequality based strategies. Furthermore, the optimal measurements coincide with the rank-one projective measurements

optimal for violating the CGLMP inequalities. However, when $d \geq 6$ our numerical calculations for $d = 6, \ldots, 10$ suggest that (I) the states $\{\rho_{x_0 x}\}$ cannot be prepared remotely with entanglement in a test of the CGLMP inequalities, and that some inputs may be associated to the same state, and (II) one of Bob's measurements is rank-one projective, whereas the other is higher-rank projective; i.e., some measurement operators are zero operators, meaning that the associated outcomes never can occur. Degenerate measurements are known to be optimal for some quantum information problems [50,51]. They can be viewed as rank-one projective measurements with additional postprocessing by which some outcomes remain untouched and other outcomes are given new labels. To further evidence the suboptimality of rank-one projective measurements (without postprocessing), we have numerically optimized Δ_d^{QS} over measurements in which all measurement operators are of trace one. Since all rank-one projectors are of trace one, and we always find the optimal measurement to be rank-one projective, the results constitute a lower bound valid for such measurements. The results (see Table I) show that, although rank-one projective measurements are sufficient to outperform strategies based violating the CGLMP inequalities, they are not optimal.

Experimental demonstration of high-dimensional quantum communication advantage.—We present an experimental demonstration of the advantages of single-system quantum communication in the considered CCPs for $d = 6, \ldots, 10$. In our experiment, d-dimensional quantum systems are encoded into the linear transverse momentum of single photons transmitted by programmable diffractive apertures [52–64].

The experimental setup is presented in Fig. 2. It is composed of two main parts: one for the state preparation and another for performing projective measurements on the prepared system. Each part is controlled by a field programmable gate array (FPGA) electronics. In the state preparation, a 690 nm single mode laser modulated with an acousto-optic modulator (AOM) and optical attenuators (not shown in Fig. 2) prepare weak coherent states with an average number of 0.9 photons per pulse. This source can be seen as an approximation to a nondeterministic single-photon source, since pulses with a single-photon account for 62.3% of the generated non-null pulses. Accidental counts are strongly suppressed by using a detection window that matches the optical pulse duration of 45 μs.

To encode the quantum states in the linear transverse momentum of single photons, we exploit the pixel programmability of spatial light modulators (SLMs) [52,53]. The state preparation and measurement stages have two fundamental blocks: an amplitude-modulation only SLM1 (SLM3), built with two linear polarizers and a liquid crystal display (LCD), and a phase-modulation only SLM2 (SLM4), composed of two linear polarizers, two quarter wave plates, and an LCD [65]. Each SLM is placed in the

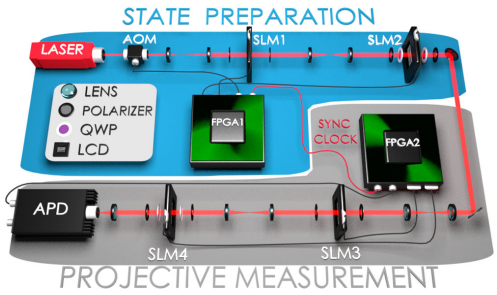

FIG. 2. Experimental setup for implementing the CCPs with quantum communication. d-dimensional quantum systems are encoded into the linear transverse momentum of single photons. The experiment is composed of two main parts: one for the state preparation and another for performing measurements on the prepared system. Both parts rely on the programmability of spatial light modulators for preparing the required states and measurements.

image plane of its predecessor. In order to experimentally generate $\rho_{x_0 x} = |\psi_{x,x_0}\rangle\langle\psi_{x,x_0}|$, a set of d slits with a width of 64 μm and equal center to center separation are displayed on SLM1 and SLM2. The individual transmittances t_l and phases ϕ_l of each slit "l" are set to reconstruct the real and imaginary parts of $|\psi_{x,x_0}\rangle$. The state vector of the transmitted photon after the SLM2 is given by $|\psi\rangle = (1/\sqrt{N}) \sum_{l=-d/2}^{d/2} \sqrt{t_l}\, e^{i\phi_l} |l\rangle$, where N is a normalization constant. The coefficients t_l and the phases ϕ_l are independently controlled by the SLM 1 and SLM 2, respectively. To implement the desired measurements, different amplitude and phase sets of the d slits are used at the SLM3 and SLM4. The transmittances and phases of each set are chosen to postselect for detection one of the required state vectors $|\varphi_{y,b_y}\rangle$. In the final part of the setup, a "pointlike" avalanche single-photon detector (APD) with a 10 μm pinhole is placed at the center of the far field plane of the SLM4. In this case, the probability of single-photon detection $P(x, x_0, y, b)$ is proportional to $|\langle\varphi_{y,b}|\psi_{x,x_0}\rangle|^2$ [53–56]. However, since one of the targeted protocol measurements is rank-two projective (see Supplemental Material [43]), we postprocess the experimental data to emulate the statistics of such a measurement. This is done by suitably relabeling the outcomes of the relevant measurements whenever, in the raw data, it is associated to an outcome which never occurs in the desired rank-two projective measurement.

After several rounds of the experiment, an experimental value of Δ_d^{QS} is obtained, namely Δ_d^{Exp}. Since the measurement uncertainty of Δ_d^{Exp} decreases with the total number of counts, the repetition of the experimental rounds for each dimension were chosen such that Δ_d^{Exp} violates the bounds for Bell inequality based strategies with at least 6

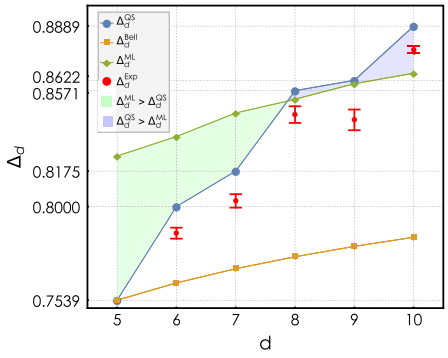

FIG. 3. Experimental results. Δ_d^{Exp} is represented by red points. The yellow points represent Δ_d^{Bell}. The blue points are the theoretical predictions of Δ_d^{QS}. The green points represent Δ_d^{ML}.

standard deviations for each d considered. Hence, any explanation in terms of an arbitrary entangled quantum system is excluded by at least 6 standard deviations, which corresponds to a p value of 1×10^{-9}.

For $d = 6, \ldots, 10$, we obtain the results

$$\Delta_6^{\mathrm{Exp}} = 0.7893 \pm 0.0026, \qquad \Delta_7^{\mathrm{Exp}} = 0.8082 \pm 0.0034,$$
$$\Delta_8^{\mathrm{Exp}} = 0.8453 \pm 0.0041, \qquad \Delta_9^{\mathrm{Exp}} = 0.8427 \pm 0.0051,$$
$$\Delta_{10}^{\mathrm{Exp}} = 0.8773 \pm 0.0018. \tag{2}$$

In Fig. 3, we compare the experimental results to the theoretical predictions for quantum communication, as well as with the limitations of both quantum and macroscopically local Bell correlations. The results are in good agreement with the theoretical predictions, surpassing the values associated to the maximal violation of the CGLMP inequalities. In particular, in the case of $d = 10$, the results also surpass the limitations of the postquantum Bell correlations obeying only macroscopic locality.

Finally, we revisit the previously numerically evidenced hypothesis of rank-one projective measurements being suboptimal. Focusing on the case of $d = 6$, we have considered whether the experimental data can be reproduced by some quantum communication strategy utilizing only such measurements. To this end, we have used an intermediate level [66] of the hierarchy of dimensionally bound quantum correlations [48], and additionally imposed upper and lower bounds on the particular probabilities measured in the lab corresponding to $(x_0, x) = (4, 0)$ and $y = 0$. In order to respect the errors of the measured probabilities, they were constrained to an interval twice larger than the experimental errors of each measurement outcome. In this manner, we have obtained the bound 0.7830 on Δ_6^{QS} which is smaller than the experimentally measured value. This demonstrates that, under the assumption of a six dimensional Hilbert space, there exists no quantum communication strategy based on rank-one projective measurements which can reproduce the experimental results.

Conclusion.—We have theoretically and experimentally studied the efficiency of high-dimensional quantum communication in a family of CCPs, as opposed to classical communication assisted by nonlocal correlations violating the facet Bell inequality to which the CCPs were originally tailored. We demonstrated significant advantages of quantum communication which increase with Hilbert space dimension and showed that they stem from degenerate measurements. Our work shows the usefulness and strength of quantum correlations generated via the communication of a high-dimensional quantum system and the practicality of experimentally realizing them.

This work was supported by Fondo Nacional de Desarrollo Científico y Tecnológico, Fondecyt, Grants No. 1160400, No. Fondecyt 11150324 and the Millennium Institute for Research in Optics (MIRO). D. M. acknowledges support from Comisión Nacional de Investigación Científica y Tecnológica, Grant Doctorado Nacional No. 2116050. A. T. acknowledges support from the Swiss National Science Foundation (Starting Grant DIAQ, NCCR-QSIT). B. M. was supported by São Paulo Research Foundation-FAPESP Project No. 2014/27223-2. A. T. and B. M. thank G. L. and G. C. for their hospitality in Concepción.

[1] E. Kushilevitz and N. Nisan, *Communication Complexity* (Cambridge University Press, Cambridge, England, 1997).
[2] G. Brassard, H. Buhrman, N. Linden, A. A. Méthot, A. Tapp, and F. Unger, Limit on Nonlocality in Any World in Which Communication Complexity Is Not Trivial, Phys. Rev. Lett. **96**, 250401 (2006).
[3] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski, Information causality as a physical principle, Nature (London) **461**, 1101 (2009).
[4] R. Landauer, The physical nature of information, Phys. Lett. A **217**, 188 (1996).
[5] H. Buhrman, R. Cleve, S. Massar, and R. de Wolf, Nonlocality and communication complexity, Rev. Mod. Phys. **82**, 665 (2010).
[6] H. Buhrman, L. Czekaj, A. Grudka, M. Horodecki, P. Horodecki, M. Markiewicz, F. Speelman, and S. Strelchuk, Quantum communication complexity advantage implies violation of a Bell inequality, Proc. Natl. Acad. Sci. U.S.A. **113**, 3191 (2016).
[7] C. Brukner, M. Zukowski, J-W Pan, and A. Zeilinger, Bells Inequalities and Quantum Communication Complexity, Phys. Rev. Lett. **92**, 127901 (2004).
[8] A. Tavakoli and M. Żukowski, Higher-dimensional communication complexity problems: Classical protocols versus quantum ones based on Bell's theorem or prepare-transmit-measure schemes, Phys. Rev. A **95**, 042305 (2017).

[9] R. Cleve and H. Buhrman, Substituting quantum entanglement for communication, Phys. Rev. A **56**, 1201 (1997).

[10] H. Buhrman, W. van Dam, P. Høyer, and A. Tapp, Multiparty quantum communication complexity, Phys. Rev. A **60**, 2737 (1999).

[11] L. Hardy and W. van Dam, Quantum whispers, Phys. Rev. A **59**, 2635 (1999).

[12] C. Brukner, M. Żukowski, and A, Zeilinger, Quantum Communication Complexity Protocol with Two Entangled Qutrits, Phys. Rev. Lett. **89**, 197901 (2002).

[13] C. Brukner, T. Paterek, and M. Żukowski, Quantum communication complexity protocols based on higher-dimensional entangled systems, Int. J. Quantum. Inform. **01**, 519 (2003).

[14] M. Epping and C. Brukner, Bound entanglement helps to reduce communication complexity, Phys. Rev. A **87**, 032305 (2013).

[15] S. Muhammad, A. Tavakoli, M. Kurant, M. Pawłowski, M. Żukowski, and M. Bourennane, Quantum Bidding in Bridge, Phys. Rev. X **4**, 021047 (2014).

[16] A. S. Holevo, Bounds for the quantity of information transmitted by a quantum communication channel, Probl. Inf. Transm. **9**, 177 (1973).

[17] A. Ambainis, A. Nayak, A. Ta-Shma, and U. Vazirani, Dense quantum coding and a lower bound for 1-way quantum automata, in *Proceedings of the Thirty-first Annual ACM Symposium on Theory of Computing, STOC 99, Atlanta, 1999* (ACM, New York, 1999), pp. 376–383.

[18] A. Nayak, Optimal lower bounds for quantum automata and random access codes, in *Proceedings of the 40th IEEE Symposium on Foundations of Computer Science (FOCS99), New York, 1999* (IEEE Computer Society, Los Alamitos, 1999), pp. 369–376.

[19] E. F. Galvão, Feasible quantum communication complexity protocol, Phys. Rev. A **65**, 012318 (2001).

[20] A. Ambainis, D. Leung, L. Mancinska, and M. Ozols, Quantum random access codes with shared randomness, arXiv:0810.2937.

[21] A. Casaccino, E. F. Galvão, and S. Severini, Extrema of discrete Wigner functions and applications, Phys. Rev. A **78**, 022310 (2008).

[22] P. Trojek, C. Schmid, M. Bourennane, C. Brukner, M. Żukowski, and H. Weinfurter, Experimental quantum communication complexity, Phys. Rev. A **72**, 050305(R) (2005).

[23] A. Tavakoli, A. Hameedi, B. Marques, and M. Bourennane, Quantum Random Access Codes Using Single *d*-Level Systems, Phys. Rev. Lett. **114**, 170502 (2015).

[24] P. A. Guérin, A. Feix, M. Araújo, and C. Brukner, Exponential Communication Complexity Advantage from Quantum Superposition of the Direction of Communication, Phys. Rev. Lett. **117**, 100502 (2016).

[25] M. Smania, A. M Elhassan, A. Tavakoli, and M. Bourennane, Experimental quantum multiparty communication protocols, npj Quantum Inf. **2**, 16010 (2016).

[26] A. K. Ekert, Quantum Cryptography Based on Bells Theorem, Phys. Rev. Lett. **67**, 661 (1991); C. H. Bennett, G. Brassard, and N. D. Mermin, Quantum Cryptography without Bells Theorem, Phys. Rev. Lett. **68**, 557 (1992).

[27] M. Fitzi, N. Gisin, and U. Maurer, Quantum Solution to the Byzantine Agreement Problem, Phys. Rev. Lett. **87**, 217901 (2001); A. Tavakoli, A. Cabello, M. Żukowski, and M. Bourennane, Quantum clock synchronization with a single qudit, Sci. Rep. **5**, 7982 (2015).

[28] M. Żukowski, A. Zeilinger, M. A. Horne, and H. Weinfurter, Quest for GHZ States, Acta Phys. Pol. A **93**, 187 (1998); M. Hillery, V. Bužek, and A Berthiaume, Quantum secret sharing, Phys. Rev. A **59**, 1829 (1999); C. Schmid, P. Trojek, M. Bourennane, C. Kurtsiefer, M. Żukowski, and H. Weinfurter, Experimental Single Qubit Quantum Secret Sharing, Phys. Rev. Lett. **95**, 230505 (2005).

[29] M. Pawłowski and A. Winter, Hyperbits: The information quasiparticles, Phys. Rev. A **85**, 022331 (2012).

[30] M. Pawłowski and M. Żukowski, Entanglement-assisted random access codes, Phys. Rev. A **81**, 042326 (2010).

[31] A. Hameedi, D. Saha, P. Mironowicz, M. Pawłowski, and M. Bourennane, Complementarity between entanglement-assisted and quantum distributed random access code, Phys. Rev. A **95**, 052345 (2017).

[32] A. Tavakoli, B. Marques, M. Pawłowski, and M. Bourennane, Spatial versus sequential correlations for random access coding, Phys. Rev. A **93**, 032336 (2016).

[33] A. Tavakoli, M. Pawłowski, M. Żukowski, and M. Bourennane, Dimensional discontinuity in quantum communication complexity at dimension seven, Phys. Rev. A **95**, 020302(R) (2017).

[34] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).

[35] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu, Bell Inequalities for Arbitrarily High-Dimensional Systems, Phys. Rev. Lett. **88**, 040404 (2002).

[36] L. Masanes, Tight Bell inequality for *d*-outcome measurements correlations, Quantum Inf. Comput. **3**, 345 (2002).

[37] M. Navascués and H. Wunderlich, A glance beyond the quantum model, Proc. R. Soc. A **466**, 88190 (2009).

[38] M. Pawłowski and N. Brunner, Semi-device-independent security of one-way quantum key distribution, Phys. Rev. A **84**, 010302(R) (2011).

[39] S. Zohren and R. D. Gill, Maximal Violation of the Collins-Gisin-Linden-Massar-Popescu Inequality for Infinite Dimensional States, Phys. Rev. Lett. **100**, 120406 (2008).

[40] R. T. Thew, A. Acín, H. Zbinden, and N. Gisin, Bell-Type Test of Energy-Time Entangled Qutrits, Phys. Rev. Lett. **93**, 010503 (2004).

[41] A. C. Dada, J. Leach, G. S. Buller, M. J. Padgett, and E. Andersson, Experimental high-dimensional two-photon entanglement and violations of generalized Bell inequalities, Nat. Phys. **7**, 677 (2011).

[42] H.-P. Lo, C.-M. Li, A. Yabushita, Y.-N. Chen, C.-W. Luo, and T. Kobayashi, Experimental violation of Bell inequalities for multi-dimensional systems, Sci. Rep. **6**, 22088 (2016).

[43] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.121.150504, for numerical findings for quantum communication strategies and a discussion comparing quantum CCP implementations.

[44] A. Hameedi, A. Tavakoli, B. Marques, and M. Bourennane, Communication Games Reveal Preparation Contextuality, Phys. Rev. Lett. **119**, 220402 (2017).

[45] R. F. Werner and M. M. Wolf, Bell inequalities and entanglement, Quantum Inf. Comput. **1**, 1 (2001).

[46] K. F. Pál and T. Vértesi, Maximal violation of a bipartite three-setting, two-outcome Bell inequality using infinite-dimensional quantum systems, Phys. Rev. A **82**, 022116 (2010).
[47] L. Vandenberghe and S. Boyd, Semidefinite programming, SIAM Rev. **38**, 49 (1996).
[48] M. Navascués and T. Vértesi, Bounding the Set of Finite Dimensional Quantum Correlations, Phys. Rev. Lett. **115**, 020501 (2015).
[49] A. Tavakoli, D. Rosset, and M-O. Renou, Enabling computation of correlation bounds for finite-dimensional quantum systems via symmetrisation, arXiv:1808.02412.
[50] K. F. Pál and T. Vértesi, Quantum bounds on Bell inequalities, Phys. Rev. A **79**, 022120 (2009).
[51] S. Luo, Quantum discord for two-qubit systems, Phys. Rev. A **77**, 042303 (2008).
[52] G. Lima, A. Vargas, L. Neves, R. Guzmán, and C. Saavedra, Manipulating spatial qudit states with programmable optical devices, Opt. Express **17**, 10688 (2009).
[53] G. Lima, L. Neves, R. Guzmán, E. S. Gómez, W. A. T. Nogueira, A. Delgado, A. Vargas, and C. Saavedra, Experimental quantum tomography of photonic qudits via mutually unbiased basis, Opt. Express **19**, 3542 (2011).
[54] D. Goyeneche, G. Cañas, S. Etcheverry, E. S. Gómez, G. B. Xavier, G. Lima, and A. Delgado, Five Measurement Bases Determine Pure Quantum States on Any Dimension, Phys. Rev. Lett. **115**, 090401 (2015).
[55] G. Cañas, S. Etcheverry, E. S. Gómez, C. Saavedra, G. B. Xavier, G. Lima, and A. Cabello, Experimental implementation of an eight-dimensional Kochen-Specker set and observation of its connection with the Greenberger-Horne-Zeilinger theorem, Phys. Rev. A **90**, 012119 (2014).
[56] E. A. Aguilar, M. Farkas, D. Martínez, M. Alvarado, J. Carine, G. B. Xavier, J. F. Barra, G. Cañas, M. Pawłowski, and G. Lima, Certifying an Irreducible 1024-Dimensional Photonic State Using Refined Dimension Witnesses, Phys. Rev. Lett. **120**, 230503 (2018).
[57] Q. P. Stefano, L. Rebón, S. Ledesma, and C. Iemmi, Determination of any pure spatial qudits from a minimum number of measurements by phase-stepping interferometry, Phys. Rev. A **96**, 062328 (2017).
[58] M. A. Solís-Prosser, M. F. Fernandes, O. Jiménez, A. Delgado, and L. Neves, Experimental Minimum-Error Quantum-State Discrimination in High Dimensions, Phys. Rev. Lett. **118**, 100501 (2017).
[59] B. Marques, A. A. Matoso, W. M. Pimenta, A. J. Gutiérrez-Esparza, M. F. Santos, and S. Pádua, Experimental simulation of decoherence in photonics qudits, Sci. Rep. **5**, 16049 (2015).
[60] M. Mirhosseini, O. S Magña-Loaiza, M. N. OSullivan, B. Rodenburg, M. Malik, M. P. J. Lavery, M. J. Padgett, D. J Gauthier, and R. W. Boyd, High-dimensional quantum cryptography with twisted light, New J. Phys. **17**, 033033 (2015).
[61] M. Mafu, A. Dudley, S. Goyal, D. Giovannini, M. McLaren, M. J. Padgett, T. Konrad, F. Petruccione, N. Lutkenhaus, and A. Forbes, Higher-dimensional orbital-angular-momentum-based quantum key distribution with mutually unbiased bases, Phys. Rev. A **88**, 032305 (2013).
[62] V. D'Ambrosio, F. Bisesto, F. Sciarrino, J. F. Barra, G. Lima, and A. Cabello, Device-Independent Certification of High-Dimensional Quantum Systems, Phys. Rev. Lett. **112**, 140503 (2014).
[63] F. Flamini, N. Spagnolo, and F. Sciarrino, Photonic quantum information processing: A review, arXiv:1803.02790.
[64] G. Cañas *et al.*, High-dimensional decoy-state quantum key distribution over multicore telecommunication fibers, Phys. Rev. A **96**, 022317 (2017).
[65] I. Moreno, P. Velásquez, C. R. Fernández-Pousa, M. M. Sánchez-López, and F. Mateos, Jones matrix method for predicting and optimizing the optical modulation properties of a liquid-crystal display, J. Appl. Phys. **94**, 3697 (2003).
[66] The hierarchy level is sometimes referred to as $1 + \mathrm{AB} + \mathrm{BB}$ and corresponds to monomials of the form $\{\mathbb{1}, \rho_{x_0,x}, M_y^b, \rho_{x_0,x} M_y^b, M_y^b M_{y'}^{b'}\}$.

Physica Scripta

PAPER

Improving autonomous thermal entanglement generation using a common reservoir

To cite this article: Zhong-Xiao Man *et al* 2019 *Phys. Scr.* **94** 075101

IOP Publishing Physica Scripta
Phys. Scr. 94 (2019) 075101 (12pp) https://doi.org/10.1088/1402-4896/ab0c51

Improving autonomous thermal entanglement generation using a common reservoir

Zhong-Xiao Man[1,3], Armin Tavakoli[2], Jonatan Bohr Brask[2] and Yun-Jie Xia[1]

[1] School of Physics and Physical Engineering, Shandong Provincial Key Laboratory of Laser Polarization and Information Technology, Qufu Normal University, 273165, Qufu, People's Republic of China
[2] Department of Applied Physics, University of Geneva, 1211 Geneva, Switzerland

E-mail: zxman@qfnu.edu.cn

Received 21 November 2018, revised 1 February 2019
Accepted for publication 4 March 2019
Published 16 April 2019

Abstract

We study the entanglement generated in the steady state of two interacting qubits coupled to thermal reservoirs. We show that the amount of steady-state entanglement can be enhanced by the presence of a third thermal reservoir which is common to both qubits. Specifically, we find that entanglement can be enhanced as long as the temperature of the common reservoir is below the thermalisation temperature of the qubits, whenever a single temperature can be assigned to the steady state of the qubits in the absence of the common reservoir. Moreover, the amount of entanglement generated with the common reservoir present can be significantly larger than that which can be obtained without it for any temperature of the individual reservoirs. From the perspective of thermodynamics, we find that enhancement of entanglement is associated with heat absorption by the common reservoir. We propose a possible implementation of our scheme in superconducting circuits and find that a significant enhancement of steady-state entanglement should be observable under experimentally realistic conditions.

Keywords: quantum thermodynamics, entanglement enhancement, steady-state entanglement

(Some figures may appear in colour only in the online journal)

1. Introduction

Quantum entanglement is a fundamental concept in quantum mechanics as well as a key resource in quantum information science, e.g. for quantum communication, computation, and metrology [1, 2]. Entanglement is notoriously fragile in the presence of environmental noise, complicating the realisation of practical applications. Hence, understanding how to generate, protect, and enhance entanglement in different environments is important both fundamentally and for enabling quantum information technologies.

A large body of work has been devoted to enhancing and protecting entanglement via direct manipulation, for example through entanglement purification [3–5], quantum error correction [6, 7], dynamical decoupling [8–10], or exploiting the quantum Zeno effect [11, 12], or weak measurements [13–16]. In addition to these strategies, which aim to counter the effects of noise, it turns out that dissipation can also be beneficial under certain conditions, and can be exploited for entanglement generation in both transient and steady regimes [17–22] in various physical contexts [23–29]. Driven dissipative preparation of entangled states has been demonstrated experimentally for atomic ensembles [30], trapped ions [31, 32], and superconducting qubits [33].

Entanglement can also be generated thermally, without any driving. In a composite, interacting quantum system, the energetic ground state may be entangled, and hence cooling the system sufficiently will generate entanglement. In thermal equilibrium at higher temperatures, entanglement may still be present. In fact, the topic of how entanglement varies with temperature has long been a concern of condensed-matter physicists [34–40]. In particular [34–37] studied the variation

[3] Author to whom any correspondence should be addressed.

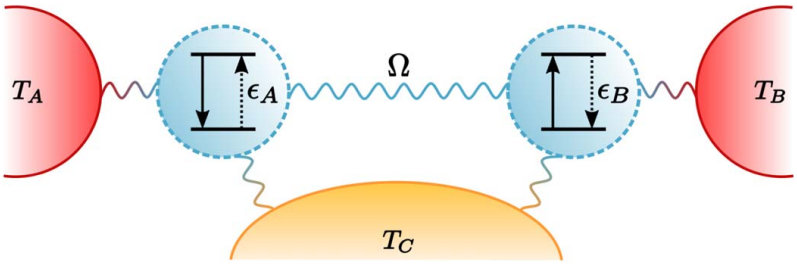

Figure 1. Schematic diagram of the physical model under consideration. Two qubits are coupled to each other with a strength Ω and to two independent heat reservoirs with temperatures T_A and T_B, respectively. A common reservoir with temperature T_C is introduced to improve the entanglement of the qubits.

of entanglement with temperature and magnetic field in spin chains in thermal equilibrium.

Interestingly, entanglement can be enhanced by moving out of thermal equlibrium where temperature gradients induce energy currents among the interacting subsystems. [41] found increase in entanglement due to an energy current in a spin chain. [42] studied changes in steady-state entanglement in a model of two interacting qubits coupled to different heat baths. The temperature gradient was shown to enhance or suppress entanglement depending on the internal coupling strength between the qubits. The dynamics of nonequilibrium thermal entanglement in a similar model was studied in [43] with particular attention to the case of non-resonant qubits, and [44, 45] studied chains of three qubits out of equilibrium. [46–49] demonstrated that entanglement can enhance the performance of quantum thermal machines, and that such machines can be harnessed for entanglement generation. In particular, in [48] a simple two-qubit thermal machine was presented which generates steady-state entanglement by operating between heat reservoirs at different temperatures. A similar two-qudit machine combined with filtering enables generation of maximal entanglement in any dimension when the temperature gradient is maximal [49]. All of these works confirm that there are strong connections between thermal entanglement and quantum thermodynamics.

While a lot can be learned from and achieved with coupled qubits in contact with independent heat reservoirs, in practical situations there will often be coupling to a common environment as well, and it is also interesting theoretically to understand the effects of such a shared reservoir. In fact, a common reservoir may itself enable entanglement generation. It was shown that entanglement between two qubits could be induced by a common, thermal, single-mode field [50]. Similarly, qubits in a common heat bath can become entangled when evolving through a purely noisy mechanism [51, 52], and steady-state entanglement is found for qubits immersed in a common thermal reservoir [40]. A common environment out of thermal equilibrium could lead to many-body entangled steady states [53, 54] and protect entanglement during evolution [55].

Here, we study thermal entanglement generation when both independent and common heat reservoirs are involved. We consider two interacting qubits coupled to individual heat reservoirs, as in the thermal machine of [48], as well as to a common reservoir. We show that the steady-state entanglement can be enhanced by the presence of this common reservoir, and that the lower the temperature of the common reservoir, the larger the enhancement. The maximal critical temperature of the common reservoir enabling entanglement growth is the thermalized temperature of the coupled qubits if thermalization is achieved. Entanglement enhancement is accompanied by a thermodynamics process where heat is dissipated into the common reservoir. We also present a possible implementation of our scheme in superconducting circuits. We find that for experimentally accessible parameter settings, a significant improvement of steady-state entanglement can be realized.

2. Model

The system we consider, as depicted in figure 1, consists of two coupled qubits A and B interacting with two independent heat reservoirs R_A and R_B, respectively, and potentially also to a common heat reservoir R_C. The Hamiltonian of the two qubits $\hat{H}_S = \hat{H}_0 + \hat{H}_{\rm int}$ with the free Hamiltonian

$$\hat{H}_0 = \epsilon_A |1\rangle_A \langle 1| \otimes \mathbb{I}_B + \epsilon_B \, \mathbb{I}_A \otimes |1\rangle_B \langle 1|, \tag{1}$$

and the interaction Hamiltonian

$$\hat{H}_{\rm int} = \Omega(\hat{\sigma}_+^A \otimes \hat{\sigma}_-^B + \hat{\sigma}_-^A \otimes \hat{\sigma}_+^B), \tag{2}$$

where $|0\rangle_\mu$ and $|1\rangle_\mu$ are the ground and excited states of qubit $\mu \in \{A, B\}$ with energy gap ϵ_μ, $\mathbb{I}_\mu$ denotes the identity operator, $\hat{\sigma}_+^\mu = |1\rangle_\mu\langle 0|$ and $\hat{\sigma}_-^\mu = |0\rangle_\mu\langle 1|$ are the raising and lowering operators for qubit μ, and Ω is qubit–qubit coupling strength.

The bosonic reservoirs are assumed to be thermal at temperatures T_A, T_B, and T_C. They are described by the Hamiltonian

$$\hat{H}_R = \sum_l \omega_{a,l} \hat{a}_l^\dagger \hat{a}_l + \sum_m \omega_{b,m} \hat{b}_m^\dagger \hat{b}_m + \sum_n \omega_{c,n} \hat{c}_n^\dagger \hat{c}_n. \tag{3}$$

Here, $\hat{a}_l^\dagger$ and $\hat{a}_l$ are creation and annihilation operators for mode l of reservoir R_A, with frequency $\omega_{a,l}$, and similarly for R_B and R_C. The interaction between the qubits and reservoirs

is given by

$$\hat{H}_{SR} = \sum_l g_{A,l}(\hat{\sigma}_+^A \hat{a}_l + \hat{\sigma}_-^A \hat{a}_l^\dagger) + \sum_m g_{B,m}(\hat{\sigma}_+^B \hat{b}_m + \hat{\sigma}_-^B \hat{b}_m^\dagger) + \sum_n [(g_{A,n}\hat{\sigma}_+^A + g_{B,n}\hat{\sigma}_+^B)\hat{c}_n + (g_{A,n}\hat{\sigma}_-^A + g_{B,n}\hat{\sigma}_-^B)\hat{c}_n^\dagger], \tag{4}$$

where $g_{A,l}$, $g_{B,m}$ are the coupling strengths of qubit A, B with mode l, m of reservoir R_A, R_B respectively, while $g_{A,n}$ and $g_{B,n}$ denote that of qubit A and B respectively with mode n of R_C. Here, we have used the rotating-wave approximation in equation (4) since the system-reservoir coupling strengthes are assumed to be much smaller than the system energy scale.

Based on the model given by $\hat{H}_S$, $\hat{H}_R$, and $\hat{H}_{SR}$, we proceed to construct a master equation for the evolution of the system qubits in the presence of the thermal reservoirs. We will work in the regime of weak system-baths interaction, where all system transition frequencies are large compared to the bath couplings. The reservoirs then couple to the delocalized eigenstates of the total system Hamiltonian H_S, and we will obtain a global master equation where each reservoir affects both qubits. For weak inter-system coupling one should instead employ a local master equation when each qubit is affected only by its local baths, as used e.g. in [48]. The global approach is valid as long as the secular approximation holds, as detailed in [56] where the validity regime for local and global master equations for a thermal machine of two qubits or two harmonic oscillator was studied.

We construct the master equation in the basis of the eigenstates $\hat{H}_S$. In terms of the free Hamiltonian eigenstates, i.e. $|\eta_1\rangle = |11\rangle$, $|\eta_2\rangle = |10\rangle$, $|\eta_3\rangle = |01\rangle$, and $|\eta_4\rangle = |00\rangle$, the eigenstates of $\hat{H}_S$ can be expressed as $|\lambda_1\rangle = |\eta_1\rangle$, $|\lambda_2\rangle = \cos\frac{\theta}{2}|\eta_2\rangle + \sin\frac{\theta}{2}|\eta_3\rangle$, $|\lambda_3\rangle = -\sin\frac{\theta}{2}|\eta_2\rangle + \cos\frac{\theta}{2}|\eta_3\rangle$, and $|\lambda_4\rangle = |\eta_4\rangle$, and the corresponding eigenvalues as $E_1 = \epsilon_A + \epsilon_B$, $E_2 = \epsilon_m + \sqrt{\Delta\epsilon^2/4 + \Omega^2}$, $E_3 = \epsilon_m - \sqrt{\Delta\epsilon^2/4 + \Omega^2}$, $E_4 = 0$ with $\epsilon_m = (\epsilon_A + \epsilon_B)/2$ and $\Delta\epsilon = \epsilon_A - \epsilon_B$. The parameter θ is defined by $\tan\theta = 2\Omega/\Delta\epsilon$.

In terms of eigenstates of $\hat{H}_S$, the total Hamiltonian $\hat{H}_{\rm tot} = \hat{H}_S + \hat{H}_R + \hat{H}_{SR}$ can be rewritten using

$$\hat{H}_S = \sum_{i=1}^4 E_i|\lambda_i\rangle\langle\lambda_i|, \qquad \text{and} \qquad \hat{H}_{SR} = \sum_{j=1}^2 \hat{\tilde{H}}_{SR,j}, \tag{5}$$

where

$$\hat{\tilde{H}}_{SR,j} = \sum_l g_{A,l}(\hat{V}_{A,j}^+ \hat{a}_l + \hat{V}_{A,j}\hat{a}_l^\dagger) + \sum_m g_{B,m}(\hat{V}_{B,j}^+ \hat{b}_m + \hat{V}_{B,j}\hat{b}_m^\dagger) + \sum_n [(g_{A,n}\hat{V}_{A,j}^+ + g_{B,n}\hat{V}_{B,j}^+)\hat{c}_n + (g_{A,n}\hat{V}_{A,j} + g_{B,n}\hat{V}_{B,j})\hat{c}_n^\dagger]. \tag{6}$$

In this expression, $\hat{V}_{\mu,j}$ and $\hat{V}_{\mu,j}^\dagger$ are jump operators corresponding respectively to processes where the system looses an excitation to a bath or receives one from it. They are eigenoperators of $\hat{H}_S$, such that $[\hat{H}_S, \hat{V}_{\mu,j}] = -\omega_j \hat{V}_{\mu,j}$ where the eigenfrequencies ω_j determine the energy lost or recieved by the system. They are given by $\omega_1 = E_3 - E_4 = E_1 - E_2 = \epsilon_m - \sqrt{\Delta\epsilon^2/4 + \Omega^2}$, corresponding to transitions $|\lambda_1\rangle \leftrightarrow |\lambda_2\rangle$ and $|\lambda_3\rangle \leftrightarrow |\lambda_4\rangle$, and $\omega_2 = E_1 - E_3 = E_2 - E_4 = \epsilon_m + \sqrt{\Delta\epsilon^2/4 + \Omega^2}$ corresponding to transitions $|\lambda_1\rangle \leftrightarrow |\lambda_3\rangle$ and $|\lambda_2\rangle \leftrightarrow |\lambda_4\rangle$. Explicitly, the $\hat{V}_{\mu,j}$ are constructed as follows

$$\begin{aligned}
\hat{V}_{A,1} &= \sin\frac{\theta}{2}(|\lambda_2\rangle\langle\lambda_1| - |\lambda_4\rangle\langle\lambda_3|),\\
\hat{V}_{A,2} &= \cos\frac{\theta}{2}(|\lambda_3\rangle\langle\lambda_1| + |\lambda_4\rangle\langle\lambda_2|),\\
\hat{V}_{B,1} &= \cos\frac{\theta}{2}(|\lambda_2\rangle\langle\lambda_1| + |\lambda_4\rangle\langle\lambda_3|),\\
\hat{V}_{B,2} &= \sin\frac{\theta}{2}(-|\lambda_3\rangle\langle\lambda_1| + |\lambda_4\rangle\langle\lambda_2|).
\end{aligned} \tag{7}$$

With the jump operators $\hat{V}_{\mu,j}$, one can derive a master equation on standard Lindblad form in the Born–Markov regime of weak coupling to the thermal reservoirs combined with a secular approximation, valid for strong inter-system coupling. Details of deriving a master equation from the system, bath, and interaction Hamiltonians can be found e.g. in [57] chapter 3 and [56, 58, 59]. A derivation with a common reservoir, as considered here, can be found in [40]. One assumes a stationary state of the baths—i.e. that the baths are sufficiently large to remain unaffected by the interaction with the system—and neglects the Lamb shift, which is small compared to the qubit–qubit coupling strength Ω [57, 58, 60] (see also the appendix). We arrive at

$$\dot{\rho} = -\mathrm{i}[\hat{H}_S, \rho] + \mathcal{L}_A[\rho] + \mathcal{L}_B[\rho] + \mathcal{L}_C[\rho], \tag{8}$$

where $\mathcal{L}_A[\rho]$, $\mathcal{L}_B[\rho]$, and $\mathcal{L}_C[\rho]$ describe the dissipative effect on the qubits' dynamics due to coupling with the reservoirs R_A, R_B, and R_C respectively. We note that, as we are working in the Born–Markov regime, the dissipators are additive [61] and so we can obtain the dynamics in the absence of R_C simply by omitting the last term above. The dissipators arising from the independent baths are given by

$$\mathcal{L}_A[\rho] = \sum_j \Gamma_A(\omega_j)[(\bar{n}_A(\omega_j) + 1)(2\hat{V}_{A,j}\rho\hat{V}_{A,j}^\dagger - \{\hat{V}_{A,j}^\dagger\hat{V}_{A,j}, \rho\}) + \bar{n}_A(\omega_j)(2\hat{V}_{A,j}^\dagger\rho\hat{V}_{A,j} - \{\hat{V}_{A,j}\hat{V}_{A,j}^\dagger, \rho\})], \tag{9}$$

and

$$\mathcal{L}_B[\rho] = \sum_j \Gamma_B(\omega_j)[(\bar{n}_B(\omega_j) + 1)(2\hat{V}_{B,j}\rho\hat{V}_{B,j}^+ - \{\hat{V}_{B,j}^+\hat{V}_{B,j}, \rho\}) + \bar{n}_B(\omega_j)(2\hat{V}_{B,j}^+\rho\hat{V}_{B,j} - \{\hat{V}_{B,j}\hat{V}_{B,j}^+, \rho\})]. \tag{10}$$

In each case, the first line corresponds to stimulated and spontaneous emission, while the second line corresponds to absorption. $\Gamma_A(\omega_j)$ and $\Gamma_B(\omega_j)$, characterize the damping rates due to interactions with the reservoirs R_A and R_B respectively. Their exact forms depend on the spectral densities of the reservoirs. Each reservoir is assumed to be in a thermal state, and the occupation number (the average number of photons) at energy ω_j of reservoir R_ν ($\nu \in \{A, B, C\}$) is given by the Bose–Einstein distribution

$$\bar{n}_\nu(\omega_j) = \frac{1}{\exp[\frac{\omega_j}{T_\nu}] - 1}. \tag{11}$$

In contrast to R_A and R_B, the reservoir R_C is common to the two qubits A and B, and we see from (6) that it will introduce dissipative terms both of the forms (9) and (10) as well as cross terms Therefore, we have $\mathcal{L}_C[\rho] = \mathcal{L}_C^{(A)}[\rho] + \mathcal{L}_C^{(B)}[\rho] + \mathcal{L}_C^{(AB)}[\rho]$, in which $\mathcal{L}_C^{(A)}[\rho]$ and $\mathcal{L}_C^{(B)}[\rho]$ indicate dissipative effects due to qubit A and B coupling individually to R_C, while the term $\mathcal{L}_C^{(AB)}[\rho]$ reflects the collective coupling. Thanks to the collective effect of the common reservoir, the steady-state entanglement induced by independent reservoirs can be further enhanced. Explicitly

$$\begin{aligned}\mathcal{L}_C^{(A)}[\rho] = &\sum_j \Gamma_C^{(A)}(\omega_j)[(\bar{n}_C(\omega_j)+1)(2\hat{V}_{A,j}\rho\hat{V}_{A,j}^+ \\ &- \{\hat{V}_{A,j}^+\hat{V}_{A,j},\,\rho\}) \\ &+ \bar{n}_c(\omega_j)(2\hat{V}_{A,j}^+\rho\hat{V}_{A,j} - \{\hat{V}_{A,j}\hat{V}_{A,j}^+,\,\rho\})],\end{aligned} \tag{12}$$

$$\begin{aligned}\mathcal{L}_C^{(B)}[\rho] = &\sum_j \Gamma_C^{(B)}(\omega_j)[(\bar{n}_C(\omega_j)+1)(2\hat{V}_{B,j}\rho\hat{V}_{B,j}^+ \\ &- \{\hat{V}_{B,j}^+\hat{V}_{B,j},\,\rho\}) \\ &+ \bar{n}_C(\omega_j)(2\hat{V}_{B,j}^+\rho\hat{V}_{B,j} - \{\hat{V}_{B,j}\hat{V}_{B,j}^+,\,\rho\})],\end{aligned} \tag{13}$$

and

$$\begin{aligned}\mathcal{L}_C^{(AB)}[\rho] = &\sum_j \Gamma_C^{(AB)}(\omega_j)[(\bar{n}_C(\omega_j)+1)(2\hat{V}_{A,j}\rho\hat{V}_{B,j}^+ \\ &- \{\hat{V}_{B,j}^+\hat{V}_{A,j},\,\rho\}) \\ &+ \bar{n}_C(\omega_j)(2\hat{V}_{A,j}^+\rho\hat{V}_{B,j} - \{\hat{V}_{B,j}\hat{V}_{A,j}^+,\,\rho\}) \\ &+ (\bar{n}_C(\omega_j)+1)(2\hat{V}_{B,j}\rho\hat{V}_{A,j}^+ - \{\hat{V}_{A,j}^+\hat{V}_{B,j},\,\rho\}) \\ &+ \bar{n}_C(\omega_j)(2\hat{V}_{B,j}^+\rho\hat{V}_{A,j} - \{\hat{V}_{A,j}\hat{V}_{B,j}^+,\,\rho\})].\end{aligned} \tag{14}$$

The collective damping rate fulfils $\Gamma_C^{(AB)}(\omega_j) = \sqrt{\Gamma_C^{(A)}(\omega_j)\Gamma_C^{(B)}(\omega_j)}$. For simplicity, in the remainder of the paper we will suppose that all the spectral densitites can be taken to be flat in the relevant energy range such that the damping rates are frequency independent, $\Gamma_A(\omega_j) = \Gamma_A$, $\Gamma_B(\omega_j) = \Gamma_B$, $\Gamma_C^A(\omega_j) = \Gamma_C^{(A)}$ and $\Gamma_C^{(B)}(\omega_j) = \Gamma_C^{(B)}$.

We are interested in the steady-state entanglement between the two qubits. The steady state is found by setting the left hand side of equation (8) to zero, i.e. by solving $\dot{\rho}^S = 0$. The entanglement of the resulting two-qubit state can then be quantified by the concurrence [62]. We obtain the steady state in the eigenbasis of H_S with the density matrix elements $\lambda_{ii'}^S = \langle\lambda_i|\,\rho^S|\lambda_{i'}\rangle$. The state can then be reexpressed in the eigenbasis of the free Hamiltonian with density matrix elements $\eta_{ii'}^S = \langle\eta_i|\,\rho^S|\eta_{i'}\rangle$ using

$$\begin{aligned}\eta_{11}^S &= \lambda_{11}^S, \\ \eta_{22}^S &= \cos^2\frac{\theta}{2}\lambda_{22}^S + \sin^2\frac{\theta}{2}\lambda_{33}^S, \\ \eta_{33}^S &= \sin^2\frac{\theta}{2}\lambda_{22}^S + \cos^2\frac{\theta}{2}\lambda_{33}^S, \\ \eta_{44}^S &= \lambda_{44}^S, \\ \eta_{23}^S &= \eta_{32}^S = \frac{1}{2}\sin\theta(\lambda_{22}^S - \lambda_{33}^S).\end{aligned} \tag{15}$$

The only mechanism, which can generate coherence, is the inter-qubit interaction described by the Hamiltonian (2). In fact, if there were no interaction between the qubits, there would be no process generating off-diagonal terms in the free Hamiltonian eigenbasis. The bath-induced dissipation tends to destroy coherence. Thus, only coherences induced by the interaction can survive in the steady state, and ρ^S will therefore be of the form

$$\rho^S = \begin{pmatrix}\eta_{11}^S & 0 & 0 & 0 \\ 0 & \eta_{22}^S & \eta_{23}^S & 0 \\ 0 & \eta_{32}^S & \eta_{33}^S & 0 \\ 0 & 0 & 0 & \eta_{44}^S\end{pmatrix}. \tag{16}$$

This is a so-called 'X state' for which the concurrence reduces to the simple expression [63]

$$\mathcal{C}(\rho^S) = 2\max\{0,\,|\eta_{23}^S| - \sqrt{\eta_{11}^S\eta_{44}^S}\}. \tag{17}$$

The state is entangled whenever $\mathcal{C}(\rho^S) > 0$ and maximally entangled for $\mathcal{C}(\rho^S) = 1$

In addition to the entanglement, it also interesting to look at the heat currents in the system. The introduction of a common reservoir with its own associated temperature will influence both the entanglement and heat current, and we will investigate this link below. The heat current associated with reservoir R_ν can be defined as[4] [57, 65]

$$Q_\nu = Tr\{\mathcal{L}_\nu[\rho^S]\hat{H}_S\}. \tag{18}$$

In equation (18), $\mathcal{L}_\nu[\rho^S]$ represents the change in the system state induced by the bath ν and the trace of it with the system Hamiltonian hence represents the associated change in the system energy. From the perspective of reservoir, a positive heat current means heat release from the reservoir, while a negative value implies heat absorption by the reservoir. Therefore, a sign change of the heat current indicates a cross-over between heat absorption and heat release or vice versa.

3. Results

We now analyse how steady-state entanglement generation and heat currents are influenced by the introduction of the common heat reservoir R_C. We first consider the case where the two independent reservoirs R_A and R_B are in thermal equilibrium, i.e. $T_A = T_B$, and then turn to the out-of-equilbrium case below. We will compare the amount of steady-state entanglement with R_C present with the amount when the system is decoupled from R_C, and also examine the heat currents.

3.1. Independent reservoirs at thermal equilibrium

In this section, we consider R_A and R_B to be in the thermal equilibrium with $T_A = T_B = T$. We will also focus on the

[4] Note that while a master equation of global type and employing a secular approximation, as we do here, can lead to zero predictions for the inter-system energy current between the qubits [64], there are no inconsistencies associated with evaluating the currents between the system and reservoirs [56].

case where the qubits are resonant, $\epsilon_A = \epsilon_B = \epsilon$ (i.e. $\theta = \pi/2$). In the absence of the third reservoir R_C, the system will relax into a thermal equilibrium state with temperature T which may contain thermal entanglement [34–40]. We are interested in how the amount of entanglement varies when the common reservoir with temperature T_C is introduced and the entanglement depends on T and T_C. That is, denoting the concurrence what the system is decoupled from R_C by $\mathcal{C}_{AB}$ and the concurrence in the presence of R_C by $\mathcal{C}_{ABC}$, we want to compare $\mathcal{C}_{AB}(T)$ with $\mathcal{C}_{ABC}(T, T_C)$.

In figure 2(a), we plot the difference $\Delta\mathcal{C} = \mathcal{C}_{ABC}(T, T_C) - \mathcal{C}_{AB}(T)$ as a function of the temperatures. This is the change in steady-state entanglement induced by introducing the common reservoir at temperature T_C. As might be expected, we observe that when $T_C = T$ (red, dashed line in the figure) there is no change, because the system retains the same thermal equilibrium state with temperature T. The concurrence increases when $T_C < T$, while it decreases for $T_C > T$. Thus, the common reservoir enhances the steady-state entanglement when it effectively cooling the system. In figure 2(b), we plot the steady-state entanglement as a function of T without the common reservoir, i.e. $\mathcal{C}_{AB}(T)$ as well as with, $\mathcal{C}_{ABC}(T, T_C)$ for different T_C. We clearly see that the maximum of $\mathcal{C}_{ABC}$ can be significantly higher than that of $\mathcal{C}_{AB}$. While entanglement vanishes for large T without R_C, it can be recovered by adding the common reservoir. When the common reservoir is cold (low T_C), the steady-state entanglement can retain a finite nonzero value for larger T, indicating that the thermal gradient induced by different T and T_C assists the entanglement generation. This can be further corroborated by studying the heat current Q_C out of the reservoir R_C, which we plot in figure 3. When T_C goes from being larger than T to being smaller, the current changes sign from positive to negative meaning that the reservoir R_C begins to absorb heat. The enhancement of steady-state entanglement is thus accompanied by heat absorption of the common reservoir from the independent ones.

The enhancement of entanglement is due to the collective effect of the common reservoir, represented by the collective dissipator $\mathcal{L}_C^{(AB)}[\rho]$ in equation (14). To visualise this, in figure 4 we compare the steady-state concurrence when $\mathcal{L}_C^{(AB)}[\rho]$ is removed from the qubits' dynamics (blue curve) to that with it (black curve) as well as to that when the common reservoir is completely decoupled (red curve). Clearly, in absence of $\mathcal{L}_C^{(AB)}[\rho]$, the introduction of a common reservoir instead suppresses the concurrence for most T compared to the situation without a common reservoir. Hence we see that it is the collective action which generates the entanglement enhancement. As we have shown above, to efficiently exert the collective effect, the temperature of the common reservoir should be lower than that of the independent reservoirs in the thermal equilibrium case. Intuitively, the dissipators $\mathcal{L}_C^{(A)}$ and $\mathcal{L}_C^{(B)}$ have a similar effect as the independent reservoirs. If the common reservoirs is warmer than the independent ones, they tend to heat the system, making it more mixed and destroying entanglement. This effect competes with the enhancement induced by $\mathcal{L}_C^{(AB)}$. Also, note that the enhancement cannot be obtained without the common bath by simple cooling using the individual baths, i.e. by lowering their temperature. This

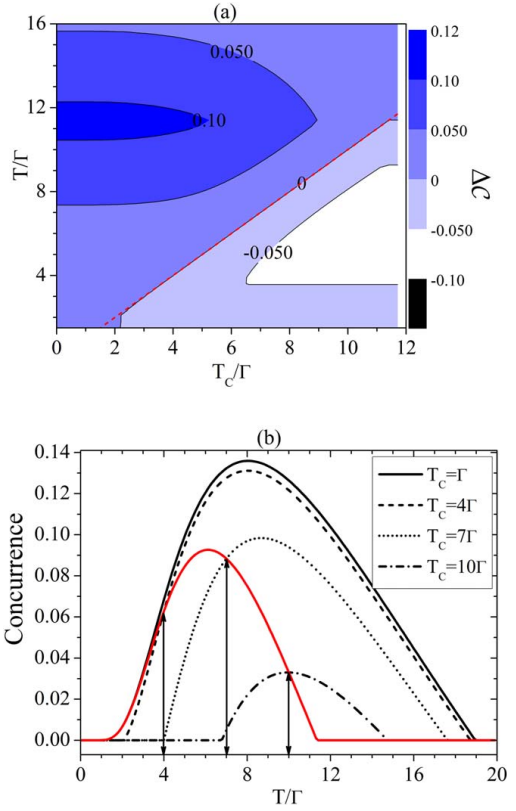

Figure 2. (a) Difference in steady-state concurrence $\Delta\mathcal{C}$ with and without the common bath versus the temperatures $T = T_A = T_B$ and T_C. The red dashed curve indicates $T_C = T$ where $\Delta\mathcal{C} = 0$. (b) Steady-state concurrence versus T without the common reservoir $\mathcal{C}_{AB}$ (red curve) and with $\mathcal{C}_{ABC}$ for different T_c (black curves). Arrows label the crossing points where $T = T_C$. In both plots, the remaining parameters are given by $\Gamma_A = \Gamma_B = \Gamma_C^{(A)} = \Gamma_C^{(B)} = \Gamma$, $\epsilon_A = \epsilon_B = 20\Gamma$, and $\Omega = 10\Gamma$.

can be seen since it is also present for temperatures below the peak of the red curve in figure 4.

3.2. Independent reservoirs out of thermal equilibrium

We now turn to the case where the two independent reservoirs R_A, R_B are not necessarily at thermal equilibrium, $T_A \neq T_B$. In the regime of weak qubit–qubit interaction, where there is negligible entanglement at equilibrium, such a temperature gradient can be harnessed for entanglement generation, as shown for thermal machines [46–49]. In [48, 49] entanglement was maximised when the temperature difference was as large as possible, e.g. for T_A approaching zero and T_B large. Here, we are interested in how the addition of a common reservoir R_C affects the amount of steady-state entanglement. In particular, we saw above that in equilibrium $T_A = T_B = T$,

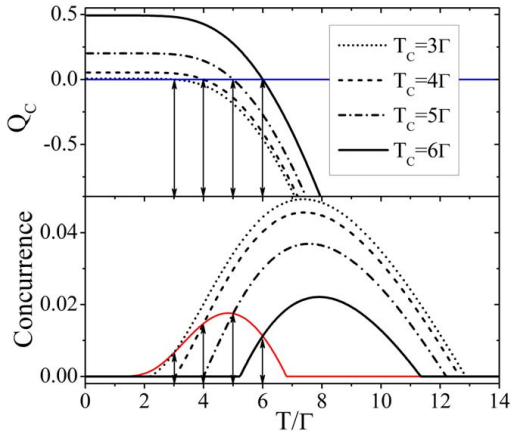

Figure 3. Heat current Q_C out of the common reservoir R_C (top panel) and steady-state concurrence (bottom panel) versus $T_A = T_B = T$ for different T_c (black curves). The red line indicates the concurrence in the absence of R_C. Arrows label the points where $T = T_C$. The remaining parameters are $\Omega = 6\Gamma$, $\epsilon_A = \epsilon_B = 20\Gamma$, $\Gamma_A = \Gamma_B = \Gamma_C^{(A)} = \Gamma_C^{(B)} = \Gamma$.

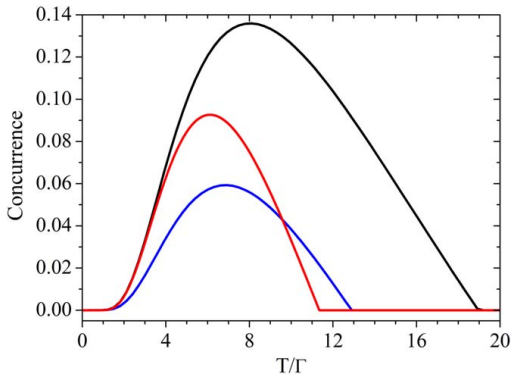

Figure 4. A comparison of the steady-state concurrence versus $T = T_A = T_B$ when the collective dissipator $\mathcal{L}_C^{(AB)}[\rho]$ (14) is removed from the qubits' dynamics (blue curve) to that with it (black curve) as well as to that when the common reservoir is decoupled (red curve). The parameters are set as $T_C = \Gamma$, $\epsilon_A = \epsilon_B = 20\Gamma$, $\Omega = 10\Gamma$ and $\Gamma_A = \Gamma_B = \Gamma_C^{(A)} = \Gamma_C^{(B)} = \Gamma$.

the addition of R_C enhances the entanglement whenever $T_C < T$. We would like to understand how this finding generalises to the nonequilibrium setting.

Since the qubit–qubit interaction is strong, where the reservoirs coupled to the delocalised eigenstates of the system Hamiltonian H_S, we can be regard our model as describing an effective four-level system connected with two independent reservoirs (in the absence of R_C). Out of equilibrium, the steady state of this system is not generally a Gibbs state, and so it is not possible to assign it a temperature in an unambiguous manner. Nevertheless, we can characterize the state of the effective four-level system via the following two effective temperatures [40, 66–68] as

$$T_{\rm eff}(\omega_1) = \frac{\omega_1}{\ln(\Gamma_1^-/\Gamma_1^+)},\ T_{\rm eff}(\omega_2) = \frac{\omega_2}{\ln(\Gamma_2^-/\Gamma_2^+)}, \tag{19}$$

where

$$\begin{aligned}
\Gamma_1^- &= \sin^2\left(\frac{\theta}{2}\right)\Gamma_A(\omega_1)[\bar{n}_A(\omega_1)+1] \\
&\quad + \cos^2\left(\frac{\theta}{2}\right)\Gamma_B(\omega_1)[\bar{n}_B(\omega_1)+1], \\
\Gamma_1^+ &= \sin^2\left(\frac{\theta}{2}\right)\Gamma_A(\omega_1)\bar{n}_A(\omega_1) + \cos^2\left(\frac{\theta}{2}\right)\Gamma_B(\omega_1)\bar{n}_B(\omega_1), \\
\Gamma_2^- &= \cos^2\left(\frac{\theta}{2}\right)\Gamma_A(\omega_2)[\bar{n}_A(\omega_2)+1] \\
&\quad + \sin^2\left(\frac{\theta}{2}\right)\Gamma_B(\omega_2)[\bar{n}_B(\omega_2)+1], \\
\Gamma_2^+ &= \cos^2\left(\frac{\theta}{2}\right)\Gamma_A(\omega_2)\bar{n}_A(\omega_2) + \sin^2\left(\frac{\theta}{2}\right)\Gamma_B(\omega_2)\bar{n}_B(\omega_2),
\end{aligned} \tag{20}$$

denote effective transition rates between the eigenstates of H_S (see e.g. [40] for details). When the two independent reservoirs are in thermal equilibrium, $T_A = T_B = T$, both effective temperatures reduce to T, consistent with the fact that the two coupled qubits eventually reach a thermal equilibrium state. By contrast, out of equilibrium, $T_A \neq T_B$, the effective temperatures are generally different, both in the range between $\min\{T_A, T_B\}$ and $\max\{T_A, T_B\}$. Depending on the reservoir temperature and the detuning between the qubits $\Delta\epsilon$, one or the other effective temperature may be larger. There thus exist some special conditions under which the two effective temperatures become equal even when reservoirs R_A and R_B are not in equilibrium, in which case we *can* assign a definite temperature to the system. We will now see that in these special situations, the result obtained in the equilibrium case above still holds.

From (19) and (20), for a given temperature gradient, we can derive a suitable energy detuning $\Delta\epsilon$ of the two qubits such that $T_{\rm eff}(\omega_1) = T_{\rm eff}(\omega_2) = T_{\rm eff}$. This is illustrated in figure 5, where we plot the two effective temperatures $T_{\rm eff}(\omega_1)$ and $T_{\rm eff}(\omega_2)$ as functions of the detuning $\Delta\epsilon$ for different temperature gradients. For the points of thermalisation in figure 5, where $T_{\rm eff}(\omega_1) = T_{\rm eff}(\omega_2) = T_{\rm eff}$, we now consider the effect of adding the common reservoir R_C. In figure 6 we show the concurrence and heat current Q_C as functions of T_C. The values in the absence of R_C are also indicated. We see that the amount of entanglement is enhanced with respect to that obtained in the absence of R_C whenever $T_C < T_{\rm eff}$. Thus our statement from the equilibrium case above generalises with T replaced by $T_{\rm eff}$. As before, the lower T_C the higher the concurrence. Again, enhancement of entanglement is associated with heat absorption by the common reservoir (Q_C becomes negative).

For the particular nonequilibrium conditions under which the qubits can be assigned a single effective temperature, we have thus shown that entanglement can be improved for T_C up to $T_{\rm eff}$. In the general nonequilibrium case where there is no

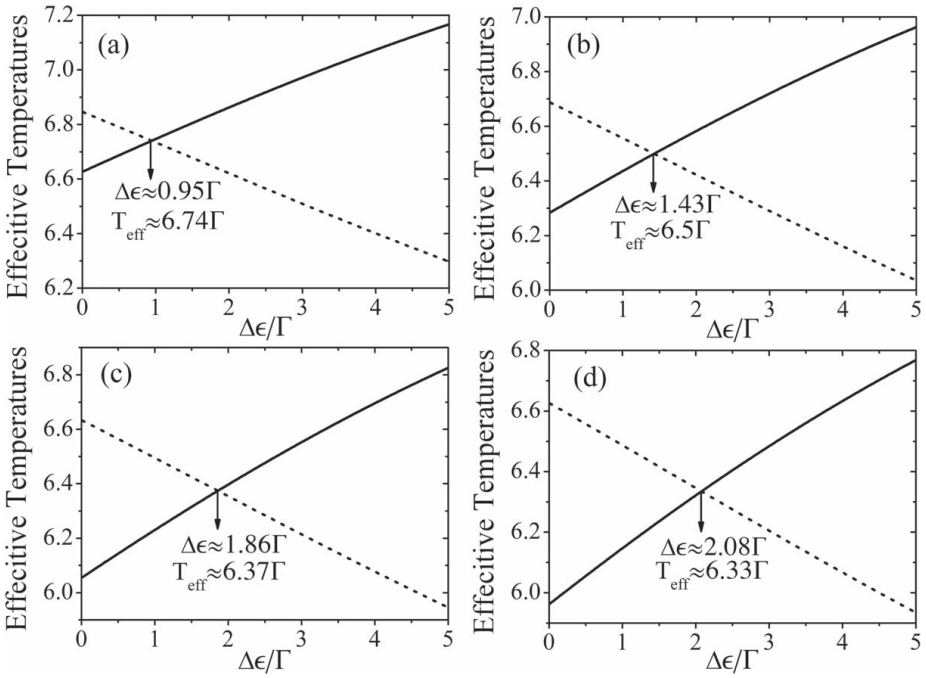

Figure 5. Plot of the effective temperatures $T_{\rm eff}$ (ω_1) (solid lines) and $T_{\rm eff}$ (ω_2) (dashed lines) in the absence of the common reservoir versus the detuning $\Delta\epsilon/\Gamma$ for $T_B = 8\Gamma$ and (a) $T_A = 5\Gamma$, (b) $T_A = 4\Gamma$, (c) $T_A = 3\Gamma$, and (d) $T_A = 2\Gamma$. The remaining parameters are $\Omega = 6\Gamma$, $\epsilon_m = 20\Gamma$, and $\Gamma_A = \Gamma_B = \Gamma$.

single thermalisation temperature, $T_{\rm eff}(\omega_1) \neq T_{\rm eff}(\omega_2)$, the addition of a common reservoir with suitable temperature can still improve the steady-state entanglement. Although we could not explicitly determine an upper bound on T_C below which entanglement is increased in this general case, we numerically verify that such upper bound lies in between the two effective temperatures $T_{\rm eff}(\omega_1)$ and $T_{\rm eff}(\omega_2)$ of the system. To be visualized, in figure 7 we exhibit the concurrence versus the temperature of the common reservoir R_C for this case. Based on the findings in figure 5 on the relations of effective temperatures and the detuning $\Delta\epsilon$, here we choose $\Delta\epsilon = 3\Gamma$ so that an effective temperature cannot be assigned to the system being contrary to the choice in figure 6. We can see from figure 7 that the entanglement promotion can still be achieved when the temperature T_C of the common reservoir is less than a value bing in between the two effective temperatures $T_{\rm eff}$ (ω_1) and $T_{\rm eff}$ (ω_2).

4. Implementation

Before we conclude, in this section we propose a possible implementation of our scheme in circuit quantum electrodynamics (QED) and compute the achievable improvement in steady-state entanglement for experimentally accessible values of the coupling parameters. A number of physical platforms could potentially enable implementations of the scheme, including trapped atoms, ions, and solid-state artificial atoms such as nitrogen-vacancy centres in diamond. However, here we focus on superconducting systems in which experimental studies of quantum thermodynamics have already been realised [69–72] and which are good candidates for implementing quantum thermal machines [48, 49, 73–75].

In circuit QED, a Hamiltonian of the form $\hat{H}_S$ can be realised by two transmon or fluxonium qubits [76], as shown in figure 8. The level spacing of fluxonium qubits is accurately tunable in a wide range from hundreds of MHz to tens of GHz. Several coupling mechanisms are available for realising the qubit–qubit interaction. Both transmon qubits [77–79] and fluxonium qubits can be coupled capacitively or inductively via a cavity in the dispersive regime (of strong detuning of the qubits and cavity from the strength of the qubit–cavity coupling) [76]. Second, an alternative is direct mutual inductive coupling as described in [80] and proposed for fluxonium qutrits in [76]. We note that while achieving strong inter-qubit coupling as considered here is certainly challenging, the dispersive regime is not a strict limitation [81, 82]. It only requires that the detuning between the qubit frequency and the cavity frequency be larger than each qubit-resonator frequency.

The qubits are naturally coupled to thermal baths due to the presence of thermal Johnson Nyquist noise in the surrounding circuitry. Effective thermal baths for each qubit can

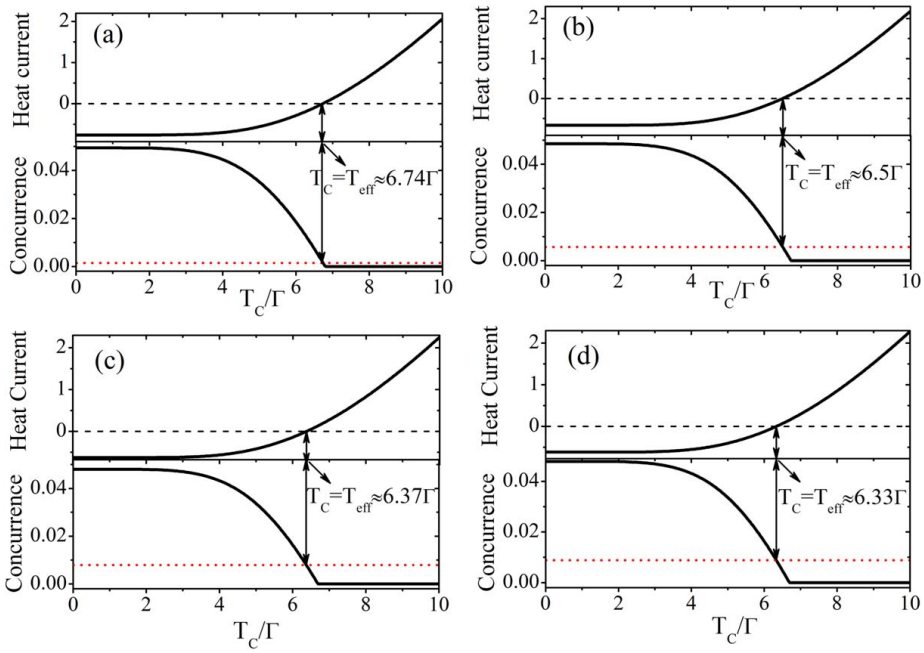

Figure 6. Each panel shows the heat current Q_C (top) and concurrence (bottom) versus the temperature of the common reservoir R_C for the thermalisation points found in figure 5, namely (a) $T_A = 5\Gamma$, $\Delta\epsilon = 0.95\Gamma$, (b) $T_A = 4\Gamma$, $\Delta\epsilon = 1.43\Gamma$, (c) $T_A = 3\Gamma$, $\Delta\epsilon = 1.86\Gamma$, and (d) $T_A = 2\Gamma$, $\Delta\epsilon = 2.08\Gamma$. Red, dotted lines indicate the concurrence in the absence of R_C. Dashed lines indicate the zero point for the heat currents, and arrows mark the points where the current changes direction. At these points, the concurrence in the presence of R_C is the same as that without it. When $T_C < T_{\text{eff}}$ the concurrence is enhanced due to the involvement of the common reservoir. The remaining parameters are the same as that in figure 5.

be implemented by controling the electronic noise coupling to each qubit. E.g. the effective temperature can be increased by increasing the noise level in particular transmission lines. A common bath coupling to both qubits can be realised similarly. Specifically, a transmission line coupling to the cavity which mediates the qubit–qubit interaction will couple to both qubits and can provide a common bath. If we model the thermal environments of the qubits by bosonic thermal reservoirs, their effect on the system is already captured by the Lindblad-type master equation (8). The system-bath coupling strengths can vary in a range of about 0.1–10 MHz. Imperfections in external control parameters, such as magnetic flux noise, will lead to additional pure dephasing [83]. We account for this phenomenologically by adding another dissipative term on the right-hand-side of (8), given by

$$\mathcal{L}_{\text{dep}}[\rho] = \gamma(\hat{D}_A\rho\hat{D}_A^\dagger + \hat{D}_B\rho\hat{D}_B^\dagger - 2\rho) \tag{21}$$

where $\hat{D}_A = \hat{\sigma}_z \otimes \mathbb{I}$ and $\hat{D}_B = \mathbb{I} \otimes \hat{\sigma}_z$ with $\hat{\sigma}_z$ the Pauli operator, and γ is the pure dephasing rate which we take to be the same for both qubits. Based on the relaxation (T_1) and Ramsey dephasing (T_2) times measured for fluxonium qubits in [84], we take the pure dephasing rate to be $\gamma = T_2^{-1} - (2T_1)^{-1} \approx 3.5 \times 10^{-2}$ MHz.

Solving for the steady state of the modified master equation, we can compute the attainable concurrence for experimentally relevant parameter settings. We obtain the result shown in figure 9, where the temperatures are given in units of the qubit transition frequency, which is set to 1 GHz. As can be seen from figure 9(a), a significant amount of entanglement can be generated in a realistic setting, even in the presence of pure dephasing. From figure 9(b) we also note that the conclusion from above is still valid: More entanglement can be generated when the common bath temperature is below the individual bath temperatures, and hence the system is out of equilibrium. The improvement in steady-state entanglement depends on the qubit–qubit interaction strength, as well as the strength of the bath couplings, as shown in figure 9(c). When the bath coupling is weaker, the improvement peaks at higher interaction strengths. Substantial improvements can be obtained for accessible parameter values.

5. Conclusion

In conclusion, we have shown that it is possible to improve the steady-state entanglement of two interacting qubits coupled to independent thermal reservoirs by simply introducing common thermal reservoir coupling to both qubits. We find that it is advantageous for the common reservoir to be cold,

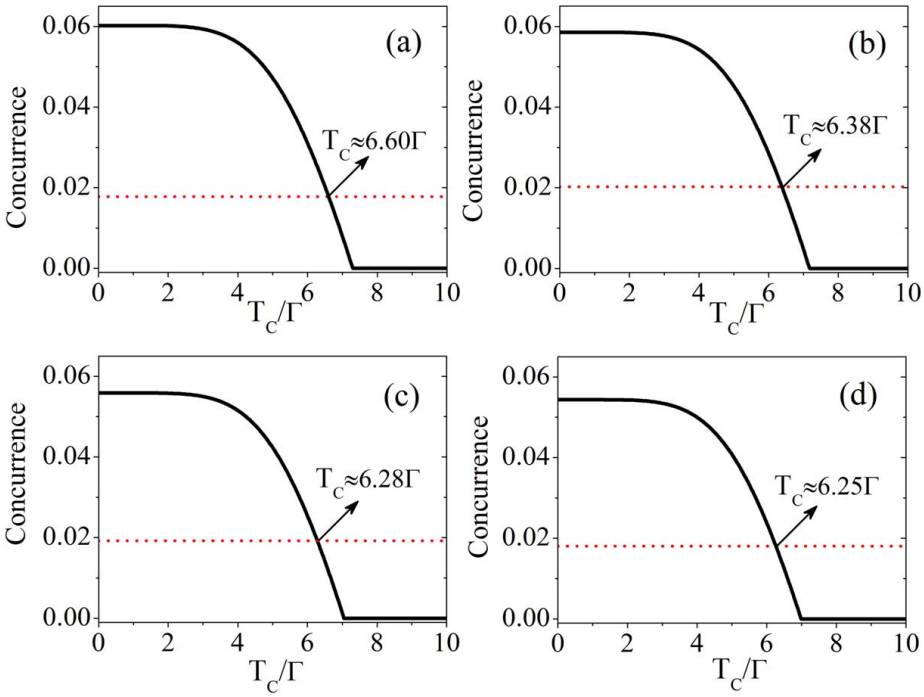

Figure 7. Each panel shows the concurrence versus the temperature of the common reservoir R_C when an effective temperature cannot be assigned to the system. The parameters in the present four panels are the same as that in figures 5 and 6 but with a choice of $\Delta\epsilon = 3\Gamma$ so that an effective temperature cannot be reached. The entanglement promotion can still be achieved when $T_C < 6.60\Gamma \in \{T_{\rm eff}(\omega_1) = 6.97\Gamma, T_{\rm eff}(\omega_2) = 6.51\Gamma\}$ in (a), $T_C < 6.38\Gamma \in \{T_{\rm eff}(\omega_1) = 6.72\Gamma, T_{\rm eff}(\omega_2) = 6.29\Gamma\}$ in (b), $T_C < 6.28\Gamma \in \{T_{\rm eff}(\omega_1) = 6.55\Gamma, T_{\rm eff}(\omega_2) = 6.21\Gamma\}$ in (c) and $T_C < 6.25\Gamma \in \{T_{\rm eff}(\omega_1) = 6.48\Gamma, T_{\rm eff}(\omega_2) = 6.20\Gamma\}$, namely, the temperature T_C of the common reservoir is less than a value bing in between the two effective temperatures $T_{\rm eff}$ (ω_1) and $T_{\rm eff}$ (ω_2).

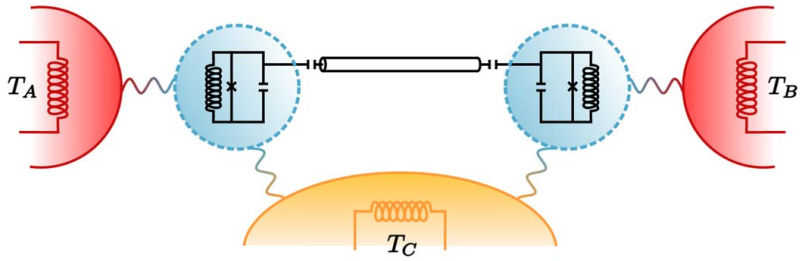

Figure 8. Possible circuit QED implementation of the scheme. Two fluxonium qubits are coupled via a microwave resonater detuned from the energy spacing of the qubits, to realise a system Hamiltonian H_S given in (1)–(2). Each qubit is coupled to effective baths with variable temperatures, corresponding to noise in external circuits which have a finite impedance. A common bath affecting both qubits is realised in the same manner. Imperfect control over external control parameters and other noise sources leads to additional pure dephasing.

and that there is a maximal temperature of this reservoir up to which entanglement is enhanced. When the two qubits in the absence of the common reservoir thermalise to a definite temperature—either because the two independent reservoirs are at thermal equilibrium or because an effective temperature can be assigned to the qubits in the steady state—then this upper bound is simply equal to the thermalisation temperature. In all cases where entanglement is enhanced, the enhancement is associated with heat absorption by the common reservoir which is thus effectively cooling the system. In the equilibrium case, we observe that with the common reservoir present, entanglement can be generated for larger temperatures of the individual reservoirs than otherwise possible. We have proposed and analysed an implementation of our scheme using superconducting qubits and have seen that even in the presence of additional dephasing and for

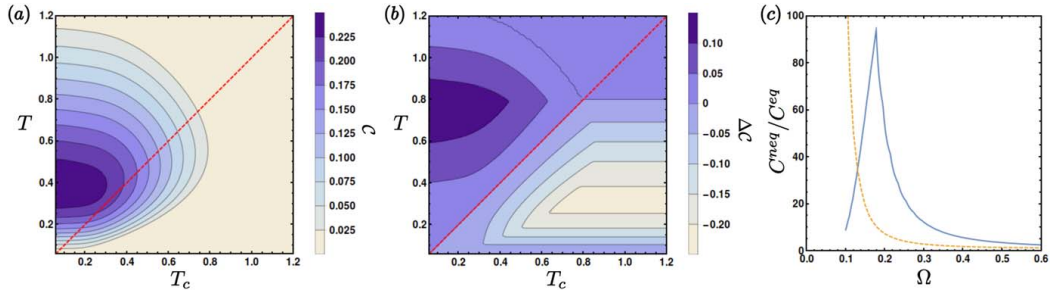

Figure 9. (a) Steady-state entanglement as measured by the concurrence versus the temperatures of the individual (taken equal $T = T_A = T_B$) and common baths. The qubit transition and interaction frequencies are $\epsilon_A = \epsilon_B = 1$ GHz and $\Omega = 0.7$ GHz. The bath coupling strengths are $\Gamma_A = \Gamma_B = \Gamma_C^{(A)} = \Gamma_C^{(B)} = 10$ MHz, and the pure dephasing rate is $\gamma = 3.5 \times 10^{-2}$ MHz. The red, dashed line indicates thermal equilibrium $T = T_C$. (b) The relative improvement in concurrence $\Delta\mathcal{C}$ for the same parameters as in (a). (c) Ratio of the maximal steady-state concurrence out of equilibrium $C^{\rm neq}$ to the maximum in equilibrium $C^{\rm eq}$ for the same energy gaps and pure dephasing rate as in (a). For the orange, dashed curve, the bath couplings are as in (a) while for the solid, blue curve $\Gamma_A = \Gamma_B = \Gamma_C^{(A)} = \Gamma_C^{(B)} = 0.1$ MHz.

experimentally accessible parameter settings, a pronounced improvement of steady-state entanglement is possible, and a signficant amount of entanglement can be generated.

Acknowledgments

We acknowledge helpful discussions with G Haack on implementations in superconducting systems. In this work ZXM and YJX are supported by National Natural Science Foundation (China) under Grant Nos. 11574178 and 61675115, Shandong Provincial Natural Science Foundation (China) under Grant No. ZR2016JL005, and Taishan Scholar Project of Shandong Province (China) under Grant No. tsqn201812059. AT and JBB acknowledge funding from the Swiss National Science Foundation starting Grant DIAQ, Grant No. 200021 169002.

Appendix: Sketch of master equation derivation

In this appendix, we sketch the derivation of the master equation (8) in the presence of a common bath. A detailed derivation can be found in [40] (see in particular appendix).

From equation (4), the interaction between the systems and the common reservoir is

$$\hat{H}_{SR}^{\rm com} = \sum_n [(g_{A,n}\hat{\sigma}_+^A + g_{B,n}\hat{\sigma}_+^B)\hat{c}_n + (g_{A,n}\hat{\sigma}_-^A + g_{B,n}\hat{\sigma}_-^B)\hat{c}_n^\dagger], \tag{A.1}$$

which can be reformulated in the interacting picture with respect to free Hamiltonians of the system and the common reservoir as

$$\begin{aligned}\hat{H}_{SR,I}^{\rm com}(t) = {} & [\tau_{12}T_{12}(t) + \tau_{34}T_{34}(t)]e^{i\omega_1 t} \\ & + [\tau_{13}T_{13}(t) + \tau_{24}T_{24}(t)]e^{i\omega_2 t} + \text{h.c.}\,,\end{aligned} \tag{A.2}$$

where $\tau_{ij} = |\lambda_i\rangle\langle\lambda_j|$ and the noise operators

$$\begin{aligned}T_{12}(t) &= \sin\frac{\theta}{2}A(t) + \cos\frac{\theta}{2}B(t),\\ T_{34}(t) &= -\sin\frac{\theta}{2}A(t) + \cos\frac{\theta}{2}B(t),\\ T_{13}(t) &= \cos\frac{\theta}{2}A(t) - \sin\frac{\theta}{2}B(t),\\ T_{24}(t) &= \cos\frac{\theta}{2}A(t) + \sin\frac{\theta}{2}B(t)\end{aligned} \tag{A.3}$$

with $A(t) = \sum_n g_{A,n}\hat{c}_n e^{-i\omega_{\hat{c},n}t}$ and $B(t) = \sum_n g_{B,n}\hat{c}_n e^{-i\omega_{\hat{c},n}t}$.

Under the standard Born–Markov approximation, we obtain the master equation for the systems as

$$\dot{\rho}_S = -\int_0^\infty dt' \,\text{Tr}_B[\hat{H}_{SR,I}^{\rm com}(t), [\hat{H}_{SR,I}^{\rm com}(t-t'), \rho_S(t)\otimes\rho_B]], \tag{A.4}$$

where ρ_B is the state of the common reservoir (which we will take to be a thermal state) and Tr_B denotes the trace over this reservoir. If we further adopt a rotating-wave approximation, we find

$$\begin{aligned}\dot{\rho}_S = \sum_{(i,j)}\Bigg[& \tau_{ij}\rho_S\tau_{ji}\int_0^\infty dt' e^{i(E_i-E_j)t'}\langle T_{ij}^\dagger(-t')T_{ij}(0)\rangle \\ & - \tau_{jj}\rho_S\int_0^\infty dt' e^{-i(E_i-E_j)t'}\langle T_{ij}^\dagger(0)T_{ij}(-t')\rangle \\ & + \tau_{ji}\rho_S\tau_{ij}\int_0^\infty dt' e^{-i(E_i-E_j)t'}\langle T_{ij}(-t')T_{ij}^\dagger(0)\rangle \\ & - \tau_{ii}\rho_S\int_0^\infty dt' e^{i(E_i-E_j)t'}\langle T_{ij}(0)T_{ij}^\dagger(-t')\rangle\Bigg] \\ & + \sum_{(ij,kl)}\Bigg[\tau_{ij}\rho_S\tau_{kl}\int_0^\infty dt' e^{i(E_i-E_j)t'}\langle T_{lk}^\dagger(-t')T_{ij}(0)\rangle \\ & + \tau_{lk}\rho_S\tau_{ji}\int_0^\infty dt' e^{i(E_i-E_j)t'}\langle T_{ij}^\dagger(-t')T_{lk}(0)\rangle\Bigg] + \text{h.c.},\end{aligned} \tag{A.5}$$

where the first sum runs over $(i, j) = (1, 2)$, $(1, 3)$, $(2, 3)$, and $(2, 4)$, while the second runs over $(ij, kl) = (12, 34)$, $(13, 42)$, $(31, 24)$, and $(43, 12)$. The bath correlation functions,

appearing under the integrals, are defined as $\langle X(t)Y(t')\rangle = \mathrm{Tr}_B[X(t)Y(t')\rho_B]$.

The real part of the integrals of the correlation functions determine the dissipation rates entering in the final master equation, while the imaginary parts contribute Lamb-type shifts to the Hamiltonian entering in the master equation. As we argue below, the latter are small and can be neglected, as we do in the main text. First, we give an example of the derivation of a dissipation rate.

Consider the correlation function $\langle T_{12}^{\dagger}(-t')T_{12}(0)\rangle$. We have

$$\begin{aligned}\int_0^{\infty} \mathrm{e}^{\mathrm{i}\omega_1 t'}\langle T_{12}^{\dagger}(-t')T_{12}(0)\rangle \mathrm{d}t' &= \sin^2(\theta/2)\int_0^{\infty} \mathrm{e}^{\mathrm{i}\omega_1 t'}\langle A^{\dagger}(-t')A(0)\rangle \mathrm{d}t' \\ &+ \cos^2(\theta/2)\int_0^{\infty} \mathrm{e}^{\mathrm{i}\omega_1 t'}\langle B^{\dagger}(-t')B(0)\rangle \mathrm{d}t' \\ &+ \frac{1}{2}\sin\theta\int_0^{\infty} \mathrm{e}^{\mathrm{i}\omega_1 t'}\langle A^{\dagger}(-t')B(0)\rangle \mathrm{d}t' \\ &+ \frac{1}{2}\sin\theta\int_0^{\infty} \mathrm{e}^{\mathrm{i}\omega_1 t'}\langle B^{\dagger}(-t')A(0)\rangle \mathrm{d}t'. \end{aligned} \tag{A.6}$$

As shown in [40], the real parts of the four integrals on the right-hand side of equation (A.6) can be obtained by using the formula

$$\int_0^{\infty} \mathrm{d}t' \mathrm{e}^{\pm \mathrm{i}\omega t'} = \pi\delta(\omega) \pm \mathrm{i}P\frac{1}{\omega}, \tag{A.7}$$

where P denotes the Cauchy principal value integral. For a reservoir in a thermal state, one obtains

$$\begin{aligned}\mathrm{Re}\left[\int_0^{\infty} \mathrm{e}^{\mathrm{i}\omega_1 t'}\langle A^{\dagger}(-t')A(0)\rangle \mathrm{d}t'\right] &= \Gamma_C^{(A)}(\omega_1)\bar{n}(\omega_1), \\ \mathrm{Re}\left[\int_0^{\infty} \mathrm{e}^{\mathrm{i}\omega_1 t'}\langle B^{\dagger}(-t')B(0)\rangle \mathrm{d}t'\right] &= \Gamma_C^{(B)}(\omega_1)\bar{n}(\omega_1), \\ \mathrm{Re}\left[\int_0^{\infty} \mathrm{e}^{\mathrm{i}\omega_1 t'}\langle A^{\dagger}(-t')B(0)\rangle \mathrm{d}t'\right] &= \Gamma_C^{(AB)}(\omega_1)\bar{n}(\omega_1), \\ \mathrm{Re}\left[\int_0^{\infty} \mathrm{e}^{\mathrm{i}\omega_1 t'}\langle B^{\dagger}(-t')A(0)\rangle \mathrm{d}t'\right] &= \Gamma_C^{(AB)}(\omega_1)\bar{n}(\omega_1). \end{aligned} \tag{A.8}$$

Here, the rates on the right-hand side fulfil $\Gamma_C^{(AB)}(\omega_1) = \sqrt{\Gamma_C^{(A)}(\omega_1)\Gamma_C^{(B)}(\omega_1)}$ and are determined by the reservoir density of states and the system-bath coupling coefficients $g_{A,n}$, $g_{B,n}$.

The imaginary part of the one-sided Fourier transform integrals is related to the real part by a Cauchy principal value integral of the form

$$f_I(\omega) = \int_0^{\infty} \frac{f_R(\omega')}{\omega - \omega'}\mathrm{d}\omega', \tag{A.9}$$

where f_R and f_I stand for the real and imaginary parts of integrals over the correlation functions as in the expressions above. In the master equation, they enter in the Hamiltonian part, as shifts of the system energy levels (i.e. Lamb shifts). Provided that the system-bath couplings (and hence the dissipation rates γ_k) are small, these imaginary parts will be small relative to the system energy splittings, and so can be neglected.

ORCID iDs

Zhong-Xiao Man https://orcid.org/0000-0003-1906-5923

References

[1] Nielsen M A and Chuang I L 2007 *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press)
[2] Giovannetti V, Lloyd S and Maccone L 2011 *Nat. Photon.* **5** 222
[3] Bennett C H, Brassard G, Popescu S, Schumacher B, Smolin J A and Wootters W K 1996 *Phys. Rev. Lett.* **76** 722
[4] Pan J W, Gasparoni S, Ursin R, Weihs G and Zeilinger A 2003 *Nature* **423** 417
[5] Kwiat P G, Barrazalopez S, Stefanov A and Gisin N 2001 *Nature* **409** 1014
[6] Shor P W 1995 *Phys. Rev.* A **52** R2493
[7] Steane A M 1996 *Phys. Rev. Lett.* **77** 793
[8] Mukhtar M, Saw T B, Soh W T and Gong J 2010 *Phys. Rev.* A **81** 012331
[9] Wang Z Y and Liu R B 2011 *Phys. Rev.* A **83** 022306
[10] Lo Franco R, D'Arrigo A, Falci G, Compagno G and Paladino E 2014 *Phys. Rev.* B **90** 054304
[11] Maniscalco S, Francica F, Zaffino R L, Lo Gullo N and Plastina F 2008 *Phys. Rev. Lett.* **100** 090503
[12] An N B, Kim J and Kim K 2010 *Phys. Rev.* A **82** 032316
[13] Sun Q, Al-Amri M, Davidovich L and Zubairy M S 2010 *Phys. Rev.* A **82** 052323
[14] Kim Y S, Lee J C, Kwon O and Kim Y H 2011 *Nat. Phys.* **8** 117
[15] Man Z X, Xia Y J and An N B 2012 *Phys. Rev.* A **86** 012325
[16] Man Z X, Xia Y J and An N B 2012 *Phys. Rev.* A **86** 052322
[17] Plenio M B, Huelga S F, Beige A and Knight P L 1999 *Phys. Rev.* A **59** 2468
[18] Bellomo B, Lo Franco R, Maniscalco S and Compagno G 2008 *Phys. Rev.* A **78** 060302
[19] Diehl S, Micheli A, Kantian A, Kraus B, Buchler H P and Zoller P 2008 *Nat. Phys.* **4** 878
[20] Verstraete F, Wolf M M and Cirac J I 2009 *Nat. Phys.* **5** 633
[21] Kraus B, Büchler H P, Diehl S, Kantian A, Micheli A and Zoller P 2008 *Phys. Rev.* A **78** 042307
[22] Ticozzi F and Viola L 2014 *Quantum Inf. Comput.* **14** 265
[23] Plenio M B and Huelga S F 2002 *Phys. Rev. Lett.* **88** 197901
[24] Schneider S and Milburn G J 2002 *Phys. Rev.* A **65** 042107
[25] Kastoryano M J, Reiter F and Sørensen A S 2011 *Phys. Rev. Lett.* **106** 090502
[26] Reiter F, Tornberg L, Johansson G and Sørensen A S 2013 *Phys. Rev.* A **88** 032317
[27] Schuetz M J A, Kessler E M, Vandersypen L M K, Cirac J I and Giedke G 2013 *Phys. Rev. Lett.* **111** 246802
[28] Cai J M, Popescu S and Briegel H J 2010 *Phys. Rev.* E **82** 021921
[29] Walter S, Budich J C, Eisert J and Trauzettel B 2013 *Phys. Rev.* B **88** 035441
[30] Krauter H, Muschik C A, Jensen K, Wasilewski W, Petersen J M, Cirac J I and Polzik E S 2011 *Phys. Rev. Lett.* **107** 080503
[31] Barreiro J T, Muller M, Schindler P, Nigg D, Monz T, Chwalla M, Hennrich M, Roos C, Zoller P and Blatt R 2011 *Nature* **470** 486
[32] Lin Y J, Gaebler J P, Reiter F, Tan T R, Bowler R S, Sorensen A S, Leibfried D and Wineland D J 2013 *Nature* **504** 415
[33] Shankar S, Hatridge M, Leghtas Z, Sliwa K, Narla A, Vool U, Girvin S M, Frunzio L, Mirrahimi M and Devoret M H 2013 *Nature* **504** 419

[34] Arnesen M C, Bose S and Vedral V 2001 *Phys. Rev. Lett.* **87** 017901
[35] Wang X 2001 *Phys. Rev.* A **64** 012313
[36] Wang X 2001 *Phys. Lett.* A **281** 101
[37] Gunlycke D, Kendon V M, Vedral V and Bose S 2001 *Phys. Rev.* A **64** 042302
[38] Lagmago Kamta G and Starace A F 2002 *Phys. Rev. Lett.* **88** 107901
[39] Canosa N and Rossignoli R 2006 *Phys. Rev.* A **73** 022347
[40] Liao J Q, Huang J F and Kuang L M 2011 *Phys. Rev.* A **83** 052110
[41] Eisler V and Zimborás Z 2005 *Phys. Rev.* A **71** 042318
[42] Quiroga L, Rodríguez F J, Ramírez M E and París R 2007 *Phys. Rev.* A **75** 032308
[43] Sinaysky I, Petruccione F and Burgarth D 2008 *Phys. Rev.* A **78** 062301
[44] Huang X L, Guo J L and Yi X X 2009 *Phys. Rev.* A **80** 054301
[45] Pumulo N, Sinayskiy I and Petruccione F 2011 *Phys. Lett.* A **375** 3157
[46] Brunner N, Huber M, Linden N, Popescu S, Silva R and Skrzypczyk P 2014 *Phys. Rev.* E **89** 032115
[47] Brask J B and Brunner N 2015 *Phys. Rev.* E **92** 062101
[48] Brask J B, Haack G, Brunner N and Huber M 2015 *New J. Phys.* **17** 113029
[49] Tavakoli A, Haack G, Huber M, Brunner N and Brask J B 2018 *Quantum* **2** 73
[50] Kim M S, Lee J, Ahn D and Knight P L 2002 *Phys. Rev.* A **65** 040101
[51] Braun D 2002 *Phys. Rev. Lett.* **89** 277901
[52] Benatti F, Floreanini R and Piani M 2003 *Phys. Rev. Lett.* **91** 070402
[53] Bellomo B and Antezza M 2013 *Europhys. Lett.* **104** 10006
[54] Bellomo B and Antezza M 2015 *Phys. Rev.* A **91** 042124
[55] Bellomo B and Antezza M 2013 *New J. Phys.* **15** 113052
[56] Hofer P P, Perarnau-Llobet M, Miranda L D M, Haack G, Silva R, Brask J B and Brunner N 2017 *New J. Phys.* **19** 123037
[57] Breuer H P and Petruccione F 2002 *The Theory of Open Quantum Systems* (Oxford: Oxford University Press)
[58] Rivas A, Plato A D K, Huelga S F and Plenio M B 2010 *New J. Phys.* **12** 113032
[59] Schaller G 2015 *Lecture notes: Non-Equilibrium Master equations* (Berlin: Technische Universität Berlin)
[60] Higgins K D B, Benjamin S C, Stace T, Milburn G, Lovett B and Gauger E 2014 *Nat. Commun.* **5** 4705
[61] Kołodyński J, Brask J B, Perarnau-Llobet M and Bylicka B 2018 *Phys. Rev.* A **97** 062124
[62] Wootters W K 1998 *Phys. Rev. Lett.* **80** 2245
[63] Yu T and Eberly J H 2007 *Quantum Inf. Comput.* **7** 459
[64] Wichterich H, Henrich M J, Breuer H P, Gemmer J and Michel M 2007 *Phys. Rev.* E **76** 031115
[65] Nieuwenhuizen T M and Allahverdyan A E 2002 *Phys. Rev.* E **66** 036102
[66] Quan H T, Zhang P and Sun C P 2005 *Phys. Rev.* E **72** 056110
[67] Quan H T, Liu Y X, Sun C P and Nori F 2007 *Phys. Rev.* E **76** 031105
[68] Quan H T, Wang Y D, Liu Y X, Sun C P and Nori F 2006 *Phys. Rev. Lett.* **97** 180402
[69] Cottet N, Jezouin S, Bretheau L, Campagne-Ibarcq P, Ficheux Q, Anders J, Auffèves A, Azouit R, Rouchon P and Huard B 2017 *Proc. Natl Acad. Sci.* **114** 7561
[70] Koski J V, Sagawa T, Saira O P, Yoon Y, Kutvonen A, Solinas P, Mottonen M, Ala-Nissila T and Pekola J P 2013 *Nat. Phys.* **9** 644
[71] Koski J V, Maisi V F, Pekola J P and Averin D V 2014 *Proc. Natl Acad. Sci.* **111** 13786
[72] Pekola J P 2015 *Nat. Phys.* **11** 118
[73] Chen Y X and Li S W 2012 *Europhys. Lett.* **97** 40003
[74] Hofer P P, Souquet J R and Clerk A A 2016 *Phys. Rev.* B **93** 041418
[75] Hofer P P, Perarnau-Llobet M, Brask J B, Silva R, Huber M and Brunner N 2016 *Phys. Rev.* B **94** 235420
[76] Manucharyan V E 2012 Superinductance *PhD Thesis* (New Haven, CT: Yale University)
[77] Majer J *et al* 2007 *Nature* **449** 443
[78] Sillanpää M A, Park J I and Simmonds R W 2007 *Nature* **449** 438
[79] DiCarlo L *et al* 2009 *Nature* **460** 240
[80] Chen Y *et al* 2014 *Phys. Rev. Lett.* **113** 220502
[81] Filipp S, Göppl M, Fink J M, Baur M, Bianchetti R, Steffen L and Wallraff A 2011 *Phys. Rev.* A **83** 063827
[82] Kim M D 2015 *Quantum Inf. Process.* **14** 3677
[83] O'Malley P J J 2016 Superconducting qubits: dephasing and quantum chemistry *PhD Thesis* (Santa Barbara, CA: University of California, Santa Barbara)
[84] Kou A, Smith W C, Vool U, Pop I M, Sliwa K M, Hatridge M H, Frunzio L and Devoret M H 2017 *Phys. Rev.* Applied **9** 064022

Enabling Computation of Correlation Bounds for Finite-Dimensional Quantum Systems via Symmetrization

Armin Tavakoli,[1,*] Denis Rosset,[2,*] and Marc-Olivier Renou[1]
[1]Department of Applied Physics, University of Geneva, 1211 Geneva, Switzerland
[2]Perimeter Institute for Theoretical Physics, Waterloo, Ontario, Canada, N2L 2Y5

(Received 29 August 2018; revised manuscript received 11 December 2018; published 20 February 2019)

We present a technique for reducing the computational requirements by several orders of magnitude in the evaluation of semidefinite relaxations for bounding the set of quantum correlations arising from finite-dimensional Hilbert spaces. The technique, which we make publicly available through a user-friendly software package, relies on the exploitation of symmetries present in the optimization problem to reduce the number of variables and the block sizes in semidefinite relaxations. It is widely applicable in problems encountered in quantum information theory and enables computations that were previously too demanding. We demonstrate its advantages and general applicability in several physical problems. In particular, we use it to robustly certify the nonprojectiveness of high-dimensional measurements in a black-box scenario based on self-tests of d-dimensional symmetric informationally complete positive-operator-valued measurements.

DOI: 10.1103/PhysRevLett.122.070501

Introduction.—Finite-dimensional quantum systems are common in quantum information theory. They are standard in the broad scope of quantum communication complexity problems (CCPs) [1] in which quantum correlations are studied under limited communication resources. Furthermore, they are widely used in semi-device-independent quantum information protocols [2] in which systems are fully uncharacterized up to their Hilbert space dimension. Also, studying correlations obtainable from finite-dimensional systems is critical for device-independent dimension witnessing [3,4].

In view of their diverse relevance, it is important to bound quantum correlations arising from dimension-bounded Hilbert spaces. To this end, semidefinite programs (SDPs) [5] constitute a powerful tool. Lower bounds on quantum correlations are straightforwardly obtained using alternating convex searchers (SDPs in see-saw) [6,7]. However, obtaining upper bounds valid for *any* quantum states and measurements is more demanding. A powerful approach to this problem is to relax some well-chosen constraints of quantum theory so that the resulting super-quantum correlations easily can be computed with SDPs, thus returning upper bounds on quantum correlations. Such approaches are commonplace in various problems in quantum information theory [8–10]. A hierarchy of semidefinite relaxations for upper-bounding quantum correlations on dimension-bounded Hilbert spaces was introduced by Navascués and Vértesi (NV) [10,11]. This is an effective tool for problems involving a small number of states and measurements, and low Hilbert space dimensions. However beyond simple scenarios, the computational requirements of evaluating the relaxations quickly become too demanding.

It is increasingly relevant to overcome the practical limitations of the NV hierarchy, i.e., to provide efficient computational tools for bounding quantum correlations in problems beyond small sizes and low Hilbert space dimensions. This is motivated by both theoretical and experimental advances. Dimension witnessing has been experimentally realized far beyond the lowest Hilbert space dimensions [12,13]. Furthermore, increasing the dimension can activate unexpectedly strong quantum correlations [14], a phenomenon that has been experimentally demonstrated [15]. Also, quantum correlations obtained from a sizable number of states and measurements are interesting for studying mutually unbiased bases [16]. Moreover, large problem sizes naturally appear in multipartite CCPs involving single particles [17–19]. Similarly sized problems also appear in multipartite CCPs for the characterization of entangled states and measurements [20]. In addition, efficiently evaluating the NV hierarchy many times can improve randomness extraction from experimental data [21].

In this work we develop techniques for efficiently bounding quantum correlations under dimension constraints. The technique is powered by the exploitation of *symmetries*, i.e., relabelings of optimization variables that leave a figure of merit invariant. The use of symmetries for reducing the complexity of SDPs was first introduced in Ref. [22] and was shown to lead to remarkable efficiency gains. These efficiency gains have also been harvested in several specific quantum information problems relying on SDPs. These include finding bounds on classical [23] and quantum [24,25] Bell correlations, quantifying entanglement [9,26], and finding symmetric Bell inequalities [27].

Note that symmetries in Bell scenarios also have been studied without application to SDPs [28–31]. In dimension-bounded scenarios, symmetries have been considered for CCPs tailored for studying the existence of mutually unbiased bases [16].

We describe a powerful, generally applicable, and easy-to-use technique for symmetrized semidefinite relaxations for dimension-bounded quantum correlations. We show how to automatize searches for symmetries in general Bell scenarios and CCPs, and how these can be exploited to reduce computational requirements in all parts of the NV hierarchy. This amounts to reducing the number of variables in an optimization, and reducing block sizes beyond previous approaches. We make these techniques readily available via a user-friendly software package supporting general correlation scenarios. Subsequently, we give examples of problems that can be solved faster (several orders of magnitude), and other previously unattainable problems that can now be computed. We focus on the usefulness of symmetrization for the problem of certifying that an uncharacterized device implements a nonprojective measurement using only the observed correlations. To this end, we introduce a family of CCPs, prove that they enable self-tests of d-dimensional symmetric informationally complete (SIC) positive-operator-valued measurements (POVMs), then use symmetrized semidefinite relaxations to bound the correlations attainable under projective measurements. This allows us to go beyond previously studied qubit systems [32–36] and robustly certify the nonprojectiveness of SIC-POVMs subject to imperfections.

Bounding finite-dimensional quantum correlations.—We begin by summarizing the NV hierarchy [10,11] for optimizing dimensionally constrained quantum correlations. For simplicity, we first describe CCPs, and later consider Bell scenarios.

Consider a CCP in which a party, Alice, holds a random input x and another party, Bob, holds a random input y. Alice encodes her input into a quantum state ρ_x of dimension d and sends it to Bob. Bob performs a measurement $\{M_y^b\}_b$ with outcome b. The resulting probability distribution is used to evaluate a functional $F(P)=\sum_{x,y,b} c_{x,y}^b P(b|x,y)$, where $c_{x,y}^b$ are real coefficients. The problem of interest is to compute the maximal quantum value of F when the probabilities are given by the Born rule $P(b|x,y)=\mathrm{tr}(\rho_x M_y^b)$, where the measurement operators are taken to be projectors. The NV hierarchy presents the following semidefinite relaxations. Sample a random set of states and measurements $\{\rho_x\}$ and $\{M_y^b\}$ of dimension d, which we collect in the set of operator variables $\{X_i\}$. Then, generate all strings, $\{s_j(X)\}_j$, of products of at most L of these operators. The choice of L determines the degree of relaxation, i.e., the level of the hierarchy. Construct a moment matrix

$$\Gamma_{j,k}=\langle s_j(X)^\dagger s_k(X)\rangle, \tag{1}$$

where, for the present CCP, the expectation value of an operator product S is $\langle S\rangle=\mathrm{tr}S$. Repeat this process many times, each time obtaining a new moment matrix. Terminate the process when the sampled moment matrix is linearly dependent on the collection of those previously generated. Hence, $\{\Gamma^{(1)},\ldots,\Gamma^{(m)}\}$ identifies a basis for the feasible affine subspace $\mathcal{F}$ of such matrices under the given dimensional constraint. The semidefinite relaxation amounts to finding an affine combination $\Gamma=\sum_{\ell=1}^m c_\ell\Gamma^{(\ell)}\in\mathcal{F}$, with $\Gamma\geq 0$, that maximizes the functional F (which can be expressed as a linear combination of entries of Γ). Hence, the relaxation reads

$$\max_{\vec{c}\in\mathbb{R}^m} F(\Gamma) \quad \text{s.t.} \quad \Gamma\geq 0, \quad \sum_{\ell=1}^m c_\ell = 1. \tag{2}$$

In summary, the problem consists in first sampling a basis enforcing the dimensional constraint and then evaluating a SDP. Crucially, the complexity of solving the SDP hinges on the number of basis elements, m, needed to complete the basis and the size of the final SDP matrix, n. For a single iteration of primal-dual interior point solvers, the required memory scales as $\mathcal{O}(m^2+mn^2)$ while the CPU time scales as $\mathcal{O}(m^3+n^3+mn^3+m^2n^2)$ [37]. Without exploitation of the problem structure, medium-sized physical scenarios, as well as small-sized scenarios with high relaxation degree, practically remain out of reach for current desktop computers. We have performed all computations using a machine of 32 GB RAM and i5-6500 3.2 GHz CP.

Symmetric relaxations.—The key to reducing the computational requirements for the NV hierarchy is twofold: First reducing the number of elements needed to form the basis in the sampling step, i.e., decreasing the dimension of $\mathcal{F}$, and then shrinking the size of the positivity constraints in the subsequent SDP by block-diagonalizing Γ. Here, we show how such a reduction can be systematically achieved by identifying and exploiting the set of symmetries of the problem.

Recall that $\{X_i\}$ collects all the operators (states, measurements etc.) present in the formulation of the problem, where $i\in\mathcal{I}$ is an index. Consider a permutation of elements of $\mathcal{I}$, i.e., a bijective function $\pi\colon\mathcal{I}\to\mathcal{I}$. We write $\pi(X_i)=X_{\pi(i)}$ and define the action of the permutation on the strings $s=X_iX_j\ldots$ of products of operators X_i appearing in the NV hierarchy as $\pi(X_iX_j\ldots)=X_{\pi(i)}X_{\pi(j)}\ldots$. We call π an *ambient symmetry* if it is a transformation of the scenario which preserves its structure, as expressed by implicit or explicit constraints on the operators $\{X_i\}$. The set of those symmetries form the *ambient group* $\mathcal{A}=\{\pi\}$. In the Supplemental Material [38] (SM, including Refs. [39–54]), we describe the ambient groups for general Bell scenarios and CCPs. Given a moment matrix Γ and $\pi\in\mathcal{A}$, we consider the relabeled

matrix $\pi(\Gamma)$ where $(\pi(\Gamma))_{j,k} = \Gamma_{\pi^{-1}(j),\pi^{-1}(k)}$, according to the convention of Eq. (1). By construction, π preserves the constraints of the problem: for a feasible moment matrix $\Gamma \in \mathcal{F}$ we have $\pi(\Gamma) \in \mathcal{F}$ for any $\pi \in \mathcal{A}$. Moreover, the feasible set $\mathcal{F}$ is convex, so any convex combination of those $\pi(\Gamma)$ is feasible as well.

However, not all elements of $\mathcal{A}$ leave the objective $F(\Gamma)$ invariant. We write $\mathcal{G} = \{\pi \in \mathcal{A} : F(\pi(\Gamma)) = F(\Gamma)\}$ the *symmetry group* of the optimization problem. One can straightforwardly find the elements of $\mathcal{G}$ by enumerating the elements of $\mathcal{A}$ and filtering those that leave $F(\pi(\Gamma)) = F(\Gamma)$ invariant. Then, following a standard procedure [16,22,24,27] we can average any optimal solution Γ under the Reynolds operator, defined as

$$\Gamma' \equiv \mathcal{R}(\Gamma) = \frac{1}{|\mathcal{G}|}\sum_{\pi\in\mathcal{G}}\pi(\Gamma), \tag{3}$$

where $|\mathcal{G}|$ is the size of $\mathcal{G}$ and obtain an optimal solution of the problem, which now satisfies $\pi(\Gamma') = \Gamma'$ for all $\pi \in \mathcal{G}$. Since the set Γ' is characterized by the relation $\mathcal{R}(\Gamma') = \Gamma'$, instead of searching the optimal Γ in the full feasible set, it is sufficient to only consider the symmetric subspace $\mathcal{R}(\mathcal{F})$ given by the image of the feasible set under $\mathcal{R}$. As discussed above, the basis of $\mathcal{F}$ is found by sampling. To sample $\mathcal{R}(\mathcal{F})$ instead, we simply apply $\mathcal{R}$ on each sample during the construction of the basis, thus obtaining $\{\Gamma'^{(1)}, \ldots, \Gamma'^{(m')}\}$. As a result, the size of the basis, m', decreases due to the smaller dimension of $\mathcal{R}(\mathcal{F})$. In the SM, we discuss methods for speeding up the computation of $\mathcal{R}$.

Moreover, a second major reduction is obtained: As the symmetrized moment matrices Γ' commute with a representation of the group $\mathcal{G}$, there exists [22] a unitary matrix that block diagonalizes the moment matrix. This reduces the size of the positivity constraint on the final SDP matrix. A complete symmetry exploitation is obtained when the decomposition of the representation of $\mathcal{G}$ into irreducible components with multiplicities is known. We achieve this via an efficient general block diagonalization method detailed in the SM. Moreover, we make available a user-friendly MATLAB package [55] for symmetrization of semidefinite relaxations in the NV hierarchy applicable to general correlation scenarios encountered in quantum information. The package automates both a search for the symmetries of a problem (if these are unknown) and the construction of symmetry-adapted relaxation.

Robust certification of nonprojective measurements based on SIC-POVMs.—We now exemplify the usefulness of symmetrization in a physical application. We certify, solely from observed data, that an uncharacterized device ("black-box") implements a nonprojective measurement. Nonprojective measurements have diverse applications in quantum theory [32,56–62]. This has motivated interest in their black-box certification [32–36]. Using semidefinite relaxations (whose complexity scales quickly with dimension) as a primary tool, these works limit themselves to qubits. We use symmetrization to overcome this limitation and certify the nonprojectiveness of higher-dimensional measurements of physical interest. Since such certificates are typically only useful for nonprojective measurements that are close (e.g., in fidelity) to a particular targeted nonprojective measurement (corresponding to the optimal quantum correlations) [33], it is important to ensure that the targeted measurement is well motivated.

One of the most celebrated nonprojective measurements is the SIC-POVM. These are sets of d^2 subnormalized rank-one projectors $\{(1/d)|\psi_x\rangle\langle\psi_x|\}_{x=1}^{d^2}$ with $|\langle\psi_x|\psi_{x'}\rangle|^2 = 1/(d+1)$ when $x \neq x'$. Higher-dimensional SIC-POVMs have been of substantial interest for both fundamental (see, e.g., Ref. [63] for a review) and practical considerations [64–68] in quantum information theory. We introduce a family of CCPs and prove that optimal quantum correlations imply a d-dimensional SIC-POVM. However, due to unavoidable experimental imperfections, such optimal correlations will never occur in practice. Therefore, we use symmetrization to certify the nonprojectiveness of measurements close to SIC-POVMs, that achieve nearly optimal correlations. Moreover, as noted in Ref. [33], the dimension-bounded scenario is well-suited for black-box studies of nonprojective measurements since said property is only well-defined on Hilbert spaces of fixed dimension.

Consider a CCP in which Alice encodes her input x into a d-dimensional system sent to Bob, who associates his input y to a measurement producing an outcome b. A general witness can be written

$$W = \sum_{x,y,b}\alpha_{xyb}P(b|x,y), \tag{4}$$

where α_{xyb} are real coefficients. By tuning the coefficients, one can construct CCPs in which the optimal correlations W^Q are uniquely realized with a particular nonprojective measurement. This is known as a self-test [69]. Consequently, there must exist some $W^P < W^Q$ which bounds the correlations under all projective measurements. Thus, observing $W > W^P$ certifies that Bob implements a nonprojective measurement.

We construct a family of CCPs (inspired by Refs. [33,70]) tailored to self-test d-dimensional SIC-POVMs. Alice and Bob each receive inputs $x \in [N]$ and $(y, y') \in [N]$ with $y < y'$, respectively, for some $N > d$ and $[N] = \{1, \ldots, N\}$. Bob outputs $b \in \{0, 1\}$. Bob also possesses another measurement setting labeled $povm$ which returns an outcome $o \in [N]$. The witness of interest is

$$W_d = \sum_{x<x'} P(b=0|x,(x,x')) + P(b=1|x',(x,x')) + \sum_{x=1}^{N} P(o=x|x,povm). \tag{5}$$

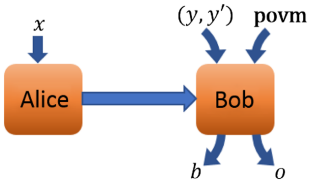

FIG. 1. Illustration of the CCP [Eq. (5)]. Bob has $\binom{N}{2}$ settings labeled by (y, y') and one additional setting labeled $povm$. Alice and Bob aim to satisfy the following relations: $o = x$ for the setting $povm$, and $b = 0$ when $x = y$ and $b = 1$ when $x = y'$, respectively, for the settings (y, y').

The scenario is illustrated in Figure 1.

Theorem 1: For $N = d^2$, the maximal quantum value of the witness is

$$W_d^Q = \frac{1}{2}\sqrt{d^5(d-1)^2(d+1)} + \binom{d^2}{2} + d. \quad (6)$$

This value self-tests that Alice prepares a SIC ensemble and that Bob's setting $povm$ corresponds to a SIC-POVM (Note that SIC-POVMs are not proven to exist in all dimensions).

The proof is given in the SM. To enable the certification of a nonprojective measurement producing nearly optimal correlations, we must obtain a bound W_d^P on W_d respected by all projective measurements. To this end, we use symmetrized semidefinite relaxations.

The symmetries of the witness [Eq. (5)] correspond to coordinated permutations of the inputs of Alice and inputs and outputs of Bob. We permute x among its N possible values. This requires us to compensate the permutation by also applying it to o. Furthermore, to preserve the probabilities appearing in the first summand of Eq. (5), we must apply a permutation to the indices (y, y') and the outcome b. Moreover, since we are interested in bounding W_d under projective measurements, said property must be explicitly imposed on Bob's setting $povm$. This means that at most d of the POVM elements $\{M_{povm}^x\}_{x=1}^{d^2}$ are nonzero, corresponding to rank-one projectors. This must be accounted for in the symmetries of the problem. In the SM we discuss the symmetries in detail.

TABLE I. Upper bounds (UBs) and lower bounds (LBs) on quantum correlations under projective measurements with $N = d^2$. The lower bounds are obtained via SDPs in an alternate convex search and the upper bounds via symmetrized semidefinite relaxations.

d	2	3	4	5	6
LB: W_d^P	12.8484	70.0961	231.2685	578.7002	1219.0129
UB: W_d^P	12.8484	70.1133	231.2685	578.7987	1219.2041
W_d^Q	12.8990	70.1769	231.3313	578.8613	1219.2667

Using the general recipe, we have implemented the symmetrized NV hierarchy. We use the relaxation degree corresponding to monomials $\{\mathbb{1}, \rho, M, M_{povm}, \rho\rho\}$ and also all the monomials $\rho_x M_{(x,x')}^b$ appearing in the first summand of Eq. (5). In Table I we present the upper bounds W_d^P. We have also obtained lower bounds for W_d under projective measurements by considering SDPs in an alternate convex search, enforcing only d nonzero elements of trace one. These lower bounds were verified to be achieved with projective measurements up to machine precision. The results show that the obtained upper bounds are either optimal or close to optimal, depending on d. In analogy with previous works [32–36], we find that the gap between optimal quantum correlations and those obtained under projective measurements is small.

Consider the role of symmetrization in obtaining the above results. In Table II we present the number of samples needed to complete the basis in the NV hierarchy, the size of the final SDP matrix, and the time required to evaluate the SDPs. We compare these parameters for a

TABLE II. Comparison between computational parameters for the task of bounding W_d under projective measurements using a standard implementation, symmetrization to reduce the number of samples [using only Eq. (3)], and symmetrization to also perform block diagonalization (BD). The notation $D[a, b]$ means that there are D blocks with the smallest being of size a and the largest of size b.

	d	2	3	4	5	6
Non-sym	No. of samples	221	>12 000	⋯	⋯	⋯
	Block sizes	1[43]	1[229]	1[741]	1[1831]	1[3823]
	SDP [s]	2.0	⋯	⋯	⋯	⋯
Sym no BD	No. of samples	65	134		137	
	Block sizes	1[43]	1[229]	1[741]	1[1831]	1[3823]
	SDP [s]	0.5	19	500	⋯	⋯
Sym + BD	No. of samples	65	134		137	
	Block sizes	4[6,16]	7[3,16]		8[3,16]	
	SDP [s]	0.3	0.6		1.2	

standard implementation, a symmetrized implementation only reducing the number of samples, and a the full symmetrization developed to also exploit block diagonalization of the SDP matrix. Without symmetries, we are unable to go beyond qubit systems ($d = 2$), since already for $d = 3$ we have over 12 000 samples. Interestingly, this rapid increase in complexity can be completely overcome via symmetrization: The number of samples becomes constant when $d = 4$, 5, 6. In addition, the size of the SDP matrix increases polynomially in d, causing symmetrization that only addresses the number of samples to still be too demanding already when $d > 4$. However, using the block-diagonalization methods detailed in the SM, we can reduce the size of the SDP matrix to be constant for $d = 4$, 5, 6. This allows us to straightforwardly solve the semidefinite relaxations in less than two seconds.

Further applications.—The general symmetrization technique applies to many problems in quantum information theory. In the SM, we consider four different examples. For each, we demonstrate the remarkable computational advantages of symmetrization, both in terms of reducing the number of basis elements and in terms of block diagonalization. This enables us to obtain improved bounds on previously studied physical quantities. The problems we consider are (high-dimensional and many-input) random access codes [71,72], I_{3322}-like Bell inequalities [11,73], a sequential communication in multipartite CCPs (in the spirit of Refs. [17,18]), and CCPs exhibiting dimensional discontinuities [14,15]. In the latter, we also exemplify the advantages in automatizing the search for the symmetries in problems in which these are not easily spotted by inspection.

Moreover, we previously observed that the complexity of the evaluation for bounding W_d^P can be reduced to be constant for $d = 4$, 5, 6 via symmetries. This suggests that similar reductions may occur for other CCPs as well. In the SM we have focused on the CCPs known as random access codes and proven that symmetries enable us to evaluate the NV hierarchy with constant complexity for any Hilbert space dimension. In this sense, the computational advantages over standard implementations, as well as over symmetrization that does not utilize block diagonalization, increase with d.

Conclusions.—We presented a technique for efficiently evaluating semidefinite relaxations of finite-dimensional quantum correlations using symmetries present in the problem. We applied it to robustly certify higher-dimensional nonprojective measurements by considering CCPs that self-test d-dimensional SIC-POVMs. The scheme could be implemented in photonics experiments using, e.g., encodings in path [64,68], path and polarization [65], and orbital angular momentum [66,67]. Measuring a value of W_d above the upper bounds (UBs) stated in Table I completes the certification. A broadly relevant open problem in this topic [32–36] is making the certification more tolerant to experimental imperfections (i.e., larger gaps between W_d^P (UB) and W_d^Q in Table I).

We conclude with two more open problems. Can the sampling approach be adapted to semidefinite relaxations in Bell inequalities without dimensional bounds? How does the symmetrization technique adapt to physical problems that do not concern quantum resources; e.g., cardinality of hidden variables [74] and the dimension of post-quantum resources?

We are thankful for useful discussions with Jean-Daniel Bancal. This work was supported by the Swiss National Science Foundation (Starting Grant DIAQ, NCCR-QSIT). Research at Perimeter Institute is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Research and Innovation. This publication was made possible through the support of a grant from the John Templeton Foundation.

Note added.—Recently, we became aware of a work-in-preparation by E. Aguilar and P. Mironowicz to generalize the results of Ref. [16].

*A. T. and D. R. contributed equally for this project.

[1] H. Buhrman, R. Cleve, S. Massar, and R. de Wolf, Nonlocality and communication complexity, Rev. Mod. Phys. **82**, 665 (2010).
[2] M. Pawłowski and N. Brunner, Semi-device-independent security of one-way quantum key distribution, Phys. Rev. A **84**, 010302(R) (2011).
[3] R. Gallego, N. Brunner, C. Hadley, and A. Acín, Device-Independent Tests of Classical and Quantum Dimensions, Phys. Rev. Lett. **105**, 230501 (2010).
[4] N. Brunner, S. Pironio, A. Acín, N. Gisin, A. A. Méthot, and V. Scarani, Testing the Dimension of Hilbert Spaces, Phys. Rev. Lett. **100**, 210503 (2008).
[5] L. Vandenberghe and S. Boyd, Semidefinite programming, SIAM Rev. **38**, 49 (1996).
[6] R. E. Wendell and A. P. Hurter, Jr., Minimization of a nonseparable objective function subject to disjoint constraints, Oper. Res. **24**, 643 (1976).
[7] K. F. Pál and T. Vértesi, Maximal violation of a bipartite three-setting, two-outcome Bell inequality using infinite-dimensional quantum systems, Phys. Rev. A **82**, 022116 (2010).
[8] M. Navascués, S. Pironio, and A. Acín, Bounding the Set of Quantum Correlations, Phys. Rev. Lett. **98**, 010401 (2007).
[9] T. Moroder, J.-D. Bancal, Y.-C. Liang, M. Hofmann, and O. Gühne, Device-Independent Entanglement Quantification and Related Applications, Phys. Rev. Lett. **111**, 030501 (2013).
[10] M. Navascués and T. Vértesi, Bounding the Set of Finite Dimensional Quantum Correlations, Phys. Rev. Lett. **115**, 020501 (2015).

[11] M. Navascués, A. Feix, M. Araújo, and A. Vértesi, Characterizing finite-dimensional quantum behavior, Phys. Rev. A **92**, 042117 (2015).

[12] V. D'Ambrosio, F. Bisesto, F. Sciarrino, J. F. Barra, G. Lima, and A. Cabello, Device-Independent Certification of High-Dimensional Quantum Systems, Phys. Rev. Lett. **112**, 140503 (2014).

[13] E. A. Aguilar, M. Farkas, D. Martínez, M. Alvarado, J. Cariñe, G. B. Xavier, J. F. Barra, G. Cañas, M. Pawłowski, and G. Lima, Certifying an Irreducible 1024-Dimensional Photonic State Using Refined Dimension Witnesses, Phys. Rev. Lett. **120**, 230503 (2018).

[14] A. Tavakoli, M. Pawłowski, M. Żukowski, and M. Bourennane, Dimensional discontinuity in quantum communication complexity at dimension seven, Phys. Rev. A **95**, 020302(R) (2017).

[15] D. Martínez, A. Tavakoli, M. Casanova, G. Cañas, B. Marques, and G. Lima, High-Dimensional Quantum Communication Complexity beyond Strategies Based on Bells Theorem, Phys. Rev. Lett. **121**, 150504 (2018).

[16] E. A. Aguilar, J. J. Borkała, P. Mironowicz, and M. Pawłowski, Connections Between Mutually Unbiased Bases and Quantum Random Access Codes, Phys. Rev. Lett. **121**, 050501 (2018).

[17] E. F. Galvão, Feasible quantum communication complexity protocol, Phys. Rev. A **65**, 012318 (2001).

[18] P. Trojek, C. Schmid, M. Bourennane, C. Brukner, M. Żukowski, and H. Weinfurter, Experimental quantum communication complexity, Phys. Rev. A **72**, 050305(R) (2005).

[19] M. Smania, A. M. Elhassan, A. Tavakoli, and M. Bourennane, Experimental quantum multiparty communication protocols, npj Quantum Inf. **2**, 16010 (2016).

[20] A. Tavakoli, A. A. Abbott, M.-O. Renou, N. Gisin, and N. Brunner, Semi-device-independent characterization of multipartite entanglement of states and measurements, Phys. Rev. A **98**, 052333 (2018).

[21] P. Mironowicz, A. Tavakoli, A. Hameedi, B. Marques, P. Pawłowski, and M. Bourennane, Increased certification of semi-device independent random numbers using many inputs and more postprocessing, New J. Phys. **18**, 065004 (2016).

[22] K. Gatermann and P. A. Parrilo, Symmetry groups, semidefinite programs, and sums of squares, J. Pure Appl. Algebra **192**, 95 (2004).

[23] M. Fadel and J. Tura, Bounding the Set of Classical Correlations of a Many-Body System, Phys. Rev. Lett. **119**, 230402 (2017).

[24] D. Rosset, Characterization of correlations in quantum networks, Ph.D. thesis, 2015.

[25] C. Bamps and S. Pironio, Sum-of-squares decompositions for a family of Clauser-Horne-Shimony-Holt-like inequalities and their application to self-testing, Phys. Rev. A **91**, 052111 (2015).

[26] Y. Cai, J.-D. Bancal, J. Romero, and V. Scarani, A new device-independent dimension witness and its experimental implementation, J. Phys. A **49**, 305301 (2016).

[27] J.-D. Bancal, N. Gisin, and S. Pironio, Looking for symmetric Bell inequalities, J. Phys. A **43**, 385303 (2010).

[28] C. Śliwa, Symmetries of the Bell correlation inequalities, Phys. Lett. A **317**, 165 (2003).

[29] D. Collins and N. Gisin, A relevant two qubit Bell inequality inequivalent to the CHSH inequality, J. Phys. A **37**, 1775 (2004).

[30] M.-O. Renou, D. Rosset, A. Martin, and N. Gisin, On the inequivalence of the CH and CHSH inequalities due to finite statistics, J. Phys. A **50**, 255301 (2017).

[31] D. Rosset, J.-D. Bancal, and N. Gisin, Classifying 50 years of Bell inequalities, J. Phys. A **47**, 424022 (2014).

[32] A. Acín, S. Pironio, T. Vértesi, and P. Wittek, Optimal randomness certification from one entangled bit, Phys. Rev. A **93**, 040102(R) (2016).

[33] A. Tavakoli, M. Smania, T. Vértesi, N. Brunner, and M. Bourennane, Self-testing non-projective quantum measurements, arXiv:1811.12712.

[34] E. S. Gómez *et al.*, Device-Independent Certification of a Nonprojective Qubit Measurement, Phys. Rev. Lett. **117**, 260401 (2016).

[35] P. Mironowicz and M. Pawłowski, Experimentally feasible semi-device-independent certification of 4 outcome POVMs, arXiv:1811.12872.

[36] M. Smania, P. Mironowicz, M. Nawareg, M. Pawlowski, A. Cabello, and M. Bourennane, Experimental device-independent certification of a symmetric, informationally complete, positive operator-valued measure, arXiv:1811.12851.

[37] M. Yamashita *et al.*, Personal communication from the SDPA collaboration.

[38] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.122.070501 for general symmetrization techniques, explicit examples and the proof of Theorem 1.

[39] S. Burgdorf and I. Klep, The truncated tracial moment problem, J. Oper. Theory, **68**, 141 (2012).

[40] D. Rosset *et al.* (to be published).

[41] C. Brukner, M. Żukowski, and A. Zeilinger, Quantum Communication Complexity Protocol with Two Entangled Qutrits, Phys. Rev. Lett. **89**, 197901 (2002).

[42] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu, Bell Inequalities for Arbitrarily High-Dimensional Systems, Phys. Rev. Lett. **88**, 040404 (2002).

[43] J. J. Benedetto and M. Fickus, Finite normalized tight frames, Adv. Comput. Math. **18**, 357 (2003).

[44] H. Derksen and G. Kemper, *Computational Invariant Theory* (Springer-Verlag, Berlin, Heidelberg, 2002), Vol. 130.

[45] D. F. Holt, B. Eick, and E. A. O'Brien, *Handbook of Computational Group Theory* (CRC Press, 2005).

[46] J. S. Leon, On an algorithm for finding a base and strong generating set for a group given by generating permutations, Math. Comput. **35**, 941 (1980).

[47] J.-P. Serre, *Linear Representations of Finite Groups, Graduate texts in Mathematics* (Springer, 1977).

[48] J. C. Gilbert and J. Cédric, *Plea for a Semidefinite Optimization Solver in Complex Numbers–The Full Report* (LAAS, INRIA Paris, 2017).

[49] The MOSEK optimization toolbox for MATLAB manual, Published by MOSEK ApS, Denmark. Available at http://docs.mosek.com/7.0/toolbox/index.html.

[50] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, England, 2004).

[51] I. Armeanu, About ambivalent groups, Ann. Math. Blaise Pascal **3**, 17 (1996).
[52] T. Maehara and K. Murota, A numerical algorithm for block-diagonal decomposition of matrix *-algebras with general irreducible components, Jpn. J. Indust. Appl. Math. **27**, 263 (2010).
[53] G. W. Anderson, A. Guionnet, and O. Zeitouni, *An Introduction to Random Matrices* (Cambridge University Press, Cambridge, England, 2009).
[54] K. Murota, Y. Kanno, M. Kojima, and S. Kojima, A numerical algorithm for blockdiagonal decomposition of matrix $f * g$-algebras with application to semidefinite programming, Jpn. J. Indust. Appl. Math. **27**, 125 (2010).
[55] The MATLAB package is available at https://denisrosset.github.io/qdimsum/.
[56] D. Dieks, Overlap and distinguishability of quantum states, Phys. Lett. A **126**, 303 (1988).
[57] A. Peres, How to differentiate between non-orthogonal states, Phys. Lett. A **128**, 19 (1988).
[58] R. Derka, V. Buzek, and A. K. Ekert, Universal Algorithm for Optimal Estimation of Quantum States from Finite Ensembles via Realizable Generalized Measurement, Phys. Rev. Lett. **80**, 1571 (1998).
[59] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, Symmetric informationally complete quantum measurements, J. Math. Phys. (N.Y.) **45**, 2171 (2004).
[60] J. Shang, A. Asadian, H. Zhu, and O. Gühne, Enhanced entanglement criterion via symmetric informationally complete measurements, Phys. Rev. A **98**, 022309 (2018).
[61] S. Gómez, A. Mattar, E. S. Gómez, D. Cavalcanti, O. Jiménez Farías, A. Acín, and G. Lima, Experimental nonlocality-based randomness generation with non-projective measurements, Phys. Rev. A **97**, 040102(R) (2018).
[62] J. B. Brask, A. Martin, W. Esposito, R. Houlmann, J. Bowles, H. Zbinden, and N. Brunner, Megahertz-Rate Semi-Device-Independent Quantum Random Number Generators Based on Unambiguous State Discrimination, Phys. Rev. Applied **7**, 054018 (2017).
[63] C. A. Fuchs, M. C. Hoang, and B. C. Stacey, The SIC question: History and state of play, Axioms **6**, 21 (2017).
[64] G. N. M. Tabia, Experimental scheme for qubit and qutrit symmetric informationally complete positive operator-valued measurements using multiport devices, Phys. Rev. A **86**, 062107 (2012).
[65] Z. E. D. Medendorp, F. A. Torres-Ruiz, L. K. Shalm, G. N. M. Tabia, C. A. Fuchs, and A. M. Steinberg, Experimental characterization of qutrits using symmetric informationally complete positive operator-valued measurements, Phys. Rev. A **83**, 051801(R) (2011).
[66] N. Bent, H. Qassim, A. A. Tahir, D. Sych, G. Leuchs, L. L. Sánchez-Soto, E. Karimi, and R. W. Boyd, Experimental Realization of Quantum Tomography of Photonic Qudits via Symmetric Informationally Complete Positive Operator-Valued Measures, Phys. Rev. X **5**, 041006 (2015).
[67] F. Bouchard, K. Heshami, D. England, R. Fickler, R. W. Boyd, B.-G. Englert, L. L. Sánchez-Soto, and E. Karimi, Experimental investigation of high-dimensional quantum key distribution protocols with twisted photons, Quantum **2**, 111 (2018).
[68] W. M. Pimenta, B. Marques, T. O. Maciel, R. O. Vianna, A. Delgado, C. Saavedra, and S. Pádua, Minimum tomography of two entangled qutrits using local measurements of one-qutrit symmetric informationally complete positive operator-valued measure, Phys. Rev. A **88**, 012112 (2013).
[69] A. Tavakoli, J. Kaniewski, T. Vértesi, D. Rosset, and N. Brunner, Self-testing quantum states and measurements in the prepare-and-measure scenario, Phys. Rev. A **98**, 062307 (2018).
[70] N. Brunner, M. Navascués, and T. Vértesi, Dimension Witnesses and Quantum State Discrimination, Phys. Rev. Lett. **110**, 150501 (2013).
[71] A. Ambainis, A. Nayak, A. Ta-Shama, and U. Varizani, Dense quantum coding and quantum finite automata, J. ACM, **49**, 496 (2002).
[72] A. Tavakoli, A. Hameedi, B. Marques, and M. Bourennane, Quantum Random Access Codes Using Single d-Level Systems, Phys. Rev. Lett. **114**, 170502 (2015).
[73] M. Froissart, Constructive generalization of Bells inequalities, Il Nuovo Cimento B **64**, 241 (1981).
[74] D. Rosset, N. Gisin, and E. Wolfe, Universal bound on the cardinality of local hidden variables in networks, Quantum Inf. Comput. **18**, 0910 (2018).

PAPER • OPEN ACCESS

Sequential random access codes and self-testing of quantum measurement instruments

IOP Publishing New J. Phys. 21 (2019) 083034 https://doi.org/10.1088/1367-2630/ab3773

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

OPEN ACCESS

RECEIVED
21 May 2019

REVISED
12 July 2019

ACCEPTED FOR PUBLICATION
31 July 2019

PUBLISHED
20 August 2019

PAPER

Sequential random access codes and self-testing of quantum measurement instruments

Karthik Mohan, Armin Tavakoli and Nicolas Brunner

Département de Physique Appliquée, Université de Genève, CH-1211 Genève, Switzerland

E-mail: armin.tavakoli@unige.ch

Keywords: random access code, quantum correlations, self-testing

Abstract

Quantum random access codes (QRACs) are key tools for a variety of protocols in quantum information theory. These are commonly studied in prepare-and-measure scenarios in which a sender prepares states and a receiver measures them. Here, we consider a three-party prepare-transform-measure scenario in which the simplest QRAC is implemented twice in sequence based on the same physical system. We derive optimal trade-off relations between the two QRACs. We apply our results to construct semi-device independent self-tests of quantum instruments, i.e. measurement channels with both a classical and quantum output. Finally, we show how sequential QRACs enable inference of upper and lower bounds on the sharpness parameter of a quantum instrument.

1. Introduction

Random access codes (RACs) are an important class of communication tasks with a broad scope of applications. In a RAC, a party Alice holds a set of randomly sampled data and another party Bob attempts to recover some randomly chosen subset of Alice's data. This is made possible by Alice communicating with Bob. Therefore, this corresponds to a *prepare-and-measure scenario* in which Alice encodes her data into a message that she sends to Bob who aims to decode the relevant information. Naturally, this task would be trivial if Alice is allowed to send unlimited information. Therefore, a RAC requires that the message is restricted in its alphabet, so that it cannot encode all of Alice's data. Interestingly however, the probability of Bob to access the desired information can be increased if Alice substitutes her classical message with a quantum message of the same alphabet. Such quantum random access codes (QRACs) have been introduced and developed for qubit systems [1, 2] as well as higher-dimensional quantum systems [3]. They are primitives for network coding [4], random number generation [5] and quantum key distribution [6]. QRACs are also common in foundational aspects of quantum theory; examples include the comparison of different quantum resources [7, 8], dimension witnessing [9], self-testing [10–12] and attempts at characterising quantum correlations from information-theoretic principles [13].

Here, we present RACs beyond standard prepare-and-measure scenarios. Specifically, we consider a 'prepare-transform-measure' scenario involving three parties, Alice, Bob and Charlie, in a line configuration. In our scenario, both Bob and Charlie are interested in randomly accessing some information held by Alice, i.e. they individually implement a RAC with Alice. In a classical picture, such sequential RACs are trivial since any information made available to Bob via Alice's communication also can be relayed by Bob to Charlie. In this sense, there is no trade-off between how well Bob and Charlie can perform their RACs. In a quantum picture however, Alice communicates a qubit system that is first sent to Bob who applies a quantum instrument (a completely positive trace-preserving map with both a classical and quantum output) whose classical output is recorded and whose quantum output is relayed to Charlie who performs a measurement. Importantly, Bob's instrument disturbs the physical state of Alice's qubit, and therefore he cannot relay Alice's original quantum message to Charlie. In other words, Charlie's ability to access the desired information depends on Bob's preceding interaction. Consequently, one expects a trade-off in the ability of Bob and Charlie to perform their separate QRACs. Here, we consider Bob and Charlie the simplest RAC for qubits (sometimes referred to as a $2 \rightarrow 1$ RAC) in sequence, and derive the optimal trade-off relation between the two QRACs. In particular, we find that both QRACs can outperform the best possible classical RAC.

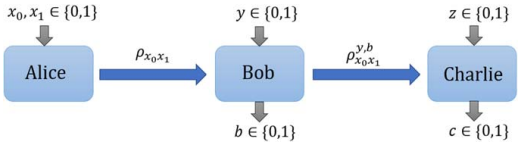

Figure 1. A three-party prepare-transform-measure scenario. Alice samples qubit states from an ensemble of four preparations. Bob performs one of two instruments with a binary classical register and qubit output. Charlie performs one of two binary-outcome measurements.

Subsequently, we apply our results to self-test a quantum instrument. Self-testing [14] is the task of inferring physical entities (states, channels, measurements) solely from correlations produced in experiments i.e. identifying the unique physical entities that are compatible with observed data. Self-testing is typically studied in Bell experiments where notably methods for self-testing quantum instruments have been developed [15, 16]. Recently however, self-testing was introduced in the broad scope of prepare-and-measure scenarios [10], and was further developed using QRACs to robustly self-test both preparations and measurements [10–12]. Notably however, prepare-and-measure scenarios do not enable self-tests of general quantum operations. In particular, it does not enable self-tests of quantum instruments since the quantum system after the measurement is irrelevant to the outcome statistics produced in the experiment. We show that our prepare-transform-measure scenario overcomes this conceptual limitation. We find that optimal pairs of sequential QRACs self-test quantum instruments. However, such optimal correlations require idealised (noiseless) scenarios which are never the case in a practical implementation. Therefore, we also show how sequential QRACs allow for inference of noise-robust bounds on the sharpness parameter in a quantum instrument. This is makes our results applicable to experimental demonstrations. Finally, we discuss relevant generalisations of our results.

2. Sequential RACs

We focus on a prepare-transform-measure scenario that involes three parties. The first party (Alice) receives a uniformly random four-valued input $x = (x_0, x_1) \in \{0, 1\}^2$. For a given input, she prepares a quantum state ρ_x. This state is uncharacterised, up to the assumption of it being of Hilbert space dimension two, i.e. it is a qubit. The state is transmitted to the second party (Bob) who receives a random binary input $y \in \{0, 1\}$. Depending on his input, Bob applies an instrument characterised by Kraus operators $\{K_{b|y}\}$ to ρ_x which produces a classical binary outcome $b \in \{0, 1\}$ and a qubit post-measurement state

$$\rho_x^{y,b} = \frac{K_{b|y}\rho_x K_{b|y}^\dagger}{\mathrm{tr}(\rho_x K_{b|y}^\dagger K_{b|y})}. \tag{1}$$

Notably, since the instrument realises a measurement, the Kraus operators of Bob must satisfy the completeness relation $\forall y : M_{0|y} + M_{1|y} = \mathbb{1}$, where $M_{b|y} = K_{b|y}^\dagger K_{b|y}$ are the corresponding elements of the positive operator-valued measures (POVMs). The post-measurement state $\rho_x^{y,b}$ is relayed to the third party (Charlie) who receives a random binary input $z \in \{0, 1\}$ to which he associates POVMs $\{C_{c|z}\}$ with a binary outcome $c \in \{0, 1\}$. The scenario is illustrated in figure 1.

In the limit of repeating the experiment many times, the results are described by the probability distribution

$$p(b, c|x, y, z) = \mathrm{tr}[K_{b|y}\rho_x K_{b|y}^\dagger C_{c|z}]. \tag{2}$$

To enable a simple and qualitative treatment of the information stored in the distribution, one may employ a correlation witness, i.e. a map from $p(b, c|x, y, z)$ to a single real number. We are interested in two separate correlation witnesses, each corresponding to a RAC The first RAC is considered between Alice and Bob. In this task, the partners are collectively awared a point if and only if Bob can guess the y'th bit of Alice input (x_0, x_1). The correlation witness is the average success probability. It reads

$$W_{\mathrm{AB}} = \frac{1}{8}\sum_{x,y} p(b = x_y|x, y) = \frac{1}{8}\sum_{x,y} \mathrm{tr}[\rho_x M_{x_y|y}], \tag{3}$$

where in the second step we have assumed a quantum description. In a classical picture (in which all states are diagonal in the same basis), this witness obeys $W_{\mathrm{AB}} \leqslant 3/4$ (which we further discuss later). The physical properties of $\{\rho_x\}$ and $\{M_{b|y}\}$ when the QRAC exceeds its classical bound were studied in [10]. It was shown that an optimal QRAC for qubits

$$W_{\mathrm{AB}} = \frac{1}{2}\left(1 + \frac{1}{\sqrt{2}}\right) \approx 0.854 \tag{4}$$

self-tests that Alice's four preparations form a square in some disk of the Bloch sphere. Up to a choice of reference frame these are written

$$\rho_{00} = \frac{1}{2}\left(\mathbb{1} + \frac{\sigma_x + \sigma_z}{\sqrt{2}}\right), \quad \rho_{11} = \frac{1}{2}\left(\mathbb{1} - \frac{\sigma_x + \sigma_z}{\sqrt{2}}\right),$$
$$\rho_{01} = \frac{1}{2}\left(\mathbb{1} + \frac{\sigma_x - \sigma_z}{\sqrt{2}}\right), \quad \rho_{10} = \frac{1}{2}\left(\mathbb{1} - \frac{\sigma_x - \sigma_z}{\sqrt{2}}\right), \tag{5}$$

where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ denotes the Pauli matrices. Moreover, an optimal QRAC also self-tests Bob's observables (defined as $M_y = M_{0|y} - M_{1|y}$) to be anticommuting. In the stated frame, the observables are written

$$M_0 = \sigma_x \qquad M_1 = \sigma_z. \tag{6}$$

Evidently however, the QRAC (3) is independent of both Charlie and of the choice of instrument for realising the POVMs $\{M_{b|y}\}$. To also take these into account, we consider an additional QRAC implemented between Alice and Charlie. Analogously, the partners are awarded a point if and only if Charlie can guess the z'th bit of Alice's input (x_0, x_1). The correlation witness corresponding to this QRAC reads

$$W_{\mathrm{AC}} = \frac{1}{8}\sum_{x,z} p(c = x_z|x, z). \tag{7}$$

This QRAC is not independent of Bob since he applies an instrument to the preparation of Alice before they arrive to Charlie. In a quantum model, the effective state $\tilde{\rho}_x$ received by Charlie is the post-measurement state of Bob averaged over Bob's inputs and classical outputs, i.e.

$$\tilde{\rho}_x = \frac{1}{2}\sum_{y,b} p(b|y)\rho_x^{y,b} = \frac{1}{2}\sum_{y,b} K_{b|y}\rho_x K_{b|y}^{\dagger}. \tag{8}$$

Therefore, we have

$$W_{\mathrm{AC}} = \frac{1}{8}\sum_{x,z} \mathrm{tr}[\tilde{\rho}_x C_{x_z|z}] = \frac{1}{16}\sum_{x,y,b,z} \mathrm{tr}[K_{b|y}\rho_x K_{b|y}^{\dagger} C_{x_z|z}]. \tag{9}$$

We are interested in the values attainable for the pair of QRACs $(W_{\mathrm{AB}}, W_{\mathrm{AC}})$. We remark that the interesting range is when both W_{AB} and W_{AC} are confined to the interval $[1/2, (1 + 1/\sqrt{2})/2]$ since either witness being $1/2 - \epsilon$ for some $\epsilon > 0$ is equivalent to a witness value of $1/2 + \epsilon$ by classically bit-flipping the outcomes.

Typically, we expect there to be a trade-off between the two QRACs. The reason is as follows. In order for W_{AB} to be large, Alice must prepare states that are close to the ones in equation (5) and Bob must implement instruments that realise POVMs that are close to the ones in equation (6). This means that Bob's measurements must be reasonably sharp. This leads to a large disturbance in the state of the measured system which causes the effective ensemble of states $\{\tilde{\rho}_x\}$ arriving to Charlie to lesser reflect the ensemble $\{\rho_x\}$ originally prepared by Alice. Therefore, the value of W_{AC} is expected to be small. Conversely, if Bob makes a very unsharp measurement (almost noninteracting), he could almost completely avoid disturbing the state of Alice's system and thus we would find that $\{\tilde{\rho}_x\}$ closely approximates $\{\rho_x\}$ which allows Charlie to find a large value of W_{AC}. However, the weak interaction of Bob then would imply a correspondingly small value of W_{AB}.

In view of the above, characterising the set of pairs $(W_{\mathrm{AB}}, W_{\mathrm{AC}})$ that can be attained in quantum theory is a nontrivial matter. By finding such a characterisation and by understanding the trade-off between the two QRACs, we enable self-tests of Bob's instrument, along with self-tests of Alice's preparations and Charlie's measurements. Note that one may also consider alternative generalisations of QRACs to sequential scenarios [17].

3. Quantum correlations in sequential RACs

Which values of the pair of QRACs $(W_{\mathrm{AB}}, W_{\mathrm{AC}})$ can be realised in a quantum model based on qubit systems? Before addressing this matter, let us first examine the substantially simpler situation in which the physical devices are classical, i.e. the state at all times is diagonal in the same basis. In such situations, Bob can interact with the preparations of Alice without disturbing their state. Therefore, a large value of W_{AB} constitutes no obstacle for also finding a large value of W_{AC}. Classically, one can optimally achieve $W_{\mathrm{AB}} = 3/4$. Clearly, as the interaction with Bob cannot contribute towards increasing the value of W_{AC}, it also holds that $W_{\mathrm{AC}} \leqslant 3/4$. This value is saturated by Alice sending x_0 to Bob who outputs $b = x_0$ and relays x_0 to Charlie who outputs $c = x_0$.

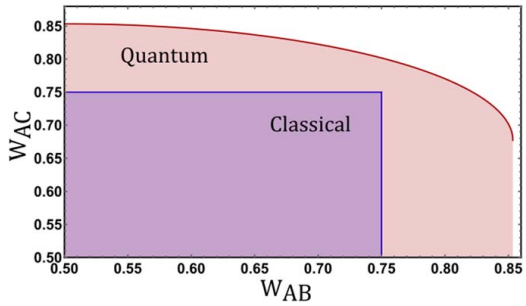

Figure 2. The correlations attainable in the space of the two witnesses (W_{AB}, W_{AC}) in a classical and quantum model respectively. The nontrivial part of the boundary of the quantum set is highligted by a solid red line. Its right-end extremal point is (W_{AB}, W_{AC}) = $\left(\frac{2+\sqrt{2}}{4}, \frac{4+\sqrt{2}}{8}\right)$. The extremal point for equal witnesses is $W_{AB} = W_{AC} = \frac{5+2\sqrt{2}}{10} > 3/4$.

Thus, the set of classically attainable correlations is $1/2 \leqslant (W_{AB}, W_{AC}) \leqslant 3/4$. This classically attainable set is illustrated in figure 2. Notice that there is no trade-off between W_{AB} and W_{AC} in a classical picture.

In a quantum model, the characterisation of the attainable set of witnesses is less straightforward. We phrase the problem as follows: for a given value (denoted α) of W_{AB}, what is the maximal value of W_{AC} possible in a quantum model? Answering this question for every $\alpha \in [1/2, (1 + 1/\sqrt{2})/2]$ provides the optimal trade-off between the two QRACs. Equivalently, it can be viewed as the nontrivial part of the boundary of the quantum set of correlations in the space of (W_{AB}, W_{AC}). Formally, the optimisation problem reads

$$\begin{aligned} &W^{\alpha}_{AC} = \max_{\rho,U,M,C} W_{AC} \\ &\text{such that } \forall x : \rho_x \in \mathbb{C}^2, \quad \rho_x \geqslant 0, \quad \mathrm{tr}\rho_x = 1, \\ &\forall z, c : C_{c|z} \geqslant 0, \quad C_{0|z} + C_{1|z} = \mathbb{1} \\ &\forall y, b : U_{yb} \in \mathrm{SU}(2), \quad M_{b|y} \geqslant 0, \quad M_{0|y} + M_{1|y} = \mathbb{1}, \\ &\text{and } W_{AB} = \alpha, \end{aligned} \tag{10}$$

i.e. it is an optimisation of Charlie's witness over all preparations, instruments and measurements that can model the observation of $W_{AB} = \alpha$. In the above, we have used the polar decomposition to write the Kraus operators as $K_{b|y} = U_{yb}\sqrt{M_{b|y}}$ for some unitary operator U_{yb} and some POVM $\{M_{b|y}\}$. Kraus operators of this form correspond to extremal quantum instruments in the considered scenario [18].

We solve the problem (10) by first giving a lower bound on W^{α}_{AC} and then matching it with an upper bound. To this end, consider a quantum strategy in which Alice prepares the ensemble of states given in equation (5) and Charlie performs the measurements in equation (6). We let Bob perform an unsharp Lüders measurement (the Kraus operators have $U_{yb} = \mathbb{1}$) of the observables in equation (6), i.e. his observables correspond to $M_0 = \eta\sigma_x$ and $M_1 = \eta\sigma_z$ for some *sharpness parameter* $\eta \in [0, 1]$ (which we will later self-test). Evaluating the pair of witnesses with this quantum strategy gives

$$\begin{aligned} W_{AB} &= \frac{1}{4}(2 + \eta\sqrt{2}) \\ W_{AC} &= \frac{1}{8}(4 + \sqrt{2} + \sqrt{2 - 2\eta^2}). \end{aligned} \tag{11}$$

Parameterising the latter in terms of the former returns a lower bound on W^{α}_{AC}. Importantly, this bound is optimal since it can be saturated with an upper bound on W^{α}_{AC}, thus solving the optimisation problem (10). This leads us to our first result.

Result 1. The optimal trade-off between the pair of QRACs (W_{AB}, W_{AC}) corresponds to

$$W^{\alpha}_{AC} = \frac{1}{8}(4 + \sqrt{2} + \sqrt{16\alpha - 16\alpha^2 - 2}), \tag{12}$$

where $\alpha \in [1/2, (1 + 1/\sqrt{2})/2]$. That is, the optimal witness pairs are of the form $(W_{AB}, W_{AC}) = (\alpha, W^{\alpha}_{AC})$. This characterises the nontrivial boundary of the quantum set in the space of witness pairs.

The proof is analytical, of technical character and detailed in appendix B. It relies on (i) treating the maximisation in (10) over the unitaries $\{U_{yb}\}$ and measurements $\{C_{c|z}\}$ as an eigenvalue problem, (ii) using the

Bloch sphere parameterisation for the preparations and instruments, and (iii) noticing that the maximisation over the preparations can be relaxed to a maximisation over two pairs of antipodal pure states in some disk of the Bloch sphere.

In figure 2, we have illustrated the set of sequential QRACs attainable in a quantum model. Notice that a maximal value (4) of W_{AB} does not imply that W_{AC} is no better than what is obtained by random guessing. In contrast, one can achieve $(W_{\mathrm{AB}},\ W_{\mathrm{AC}}) = \left(\frac{2+\sqrt{2}}{4}, \frac{4+\sqrt{2}}{8}\right)$. The reason is that the ensemble relayed to Charlie corresponds to that originally prepared by Alice but with Bloch vectors of half the original length. In addition, there exists a subset of the quantum set in which both W_{AB} and W_{AC} exceed the classical bound.

4. Self-testing

Finding the optimal trade-off between the two QRACs (Result 1) allows for self-testing. To obtain a self-test, one must additionally show that the optimal QRAC pairs only admit a realisation with unique preparations, instruments and measurements (up to collective unitary transformations). That is, we need to identify the unique physical entities $\{\rho_x\}$, $\{K_{b|y}\}$, and $\{C_{c|z}\}$ necessary for optimal correlations.

Such a self-testing argument can be established largely from the proof of Result 1 (see appendix B). The reason is that our approach to deriving Result 1 successively identifies the form of the physical entities required for optimality. To turn the statement into a self-test, we identify key inequalities used to upper bound W^{α}_{AC} and instead impose strict equality constraints. This allows us to pinpoint the states, measurements and instruments one by one. These additional arguments are discussed in appendix B. This leads us to the following self-test statement based on optimal sequential QRACs.

Result 2. An optimal pair of QRACs $(W_{\mathrm{AB}},\ W_{\mathrm{AC}}) = (\alpha,\ W^{\alpha}_{\mathrm{AC}})$, as in equation (10), self-tests that

- Alice's states are pure and pairwise antipodal on the Bloch sphere, on which they form a square. These correspond to the states given in equation (5).
- Bob's instruments are Kraus operators $K_{b|y} = U_{yb}\sqrt{M_{b|y}}$ that correspond to unsharp measurements along the diagonals of Alice's square of preparations followed by a collective unitary. Specifically, $\forall\, y,\ b$: $U_{yb} = U$, $M_0 = \eta\sigma_x$ and $M_1 = \eta\sigma_z$ where $\eta = \sqrt{2}\,(2W_{\mathrm{AB}} - 1)$.
- Charlie's measurements are rank-one projective along the diagonals of the square formed by Alice's preparations, up to the unitary of Bob. That is, $C_0 = U\sigma_x U^{\dagger}$ and $C_1 = U\sigma_z U^{\dagger}$.

The self-tests are valid up to a collective choice of reference frame.

This result applies to optimal pairs of QRACs (highlighted by a solid red line in figure 2). An interesting question is how to make this result noise-tolerant so that it applies to suboptimal pairs of QRACs that nevertheless lack a classical model. Naturally, when the QRACs are suboptimal, one can no longer pinpoint the physical entities as done in Result 2. However, it is possible to give qualitative statements about the quantum strategies that in principle could model the observed correlations. We consider this matter for the sharpness parameter in Bob's instruments. Since any binary-outcome qubit observable can be written on the form $M_y = c_{y0}\mathbb{1} + \vec{c}_y \cdot \vec{\sigma}$, we define the sharpness parameter of Bob's instrument as the length of the Bloch vector $\vec{c}_y$. For simplicity, we take both his instruments to have the same sharpness $\eta \equiv |\vec{c}_0| = |\vec{c}_1|$.

We can place a lower bound on η from the witness W_{AB}; it corresponds to the smallest η for which there exists preparations and instruments that can model W_{AB}. In appendix C, we show that this lower bound reads

$$\eta \geqslant \sqrt{2}\,(2W_{\mathrm{AB}} - 1). \tag{13}$$

This lower bound is nontrivial whenever $W_{\mathrm{AB}} > 1/2$. Notice also that an optimal QRAC (4) necessitates a sharp measurement ($\eta = 1$). Similarly, we can place an upper bound on η from the witness W_{AC}, corresponding to the largest η for which there exists preparations, instruments and measurements that can model W_{AC}. In appendix C we show that such a bound reads

$$\eta \leqslant 2\sqrt{(2 + \sqrt{2} - 4W_{\mathrm{AC}})(2W_{\mathrm{AC}} - 1)}\,, \tag{14}$$

when $\frac{4+\sqrt{2}}{8} \leqslant W_{\mathrm{AC}} \leqslant \frac{2+\sqrt{2}}{4}$ (otherwise the bound is trivial). The lower bound (13) and the upper bound (14) are tight, i.e. they can be saturated with an explicit quantum strategy. Notice that the upper bound (14) conincides with the lower bound (13) for optimal W_{AC} (i.e. when $W_{\mathrm{AC}} = W^{\alpha}_{\mathrm{AC}}$) as given in equation (12). In addition, the bound (14) reduces to the trivial $\eta \leqslant 1$ when $W_{\mathrm{AC}} = (4 + \sqrt{2})/8 \approx 0.6767$.

As a simple example, consider an experiment that attempts to implement the quantum strategy (11) for the optimal witness pair $(W_{AB},\ W_{AC})$ corresponding to $\eta = 1/\sqrt{2}$. However the experiment is subject to losses. For example, take a 95% visibility[1] in Alice's preparations, 90% visibility in Bob's instruments, and 95% visibility in Charlie's measurements. Instead of finding the optimal witness pair $(W_{AB},\ W_{AC}) = (3/4,\ (5+\sqrt{2})/8)$, one finds $(W_{AB},\ W_{AC}) \approx (0.7138,\ 0.7826)$. Therefore, we find that η must be confined to the interval $0.6047 \leqslant \eta \leqslant 0.8010$. The interval is fairly wide, which emphasises the need for high-quality practical realisations in order to confine η to a reasonably small interval.

5. Generalisations

Above, we have thoroughly considered the scenario in which a sequence of three observers implement a pair of the simplest QRAC. This is arguably the simplest scenario in which to study sequential QRACs. It would be interesting to consider more general scenarios; both involving higher-dimensional [3] and many-input QRACs, as well as sequences of more than three observers.

Consider for example the above considered RAC played between Alice and a sequence of N parties. We denote the RAC between Alice and sequential party number k by W_k. Let Alice prepare the optimal states in equation (5). We know that if the first party performs optimal projective measurements (6) (with Kraus operators $K_{b|y} = M_{b|y}$), he will find the optimal QRAC given in equation (4). Moreover, if the second party performs the same Kraus operators we find $W_3 = (1 + 1/(2\sqrt{2}))/2$. The reason is that the effective state ensemble (8) relayed by the first party is identical the the preparations of Alice except that their Bloch vectors have shrunk to half the unit lenght. Similarly, the effective ensemble relayed by the second party will be identical to that relayed by the first party, except that the Bloch vectors will again by shrunk to a quarter of unit length. Continuing the sequence in this manner, the square formed in the Bloch sphere by the effective post-measurement ensemble will at each step have its half-diagonal reduced by a factor $1/2$, and we find

$$W_k = \frac{1}{2}\left(1 + \frac{\sqrt{2}}{2^k}\right). \tag{15}$$

Moreover, one can ask what is the longest sequence of QRACs such that all of them can exceed the classical bound. The number is at least two, since we found $W_{AB} = W_{AC} = \frac{5+2\sqrt{2}}{10} \approx 0.7828 > 3/4$. However, a third sequential violation is unlikely to be possible, i.e. to find $W_1 = W_2 = W_3 > 3/4$. The reason is based on the possibility of relating witnesses in dimension-bounded prepare-and-measure scenarios to Bell inequalities [19–21]. Via such methods, the considered RAC can be related to the CHSH inequality [19]. However, sequential violations of the CHSH inequality were studied in [22] and it was found that no more than two CHSH inequality violations are possible when inputs are uniformly distributed [23, 24].

6. Conclusions

We have studied sequential QRACs and characterised their optimal trade-off. This ties in with the recent interest in sequential quantum correlations obtained in various forms of tests of nonclassicality [22, 24–30]. We applied our results to show that quantum instruments can be semi device-independently self-tested. Notably, since all quantum instruments also realise some POVM, our results trivially implies a certification of unsharp measurements. Our results complement the many recent self-tests of preparations and measurements in standard prepare-and-measure scenarios with a method for self-testing quantum instruments. In addition, we showed how to robustly certify the sharpness parameter of quantum instruments based on noisy correlations. This makes our results readily applicable to experimental applications. Such tests are well within the state-of-the-art experiments [30–32]. Moreover, we notice that the class of quantum instruments self-tested in this work are precisely those implemented by the experimental realisations in [30–33].

We conclude with some open questions. Firstly, it would be interesting to generalise our results to cover higher-dimensional QRACs and longer sequences of observers. Secondly, a possible further development is to characterise the optimal trade-off between sequential QRACs encountered in tests of preparation contextuality [30]. Thirdly, in the spirit of [15], it would be interesting to develop noise-robust self-testing of quantum instruments. Typically, such a robust self-test address the closeness (based on observed witness values) between the unknown laboratory instrument and the ideal instrument that would have been self-tested in case correlations were optimal. Finally, one could consider the task of self-testing quantum instruments based on the sequential correlation experiments in the fully device-independent scenario (see [22]).

[1] Here, visibility corresponds to a parameter $v \in [0,\ 1]$ and means that the ideal physical entity is implemented with probability v and with probability $(1 - v)$ the implemented physical entity is maximally mixed.

Acknowledgments

We thank Denis Rosset and Jędrzej Kaniewski for discussions. This work was funded by the Swiss National Science Foundation (Starting grant DIAQ, NCCR-QSIT).

Note added

During the completion of this work, we became aware of the related work of [34].

Appendix A. Proof of results 1 and 2

We first prove Result 1 and then develop the argument further to also prove Result 2.

Consider the maximisation of the witness

$$W_{\mathrm{AC}} = \frac{1}{16}\sum_{x,y,b,z} \mathrm{tr}[K_{b|y}\rho_x K_{b|y}^{\dagger} C_{x_z|z}] \tag{A1}$$

under the constraint that

$$\alpha \equiv W_{\mathrm{AB}} = \frac{1}{8}\sum_{x,y} \mathrm{tr}[\rho_x K_{x_y|y}^{\dagger} K_{x_y|y}]. \tag{A2}$$

The optimisation is relevant for every $\alpha \in [1/2,\ (1+1/\sqrt{2})/2]$, ranging from the trivial witness value to the maximal witness value.

To contend with this, we first use the polar decomposition $K_{b|y} = U_{yb}\sqrt{M_{b|y}}$, where U_{yb} are arbitrary unitary operators. We can then use the cyclicity of the trace along with the substitution $C_{1|z} = \mathbb{1} - C_{0|z}$ to write equation (A1) as

$$W_{\mathrm{AC}} = \frac{1}{2} + \frac{1}{16}\sum_{x,y,b,z} (-1)^{x_z}\mathrm{tr}[\sqrt{M_{b|y}}\rho_x\sqrt{M_{b|y}}U_{yb}^{\dagger}C_{0|z}U_{yb}]. \tag{A3}$$

The sum over x can be moved inside the trace; we define $\gamma_z = \sum_x (-1)^{x_z}\rho_x$. Moreover, we also define $A_{zyb} = U_{yb}^{\dagger}C_{0|z}U_{yb}$. We can now consider the optimisation over $\{U_{yb}\}$ and $\{C_{c|z}\}$ as a single optimisation over A_{zyb}. To this end, we note that the set of measurements $\{C_{c|z}\}$ is convex. Therefore, every nonextremal (interior point) measurement can be written as a convex combination of extremal measurements (on the boundary). Due to linearity, no nonextremal POVM can lead to a larger value of W_{AC} than some extremal POVM. The extremal binary-outcome qubit measurements are rank-one projectors. Therefore, we can consider the optimisation over A_{zyb} as an optimisation over general rank-one projectors. This gives

$$\begin{aligned}\max W_{\mathrm{AC}} &= \frac{1}{2} + \max_{\rho,A,M}\frac{1}{16}\sum_{y,b,z}\mathrm{tr}[\sqrt{M_{b|y}}\gamma_z\sqrt{M_{b|y}}A_{zyb}] \\ &= \frac{1}{2} + \max_{\rho,M}\frac{1}{16}\sum_{y,b,z}\lambda_{\max}[\sqrt{M_{b|y}}\gamma_z\sqrt{M_{b|y}}],\end{aligned} \tag{A4}$$

where we have made the optimal choice of letting A_{zyb} project onto the eigenvector of $\sqrt{M_{b|y}}\gamma_z\sqrt{M_{b|y}}$ with the largest eigenvalue (denoted by $\lambda_{\max}$).

To proceed further, we make use of the fact that qubit operations can be parameterised on the Bloch sphere. We write the preparations as $\rho_x = (\mathbb{1} + \vec{n}_x\cdot\vec{\sigma})/2$ for some Bloch vectors $\vec{n}_x \in \mathbb{R}^3$ with $|\vec{n}_x| \leqslant 1$. This leads to

$$\gamma_z = [(\vec{n}_{00} - \vec{n}_{11}) + (-1)^z(\vec{n}_{01} - \vec{n}_{10})]\cdot\vec{\sigma}. \tag{A5}$$

We define the effective (unnormalised) Bloch vectors $\vec{m}_z = (\vec{n}_{00} - \vec{n}_{11}) + (-1)^z(\vec{n}_{01} - \vec{n}_{10})$. Consequently, the dependence of W_{AC} on the preparations can be reduced to its dependence on $(\vec{m}_0,\ \vec{m}_1)$. However, given any set of preparations $\{\vec{n}_x\}$, we can consider other preparations $\{\vec{n}'_x\}$ choosen such that $\vec{n}'_{00} = -\vec{n}'_{11}$ and $\vec{n}'_{01} = -\vec{n}'_{10}$ with $2\vec{n}'_{00} = \vec{n}_{00} - \vec{n}_{11}$ and $2\vec{n}'_{01} = \vec{n}_{01} - \vec{n}_{10}$. The both ensembles $\{\vec{n}_x\}$ and $\{\vec{n}'_x\}$ imply the same vectors $(\vec{m}_0,\ \vec{m}_1)$. Moreover, it is evident that if not all preparations are pure, one cannot obtain optimal correlations (since impurity corresponds to decreasing the magnitude of $(\vec{m}_0,\ \vec{m}_1)$). This means that the Bloch vectors are of unit lenght and therefore that the optimal preparations *must* be of the type $\{\vec{n}'_x\}$ (i.e. two antipodal pairs). Notice that purity also implies that $\vec{m}_0\cdot\vec{m}_1 = 0$.

W.l.g. we can choose a reference frame in which $\vec{m}_0 \propto (1,\ 0,\ 0)$ and $\vec{m}_1 \propto (0,\ 0,\ 1)$. We denote the relative angle between the two pairs of antipodal preparation pairs by $\theta \in [0, \pi/2]$. This gives

$$|\vec{m}_0| = \sqrt{2(1+\cos\theta)} \quad \text{and} \quad |\vec{m}_1| = \sqrt{2(1-\cos\theta)}.$$

We can further place an upper bound on equation (A4) by using the following relation

$$\forall M, \forall \vec{a} \in \mathbb{R}^3 : \sum_{b=0,1} \lambda_{\max}[\sqrt{M_b}(\vec{a}\cdot\vec{\sigma})\sqrt{M_b}] \leqslant |\vec{a}|, \tag{A6}$$

with equality if and only if $\vec{a}$ is aligned with the Bloch vector of the POVM. Identifying $\vec{a}$ with $\vec{m}_z$, we apply it twice to equation (A4) corresponding to the terms in which $z = y$. This gives

$$W_{\mathrm{AC}} \leqslant \frac{1}{2} + \frac{1}{16}\left(|\vec{m}_0| + |\vec{m}_1| + \underbrace{\sum_{y,b} \lambda_{\max}[\sqrt{M_{b|y}}(\vec{m}_{\bar{y}}\cdot\vec{\sigma})\sqrt{M_{b|y}}]}_{=S}\right), \tag{A7}$$

where $\bar{y}$ denotes a bit-flip. We turn our attention to the sum denoted by S in equation (A7). We define the observable $M_y = M_{0|y} - M_{1|y}$ and apply the Bloch sphere parameterisation. We may write $M_y = c_{y0}\mathbb{1} + \vec{c}_y\cdot\vec{\sigma}$ where $\vec{c}_y = (c_{y1}, c_{y2}, c_{y3})$ with $|\vec{c}_y| \leqslant 1$ and $|\vec{c}_y| - 1 \leqslant c_{y0} \leqslant 1 - |\vec{c}_y|$. These constraints ensure positivity. Hence

$$M_{b|y} = f_{yb}|\vec{c}_y\rangle\langle\vec{c}_y| + h_{yb}|-\vec{c}_y\rangle\langle-\vec{c}_y|, \tag{A8}$$

where $|\vec{c}_y\rangle$ is the pure state corresponding to the Bloch sphere direction of $\vec{c}_y$, and

$$\begin{aligned} f_{yb} &= \frac{1}{2}(1 + (-1)^b c_{y0} + (-1)^b|\vec{c}_y|) \\ h_{yb} &= \frac{1}{2}(1 + (-1)^b c_{y0} - (-1)^b|\vec{c}_y|). \end{aligned} \tag{A9}$$

Firstly, this allows us to write the constraint (A2) as

$$\alpha = \frac{1}{8}(4 + |\vec{m}_0|c_{01} + |\vec{m}_1|c_{13}). \tag{A10}$$

Secondly, we can now solve the characteristic equation $\det(\sqrt{M_{b|y}}(\vec{m}_{\bar{y}}\cdot\vec{\sigma})\sqrt{M_{b|y}} - \mu\mathbb{1}) = 0$, and after some simplifications obtain

$$S = \sum_{y,b}\frac{|\vec{m}_{\bar{y}}|}{2}\sqrt{(1 + (-1)^b c_{y0})^2 - |\vec{c}_y|^2(1 - \langle\vec{c}_y|\hat{m}_{\bar{y}}\cdot\vec{\sigma}|\vec{c}_y\rangle^2)}, \tag{A11}$$

where $\hat{m} = \vec{m}/|\vec{m}|$. We can now consider the optimisation over c_{y0} by separately considering the two terms corresponding to $y = 0$ and $y = 1$ respectively. This amounts to maximising expressions of the form $\sqrt{(1+x)^2 - K} + \sqrt{(1-x)^2 - K}$, for some positive constant K. It is easily shown that such functions are uniquely maximised by setting $x = 0$. Thus, we require $c_{00} = c_{10} = 0$. Moreover, since $(\vec{m}_0, \vec{m}_1)$ have no component along the y-axis, it is seen from (A10) and (A11) that one optimally chooses $c_{02} = c_{12} = 0$. This simplifies matters to

$$\max S = |\vec{m}_0|\sqrt{1 - (c_{11}^2 + c_{13}^2)(1 - c_{11}^2)} + |\vec{m}_1|\sqrt{1 - (c_{01}^2 + c_{03}^2)(1 - c_{03}^2)}. \tag{A12}$$

Note that c_{03} and c_{11} do not appear in the constraint (A10), that they are associated to different settings of Bob and that they appear in different terms in equation (A12). Therefore, we can separately maximise search square-root expression above by standard differentiation. This returns that the unique maximum is attained for $c_{03} = c_{11} = 0$. Hence, we have

$$W_{\mathrm{AC}} \leqslant \frac{1}{2} + \frac{1}{16}(|\vec{m}_0| + |\vec{m}_1| + |\vec{m}_0|\sqrt{1 - c_{13}^2} + |\vec{m}_1|\sqrt{1 - c_{01}^2}) \equiv W. \tag{A13}$$

Denoting $c_{01} = \cos\phi_0$ and $c_{13} = \cos\phi_1$ for $\phi_1, \phi_2 \in [0, \pi/2]$, we can re-write the right hand side on the more convenient form

$$W = \frac{1}{2} + \frac{1}{8}\left(\cos\frac{\theta}{2} + \sin\frac{\theta}{2} + \cos\frac{\theta}{2}\sin\phi_1 + \sin\frac{\theta}{2}\sin\phi_0\right) \tag{A14}$$

and the constraint (A10) as

$$\alpha = \frac{1}{8}\left(4 + \cos\frac{\theta}{2}\cos\phi_0 + \sin\frac{\theta}{2}\cos\phi_1\right). \tag{A15}$$

To maximise W over (θ, ϕ_0, ϕ_1), we use the following lemma.

Lemma 1. *For every tuple (θ, ϕ_0, ϕ_1) corresponding to (α, W), there exists another tuple $(\theta, \phi_0, \phi_1) = (\pi/2, \phi, \phi)$ that produces (α, W') with $W' \geqslant W$. Moreover, $\theta = \pi/2$ and $\phi_0 = \phi_1$ is necessary for an optimal W'.*

To prove this statement, we must show that for all θ, ϕ_0, $\phi_1 \in [0, \pi/2]$ there exists a $\phi \in [0, \pi/2]$ such that

$$\cos\frac{\theta}{2}\cos\phi_0 + \sin\frac{\theta}{2}\cos\phi_1 = \sqrt{2}\cos\phi$$
$$\cos\frac{\theta}{2} + \sin\frac{\theta}{2} + \cos\frac{\theta}{2}\sin\phi_1 + \sin\frac{\theta}{2}\sin\phi_0 \leqslant \sqrt{2} + \sqrt{2}\sin\phi. \tag{A16}$$

Proof. It trivially holds that $\cos\frac{\theta}{2} + \sin\frac{\theta}{2} \leqslant \sqrt{2}$ with equality if and only if $\theta = \pi/2$. We eliminate this part from the second equation in (A16). Then, by squaring both equations, we can combine them into a single equation in which ϕ is eliminated. The statement reduces to the inequality

$$\cos\theta(\cos^2\phi_0 - \cos^2\phi_1) + \sin\theta\cos(\phi_0 - \phi_1) \leqslant 1. \tag{A17}$$

Using differentiation w.r.t. ϕ_0 one finds that the optimum of the left hand side is attained for $\phi_1 = \phi_0$, which proves the relation (A17). ■

By virtue of lemma 1, we can reduce our consideration of (A14) and (A15) to $\theta = \pi/2$ and $c_{01} = c_{13} \equiv c$. Therefore equation (A10) reduces to

$$c = \sqrt{2}(2\alpha - 1) \tag{A18}$$

and we also have $W = \frac{1}{2} + \frac{1}{4\sqrt{2}}(1 + \sqrt{1 - c^2})$. Thus, we have arrived to the upper bound

$$W_{\mathrm{AC}}^{\alpha} \leqslant \frac{1}{8}(4 + \sqrt{2} + \sqrt{16\alpha - 16\alpha^2 - 2}). \tag{A19}$$

As shown in the main text, this upper bound could be saturated with an explicit quantum strategy. This proves Result 1.

Let us now extend this to a proof of Result 2 by more closely examining the above steps needed to arrive at equation (A19). Firstly, we have already shown that the preparations must be pure, pairwise antipodal and by lemma 1 they must have a relative angle of $\pi/2$. Thus, this corresponds to a square in a disk of the Bloch sphere. The above arguments fully characterise Alice's preparations up to a reference frame.

For Bob's instrument, we have shown that the Bloch vectors ($\vec{c}_0$, $\vec{c}_1$) only can have non-zero components in the x- and z-directions respectively and that the length of the Bloch vector is given by equation (A18). This fully characterises the Bloch vectors. Moreover, in equation (A4) we required that A_{zyb} is aligned with the eigenvector of $\sqrt{M_{b|y}}\gamma_z\sqrt{M_{b|y}}$ corresponding to the largest eigenvalue. However, we now have that $\gamma_0 = \sigma_x$ and $\gamma_1 = \sigma_z$ whereas $M_0 \propto \sigma_x$ and $M_1 \propto \sigma_z$. Therefore, we have that $\forall\, y,\ b : A_{0yb} = |+\rangle\langle+|$ and $\forall\, y,\ b$: $A_{1yb} = |0\rangle\langle 0|$. Therefore, we have that

$$\forall\, yb : U_{yb}^{\dagger} C_{0|0} U_{yb} = |+\rangle\langle+|, \tag{A20}$$
$$\forall\, yb : U_{yb}^{\dagger} C_{0|1} U_{yb} = |0\rangle\langle 0|. \tag{A21}$$

This implies that all unitaries are equal; $U_{yb} = U$. Therefore, Charlie's observables $C_z = C_{0|z} - C_{1|z}$ satisfy $C_0 = U\sigma_x U^{\dagger}$ and $C_1 = U\sigma_z U^{\dagger}$.

Appendix B. Bounding the sharpness parameter from noisy correlations

In order to bound the sharpness of Bob's instrument, consider first the witness W_{AB}. Using the notations from the previous appendix, we have that

$$W_{\mathrm{AB}} = \frac{1}{8}(4 + |\vec{c}_0||\vec{m}_0|\hat{m}_0 \cdot \hat{c}_0 + |\vec{c}_1||\vec{m}_1|\hat{m}_1 \cdot \hat{c}_1). \tag{B1}$$

We focus on the simplified case in which the sharpness parameter is the same in Bob's two settings, i.e. $\eta \equiv |\vec{c}_0| = |\vec{c}_1|$. Re-arranging gives

$$\eta = \frac{8W_{\mathrm{AB}} - 4}{|\vec{m}_0|\hat{m}_0 \cdot \hat{c}_0 + |\vec{m}_1|\hat{m}_1 \cdot \hat{c}_1}. \tag{B2}$$

To find the smallest possible η, we maximise the denominator. That corresponds to setting $\hat{m}_0 \cdot \hat{c}_0 = \hat{m}_1 \cdot \hat{c}_1 = 1$ and $|\vec{m}_0| = |\vec{m}_1| = \sqrt{2}$. That gives the lower bound

$$\eta \geqslant \sqrt{2}(2W_{\mathrm{AB}} - 1). \tag{B3}$$

Consider now the witness W_{AC}. In the previous appendix, we have shown that its optimal value for a given choice of $\eta \equiv |\vec{c}_0| = |\vec{c}_1|$ is upper bounded as follows

$$W_{\mathrm{AC}} \leqslant \frac{1}{2} + \frac{1}{4\sqrt{2}}(1 + \sqrt{1 - \eta^2}). \tag{B4}$$

Solving this inequality for η gives

$$\eta \leqslant 2\sqrt{(2 + \sqrt{2} - 4W_{\mathrm{AC}})(2W_{\mathrm{AC}} - 1)}\,. \tag{B5}$$

References

[1] Ambainis A, Nayak A, Ta-Shma A and Vazirani U 1999 Dense quantum coding and a lower bound for 1-way quantum automata *Proc. 31st Annual ACM Symp. on Theory of Computing (STOC'99)* pp 376–83
[2] Ambainis A, Leung D, Mancinska L and Ozols M 2008 *Quantum Random Access Codes with Shared Randomness* arXiv:0810.2937
[3] Tavakoli A, Hameedi A, Marques B and Bourennane M 2015 Quantum random access codes using single d-Level systems *Phys. Rev. Lett.* **114** 170502
[4] Hayashi M, Iwama K, Nishimura H, Raymond R and Yamashita S 2007 Quantum network coding *Proc. 24th Int. Symp. on Theoretical Aspects of Computer Science (STACS 2007), Lecture Notes in Computer Science 4393* pp 610–21
[5] Li H-W, Yin Z-Q, Wu Y-C, Zou X-B, Wang S, Chen W, Guo G-C and Han Z-F 2011 Semi-device-independent random-number expansion without entanglement *Phys. Rev.* A **84** 034301
[6] Pawłowski M and Brunner N 2011 Semi-device-independent security of one-way quantum key distribution *Phys. Rev.* A **84** 010302(R)
[7] Tavakoli A, Marques B, Pawłowski M and Bourennane M 2016 Spatial versus sequential correlations for random access coding *Phys. Rev.* A **93** 032336
[8] Hameedi A, Saha D, Mironowicz P, Pawłowski M and Bourennane M 2017 Complementarity between entanglement-assisted and quantum distributed random access code *Phys. Rev.* A **95** 052345
[9] Wehner S, Christandl M and Doherty A C 2008 Lower bound on the dimension of a quantum system given measured data *Phys. Rev.* A **78** 062112
[10] Tavakoli A, Kaniewski J, Vértesi T, Rosset D and Brunner N 2018 Self-testing quantum states and measurements in the prepare-and-measure scenario *Phys. Rev.* A **98** 062307
[11] Farkas M and Kaniewski J 2019 Self-testing mutually unbiased bases in the prepare-and-measure scenario *Phys. Rev.* A **99** 032316
[12] Tavakoli A, Smania M, Vértesi T, Brunner N and Bourennane M 2018 *Self-Testing Non-Projective Quantum Measurements in Prepare-and-Measure Experiments* arXiv:1811.12712
[13] Pawłowski M, Paterek T, Kaszlikowski D, Scarani V, Winter A and Żukowski M 2009 Information causality as a physical principle *Nature* **461** 1101
[14] Mayers D and Yao A 2004 Quantum cryptography with imperfect apparatus *Quantum Inf. Comput.* **4** 273
[15] Wagner S, Bancal J-D, Sangouard N and Sekatski P 2018 *Device-Independent Characterization of Generalized Measurements* arXiv:1812.02628
[16] Sekatski P, Bancal J-D, Wagner S and Sangouard N 2018 Certifying the building blocks of quantum computers from Bell's theorem *Phys. Rev. Lett.* **121** 180505
[17] Wang Y, Primaatmaja I W, Lavie E, Varvitsiotis A and Lim C C W 2019 Characterising the correlations of prepare-and-measure quantum networks *NPJ Quantum Inf.* **5** 17
[18] Pellonpää J-P 2013 Quantum instruments: I. Extreme instruments *J. Phys. A: Math. Theor.* **46** 025302
[19] Buhrman H, Cleve R and van Dam W 2001 Quantum entanglement and communication complexity *SIAM J. Comput.* **30** 1829
[20] Brukner C, Żukowski M, Pan J-W and Zeilinger A 2004 Bell's inequalities and quantum communication complexity *Phys. Rev. Lett.* **92** 127901
[21] Tavakoli A and Żukowski M 2017 Higher-dimensional communication complexity problems: classical protocols versus quantum ones based on Bell's theorem or prepare-transmit-measure schemes *Phys. Rev.* A **95** 042305
[22] Silva R, Gisin N, Guryanova Y and Popescu S 2015 Multiple observers can share the nonlocality of half of an entangled pair by using optimal weak measurements *Phys. Rev. Lett.* **114** 250401
[23] Mal S, Majumdar A S and Home D 2016 Sharing of nonlocality of a single member of an entangled pair of qubits is not possible by more than two unbiased observers on the other wing *Mathematics* **4** 48
[24] Shenoy A H, Designolle S, Hirsch F, Silva R, Gisin N and Brunner N 2019 Unbounded sequence of observers exhibiting Einstein–Podolsky–Rosen steering *Phys. Rev.* A **99** 022317
[25] Gallego R, Würflinger L E, Chaves R, Acín A and Navascués M 2014 Nonlocality in sequential correlation scenarios *New J. Phys.* **16** 033037
[26] Curchod F J, Johansson M, Augusiak R, Hoban M J, Wittek P and Acín A 2017 Unbounded randomness certification using sequences of measurements *Phys. Rev.* A **95** 020102(R)
[27] Tavakoli A and Cabello A 2018 Quantum predictions for an unmeasured system cannot be simulated with a finite-memory classical system *Phys. Rev.* A **97** 032131
[28] Bera A, Mal S, Sen De A and Sen U 2018 Witnessing bipartite entanglement sequentially by multiple observers *Phys. Rev.* A **98** 062304
[29] Sasmal S, Das D, Mal S and Majumdar A S 2018 Steering a single system sequentially by multiple observers *Phys. Rev.* A **98** 012305
[30] Anwer H, Wilson N, Silva R, Muhammad S, Tavakoli A and Bourennane M 2019 *Noise-Robust Contextuality Shared Between Anynumber of Observers* arXiv:1904.09766
[31] Schiavon M, Calderaro L, Pittaluga M, Vallone G and Villoresi P 2017 Three-observer Bell inequality violation on a two-qubit entangled state *Quantum Sci. Technol.* **2** 015010
[32] Hu M-J, Zhou Z-Y, Hu X-M, Li C-F, Guo G-C and Zhang Y-S 2018 Observation of non-locality sharing among three observers with one entangled pair via optimal weak measurement *NPJ Quantum Inf.* **4** 63
[33] Foletto G, Calderaro L, Tavakoli A, Schiavon M, Picciariello F, Cabello A, Villoresi P and Vallone G 2019 arXiv:1906.07412 [quant-ph]
[34] Miklin N, Borkała J J and Pawłowski M 2019 *Self-Testing of Unsharp Measurements* arXiv:1903.12533

Measurement incompatibility and steering are necessary and sufficient for operational contextuality

Armin Tavakoli* and Roope Uola*
Département de Physique Appliquée, Université de Genève, CH-1211 Genève, Switzerland

(Received 7 June 2019; revised manuscript received 14 August 2019; published 6 January 2020)

Contextuality is a signature of operational nonclassicality in the outcome statistics of an experiment. This notion of nonclassicality applies to a breadth of physical phenomena. Here, we establish its relation to two fundamental nonclassical entities in quantum theory; measurement incompatibility and steering. We show that each is necessary and sufficient the failure of operational contextuality. We exploit the established connection to contextuality to provide a novel approach to problems in measurement incompatibility and steering.

DOI: 10.1103/PhysRevResearch.2.013011

I. INTRODUCTION

The nonclassical nature of quantum theory has a variety of different manifestations. On the one hand, quantum theory postulates theoretical entities with properties that lack a counterpart in classical physics. On the other hand, the nonclassicality of quantum theory is also present on the observable level, i.e., in the outcome statistics of experiments. Evidently, if an experiment takes the reality of the quantum formalism for granted, every nonclassical entity of quantum theory can be experimentally detected. However, if the assumption of nature being quantum is dropped, the outcome statistics can frequently be reproduced with some classical model. Matters become more interesting when the nonclassicality of outcome statistics can be operationally determined in the spirit of device independence [1], that is, in experiments that demonstrate nonclassicality while making weak assumptions on the underlying physical nature.

The strongest form of operational inference is encountered in tests of Bell inequalities [2]. These experiments statistically analyze the correlations between the outcomes of measurements performed in space-like separated events. If the correlations violate a Bell inequality, it follows that the outcome statistics cannot be explained by *any* classical (local hidden variable) theory. Famously, by sharing entangled states and performing incompatible measurements that together steer the remote partner system, quantum theory can violate these inequalities and therefore provide an unequivocal demonstration of nonclassicality [3,4]. Surprisingly however, not all incompatible measurements, nor all steerable ensembles, enable Bell inequality violations [5–7]. This motivates the question: Is nonclassicality at the level of theoretical entities both necessary and sufficient for some form of operational nonclassicality?

We focus on two fundamental theoretical entities in quantum theory: the incompatibility of quantum measurements and the ability to steer another system by local measurements and classical communication (a feature of quantum theory originating from Schrödinger's remarks [8] on the Einstein-Podolsky-Rosen paradox [9]). These two features of quantum theory have been thoroughly researched, see, e.g., Refs. [10–13] and Refs. [14,15] respectively. We show that both measurement incompatibility and steering admit a generally valid one-to-one connection with a family of physical tasks which in turn correspond to tests of operational *contextuality*.[1] Contextuality in quantum theory has for long been researched in its own interest and is closely related to, e.g., advantages in quantum computation [19–22], advantages in particular communication tasks [23–26], quantum zero-error communication [27], and anomalous weak values [28].

The established general connection between the theoretical entities of measurement incompatibility and steering on the one hand and operational contextuality on the other, enables us to approach relevant problems in the former ones using tools originally developed for the latter. We exploit this to present a family of noncontextuality inequalities and provide numerical evidence that these are necessary and sufficient conditions for certifying the incompatibility of any set of binary qubit observables, and that they also constitute optimal tests of the steerability of a pair of qubits in a singlet state subject to noisy environments. Moreover, since our task-oriented characterization of measurement incompatibility and steering makes reasonably weak assumptions on the characterization of physical devices, such applications make possible more stringent experimental certificates of all incompatible measurements and steerable states via experimental proofs of contextuality.

*These authors have contributed equally to this work.

[1]As originally introduced by Bell, Kochen, and Spekker [16,17], contextuality is a property of projective measurements in quantum theory. However, the concept has seen a generalization that applies on the level of ontological models, and therefore to general operational theories used to model outcome statistics [18].

II. CONTEXTUALITY

Contextuality in an operational theory [18] is developed within the framework of ontological models [29]. An ontological model describes a preparation procedure $\mathbf{P}$ by an ontic state (hidden variable) λ with a distribution $p(\lambda|\mathbf{P})$. When a measurement procedure $\mathbf{M}$ is applied, the ontological model determines the probability of an outcome b with some response function $p(b|\mathbf{M}, \lambda)$. Therefore, the outcome statistics reads

$$p(b|\mathbf{P}, \mathbf{M}) = \sum_{\lambda} p(\lambda|\mathbf{P})p(b|\mathbf{M}, \lambda). \tag{1}$$

Furthermore, ontological models are linear in the sense that convex combinations of preparation and measurement procedures are represented by convex sums of the relevant ontic state distributions and response functions respectively. See Ref. [18] for a discussion of this property.

In quantum theory, preparations are represented by density matrices ρ and measurements are positive operator-valued measures (POVMs) $M = \{M_b\}$, i.e., $M_b \geqslant 0$ and $\sum_b M_b = \mathbb{1}$. Outcome statistics is given by the Born rule $p(b|\rho, M) = \mathrm{tr}\,[\rho M_b]$. A state (measurement) can be realized in as many ways as it can be decomposed into mixtures of other states (decomposed into element-wise mixtures or coarse-graining of other measurements). Different ways of preparing the same state (performing the same measurement) are called contexts for ρ (M). An ontological model is said to be preparation noncontextual if the ontic state distribution is independent of the context, i.e., if $p(\lambda|\mathbf{P}) = p(\lambda|\rho)$. Similarly, an ontological model is said to be measurement noncontextual if the response functions are context independent, i.e., if $p(b|\mathbf{M}, \lambda) = p(b|M, \lambda)$. These notions embody the idea that if two laboratory procedures are indistinguishable, then they are also indistinguishable on the level of ontic states. We remark that to ensure that two procedures truly are indistinguishable, one needs to be able to perform measurements (prepare states) that span the measurement (state) space. In contrast, if outcome statistics cannot be reproduced with any preparation (measurement) noncontextual model, it is said to be preparation (measurement) contextual. See Ref. [18] for a detailed discussion of operational contextuality.

III. MEASUREMENT INCOMPATIBILITY

Measurement incompatibility [10–13] is the impossibility of jointly measuring a set of (at least two) POVMs by employing only a single measurement and classical postprocessing of its outcomes. More precisely, let $\{A_{a|x}\}$ be a set of POVMs, with a labeling the outcome and x labeling the measurement. The set is called compatible (jointly measurable) if there exists a POVM $\{G_\lambda\}$ which allows us to recover the set $\{A_{a|x}\}$ via some postprocessing probability distribution $p(a|x, \lambda)$:

$$A_{a|x} = \sum_{\lambda} p(a|x, \lambda)G_\lambda. \tag{2}$$

If such a model does not exist, the set $\{A_{a|x}\}$ is called incompatible (not jointly measurable). This extends the textbook concept of commutativity in the sense that mutually commuting POVMs are jointly measurable, but the converse does not hold in general. The converse holds, however, for textbook observables, i.e., projective measurements. It is worth noting that joint measurability can be characterized as the existence of a common Naimark dilation in which the projective measurements commute.

IV. STEERING

Steering [7] is a qualitative property of some entangled quantum states regarding the set of ensembles that can be remotely prepared with local measurements and classical communication. Specifically, one considers a pair of entangled systems in state ρ and performs a set of measurements $\{A_{a|x}\}$ on the first system. Given the choice of x, this renders the second system in the state $\rho_{a|x} = \mathrm{tr_A}\,[A_{a|x} \otimes \mathbb{1}\rho]/\,\mathrm{tr}\,[A_{a|x} \otimes \mathbb{1}\rho]$ with probability $p(a|x) = \mathrm{tr}\,[A_{a|x} \otimes \mathbb{1}\rho]$. It is important to underline the fact that classical communication is necessary for the steered party to be able to distinguish between different local states $\rho_{a|x}$. These local states can be effectively described with a set of unnormalized states (called an assemblage) $\{\sigma_{a|x}\}$ where $\sigma_{a|x} = \mathrm{tr_A}\,[A_{a|x} \otimes \mathbb{1}\rho]$. Such assemblages are no-signaling, i.e., $\sum_a \sigma_{a|x} = \sum_a \sigma_{a|x'}$. In this work all assemblages are assumed to be no-signaling. We remark that the Gisin-Hughston-Josza-Wootters theorem [30,31] ensures that every assemblage can be prepared by a distant party's local measurements (supported by classical communication) on a properly chosen entangled state. An assemblage is said to be unsteerable if it admits a so-called local hidden state model. Such models use (a, x) as information toward a postprocessing $p(a|x, \lambda)$ of a set of local states ρ_λ appearing with probability $p(\lambda)$ to explain the assemblage $\{\sigma_{a|x}\}$. Hence, if the state is unsteerable, it can be written as

$$\sigma_{a|x} = \sum_{\lambda} p(\lambda)p(a|x, \lambda)\rho_\lambda. \tag{3}$$

If no local hidden state model is possible, the assemblage is called steerable.

V. MAIN RESULTS

We begin by proving a one-to-one relation between measurement incompatibility and preparation contextuality.

Theorem 1. A set of measurements is compatible if and only if their statistics admit a preparation noncontextual model for all states.

Proof. Assume that the set of POVMs $\{A_{a|x}\}$ when applied to any quantum state ρ returns outcome statistics that is preparation noncontextual. We denote as $\mathcal{P}_\rho$ the set of preparation procedures (contexts) in which ρ can be prepared. Then, using the label x to denote the measurement procedure, it holds that

$$\forall\, \mathbf{P} \in \mathcal{P}_\rho : p(a|x, \mathbf{P}) = \sum_{\lambda} p(\lambda|\rho)p(a|x, \lambda). \tag{4}$$

For each λ, the object $p(\lambda|\rho)$ is a convexity-preserving map from the space of quantum states to the interval [0,1]. The Riesz representation theorem [13,32] asserts that such maps can be written as an inner product $p(\lambda|\rho) = \mathrm{tr}\,[G_\lambda \rho]$ for some unique operator $0 \leqslant G_\lambda \leqslant \mathbb{1}$. Moreover, since $\forall\ \rho$: $\sum_\lambda p(\lambda|\rho) = 1$, it follows that $\sum_\lambda G_\lambda = \mathbb{1}$. Inserting this into

Eq. (4), we have

$$\forall\, \mathbf{P} \in \mathcal{P}_\rho : p(a|x, \mathbf{P}) = \sum_\lambda p(a|x, \lambda)\, \mathrm{tr}\,[G_\lambda \rho]. \qquad (5)$$

We have recovered the outcome statistics obtained from measuring ρ with a compatible set of POVMs.

Conversely, assume that $\{A_{a|x}\}$ is a compatible set of POVMs. Then, the statistics obtained from measuring any state ρ prepared with a procedure $\mathbf{P}$ is given by Eq. (5). By defining $p(\lambda|\rho) = \mathrm{tr}\,[G_\lambda \rho]$, we recover the definition of outcome statistics being preparation noncontextual. ■

It is interesting to note that tests of preparation contextuality can be formulated as communication tasks between two separated parties, in which the receiver is kept oblivious about parts of the sender's input [23–25]. Such obliviousness corresponds to different contexts for the states. From Theorem 1, we can therefore infer the following corollary:

Corollary. Every set of incompatible measurements enables a quantum-over-classical advantage in an oblivious communication task.

We remark that the advantages of all incompatible sets of measurements have recently been shown in various measurement-device-independent communication tasks [33–36].

In a spirit similar to that of Theorem 1, we prove a one-to-one relation between steering and measurement contextuality.

Theorem 2. An assemblage is unsteerable if and only if its statistics admits a preparation and measurement noncontextual model for all measurements.

Proof. Assume that the assemblage $\{\sigma_{a|x}\}$ when measured with any POVM M returns outcome statistics that is measurement noncontextual. We denote the set of measurement procedures (contexts) in which M can be realized by $\mathcal{M}_M$. Due to assemblages being no-signaling, we have that $p(a, b|x, \mathbf{M}) = p(b|a, x, \mathbf{M})p(a|x)$ and that

$$\forall\, \mathbf{M} \in \mathcal{M}_M : p(b|a, x, \mathbf{M})p(a|x) = p(a|x) \sum_\lambda p(\lambda|a, x)p(b|M, \lambda), \qquad (6)$$

where (a, x) labels the preparation procedure. For every λ, the object $p(b|M, \lambda)$ is a map from the space of POVMs to the space of probability distributions. Such maps are characterized by the works of Gleason [37] and Busch [38]. The Gleason-Busch theorem asserts that $p(b|M, \lambda) = \mathrm{tr}\,[\rho_\lambda M_b]$ for some unique state ρ_λ. Inserting this into Eq. (6), we have

$$\forall\, \mathbf{M} \in \mathcal{M}_M : p(b|a, x, \mathbf{M})p(a|x) = p(a|x) \sum_\lambda p(\lambda|a, x)\, \mathrm{tr}\,[\rho_\lambda M_b]. \qquad (7)$$

Using Bayes' rule and the fact that x and λ are independent,[2] one straightforwardly finds that $p(a|x)p(\lambda|a, x) = p(\lambda)p(a|x, \lambda)$. Inserting this in (7), we recover the outcome statistics obtained from applying M to an unsteerable assemblage (3).

Conversely, if the assemblage has a local hidden state model, then for every POVM the outcome statistics reads

$$\forall\, \mathbf{M} \in \mathcal{M}_M : p(b|a, x, \mathbf{M}) = \frac{1}{p(a|x)} \sum_\lambda p(\lambda)p(a|x, \lambda)\, \mathrm{tr}\,[\rho_\lambda M_b]. \qquad (8)$$

From Bayes' rule and the independence of x and λ, we have that $p(\lambda)p(a|x, \lambda)/p(a|x) = p(\lambda|a, x)$. Note that said independence implies preparation noncontextuality. Inserted into Eq. (8) we find the outcome statistics obtained in a measurement noncontextual model. ■

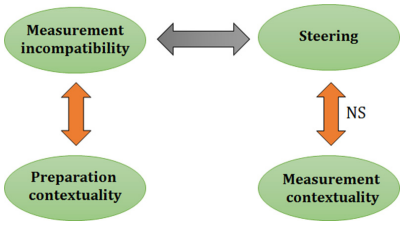

FIG. 1. Orange arrows illustrate Theorems 1 and 2. Measurement incompatibility and steering each enable a proof of a form of contextuality provided that one possesses a proper catalyst preparation or measurement procedure. Conversely, having observed preparation contextuality, one infers measurement incompatibility. Similarly, having observed measurement contextuality in a bipartite no-signaling scenario, one infers steerability of the shared state. The grey arrow indicates the previously known relation that a set of measurements is incompatible if and only if it enables steering with a proper catalyst state [39,40]. It is worth noting that this connection can also be seen as a mapping between the problems in measurement incompatibility and steerability [41].

We have illustrated the theorems in Fig. 1. Notice that Theorems 1 and 2 give a characterization of the ontic variables using quantum theory. Whereas this characterization is relevant for noncontextual models covering all states or measurements, it would be interesting to see whether such characterization exists in the case of fragments of quantum theory, i.e., for noncontextual models covering subsets of states and measurements.

Also, it is worth noting that a number of works have (in different ways) shown that outcome statistics that violate a Bell inequality is proof of preparation contextuality [18,23,24,42]. In Appendix A, we note that this fact follows immediately from ontological models and the no-signaling principle (see also Ref. [43] for similar results).[3]

VI. NONCONTEXTUALITY INEQUALITIES FOR QUBIT MEASUREMENT INCOMPATIBILITY AND STEERING

We proceed to use the established connection to contextuality to address two relevant problems in measurement

[2]The independence of x and λ follows from the fact that λ cannot carry information about the oblivious variable x (the obliviousness comes from no-signaling), i.e., the assumption of preparation noncontextuality. See, e.g., Refs. [23,25] for an elaboration.

[3]This result was shown originally in the unpublished note Ref. [44].

incompatibility and steering: (i) Can one find a preparation noncontextuality inequality whose violation is both necessary and sufficient for certifying measurement incompatibility for interesting families of measurements? Note that despite Theorem 1 this is a nontrivial matter since any *single* test of preparation contextuality is only a sufficient condition for measurement incompatibility (and the full characterization of all tests of preparation contextuality is a demanding problem). (ii) Can one find a measurement noncontextuality inequality (in a no-signaling scenario) that for an interesting class of states optimally certifies their steerability? In analogy to the previous, this is nontrivial since any single test of measurement contextuality in a no-signaling scenario is only a sufficient condition for steering.

To answer these questions, we present a family of correlation inequalities (parametrized by an integer $n \geqslant 2$) inspired by the works of Refs. [25,45,46]. Consider a Bell-like (no-signaling) experiment in which two separated observers, Alice and Bob, share parts of a physical system. Alice (Bob) performs measurements labeled by her (his) uniformly random input $x \in \{0,1\}^{n-1}$ ($y \in \{1,\ldots,n\}$). The outcome is denoted by $a \in \{0,1\}$ ($b \in \{0,1\}$). Alice's measurement procedures are constrained by operational equivalences. That is, her outcome statistics always upholds suitable indistinguishability relations which enable us to consider the statistics of different contexts for her measurements. Specifically, for every bit string $r \in \{0,1\}^n$, we require that the measurement procedures $\mathbf{M}_{r,0}$ and $\mathbf{M}_{r,1}$ corresponding to a uniform mixing of all (a,x) satisfying $r \cdot \bar{x} = 0$ and $r \cdot \bar{x} = 1$ respectively [where $\bar{x} = (a, x+a)$], are indistinguishable from each other. In quantum theory, this means that

$$\sum_{a,x|r\cdot\bar{x}=0} M_{a|x} = \sum_{a,x|r\cdot\bar{x}=1} M_{a|x}. \tag{9}$$

Note that whenever r has an even number of ones, this condition is always satisfied since $M_{0|x} + M_{1|x} = \mathbb{1}$. For odd strings r, Eq. (9) is a nontrivial constraint. Now, let Alice and Bob play a game in which they aim to maximize the probability of finding $a + b = \bar{x}_y \mod 2$. When Alice is considered the sender of Bob's remotely prepared local states, we can consider the scenario as a test of preparation contextuality. In contrast, when Bob is considered the sender of Alice's remotely prepared local states, we can consider the scenario as a test of preparation and measurement contextuality. In the case of either being noncontextual, the average success probability is bounded by

$$\mathcal{A}_n \equiv \frac{1}{n2^{n-1}} \sum_{x,y} p(a+b=\bar{x}_y|x,y) \leqslant \frac{n+1}{2n}. \tag{10}$$

The proof of this result is a simple modification of the arguments presented in Ref. [25] and is discussed in Appendix B. A violation of the inequality (10) means that Bob's measurements (which are unconstrained) are incompatible (by Theorem 1) and that Alice's local assemblage (prepared by Bob) is steerable (by Theorem 2). We now study the usefulness of the inequality (10) for certifying qubit measurement incompatibility and two-qubit steerability.

For the case of $n = 2$ the inequality (10) reduces to the Clauser-Horne-Shimony-Holt Bell inequality [4] for which it is known that all pairs of incompatible measurements enable a violation [47]. For $n > 2$ (specifically studying $n = 3,\ldots,7$) we have numerically obtained support (10 000 examples for each n) for the following conjecture:

Conjecture 1. Every set of n incompatible two-outcome qubit measurements enable a proof of preparation contextuality by a violation of the inequality (10).

In Appendix C, we describe the numerical procedure employed to motivate this conjecture.

Consider now the case of steering. For simplicity, let Alice and Bob share the noisy singlet state $\rho_v = v|\psi^-\rangle\langle\psi^-| + (1-v)\mathbb{1}/2$, where $|\psi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$ for some visibility $v \in [0,1]$. What is the critical value of $v = v_n$ so that Bob can steer Alice using n projective measurements? Although this question is well studied (see, e.g., Refs. [48–50]) an analytical formula is lacking. However, Ref. [48] presented nearly matching upper and lower bounds on v_n for $n = 2,\ldots,13$ and $n = 2,\ldots,5$ respectively. Using our inequality (10), we have numerically implemented alternating convex searches to find an upper bound on the critical v_n (below which we can no longer find a quantum violation). This returns

$$\begin{aligned} &v_2 = 0.7071, \quad v_3 = 0.5774, \quad v_4 = 0.5547, \\ &v_5 = 0.5422, \quad v_6 = 0.5270, \quad v_7 = 0.5234. \end{aligned} \tag{11}$$

Interestingly, these numbers coincide precisely with those presented in Ref. [48] (up to the number of decimals presented in Ref. [48]). This motivates the conjecture

Conjecture 2. The inequality (10) is a tight steering inequality for the noisy singlet state under n projective measurements.

Finding a conclusive proof of Conjectures 1 and 2 would be interesting. We remark that although the above considerations are straightforwardly analyzed with a computer, the criterion (10) can be treated in a fully analytical manner.

VII. DISCUSSION

We have shown that every set of incompatible measurements and every steerable assemblage can be operationally certified as nonclassical in a test of operational contextuality, and that the latter also implies the formers. A direct consequence is that problems of joint measurability and steering can be viewed through the lens of contextuality, as we illustrated through our conjectures. In this sense, our results bridge the two research directions of quantum measurements and quantum steering with the line of research focused on quantum contextuality.

Moreover, since tests of operational contextuality only rely on weak characterization of the experimental devices [51], our results can also be considered as semi-device-independent certificates of measurement incompatibility and steering. Naturally, fully device-independent certificates are found by violating a Bell inequality. However, in addition to such tests being experimentally demanding, it is importantly also the case that not all incompatible measurements nor all steerable ensembles violate any Bell inequality [5–7]. This makes tests of operational contextuality relevant for practical considerations when no fully device-independent certificate is either possible or known.

ACKNOWLEDGMENTS

We thank Alastair Abbott, Costantino Budroni, Matthew Pusey, Tom Bullock, Chau Nguyen, Marco Quintino, and Leonardo Guerini for comments. This work was supported by the Swiss National Science Foundation (Starting grant DIAQ, NCCR-QSIT) and by the Finnish Cultural Foundation.

APPENDIX A: BELL NONLOCALITY IMPLIES PREPARATION CONTEXTUALITY

We give a simple argument for every probability distribution that violates a Bell inequality also being a proof of preparation contextuality (see also Ref. [43]). We show this immediately from ontological models supplemented with the no-signaling principle encountered in Bell inequality tests.

To see this, we write a general ontological model for a Bell experiment as

$$p(a, b|x, y) = \sum_{\lambda} p(a|x, y)p(\lambda|a, x)p(b|y, \lambda). \tag{A1}$$

If we also impose no-signaling, then Alice's local marginals are independent of Bob's input. Therefore,

$$p(a, b|x, y) = \sum_{\lambda} p(a|x)p(\lambda|a, x)p(b|y, \lambda). \tag{A2}$$

Bayes' rule together with the independence of x and λ gives that $p(a|x)p(\lambda|a, x) = p(\lambda)p(a|x, \lambda)$. Inserting this into Eq. (A2), we obtain

$$p(a, b|x, y) = \sum_{\lambda} p(\lambda)p(a|x, \lambda)p(b|y, \lambda). \tag{A3}$$

This is a local hidden variable model, i.e., the notion of classicality in Bell inequality tests. The assumption of preparation noncontextuality is enforced due to the assignment of the same ontic-state distribution for the preparation procedures corresponding to the remotely prepared state on Bob's side when averaged over Alice's outcomes, i.e., the principle of no-signaling. Therefore, whenever $p(a, b|x, y)$ has no local hidden variable model, it also has no preparation noncontextual model.

APPENDIX B: NONCONTEXTUAL BOUND

In the main text, we considered a scenario in which separated parties Alice and Bob share a state and perform local measurements with binary outcomes $a, b \in \{0, 1\}$. Alice's measurement settings are labeled by a bit string $x \in \{0, 1\}^{n-1}$ and Bob's measurement settings are labeled by $y \in \{1, \ldots, n\}$. Alice and Bob aim to satisfy the relation $a + b = \bar{x}_y \mod 2$ where $\bar{x} = (a, a + x)$ is an n-bit string. The notation $\bar{x}_y$ labels the yth bit in the string $\bar{x}$. Their average success probability is

$$\mathcal{A}_n \equiv \frac{1}{n2^{n-1}} \sum_{x,y} p(a + b = \bar{x}_y|x, y). \tag{B1}$$

Alice and Bob are restricted by two constraints. First, they obey the no-signaling principle. This means that the preparations of Alice on Bob's side (denoted $\mathbf{P}_{a,x}$), effectively achieved by a local measurement on her system, realize the same preparation in different contexts. That is, the following operational equivalences hold: $\sum_a \mathbf{P}_{a,x} \sim \sum_a \mathbf{P}_{a,x'}$. The analogy holds in the other direction, i.e., by the preparations of Bob on Alice's side achieved by him locally measuring his system. Second, Alice's measurements are required to uphold certain operational equivalences. In quantum theory, these are written

$$\sum_{a,x|r\cdot\bar{x}=0} M_{a|x} = \sum_{a,x|r\cdot\bar{x}=1} M_{a|x}, \tag{B2}$$

for every n-bit string $r \in \{0, 1\}^n$ with at least two instances of $'1'$. For clarity, we give as an example the case of $n = 3$. There exists eight three-bit strings of which four have at least two instances of $'1'$. Those are $r = 011$, $r = 101$, $r = 110$, and $r = 111$. For each r we have the relation in Eq. (B2). In the case of, for example, $r = 011$ we find

$$\begin{aligned} &M_{0|00} + M_{1|00} + M_{0|11} + M_{1|11} \\ &\quad = M_{0|01} + M_{1|01} + M_{0|10} + M_{1|10}. \end{aligned} \tag{B3}$$

However, this is trivially satisfied since $\forall\, x : M_{0|x} + M_{1|x} = \mathbb{1}$. Similarly, one finds that the constraint (B2) is trivial also for $r = 101$ and $r = 110$. However, for $r = 111$ we obtain

$$\begin{aligned} &M_{0|00} + M_{0|11} + M_{1|01} + M_{1|10} \\ &\quad = M_{0|01} + M_{0|10} + M_{1|00} + M_{1|11}, \end{aligned} \tag{B4}$$

which is a nontrivial constraint.

Imagine now that instead of performing local measurements on a shared state, Alice directly prepares the would-have-been post-measurement states of Bob's system [labeled by the pair (a, x)] and sends them to Bob, who then measures the system and records $b \in \{0, 1\}$. This represents a prepare-and-measure scenario in which Alice has 2^n inputs (a, x) with some prior distribution $p(a, x) = p(a|x)/2^{n-1}$. Alice's preparations are required to satisfy the operational equivalence which in quantum theory reads

$$\forall\, r : \quad \sum_{a,x|r\cdot\bar{x}=0} p(a|x)\rho_{a,x} = \sum_{a,x|r\cdot\bar{x}=1} p(a|x)\rho_{a,x}. \tag{B5}$$

Notice first that in the original scenario, every assemblage prepared by Alice on Bob's side can also be directly sent in this prepare-and-measure model; simply define $\rho_{a,x} = \mathrm{tr}_A\,[M_{a|x} \otimes \mathbb{1}\rho]/\,\mathrm{tr}\,[M_{a|x} \otimes \mathbb{1}\rho]$, and the prior distribution as $p(a|x) = \mathrm{tr}\,[M_{a|x} \otimes \mathbb{1}\rho]$. Conversely, every ensemble that Alice can communicate to Bob in the prepare-and-measure scenario can also be realized in the original scenario via local measurements on an entangled state and classical communication. This follows from the Gisin-Hughston-Josza-Wootters theorem [30,31] and the fact that Eq. (B5) enforces a no-signaling-like preparation ensemble.

In Ref. [25] it was shown that when $p(a|x) = 1/2$, the considered prepare-and-measure scenario serves as the following test of preparation contextuality: the inequality

$$\frac{1}{n2^{n-1}} \sum_{a,x,y} p(a|x)p(b = (a, x)_y|a, x, y) \leqslant \frac{n + 1}{2n} \tag{B6}$$

holds for every preparation noncontextual model. Moreover, it is a trivial modification of the arguments of Ref. [25] to show that the same bound holds regardless of the prior distribution

$p(a|x)$. Therefore, due to the connection between the prepare-and-measure scenario and the original scenario, it also holds that

$$\mathcal{A}_n \leqslant \frac{n+1}{2n} \tag{B7}$$

in a preparation noncontextual model, in which we view Alice as effectively preparing the local states of Bob.

Moreover, in the original scenario, we can equally well consider Bob as the effective sender of Alice's local states. If we impose measurement noncontextuality, the response function of Alice takes no regard of the different contexts of her measurement (related to r). Note that preparation noncontextuality is still present due to Alice and Bob being no-signaling.

APPENDIX C: NUMERICAL EVIDENCE IN SUPPORT OF CONJECTURE 1

The numerical evidence behind Conjecture 1 was obtained as follows. We used the prepare-and-measure variant (discussed in Appendix B, based on Ref. [25]) for the numerics. We sample a set of n random two-outcome qubit POVMs $\mathcal{M} = \{B_{b|y}\}_{y=1}^{n}$. The sampling is done by using the Bloch sphere parametrization of the most general two-outcome qubit measurement, i.e.,

$$\forall\, y : B_{0|y} = \frac{\alpha_y \mathbb{1} + \eta_y \vec{n}_y \cdot \vec{\sigma}}{2}, \tag{C1}$$

$$B_{1|y} = \frac{(2-\alpha_y)\mathbb{1} - \eta_y \vec{n}_y \cdot \vec{\sigma}}{2} \tag{C2}$$

for some random unit vectors $\vec{n}_y$, some random numbers $\eta_y \in [0, 1]$, and some random numbers $\eta_y \leqslant \alpha_y \leqslant 2 - \eta_y$.

For the sampled $\mathcal{M}$, we evaluate the largest possible value of the witness $\mathcal{A}_n$ via a semidefinite program optimizing over the state ensemble of Alice. This returns the maximal value of $\mathcal{A}_n(\mathcal{M})$ attainable with $\mathcal{M}$. We denote the optimal ensemble returned by the semidefinite program by $\mathcal{P}$. Provided that $\mathcal{A}_n(\mathcal{M})$ violates the noncontextuality inequality (in its prepare-and-measure form), we construct new measurements $B'_{b|y} = vB_{b|y} + (1-v)\mathbb{1}/2$ where $v \in [0, 1]$. We write $\mathcal{M}' = \{B'_{b|y}\}$. For the states $\mathcal{P}$ we have that

$$\mathcal{A}_n(\mathcal{M}', \mathcal{P}) = v\mathcal{A}_n(\mathcal{M}) + (1-v)\mathcal{A}_n(\{\mathbb{1}/2\}, \mathcal{P}). \tag{C3}$$

We choose the value of v for which $\mathcal{A}_n(\mathcal{M}', \mathcal{P})$ saturates the noncontextual bound, i.e.,

$$v = \frac{\mathcal{C}_n - \mathcal{A}_n(\{\mathbb{1}/2\}, \mathcal{P})}{\mathcal{A}_n(\mathcal{M}) - \mathcal{A}_n(\{\mathbb{1}/2\}, \mathcal{P})}, \tag{C4}$$

where $\mathcal{C}_n = (n+1)/(2n)$ is the noncontextual bound. Then, via a semidefinite program, we check whether $\mathcal{M}'$ is jointly measurable. Evidently, any perturbation of v to the positive renders $\mathcal{M}'$ incompatible since it implies a violation of the preparation noncontextuality inequality. We have repeated the procedure 10 000 times]postselected on the cases in which $\mathcal{A}_n(\mathcal{M})$ constitutes a proof of preparation contextuality] for $n = 3, 4, 5, 6, 7$ respectively. Without exception, we have found that $\mathcal{M}'$ is jointly measurable.

[1] S. Pironio, V. Scarani, and T. Vidick, Focus on device independent quantum information, New J. Phys. **18**, 100202 (2016).

[2] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).

[3] J. S. Bell, On the Einstein Podolsky Rosen paradox, Physics **1**, 195 (1964).

[4] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, Phys. Rev. Lett. **23**, 880 (1969).

[5] E. Bene and T. Vértesi, Measurement incompatibility does not give rise to Bell violation in general, New J. Phys. **20**, 013021 (2018).

[6] F. Hirsch, M. T. Quintino, and N. Brunner, Quantum measurement incompatibility does not imply Bell nonlocality, Phys. Rev. A **97**, 012129 (2018).

[7] H. M. Wiseman, S. J. Jones, and A. C. Doherty, Steering, Entanglement, Nonlocality, and the Einstein-Podolsky-Rosen Paradox, Phys. Rev. Lett. **98**, 140402 (2007).

[8] E. Schrödinger, Discussion of probability relations between separated systems, Proc. Cambridge Philos. Soc. **31**, 555 (1935).

[9] A. Einstein, B. Podolsky, and N. Rosen, Can quantum-mechanical description of physical reality be considered complete? Phys. Rev. **47**, 777 (1935).

[10] T. Heinosaari, T. Miyadera, and M. Ziman, An invitation to quantum incompatibility, J. Phys. A: Math. Theor. **49**, 123001 (2016).

[11] P. Busch, P. J. Lahti, and P. Mittelstaedt, *The Quantum Theory of Measurement* (Springer, New York, 1996).

[12] P. Busch, M. Grabowski, and P. J. Lahti, *Operational Quantum Physics* (Springer, New York, 1995).

[13] P. Busch, P. Lahti, J-P. Pellonpää, and K. Ylinen, *Quantum Measurement* (Springer, New York, 2016).

[14] D. Cavalcanti and P. Skrzypczyk, Quantum steering: A review with focus on semidefinite programming, Rep. Prog. Phys. **80**, 024001 (2017).

[15] R. Uola, A. C. S. Costa, H. C. Nguyen, and O. Gühne, Quantum steering, arXiv:1903.06663.

[16] J. S. Bell, On the problem of hidden variables in quantum mechanics, Rev. Mod. Phys. **38**, 447 (1966).

[17] S. Kochen and E. P. Specker, The problem of hidden variables in quantum mechanics, J. Math. Mech. **17**, 59 (1967).

[18] R. W. Spekkens, Contextuality for preparations, transformations, and unsharp measurements, Phys. Rev. A **71**, 052108 (2005).

[19] R. Raussendorf, Contextuality in measurement-based quantum computation, Phys. Rev. A **88**, 022322 (2013).

[20] M. Howard, J. Wallman, V. Veitch, and J. Emerson, Contextuality supplies the 'magic' for quantum computation, Nature **510**, 351 (2014).

[21] N. Delfosse, P. A. Guerin, J. Bian, and R. Raussendorf, Wigner Function Negativity and Contextuality in Quantum Computation on Rebits, Phys. Rev. X **5**, 021003 (2015).

[22] J. Bermejo-Vega, N. Delfosse, D. E. Browne, C. Okay, and R. Raussendorf, Contextuality as a Resource for Models of Quantum Computation with Qubits, Phys. Rev. Lett. **119**, 120505 (2017).
[23] A. Hameedi, A. Tavakoli, B. Marques, and M. Bourennane, Communication Games Reveal Preparation Contextuality, Phys. Rev. Lett. **119**, 220402 (2017).
[24] D. Saha, and A. Chaturvedi, Preparation contextuality: The ground of quantum communication advantage? Phys. Rev. A **100**, 022108 (2019).
[25] R. W. Spekkens, D. H. Buzacott, A. J. Keehn, B. Toner, and G. J. Pryde, Preparation Contextuality Powers Parity-Oblivious Multiplexing, Phys. Rev. Lett. **102**, 010401 (2009).
[26] D. Schmid and R. W. Spekkens, Contextual Advantage for State Discrimination, Phys. Rev. X **8**, 011015 (2018).
[27] T. S. Cubitt, D. Leung, W. Matthews, and A. Winter, Improving Zero-Error Classical Communication with Entanglement, Phys. Rev. Lett. **104**, 230503 (2010).
[28] M. F. Pusey, Anomalous Weak Values Are Proofs of Contextuality, Phys. Rev. Lett. **113**, 200401 (2014).
[29] N. Harrigan and R. W. Spekkens, Einstein, incompleteness, and the epistemic view of quantum states, Found. Phys. **40**, 125 (2010).
[30] N. Gisin, Stochastic quantum dynamics and relativity, Helv. Phys. Acta **62**, 363 (1989).
[31] L. P. Hughston, R. Jozsa, and W. K. Wootters, A complete classification of quantum ensembles having a given density matrix, Phys. Lett. A **183**, 14 (1993).
[32] T. Heinosaari and M. Ziman, *The Mathematical Language of Quantum Theory* (Cambridge University, New York, 2011).
[33] P. Skrzypczyk, I. Šupić, and D. Cavalcanti, All Sets of Incompatible Measurements give an Advantage in Quantum State Discrimination, Phys. Rev. Lett. **122**, 130403 (2019).
[34] R. Uola, T. Kraft, J. Shang, X-D. Yu, and O. Gühne, Quantifying Quantum Resources with Conic Programming, Phys. Rev. Lett. **122**, 130404 (2019).
[35] C. Carmeli, T. Heinosaari, and A. Toigo, Quantum Incompatibility Witnesses, Phys. Rev. Lett. **122**, 130402 (2019).
[36] L. Guerini, M. T. Quintino, and L. Aolita, Distributed sampling, quantum communication witnesses, and measurement incompatibility, Phys. Rev. A **100**, 042308 (2019).
[37] A. M. Gleason, Measures on the closed subspaces of a Hilbert space, J. Math. Mech. **6**, 885 (1957).
[38] P. Busch, Quantum States and Generalized Observables: A Simple Proof of Gleason's Theorem, Phys. Rev. Lett. **91**, 120403 (2003).
[39] M. T. Quintino, T. Vértesi, and N. Brunner, Joint Measurability, Einstein-Podolsky-Rosen Steering, and Bell Nonlocality, Phys. Rev. Lett. **113**, 160402 (2014).
[40] R. Uola, T. Moroder, and O. Gühne, Joint Measurability of Generalized Measurements Implies Classicality, Phys. Rev. Lett. **113**, 160403 (2014).
[41] R. Uola, C. Budroni, O. Gühne, and J-P. Pellonpää, One-to-One Mapping between Steering and Joint Measurability Problems, Phys. Rev. Lett. **115**, 230402 (2015).
[42] M. S. Leifer and O. J. E. Maroney, Maximally Epistemic Interpretations of the Quantum State and Contextuality, Phys. Rev. Lett. **110**, 120401 (2013).
[43] M. F. Pusey, Robust preparation noncontextuality inequalities in the simplest scenario, Phys. Rev. A **98**, 022112 (2018).
[44] J. Barrett (unpublished).
[45] A. Chailloux, I. Kerenidis, S. Kundu, and J. Sikora, Optimal bounds for parity-oblivious random access codes, New J. Phys. **18**, 045003 (2016).
[46] A. Tavakoli, B. Marques, M. Pawłowski, and M. Bourennane, Spatial versus sequential correlations for random access coding, Phys. Rev. A **93**, 032336 (2016).
[47] M. M. Wolf, D. Perez-Garcia, and C. Fernandez, Measurements Incompatible in Quantum Theory Cannot Be Measured Jointly in Any Other No-Signaling Theory, Phys. Rev. Lett. **103**, 230402 (2009).
[48] J. Bavaresco, M. T. Quintino, L. Guerini, T. O. Maciel, D. Cavalcanti, and M. T. Cunha, Most incompatible measurements for robust steering tests, Phys. Rev. A **96**, 022110 (2017).
[49] R. Uola, K. Luoma, T. Moroder, and T. Heinosaari, Adaptive strategy for joint measurements, Phys. Rev. A **94**, 022109 (2016).
[50] S. J. Jones and H. M. Wiseman, Nonlocality of a single photon: Paths to an Einstein-Podolsky-Rosen-steering experiment, Phys. Rev. A **84**, 012110 (2010).
[51] M. D. Mazurek, M. F. Pusey, R. Kunjwal, K. J. Resch, and R. W. Spekkens, An experimental test of noncontextuality without unphysical idealizations, Nat. Commun. **7**, 11780 (2016).

Autonomous multipartite entanglement engines

Armin Tavakoli,[1] Géraldine Haack,[1] Nicolas Brunner,[1] and Jonatan Bohr Brask[1,2]
[1]*Department of Applied Physics, University of Geneva, 1211 Geneva, Switzerland*
[2]*Department of Physics, Technical University of Denmark, Fysikvej, 2800 Kongens Lyngby, Denmark*

(Received 11 June 2019; published 13 January 2020)

The generation of genuine multipartite entangled states is challenging in practice. Here we explore an alternative route to this task, via autonomous entanglement engines which use only incoherent coupling to thermal baths and time-independent interactions. We present a general machine architecture, which allows for the generation of a broad range of multipartite entangled states in a heralded manner. Specifically, given a target multiple-qubit state, we give a sufficient condition ensuring that it can be generated by our machine. We discuss the cases of Greenberger-Horne-Zeilinger, Dicke, and cluster states in detail. These results demonstrate the potential of purely thermal resources for creating multipartite entangled states useful for quantum information processing.

DOI: 10.1103/PhysRevA.101.012315

I. INTRODUCTION

Quantum thermal machines combine quantum systems with thermal reservoirs at different temperatures and exploit the resulting heat flows to perform useful tasks. These can be work extraction or cooling, in analogy with classical heat engines and refrigerators, but may also be of a genuinely quantum nature. In particular, it is possible to devise entanglement engines—thermal machines generating entangled quantum states. Entanglement is a key resource for quantum information processing but is generally very fragile and easily destroyed by environmental noise. It is nevertheless possible to exploit dissipation to create and stabilize entanglement [1–13]. This was studied in a variety of settings and physical systems [14–24] and dissipative entanglement generation using continuous driving was experimentally demonstrated, mainly for bipartite states [25–28].

Autonomous entanglement engines represent a particularly simple case. Here, entanglement can be generated dissipatively with minimal resources, using only time-independent interactions and contact to thermal reservoirs at different temperatures. No driving, coherent control, or work input is required. For the bipartite case, a two-qubit entangled state can be generated in a steady-state, out-of-thermal-equilibrium regime [29]. Although the entanglement produced by such machines is typically weak, it can be boosted via entanglement distillation [30], or by coupling to negative-temperature [31] or joint baths [32]. In fact, applying a local filtering operation to the steady state of a bipartite entanglement engine can herald maximal entanglement between two systems of arbitrary dimension [33].

These first results show that using dissipative, out-of-equilibrium thermal resources offers an interesting perspective on entanglement generation. A natural question is whether this setting could also be used to generate more complex forms of entanglement, in particular entanglement between a large number of subsystems. It is of fundamental interest to understand the possibilities and limits of thermal entanglement generation. In addition, such multipartite entangled states represent key resources, e.g., for measurement-based quantum computation, quantum communications, and quantum-enhanced sensing and metrology. The creation and manipulation of complex entangled states is therefore of strong interest for many experimental platforms, although typically very challenging in practice.

Here, we propose autonomous entanglement engines as an alternative route to the generation of multipartite entanglement and explore their potential. A first question concerns which types of multipartite entangled states can be created. We present a sufficient condition for a given target N-qubit state to be obtainable. Specifically, for any target state satisfying our criterion, we construct an autonomous entanglement engine that will generate this state. The engine consists of N interacting qutrits (three-level systems), each qutrit being locally connected to a thermal bath. From the resulting steady state, a local filtering operation then leads to the desired target state. In particular, our scheme can generate important classes of genuine multipartite entangled states, including Greenberger-Horne-Zeilinger (GHZ), Dicke, and cluster states, which we discuss in detail. We show that these states can be generated with high fidelities and good heralding probabilities.

II. ENTANGLEMENT ENGINE

We begin by describing the entanglement engine. The structure of the machine is determined by the choice of subspace, energy spectrum, and bath temperature for each qutrit, as well as the form of the interaction, all of which generally depend on the N-qubit target state $|\psi\rangle$. This state is obtained in a heralded manner from the steady state of the machine by projection of each qutrit to a qubit subspace. Figure 1 shows an example targeting a GHZ state.

The machine evolution consists of a Hamiltonian contribution and a dissipative contribution due to the heat baths. The evolution is autonomous in the sense that both the

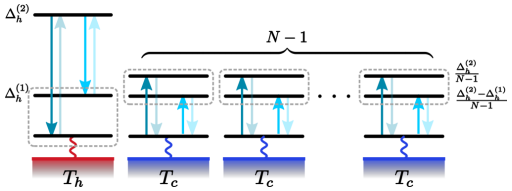

FIG. 1. Autonomous thermal machine for the generation of N-qubit GHZ states. One qutrit is coupled to a hot thermal bath, while $N-1$ qutrits are coupled to cold thermal baths at equal temperatures. The energy-level structure is such that transitions in the hot qutrit are resonant with collective transitions of the cold qutrits, as indicated by arrows. All the cold systems have the same structure, i.e., $\Delta_k^{(1)} = \Delta_c^{(1)}$ and $\Delta_k^{(2)} = \Delta_c^{(2)}$ for $k = 2, \ldots, N$, and $\Delta_c^{(1)} = (\Delta_h^{(2)} - \Delta_h^{(1)})/(N-1)$ and $\Delta_c^{(2)} = \Delta_h^{(2)}/(N-1)$. Local filters project the qutrits onto the qubit subspaces enclosed in dashed, gray boxes.

Hamiltonians and the bath couplings are time independent, and the machine thus requires no work input to run. Denoting the energy basis states of qutrit k by $\{|0\rangle_k, |1\rangle_k, |2\rangle_k\}$ and taking the corresponding energies to be $\{0, \Delta_k^{(1)}, \Delta_k^{(2)}\}$, the free Hamiltonian of each qutrit is $H_k = \Delta_k^{(1)}|1\rangle_k\langle 1| + \Delta_k^{(2)}|2\rangle_k\langle 2|$. The free Hamiltonian of the machine is

$$H_{\text{free}} = \sum_{k=1}^{N} H_k = \sum_{k=1}^{N}\left(\sum_{l=1}^{2}\Delta_k^{(l)}|l\rangle_k\langle l|\right). \tag{1}$$

In addition, the qutrits interact via a time-independent Hamiltonian H_{int}, specified below.

We model the machine evolution including the heat-bath induced dissipation with a master equation of the form

$$\frac{d\rho}{dt} = -i[H_{\text{free}} + H_{\text{int}}, \rho] + \mathcal{L}(\rho). \tag{2}$$

For simplicity, we adopt a local reset model in which the dissipator $\mathcal{L}$ corresponds to spontaneous, probabilistic, independent resets of each qutrit to a thermal state at the corresponding temperature [8,34]. That is,

$$\mathcal{L}(\rho) = \mathcal{L}_k(\rho) = \sum_{k=1}^{N}\gamma_k[\tau_k \otimes_k \text{Tr}_k(\rho) - \rho], \tag{3}$$

where γ_k is the reset rate for qutrit k, $\tau_k = \exp(-H_k/T_k)/\text{Tr}[\exp(-H_k/T_k)]$ is a thermal state of qutrit k, and $\otimes_k$ denotes tensoring at position k. For such a Markovian master equation description to be valid, the system-bath couplings γ_k must be small relative to the system energy scale $\Delta_k^{(l)}$. In addition, each dissipator acts only on the corresponding qutrit, i.e., they are local. This requires that the strength of the interaction between the qutrits is at most comparable to the bath couplings γ_k [35,36]. We note that the reset model can be mapped to a standard Lindblad-type model which can be derived from a microscopic, physical model of the baths [33].

The goal of the machine is to produce the N-qubit target state by local filtering of the N-qutrit steady state of (2). The steady state ρ_∞ is obtained by solving $d\rho/dt = 0$, and the filter is defined by a local projection $\Pi_k = \mathbb{1} - |R_k\rangle\langle R_k|$ of each qutrit onto the chosen qubit subspace. The state of the machine after filtering and the probability for the filtering to succeed are given by

$$\rho' = \frac{\Pi\rho_\infty\Pi}{\text{Tr}(\rho_\infty\Pi)}, \quad p_{\text{suc}} = \text{Tr}(\rho_\infty\Pi), \tag{4}$$

where $\Pi = \bigotimes_{k=1}^{N}\Pi_k$. The temperatures, filters, bath couplings γ_k, and interaction must be chosen appropriately for the heralded state ρ' to approach the target state.

Here, for a given N-qubit target $|\psi\rangle$, we focus on the following choice for the interaction

$$H_{\text{int}} = g(|\bar{\psi}\rangle\langle R| + |R\rangle\langle\bar{\psi}|), \tag{5}$$

where $g > 0$ is the interaction strength, and the states $|\bar{\psi}\rangle$ and $|R\rangle$ are defined by the choices of filtered qubit subspace for each qutrit. For qutrit k, we let $R_k = 0, 1, 2$ label the level which is *not* part of the qubit, i.e., qubit k is spanned by the two levels complementary to $|R_k\rangle$. Then $|\bar{\psi}\rangle$ is the embedding of the target $|\psi\rangle$ into these qubit subspaces, and $|R\rangle = |R_1 \ldots R_N\rangle$. That is, H_{int} swaps the target state and the state in which every qutrit is outside the filtered subspace.

We furthermore focus on the regime of weak intersystem coupling, where g is small relative to the free energies $\Delta_k^{(l)}$ (where the local master equation is valid). For there to be any nontrivial evolution in this regime, the interaction needs to be energy conserving, i.e., $[H_{\text{int}}, H_{\text{free}}] = 0$. This restricts which target states can be generated. However, that is the only restriction. Our main result is that

> any state $|\psi\rangle$, for which the Hamiltonians H_{free} and H_{int} of Eqs. (1) and (5) can be constructed to satisfy $[H_{\text{int}}, H_{\text{free}}] = 0$, can be generated by an entanglement engine as described above.

Specifically, one may choose a single qutrit to be connected with coupling strength γ_h to a hot bath at temperature T_h and all other qubits to be connected with coupling strength γ_c to cold baths at T_c. For the hot qutrit, one chooses $R_k = 2$, while for all the cold qutrits $R_k = 0$. The target $|\psi\rangle$ is then obtained in the limit of extremal temperatures $T_c = 0$, $T_h \to \infty$, and small coupling-strength ratios $g \lesssim \gamma_h \ll \gamma_c$. A full proof is given in Appendix A. However, one can intuitively understand why the machine works well in this regime. When $T_c = 0$, resets of the cold qutrits will take them to the ground state $|0\rangle_k$. Since for the cold qutrits $R_k = 0$, the ground state is not part of the filtered subspace. Therefore, cold resets will only lower the filtering success probability but will not affect the overlap of the filtered state with the target state $|\psi\rangle$. Once a cold qutrit is in the ground state, the only process which can bring it back into the filtered subspace is H_{int}, and this can only happen once all qutrits are in the state $|R_k\rangle$. The hot qutrit must then be in state $|2\rangle$, which can happen via a hot reset. Hot resets also degrade the quality of the filtered state, and hence must be much less frequent than cold reset. This way, the system is most likely to be found outside the filtered subspace (making p_{suc} small), but if found inside it is likely to be in state $|\psi\rangle$ (because it is unlikely a hot reset happens before a cold one drives the system back out).

We note that, even if a given target $|\psi\rangle$ does not admit any choice of H_{free} and H_{int} satisfying $[H_{\text{int}}, H_{\text{free}}] = 0$, it may happen that by applying local unitaries to each qubit one can obtain another state $|\psi'\rangle$ which does. Since entanglement is

preserved under local unitaries, one may then first generate $|\psi'\rangle$ and simply apply the inverse local unitaries to obtain $|\psi\rangle$. Thus, effectively, the set of states which can be generated using the entanglement engine above consists of all states within the local unitary orbit of those $|\psi\rangle$ for which energy conservation can be satisfied.

III. ENERGY CONSERVATION

We now derive conditions for $|\psi\rangle$ to admit choices of $H_{\rm free}$ and $H_{\rm int}$ such that $[H_{\rm int}, H_{\rm free}] = 0$. This holds if and only if every transition generated by $H_{\rm int}$ is energy conserving with respect to $H_{\rm free}$. From (5), these transitions depend on the target state and on the choice of $|R\rangle$. We can write the target N-qubit state as $|\psi\rangle = \sum_{\mathbf{n}\in S_\psi} c_{\mathbf{n}}|\mathbf{n}\rangle$ where $S_\psi = \{\mathbf{n} \in \{0,1\}^N \,|\, \langle\psi|\mathbf{n}\rangle \neq 0\}$ determines the set of basis states on which $|\psi\rangle$ has support, and $c_{\mathbf{n}} \in \mathbb{C}$. Denoting the embedding of $|\mathbf{n}\rangle$ into the N qutrits by $|\bar{\mathbf{n}}\rangle$, both $|\bar{\mathbf{n}}\rangle$ and $|R\rangle$ are eigenstates of $H_{\rm free}$ with respective eigenvalues $E_{\bar{\mathbf{n}}}$ and E_R. The conditions for energy conservation are then $E_{\bar{\mathbf{n}}} = E_R$ for every $\mathbf{n} \in S_\psi$. This can be expressed as

$$\frac{1}{2}\sum_{k=1}^{N}\left\{R_k n_k \Delta_k^{(1)} + (2-R_k)\left[(1-n_k)\Delta_k^{(1)} + n_k\Delta_k^{(2)}\right]\right\} - \frac{1}{2}\sum_{k=1}^{N}\left[R_k\Delta_k^{(2)}\right] = 0, \tag{6}$$

where we have restricted ourselves to cases where the qubit states are either $\{|1\rangle_k, |2\rangle_k\}$ or $\{|0\rangle_k, |1\rangle_k\}$ for each qutrit (i.e., $R_k = 0$ or 2) [37]. Given a target state $|\psi\rangle$, the question is thus whether there exist choices of R_k, $\Delta_k^{(1)}$, and $\Delta_k^{(2)}$ which fulfill (6) for all $\mathbf{n} \in S_\psi$.

Although (6) depends only on S_ψ and not on the coefficients $c_{\mathbf{n}}$, a general solution is not easy to obtain, because the number of variables increases with N. Nevertheless, (6) can be significantly simplified. In Appendix B, we show that whenever (6) has a solution it has a solution with $R_k = 0$ for all but a single k. For a given $|\psi\rangle$ it is thus sufficient to check whether there exist choices of $k' \in \{1,\ldots,N\}$, $\Delta_k^{(1)}$, and $\Delta_k^{(2)}$ fulfilling

$$n_{k'}\Delta_{k'}^{(1)} + \sum_{k\neq k'}\left[(1-n_k)\Delta_k^{(1)} + n_k\Delta_k^{(2)}\right] - \Delta_{k'}^{(2)} = 0. \tag{7}$$

If there do, then it follows from the proof in Appendix A that the machine defined by these choices, with bath k' hot and all other baths cold, can generate states arbitrarily close to $|\psi\rangle$.

Below, we consider several families of genuine multipartite entangled states, important in quantum information processing, namely, GHZ, Dicke, and cluster states. We show that they admit solutions to (7) and hence can be generated. Furthermore, we consider the tradeoff between heralding success probability and the quality of the generated states, as well as the effect of finite temperatures, and show that they can be robustly generated also away from the ideal limit of the entanglement engine.

IV. GHZ STATES

We start with N-qubit GHZ state $|\mathrm{GHZ}\rangle = \frac{1}{\sqrt{2}}(|10\ldots0\rangle + |01\ldots1\rangle)$. This state admits a solution to (7) (see Fig. 1). We take the first bath to be hot and the rest cold, and let the free Hamiltonians of the hot qutrit and each of the $N-1$ cold qutrits be $H_h = \Delta_h^{(1)}|1\rangle\langle 1| + \Delta_h^{(2)}|2\rangle\langle 2|$ and $H_c = [(\Delta_h^{(2)} - \Delta_h^{(1)})|1\rangle\langle 1| + \Delta_h^{(2)}|2\rangle\langle 2|]/(N-1)$, respectively. To construct an energy-conserving interaction Hamiltonian, we follow the recipe above. Writing $\bar{0}$ for a string of $N-1$ zeros $0\ldots0$, and similarly for $\bar{1}$ and $\bar{2}$, we have $|R\rangle = |2\bar{0}\rangle$. Embedding $|\mathrm{GHZ}\rangle$ in the qutrit space, from (5) we get

$$H_{\rm int} = g(|2\bar{0}\rangle\langle 1\bar{1}| + |2\bar{0}\rangle\langle 0\bar{2}| + |1\bar{1}\rangle\langle 2\bar{0}| + |0\bar{2}\rangle\langle 2\bar{0}|). \tag{8}$$

Once the steady state of the dynamics (2) is obtained, we apply the filter $\Pi_h = |0\rangle\langle 0| + |1\rangle\langle 1|$ to the hot system and the filter $\Pi_c = |1\rangle\langle 1| + |2\rangle\langle 2|$ to each of the cold systems. Successful filtering heralds the generation of $|\mathrm{GHZ}\rangle$.

As explained above, the perfect GHZ state is obtained only under idealized conditions (maximal temperature gradient and coupling strength ratios tending to zero). We now consider the quality of the generated state in case of varying filtering success probabilities (4) and then for finite temperatures.

As argued above, in the ideal limit, $\gamma_h \ll \gamma_c$, the system is most likely found outside the filtered subspace, causing $p_{\rm suc} \to 0$ as $\gamma_h/\gamma_c \to 0$. However, away from this idealized limit, we find that the state ρ' after filtering (considered as an N-qubit state) may still have a high fidelity $F = \langle\mathrm{GHZ}|\rho'|\mathrm{GHZ}\rangle$ with the GHZ state. Figure 2 shows the tradeoff between F and $p_{\rm suc}$ for $N = 2, 3, 4$ systems. We see that fidelities above 90% are obtained for $p_{\rm suc}$ at the 5% level. Note that $p_{\rm suc}$ is bounded, even when the fidelity is allowed to degrade. The maximal $p_{\rm suc}$ decreases with increasing N, however the corresponding fidelity also increases. E.g., for $N = 4$, the fidelity does not reach $F = 1/2$ before $p_{\rm suc}$ reaches its maximal value of $p_{\rm suc} = 1/9$. This suggests that, as N grows, the fidelity achievable up to the maximal $p_{\rm suc}$ increases. In Appendix C, we derive the maximal value of $p_{\rm suc}$ for any N. Finally, we have also considered an analogous autonomous entanglement engine for $N = 3$ with two hot systems and one cold system, but found that the performance is worse (see Fig. 2).

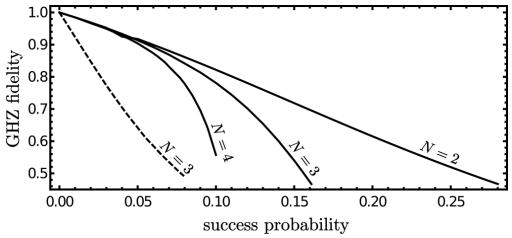

FIG. 2. Fidelity of the generated state with the GHZ state vs the probability of successful filtering for different numbers of qutrits with one hot bath (solid lines) and two hot baths (dashed line). The curves are obtained by numerical optimization over the coupling parameters under the constraint g, $\gamma_k \leqslant 10^{-2}\Delta_{\rm min}$ where $\Delta_{\rm min}$ is the smallest energy gap in each case.

We remark that, for the states considered here which have only two nonzero off-diagonal elements, a GHZ fidelity $F > 1/2$ implies genuinely multipartite entanglement [38] which can also be semi-device-independently certified via

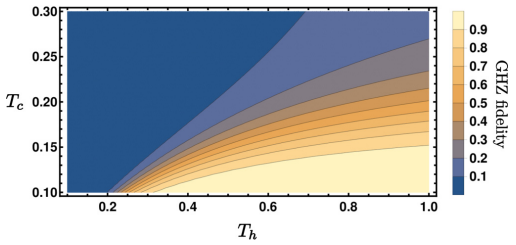

FIG. 3. Fidelity of the filtered state with the GHZ state vs the bath temperatures, for $N = 3$ and $g = 1.6 \times 10^{-3}$, $\gamma_h = 10^{-4}$, $\gamma_c = 5 \times 10^{-3}$, $\Delta^{(1)} = 1$, $\Delta^{(2)} = 2.5$. The units are dimensionless (Boltzmann's constant $k_B = 1$).

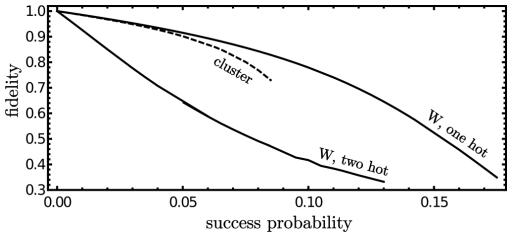

FIG. 4. Fidelity vs the filtering success probability for generation of W states using one and two hot baths (solid) and cluster states using one hot bath (dashed). The results are obtained by constrained optimization over $\gamma_h, \gamma_c, g \leqslant 10^{-2}\Delta_{\min}$, where $\Delta_{\min}$ is the smallest energy gap in each case.

the scheme of Ref. [39]. In Appendix D we have studied when the entanglement of the generated state can be device-independently certified by violating a Bell inequality.

Next, we consider the effect of finite temperatures, i.e., $T_c > 0$ and $T_h < \infty$. We keep the interaction and bath coupling strengths fixed (thus also avoiding the idealized limit of vanishing couplings). The results are presented in Fig. 3. We note that even for temperatures far from the ideal limit fidelities close to unity are possible.

Thus, our entanglement engine functions well not only in the ideal limit but also for finite temperatures and coupling strengths. In Appendix E, we further show that qualitatively similar results can be obtained when the simple reset model is replaced by a master equation in standard Lindblad form.

V. DICKE STATES

As a second example, we consider N-qubit Dicke states. The Dicke state with l excitations is given by $|D_l^N\rangle = (\binom{N}{l})^{-1/2} \sum_s \sigma_s[|1\rangle^l \otimes |0\rangle^{N-l}]$, where the sum is over all permutations σ_s of the subsystems. Notably, setting $l = 1$ returns the well-known W states. Again, one finds that all such states admit solutions to (7). Hence, every Dicke state can be generated by an autonomous entanglement engine. For instance, we choose the first qutrit hot (H_h) and the rest cold (H_c), and energies $\Delta_h^{(1)} = \Delta_c^{(2)} - \Delta_c^{(1)}$ and $\Delta_h^{(2)} = l\Delta_c^{(2)} + (N - l - 1)\Delta_c^{(1)}$. For the case $(N, l) = (3, 1)$, we have analytically solved the reset master equation in terms of g, γ_h, and γ_c and computed the fidelity $F = \langle D_1^3|\rho'|D_1^3\rangle$. Similarly, we have analytically evaluated p_{suc} in Eq. (4). The tradeoff between F and p_{suc} is shown in Fig. 4. As for the GHZ case, we find that high fidelities can be reached with success probabilities at the few-percent level. We have also checked that increasing the number of hot systems (to two) does not improve performance.

VI. CLUSTER STATE

Finally, we consider a linear four-qubit cluster state $|C\rangle = \frac{1}{2}(|0110\rangle + |0101\rangle + |1010\rangle - |1001\rangle)$. A solution to (7) is obtained with one hot system and three cold systems by choosing $\Delta_h^{(1)} = \Delta_c^{(2)} - \Delta_c^{(1)}$ and $\Delta_h^{(2)} = 2\Delta_c^{(2)} + \Delta_c^{(1)}$. In analogy with the previous, we consider the tradeoff between the $F = \langle C|\rho'|C\rangle$ of the generated state ρ' with the cluster state and filtering success probability p_{suc}. We have evaluated both F and p_{suc} analytically for a single hot bath, and optimized over the couplings g, γ_h, and γ_c to obtain the results in Fig. 4. Again, high-fidelity cluster states can be generated with success probabilities at the few-percent level. Furthermore, in Appendix D, we have considered the device-independent certification of ρ' via Bell inequalities tailored for cluster states [40] at varying p_{suc}. We find that large Bell inequality violations can be obtained for every p_{suc} up to its maximal value of $p_{\text{suc}} \approx 0.085$, demonstrating that the entanglement engine works well over a wide regime.

VII. CONCLUSION

We have given a general recipe for autonomous entanglement engines which enable heralded generation of multipartite entangled states between any number of parties. As demonstrated by several examples, a wide range of states can be targeted, including GHZ, Dicke, and cluster states. While pure target states are only generated perfectly for infinite temperature gradients and vanishing heralding success probabilities, we have explored finite temperatures and heralding probabilities as well and have found that high fidelities can be attained also away from the ideal regime.

Thus, probabilistic generation of high-quality multipartite entanglement is possible using only incoherent, thermal processes and energy-preserving interactions, requiring no work input. It would be interesting to understand if strong entanglement could be generated by an autonomous engine in a deterministic manner, i.e., without filtering. Finally, perspectives for experimental implementation could be explored. In that context, a natural question is whether genuine multipartite entangled states can be generated autonomously using only two-body Hamiltonians.

ACKNOWLEDGMENTS

We thank Marcus Huber for discussions. J.B.B. was supported by the Independent Research Fund Denmark, A.T. and N.B. were supported by the Swiss National Science Foundation (Grant No. 200021_169002 and National Centre of Competence in Research (NCCR) Quantum Science and Technology (QSIT), and G.H. was supported by the Swiss National Science Foundation through Starting Grant No. PRIMA PR00P2_179748.

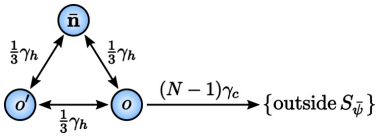

FIG. 5. Flow diagram for population entering and leaving the state $|o\rangle$. Hot resets take the system from $|o\rangle$ to state $|o'\rangle$ or $|\bar{\mathbf{n}}\rangle$, while cold resets take it to other states outside the support $S_{\bar{\psi}}$. The transition rates due to hot and cold resets are indicated.

APPENDIX A: AUTONOMOUS GENERATION OF TARGET STATES

We prove that any state $|\psi\rangle$ which admits a solution to the energy-conservation condition can be generated by an autonomous entanglement engine. Following the main text, we write the target state as

$$|\psi\rangle = \sum_{\mathbf{n}\in S_\psi} c_{\mathbf{n}}|\mathbf{n}\rangle \tag{A1}$$

where $c_{\mathbf{n}} \in \mathbb{C}$, $\sum_{\mathbf{n}} |c_{\mathbf{n}}|^2 = 1$, and S_ψ is the set of binary strings $s = \{0, 1\}^N$ such that $|\psi\rangle$ has support of $|s\rangle$. We show that the state ρ' returned by the machine described in the main text (after heralding) is indeed the target state. To this end, we must characterize ρ'. For simplicity, we will first focus on the diagonal elements of ρ' and then on its off-diagonal elements.

1. Diagonal elements

We aim to show that the diagonal elements of ρ' correspond to the populations $|c_{\bar{\mathbf{n}}}|^2$, where $|\bar{\mathbf{n}}\rangle$ are the computational basis states on which the embedded target state $|\bar{\psi}\rangle$ has support. To enable the characterization of the diagonal elements of ρ', we use flow diagrams as illustrated in Fig. 5. Such a diagram represents the transitions induced by the influence of hot and cold resets, along with the rate of said transitions, on a given support state $|\bar{\mathbf{n}}\rangle$. As illustrated, by a hot reset on $|\bar{\mathbf{n}}\rangle$ one can reach two other states, denoted by $|o\rangle$ and $|o'\rangle$. Importantly, neither of these two states can be members of $S_{\bar{\psi}}$ since it is otherwise at odds with the conditions for an autonomous Hamiltonian. From the flow diagram, we obtain the following steady-state condition when considering the flow into and out of the state $|o\rangle$:

$$P_o\left[2\frac{\gamma_h}{3} + \gamma_c(N-1)\right] = \frac{\gamma_h}{3}(P_{\bar{\mathbf{n}}} + P_{o'}), \tag{A2}$$

where we have adopted the simplified notation $P_s = \langle s|\rho|s\rangle$. However, since $|o\rangle, |o'\rangle \notin S_{\bar{\psi}}$ (nor do they equal the state $|R\rangle$), they do not appear in the interaction Hamiltonian and are treated equally by the dissipation. Hence, it follows that $P_o = P_{o'}$. This leads us to rewrite (A2) as

$$\frac{P_o}{P_{\bar{\mathbf{n}}}} = \frac{\gamma_h}{3(N-1)\gamma_c + \gamma_h}. \tag{A3}$$

Let us now consider the filtered subspace, i.e., the space in which the heralded state ρ' lives. Since the filtering corresponds to projecting each qutrit onto a qubit subspace, there are consequently 2^N computational basis states spanning the filtered subspace. Of these, $\nu = |S_{\bar{\psi}}|$ are members of $S_{\bar{\psi}}$, whereas another ν are reachable by a hot reset to each element in $S_{\bar{\psi}}$. Denote the latter set of states by G_h. The remaining $2^N - 2\nu$ states have no population (diagonal element equal to zero) since they cannot be reached either via the interaction Hamiltonian or via resets. Let $\bar{P}_o$ denote renormalized P_o after filtering, i.e., $\bar{P}_o = \langle o|\rho'|o\rangle$. Normalization requires that

$$\sum_{o\in S_{\bar{\psi}}} \bar{P}_o + \sum_{o\in G_h} \bar{P}_o = 1. \tag{A4}$$

However, due to the symmetries of the interaction Hamiltonian and the linearity of the dynamics, we may write $\bar{P}_o = |c_o|^2\bar{P}_{\mathrm{S}}$ for $o \in S_{\bar{\psi}}$ for some constant population $\bar{P}_{\mathrm{S}}$ independent of o. Similarly, we may write $\bar{P}_o = |c_o|^2\bar{P}_{\mathrm{G}}$ for $o \in G_h$ for some constant population $\bar{P}_{\mathrm{G}}$ independent of o. The normalization condition reduces to

$$\bar{P}_{\mathrm{S}}\left(1 + \frac{\bar{P}_{\mathrm{G}}}{\bar{P}_{\mathrm{S}}}\right) = 1, \tag{A5}$$

which together with (A3) gives

$$\bar{P}_{\mathrm{S}} = \left(1 + \frac{\bar{P}_{\mathrm{G}}}{\bar{P}_{\mathrm{S}}}\right) = \left[1 + \frac{\gamma_h}{3(N-1)\gamma_c + \gamma_h}\right]^{-1}. \tag{A6}$$

In the limit $\gamma_h \ll \gamma_c$ we have $\bar{P}_{\mathrm{S}} \to 1$, and therefore also $\bar{P}_{\mathrm{G}} \to 0$. Consequently, we have found that in the given limit, for $\bar{\mathbf{n}} \in S_{\bar{\psi}}$,

$$\bar{P}_{\bar{\mathbf{n}}} = \langle\bar{\mathbf{n}}|\rho'|\bar{\mathbf{n}}\rangle = |c_{\bar{\mathbf{n}}}|^2. \tag{A7}$$

These are the desired diagonal elements.

2. Off-diagonal elements

We now aim to show that the off-diagonal elements of ρ' correspond to $c_n c_{n'}^*$. Due to hermiticity, it is sufficient to consider the upper triangle in the matrix of ρ'. Among these off-diagonal entries, there are $\binom{\nu}{2}$ that correspond to coherences generated between the computational basis states associated to $n, n' \in S_{\bar{\psi}}$ [we have dropped the notation in bold ($\bar{\mathbf{n}}$) since in this section n will sometimes be a member of $S_{\bar{\psi}}$]. Another ν off-diagonals correspond to coherences generated between the computational basis states associated to $n \in S_{\bar{\psi}}$ and the state $|R\rangle$. The remaining off-diagonal elements are not reachable by the dynamics (neither via resets nor via the Hamiltonian) and are therefore equal zero. We use the short-hand notation $\rho_{n,n'} = \langle n|\rho|n'\rangle$ to write the reset master equation in the steady state as

$$\begin{aligned} 0 &= \dot{\rho}_{n,n'} \\ &= -i\langle n|[H,\rho]|n'\rangle + \frac{\gamma_h}{3}\langle n|\mathbb{1}\otimes \mathrm{Tr}_1(\rho)|n'\rangle \\ &\quad + \sum_{k=2}^{N} \gamma_c\langle n|[|0\rangle\langle 0|\otimes_k \mathrm{Tr}_k(\rho)]|n'\rangle - (\gamma_h+\gamma_c)\rho_{n,n'}. \end{aligned} \tag{A8}$$

For the first term in Eq. (A8) we have that

$$\begin{aligned} &\langle n|[H,\rho]|n'\rangle \\ &= g\langle n|(|\bar{\psi}\rangle\langle R| + |R\rangle\langle\bar{\psi}|)\rho - \rho(|\bar{\psi}\rangle\langle R| + |R\rangle\langle\bar{\psi}|)|n'\rangle \\ &= g(\langle n|\bar{\psi}\rangle\langle R|\rho|n'\rangle + \langle n|R\rangle\langle\bar{\psi}|\rho|n'\rangle \\ &\quad - \langle n|\rho|\bar{\psi}\rangle\langle R|n'\rangle - \langle n|\rho|R\rangle\langle\bar{\psi}|n'\rangle). \end{aligned} \tag{A9}$$

Taking $n, n' \neq R$, the two middle terms vanish. Moreover, if $n, n' \notin S_{\bar{\psi}}$ also the first and fourth terms vanish. If $n, n' \in S_{\bar{\psi}}$ then we have $\langle n|\bar{\psi}\rangle = c_n$ and $\langle \bar{\psi}|n'\rangle = c^*_{n'}$ and therefore $\langle n|[H,\rho]|n'\rangle = g(c_n\rho_{R,n'} - c^*_{n'}\rho_{n,R})$. Thus,

$$\langle n|[H,\rho]|n'\rangle = \begin{cases} g(c_n\rho_{R,n'} - c^*_{n'}\rho_{n,R}) & \text{if } n, n' \in S_{\bar{\psi}} \\ 0 & \text{if } n, n' \notin S_{\bar{\psi}} \text{ and } n, n' \neq R \end{cases}. \tag{A10}$$

For the second term in Eq. (A8) a direct calculation gives

$$\langle n|\mathbb{1} \otimes \mathrm{Tr}_1(\rho)|n'\rangle = \delta_{n_1,n'_1} \sum_j \rho_{j\bar{n},j\bar{n}'}, \tag{A11}$$

where the bar sign denotes $\bar{s} = s_2 \ldots s_N$. Moreover, the third term in Eq. (A8) straightforwardly evaluates to

$$\langle n|[|0\rangle\langle 0| \otimes_k \mathrm{Tr}_k(\rho)]|n'\rangle = \delta_{n_k,0}\delta_{n'_k,0} \sum_{j_k} \rho_{\overleftarrow{n}\, j_k \overrightarrow{n},\, \overleftarrow{n}'\, j_k \overrightarrow{n}'}, \tag{A12}$$

where $\overleftarrow{s} = s_1 \ldots s_{k-1}$ and $\overrightarrow{s} = s_{k+1} \ldots s_N$. Notice that this term vanishes for $k = 2, \ldots, N$ if either n or n' is a member of $S_{\bar{\psi}}$. In conclusion, for $n, n' \in S_{\bar{\psi}}$, we can rewrite (A8) as

$$\begin{aligned} 0 &= \dot{\rho}_{n,n'} \\ &= -ig(c_n\rho_{R,n'} - c^*_{n'}\rho_{n,R}) \\ &\quad + \frac{\gamma_h}{3}\delta_{n_1,n'_1} \sum_j \rho_{j\bar{n},j\bar{n}'} - (\gamma_h + \gamma_c)\rho_{n,n'}. \end{aligned} \tag{A13}$$

When $n \neq n'$ (since one cannot transition between two support states by a hot reset) Eq. (A11) becomes $\delta_{n_1,n'_1}\sum_j \rho_{j\bar{n},j\bar{n}'} = \delta_{n_1,n'_1}\rho_{n,n'}$. Furthermore, by hermiticity we have that $\rho_{R,n'} = \rho^*_{n',R}$, and due to the symmetries of the Hamiltonian it also holds that $\rho_{n,R} = c_n L$ where L is a constant related to the population in the steady state that is independent of n. With this in hand, we consider the three equations obtained from (A13):

$$\begin{aligned} 0 &= \dot{\rho}_{n,n'} \\ &= -igc_nc^*_{n'}(L^* - L) + \frac{\gamma_h}{3}\delta_{n_1,n'_1}\rho_{n,n'} - (\gamma_h + \gamma_c)\rho_{n,n'}, \end{aligned} \tag{A14}$$

$$0 = \dot{\rho}_{n,n} = -ig|c_n|^2(L^* - L) + \frac{\gamma_h}{3}\sum_j \rho_{j\bar{n},j\bar{n}} - (\gamma_h + \gamma_c)\rho_{n,n}, \tag{A15}$$

$$0 = \dot{\rho}_{i\bar{n},i\bar{n}} = \frac{\gamma_h}{3}\sum_j \rho_{j\bar{n},j\bar{n}} - (\gamma_h + \gamma_c)\rho_{i\bar{n},i\bar{n}}, \tag{A16}$$

where in the first equation we have taken $n, n' \in S_{\bar{\psi}}$ with $n \neq n'$; in the second equation we have taken $n, n' \in S_{\bar{\psi}}$ with $n = n'$; and in the third equation we have taken $n, n' \in S_{\bar{\psi}}$ with $n = n'$ but then replaced n_1 with the index i which runs over the two values $i \neq n_1$. Summing over i in Eq. (A16) gives

$$\sum_{i\neq n_1} \rho_{i\bar{n},i\bar{n}} = \frac{2\gamma_h}{3\gamma_c + \gamma_h}\rho_{n,n}. \tag{A17}$$

Inserted into Eq. (A15) we obtain

$$ig(L^* - L) = -\frac{\rho_{n,n}}{|c_n|^2}\frac{3\gamma_c(\gamma_h + \gamma_c)}{3\gamma_c + \gamma_h}. \tag{A18}$$

Finally, when inserted into Eq. (A14), we can obtain the off-diagonal elements from the diagonal elements of ρ'. We obtain

$$\rho_{n,n'} = -\frac{3\gamma_c(\gamma_h + \gamma_c)}{3\gamma_c + \gamma_h}\left[\frac{\gamma_h}{3}\delta_{n_1,n'_1} - (\gamma_h + \gamma_c)\right]^{-1}\frac{c_nc^*_{n'}}{|c_n|^2}\rho_{n,n}. \tag{A19}$$

However, the ratios between the off-diagonal terms are conserved after filtering if they belong to the filtered subspace. We use the notation $\bar{\rho}_{s,s'} = \langle s|\rho'|s'\rangle$. Then, taking the relevant limit of $\gamma_h \ll \gamma_c$, we obtain

$$\lim_{\gamma_h \ll \gamma_c} \bar{\rho}_{n,n'} = \frac{c_nc^*_{n'}}{|c_n|^2} \lim_{\gamma_h \ll \gamma_c} \bar{\rho}_{n,n}. \tag{A20}$$

The right-hand side features a diagonal element which was evaluated in Eq. (A7). In the relevant limit, we obtain the final result:

$$\lim_{\gamma_h \ll \gamma_c} \bar{\rho}_{n,n'} = c_nc^*_{n'}. \tag{A21}$$

In conclusion, we have shown that the heralded state ρ' is the target state.

APPENDIX B: SIMPLIFIED CONDITIONS FOR ENERGY CONSERVATION

1. A single hot system is sufficient

Here, we show that if the conditions for the interaction to be energy conserving can be solved using q hot systems (i.e., systems with $R_k = 2$) and $N - q$ cold systems (i.e., systems with $R_k = 0$) then there also exists a solution with just a single hot system and $N - 1$ cold systems.

To prove this, we show that any set of valid energies $\Delta^{(1)}_k$, $\Delta^{(2)}_k$ fulfilling the energy-conservation condition for q hot systems allows one to define another set of energies $\{\varepsilon^{(1)}_k, \varepsilon^{(2)}_k\}$ which fulfill the corresponding condition with a single hot system. Without loss of generality (as one may always permute the parties), we can take the hot systems to be the first ones. Then the energy-conservation condition with q hot systems reads

$$\begin{aligned} \forall \mathbf{n} \in S_\psi : \ &\sum_{k=1}^{q}\left(n_k\Delta^{(1)}_k - \Delta^{(2)}_k\right) \\ &+ \sum_{k=q+1}^{N}\left[(1-n_k)\Delta^{(1)}_k + n_k\Delta^{(2)}_k\right] = 0, \end{aligned} \tag{B1}$$

while the corresponding condition with a single hot system ($q = 1$) becomes

$$\forall \mathbf{n} \in S_\psi : \ \left(n_1\varepsilon^{(1)}_1 - \varepsilon^{(2)}_1\right) + \sum_{k=2}^{N}\left[(1-n_k)\varepsilon^{(1)}_k + n_k\varepsilon^{(2)}_k\right] = 0. \tag{B2}$$

Note that the energies must satisfy $\Delta^{(2)}_k > \Delta^{(1)}_k > 0$ and similarly $\varepsilon^{(2)}_k > \varepsilon^{(1)}_k > 0$. To construct a solution to (B2) given a solution to (B1), we choose

$$\varepsilon^{(1)}_k = \Delta^{(1)}_k \quad \text{for} \quad k = q+1, \ldots, N, \tag{B3}$$

$$\varepsilon^{(2)}_k = \Delta^{(2)}_k \tag{B4}$$

and

$$\varepsilon_k^{(1)} = t_k - \Delta_k^{(2)} \quad \text{for} \quad k = 2, \ldots, q, \tag{B5}$$

$$\varepsilon_k^{(2)} = t_k - \Delta_k^{(2)} + \Delta_k^{(1)} \tag{B6}$$

for some t_k satisfying $t_k > \Delta_k^{(2)}$. Note that with these choices we have $\varepsilon_k^{(2)} > \varepsilon_k^{(1)} > 0$ for $k = 2, \ldots, N$, as desired. Inserting in Eq. (B2), we get

$$\forall \mathbf{n} \in S_\psi : \left(n_1 \varepsilon_1^{(1)} - \varepsilon_1^{(2)} + \sum_{k=2}^{q} t_k \right) + \sum_{k=2}^{q} \left(n_k \Delta_k^{(1)} - \Delta_k^{(2)} \right) + \sum_{k=q+1}^{N} \left[(1 - n_k) \Delta_k^{(1)} + n_k \Delta_k^{(2)} \right] = 0. \tag{B7}$$

This reduces to (B1) provided that

$$\forall n_1 : \left(n_1 \varepsilon_1^{(1)} - \varepsilon_1^{(2)} + \sum_{k=2}^{q} t_k \right) = n_1 \Delta_1^{(1)} - \Delta_1^{(2)}, \tag{B8}$$

which is solved by

$$\varepsilon_1^{(1)} = \Delta_1^{(1)}, \tag{B9}$$

$$\varepsilon_1^{(2)} = \Delta_1^{(2)} + \sum_{k=2}^{q} t_k. \tag{B10}$$

It is easy to see that $\varepsilon_1^{(2)} > \varepsilon_1^{(1)} > 0$. We thus have a valid choice of energies $\varepsilon_k^{(1)}$, $\varepsilon_k^{(1)}$ for which (B2) reduces (B1). Hence, any solution with q hot systems also implies the existence of a solution with a single hot system, as claimed.

2. Identical energy structures for all hot and all cold systems

If the energy spectra of all hot systems (i.e., all systems with $R_k = 2$) are identical, and similarly those of cold systems (with $R_k = 0$) are identical, then the energy-conservation conditions can be simplified. Note that all the examples given in the main text (for GHZ, Dicke, and cluster states) belong to this setting.

Specifically, here we show that if $\Delta_k^{(1)}$ and $\Delta_k^{(2)}$ depend only on R_k then the existence of $R \in \{0, 2\}^N$ and a choice of energies fulfilling the energy-conservation conditions in the main text is equivalent to the existence of a vector $\mathbf{r} \in \{0, 1\}^N$ such that $\mathbf{r} \neq \mathbf{0}, \mathbf{1}$ and for each pair of vectors $\mathbf{n}, \mathbf{n}' \in S_\psi$ either

$$(\mathbf{n} - \mathbf{n}') \cdot \mathbf{r} = (\mathbf{n} - \mathbf{n}') \cdot (\mathbf{1} - \mathbf{r}) = 0 \tag{B11}$$

or

$$\frac{(\mathbf{n} - \mathbf{n}') \cdot \mathbf{r}}{(\mathbf{n} - \mathbf{n}') \cdot (\mathbf{1} - \mathbf{r})} = c \tag{B12}$$

where $c < 0$ is a constant independent of $\mathbf{n}, \mathbf{n}'$, and $\mathbf{0} = (0, 0, \ldots, 0)$ and $\mathbf{1} = (1, 1, \ldots, 1)$. That is, the interaction Hamiltonian can be made energy conserving if and only if an $\mathbf{r} \neq \mathbf{0}, \mathbf{1}$ exists fulfilling (B11) and (B12).

Before we proceed with the proof, we illustrate (B11) and (B12) and the notation introduced above in the simplest setting of two parties. We take the maximally entangled state $|\Psi^+\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$ as the target and choose $|R\rangle = |20\rangle$. The target has support on just two states, $S_\psi = \{\mathbf{n}, \mathbf{n}'\}$, where

$$\mathbf{n} = (0, 1) \quad \text{and} \quad \mathbf{n}' = (1, 0). \tag{B13}$$

It is straightforward to verify that (B12) is satisfied for $\mathbf{r} = (1, 0)$, with $c = -1$. Hence, $|\Psi^+\rangle$ can indeed be generated autonomously. Looking at the energy conditions in the main text, we see that the conditions on the energies coming from $\mathbf{n}$ and $\mathbf{n}'$ are, respectively,

$$\Delta_2^{(2)} = \Delta_1^{(2)} \tag{B14}$$

and

$$\Delta_2^{(1)} = \Delta_1^{(2)} - \Delta_1^{(1)}. \tag{B15}$$

Thus, the two qutrits have the same maximal energy but inverted level structures. The gap between the two lower levels for the second qutrit equals the gap between the upper two levels for the first qutrit. This corresponds exactly to the entanglement engine of Ref. [33].

The conditions (B11) and (B12) can be defined as follows. If we define a vector $\mathbf{r} \in \{0, 1\}^N$ such that $r_k = 0$ if $R_k = 0$ and $r_k = 1$ for $R_k = 2$, then for each $\mathbf{n} \in S_\psi$ the condition $E_{\mathbf{n}} = E_{\bar{R}}$ from the main text can be expressed as

$$\sum_{k=1}^{N} \left\{ r_k n_k \Delta_k^{(1)} + (1 - r_k) \left[(1 - n_k) \Delta_k^{(1)} + n_k \Delta_k^{(2)} \right] - r_k \Delta_k^{(2)} \right\} = 0. \tag{B16}$$

The question is whether there exist choices of $\mathbf{r}$, $\Delta_k^{(1)}$, and $\Delta_k^{(2)}$ which fulfill this. Rewriting, we have

$$\sum_{k \text{ s.t. } r_k = 0} \left[(1 - n_k) \Delta_k^{(1)} + n_k \Delta_k^{(2)} \right] + \sum_{k \text{ s.t. } r_k = 1} \left[n_k \Delta_k^{(1)} - \Delta_k^{(2)} \right] = 0. \tag{B17}$$

Now, if the energy structures of all qutrits with the same R_k are the same, then the energies appearing under each sum become independent of k. Let us denote the energy gaps of qutrits with $R_k = 0$ by $\delta_1 = \Delta_k^{(1)}$ and $\delta_2 = \Delta_k^{(2)} - \Delta_k^{(1)}$ and those of qutrits with $R_k = 2$ by $\delta_3 = \Delta_k^{(1)}$ and $\delta_4 = \Delta_k^{(2)} - \Delta_k^{(1)}$. Then (B17) becomes

$$\sum_{k \text{ s.t. } r_k = 0} [\delta_1 + n_k \delta_2] + \sum_{k \text{ s.t. } r_k = 1} [(n_k - 1)\delta_3 - \delta_4] = 0, \tag{B18}$$

which is equivalent to

$$(N - |\mathbf{r}|)\delta_1 + (\mathbf{1} - \mathbf{r}) \cdot \mathbf{n}\, \delta_2 - \mathbf{r} \cdot (\mathbf{1} - \mathbf{n})\delta_3 - |\mathbf{r}|\delta_4 = 0, \tag{B19}$$

where $\mathbf{1} = (1, \ldots, 1)$ and $|\mathbf{r}|$ is the number of 1's in $\mathbf{r}$. This must hold for every $\mathbf{n} \in S_\psi$, and thus we have a set of linear

equations:

$$\begin{pmatrix} N-|\mathbf{r}| & (\mathbf{1}-\mathbf{r})\cdot\mathbf{n}^{(1)} & -\mathbf{r}\cdot(\mathbf{1}-\mathbf{n}^{(1)}) & -|\mathbf{r}| \\ N-|\mathbf{r}| & (\mathbf{1}-\mathbf{r})\cdot\mathbf{n}^{(2)} & -\mathbf{r}\cdot(\mathbf{1}-\mathbf{n}^{(2)}) & -|\mathbf{r}| \\ \vdots & & & \\ N-|\mathbf{r}| & (\mathbf{1}-\mathbf{r})\cdot\mathbf{n}^{(\nu)} & -\mathbf{r}\cdot(\mathbf{1}-\mathbf{n}^{(\nu)}) & -|\mathbf{r}| \end{pmatrix} \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{pmatrix} = \mathbf{0}, \tag{B20}$$

where ν is the number of elements of S_ψ. Regarding $\delta = (\delta_1, \ldots, \delta_4)$ as a variable, we would like to know when there exists $\mathbf{r} \in \{0,1\}^N$ such that (B20) has a solution over $(\mathbb{R}^+)^4$, i.e., a positive solution. Given such a solution, for any $l = 1, \ldots, \nu$ we must have

$$(\mathbf{1}-\mathbf{r})\cdot\mathbf{n}^{(l)}\delta_2 - \mathbf{r}\cdot(\mathbf{1}-\mathbf{n}^{(l)})\delta_3 = |\mathbf{r}|\delta_4 - (N-|\mathbf{r}|)\delta_1, \tag{B21}$$

where the right-hand side is independent of l. Note that this condition can never be fulfilled if $\mathbf{r} = \mathbf{0}$ or 1, because the two sides of the equation then have opposite signs. However, if the condition is satisfied, then for any pair $l, l' = 1, \ldots, \nu$ we have

$$(\mathbf{1}-\mathbf{r})\cdot(\mathbf{n}^{(l)}-\mathbf{n}^{(l')})\delta_2 - \mathbf{r}\cdot(\mathbf{n}^{(l')}-\mathbf{n}^{(l)})\delta_3 = 0. \tag{B22}$$

Hence, for a positive solution to exist, for each pair of support states either $\mathbf{n}^{(l)}$ and $\mathbf{n}^{(l')}$ has an equal number of 1's in positions where $\mathbf{r}$ has zero and an equal number of 1's in positions where $\mathbf{r}$ has 1, or

$$\frac{\mathbf{r}\cdot(\mathbf{n}^{(l)}-\mathbf{n}^{(l')})}{(\mathbf{1}-\mathbf{r})\cdot(\mathbf{n}^{(l)}-\mathbf{n}^{(l')})} = -\frac{\delta_2}{\delta_3} < 0 \tag{B23}$$

is a negative constant independent of l, l'. On the other hand, if an $\mathbf{r} \neq \mathbf{0}, \mathbf{1}$ exists fulfilling these conditions, then a positive solution of (B20) is guaranteed to exist. This is because the left-hand side of (B21) is then independent of l and thus one can always find positive δ_1 and δ_4 which make the equality true.

APPENDIX C: MAXIMAL FILTERING PROBABILITY IN THE GHZ-STATE MACHINE

Naturally, since N local filters are performed on the steady state of an N-qutrit autonomous thermal machine, the probability of a successful filtering decreases with N. It is therefore reasonable to ask what this maximal possible success probability is. This can be determined analytically by considering the flow of population in the steady state of the GHZ machine.

Since a cold reset always takes a system out of the filtered subspace, the maximal success probability is obtained in the limit $\gamma_h \gg \gamma_c$, i.e., the opposite of the limit maximizing the fidelity of the generated state with the target state. To determine p_{suc} in this limit, let S_k denote the set of all eigenstates of the joint free Hamiltonian where k cold qutrits are in one of the excited states (all in the same one), while the remaining $N-k-1$ cold qutrits are in the ground state. For instance, in S_{N-1} we have the states $S_{N-1} = \{|0,\bar{1}\rangle, |1,\bar{1}\rangle, |2,\bar{1}\rangle, |0,\bar{2}\rangle, |1,\bar{2}\rangle, |2,\bar{2}\rangle\}$ while S_0 consists of the states $S_0 = \{|0,\bar{0}\rangle, |1,\bar{0}\rangle, |2,\bar{0}\rangle\}$. We will compare the flows of population into and out of the S_k. However, first we argue that within each S_k the populations on each of the states are equal in the steady state. We first note that all processes (the evolution driven by the H_{int} of the GHZ machine, as well as hot and cold resets) are symmetric in the states $|\bar{1}\rangle$ and $|\bar{2}\rangle$ of the cold qutrits. The populations of states with the hot qutrit in a fixed state and a fixed number of cold qutrits excited to the same excited state, and which differ only in whether this state is $|1\rangle$ or $|2\rangle$, must therefore be equal in the steady state. In contrast, H_{int} is not symmetric in the states $|0\rangle$, $|1\rangle$, and $|2\rangle$ of the hot qutrit, and hence populations of states with the hot qutrit in different levels are not expected to be equal in the steady state in general. However, in the limit $\gamma_h \gg \gamma_c$, there are many hot resets between each cold one. This will then equalize the populations within each set S_k before a cold reset causes a transition to S_{k-1}. Hence, all populations with each S_k are equal in the steady state.

We can now draw the flow diagram shown in Fig. 6 for population transfer between the S_k. In the steady state, the flow into each set S_k must equal the flow out. If we denote the population per state in S_k by P_k, we therefore have, for $k = 2, \ldots, N-1$,

$$k\gamma_c|S_k|P_k = (k-1)\gamma_c|S_{k-1}|P_{k-1}. \tag{C1}$$

The number of states in the set S_k is given by

$$\begin{aligned} |S_0| &= 3, \\ |S_k| &= 6\binom{N-1}{k}, \quad k > 0. \end{aligned} \tag{C2}$$

Inserting in Eq. (C1) and rearranging, one finds that

$$\begin{aligned} P_k &= \frac{k-1}{k}\binom{N-1}{k}^{-1}\binom{N-1}{k-1}P_{k-1} \\ &= \frac{k-1}{N-k}P_{k-1}, \quad k = 2, \ldots, N-1. \end{aligned} \tag{C3}$$

From this it follows that $P_{N-1} = P_1$ and $P_{N-2} = P_2$, etc. That is,

$$P_{N-k} = P_k, \quad k = 1, \ldots, N-1. \tag{C4}$$

To determine the relation with P_0, we note that H_{int} drives swaps between the states $|2\bar{0}\rangle \leftrightarrow \frac{1}{\sqrt{2}}(|1\bar{1}\rangle + |0\bar{2}\rangle)$ and hence

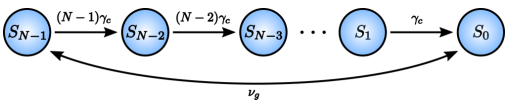

FIG. 6. Flow diagram for population entering and leaving the sets of states S_k with k cold qubits excited. The rates per state in the set of origin are indicated.

between S_0 and S_{N-1}. This process is a unitary rotation. Nevertheless, in the steady state it still results in a flow of population with a constant rate, which we can denote v_g. Focusing on the flow in and out of S_{N-1}, we can write

$$v_g P_{2\bar{0}} = v_g(P_{0\bar{2}} + P_{1\bar{1}}) + (N-1)\gamma_c |S_{N-1}|P_{N-1}. \tag{C5}$$

As argued above, when $\gamma_h \gg \gamma_c$, all states in each S_k are equally probable, and so

$$\tfrac{1}{3} v_g P_0 = \tfrac{1}{3} v_g P_{N-1} + 6(N-1)\gamma_c P_{N-1}. \tag{C6}$$

Now, if further $v_g \gg (N-1)\gamma_c$, then

$$P_0 = P_{N-1}. \tag{C7}$$

Finally, normalization of the steady state requires that

$$1 = \sum_{k=0}^{N-1} |S_k|P_k = 3P_0 + 6\sum_{k=1}^{N-1} \binom{N-1}{k} P_k. \tag{C8}$$

Together, Eqs. (C3), (C7), and (C8) provide N independent equations from which the populations P_k, $k = 0, \ldots, N-1$ can be determined. Explicitly, we can first express everything in terms of P_0. For $k \geqslant 1$,

$$P_k = \prod_{s=1}^{k-1} \frac{s}{N-s-1} P_1 = \binom{N-2}{k-1}^{-1} P_0 \tag{C9}$$

where we used that $P_1 = P_{N-1} = P_0$. Then, from (C8),

$$1 = \left[3 + 6\sum_{k=1}^{N-1} \binom{N-1}{k}\binom{N-2}{k-1}^{-1}\right] P_0 \tag{C10}$$

$$= \left[3 + 6(N-1)\sum_{k=1}^{N-1} \frac{1}{k}\right] P_0 \tag{C11}$$

$$= 3[1 + 2(N-1)h_{N-1}]P_0, \tag{C12}$$

and hence

$$P_0 = \frac{1}{3[1 + 2(N-1)h_{N-1}]} \tag{C13}$$

where h_n is the nth harmonic number. We can now compute the probability for successful filtering, given the steady-state populations (C9) and (C13). The success probability becomes

$$p_{\text{suc}} = P(\text{hot qutrit not in } |2\rangle, \text{no cold in } |0\rangle) \tag{C14}$$

$$= 4P_{N-1} = 4P_0 = \frac{4}{3[1 + 2(N-1)h_{N-1}]} \tag{C15}$$

$$\approx \frac{4}{3N\log(N)}, \tag{C16}$$

where the last line is valid for large N. We note that the assumption $v_g \gg (N-1)\gamma_c$ leading to (C7) may not formally be justified for the local master equation. However, we have checked that the final expression (C15) is consistent with solutions obtained for $N \leqslant 8$ without making this assumption.

It is interesting to observe that the critical p_{suc} for obtaining a nontrivial GHZ-state fidelity approaches the above maximal value (C14) of p_{suc} rapidly already for $N = 3$ and 4 displayed in the main text. Provided that this observation extends to larger N, it is interesting to note that genuinely multipartite entanglement can be generated with a success probability which decreases only log linearly with N.

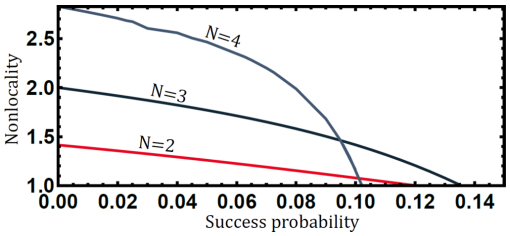

FIG. 7. Nonlocality vs filtering success probability for $N = 2, 3, 4$ in a GHZ machine with one hot system and $N-1$ cold systems. The results are obtained numerically by optimizing over $\gamma_h, g \leqslant 10^{-2}\Delta_{\min}$.

APPENDIX D: NONLOCALITY VERSUS FILTERING PROBABILITY IN THE GHZ-STATE AND CLUSTER STATE MACHINES

A particularly strong form of entanglement is that which can violate a Bell inequality. Therefore, we have considered whether the states generated by the GHZ machine at fixed success probabilities have the ability of violating Bell inequalities. To this end, we have focused on the Mermin inequalities [41], which is a family of Bell inequalities applicable to scenarios in which N observers share a state and perform one of two local measurements with binary outcomes. These inequalities are known to be maximally violated by a GHZ state. Let the input of the kth observer in the Bell scenario be $x_k \in \{0, 1\}$ and the corresponding output be $a_k \in \{0, 1\}$. We use a somewhat modified variant [42] of the Mermin inequalities which reads

$$\frac{1}{2^N} \sum_{x_1 \ldots x_N \in \{0,1\}} \left| \left\langle \prod_{k=1}^{N} \left(A_0^{(k)} + (-1)^{x_k} A_1^{(k)}\right) \right\rangle \right| \leqslant 1, \tag{D1}$$

where

$$\left\langle A_{x_1}^{(1)} \ldots A_{x_N}^{(N)} \right\rangle = \sum_{a_1 \ldots a_N} (-1)^{a_1 + \cdots + a_N} P(a_1 \ldots a_N | x_1 \ldots x_N). \tag{D2}$$

We have fixed the measurements of each observer to be those required for a maximal violation with a GHZ state. For $N = 2$, the optimal measurements are σ_x and σ_z for one observer, and $(\sigma_z + \sigma_x)/\sqrt{2}$ and $(\sigma_z - \sigma_x)/\sqrt{2}$ for the other observer. For $N = 3$ we have let all three observers perform either σ_x or σ_y, and for $N = 4$ one observer performs either σ_x or σ_y whereas the remaining three choose between $(\sigma_x + \sigma_y)/\sqrt{2}$ and $(\sigma_x - \sigma_y)/\sqrt{2}$. We have numerically obtained the tradeoff between nonlocality and the filtering success probability. The results are illustrated in Fig. 7. We conclude that the states generated by the GHZ machine can violate Bell inequalities for reasonable p_{suc}.

We have also performed an analogous analysis for the states generated at fixed success probabilities in the cluster state machine. Specifically, we have considered whether

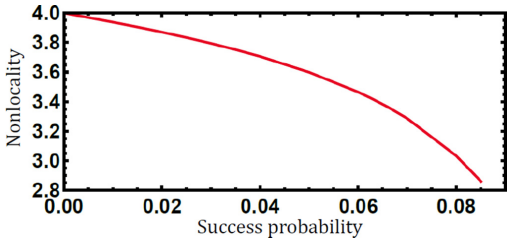

FIG. 8. Nonlocality vs filtering success probability for the cluster state machine. The results are obtained by constrained optimization over $\gamma_h, \gamma_c, g \leqslant 10^{-2}\Delta_{\min}$.

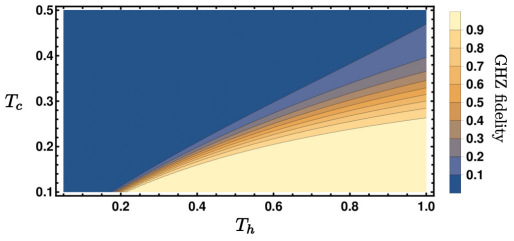

FIG. 9. Fidelity of the filtered state with the GHZ state vs the bath temperatures when using the Lindblad-type master equation (E1). The plot is for $N = 3$ parties and the parameter settings are $\Gamma_1 = 10^{-4}$, $\Gamma_2 = \Gamma_3 = 5 \times 10^{-3}$, $g = 1.6 \times 10^{-3}$, $\Delta^{(1)} = 1$ $\Delta^{(2)} = 2.5$.

these states can violate a Bell inequality tailored for cluster states [40]. We have restricted ourselves to the measurements optimal for a cluster state $1/2(|0000\rangle + |0011\rangle + |1100\rangle - |1111\rangle)$ which is unitarily equivalent to the target state. Hence, after a suitable local unitary, the Bell expression reads

$$B = \langle \sigma_x\sigma_y\sigma_y\sigma_x + \sigma_x\sigma_y\sigma_x\sigma_y + \mathbb{1}\sigma_z\sigma_x\sigma_x - \mathbb{1}\sigma_z\sigma_y\sigma_y\rangle, \quad \text{(D3)}$$

which is bounded by $B \leqslant 2$ in all local hidden variable models. With a cluster state, one can achieve $B = 4$. The tradeoff between B and the success probability of filtering is displayed in Fig. 8. We find that the generated states are nonlocal for any p_{suc} up to its maximal value.

APPENDIX E: LINDBLAD-TYPE MASTER EQUATION

To demonstrate that our results are not restricted to the simple reset model employed in the main text, here we provide a Lindblad-type master equation, which can be derived from a microscopic model with bosonic baths. The reset model is replaced by

$$\frac{d}{dt}\rho = -i[H_{\text{free}} + H_{\text{int}}, \rho] + \sum_k \Gamma_k n_B(E_k, T_k)\mathcal{D}[A_k^+]\rho(t) + \sum_k \Gamma_k[1 + n_B(E_k, T_k)]\mathcal{D}[A_k^-]\rho(t), \quad \text{(E1)}$$

where Γ_k denotes the rate of a transition, $n_B(E, T) = 1/(e^{E/T} - 1)$ is the Bose-Einstein distribution, and $\mathcal{D}$ denotes the dissipator [43].

Results from the Lindblad-type model qualitatively agree with those of the reset model. As an example, we again consider a GHZ target state for three parties ($N = 3$), solve for the steady state, and find the GHZ fidelity of the filtered state as a function of T_h and T_c. The result is shown in Fig. 9. Just as in the analogous figure in the main text based on the reset model, we see that high fidelities can be attained with reasonably low-temperature gradients. Parameter values are chosen based on recent experimental results in circuit QED [44–47] (see also [33]).

[1] M. B. Plenio, S. F. Huelga, A. Beige, and P. L. Knight, Phys. Rev. A **59**, 2468 (1999).
[2] M. B. Plenio and S. F. Huelga, Phys. Rev. Lett. **88**, 197901 (2002).
[3] S. Schneider and G. J. Milburn, Phys. Rev. A **65**, 042107 (2002).
[4] M. S. Kim, J. Lee, D. Ahn, and P. L. Knight, Phys. Rev. A **65**, 040101(R) (2002).
[5] L. Jakóbczyk, J. Phys. A: Math. Gen. **35**, 6383 (2002).
[6] D. Braun, Phys. Rev. Lett. **89**, 277901 (2002).
[7] F. Benatti, R. Floreanini, and M. Piani, Phys. Rev. Lett. **91**, 070402 (2003).
[8] L. Hartmann, W. Dür, and H.-J. Briegel, Phys. Rev. A **74**, 052304 (2006).
[9] L. Quiroga, F. J. Rodríguez, M. E. Ramírez, and R. París, Phys. Rev. A **75**, 032308 (2007).
[10] D. Burgarth and V. Giovannetti, Phys. Rev. A **76**, 062307 (2007).
[11] B. Kraus, H. P. Büchler, S. Diehl, A. Kantian, A. Micheli, and P. Zoller, Phys. Rev. A **78**, 042307 (2008).
[12] S. Diehl, A. Micheli, A. Kantian, B. Kraus, H. P. Buchler, and P. Zoller, Nat. Phys. **4**, 878 (2008).
[13] F. Verstraete, M. M. Wolf, and I. J. Cirac, Nat. Phys. **5**, 633 (2009).
[14] J. Cai, S. Popescu, and H. J. Briegel, Phys. Rev. E **82**, 021921 (2010).
[15] M. J. Kastoryano, F. Reiter, and A. S. Sørensen, Phys. Rev. Lett. **106**, 090502 (2011).
[16] M. Žnidarič, Phys. Rev. A **85**, 012324 (2012).
[17] B. Bellomo and M. Antezza, New J. Phys. **15**, 113052 (2013).
[18] F. Reiter, L. Tornberg, G. Johansson, and A. S. Sørensen, Phys. Rev. A **88**, 032317 (2013).
[19] M. J. A. Schuetz, E. M. Kessler, L. M. K. Vandersypen, J. I. Cirac, and G. Giedke, Phys. Rev. Lett. **111**, 246802 (2013).
[20] S. Walter, J. C. Budich, J. Eisert, and B. Trauzettel, Phys. Rev. B **88**, 035441 (2013).
[21] F. Ticozzi and L. Viola, Quantum Inf. Comput. **14**, 0265 (2014).
[22] D. Boyanovsky and D. Jasnow, Phys. Rev. A **96**, 012103 (2017).
[23] A. Hewgill, A. Ferraro, and G. De Chiara, Phys. Rev. A **98**, 042102 (2018).

[24] C. K. Lee, M. S. Najafabadi, D. Schumayer, L. C. Kwek, and D. A. W. Hutchinson, Sci. Rep. **9**, 9147 (2019).
[25] H. Krauter, C. A. Muschik, K. Jensen, W. Wasilewski, J. M. Petersen, J. I. Cirac, and E. S. Polzik, Phys. Rev. Lett. **107**, 080503 (2011).
[26] J. T. Barreiro, M. Muller, P. Schindler, D. Nigg, T. Monz, M. Chwalla, M. Hennrich, C. F. Roos, P. Zoller, and R. Blatt, Nature (London) **470**, 486 (2011).
[27] S. Shankar, M. Hatridge, Z. Leghtas, K. M. Sliwa, A. Narla, U. Vool, S. M. Girvin, L. Frunzio, M. Mirrahimi, and M. H. Devoret, Nature (London) **504**, 419 (2013).
[28] Y. Lin, J. P. Gaebler, F. Reiter, T. R. Tan, R. Bowler, A. S. Sorensen, D. Leibfried, and D. J. Wineland, Nature (London) **504**, 415 (2013).
[29] J. B. Brask, G. Haack, N. Brunner, and M. Huber, New J. Phys. **17**, 113029 (2015).
[30] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Phys. Rev. Lett. **76**, 722 (1996).
[31] F. Tacchino, A. Auffèves, M. F. Santos, and D. Gerace, Phys. Rev. Lett. **120**, 063604 (2018).
[32] Z.-X. Man, A. Tavakoli, J. B. Brask, and Y.-J. Xia, Phys. Scr. **94**, 075101 (2019).
[33] A. Tavakoli, G. Haack, M. Huber, N. Brunner, and J. B. Brask, Quantum **2**, 73 (2018).
[34] N. Linden, S. Popescu, and P. Skrzypczyk, Phys. Rev. Lett. **105**, 130401 (2010).
[35] P. P. Hofer, M. Perarnau-Llobet, L. D. M. Miranda, G. Haack, R. Silva, J. B. Brask, and N. Brunner, New J. Phys. **19**, 123037 (2017).
[36] J. O. González, L. A. Correa, G. Nocerino, J. P. Palao, D. Alonso, and G. Adesso, Open Sys. Inform. Dyn. **24**, 1740010 (2017).
[37] Thermal resets on a given qutrit destroy entanglement with the other qutrits. For cold baths, thermal resets tend to drive the corresponding qutrit to the ground state. To suppress the effect of reset, it is therefore beneficial to choose $R_k = 0$ when the bath temperature is cold. For infinitely hot baths, resets equalize the populations on the three levels, and it thus does not matter which subspace is filtered.
[38] O. Gühne and M. Seevinck, New J. Phys. **12**, 053002 (2010).
[39] A. Tavakoli, A. A. Abbott, M.-O. Renou, N. Gisin, and N. Brunner, Phys. Rev. A **98**, 052333 (2018).
[40] V. Scarani, A. Acín, E. Schenck, and M. Aspelmeyer, Phys. Rev. A **71**, 042325 (2005).
[41] N. D. Mermin, Phys. Rev. Lett. **65**, 1838 (1990).
[42] A. Tavakoli, J. Phys. A: Math. Theor. **49**, 145304 (2016).
[43] We remark that this model suppresses one transition, in analogy with the Lindblad-type master equation considered in Ref. [33].
[44] I. M. Pop, K. Geerlings, G. Catelani, R. J. Schoelkopf, L. I. Glazman, and M. H. Devoret, Nature (London) **508**, 369 (2014).
[45] M. Jerger, P. Macha, A. R. Hamann, Y. Reshitnyk, K. Juliusson, and A. Fedorov, Phys. Rev. Appl. **6**, 014014 (2016).
[46] N. Cottet, S. Jezouin, L. Bretheau, P. Campagne-Ibarcq, Q. Ficheux, J. Anders, A. Auffèves, R. Azouit, P. Rouchon, and B. Huard, Proc. Natl. Acad. Sci. USA **114**, 7561 (2017).
[47] Y.-H. Lin, L. B. Nguyen, N. Grabon, J. San Miguel, N. Pankratova, and V. E. Manucharyan, Phys. Rev. Lett. **120**, 150503 (2018).

PHYSICS

Self-testing nonprojective quantum measurements in prepare-and-measure experiments

Armin Tavakoli[1]*, Massimiliano Smania[2], Tamás Vértesi[3], Nicolas Brunner[1], Mohamed Bourennane[2]

Self-testing represents the strongest form of certification of a quantum system. Here, we theoretically and experimentally investigate self-testing of nonprojective quantum measurements. That is, how can one certify, from observed data only, that an uncharacterized measurement device implements a desired nonprojective positive-operator valued measure (POVM). We consider a prepare-and-measure scenario with a bound on the Hilbert space dimension and develop methods for (i) robustly self-testing extremal qubit POVMs and (ii) certifying that an uncharacterized qubit measurement is nonprojective. Our methods are robust to noise and thus applicable in practice, as we demonstrate in a photonic experiment. Specifically, we show that our experimental data imply that the implemented measurements are very close to certain ideal three- and four-outcome qubit POVMs and hence non-projective. In the latter case, the data certify a genuine four-outcome qubit POVM. Our results open interesting perspective for semi–device-independent certification of quantum devices.

INTRODUCTION

Measurements in quantum theory were initially represented by complete sets of orthogonal projectors on a Hilbert space. Such measurements are standard in a multitude of applications. Nevertheless, in a modern understanding of quantum theory, measurements are described by positive-operator valued measures (POVMs), i.e., a set of positive semi-definite operators summing to identity. POVMs are the most general notion of a quantum measurement; all projective measurements are POVMs, but not all POVMs need be projective.

Nonprojective measurements are widely useful in both conceptual and applied aspects of quantum theory, as well as in quantum information processing. In several practically motivated tasks, they present concrete advantages over projective measurements. Nonprojective measurements enhance estimation and tomography of quantum states (*1*, *2*), as well as entanglement detection (*3*) and unambiguous state discrimination of nonorthogonal states (*4*, *5*). They have also found applications in quantum cryptography (*6*, *7*) and randomness generation (*8*). In addition, nonprojective measurements can be used to maximally violate particular Bell inequalities (*9*) (assuming a bound on the Hilbert space dimension), a fact that has been applied to improve randomness extraction beyond what is achievable with projective measurements (*10*, *11*).

In view of their diverse and growing applicability, it is important to develop tools for certifying and characterizing nonprojective measurements under minimal assumptions. The strongest possible form of certification involves a "black-box" scenario, where the quantum devices are a priori uncharacterized. Astonishingly, it is possible in certain cases to completely characterize both the quantum state and the measurements based only on observed data, which is referred to as "self-testing" (*12*). A well-known example is that the maximal violation of the Clauser-Horne-Shimony-Holt Bell inequality (*13*) implies (self-tests) a maximally entangled two-qubit state and pairs of anticommuting local projective measurements (*14*–*16*). Self-testing can also be made robust to noise (*17*, *18*).

However, for the purpose of characterizing nonprojective measurements in the black-box scenario, methods based on Bell inequalities encounter a challenge. Because of Neumark's theorem, every nonprojective measurement can be recast as a projective measurement in a larger Hilbert space. That is, any nonprojective measurement on a given system is equivalent to projective measurement applied to the joint state of the system and an ancilla of a suitable dimension [see, e.g., (*19*)]. Since one usually considers no restriction on Hilbert space dimension in the Bell scenario, it is nontrivial to characterize a nonprojective measurement based on a Bell inequality. While this is possible in theory (in the absence of noise) (*10*), it appears challenging in the more realistic scenario where the experiment features imperfections. To the best of our knowledge, robust self-testing methods for nonprojective measurements in Bell scenarios have not yet been developed. A possible way to circumvent the problem is to consider a Bell scenario with quantum systems of bounded Hilbert space dimension. In particular, Gómez *et al.* (*11*) and Gómez *et al.* (*20*) recently reported the experimental certification of a nonprojective measurement in a Bell experiment assuming qubits. However, these experiments do not represent self-tests, as they certify the nonprojective character of a measurement, but not how it relates to a specific target POVM.

Here, we investigate the problem of self-testing nonprojective measurements under the assumption of bounded Hilbert space dimension. We follow a different approach, by considering a prepare-and-measure scenario instead of a Bell scenario. First, this scenario offers a natural framework for certifying and characterizing nonprojective measurements. The reason is that, as argued above, the notion of nonprojectiveness almost inherently involves a notion of Hilbert spaces of fixed dimension. Then, the prepare-and-measure scenario is arguably the simplest scenario in which the problem can be studied without further assumptions. Second, the prepare-and-measure scenario offers a very significant practical advantage as compared to Bell experiments. The reason is that there is no need to involve distant observers and entangled states. This makes prepare-and-measure scenarios simpler to implement (*21*–*26*). Moreover, prepare-and-measure scenarios are easier to analyze theoretically, which allows

[1]Département de Physique Appliquée, Université de Genève, CH-1211 Genève, Switzerland. [2]Department of Physics, Stockholm University, S-10691 Stockholm, Sweden. [3]MTA Atomki Lendület Quantum Correlations Research Group, Institute for Nuclear Research, Hungarian Academy of Sciences, H-4001 Debrecen, P.O. Box 51, Hungary. *Corresponding author. Email: armin.tavakoli@unige.ch

us to develop self-testing methods that are versatile and highly robust to noise. Third, the assumption of a dimension bound is reasonable for characterization schemes. This is due to the fact that characterization schemes are not adversarial; i.e., they do not involve malicious devices. The experimenter typically knows which degrees of freedom are relevant; for example, the polarization of photons. However, every experiment is subject to unavoidable errors due, e.g., to technical noise and alignment errors. Characterization of quantum devices in this realistic setting is well captured by our assumption of a dimension bound.

In the first part of the paper, we present methods for characterizing nonprojective measurements. First, we present a method for self-testing a targeted nonprojective measurement in noiseless scenarios. Second, since noiseless statistics never occur in practice, we present methods for inferring a lower bound on the closeness of the uncharacterized measurement and a given target POVM, based on the observed noisy statistics; specifically, we lower-bound the worst-case fidelity between the real measurement and the ideal target one. Third, we introduce a method for determining whether the observed statistics could have arisen from some (unknown) projective measurements. If not, the measurement is certified as nonprojective. These methods have twofold relevance. On the one hand, they enable foundational insights into physical inference of nonprojective measurements in a semi–device-independent setting. On the other hand, they provide tools for assessing and certifying the quality of an experimental setup. We demonstrate the practicality of these self-testing methods in two experiments. In the first, we target a symmetric informationally complete (SIC) qubit POVM and demonstrate an estimated 98% worst-case fidelity. In addition, our data certify a genuine four-outcome qubit POVM. In the second experiment, we target a symmetric three-outcome qubit POVM and certify a worst-case fidelity of at least 96%. Last, we discuss some open questions.

THE SELF-TESTING PROBLEM, THE SCENARIO, AND OVERVIEW OF RESULTS

Self-testing is the task of characterizing a quantum system based only on observed data. In other words, it is about gaining knowledge of the physical properties of initially unknown states and/or measurements present in an experiment by studying the correlations observed in the laboratory.

In this work, we focus on prepare-and-measure scenarios. They differentiate themselves from Bell scenarios in two important ways. First, prepare-and-measure scenarios involve communicating observers and thus no space-like separation. Second, they do not involve entanglement, whereas Bell scenarios do. Prepare-and-measure scenarios can generally be modeled by two separated parties, Alice and Bob, who receive random inputs x and y, respectively. Alice prepares and sends a quantum state ρ_x to Bob who performs a measurement y with outcome b, represented by a POVM $\{M_y^b\}_b$ with

$$M_y^b \geq 0 \text{ and } \sum_b M_y^b = \mathbb{1}\ \forall y \tag{1}$$

This generates a probability distribution

$$P(b\,|\,x,y) = \mathrm{tr}(\rho_x M_y^b) \tag{2}$$

To make the problem nontrivial, an assumption on Alice's preparations is required; otherwise, Alice could simply send x to Bob and any probability distribution $P(b\,|\,x,\,y)$ would be achievable. The assumption we consider in this work is that Alice's preparations, i.e., the set of states ρ_x, can be represented in Hilbert space of given dimension d. By choosing $d < |\,x\,|$, we prevent Alice from communicating all information about her input x to Bob. There exist distributions obtained from quantum systems of a dimension d that cannot be simulated classically [see, e.g., (*27*)]. That is, no strategy in which Alice communicates a classical d-valued message to Bob can possibly reproduce the observed data. Such distributions that cannot be classically simulated are candidates for self-testing considerations.

The problem of self-testing consists in characterizing the set of states $\{\rho_x\}$ and/or the set of measurements $\{M_y^b\}$ based only on the distribution $P(b|x,\,y)$. This characterization can usually be done only up to a unitary transformation and possibly a relabeling. In a recent work (*28*), methods were presented for self-testing sets of pure quantum states and sets of projective measurements in the qubit case. These were subsequently extended to higher dimensional systems in (*29*, *30*).

Formally, a self-test can be made via a witness, which is a linear function of the probability distribution $P(b|x,\,y)$

$$\mathcal{A}[P(b\,|\,x,y)] = \sum_{x,y,b} \alpha_{xyb} P(b\,|\,x,y) \tag{3}$$

where α_{xyb} are real coefficients. Moreover, given a witness, one can determine its maximal witness value $\mathcal{A}^Q$ achievable under quantum distributions (Eq. 2) in a bounded Hilbert space. The witness can then be used for self-testing a set of quantum states and/or measurements, whenever there is a unique combination of states and/or measurements that achieves $\mathcal{A}^Q$. Then, it is clear that when the observed distribution $P(b\,|\,x,\,y)$ leads to $\mathcal{A}^Q$, a specific set of states and/or measurements is identified (up to a simple class of transformations). A necessary condition for a witness to be useful for self-testing is that, for a given dimension d, quantum systems outperform classical ones; if not, several strategies would generally be compatible with the data [see (*21*, *21*, *27*) for examples of such witnesses]. In the "Self-testing nonprojective measurements: Noiseless case" section, we present a method for constructing witnesses whose maximal value can self-test a targeted nonprojective qubit measurement $\mathcal{M}^{\text{target}}$.

Next, we turn to robust self-tests, i.e., self-tests that can be applied even when the statistics is not ideal, causing the witness value to be less than $\mathcal{A}^Q$. This is fundamental to make our methods applicable in practice, as any realistic experiment is prone to noise. The influence of noise makes it impossible to perfectly pinpoint the states and measurements. This motivates the following question. Given an observation of a witness value $\mathcal{A} < \mathcal{A}^Q$, how close are the states and measurements to the ideal ones, i.e., those that would have been perfectly self-tested if we had observed $\mathcal{A} = \mathcal{A}^Q$? In the "Robust self-testing of nonprojective measurements" section, we develop methods for robustly self-testing nonprojective qubit measurements by lower-bounding the fidelity between the implemented measurement and the ideal one. A tight robust self-testing would give the fidelity between the measurement that is most distant from the ideal one and that could have generated a witness value $\mathcal{A} < \mathcal{A}^Q$. Since the presented method does not apply to all types of self-tests, we complement it with a numerical method based on random sampling, which efficiently estimates the robustness of self-testing nonprojective qubit measurements.

Whereas robust self-testing represents a quantitative physical inference, it is also relevant to consider a more qualitative inference. On the basis of the witnesses we develop for self-testing, we show

how to certify that the uncharacterized measurement is nonprojective. In the "Certification methods for nonprojective measurements" section, we determine the largest value of our witness that is compatible with qubit projective measurements. When observing a larger value, the nonprojective character of the measurement is certified. In a similar spirit, we determine a bound on our witness above which a genuine four-outcome (nonprojective) qubit measurement is certified.

An overview of all the self-testing methods developed in this work is illustrated in Fig. 1. The methods will be applied in the "Qubit SIC-POVM" section to self-test particularly relevant nonprojective qubit measurements. For these examples, we will demonstrate the usefulness of our methods by implementing them in a photonic experiment. Specifically, our experimental data imply that the implemented measurements are very close to certain ideal three- and four-outcome qubit POVMs and hence are nonprojective. In the latter case, the data certify a genuine four-outcome qubit POVM.

RESULTS

This section presents how to certify and characterize nonprojective measurements in prepare-and-measure scenarios with both noiseless and noisy statistics. The focus will be on qubit systems. Therefore, we begin by summarizing the properties of qubit POVMs.

A POVM with O outcomes is a set of operators $\{E_i\}_{i=1}^{O}$ with the property that $E_i \geq 0$ and that $\sum_i E_i = \mathbb{1}$. In the case of qubits, E_i can be represented on the Bloch sphere as

$$E_i = \lambda_i(\mathbb{1} + \vec{n}_i \cdot \vec{\sigma}) \tag{4}$$

where $\vec{n}_i$ (with $\vec{n}_i \leq 1$) is the Bloch vector, $\lambda_i \geq 0$, and $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the Pauli matrices. Positivity and normalization imply that

$$\sum_{i=1}^{O} \lambda_i = 1 \text{ and } \sum_{i=1}^{O} \lambda_i \vec{n}_i = 0 \tag{5}$$

The set of POVMs is convex, and a POVM is called extremal if it cannot be decomposed as a convex mixture of other POVMs. For qubits, extremal POVMs have either $O = 2,3,4$ outcomes (*31*). In the case $O = 2$, extremal POVMs are simply projective, whereas for $O = 3$ and $O = 4$, they are nonprojective; an extremal three-outcome qubit POVM has three unit Bloch vectors in a plane, and an extremal four-outcome qubit POVM has four unit Bloch vectors of which no choice of three are in the same plane (*31*). An extremal qubit POVM is therefore characterized by its Bloch vectors. As the statistics of nonextremal POVMs can always be simulated by stochastically implementing extremal POVMs, it is clear that only extremal POVMs can be self-tested.

Self-testing nonprojective measurements: Noiseless case

Consider a target extremal nonprojective qubit POVM $\mathcal{M}^{\text{target}}$, with $O = 3$ or $O = 4$ outcomes, for which we associate the outcome b to the unit Bloch vector $\vec{v}_b$. Our goal is now to construct a witness $\mathcal{A}$ such that its maximal value self-tests $\mathcal{M}^{\text{target}}$. The method consists of two steps summarized in Fig. 2.

Step 1. First, we construct a simpler witness $\mathcal{A}'$ featuring O preparations; i.e., Alice has O inputs. Bob receives an input $y = 1, \ldots, Y$ and provides a binary outcome. The goal of this simpler witness is to self-test a particular relation among the prepared states $|\psi_x\rangle$. Specifically, we would like to certify that their unit Bloch vectors $\vec{u}_x$ point in opposite direction (on the Bloch sphere) to those of the target POVM $\mathcal{M}^{\text{target}}$; i.e., $\vec{u}_x = -\vec{v}_x$ for $x = 1, \ldots, O$. Let us define

$$\mathcal{A}' = \sum_{x,y,b} c_{xyb} P(b \mid x, y) \tag{6}$$

with real coefficients c_{xyb} chosen such that the maximal value $\mathcal{A}'^{Q}$ of the witness for qubits self-tests the desired set of prepared states $\{|\psi_x\rangle\}$ (up to a global unitary and relabelings). In general, we believe that it is always possible to find such a self-test by considering enough inputs for Bob, corresponding to well-chosen projective measurements, and suitable coefficients c_{xyb} [see (*28*) for examples]. Furthermore,

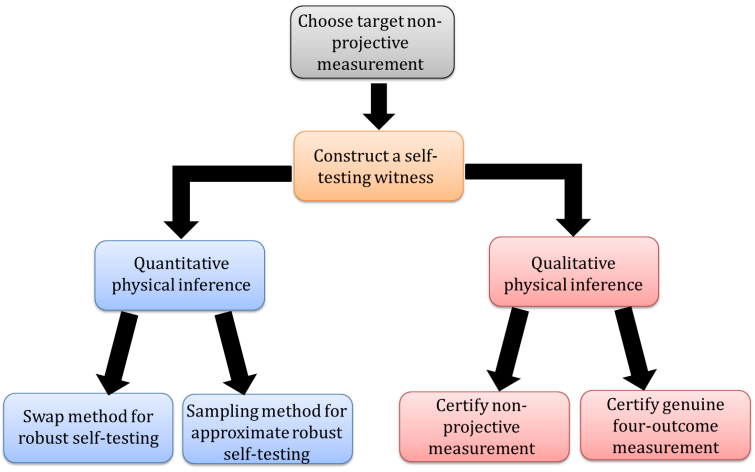

Fig. 1. Graphical overview of the self-testing methods and steps presented in Results.

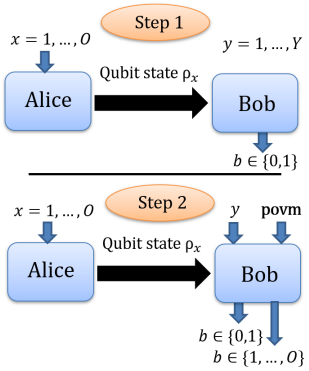

Fig. 2. Method for self-testing a targeted nonprojective qubit measurement by exploiting simpler self-tests of preparations. Step 1: tailor scenario and witness such that a maximal $\mathcal{A}'$ self-tests Alice's preparations to have Bloch vectors that are anti-aligned with those of the target measurement. Step 2: Add an extra setting to Bob and modify the witness to self-test the target non-projective measurement.

note that one could also, in principle, have more than O preparations for Alice and then self-test that O of them have the desired relation to $\mathcal{M}^{\text{target}}$. In addition, we remark that the construction of an adequate witness $\mathcal{A}'$ is not unique in general.

Step 2. We construct our final witness $\mathcal{A}$ from $\mathcal{A}'$. Specifically, we supply Bob with one additional measurement setting called **povm**. This setting corresponds to a measurement with O outcomes. Since the intention is to self-test the measurement corresponding to this setting as $\mathcal{M}^{\text{target}}$, we associate the setting **povm** to O outcomes. We define

$$\mathcal{A} = \mathcal{A}' - k\sum_{x=1}^{O} P(b = x \mid x, \textbf{povm}) \tag{7}$$

for some positive constant k. A maximal witness value $\mathcal{A}^Q = \mathcal{A}'^Q$ now implies that the setting **povm** corresponds to $\mathcal{M}^{\text{target}}$ (up to a unitary and relabelings). This is because a maximal witness value implies that (i) the set of prepared states $\{|\psi_x\rangle\}$ have Bloch vectors anti-aligned with those of $\mathcal{M}^{\text{target}}$ and (ii) $P(b = x \mid x, \textbf{povm}) = 0$ for all x; hence, the Bloch vectors of the setting **povm** are of unit length and aligned with those of $\mathcal{M}^{\text{target}}$. Moreover, as a qubit POVM is characterized by its Bloch vectors, we see that $\mathcal{M}^{\text{target}}$ is the only POVM that can attain the maximal witness value $\mathcal{A}^Q$. Therefore, we obtain a self-test of the target POVM $\mathcal{M}^{\text{target}}$.

In the "Qubit SIC-POVM" section, we will apply this method to self-test symmetric qubit POVMs with three and four outcomes.

Robust self-testing of nonprojective measurements

No experiment can achieve the noiseless conditions needed to obtain exactly a maximal value of $\mathcal{A}$. Therefore, it is paramount to discuss the case when a nonmaximal value of $\mathcal{A}$ is observed. We will show that, in this case, one can nevertheless make a statement about how close the uncharacterized measurement $\mathcal{E}$ performed in the laboratory (corresponding to the setting **povm**) is to the target POVM $\mathcal{M}^{\text{target}}$.

To address this question, we must first define a measure of closeness between two measurements. A natural and frequently used distance measure in quantum information is the fidelity, F, between two operators. We consider a measure of closeness amounting to the best possible weighted average fidelity between the extremal qubit target POVM elements $M^{\text{target}} = \{M_i\}$ and the actual POVM elements $\mathcal{E} = \{E_i\}$. That is, we allow for a quantum extraction channel Λ to be applied to the actual POVM. The set of allowed extraction channels is the set of unital channels in the relevant Hilbert space dimension. This is understood from the fact that the extraction channel must map O-outcome POVMs to O-outcome POVMs in the given Hilbert space dimension. Because of linearity, this implies that the channel is unital. Conversely, since every channel preserves positivity, every unital channel in the relevant Hilbert space dimension maps POVMs to POVMs. We look for the best possible extraction channel. We thus define the quantity

$$F(E, M^{\text{target}}) = \max_{\Lambda} \frac{1}{2}\sum_{i=1}^{O} \frac{tr\,(\Lambda[E_i]M_i)}{tr\,(M_i)} \tag{8}$$

Since the target measurement is extremal, the POVM elements are proportional to rank-one projectors; $M_i \propto P_i$. Because of Eq. 4, we can write $\Lambda[E_i] = \lambda_i(\mathbb{1} + \vec{n}_i \cdot \vec{\sigma})$ subject to the constraints (Eq. 5). By evaluating Eq. 8, we find that $F = 1/2 + 1/2\sum_i \lambda_i \mathrm{tr}\,(P_i \vec{n}_i \cdot \vec{\sigma}_i) \leq 1$. To saturate the inequality, each Bloch vector $\vec{n}_i$ must be of unit length, i.e., $|\vec{n}_i| = 1$, and aligned with the Bloch vector of P_i. Hence, M_i and $\Lambda[E_i]$ are both proportional to the same rank-one projector. Since a POVM with Bloch vectors of unit length is fully characterized, i.e., all coefficients λ_i are fixed by the conditions (Eq. 5), this implies that $M_i = \Lambda[E_i]$. Thus, a maximal fidelity of $F = 1$ is uniquely achieved when the actual POVM is equal to the target measurement.

In general, a nonmaximal value of the witness $\mathcal{A}$ can arise from many different possible choices of states and measurements. We denote by $S(\mathcal{A})$ the set of all O-outcome POVMs that are compatible with a given observed value $\mathcal{A}$. Our goal is now to find a lower-bound on the average fidelity F that holds for every measurement $\mathcal{E}' \in S(\mathcal{A})$. Therefore, the quantity of interest is the worst-case average fidelity:

$$\mathcal{F}(\mathcal{A}) = \min_{\mathcal{E}' \in S(\mathcal{A})} F(\mathcal{E}', \mathcal{M}^{\text{target}}) \tag{9}$$

Calculating this quantity, or even lower-bounding it, is typically a nontrivial problem even in the simplest case. We proceed with presenting two methods for this task.

We remark that the definition (Eq. 8), given for qubits, could potentially be extended to higher-dimensional systems (replacing the factor 1/2 by $1/d$). This could work for POVMs where all elements are proportional to rank-one projectors. However, the latter are only a strict subset of general extremal POVMs. Finding a more general figure of merit is thus an interesting open question.

Robust self-testing with the swap method

A lower-bound on the worst-case average fidelity can be obtained via semidefinite programming (*32*). The method combines the so-called swap method (*33*, *34*), introduced for self-testing in the Bell scenario, and the hierarchy of dimensionally bounded quantum correlations (*35*). Such adaptations of the swap method to prepare-and-measure scenarios were introduced in (*28*) to self-test pure state and projective measurements. In section S1, we outline the details of how the swap method is adapted to robustly self-test nonprojective measurements. This method benefits from being applicable in a variety of scenarios and for returning rigorous lower bounds on $\mathcal{F}$. Nevertheless, it suffers from two drawbacks. First, the method only overcomes the fact that self-tests are valid up to a global unitary, but

not that they may be valid up to relabelings. Thus, it is only useful for target measurements that are self-tested up to a unitary. Second, while rarely producing tight bounds on $\mathcal{F}$, the computational requirements scale rapidly with the number of inputs, the number of outputs, and the chosen level of the hierarchy. In the "Qubit SIC-POVM" section, we will show that the method can be efficiently applied for robustly self-testing a three-outcome qubit POVM.

Numerically approximating robust self-testing

To also address cases in which self-tests are valid up to both a unitary transformation and relabelings, we can estimate $\mathcal{F}$ based on random sampling. The approximation method benefits from being straightforward and broadly useful, while it suffers from the fact that it merely estimates the value of $\mathcal{F}$ instead of providing a strict lower bound. The key feature is that the minimization appearing in Eq. 9 is replaced by a minimization taken over data obtained from many random samples of the setting **povm**. We detail this method in section S2 and apply it to an example in the "Qubit SIC-POVM" section.

Certification methods for nonprojective measurements

Whereas robust self-testing considers quantitative aspects of physical inference from noisy data, it is important to also consider the qualitative inference. An important qualitative statement is to prove that the uncharacterized measurement is nonprojective or, more generally, that it cannot be simulated by projective measurements. It is known that when POVMs are sufficiently noisy, they become perfectly simulable via projective measurements (*19*, *36*, *37*). The witnesses we construct can address this question. We will see that whenever the observed value of the witness $\mathcal{A}$ is sufficiently large, one can certify that the setting **povm** necessarily corresponds to some nonprojective measurement and could not have been simulated via projective measurements. Specifically, we derive an upper bound on $\mathcal{A}$ for projective measurements (or convex combination of them). The violation of such a bound thus certifies a nonprojective measurement or, more precisely, a genuine three-outcome (or four-outcome) POVM. At the end of this subsection, we also show how to certify a genuine four-outcome POVM.

A projective qubit measurement has binary outcomes and can therefore be represented by an observable $M \equiv M_0 - M_1$, where M_i is the measurement operator corresponding to outcome $i = 0,1$. Let us consider the case where the O-outcome measurement **povm** is projective. One may assign two outcomes to rank-one projectors and the rest to trivial zero operators. Note that it is enough here to consider these cases, as the witness $\mathcal{A}$ is linear in terms of the measurement operators. Projectors can thus be assigned in three ($O = 3$) or six ($O = 4$) different ways, of which the optimal instance must be chosen. Let the outcomes in the optimal instance be $o_{0|\mathbf{povm}}$ and $o_{1|\mathbf{povm}}$ and associate the observable $M_{\mathbf{povm}} \equiv M_{Y+1} = M_{\mathbf{povm}}^{o_{0|\mathbf{povm}}} - M_{\mathbf{povm}}^{o_{1|\mathbf{povm}}}$. The witness (Eq. 7) can be written as

$$\mathcal{A} = C(k) + \sum_x \mathrm{tr}[\rho_x \mathcal{L}_x^{(k)}(\{M_y\})] \tag{10}$$

where $C(k)$ is a constant and $\mathcal{L}_x^{(k)}(\{M_y\})$ is a linear combination of the observables $\{M_1, \ldots, M_{Y+1}\}$. Note that $\mathcal{L}_x^{(k)}(\{M_y\})$ does not depend on the index y but on the collection of observables. Using the Cauchy-Schwarz inequality for operators, we obtain

$$\mathcal{A} \leq C(k) + \sum_x \sqrt{\mathrm{tr}[\rho_x \mathcal{L}_x^{(k)}(\{M_y\})^2]} \tag{11}$$

Because of projectivity, we have $M_y = \vec{n}_y \cdot \vec{\sigma}$, where $\vec{n}_y$ is of unit length. Using $\{M_k, M_l\} = 2\vec{n}_k \cdot \vec{n}_l \mathbb{1}$, one finds $\mathcal{L}_x^{(k)}(\{M_y\})^2 = t_x^{(k)}(\{\vec{n}_y\})\mathbb{1}$, for some function t, which is a weighted sum of scalar products of the Bloch vectors of the observables. Consequently, to bound $\mathcal{A}$ under all projective measurements, we have

$$\mathcal{A} \overset{\text{Proj}}{\leq} C(k) + \max_{\{\vec{n}_y\}} \sum_x \sqrt{t_x^{(k)}(\{\vec{n}_y\})} \equiv \mathcal{B}(k) \tag{12}$$

Thus, $\mathcal{B}(k)$ bounds the value of $\mathcal{A}$ for projective measurements. The evaluation of this bound only depends on Bob's Bloch vectors and is further simplified by their parameterization in terms of two angles. The effort needed to evaluate the bound depends on the chosen prepare-and-measure scenario. Typically, considering scenarios with some symmetry properties is beneficial.

Moreover, when targeting a four-outcome qubit POVM, we consider also a finer form of qualitative characterization by considering whether $\mathcal{A}$ can be simulated by the setting **povm** being some three-outcome POVM. If not, the measurement is certified as a genuine four-outcome measurement. This amounts to bounding the value of $\mathcal{A}$ achievable under any two- or three-outcome qubit POVM and then observing a violation of that bound. For this purpose, one may use the hierarchy of dimensionally bounded quantum correlations (*35*), which can be used to upper-bound $\mathcal{A}$ under three-outcome POVMs. Since the hierarchy is built on projective measurements, one must embed Alice's preparations in a larger Hilbert space with the dimension chosen such that three-outcome POVMs can be recast as projective measurement following Neumark's theorem. To obtain tight bounds, one may need a reasonably high hierarchy level, which can be efficiently implemented using the methods of (*30*).

Next, in the "Qubit SIC-POVM" section, we will apply the outlined methods to specific nonprojective measurements and experimentally demonstrate the certification of both nonprojective and genuine four-outcome measurements.

Relevant examples and their experimental realization

In the above, we have discussed methods for self-testing a target nonprojective measurement. Here, we put these methods in practice in a photonic experiment. We implement three- and four-outcome symmetric qubit POVMs, with Bloch vectors forming a star (trine-POVM) and a tetrahedron (SIC-POVM), respectively. In the first case, we certify a nonprojective measurement and apply our methods for robust self-testing, demonstrating worst-case average fidelity of at least 96% compared to an ideal trine-POVM. In the second case, we certify a genuine four-outcome qubit POVM and demonstrate worst-case average fidelity of approximately 98% with respect to an ideal SIC-POVM. We consider each example separately by first applying the methods of Results to obtain adequate witnesses and then present the corresponding experimental realization. The setup common to both experiments is presented in Materials and Methods.

Qubit SIC-POVM

We begin by illustrating the self-testing methods for a frequently used nonprojective measurement, namely, the qubit SIC-POVM, which we denote $\mathcal{M}_{\mathrm{SIC}}$. This measurement has four outcomes, and its four unit Bloch vectors $\{\vec{v}_b\}_b$ form a regular tetrahedron on the Bloch sphere, with weights $\lambda_b = 1/4$. Such a regular tetrahedron construction can be achieved via two different labelings of the four outcomes that are not equivalent under unitary transformations.

Up to a unitary transformation, each such SIC-POVM can be written with Bloch vectors

$$\vec{v}_1 = [1,1,1]/\sqrt{3} \qquad \vec{v}_2 = [1,-1,-1]/\sqrt{3}$$
$$\vec{v}_3 = [-1,1,-1]/\sqrt{3} \quad \vec{v}_4 = [-1,-1,1]/\sqrt{3} \tag{13}$$

and the set of Bloch vectors $\{-\vec{v}_l\}_l$, respectively.

Noiseless self-test

We find a prepare-and-measure scenario for self-testing $\mathcal{M}_{\mathrm{SIC}}$. Following step 1 in the "Self-testing nonprojective measurements: Noiseless case" section, we introduce a prepare-and-measure scenario in which Alice has four preparations, $x \in \{1,2,3,4\}$, and Bob has three binary-outcome measurements, $y \in \{1,2,3\}$. The witness is chosen as

$$\mathcal{A}'_{\mathrm{SIC}} = \frac{1}{12}\sum_{x,y} P(b = S_{x,y}\,|\,x,y) \tag{14}$$

where $S_{1,y} = [0,0,0]$, $S_{2,y} = [0,1,1]$, $S_{3,y} = [1,0,1]$, and $S_{4,y} = [1,1,0]$. The maximal value, $\mathcal{A}'_{\mathrm{SIC}} = 1/2(1 + 1/\sqrt{3})$, can be achieved by Alice preparing her four states forming a regular tetrahedron, e.g., with the Bloch vectors in Eq. 13, and Bob performing the measurements σ_x, σ_y, and σ_z. In section S3, we prove the maximal witness value and show that it self-tests that Alice's preparations indeed must form a regular tetrahedron on the Bloch sphere. By step 2 in the "Self-testing nonprojective measurements: Noiseless case" section, we supply Bob with an additional four-outcome measurement **povm** and consider the modified witness

$$\mathcal{A}_{\mathrm{SIC}} = \frac{1}{12}\sum_{x,y} P(b = S_{x,y}\,|\,x,y) - k\sum_{x=1}^{4} P(b = x\,|\,x,\mathbf{povm}) \tag{15}$$

Thus, we conclude that $\mathcal{A}_{\mathrm{SIC}} = 1/2(1 + 1/\sqrt{3})$ self-tests $\mathcal{M}_{\mathrm{SIC}}$.

We note that there also exist other prepare-and-measure scenarios fulfilling the requirements of step 1. For example, one may achieve the desired self-test using the so-called $3 \to 1$ random access code whose self-testing properties were considered in (*28*). However, this prepare-and-measure scenario requires more preparations than the one presented here.

Robust self-test

Next, we consider the worst-case fidelity (given in Eq. 9) of the measurement corresponding to the setting **povm** with $\mathcal{M}_{\mathrm{SIC}}$. Since the self-test of $\mathcal{M}_{\mathrm{SIC}}$ is valid up to a relabeling and a collective unitary, we cannot use the swap method to lower-bound $\mathcal{F}$. Instead, we use the numerical approximation method (see section S2 for details). Figure 3 displays roughly 3×10^5 optimal pairs $(\mathcal{A}_{\mathrm{SIC}}, F)$ each evaluated from a randomly sampled measurement for the setting **povm**. The evaluation was done for $k = 1/5$ (which, as will soon be shown, turns out to be the most noise-resilient choice of k). We see that the minimal sampled fidelity as a function of $\mathcal{A}_{\mathrm{SIC}}$ describes a curve, which constitutes the approximation of $\mathcal{F}$.

Certifying nonprojective and genuine four-outcome POVMs

Last, we derive a tight bound valid for all qubit projective measurements on the value of $\mathcal{A}_{\mathrm{SIC}}$. Because of the symmetries of $\mathcal{A}_{\mathrm{SIC}}$, we can, without loss of generality, let the nontrivial (nonzero measurement operator) outcomes of the measurement **povm** be the outcomes $b = 1,2$. Hence, we define the observable $M_{\mathbf{povm}} \equiv M_4 = M^1_{\mathbf{povm}} - M^2_{\mathbf{povm}}$. Then, we follow the steps outlined in the "Certification methods for nonprojective measurements" section. First, we re-write $\mathcal{A}_{\mathrm{SIC}}$ in the form of Eq. 10. We find $C(k) = (1 - 2k)/2$ and

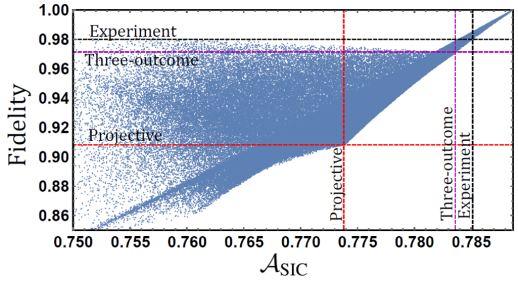

Fig. 3. Numerical approximation of the worst-case fidelity of the unknown measurement (setting povm) with the qubit SIC-POVM by roughly 3×10^5 random three- and four-outcome POVM samples for which the optimal values of $(\mathcal{A}, \mathcal{F})$ were calculated. The figure also displays the critical limits on $\mathcal{A}_{\mathrm{SIC}}$ and $\mathcal{F}$ for projective and three-outcome POVMs, respectively, as well as the experimentally measured values.

$$\mathcal{L}^{(k)}_{x=0,1}(\{M_y\}) = \frac{1}{24}[\,1,(-1)^x,(-1)^x,(-1)^{x+1}\,12k]\cdot\vec{M}$$
$$\mathcal{L}^{(k)}_{x=2,3}(\{M_y\}) = \frac{1}{24}[\,-1,(-1)^x,(-1)^{x+1},0]\cdot\vec{M} \tag{16}$$

where $\vec{M} = [M_1, M_2, M_3, M_4]$, with $M_y = \vec{n}_y \cdot \vec{\sigma}$. After applying the Cauchy-Schwarz inequality, we obtain a cumbersome expression of the form of Eq. 11. To evaluate its maximal value (following Eq. 12), we use the following concavity inequality: $\sqrt{r} + \sqrt{s} \le \sqrt{2(r+s)}$ for $r, s \ge 0$, with equality if and only if $r = s$. Apply this inequality twice to the expression (Eq. 12), first to the two terms associated to $x = 0,1$, and then to the two terms associated to $x = 2,3$. After a simple optimization over $\vec{n}_3$ and denoting $x = \vec{n}_1 \cdot \vec{n}_2$, one arrives at

$$\mathcal{A}_{\mathrm{SIC}} \le \frac{1-2k}{2} + \frac{\sqrt{2}}{24}\sqrt{6-4x} + \frac{\sqrt{2}}{24}\sqrt{2\,r_k + 4x + 48k\sqrt{2}\,\sqrt{1+x}} \equiv f_k(x)$$

where $r_k = 3 + 144k^2$. This bound is valid for a particular value of x. To hold for all projective measurements, we simply maximize $f_k(x)$ over x. This requires only an optimization in a single real variable $x \in [-1,1]$, which is straightforward. The optimal choice is denoted x^*. Setting $\mathcal{B}(k) = f_k(x^*)$, we have $\mathcal{A}_{\mathrm{SIC}} \le \mathcal{B}(k)$ for all projective measurements. Although the expressions involved are cumbersome, the analysis is simple and straightforward. We have considered the tightness of the projective bound for $k \in \{1/100, 2/100, \ldots, 1\}$ by numerically optimizing $\mathcal{A}_{\mathrm{SIC}}$ under unit-trace measurements (which includes all rank-one projective measurements). In all cases, we saturate the bound $\mathcal{B}(k)$ up to machine precision with a projective measurement.

Furthermore, we have also considered bounding $\mathcal{A}_{\mathrm{SIC}}$ under three-outcome qubit POVMs using the hierarchy of dimensionally bounded quantum correlations (as described in the "Certification methods for nonprojective measurements" section). In our implementation of (*35*), we have embedded the qubit preparations into a three-dimensional Hilbert space and optimized $\mathcal{A}_{\mathrm{SIC}}$ under projective measurements of the only existing nontrivial rank combination. The relaxation level involved some monomials from both the second and third level, and the size of the moment matrix was 126. This was

done for all $k \in \{1/100, 2/100 \ldots, 1\}$, and each upper bound was saturated up to numerical precision using lower bounds numerically obtained via semidefinite programs.

To study the robustness of both the nonprojective and the genuine four-outcome certification, we have considered the critical visibility of the system needed when exposed to noise. This is modeled by the preparations taking the form $\rho_x(v) = v\rho_x + (1 - v)\rho_{noise}$, where $v \in [0,1]$ is the visibility and ρ_{noise} is some arbitrary qubit state. A straightforward calculation shows that the critical visibility for violating some given bound $\mathcal{B}$ is

$$v_{crit}(k) = \frac{\mathcal{B}(k) - \mathcal{A}^{rand} + k}{\mathcal{A}^{Q} - \mathcal{A}^{rand} + k} \tag{17}$$

where $\mathcal{A}^{rand}$ is the witness value obtained from the optimal measurements performed on the maximally mixed state. Notably, this expression is independent of the specific form of ρ_{noise}. We have applied this to $\mathcal{A}_{SIC}$ with $\mathcal{B}(k)$, corresponding to the bounds on projective and three-outcome measurements, respectively. The corresponding critical visibilities are plotted in section S5. In both cases, we find that the largest amount of noise is tolerated for $k = 1/5$, corresponding to $v_{crit} = 0.970$ and $v_{crit} = 0.990$, respectively.

Experimental result

Wave-plate settings for Alice's prepared states in Eq. 13 and Bob's measurements σ_x, σ_y, σ_z, and the four-outcome SIC-POVM anti-aligned to the vectors in Eq. 13, are reported in section S5. In section S5, we also report a state tomography of Alice's preparations.

Optimally choosing $k = 1/5$, the measured value of the witness as compared to the relevant bounds is

$$\begin{aligned} A_{SIC} &\overset{\text{projective}}{\leq} 0.7738 \overset{\text{3-outcome}}{\leq} 0.7836 \overset{\text{qubit}}{\leq} 0.7887. \\ A_{SIC}^{Lab} &= 0.78514 \pm 5 \times 10^{-5}_{stat} \pm 1.0 \times 10^{-4}_{syst} \end{aligned} \tag{18}$$

The statistical error originates from Poissonian statistics, and the systematic error originates from the precision of the wave-plate settings. More details about the errors are discussed in section S5.

We observe a substantial violation of both the projective measurement and the three-outcome measurement bounds. Thus, we can certify that Bob's measurement **povm** is a genuine four-outcome qubit POVM. Furthermore, as illustrated by the results in Fig. 3, we certify approximately a 98% worst-case fidelity with the qubit SIC-POVM.

Qubit trine-POVM

We consider a second example in which the target POVM is the so-called trine-POVM. This measurement has three outcomes, and its Bloch vectors form an equilateral triangle on a disk of the Bloch sphere, with $\lambda_l = 1/3$. The Bloch vectors are hence defined by

$$\vec{v}_1 = [0, 0, -1], \vec{v}_2 = \frac{1}{2}[-\sqrt{3}, 0, 1], \vec{v}_3 = \frac{1}{2}[\sqrt{3}, 0, 1] \tag{19}$$

Noiseless self-test

We introduce a prepare-and-measure scenario in which Alice has three inputs $x \in \{1,2,3\}$, and Bob has two binary-outcome measurements labeled by $y \in \{1,2\}$, and consider the witness

$$\mathcal{A}'_{tri} = \sum_{x,y,b} T_{x,y}(-1)^b P(b \,|\, x, y) \tag{20}$$

where $T_{x,1} = [1, 1, -1]$ and $T_{x,2} = [\sqrt{3}, -\sqrt{3}, 0]$. In section S3, we show that its maximal value is $\mathcal{A}'_{tri} = 5$, and that this value implies that Alice's three preparations form an equilateral triangle on the Bloch sphere. Then, we add an additional input **povm** for Bob and consider the witness

$$\mathcal{A}_{tri} = \sum_{x,y,b} T_{x,y}(-1)^b P(b \,|\, x, y) - k \sum_{x=1}^{3} P(b = x \,|\, x, \mathbf{povm}) \tag{21}$$

for some $k > 0$. Then, $\mathcal{A}_{tri} = 5$ self-tests the setting **povm** as the trine-POVM up to a unitary.

Robust self-test

We now turn to considering its robust self-testing properties, i.e., lower-bounding the worst-case fidelity of the unknown measurement (setting **povm**) with the target measurement for a given value of $\mathcal{A}_{tri}$. Since the above self-test is achieved only up to unitary transformations, we may find rigorous lower bounds on the worst-case fidelity $\mathcal{F}$ using semidefinite programming. In accordance with the "Robust self-testing of nonprojective measurements" section, we have performed the swap-operation on Bob's side and used the hierarchy of finite-dimensional correlations to lower-bound $\mathcal{F}$. The hierarchy level was an intermediate level containing some higher-order moments corresponding to an SDP matrix of size 105. In addition, for the sake of comparison, we have implemented the numerical approximation method for robust self-testing to estimate the accuracy of the bound obtained via the swap method. The results are shown in Fig. 4. A comparison suggests that the swap method returns a suboptimal bound. Its accuracy could potentially be improved by using a higher hierarchy level. Nevertheless, the obtained bound will prove sufficient for the practical purpose of experimentally certifying the targeted POVM with high accuracy.

Last, we have also self-tested the trine-POVM in a different prepare-and-measure scenario (see section S3). In section S4, we use this prepare-and-measure scenario to derive a tight bound on projective measurements by evaluating the right-hand side of Eq. 12.

Experimental realization

The witness in Eq. 21 is maximized if Alice's three Bloch vectors point to the vertices of an equilateral triangle on a disk of the Bloch sphere. We take that disk to be the xz plane, taking $\vec{t}_i = -\vec{v}_i$ (from Eq. 19), and Bob performs one of three measurements σ_z, σ_x, and the three-outcome POVM with vectors anti-aligned to Alice's states. See section S5 for state tomography of Alice's preparations. In contrast to the previous experiment, output 2 of Bob's measurement station only consists of one detector (D3) and no wave plate or polarizing beamsplitter (PBS) (see Fig. 5). The wave-plate settings corresponding to the above states and measurements are reported in section S5.

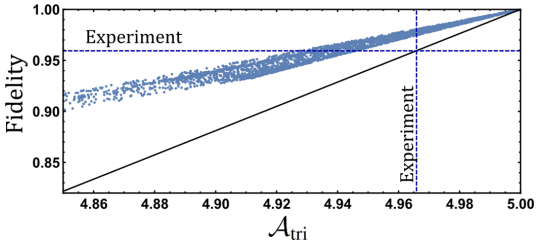

Fig. 4. Lower bound on $\mathcal{F}(\mathcal{A}_{tri})$ for k=1 obtained from the swap method, together with roughly 3000 points $(\mathcal{A}_{tri}, \mathcal{F})$ obtained via the numerical approximation method. This is displayed next to the experimentally achieved results.

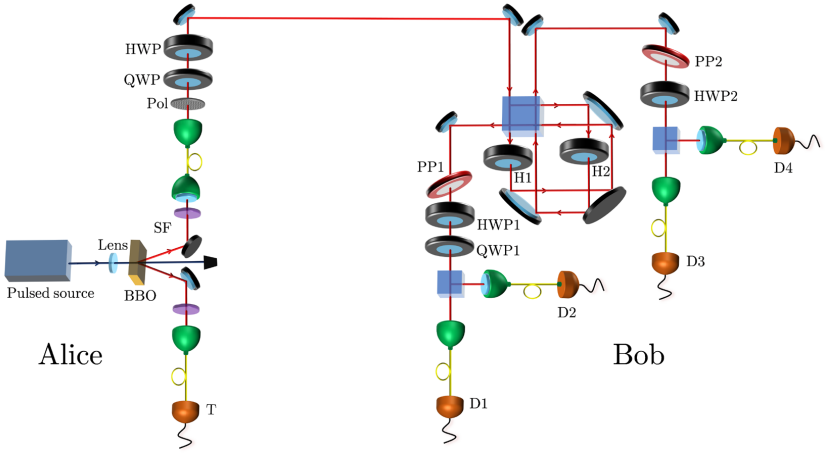

Fig. 5. Experimental setup. More details, including labeling, can be found in the main text. Pol, polarizer.

With the said settings, we have obtained the experimentally measured value of $\mathcal{A}_{\text{tri}}$ as a function of k. Since we aim to demonstrate a large worst-case fidelity with the trine-POVM, we have computed the lower bound on $\mathcal{F}(\mathcal{A}_{\text{tri}})$ for many different values of k and found that choosing $k = 1$ leads to the optimal result. The corresponding experimentally measured witness is

$$\mathcal{A}_{\text{tri}}(k = 1) \overset{\text{projective}}{\leq} 4.89165 \overset{\text{qubit}}{\leq} 5 \tag{22}$$

$$\mathcal{A}_{\text{tri}}^{\text{Lab}}(k = 1) = 4.9659 \pm 7 \times 10_{\text{stat}}^{-4} \pm 1.7 \times 10_{\text{syst}}^{-3} \tag{23}$$

This data point and its relation to the worst-case fidelity of the laboratory measurement with the targeted POVM are depicted in Fig. 4. From $\mathcal{A}_{\text{tri}}^{\text{Lab}}$, we infer a closeness of at least 96%. This can be compared to the largest possible fidelity between a projective measurement and the trine-POVM, which is straightforwardly found to be $(2 + \sqrt{3})/4 \approx 0.933$. However, as indicated by the results of the sampling-based numerical approximation method for robust self-testing (presented in Fig. 4), a better bound of $\mathcal{F}$ may allow us to rigorously infer a worst-case fidelity of at least 97.3%.

Furthermore, we have considered the possibility of the experimental data certifying a nonprojective qubit measurement. However, to this end, we found that another choice of k is optimal with respect to the witness value that is achievable under projective measurements. We found that the optimal choice is $k \approx 4.5$. The corresponding experimentally measured value becomes

$$\begin{aligned} &\mathcal{A}_{\text{tri}}(k = 4.5) \overset{\text{projective}}{\leq} 4.71139 \overset{\text{qubit}}{\leq} 5 \\ &\mathcal{A}_{\text{tri}}^{\text{Lab}}(k = 4.5) = 4.93613 \pm 5 \times 10_{\text{stat}}^{-5} \pm 1.0 \times 10_{\text{syst}}^{-4} \end{aligned} \tag{24}$$

We conclude that our experimental data certifies a nonprojective qubit measurement.

DISCUSSION

We investigated the problem of self-testing nonprojective measurements. We argued that a prepare-and-measure scenario with an upper bound on the Hilbert space dimension represents a natural framework for investigating this problem. We considered both the qualitative certification of a measurement being nonprojective and/or genuine four-outcome, as well as a quantitative characterization in terms of worst-case fidelity to a given target POVM. We demonstrate the practical relevance of these methods in two experiments in which we both certify a genuine four-outcome POVM and infer a high worst-case fidelity with respect to target symmetric qubit POVMs.

It would be interesting to overcome the limitation of the swap method and develop a rigorous robust self-testing method for general four-outcome qubit POVMs. Also extending these methods to high-dimensional POVMs would be relevant since there exist extremal nonprojective measurements that feature the same number of outcomes as projective measurements (contrary to the qubit case). Moreover, it would be interesting to investigate self-testing of nonprojective measurements using different assumptions as in our work. One could consider for instance prepare-and-measure scenarios with a bound on the entropy (*38*), the overlap between the prepared states (*8*), or their mean energy (*39*). Last, one may ask whether it would be possible to robustly self-test a nonprojective measurement in the fully device-independent case, i.e., returning to the Bell scenario without any assumption on the dimension.

MATERIALS AND METHODS

In the experiment, the qubit states are encoded in the polarization degree of freedom of a single photon, with the convention of $|H\rangle \equiv |0\rangle$ and $|V\rangle \equiv |1\rangle$. The setup is depicted in Fig. 5.

Alice's station includes a heralded single-photon source where femtosecond laser pulses at 390 nm are converted into pairs of photons at 780 nm, through type I spontaneous parametric down-conversion in two orthogonally oriented beta-barium borate crystals. Photon

pairs go through 3-nm spectral filters and are then coupled into two single-mode fibers for spatial mode filtering. The idler photon is sent to the trigger avalanche photodiode (APD) detector (T) and heralds the presence of a signal photon. The latter is then emitted again into free space and undergoes Alice's state preparation, consisting of a fixed linear polarizer, a λ/4 wave plate [or quarter–wave plate (QWP)], and a λ/2 wave plate [or half–wave plate (HWP)].

Upon preparing the required qubit state, Alice forwards the signal photon to Bob's measurement station, where it goes through a double-path Sagnac interferometer, each path of which contains an HWP. The interferometer mixes the polarization degree of freedom with path, effectively enabling Bob to perform either projective or nonprojective measurements in the original polarization Hilbert space where the qubit was prepared, thanks to the two polarization analyzers at the outputs. Each of these consists of a phase plate, an HWP, and (in output 1) a QWP, a polarizing beam splitter and two single-photon detectors. Outputs from all detectors (T and D1 to D4) are sent to a coincidence unit connected to a computer.

All measurements were performed with heralded photon rates of approximately 1×10^4 counts per second, while each setting was measured for 500 s. We have made an assumption of fair sampling, i.e., that the detection events are representative of the total number of signal photons. This assumption is reasonable for tasks that do not include a notion of an adversary. The quality of state preparation and measurement can be estimated by preparing states $|H\rangle$, $|+\rangle = (|H\rangle + |V\rangle)/\sqrt{2}$, and $|R\rangle = (|H\rangle + i\,|V\rangle)/\sqrt{2}$ and measuring them in the Pauli bases σ_z, σ_x, and σ_y, respectively. The three visibilities obtained in our setup with this characterization measurement were

$$\begin{aligned} V_{\sigma_z} &= (99.91 \pm 0.02)\,\% \\ V_{\sigma_x} &= (99.31 \pm 0.01)\,\% \\ V_{\sigma_y} &= (99.23 \pm 0.02)\,\% \end{aligned} \tag{25}$$

While the almost optimal V_{σ_z} is a direct consequence of the high extinction ratios of the PBSs used, the lower visibilities in the interference bases are mainly due to the double-path Sagnac interferometer, which showed a visibility of around 99.4%, therefore effectively bounding from above the results we can achieve in the experiments.

Note added. During the completion of this manuscript, we became aware of an independent work (*40*) discussing the certification of qubit POVMs.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/6/16/eaaw6664/DC1

REFERENCES AND NOTES

1. R. Derka, V. Bužek, A. K. Ekert, Universal algorithm for optimal estimation of quantum states from finite ensembles via realizable generalized measurement. *Phys. Rev. Lett.* **80**, 1571–1575 (1998).
2. J. M. Renes, R. Blume-Kohout, A. J. Scott, C. M. Caves, Symmetric informationally complete quantum measurements. *J. Math. Phys.* **45**, 2171–2180 (2004).
3. J. Shang, A. Asadian, H. Zhu, O. Gühne, Enhanced entanglement criterion via symmetric informationally complete measurements. *Phys. Rev. A* **98**, 022309 (2018).
4. D. Dieks, Overlap and distinguishability of quantum states. *Phys. Lett. A* **126**, 303–306 (1988).
5. A. Peres, How to differentiate between non-orthogonal states. *Phys. Lett. A* **128**, 19 (1988).
6. C. H. Bennett, Quantum cryptography using any two nonorthogonal states. *Phys. Rev. Lett.* **68**, 3121–3124 (1992).
7. J. M. Renes, Spherical-code key-distribution protocols for qubits. *Phys. Rev. A* **70**, 052314 (2004).
8. J. B. Brask, A. Martin, W. Esposito, R. Houlmann, J. Bowles, H. Zbinden, N. Brunner, Megahertz-rate semi-device-independent quantum random number generators based on unambiguous state discrimination. *Phys. Rev. Appl.* **7**, 054018 (2017).
9. T. Vértesi, E. Bene, Two-qubit Bell inequality for which positive operator-valued measurements are relevant. *Phys. Rev. A* **82**, 062115 (2010).
10. A. Acín, S. Pironio, T. Vértesi, P. Wittek, Optimal randomness certification from one entangled bit. *Phys. Rev. A* **93**, 040102 (2016).
11. S. Gómez, A. Mattar, E. S. Gómez, D. Cavalcanti, O. Jiménez Farías, A. Acín, G. Lima, Experimental nonlocality-based randomness generation with non-projective measurements. *Phys. Rev. A* **97**, 040102 (2018).
12. D. Mayers, A. Yao, Self testing quantum apparatus. *Quantum Inform. Comput.* **4**, 273–286 (2004).
13. J. F. Clauser, M. A. Horne, A. Shimony, R. A. Holt, Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.* **23**, 880–884 (1969).
14. S. J. Summers, R. Werner, Bell's inequalities and quantum field theory. II. Bell's inequalities are maximally violated in the vacuum. *J. Math. Phys.* **28**, 2448–2456 (1987).
15. S. Popescu, D. Rohrlich, Which states violate Bell's inequality maximally? *Phys. Lett. A* **169**, 411–414 (1992).
16. B. S. Tsirelson, Some results and problems on quantum Bell-type inequalities. *Hadronic J. Suppl.* **8**, 329–345 (1993).
17. M. McKague, T. H. Yang, V. Scarani, Robust self-testing of the singlet. *J. Phys. A Math. Theor.* **45**, 455304 (2012).
18. B. W. Reichardt, F. Unger, U. Vazirani, Classical command of quantum systems. *Nature* **496**, 456–460 (2013).
19. M. Oszmaniec, L. Guerini, P. Wittek, A. Acín, Simulating positive-operator-valued measures with projective measurements. *Phys. Rev. Lett.* **119**, 190501 (2017).
20. E. S. Gómez, S. Gómez, P. González, G. Cañas, J. F. Barra, A. Delgado, G. B. Xavier, A. Cabello, M. Kleinmann, T. Vértesi, G. Lima, Device-independent certification of a nonprojective qubit measurement. *Phys. Rev. Lett.* **117**, 260401 (2016).
21. M. Hendrych, R. Gallego, M. Mičuda, N. Brunner, A. Acín, J. P. Torres, Experimental estimation of the dimension of classical and quantum systems. *Nat. Phys.* **8**, 588–591 (2012).
22. J. Ahrens, P. Badziąg, A. Cabello, M. Bourennane, Experimental device-independent tests of classical and quantum dimensions. *Nat. Phys.* **8**, 592–595 (2012).
23. M. Smania, A. M. Elhassan, A. Tavakoli, M. Bourennane, Experimental quantum multiparty communication protocols. *npj Quant. Inf.* **2**, 16010 (2016).
24. T. Lunghi, J. B. Brask, C. C. W. Lim, Q. Lavigne, J. Bowles, A. Martin, H. Zbinden, N. Brunner, Self-testing quantum random number generator. *Phys. Rev. Lett.* **114**, 150501 (2015).
25. A. Tavakoli, A. Hameedi, B. Marques, M. Bourennane, Quantum random access codes using single *d*-level systems. *Phys. Rev. Lett.* **114**, 170502 (2015).
26. D. Martínez, A. Tavakoli, M. Casanova, G. Cañas, B. Marques, G. Lima, High-dimensional quantum communication complexity beyond strategies based on Bell's theorem. *Phys. Rev. Lett.* **121**, 150504 (2018).
27. R. Gallego, N. Brunner, C. Hadley, A. Acín, Device-independent tests of classical and quantum dimensions. *Phys. Rev. Lett.* **105**, 230501 (2010).
28. A. Tavakoli, J. Kaniewski, T. Vértesi, D. Rosset, N. Brunner, Self-testing quantum states and measurements in the prepare-and-measure scenario. *Phys. Rev. A* **98**, 062307 (2018).
29. M. Farkas, J. Kaniewski, Self-testing mutually unbiased bases in the prepare-and-measure scenario. *Phys. Rev. A* **99**, 032316 (2019).
30. A. Tavakoli, D. Rosset, M.-O. Renou, Enabling computation of correlation bounds for finite-dimensional quantum systems via symmetrisation. *Phys. Rev. Lett.* **122**, 070501 (2019).
31. G. M. D'Ariano, P. L. Presti, P. Perinotti, Classical randomness in quantum measurements. *J. Phys. A Math. Gen.* **38**, 5979–5991 (2005).
32. L. Vandenberghe, S. Boyd, Semidefinite programming. *SIAM Rev.* **38**, 49–95 (1996).
33. T. H. Yang, T. Vértesi, J.-D. Bancal, V. Scarani, M. Navascués, Robust and versatile black-box certification of quantum devices. *Phys. Rev. Lett.* **113**, 040401 (2014).
34. J.-D. Bancal, M. Navascués, V. Scarani, T. Vértesi, T. H. Yang, Physical characterization of quantum devices from nonlocal correlations. *Phys. Rev. A* **91**, 022115 (2015).
35. M. Navascués, T. Vértesi, Bounding the set of finite dimensional quantum correlations. *Phys. Rev. Lett.* **115**, 020501 (2015).
36. F. Hirsch, M. T. Quintino, T. Vertesi, M. Navascues, N. Brunner, Better local hidden variable models for two-qubit Werner states and an upper bound on the Grothendieck constant $K_G(3)$. *Quantum* **1**, 3 (2017).
37. L. Guerini, J. Bavaresco, M. Terra Cunha, A. Acin, Operational framework for quantum measurement simulability. *J. Math. Phys.* **58**, 092102 (2017).
38. R. Chaves, J. Bohr Brask, N. Brunner, Device-independent tests of entropy. *Phys. Rev. Lett.* **115**, 110501 (2015).
39. T. Van Himbeeck, E. Woodhead, N. J. Cerf, R. García-Patrón, S. Pironio, Semi-device-independent framework based on natural physical assumptions. *Quantum* **1**, 33 (2017).

40. P. Mironowicz, M. Pawłowski, Experimentally feasible semi-device-independent certification of 4 outcome POVMs. *Phys. Rev. A* **100**, 030301 (2019).
41. R. F. Werner, M. M. Wolf, Bell inequalities and entanglement. *Quantum Inf. Comput.* **1**, 1–25 (2001).

Acknowledgments: We thank J. Kaniewski for insightful comments. **Funding:** This work was supported by the Swiss National Science Foundation (starting grant DIAQ, NCCR-QSIT), the Swedish Research Council, and Knut and Alice Wallenberg Foundation. T.V. was supported by the National Research, Development and Innovation Office NKFIH (grant nos. K111734 and KH125096). **Author contributions:** A.T. and T.V. proposed the basic concept. A.T., T.V., and N.B. developed the theory. M.S. performed the experiments and the data analysis supported by M.B. All authors discussed the results and participated in the writing of the manuscript.

Competing interests: The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 15 January 2019
Accepted 22 January 2020
Published 17 April 2020
10.1126/sciadv.aaw6664

Citation: A. Tavakoli, M. Smania, T. Vértesi, N. Brunner, M. Bourennane, Self-testing nonprojective quantum measurements in prepare-and-measure experiments. *Sci. Adv.* **6**, eaaw6664 (2020).

Noise-robust preparation contextuality shared between any number of observers via unsharp measurements

Hammad Anwer,[1,*] Natalie Wilson,[1,*] Ralph Silva,[2] Sadiq Muhammad,[1] Armin Tavakoli,[3] and Mohamed Bourennane[1]

[1]*Department of Physics, Stockholm University, S-10691 Stockholm, Sweden*
[2]*Institute for Theoretical Physics, ETH Zurich, Switzerland*
[3]*Département de Physique Appliquée, Université de Genève, CH-1211 Genève, Switzerland*

Multiple observers who independently harvest nonclassical correlations from a single physical system share the system's ability to enable quantum correlations. We show that any number of independent observers can share the preparation contextual outcome statistics enabled by state ensembles in quantum theory. Furthermore, we show that even in the presence of any amount of white noise, there exists quantum ensembles that enable such shared preparation contextuality. The findings are experimentally realised by applying sequential unsharp measurements to an optical qubit ensemble which reveals three shared demonstrations of preparation contextuality.

Introduction.— Quantum correlations can surpass the limitations of corresponding classical models. In their most common form, quantum correlations are obtained from the outcomes of *single* (albeit randomly chosen) measurements performed on a physical system. After the measurement, the physical system can be discarded, or even demolished by the measurement apparatus. Therefore, since one does not need to consider the measurement-induced decoherence in the state of the physical system, optimal quantum correlations are often obtained from sharp (projective) measurements that extract a maximal amount of information from the physical system while also inducing a maximal disturbance in its state [1].

Arguably, the fact that measurements disturb physical states should have interesting consequences for more general quantum correlations. To reveal the influence of measurement-induced disturbances on observed outcome statistics, one requires systems to undergo more than a single measurement. A simple scenario for studying the trade-off between the strength of quantum correlations and the disturbance induced by extracting them is one in which quantum correlations are *shared* between many observers. Sharing quantum correlations means that a physical system is measured by a sequence of independent observers, each of whom are tasked with falsifying the existence of a classical model for their observed correlations. Hence, the stronger the correlations extracted by the first observer, the larger the disturbance induced in the state of the system, and thus the weaker the correlations that can possibly be extracted by a second observer. Sharing quantum correlations requires the first observer to measure in such a way that the outcome correlations are strong enough to elude all classical models while the induced disturbance is small enough to enable a second observer independently repeat the same feat. Understanding and characterising quantum correlations obtained via sequential measurements is a conceptually interesting problem [2–5] which has promising applications in quantum information protocols [6, 7].

Sharing quantum correlations was first studied in the context of Bell inequality tests [4] where it was found that a pair of qubits in a singlet state can enable two sequential Bell inequality violations. This has also been experimentally demonstrated [8, 9]. In addition, the number of sequential Bell inequality violations can be indefinitely extended at the price of all observers strongly biasing their choice of measurement and therefore rendering the quantum correlations super-exponentially fragile to noise [4]. Moreover, the shared quantum correlations have recently also been studied in other entanglement-based tasks such as entanglement witnessing [10] and quantum steering [11, 12].

Here, we theoretically and experimentally study the sharing of quantum correlations that demonstrate preparation contextuality. These are correlations that cannot be reproduced in a hidden variable theory that ascribes equivalent representations to indistinguishable preparations, i.e. it disregards the context (specific procedure) underlying a state preparation [13]. Such quantum contextuality does not require entanglement but only single quantum systems, and is well-studied both in theory (see e.g. Refs.[13–19]) and experiment (see e.g. Refs. [15, 16, 20]). In our scenario, states are sampled from an ensemble and communicated sequentially between independent observers, each of whom performs a measurement with the aim of obtaining preparation contextual outcome statistics. We show that preparation contextuality can be shared between any number of sequential observers. Furthermore, we show that the sharing is robust to noise, in the sense that for any given number of independent observers and exposure to any nontrivial amount of white noise, one can find an ensemble whose contextuality can be shared between all the observers. We proceed to experimentally demonstrate the sharing of preparation contextuality. We realise a four-observer scenario in which the first observer prepares an optical qubit ensemble and the remaining three observers perform sequential unsharp (non-maximally disturbing) measurements. Thus, we obtain three shared demonstrations of preparation contextuality.

Nonclassicality via preparation contextuality.— The impossibility of describing the set of observables in quantum theory by underlying classical (noncontextual) quantities originates in the arguments of Bell, Kochen and Specker [21]. More recently, the notion of contextuality has seen a generalisation formulated in operational terms (i.e., in terms of probabilities) applying to measurements, transformations and

* H. A. and N. W. contributed equally to this work.

preparations [13]. Here, we are interested in contextuality in terms of preparations.

The predictions of an operational theory (e.g. quantum theory) may be explained by an ontological model [22] that ascribes each physical system S to a set Λ of ontic (objective) states λ. A particular preparation P of the system is associated to a distribution $\mu_P(\lambda)$ over the ontic state space. Similarly, the probability of outcome b of a measurement M is described by a response function $\xi_{b,M}(\lambda)$. The ontological model thus seeks a μ and a ξ to explain the observed statistics by $p(b|P,M) = \int_\Lambda \mu_P(\lambda)\xi_{b,M}(\lambda)d\lambda$. The model is said to be *preparation noncontextual* if two different preparations P and P' that cannot be distinguished by the statistics generated by any measurement (that is; $\forall M : p(b|P,M) = p(b|P',M)$) are associated to the *same* distribution over ontic states, i.e., $\mu_P = \mu_{P'}$. If observed statistics falsify this assumption, then it is said to be preparation contextual. Quantum state ensembles are known to enable preparation contextuality.

A family of preparation noncontextuality inequalities.— In order to prove preparation contextuality, it is sufficient to violate an inequality bounding the correlations attainable in a preparation noncontextual model. We focus on a family of such inequalities introduced in Ref. [15] related to a variant of Random Access Coding [23, 24]. Consider a party Alice receiving a random input string $x = x_1 \dots x_n \in \{0,1\}^n$. Her input is associated to a preparation P_x (one of 2^n possible) which is sent to a receiver Bob. Her preparations are constrained to satisfy certain indistinguishability relations: there must exist no measurement that can reveal any information about the parity of the string $r \cdot x$ for every $r \in \{0,1\}^n$ with $|r| \geq 2$. Bob receives a random input $y \in \{1,\dots,n\}$, and performs a measurement $\{M_y^b\}$ with outcome $b \in \{0,1\}$. The partnership is awarded a point if the outcome of Bob coincides with the yth entry in Alice's string. In any preparation noncontextual theory, the probability of winning obeys the following bound [15]:

$$\mathcal{A}^{(n)} \equiv \frac{1}{n2^n}\sum_{x,y} p(b = x_y|x,y) \leq \frac{n+1}{2n}. \qquad (1)$$

Due to the contextual nature of quantum theory, these inequalities can be violated. Maximal quantum violations for any $n \geq 2$ are known [25]. Bob performs dichotomic measurements characterised by an observable $G^T_{n,y}$. These are recursively defined from $G_{2,1} = \sigma_x$, $G_{2,2} = \sigma_y$, and $G_{3,1} = \sigma_x$, $G_{3,2} = \sigma_y$ and $G_{3,3} = \sigma_z$, and

$$\begin{aligned} &n \text{ even:} \quad G_{n,k} = G_{n-1,k} \otimes \sigma_x \qquad \forall i \in \{1,\dots,n-1\}, \\ &n \text{ odd:} \quad\ G_{n,k} = G_{n-2,k} \otimes \sigma_x \qquad \forall i \in \{1,\dots,n-2\} \end{aligned} \qquad (2)$$

with $G_{n,n} = \mathbb{1} \otimes \sigma_y$ if n is even, and $G_{n,n} = \mathbb{1} \otimes \sigma_z$ and $G_{n,n-1} = \mathbb{1} \otimes \sigma_y$ of n is odd. The optimal preparations are states of $\lfloor n/2 \rfloor$ qubits specified by

$$\rho_x = \mathrm{tr}_\mathrm{A}\left[(\mathbb{1} + A_x) \otimes \mathbb{1}\phi_{\max}^{\otimes \lfloor n/2 \rfloor}\right], \qquad (3)$$

where $A_x = \frac{1}{\sqrt{n}}\sum_{i=1}^n (-1)^{x_i} G_{n,i}$, $\phi_{\max}$ corresponds to the maximally entangled state $(|0,0\rangle + |1,1\rangle)/\sqrt{2}$, and the trace

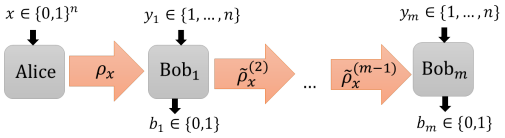

FIG. 1. Alice's preparations are sent from one observer to the next, each performing a measurement aiming to independently reveal preparation contextual statistics. To this end, only the average post-measurement state $\tilde{\rho}_x^{(k)}$ is relevant.

is taken over the first system in every entangled pair. Note that Alice's preparations are single quantum systems, and only for simplicity written in terms of post-measurement states of a collection of entangled states. The presented strategy leads to the maximal quantum value $\mathcal{A}^{(n)} = 1/2(1+1/\sqrt{n})$ for every n [25].

Sequential scenario.— We consider a scenario in which the ability to violate the inequality (1) is shared between many independent observers, named Bob_1,..., Bob_m, each of whom receive an independent random input $y_k \in \{1,\dots,n\}$ and output $b_k \in \{0,1\}$. Alice's randomly chosen preparation is sent to Bob_1 who performs a measurement and passes the post-measurement state to Bob_2 who performs a measurement and passes the post-measurement state to Bob_3 etc. The scenario is illustrated in Fig. 1. The pair Alice-Bob_k uses the marginal distribution $p(b_k|x,y_k)$ to compute the witness (1) (here labelled $\mathcal{A}_k^{(n)}$) to check for preparation contextuality.

In a quantum approach, we may denote Alice's preparations by ρ_x which must satisfy the indistinguishability relation $\sum_{r\cdot x=0}\rho_x = \sum_{r\cdot x=1}\rho_x$ for every string r with $|r| \geq 2$. Since one has to keep track of both the statistics and the post-measurement states of each Bob, we require the detailed set of Kraus operators for each measurement. By $K_{y_k}^{b_k}$ we denote the Kraus operators of Bob_k associated to the y_kth measurement and b_kth outcome. The state received by Bob_k is specified by Alice's input x, and the strings of inputs $(y_1,\dots,y_{k-1})$ and outputs $(b_1,\dots,b_{k-1})$ of all previous Bobs. However, we treat each Bob in the sequence as independent from the rest, meaning that they do not know the specific inputs or outputs of the other Bobs in each run of the experiment. Thus, in order to calculate the relevant marginal distributions $p(b_k|x,y_k)$, only the average state $\tilde{\rho}_x^{(k)}$ received by Bob_k is required, i.e., the state obtained from averaging a preparation ρ_x of Alice over all the inputs and outputs of all previous Bobs:

$$\tilde{\rho}_x^{(k)} = \frac{1}{n}\sum_{y_{k-1},b_{k-1}} K_{y_{k-1}}^{b_{k-1}}\tilde{\rho}_x^{(k-1)}(K_{y_{k-1}}^{b_{k-1}})^\dagger, \qquad (4)$$

with $\tilde{\rho}_x^{(1)} = \rho_x$. Consequently, the desired marginal statistics for Bob_k are $p(b_k|x,y_k) = \mathrm{tr}\left(\tilde{\rho}_x^{(k)}(K_{y_k}^{b_k})^\dagger K_{y_k}^{b_k}\right)$. This constitutes a description of general quantum strategies in the sequential scenario.

Sharing preparation contextuality.— We apply the above general description to construct a specific family of quantum strategies for sharing preparation contextuality, that is

inspired by the previously described optimal quantum strategy for the maximal violation of the inequalities (1). Alice prepares the states (3) while each Bob performs an unsharp variant of the measurements optimal for violating (1). In that strategy the measurements of Bob are the dichotomic observables G^T_{n,y_k} defined in (2), corresponding to the projectors $\Pi^b_{n,y} = (\mathbb{1} + (-1)^b G^T_{n,y})/2$ that are both the Kraus operators and POVM elements. For a weaker measurement, one modifies the POVM element to $(\mathbb{1} + (-1)^b\ \eta_k\ G^T_{n,y})/2$, for some $\eta_k \in [0,1]$. If $\eta_k = 1$ ($\eta_k = 0$), the measurement is sharp (non-interacting). Choosing $0 < \eta_k < 1$ corresponds to an unsharp measurement. The corresponding Kraus operator is given by

$$K^{b_k}_{y_k} = \sqrt{\frac{1+\eta_k}{2}}\Pi^{b_k}_{n,y_k} + \sqrt{\frac{1-\eta_k}{2}}\Pi^{\bar{b}_k}_{n,y_k}, \tag{5}$$

where the bar-sign denotes a bit-flip. This class of strategies has the following convenient property.

Lemma 1. *If Alice prepares the states in Eq. (3) and the Bobs each measure G^T_{n,y_k} with sharpness η_k, the average state received by Bob$_k$ is*

$$\tilde{\rho}^{(k)}_x = v_k \rho_x + (1 - v_k)\, \rho_{mix}, \tag{6}$$

where ρ_{mix} is the maximally mixed state and the visibility $v_k \in [0,1]$ is given recursively by

$$v_k = v_{k-1} f_{k-1} = \prod_{j=1}^{k-1} f_j, \tag{7}$$

where $v_1 = 1$ by definition, and the "quality factor" f_k of the measurement of Bob$_k$ is defined from the sharpness η_k as $f_k = (1 + (n-1)\sqrt{1-\eta_k^2})/n$.

Proof. The proof is technical in character and is given in Appendix (section A). ■

Using Eq. (6), the figure of merit (1) for the pair Alice and Bob$_k$ reads

$$\mathcal{A}^{(n)}_k = \frac{1}{2}\left(1 + \frac{v_k \eta_k}{\sqrt{n}}\right). \tag{8}$$

This leads to preparation contextuality whenever $\eta_k > 1/(v_k\sqrt{n})$. This can be used to recursively calculate the critical pairs (η_k, v_k). Thusly, we arrive at the following result.

Result 1. *The number of observers who can independently share the preparation contextuality enabled by Alice's ensemble is at least n.*

Proof. Consider that each Bob tunes the sharpness of his measurement so as to just violate the inequality (1), but not more. Expressing the measurement sharpness $\eta_k = \sin\theta_k$, where $\theta_k \in [0, \pi/2]$, we thus require $\sin\theta_k = 1/(v_k\sqrt{n})$. On the other hand, a trivial lower bound on the quality factor of Bob$_k$'s measurement is $f_k = (1 + (n-1)\cos\theta_k)/n \geq \cos\theta_k$. Squaring, and using the expression for the critical value of $\sin\theta_k$ above, we find that $f_k^2 \geq 1 - 1/(v_k^2 n)$. Since the visibility of the next Bob is $v_{k+1} = v_k f_k$, we have $v^2_{k+1} = v_k^2 f_k^2 \geq v_k^2\left(1 - 1/(v_k^2 n)\right)$. Hence, the decrease in visibility from each Bob to the next is bounded by $v_k^2 - v^2_{k+1} \leq 1/n$ which together with $v_1 = 1$ gives $v^2_{k+1} \geq 1 - k/n$. This implies that the visibility of the nth Bob is at least $v_n \geq 1/\sqrt{n}$, which is precisely the condition for violating the preparation noncontextuality inequality. ■

Thus by suitably choosing n, an arbitrary long sequence of observers can share the preparation contextual correlations enabled by Alice's ensemble. Moreover, we show in Appendix (section B) that for the considered class of quantum strategies, the number of observers who share preparation contextuality can be no more than n. Also, as shown in Appendix (section C), one can share preparation contextuality between any number of observers also in a scenario in which none of the Bob's knows his position in the sequence.

Noise-robustness.— The scenario we have considered so far is an idealisation in which no noise appears. In addition to this not being realistic in any experiment, it is interesting to consider whether the noiseless scenario is distinctive, or also significantly noisy ensembles [26] enable shared preparation contextuality. To address this matter, we let Alice's preparations be mixtures of the intended state ρ_x with the maximally mixed state: $\rho_x(q) = q\rho_x + (1-q)\rho_{\text{mix}}$ for some visibility $q \in [0,1]$. For a given number of observers, what is the smallest q such that preparation contextuality can be shared between all observers?

Result 2. *For any given number of independent observers m, there exists an ensemble whose contextuality can be shared between all observers for any $q > 0$.*

Proof. We substitute ρ_x for $\rho_x(q)$ in the proof of Result 1. This means $v_1 = q$, and leads to $v^2_{k+1} \geq q - k/n$. Thus, in order to observe m violations, one must choose $n \geq \lceil \frac{m}{q} \rceil$. ■

Hence, preparation contextuality can be shared between any number of observers using ensembles with an arbitrarily large noise-component by choosing a sufficiently large n. The price to pay for this property is that when $q \to 0$, both the Hilbert space dimension of Alice's ensemble and the number of preparations and measurements diverge.

Experiment.— We demonstrate the theoretical findings in an experiment with three ($n = 3$) sequential tests of preparation contextuality. Alice prepares the eight qubit states (3) with Bloch vectors $\vec{a}_x = [(-1)^{x_1}, (-1)^{x_2}, (-1)^{x_3}]/\sqrt{3}$. Bob$_1$ and Bob$_2$ perform unsharp measurements (5) of σ_x, σ_y and σ_z whereas Bob$_3$ performs projective (sharp) measurements of the same observables.

In the experiment we peform unsharp measurements on the polarisation state of a single photon using shifted Sagnac interferometers, as shown in Bob$_1$ and Bob$_2$ in Fig. (2). A HWP is placed in each path of the interferometer, rotated to $\theta_i/2$ in the horizontal path and $\pi/4 - \theta_i/2$ in the vertical path to control the sharpness of the measurement. A HWP and QWP before and after the interferometer are used to select the basis of the measurement. The measurement outcome is encoded in the output path, i.e. outcome $b_i = 0$ ($b_i = 1$) corresponds to the detection of the photon in output path 1 (2, beam blocked in figure). In the sequential scenario we choose to consider

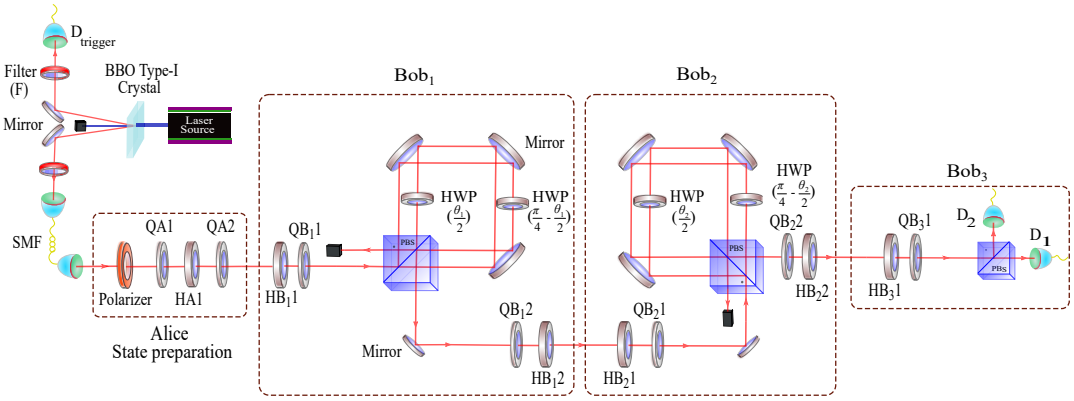

FIG. 2. Optical set-up used to reveal contextuality sharing. See text for details. Q and H represent quarter-wave plates (QWPs) and half-wave plates (HWPs).

only one path at a time for Bob_1 and Bob_2 to simplify the set-up. By adding an additional rotation to the HWPs or QWPs before and after Bob, we can select the output we want to analyse [8, 9]. The results of Bob_1 and Bob_2's unsharp measurements are therefore obtained at Bob_3, comprised of a PBS and single photon detectors D_1 and D_2. For example, if we consider output 1 at Bob_1 and Bob_2, a click in either detector at Bob_3 tells us that Bob_1 and Bob_2 had the outcome $b_1 = 0$ and $b_2 = 0$. We analyse the counts in Bob_3 corresponding to all possible combinations of output ports to realise a full measurement. This protocol relies on a stable photon generation rate. Details of measurement angles are given in Appendix (section D). This set-up can be used to perform projective measurements ($\eta = 1$, $\theta_i = 0$), no measurement ($\eta = 0$, $\theta_i = \pi/4$), or an intermediate-strength measurement, where the the sharpness (strength) of the measurement is tuned by varying θ_i.

The full set-up is shown in Fig. 2. We generate heralded single photons at 780 nm via spontaneous parametric down-conversion (SPDC) using a single type-I beta barium borate (BBO) crystal of thickness 2 mm pumped by 390 nm femtosecond laser pulses. The idler photon is detected by an APD single-photon detector, $\text{D}_{\text{trigger}}$, and is used as a trigger. The single photons are coupled into single-mode fibres (SMF) after passing through a narrowband 3 nm interference filter (F) to define the spatial and spectral properties of the photons. After filtering, the signal photon is prepared into one of Alice's eight states, using a polariser, two QWPs and a HWP (angles given in Appendix (section D). The unsharp measurements of Bob_1 and Bob_2 correspond to $\theta_1 = 24.95^\circ$ ($\eta_1 = 0.6441$) and $\theta_2 = 20.10^\circ$ ($\eta_2 = 0.7637$) respectively, which ideally produce $\mathcal{A}_1 = \mathcal{A}_2 = \mathcal{A}_3 = 0.6859 > 2/3$ with $\mathcal{A}_k = \mathcal{A}_k^{(3)}$.

Results.— In order to test each of the three preparation non-contextuality inequalities (between Alice and each of the three Bobs), we require 24 marginal probabilities (the 'winning' answers $b_k = x_{y_k}$) corresponding to the three measurement bases and Alice's eight preparations. To reduce the Poissonian error, each Bob collects approximately 34 million counts for each of these 24 settings. Our experimental values can be found in Appendix (section E). These lead to three preliminary values of $\mathcal{A}_1^{\text{pre}} = 0.687 \pm 0.001$, $\mathcal{A}_2^{\text{pre}} = 0.675 \pm 0.001$, and $\mathcal{A}_3^{\text{pre}} = 0.681 \pm 0.001$.

Data analysis.— Due to small yet unavoidable experimental imperfections, e.g. waveplate imperfections and offsets in the rotation of the waveplates, it is impossible to perfectly satisfy the operational indistinguishability relations required to test preparation contextuality. This problem can be overcome by suitable post-processing methods [20]. As described in Appendix (section F), we have used a relaxed variant of these methods to enforce the indistinguishability relations relevant to a test of inequality (1) on our experimental data. This comes at the cost of the observed values $\left(\mathcal{A}_1^{\text{pre}}, \mathcal{A}_2^{\text{pre}}, \mathcal{A}_3^{\text{pre}}\right)$ decreasing in a manner corresponding to how well the statistics approximates said relations. Due to the high visibility and precision of the experimental set-up, we find only a small decrease in the three correlation witnesses:

$$\begin{aligned}\mathcal{A}_1^{\text{post}} &= 0.683 \pm 0.001\\ \mathcal{A}_2^{\text{post}} &= 0.670 \pm 0.001\\ \mathcal{A}_3^{\text{post}} &= 0.677 \pm 0.001\end{aligned}$$

all of which violate inequality (1).

Conclusions.— We have theoretically developed and experimentally demonstrated the sharing of preparation contextual correlations in scenarios that require no entanglement. In addition to such correlations being possible to share between any number of observers, we found that this can be done in a strongly noise-robust manner. This distinguishes shared preparation contextuality from known results in e.g. shared Bell nonlocality in which the fragility to noise of sequential demonstrations scales super-exponentially [4]. This fragility poses a significant experimental hurdle and has hitherto limited demonstrations to two sequential violations of Bell inequalities [8, 9]. We experimentally observed three sequential

demonstrations of preparation contextuality. Optical set-ups of this spirit (see also Refs [8, 9]) are promising candidates for a variety of sequential correlation tests. Finally, an interesting question is to understand which forms of quantum correlations can be shared between indefinitely many observers in a noise-robust manner.

Acknowledgements.— This work was supported by the project "Photonic Quantum Information" (Knut and Alice Wallenberg Foundation, Sweden), the Swedish Research Council, and the Swiss National Science Foundation (Starting grant DIAQ, grant No. 200020_165843, and the National Centre of Competence in Research "Quantum Systems and Technology").

[1] C. A. Fuchs, and A. Peres, Quantum-state disturbance versus information gain: Uncertainty relations for quantum information Phys. Rev. A **53**, 2038 (1996).

[2] R. Gallego, L. E. Würflinger, R. Chaves, A. Acín, M. Navascués, Nonlocality in sequential correlation scenarios, New J. Phys. **16**, 033037 (2014).

[3] C. Budroni, T. Moroder, M. Kleinmann, and O. Gühne, Bounding Temporal Quantum Correlations, Phys. Rev. Lett. **111**, 020403 (2013).

[4] R. Silva, N. Gisin, Y. Guryanova, and S. Popescu, Multiple Observers Can Share the Nonlocality of Half of an Entangled Pair by Using Optimal Weak Measurements, Phys. Rev. Lett. **114**, 250401 (2015).

[5] A. Tavakoli, A. Cabello, Quantum predictions for an unmeasured system cannot be simulated with a finite-memory classical system, Phys. Rev. A **97**, 032131 (2018).

[6] F. J. Curchod, M. Johansson, R. Augusiak, M. J. Hoban, P. Wittek, and A. Acín, Unbounded randomness certification using sequences of measurements, Phys. Rev. A **95**, 020102(R) (2017).

[7] B. Coyle, M. J. Hoban, and E. Kashefi, One-Sided Device-Independent Certification of Unbounded Random Numbers, EPTCS **273**, 14-26 (2018).

[8] M. Schiavon, L. Calderaro, M. Pittaluga, G. Vallone, and P. Villoresi, Three-observer Bell inequality violation on a two-qubit entangled state, Quantum Sci. Technol. **2** 015010 (2017).

[9] M-J. Hu, Z-Y. Zhou, X-M. Hu, C-F. Li, G-C. Guo, and Y-S. Zhang, Observation of non-locality sharing among three observers with one entangled pair via optimal weak measurement, npj Quantum Information **4**, 63 (2018).

[10] A. Bera, S. Mal, A. Sen De, and U. Sen, Witnessing bipartite entanglement sequentially by multiple observers, Phys. Rev. A **98**, 062304 (2018).

[11] S. Sasmal, D. Das, S. Mal, and A.S. Majumdar, Steering a single system sequentially by multiple observers, Phys. Rev. A **98**, 012305 (2018).

[12] A. Shenoy H, S. Designolle, F. Hirsch, R. Silva, N. Gisin, and N. Brunner, Unbounded sequence of observers exhibiting Einstein-Podolsky-Rosen steering, Phys. Rev. A **99**, 022317 (2019).

[13] R. W. Spekkens, Contextuality for preparations, transformations, and unsharp measurements Phys. Rev. A **71**, 052108 (2005).

[14] R. W. Spekkens, Negativity and Contextuality are Equivalent Notions of Nonclassicality, Phys. Rev. Lett. **101**, 020401 (2008).

[15] R. W. Spekkens, D. H. Buzacott, A. J. Keehn, B. Toner, and G. J. Pryde, Preparation Contextuality Powers Parity-Oblivious Multiplexing Phys. Rev. Lett. **102**, 010401 (2009).

[16] A. Hameedi, A. Tavakoli, B. Marques, and M. Bourennane, Communication games reveal preparation contextuality, Phys. Rev. Lett. **119**, 220402 (2017)

[17] M. S. Leifer, and O. J. E. Maroney, Maximally Epistemic Interpretations of the Quantum State and Contextuality, Phys. Rev. Lett. **110**, 120401 (2013).

[18] M. Banik, S. S. Bhattacharya, A. Mukherjee, A. Roy, A. Ambainis, and A. Rai, Limited preparation contextuality in quantum theory and its relation to the Cirel'son bound, Phys. Rev. A **92**, 030103(R) (2015).

[19] D. Saha, and A. Chaturvedi, Preparation contextuality: the ground of quantum communication advantage, arXiv:1802.07215.

[20] M. D. Mazurek, M. F. Pusey, R. Kunjwal, K. J. Resch, and R. W. Spekkens, An experimental test of noncontextuality without unphysical idealizations, Nature Communications **7**, 11780 (2016).

[21] S. Kochen, and E. P. Specker, The Problem of Hidden Variables in Quantum Mechanics, Indiana University Mathematics Journal, **17**, 59 (1967).

[22] N. Harrigan, and R. W. Spekkens, Einstein, Incompleteness, and the Epistemic View of Quantum States, Found Phys (2010) **40**, 125 (2010).

[23] A. Ambainis, A. Nayak, A. Ta-Shma, U. Vazirani, Dense quantum coding and a lower bound for 1-way quantum automata, Proceedings of the 31st Annual ACM Symposium on Theory of Computing (STOC'99), 376-383 (1999).

[24] A. Tavakoli, A. Hameedi, B. Marques, and M. Bourennane, Quantum random access codes using single d-Level systems, Phys. Rev. Lett. **114**, 170502 (2015).

[25] A. Chailloux, I. Kerenidis, S. Kundu, and J. Sikora, Optimal bounds for parity-oblivious random access codes, New J. Phys. **18**, 045003 (2016).

[26] One could alternatively consider the Bobs' measurement devices inducing the noise. However, this is less detrimental than noisy preparations. The reason is that if Alice's preparations are noisy the correlations due to all Bobs' measurements are weaker, whereas if instead one (or many) of the Bobs sometimes fail to perform the intended measurement, the state relayed to the next Bob retains a higher degree of coherence and leads to him observering stronger correlations.

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-020-16137-4 OPEN

Constraints on nonlocality in networks from no-signaling and independence

Nicolas Gisin[1✉], Jean-Daniel Bancal[1✉], Yu Cai[1✉], Patrick Remy[1], Armin Tavakoli[1], Emmanuel Zambrini Cruzeiro[1,2], Sandu Popescu[3,4] & Nicolas Brunner[1]

The possibility of Bell inequality violations in quantum theory had a profound impact on our understanding of the correlations that can be shared by distant parties. Generalizing the concept of Bell nonlocality to networks leads to novel forms of correlations, the characterization of which is, however, challenging. Here, we investigate constraints on correlations in networks under the natural assumptions of no-signaling and independence of the sources. We consider the triangle network with binary outputs, and derive strong constraints on correlations even though the parties receive no input, i.e., each party performs a fixed measurement. We show that some of these constraints are tight, by constructing explicit local models (i.e. where sources distribute classical variables) that can saturate them. However, we also observe that other constraints can apparently not be saturated by local models, which opens the possibility of having nonlocal (but non-signaling) correlations in the triangle network with binary outputs.

[1] Département de Physique Appliquée, Université de Genève, Genève, Switzerland. [2] Laboratoire d'Information Quantique (LIQ), Université Libre de Bruxelles, Bruxelles, Belgium. [3] H.H. Wills Physics Laboratory, Tyndall Avenue, BS8 1TL Bristol, UK. [4] Institute for Theoretical Studies, ETH, Zurich, Switzerland. ✉email: Nicolas.Gisin@unige.ch; Jean-Daniel.Bancal@unige.ch; Yu.Cai@unige.ch

The no-signaling principle states that instantaneous communication at a distance is impossible. This imposes constraints on the possible correlations between distant observers. Consider the so-called Bell scenario[1], where each party performs different local measurements on a shared physical resource distributed by a single common source. In this case, the no-signaling principle implies that the choice of measurement (the input) of one party cannot influence the measurement statistics observed by the other parties (their outputs). In other words, the marginal probability distribution of each party (or subset of parties) must be independent of the input of any other party. These are the well-known no-signaling conditions, which represent the weakest conditions that correlations must satisfy in any reasonable physical theory[2], in the sense of being compatible with relativity. More generally, the no-signaling principle ensures that the information cannot be transmitted without any physical carrier. This provides a useful framework to investigate quantum correlations (which obviously satisfy the no-signaling conditions, but do not saturate them in general[2]) within a larger set of physical theories satisfying no-signaling; see e.g., refs. [2–9].

Recently, the concept of Bell nonlocality has been generalized to networks, where separated sources distribute physical resources to subsets of distant parties (Fig. 1). Assuming the sources to be independent from each other[10,11], arguably a natural assumption in this context, leads to many novel effects. Notably, it becomes now possible to demonstrate quantum nonlocality without the use of measurement inputs[11–15], but only by considering the output statistics of fixed measurements. Just recently, a first example of such nonlocality genuine to networks was proposed[15,16]. This radically departs from the standard setting of Bell nonlocality, and opens many novel questions. Characterizing correlations in networks (local or quantum) is however still very challenging at the moment, despite recent progress[17–28].

Moving beyond quantum correlations, this naturally raises the question of finding the limits of possible correlations in networks, assuming only no-signaling and independence (NSI) of the sources[22,29–33]. Here, we investigate this question and derive limits on correlations, which we refer to as NSI constraints. While our approach can in principle be applied to any network, we focus here on the well-known triangle network with binary outputs and no inputs, for which we obtain strong, and even tight NSI constraints. Specifically, we show that, despite the absence of an input, some statistics imply the possibility for one party to signal to others by locally changing (or not changing) the structure of the network. Formally, this amounts to considering a specific class of so-called network inflations, as introduced in ref. [22], which we show can lead to general and strong NSI constraints. Moreover, we prove that some of our NSI constraints are in fact tight, by showing that they can be saturated by correlations from explicit trilocal models, in which the sources distribute classical variables. Interestingly, however, it appears that not all of our NSI constraints can be saturated by trilocal models, which opens the possibility of having nonlocal (but nevertheless non-signaling) correlations in the triangle network with binary outputs. Finally, we conclude with a list of open questions.

Results

NSI constraints. The triangle network (sketched in Fig. 1a) features three observers: Alice, Bob, and Charlie. Every pair of observers is connected by a (bipartite) source, providing a shared physical system. Importantly, the three sources are assumed to be independent from each other. Hence, the three observers share no common (i.e., tripartite) piece of information. Based on the received physical resources, each observer provides an output (a, b, and c, respectively). Note that the observers receive no input

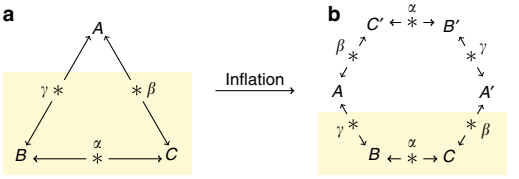

Fig. 1 Inflation of the triangle network to the hexagon network. In order to capture NSI constraints in the triangle network **a**, we consider an inflation to the hexagon network **b**. Importantly, from the point of view of Bob and Charlie, the two situations must be indistinguishable. If not, then Alice could (instantaneously) signal to Bob and Charlie, simply by locally modifying the network structure.

in this setting, contrary to standard Bell nonlocality tests. The statistics of the experiment are thus given by the joint probability distribution $p(a, b, c)$. We focus on the case of binary outputs: $a, b, c \in \{+1, -1\}$. It is then convenient to express the joint distribution as follows:

$$p(a,b,c) = \frac{1}{8}(1 + aE_{\mathrm{A}} + bE_{\mathrm{B}} + cE_{\mathrm{C}} + abE_{\mathrm{AB}} + acE_{\mathrm{AC}} + bcE_{\mathrm{BC}} + abcE_{\mathrm{ABC}}), \quad (1)$$

where E_{A}, E_{B}, and E_{C} are the single-party marginals, E_{AB}, E_{BC}, and E_{AC} are the two-party marginals, and E_{ABC} is the three-body correlator. Note that the positivity of $p(a, b, c)$ implies constraints on marginals, in particular $p(+++) + p(---) \geq 0$ implies

$$E_{\mathrm{AB}} + E_{\mathrm{AC}} + E_{\mathrm{BC}} \geq -1\ . \quad (2)$$

In the following, we will derive nontrivial constraints bounding and relating the single-party and two-party marginals of $p(a, b, c)$ under the assumption of NSI. While it seems a priori astonishing that the no-signaling principle can impose constraints in a Bell scenario, featuring no inputs for the parties, we will see that this is nevertheless the case in the triangle network.

The main idea is the following. Although one party (say Alice) receives no input, she could still potentially signal to Bob and Charlie by locally modifying the structure of the network. To see this, consider the hexagon network depicted in Fig. 1b, and focus on parties Bob and Charlie. From their point of view, the two networks (triangle and hexagon) should be indistinguishable. This is because all the modification required to bring the triangle network to the hexagon (e.g., by having Alice adding extra parties and sources) occurs on Alice's side, and can therefore be space-like separated from Bob and Charlie. If Alice, by deciding which network to use, could remotely influence the statistics of Bob and Charlie, this would clearly lead to signaling. Hence, we conclude that the local statistics of Bob and Charlie (i.e., the single-party marginals E_{B} and E_{C}, as well as the two-party marginals E_{BC}) must be the same in the triangle and in the hexagon. To see that this condition really captures the possibility to signal, we could imagine a thought experiment in which we would give an input to Alice, which determines whether she modifies her network structure or not. If she does so and this has an incidence on the E_{BC} marginal, then Bob and Charlie can learn about Alice's input, hence breaking the usual notion no-signaling condition. Note that the input considered here is however purely fictional, Alice's input is not present in the actual experiment.

From the above reasoning, we conclude that the joint output probability distribution for the hexagon, i.e., $p(a, b, c, a', b', c')$, must satisfy several constraints. In particular, one should have that

$$\sum b\ p(a,b,c,a',b',c') = \sum b'\ p(a,b,c,a',b',c') = E_{\mathrm{B}} \quad (3)$$

$$\sum c\; p(a,b,c,a',b',c') = \sum c'\; p(a,b,c,a',b',c') = E_{\rm C} \quad (4)$$

$$\sum bc\; p(a,b,c,a',b',c') = \sum b'c'\; p(a,b,c,a',b',c') = E_{\rm BC}, \quad (5)$$

where all sums go over all outputs a, b, c, a', b', c'. From the independence of the sources, we obtain additional constraints, namely

$$\sum bb'\; p(a,b,c,a',b',c') = E_{\rm B}^2 \quad (6)$$

$$\sum cc'\; p(a,b,c,a',b',c') = E_{\rm C}^2 \quad (7)$$

$$\sum bb'c\; p(a,b,c,a',b',c') = E_{\rm BC}E_{\rm B} \quad (8)$$

$$\sum bcc'\; p(a,b,c,a',b',c') = E_{\rm BC}E_{\rm C} \quad (9)$$

$$\sum bcb'c'\; p(a,b,c,a',b',c') = E_{\rm BC}^2\ . \quad (10)$$

Clearly, we also get similar constraints when considering signaling between any other party (Bob or Charlie) to the remaining two.

Altogether, we see that NSI imposes many constraints on $p(a,b,c,a',b',c')$. Obviously, we also require that

$$p(a,b,c,a',b',c') \geq 0 \quad \text{and} \quad \sum p(a,b,c,a',b',c') = 1\ . \quad (11)$$

Now reversing the argument, we see that the non-negativity of $p(a,b,c,a',b',c')$ imposes nontrivial constraints relating the single- and two-party marginals of the triangle distribution $p(a, b, c)$. To illustrate this, let us proceed with an example in a slightly simplified scenario, assuming all single-party marginals to be uniformly random, i.e., $E_{\rm A} = E_{\rm B} = E_{\rm C} = 0$. In this case, we obtain

$$\begin{aligned} 64\; p(a,b,c,a',b',c') = {} & 1 + (ab + a'b')E_{\rm AB} + (bc + b'c')E_{\rm BC} + (ca' + c'a)E_{\rm AC} \\ & + (abc + a'b'c')F_3 + (bca' + b'c'a)F_3' + (ca'b' + c'ab)F_3'' \\ & + aa'bb'E_{\rm AB}^2 + bb'cc'E_{\rm BC}^2 + aa'cc'E_{\rm AC}^2 + aa'(bc + b'c')F_4 \\ & + bb'(ca' + c'a)F_4' + cc'(ab + a'b')F_4'' + aa'bb'(c + c')F_5 \\ & + bb'cc'(a + a')F_5' + aa'cc'(b + b')F_5'' + aa'bb'cc'F_6 \geq 0 \end{aligned} \quad (12)$$

Importantly, notice that the above expression contains a number of variables (of the form $F_{\rm X}$) that are uncharacterized; these represent X-party correlators in the hexagon network, see Supplementary Note 1 for more details. Hence, we obtain a set of inequalities imposing constraints on our variables of interest (i.e., $E_{\rm AB}$, $E_{\rm BC}$, and $E_{\rm AC}$), but containing also additional variables that we would like to discard. This can be done systematically via the algorithm of Fourier–Motzkin elimination[34]. Note that here we need to treat the squared terms, such as $E_{\rm AB}^2$, as new variables, independent from $E_{\rm AB}$, so that we get a system of linear inequalities. Solving the latter, and taking into account positivity constraints as in Eq. (2), we obtain a complete characterization of the set of two-body marginals (i.e., $E_{\rm AB}$, $E_{\rm BC}$, and $E_{\rm AC}$) that are compatible with NSI in the triangle network (for a hexagon inflation and uniform single-party marginals), in terms of a single inequality

$$(1 - E_{\rm AB})^2 - E_{\rm BC}^2 - E_{\rm AC}^2 \geq 0\ , \quad (13)$$

and its symmetries (under relabeling of the parties and of the outputs). This implies a more symmetric, but slightly weaker inequality:

$$(1 + E_{\rm AB})^2 + (1 + E_{\rm BC})^2 + (1 + E_{\rm AC})^2 \leq 6\ . \quad (14)$$

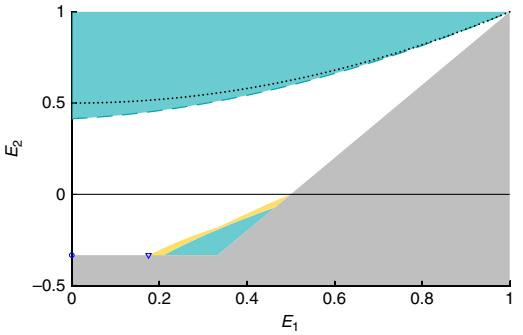

Fig. 2 Region of allowed correlations for symmetric distributions; projection in the plane E_2 vs E_1. The turquoise region is ruled out by NSI constraints, while the gray region is excluded from simple positivity constraints. The white region is accessible via trilocal models. Correlations in the yellow region satisfy NSI constraints (from the hexagon inflation), but we could not find a trilocal model for them. The constraint Eq. (34) of ref. [22] is shown in dotted black. The dashed turquoise curve corresponds to the NSI inequality Eq. (15), which turns out to be tight. Explicit trilocal models are also obtained for the correlations marked by blue dots (Supplementary Note 2).

Note that when $E_{\rm AB} = E_{\rm BC} = E_{\rm AC} \equiv E_2$, we get simply $E_2 \leq \sqrt{2} - 1 \approx 0.41$.

Next, we consider the symmetric case (i.e., $E_{\rm A} = E_{\rm B} = E_{\rm C} \equiv E_1$ and $E_{\rm AB} = E_{\rm BC} = E_{\rm AC} \equiv E_2$) and obtain nontrivial NSI constraints on the possible values of E_1 and E_2 (Fig. 2). In particular, correlations compatible with NSI must satisfy the following inequality

$$(1 + 2|E_1| + E_2)^2 \leq 2(1 + |E_1|)^3\ . \quad (15)$$

Let us move now to the most general case, with arbitrary values for single- and two-party marginals. For a given set of values $E_{\rm A}$, $E_{\rm B}$, $E_{\rm C}$, $E_{\rm AB}$, $E_{\rm BC}$, and $E_{\rm AC}$, it is possible here to determine via a linear program whether this set is compatible with NSI or not (Supplementary Note 1). More generally, obtaining a characterization of the NSI constraints in terms of explicit inequalities (as above) is challenging, due mainly to the number of parameters and nonlinear constraints. We nevertheless obtain that the following inequality represents an NSI constraint

$$\begin{aligned} & (1 + |E_{\rm A}| + |E_{\rm B}| + E_{\rm AB})^2 \\ & + (1 + |E_{\rm A}| + |E_{\rm C}| + E_{\rm AC})^2 \\ & + (1 + |E_{\rm B}| + |E_{\rm C}| + E_{\rm BC})^2 \\ & \leq 6(1 + |E_{\rm A}|)(1 + |E_{\rm B}|)(1 + |E_{\rm C}|)\ . \end{aligned} \quad (16)$$

A proof of this general inequality is given in Supplementary Note 1. Note that this inequality reduces to Eq. (14) when $E_{\rm A} = E_{\rm B} = E_{\rm C} = 0$, as well as to Eq. (15) for the symmetric case.

It is worthwhile discussing the connection between our approach and the inflation technique presented in refs. [22,25]. There, the main focus is on using inflated networks for deriving constraints on correlations achievable, with classical resources. In that case, information can be readily copied, so that sources can send the same information to several parties. Ultimately, this allows for a full characterization of correlations achievable with classical resources[22]. Copying information is however not possible in our case, as no-signaling resources cannot be perfectly cloned in general[6]. Hence only inflated networks with bipartite

sources can be considered in our case, such as the hexagon. A discussion of these ideas can be found in Section V.D of ref. [22], where the idea of using inflation to limit no-signaling correlations in networks is mentioned. Here, we derive explicitly bounds that all correlations satisfying the NSI constraints, whether quantum of post-quantum, have to satisfy, and identify the physical principle behind them.

Finally, the choice of the hexagon inflation deserves a few words. As seen from Fig. 1b, it is judicious to consider inflated networks forming a ring, with a number of parties that is a multiple of three. Intuitively, this should enforce the strongest constraints on the correlations of the inflated network; in particular, all single- and two-body marginals are fixed by the correlations of the triangle. This would not be the case when considering inflations to ring networks, with a number of parties that is not divisible by three.

Tightness. A natural question is whether the constraints we derived above, that are necessary to satisfy NSI, are also sufficient. There is a priori no reason why this should be the case. Of course, starting from the triangle network, there are many (in fact infinitely many) possible extended networks that can be considered, and no-signaling must be enforced in all cases. For instance, instead of extending the network to a hexagon (as in Fig. 1), Alice could consider an extension to a ring network featuring 9, 12, or more parties. Clearly, such extensions could lead to stronger constraints than those derived here for the hexagon network.

Nevertheless, we show that some of the constraints we obtain above are in fact tight, i.e., necessary and sufficient for NSI. We prove this by presenting explicit correlations (constructed within a generalized probabilitic theory satisfying NSI) that saturate these constraints. In fact, we consider simply the case where all sources distribute classical variables to each party, which we refer to as trilocal models. The latter give rise to correlations of the form

$$p(a,b,c) = \int \mu(\alpha)d\alpha \int \nu(\beta)d\beta \int \omega(\gamma)d\gamma \; p_A(a|\beta,\gamma)\; p_B(b|\alpha,\gamma)\; p_C(c|\alpha,\beta)\,, \tag{17}$$

where α, β, and γ represent the three local variables distributed by each source, with arbitrary probability densities $\mu(\alpha)$, $\nu(\beta)$, and $\omega(\gamma)$. Also, $p_A(a|\beta,\gamma)$ represents an arbitrary response function for Alice, and similarly for $p_B(b|\alpha,\gamma)$ and $p_C(c|\alpha,\beta)$. Note that such trilocal models represents a natural extension of the concept of Bell locality to networks (see e.g., refs. [10,19]).

We first consider the case of symmetric distributions, i.e., characterized by the two parameters E_1 and E_2, and seek to determine the set of correlations that can be achieved with trilocal models. As shown in Fig. 2, it turns out that almost all NSI constraints can be saturated in this case, in particular the inequality (15). After performing a numerical search, we could construct explicitly some of these trilocal models, which involve up to ternary local variables (see Supplementary Note 2 for details). Moreover, we compare our NSI constraint (15) to the one derived in ref. [22] (see Eq. (34)), and find that the present one is stronger, and in fact tight (Fig. 2). Note also that a previous work derived an NSI constraint based on entropic quantities[29]; such constraints are however known to be generally weak, as entropies are a coarse-graining of the statistics, which no longer distinguishes between correlations and anticorrelations.

As seen from Fig. 2, there is however a small region (in yellow) that is compatible with NSI (considering the hexagon inflation), but for which we could not construct a trilocal model. Whether this gap can be closed by considering more sophisticated local models (using variables of larger alphabet) or whether stronger no-signaling bounds can be obtained is an interesting open question.

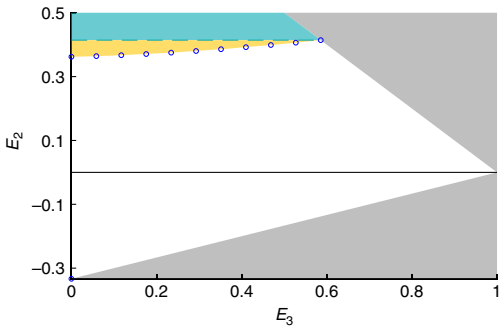

Fig. 3 Region of allowed correlations for symmetric distributions with $E_1 = 0$; represented in the plane E_2 vs E_3. The turquoise region is ruled out by NSI constraints (dashed turquoise line given by Eq. (15)), while the gray region is excluded from simple positivity constraints. The white region is accessible via trilocal models. Correlations in the yellow region satisfy NSI constraints (from the hexagon inflation), but we could not find a trilocal model for them. Explicit trilocal models are also obtained for the correlations marked by blue dots (see Supplementary Note 2).

For the triangle network with binary outcomes, any trilocal distribution can be obtained by considering shared variables of dimension (at most) six, and deterministic response functions[24].

In fact, another (and arguably much more interesting) possibility would be that this gap cannot be closed, as it would feature correlations with binary outcomes satisfying NSI, but that are nevertheless non-trilocal. To further explore this question, let us now focus on the case where single-party marginals vanish, i.e., $E_1 = 0$. We investigate the relation between two-party marginals E_2 and the three-party correlator $E_3 = E_{ABC}$, comparing NSI constraints and trilocal models. Notice that the NSI constraints we obtain here do not involve E_3 (as the latter cannot be recovered within the analysis of the hexagon). Hence NSI imposes only $E_2 \leq \sqrt{2} - 1$, while positivity of $p(a, b, c)$ imposes other constraints. This is shown in Fig. 3, where we also seek to characterize the set of correlations achievable via trilocal models (proceeding as above). Interestingly, we find again a potential gap between trilocal correlations and NSI constraints. This should however be considered with care. First, the NSI constraints obtained from the hexagon may not be optimal (see Discussion section). Second, there could exist more sophisticated trilocal models (e.g., involving higher-dimensional variables) that could lead to a stronger correlations (i.e., cover a larger region in Fig. 3). Note also that we investigated whether quantum distributions satisfying the independence assumption exist outside of the trilocal region, but we could not find any example (we performed a numerical search, considering entangled states of dimension up to 4×4).

Finally, note that we also performed a similar analysis for the case where single-party marginals vanish, but two-body marginals are not assumed to be identical to each other. Here, we find that inequality (13) can be saturated in a few specific cases. However, there also exist correlations satisfying the NSI bounds that do not seem to admit a trilocal model; details in Supplementary Note 1.

Discussion

We discussed the constraints arising on correlations in networks, under the assumption of NSI of the sources. We focused our attention on the triangle network with binary outputs for which we derived strong constraints, including tight ones. Our work raises a number of open questions that we now discuss further.

A first question is whether the constraints we derive (necessary under NSI), could also be sufficient. We believe this not to be the case, as stronger NSI constraints could arise from inflations of the triangle to more complex networks (e.g., loop networks with an arbitrary number of parties). Note that there could also exist different forms of no-signaling constraints, that cannot be enforced via inflation. In this respect, we compare in Supplementary Note 1 our NSI constraints with the recent work of ref. [32] proposing a very different approach to this problem, using the Finner inequality. A notable difference is that the latter imposes constraints on tripartite correlations, which is not the case here.

Another important question is whether there could exist nonlocality in the simplest triangle network with binary outcomes. That is, can we find a $p(a, b, c)$ that satisfies NSI, but that is nevertheless non-trilocal? While we identified certain potential candidate distributions for this, we could not prove any conclusive result at this point. We cannot exclude the possibilities that (i) these correlations are in fact not compatible with NSI (as there exist stronger NSI constraints) or (ii) these correlations can in fact be reproduced by a trilocal model. In order to address point (i), one could try to reproduce these correlations via an explicit NSI model, for instance considering that all sources emit no-signaling resources (such as nonlocal boxes[2]) which could then be wired together by the parties. To address point (ii), one could show that these correlations violate a multilocality inequality for the triangle network. Of course finding such inequalities is notably challenging, see e.g., ref. [13].

Furthermore, it would be interesting to derive NSI constraints for other types of networks. Indeed, the approach developed here can be straightforwardly used. Cases of high interest are general loop networks, as well as the triangle network with larger output alphabet (where examples of quantum nonlocality are proven to exist[11,15]).

Finally, a more fundamental question is whether any correlation satisfying the complete NSI constraints can be realized within an explicit physical theory satisfying no-signaling (the latter are usually referred to as generalized probabilistic theories[6]). While this is the case in the standard Bell scenario (where all parties share a common resource), it is not clear if that would also be the case in the network scenario.

Received: 28 October 2019; Accepted: 16 April 2020;
Published online: 13 May 2020

References

1. Bell, J. S. On the Einstein Podolsky Rosen paradox. *Physics* **1**, 195 (1964).
2. Popescu, S. & Rohrlich, D. Quantum nonlocality as an axiom. *Found. Phys.* **24**, 379–385 (1994).
3. Barrett, J. et al. Nonlocal correlations as an information-theoretic resource. *Phys. Rev. A* **71**, 022101 (2005).
4. Van Dam, W. Implausible consequences of superstrong nonlocality. *Nat. Comput.* **12**, 9–12 (2013).
5. Brassard, G. et al. Limit on nonlocality in any world in which communication complexity is not trivial. *Phys. Rev. Lett.* **96**, 250401 (2006).
6. Barrett, J. Information processing in generalized probabilistic theories. *Phys. Rev. A* **75**, 032304 (2007).
7. Pawłowski, M. et al. Information causality as a physical principle. *Nature* **461**, 1101–1104 (2009).
8. Brunner, N., Cavalcanti, D., Pironio, S. & Scarani, V. Wehner, S. Bell nonlocality. *Rev. Mod. Phys.* **86**, 419 (2014).
9. Popescu, S. Nonlocality beyond quantum mechanics. *Nat. Phys.* **10**, 264–270 (2014).
10. Branciard, C., Gisin, N. & Pironio, S. Characterizing the nonlocal correlations created via entanglement swapping. *Phys. Rev. Lett.* **104**, 170401 (2010).
11. Fritz, T. Beyond Bell's theorem: correlation scenarios. *New J. Phys.* **14**, 103001 (2012).
12. Branciard, C., Rosset, D., Gisin, N. & Pironio, S. Bilocal versus nonbilocal correlations in entanglement-swapping experiments. *Phys. Rev. A* **85**, 032119 (2012).
13. Gisin, N. Entanglement 25 years after quantum teleportation: testing joint measurements in quantum networks. *Entropy* **21**, 325 (2019).
14. Fraser, T. C. & Wolfe, E. Causal compatibility inequalities admitting quantum violations in the triangle structure. *Phys. Rev. A* **98**, 022113 (2018).
15. Renou, M.-O. et al. Genuine quantum nonlocality in the triangle network. *Phys. Rev. Lett.* **123**, 140401 (2019).
16. Pusey, M. F. Quantum correlations take a new shape. *Physics* **12**, 106 (2019).
17. Chaves, R. & Fritz, T. Entropic approach to local realism and noncontextuality. *Phys. Rev. A* **85**, 032113 (2012).
18. Tavakoli, A., Skrzypczyk, P., Cavalcanti, D. & Acín, A. Nonlocal correlations in the star-network configuration. *Phys. Rev. A* **90**, 062109 (2014).
19. Rosset, D. et al. Nonlinear Bell inequalities tailored for quantum networks. *Phys. Rev. Lett.* **116**, 01040 (2016).
20. Chaves, R. Polynomial Bell inequalities. *Phys. Rev. Lett.* **116**, 010402 (2016).
21. Tavakoli, A. Quantum correlations in connected multipartite Bell experiments. *J. Phys. A Math. Theor.* **49**, 145304 (2016).
22. Wolfe, E., Spekkens, R. W. The inflation technique for causal inference with latent variables. *J. Causal Inference* **7** https://doi.org/10.1515/jci-2017-0020 (2019).
23. Lee, C. M. & Spekkens, R. W. Causal inference via algebraic geometry: feasibility tests for functional causal structures with two binary observed variables. *J. Causal Inference* **5** https://doi.org/10.1515/jci-2016-0013 (2017).
24. Rosset, D., Gisin, N. & Wolfe, E. Universal bound on the cardinality of local hidden variables in networks. *Quantum Inf. Comput.* **18**, 910–926 (2018).
25. Navascues, M. & Wolfe, E. The inflation technique completely solves the causal compatibility problem. Preprint at https://arxiv.org/abs/1707.06476 (2017).
26. Luo, M.-X. Computationally efficient nonlinear Bell inequalities for quantum networks. *Phys. Rev. Lett.* **120**, 140402 (2018).
27. Canabarro, A., Brito, S. & Chaves, R. Machine learning nonlocal correlations. *Phys. Rev. Lett.* **122**, 200401 (2019).
28. Pozas-Kerstjens, A. et al. Bounding the sets of classical and quantum correlations in networks. *Phys. Rev. Lett.* **123**, 140503 (2019).
29. Henson, J., Lal, R. & Pusey, M. F. Theory-independent limits on correlations from generalized bayesian networks. *New. J. Phys.* **16**, 113043 (2014).
30. Fritz, T. Beyond Bell's theorem II: scenarios with arbitrary causal structure. *Commun. Math. Phys.* **341**, 391–434 (2016).
31. Chaves, R. & Budroni, C. Entropic nonsignaling correlations. *Phys. Rev. Lett.* **116**, 240501 (2016).
32. Renou, M.-O. et al. Limits on correlations in networks for quantum and no-signaling resources. *Phys. Rev. Lett.* **123**, 070403 (2019).
33. Weilenmann, M. & Colbeck, R. Analysing causal structures in generalised probabilistic theories. *Quantum* **4**, 2020 (2020).
34. Ziegler, G. *Lectures on Polytopes* (Springer, New York, 1998).

Acknowledgements

We thank Stefano Pironio, Marc-Olivier Renou, Denis Rosset, and Elie Wolfe for discussions. We acknowledge financial support from the Swiss national science foundation (Starting grant DIAQ, NCCR-QSIT, and NCCR-Swissmap). E.Z.C. acknowledges support by the Swiss National Science Foundation via the Mobility Fellowship P2GEP2_188276.

Author contributions

N.G., S.P., and N.B. came up with the idea of the method. N.G., J.-D.B., Y.C., P.R., A.T., E.Z.C., S.P., and N.B. participated in deriving the results, and writing and editing the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-16137-4.

Correspondence and requests for materials should be addressed to N.G., J.-D.B. or Y.C.

Peer review information *Nature Communications* thanks Gilles Brassard and the other, anonymous, reviewer for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Experimental Certification of Sustained Entanglement and Nonlocality after Sequential Measurements

Giulio Foletto,[1] Luca Calderaro,[1] Armin Tavakoli,[2] Matteo Schiavon,[1] Francesco Picciariello,[1] Adán Cabello,[3,4] Paolo Villoresi,[1,5] and Giuseppe Vallone[1,6,*]

[1] *Dipartimento di Ingegneria dell'Informazione, Università di Padova, IT-35131 Padova, Italy*
[2] *Département de Physique Appliquée, Université de Genève, CH-1211 Genève, Switzerland*
[3] *Departamento de Física Aplicada II, Universidad de Sevilla, E-41012 Sevilla, Spain*
[4] *Instituto Carlos I de Física Teórica y Computacional, Universidad de Sevilla, E-41012 Sevilla, Spain*
[5] *Istituto di Fotonica e Nanotecnologie, CNR, IT-35131 Padova, Italy*
[6] *Dipartimento di Fisica e Astronomia, Università di Padova, IT-35131 Padova, Italy*

(Received 20 June 2019; revised manuscript received 12 February 2020; accepted 28 February 2020; published 3 April 2020)

Entanglement is a fundamental resource for quantum information science. However, bipartite entanglement is destroyed when one particle is observed via projective (sharp) measurements, as is typically the case in most experiments. Here, we show experimentally that, if instead of sharp measurements, one performs many sequential unsharp measurements on one particle that are suitably chosen depending on the previous outcomes, then entanglement is preserved and it is possible to reveal quantum correlations through measurements on the second particle at any step of the sequence. Specifically, we observe that pairs of photons entangled in polarization maintain their entanglement when one particle undergoes three sequential measurements and that each of these can be used to violate a Clauser-Horne-Shimony-Holt inequality. This proof-of-principle experiment demonstrates the possibility of repeatedly harnessing two crucial resources, entanglement and Bell nonlocality, that, in most quantum protocols, are destroyed after a single measurement. The protocol we use, which in principle works for an unbounded sequence of measurements, can be useful for randomness extraction.

DOI: 10.1103/PhysRevApplied.13.044008

I. INTRODUCTION

Entanglement is at the heart of foundational and applied aspects of quantum theory [1]. Its paradigmatic applications include cryptography [2], teleportation [3], metrology [4], and device-independent quantum information [5]. However, it is also a fragile resource. The prolonged exposure of an entangled system to spontaneous decohering influences from the surrounding environment leads to its decay and eventual disappearance [6,7]. Furthermore, entanglement can vanish due to local measurements performed on one or several of the entangled systems. In particular, bipartite entanglement is completely destroyed as soon as a sharp measurement (i.e., a nondegenerate projective measurement) is performed on one of the two entangled systems [8]. For example, a sharp measurement of the spin along the x direction on one of the two spin qubits in a maximally entangled state leaves the qubits in a product state. Nonetheless, such entanglement-breaking measurements are commonplace in entanglement-based applications of quantum theory. Moreover, when applied to suitable entangled states, they typically give rise to the strongest quantum correlations in tests of Bell inequalities [9]. This certifies the presence of entanglement in a device-independent manner.

Recently, however, a number of works have considered the generation of entanglement-based quantum correlations in scenarios in which physical systems undergo several sequential measurements [10–12]. It has been found possible to perform local measurements on an entangled state such that the resulting correlations violate a Bell inequality but the postmeasurement state nevertheless remains entangled enough to make yet another Bell-inequality violation achievable [10]. Naturally, this feat is impossible with projective measurements. The measurements must be sharp enough to generate correlations that cannot be classically modeled but, nevertheless, unsharp enough so that some entanglement is still preserved after the measurement to make another Bell-inequality violation possible. These sequential unsharp measurements have been applied in studies of incompatible observables [13], state tomography [14], contextuality [15], and self-testing of quantum instruments [16,17]. Entanglement-based protocols using them have been proposed for certifying an

*vallone@dei.unipd.it

unbounded amount of device-independent [18] and one-sided device-independent random numbers [19], as well as for tests of finite-memory classical systems [20].

These advances make it relevant to develop experimental tools for sustaining entanglement over sequential measurements. While it has already been shown that appropriately chosen unsharp measurements do not destroy entanglement [21] and that others are capable of certifying it [22,23], proving experimentally that it is possible to do both things in a sequential manner remains a challenge. Notably, two sequential violations of the Clauser-Horne-Shimony-Holt (CHSH) Bell inequality [24] have been demonstrated [25,26]. However, extending the sequence to three and more measurements is demanding due to the sensitivity to noise [10]. Here, we demonstrate the ability to sustain entanglement over sequential measurements in a scenario in which the measurement choices depend on the history of previously performed measurements and observed outcomes. Since a given sequence of measurement choices and observed outcomes determines the evolution of the original state, one is faced with the task of demonstrating sustained entanglement along every possible branch of the resulting treelike structure of possible evolutions. We accomplish this for three sequential measurements on an entangled state, either by observing a violation of the CHSH inequality [24] or with a suitable entanglement witness. In principle, the protocol we use works for an unbounded sequence of measurements and can be useful for randomness extraction [18,19,27–31].

II. THEORETICAL MODEL

Consider a scenario in which two separated parties, Alice and Bob, share a two-qubit maximally entangled state $|\psi_1\rangle = 1/\sqrt{2}\,(|00\rangle + |11\rangle)$. Alice performs sequential measurements on her part of the state. In the first step, she randomly selects one of two dichotomic observables A_0 and A_1,

$$A_m(\mu_1) = K^\dagger_{+1|m}(\mu_1)K_{+1|m}(\mu_1) - K^\dagger_{-1|m}(\mu_1)K_{-1|m}(\mu_1), \tag{1}$$

where $m \in \{0,1\}$ and the Kraus operators $K_{\pm 1|m}$ are defined by

$$\begin{aligned} K_{+1|m}(\mu_1) &= \cos(\mu_1)\Pi^+_m + \sin(\mu_1)\Pi^-_m, \\ K_{-1|m}(\mu_1) &= \sin(\mu_1)\Pi^+_m + \cos(\mu_1)\Pi^-_m. \end{aligned} \tag{2}$$

Here, Π^+_0 and Π^-_0 (Π^+_1 and Π^-_1) are the projectors onto the positive and negative eigenvectors of σ_Z (σ_X), respectively. Moreover, the parameter $\mu_1 \in [0, \pi/4]$ can be used to tune the sharpness of her measurement [32]. On the one end, choosing $\mu_1 = 0$ means that the measurement is sharp (projective) and therefore consumes all the entanglement of the shared state. On the other end, choosing $\mu_1 = \pi/4$ means that the measurement is noninteractive ($K_{\pm 1|m} = \mathbb{1}/\sqrt{2}$) and therefore produces random outcomes, leaving the shared state unaltered. Choosing $\mu_1 \in (0, \pi/4)$ corresponds to an unsharp but nevertheless interactive measurement. Depending on Alice's choice of measurement and her observed outcome, the postmeasurement state ends up in one of four possible configurations. Since it is necessarily pure, it can be written in the form

$$\begin{aligned} |\psi_2\rangle &= \frac{K_{\pm 1|m}(\mu_1)}{\sqrt{\langle\psi_1|\, K^\dagger_{\pm 1|m}(\mu_1)K_{\pm 1|m}(\mu_1)\,|\psi_1\rangle}}\,|\psi_1\rangle \\ &= U_{A,2} \otimes U_{B,2}[\cos(\eta_2)\,|00\rangle + \sin(\eta_2)\,|11\rangle] \end{aligned} \tag{3}$$

for some angle $\eta_2 \in (0, \pi/4]$ that quantifies the entanglement in the state and some unitary transformations $U_{A,2}$ and $U_{B,2}$ that depend on Alice's choice of measurement and observed outcome.

In the second step in the sequence, Alice uses her knowledge of the measurement choice and the recorded outcome to apply $U^\dagger_{A,2}$ to her system. Then, she again randomly chooses between the measurements in Eq. (2), with the sharpness parameter denoted by μ_2. Again, the global state $|\psi_3\rangle$ after Alice's second measurement can end up in one of four possible configurations (given knowledge of the postmeasurement state after the first step of the protocol) and it can again be written on the form of Eq. (3), with suitable angles and unitary operations.

Acting in analogy with the second step, Alice can indefinitely continue the protocol and hence perform an arbitrarily long sequence of measurements. At the generic step k, the state is described by

$$|\psi_k\rangle = U_{A,k} \otimes U_{B,k}[\cos(\eta_k)\,|00\rangle + \sin(\eta_k)\,|11\rangle]. \tag{4}$$

In Table III (Appendix A), we give exact expressions for unitary operations $U_{A,k}$, $U_{B,k}$ and parameter η_k, which depend on the history of Alice's measurements and outcomes up to step $k-1$. Alice applies $U^\dagger_{A,k}$ to her subsystem; she performs either measurement A_0 or A_1 with strength parameter μ_k and the state takes again the form of Eq. (4), with k replaced by $k+1$ so that step $k+1$ can begin. We note that if Alice chooses $\mu_j > 0\ \forall j \le k$, then $\eta_{k+1} > 0$, meaning that $|\psi_{k+1}\rangle$ is still entangled. Not only this: if she uses measurement A_0 with strength parameter $\mu_k > \arctan(\tan^2\eta_k)$ and finds outcome -1, the new entanglement parameter is $\eta_{k+1} = \arctan(\tan\mu_k/\tan\eta_k) > \eta_k$ and therefore entanglement has been amplified.

At any step k, the protocol can be interrupted for the purpose of certifying that entanglement is still present via a violation of the CHSH inequality. Bob must apply $U^\dagger_{B,k}$, projectively measure either observable $B_{0,k} = \cos(\theta_k)\sigma_X + \sin(\theta_k)\sigma_Z$ or $B_{1,k} = -\cos(\theta_k)\sigma_X +$

$\sin(\theta_k)\sigma_Z$, where $\theta_k = \text{arccot}[\sin(2\eta_k)]$, and finally record outcome ± 1. Then, he can correlate his results with those of Alice at the same step k and calculate the CHSH quantity

$$S_{\text{CHSH}} \equiv \langle A_0 B_0 \rangle + \langle A_0 B_1 \rangle + \langle A_1 B_0 \rangle - \langle A_1 B_1 \rangle. \quad (5)$$

A violation of the CHSH inequality ($S_{\text{CHSH}} \leq 2$) certifies that the experiment involves inherently quantum effects. With the choice of parameters that we outline, Bob finds that

$$S_{\text{CHSH},k} = 2\cos(2\mu_k)\sqrt{1+\sin^2(2\eta_k)}, \quad (6)$$

meaning that this violation ($S_{\text{CHSH},k} > 2$) happens for any choice of k if Alice uses $\mu_k < \mu_{k,\max} = \frac{1}{2}\arctan(\sin 2\eta_k)$. In order to find θ_k and $U_{B,k}$, Bob needs to know Alice's history of measurements and outcomes at steps $1 \ldots k-1$. We can imagine that Alice feeds them back to him after each step or that he selects his operations at random from all his possibilities and then correlates his results only with those of Alice for which his choice was right. When Bob acts, he only certifies the entanglement of the state after one specific history of Alice's measurements and outcomes. However, in sufficiently many runs of the protocol, he can cover many different histories.

This protocol underlines two points:

(a) If they are weak, Alice's measurements do not destroy entanglement ($\mu_k > 0 \rightarrow \eta_{k+1} > 0$).

(b) If they are not too weak, the same weak measurements that preserve entanglement are able to extract enough information to violate a CHSH inequality and thus certify the presence of entanglement in the premeasurement state ($\mu_k < \mu_{k,\max} \rightarrow S_{\text{CHSH},k} > 2$).

This means that Alice can perform an arbitrary number of measurements on every single entangled system, each time fulfilling the certification requirement and without ever destroying entanglement. For random-number generation, she could extract more than 1 bit of certified local randomness from the sequences in which she only measures σ_X, thus beating the limit imposed by projective measurements. For instance, with a short sequence of two measurements, with $\mu_1 = 0.13$ and $\mu_2 = 0$, she would certify 1.026 bits.

III. EXPERIMENTAL METHOD

We describe here our proof-of-concept implementation aimed at verifying the two points above. Alice makes at most three sequential measurements and the protocol can be stopped at step 1, 2, or 3. We choose $\mu_1 \approx 0.34$, $\mu_2 \approx 0.19$ and $\mu_3 = 0$. The former two parameters optimize the expected values of S_{CHSH} around 2.2 at steps 1 and 2, enough to grant a significant experimental observation without sacrificing the value at step 3. We can set $\mu_3 = 0$ because we are sure that the protocol will not continue after step 3 and therefore there is no need to preserve entanglement. Just before step 3, the shared system can be in 16 possible states, depending on Alice's previous choices and outcomes. Although a CHSH-inequality violation is possible for all of them, only in four cases is the achievable value of S_{CHSH} sufficiently greater than 2, to admit the experimental detection. In the remaining 12 cases, we verify entanglement using a different strategy: Alice and Bob apply the operation $U_A^\dagger \otimes U_B^\dagger$ and then measure the entanglement witness $W = \mathbb{1}\otimes\mathbb{1} - \sigma_Z\otimes\sigma_Z - \sigma_X\otimes\sigma_X$. It is easy to prove that the mean value of this witness is negative on the state of interest, whereas it would be positive or zero on any separable state [33]. In total, we measure nine independent values of S_{CHSH} (one when we stop the protocol at step 1, four at step 2, and four at step 3), plus 12 values of $\langle W \rangle$.

We encode two qubits in the polarization degree of freedom of two separated photons. Polarization-entangled photon pairs are generated by a custom-built source [14,25] based on a Sagnac interferometer. It prepares the entangled state $|\psi_1\rangle$, where $|0\rangle$ and $|1\rangle$ refer to the horizontal and vertical polarizations. The pairs are sent to the two arms of our experimental setup, which correspond to Alice and Bob in the theoretical protocol. Figure 1 schematizes the optical implementation for each of Alice's measurement steps: two half-wave plates (HWPs) apply $U_A^\dagger$, which is always a rotation in the space of linear polarizations; then, another HWP represents the choice between the measurements A_0 and A_1; and, finally, a

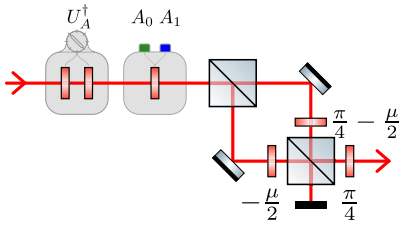

FIG. 1. A conceptual optical scheme for each of Alice's steps. Angle μ corresponds to the sharpness parameter in Alice's measurements. The two states obtained at the two outputs correspond to the Kraus operators $K_{+1|m}$ and $K_{-1|m}$ applied to the input state, meaning that each measurement outcome is mapped to an output. In this model, mirrors apply the σ_Z operation to incident polarization, whereas polarizing beam splitters simply separate two orthogonal polarizations without introducing any relative phase. In our implementation, one of the exits is blocked and we change the outcome corresponding to the active one by rotating the external wave plates. For a key to the optical elements, see Fig. 2.

polarization-based Mach-Zehnder interferometer (MZI) implements the unsharp measurement. It entangles the polarization with the path degree of freedom, while the sharpness parameter is set by the angles $-\mu/2$ and $\pi/4 - \mu/2$ of the internal HWPs. The two exit paths correspond to the two outcomes of the polarization measurement: the probability of a photon taking each of them is equal to the probability of each outcome, while the polarization state of the photon (if observed in each path) is the expected postmeasurement state. One can imagine putting many of these devices in a treelike structure that, in principle, can grow unlimited, but in our experiment we stop after three of them. Every branch of the tree corresponds to a particular history of outcomes: detecting a photon at the end of a branch allows us to retrieve this history, attesting that the photon has taken the corresponding sequence of exits. Alice does not need to retrieve the outcome after each measurement, because the wave plates are set to execute the correct unitary operation $U_A^\dagger$, depending on which branch they are in. This encoding of measurement outcomes in the path degree of freedom is common in experiments involving sequential measurements on single photons [34–39]. Figure 2 depicts this idea.

Bob makes only projective measurements in the space of linear polarization; hence his scheme can be simplified to a HWP that selects the observable and a polarization beam splitter (PBS) that separates the two outcomes.

In practice, our implementation is simplified with respect to Fig. 2 and only uses one detector (a single-photon avalanche diode, SPAD) for Alice and one for Bob. Since we set Alice's third measurement to be projective, we need only two MZIs in a sequence on her side. One exit of each is blocked, so that there is only one path from the source to Alice's detector. Each interferometer is set to change the input polarization according to the Kraus operator $K_{+1|0}(\mu) = \cos(\mu)\Pi_0^+ + \sin(\mu)\Pi_0^-$, while two HWPs, one before and one after it, can change any of $\{\Pi_0^+, \Pi_0^-, \Pi_1^+, \Pi_1^-\}$ into another, thus selecting the basis and outcome of the measurement. This means that depending on the orientation of these plates, the interferometer can carry out each of the four Kraus operators required by the protocol. We mechanically rotate the HWPs in different configurations, each corresponding to one measurement-outcome combination. By orientating all of Alice's HWPs properly, we select which branch of the tree is implemented by the one path of our setup, thus setting her complete history of measurements and outcomes. The choice of measurement bases is not made in real time, photon by photon, as a faithful realization of the protocol would require, but, rather, at fixed temporal intervals, the length of which is limited by the speed of rotation of the plates and the integration time needed to keep the statistical error small enough. Moreover, the setup cannot evaluate different measurement outcomes simultaneously but we have to check their relative frequencies one by one. We evaluate sequentially all the combinations of plate orientations, thus reconstructing the entire tree one branch at a time. For each combination, we count coincident detections between Alice and Bob for a fixed exposure time. These counts are proportional to the joint probability of obtaining the combination of outcomes under test and hence allow us to find S_{CHSH} and $\langle W \rangle$. We note that Alice never communicates her previous history of measurements and outcomes to Bob: we externally choose it and then select the same plate orientations that Bob would use if he received such a message.

We operate under the fair-sampling assumption that coincident detection events faithfully represent the photon pairs produced by the crystal. Moreover, our setup is affected by the "locality loophole," i.e., classical communication between Alice and Bob during the measurement of S_{CHSH} cannot be physically excluded.

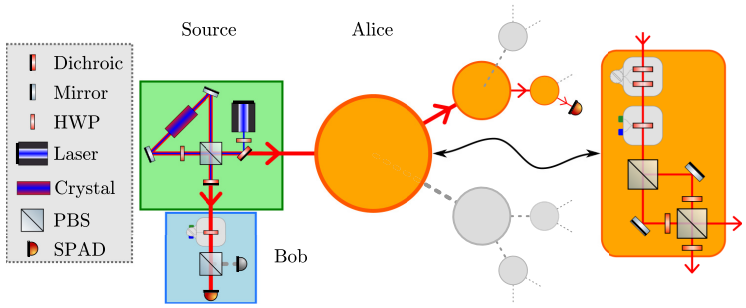

FIG. 2. A conceptual treelike structure of the protocol. The two positions of Bob's HWP if he stops the protocol at the kth step depend on Alice's choices and outcomes of all the previous $k-1$ steps. Also, those of the HWPs that implement $U_A^\dagger$ inside each of Alice's blocks depend on her previous history. In our implementation, Alice stops at most at the third measurement. Moreover, we do not build the entire tree but only one branch and we change the combination of outcomes to which it corresponds by rotating wave plates.

Finally, avoiding the tree structure increases the experiment duration, because the probabilities of outcome are not all recorded at once. Furthermore, it prevents our setup from being straightforwardly adapted for applications such as randomness extraction. However, it greatly simplifies the implementation for the goal of certifying entanglement.

IV. RESULTS

We use a coincidence window of ± 1 ns and an exposure time of 20 s for all measurements. Given the production rate of our source and the losses in the setup, the total number of photon pairs that contribute to our measurements is approximately 3×10^4. The detection efficiency of Alice's channel is approximately 1%, while Bob's is approximately 8%, with the difference being due to the multimode fiber on Bob's side (details in Appendix C). Before the experiment, we verify the quality of the initial entangled state using the visibility figure of merit and we obtain 99% and 98% when measuring the $\sigma_Z \otimes \sigma_Z$ and $\sigma_X \otimes \sigma_X$ correlations, respectively. The visibility in the former basis depends on the extinction ratio of the polarizing elements in the measurement setup, whereas in the latter basis it is limited by the quality of the Sagnac interferometer.

Table I shows the experimental results for the nine values of S_{CHSH}, whereas Table II shows the 12 mean values for the entanglement witnesses. For completeness, we also report in Table IV (Appendix B) the witnesses in the other four cases at step 3, for which the CHSH violation is a stronger certification of entanglement because it requires fewer assumptions [40]. We observe the violation of all the nine CHSH inequalities with several standard deviations of statistical significance, proving that Alice's sequential measurements do not destroy entanglement and at the same time can certify its presence. The former point is also corroborated by the results of $\langle W \rangle$, which are always significantly negative. We also note that the value of S_{CHSH} at step 3 is greater than those at steps 1 and 2. This is expected given the particular sharpness parameters that we use in the experiment and proves that the protocol can be used for entanglement amplification, although only for a subset of measurement choices and outcomes.

We still observe small deviations from the expected values and we attribute them to systematic alignment errors in our setup. Imperfections in one of Alice's interferometers might make the measurement that we perform suboptimal, thus reducing S_{CHSH} at the corresponding step. Moreover, they might degrade the entanglement in the output state, thus also decreasing S_{CHSH} at the steps that follow. Finally, the results at the first two steps can be influenced by defects in parts of the setup that ensue the corresponding interferometers, because photons must still go through these parts before they reach the detectors. The main sources of error are the phase between the arms of the MZIs, which has to be carefully regulated by tilting the PBDs, and rotation of the wave plates, which must be accurate. These rotations can also deviate the photons out of the detectors' entrance, thus invalidating the polarization measurements. Alignment difficulties such as these are the reason why simplification of the experimental setup is of paramount importance. Regarding statistical errors, we verify that the repeatability of the motorized rotators used to set the orientation of the wave plates is good enough that its contribution is negligible; hence the standard deviations reported in Tables I and II are derived only from the Poissonian error on the photon counts.

TABLE I. The experimental values of S_{CHSH}. The second column reports the history of measurements and outcomes that precede the one that yields S_{CHSH} on Alice's side. The notation is as follows: outcome at step 1 | measurement choice at step 1; outcome at step 2 | measurement choice at step 2. The violation (i.e., $S_{\mathrm{CHSH}} - 2$) is expressed in units of the standard deviation on S_{CHSH}, derived from Poissonian error on the counts and error propagation.

Final step	Alice's history	S_{CHSH}	SD	Violation (units of SD)
1	not applicable	2.15	0.01	20
2	$+1\|0$	2.13	0.01	12
2	$-1\|0$	2.07	0.01	6
2	$+1\|1$	2.12	0.01	10
2	$-1\|1$	2.09	0.01	7
3	$+1\|0; -1\|0$	2.48	0.03	16
3	$-1\|0; -1\|0$	2.53	0.03	17
3	$+1\|1; -1\|0$	2.47	0.03	15
3	$-1\|1; -1\|0$	2.46	0.03	15

TABLE II. The experimental mean values of the entanglement witness. The final step is the third for all results. The first column reports the history of measurements and outcomes that precede the one that yields $\langle W \rangle$ on Alice's side. The notation is as follows: outcome at step 1 | measurement choice at step 1; outcome at step 2 | measurement choice at step 2. The last column reports $-\langle W \rangle$, expressed in units of its standard deviation, derived from Poissonian error on the counts and error propagation.

Alice's history	$\langle W \rangle$	SD	Confirmation (units of SD)
$+1\|0; +1\|0$	-0.12	0.01	13
$+1\|0; +1\|1$	-0.17	0.01	14
$+1\|0; -1\|1$	-0.20	0.01	17
$-1\|0; +1\|0$	-0.07	0.01	8
$-1\|0; +1\|1$	-0.12	0.01	11
$-1\|0; -1\|1$	-0.14	0.01	13
$+1\|1; +1\|0$	-0.06	0.01	7
$+1\|1; +1\|1$	-0.13	0.01	10
$+1\|1; -1\|1$	-0.18	0.01	14
$-1\|1; +1\|0$	-0.07	0.01	8
$-1\|1; +1\|1$	-0.17	0.01	13
$-1\|1; -1\|1$	-0.16	0.01	13

V. CONCLUSIONS

In this work, we show that it is experimentally feasible to sustain entanglement over a sequence of unsharp measurements while being able to generate correlations that are strong enough to violate a Bell inequality through the same measurements. We report strong violations of the CHSH inequality, backed by more than 10 standard deviations of statistical significance, even at the third step of the sequence (albeit only for some of the possible histories of previous measurements and outcomes). This is important for protocols that require certified entanglement for quantum information tasks, such as the extraction of random bits from measurement outcomes. Our proof-of-principle experiment is based on entangled photon pairs and exploits only three well-controlled sequential measurements. It would be of evident interest to extend these ideas to other relevant physical systems that make substantially longer sequences of unsharp measurements possible, allowing one to harness entanglement many times for quantum information applications.

ACKNOWLEDGMENTS

Part of this work was supported by the Italian Ministry of Education, University and Research (MIUR) under the initiative "Departments of Excellence" (Law 232/2016). In addition, G.F. would like to thank M. Zahidy for the useful discussions about the quantum measurement problem. A.T. was supported by the Swiss National Science Foundation (Starting Grant DIAQ, NCCR-QSIT).

APPENDIX A: DETAILED DESCRIPTION OF THE PROTOCOL

At the beginning of step k, Alice and Bob share the pure and entangled state

$$|\psi_k\rangle = U_{A,k} \otimes U_{B,k}[\cos(\eta_k)\,|00\rangle + \sin(\eta_k)\,|11\rangle], \quad \text{(A1)}$$

where $U_{A,k}$ and $U_{B,k}$ are local unitary operations and $\eta_k \in (0, \pi/4]$. Alice has perfect knowledge of the state; hence she can apply $U_{A,k}^{\dagger}$ to her subsystem. The shared state becomes

$$|\psi_k'\rangle = \mathbb{1}_A \otimes U_{B,k}[\cos(\eta_k)\,|00\rangle + \sin(\eta_k)\,|11\rangle]. \quad \text{(A2)}$$

She chooses the strength of her measurement, in the form of parameter $\mu_k \in (0, \mu_{k,\max})$, where

$$\mu_{k,\max} = \frac{1}{2}\arctan(\sin 2\eta_k). \quad \text{(A3)}$$

We require $\mu_k > 0$ to preserve entanglement at step $k+1$ (indeed, Alice is allowed to choose $\mu_k = 0$ if she agrees with Bob to stop the protocol at step k). The upper bound is required to make the violation of the CHSH inequality at step k possible. Note that this implies that $\tan(2\mu_k) \leq \sin(2\eta_k) \leq \tan(2\eta_k)$ and hence $\mu_k \leq \eta_k$.

Then, she chooses between the two observables $A_0(\mu_k) = E_{+1|0}(\mu_k) - E_{-1|0}(\mu_k)$ and $A_1(\mu_k) = E_{+1|1}(\mu_k) - E_{-1|1}(\mu_k)$, where

$$\begin{aligned} E_{+1|0}(\mu_k) &= \frac{1}{2}[\mathbb{1} + \cos(2\mu_k)\sigma_Z], \\ E_{-1|0}(\mu_k) &= \frac{1}{2}[\mathbb{1} - \cos(2\mu_k)\sigma_Z], \\ E_{+1|1}(\mu_k) &= \frac{1}{2}[\mathbb{1} + \cos(2\mu_k)\sigma_X], \\ E_{-1|1}(\mu_k) &= \frac{1}{2}[\mathbb{1} - \cos(2\mu_k)\sigma_X]. \end{aligned} \quad \text{(A4)}$$

$A_0(\mu_k)$ and $A_1(\mu_k)$ are noisy measurements of σ_Z and σ_X, respectively. Moreover, $E_{\pm1|m}(\mu_k) = K_{\pm1|m}(\mu_k)^{\dagger}K_{\pm1|m}(\mu_k)$ where $K_{\pm1|m}(\mu_k)$ are the Kraus operators mentioned in the main text:

$$\begin{aligned} K_{+1|0}(\mu_k) &= \cos(\mu_1)\,|0\rangle\langle 0| + \sin(\mu_k)\,|1\rangle\langle 1|, \\ K_{-1|0}(\mu_k) &= \sin(\mu_1)\,|0\rangle\langle 0| + \cos(\mu_k)\,|1\rangle\langle 1|, \\ K_{+1|1}(\mu_k) &= \cos(\mu_1)\,|+\rangle\langle +| + \sin(\mu_k)\,|-\rangle\langle -|, \\ K_{-1|1}(\mu_k) &= \sin(\mu_1)\,|+\rangle\langle +| + \cos(\mu_k)\,|-\rangle\langle -|, \end{aligned} \quad \text{(A5)}$$

where $|+\rangle$ and $|-\rangle$ are the two eigenstates of σ_X.

After performing the measurement and recording the outcome, the shared state becomes

$$|\psi_{k+1}\rangle = U_{A,k+1} \otimes U_{B,k+1}[\cos(\eta_{k+1})|00\rangle + \sin(\eta_{k+1})|11\rangle] \quad \text{(A6)}$$

and step $k+1$ can begin.

The unitary operations $U_{A,k+1}$ and $U_{B,k+1}$ and the new parameter η_{k+1} can be found from their corresponding values at step k. In particular,

$$\begin{aligned} U_{A,k+1} &= e^{-i\alpha_{k+1}\sigma_Y}, \\ U_{B,k+1} &= e^{-i\beta_{k+1}\sigma_Y}U_{B,k}, \end{aligned} \quad \text{(A7)}$$

where angles α_{k+1} and β_{k+1} depend on the choice of measurement and outcome at step k, as summarized in Table III.

We emphasize that if the measurement choice is A_0, the outcome is -1 and $\tan(\mu_k) > \tan^2(\eta_k)$, then $\eta_{k+1} > \eta_k$, which means that entanglement has been amplified; this cannot happen in the other cases. To find the expressions of Table III, one should write $|\psi_{k+1}\rangle = K_{\pm1|m}(\mu_k)/\sqrt{\langle\psi_k'|\,E_{\pm1|m}(\mu_k)\,|\psi_k'\rangle}\,|\psi_k'\rangle$ and then perform the Schmidt decomposition on this state. The singular vectors (which form the columns of $U_{A,k+1}$ and $U_{B,k+1}$) should be ordered according to decreasing singular values. Then, $\tan(\eta_{k+1})$ is simply the ratio between the smaller and larger singular values. This sequence begins at step 1 with

TABLE III. The properties of step $k+1$ given those of step k.

Kraus at step k	α_{k+1}	β_{k+1}	η_{k+1}
$K_{+1\|0}$	0	0	$\arctan[\tan(\mu_k)\tan(\eta_k)]$
$K_{-1\|0}$	$\pi/2$	$\pi/2$	$\arctan[\tan(\mu_k)/\tan(\eta_k)]$
$K_{+1\|1}$	$\frac{1}{2}\text{arccot}[\tan(2\mu_k)\cos(2\eta_k)]$	$\frac{1}{2}\arctan[\tan(2\eta_k)\cos(2\mu_k)]$	$\frac{1}{2}\arcsin[\sin(2\mu_k)\sin(2\eta_k)]$
$K_{-1\|1}$	$-\frac{1}{2}\text{arccot}[\tan(2\mu_k)\cos(2\eta_k)]$	$-\frac{1}{2}\arctan[\tan(2\eta_k)\cos(2\mu_k)]$	$\frac{1}{2}\arcsin[\sin(2\mu_k)\sin(2\eta_k)]$

$U_{A,1} = U_{B,1} = \mathbb{1}$ and $\eta_1 = \pi/4$. With this information and the above updating rules, it is possible to find all parameters at all steps.

If Alice and Bob decide to interrupt the protocol at step k, Bob must apply $U^{\dagger}_{B,k}$ and measure projectively the two observables $B_{0,k} = \cos(\theta_k)\sigma_X + \sin(\theta_k)\sigma_Z$ and $B_{1,k} = -\cos(\theta_k)\sigma_X + \sin(\theta_k)\sigma_Z$, where $\theta_k = \text{arccot}[\sin(2\eta_k)]$. Inserting these expressions in the definition of S_{CHSH} yields Eq. (6). From this, one can prove that in order to violate the CHSH inequality, μ_k must be chosen such that $\tan(2\mu_k) < \sin(2\eta_k)$, as stated in the main text.

APPENDIX B: VALUES FOR THE THREE-STEPS IMPLEMENTATION

Table IV contains the numerical values for all the parameters of the protocol, restricted to our three-steps implementation.

APPENDIX C: DETAILED DESCRIPTION OF THE EXPERIMENTAL SETUP

The heart of our entangled-photons source is a 30-mm-long periodically poled potassium titanyl phosphate (PPKTP) crystal, which lies inside a Sagnac interferometer. A continuous-wave laser at 404 nm sends diagonally polarized light to the PBS of the interferometer so that the crystal is illuminated from both directions. By a spontaneous parametric down-conversion process in a quasiphase-matching configuration, pairs of orthogonally polarized photons at 808 nm are generated. Due to a dual wavelength half-wave plate (that works both at 404 nm and 808 nm) inside the Sagnac interferometer, the quantum state just after it is $1/\sqrt{2}(|01\rangle + |10\rangle)$, where the horizontal ($|0\rangle$) and vertical ($|1\rangle$) polarizations are defined by the aforementioned PBS. Two single-mode fibers collect the photons and bring them to the two arms of the measurement setup, Alice and Bob. In each of them, a HWP

TABLE IV. The values of the parameters for our three-steps implementation and comparison with the observed values for S_{CHSH} and $\langle W\rangle$. The notation for the second column is as follows: outcome at step 1 | measurement choice at step 1; outcome at step 2 | measurement choice at step 2. The standard deviations in the last two columns are derived from Poissonian error on the counts and error propagation.

Step k	Alice's history	η_k	α_k	β_k	θ_k	μ_k	S_{CHSH}	$\langle W\rangle$	S_{CHSH} (observed)	$\langle W\rangle$ (observed)
1	not applicable	$\pi/4$	0	0	$\pi/4$	0.34	2.20	-1	2.15 ± 0.01	no data
2	$+1\|0$	0.34	0	0	1.01	0.19	2.19	-0.63	2.13 ± 0.01	no data
2	$-1\|0$	0.34	$\pi/2$	$\pi/2$	1.01	0.19	2.19	-0.63	2.07 ± 0.01	no data
2	$+1\|1$	0.34	$\pi/4$	$\pi/4$	1.01	0.19	2.19	-0.63	2.12 ± 0.01	no data
2	$-1\|1$	0.34	$-\pi/4$	$-\pi/4$	1.01	0.19	2.19	-0.63	2.09 ± 0.01	no data
3	$+1\|0;+1\|0$	0.07	0	0	1.44	0	2.02	-0.14	no data	-0.12 ± 0.01
3	$+1\|0;-1\|0$	0.50	$\pi/2$	$\pi/2$	0.87	0	2.61	-0.84	2.48 ± 0.03	-0.75 ± 0.01
3	$+1\|0;+1\|1$	0.12	0.63	0.32	1.34	0	2.05	-0.23	no data	-0.17 ± 0.01
3	$+1\|0;-1\|1$	0.12	-0.63	-0.32	1.34	0	2.05	-0.23	no data	-0.20 ± 0.01
3	$-1\|0;+1\|0$	0.07	0	0	1.44	0	2.02	-0.14	no data	-0.07 ± 0.01
3	$-1\|0;-1\|0$	0.50	$\pi/2$	$\pi/2$	0.87	0	2.61	-0.84	2.53 ± 0.03	-0.79 ± 0.01
3	$-1\|0;+1\|1$	0.12	0.63	0.32	1.34	0	2.05	-0.23	no data	-0.12 ± 0.01
3	$-1\|0;-1\|1$	0.12	-0.63	-0.32	1.34	0	2.05	-0.23	no data	-0.14 ± 0.01
3	$+1\|1;+1\|0$	0.07	0	0	1.44	0	2.02	-0.14	no data	-0.06 ± 0.01
3	$+1\|1;-1\|0$	0.50	$\pi/2$	$\pi/2$	0.87	0	2.61	-0.84	2.47 ± 0.03	-0.78 ± 0.01
3	$+1\|1;+1\|1$	0.12	0.63	0.32	1.34	0	2.05	-0.23	no data	-0.13 ± 0.01
3	$+1\|1;-1\|1$	0.12	-0.63	-0.32	1.34	0	2.05	-0.23	no data	-0.18 ± 0.01
3	$-1\|1;+1\|0$	0.07	0	0	1.44	0	2.02	-0.14	no data	-0.07 ± 0.01
3	$-1\|1;-1\|0$	0.50	$\pi/2$	$\pi/2$	0.87	0	2.61	-0.84	2.46 ± 0.03	-0.68 ± 0.02
3	$-1\|1;+1\|1$	0.12	0.63	0.32	1.34	0	2.05	-0.23	no data	-0.17 ± 0.01
3	$-1\|1;-1\|1$	0.12	-0.63	-0.32	1.34	0	2.05	-0.23	no data	-0.16 ± 0.01

and a QWP correct the unitary operations applied by the fibers. Bob also uses a liquid-crystal retarder (LCR) to fine tune the phase between different polarization components. These optical elements change the state to

$$|\psi_1\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \tag{C1}$$

where $|0\rangle$ and $|1\rangle$ are now defined by Alice's and Bob's polarizers.

The principle of the two measurement setups is that polarizing and birefringent optical elements select one measurement effect and then their axes are rotated to evaluate sequentially all the effects of interest. The number of coincident detections in each configuration is counted and associated with the corresponding effect. Bob has to measure only linear polarizations; therefore his measurement setup consists of a rotating HWP and a fixed linear polarizer (LP). A multimode fiber then collects the photons and brings them to a SPAD. On Alice's side, two Mach-Zehnder interferometers in a series implement the two weak measurements. They separate the horizontal- and vertical-polarization components using PBDs. For convenience, we use three HWPs instead of two in our MZIs: in this way, we can regulate the sharpness parameter μ by rotating a single plate, while the others are fixed. Indeed, the arm carrying the $|0\rangle$ polarization encounters a HWP with an axis at $-\pi/8$, while the other encounters a HWP at $\pi/8$. Then, a HWP at angle $\pi/8 - \mu/2$ spans across both. The interferometer, followed by a HWP at angle $\pi/4$ that swaps $|0\rangle$ and $|1\rangle$ implements the Kraus operator:

$$K_{+1|0}(\mu) = \cos(\mu)\Pi_0^+ + \sin(\mu)\Pi_0^-. \tag{C2}$$

Two HWPs, one before and one after the MZI, can change any of $\{\Pi_0^+, \Pi_0^-, \Pi_1^+, \Pi_1^-\}$ into another, thus selecting the basis and outcome of the measurement. This means that depending on the orientation of these plates, the interferometer can carry out each of the four Kraus operators required by the protocol, shown in Eq. (A5).

The unitary operations needed before each weak measurement are realized by the HWP at the beginning of the next step. The total number of HWPs needed between the measurement steps would be five (one to select the measurement-outcome combination of the previous measurement, one to do the same for the next one, one to swap $|0\rangle$ and $|1\rangle$, and two for the unitary operation), but this can be reduced to one, as is true for any odd number of HWPs. Since the third measurement is strong, it is achieved by a HWP and a LP. A single-mode fiber finally collects Alice's photons and brings them to a SPAD, the signal of which is correlated with Bob's signal by a time tagger with 80-ps resolution, that then returns coincidence counts within a ± 1-ns window. A faithful representation of our implementation is shown in Fig. 3.

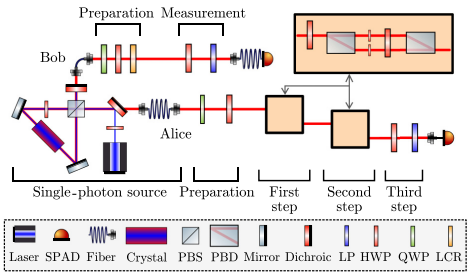

FIG. 3. The actual optical implementation.

The total rate of coincidences summed over the outcomes of a polarization measurement is about 1500 Hz. We measure the efficiency of Alice's detection system as the ratio between the rate of coincidences and that of single counts in Bob's channel and we obtain approximately 1%. This value includes the quantum efficiency of Alice's SPAD and losses in the optical system, but is mostly limited by the two couplings into single-mode fiber that photons must endure in their path from the crystal to the detector. Bob's efficiency is comparatively much better (approximately 8%) because of the multimode fiber (with a higher collection probability) that we use at the detection stage.

The coincidence rate, integrated over an exposure time of 20 s, makes the total number of coincident events contributing to a complete measurement about 3×10^4, sufficient to make statistical errors small. Systematic misalignments of the setup are the main source of error. In particular, imperfections in the wave plates can cause imbalances in the photon counts, which are critical for the final results. Rotating plates can slightly deviate the beam out of the fiber entrance, hindering the accuracy of the polarization measurements. Preparation of the entangled state is also important and needs precise alignment of the source. Finally, the Mach-Zehnder interferometers need to be perfectly balanced to achieve sufficient visibility.

[1] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, Rev. Mod. Phys. **81**, 865 (2009).
[2] A. K. Ekert, Quantum Cryptography Based on Bell's Theorem, Phys. Rev. Lett. **67**, 661 (1991).
[3] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Teleporting an Unknown Quantum State via Dual Classical and Einstein-Podolsky-Rosen Channels, Phys. Rev. Lett. **70**, 1895 (1993).
[4] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum-enhanced measurements: Beating the standard quantum limit, Science **306**, 1330 (2004).

[5] S. Pironio, V. Scarani, and T. Vidick, Focus on device independent quantum information, New J. Phys. **18**, 100202 (2016).

[6] T. Yu and J. H. Eberly, Finite-Time Disentanglement via Spontaneous Emission, Phys. Rev. Lett. **93**, 140404 (2004).

[7] M. P. Almeida, F. de Melo, M. Hor-Meyll, A. Salles, S. P. Walborn, P. H. Souto Ribeiro, and L. Davidovich, Environment-induced sudden death of entanglement, Science **316**, 579 (2007).

[8] C. A. Fuchs and A. Peres, Quantum-state disturbance versus information gain: Uncertainty relations for quantum information, Phys. Rev. A **53**, 2038 (1996).

[9] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).

[10] R. Silva, N. Gisin, Y. Guryanova, and S. Popescu, Multiple Observers Can Share the Nonlocality of Half of an Entangled Pair by Using Optimal Weak Measurements, Phys. Rev. Lett. **114**, 250401 (2015).

[11] S. Sasmal, D. Das, S. Mal, and A. S. Majumdar, Steering a single system sequentially by multiple observers, Phys. Rev. A **98**, 012305 (2018).

[12] A. Shenoy H, S. Designolle, F. Hirsch, R. Silva, N. Gisin, and N. Brunner, Unbounded sequence of observers exhibiting Einstein-Podolsky-Rosen steering, Phys. Rev. A **99**, 022317 (2019).

[13] F. Piacentini, A. Avella, M. P. Levi, M. Gramegna, G. Brida, I. P. Degiovanni, E. Cohen, R. Lussana, F. Villa, A. Tosi, F. Zappa, and M. Genovese, Measuring Incompatible Observables by Exploiting Sequential Weak Values, Phys. Rev. Lett. **117**, 170402 (2016).

[14] L. Calderaro, G. Foletto, D. Dequal, P. Villoresi, and G. Vallone, Direct Reconstruction of the Quantum Density Matrix by Strong Measurements, Phys. Rev. Lett. **23**, 230501 (2018).

[15] H. Anwer, N. Wilson, R. Silva, S. Muhammad, A. Tavakoli, and M. Bourennane, arXiv:1904.09766.

[16] K. Mohan, A. Tavakoli, and N. Brunner, Sequential random access codes and self-testing of quantum measurement instruments, New J. Phys. **21**, 083034 (2019).

[17] N. Miklin, J. J. Borkała, and M. Pawłowski, arXiv:1903.12533.

[18] F. J. Curchod, M. Johansson, R. Augusiak, M. J. Hoban, P. Wittek, and A. Acín, Unbounded randomness certification using sequences of measurements, Phys. Rev. A **95**, 020102(R) (2017).

[19] B. Coyle, M. J. Hoban, and E. Kashefi, One-sided device-independent certification of unbounded random numbers, EPTCS **273**, 14 (2018).

[20] A. Tavakoli and A. Cabello, Quantum predictions for an unmeasured system cannot be simulated with a finite-memory classical system, Phys. Rev. A **97**, 032131 (2018).

[21] Y. S. Kim, J. C. Lee, O. Kwon, and Y. H. Kim, Protecting entanglement from decoherence using weak measurement and quantum measurement reversal, Nat. Phys. **8**, 117 (2012).

[22] B. L. Higgins, M. S. Palsson, G. Y. Xiang, H. M. Wiseman, and G. J. Pryde, Using weak values to experimentally determine "negative probabilities" in a two-photon state with Bell correlations, Phys. Rev. A **91**, 012113 (2015).

[23] T. C. White *et al.*, Preserving entanglement during weak measurement demonstrated with a violation of the Bell-Leggett-Garg inequality, npj Quantum Inf. **2**, 15022 (2016).

[24] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, Phys. Rev. Lett. **23**, 880 (1969).

[25] M. Schiavon, L. Calderaro, M. Pittaluga, G. Vallone, and P. Villoresi, Three-observer Bell inequality violation on a two-qubit entangled state, Quantum Sci. Technol. **2**, 015010 (2017).

[26] M.-J. Hu, Z.-Y. Zhou, X.-M. Hu, C.-F. Li, G.-C. Guo, and Y.-S. Zhang, Observation of non-locality sharing among three observers with one entangled pair via optimal weak measurement, npj Quantum Inf. **4**, 63 (2018).

[27] S. Pironio *et al.*, Random numbers certified by Bell's theorem, Nature **464**, 1021 (2010).

[28] A. Acín, S. Massar, and S. Pironio, Randomness versus Nonlocality and Entanglement, Phys. Rev. Lett. **108**, 100402 (2012).

[29] U. Vazirani and T. Vidick, Certifiable Quantum Dice, Philos. Trans. R. Soc. A **370**, 3432 (2012).

[30] J.-D. Bancal, L. Sheridan, and V. Scarani, More randomness from the same data, New J. Phys. **16**, 033011 (2014).

[31] A. Acín, S. Pironio, T. Vértesi, and P. Wittek, Optimal randomness certification from one entangled bit, Phys. Rev. A **93**, 040102(R) (2016).

[32] M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2000).

[33] G. Tóth and O. Gühne, Detecting Genuine Multipartite Entanglement with Two Local Measurements, Phys. Rev. Lett. **94**, 060501 (2005).

[34] E. Amselem, M. Rådmark, M. Bourennane, and A. Cabello, State-Independent Quantum Contextuality with Single Photons, Phys. Rev. Lett. **103**, 160405 (2009).

[35] E. Amselem, L. E. Danielsen, A. J. López-Tarrida, J. R. Portillo, M. Bourennane, and A. Cabello, Experimental Fully Contextual Correlations, Phys. Rev. Lett. **108**, 200405 (2012).

[36] V. D'Ambrosio, I. Herbauts, E. Amselem, E. Nagali, M. Bourennane, F. Sciarrino, and A. Cabello, Experimental Implementation of a Kochen-Specker Set of Quantum Tests, Phys. Rev. X **3**, 011012 (2013).

[37] B. H. Liu, X. M. Hu, J. S. Chen, Y. F. Huang, Y. J. Han, C. F. Li, G. C. Guo, and A. Cabello, Nonlocality from Local Contextuality, Phys. Rev. Lett. **117**, 220402 (2016).

[38] X. Zhan, E. G. Cavalcanti, J. Li, Z. Bian, Y. Zhang, H. M. Wiseman, and P. Xue, Experimental generalized contextuality with single-photon qubits, Optica **4**, 966 (2017).

[39] A. Crespi, M. Bentivegna, I. Pitsios, D. Rusca, D. Poderini, G. Carvacho, V. D'Ambrosio, A. Cabello, F. Sciarrino, and R. Osellame, Single-photon quantum contextuality on a chip, ACS Photonics **4**, 2807 (2017).

[40] O. Gühne and G. Tóth, Entanglement detection, Phys. Rep. **474**, 1 (2009).

Does violation of a Bell inequality always imply quantum advantage in a communication complexity problem?

Armin Tavakoli[1], Marek Żukowski[2], and Časlav Brukner[3,4]

[1]Département de Physique Appliquée, Université de Genève, CH-1211 Genève, Switzerland
[2]International Centre for Theory of Quantum Technologies (ICTQT), University of Gdansk, 80-308 Gdansk, Poland
[3]Faculty of Physics, University of Vienna, Boltzmanngasse 5, A-1090 Vienna, Austria
[4]Institute of Quantum Optics and Quantum Information, Austrian Academy of Sciences, Boltzmanngasse 3, A-1090 Vienna, Austria

Quantum correlations which violate a Bell inequality are presumed to power better-than-classical protocols for solving communication complexity problems (CCPs). How general is this statement? We show that violations of correlation-type Bell inequalities allow advantages in CCPs, when communication protocols are tailored to emulate the Bell no-signaling constraint (by not communicating measurement settings). Abandonment of this restriction on classical models allows us to disprove the main result of, inter alia, [Brukner et. al., Phys Rev. Lett. 89, 197901 (2002)]; we show that quantum correlations obtained from these communication strategies assisted by a small quantum violation of the CGLMP Bell inequalities do not imply advantages in any CCP in the input/output scenario considered in the reference. More generally, we show that there exists quantum correlations, with nontrivial local marginal probabilities, which violate the I_{3322} Bell inequality, but do not enable a quantum advantange in any CCP, regardless of the communication strategy employed in the quantum protocol, for a scenario with a fixed number of inputs and outputs

1 Introduction

Entanglement in itself cannot be used for information transfer. However, when combined with classical communication, it becomes a paradigmatic resource for quantum information transfer. It can amplify the capacity of a channel [1, 2], most famously in superdense coding [3]. Also, it can be used as a resource for better-than-classical communication complexity [4, 5]. Such reductions of communication complexity have a range of applications (see e.g. [7–11]) and are central tools for understanding the power of entanglement as a resource, both in terms of the extent to which it can outperform classical approaches [12–14] and how it compares to other quantum resources [15–17].

Communication complexity problems (CCPs) are tasks in which separated parties collaborate to compute a function dependent on inputs distributed among them, while only being allowed a limited amount of communication[1]. In their simplest form, such tasks can be viewed as games in which two parties Alice and Bob hold random inputs X and Y respectively and collaborate so that one of them (say Bob) can compute a function $f(X,Y)$. Alice communicates a classical message $m(X)$ to Bob who outputs a guess $g(m,Y)$ for the value of f. If the guess is correct, the partnership earns a "point". Importantly, the communication is limited so that the alphabet of m is smaller than that of X, typically rendering perfect evaluations of f impossible. The CCP is to find for Alice and Bob a communication strategy maximising the score, i.e. the averaged (over the distribution of inputs) number of points.

By sharing entanglement, Alice and Bob can sometimes increase their score beyond what is classically achievable [5]. To this end, it is necessary that they exploit entanglement to distribute strong correlations that violate a Bell

[1]An alternative approach to CCP considers the minimal amount of communication required to compute a function with distributed inputs. However, in this work, our focus is scenarios in which tasks are performed with limited communication.

inequality. We illustrate this with an example [18]. Alice (Bob) has fully random inputs $X = (x_0, x) \in [2]^2$ ($Y = y \in [2]$), where $[s]$ denotes the set $\{0, \ldots, |s| - 1\}$. The CCP (game) is: Bob earns a point if he gives his guess g equal to $f = x_0 + xy \mod 2$, while Alice can send him only a binary (bit) message $m(x_0, x)$. The score in the CCP is written $\mathcal{S} = 1/8 \sum_{x_0,x,y} P(g = f|x_0, x, y)$. The optimal classical score, which is $\mathcal{S} = \frac{3}{4}$, is achieved with some deterministic encoding/decoding procedure. Due to the small number of inputs, one can easily consider all possible messaging and guessing strategies. One finds that there are several different strategies achieving the optimal classical score in the CCP. It can be shown, see [5] and e.g. [6], that among other ones, there is an optimal strategy which runs as follows: Alice sends $m(x_0, x) = a(x) + x_0 \mod 2$ and Bob guesses $g = m + b(y) \mod 2$, where $a(x)$ and $b(y)$ are binary-valued functions of the inputs x and y respectively. The winning condition $f = g$ now reduces to $a(x) + b(y) = xy$, which allows us to put the score in the form of the Clauser-Horne-Shimony-Holt (CHSH) [19] Bell inequality: $\mathcal{S} = 1/4 \sum_{x,y} P(a + b = xy|x, y) \leq 3/4$. Thus immediately the following entanglement-assisted strategy becomes relevant: Alice and Bob use their inputs (x, y) as settings in a quantum test of the CHSH inequality where $a, b \in [2]$ are their respective local outcomes. Having obtained her outcome in the CHSH test, Alice sends the message $m(x_0, x) = a + x_0 \mod 2$ to Bob. Notice that this message emulates the Bell no-signaling constraint in the sense that it does not allow Bob to read out the value of x, which was used as a setting in the CHSH test. Bob uses his outcome in the CHSH test to construct the guess $g = m + b \mod 2$. Since shared entanglement enables Alice and Bob to violate the CHSH inequality, this quantum strategy leads to a score of up to $\mathcal{S} = \frac{1}{2} + \frac{1}{2\sqrt{2}} \approx 0.854$, which is the largest possible violation of the CHSH inequality. Thus, the entanglement-assisted strategy holds an advantage over all possible classical strategies in the CCP.

There are many more results showing that every probability distribution that violates *specific* Bell inequalities has the ability of enhancing a CCP beyond classical protocols. Examples include the Mermin inequalities [5, 20], the Collins-Gisin-Linden-Massar-Popescu (CGLMP) inequalities [16, 21–23], the elegant Bell inequality [24], Bell inequalities for random access coding [15, 25–27], the biased CHSH inequalities [11, 28] and a large class of bipartite many-outcome Bell inequalities [29]. More generally, Ref. [30] showed that the violation of *every* multipartite correlation Bell inequality with binary outcomes implies beating the best possible classical score in a corresponding CCP constructed by generalising the above example based on the CHSH inequality. See the topical review [31] for further discussions. This fauna of results begs the question: does every nonlocal probability distribution (i.e. a probability distribution that violates a Bell inequality) lead to an advantage in a CCP? To show such advantages, one requires only an example of a CCP in which access to the nonlocal probability distribution is advantageous. However, proving that no such advantage is possible is significantly more challenging; one must rule out the possibility of an advantage in *every* possible CCP, i.e. no matter the number of inputs and outputs, the choice of score and the chosen classical communication strategy.

Whereas we do not provide a decisive answer to whether Bell nonlocality always implies advantages in CCPs, we show that there exists a natural input/output scenario in which Bell nonlocality does not enable a quantum advantage in any CCP. We first formalise classical and entanglement-assisted CCPs. Then, we show how to map multipartite d-outcome correlation Bell inequalities to corresponding CCPs. This method allows to e.g. reproduce the examples studied in the literature, listed in the previous paragraph. We prove that a violation of a Bell inequality, together with restricted communication strategies, which do not reveal the input the sender would use to define her measurement setting the Bell inequality test, implies beating an analogous classical protocol for the corresponding CCP.

This restriction on classical strategies is tacitly used in several previous works (see e.g. Refs. [16, 22, 23]) which enables a quantum advantage. Our more complete analysis of classical strategies no longer sustains the generality of e.g. the main result of Ref. [22], that every violation of the CGLMP inequality combined with the above mentioned communication strategies implies an advantage in some CCP for a fixed number of inputs and outputs. This leads us to consider

the classical simulation of entanglement-assisted CCPs. We consider a situation with fixed number of inputs and outputs and show that there exists a quantum nonlocal probability distribution that does not enable better-than-classical communication complexity, regardless of the communication strategy and the choice of score. Our results are in opposition to the common belief that Bell nonlocality always is useful for better-than-classical communication complexity.

2 Formal scheme of the communication complexity problems analysis

We mainly consider two-party protocols. These are formulated as *games*. Alice and Bob each receive random inputs, respectively $X \in [N_{\mathrm{A}}]$ and $Y \in [N_{\mathrm{B}}]$. Alice sends a message $m \in [M]$ (with $M < N_{\mathrm{A}}$) to Bob who outputs $g \in [G]$, which is awarded with $t^g_{X,Y}$ points. The tuple $(N_{\mathrm{A}}, N_{\mathrm{B}}, M, G)$ corresponds to a choice of *scenario*. The score of a specific CCP within the chosen scenario is written as

$$\mathcal{S}[p(g|X,Y)] = \sum_{g,X,Y} t^g_{X,Y} p(g|X,Y) \qquad (1)$$

where $p(g|X,Y)$ is the probability of Bob's output for local inputs X, Y. Notice that the scoring function can always absorb prior probabilities $p(X,Y)$.

In a classical picture, Alice encodes her message with a function $E : [N_{\mathrm{A}}] \to [M]$ and Bob constructs his guess with a function $D : [M] \times [N_{\mathrm{B}}] \to [G]$. The choice of (E, D) can be coordinated via a shared random variable λ, with some probability distribution $p(\lambda)$. Therefore, a classical model is of the form

$$p^{\mathrm{C}}(g|X,Y) = \sum_{\lambda} p(\lambda) p_\lambda(g|X,Y). \qquad (2)$$

where the deterministic distribution is $p_\lambda(g|X,Y) = \sum_m p(m|X,\lambda) p(g|m,Y,\lambda)$ with $p(m|X,\lambda) = \delta_{m,E_\lambda(X)}$ and $p(g|m,Y,\lambda) = \delta_{g,D_\lambda(m,Y)}$. Due to linearity in Eq. (1), the largest score is found with a deterministic communication strategy. We therefore define the optimal classical score in a CCP as

$$\mathcal{S}^{\mathrm{C}} = \max_{\lambda} \mathcal{S}[p_\lambda(g|X,Y)]. \qquad (3)$$

In contrast, if Alice and Bob share an entangled state ρ, they may use their inputs to select measurement settings with associated outcomes a and b respectively. The statistics reads $p(a,b|X,Y) = \mathrm{tr}\left[A^a_X \otimes B^b_Y \rho\right]$ where A^a_X and B^b_Y are measurement operators. Subsequently, Alice sends $m = E(a,X)$ for some function $E : [|a|] \times [N_{\mathrm{A}}] \to [M]$ and Bob guesses $g = D(m,b,y)$ for some function $D : [M] \times [|b|] \times [N_{\mathrm{B}}] \to [G]$, where $|a|$ and $|b|$ denote the cardinality of the respective output spaces. Here, we have assumed that the Bell inequality test is performed before Bob receives Alice's message (in line with spacelike separation). Moreover, although shared randomness could be absorbed into the shared entangled state, we treat it separately in order to emphasise that it is a classical resource. Therefore, a quantum model is of the form

$$p^{\mathrm{Q}}(g|X,Y) = \sum_{\lambda} p(\lambda) p^{\mathrm{Q}}_\lambda(g|X,Y), \qquad (4)$$

where

$$p^{\mathrm{Q}}_\lambda(g|X,Y) = \sum_{a,b,m} p(a,b|X,Y) p(m|a,X,\lambda) p(g|m,b,Y,\lambda).$$

3 All violations of correlation Bell inequalities power advantages in constrained CCPs

Consider a Bell scenario with N parties $O_1, \ldots, O_N$ who perform measurements labelled $x_1, \ldots, x_N$ with outcomes $o_1, \ldots, o_N \in [d]$ for some $d \geq 2$, and any Bell inequality of the form

$$\mathcal{B} = \sum_{\vec{x}} \sum_{r} t^r_{\vec{x}} P_{\vec{x}} \left(\sum_{i=1}^{N} o_i = f^r_{\vec{x}} \right) \overset{\mathrm{LHV}}{\leq} C, \qquad (5)$$

where $\vec{x} = (x_1, \ldots, x_N)$, C is the LHV bound, r is an additional integer-valued indexing parameter which allows for more general Bell inequalities[2], $f^r_{\vec{x}} \in [d]$ and $t^r_{\vec{x}}$ are real coefficients. The relation $\sum_i o_i = f^r_{\vec{x}}$ is evaluated modulo d. The Bell inequalities in Eq (5) are sometimes referred to as full correlation Bell inequalities. To map (5) to a CCP, let O_i (for $i = 1, \ldots, N-1$) have an

[2]For instance the CHSH inequality requires only one value of r but in order to write the CGLMP inequalities on the form (5) one requires several values of r.

input $X_i = (x_i, x_0^{(i)})$ where $x_0^{(i)} \in [d]$. Each of these parties may send a message $m_i \in [d]$ after which O_N, who has input $X_N = x_N$, produces a guess $g \in [d]$ and earns $t^r_{\vec{x}}/d^{N-1}$ points whenever $g = f^r_{\vec{x}} + \sum_{i=1}^{N-1} x_0^{(i)}$. The score is defined as

$$\mathcal{S} = \frac{1}{d^{N-1}} \sum_{\vec{x},\vec{x}_0} \sum_r t^r_{\vec{x}} P_{\vec{x}} \left(g = f^r_{\vec{x}} + \sum_{i=1}^{N-1} x_0^{(i)} \right), \tag{6}$$

where $\vec{x}_0 = (x_0^{(1)}, \ldots, x_0^{(N-1)})$. Notice that the coefficients $t^r_{\vec{x}}/d^{N-1}$ in the CCP do not depend on $\vec{x}_0$. To put the Bell inequality and the CCP on equal footing, let the N parties share an entangled state and use their inputs $\vec{x}$ to perform a measurement with outcome $o_i \in [d]$. Then, the parties O_i for $i = 1, \ldots, N-1$ send $m_i = o_i + x_0^{(i)}$ to O_N who outputs $g = o_N + \sum_i m_i$. Here, in analogy with the previous CHSH-inspired example, the addition of $x_0^{(i)}$ in the message ensures that O_N cannot learn the input x_i which was used as a setting in a Bell inequality test. Also, notice that the parties $O_1, \ldots, O_{N-1}$ only use part of their inputs for choosing a measurement setting. This is in analogy with previous litterature (e.g. Refs. [15, 22, 29, 30]). We then find $\mathcal{S} = \mathcal{B}$. In comparison, consider a classical situation that is restricted to the same type of communication strategies, i.e. "additive" messages on the form $m_i = o_i(x_i) + x_0^{(i)}$, where $o_i(x_i)$ is a function of x_i, that do not reveal the value of x_i. Naturally, this leads to $\mathcal{S} = \mathcal{B} \leq C$. Therefore, when restricting to additive communication strategies for both the quantum and classical situation, one finds that violation of Eq. (5) implies $\mathcal{S} > C$, i.e. a quantum advantage in the CCP. The above construction generalises results in Refs. [5, 11, 15, 16, 22, 23, 26, 27, 30]. Communication which does not allow one to reveal measurement settings (as above) is important in scenarios in which the task function should be calculated in a way which does not allow an eavesdropper to learn the inputs of a sender, or even in a more subtle situation in which a sender does not want the receiver to know her inputs.

For $d = 2$, the scenario reduces to that of Ref. [30], in which it was shown that messages of the form $m_i = o_i + x_0^{(i)}$ lead to the optimal classical score (3). However, the same does not have to be true for $d > 2$. We shall explicitly show that such is not the case using the specific example of Ref. [22]. Ref. [22] showed (via the above map) that every violation of the CGLMP inequality [21] implies an advantage in a corresponding CCP in which the communication is restricted to the additive communication strategies defined above. As we show next, this constraint effectively excludes the optimal classical strategy.

4 The CGLMP inequality and communication complexity

Let us consider the CCP of Ref. [22] obtained by choosing (5) as the CGLMP inequality. This inequality is a facet Bell inequality when Alice and Bob have two settings $x, y \in [2]$ and three possible outcomes $a, b \in [3]$;

$$\mathcal{B}_{\text{cglmp}} = \frac{1}{4} \sum_{x,y} \left[P_{xy}(a+b = f^1) - P_{xy}(a+b = f^2) \right] \overset{\text{LHV}}{\leq} \frac{1}{2}, \tag{7}$$

where $f^1 = -xy$ and $f^2 = -xy + (-1)^{x+y}$. Using two entangled qutrits, one can reach the maximal quantum violation $\mathcal{B}^{\text{Q}}_{\text{cglmp}} \approx 0.7287$ [32]. In the corresponding CCP, Alice (Bob) has random inputs $x_0 \in [3]$ and $x \in [2]$ ($y \in [2]$). Alice may send a ternary message $m \in [3]$ to Bob who outputs a guess $g \in [3]$. The score (6) is

$$\mathcal{S}_{\text{cglmp}} = \frac{1}{12} \sum_{x_0,x,y} \left[P_{xy}(g = x_0 + f^1) - P_{xy}(g = x_0 + f^2) \right]. \tag{8}$$

For an additive communication strategy, the violation of (7) is necessary and sufficient for outperforming the corresponding classical value $\mathcal{S}_{\text{cglmp}} = 1/2$. Ref. [22] restricted itself to such additive communication strategies (for both classical and quantum models) to prove that a quantum enhancement of classical protocols in this CCP is possible if and only if the CGLMP inequality is violated. Since additive communication strategies assisted by Bell nonlocal correlations allow for a strong link between CCPs and Bell inequality violations, it appears natural that such communication strategies would power quantum-over-classical advantages in CCPs. However, as we show next, this does not necessarily mean that the best classical strategies are of the additive type, i.e. it is not

always optimal to tailor a classical strategy to a Bell inequality.

We now relax the assumption of additive communication strategies ($m = x_0 + a \mod 3$) and compute the optimal classical score (3). There are 3^{12} deterministic encoding and guessing strategies. Interestingly, by separately considering all of them, we find that

$$\mathcal{S}^{\mathrm{C}}_{\mathrm{cglmp}} = \frac{2}{3}. \tag{9}$$

This can be saturated by Alice sending $m(x_0, x) = \delta_{x,0}\delta_{x_0,2} + 2\delta_{x,1}\delta_{x_0,1} \mod 3$ and Bob guessing $g(m, y) = 2\delta_{y,0}m + \delta_{y,1}(m+1) \mod 3$. Note that $m = 1, 2$ informs Bob of the value of x.

Thus, the broad class of communication strategies considered in Ref. [22] is insufficient to find the optimal classical score. A large violation of the CGLMP inequality, $\mathcal{B}_{\mathrm{cglmp}} > 2/3$, indeed does imply an advantage over general classical protocols for the CCP. However, weaker violations, $1/2 < \mathcal{B}_{\mathrm{cglmp}} \leq 2/3$, are insufficient to achieve the same feat. In Appendix. A we show that analogous criticism applies to the CCPs of Refs. [16, 23]. Also, we have numerically considered whether the limitation $\mathcal{S}_{\mathrm{cglmp}} \leq \mathcal{B}^{\mathrm{Q}}_{\mathrm{cglmp}}$ can be overcome by using a more general message (see Appendix. B for details). However, we have found no improvement over the strategy in which Alice and Bob maximally violate the CGLMP inequality and the message is additive.

Departing from the particular CCP (8), we remind ourselves that entanglement-assisted advantages originate from the probability distribution $p(g|x_0, x, y)$. Evidently $p(g|x_0, x, y)$ may lack a classical model even when the *specific* score (8) does not exceed the classical bound (9). Is it the case that for every probability distribution $p(g|x_0, x, y)$ obtained by a trit-communication and a violation of the CGLMP inequalities, there exists some other CCP in which a higher score is obtained than is classically possible? Presently, we answer this for the case of the additive communication strategy. Let Alice and Bob use their shared entanglement to generate a probability distribution of the form

$$p(a, b|x, y) = vp^{\mathrm{cglmp}}(a, b|x, y) + \frac{1-v}{9}, \tag{10}$$

where $p^{\mathrm{cglmp}}(a, b|x, y)$ maximally violates the CGLMP inequality and $v \in [0, 1]$ is the protocol visibility parameter. This violates the CGLMP inequality when $v > 0.6861$. The probability distribution $p^{\mathrm{Q}}_v(g|x_0, x, y)$, obtained via shared entanglement and an additive communication strategy, beats the classical bound (9) when $v > 0.9149$. We seek the largest v for which p^{Q}_v can be simulated by a classical model. This means solving the linear program

$$\begin{aligned} \max_{p(\lambda)} v \quad & \text{s.t.} \quad p(\lambda) \geq 0, \quad \textstyle\sum_\lambda p(\lambda) = 1, \\ \text{and} \quad & p^{\mathrm{Q}}_v(g|x_0, x, y) = \textstyle\sum_\lambda p(\lambda) p_\lambda(g|x_0, x, y). \end{aligned} \tag{11}$$

By considering $p_\lambda(g|x_0, x, y)$ for all possible deterministic strategies, we have found that the corresponding polytope of classical probability distributions has 47601 vertices. We have evaluated the linear program and found $v \approx 0.7943$. Hence, probability distributions $p^{\mathrm{Q}}_v(g|x_0, x, y)$ for $0.7943 < v \leq 0.9149$ indeed imply an advantage over classical protocols in some CCP despite the particular CCP (8) failing to detect it. However, when $0.6861 < v \leq 0.7942$ the CGLMP inequality is violated, but the probability distribution $p^{\mathrm{Q}}_v(g|x_0, x, y)$ can be classically modelled.

5 Bell nonlocality without CCP advantages in fixed scenarios

The above classical simulation focuses on entanglement-assisted correlations obtained via an additive communication strategy. Here, we prove a more general statement: that for a given scenario (that is, a fixed number of inputs and outputs) there exists a nonlocal probability distribution which cannot be used to improve any CCP beyond classical constraints, regardless of the choice of communication strategy. Specifically, we find a nonlocal probability distribution that when combined with *any* communication strategy gives rise to a $p(g|X, Y)$ which can be simulated in a classical model for the given scenario.

To this end, we focus on the a simple Bell scenario going beyond that of the CHSH inequality. In order to work with the smallest number of outcomes possible ($a, b \in [2]$), we must consider two parties with ternary settings settings $x, y \in [3]$. Three settings are needed, as the two-setting scenario is fully characterised by the CHSH inequality, which is a correlation inequality and therefore implies advantages in a CCP whenever violated [18, 30]. Alternatively, one could also consider

the previously discussed Bell scenario with two settings and three outcomes. However, we focus on the former due to its conceptual simplicity and (as it turns out) computational advantages. In the three-setting scenario, the facet Bell inequalities are the (lifted) CHSH inequality and the I_{3322} inequality [35, 36]. The I_{3322} inequality reads

$$I = -P_{\mathrm{A}}(0) - 2P_{\mathrm{B}}(0) - P_{\mathrm{B}}(1) + \sum_{x,y} T_{x,y} P(x,y) \leq 0, \tag{12}$$

where $P(x,y)$ is the probability of outputting $a = b = 0$, $P_{\mathrm{A}}(x) = p(a=0|x)$, $P_{\mathrm{B}}(y) = p(b=0|y)$ and $T = \{[1,1,1],[1,1,-1],[1,-1,0]\}$. Notably, this inequality is not a correlation Bell inequality and is therefore not in the broad class of Bell inequalities whose violation necessarily implies advantages in CCPs [30].

Motivated by the choice of scenario in previous discussions for translating Bell inequality violations to advantages in CCPs, we consider a communication scenario in which Alice has inputs $x_0 \in [2]$ and $x \in [3]$ and Bob has an input $y \in [3]$. Alice sends $m \in [2]$ to Bob who outputs $g \in [2]$. To further motivate that this scenario is a good choice for revealing the CCP advantages of probability distributions that violate the I_{3322} inequality, we have shown in Appendix. C that a maximal violation of (12) (for qubits) implies better-than-classical communication complexity, and also that *every* probability distribution violating (12) obtained from mixing the optimal one with a uniform probability distribution (in analogy with Eq. (10)) also implies such an advantage. Note that such a scenario is a natural extension of the ones studied in Ref. [30].

Nevertheless, we show that there exists a non-maximally entangled state and local measurements that give rise to a probability distribution that violates I_{3322} inequality, that however is not advantageous in any CCP in the stated scenario. To this end, Alice and Bob can generate a Bell-like distribution of the form $p(a,b|x_0,x,y)$. Notice that only three of Alice's six inputs are required to create Bell nonlocal correlations that violate the I_{3322} inequality. We can without loss of generality label these three inputs by x. Thus, with these labels, the dependence on x_0 in the Bell nonlocal distribution becomes trivial, i.e. $p(a,b|x_0,x,y) = p(a,b|x,y)$. Now, we can choose our *candidate probability distribution* $p^{\mathrm{cand}}(a,b|x,y)$. This distribution has a quantum realisation. It also weakly violates Eq. (12) ($I \approx 0.0129$), but importantly does not violate the CHSH inequality and hence cannot lead a better-than-classical score in a CCP based on the CHSH inequality. The candidate probability distribution was originally proposed in Ref. [36] and we detail it and its quantum realisation in Appendix. D. We show that for every possible communication strategy within the scenario, there exists no CCP in which p^{cand} enables an advantage over classical protocols. We first note that since we have fixed the probability distribution in the Bell scenario to p^{cand}, the set of distributions (4) that Alice and Bob can generate in the communication scenario forms a polytope. Therefore, it suffices to show that all deterministic communication strategies with access to p^{cand} can be classically modelled. Since Alice maps the twelve possible values of (a,x_0,x) to her binary message m, and Bob maps the twelve values (m,b,y) to his binary output g, there exists a total of 2^{24} deterministic communication strategies. For each of these (indexed by μ), we have evaluated the corresponding probability distribution $p_\mu(g|x_0,x,y) = \sum_{a,b,m} p(a,b|x,y) p_\mu(m|a,x_0,x) p_\mu(g|m,b,y)$. We have found that the relevant polytope of probability distributions in the communication scenario has 8192992 vertices. To show that the probability distribution $p_\mu(g|x_0,x,y)$ can be simulated by a classical model for all vertices, we consider the mixture of each vertex probability distribution with random outcomes; $p_\mu^{\mathrm{Q},v}(g|x_0,x,y) = v p_\mu(g|x_0,x,y) + (1-v)/2$. Then, for each of the roughly eight million values of μ, we decide the possibility of a classical model by running a linear program algorithm[3] analogous to Eq. (11). We find that for every choice of μ, the value of v is never smaller than one (up to machine precision). That is, every $p_\mu(g|x_0,x,y)$ can be classically modelled. Thus, we conclude that in the scenario in which Alice has $X \in [6]$ and Bob has $Y \in [3]$ and m and g are bit valued, there exists no CCP that can be improved beyond classical

[3] Since each run of the linear program takes a few seconds to complete, it would require months to complete all 8192992 cases on a standard computer. To contend with this problem, we have distributed the computation; roughly 40% of it to the high-performance computing cluster Baobab at the University of Geneva, another 20% to two workstation computers, and the remaining 40% to five standard desktop computers. This allowed us to complete the full computation in less than three weeks.

constraints by the parties sharing the nonlocal probability distribution $p^{\text{cand}}(a, b|x, y)$.

6 Conclusions

A substantial number of examples of quantum advantages in CCPs being powered by Bell inequality violations can be understood as different instances of a single map from Bell inequalities to CCPs. We found that a violation of the former implies an advantage in the latter for a simple class of communication strategies. As we explicitly showed, a complete analysis of classical communication complexity requires the revision of several previous claims in which violations of particular Bell inequalities where thought to imply advantages in CCPs. Going beyond that, we found that there exists nonlocal distributions for which the statistics of every possible communication strategy in any possible CCP can be simulated by classical models in an input/output scenario that naturally extends previous works. This suggests that not all forms of Bell nonlocality are useful for better-than-classical communication complexity. A definite proof of this statement would require an extension of our results to CCPs with any number of inputs and outputs. Our results motivate a characteriation of the (now seemingly nontrivial) relation between Bell nonlocality and entanglement-assisted communication complexity. Which nonlocal probability distributions are useful for outperforming classical limitations in CCPs and which are not?

Acknowledgments

We thank Fabian Bernards, Nicolas Brunner and Nicolas Gisin for discussions. We also thank Dmitry Tabakaev, Marc-Olivier Renou, Cai Yu, Sebastien Designolle and Davide Rusca for lending us their computers. Finally, we thank Emmanuel Zambrini for much appreciated help with the Baobab computer system. AT acknowledges support from the Swiss National Science Foundation (Starting grant DIAQ, NCCR-QSIT). MZ acknowledges the ICTQT IRAP project of FNP, financed by structural funds of EU, and the COPERNICUS FNP/DFG grant-award. CB acknowledges the support of the Austrian Science Fund (FWF) through the SFB project "BeyondC" and the project I-2526-N27.

A The optimal classical score in the communication complexity problems of Refs. [16, 23]

Ref. [22] introduced CCPs based on the ternary-outcome CGLMP inequality. These were extended in Refs. [16, 23] to CGLMP inequalities with any number (d) of outcomes. In a spirit similar to that of Ref. [22], these subsequent works restrict themselves to considering classical communication strategies of the additive type. Here, we show that also for $d > 3$ this fails to capture the optimal classical score of the CCP. For simplicity, we focus on the case of four outcomes.

Refs. [16, 23] present the following CCPs (up to minor modifications). Alice has eight possible inputs written in terms of a bit $x \in [2]$ and a quart $x_0 \in [4]$. Bob has two possible inputs $y \in [2]$. Alice may communicate at most a four-valued message to Bob who aims to maximise the score

$$\mathcal{S}_{\text{cglmp}} = \frac{1}{16} \sum_{x_0,x,y} P_{x,y}(g = x_0+f_1^1) - P_{x,y}(g = x_0+f_2^1) + \frac{1}{3} \left(P_{x,y}(g = x_0 + f_1^2) - P_{x,y}(g = x_0 + f_2^2) \right), \tag{13}$$

where

$$f_1^1 = x_0 - xy \qquad f_2^1 = x_0 - xy + (-1)^{x+y}$$
$$f_1^2 = x_0 - xy - (-1)^{x+y} \qquad f_2^2 = x_0 - xy + 2(-1)^{x+y},$$

computed modulo four. Alice (Bob) uses x (y) to measure an entangled pair with possible outcomes $a \in [4]$ ($b \in [4]$). Then, using an additive communication strategy, i.e. $m = x_0 + a \mod 4$ and $g = m + b \mod 4$, one finds that $\mathcal{S}_{\text{cglmp}}$ becomes identical to the Bell expression in the four-outcome CGLMP inequality [16, 23] which has an LHV bound (in this form and normalisation) of $1/2$. Therefore, under additive communication strategies Refs. [16, 23] found $\mathcal{S}_{\text{cglmp}} \leq 1/2$.

However, the optimal classical score is not saturated with such a communication strategy. Since Alice maps eight inputs to four outputs, and similarly for Bob, there is a total of 4^{16} pairs of encoding and guessing functions. We have evaluated the score for all such pairs and found that the optimal classical score is

$$\mathcal{S}^{\text{C}}_{\text{cglmp}} = \frac{2}{3}. \tag{14}$$

An encoding/decoding strategy that saturates this bound is

$$m = [0, 0, 0, 1, 0, 2, 3, 0] \tag{15}$$
$$g = [0, 3, 1, 2, 2, 3, 0, 1], \tag{16}$$

where the tuple represents the response to the pair (x, x_0) and (y, m) respectively (ordered as $(0,0), (0,1), \ldots, (1,2), (1,3)$). Thus, in full analogy with the discussion in the main text focused on ternary-outcome CGLMP inequalities, corrections also apply to its generalisation to more than three outcomes.

B Numerical search for the optimal quantum score in the CCP based on the CGLMP inequality

We present numerical methods which we used in support of the conjecture that an additive communication strategy and a quantum probability distribution that maximally violates the CGLMP inequality give the optimal entanglement-assisted score in the CCP based on the ternary-outcome CGLMP inequality (8).

The joint state of Bob's local system (after Alice's measurement) and the classical message when averaged over Alice's outcome can be written

$$\rho_{x_0 x} = \sum_a \mathrm{tr}_\mathrm{A} \left[A_x^a \otimes \mathbb{1}\rho\right] \otimes |m\rangle\langle m|, \tag{17}$$

where we encode the classical message in the computational basis state $|m\rangle\langle m|$. For all deterministic messages of the restricted class $m = m(a, x_0)$, we have evaluated the score and found that only those of the additive type lead to a better-than-classical score. However, for a general deterministic message $m = m(a, x_0, x)$ such a brute-force approach is too time-consuming. To address the general case, we can obtain upper bounds on the score by substituting $|m\rangle\langle m|$ in Eq. (17) with a quantum system $\sigma_{a,x_0,x} \in \mathbb{C}^3$. Notice that this only serves as a tool towards treating the relevant problem in which the message is classical. Moreover, this substitution is far more constraining than allowing for general quantum communication assisted by shared entanglement. The substitution of the classical message for a quantum one comes with the advantage that one can efficiently run alternating convex searches for lower bounding the quantity $\mathcal{S}^{\mathrm{QC}}_{\mathrm{cglmp}} = \max_{\rho,A,B,\sigma} \mathcal{S}_{\mathrm{cglmp}}$. The searches are alternating in the sense that we first considers a semidefinite program optimising over the shared state, then a second one optimising over Alice's measurements, then a third one optimising over Bob's measurements and finally a fourth one optimising over the quantum message. This procedure of four semidefinite programs is iterated until the results appear to converge. For three respectively four dimensional entangled systems, we implemented the procedure by alternating semidefinite programs each optimising over the state, Alice's measurements, Bob's measurements and the quantum message respectively. We have implemented this procedure for 10000 randomly chosen starting points. Each of these 10000 trials involves ten iterations of the described procedure (that is, 40 evaluations of a semidefinite program). In all 10000 cases we find that the optimisation converges to the value $\mathcal{B}^{\mathrm{Q}}_{\mathrm{cglmp}}$, which is what is obtained by maximally violating the CGLMP inequality and then using an additive communication strategy.

C Communication complexity advantages via violation of the I_{3322} inequality

Ref. [36] found that the maximal quantum violation of the I_{3322} inequality with a shared entangled pair of qubits is $I^{\mathrm{Q}} = 1/4$ and is achieved with the singlet state $|\psi^-\rangle = (|0,1\rangle - |1,0\rangle)/\sqrt{2}$ and measurements in the xz-plane of the Bloch sphere whose Bloch vectors (including the antipodal vectors) form a hexagon on both Alice's and Bob's side. Specifically, the Bloch vectors read

$$\begin{aligned} \vec{a}_1 &= [0, 0, 1] & \vec{b}_1 &= -[\sqrt{3}, 0, 1]/2 \\ \vec{a}_2 &= [\sqrt{3}, 0, 1]/2 & \vec{b}_2 &= -[0, 0, 1] \\ \vec{a}_3 &= [\sqrt{3}, 0, -1]/2 & \vec{b}_3 &= [\sqrt{3}, 0, -1]/2. \end{aligned} \tag{18}$$

Thus the resulting probability distribution is

$$p^{3322}(a, b|x, y) = \frac{1}{4}\left[1 - (-1)^{a+b}\vec{a}_x \cdot \vec{b}_y\right]. \tag{19}$$

We have considered the mixture of this probability distribution with a uniformly random probability distribution, i.e.

$$p^v(a, b|x, y) = v p^{3322}(a, b|x, y) + \frac{1 - v}{4}. \tag{20}$$

This probability distribution violates the I_{3322} inequality only when $v > 4/5$. We consider its usefulness in CCPs in a scenario (same as in the main text) in which Alice has six inputs, Bob has three inputs, and Alice's and Bob's outputs both are binary. We choose an additive communication strategy in which Alice sends $m = a + x_0 \mod 2$ and Bob outputs $g = m + b \mod 2$. This leads to a specific probability distribution in the CCP (dependent on v). We then run a linear program of the type presented in the main text to determine the largest v for which the entanglement-assisted probability distribution in the CCP has a classical model. We find that it returns $v = 4/5$, thus showing that every probability distribution of the form (20) that violates the I_{3322} inequality implies advantages in a CCP.

D The candidate probability distribution

There exists probability distributions in the Bell scenario with three setting and two outcomes that violate the I_{3322} inequality but not the CHSH inequality. Ref. [36] provided an example, which we in the main text used as the candidate probability distribution $p^{\text{cand}}(a, b|x, y)$. It is obtained by Alice and Bob sharing the noisy state

$$\rho = \frac{17}{20}|\phi\rangle\langle\phi| + \frac{3}{20}|0,1\rangle\langle 0,1| \tag{21}$$

where $|\phi\rangle = (2|0,0\rangle + |1,1\rangle)\sqrt{5}$. This state cannot violate the CHSH inequality for any choice of measurements (which can be checked via the Horodecki criterion [37]), but it can violate the I_{3322} inequality as follows. Write Alice's (Bob's) Bloch vectors in the xz-plane as $\vec{a}_x = [\sin\theta_x, \cos\theta_x]$ ($\vec{b}_y = [\sin\phi_y, \cos\phi_y]$) with

$$\theta_1 = \eta \qquad \theta_2 = -\eta \qquad \theta_3 = -\frac{\pi}{2}$$
$$\phi_1 = -\chi \qquad \phi_2 = \chi \qquad \phi_3 = \frac{\pi}{2}$$

with $\eta = \arccos\left(\sqrt{7/8}\right)$ and $\chi = \arccos\left(\sqrt{2/3}\right)$. This defines the candidate probability distribution

$$p^{\text{cand}}(a, b|x, y) = \frac{1}{4}\,\text{tr}\left[(\mathbb{1} + (-1)^a \vec{a}_x \cdot \vec{\sigma}) \otimes (\mathbb{1} + (-1)^b \vec{b}_y \cdot \vec{\sigma})\rho\right] \tag{22}$$

which achieves $I \approx 0.0129$.

References

[1] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem, IEEE Trans. Inf. Theory **48**, 2637 (2002).

[2] Q. Zhuang, E. Y. Zhu, and P. W. Shor, Additive Classical Capacity of Quantum Channels Assisted by Noisy Entanglement, Phys. Rev. Lett. **118**, 200503 (2017).

[3] C. H. Bennett, and S. J. Wiesner Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states, Phys. Rev. Lett. **69**, 2881 (1992).

[4] G. Brassard, Quantum Communication Complexity, Foundations of Physics **33**, 11 (2003).

[5] R. Cleve and H. Buhrman, Substituting Quantum Entanglement for Communication, Phys. Rev. A **56**, 1201 (1997).

[6] P. Trojek, C. Schmid, M. Bourennane, C. Brukner, M Żukowski, H. Weinfurter, Experimental quantum communication complexity, Phys. Rev. A **72**, 050305 (2005).

[7] C. H. Bennett, D. P. DiVincenzo, P. W. Shor, J. A. Smolin, B. M. Terhal, and W. K. Wootters, Remote state preparation, Phys. Rev. Lett. **87**, 077902 (2001).

[8] D. Gavinsky, J. Kempe, O. Regev, and R. de Wolf, Bounded-error quantum state identification and exponential separations in communication complexity, In Proceedings of 38th ACM STOC, 594 (2006).

[9] S. Laplante, M. Laurière, A. Nolin, J. Roland, and G. Senno, Robust Bell inequalities from communication complexity, Quantum **2**, 72 (2018).

[10] A. Tavakoli, A. A. Abbott, M-O. Renou, N. Gisin, and N. Brunner, Semi-device-independent characterization of multipartite entanglement of states and measurements, Phys. Rev. A **98**, 052333 (2018).

[11] S. Muhammad, A. Tavakoli, M. Kurant, M. Pawłowski, M. Żukowski, and M. Bourennane, Quantum Bidding in Bridge, Phys. Rev. X **4**, 021047 (2014).

[12] H. Buhrman, R. Cleve and A. Wigderson, Quantum vs. classical communication and

computation, Proceedings of the 30th Annual ACM Symposium on Theory of Computin, 63 (1998).

[13] R. Raz, Exponential separation of quantum and classical communication complexity, In Proceedings of 31st ACM STOC, 358 (1999).

[14] G. Brassard, R. Cleve and A. Tapp, Cost of exactly simulating quantum entanglement with classical communication, Phys. Rev. Lett. **83**, 1874 (1999).

[15] M. Pawłowski and M. Żukowski, Entanglement assisted random access codes, Phys. Rev. A **81**, 042326 (2010).

[16] A. Tavakoli, M. Pawłowski, M. Żukowski, and M. Bourennane, Dimensional discontinuity in quantum communication complexity at dimension seven, Phys. Rev. A **95**, 020302(R) (2017).

[17] H. Buhrman, Ł. Czekaj, A. Grudka, M. Horodecki, P. Horodecki, M. Markiewicz, F. Speelman, and S. Strelchuk, Quantum communication complexity advantage implies violation of a Bell inequality, PNAS **113**, 3191 (2016).

[18] H. Buhrman, R. Cleve and W. van Dam, Quantum entanglement and communication complexity, SIAM J. Comput. **30**, 1829 (2001).

[19] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, Phys. Rev. Lett. **23**, 880 (1969).

[20] N. D. Mermin, Extreme quantum entanglement in a superposition of macroscopically distinct states, Phys. Rev. Lett. **65**, 1838 (1990).

[21] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu, Bell Inequalities for Arbitrarily High-Dimensional Systems, Phys. Rev. Lett. **88**, 040404 (2002).

[22] C. Brukner, M. Żukowski, and A. Zeilinger, Quantum Communication Complexity Protocol with Two Entangled Qutrits, Phys. Rev. Lett. **89**, 197901 (2002).

[23] C. Brukner, T. Paterek, and M. Żukowski, Quantum communication complexity protocols based on higher-dimensional entangled systems, Int J of Quant Inf. **1**, 4 (2003).

[24] N. Gisin, Bell inequalities: Many questions, a few answers, in *Quantum Reality, Relativistic Causality, and Closing the Epistemic Circle: Essays in Honour of Abner Shimony.* The Western Ontario Series in Philosophy of Science (Springer, Berlin, 2009), Vol. 73, p. 125.

[25] J. Oppenheim and S. Wehner, The Uncertainty Principle Determines the Nonlocality of Quantum Mechanics Science **19**, 330 (2010).

[26] A. Tavakoli, B. Marques, M. Pawłowski, and M. Bourennane, Spatial versus sequential correlations for random access coding, Phys. Rev. A **93**, 032336 (2016).

[27] A. Hameedi, D. Saha, P. Mironowicz, M. Pawłowski, and M. Bourennane, Complementarity between entanglement-assisted and quantum distributed random access code, Phys. Rev. A **95**, 052345 (2017).

[28] T. Lawson, N. Linden, and S. Popescu, Biased nonlocal quantum games, arXiv:1011.6245

[29] A. Tavakoli and M. Żukowski, Higher-dimensional communication complexity problems: Classical protocols versus quantum ones based on Bell's theorem or prepare-transmit-measure schemes, Phys. Rev. A **95**, 042305 (2017).

[30] C. Brukner, M. Żukowski, J-W. Pan, and A. Zeilinger, Bell's Inequalities and Quantum Communication Complexity, Phys. Rev. Lett. **92**, 127901 (2004).

[31] H. Buhrman, R. Cleve, S. Massar, and R. de Wolf, Nonlocality and communication complexity, Rev. Mod. Phys. **82**, 665 (2010).

[32] M. Navascués, S. Pironio, and A. Acín, Bounding the Set of Quantum Correlations, Phys. Rev. Lett. **98**, 010401 (2007).

[33] D. Martínez, A. Tavakoli, M. Casanova, G. Cañas, B. Marques, and G. Lima, High-Dimensional Quantum Communication Complexity beyond Strategies Based on Bell's Theorem, Phys. Rev. Lett. **121**, 150504 (2018).

[34] L. Vandenberghe and S. Boyd, SIAM Review **38**, 49 (1996).

[35] M. Froissart, Constructive generalization of Bells inequalities, Il Nuovo Cimento B **64**, 241 (1981).

[36] D. Collins, and N. Gisin, A Relevant Two Qubit Bell Inequality Inequivalent to the CHSH Inequality, J. Phys. A: Math. Gen. **37** 1775 (2004).

[37] R. Horodecki, P. Horodecki and M. Horodecki, Violating Bell inequality by mixed states: necessary and sufficient condition, Phys. Lett. A **200**, 340 (1995).

Informationally restricted quantum correlations

Armin Tavakoli[1], Emmanuel Zambrini Cruzeiro[1], Jonatan Bohr Brask[2], Nicolas Gisin[1], and Nicolas Brunner[1]

[1]Département de Physique Appliquée, Université de Genève, CH-1211 Genève, Switzerland
[2]Department of Physics, Technical University of Denmark, Fysikvej, 2800 Kongens Lyngby, Denmark

Quantum communication leads to strong correlations, that can outperform classical ones. Complementary to previous works in this area, we investigate correlations in prepare-and-measure scenarios assuming a bound on the information content of the quantum communication, rather than on its Hilbert-space dimension. Specifically, we explore the extent of classical and quantum correlations given an upper bound on the one-shot accessible information. We provide a characterisation of the set of classical correlations and show that quantum correlations are stronger than classical ones. We also show that limiting information rather than dimension leads to stronger quantum correlations. Moreover, we present device-independent tests for placing lower bounds on the information given observed correlations. Finally, we show that quantum communication carrying $\log d$ bits of information is at least as strong a resource as d-dimensional classical communication assisted by pre-shared entanglement.

1 Introduction

Separated parties, initially independent, can become correlated via communication. Intuitively, more communication enables stronger correlations. Also, the strength of the correlations may vary depending on the nature of the communication; for example if the message is carried by a quantum system rather than a classical one. In general, understanding the relation between communication and correlations is a fundamental question, at the intersection of information theory and physics.

Consider a simple scenario (see Fig. 1) with two separated parties. A first party, Alice, receives an input x and sends a message to a second party, Bob. Upon receiving this message, as well as some input y, Bob produces an output b. When repeated many times (with inputs x and y randomly sampled), this experiment is described by the conditional probability distribution $p(b|x,y)$ which characterises the correlations between Alice and Bob. Clearly, the amount of information about x encoded in Alice's message determines the strength of the possible correlations. If Alice sends no message at all (or if the message is independent of x), then no correlations are generated, i.e. $p(b|x,y) = p(b|y)$. On the other hand, if the message perfectly encodes x, then maximal correlations can be established; any distribution $p(b|x,y)$ is possible. Thus the main question is: how strong correlations can be established provided that the amount of communication from Alice to Bob is quantitatively limited?

Naturally, the answer depends on how exactly communication is quantified. The most common approach consists in measuring communication via the dimension of the message, i.e. the number of bits the message could carry. This is used in the field of communication complexity (see e.g. [1]), where the goal is to find out how the minimum dimension required to solve a problem (i.e. demanding that the output b corresponds to a certain function of the inputs x and y) scales with the problem size. Notably, the use of quantum communication is advantageous since it allows one to solve certain problems with exponentially smaller dimension [2, 3]. In parallel, there has been interest in characterising the set of possible correlations $p(b|x,y)$ for classical and quantum systems of bounded dimension [4–7]. Again, quantum correlations turn out to be stronger than classical ones. This led to a novel framework for quantum information processing termed "semi-device-independent" [8–12], where devices are assumed to process quantum

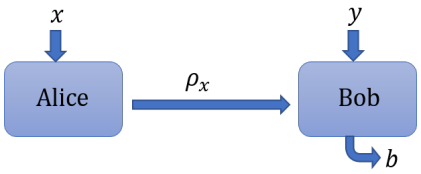

Figure 1: Prepare-and-measure scenario. In this work we investigate the strength of possible correlations $p(b|x, y)$ given a limit on the information carried by the quantum message ρ_x.

systems of bounded dimension, but are otherwise uncharacterised.

However, measuring communication via the dimension provides only a partial characterisation. Information-theoretic concepts are typically better suited to get a complete picture. This raises a natural question, namely to understand the relation between the strength of correlations and the amount of information that the communication contains. But then, information about what? In correlation experiments, the answer is very natural: we are interested in the information that the message contains about Alice's input x.

Here we formalise this problem and investigate classical and quantum correlations for informationally restricted communication. Naturally, however, there are many different ways of quantifying information based on entropies. We quantify the information content of an ensemble of prepared states (classical or quantum) via a one-shot version of accessible information based on min-entropies [13]. This choice of information measure has a two-fold motivation. Firstly, it admits a simple operational interpretation in terms of how well one could determine Alice's input from her message, via the best possible measurement. Secondly, it proves convenient to work with. Our approach is clearly complementary to previous works based on dimension. Firstly, information is a continuous quantity, while dimension is discrete; one can consider ensembles of states carrying only half a bit of information about Alice's input, which would have no analogue using dimension. Secondly, even when considering ensembles of states carrying $\log d$ bits of information (for some dimension d), there exist ensembles of dimension $d' > d$ that carry no more than $\log d$ bits of information, e.g. certain ensembles of non-orthogonal quantum states.

In this work, we develop a framework for characterising informationally restricted correlations. For the case of classical systems, we show that the relevant set of correlations forms a convex polytope, which can be fully characterised. This allows one to find the minimal amount of information required to reproduce a given correlation using classical communication. In turn, we prove that quantum correlations can be stronger than classical ones. Moreover, we derive device-independent lower bounds on the information, given observed correlations. These ideas are illustrated in a simple scenario. We also show that ensembles of higher-dimensional quantum states carrying no more than one bit of information can generate stronger correlations than two-dimensional quantum systems (i.e. qubits). Finally, we show that any correlations achievable with classical communication (of a d-dimensional message) assisted by pre-shared entanglement can also be achieved using quantum communication carrying $\log d$ bits of information.

2 Setting

We start by defining informationally restricted correlations in a quantum prepare-and-measure scenario. The sender, Alice, receives an input $x \in [n]$ sampled from a random variable X (where $[s] = \{1, \ldots, s\}$) which she encodes into a quantum state ρ_x that she relays to the receiver, Bob. Bob also receives a random input $y \in [l]$ and then measures the received state with some generalised measurement (positive operator-valued measure, POVM) $\{M_{b|y}\}$ with outcome $b \in [k]$. The observed correlations are

$$p(b|x, y) = \mathrm{tr}\left(\rho_x M_{b|y}\right). \tag{1}$$

Let us now characterise the information in Alice's message about her input x. Since x is random, sampled from some distribution $p_X(x)$, the ensemble of messages is given by $\mathcal{E} = \{p_X(x), \rho_x\}$. How well could an observer, via any possible POVM $\{N_z\}$, guess x from $\mathcal{E}$? The *guessing probability* is

$$P_g(X|\mathcal{E}) = \max_{\{N_z\}} \sum_{x=1}^{n} p_X(x)\, \mathrm{tr}\left[\rho_x N_x\right]. \tag{2}$$

Note that the optimal POVM, $\{N_z^*\}$, does not need to be part of set of POVMs $\{M_{b|y}\}$. Hence

the statistics obtained from measuring $\{N_z^*\}$ do not necessarily appear in the correlations $p(b|x,y)$.

The observer's minimal uncertainty about X when provided $\mathcal{E}$, i.e. the conditional min-entropy, is $H_{\min}(X|\mathcal{E}) = -\log[P_g(X|\mathcal{E})]$. The amount of information gained by learning $\mathcal{E}$, i.e. the information carried by $\mathcal{E}$, is then the difference in uncertainty without and with the communication [13];

$$\mathcal{I}_X(\mathcal{E}) = H_{\min}(X) - H_{\min}(X|\mathcal{E}), \tag{3}$$

where $H_{\min}(X) = -\log[\max_x p_X(x)]$ is the min-entropy. The quantity $\mathcal{I}_X(\mathcal{E})$ can be viewed as a single-shot version of accessible information [14, 15]. Note that for any given ensemble $\mathcal{E}$, the guessing probability (and hence the information) can be computed via a semidefinite program [16].

We can now define the set of possible correlations $p(b|x,y)$ when the information of the message is upper bounded. Importantly, we do not limit the Hilbert-space dimension for representing the set of the quantum states $\{\rho_x\}$. We also allow for shared randomness between Alice's and Bob's devices. This makes the model more general, and at the same time simplifies the characterisation of the sets of correlations (as these sets are now convex). Formally, we define the set $\mathcal{S}_\alpha^{\mathrm{Q}}$ of correlations of the form

$$p(b|x,y) = \sum_\lambda p(\lambda)\,\mathrm{tr}\left(\rho_x^{(\lambda)} M_{b|y}^{(\lambda)}\right), \tag{4}$$

where λ denotes the shared classical variable, distributed according to $p(\lambda)$, and the information is bounded by $\mathcal{I}_X \leq \alpha$. The quantity $\mathcal{I}_X$ is computed via Eq. (3), considering the average guessing probability of the ensemble $\mathcal{E} = \{p(\lambda), \mathcal{E}_\lambda\}$:

$$P_g(X|\mathcal{E}) = \sum_\lambda p(\lambda) P_g(X|\mathcal{E}_\lambda), \tag{5}$$

where $P_g(X|\mathcal{E}_\lambda)$ denotes the guessing probability for the subensemble $\mathcal{E}_\lambda = \{p_X(x), \rho_x^{(\lambda)}\}$.

3 Classical correlations

Similarly to above, we can characterise the set of classical correlations, $\mathcal{S}_\alpha^{\mathrm{C}}$, subject to an information bound. In this setting, Alice encodes x into a classical message $m \in [d]$. Bob then provides an output based on his input y and the message m. Considering again shared randomness, the resulting correlations take the form

$$p(b|x,y) = \sum_\lambda p(\lambda) \sum_{m=1}^{d} p_{\mathrm{A}}(m|x,\lambda) p_{\mathrm{B}}(b|m,y,\lambda). \tag{6}$$

In order to characterise correlations of the above form such that $\mathcal{I}_X \leq \alpha$, we proceed as follows. First, notice that the dimension d of the message may a priori be unbounded. However, it turns out that, without loss of generality, one can restrict to the case $d = n$. Next, notice that each encoding of the message $p_{\mathrm{A}}(m|x,\lambda)$ can be taken to be deterministic, i.e. m is a deterministic function of x and λ. Finally, to each of these deterministic encodings, we can associate a guessing probability $P_g^{(\lambda)}$. A detailed discussion is given in Appendix.

With these in hand, we notice that the constraint $\mathcal{I}_X \leq \alpha$ is equivalent to $\sum_\lambda p(\lambda) P_g^{(\lambda)} \leq 2^{\alpha - H_{\min}(X)}$, which is linear in $p(\lambda)$. Therefore, the set $\mathcal{S}_\alpha^{\mathrm{C}}$ forms a convex polytope. The facets of the polytope correspond to linear inequalities

$$\sum_{x,y,b} r_{xyb}\, p(b|x,y) \leq \beta \tag{7}$$

where r_{xyb} and β are real coefficients, which give a complete characterisation of $\mathcal{S}_\alpha^{\mathrm{C}}$.

We have explicitly characterised $\mathcal{S}_\alpha^{\mathrm{C}}$ for scenarios featuring a small number[1] of inputs and outputs. We find three types of facet inequalities: (i) positivity conditions, e.g. $p(b|x,y) \geq 0$, (ii) inequalities ensuring the information bound on the observed correlations, e.g. $\sum_x p(b = x|x,y) \leq 2^{\alpha - H_{\min}(X)}$ (assuming here $n = k$), and (iii) other inequalities. Inequalities (i) and (ii) are in a sense trivial, as they must be satisfied by all physical correlations (when assuming $\mathcal{I}_X \leq \alpha$). On the contrary, inequalities (iii) are non-trivial, and thus capture limits of classical correlations. These inequalities do not necessarily hold for quantum correlations, as we show below.

Finally, note that the problem of determining whether some observed correlations $p(b|x,y)$ can be obtained classically with $\mathcal{I}_X \leq \alpha$ bits of infor-

[1] Typically, characterising $\mathcal{S}_\alpha^{\mathrm{C}}$ quickly becomes computationally demanding as we increase the number of inputs and outputs (the number of vertices grows rapidly). While we could solve cases with $n = 2, 3$ efficiently and the case of $n = 4$ preparations within reasonable time, evaluating $n = 5$ preparations becomes time-consuming on a standard desktop computer.

mation is a linear program. One can thus determine the minimal amount of information required to produce $p(b|x,y)$ in a classical protocol.

4 Quantum advantage

A critical question is whether informationally restricted quantum correlations can outperform their classical counterparts (and thereby provide a quantum advantage). To answer this question, we have considered simple scenarios – labelled by the number of inputs and outputs, i.e. (n,l,k) – and characterised their classical polytope $\mathcal{S}_\alpha^{\mathrm{C}}$. Alice's input is always chosen to be uniformly distributed, i.e. $p_X(x) = 1/n$. The simplest scenario where we could find a non-trivial facet inequality is (3,2,2). We conjecture that $n \geq 3$ is necessesary (we have checked that no quantum advantage is possible for $(2,2,2)$ and $(2,2,3)$).

The scenario $(3,2,2)$ features two non-trivial facets showing a quantum advantage (see Appendix) . Here we focus on one of them:

$$F_1 \equiv -E_{11}-E_{12}-E_{21}+E_{22}+E_{31} \leq 2^{\alpha+1}-1 \quad (8)$$

where $E_{xy} = p(0|x,y) - p(1|x,y)$ and $\mathcal{I}_X \leq \alpha \in [0, \log 3]$. Notice that for $\alpha = 1$, this inequality is identical to the simplest dimension witness of Ref. [4] for classical bits.

Importantly, the above inequality can be violated in quantum theory whenever[2] $\mathcal{I}_X \in (0, \log 3)$, as illustrated in Fig. 2. Let Alice and Bob share one bit of randomness ($\lambda \in \{0,1\}$) with distribution $q \equiv p(\lambda = 0)$. When $\lambda = 0$, Alice prepares the qubit ensemble $\mathcal{E}_0 = \{\frac{1}{3}, |\psi_x\rangle\}$ with $|\psi_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, $|\psi_2\rangle = |0\rangle$ and $|\psi_3\rangle = \sin\frac{\pi}{8}|0\rangle - \cos\frac{\pi}{8}|1\rangle$. Bob measures the observables $-\frac{\sigma_x+\sigma_z}{\sqrt{2}}$ and $\frac{\sigma_z-\sigma_x}{\sqrt{2}}$, where $(\sigma_x, \sigma_y, \sigma_z)$ are the Pauli matrices. When $\lambda = 1$, Alice sends no information and Bob outputs $b = 1$ regardless of y. This strategy results in the witness value $F_1 = 1 + 2\sqrt{2}q$, while the information is $\mathcal{I}_X = \log(1+q)$. Thus, this strategy is relevant in the range $\mathcal{I}_X \in [0,1]$. When $\mathcal{I}_X \in [1, \log(3)]$, we consider another mixed strategy. For $\lambda = 0$ we use again the ensemble $\mathcal{E}_0$ and associated measurements, and for $\lambda = 1$ a qutrit strategy in which Alice sends x to Bob, thus attaining the maximal value of $F_1 = 5$. We get $F_1 = \left(1 + 2\sqrt{2}\right) q + 5\,(1-q)$ and $\mathcal{I}_X = \log(3-q)$.

[2]The extremal cases $\mathcal{I}_X \in \{0, \log 3\}$ are trivial since they correspond to no information and relaying x respectively.

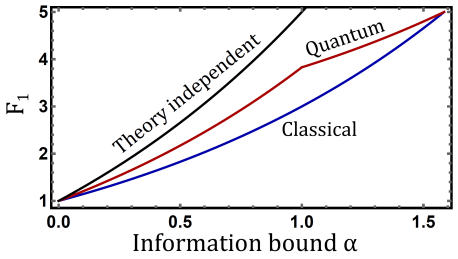

Figure 2: Witness value F_1 as a function of the information bound $\mathcal{I}_X \leq \alpha$. Classical correlations necessarily satisfy the inequality (8) (blue curve). Quantum correlations outperform classical ones for $\alpha \in (0, \log 3)$; the red curve is obtained by a family of quantum protocols and therefore constitutes a lower bound on the optimal quantum correlations. The black curve represents a lower bound on theory-independent correlations, i.e. a lower bound on the information needed for the value F_1.

An interesting question is to find the optimal value of F_1 for any possible quantum strategy with bounded information. This is a non-trivial question as one should consider quantum systems of arbitrarily large Hilbert-space dimension. Based on numerical search, we show in Appendix the existence of slightly better quantum strategies than the above one, but we did not find a simple parameterisation for them.

5 Device-independent bounds on information

While determining the limits of quantum correlations for limited information is challenging, we can nevertheless infer a general, theory-independent, lower bound on information given observed correlations $p(b|x,y)$.

The assumption $\mathcal{I}_X \leq \alpha$ implies that, from any of the distributions $\{p(b|x,1), \ldots, p(b|x,l)\}$, one cannot extract more than α bits of information about x. Allowing for an arbitrary post-processing of the data (Bob creating a new output b' from y and b), i.e. $p(b'|y,b) \geq 0$ with $\sum_{b'} p(b'|y,b) = 1$ where $b' \in [n]$, we obtain the constraints

$$\forall y: \quad \sum_{x,b} p_X(x) p(b|x,y) p(b'=x|y,b) \leq 2^{\alpha - H_{\min}(X)}.$$

Determining whether a given correlation $p(b|x,y)$ is compatible with the above constraints can be cast as a linear program. If the program admits no feasible solution, then an information $\mathcal{I}_X > \alpha$ is necessary to reproduce $p(b|x,y)$. Note that, while the above constraints are necessary to ensure that $\mathcal{I}_X \leq \alpha$, they are most likely not sufficient in general. How to derive stronger constraints on information is an interesting open problem.

To illustrate the relevance of these ideas, we have derived a lower bound on $\mathcal{I}_X$ given an observed value of the witness F_1. The results are illustrated in Fig. 2 and demonstrate the possibility of certifying a device-independent lower bound on the information. Note that the bound applies to quantum correlations, and more generally to any operational theory.

6 Information versus dimension

Another relevant question is to compare quantum correlations with bounded information to those achievable with bounded dimension. Such comparison makes sense when $\mathcal{I}_X \leq \log d$, where d is the Hilbert-space dimension of the quantum systems. Clearly, any correlation achieved via d-dimensional systems (qudits) requires at most $\mathcal{I}_X = \log d$, as any ensemble of qudits carries no more than $\log d$ bits of information [14]. However, it turns out that there are quantum correlations not achievable via qudits that can nevertheless be obtained with information $\mathcal{I}_X = \log d$.

Specifically, we consider the case $d = 2$ and exhibit quantum correlations achievable with $\mathcal{I}_X = 1$ that cannot be obtained from qubits. Consider a Random Access Code [17–19] in which Alice receives a uniformly random four-bit input $x = (x_1, x_2, x_3, x_4) \in [2]^4$. Bob has settings $y \in [4]$, and returns a binary output b with which he aims to guess x_y. The score is

$$F_{\text{RAC}} = \frac{1}{64} \sum_{x,y} p(b = x_y|x,y). \tag{9}$$

Qubit strategies must satisfy $F_{\text{RAC}} < 3/4$; this follows from the impossibility of having four mutually unbiased bases for qubits [12, 18]. Moreover, numerical optimisation strongly suggests that $F_{\text{RAC}} \leq 0.741$ for qubits [18].

It is nevertheless possible to obtain the score $F_{\text{RAC}} = 3/4$ using quantum ensembles with $\mathcal{I}_X = 1$. The strategy employs 16 four-dimensional quantum states of the form

$$\rho_x = \frac{1}{8}\Big(2\mathbb{1}\otimes\mathbb{1} - (-1)^{x_4}\mathbb{1}\otimes\sigma_y - (-1)^{x_1}\sigma_x\otimes\sigma_x - (-1)^{x_2}\sigma_y\otimes\sigma_x - (-1)^{x_3}\sigma_z\otimes\sigma_x\Big), \tag{10}$$

and Bob measures the observables $B_1 = \sigma_x \otimes \sigma_x$, $B_2 = \sigma_y \otimes \sigma_x$, $B_3 = \sigma_z \otimes \sigma_x$ and $B_4 = \mathbb{1} \otimes \sigma_y$. Note that, despite being four-dimensional, these states are noisy (with purity $\text{tr}\,(\rho_x^2) = 1/2\ \forall x$) and carry only one bit of information. Since all states have the same spectrum, $(1/2, 1/2, 0, 0)$, this can be checked analytically as follows. For any quantum ensemble, the information is upper bounded by

$$I_X \leq \log(d) + \log\left(\frac{\max_x p_X(x)\lambda_{\max}(\rho_x)}{\max_x p_X(x)}\right), \tag{11}$$

where $\lambda_{\max}(\rho_x)$ is the largest eigenvalue of ρ_x, and d the Hilbert-space dimension. The bound is obtained from using the relation $\text{tr}\,[\rho_x N_x] \leq \lambda_{\max}(\rho_x)\,\text{tr}\,[N_x]$ in Eq. (2) and then $\sum_x N_x = \mathbb{1}_d$. The bound Eq. (11) is tight when (i) for each x, ρ_x only has one non-zero eigenvalue (with possible multiplicity) and (ii) $p_X(x)\lambda_{\max}(\rho_x)$ is constant in x. The ensemble in Eq. (10) satisfies this criteria.

An interesting question is whether larger separation is possible. That is, how much stronger can quantum correlations with $\mathcal{I}_X = \log d$ bits of information become compared to quantum correlations using d-dimensional quantum systems. In Appendix, we show that, in a scenario without shared randomness, this advantage can be made arbitrarily large. Specifically, we construct quantum correlations achievable with $\mathcal{I}_X = 1$ bit of information, that can only be reproduced using an arbitrary large Hilbert-space dimension.

7 Quantum communication versus entanglement-assisted classical communication

Informationally restricted quantum systems also have interesting implications when comparing quantum resources in different communication scenarios. On the one hand, Alice may send an amount of quantum communication to Bob

(as in Fig. 1). On the other hand, Alice and Bob may share unlimited entanglement and Alice communicates the same amount classically. These two approaches are generally not equivalent. In fact, for dimensionally restricted classical and quantum messages, there is no strict hierarchy. In some cases, quantum communication outperforms entanglement-assisted classical communication [20–22] and vice versa in others [22–24]. Given this seemingly complicated picture, no generally valid criterion is known for determining which quantum resource is more efficient for a given communication task. Interestingly, we show that every correlation obtained via entanglement-assisted classical communication of a d-dimensional message can also be obtained via quantum communication carrying at most $\log d$ bits of information. That is, in this setting, quantum communication is the stronger resource.

Consider a scenario with classical communication, where Alice and Bob can use a pre-shared entangled state ρ_{AB}. Upon receiving input x, Alice performs a measurement $\{A_{a|x}\}$ with outcome a on her half of ρ_{AB}, which projects Bob's system onto the state $\sigma_{a|x} = \mathrm{tr}_{\mathrm{A}}([A_{a|x}\otimes \mathbb{1}_{\mathrm{B}}]\rho_{\mathrm{AB}})/p(a|x)$, where $p(a|x) = \mathrm{tr}([A_{a|x}\otimes \mathbb{1}_{\mathrm{B}}]\rho_{\mathrm{AB}})$. Alice then sends a classical message to Bob; which we represent as a collection of d-dimensional quantum states $\mu_{a|x}$ diagonal in the same basis. Thus, Bob holds the classical-quantum state $\mu_{a|x}\otimes\sigma_{a|x}$, on which he can perform some measurements in order to establish correlations $p(b|x,y)$. The information cost of this protocol originates only from the classical message, as the entanglement is pre-shared.

Now, we construct a quantum communication protocol to simulate the above correlations using at most $\log d$ bits of information. Upon receiving x, Alice samples from $p(a|x)$, and sends to Bob the classical-quantum state $\mu_{a|x}\otimes\sigma_{a|x}$. Evidently, Bob can now reproduce the same correlations $p(b|x,y)$. The key point is now to show that this protocol does not require more information than above. The ensemble (averaged over a) can be written $\mathcal{E}_{\mathrm{QC}} = \{p_X(x),\tau_x\}$ where $\tau_x = \sum_a p(a|x)\mu_{a|x}\otimes\sigma_{a|x}$. The corresponding guessing probability is

$$P_g^{\mathrm{QC}} = \max_{\{N_z\}} \sum_{a,x} p_X(x)p(a|x)\,\mathrm{tr}\left(\mu_{a|x}\otimes\sigma_{a|x}N_x\right) \tag{12}$$

where the POVM $\{N_z\}$ acts jointly on the classical message space and on the quantum state space. We can place the following upper bound on the guessing probability

$$P_g^{\mathrm{QC}} \leq \max_{\{N_z\}} \sum_x p_X(x)\,\mathrm{tr}\left(\left(\sum_a p(a|x)\sigma_{a|x}\right)N_x^{\mathrm{B}}\right), \tag{13}$$

where we have used that $\mathrm{tr}(\mu_{a|x}\otimes\sigma_{a|x}N_x) \leq \mathrm{tr}(\sigma_{a|x}N_x^{\mathrm{B}})$, where N_x^{B} is the partial trace of N_x over the first system (the classical message space). Importantly, since for every x the ensemble $\{p(a|x),\sigma_{a|x}\}$ is remotely prepared by Alice on Bob's side, it follows that

$$\sum_a p(a|x)\sigma_{a|x} = \sum_a \mathrm{tr}_{\mathrm{A}}\left(A_{a|x}\otimes\mathbb{1}_{\mathrm{B}}\rho_{\mathrm{AB}}\right) = \mathrm{tr}_{\mathrm{A}}\left(\rho_{\mathrm{AB}}\right) = \rho_{\mathrm{B}}. \tag{14}$$

Therefore, the guessing probability obeys

$$P_g^{\mathrm{QC}} \leq \max_{\{N_z\}} \sum_x p_X(x)\,\mathrm{tr}\left(N_x^{\mathrm{B}}\rho_{\mathrm{B}}\right) \tag{15}$$

$$\leq \left(\max_x p_X(x)\right) \max_{\{N_z\}} \mathrm{tr}\left(\sum_x N_x^{\mathrm{B}}\rho_{\mathrm{B}}\right). \tag{16}$$

Finally, we use the completeness relation of POVMs to obtain

$$\sum_x N_x^{\mathrm{B}} = \sum_x \mathrm{tr}_1\left(N_x\right) = \mathrm{tr}_1\left(\mathbb{1}_d\otimes\mathbb{1}\right) = d\mathbb{1}, \tag{17}$$

where we have used that the identity operator on the classical message space is d-dimensional. Thus, we conclude that

$$P_g^{\mathrm{QC}} \leq d\max_x p_X(x). \tag{18}$$

Consequently, the information is bounded by

$$\begin{aligned}\mathcal{I}_X &= -\log\left(\max_x p_X(x)\right) + \log\left(P_g^{\mathrm{QC}}\right)\\ &\leq -\log\left(\max_x p_X(x)\right) + \log\left(d\max_x p_X(x)\right) = \log d.\end{aligned} \tag{19}$$

This concludes the proof: quantum communication of $\log d$ bits of information is a stronger resource than classical communication of a d-dimensional message assisted by any amount of entanglement.

Finally, we also note that this proof remains valid also if Alice uses her classical outcome a and her input x to encode a quantum d-dimensional message $\mu_{a|x}$. This is, however, not the most general quantum operation that may be considered.

8 Conclusions

We have investigated correlations in prepare-and-measure scenarios under the assumption of an upper bound on the information. We have shown how to fully characterise correlations in the case of classical systems and proved a quantum advantage. Moreover, we showed that stronger quantum correlations can be obtained when bounding the information rather than the dimension, and devised device-independent tests of information.

An outstanding open question is to characterise quantum correlations when the transmitted information is bounded. Is it sufficient to consider quantum ensembles of finite dimension, as in the classical case? Or are there correlations that require infinite-dimensional quantum systems? Another point is to understand how much stronger quantum correlations can be compared to classical ones. For the case where shared randomness is not allowed, we could show a diverging advantage. Is it also the case in a scenario including shared randomness? In addition, it would be interesting to consider informationally restricted correlations based on other information measures than the one we consider here; for instance based on Shannon entropies. Another possible direction is to explore connections between our approach and other scenarios in information theory, for instance the (quantum) information bottleneck function [25, 26]. Furthermore, it would also be relevant to explore the role of informationally restricted correlations with respect to the line of research focused on operational contextuality [27] in which one considers prepare-and-measure scenarios featuring an assumption of oblivious communication (see e.g. [28, 29]).

Finally, we briefly discuss the prospects of using our approach in experiments, notably towards possible applications in semi-device-independent (SDI) quantum information processing. In this area, protocols so far were mostly based on a dimension assumption, see e.g. [8–12], which is usually justified from the physics of the experiment. For instance, a setup where the relevant degree of freedom is the polarization of a single photon motivates the assumption that the prepared states can be described as qubits. In practice, however, single-photon sources feature imperfections which result in unavoidable multi-photon emissions, which clearly no longer satisfy the qubit assumption. Taking these into account is typically cumbersome and inefficient (see for instance [11]). In comparison, the information approach might be much better adapted here. From a physical model of the source, the rate of multi-photon events can be estimated. For instance, a weak laser source will exhibit Poisson statistics. For each photon number the carried information can be estimated, which in turn results in an overall bound on the carried information. In this way, one could continuously tune the information bound, taking into account the relevant degrees of freedom and photon statistics. Bounding the information rather than the dimension may therefore represent a more natural assumption, better motivated by the physics of the source. It would be interesting to explore these ideas in practice, as well as to understand the relation between the information approach and other SDI approaches recently developed, based on bounding the energy [30], the overlap [31, 32] or the entropy [33] of the quantum communication.

Acknowledgements

We thank Jean-Daniel Bancal, Joseph Renes, Renato Renner, Joseph Bowles, Alastair Abbott, Francesco Buscemi, Michele Dall'Arno, Stefano Pironio and Erik Woodhead for discussions. This work was supported by the Swiss National Science Foundation (Starting grant DIAQ, NCCR-QSIT) and the Independent Research Fund Denmark.

References

[1] H. Buhrman, R. Cleve, S. Massar, and R. de Wolf, Nonlocality and communication complexity, Rev. Mod. Phys. **82**, 665 (2010).

[2] H. Buhrman, R. Cleve and A. Wigderson, Quantum vs. classical communication and computation, Proceedings of the 30th Annual ACM Symposium on Theory of Computin, 63 (1998).

[3] R. Raz, Exponential separation of quantum and classical communication complexity, In Proceedings of 31st ACM STOC, 358 (1999).

[4] R. Gallego, N. Brunner, C. Hadley, and A. Acín, Device-Independent Tests of Classical and Quantum Dimensions, Phys. Rev. Lett. **105**, 230501 (2010).

[5] J. Ahrens, P. Badziag, A. Cabello, and M. Bourennane, Experimental Device-independent Tests of Classical and Quantum Dimensions, Nature Physics **8**, 592 (2012).

[6] M. Hendrych, R. Gallego, M. Micůda, N. Brunner, A. Acín, J. P. Torres, Experimental estimation of the dimension of classical and quantum systems, Nature Physics **8**, 588 (2012).

[7] M. Navascués, and T. Vértesi, Bounding the Set of Finite Dimensional Quantum Correlations, Phys. Rev. Lett. **115**, 020501 (2015).

[8] M. Pawłowski, and N. Brunner, Semi-device-independent security of one-way quantum key distribution, Phys. Rev. A **84**, 010302(R) (2011).

[9] H-W. Li, Z-Q. Yin, Y-C. Wu, X-B. Zou, S. Wang, W. Chen, G-C. Guo, and Z-F. Han, Semi-device-independent random-number expansion without entanglement, Phys. Rev. A **84**, 034301 (2011).

[10] E. Woodhead, S. Pironio, Secrecy in Prepare-and-Measure Clauser-Horne-Shimony-Holt Tests with a Qubit Bound, Phys. Rev. Lett. **115**, 150501 (2015).

[11] T. Lunghi, J. B. Brask, C. C. W. Lim, Q. Lavigne, J. Bowles, A. Martin, H. Zbinden, and N. Brunner, Self-Testing Quantum Random Number Generator, Phys. Rev. Lett. **114**, 150501 (2015).

[12] A. Tavakoli, J. Kaniewski, T. Vértesi, D. Rosset, and N. Brunner, Self-testing quantum states and measurements in the prepare-and-measure scenario, Phys. Rev. A **98**, 062307 (2018).

[13] N. Ciganović, N. J. Beaudry, and Renato Renner, Smooth Max-Information as One-Shot Generalization for Mutual Information, IEEE Transactions on Information Theory **60**, 1573 (2014).

[14] A. S. Holevo, Bounds for the quantity of information transmitted by a quantum communication channel, Problems of Information Transmission. **9**, 177 (1973).

[15] R. Jozsa, D. Robb, and W. K. Wootters, Lower bound for accessible information in quantum mechanics, Phys. Rev. A **49**, 668 (1994).

[16] Convex Optimization, S. Boyd and L. Vandenberghe, Cambridge University Press, 2004.

[17] A. Ambainis, A. Nayak, A. Ta-Shma, U. Vazirani, Dense quantum coding and a lower bound for 1-way quantum automata, Proceedings of the 31st Annual ACM Symposium on Theory of Computing (STOC'99), 376-383 (1999).

[18] A. Ambainis, D. Leung, L. Mancinska, M. Ozols, Quantum Random Access Codes with Shared Randomness, arXiv:0810.2937.

[19] A. Tavakoli, A. Hameedi, B. Marques, and M. Bourennane, Quantum random access codes using single d-Level systems, Phys. Rev. Lett. **114**, 170502 (2015).

[20] A. Tavakoli, M. Pawłowski, M. Żukowski, and M. Bourennane, Dimensional discontinuity in quantum communication complexity at dimension seven, Phys. Rev. A 95, 020302(R) (2017).

[21] A. Tavakoli, B. Marques, M. Pawłowski, and M. Bourennane, Spatial versus sequential correlations for random access coding, Phys. Rev. A **93**, 032336 (2016).

[22] A. Hameedi, D. Saha, P. Mironowicz, M. Pawłowski, and M. Bourennane, Complementarity between entanglement-assisted and quantum distributed random access code, Phys. Rev. A **95**, 052345 (2017).

[23] M. Pawłowski, and M Żukowski, Entanglement-assisted random access codes, Phys. Rev. A **81**, 042326 (2010).

[24] A. Tavakoli, and M. Zukowski, Higher-dimensional communication complexity problems: Classical protocols versus quantum ones based on Bell's theorem or prepare-transmit-measure schemes, Phys. Rev. A **95**, 042305 (2017).

[25] N. Tishby, F. C. Pereira and W. Bialek, The information bottleneck method, Proc. of the 37th Annual Allerton Conference on Communication, Control and Computing, pages 368-377, (1999)

[26] N. Datta, C. Hirche and A. Winter, Convexity and Operational Interpretation of the Quantum Information Bottleneck Function, Proc. ISIT 2019, 7-12 July 2019, Paris, pp. 1157-1161.

[27] R. W. Spekkens, Contextuality for preparations, transformations, and unsharp measurements, Phys. Rev. A **71**, 052108 (2005).

[28] R. W. Spekkens, D. H. Buzacott, A. J. Keehn, B. Toner, and G. J. Pryde, Prepara-

tion Contextuality Powers Parity-Oblivious Multiplexing, Phys. Rev. Lett. **102**, 010401 (2009).

[29] A. Hameedi, A. Tavakoli, B. Marques and M. Bourennane, Communication Games Reveal Preparation Contextuality, Phys. Rev. Lett. **119**, 220402 (2017).

[30] T. V. Himbeeck, E. Woodhead, N. J. Cerf, R. Garcia-Patron, and S. Pironio, Semi-device-independent framework based on natural physical assumptions, Quantum **1**, 33 (2017).

[31] J. B. Brask, A. Martin, W. Esposito, R. Houlmann, J. Bowles, H. Zbinden, and N. Brunner, Megahertz-Rate Semi-Device-Independent Quantum Random Number Generators Based on Unambiguous State Discrimination, Phys. Rev. Applied **7**, 054018 (2017).

[32] Y. Wang, I. W. Primaatmaja, E. Lavie, A. Varvitsiotis, C. C. W. Lim, Characterising the correlations of prepare-and-measure quantum networks, npj Quantum Information **5**, 17 (2019).

[33] R. Chaves, J. B. Brask, and N. Brunner, Device-Independent Tests of Entropy, Phys. Rev. Lett. **115**, 110501 (2015).

[34] M. Hayashi1, K. Iwama, H. Nishimura, R. Raymond, and S. Yamashita, (4,1)-Quantum random access coding does not exist - one qubit is not enough to recover one of four bits, New J. Phys. **8** 129 (2006).

[35] A. Chailloux, I. Kerenidis, S. Kundu, and J. Sikora, Optimal bounds for parity-oblivious random access codes, New J. Phys. **18** 045003 (2016).

A Characterisation of classical correlations

We describe a classical scheme, starting with deterministic strategies. Alice uses an encoding function $E : [n] \to [d]$ to associate her input to a d-valued message $m = E(x)$ and sends it to Bob. No limitation on d is assumed. Bob uses a decoding function $D : [d] \times [l] \to [k]$ to map the pair (m, y) into an k-valued output $b = D(m, y)$. Since there are $Z_{\mathrm{A}} = d^n$ ($Z_{\mathrm{B}} = k^{dl}$) possible encoding (decoding) functions, the number of deterministic strategies is $Z = Z_{\mathrm{A}} Z_{\mathrm{B}}$. We index them by $(E_{\lambda_{\mathrm{A}}}, D_{\lambda_{\mathrm{B}}})$ for $\lambda_{\mathrm{A}} \in [Z_{\mathrm{A}}]$ and $\lambda_{\mathrm{B}} \in [Z_{\mathrm{B}}]$ respectively. Via the shared randomness $\lambda = (\lambda_{\mathrm{A}}, \lambda_{\mathrm{B}})$, classical correlations are written

$$p^{\mathrm{C}}(b|x, y) = \sum_{\lambda} p(\lambda) \sum_{m=1}^{d} \delta_{m, E_{\lambda_{\mathrm{A}}}(x)} \delta_{b, D_{\lambda_{\mathrm{B}}}(m,y)}. \tag{20}$$

We now characterise $p^{\mathrm{C}}(b|x, y)$ when $\mathcal{I}_X \leq \alpha$ for some real $\alpha \geq 0$. To this end, we need to eliminate the dimension d. Below, in section A.3 we show that without loss of generality one can choose $d = n$ (i.e. the dimension equal to the number of inputs for Alice). We will use this fact to characterise the polytope of classical correlations and leave the proof for the end of this section.

A.1 The classical polytope

We use that classical messages of dimension $d = n$ are sufficient. Therefore, we can denote all encoding functions and decoding functions $(E_{\lambda_{\mathrm{A}}}, D_{\lambda_{\mathrm{B}}})$ where the index $\lambda = (\lambda_{\mathrm{A}}, \lambda_{\mathrm{B}})$ acts as a shared random variable (whose cardinality is now finite) allowing the coordination of deterministic encoding and decoding strategies. For a fixed deterministic strategy, we obtain a distribution $p'_\lambda(b|x, y)$. This distribution is a vertex of the polytope $\mathbb{P}$ which is the space of all probabilities $p(b|x, y)$. However, many deterministic strategies give rise to the same vertex in the probability space. Therefore, we write $\{p_\gamma(b|x, y)\}_\gamma$ for the unique elements in the set $\{p'_\lambda(b|x, y)\}_\lambda$. We define

$$E_\gamma = \{\lambda = (\lambda_{\mathrm{A}}, \lambda_{\mathrm{B}}) | p_\gamma(b|x, y) = p'_\lambda(b|x, y)\}, \tag{21}$$

where $\{p_\gamma(b|x, y)\}$ is the list of vertices of $\mathbb{P}$ (without duplicates). In other words, E_γ is the set of all deterministic strategies that generate the vertex $p_\gamma(b|x, y)$.

To each vertex of $\mathbb{P}$ we associate the smallest amount of information needed to generate it (for simplicity, we work with the guessing probability). That is,

$$P_g^{(\gamma)} = \min_{\lambda \in E_\gamma} P_g^{(\lambda_{\mathrm{A}})} \tag{22}$$

where the guessing probability of the determinis-

tic strategy is given by

$$P_g^{(\lambda_A)} = \max_{\mu} \sum_x p_X(x) \sum_{m=1}^{d} \delta_{m,E_{\lambda_A}(x)} \delta_{x,\tilde{D}_\mu(m)}, \tag{23}$$

where the maximisation is over all the deterministic decoding strategies $\tilde{D} : [d] \to [n]$ (of which there are n^d).

We now impose the information restriction, $I_X \leq \alpha$. This can be formulated as a linear constraint in the shared randomness. The characterisation of the set of information restricted classical correlations reads

$$p(b|x,y) = \sum_\gamma p(\gamma) p_\gamma(b|x,y) \tag{24}$$

$$\sum_\gamma p(\gamma) P_g^{(\gamma)} \leq 2^{\alpha - H_{\min}(X)} \tag{25}$$

$$\sum_\lambda p(\gamma) = 1 \tag{26}$$

$$p(\gamma) \geq 0. \tag{27}$$

This defines a convex polytope. Its facets can be obtained using standard polytope software. We label this polytope $\mathbb{P}_\alpha$ and note that it is contained inside $\mathbb{P}$.

As an illustration of how the polytope $\mathbb{P}_\alpha$ may look, we have displayed in Fig. 3 a schematic of the polytope in the simplest case of Alice having two inputs and Bob performing a single binary outcome measurement ($n = k = 2$, $l = 1$), for which the polytopes $\mathbb{P}$ and $\mathbb{P}_\alpha$ are polygons.

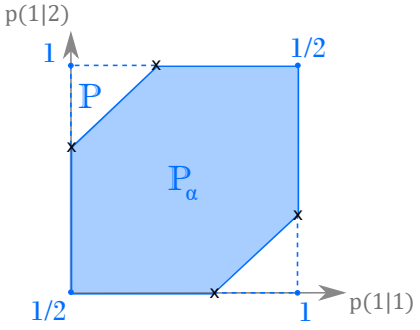

Figure 3: The classical set of correlations for a scenario with two preparations and one binary outcome measurement $(n, l, k) = (2, 1, 2)$. The polytope $\mathbb{P}$ has four vertices, each corresponding to a guessing probability of either one or one half (written in blue). The facets are lines. Therefore there is only one pair of vertices per facet, for each of which we inscribe a new vertex (represented by a tick) as imposed by limiting the guessing probability. Thus, the blue region is the polytope $\mathbb{P}_\alpha$.

A.2 Optimal classical correlations via linear programming

Since the set of classical correlations forms a convex polytope for $\mathcal{I}_X \leq \alpha$, one can determine whether a given $p(b|x,y)$ belongs to said polytope via a linear program. This allows one to determine whether $p(b|x,y)$ is classically realisable with information no more than α.

Moreover, given any linear functional of probabilities,

$$F = \sum_{x,y,b} r_{xyb}\, p(b|x,y), \tag{28}$$

one can determine the exact classical bound through the evaluation of the linear program

$$\begin{aligned} F^{\mathrm{C}} &= \max_{p(\lambda)} F[p(b|x,y)] \\ &\text{such that } \sum_\lambda p(\lambda) P_g^{(\lambda_A)} \leq 2^{\alpha - H_{\min}(X)}, \\ &\sum_\lambda p(\lambda) = 1, \quad \text{and} \quad p(\lambda) \geq 0. \end{aligned} \tag{29}$$

This allows to obtain witnesses for classical correlations.

A.3 Dimension n is sufficient for classical messages

Here, we show that the optimal classical correlations, for any correlation witness constrained by bounded guessing probability (or equivalently, bounded information) with shared randomness, is obtained with a message dimension not larger than the cardinality of the input of Alice, i.e. $d = n$.

Any classical strategy can be decomposed as a mixture of deterministic strategies, as given by Eq. (20). For a fixed value of the shared variable λ, the encoding strategy E_{λ_A} is fixed. Since x can take at most n different values, there is then at most n different values of $E_{\lambda_A}(x)$. Thus, for a fixed λ, at most n message symbols are used. Whether there is any advantage in using message dimensions $d > n$ thus becomes a question of whether there is any advantage in using different sets of message symbols for different λ.

We first show that any value of the maximum in Eq. (29) obtained with different sets of message symbols for different λ can also be achieved using the same set of n symbols for all λ. This can be seen from Eq. (20). For each value of λ, the factor $\delta_{m,E_{\lambda_A}(x)}$ is nonzero for at most n different values

of m. The decoding function $D_{\lambda_B}(m)$ hence needs to be defined only on these values. If any of these values lie outside $[n] = \{1, \ldots, n\}$ then there must be corresponding values in $[n]$ which are not used. We can then redefine E_{λ_A} and D_{λ_B} to use these values instead.

Specifically, for some fixed λ, say that $E_{\lambda_A}(x_0) = \nu \notin [n]$ for some x_0. Then there exists $\nu' \in [n]$ such that $E_{\lambda_A}(x) \neq \nu'$ for all x. We then define

$$E'_{\lambda_A}(x) = \begin{cases} \nu' & \text{if } x = x_0, \\ E_{\lambda_A}(x) & \text{otherwise,} \end{cases} \tag{30}$$

$$D'_{\lambda_B}(m) = \begin{cases} D_{\lambda_B}(\nu) & \text{if } m = \nu', \\ D_{\lambda_B}(m) & \text{otherwise.} \end{cases} \tag{31}$$

Substituting $E_{\lambda_A} \to E'_{\lambda_A}$ and $D_{\lambda_B} \to D'_{\lambda_B}$ in (20) leaves the probabilities $p(b|x, y)$ unchanged. Repeating this process, the message symbols can be chosen in $[n]$ for every λ, without changing the probabilities and hence a distribution achieving the optimum in Eq. (29) remains optimal.

The only remaining question is now, whether this remapping to a strategy using the same n symbols for all λ can lead to violation of the information constraint. From (23), we can see that this is not the case. Let $\tilde{D}_{\mu^*}$ be the optimal decoding function which achieves the maximum on the right-hand side of (23), for some fixed λ. When E_{λ_A} is replaced by E'_{λ_A} as above, the maximum remains unchanged and is achieved by

$$\tilde{D}'_{\mu^*}(m) = \begin{cases} \tilde{D}_{\mu^*}(\nu) & \text{if } m = \nu', \\ \tilde{D}_{\mu^*}(m) & \text{otherwise.} \end{cases} \tag{32}$$

Thus, following the recipe above, we can replace all the encoding and decoding functions $E_{\lambda_A} : [n] \to [d]$, $D_{\lambda_B} : [d] \to [n]$, and $\tilde{D}_\mu : [d] \to [n]$ by other functions $E_{\lambda_A} : [n] \to [n]$, $D_{\lambda_B} : [n] \to [n]$, and $\tilde{D}_\mu : [n] \to [n]$ without changing the probabilities $p(b|x, y)$ or the guessing probabilities $P_g^{(\lambda_A)}$. It follows that the optimum of Eq. (29) can always be attained using a message dimension of at most n.

B Case study for $(n, l, k) = (3, 2, 2)$

We have obtained the facets of the polytope for several simple scenarios. The simplest scenario in which we have found non-trivial facets is $(n, l, k) = (3, 2, 2)$. One can consider different values for the information bound $\mathcal{I}_X \leq \alpha$. We have considered different values of α for each of which we have found two non-trivial inequalities (i.e. they are not positivity nor the information restriction). More precisely, we considered eleven evenly spaced values of the guessing probability in the range $(1/3, 1)$. The facets are

$$F_1 = \sum_{x,y} t^1_{x,y} E(x, y) \leq 6P_g - 1 \tag{33}$$

$$F_2 = \sum_{x,y} t^2_{x,y} E(x, y) \leq 12P_g - 4. \tag{34}$$

where $t^1_{x,y} = \{[-1, -1], [-1, 1], [1, 0]\}$ and $t^2_{x,y} = \{[-1, -1], [-1, 1], [2, 0]\}$. Note that for convenience, we have expressed the upper bounds in terms of the guessing probability instead of the information. Both inequalities can be violated in quantum theory. For the first inequality, a violation valid for any non-trivial information was presented in the main text using a quantum strategy with one bit of shared randomness. Notably, said strategy also violates the second inequality but not in the entire range $\mathcal{I}_X \in (0, \log 3)$.

Moreover, we have numerically explored whether larger violations of the first inequality are possible. We considered the case in which Alice prepares general qutrit states and found it to be advantageous. We have employed a brute-force numerical search using the function "fmincon" in MATLAB. We employ an effective Lagrange multiplier λ and seek to maximise the function

$$\tilde{F}_1 = F_1 - \lambda|\mathcal{I}_X - \alpha|, \tag{35}$$

for a given information bound α. We have chosen $\lambda = 100$. In every step, we evaluate the information $\mathcal{I}_X$ in the three preparations via a semidefinite program. Then, we evaluate the largest possible value of F_1 for the given preparations, which thanks to the binary outcomes can be cast as an eigenvalue problem. We then ask MATLAB to maximise $\tilde{F}_1$. In Fig 4 the results are compared to those of the strategy in the main text. In the range $\mathcal{I}_X \in (0, 1)$ we find an improvement, but not in the range $\mathcal{I}_X \in [1, \log 3]$. However, we have not found a simple parameterisation of these quantum strategies. Also, it could be possible that even better results can be obtained with higher-dimensional preparations.

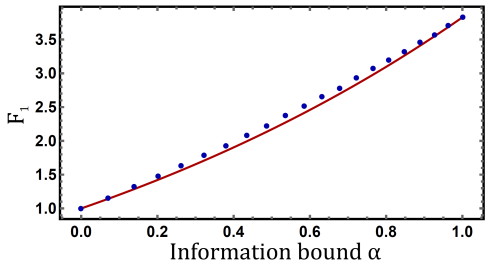

Figure 4: Witness value F_1 as a function of the information $\mathcal{I}_X \leq \alpha$. The quantum strategy from the main text is displayed (red curve) and the numerically obtained quantum violations based on qutrits are displayed in blue. In the range $\alpha \in (0, 1)$ these improve on the first quantum strategy. Notably, numerics showed that an improvement on the red curve is possible already with qubit preparations.

C Arbitrary large advantage over dimension-bounded quantum ensembles without shared randomness

In the main text, we showed that one bit of communication is not always optimally encoded in a qubit ensemble but sometimes in an ensemble of higher-dimensional quantum systems. Here, we show that such advantages over dimension-bounded systems can become more significant in scenarios without shared randomness.

Consider the following variant of a quantum Random Access Code (without shared randomness). Alice has a uniformly random variable $X \in [2^n]$ with values $x = x_1 \dots x_n \in [2]^n$. She sends m bits of information to Bob, who has a random variable $Y \in [n]$ with values y from which he produces an outcome $b \in [2]$. The aim is to maximise the *worst-case* success probability of finding $b = x_y$, i.e.,

$$\mathcal{A}_n^m = \min_{x,y} p(b = x_y|x, y). \tag{36}$$

Let us first choose $m = 1$. It is known that with two-valued classical messages or with two-dimensional quantum systems, it is impossible to achieve a better result than that obtained with random guessing, i.e. $\mathcal{A}_n^1 = 1/2$, when $n > 3$ [34]. In contrast, for $n = 2$ and $n = 3$, qubits hold an advantage over classical two-valued messages. The reason is that for $n = 4$ (and analogously for $n > 4$) it is impossible to cut the Bloch sphere into $2^4 = 16$ symmetric parts with four planes passing through the origin. By a similar argument using the generalised higher-dimensional Bloch sphere, it has been shown [34] that for general integers $m \geq 1$, sending m classical two-valued messages or sending m qubits (2^m-dimensional quantum systems) cannot achieve a better result than $\mathcal{A} = 1/2$ when n is choosen as at least 2^{2m}.

We compare this with sending a general quantum ensemble of limited information. Again, we first choose $m = 1$ and $n = 4$. Using the ensemble and measurements specified in the main text for four-bit Random Access Code (average success probability variant), one immediately finds that $\forall x, y : p(b = x_y|x, y) = 3/4$, and therefore that $\mathcal{A}_4^1 = 3/4$. Thus, the ensemble of mixed four-dimensional systems provides an advantage over two-valued classical messages when qubit ensembles fail to provide any better-than-classical result.

Refs. [21, 35] derived Bell inequalities for Random Access Codes. Using the results of Ref. [35], Alice and Bob can share an entangled state of local dimension $D = 2^{\lfloor \frac{n}{2} \rfloor}$ and use their inputs as settings for testing the Bell inequalities of [21, 35]. Then, if Alice communicates her binary outcome to Bob, he can satisfy the relation $b = x_y$ with probability

$$\forall x, y : \qquad p(b = x_y|x, y) = \frac{1}{2} + \frac{1}{2\sqrt{n}}. \tag{37}$$

In the main text we showed that any correlations achievable by means of entanglement-assisted classical communication also is achievable by means of quantum communication without sending more information (and without the need of share randomness). Therefore, we can obtain the correlations (37) using the quantum communication model discussed in the main text. Consequently, using only a single bit of quantum information (encoded in a general ensemble), we can achieve

$$\mathcal{A}_n^1 = \frac{1}{2} + \frac{1}{2\sqrt{n}}. \tag{38}$$

Note that this is strictly greater than $1/2$ for all $n \geq 2$. Therefore, if we choose $n \geq 2^{2m}$ but use only a single bit of information, we outperform the best possible quantum protocols in which the allowed m bits are encoded in 2^m-dimensional quantum systems. Thus, the advantage is unbounded in the sense that a fixed amount (one

bit) of general quantum information holds an advantage over the m bits carried by m qubits, for any (potentially) arbitrarily large choice of m.

Mutually unbiased bases and symmetric informationally complete measurements in Bell experiments: Bell inequalities, device-independent certification and applications

Armin Tavakoli,[1] Máté Farkas,[2,3] Denis Rosset,[4] Jean-Daniel Bancal,[1] and Jędrzej Kaniewski[5]

[1]*Department of Applied Physics, University of Geneva, 1211 Geneva, Switzerland*
[2]*Institute of Theoretical Physics and Astrophysics, National Quantum Information Centre, Faculty of Mathematics, Physics and Informatics, University of Gdansk, 80-952 Gdansk, Poland*
[3]*International Centre for Theory of Quantum Information, University of Gdansk, 80-308 Gdansk, Poland*
[4]*Perimeter Institute for Theoretical Physics, 31 Caroline St. N, Waterloo, Ontario, N2L 2Y5, Canada*
[5]*Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warsaw, Poland*

(Dated: December 9, 2019)

Mutually unbiased bases (MUBs) and symmetric informationally complete projectors (SICs) are central to many conceptual and practical aspects of quantum theory. In this work, we investigate their role in quantum nonlocality. For every integer $d \geq 2$, we introduce Bell inequalities for which pairs of d-dimensional MUBs and SICs, respectively, produce the largest violations allowed in quantum mechanics. To investigate whether these inequalities can be used for the purpose of device-independent certification of measurements, we show that the concepts of MUBs and SICs admit a natural operational interpretation which does not depend on the dimension of the underlying Hilbert space. We prove that the maximal quantum violations certify precisely these operational notions. In the case of MUBs we also show that the maximal violation certifies the presence of a maximally entangled state of local dimension d and that the maximal violation is achieved by a unique probability distribution. This constitutes the first example of an extremal point of the quantum set which admits physically inequivalent quantum realisations, i.e. is not a self-test. Finally, we investigate the performance of our Bell inequalities in two tasks of practical relevance. We show that the Bell inequalities for MUBs guarantee the optimal key rate in a device-independent quantum key distribution protocol with d outcomes. Moreover, using the Bell inequalities for SICs, we show how qubit and qutrit systems can generate more device-independent randomness than higher-dimensional implementations based on standard projective measurements. We also investigate the robustness of the key and randomness generation schemes to noise. The results establish the relevance of MUBs and SICs for both fundamental and applied considerations in quantum nonlocality.

I. INTRODUCTION

Mutually unbiased bases (MUBs) and symmetric informationally complete projectors (SICs) are widely celebrated, intensively studied and broadly useful concepts in quantum theory. Two bases of a finite-dimensional Hilbert space are called *mutually unbiased* if the inner product between any two elements belonging to different bases has constant magnitude. In other words, if a system is prepared in a state belonging to the first basis, then when a measurement is performed in the second basis, all the outcomes are equally probable [1]. In a similar spirit, a set of rank-one projectors is called *symmetric informationally complete* when the magnitude of all inner products between different projectors is equal and the projectors are tomographically complete [2, 3]. More formally,

- Let $\{|e_j\rangle\}_{j=1}^{d}$ and $\{|f_k\rangle\}_{k=1}^{d}$ be two orthonormal bases of the d-dimensional Hilbert space $\mathbb{C}^d$. The two bases are mutually unbiased if
$$|\langle e_j|f_k\rangle|^2 = \frac{1}{d} \tag{1}$$
for all j and k. The constant on the right-hand-side is merely a consequence of the two bases being normalised.
- Let $\{|r_j\rangle\}_{j=1}^{d^2}$ be a set of unit vectors in $\mathbb{C}^d$. The set is called symmetric informationally complete if
$$|\langle r_j|r_k\rangle|^2 = \frac{1}{d+1} \tag{2}$$
for all $j \neq k$. Again, the constant on the right-hand-side is fixed by normalisation. Moreover, the reason for there being precisely d^2 elements in a SIC[1] is that this is the largest number of unit vectors in $\mathbb{C}^d$ that could possibly admit the uniform overlap property (up to normalisation identical to Eq. (2)).

Whereas MUBs and SICs are inherently different objects, they are frequently studied jointly [4–9]. This is in part due to both being highly symmetric and elegant algebraic structures, and in part due to the interesting connections that exist between them. Their conceptually appealing properties make them important in the general study of quantum theory, encompassing both foundational matters and applications.

MUBs are central to the understanding of quantum complementarity and its many applications; see e.g. Ref. [10] for a review of MUBs. The former is manifested in the fact that MUBs give rise to the strongest entropic uncertainty relations (among projective measurements acting on a fixed dimension) [11]. Moreover, MUBs play a prominent role in quantum cryptography, where they are employed in many of the most well-known quantum key distribution protocols [12–16] as well as in secret sharing protocols [17–19]. In addition, complete sets of MUBs are known to be statistically optimal for quantum state tomography [20, 21]. Also, MUBs are instrumental for quantum random access coding [22–26]. Two

[1] By SIC (in singular) we refer to one set of symmetric informationally complete projectors. By SICs (in plural), we refer to all such sets.

other interesting applications are quantum error correction [27, 28] and entanglement detection in both high-dimensional and multipartite systems [29]. Notably, MUBs are also at the heart of the Mean King's Problem [30, 31]. Much attention has been directed at determining the number of MUBs that exist in general Hilbert space dimensions [10].

In a similar spirit, SICs are widely studied for both fundamental and practical reasons; see e.g. Ref. [32] for a recent review of SICs. This has motivated substantial research effort directed towards proving their existence in all Hilbert space dimensions (presently known, at least, up to dimension 121) [2, 3, 33, 34]. Every SIC can be suitably normalised such that it forms a single quantum measurement with d^2 outcomes. This is clearly not a projective measurement but a positive operator-valued measure (POVM) and hence we refer to the resulting object as a SIC-POVM. It has been shown that SIC-POVMs are optimal POVMs for (single-measurement) quantum state tomography [35–38]. Furthermore, they are useful for entanglement witnessing [39], have found applications in quantum key distribution [40, 41] and enable optimal random number generation from a singlet state [42]. In addition, SICs are at the heart of many protocols for certifying the non-projective nature of a measurement [43–46]. Moreover, SICs exhibit interesting connections to several areas of mathematics, for instance Lie and Jordan algebras [47] and algebraic number theory [48, 49].

Due to their highly symmetric properties and breadth of relevance, it is important to study the role of MUBs and SICs in the context of generating correlations that do not admit a classical description. The strongest form of such correlations are those that are nonlocal, i.e. correlations that violate a Bell inequality [50]. For instance, qubit MUBs occur commonly in the simplest Bell scenarios [51, 52] and SIC-POVMs were used to reveal the relevance of non-projective measurements in quantum nonlocality [53]. There is also an example of a Bell inequality in which three three-dimensional MUBs are required to produce the maximal quantum violation [54]. While attempts have been made at establishing more general relations between quantum nonlocality and MUBs [55, 56], results of substantial generality are lacking. Nevertheless, two questions appear particularly natural. Firstly, can one construct Bell inequalities in which MUBs and SICs of any given Hilbert space dimension generate the largest quantum violations? Secondly, and conversely, could one determine, by only observing some form of quantum nonlocality, that an initially uncharacterised measurement obeys some operational notion of mutual unbiasedness or symmetric informational completeness? While both these questions are foundationally important, positive answers would also pave the way for device-independent quantum information protocols for the many practical applications for which MUBs and SICs are desirable.

In this work we solve these challenges for both MUBs and SICs. We show how to construct Bell inequalities that are maximally violated in quantum theory using a maximally entangled state of local dimension d and, respectively, a pair of d-dimensional MUBs and a d-dimensional SIC. Then, we ask what can be inferred if the maximal Bell inequality violation is observed. In the case of MUBs, we show that the maximal quantum violation of the proposed Bell inequality implies that the measurements satisfy an operational definition of mutual unbiasedness, and that the shared state is essentially a maximally entangled state of local dimension d. Similarly, in the case of SICs, we find that the maximal quantum violation implies that the measurements satisfy an analogous operational definition of symmetric informational completeness.

Before proceeding any further let us explain how our results are related to the phenomenon of self-testing (rigidity), in which the unknown state and measurements are identified up to additional degrees of freedom, local isometries and possibly a transposition (see Ref. [57] for a review on self-testing). While the state certification for the MUB inequalities coincides with the notion used in self-testing, the conclusions we draw regarding the measurements constitute a weaker form of certification. To stress this point in this work we have chosen to consistently use the term "certification" over "self-testing".

Finally, we show that our Bell inequalities are useful in two practically relevant tasks. For the case of MUBs, we consider a scheme for device-independent quantum key distribution and prove a key rate of $\log d$ bits, which is optimal for any protocol that extracts key from a d-outcome measurement. By conducting numerical studies for the case of qutrit systems we show that the protocol is robust to noise. For SICs, we construct a scheme for device-independent random number generation. For two-dimensional SIC-POVMs, we obtain the largest amount of randomness possible for any protocol based on qubits. For three-dimensional SIC-POVMs, we obtain more randomness than can be obtained in any protocol based on projective measurements and quantum systems of dimension up to seven. In addition, we investigate the robustness of these schemes to noise.

II. BELL INEQUALITIES FOR MUTUALLY UNBIASED BASES

We present a family of Bell inequalities in which the maximal quantum violation is achieved with any pair of d-dimensional MUBs and a maximally entangled state. To this end, consider a bipartite Bell scenario parameterised by an integer $d \geq 2$. Alice randomly receives one of d^2 possible inputs labelled by $x \equiv x_1 x_2 \in [d]^2$ (where $[s] \equiv \{1, \ldots, s\}$) and produces a ternary output labelled by $a \in \{1, 2, \perp\}$. Bob receives a random binary input labelled by $y \in \{1, 2\}$ and produces a d-valued output labelled by $b \in [d]$. The joint probability distribution in the Bell scenario is denoted by $p(a, b|x, y)$ and the scenario is illustrated in Figure 1.

To make our choice of Bell functional transparent, we will phrase it as a game in which Alice and Bob collectively win or lose points. If Alice outputs $a = \perp$, no points will be won or lost. If she outputs $a \in \{1, 2\}$, points will be won or lost if $b = x_y$. More specifically, Alice and Bob win a point if $a = y$ and lose a point if $a = \bar{y}$, where the bar-sign flips the value of

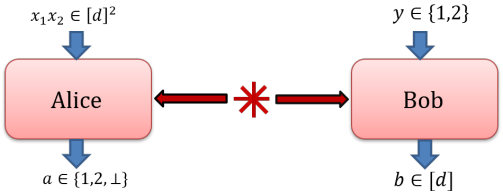

FIG. 1. Bell scenario for two MUBs of dimension d. Alice receives one of d^2 inputs and produces a ternary output while Bob receives a binary input and produces a d-valued output.

$y \in \{1,2\}$. This leads to the score

$$\mathcal{R}_d^{\text{MUB}} \equiv \sum_{x,y} p(a=y, b=x_y|x,y) - p(a=\bar{y}, b=x_y|x,y), \tag{3}$$

where the sum goes over $x = x_1x_2 \in [d]^2$ and $y \in \{1,2\}$.

At this point the outcome $a = \perp$ might seem artificial, so let us show why it plays a crucial role in the construction of the game. To this end, we use intuition based on the hypothetical case in which Alice and Bob share a maximally entangled state

$$|\psi_d^{\max}\rangle = \frac{1}{\sqrt{d}} \sum_{k=1}^{d} |k,k\rangle. \tag{4}$$

The reason we consider the maximally entangled state is that we aim to tailor the Bell inequalities so that this state is optimal. Then, we would like to ensure that Alice, via her measurement and for her outcomes $a \in \{1,2\}$, remotely prepares Bob in a pure state. This would allow Bob to create stronger correlations as compared to the case of Alice remotely preparing his system is a mixed state. Hence, this corresponds to Alice's outcomes $a \in \{1,2\}$ being represented by rank-one projectors. Since the subsystems of $|\psi_d^{\max}\rangle$ are maximally mixed, it follows that $p(a=1|x) = p(a=2|x) = 1/d \; \forall x$. Thus, we want to motivate Alice to employ a strategy in which she outputs $a = \perp$ with probability $p(a = \perp |x) = 1 - 2/d$. Our tool for this purpose is to introduce a penalty. Specifically, whenever Alice decides to output $a \in \{1,2\}$, she is penalised by losing γ_d points. Thus, the total score (the Bell functional) reads

$$\mathcal{S}_d^{\text{MUB}} \equiv \mathcal{R}_d^{\text{MUB}} - \gamma_d \sum_x \big(p(a=1|x) + p(a=2|x)\big). \tag{5}$$

Now, outputting $a \in \{1,2\}$ contributes towards $\mathcal{R}_d^{\text{MUB}}$ but also causes a penalty γ_d. Therefore, we expect to see a trade-off between γ_d and the rate at which Alice outputs $a = \perp$. We must suitably choose γ_d such that Alice's best strategy is to output $a = \perp$ with (on average over x) the desired probability $p(a = \perp |x) = 1 - 2/d$. This accounts for the intuition that leads us to the following Bell inequalities for MUBs.

Theorem II.1 (Bell inequalities for MUBs)**.** *The Bell functional $\mathcal{S}_d^{\text{MUB}}$ in Eq. (5) with*

$$\gamma_d = \frac{1}{2}\sqrt{\frac{d-1}{d}}, \tag{6}$$

obeys the tight local bound

$$\mathcal{S}_d^{\text{MUB}} \overset{\text{LHV}}{\leq} 2(d-1)\left(1 - \frac{1}{2}\sqrt{\frac{d-1}{d}}\right), \tag{7}$$

and the quantum bound

$$\mathcal{S}_d^{\text{MUB}} \overset{\text{Q}}{\leq} \sqrt{d(d-1)}. \tag{8}$$

Moreover, the quantum bound can be saturated by sharing a maximally entangled state of local dimension d and Bob performing measurements in any two mutually unbiased bases.

Proof. A complete proof is presented in Appendix A 1. The essential ingredient to obtain the bound in Eq. (8) is the Cauchy–Schwarz inequality. Furthermore, for local models, by inspecting the symmetries of the Bell functional $\mathcal{S}_d^{\text{MUB}}$, one finds that the local bound can be attained by Bob always outputting $b = 1$. This greatly simplifies the evaluation of the bound in Eq. (7).

To see that the bound in Eq. (8) can be saturated in quantum theory, let us evaluate the Bell functional for a particular quantum realisation. Let $|\psi\rangle$ be the shared state, $\{P_{x_1}\}_{x_1=1}^d$ and $\{Q_{x_2}\}_{x_2=1}^d$ be the measurement operators of Bob corresponding to $y = 1$ and $y = 2$ respectively and A_x be the observable of Alice defined as the difference between Alice's outcome-one and outcome-two measurement operators, i.e. $A_x = A_x^1 - A_x^2$. Then, the Bell functional reads

$$\mathcal{S}_d^{\text{MUB}} = \sum_x \langle\psi| A_x \otimes (P_{x_1} - Q_{x_2}) - \gamma_d \left(A_x^1 + A_x^2\right) \otimes \mathbb{1} |\psi\rangle. \tag{9}$$

Now, we choose the maximally entangled state of local dimension d, i.e. $|\psi\rangle = |\psi_d^{\max}\rangle$, and define Bob's measurements as rank-one projectors $P_{x_1} = |\phi_{x_1}\rangle\langle\phi_{x_1}|$ and $Q_{x_2} = |\varphi_{x_2}\rangle\langle\varphi_{x_2}|$ which correspond to MUBs, i.e. $|\langle\phi_{x_1}|\varphi_{x_2}\rangle|^2 = 1/d$. Finally, we choose Alice's observables as $A_x = \sqrt{d/(d-1)}(P_{x_1} - Q_{x_2})^{\text{T}}$, where the pre-factor ensures the correct normalisation and $^{\text{T}}$ denotes the transpose in the standard basis. Note that A_x is a rank-two operator; the corresponding measurement operator A_x^1 (A_x^2) is a rank-one projector onto the eigenvector of A_x associated to the positive (negative) eigenvalue. Since the subsystems of $|\psi_d^{\max}\rangle$ are maximally mixed, this implies $\langle\psi_d^{\max}|(A_x^1 + A_x^2) \otimes \mathbb{1}|\psi_d^{\max}\rangle = 2/d$. Inserting all this into the above quantum model and exploiting the fact that for any linear operator O we have $O \otimes \mathbb{1}|\psi_d^{\max}\rangle = \mathbb{1} \otimes O^{\text{T}}|\psi_d^{\max}\rangle$, we straightforwardly saturate the bound in Eq. (8). ■

We remark that for the case of $d = 2$ one could also choose $\gamma_2 = 0$ and retain the property that qubit MUBs are optimal. In this case the marginal term is not necessary, because in the

optimal realisation Alice never outputs $\perp$. Then, the quantum bound becomes $2\sqrt{2}$ and the local bound becomes 2. The resulting Bell inequality resembles the Clauser–Horne–Shimony–Holt (CHSH) inequality [51], not just because it gives the same local and quantum values, but also because the optimal realisations coincide. More specifically, the measurements of Bob are precisely the optimal CHSH measurements, whereas the four measurements of Alice correspond to two pairs of optimal CHSH measurements.

III. DEVICE-INDEPENDENT CERTIFICATION OF MUTUAL UNBIASEDNESS

Theorem II.1 establishes that a pair of MUBs of any dimension can generate a maximal quantum violation in a Bell inequality test. We now turn to the converse matter, namely that of device-independent certification. Specifically, given that we observe the maximal quantum violation, i.e. equality in Eq. (8), what can be said about the shared state and the measurements? Since the measurement operators can only be characterised on the support of the state, to simplify the notation let us assume that the marginal states of Alice and Bob are full-rank.

Theorem III.1 (Device-independent certification)**.** *The maximal quantum value of the Bell functional $\mathcal{S}_d^{\text{MUB}}$ in Eq. (5) implies that*

- *There exist local isometries which allow Alice and Bob to extract a maximally entangled state of local dimension d.*
- *Under the assumption that the marginal state of Bob is full-rank, the two d-outcome measurements he performs satisfy the relations*

$$P_a = dP_aQ_bP_a \qquad \text{and} \qquad Q_b = dQ_bP_aQ_b, \tag{10}$$

for all a and b.

Proof. The proof is detailed in Appendix A 2. Here, we briefly summarise the part concerning Bob's measurements. Since the Cauchy–Schwarz inequality is the main tool for proving the quantum bound in Eq. (8), saturating it implies that also the Cauchy–Schwarz inequality is saturated. This allows us to deduce that the measurements of Bob are projective and moreover we obtain the following optimality condition:

$$A_x \otimes \mathbb{1}|\psi\rangle = \mathbb{1} \otimes \sqrt{\frac{d}{d-1}}\,(P_{x_1} - Q_{x_2})\,|\psi\rangle, \tag{11}$$

for all $x_1, x_2 \in [d]$ where the factor $\sqrt{d/(d-1)}$ can be regarded as a normalisation. Since we do not attempt to certify the measurements of Alice, we can without loss of generality assume that they are projective. This implies that the spectrum of A_x only contains $\{+1, -1, 0\}$ and therefore $(A_x)^3 = A_x$. This allows us to obtain a relation that only contains Bob's operators. Tracing out Alice's system and subsequently eliminating the marginal state of Bob (it is assumed to be full-rank) leads to

$$P_{x_1} - Q_{x_2} = \frac{d}{d-1}\,(P_{x_1} - Q_{x_2})^3\,. \tag{12}$$

Expanding this relation and then using projectivity and the completeness of measurements, one recovers the result in Eq. (10). ■

We have shown that observing the maximal quantum value of $\mathcal{S}_d^{\text{MUB}}$ implies that the measurements of Bob satisfy the relations given in Eq. (10). It is natural to ask whether a stronger conclusion can be derived, but the answer turns out to be negative. In Appendix A 2 c we show that any pair of d-outcome measurements (acting on a finite-dimensional Hilbert space) satisfying the relations in Eq. (10) is capable of generating the maximal Bell inequality violation. For $d = 2, 3$ the relations given in Eq. (10) imply that the unknown measurements correspond to a direct sum of MUBs (see Appendix B 3 a) and since in these dimension there exists only a single pair of MUBs (up to unitaries and complex conjugation), our results imply a self-testing statement of the usual kind. However, since in higher dimensions not all pairs of MUBs are equivalent [58], our certification statement is less informative than the usual formulation of self-testing. In other words, our inequalities allow us to self-test the quantum state, but we cannot completely determine the measurements (see Refs. [59, 60] for related results). Note that we could also conduct a device-independent characterisation of the measurements of Alice, but since these are not relevant for the scope of this work (namely MUBs and SICs), we do not do it explicitly.

The certification provided in Theorem III.1 turns out to be sufficient to determine all the probabilities $p(a, b|x, y)$ that arise in the Bell experiment (see Appendix A 3), which means that the maximal quantum value of $\mathcal{S}_d^{\text{MUB}}$ is achieved by a single probability distribution. Due to the existence of inequivalent pairs of MUBs in certain dimensions (e.g. for $d = 4$), this constitutes the first example of an extremal point of the quantum set which admits inequivalent quantum realisations.[2]

It is important to understand the relation between the condition given in Eq. (10) and the concept of MUBs. Naturally, if $\{P_a\}_{a=1}^d$ and $\{Q_b\}_{b=1}^d$ are d-dimensional MUBs, the relations (10) are satisfied. Interestingly, however, there exist solutions to Eq. (10) which are neither MUBs nor direct sums thereof. While, as mentioned above, for $d = 2, 3$ one can show that any measurements satisfying the relations (10) must correspond to a direct sum of MUBs, this is not true in general. For $d = 4, 5$ we have found explicit examples of measurement operators satisfying Eq. (10) which cannot be written as a direct sum of MUBs. In fact, they cannot even be transformed into a pair of MUBs via a completely positive unital map (see Appendix B for details). These results beg the crucial question: how should one interpret the condition given in Eq. (10)?

[2] Recall that the notion of equivalence we employ is precisely the one that appears in the context of self-testing, i.e. we allow for additional degrees of freedom, local isometries and a transposition.

To answer this question we resort to an operational formulation of what it means for two measurements to be mutually unbiased. An operational approach must rely on observable quantities (i.e. probabilities), as opposed to algebraic relations between vectors or operators. This leads to the following natural definition of mutually unbiased measurements (MUMs)[3].

Definition III.2 (Mutually unbiased measurements)**.** *We say that two n-outcome measurements $\{P_a\}_{a=1}^n$ and $\{Q_b\}_{b=1}^n$ are mutually unbiased if they are projective and the following implications hold:*

$$\begin{aligned}\langle\psi|P_a|\psi\rangle = 1 \Rightarrow \langle\psi|Q_b|\psi\rangle = \frac{1}{n}\\ \langle\psi|Q_b|\psi\rangle = 1 \Rightarrow \langle\psi|P_a|\psi\rangle = \frac{1}{n},\end{aligned} \tag{13}$$

for all a and b. That is, two projective measurements are mutually unbiased if the eigenvectors of one measurement give rise to a uniform outcome distribution for the other measurement.

Note that this definition captures precisely the intuition behind MUBs without the need to specify the dimension of the underlying Hilbert space. Interestingly enough, MUMs admit a simple algebraic characterisation.

Theorem III.3. *Two n-outcome measurements $\{P_a\}_{a=1}^n$ and $\{Q_b\}_{b=1}^n$ are mutually unbiased if and only if*

$$P_a = nP_aQ_bP_a \qquad \textit{and} \qquad Q_b = nQ_bP_aQ_b, \tag{14}$$

for all a and b.

Proof. Let us first assume that the algebraic relations hold. By summing over the middle index, one finds that both measurements are projective. Moreover, if $|\psi\rangle$ is an eigenvector of P_a, then $\langle\psi|Q_b|\psi\rangle = \langle\psi|P_aQ_bP_a|\psi\rangle = \frac{1}{n}\langle\psi|P_a|\psi\rangle = \frac{1}{n}$. By symmetry, the analogous property holds if $|\psi\rangle$ is an eigenvector of Q_b.

Conversely, let us show that MUMs must satisfy the above algebraic relations. Since $\sum_a P_a = \mathbb{1}$ we can choose an orthonormal basis of the Hilbert space composed only of the eigenvectors of the measurement operators. Let $\{|e_j^a\rangle\}_{a,j}$ be an orthonormal basis, where $a \in [n]$ tells us which projector the eigenvector corresponds to and j labels the eigenvectors within a fixed projector (if P_a has finite rank, then $j \in [\operatorname{tr} P_a]$, otherwise $j \in \mathbb{N}$). By construction for such a basis we have $P_a|e_j^{a'}\rangle = \delta_{aa'}|e_j^a\rangle$. To show that $P_a = nP_aQ_bP_a$ it suffices to show that the two operators have the same coefficients in this basis. Since

$$\langle e_j^{a'}|nP_aQ_bP_a|e_k^{a''}\rangle = n\delta_{aa'}\delta_{aa''}\langle e_j^a|Q_b|e_k^a\rangle, \tag{15}$$

$$\langle e_j^{a'}|P_a|e_k^{a''}\rangle = \delta_{aa'}\delta_{aa''}\delta_{jk} \tag{16}$$

it suffices to show that $n\langle e_j^a|Q_b|e_k^a\rangle = \delta_{jk}$. For $j = k$ this is a direct consequence of the definition in Eq. (13). To prove the other case, define $|\phi_\theta\rangle = \left(|e_j^a\rangle + \mathrm{e}^{\mathrm{i}\theta}|e_k^a\rangle\right)/\sqrt{2}$, for $\theta \in [0, 2\pi)$. Since $P_a|\phi_\theta\rangle = |\phi_\theta\rangle$, we have $\langle\phi_\theta|Q_b|\phi_\theta\rangle = 1/n$. Writing this equality out gives

$$\frac{1}{n} = \frac{1}{2}\left(\frac{2}{n} + \mathrm{e}^{\mathrm{i}\theta}\langle e_j^a|Q_b|e_k^a\rangle + \mathrm{e}^{-\mathrm{i}\theta}\langle e_k^a|Q_b|e_j^a\rangle\right). \tag{17}$$

Choosing $\theta = 0$ implies that the real part of $\langle e_j^a|Q_b|e_k^a\rangle$ vanishes, while $\theta = \pi/2$ implies that the imaginary part vanishes. Proving the relation $Q_b = nQ_bP_aQ_b$ proceeds in an analogous fashion. ∎

Theorem III.3 implies that the maximal violation of the Bell inequality for MUBs certifies precisely the fact the Bob's measurements are mutually unbiased. To provide further evidence that MUMs constitute the correct device-independent generalisation of MUBs, we give two specific situations in which the two objects behave in the same manner.

Maassen and Uffink considered a scenario in which two measurements (with a finite number of outcomes) are performed on an unknown state. Their famous uncertainty relation provides a state-independent lower bound on the sum of the Shannon entropies of the resulting distributions [11]. While the original result only applies to rank-one projective measurements, a generalisation to non-projective measurements reads [61]

$$H(P) + H(Q) \geq -\log c, \tag{18}$$

where H denotes the Shannon entropy and $c = \max_{a,b}\|\sqrt{P_a}\sqrt{Q_b}\|^2$ where $\|\cdot\|$ is the operator norm. If we restrict ourselves to rank-one projective measurements on a Hilbert space of dimension d, one finds that the largest uncertainty, corresponding to $c = 1/d$, is obtained only by MUBs. It turns out that precisely the same value is achieved by any pair of MUMs with d outcomes regardless of the dimension of the Hilbert space:

$$\begin{aligned}c = \max_{a,b}\|\sqrt{P_a}\sqrt{Q_b}\|^2 = \max_{a,b}\|P_aQ_b\|^2\\ = \max_{a,b}\|P_aQ_bP_a\| = \max_a\|P_a/d\| = \frac{1}{d}.\end{aligned} \tag{19}$$

A closely related concept is that of measurement incompatibility, which captures the phenomenon that two measurements cannot be performed simultaneously on a single copy of a system. The extent to which two measurements are incompatible can be quantified e.g. by so-called incompatibility robustness measures [62]. In Appendix B 4, we show that according to these measures MUMs are exactly as incompatible as MUBs. Moreover, we can show that for the so-called generalised incompatibility robustness [63], MUMs are among the most incompatible pairs of d-outcome measurements.

IV. APPLICATION: DEVICE-INDEPENDENT QUANTUM KEY DISTRIBUTION

The fact that the maximal quantum violation of the Bell inequalities introduced above requires a maximally entangled

[3] Note that in what follows we use the term "eigenvector" to refer to eigenvectors corresponding to non-zero eigenvalues.

state and MUMs, and moreover that it is achieved by a unique probability distribution, suggests that these inequalities might be useful for device-independent quantum information processing. In the task of quantum key distribution [12, 13, 64] Alice and Bob aim to establish a shared data set (a key) that is secure against a malicious eavesdropper. Such a task requires the use of incompatible measurements, and MUBs in dimension $d = 2$ constitute the most popular choice. Since in the ideal case the measurement outcomes of Alice and Bob that contribute to the key should be perfectly correlated, most protocols are based on maximally entangled states. In the device-independent approach to quantum key distribution, the amount of key and its security is deduced from the observed Bell inequality violation.

We present a proof-of-principle application to device-independent quantum key distribution based on the quantum nonlocality witnessed through the Bell functional in Eq. (5). In the ideal case, Alice and Bob follow the strategy that gives them the maximal violation, i.e. they share a maximally entangled state of local dimension d and Bob measures two MUBs. To generate the key we provide Alice with an extra setting that produces outcomes which are perfectly correlated with the outcomes of the first setting of Bob. This will be the only pair of settings from which the raw key will be extracted and let us denote them by $x = x^*$ and $y = y^* = 1$. In most rounds of the experiment, Alice and Bob choose these settings and therefore contribute towards the raw key. However, to ensure security, a small number of rounds is used to evaluate the Bell functional. In these rounds, which are chosen at random, Alice and Bob randomly choose their measurement settings. Once the experiment is complete, the resulting value of the Bell functional is used to infer the amount of secure raw key shared between Alice and Bob. The raw key can then be turned into the final key by standard classical post-processing. For simplicity, we consider only individual attacks and moreover we focus on the limit of asymptotically many rounds in which fluctuations due to finite statistics can be neglected.

The key rate, K, can be lower bounded by [65]

$$K \geq -\log\left(P_g^\beta\right) - H(B_{y^*}|A_{x^*}), \tag{20}$$

where P_g^β denotes the highest probability that the eavesdropper can correctly guess Bob's outcome when his setting is y^* given that the Bell inequality value β was observed, and $H(\cdot|\cdot)$ denotes the conditional Shannon entropy. The guessing probability P_g^β is defined as

$$P_g^\beta \equiv \sup\left\{\sum_{c=1}^{d}\langle\psi_{\text{ABE}}|\mathbb{1}\otimes P_c\otimes E_c|\psi_{\text{ABE}}\rangle\right\}, \tag{21}$$

where $\{E_c\}_{c=1}^d$ is the measurement employed by the eavesdropper to produce her guess, the expression inside the curly braces is the probability that her outcome is the same as Bob's for a particular realisation and the supremum is taken over all quantum realisations (the tripartite state and measurements of all three parties) compatible with the observed Bell inequality value β.

Let us first focus on the key rate in a noise-free scenario, i.e. in a scenario in which $\mathcal{S}_d^{\text{MUB}}$ attains its maximal value. Then, one straightforwardly arrives at the following result.

Theorem IV.1 (Device-independent key rate)**.** *In the noiseless case the quantum key distribution protocol based on $\mathcal{S}_d^{\text{MUB}}$ achieves the key rate of*

$$K = \log d \tag{22}$$

for any integer $d \geq 2$.

Proof. In the noiseless case, Alice and Bob observe exactly the correlations predicted by the ideal setup. In this case the outcomes for settings (x^*, y^*) are perfectly correlated which implies that $H(B_{y^*}|A_{x^*}) = 0$. Therefore, the only non-trivial task is to bound the guessing probability.

Since the actions of the eavesdropper commute with the actions of Alice and Bob, we can assume that she performs her measurement first. If the probability of the eavesdropper observing outcome $c \in [d]$, which we denote by $p(c)$, is non-zero, then the (normalised) state of Alice and Bob conditioned on the eavesdropper observing that outcome is given by:

$$\rho_{\text{AB}}^{(c)} = \frac{1}{p(c)}\,\text{tr}_C\left[(\mathbb{1}\otimes\mathbb{1}\otimes E_c)|\psi_{\text{ABE}}\rangle\langle\psi_{\text{ABE}}|\right]. \tag{23}$$

Now Alice and Bob share one of the post-measurement states $\rho_{\text{AB}}^{(c)}$ and when they perform their Bell inequality test, they will obtain different distributions depending on c, which we write as $p_c(a, b|x, y)$. However, since the statistics achieve the maximal quantum value of $\mathcal{S}_d^{\text{MUB}}$ and we have previously shown that the maximal quantum value is achieved by a single probability point, all the probability distributions $p_c(a, b|x, y)$ must be the same. Moreover, we have shown that for this probability point, the marginal distribution of outcomes on Bob's side is uniform over $[d]$ for both inputs. This implies that

$$P_g = \sum_{c=1}^{d} p(c)p_c(b = c|y = 1) = \frac{1}{d}, \tag{24}$$

because $p_c(b = c|y = 1) = p(b = c|y = 1) = \frac{1}{d}$ for all c. ■

We remark that the argument above is a direct consequence of a more general result which states that if a bipartite probability distribution is a nonlocal extremal point of the quantum set, then no external party can be correlated with the outcomes [66].

It is interesting to note that the obtained key rate is the largest possible for general setups in which the key is generated from a d-outcome measurement. Also, the key rate is optimal for all protocols based on a pair of entangled d-dimensional systems subject to projective measurements. This follows from the fact that projective measurements in $\mathbb{C}^d$ cannot have more than d outcomes. Note that it has recently been shown that the same amount of randomness can be generated using a modified version of the Collins–Gisin–Linden–Massar–Popescu inequalities [67].

Let us now depart from the noise-free case and estimate the key rate in the presence of noise. To ensure that both the

guessing probability and the conditional Shannon entropy can be computed in terms of a single noise parameter, we have to introduce an explicit noise model. We employ the standard approach in which the measurements remain unchanged, while the maximally entangled state is replaced with an isotropic state given by

$$\rho_v = v|\psi_d^{\max}\rangle\langle\psi_d^{\max}| + \frac{1-v}{d^2}\mathbb{1}, \tag{25}$$

where $v \in [0,1]$ is the visibility of the state. Using this state and the ideal measurements for Alice and Bob, the relation between v and $\mathcal{S}_d^{\text{MUB}}$ can be easily derived from (9), namely,

$$v = \frac{1}{2}\left(1 + \frac{\mathcal{S}_d^{\text{MUB}}}{\sqrt{d(d-1)}}\right). \tag{26}$$

Utilising this formula, we also obtain the value of $H(B_{y^*}|A_{x^*})$ as a function of the Bell violation. The remaining part of (20) is the guessing probability (21). In the case of $d = 3$, we proceed to bound this quantity through semidefinite programming.

Concretely, we implement the three-party semidefinite relaxation [68] of the set of quantum correlations at local level one[4]. This results in a moment matrix of size 532×532 with 15617 variables. The guessing probability is directly given by the sum of three elements of the moment matrix. It can then be maximised under the constraints that the value of the Bell functional $\mathcal{S}_3^{\text{MUB}}$ is fixed and the moment matrix is positive semidefinite. However, we notice that this problem is invariant under the following relabelling: $b \to \pi(b)$ for $y = 1$, $c \to \pi(c)$, $x_1 \to \pi(x_1)$, where $\pi \in S_3$ is a permutation of three elements. Therefore, it is possible to simplify this semidefinite program by requiring the matrix to be invariant under the group action of S_3 on the moment matrix (i.e. it is a Reynolds matrix) [43, 69, 70]. This reduces the number of free variables in the moment matrix to 2823. With the SeDuMi [71] solver, this lowers the precision (1.1×10^{-6} instead of 8.4×10^{-8}), but speeds up the computation (155s instead of 8928s) and requires less memory (0.1GB instead of 5.5GB). For the maximal value of $\mathcal{S}_d^{\text{MUB}}$, we recover the noise-free result of $K = \log 3$ up to the fifth digit. Also, we have a key rate of at least one bit when $\mathcal{S}_d^{\text{MUB}} \gtrsim 2.432$ and a non-zero key rate when $\mathcal{S}_d^{\text{MUB}} \gtrsim 2.375$. The latter is close to the local bound, which is $\mathcal{S}_d^{\text{MUB}} \approx 2.367$. The resulting lower bound on the key rate as a function of the Bell inequality violation is plotted in Fig. 2.

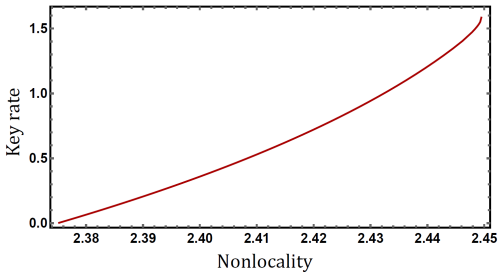

FIG. 2. Lower bound on the key rate K in the asymptotic limit versus the value of the Bell functional $\mathcal{S}_3^{\text{MUB}}$.

V. NONLOCALITY FOR SYMMETRIC INFORMATIONAL COMPLETENESS

We now shift our focus from MUBs to SICs. We construct Bell inequalities whose maximal quantum violations are achieved with SICs. Since this turns out to be more challenging than for the case of MUBs, we first establish the relevance of SICs in a simplified Bell scenario subject to additional constraints. This serves as a stepping stone to a subsequent relaxation which gives a standard (unconstrained) Bell inequality for SICs. We then focus on the device-independent certification power of these inequalities, which leads us to an operational notion of symmetric informational completeness. Finally, we extend the Bell inequalities so that their maximal quantum violations are achieved with both projectors forming SICs and a single generalised measurement corresponding to a SIC-POVM.

A. Stepping stone: quantum correlations for SICs

Consider a Bell scenario, parameterised by an integer $d \geq 2$, involving two parties Alice and Bob who share a physical system. Alice receives an input labelled by a tuple (x_1, x_2) representing one of $\binom{d^2}{2}$ possible inputs, which we collectively refer to as $x = x_1x_2$. The tuple is randomly taken from the set $\text{Pairs}(d^2) \equiv \{x|x_1, x_2 \in [d^2] \text{ and } x_1 < x_2\}$. Alice performs a measurement on her part of the shared system and produces a ternary output labelled by $a \in \{1, 2, \perp\}$. Bob receives an input labelled by $y \in [d^2]$ and the associated measurement produces a binary outcome labelled by $b \in \{1, \perp\}$. The joint probability distribution is denoted by $p(a, b|x, y)$, and the Bell scenario is illustrated in Fig. 3.

Similar to the case of MUBs, in order to make our choice of Bell functional transparent, we phrase it as a game played by Alice and Bob. We imagine that their inputs are supplied by a referee, who promises to provide $x = x_1x_2$ and y such that either $y = x_1$ or $y = x_2$. Similar to the previous game Alice can output $a = \perp$ to ensure that no points are won or lost. However, in this game also Bob can ensure that no points are won or lost by outputting $b = \perp$. If neither of them outputs $\perp$, a point is either won or lost. Specifically, when $a = 1$ a point is won if $y = x_1$ (and lost otherwise), whereas if $a = 2$ then a point is won if $y = x_2$ (and lost otherwise). Let us remark that in this game Bob's only role is to decide whether in a given round points can be won/lost or not. For this game the total

[4] We attribute one operator to each outcome of Bob and the eavesdropper, but only take into account the first two outcomes of Alice.

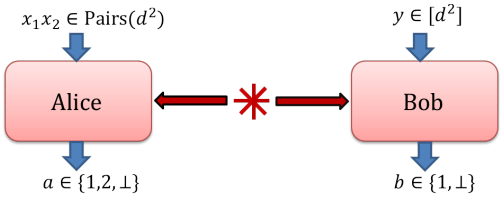

FIG. 3. Bell scenario for SICs of dimension d. Alice receives one of $\binom{d^2}{2}$ inputs and returns a ternary outcome while Bob receives one of d^2 inputs and returns a binary outcome.

number of points (the Bell functional) reads

$$\mathcal{R}_d^{\text{SIC}} \equiv \sum_{x_1<x_2} \Big(p(1,1|x,x_1) - p(1,1|x,x_2) + p(2,1|x,x_2) - p(2,1|x,x_1) \Big), \tag{27}$$

where the sum is taken over all $x \in \text{Pairs}(d^2)$.

Let us now impose additional constraints on the marginal distributions of the outputs. More specifically, we require that

$$\begin{aligned} &\forall x: \quad p(a=1|x) + p(a=2|x) = \frac{2}{d}, \\ &\forall y: \quad p(b=1|y) = \frac{1}{d}. \end{aligned} \tag{28}$$

The intuition behind these constraints is analogous to that discussed for the case of MUBs. Namely, we imagine that Alice and Bob perform measurements on a maximally entangled state of local dimension d. Then, we wish to fix the marginals such that the measurements of Alice (Bob) for the outcomes $a \in \{1,2\}$ ($b = 1$) remotely prepare Bob's (Alice's) subsystem in a pure state. This corresponds to the marginals $p(a=1|x) = p(a=2|x) = p(b=1|x) = 1/d$ which is reflected in the marginal constraints in Eq. (28). We remark that imposing these constraints simplifies both the intuitive understanding of the game and the derivation of the results below. However, it merely serves as a stepping stone to a more general subsequent treatment in which the constraints (28) will be removed.

To write the value of the Bell functional of a quantum realisation, let us introduce two simplifications. The measurement operators of Alice are denoted by $\{A_x^a\}$ and as before it is convenient to work with the observables defined as $A_x = A_x^1 - A_x^2$. The measurements of Bob are denoted by $\{B_y^b\}$, but since they only have two outcomes, all the expressions can be written in terms of a single operator from each input y. In our case it is convenient to use the outcome-one operator and for convenience we will skip the superscript, i.e. we will write $B_y \equiv B_y^1$ for all y. Then, the Bell functional evaluated on a specific quantum realisation reads

$$\mathcal{R}_d^{\text{SIC}} = \sum_{x_1<x_2} \langle\psi| A_x \otimes (B_{x_1} - B_{x_2}) |\psi\rangle. \tag{29}$$

Note that the Bell functional, in particular when written in a quantum model, is much reminiscent of the expression $\mathcal{R}_d^{\text{MUB}}$ (3) encountered for MUBs, with the key difference that the roles of the inputs and outputs of Bob are swapped.

Let us consider a quantum strategy in which Alice and Bob share a maximally entangled state $|\psi_d^{\max}\rangle$. Moreover, Bob's measurements are defined as $B_y = |\phi_y\rangle\langle\phi_y|$, where $\{|\phi_y\rangle\}_{y=1}^{d^2}$ is a set of unit vectors forming a SIC (assuming it exists in dimension d), i.e. $|\langle\phi_y|\phi_{y'}\rangle|^2 = 1/(d+1)$ for all $y \neq y'$. Also, we define Alice's observables as $A_x = \sqrt{(d+1)/d}\,(B_{x_1} - B_{x_2})^{\text{T}}$, where the pre-factor ensures normalisation. Firstly, since the subsystems of Alice and Bob are maximally mixed, and the outcomes $a \in \{1,2\}$ and $b = 1$ each correspond to rank-one projectors, the marginal constraints in Eq. (28) are satisfied. Using the fact that for any linear operator O we have $O \otimes \mathbb{1}|\psi_d^{\max}\rangle = \mathbb{1} \otimes O^{\text{T}}|\psi_d^{\max}\rangle$, we find that

$$\begin{aligned} &\mathcal{R}_d^{\text{SIC}} = \\ &\sqrt{\frac{d+1}{d}} \sum_{x_1<x_2} \langle\psi_d^{\max}| \mathbb{1} \otimes (|\phi_{x_1}\rangle\langle\phi_{x_1}| - |\phi_{x_2}\rangle\langle\phi_{x_2}|)^2 |\psi_d^{\max}\rangle \\ &= \sqrt{\frac{d+1}{d}} \sum_{x_1<x_2} \left(\frac{2}{d} - \frac{2}{d(d+1)} \right) = d(d-1)\sqrt{d(d+1)}. \end{aligned} \tag{30}$$

In fact, this strategy relying on a maximally entangled state and a SIC achieves the maximal quantum value of $\mathcal{R}_d^{\text{SIC}}$ under the constraints of Eq. (28). In Appendix C 1 we prove that under these constraints the tight quantum and no-signaling bounds on $\mathcal{R}_d^{\text{SIC}}$ read

$$\mathcal{R}_d^{\text{SIC}} \overset{\text{Q}}{\leq} d(d-1)\sqrt{d\,(d+1)} \tag{31}$$

$$\mathcal{R}_d^{\text{SIC}} \overset{\text{NS}}{\leq} d\left(d^2-1\right). \tag{32}$$

We remark that SICs are not known to exist in all Hilbert space dimensions. However, their existence in all dimensions is strongly conjectured and explicit SICs have been found in all dimensions up to 121 [34].

B. Bell inequalities for SICs

The marginal constraints in Eq. (28) allowed us to prove that the quantum realisation based on SICs achieves the maximal quantum value of $\mathcal{R}_d^{\text{SIC}}$. Our goal now is to remove these constraints to obtain a standard Bell functional. Analogously to the case of MUBs we add marginal terms to the original functional $\mathcal{R}_d^{\text{SIC}}$.

To this end, we introduce penalties for both Alice and Bob. Specifically, if Alice outputs $a \in \{1,2\}$ they lose α_d points, whereas if Bob outputs $b = 1$, they lose β_d points. The total number of points in the modified game constitutes our final

Bell functional

$$\mathcal{S}_d^{\text{SIC}} \equiv \mathcal{R}_d^{\text{SIC}} - \alpha_d \sum_{x_1<x_2} \left(p\left(a=1|x\right)+p\left(a=2|x\right)\right) - \beta_d \sum_y p\left(b=1|y\right). \quad (33)$$

Hence, our aim is to suitably choose the penalties α_d and β_d so that the maximal quantum value of $\mathcal{S}_d^{\text{SIC}}$ is achieved with a strategy that closely mimics the marginal constraints (28) and thus maintains the optimality of Bob performing a SIC.

Theorem V.1 (Bell inequalities for SICs)**.** *The Bell functional $\mathcal{S}_d^{\text{SIC}}$ in Eq. (33) with*

$$\begin{aligned} \alpha_d &= \frac{1-\delta_{d,2}}{2}\sqrt{\frac{d}{d+1}} \\ \beta_d &= \frac{d-2}{2}\sqrt{d(d+1)}, \end{aligned} \quad (34)$$

obeys the tight local bound

$$\mathcal{S}_d^{\text{SIC}} \overset{\text{LHV}}{\leq} \begin{cases} 4 & \text{for} \quad d=2, \\ d^2(d-1)-d(d^2-d-1)\sqrt{\frac{d}{d+1}} & \text{for} \quad d\geq 3, \end{cases} \quad (35)$$

and the quantum bound

$$\mathcal{S}_d^{\text{SIC}} \overset{\text{Q}}{\leq} \frac{d+2\delta_{d,2}}{2}\sqrt{d\left(d+1\right)}. \quad (36)$$

Moreover, the quantum bound is tight and can be saturated by sharing a maximally entangled state of local dimension d and choosing Bob's outcome-one projectors to form a SIC.

Proof. The proof is presented in Appendix C 2. In order to obtain the quantum bound in Eq. (36), the key ingredients are the Cauchy–Schwarz inequality and semidefinite relaxations of polynomial optimisation problems. To derive the local bound in Eq. (35), the key observation is that the symmetries of the Bell functional allow us to significantly simplify the problem.

The fact that the quantum bound is saturated by a maximally entangled state and Bob performing a SIC can be seen immediately from the previous discussion that led to Eq. (30). With that strategy, we find $\mathcal{R}_d^{\text{SIC}} = d(d-1)\sqrt{d(d+1)}$. Since it also respects $p(a=1|x)+p(a=2|x) = 2/d\;\forall x$, as well as $p(b=1|y) = 1/d\;\forall y$, a direct insertion into Eq. (33) saturates the bound in Eq. (36). ∎

Note that in the limit of $d\to\infty$ both the local bound and the quantum bound grow as $\sim d^2$.

We remark that for the special case of $d=2$, no penalties are needed to maintain the optimality of SICs (which is why the delta function appears in Eq. (34)). The derived Bell inequality for a qubit SIC (which corresponds to a tetrahedron configuration on the Bloch sphere) can be compared to the so-called elegant Bell inequality [52] whose maximal violation is also achieved using the tetrahedron configuration. While we require six settings of Alice and four settings of Bob, the elegant Bell inequality requires only four settings of Alice and three settings of Bob. However, the additional complexity in our setup carries an advantage when considering the critical visibility of the shared state; i.e. the smallest value of v in Eq. (25) (defining an isotropic state) for which the Bell inequality is violated. The critical visibility for violating the elegant Bell inequality is 86.6%, whereas for our Bell inequality it is lowered to 81.6%. We remark that on the Bloch sphere, the anti-podal points corresponding to the four measurements of Alice and the six measurements of Bob form a cube and a cuboctahedron respectively, which constitutes an instance of the type of Bell inequalities proposed in Ref. [72].

C. Device-independent certification

Theorem V.1 shows that for any dimension $d\geq 2$ we can construct a Bell inequality which is maximally violated by a SIC in that dimension (provided a SIC exists). Let us now consider the converse question, namely that of device-independent certification. In analogy with the case of MUBs (Eq. (10)), we find a simple description of Bob's measurements.

Theorem V.2 (Device-independent certification)**.** *The maximal quantum value of the Bell functional $\mathcal{S}_d^{\text{SIC}}$, provided the marginal state of Bob is full-rank, implies that his measurement operators $\{B_y\}_{y=1}^{d^2}$ are projective and satisfy*

$$\sum_y B_y = d\,\mathbb{1} \quad (37)$$

and

$$B_y = (d+1)B_y B_{y'} B_y \quad (38)$$

for all $y\neq y'$.

A complete proof, which is similar in spirit to the proof of Theorem III.1, can be found in Appendix C 3.

For the special case of $d=2$, the conclusion can be made even more accurate: the maximal quantum violation of $\mathcal{S}_2^{\text{SIC}}$ implies that Bob's outcome-one projectors are rank-one projectors acting on a qubit whose Bloch vectors form a regular tetrahedron (up to the three standard equivalences used in self-testing).

Similar to the case of MUBs, we face the key question of interpreting the condition in Eq. (38) and its relation to SICs. Again in analogy with the case of MUBs, we note that the concept of a SIC references the dimension of the Hilbert space, which should not appear explicitly in a device-independent scenario. Hence we consider an operational approach to SICs, which must rely on observable quantities (i.e. probabilities). This leads us to the following natural definition of a set of projectors being *operationally symmetric informationally complete* (OP-SIC).

Definition V.3 (Operational SIC)**.** *We say that a set of projectors $\{B_a\}_{a=1}^{n^2}$ is operationally symmetric informationally complete (OP-SIC) if*

$$\sum_a B_a = n\,\mathbb{1} \quad (39)$$

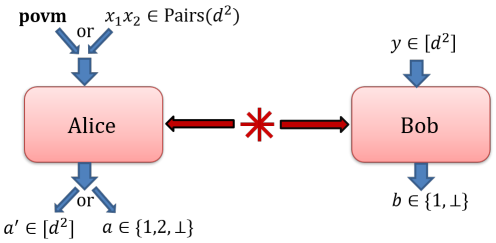

FIG. 4. Bell scenario for SICs and SIC-POVMs of dimension d. This scenario modifies the original Bell scenario for SICs (see Figure 3) by supplying Alice with an extra setting labelled by $\mathbf{povm}$ which has d^2 possible outcomes.

and

$$\langle\psi|B_a|\psi\rangle = 1 \Rightarrow \langle\psi|B_b|\psi\rangle = \frac{1}{n+1}, \tag{40}$$

for all $a \neq b$.

This definition trivially encompasses SICs as special instances of OP-SICs. More interestingly, an argument analogous to the proof of Theorem III.3 shows that this definition is in fact equivalent to the relations given in Eqs. (37) and (38). Hence, in analogy with the case of MUBs, the property of Bob's measurements certified by the maximal violation of our Bell inequality is precisely the notion of OP-SICs.

D. Adding a SIC-POVM

The Bell inequalities proposed above (Bell functional $\mathcal{S}_d^{\text{SIC}}$) are tailored to sets of rank-one projectors forming a SIC. However, it is also interesting to consider a closely related entity, namely a SIC-POVM, which is obtained simply by normalising these projectors, so that they can be collectively interpreted as arising from a single measurement. That is, a SIC-POVM on $\mathbb{C}^d$ is a measurement $\{E_a\}_{a=1}^{d^2}$ in which every measurement operator can be written as $E_a = \frac{1}{d}|\phi_a\rangle\langle\phi_a|$, where the set of rank-one projectors $\{|\phi_a\rangle\langle\phi_a|\}_a$ forms a SIC. Due to the simple relation between SICs and SIC-POVMs, we can extend the Bell inequalities for SICs proposed above such that they are optimally implemented with both a SIC (as before) and a SIC-POVM.

It is clear that in order to make SIC-POVMs relevant to the Bell experiment, it must involve at least one setting which corresponds to a d^2-outcome measurement. For the Bell scenario previously considered for SICs (see Figure 3), no such measurement is present. Therefore, we supplement the original Bell scenario by introducing a single additional measurement setting of Alice, labelled by $\mathbf{povm}$, which has d^2 outcomes labelled by $a' \in [d^2]$. The modified Bell scenario is illustrated in Figure 4. We construct the Bell functional $\mathcal{T}_d^{\text{SIC}}$ for this scenario by modifying the previously considered Bell functional $\mathcal{S}_d^{\text{SIC}}$:

$$\mathcal{T}_d^{\text{SIC}} = \mathcal{S}_d^{\text{SIC}} - \sum_{y=1}^{d^2} p(a'=y, b=\perp|\mathbf{povm}, y). \tag{41}$$

Hence, whenever Bob outputs "$\perp$" and the outcome associated to the setting $\mathbf{povm}$ coincides with the input of Bob, a point is lost.

Evidently, the largest quantum value of $\mathcal{T}_d^{\text{SIC}}$ is no greater than the largest quantum value of $\mathcal{S}_d^{\text{SIC}}$. In order for the former to equal the latter, we require that: i) $\mathcal{S}_d^{\text{SIC}}$ reaches its maximal quantum value (which is given in Eq. (36)) and ii) that $p(a'=y, b=\perp|\mathbf{povm}, y) = 0\ \forall y$. We have already seen that by sharing a maximally entangled state and Bob's outcome-one projectors $\{B_y\}_y$ forming a SIC, the condition i) can be satisfied. By normalisation, we have that Bob's outcome-$\perp$ projectors are $B_y^\perp = \mathbb{1} - B_y$. Again noting that for any linear operator O we have $O \otimes \mathbb{1}|\psi_d^{\max}\rangle = \mathbb{1} \otimes O^{\text{T}}|\psi_d^{\max}\rangle$, observe that if Bob applies $B_y^\perp$, then Alice's local state is orthogonal to B_y. Hence, if Alice chooses her POVM $\{E_{a'}\}$, corresponding to the setting $\mathbf{povm}$, as the SIC-POVM defined by $E_{a'} = \frac{1}{d}B_{a'}^{\text{T}}$, the probability of finding $a' = y$ vanishes. This satisfies condition ii). Hence, we conclude that in a general quantum model

$$\mathcal{T}_d^{\text{SIC}} \overset{\text{Q}}{\leq} \frac{d + 2\delta_{d,2}}{2}\sqrt{d\,(d+1)}, \tag{42}$$

and that the bound can be saturated by supplementing the previous optimal realisation with a SIC-POVM on Alice's side.

VI. APPLICATION: DEVICE-INDEPENDENT QUANTUM RANDOM NUMBER GENERATION

The fact that the Bell functionals $\mathcal{S}_d^{\text{SIC}}$ and $\mathcal{T}_d^{\text{SIC}}$ achieve their maximal quantum values with a SIC and a SIC-POVM respectively, opens up the possibility for device-independent quantum information protocols for tasks in which SICs and SIC-POVMs are desirable. We focus on one such application, namely that of device-independent quantum random number generation [73]. This is the task of certifying that the data generated by a party cannot be predicted by a malicious eavesdropper. In the device-independent setting, both the amount of randomness and its security is derived from the violation of a Bell inequality.

Non-projective measurements, such as SIC-POVMs, are useful for this task. The reason is that a Bell experiment implemented with entangled systems of local dimension d and standard projective measurements cannot have more than d outcomes. Consequently, one cannot hope to certify more than $\log d$ bits of local randomness. However, Bell experiment relying on d-dimensional entanglement implemented with (extremal) non-projective measurements can have up to d^2 outcomes [74]. This opens the possibility of generating up to $2\log d$ bits of local randomness without increasing the dimension of the shared entangled state. Notably, for the case of $d = 2$, such optimal quantum random number generation has been shown using a qubit SIC-POVM [42].

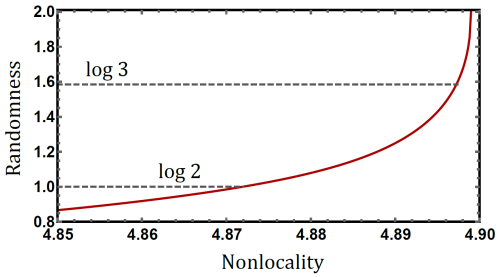

FIG. 5. Lower bound on the amount of device-independent randomness versus the value of $\mathcal{T}_2^{\text{SIC}}$.

Here, we employ our Bell inequalities for SIC-POVMs to significantly outperform standard protocols relying on projective measurements on d-dimensional entangled states. To this end, we briefly summarise the scenario for randomness generation. Alice and Bob perform many rounds of the Bell experiment illustrated in Figure 4. Alice will attempt to generate local randomness from the outcomes of her setting labelled by $\mathbf{povm}$. In most rounds of the Bell experiment, Alice performs $\mathbf{povm}$ and records the outcome a'. In a smaller number of rounds, she randomly chooses her measurement setting and the data is used towards estimating the value of the Bell functional $\mathcal{T}_d^{\text{SIC}}$ defined in Eq. (41). A malicious eavesdropper may attempt to guess Alice's relevant outcome a'. To this end, the eavesdropper may entangle her system with that of Alice and Bob, and perform a well-chosen POVM $\{E_c\}_c$ to enhance her guess. In analogy to Eq. (21), the eavesdropper's guessing probability reads

$$P_g^\beta \equiv \sup\left\{\sum_{c=1}^{d^2}\langle\psi_{\text{ABE}}|A^c_{\mathbf{povm}}\otimes\mathbb{1}\otimes E_c|\psi_{\text{ABE}}\rangle\right\}, \quad (43)$$

where $\{E_c\}_{c=1}^{d^2}$ is the measurement employed by the eavesdropper to produce her guess, the expression inside the curly braces is the probability that her outcome is the same as Alice's outcome for the setting $\mathbf{povm}$ for a particular realisation and the supremum is taken over all quantum realisations (the tripartite state and measurements of all three parties) compatible with the observed Bell inequality violation $\beta = \mathcal{T}_d^{\text{SIC}}$.

We quantify the randomness generated by Alice using the conditional min-entropy $H_{\min}(A_{\mathbf{povm}}|E) = -\log\left(P_g^\beta\right)$. To obtain a device-independent lower bound on the randomness, we must evaluate an upper bound on P_g^β for a given observed value of the Bell functional. We saw in Section IV that if the eavesdropper is only trying to guess the outcome of a single measurement setting, we can without loss of generality assume that they are only classically correlated with the systems of Alice and Bob. As before, we restrict ourselves to the asymptotic limit of many rounds, in which fluctuations due to finite statistics can be neglected.

In order to bound the randomness for some given value of $\mathcal{T}_d^{\text{SIC}}$, we use the hierarchy of quantum correlations [68]. We restrict ourselves to the cases of $d = 2$ and $d = 3$. For the case of $d = 2$, we construct a moment matrix with the operators $\{(\mathbb{1}, A_x)\otimes(\mathbb{1}, B_y)\otimes(\mathbb{1}, E)\}\cup\{A_{\mathbf{povm}}\otimes(\mathbb{1}, B_y, E)\}$, neglecting the $\perp$ outcome. The matrix is of size 361×361 with 10116 variables. Again, we can make use of symmetry to simplify the semidefinite program. In this case, the following permutation leaves the problem invariant: $x_1 \to \pi(x_1)$, $x_2 \to \pi(x_2)$, $a \to f_\pi(a, x_1, x_2)$, $a' \to \pi(a')$, $y \to \pi(y)$, $c \to \pi(c)$, where

$$f_\pi(a, x_1, x_2) = \begin{cases} a & \pi(x_1) < \pi(x_2) \\ 2 & \pi(x_1) \geq \pi(x_2) \text{ and } a = 1 \\ 1 & \pi(x_1) \geq \pi(x_2) \text{ and } a = 2 \\ \perp & \pi(x_1) \geq \pi(x_2) \text{ and } a = \perp \end{cases} \quad (44)$$

and $\pi \in S_4$. Using this symmetry reduces the number of free variables to 477. The trade-off between the amount of certified randomness and the nonlocality is illustrated in Figure 5. We find that for sufficiently large values of $\mathcal{T}_2^{\text{SIC}}$ (roughly $\mathcal{T}_2^{\text{SIC}} \geq 4.8718$), we outperform the one-bit limitation associated to projective measurements on entangled qubits. Notably, for even larger values of $\mathcal{T}_2^{\text{SIC}}$, we also outperform the restriction of $\log 3$ bits associated to projective measurements on entangled systems of local dimension three. For the optimal value of $\mathcal{T}_2^{\text{SIC}}$ we find $H_{\min}(A_{\mathbf{povm}}|E) \gtrsim 1.999$, which is compatible up to numerical precision with the largest possible amount of randomness obtainable from qubit systems under general measurements, namely two bits.

For the case of $d = 3$ we bound the guessing probability following the method of Ref [73]. This has the advantage of requiring only a bipartite, and hence smaller, moment matrix than the tripartite formulation. However, the amount of symmetry leaving the problem invariant is reduced, because the objective function only involves one outcome. Concretely, we construct a moment matrix of size 820×820 with 263549 variables. We then write the guessing probability as $P(a' = 1|\mathbf{povm})$ and identify the following group of permutations leaving the problem invariant: $x_1 \to \pi(x_1)$, $x_2 \to \pi(x_2)$, $a \to f_\pi(a, x_1, x_2)$, $a' \to \pi(a')$, $y \to \pi(y)$, where $\pi \in S_9$ leaves element 1 invariant and permutes elements $2, \ldots, 9$ in all possible ways. Taking this symmetry into account reduces the number of free variables to 460. In order to further simplify the problem we make use of RepLAB, a recently developed tool which decomposes representations of finite groups into irreducible representations [75, 76]. This allows us to write the moment matrix in a preferred basis in which it is block diagonal. The semidefinite constraint can then be imposed on each block independently, with the largest block of size 28×28 instead of 820×820. Solving one semidefinite program with SeDuMi [71] then takes 0.7s with $<$ 0.1GB of memory instead of 162s/0.2GB without block-diagonalisation, and fails due to lack of memory without any symmetrisation ($>$ 400GB required).

Using entangled states of dimension three and corresponding SIC-POVMs, one can attain the full range of values for $\mathcal{T}_3^{\text{SIC}}$. Importantly, the guessing probability is independent of the outcome guessed by the eavesdropper, and we can verify that the bound we obtain is convex, hence guaranteeing that no mixture of strategy by the eavesdropper must be consid-

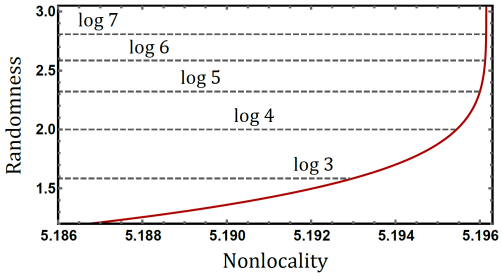

FIG. 6. Lower bound on the amount of device-independent randomness versus the value of $\mathcal{T}_3^{\text{SIC}}$.

ered [73]. The randomness is then given in Figure 6, which indicates that by increasing the value of $\mathcal{T}_3^{\text{SIC}}$, we can obtain more randomness than the best possible schemes relying on standard projective measurements and entangled systems of dimensions $3, 4, 5, 6, 7$. Especially, in the case of $\mathcal{T}_3^{\text{SIC}}$ being maximal, we find that $H_{\min}(A_{\mathbf{povm}}|E) \approx 3.03$ bits. This is larger than what can be obtained by performing projective measurements on eight dimensional systems (since $\log 8 = 3$ bits). It is, however, worth noting that this last value is obtained at the boundary of the set of quantum correlations where the precision of the solver is significantly reduced[5]. It is not straightforward to estimate the extent to which this reduced precision may influence the guessing probability, so it would be interesting to reproduce this computation with a more precise solver such as SDPA [77].

VII. CONCLUSIONS

MUBs and SICs are conceptually elegant, fundamentally important and practically useful features of quantum theory. We investigated their role in quantum nonlocality. For both MUBs and SICs (of any Hilbert space dimension) we presented families of Bell inequalities for which they produce the maximal quantum violations. Moreover, we showed that these maximal quantum violations certify natural operational notions of mutual unbiasedness and symmetric informational completeness. Then, we considered applications of both families of Bell inequalities in practically relevant tasks. The Bell inequalities for MUBs turn out to be useful for the task of device-independent quantum key distribution and give the optimal key rate for measurements with d outcomes. Moreover, for the case of qutrit systems we investigated the noise robustness of the protocol. For the Bell inequalities for SICs, we considered device-independent random number generation for qubits and qutrits based on SIC-POVMs. We showed (up to numerical precision) optimal randomness generation for qubit systems. For qutrit systems, we showed that more randomness can be generated than in any scheme using standard projective measurements and entanglement of up to dimension seven. These results were obtained using the RepLAB package, which helped to significantly reduce the complexity of the corresponding semidefinite programs by taking advantage of their symmetry.

This work opens many new research directions, so let us mention just a few of them. We showed that a maximal quantum violation of the Bell inequality for MUBs self-tests a maximally entangled state of local dimension d. In the case of the Bell inequality for SICs we have managed to certify the measurements of Bob, but we do not have a self-testing result for the state. If a self-test of the state is possible, what are the implications for the device-independent certification of the SIC-POVM setting? This may prove helpful towards solving another interesting question, namely that of proving optimal local randomness generation (i.e. $2 \log d$ bits) for any d based on the Bell inequality for SIC-POVMs. Another avenue of exploration regards the concept of mutually unbiased measurements (MUMs). In this work, we have shown some of their basic properties with regard to MUBs as well as examples of how they are relevant in quantum information theory. However, a more systematic exploration of MUMs would be desirable. Similarly, a general exploration of operational SICs (OP-SICs) in quantum information theory, as well as their relation to SICs, would be of similar interest. Finally, we note that our noise-robust results for quantum key distribution and quantum random number generation may be relevant for experimental implementations.

ACKNOWLEDGMENTS

We would like to thank Thais de Lima Silva and Nicolas Gisin for fruitful discussions. This work was supported by the Swiss National Science Foundation (Starting grant DIAQ, NCCR-QSIT). JK acknowledges funding from the HOMING programme of the Foundation for Polish Science (grant agreement no. POIR.04.04.00-00-5F4F/18-00) co-financed by the European Union under the European Regional Development Fund. MF acknowledges support from the Polish NCN grant Sonata UMO-2014/14/E/ST2/00020.

[1] J. Schwinger, Unitary operator bases, PNAS **46**, 570-579 (1960).

[2] G. Zauner, Quantendesigns, Grundzüge einer nichtkommutativen Designtheorie, Doctoral thesis at University of Vienna (1999).

[3] J. M. Renes, R. Blume-Kohout, A. J. Scott, C. M. Caves, Symmetric Informationally Complete Quantum Measurements, J. Math. Phys. **45**, 2171 (2004).

[5] In particular, the DIMACS errors at this point are of the order of 10^{-4}.

[4] W. K. Wooters, Quantum Measurements and Finite Geometry, Found Phys **36**, 112 (2006).

[5] M. Grassl, On SIC-POVMs and MUBs in Dimension 6, arXiv:quant-ph/0406175

[6] R. Beneduci, T. J. Bullock, P. Busch, C. Carmeli, T. Heinosaari, and A. Toigo, Operational link between mutually unbiased bases and symmetric informationally complete positive operator-valued measures, Phys. Rev. A **88**, 032312 (2013).

[7] I. Bengtsson, From SICs and MUBs to Eddington, J. Phys. Conf. Ser. **254** 012007 (2010).

[8] I. Bengtsson, K. Blanchfield, and A. Cabello, A Kochen-Specker inequality from a SIC, Phys. Lett. A **376**, 374 (2012).

[9] A. E. Rastegin, Uncertainty relations for MUBs and SIC-POVMs in terms of generalized entropies, Eur. Phys. J. D **67**, 269 (2013).

[10] T. Durt, B-G. Englert, I. Bengtsson and K. Życzkowski, On mutually unbiased bases, Int. J. Quantum Information 8, 535 (2010).

[11] H. Maassen and J. B. M. Uffink, Generalized entropic uncertainty relations, Phys. Rev. Lett. **60**, 1103 (1988).

[12] C. H. Bennett and G. Brassard, Quantum cryptography: Public key distribution and coin tossing, In Proceedings of IEEE International Conference on Computers, Systems and Signal Processing, volume 175, page 8. New York, 1984.

[13] A. K. Ekert, Quantum cryptography based on Bell's theorem, Phys. Rev. Lett. **67**, 661 (1991).

[14] D. Bruß, Optimal Eavesdropping in Quantum Cryptography with Six States, Phys. Rev. Lett. **81**, 3018 (1998).

[15] N. J. Cerf, M. Bourennane, A. Karlsson and N. Gisin, Security of quantum key distribution using d-level systems, Phys. Rev. Lett. **88**, 127902 (2002).

[16] V. Scarani, A. Acín, G. Ribordy, and N. Gisin, Quantum Cryptography Protocols Robust against Photon Number Splitting Attacks for Weak Laser Pulse Implementations, Phys. Rev. Lett. **92**, 057901 (2004).

[17] M. Hillery, V. Bužek, and A. Berthiaume Quantum secret sharing, Phys. Rev. A **59**, 1829 (1999).

[18] I-Ching Yu, Feng-Li Lin, and Ching-Yu Huang, Quantum secret sharing with multilevel mutually (un)biased bases, Phys. Rev. A **78**, 012344 (2008)

[19] A. Tavakoli, I. Herbauts, M. Żukowski, and M. Bourennane, Secret sharing with a single d-level quantum system, Phys. Rev. A **92**, 030302(R) (2015).

[20] W. K. Wootters and B. D. Fields, Optimal state-determination by mutually unbiased measurements, Ann. Phys. **191**, 363 (1989).

[21] R. B. A. Adamson and A. M. Steinberg, Improving Quantum State Estimation with Mutually Unbiased Bases, Phys. Rev. Lett. **105**, 030406 (2010).

[22] A. Ambainis, A. Nayak, A. Ta-Shma, U. Vazirani, Dense quantum coding and a lower bound for 1-way quantum automata, Proceedings of the 31st Annual ACM Symposium on Theory of Computing (STOC'99), pp. 376-383, 1999.

[23] A. Tavakoli, A. Hameedi, B. Marques, and M. Bourennane, Quantum Random Access Codes Using Single d-Level Systems, Phys. Rev. Lett. **114**, 170502 (2015)

[24] E. A. Aguilar, J. J. Borkała, P. Mironowicz, and M. Pawłowski, Connections between Mutually Unbiased Bases and Quantum Random Access Codes, Phys. Rev. Lett. **121**, 050501 (2018).

[25] A. Tavakoli, J. Kaniewski, T. Vértesi, D. Rosset, and N. Brunner, Self-testing quantum states and measurements in the prepare-and-measure scenario, Phys. Rev. A **98**, 062307 (2018).

[26] M. Farkas and J. Kaniewski, Self-testing mutually unbiased bases in the prepare-and-measure scenario, Phys. Rev. A **99**, 032316 (2019).

[27] D. Gottesman, Class of quantum error-correcting codes saturating the quantum Hamming bound, Phys. Rev. A **54**, 1862 (1996).

[28] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, Quantum Error Correction and Orthogonal Geometry, Phys. Rev. Lett. **78**, 405 (1997).

[29] C. Spengler, M. Huber, S. Brierley, T. Adaktylos, and B. C. Hiesmayr, Entanglement detection via mutually unbiased bases, Phys. Rev. A **86**, 022311 (2012).

[30] L. Vaidman, Y. Aharonov, and D. Z. Albert, How to ascertain the values of σ_x, σ_y and σ_z of a spin-1/2 particle, Phys. Rev. Lett. **87**, 1385 (1987).

[31] B-G. Englert and Y. Aharonov, The mean king's problem: Prime degrees of freedom, Phys. Lett. A **284**, 1 (2001).

[32] C. A. Fuchs, M. C. Hoang, and B. C. Stacey, The SIC Question: History and State of Play, Axioms **21**, 6 (2017).

[33] A. J. Scott, and M. Grassl, SIC-POVMs: A new computer study, J. Math. Phys. **51**, 042203 (2010)

[34] A. J. Scott, SICs: Extending the list of solutions, arXiv:1703.03993

[35] C. M. Caves, C. A. Fuchs, and R. Schack, Unknown quantum states: The Quantum de Finetti representation, J. Math. Phys. **43**, 4537 (2002).

[36] Z. E. D. Medendorp, F. A. Torres-Ruiz, L. K. Shalm, G. N. M. Tabia, C. A. Fuchs, and A. M. Steinberg, Experimental characterization of qutrits using symmetric informationally complete positive operator-valued measurements, Phys. Rev. A **83**, 051801(R) (2011).

[37] W. M. Pimenta, B. Marques, T. O. Maciel, R. O. Vianna, A. Delgado, C. Saavedra, and S. Pádua, Minimum tomography of two entangled qutrits using local measurements of one-qutrit symmetric informationally complete positive operator-valued measure, Phys. Rev. A **88**, 012112 (2013).

[38] N. Bent, H. Qassim, A. A. Tahir, D. Sych, G. Leuchs, L. L. Sánchez-Soto, E. Karimi, and R. W. Boyd, Experimental Realization of Quantum Tomography of Photonic Qudits via Symmetric Informationally Complete Positive Operator-Valued Measures, Phys. Rev. X **5**, 041006 (2015).

[39] J. Shang, A. Asadian, H. Zhu, and O. Gühne, Enhanced entanglement criterion via symmetric informationally complete measurements, Phys. Rev. A **98**, 022309 (2018).

[40] J. M. Renes, Equiangular Spherical Codes in Quantum Cryptography, Quant. Inf. Comput. **5**, 080 (2005).

[41] F. Bouchard, K. Heshami, D. England, R. Fickler, R. W. Boyd, B-G. Englert, L. L. Sánchez-Soto, and E. Karimi, Experimental investigation of high-dimensional quantum key distribution protocols with twisted photons, Quantum **2**, 111 (2018).

[42] A. Acín, S. Pironio, T. Vértesi, and P. Wittek, Optimal randomness certification from one entangled bit, Phys. Rev. A **93**, 040102(R) (2016).

[43] A. Tavakoli, D. Rosset, and M-O. Renou, Enabling Computation of Correlation Bounds for Finite-Dimensional Quantum Systems via Symmetrization, Phys. Rev. Lett. **122**, 070501 (2019)

[44] P. Mironowicz and M. Pawłowski, Experimentally feasible semi-device-independent certification of four-outcome positive-operator-valued measurements, Phys. Rev. A **100**, 030301(R) (2019).

[45] A. Tavakoli, M. Smania, T. Vértesi, N. Brunner, and M. Bourennane, Self-testing non-projective quantum measurements in prepare-and-measure experiments, arXiv:1811.12712

[46] M. Smania, P. Mironowicz, M. Nawareg, M. Pawłowski, A. Ca-

bello, and M. Bourennane, Experimental device-independent certification of a symmetric, informationally complete, positive operator-valued measure, arXiv:1811.12851
[47] D. M. Appleby, C. A. Fuchs, and H. Zhu, Group theoretic, lie algebraic and Jordan algebraic formulations of the sic existence problem, Quantum Inf. Comput. **15**, 61 (2015).
[48] I. Bengtsson, The Number Behind the Simplest SIC-POVM, Found Phys **47**, 1031 (2017).
[49] M. Appleby, S. Flammia, G. McConnell, and J. Yard, SICs and Algebraic Number Theory, Found Phys **47**, 1042 (2017).
[50] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).
[51] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, Phys. Rev. Lett. **23**, 880 (1969).
[52] N. Gisin, Bell inequalities: many questions, a few answers, arXiv:quant-ph/0702021
[53] T. Vértesi and E. Bene, Two-qubit Bell inequality for which positive operator-valued measurements are relevant, Phys. Rev. A **82**, 062115 (2010).
[54] J. Kaniewski, I. Šupić, J. Tura, F. Baccari, A. Salavrakos, and R. Augusiak, Maximal nonlocality from maximal entanglement and mutually unbiased bases, and self-testing of two-qutrit quantum systems, Quantum **3**, 108 (2019).
[55] H. Bechmann-Pasquinucci and N. Gisin, Intermediate states in quantum cryptography and Bell inequalities, Phys. Rev. A **67**, 062310 (2003).
[56] S-W. Ji, J. Lee, J. Lim, K. Nagata, and H-W. Lee, Multisetting Bell inequality for qudits, Phys. Rev. A **78**, 052103 (2008).
[57] I. Šupić, and J. Bowles, Self-testing of quantum systems: a review, arXiv:1904.10042
[58] S. Brierley, S. Weigert, and I. Bengtsson, All Mutually Unbiased Bases in Dimensions Two to Five, Quantum Info. & Comp. **10**, 0803 (2010).
[59] C. Jebarathinam, J-C. Hung, S-L. Chen, and Y-C. Liang, Maximal violation of a broad class of Bell inequalities and its implication on self-testing, Phys. Rev. Research **1**, 033073 (2019).
[60] J. Kaniewski, A weak form of self-testing, arXiv:1910.00706
[61] M. Krishna and K. R. Parthasarathy, An Entropic Uncertainty Principle for Quantum Measurements, Indian Journal of Statistics **64**, (3) 842 (2002).
[62] T. Heinosaari, T. Miyadera, and M. Ziman, An invitation to quantum incompatibility, Journal of Physics A: Mathematical and Theoretical **49**, (12) 123001 (2016).
[63] E. Haapasalo, Robustness of incompatibility for quantum devices, Journal of Physics A: Mathematical and Theoretical **48**, (25) 255303 (2015).
[64] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Quantum cryptography, Rev. Mod. Phys. **74**, 145 (2002).
[65] L. Masanes, S. Pironio, and A. Acín, Secure device-independent quantum key distribution with causally independent measurement devices, Nature Communications **2**, 238 (2011).
[66] T. Franz, F. Furrer, and R. F. Werner, Extremal quantum correlations and cryptographic security, Phys. Rev. Lett. **106**, 250502 (2011).
[67] S. Sarkar, D. Saha, J. Kaniewski, and R. Augusiak, Self-testing quantum systems of arbitrary local dimension with minimal number of measurements, arXiv:1909.12722
[68] M. Navascués, S. Pironio, and A. Acín, Bounding the Set of Quantum Correlations, Phys. Rev. Lett. **98**, 010401 (2007).
[69] D. Rosset, SymDPoly: symmetry-adapted moment relaxations for noncommutative polynomial optimization, arXiv:1808.09598
[70] Y. Cai, J-D. Bancal, J. Romero, and V. Scarani, A new device-independent dimension witness and its experimental implementation Journal of Physics A: Mathematical and Theoretical **49**, 305301 (2016)
[71] J. F. Sturm, Using SeDuMi 1.02, A Matlab toolbox for optimization over symmetric cones, Optimization Methods and Software **11**, 625 (1999).
[72] A. Tavakoli and N. Gisin, Platonic solids and fundamental tests of quantum mechanics, (in preparation).
[73] S. Pironio, A. Acín, S. Massar, A. Boyer de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe Random numbers certified by Bell's theorem, Nature **464**, 1021 (2010).
[74] G. M. D'Ariano, P. L. Presti, and P. Perinotti, Classical randomness in quantum measurements, J. Phys. A: Math. Gen. **38**, 5979 (2005).
[75] https://replab.github.io
[76] D. Rosset, F. Montealegre-Mora, J-D. Bancal, RepLAB: a computational/numerical approach to representation theory, arXiv:1911.09154.
[77] http://sdpa.sourceforge.net
[78] M. Navascués, S. Pironio, and A. Acín, SDP relaxations for non-commutative polynomial optimization, Handbook on semidefinite, conic and polynomial optimization. International series in operations research & management science **166**, 601 (2012).
[79] M. Horodecki, P. Horodecki, and R. Horodecki, Separability of mixed states: necessary and sufficient conditions, Physics Letters A. **223**, 1–8 (1996)
[80] S. Boyd and L. Vandenberghe, Convex optimization, Cambridge university press, 2004
[81] M. D. Choi, Completely positive linear maps on complex matrices, Linear algebra and its applications, **10**(3), 285-290 (1975)
[82] A. Jamiołkowski, Linear transformations which preserve trace and positive semidefiniteness of operators, Reports on Mathematical Physics, **3**(4), 275-278 (1972)
[83] S. Designolle, M. Farkas, and J. Kaniewski, Incompatibility robustness of quantum measurements: a unified framework, New J. Phys. **21**, 113053 (2019).
[84] J. Kaniewski, M. Tomamichel, and S. Wehner, Entropic uncertainty from effective anticommutators, Phys. Rev. A **90**, 012332 (2014).

Appendix A: Bell inequalities for mutually unbiased bases

In this appendix we fill in some details on the Bell inequalities for MUBs. We start by deriving the local and quantum bounds and proving the device-independent certification result stated in the main text. Then, we proceed to show that no stronger characterisation of Bob's measurement can be obtained from the maximal violation of our Bell inequality and, moreover, that the maximal violation is achieved by a single probability point.

The Platonic solids and fundamental tests of quantum mechanics

Armin Tavakoli[1] and Nicolas Gisin[1]

[1]*Département de Physique Appliquée, Université de Genève, CH-1211 Genève, Switzerland*

The Platonic solids is the name traditionally given to the five regular convex polyhedra, namely the tetrahedron, the octahedron, the cube, the icosahedron and the dodecahedron. Perhaps strongly boosted by the towering historical influence of their namesake, these beautiful solids have, in well over two millennia, transcended traditional boundaries and entered the stage in a range of disciplines. Examples include natural philosophy and mathematics from classical antiquity, scientific modeling during the days of the European scientific revolution and visual arts ranging from the renaissance to modernity. Motivated by mathematical beauty and a rich history, we consider the Platonic solids in the context of modern quantum mechanics. Specifically, we construct Bell inequalities whose maximal violations are achieved with measurements pointing to the vertices of the Platonic solids. These Platonic Bell inequalities are constructed only by inspecting the visible symmetries of the Platonic solids. We also construct Bell inequalities for more general polyhedra and find a Bell inequality that is more robust to noise than the celebrated Clauser-Horne-Shimony-Holt Bell inequality. Finally, we elaborate on the tension between mathematical beauty, which was our initial motivation, and experimental friendliness, which is necessary in all empirical sciences.

I. INTRODUCTION

Which physicist has never been attracted by mathematical beauty? And what is more beautiful than the Platonic solids; the five regular polyhedra in our three-dimensional space (see Fig.1)? Here, we first present the fascinating history of these solids and then use them to derive simple Bell inequalities tailored to be maximally violated for measurement settings pointing towards the vertices of the Platonic solids. In this way, we connect beautiful mathematics with foundational quantum physics. However, these *Platonic Bell inequalities* do not distinguish themselves with regard to experimental friendliness: quantum theory predicts that their violations are less robust to noise than the much simpler Clauser-Horne-Shimony-Holt (CHSH) Bell inequality [1]. In fact, Platonic Bell inequalities require more measurement settings - as many as the number of vertices of the platonic solid - than the CHSH Bell inequality, which requires only the absolute minimum of two settings per side. We also construct Bell inequalities tailored to another class of elegant polyhedra, namely the Archimedean solids, i.e. the semi-regular polyhedra. In particular we consider the famous Buckyball, a polyhedron which corresponds to the carbon-60 molecule used in the first molecular interferometer [2] , which requires even more measurement settings. However, we find that these Bell inequalities also do not offer notable experimental advantages. Finally, we depart from Bell inequalities motivated by mathematical beauty and instead focus our research on finding experimentally friendly Bell inequalities: starting from the Buckyball we iteratively search for noise robust Bell inequalities. This leads us to a Bell inequality that is somewhat more noise tolerant than the CHSH Bell inequality. However, it is remarkably inelegant. We conclude with a discussion of the danger for theoretical physics to become - and remain - too focused on mathematical beauty [3] at the expense of developing connections with experiments.

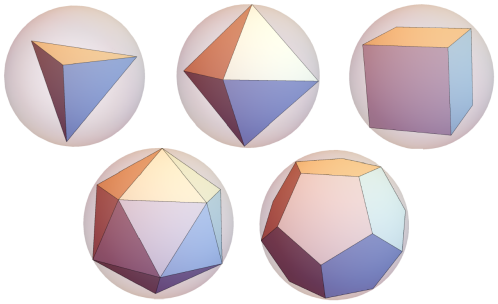

FIG. 1: The five Platonic solids inscribed in spheres. From left to right: the tetrahedron, the octahedron, the cube, the icosahedron and the dodecahedron.

II. A BRIEF HISTORY OF THE PLATONIC SOLIDS IN ARTS, PHILOSOPHY AND SCIENCE

This section provides a broader context for the Platonic solids. Readers interested exclusively in Bell inequalities may jump to the next section.

The ancient Greek civilisations laid the foundations of western natural philosophy. The development of the latter is permeated by a fascination for geometry. The magnum opus of Greek geometry, Euclid's *Elements*, remained a standard textbook until the 20'th century [4]. First printed in Venice in 1482 as one of the earliest mathematics books set in type, it has since been re-printed in at least a thousand editions[1] and is certainly the most influential mathematical work in history [5]. Geometry allowed the early natural philosophers to describe, understand and make predictions about, the physical world. In the sixth century BC, Thales of Miletus, often

[1] Ref. [5], authored in 1968, suggests that the Elements is only outdone in number of editions by the Bible.

hailed as the first scientific philosopher in western civilisation, likely used his knowledge of geometry to measure the height of the pyramids of Egypt [6]. Centuries later, in the Hellenistic period, Eratostenes accurately calculated the circumference of the Earth and Hipparcus discovered the precession of the equator. Archimedes' geometry led him to the Law of the Lever [7], still taught to every pupil in physics class.

Geometry was often ascribed a deeper meaning, beyond pure mathematics and its applications. This entails attributing spiritual, religious or philosophical meaning to certain proportions, planar shapes and solids, elevating the geometries to a tangibly sacred status. The perhaps most famous example of such metaphysical beliefs is due to the Pythagoreans[2] [8]. Their ideas of sacred geometries were influential, notably also on key figures such as Plato in the fifth century BC. In *The Republic*, Plato writes that *"geometry will draw the soul towards truth, and create the spirit of philosophy"* [9]. In *Timaeus*, Plato makes concrete the link between geometry and natural philosophy; he discusses the five regular polyhedra, i.e. the polyhedra whose vertices are identical and whose faces are identical regular polygons, namely the tetrahedron, the octahedron, the cube, the icosahedron and the dodecahedron. Today, these five solids are known as the *Platonic solids* (see Fig. 1). Plato assigned four of the solids to the four classical elements thought to be the fundamental form of all matter; the tetrahedron to fire, the octahedron to air, the cube to earth and the icosahedron to water. To the remaining fifth solid, Plato left the following mysterious comment [10] *"A fifth regular solid still exists, namely the dodecahedron, which does not form the element of any substance; but God used it as a pattern for dividing the zodiac into its twelve signs."* Later, his pupil Aristotle added a fifth element to the original four elements, namely the aether[3]. It historically became associated to the dodecahedron, perhaps due to its relevance for the golden ratio. From a purely mathematical standpoint, the Platonic solids were the focus of the 13'th book of Euclid's *Elements* which studies their construction and their proportions when inscribed in a sphere.

The Platonic solids can be appreciated by modern mathematicians for their appealing geometric properties, by modern natural scientists for their occurrence in nature, historical scientific models and metaphysical ideas, and by a broader modern audience for their historical appearance in western visual arts and natural philosophy, as well as their sheer beauty. It appears reasonable to say that the historical interest in the Platonic solids was substantially aided by the fact they were so strongly endorsed by a character as titanic as Plato.

Almost two millennia after Plato, the maintained appreciation for the Platonic solids could for instance be seen in Luca Pacioli's mathematics book *De Divina Proportione*. Published in 1509, it spends its first section motivating the divinity of the golden ratio; in particular by emphasising that the golden ratio appears in the dodecahedron, which is a representation of the aether [11]. The book's lasting success even outside mathematics circles may in part be due to its masterful illustrations of the Platonic solids and various other geometries, in drawings signed Leonardo da Vinci. In fact, the works of many artists feature the Platonic solids; ranging from the renaissance mosaics in the cathedral of San Marco in Venice to the 20'th century works of Maurits Escher, who incidentally also kept a coveted model of the nested Platonic solids in his office [12]. Salvador Dalí's 1955 painting *The Sacrament of the Last Supper* (framed in the golden ratio) sets stage inside a dodecahedron.

In the realm of natural philosophy, the Platonic solids found a new role in the 1597 publication of *Mysterium Cosmographicum* authored by Johannes Kepler. Kepler proposed a model of the heliocentric solar system in which the six known planets were modeled by nesting the five Platonic solids and inscribing and circumscribing them by spheres [13]. Although this model was later abandoned due to its inconsistencies with astronomical observations, it served as a stepping stone to Kepler's three laws of planetary motion. Albeit not in the solar system, the Platonic solids present themselves elsewhere in nature. Three of them are natural structures of crystals. A range of Boron compounds include Boron-12 which takes an icosahedral form. The icosahedron is also the structure of many species of Radiolaria and viruses, e.g. polio. Curiously, it was the discovery of the icosahedral phase in quasicrystals that led to the Nobel prize in chemistry in 2011 [14]. Notably, the most common silicates are structured as a silicon atom binding to four oxygen atoms. The silicon atom sits at the center of a tetrahedron with the oxygen atoms sitting at its vertices. Interestingly, silicates comprise the majority of Earth's crust and mantle, and they are often the dominating mineral in various forms of soil. Perhaps, had Plato ascribed the tetrahedron rather than the cube as the manifestation of earth, his metaphysical ideas might have better withstood the test of time.

III. A BRIEF HISTORY OF BELL INEQUALITIES

This section provides a non-technical introduction to Bell inequalities. Readers interested mainly in the technical considerations may proceed immediately to the next section.

Modern science, with its emphasis on empiricism, has for long left behind ideas of Euclidean geometry being fundamental to describing nature. The 19'th century saw the development of curved (non-euclidean) geometry[4] which in the early 20'th century found a fundamental role in Einstein's theory

[2] For instance, the number three was an ideal number as it was the number of vertices in a triangle, which was a symbol of Apollo. The number ten was termed a perfect number due to the number of vertices in a geometry called a tetractys. The number was therefore honoured by the Pythagoreans not gathering in groups of more than ten people.

[3] Aether theories persisted in science until the strong negative evidence put forward by the Michelson-Morley experiment, performed in 1887.

[4] Non-euclidean geometry was the climax of two millennia of mathematical discussions, first led by Greeks, then by Arabs and Persians and finally by renaissance Europeans, about Euclid's fifth postulate (parallel lines) [15].

of gravity. The 20'th century also brought with it the perhaps most radical change of scientific paradigm since the days of Newton, namely the theory of quantum mechanics, which governs nature on the scale of atoms and elementary particles. The most radical predictions of quantum mechanics defied the *principle of locality*, i.e. that events that are very far separated in space and time cannot influence each other[5] [16]. This counterintuitive feature put quantum mechanics on an apparent collision course with the famous no-signaling principle.

Quantum mechanics claimed that two objects, separated by large distances could still influence each other. Take a pair of atoms, which have a magnetic moment due the angular momentum and spin of their electrons and nucleus. We measure the direction of the atom's magnetic moment. Quantum mechanics tells us that if we were to find the magnetic moment of the first atom pointing upwards, then this can change the magnetic moment of our second atom so that it will also be found pointing upwards. This influence is *immediate*, and does not even require some carrier (e.g. a mechanical wave or light) to bring it from one atom to the other. Today, this phenomenon is famous under the name *entanglement*: the fact that the whole system is greater than the collection of its individual parts. Remarkably, however, quantum mechanics still manages to peacefully coexist with the principle of no-faster-than-light communication. The reason is that although distant systems influence each other, the influence does not carry any information from one system to the other. In the 1920s, the question of whether entanglement exists prompted an intensive series of debates between Einstein and Bohr; the former speaking of a "spooky action at a distance", and the latter in support of quantum mechanics.

Nevertheless, and most remarkably, in 1964 physicist John Bell proved that the existence of entanglement could in fact be scientifically settled [17]. Bell found a way of capturing the essence of what local theories predicted about the correlations between the magnetic moments. For example, if one finds the first magnetic moment pointing in some direction, how often does one also find the second magnetic moment pointing in the same direction? If the former points to the left, to what extent does it mean that the latter will be pointing right? Answering such questions tells us the *correlations* between the two distant magnetic moments. Bell showed that some correlations that were possible in quantum mechanics were in fact impossible in local theories; local correlations obey relations today known as *Bell inequalities*, which can be violated in quantum theory [18]. The existence of entanglement could therefore be confirmed by an experiment (see Fig. 2 for an illustration of a Bell experiment) successfully violating a Bell inequality. Early experiments strongly supported quantum mechanics [19, 20] and the matter was definitely settled by experiments in 2015 [21]. The monumental violation of Bell inequalities established entanglement as a natural phenomenon which gave rise to the today rapidly developing field of quantum information theory. This field promises things such as quantum computers, quantum cryptography and teleportation as exciting technologies in a currently unraveling "second quantum revolution" [22].

[5] It is interesting to point out that some earlier theories such as Newtonian gravity in fact did not respect the principle of locality; gravity propagates instantaneously. This was, however, generally perceived as a major drawback.

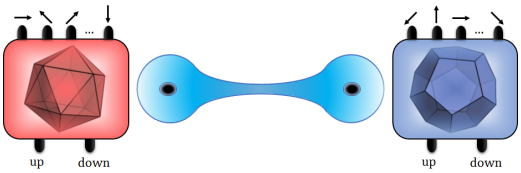

FIG. 2: Illustration of a Bell experiment. Two separate atoms that are entangled with each other are sent to different stations where their magnetic moments are measured along various directions. Each measurement answers whether the magnetic moment points up or down the axis along which it is measured. In a Platonic Bell inequality, the best measurements at each station are those that form a Platonic solid.

IV. THE PLATONIC SOLIDS

A three-dimensional solid that has sharp corners, straight edges and polygonal faces is called a polyhedron. The Platonic solids is the umbrella term for all polyhedra that are both convex and regular. In an intuitive but informal way, this means that

- **Convex polyhedron:** every two points inside the polyhedron can be connected with a straight line that itself is inside the polyhedron.
- **Regular polyhedron:** the edges, vertices and faces respectively look the same.

In two dimensions, it is easily seen that there are infinitely many regular convex polygons. Remarkably, the situation changes completely in three dimensions; Euclid proved that there are only five regular convex polyhedra. These are called the Platonic solids (see Fig. 1). Let us briefly review each of them.

- **Tetrahedron.** A triangular pyramid with four faces, four vertices and six edges.
- **Octahedron.** A triangular antiprism with eight faces, six vertices and twelve edges.
- **Cube.** A box with six square faces, eight vertices and twelve edges.
- **Icosahedron.** 20 triangular faces, twelve vertices and 30 edges. By dividing its vertices suitably in three sets of four, one can inscribe three perpendicular golden rectangles.
- **Dodecahedron.** Twelve pentagonal faces, 20 vertices and 30 edges. Its surface area, volume and distance between adjacent vertices are related to the golden ratio.

To every polyhedron, we can associate a partner polyhedron called its dual. The dual of the dual is again the original polyhedron. To construct the dual of a polyhedron, the main idea is to let the vertices of the dual pass through the midpoint of the faces of the original polyhedron. The Platonic solids exhibit particularly elegant duality relations: the tetrahedron is its own dual whereas the octahedron and cube are dual to each other and similarly for the icosahedron and the dodecahedron. Thus, the dual of a Platonic solid is always a Platonic solid.

V. BELL INEQUALITIES

The magnetic moment of an atom is a direction in three-dimensional space; we can think of it as an arrow denoted $\vec{n}$ on a unit-radius sphere. Imagine that we want to measure the magnetic moment. This can be done along any axis we want, labeled by an arrow $\vec{m}$ on our sphere. Quantum mechanics tells us how to compute the probability of our magnetic moment, initially in direction $\vec{n}$, being found up (along the positive axis) and down (along the negative axis) respectively, when measured along $\vec{m}$.

Let us now add a second atom. We separate the pair, sending one atom to Alice and one atom to Bob. Alice may measure the magnetic moment of her atom in various directions. Let us say that she has $N_{\rm A}$ different directions to choose from. We label her choice of measurement direction $x = 1, \ldots, N_{\rm A}$ and label the corresponding direction by $\vec{a}_x$. Similarly, Bob may measure his magnetic moment in one of $N_{\rm B}$ different directions. We label his choice of direction $y = 1, \ldots, N_{\rm B}$ and the specific direction by $\vec{b}_y$. For given choices of measurements, there are four possible outcomes. These are $++, +-, -+$ and $--$. If Alice and Bob have the same outcome, i.e. either $++$ or $--$, we say that Alice and Bob are correlated. If they have different outcomes, either $+-$ or $-+$, we say that Alice and Bob are anticorrelated. It is therefore handy to introduce a correlator which captures the degree of correlation or anticorrelation;

$$E(x,y) = p(+,+) + p(-,-) - p(+,-) - p(-,+). \quad (1)$$

The closer E is to one (negative one), the stronger are the correlations (anticorrelations). When $E = 0$ there are no correlations between the outcomes.

We wish to determine whether the correlations contained in the list $\{E(x,y)\}_{x,y}$ can be explained by local theories. To this end, we must construct Bell inequalities. These are inequalities of the form

$$\mathcal{B} \equiv \sum_{x=1}^{N_{\rm A}} \sum_{y=1}^{N_{\rm B}} c_{x,y} E(x,y) \overset{\text{local}}{\leq} C, \quad (2)$$

where $c_{x,y}$ are some real numbers and C is a bound that is respected by *all possible local theories*. We emphasise that the local bound holds irrespective of the measurement directions used to obtain the expectation values.

What does it mean that the correlations can be modeled with a local theory? Local models assume that when the particles were created, they were endowed with some shared property λ. A measurement simply reveals that already existing property. If Alice chooses measurement x, a local model determines whether the outcome is $+$ or $-$ given the property λ. The analogous goes for Bob. However, we do not know what specific property λ represents. Our ignorance of it is represented by a probability distribution $p(\lambda)$. Therefore, in a local model, the correlators reads

$$E(x,y) = \sum_{\lambda} p(\lambda) E^{\rm A}_{\lambda}(x) E^{\rm B}_{\lambda}(y). \quad (3)$$

Thus, to find the local bound C in Eq. (2), we must maximise $\mathcal{B}$ over $p(\lambda)$. Fortunately, this can be determined by checking a finite number of specific choices of $p(\lambda)$ (all the deterministic responses of Alice and Bob) and pick the largest one [24].

The critical point is that Bell inequalities can sometimes be violated ($\mathcal{B} > C$) if the Bell experiment is modeled within quantum mechanics, i.e. by Alice and Bob having their two magnetic moments in an entangled state. The most interesting case is when the two magnetic moments are *maximally entangled*, i.e. in the state

$$|\phi^+\rangle = \frac{|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle}{\sqrt{2}}. \quad (4)$$

This state has the remarkable property that if Alice measures her magnetic moment along direction $\vec{n}$, the magnetic moment of Bob ends up also pointing either up or down the axis $\vec{n}$ (up to a reflection in the xz-plane). This paves the way for quantum correlations that violate the Bell inequality and therefore do not admit a local model. The natural question becomes, how strong can quantum correlations be? How much can they violate a Bell inequality? In what follows, we construct Bell inequalities that achieve their maximal correlations in quantum mechanics by Alice and Bob choosing their measurement directions $\vec{a}_x$ and $\vec{b}_y$ to respectively point to the vertices of a Platonic solid.

By *Platonic Bell inequality*, we mean to say a Bell inequality that is maximally violated in quantum theory with measurements forming pairs of Platonic solids (see Fig. 2). Notably, Platonic solids have previously been used in the context of quantum mechanics, e.g. to construct correlation tests for a phenomenon known as steering [25], which is a weaker notion of a genuinely quantum phenomenon, as compared to the violation of Bell inequalities.

VI. TWO SIMPLE PLATONIC BELL INEQUALITIES

We begin by presenting two particularly simple Platonic Bell inequalities. Their simplicity stems from the fact that all the coefficients $c_{x,y}$ appearing in Eq. (2) are either $+1$, -1 or 0, and that the Bell inequalities are constructed by inspecting the symmetries between a Platonic solid and its dual Platonic solid. Our first Platonic Bell inequality gives Alice and Bob measurement settings that correspond to a cube and an octahedron respectively (being dual polyhedra). Our second Platonic Bell inequality is based on the icosahedron and the dodecahedron (again being dual polyhedra).

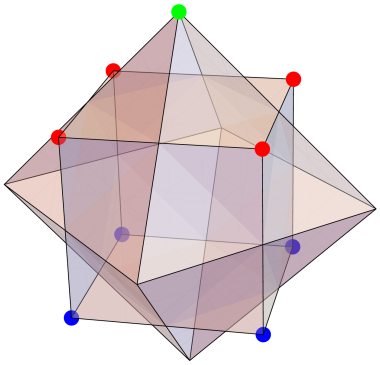

FIG. 3: A compound of two dual Platonic solids: the cube and the octahedron. For each vertex of the octahedron (for example green point), four vertices of the cube are equally close to it (red points) whereas the remaining four vertices of the cube are equally distant to it (blue points).

A. The first Platonic Bell inequality

We construct a Platonic Bell inequality for the cube and the octahedron. To this end, we consider a Bell experiment in which Alice has eight possible settings which we label by a three-bit string $x = x_1x_2x_3 \in \{0,1\}^3$ and Bob has six possible settings $y = y_1y_2$ which we label by a trit $y_1 \in \{1,2,3\}$ and a bit $y_2 \in \{0,1\}$. In order to construct the Bell inequality, we visualise a compound of a cube and an octahedron (see Fig. 3). The fact that these solids are dual to each other makes the compound highly symmetric. We exploit this to construct our Platonic Bell inequlity.

We now reason as follows. If Alice's and Bob's magnetic moments are maximally entangled, it means that if Alice measures her magnetic moment in the direction corresponding to the vertex of the octahedron (green point) and finds the outcome (say) $+$, she will remotely prepare Bob's magnetic moment in the same state (up to reflection in the xz-plane) as that into which her state has collapsed. Also, if a magnetic moment points in direction $\vec{n}$ and is measured along $\vec{m}$, the correlations (anticorrelations) are stronger the closer (more distant) the two vectors are. Four of the vertices of the cube (red points) are close, and equally close, to the vertex of the octahedron (green point). Therefore, we let the reasonably strong correlations contribute towards our Bell inequality test; specifically we put $c_{x,y} = 1$. Similarly, the other four vertices of the cube (blue points) are distant, and equally distant, from the vertex of the octahedron (green point). Hence, we let the reasonably strong anticorrelations contribute towards our Bell test; we put $c_{x,y} = -1$. Repeating this reasoning for every vertex of the octahedron, we arrive at the first Platonic Bell inequality. It reads

$$\mathcal{B}_{\text{cuboct}} = \sum_{x,y} (-1)^{x_{y_1}+y_2} E(x,y) \overset{\text{local}}{\leq} 24. \tag{5}$$

The local bound is obtained by considering all assignments of $+$ and $-$ to the outcomes of Alice and Bob. To derive it, we write $A_x, B_y \in \{\pm 1\}$ and impose the form of Eq. (3). This gives

$$\begin{aligned}\mathcal{B}_{\text{cuboct}} &= \\ &\sum_x A_x \sum_y (-1)^{x_{y_1}+y_2} B_y \leq \sum_x \left| \sum_y (-1)^{x_{y_1}+y_2} B_y \right| \\ &= \sum_x \Big| (-1)^{x_1}(B_{10} - B_{11}) + (-1)^{x_2}(B_{20} - B_{21}) \\ &\qquad + (-1)^{x_3}(B_{30} - B_{31}) \Big|. \end{aligned} \tag{6}$$

Notice that for all y_1, we have $B_{y_1 0} - B_{y_1 1} \in \{-2, 0, 2\}$. A little inspection shows that it is optimal to never choose the value zero. In fact, as long as we choose $B_{y_1 0} - B_{y_1 1} = \pm 2$, we always find the local bound $\mathcal{B}_{\text{cuboct}} = 24$. We remark that the Bell inequality (5) is closely related to the so-called Elegant Bell inequality [26]; the settings of Alice and Bob are merely doubled.

Now, in order to show that we indeed have a Platonic Bell inequality, we must derive the maximal quantum violation and show that it is achievable with a cube on Alice's side and an octahedron on Bob's side. If we let Alice and Bob share the maximally entangled state and perform measurements corresponding to these Platonic solids, we find that

$$\mathcal{B}_{\text{cuboct}} = 16\sqrt{3} \approx 27.71, \tag{7}$$

which is a violation of the Bell inequality.

Let us now prove that no larger value is possible in quantum theory i.e. there exists no entangled state (of potentially higher dimension) and no local measurements that can generate a larger Bell inequality violation. We write

$$\mathcal{B}_{\text{cuboct}} = \sum_x \langle \alpha_x | \beta_y \rangle \tag{8}$$

where

$$|\alpha_x\rangle = A_x \otimes \mathbb{1} |\psi\rangle \tag{9}$$

$$|\beta_x\rangle = \mathbb{1} \otimes \sum_y (-1)^{x_{y_1}+y_2} B_y |\psi\rangle. \tag{10}$$

Here A_x is a general observable of Alice and B_y is a general observable of Bob. We use the Cauchy-Schwarz inequality, the fact that $\langle \alpha_x | \alpha_x \rangle = 1$ and a simple concavity inequality to write

$$\mathcal{B}_{\text{cuboct}} \leq \sum_x \sqrt{\langle \beta_x | \beta_x \rangle} \leq \sqrt{8} \sqrt{\sum_x \langle \beta_x | \beta_x \rangle}. \tag{11}$$

Let us now consider the sum under the square-root on the

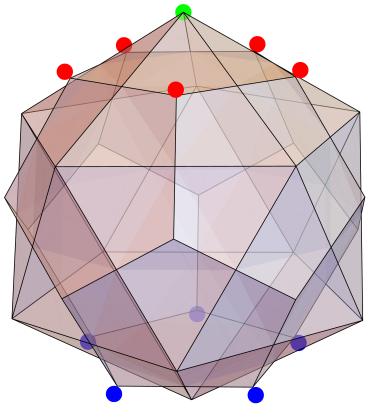

FIG. 4: A compound of two dual Platonic solids: the icosahedron and the dodecahedron. For each vertex of the icosahedron (for example green point), five vertices of the dodecahedron are (equally) close to it (red points) whereas another five vertices of the dodecahedron are (equally) distant to it (blue points).

right-hand-side. We find

$$\begin{aligned}\sum_x \langle\beta_x|\beta_x\rangle &= \sum_x \sum_{y,y'} (-1)^{x_{y_1}+x_{y'_1}+y_2+y'_2} \langle\psi|B_y B_{y'}|\psi\rangle \\ &= \sum_{y,y'} (-1)^{y_2+y'_2} \left(\sum_x (-1)^{x_{y_1}+x_{y'_1}}\right) \langle\psi|B_y B_{y'}|\psi\rangle \\ &= 8 \sum_{y_1,y_2,y'_2} (-1)^{y_2+y'_2} \langle\psi|B_{y_1 y_2} B_{y_1 y'_2}|\psi\rangle \\ &= 48 - 8 \sum_{y_1} \langle\psi|\{B_{y_1 0}, B_{y_1 1}\}|\psi\rangle \leq 96, \end{aligned} \tag{12}$$

where we have used that $B_y^2 = \mathbb{1}$ and that $\{B_{y_1 0}, B_{y_1 1}\} \geq -2\mathbb{1}$. Inserting this into Eq. (11), we recover the quantum bound $\mathcal{B}_{\text{cuboct}} \leq 16\sqrt{3}$. We conclude that our inequality (5) indeed is a Platonic Bell inequality for the cube and the octahedron.

B. The second Platonic inequality

Our first Platonic Bell inequality relied on exploiting the duality between the cube and the octahedron. The same intuition can be used to construct a simple Platonic Bell inequality for Alice performing measurements forming an icosahedron and Bob performing measurements forming a dodecahedron. Since the vector antipodal to every vector pointing to a vertex of the icosahedron and the dodecahedron respectively also points to a vertex, we can simplify the setting by only supplying Alice and Bob with a number of settings equal to half the number of vertices in the icosahedron and dodecahedron respectively. This means that we consider a Bell inequality test in which Alice has six settings and Bob has ten settings. In analogy with the previous, we visualise a compound of the icosahedron and the dodecahedron, see Fig. 4. Duality presents us with a highly symmetric compound which we exploit to construct our Bell inequality. Again, we imagine that the two magnetic moments are maximally entangled and that Alice therefore remotely prepares Bob's magnetic moment in the same direction as her own once she has measured it. Then, both the magnetic moments will (for example) point to the vertex (green point) of the icosahedron. Since this vertex is close, and equally close, to five vertices of the dodecahedron (red points) while distant, and equally distant, to five other vertices of the dodecahedron (blue points), we reward (put $c_{x,y} = 1$) correlations in the first five events and analogously reward (put $c_{x,y} = -1$) anticorrelations in the latter five events. In the event of Bob measuring in a direction corresponding to a vertex of the dodecahedron which is neither among the five close nor the five distant ones, we give no reward ($c_{x,y} = 0$). This simple reasoning leads to a list of coefficients which can straightforwardly be rearranged (permutations and global sign flips) to the coefficients

$$c^{\text{icodod}} = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 1 & 0 & 0 & 1 & -1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 & -1 & -1 & 0 & 1 & 0 & 1 \end{pmatrix} \tag{13}$$

The corresponding Bell inequality becomes

$$\mathcal{B}_{\text{icodod}} = \sum_{x=1}^{6} \sum_{y=1}^{10} c_{x,y}^{\text{icodod}} E(x,y) \overset{\text{local}}{\leq} 20, \tag{14}$$

where the local bound is obtained by considering all assignments of outcomes $(+,-)$ to Alice and Bob.

By sharing the maximally entangled state $|\phi^+\rangle$ and Alice performing measurements corresponding to an icosahedron and Bob performing measurements corresponding to a dodecahedron, we obtain the quantum value

$$\mathcal{B}_{\text{icodod}} = 2\sqrt{45 + 60\varphi} \approx 23.84, \tag{15}$$

where $\varphi = \frac{1+\sqrt{5}}{2}$ is the golden ratio. We have confirmed the optimality of this value (up to machine precision) using the hierarchy of quantum correlations [27]. This shows that Eq. (14) indeed is a Platonic Bell inequality. We note that one can attempt a more standard analytical proof of the quantum bound via the method used to derive the optimality of Eq. (7). However, this is significantly more cumbersome due to the increased number of settings.

VII. A SYSTEMATIC METHOD

Let us now outline a more general approach to the construction of Platonic Bell inequalities. Here, we choose a pair of Platonic solids for Alice and Bob and construct a Bell inequality for which the chosen solids are optimal.

	Tetrahedron	Octahedron	Cube	Icosahedron	Dodecahedron
Tetrahedron	16/3 | 16/3	7.82 | 8	9.24 | 32/3	14.78 | 16	22.82 | 80/3
Octahedron	-	12 | 12	13.86 | 16	21.96 | 24	34.40 | 40
Cube	-	-	64/3 | 64/3	29.89 | 32	47.51 | 160/3
Icosahedron	-	-	-	41.89 | 48	63.57 | 80
Dodecahedron	-	-	-	-	109.7 | 400/3

TABLE I: Local (left) and quantum (right) bounds for Bell inequalities for all pairs of Platonic solids. In all cases except that of two tetrahedra, two octahedra and two cubes we find a quantum violation. In all cases but these, we have Platonic Bell inequalities.

Let the vectors pointing to the vertices of Alice's Platonic solid be denoted $\{\vec{v}_x\}$. Similarly, the vectors $\{\vec{u}_y\}$ denote the vertices of Bob's Platonic solid. For simplicity, we let Alice have the solid with the smaller number of vertices. Consider now the following Bell inequality

$$\mathcal{B}_{\text{Plato}} \equiv \sum_{x=1}^{N_{\text{A}}} \sum_{y=1}^{N_{\text{B}}} (\vec{v}_x \cdot \vec{u}_y^*) E(x,y) \overset{\text{local}}{\leq} C, \tag{16}$$

where $\vec{u}^* = (u^1, -u^2, u^3)$. That is, we reward correlations and anticorrelations between Alice and Bob by an amount corresponding to the scalar product between the vertices of the desired Platonic solids (up to one being reflected in the xz-plane). It is worth noting that the Bell inequality depends on the relative angle between the two Platonic solids, which typically also will influence the local bound. The local bound C can straightforwardly be evaluated by considering all output strategies;

$$\begin{aligned} C &= \max_{\substack{A_1,\ldots,A_{N_{\text{A}}} \in \{\pm 1\}^{N_{\text{A}}} \\ B_1,\ldots,B_{N_{\text{B}}} \in \{\pm 1\}^{N_{\text{B}}}}} \sum_y B_y \sum_x (\vec{v}_x \cdot \vec{u}_y^*) A_x \\ &= \max_{A_1,\ldots,A_{N_{\text{A}}} \in \{\pm 1\}^{N_{\text{A}}}} \sum_y \left| \sum_x (\vec{v}_x \cdot \vec{u}_y^*) A_x \right|. \end{aligned} \tag{17}$$

Thus, we find the local bound by considering $2^{N_{\text{A}}}$ evaluations.

Let us now evaluate the value of $\mathcal{B}_{\text{Plato}}$ in a quantum model in which Alice and Bob share the maximally entangled state $|\phi^+\rangle$. We let Alice's measurements be represented by the vectors $\vec{a}_x = \vec{v}_x$ and Bob's measurements be represented by $\vec{b}_y = \vec{u}_y$. We find

$$\begin{aligned} \mathcal{B}_{\text{Plato}} &= \sum_{x,y} (\vec{v}_x \cdot \vec{u}_y^*) \langle \phi^+ | A_x \otimes B_y | \phi^+ \rangle \\ &= \sum_{x,y} (\vec{v}_x \cdot \vec{u}_y^*) \langle \phi^+ | \mathbb{1} \otimes B_y A_x^{\text{T}} | \phi^+ \rangle \\ &= \sum_{x,y} (\vec{v}_x \cdot \vec{u}_y^*) \langle \phi^+ | \mathbb{1} \otimes (\vec{u}_y^* \cdot \vec{\sigma}^{\text{T}})(\vec{v}_x \cdot \vec{\sigma}^{\text{T}}) | \phi^+ \rangle \\ &= \sum_{x,y} (\vec{v}_x \cdot \vec{u}_y^*)^2 \end{aligned} \tag{18}$$

In the second line, we have used that for any observable $R \otimes \mathbb{1} |\phi^+\rangle = \mathbb{1} \otimes R^{\text{T}} |\phi^+\rangle$ and in the penultimate line we have used that $\text{tr}\left((\vec{u}_y^* \cdot \vec{\sigma}^{\text{T}})(\vec{v}_x \cdot \vec{\sigma}^{\text{T}})\right) = 2 \vec{v}_x \cdot \vec{u}_y^*$.

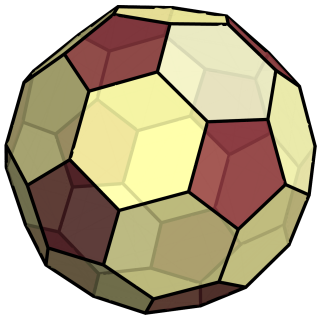

FIG. 5: The truncated icosahedron is an Archimedean solid with 32 faces, 60 vertices and 90 edges.

Let us now consider the maximal quantum correlations. We note that there are 15 possible pairs of Platonic solids (including when both solids are the same). For each of these 15 cases, we have constructed the Bell inequality (16), computed the quantum value (18) and compared it to the maximal quantum value obtained via the first level of the hierarchy of quantum correlations. We find that the quantum strategy based on the Platonic solids always is optimal. In Table I we compare the maximal quantum correlations with the local bound. We see that in all cases except for that of two tetrahedra, two octahedra and two cubes, the quantum correlations violate the local bound[6][7]. Moreover, due to the structure of the Platonic Bell inequalities, the maximal quantum value of $\mathcal{B}_{\text{Plato}}$ is a simple rational number.

[6] Since the relative angle between the two Platonic solids matters, we specify that the vertices of the Platonic solids where chosen to be the ones given by the software Mathematica's built-in function "PolyhedronData".

[7] We remark that the visibility required for a violation in the presence of white noise is the ratio between the local and quantum bounds.

A. A Buckyball Bell inequality

The Bell inequality construction (16) also works for some polyhedra that are not Platonic solids. Here, we illustrate this fact by considering a so-called Archimedean solid[8]. Specifically, we focus on the solid obtained from cutting an icosahedron symmetrically at every vertex so that each of them is replaced with a facet. Since at every vertex of the icosahedron, five of its faces meet, the cut polyhedron, called a truncated icosahedron, has five times as many vertices. The truncated icosahedron therefore has 60 vertices and its faces are either identical pentagons or identical hexagons - see Fig. 5 for an illustration. Incidentally, the truncated icosahedron is the design of the classic football and the structure of the carbon allotrope Buckminsterfullerene. The latter is often colloquially referred to as a "Buckyball".

In analogy with the Platonic Bell inequalities, we obtain a Buckyball Bell inequality using the construction in Eq. (16). To facilitate the fact that Alice and Bob will have 60 measurements each, we note that if a vector points to a vertex of the Buckyball, then the antipodal vector also points to a vertex of the Buckyball. Therefore, we only supply Alice and Bob with 30 measurements each, which are intended to point to the 30 vertices of the Buckyball which are not antipodal to each other. By choosing two perfectly aligned Buckyballs, the resulting Buckyball Bell inequality is

$$\mathcal{B}_{\text{Buckyball}} \overset{\text{local}}{\leq} \frac{20}{109}\left(461+493\varphi\right) \approx 230.952 \tag{19}$$

$$\mathcal{B}_{\text{Buckyball}} \overset{\text{quantum}}{\leq} 300, \tag{20}$$

where the quantum bound is obtained via the hierarchy of quantum correlations and saturated by choosing the Buckyball in Eq. (18). The local bound is obtained by evaluating Eq. (17).

VIII. OUTPERFORMING THE CHSH BELL INEQUALITY

The simplest Bell inequality test requires only two measurements each for Alice and Bob. The Bell inequality which describes this setting is known as the Clauser-Horne-Shimony-Holt (CHSH) inequality [1]. In fact, the CHSH inequality can straightforwardly be obtained from our general form in Eq. (16) by choosing $\vec{v}_1 = (1,0,0)$ and $\vec{v}_2 = (0,0,1)$ as well as $\vec{u}_1 = (1,0,1)/\sqrt{2}$ and $\vec{u}_2 = (1,0,-1)/\sqrt{2}$. The CHSH inequality reads

$$\mathcal{B}_{\text{CHSH}} \equiv E(1,1)+E(1,2)+E(2,1)-E(2,2) \overset{\text{local}}{\leq} 2. \tag{21}$$

Via Eq. (18), we saturate the maximal quantum violation, $\mathcal{B}_{\text{CHSH}} = 2\sqrt{2}$.

[8] The Archimedean solids are the semi-regular convex polyhedra (excluding the Platonic solids, prisms and antiprisms) of which there are 13.

An interesting question is the amount of disturbance that the quantum implementation can tolerate before ceasing to violate a Bell inequality. This is commonly modeled by mixing the desired quantum state (typically, the maximally entangled state) with white noise represented by the maximally mixed state, i.e.

$$\rho_v = v|\phi^+\rangle\langle\phi^+| + \frac{1-v}{4}\mathbb{1}, \tag{22}$$

where $v \in [0,1]$ is called the *visibility*. It is then relevant to find the critical visibility below which one can no longer violate a Bell inequality. In the case of the CHSH inequality, a simple computation shows that the critical visibility is $v = 1/\sqrt{2} \approx 0.7071$. As it has turned out, only few Bell inequalities can outperform the CHSH inequality in terms of their critical visibility for the maximally entangled state. The first example was reported in 2008; Ref. [28] constructed a Bell inequality with 465 settings on each side and showed a critical visibility of $v \approx 0.7056$. Recently, Bell inequalities with 42 settings on each side have been discovered, that further reduce the critical visibility of the maximally entangled state to $v \approx 0.7012$ [23]. The method for finding the latter Bell inequality relies on the development of an efficient algorithm for finding a separating hyperplane between a point and a convex set. In this context, the point is a quantum probability distribution measured in a Bell experiment and the convex set is the set of local correlations.

We have implemented the algorithm of Ref. [23] based on the Buckyball. Specifically, we compute the probability distribution corresponding to Alice and Bob measuring along aligned Buckyballs on the maximally entangled state. Via the algorithm, we find a hyperplane that separates it from the local set. Such a hyperplane can be written as the left-hand-side of a general Bell inequality, i.e. as in Eq. (2). We compute the local bound associated to the hyperplane as well as the maximal quantum violation. This gives us a new probability distribution. We mix it with a small amount of noise, corresponding to Eq. (22), and again run the algorithm. The procedure is repeated, and thus, noise is added and the probability distribution is perturbed, until it appears that we no longer find Bell inequalities with improved critical visibility. We illustrate the procedure in Figure 6. Implementing this procedure based on the Buckyball, we have found a 30 setting Bell inequality with a critical visibility of $v \approx 0.7054$. Whereas we used the Buckyball as our starting point, the quantum violation that corresponds to the stated visibility is achieved with other polyhedra that have more complicated structures. Unfortunately, the Bell inequality appears not to admit a simple analytical form. However, for sake of completeness, we present it in Appendix.

IX. LOST IN BEAUTY

There are different ways of reading our findings. First, there is the attractive connection established between the beautiful and historically rich Platonic solids and foundational relations

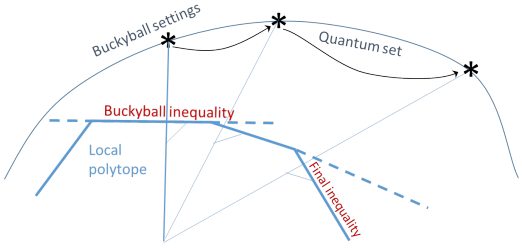

FIG. 6: Illustration of our application of the algorithm of Ref. [23]. Starting from the quantum probability distribution obtained from the Buckball, we find a Bell inequality that detects it. Then, we find the best quantum violation of that Bell inequality and repeat the procedure many times.

in our arguably most successful physics theory, quantum mechanics. But, secondly, there is a lesson to be learned here. Mathematical beauty was our initial motivation. The derived Platonic Bell inequalities are undoubtedly very elegant. However, admittedly, they are not experimentally friendly. They require many more measurement settings than necessary and in spite of the efforts going into developing an elegant construction, their resistance to noise (which is unavoidable in any experiment) is lower than in numerous simpler Bell inequalities. Naturally, it would be nice to see the Platonic Bell inequalities be violated in experiments; motivated simply by the appreciation of the Platonic solids and quantum nonlocality. However, unless the relevant technology incidentally happens to be set up and ready to use, it is unlikely that a practically minded experimenter would perform such an experiment. Indeed, only when we moved away from mathematical beauty, we eventually found a Bell inequality experiment (somewhat related to the Archimedean Buckyball) which is more noise resistant than the CHSH Bell inequality. The improvement is small, but it illustrates that searching to connect with experimental physics led us away from mathematical beauty. We believe that this carries a general lesson, namely that there is tension between mathematical beauty and experimentally friendly theoretical models [3]. Mathematical beauty can help in structuring the initial steps in new research directions, but unless theoretical models have experimental realities in mind, there is the danger of losing sight of empirical sciences.

Acknowledgments

We thank Flavien Hirsch for sharing his code for implementing the algorithm of Ref. [23]. We thank Thors Hans Hansson, Antonio Ortu and Augustin Baas for comments on the introduction. We thank Jessica Elsa Sellin for drawing our attention to Platonic solids in the structure of silicates. This work was supported by the Swiss National Science Foundation (Starting grant DIAQ, NCCR-QSIT).

[1] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, Phys. Rev. Lett. 23, 880 (1969).
[2] M. Arndt, O. Nairz, J. Vos-Andreae, C. Keller, G. van der Zouw and A. Zeilinger, WaveâĂŞparticle duality of C60 molecules, Nature 401, 680 (1999).
[3] Sabine Hossenfelder, *Lost in Math: How Beauty Leads Physics Astray*, Basic Book 2018.
[4] *Encyclopedia of Ancient Greece*, N. W. Wilson. Taylor & Francis 2010.
[5] *A History of Mathematics*, U. C. Merzbach and C. B. Boyer. John Wiley & Sons, Third edition 2011.
[6] *A Commentary on the First Book of Euclid's Elements*, Proklos Diadochos. Princeton University Press, Reprint edition 1992.
[7] *A History of Mechanical Inventions*, A. P. Usher. Harvard University Press, Revised Edition 2011.
[8] *Measuring Heaven: Pythagoras and His Influence on Thought and Art in Antiquity and the Middle Ages*, C. L. Joost-Gaugier. Cornell University Press 2007.
[9] *The Republic, VII*, Plato.
[10] *Timaeus*, Plato. Hackett Publishing Company, Second edition 2000.
[11] *The Golden Ratio: The Story of Phi, the World's Most Astonishing Number*, M. Livio. Broadway Books; Reprint edition 2003.
[12] *The Magic Mirror of M. C. Escher*, B. Ernst. Ballantine Books, 1976.
[13] E. Aiton, Johannes Kepler and the 'Mysterium Cosmographicum', Sudhoffs Archiv, Bd. 61, H. 2 (1977 2. QUARTAL), pp. 173-194.
[14] D. Monroe, Focus: Nobel Prize-Discovery of Quasicrystals, Phys. Rev. Focus 28, 14 (2011).
[15] *History of the Parallel Postulate*, F. P. Lewis, The American Mathematical Monthly, Vol. 27, No. 1. (Jan., 1920), pp. 16-23.
[16] N. Gisin, Quantum Chance, Springer 2014.
[17] J. S. Bell, On the Einstein Podolsky Rosen Paradox, Physics Vol 1, 3 pp.195-200 (1964).
[18] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. 86, 419 (2014).
[19] S. J. Freedman and J. F. Clauser, Experimental Test of Local-Hidden-Variable Theories, Phys. Rev. Lett. 28, 938 (1972).
[20] A. Aspect, P. Grangier, and G. Roger, Experimental Tests of Realistic Local Theories via Bell's Theorem, Phys. Rev. Lett. 47, 460 (1981).
[21] B. Hensen et. al., Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres, Nature 526, 682 (2015); L. K. Shalm, Strong Loophole-Free Test of Local Realism, Phys. Rev. Lett. 115, 250402 (2015); M. Giustina, Significant-Loophole-Free Test of Bell's Theorem with Entangled Photons, Phys. Rev. Lett. 115, 250401 (2015).
[22] A. Aspect, To be or not to be local, Nature 446, 866-867 (2007).
[23] S. Brierley, M. Navascues, T. Vértesi, Convex separation from convex optimization for large-scale problems, arXiv:1609.05011.
[24] A. Fine, Hidden Variables, Joint Probability, and the Bell In-

equalities Phys. Rev. Lett. 48, 291 (1982).
[25] D. J. Saunders,, S. J. Jones, H. M. Wiseman, and G. J. Pryde, Experimental EPR-Steering of Bell-local States, Nature Physics 6, 845 (2010).
[26] N. Gisin, Bell inequalities: many questions, a few answers, The Western Ontario Series in Philosophy of Science, pp 125-140, Springer 2009.
[27] M. Navascues, S. Pironio and A. Acín, Bounding the set of quantum correlations, Phys. Rev. Lett. 98, 010401 (2007).
[28] T. Vértesi, More efficient Bell inequalities for Werner states, Phys. Rev. A 78, 032112 (2008).

X. NOISE-TOLERANT BELL INEQUALITY

Below we give the coefficients $c_{x,y}$ for a Bell inequality of the form of Eq. (2) that outperforms the CHSH inequality in terms of noise tolerance. The local bound of the Bell inequality is 145.0181 and a quantum violation of 205.5873 is possible using a maximally entangled state. Notably, the critical visibility is the ratio of these two numbers, which is 0.7054. We give the coefficients in two matrices: the first one covers the values $y = 1, \ldots, 15$ and the latter covers the values $y = 16, \ldots, 30$.

$$\begin{pmatrix}
\{-0.938939, 0.158357, -0.533966, -0.23064, -0.724554, 0.244518, -0.272051, -0.503304, -0.261808, 0.424143, 0.0966467, 0.796627, 0.188715, 0.594949, -0.237468\} \\
\{0.18072, -0.768066, -0.610507, 0.60621, -0.777618, -0.237579, -0.429268, -0.464182, 0.318353, -0.327486, 0.807169, 0.443985, 0.476952, 0.139788, 0.647954\} \\
\{-0.804436, -0.644061, 0.152624, -0.0978691, 0.195717, 0.594001, 0.225163, 0.271193, -0.842739, 0.502871, -0.743338, -0.787921, -0.416596, -0.775433, 0.720186\} \\
\{-0.265651, 0.14169, -0.369637, -0.236285, 0.568644, 0.113412, -0.602475, 0.624299, -0.226567, 0.317713, -0.553164, 0.54268, -0.776341, 0.652845, -0.710988\} \\
\{-0.649483, -0.470572, 0.245036, 0.737009, 0.258046, -0.493589, 0.227192, 0.242118, 0.305222, -0.846779, -0.61715, -0.725507, -0.799294, -0.641217, 0.610627\} \\
\{0.413947, -0.307873, 0.635676, 0.160502, -0.149634, -0.314452, 0.51116, -0.145385, 0.425781, -0.0626215, 0.661908, -0.385767, 0.675485, -0.781107, 0.234186\} \\
\{-0.354744, -0.683865, 0.21573, -0.185906, 0.294178, 0.658654, 0.153385, 0.214849, -0.571555, 0.68996, -0.43329, -0.361252, -0.230952, -0.405367, 0.552939\} \\
\{-0.568539, -0.285698, 0.218428, 0.575954, 0.339942, -0.585974, 0.15821, 0.232436, 0.812193, -0.43336, -0.309785, -0.250307, -0.396192, -0.213033, 0.72279\} \\
\{-0.340875, 0.637664, -0.73432, -0.164406, 0.34447, 0.41788, -0.281528, 0.567055, -0.629244, 0.0903495, -0.714465, -0.799878, -0.779148, 0.356293, -0.241836\} \\
\{0.328608, -0.362466, 0.291346, 0.406731, -0.70639, -0.235605, 0.57929, -0.512153, 0.149479, -0.798215, -0.308225, -0.726373, 0.641391, -0.624684, 0.667658\} \\
\{0.43759, 0.89431, -0.773745, -0.548761, -0.814545, 0.562296, -0.290851, -0.245683, -0.771971, -0.737208, 0.259936, 0.114717, 0.308442, 0.248691, 0.232629\} \\
\{0.787729, 0.179406, -0.641468, 0.691865, -0.932348, -0.605349, -0.406903, -0.297606, -0.519086, -0.67114, 0.189645, 0.206028, 0.354553, 0.731856, 0.619125\} \\
\{0.220948, 0.449649, -0.667992, -0.907741, -0.716903, 0.664622, -0.197348, -0.402631, -0.692509, 0.462677, 0.781921, 0.256704, 0.330691, 0.127435, -0.620854\} \\
\{0.431043, 0.150873, -0.726035, 0.546273, -0.349899, -0.871917, -0.296315, -0.31912, 0.652939, -0.635094, 0.270925, 0.356846, 0.208299, 0.506987, 0.646933\} \\
\{-0.222491, 0.597369, 0.717729, -0.740314, 0.494043, 0.318595, 0.406368, 0.890863, -0.217487, 0.576908, 0.275939, 0.731067, -0.618467, 0.633616, 0.0314815\} \\
\{0.62162, -0.226652, 0.678117, 0.236063, 0.576352, -0.728841, 0.623365, 0.609693, 0.650236, -0.241325, 0.688331, 0.19918, 0.600824, -0.625507, 0.247508\} \\
\{-0.629058, -0.665512, 0.762582, 0.699471, 0.663632, 0.55867, 0.74883, 0.594934, -0.779332, -0.666641, 0.756054, 0.747264, -0.335007, -0.0285627, 0.696037\} \\
\{-0.673681, 0.362363, 0.764493, -0.569922, 0.415839, 0.823724, 0.424997, 0.530407, -0.332561, -0.424747, 0.576711, 0.76404, -0.395631, 0.703754, 0.823685\} \\
\{0.43872, -0.683226, 0.370403, 0.741479, 0.803973, -0.679839, 0.553606, 0.61573, -0.318631, -0.37065, 0.722129, 0.66742, 0.574685, -0.505264, 0.217202\} \\
\{-0.925456, 0.542483, -0.319801, -0.719634, 0.438413, 0.787642, 0.603134, 0.868044, 0.744695, 0.355048, -0.147578, -0.336373, -0.54689, -0.772933, -0.427955\} \\
\{0.661664, -0.711855, 0.637209, 0.599201, -0.2172, -0.921964, 0.676367, 0.688874, 0.333304, 0.476668, -0.292252, -0.267618, -0.791729, -0.213157, 0.0155498\} \\
\{-0.000473923, 0.674635, -0.621596, -0.654198, -0.771299, 0.222375, 0.344293, 0.298822, 0.59341, 0.489033, -0.703297, -0.355922, -0.663328, -0.612558, -0.629366\} \\
\{0.707855, -0.439378, -0.392686, 0.53223, -0.51318, -0.660753, 0.753103, 0.487686, 0.292849, 0.578796, -0.267075, -0.212718, -0.62204, -0.703102, -0.81733\} \\
\{-0.741289, 0.29658, -0.549618, -0.240212, -0.63298, 0.392904, -0.298802, -0.908456, -0.382104, 0.875208, 0.629615, 0.615326, 0.697973, 0.243977, -0.744801\} \\
\{0.212004, -0.585664, -0.7033, 0.32745, -0.43961, -0.258786, -0.581786, -0.370539, 0.658757, -0.363462, 0.263837, 0.709558, 0.143703, 0.42078, 0.58684\} \\
\{-0.697701, 0.446153, -0.667549, -0.265277, 0.391646, 0.394696, -0.624305, -0.659119, -0.196455, 0.625685, 0.820596, 0.356001, 0.669307, 0.293896, -0.642148\} \\
\{0.563873, -0.56564, -0.121853, 0.551754, -0.588406, -0.536469, -0.697011, -0.721448, 0.673423, -0.243212, 0.224131, 0.780953, 0.176979, 0.588214, 0.670529\} \\
\{0.694967, 0.649393, -0.277494, -0.705304, -0.215655, -0.515298, -0.688266, -0.582478, 0.66964, 0.671606, 0.375588, 0.287088, 0.806006, 0.782312, -0.381033\} \\
\{-0.222341, 0.305532, 0.7595, -0.795671, 0.193786, 0.65971, 0.734857, 0.403256, -0.485025, 0.624732, -0.709763, -0.737372, -0.389726, -0.399119, 0.209891\} \\
\{0.509022, -0.0679171, 0.259325, 0.433121, 0.692275, -0.867476, 0.300912, 0.637387, 0.672749, -0.504335, -0.491735, -0.633081, -0.638331, -0.324914, 0.244414\}
\end{pmatrix}$$

$$\begin{pmatrix}
\{0.618765, -0.662851, -0.734738, 0.484761, -0.73526, 0.727705, -0.567132, 0.671811, -0.501899, 0.219975, -0.689996, 0.409004, 0.678675, -0.173631, 0.548544\} \\
\{-0.205502, -0.620415, 0.480766, -0.702919, 0.687927, -0.760476, 0.68856, 0.0236075, 0.204386, -0.644982, 0.634582, -0.656913, 0.767748, 0.341393, -0.233116\} \\
\{0.350289, 0.701243, 0.693346, 0.330624, -0.214131, 0.325917, -0.564207, -0.752727, -0.332827, -0.521661, -0.640136, 0.169096, -0.150438, 0.746876, 0.245401\} \\
\{0.3686, 0.452042, -0.752094, 0.826602, -0.802341, 0.707194, -0.664093, 0.211129, -0.382652, 0.370586, -0.380456, 0.278798, -0.467576, -0.744942, 0.416623\} \\
\{0.647496, 0.664758, 0.446486, 0.845724, 0.698577, -0.435885, -0.453555, -0.631341, -0.668105, -0.32313, -0.0426063, -0.686365, -0.252337, 0.262152, 0.802131\} \\
\{-0.761503, 0.639096, 0.695576, -0.389685, 0.582081, -0.627134, 0.350879, -0.652563, 0.388622, -0.162314, 0.478742, -0.347407, -0.557886, 0.601113, -0.929526\} \\
\quad \{0.866332, 0.697774, 0.608627, 0.645319, 0.624741, 0.758558, 0.498415, 0.434939, -0.286139, -0.711873, -0.82185, -0.711271, -0.559926, 0.723296, 0.352421\} \\
\quad \{0.602424, 0.770311, 0.672259, 0.53567, 0.744653, 0.7174, 0.729257, 0.322235, -0.728996, -0.405124, -0.635146, -0.593506, -0.529851, 0.252734, 0.834103\} \\
\{0.597579, -0.681695, -0.255051, -0.409194, 0.611591, 0.336535, 0.626865, 0.303354, -0.221829, 0.695675, -0.269514, 0.570629, 0.603758, -0.640025, 0.616101\} \\
\{-0.151168, -0.789821, -0.375987, -0.35428, 0.55729, 0.481758, 0.313856, 0.606919, 0.647259, -0.238287, 0.631875, -0.247064, 0.634522, 0.574709, -0.439511\} \\
\quad \{0.68519, 0.620034, 0.625722, 0.718528, -0.212594, -0.368697, -0.329059, -0.431544, 0.810066, 0.406892, 0.7785, 0.378512, 0.381378, -0.602278, -0.747872\} \\
\quad \{0.194688, 0.768665, 0.865638, 0.506669, -0.467267, -0.221339, -0.38699, -0.665194, 0.272847, 0.560317, 0.24777, 0.667351, 0.469105, -0.382848, -0.810405\} \\
\{0.669664, 0.00024622, -0.451122, 0.69841, -0.251988, -0.679355, -0.695921, -0.609369, 0.590384, 0.205755, 0.631482, 0.302091, 0.859159, -0.339175, -0.433473\} \\
\{-0.716336, -0.264214, 0.584305, -0.398949, -0.857242, -0.659392, -0.48178, -0.563143, 0.236219, 0.590674, 0.22193, 0.600616, 0.85844, -0.608455, -0.22848\} \\
\{0.22872, 0.512406, 0.719643, 0.268033, -0.375038, -0.476979, -0.314531, -0.774318, -0.688045, 0.759848, -0.676097, 0.631291, -0.387913, 0.622789, 0.230733\} \\
\{0.0618877, 0.589273, 0.211638, 0.805043, -0.0651967, -0.337303, -0.728714, -0.740788, 0.708116, -0.661282, 0.614384, -0.682569, -0.422501, 0.228226, 0.261265\} \\
\quad \{0.601843, 0.22407, 0.172081, 0.225597, -0.789967, -0.710092, -0.193365, -0.229389, -0.381636, -0.279267, 0.700554, 0.660094, -0.28382, 0.285817, 0.44618\} \\
\quad \{0.222794, 0.288243, 0.615222, 0.172299, -0.449681, -0.787868, -0.246614, -0.145826, -0.67505, 0.64223, 0.467168, 0.322405, -0.285842, 0.686903, 0.348719\} \\
\quad \{0.751474, 0.198228, 0.153561, 0.712058, -0.74731, -0.315494, -0.628324, -0.128811, 0.678201, -0.724899, 0.381721, 0.347916, -0.30515, 0.269474, 0.566261\} \\
\{-0.468073, -0.794563, -0.314278, -0.776495, 0.398847, 0.0624419, 0.307419, 0.287304, -0.535865, -0.456788, -0.291028, -0.667804, 0.868123, -0.611396, 0.335802\} \\
\quad \{-0.41862, -0.830706, -0.775694, -0.469073, 0.213926, 0.457578, 0.208123, 0.597829, -0.584421, -0.39469, -0.632849, -0.15915, 0.457207, 0.604093, -0.620319\} \\
\{-0.821301, -0.136446, -0.191609, -0.213651, 0.476984, 0.294063, 0.286634, 0.231716, -0.797946, -0.457097, -0.572769, -0.845424, 0.753208, -0.663784, -0.665987\} \\
\{-0.468497, -0.279621, -0.717694, -0.283827, 0.275871, 0.308916, 0.146064, 0.22691, -0.504266, -0.774309, -0.854537, -0.355399, 0.811631, -0.402012, -0.626346\} \\
\quad \{0.98798, -0.252349, -0.713178, 0.730876, -0.314607, -0.491696, -0.797988, -0.58876, -0.606883, 0.137153, -0.460238, 0.240625, 0.794913, -0.416869, 0.564035\} \\
\{-0.684411, -0.391126, 0.642472, -0.643781, -0.454216, -0.63543, -0.348454, -0.736423, 0.166555, -0.586306, 0.176885, -0.631294, 0.573863, 0.275525, -0.307261\} \\
\quad \{0.416442, 0.603794, 0.328354, 0.280055, -0.215751, -0.646012, -0.327045, -0.757004, -0.704772, 0.21979, -0.283136, 0.156536, -0.772339, -0.64949, 0.791263\} \\
\{-0.631459, 0.595693, 0.356961, 0.407268, -0.708888, -0.237543, -0.702033, -0.680114, 0.240264, -0.408217, 0.222841, -0.274835, -0.296321, 0.765918, -0.74526\} \\
\{-0.454223, -0.293342, -0.304669, -0.372679, 0.479959, 0.857962, 0.814488, 0.739922, 0.706705, 0.668698, -0.412643, -0.464132, 0.0813487, -0.305294, -0.364783\} \\
\quad \{0.25035, 0.300207, 0.725872, 0.231932, -0.67687, 0.226395, -0.739863, -0.744072, -0.315295, 0.477261, -0.711921, 0.692551, -0.337137, 0.788336, 0.141295\} \\
\quad \{0.518217, 0.476583, 0.375773, 0.618415, 0.588361, -0.655611, -0.447799, -0.675673, 0.385371, -0.353972, 0.855368, -0.812284, -0.347944, 0.112769, 0.770458\}
\end{pmatrix}$$

Experimental Characterization of Unsharp Qubit Observables and Sequential Measurement Incompatibility via Quantum Random Access Codes

Hammad Anwer,[1] Sadiq Muhammad,[1] Walid Cherifi,[1] Nikolai Miklin,[2] Armin Tavakoli,[3] and Mohamed Bourennane[1]

[1]*Department of Physics, Stockholm University, S-10691 Stockholm, Sweden*
[2]*Institute of Theoretical Physics and Astrophysics, National Quantum Information Center, Faculty of Mathematics, Physics and Informatics, University of Gdansk, 80-952 Gdánsk, Poland*
[3]*Département de Physique Appliquée, Université de Genève, CH-1211 Genève, Switzerland*

(Received 17 January 2020; revised 30 April 2020; accepted 23 July 2020; published 19 August 2020)

Unsharp measurements are increasingly important for foundational insights in quantum theory and quantum information applications. Here, we report an experimental implementation of unsharp qubit measurements in a sequential communication protocol, based on a quantum random access code. The protocol involves three parties; the first party prepares a qubit system, the second party performs operations that return both a classical and quantum outcome, and the latter is measured by the third party. We demonstrate a nearly optimal sequential quantum random access code that outperforms both the best possible classical protocol and any quantum protocol that utilizes only projective measurements. Furthermore, while only assuming that the involved devices operate on qubits and that detected events constitute a fair sample, we demonstrate the noise-robust characterization of unsharp measurements based on the sequential quantum random access code. We apply this characterization towards quantifying the degree of incompatibility of two sequential pairs of quantum measurements.

DOI: 10.1103/PhysRevLett.125.080403

Introduction.—Textbook measurements in quantum theory are represented by complete sets of orthogonal projectors. However, general measurements in quantum theory are described by positive operator-valued measures (POVMs), i.e., an ordered set of positive operators $\{M_i\}_i$ with normalization $\sum_i M_i = \mathbb{1}$. Evidently, projective measurements are instances of POVMs but not all POVMs are projective measurements. These nonprojective measurements are well defined in Hilbert spaces of fixed dimension (otherwise they can be viewed as projective measurements in a larger space [1]). They are foundationally interesting and relevant to many phenomena and applications of quantum theory.

Some nonprojective measurements are extremal in the space of all POVMs with fixed Hilbert space dimension and number of outcomes i.e., they cannot be simulated with stochastic implementation of other measurements [2]. Whereas such POVMs have been studied in broad contexts [2–11], far from all nonprojective measurements are of this type. In fact, many interesting POVMs are *unsharp* measurements, in the sense that they are weaker (noisy) variants of projective measurements. By suitably tuning the noise parameter (sharpness), an experimenter can control the information-disturbance trade-off [12]; continuously from extracting no information and inducing no disturbance (noninteractive measurement) to extracting maximal information and inducing maximal disturbance (sharp projective measurement). Sequential unsharp measurements that individually induce only a small disturbance can be used for real-time monitoring of the evolution of single quantum systems [13–16]. When sufficiently frequent, such sequences effectively constitute continuous measurements, which have broad relevance in quantum information science (see, e.g., the review in Ref. [17]). Two key application of sequential unsharp measurements are adaptive measurement protocols [18,19] and quantum feedback protocols [20–22]. Interestingly, such sequences are also versatile as they can be used to realize the most general quantum measurements [23]. Moreover, unsharp measurements have prominent roles in a number of other topics including weak values [24], entanglement amplification [25], quantum random number generation [26], tests of the memory capacity of classical systems [27] and sequential quantum correlations [12,28–31]. This has prompted a number of experiments focused on the implementation of incompatible measurements [32–34], quantum contextuality [30], and quantum nonlocality [35–37].

Recently, Refs. [38,39] considered unsharp measurements in a sequential implementation of a frequently

encountered communication task known as a quantum random access code (QRAC) [40–42]. In a (Q)RAC, a sender, Alice, receives two input bits (x_0, x_1) which she encodes into a (qu)bit that is sent to a receiver, Bob. Bob receives an input bit y and then attempts to choose his output b such that it equals to Alice's yth bit, i.e., $b = x_y$. In an optimal classical protocol, Alice always sends x_0; thus Bob succeeds when $y = 0$ and takes a random guess when $y = 1$, leading to an average success probability of 0.75. However, a quantum advantage is obtained if Alice prepares four qubit states forming a square on the equator of the Bloch sphere and Bob measures two suitably aligned Pauli observables, resulting in a success rate of ≈ 0.85. From an alternative perspective, a QRAC can be viewed as a certification tool that allows an experimenter to characterize the involved preparation and measurement devices solely from its success rate, while assuming only that the setup operates on qubits [43].

However, unsharp measurements in standard QRACs are unremarkable as their outcome statistics can be simulated by a measurement device that stochastically implements projective measurements. Therefore, Refs. [38,39] considered a sequential scenario (see Fig. 1) in which the postmeasurement state of Bob is relayed to another receiver, Charlie, who receives an input bit z and analogously attempts to recover the zth bit of Alice. Thus, Alice sequentially implements a QRAC with Bob and Charlie in the respective order. Here, unsharp measurements become indispensable: in order for both QRACs to achieve a high success rate, Bob must interact with the incoming system in such a way that sufficient information is extracted to power his guess of x_y, while simultaneously the disturbance is limited to allow Charlie to accurately guess x_z. Furthermore, it was shown [38,39] that sequential QRACs can serve as certification tools for characterizing the unsharpness of Bob's operations while only assuming that the states are qubits.

In this Letter, we report experimental implementation of sequential QRACs using measurements of tunable unsharpness and demonstrate nearly optimal quantum correlations that outperform both all classical protocols as well as all quantum protocols based only on projective qubit measurements. We harvest these quantum communication

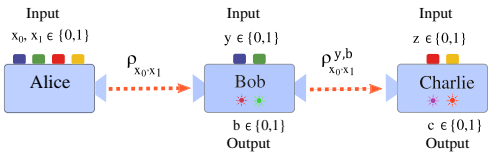

FIG. 1. Alice receives two bits x_0, x_1 and sends the qubit state ρ_{x_0,x_1} to Bob who receives an input y and produces a classical output b and a quantum output $\rho^{y,b}_{x_0,x_1}$, which is measured by Charlie according to his input z, yielding an outcome c.

advantages to certify the unsharpness parameter by confining it to a narrow interval. Subsequently, we theoretically develop and experimentally demonstrate how the sequential QRACs can be applied to quantify the degree of incompatibility [44] between two sequential pairs of quantum measurements.

Scenario and theoretical background.—Based on Refs. [38,39], we describe the sequential QRAC experiment. It involves three parties, Alice, Bob, and Charlie (see Fig. 1). Alice receives an input $x \equiv x_0, x_1 \in \{0, 1\}$ and prepares some uncharacterized qubit state denoted ρ_x, which she sends to Bob. Bob receives an input $y \in \{0, 1\}$ and performs a corresponding operation on ρ_x. This operation produces a classical output $b \in \{0, 1\}$ and some postoperation qubit state denoted $\rho_x^{y,b}$, which is sent to Charlie. Charlie receives an input $z \in \{0, 1\}$ and then measures $\rho_x^{y,b}$, yielding an outcome $c \in \{0, 1\}$. All inputs (x, y, z) are statistically independent and uniformly distributed. The limit of many rounds yields conditional probability distributions $p(b, c|x, y, z)$.

The conditional probability distributions $p(b, c|x, y, z)$ are used to evaluate the success rate of two QRACs: one between Alice and Bob, and one between Alice and Charlie. The former is successful when $b = x_y$ and the latter is successful when $c = x_z$. The two respective success rates read

$$W_{\mathrm{AB}} = \frac{1}{8}\sum_{x,y} P(b = x_y|x, y),$$
$$W_{\mathrm{AC}} = \frac{1}{8}\sum_{x,z} P(c = x_z|x, z). \tag{1}$$

Note that we can always take $W_{\mathrm{AB}}, W_{\mathrm{AC}} \in [\frac{1}{2}, 1]$. Evidently, W_{AB} is independent of Charlie. However, W_{AC} is not independent of Bob because he operates on the system before it reaches Charlie.

Bob's two operations ($y = 0$, 1) are described by the notion of a quantum instrument [45], which captures both the measurement statistics and the evolution of the measured state. A quantum instrument is defined as an ordered set of trace-non-increasing completely positive maps $\{\Lambda_{b|y}\}_b$ with the property that for any state ρ it holds that $p(b|y) = \mathrm{tr}[\Lambda_{b|y}(\rho)]$. Having observed the classical output b, the quantum output of the instrument is $\rho^{y,b} = \Lambda_{b|y}(\rho)/\mathrm{tr}[\Lambda_{b|y}(\rho)]$. Since we consider qubits and Bob has binary outcomes, the extremal quantum instruments are written as $\Lambda_{b|y}(\rho) = K_{b|y}\rho K_{b|y}^\dagger$, where $\{K_{b|y}\}_b$ are Kraus operators satisfying $\sum_b {K_{b|y}}^\dagger K_{b|y} = \mathbb{1}$, with the convenient property that ${K_{b|y}}^\dagger K_{b|y} = B_{b|y}$ where $\{B_{b|y}\}_b$ are the two POVMs of Bob [46]. For simplicity, we can represent Bob's measurements in terms of two observables which, in general, read $B_y \equiv B_{0|y} - B_{1|y} = \alpha_y \mathbb{1} + \vec{n}_y \cdot \vec{\sigma}$, where $\vec{n}_y$ are Bloch vectors, $\vec{\sigma}$ are the Pauli matrices, and $|\alpha_y| \leq 1 - |\vec{n}_y|$. The sharpness of Bob's measurements is

defined as $\eta_y = |\vec{n}_y|$. Notice that for $\eta_y \in \{0, 1\}$, the measurements are noninteractive and sharp, respectively, whereas $\eta_y \in (0, 1)$ corresponds to intermediate cases. We consider the case of $\eta \equiv \eta_0 = \eta_1$. We emphasize that one can stochastically simulate Bob's unsharp POVMs using only projective measurements, but one cannot simulate his quantum instrument in this manner. Therefore, we can distinguish a projective simulation from a genuine unsharp measurement by considering both the classical and quantum output.

By inspecting the witnesses $(W_{\mathrm{AB}}, W_{\mathrm{AC}})$, one may characterize the sharpness parameter η. References [38,39] showed that for a given value of W_{AB}, the optimal value of W_{AC} in quantum theory is given by

$$W_{\mathrm{AC}} = \frac{1}{8}\left(4 + \sqrt{2} + \sqrt{16W_{\mathrm{AB}} - 16W_{\mathrm{AB}}^2 - 2}\right), \quad (2)$$

and that such an optimal pair implies a precise value of η. However, in the experimentally realistic case in which perfectly optimal quantum correlations are not relevant, a suboptimal witness pair can be used to deduce upper and lower bounds on η,

$$\eta \geq \sqrt{2}(2W_{\mathrm{AB}} - 1) \equiv \eta_{\min},$$
$$\eta \leq 2\sqrt{(2 + \sqrt{2} - 4W_{\mathrm{AC}})(2W_{\mathrm{AC}} - 1)} \equiv \eta_{\max}. \quad (3)$$

Thus, the closer the experimentally observed correlations are to the optimal ones in Eq. (2), the narrower is the interval $I(W_{\mathrm{AB}}, W_{\mathrm{AC}}) \equiv [\eta_{\min}, \eta_{\max}]$ to which we can confine the sharpness η.

Experiment.—The optimal quantum correlations (2) are obtained with a unique quantum strategy (up to a global unitary) [38]. Alice needs to prepare four states forming a square on a great circle on the Bloch sphere. For simplicity we choose the xz plane and Alice's four states $|\psi_{x_0x_1}\rangle = \cos\alpha_{x_0x_1}|0\rangle + \sin\alpha_{x_0x_1}|1\rangle$, corresponding to the four values $\{(\pi/8), -(3\pi/8), (9\pi/8), (5\pi/8)\}$ of $\alpha_{x_0x_1}$, respectively, where $\rho_x = |\psi_{x_0x_1}\rangle\langle\psi_{x_0x_1}|$. Similarly, the optimal quantum instruments of Bob correspond to the Kraus operators $K_{b|y} = \sqrt{(\mathbb{1} + (-1)^b B_y)/2}$ for a suitably chosen η, where $B_y \in \{\eta\sigma_x, \eta\sigma_z\}$ are the corresponding observables of Bob. The quantum output is sent to Charlie whose observables are two complementary projective measurements $C_0 = \sigma_x$ and $C_1 = \sigma_z$. In an ideal experiment, for every η, we obtain the witness pair,

$$W_{\mathrm{AB}} = \frac{2 + \sqrt{2}\eta}{4}, \qquad W_{\mathrm{AC}} = \frac{4 + \sqrt{2} + \sqrt{2 - 2\eta^2}}{8}, \quad (4)$$

which satisfies the optimality condition (2).

We implemented this optimal strategy, using single-photon polarization qubits where the computational basis corresponds to horizontal (H) and vertical (V) polarization, i.e., $|H\rangle \equiv |0\rangle$ and $|V\rangle \equiv |1\rangle$. The complete optical setup is shown in Fig. 2. Alice's preparation device also encloses a heralded single photon source that produces photons at wavelength 780 nm through spontaneous parametric down conversion (SPDC) by pumping a type-I beta barium borate (BBO) single crystal of thickness 2 mm using 390 nm fs laser pulses. Time correlated idler and signal photons are spectrally and spatially purified by passing through 3 nm (FWHM) wide optical filters (F) and coupling into single mode fibers (SMF), respectively. The idler photons in mode "a" are detected by an avalanche photodiode (APD), marked as D_{Trigger}, with detection efficiency ~60%, which produces a trigger signal indicating the presence of a photon in mode "b." Alice prepares this photon in one of the four desired states $|\psi_{x_0x_1}\rangle$ using a polarizer when it only passes through $|H\rangle$ and a half-wave plate, HWP(A), at angles 11.25°, −11.25°33.75° and −33.75°, respectively, and sends it to Bob.

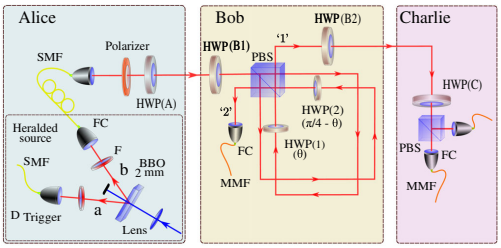

FIG. 2. Experimental setup. Alice prepares her states using a heralded photon source, a polarizer and a half-wave plate HWP (A). Bob's instrument is realized by a shifted Sagnac interferometer where the sharpness parameter $\eta = \cos(4\theta)$ is tuned using half-wave plates HWP(1) and HWP(2). HWP(B1) and HWP(B2) are used to switch between the observables B_0 and B_1 as well as selecting the output corresponding to the outcome $b = 0$ and $b = 1$. Charlie performs projective (sharp) measurements on the received qubit from Bob using a HWP(C) and a polarization beam splitter (PBS).

Bob's unsharp measurements on the received photons are performed using a shifted Sagnac interferometer as described in Refs. [30,36]. In this setup the strength of the measurement is controlled by rotating half-wave plate HWP(1) to θ and HWP(2) to $(\pi/4) - \theta$, that are placed, respectively, in the path of horizontally and vertically polarized beams after the polarization beam splitter (PBS) such that $\eta = \cos(4\theta)$. To switch between the bases B_y according to the input y, Bob rotates both his wave-plates HWP(B1) and HWP(B2) to 22.5° and 0°, respectively. The outcome of these measurements $b \in \{0, 1\}$ is encoded in the output path of the interferometer such that $b = 0$ ($b = 1$) corresponds to the detection of the photon in the output path "1" ≡ transmission ("2" ≡ reflection). In a sequential scenario, we choose to consider only one output path at a time to simplify the setup and by adding an additional rotation to the HWP(B1) and HWP(B2), we can

select the output we want to analyze at a given time. Using output 2, Bob will locally be able to learn the outcome of his measurement counterfactually when using perfect detectors. Also, when the fair sampling assumption is invoked, which is the case in this experiment, Bob can still infer his outcome of the measurement locally using average photon rates.

Finally, Charlie's projective measurement setup consists of HWP(C), PBS, a pair of fiber couplers (FC) and multimode fibers (MMF) that propagate the photons to a pair of APDs. He performs $C_z \in \{\sigma_x, \sigma_z\}$ on the received qubits according to his random input $z \in \{0, 1\}$, by rotating HWP(C) to 22.5° and 0°, respectively. The results of Charlie's marginal probabilities (for evaluating W_{AC}) are obtained by averaging out the inputs and outputs of Bob.

Results.—To evaluate $(W_{\mathrm{AB}}, W_{\mathrm{AC}})$ from the data, we require the marginal probabilities appearing in Eq (1). All parties setting are set using motorized rotation stages that are controlled by a computer program. To gather sufficient statistics we measure 60 sec in each setting with a rate of ∼20 kHz and collected at least 1.2 million events. Each measured value of $(W_{\mathrm{AB}}, W_{\mathrm{AC}})$ together with the (black color) error bars (horizontal and vertical corresponding to W_{AB} and W_{AC}, respectively) is shown in Fig. 3 and can be compared to the optimal quantum correlations (red color) and the optimal classical correlations (blue color, given by $(W_{\mathrm{AB}}, W_{\mathrm{AC}}) \leq 3/4$). Our obtained quantum correlations are nearly optimal for all considered values of η. Also, in the worst case, the classical limit is outperformed by at least 15 standard deviations. Moreover, the data reliably outperforms the optimal quantum correlations attainable when Bob uses stochastic projective measurements (green color) (see Ref. [39]). This certifies the advantage of unsharp measurements in sequential QRACs. Notably, the projective bound is not outperfromed for the two data points corresponding to $\eta \in \{0, 1\}$ since in these cases the bound coincides with the optimal quantum correlations.

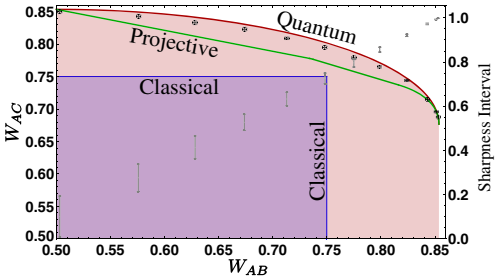

FIG. 3. Experimental results. Optimal quantum correlations (red), optimal quantum correlations from stochastic projective measurements (green), optimal classical correlations (blue), and experimentally obtained values of witness pairs $(W_{\mathrm{AB}}, W_{\mathrm{AC}})$ (black). The characterization of the sharpness parameter η is depicted by gray bars corresponding to the interval to which it is confined (y axis on the right-hand side).

From the inequalities in Eq. (3), we can determine an upper and a lower bound on the sharpness parameter. Thus, η can be confined to the interval $I(W_{\mathrm{AB}}, W_{\mathrm{AC}})$ for each of the measured values of the witness pair $(W_{\mathrm{AB}}, W_{\mathrm{AC}})$. These certified intervals are depicted by gray bars in Fig. 3 located vertically from the corresponding witness' and using the y axis on the right side. We observe that the certification is more precise (the interval is smaller) as the sharpness parameter increases. The smallest (largest) interval, corresponding to an essentially projective (noninteractive) measurement, has a width of about 10^{-3} (0.2). This is due to the fact that the bounds in Eq. (3) become more sensitive to small imperfections when W_{AC} increases. Further details about the experimental data is presented in the Supplemental Material [47]. Moreover, in Ref. [47], we also compare this characterization of unsharp measurements to a simple tomographic model with an essentially trusted preparation device subjected to comparably small errors.

Data analysis.—The experiment is influenced by systematic errors originating from, for instance, imperfect wave plates as well as offsets in their marked zero position, finite PBS extinction and cross talk, and limited interference visibility. The magnitude of these errors is revealed by the extent to which the experimental points are shifted away from the optimal quantum correlations. In order to minimize systematic errors, we carefully select and characterize all the optical components. This brings us closer to the optimal quantum correlation and the experimental points correspond to a more than 98% total visibility estimation. Nevertheless, random errors due to Poissonian statistics or due to repetition of the experimental settings with limited precision will spread the observed point on Fig. 3 to a region contained within the black bars. To keep this error low, all the settings are set by computerized controlled motors with repetition precision $< 0.02°$. Errors together with mean values are provided in the Supplemental Material [47].

Quantifying sequentual measurement incompatibility.—In order to witness quantum correlations, one requires incompatible measurements. In that sense, violating the classical constraint with W_{AB} (W_{AC}) certifies that Bob's (Charlie's) two POVMs are incompatible [48,49]. It is, however, more informative to consider a quantitative inference; is it possible to deduce from $(W_{\mathrm{AB}}, W_{\mathrm{AC}})$ a lower bound on the extent to which Bob's and Charlie's POVMs are incompatible? In order to achieve such quantification of Heisenberg uncertainty, one must first define a measure of incompatibility valid for dichotomic qubit observables. We use the *degree of incompatibility* introduced in Ref. [44];

$$D(\vec{n}_0, \vec{n}_1) = |\vec{n}_0 + \vec{n}_1| + |\vec{n}_0 - \vec{n}_1| - 2, \qquad (5)$$

where $\vec{n}_0$ and $\vec{n}_1$ are the Bloch vectors of the observables. All compatible observables obey $D \leq 0$ whereas incompatible observables obey $D \leq 2(\sqrt{2}-1)$. As expected, the bound is saturated by two Pauli observables. Since we are interested in incompatible observables, we simply reset negative values of D to 0. As shown in Supplemental Material [47], the success rate of a QRAC implies a lower bound on D:

$$D \geq 8W - 6. \tag{6}$$

Thus, whenever a QRAC exceeds the classical bound of $\frac{3}{4}$, a degree of incompatibility is certified and quantified. By choosing $W = W_{\mathrm{AB}}$, we use Eq (6) to quantify the incompatibility of Bob's unsharp measurements. The bound in Eq. (6) can also be applied to Charlie's measurements, but it would significantly underestimate their degree of incompatibility due to the sequential nature of the experiment. A more sophisticated quantification is possible when exploiting both W_{AB} and W_{AC} and the fact that $\eta \in I(W_{\mathrm{AB}}, W_{\mathrm{AC}})$. Considering unbiased observables for Bob, i.e., $B_y = \eta(\hat{n}_y \cdot \vec{\sigma})$, where $\hat{n}_y$ is the normalized Bloch vector, we show in the Supplemental Material [47] that Charlie's degree of incompatibility respects

$$D \geq \min_{\eta \in I(W_{\mathrm{AB}}, W_{\mathrm{AC}})} \frac{16W_{\mathrm{AC}} - 8}{1 + g_\eta + f_{W_{\mathrm{AB}}}(1 - g_\eta)} - 2 \tag{7}$$

where $g_\eta \equiv \sqrt{1-\eta^2}$ and $f_{W_{\mathrm{AB}}} \equiv 2(\eta_{\min}/\eta)\sqrt{1-(\eta_{\min}/\eta)^2}$.

Notice that if we choose not to exploit the certified interval $I(W_{\mathrm{AB}}, W_{\mathrm{AC}})$, we may simply take the limit of $\eta \to 0$ and recover the bound in Eq. (6) for $W = W_{\mathrm{AC}}$. In Fig. 4 we show the degree of incompatibility as obtained from the twelve experimentally measured witness pairs $(W_{\mathrm{AB}}, W_{\mathrm{AC}})$ corresponding to different targeted values of the sharpness parameter η. As expected, we see that the incompatibility of Bob's measurements decreases with η and vanishes in the vicinity of $\eta = 1/\sqrt{2}$, which is the theoretical threshold. For Charlie, we always find a high degree of incompatibility which stems from his projective measurements.

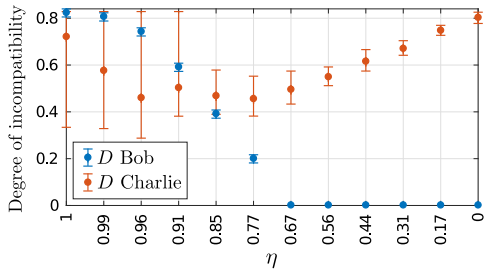

FIG. 4. Lower bound on the degree of incompatibility in Bob's (blue) and Charlie's (orange) respective pair of measurements for the twelve different targeted values of the sharpness parameter η.

Conclusions.—By precise control of unsharp quantum measurements, we demonstrated nearly optimal sequential quantum random access codes that outperform not only the best possible classical protocols but also the best possible quantum protocols based only on projective measurements. We harvested the quantum advantage in the communication task in order to certify the degree of unsharpness in the preformed measurements. Exploiting both the sequential QRACs and the certification of the unsharpness, we quantitatively demonstrated the incompatibility of two sequential pairs of measurements across a wide range of sharpness parameters. Our results demonstrate the usefulness of unsharp measurements in quantum communication tasks, the possibility of quantifying the degree of incompatibility of sequential pairs of unsharp observables and the practical feasibility of characterizing them under weak assumptions.

This work was supported by the Swedish Research Council, Knut and Alice Wallenberg Foundation, and the Swiss National Science Foundation (Starting Grant DIAQ, NCCR-QSIT). N. M. acknowledges the financial support by the Foundation for Polish Science through the First Team Grant No. 2016-1/5.

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, 10th anniversary ed. (Cambridge University Press, Cambridge, England, 2011).
[2] G. M. D'Ariano, P. L. Presti, and P. Perinotti, Classical randomness in quantum measurements, J. Phys. A **38**, 5979 (2005).
[3] S. M. Barnett and S. Croke, Quantum state discrimination, Adv. Opt. Photonics **1**, 238 (2009).
[4] R. Derka, V. Bužek, and A. K. Ekert, Universal Algorithm for Optimal Estimation of Quantum States from Finite Ensembles via Realizable Generalized Measurement, Phys. Rev. Lett. **80**, 1571 (1998).
[5] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, Symmetric informationally complete quantum measurements, J. Math. Phys. (N.Y.) **45**, 2171 (2004).
[6] J. B. Brask, A. Martin, W. Esposito, R. Houlmann, J. Bowles, H. Zbinden, and N. Brunner, Megahertz-Rate Semi-Device-Independent Quantum Random Number Generators Based on Unambiguous State Discrimination, Phys. Rev. Applied **7**, 054018 (2017).
[7] E. S. Gómez *et al.*, Device-Independent Certification of a Nonprojective Qubit Measurement, Phys. Rev. Lett. **117**, 260401 (2016).
[8] A. Tavakoli, D. Rosset, and M-O. Renou, Enabling Computation of Correlation Bounds for Finite-Dimensional Quantum Systems via Symmetrisation, Phys. Rev. Lett. **122**, 070501 (2019).
[9] J. M. Renes, Spherical-code key-distribution protocols for qubits, Phys. Rev. A **70**, 052314 (2004).

[10] M. Oszmaniec, L. Guerini, P. Wittek, and A. Acín, Simulating Positive-Operator-Valued Measures with Projective Measurements, Phys. Rev. Lett. **119**, 190501 (2017).

[11] A. Tavakoli, M. Smania, T. Vértesi, N. Brunner, and M. Bourennane, Self-testing non-projective quantum measurements in prepare-and-measure experiments, Sci. Adv. **6**, 16 (2020).

[12] R. Silva, N. Gisin, Y. Guryanova, and S. Popescu, Multiple Observers Can Share the Nonlocality of Half of an Entangled Pair by Using Optimal Weak Measurements, Phys. Rev. Lett. **114**, 250401 (2015).

[13] A. N. Korotkov, Continuous quantum measurement of a double dot, Phys. Rev. B **60**, 5737 (1999).

[14] J. Audretsch, T. Konrad, and A. Scherer, Sequence of unsharp measurements enabling a real-time visualization of a quantum oscillation, Phys. Rev. A **63**, 052102 (2001).

[15] T. Konrad, A. Rothe, F. Petruccione, and L. Diósi, Monitoring the wave function by time continuous position measurement, New J. Phys. **12**, 043038 (2010).

[16] K. W. Murch, S. J. Weber, C. Macklin, and I. Siddiqi, Observing single quantum trajectories of a superconducting quantum bit, Nature (London) **502**, 211 (2013).

[17] A. A. Clerk, M. H. Devoret, S. M. Girvin, F. Marquardt, and R. J. Schoelkopf, Introduction to quantum noise, measurement, and amplification, Rev. Mod. Phys. **82**, 1155 (2010).

[18] M. A. Armen, J. K. Au, J. K. Stockton, A. C. Doherty, and H. Mabuchi, Adaptive Homodyne Measurement of Optical Phase, Phys. Rev. Lett. **89**, 133602 (2002).

[19] J. S. Lundeen, B. Sutherland, A. Patel, C. Stewart, and C. Bamber, Direct measurement of the quantum wavefunction, Nature (London) **474**, 188 (2011).

[20] R. L. Cook, P. J. Martin, and J. M. Geremia, Optical coherent state discrimination using a closed-loop quantum measurement, Nature (London) **446**, 774 (2007).

[21] G. G. Gillett, R. B. Dalton, B. P. Lanyon, M. P. Almeida, M. Barbieri, G. J. Pryde, J. L. O'Brien, K. J. Resch, S. D. Bartlett, and A. G. White, Experimental Feedback Control of Quantum Systems Using Weak Measurements, Phys. Rev. Lett. **104**, 080503 (2010).

[22] C. Sayrin, I. Dotsenko, X. Zhou, B. Peaudecerf, T. Rybarczyk, S. Gleyzes, P. Rouchon, M. Mirrahimi, H. Amini, M. Brune, J-M. Raimond, and S. Haroche, Real-time quantum feedback prepares and stabilizes photon number states, Nature (London) **477**, 73 (2011).

[23] O. Oreshkov and T. A. Brun, Weak Measurements Are Universal, Phys. Rev. Lett. **95**, 110409 (2005).

[24] Y. Aharonov, D. Z. Albert, and L. Vaidman, How the Result of a Measurement of a Component of the Spin of a Spin-1/2 Particle can turn out to be 100, Phys. Rev. Lett. **60**, 1351 (1988).

[25] Y. Ota, S. Ashhab, and F. Nori, Entanglement amplification via local weak measurements, J. Phys. A **45**, 415303 (2012).

[26] F. J. Curchod, M. Johansson, R. Augusiak, M. J. Hoban, P. Wittek, and A. Acín, Unbounded randomness certification using sequences of measurements, Phys. Rev. A **95**, 020102(R) (2017).

[27] A. Tavakoli and A. Cabello, Quantum predictions for an unmeasured system cannot be simulated with a finite-memory classical system, Phys. Rev. A **97**, 032131 (2018).

[28] A. Bera, S. Mal, A. Sen(De), and U. Sen, Witnessing bipartite entanglement sequentially by multiple observers, Phys. Rev. A **98**, 062304 (2018).

[29] A. H. Shenoy, S. Designolle, F. Hirsch, R. Silva, N. Gisin, and N. Brunner, Unbounded sequence of observers exhibiting Einstein-Podolsky-Rosen steering, Phys. Rev. A **99**, 022317 (2019).

[30] H. Anwer, N. Wilson, R. Silva, S. Muhammad, A. Tavakoli, and M. Bourennane, Noise-robust preparation contextuality shared between any number of observers via unsharp measurements, arXiv:1904.09766.

[31] A. Palacios-Laloy, F. Mallet, F. Nguyen, P. Bertet, D. Vion, D. Esteve, and A. N. Korotkov, Experimental violation of a Bell's inequality in time with weak measurement, Nat. Phys. **6**, 442 (2010).

[32] F. Piacentini, A. Avella, M. P. Levi, M. Gramegna, G. Brida, I. P. Degiovanni, E. Cohen, R. Lussana, F. Villa, A. Tosi, F. Zappa, and M. Genovese, Measuring Incompatible Observables by Exploiting Sequential Weak Values, Phys. Rev. Lett. **117**, 170402 (2016).

[33] Y. Kim, Y-S. Kim, S-Y. Lee, S-W. Han, S. Moon, Y-H. Kim, and Y-W. Cho, Direct quantum process tomography via measuring sequential weak values of incompatible observables, Nat. Commun. **9**, 192 (2018).

[34] J.-S. Chen, M.-J. Hu, X-M. Hu, B.-H. Liu, Y.-F. Huang, C.-F. Li, C.-G. Guo, and Y.-S. Zhang, Experimental realization of sequential weak measurements of non-commuting Pauli observables, Opt. Express **27**, 6089 (2019).

[35] M. Schiavon, L. Calderaro, M. Pittaluga, G. Vallone, and P. Villoresi, Three-observer Bell inequality violation on a two-qubit entangled state, Quantum Sci. Technol. **2**, 015010 (2017).

[36] M.-J. Hu, Z.-Y. Zhou, X.-M. Hu, C.-F. Li, G.-C. Guo, and Y.-S. Zhang, Observation of non-locality sharing among three observers with one entangled pair via optimal weak measurement, Quantum Inf. **4**, 63 (2018).

[37] G. Foletto, L. Calderaro, A. Tavakoli, M. Schiavon, F. Picciariello, A. Cabello, P. Villoresi, and G. Vallone, Experimental Certification of Sustained Entanglement and Nonlocality after Sequential Measurements, Phys. Rev. Applied **13**, 044008 (2020).

[38] K. Mohan, A. Tavakoli, and N. Brunner, Sequential random access codes and self-testing of quantum measurement instruments, New J. Phys. **21**, 083034 (2019).

[39] N. Miklin, J. J. Borkała, and M. Pawłowski, Semi-device-independent self-testing of unsharp measurements, Phys. Rev. Research **2**, 033014 (2020).

[40] A. Ambainis, A. Nayak, A. Ta-Shama, and U. Varizani, Dense quantum coding and a lower bound for 1-way quantum automata, in *Proceedings of 31st ACM Symposium on Theory of Computing* (Association for Computing Machinery (ACM), New York, 1999), pp. 376–383, https://doi.org/10.1145/301250.301347.

[41] A. Ambainis, D. Leung, L. Mancinska, and M. Ozols, Quantum random access codes with shared randomness, arXiv:0810.2937.

[42] A. Tavakoli, A. Hameedi, B. Marques, and M. Bourennane, Quantum Random Access Codes using Single d-Level Systems, Phys. Rev. Lett. **114**, 170502 (2015).

[43] A. Tavakoli, J. Kaniewski, A. Vértesi, D. Rosset, and N. Brunner, Self-testing quantum states and measurements in the prepare-and-measure scenario, Phys. Rev. A **98**, 062307 (2018).

[44] P. Busch, P. Lahti, and R. F. Werner, Heisenberg uncertainty for qubit measurements, Phys. Rev. A **89**, 012129 (2014).

[45] T. Heinosaari and M. Ziman, *The Mathematical Language of Quantum Theory* (Cambridge University Press, Cambridge, England, 2011).

[46] J-P. Pellonpää, Quantum instruments: I. Extreme instruments, J. Phys. A **46**, 025302 (2013).

[47] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.125.080403 for experimental data, quantification of Bob's degree of incompatibility from a QRAC and quantification of Charlie's degree of incompatibility from two QRAC, and comparison to error-bounded detector tomography.

[48] A. Tavakoli and R. Uola, Measurement incompatibility and steering are necessary and sufficient for operational contextuality, Phys. Rev. Research **2**, 013011 (2020).

[49] C. Carmeli, T. Heinosaari, and A. Toigo, Quantum random access codes and incompatibility of measurements, Europhys. Lett. **130**, 50001 (2020).

Semi-device-independent certification of independent quantum state and measurement devices

Armin Tavakoli[1]

[1]*Département de Physique Appliquée, Université de Genève, CH-1211 Genève, Switzerland*

Certifying that quantum devices behave as intended is crucial for quantum information science. Here, methods are developed for certification of both state preparation devices and measurement devices based on prepare-and-measure experiments with independent devices. The experimenter assumes the independence of the devices and knowledge of the Hilbert space dimension. Thus no precise characterisation of any part of the experiment is required. The certification is based on a randomised version of unambiguous state discrimination and targets the class of state ensembles corresponding to quantum t-designs of any size and any dimension. These quantum designs are sets of states over which the average of any t-degree polynomial equals its average over all pure states, and they accommodate many of the most useful discrete structures in quantum information processing. Furthermore, it is shown that the same experiments also certify the detection efficiency of the measurement devices, as well as their non-projective nature. The presented methods can readily be implemented in experiments.

Introduction.— The precise control of quantum devices is crucial for the development of quantum technologies and experimental tests of foundational phenomena in quantum theory. Therefore, methods for certifying and characterising quantum devices are indispensable in quantum information science. Such methods allow one to ensure that, for example, a state preparation device indeed prepares the intended state or that a measurement device implements the desired measurements. The most common approaches are quantum state tomography [1] and quantum detector tomography [2]. In the former, the state emitted by a preparation device is measured in several different bases and the resulting outcome statistics is used to determine the state. In the latter, the measurement is determined from the outcome statistics obtained from probing it with different states. Therefore, the success of state (detector) tomography hinges on the auxiliary measurement (state preparation) device being precisely calibrated. Consequently, imperfections on experimentally relevant parameters in the auxiliary devices can undermine both state tomography [3] and detector tomography [4]. Moreover, in order to precisely calibrate the auxiliary device itself, one typically also requires a tomographic procedure which leads to an infinite regress.

The requirement of precise control in tomography can be overcome by more sophisticated certification methods. Methods have been developed for certifying states and measurements in experimental settings in which a sender prepares states and a receiver measures them but without neither device requiring a detailed characterisation [5]. Instead, the only assumption is that the Hilbert space dimension is known, which can often be justified from inspecting the specific experimental setup. This so-called semi-device-independent (SDI) approach to certification of quantum devices benefits from the fact that prepare-and-measure experiments are practical to implement [6–11] while also allowing for realistic experimental imperfections. A variety of SDI certification methods have been developed, e.g. for qubit states and measurements [5], pairs of mutually unbiased bases and symmetric informationally complete measurements [12, 13], non-projective qubit measurements [14, 15] and qubit quantum instruments [16, 17]. The practical viability of SDI certification schemes has also been experimentally demonstrated [4, 14, 18, 19].

With a few notable exceptions focused on dimension witnessing [20, 21] and random number generation [10], previous works on dimension-bounded quantum correlations in general, and SDI certification schemes in particular, have adopted models in which the involved quantum devices can be classically correlated in a stochastic and (to the experimenter) unknown manner. Such models make the set of quantum correlations convex, thereby considerably simplifying their analysis. However, these models correspond to a paranoid setting in which devices can classically conspire against the experimenter. In experiments that do not involve malicious parties, it is often natural to assume that separate quantum devices are independent.

In this work, we develop SDI certification methods for independent preparation and measurement devices. To this end, we present a SDI variant of the well-studied task of unambiguous state discrimination [22–24]. We prove that by observing optimal correlations in this task, one can certify the collection of states produced by the preparation device. The certification has a broad scope of relevance since it targets any number of states in any dimension that form a quantum t-design [25, 26]. A quantum t-design is a set of d-dimensional quantum states with the property that the average of any t-degree polynomial over the set equals the average taken over all pure states. These interesting and highly symmetric structures have broad applications in quantum information science. Examples include quantum tomography [27, 28], quantum key distribution [29, 30], Bell inequalities [31], entropic uncertainty relations [32] and entanglement detection [33, 34]. They also accommodate (as special cases) some of the most intensely researched and celebrated discrete Hilbert space structures such as rank-one generalised measurements, complete sets of mutually unbiased bases [35] and symmetric informationally complete sets of states [36], as well as the Platonic solids [37]. Moreover, we also use the same task to show that useful properties of the measurement device can be certified. Specifically, we show that one can certify the detection efficiency of the setup in a SDI manner. This is motivated by the fact that detectors do not always succeed with detecting an incoming physical system and that this is an important consideration in many quantum information protocols. Importantly, in order to make the certification of states and measurements experi-

mentally relevant, we develop its robustness to errors. Finally, we also exemplify the application of the scheme towards certification of non-projective measurements and show that it is substantially more robust to errors than established SDI certification schemes based on classically correlated devices.

Randomised unambiguous state discrimination.— Our platform for certifying quantum states and measurements is inspired by the textbook task of unambiguous state discrimination (USD). In USD, a sender (Alice) randomly chooses one of two possible states, $|\phi_1\rangle$ and $|\phi_2\rangle$, and sends it to a receiver (Bob). By measuring the incoming state, Bob tries to unambiguously decide which state he has received. Thus, he must either correctly identify the state or declare that he does not know the answer (inconclusive). His success rate is

$$p_{\text{usd}}^{\phi_1,\phi_2} = \frac{1}{2}\left(p(1|\phi_1) + p(2|\phi_2)\right), \tag{1}$$

while no errors are made, i.e. $p(1|\phi_2) = p(2|\phi_1) = 0$. Naturally, as soon as Alice's two states are not perfectly distinguishable (orthogonal), Bob cannot achieve a perfect success rate and must sometimes declare inconclusive rounds. It is well-known [22–24] that Bob's best measurement is $\{M^1, M^2, M^\perp\}$ where

$$M^1 = \frac{\mathbb{1} - |\phi_2\rangle\langle\phi_2|}{1 + |\langle\phi_1|\phi_2\rangle|}, \qquad M^2 = \frac{\mathbb{1} - |\phi_1\rangle\langle\phi_1|}{1 + |\langle\phi_1|\phi_2\rangle|}, \tag{2}$$

and $M^\perp = \mathbb{1} - M^1 - M^2$, where "$\perp$" denotes the inconclusive outcome. This leads to the optimal success rate

$$\max_M p_{\text{usd}}^{\phi_1,\phi_2} = 1 - |\langle\phi_1|\phi_2\rangle|. \tag{3}$$

In USD, the overlap of Alice's two states is assumed to be known. From this, we draw inspiration in order to construct the task of *randomised USD*, in which Alice's device requires no precise characterisation but is assumed to produce states of a known Hilbert space dimension.

In randomised USD, Alice generates a random input $x \in \{1, \ldots, N\}$ and subsequently produces a d-dimensional state ρ_x. The state is sent to Bob who generates a random pair of inputs $y \equiv (y_1, y_2)$; these can be any one of the $\binom{N}{2}$ ordered pairs ($y_1 < y_2$) of two positive integers no larger than N. He then implements a corresponding measurement $M_y = \{M_{b|y}\}_b$ with three possible outcomes $b \in \{1, 2, \perp\}$. The random input informs him that his task is to unambiguously discriminate between Alice's two states $\{\rho_{y_1}, \rho_{y_2}\}$. Thus, when $x \in \{y_1, y_2\}$, the task is to perform USD whereas otherwise the round is inconsequential. Therefore, for a given input y, the success rate is defined as

$$p_{\text{usd}}^{y} = \frac{1}{2}\left(p(1|y_1, y) + p(2|y_2, y)\right), \tag{4}$$

and the unambiguity requires $p(1|y_2, y) = p(2|y_1, y) = 0\ \forall y$, where the probabilities are given by the Born rule $p(b|x, y) = \text{tr}\left(\rho_x M_{b|y}\right)$. The performance in randomised USD is based on all the individual USD tasks (for each y). To specify the figure of merit, we first note that the rate of inconclusive rounds is simply $1 - p_{\text{usd}}^{y}$. We consider the moments of the rate of inconclusive events accumulated over all the individual USD tasks:

$$\mathcal{S}_t \equiv \sum_{y_1<y_2} \left(1 - p_{\text{usd}}^{y}\right)^{2t}, \tag{5}$$

where the integer $t \geq 1$ is the order of the moments. As shall soon become clear, this figure of merit is chosen due to its connection to quantum designs and for enabling a handy technical treatment. Thus, randomised USD is parameterised by the dimension d, the ensemble size N and the order t. Aiming to perform USD well for every y means that Alice and Bob aim to *minimise* $\mathcal{S}_t$. Importantly, we stress that when $N > d$, it is impossible for Alice to prepare her N states so that they are all pairwise distinguishable (trivialising the task).

Certifying quantum designs.— We show that randomised USD allows us to certify that Alice's states form a quantum t-design. While we presently restrict ourselves to an ideal setting in which Bob's discrimination is unambiguous for every y, we will later consider the more general case in which the discrimination features errors.

We use the fact that each of Bob's measurements apply to a single USD task (that corresponding to input y). This allows us to write

$$\min_{\text{quantum}} \mathcal{S}_t = \min_{\{\rho\}} \sum_{y_1<y_2} \left(1 - \max_{M_y} p_{\text{usd}}^{y}\right)^{2t}. \tag{6}$$

Leveraging the fact that the devices are independent, the minimal value of $\mathcal{S}_t$ is achieved with pure states. Therefore, in order to evaluate (6), we may write $\rho_x = |\psi_x\rangle\langle\psi_x|$. Although the states are d-dimensional, every pair $\{|\psi_{y_1}\rangle, |\psi_{y_2}\rangle\}$ can be viewed as an effective qubit embedded in the larger Hilbert space. Therefore, for every y, it is optimal for Bob to perform a measurement analogous to that in Eq (2). Hence, for given states of Alice and input y, the optimal success rate is analogous to Eq (3). A simple re-arrangement of the summation then gives

$$\min_{\text{quantum}} \mathcal{S}_t = -\frac{N}{2} + \frac{1}{2} \min_{\{\psi\}} \sum_{y_1,y_2} |\langle\psi_{y_1}|\psi_{y_2}\rangle|^{2t}. \tag{7}$$

At first sight, evaluating the right-hand-side seems challenging. However, the quantity subject to the minimisation is both well-studied and closely linked to quantum designs; it is commonly referred to as the (t'th-order) *frame potential* [38].

A quantum design is a set $\{|\phi_i\rangle\}$ of N pure d-dimensional states with the property that the average of any polynomial, g_t of degree t, taken over the set is identical to the average of the same polynomial taken over all pure d-dimensional states. That is,

$$\frac{1}{N}\sum_{i=1}^{N} g_t(\phi_i) = \int d\phi g_t(\phi), \tag{8}$$

where $d\phi$ is the Haar measure on the space of pure quantum states of dimension d. The polynomial g_t can be written as $g_t(\phi) = \langle\Phi|G_t|\Phi\rangle$ where $|\Phi\rangle = |\phi\rangle^{\otimes t}$ and G_t is some

bounded operator in the symmetric subspace of $(\mathbb{C}^d)^{\otimes t}$. It immediately follows that a quantum t-design also is a t'-design for $t' \leq t$. How does one determine whether a set of states is a quantum design? The answer is based on the frame potential. An ensemble of N d-dimensional states constitutes a quantum t-design if and only if it saturates the following lower bound on the frame potential [26]

$$V_t(\{\phi\}) \equiv \sum_{j,k=1}^{N} |\langle\phi_j|\phi_k\rangle|^{2t} \geq \frac{N^2 t!(d-1)!}{(t+d-1)!} \equiv J_t. \quad (9)$$

With this knowledge of quantum designs in hand, we can assert that the optimal quantum implementation of randomised USD obeys

$$\min_{\text{quantum}} \mathcal{S}_t \geq \frac{1}{2}(J_t - N) \equiv \mathcal{Q}_t. \quad (10)$$

The bound $\mathcal{Q}_t$ can be saturated if and only if Alice's states form a t-design of dimension d compsed of N states. Hence, this completes the certification. Moreover, notice that (10) also serves as a family of device-independent dimension witnesses for independent devices.

Certification under discrimination errors.— Since a small rate of failed discriminations (incorretly identifying the state) is to be expected in any realistic experiment, let us depart from the ideal situation and consider certification of the preparations when Bob's discrimination, for each input y, is subject to an error rate. We show that certification of quantum designs remains possible.

We adopt a model in which Alice's two pre-established equiprobable states $|\phi_1\rangle$ and $|\phi_2\rangle$ are to be discriminated in such a way that the rate of incorrect announcements associated to outcome 1 and 2 resepctively does not exceed $\epsilon \in [0, \frac{1}{2}]$. That is, the rate of error is bounded by $q_1 \leq \epsilon$ and $q_2 \leq \epsilon$ where

$$q_1 = \frac{p(1|\phi_2)}{p(1|\phi_1) + p(1|\phi_2)}, \quad q_2 = \frac{p(2|\phi_1)}{p(2|\phi_1) + p(2|\phi_2)}. \quad (11)$$

Evidently, standard USD corresponds to choosing $\epsilon = 0$. Interestingly, the problem of finding the optimal success rate (1) under the bounded error conditions has been solved [39]. Ref [39] found that the optimal success rate is given by

$$p_{\text{usd}}(\epsilon) = \begin{cases} \alpha_\epsilon \left(1 - |\langle\phi_1|\phi_2\rangle|\right) & \text{for} \quad \epsilon \leq \epsilon_{\text{c}} \\ \frac{1}{2}\left(1 + \sqrt{1 - |\langle\phi_1|\phi_2\rangle^2|}\right) & \text{for} \quad \epsilon_{\text{c}} \leq \epsilon, \end{cases} \quad (12)$$

where

$$\alpha_\epsilon = \frac{1-\epsilon}{(1-2\epsilon)^2}\left(1 + 2\sqrt{\epsilon(1-\epsilon)}\right) \quad (13)$$

and $\epsilon_{\text{c}} = 1/2\left(1 - \sqrt{1 - |\langle\phi_1|\phi_2\rangle|^2}\right)$.

Equipped with this, we can now certify the preparation device also when the error rate ϵ is found in the data. From the measured probabilities, one can appropriately choose ϵ. Then, based on the observed error rate, we modify the original figure of merit (5) so that it reads

$$\mathcal{S}_t^\epsilon = \sum_{y_1 < y_2} \left(\alpha_\epsilon - p_{\text{usd}}^y\right)^{2t}. \quad (14)$$

Notice that the error-free case ($\epsilon = 0$) returns the original figure of merit since $\alpha_0 = 1$. To find the optimal value of $\mathcal{S}_t^\epsilon$, we can recycle the reasoning in the previous section. One can account for the piecewise continuous feature in Eq (12) by noticing that α_ϵ is monotonically increasing and that the upper expression in (12) therefore serves as an upper bound on the lower expression when $\epsilon \geq \epsilon_{\text{c}}$. This can be applied to our problem for every y and any ϵ. Then, we arrive at the error-tolerant statement

$$\min_{\text{quantum}} \mathcal{S}_t^\epsilon \geq \alpha_\epsilon^{2t} \mathcal{Q}_t, \quad (15)$$

which generalises (10). The inequality can be saturated if and only if Alice's states form a t-design with the property that the relation $\epsilon \leq \epsilon_{\text{c}}$ is satisfied for all pairs of states. Hence, designs can be certified also in presence of discrimination errors. For example, a celebrated family of designs are known as symmetric informationally complete (SIC) [26]. They correspond to $t = 2$ and $N = d^2$ for any $d \geq 2$. Any such design can be certified through (15) as long as ϵ remains reasonably small. Specifically the bound is tight when $\epsilon \leq \frac{1}{2}\left(1 - \sqrt{\frac{d}{d+1}}\right)$. For qubits ($d = 2$), the critical error becomes $\epsilon \approx 9.2\%$ which is well above experimentally achieved error rates in USD [40].

Furthermore, consider a situation in which we observe an error rate ϵ but Alice's preparations nevertheless do not precisely form a design. Then, we can estimate how close they are to forming a design based on the measured value of $\mathcal{S}_t^\epsilon$. If we momentarily assume pure states a natural quantifier is the frame potential, which by arguments analogous to the above satisfies

$$V_t \leq N + \frac{2}{\alpha_\epsilon^{2t}} \mathcal{S}_t^\epsilon. \quad (16)$$

Hence, the closer $\mathcal{S}_t^\epsilon$ is to its optimum, the more accurate is the certification of the design structure. Notably, we can extend this to account also for the possibility of mixed states by expanding the domain of the frame potential. Define $\tilde{V}_t(\{\rho\}) \equiv \sum_{j,k=1}^{N} F(\rho_j, \rho_k)^{2t}$, where F denotes the fidelity. Since a pair of mixed states cannot be discriminated with success probability larger than $\alpha_\epsilon(1-F)$ when the error rate ϵ is allowed [41], it follows that also $\tilde{V}_t$ is bounded by the right-hand-side of Eq (16) and thus admits a robust certification.

Certifying detection efficiency.— We turn our attention to the measurement device. A realistic measurement device can be modelled as succeeding with performing the intended detection only with some probability $\eta \in [0, 1]$. As a typical example, a single-photon avalanche diode for visibile light has a detection efficiency around $\eta = 55\%$ [42]. Naturally, it is often practical to infer the efficiency by assuming a simple model for the detector and probing it with single photons.

Here, however, we consider the SDI situation in which the overall detection efficiency is bounded based solely on the statistics gathered in the considered experiments.

As before, we appropriately choose ϵ by inspecting the data $p(b|x,y)$ and accordingly consider the figure of merit (14). In order to certify the detection efficiency, we must consider the optimal value of $\mathcal{S}_t^\epsilon$ that is compatible with a hypothesised value of η. Since the measurement device is uncharacterised, it can internally map a failed detection onto the outputs $b \in \{1,2,\perp\}$ so that detection failure cannot be read out directly by the experimenter. If failed detections are outputted as $b=1$ or $b=2$, it will sometimes (in half the cases) give a wrong answer to the discrimination task. This sharply increases the error rate ϵ while making no better contribution to the discrimination than a trivial random guess. Therefore, the device optimally treats failed detections as inconclusive outcomes ($b=\perp$). From Eq (12), this causes the success probability in bounded-error discrimination to obey $p_{\text{usd}}^{\phi_1,\phi_2} \leq \eta\alpha_\epsilon\left(1-|\langle\phi_1|\phi_2\rangle|\right)$ with a possible equality when $\epsilon \leq \epsilon_{\text{c}}$.

Now, we can evaluate a bound on the optimal quantum value of $\mathcal{S}_t^\epsilon$ when subject to a given amount of detection loss. A simple calculation asserts the following useful inequality

$$(\alpha_\epsilon - \max_M p_{\text{usd}}^{\phi_1,\phi_2})^2 \geq \alpha_\epsilon^2\left((1-\eta)^2+\eta(2-\eta)|\langle\phi_1|\phi_2\rangle|^2\right). \tag{17}$$

Applying this inequality for every input y, we can bound the figure of merit as follows:

$$\begin{aligned}\mathcal{S}_t^\epsilon &\geq \min_{\{\psi\}}\alpha_\epsilon^{2t}\sum_{y_1<y_2}\left((1-\eta)^2+\eta(2-\eta)|\langle\psi_{y_1}|\psi_{y_2}\rangle|^2\right)^t = \frac{\alpha_\epsilon}{2}\\ &\times\left(-N+\min_{\{\psi\}}\sum_{n=0}^{t}\binom{t}{n}(1-\eta)^{2(t-n)}\eta^n(2-\eta)^n V_n(\{\psi\})\right)\\ &\geq \alpha_\epsilon^{2t}\left(-\frac{N}{2}+\frac{1}{2}\sum_{n=0}^{t}\binom{t}{n}(1-\eta)^{2(t-n)}\eta^n(2-\eta)^n J_n\right).\end{aligned} \tag{18}$$

In the second line we have used the binomial theorem and identified the n'th order frame potential, and in the third line we have used Eq (9). For simplicity, we write $\mathcal{S}_t^\epsilon \geq \alpha_\epsilon^{2t}\mathcal{Q}_t^\eta$ where $\mathcal{Q}_t^\eta$ denotes the bracket on the last line. Hence, for any observed $\mathcal{S}_t^\epsilon$, the following $2t$-degree polynomial must be positive: $P(\eta) \equiv \mathcal{S}_t^\epsilon - \alpha_\epsilon^{2t}\mathcal{Q}_t^\eta \geq 0$. To determine a lower bound on the detection efficiency, we must decide the values of η that respect the positivity of $P(\eta)$. This is achieved by finding the real-valued roots (in the interval $[0,1]$) of $P(\eta)$.

Application: detection efficiency based on SICs.— While the bound on η is typically not tight due to (17), it enables useful certification. We exemplify this through the previously considered one-parameter family of designs known as SICs (corresponding to $N=d^2$ with $t=2$ for any $d\geq 2$). For this family, we evaluate the relevant root of $P(\eta)$ and find that

$$\eta \geq \frac{\alpha_\epsilon d(d-1)-\sqrt{d-1}\sqrt{\sqrt{2}\sqrt{(d^2-1)\,\mathcal{S}_2^\epsilon}-\alpha_\epsilon^2 d(d-1)}}{\alpha_\epsilon d(d-1)} \tag{19}$$

Notice that an optimal implementation (possible when $\epsilon \leq \frac{1}{2}\left(1-\sqrt{\frac{d}{d+1}}\right)$) leads to $\mathcal{S}_2^\epsilon = \alpha_\epsilon^4\frac{d^2(d-1)}{2(d+1)}$ which inserted into Eq (19) implies perfect detection efficiency ($\eta=1$). In the other extreme, the bound only becomes trivial ($\eta\geq 0$) when $\mathcal{S}_2^\epsilon = \frac{\alpha_\epsilon^4}{2}(d^4-d^2)$, where the second factor is the algebraically maximal (trivial) value of the frame potential. For any intermediate value of $\mathcal{S}_2^\epsilon$, we obtain a non-trivial bound on η. Notably, this stands in contrast to certification of detection efficiency based on Bell inequality violations for which there exists a (often quite high [43]) threshold value for η below which no device-independent certification of detection efficiency can be made.

We exemplify the certification in a concrete implementation with imperfections. Take the qubit case of $d=2$ ($N=4$ and $t=2$) and consider that the ideal preparations of Alice and the ideal measurements of Bob are subject to some noise rate γ. For simplicity, let us concentrate all the noise in Bob's measurements: the optimal measurements for USD are only implemented with probability $1-\gamma$ whereas with probability γ the measurement corresponds to a random guess $b\in\{1,2\}$. In addition, we let Bob's device have a detection efficiency of $\eta_{\text{exp}}=55\%$ and let it treat failed outcomes as $b=\perp$. The combination of noise and detection loss leads to both errors (with probability $\gamma/2$) and sub-optimal correlations corresponding to $p_{\text{usd}}^y = \eta_{\text{exp}}\left((1-\gamma)\left(1-|\langle\psi_{y_1}|\psi_{y_2}\rangle|\right)+\gamma/2\right)$. Using that Alice's tetrahedral states have $|\langle\psi_{y_1}|\psi_{y_2}\rangle|^2 = 1/3$, we can establish the error ϵ through Eq (11) and evaluate the certified detection efficiency through Eq (19). For a nearly noise-free implemenetation ($\gamma=0.5\%$) we certify $\eta\geq 31.8\%$. For an order of magnitude higher noise rate ($\gamma=5\%$) we can still certify a detection efficiency of $\eta\geq 21.0\%$.

Certification of non-projective measurements.— An interesting feature of USD is that the optimal implementation uses non-projective measurements. Due to the increasing interest in non-projective measurements for quantum information applications, it is relevant to certify such measurement in SDI scenarios. In Supplementary Material we exemplify this for $(N,d,t)=(4,2,2)$ and show that a certification can be achieved for a detection efficiency of at least $\eta=\frac{3+\sqrt{3}}{6}\approx 78.9\%$. This threshold is notable since it is much lower than the nearly perfect detection efficiency required in other SDI schemes based on classically correlated quantum devices [13–15, 18, 44]

Conclusions.— I have developed methods for the certification of state preparation devices and measurement devices in prepare-and-measure experiments in which the devices are assumed to be independent. The presented scheme is versatile as it applies to three qualitatively different problems: i) certification of quantum states, ii) certification of detection efficiency and iii) certification of non-projective measurements. The certification is robust to errors and therefore applicable to experiments. Notably, small experimental deviations from the assumptions in the SDI scenario, such as memory effects in the detector or multiphoton events, can be accounted for using the method of Ref. [10].

The framework based on independent quantum devices departs from the more common setting in which devices can be

classically correlated. This is often natural when considering tasks that are not of adversarial nature. It is therefore relevant to develop such certification schemes targeting various useful properties of quantum systems. More generally, the loss of convexity that comes with the independence assumption makes it challenging to determine the limitations of quantum correlations and consequently also their applications towards various certification tasks. Here, our tool for overcoming this obstacle relied significantly on USD and the theory of quantum designs. It is of general interest to develop tools for characterising the set of quantum correlations without shared randomness. This would be both of foundational interest and a route to interesting protocols for quantum information processing.

Note added.— During the completion of this work, I became aware of the related work of Ref [45].

Acknowledgments

I thank Nicolas Brunner and Jonatan Bohr Brask for discussions and comments. This work was supported by the Swiss National Science Foundation (Starting grant DIAQ, NCCR-QSIT).

[1] K. Vogel and H. Risken, Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase, Phys. Rev. A **40**, 2847(R) (1989).
[2] A. Luis and L. L. Sánchez-Soto, Complete Characterization of Arbitrary Quantum Measurement Processes, Phys. Rev. Lett. **83**, 3573 (1999).
[3] D. Rosset, R. Ferretti-Schöbitz, J-D. Bancal, N. Gisin and Y-C. Liang, Imperfect measurement settings: Implications for quantum state tomography and entanglement witnesses, Phys. Rev. A **86**, 062325 (2012).
[4] H. Anwer, S. Muhammad, W. Cherifi, N. Miklin, A. Tavakoli and M. Bourennane, Experimental characterisation of unsharp qubit measurements in a semi-device-independent setting Phys. Rev. Lett. **125**, 080403 (2020).
[5] A. Tavakoli, J. Kaniewski, T. Vértesi, D. Rosset, and N. Brunner, Self-testing quantum states and measurements in the prepare-and-measure scenario, Phys. Rev. A **98**, 062307 (2018).
[6] M. Hendrych, R. Gallego, M. Micuda, N. Brunner, A. Acin, and J. P. Torres, Experimental estimation of the dimension of classical and quantum systems, Nature Physics **8**, 588 (2012).
[7] J. Ahrens, P. Badziag, A. Cabello, and M. Bourennane, Experimental device-independent tests of classical and quantum dimensions, Nature Physics **8**, 592 (2012).
[8] A. Tavakoli, A. Hameedi, B. Marques, and M. Bourennane, Quantum random access codes using single d-level systems, Phys. Rev. Lett. **114**, 170502 (2015).
[9] V. D'Ambrosio, F. Bisesto, F. Sciarrino, J. F. Barra, G. Lima, and A. Cabello, Device-independent certification of high-dimensional quantum systems, Phys. Rev. Lett. **112**, 140503 (2014).
[10] T. Lunghi, J. B. Brask, C. C. W. Lim, Q. Lavigne, J. Bowles, A. Martin, H. Zbinden, and N. Brunner, Self-testing quantum random number generator, Phys. Rev. Lett. **114**, 150501 (2015).
[11] D. Martínez, A. Tavakoli, M. Casanova, G. Cañas, B. Marques, and G. Lima, High-dimensional quantum communication complexity beyond strategies based on Bell's theorem, Phys. Rev. Lett. **121**, 150504 (2018).
[12] M. Farkas and J. Kaniewski, Self-testing mutually unbiased bases in the prepare-and-measure scenario, Phys. Rev. A **99**, 032316 (2019).
[13] A. Tavakoli, D. Rosset, and M-O. Renou, Enabling Computation of Correlation Bounds for Finite-Dimensional Quantum Systems via Symmetrization, Phys. Rev. Lett. **122**, 070501 (2019)
[14] A. Tavakoli, M. Smania, T. Vértesi, N. Brunner, and M. Bourennane, Self-testing non-projective quantum measurements in prepare-and-measure experiments, Science Advances **6**, 16 (2020).
[15] P. Mironowicz and M. Pawłowski, Experimentally feasible semi-device-independent certification of 4 outcome POVMs, Phys. Rev. A **100**, 030301 (2019)
[16] K. Mohan, A. Tavakoli and N. Brunner, Sequential random access codes and self-testing of quantum measurement instruments, New J. Phys. **21** 083034 (2019).
[17] N. Miklin, J. J. Borkała, and M. Pawłowski, Self-testing of unsharp measurements, Phys. Rev. Research **2**, 033014 (2020).
[18] M. Smania, P. Mironowicz, M. Nawareg, M. Pawłowski, A. Cabello, and M. Bourennane, Optica **7**, 123 (2020).
[19] G. Foletto, L. Calderaro, G. Vallone and P. Villoresi, Experimental demonstration of sequential quantum random access codes, Phys. Rev. Research **2**, 033205 (2020).
[20] J. Bowles, M. T. Quintino and N. Brunner, Certifying the dimension of classical and quantum systems in a prepare-and-measure scenario with independent devices, Phys. Rev. Lett. **112**, 140407 (2014).
[21] J. I. de Vicente, Shared randomness and device-independent dimension witnessing, Phys. Rev. A **95**, 012340 (2017).
[22] I. D. Ivanovic, How to differentiate between non-orthogonal states, Phys. Lett. A **123**, 257 (1987).
[23] D. Dieks, Overlap and distinguishability of quantum states, Phys. Lett. A **126**, 303 (1988).
[24] A. Peres, How to differentiate between non-orthogonal states, Phys. Lett. A **128**, 19 (1988).
[25] P. Delsarte, J. M. Goethals and J. J. Seidel, Spherical codes and designs, Geom. Dedicata **6**, 363 (1977).
[26] J. M. Renes, R. Blume-Kohout, A. J. Scott, C. M. Caves, Symmetric Informationally Complete Quantum Measurements, J. Math. Phys. **45**, 2171 (2004).
[27] A. Hayashi, T. Hashimoto and M. Horibe, Reexamination of optimal quantum state estimation of pure states, Phys. Rev. A **72**, 032325 (2005).
[28] A. Scott, Optimizing quantum process tomography with unitary 2-designs, J. Phys. A **41**, 055308 (2008).
[29] D. Bruß, Optimal Eavesdropping in Quantum Cryptography with Six States, Phys. Rev. Lett. **81**, 3018 (1998).
[30] J. M. Renes, Spherical-code key-distribution protocols for qubits, Phys. Rev. A **70**, 052314 (2004).
[31] A. Tavakoli, M. Farkas, D. Rosset, J-D. Bancal and J. Kaniewski, Mutually unbiased bases and symmetric informa-

tionally complete measurements in Bell experiments: Bell inequalities, device-independent certification and applications, arXiv:1912.03225
[32] A. Ketterer and O. Gühne, Entropic uncertainty relations from quantum designs, arXiv:1911.07533
[33] Z-W. Liu, S. Lloyd, E. Y. Zhu, and H. Zhu, Generalized Entanglement Entropies of Quantum Designs, Phys. Rev. Lett. **120**, 130502 (2018).
[34] J. Bae, B. C Hiesmayr and D. McNulty, Linking entanglement detection and state tomography via quantum 2-designs, New J. Phys. **21** 013012 (2019).
[35] T. Durt, B-G. Englert, I. Bengtsson and K. Życzkowski, On mutually unbiased bases, Int. J. Quantum Information 8, 535 (2010).
[36] C. A. Fuchs, M. C. Hoang, and B. C. Stacey, The SIC Question: History and State of Play, Axioms **21**, 6 (2017).
[37] A. Tavakoli and N. Gisin, The Platonic solids and fundamental tests of quantum mechanics, arXiv:2001.00188
[38] J. J. Benedetto and M. Fickus, Finite Normalized Tight Frames, Adv. Comput. Math. **18**, 357 (2003).
[39] A. Hayashi, T. Hashimoto, and M. Horibe, State discrimination with error margin and its locality, Phys. Rev. A **78**, 012333 (2008).
[40] R. B. M. Clarke, A. Chefles, S. M. Barnett and E. Riis, Experimental demonstration of optimal unambiguous state discrimination, Phys. Rev. A **63**, 040305(R) (2001).
[41] This is straightforwardly proven by first purifying the two mixed states, exploiting the known results for error-bounded discrimination of pure states [39] and lastly using Uhlmann's theorem.
[42] F. Villa, D. Bronzi, Y. Zou, C. Scarcella, G. Boso, S. Tisa, A. Tosi, F. Zappa, D. Durini, S. Weyers, U. Paschen and W. Brockherde, CMOS SPADs with up to 500 μm diameter and 55% detection efficiency at 420 nm, Journal of Modern Optics **61**, 102 (2014).
[43] N. Brunner and N. Gisin, Partial list of bipartite Bell inequalities with four binary settings, Phys. Lett. A **372**, 3162 (2008).
[44] E. S. Gómez, S. Gómez, P. González, G. Cañas, J. F. Barra, A. Delgado, G. B. Xavier, A. Cabello, M. Kleinmann, T. Vértesi, and G. Lima, Device-Independent Certification of a Nonprojective Qubit Measurement, Phys. Rev. Lett. **117**, 260401 (2016)
[45] N. Miklin and M. Oszmaniec, A universal scheme for robust self-testing in the prepare-and-measure scenario, arXiv:2003.01032 (03-03-2020).

Appendix A: Randomised USD with stochastic projective measurements

In order to bound the quantum performance of randomised USD under projective measurements, we must first remind ourselves of how projective measurements perform in standard USD. It is a well-known result that for any two pre-established equiprobable non-orthogonal pure states, the optimal USD under stochastic projective measurements is obtained by randomly measuring either the eigenbasis of the first state or the eigenbasis of the second state. This leads to

$$\max_{\text{projective}} p_{\text{usd}} = \frac{1-|\langle\phi_1|\phi_2\rangle|^2}{2}. \tag{A1}$$

Naturally, in the special case of orthogonal states, the USD has a unit success rate.

We apply this to randomised USD. For every input y, the best success rate reads

$$\max_{\text{projective}} p^y_{\text{usd}} = \frac{1}{2}\left(1-|\langle\psi_{y_1}|\psi_{y_2}\rangle|^2+\tau_{\psi_{y_1},\psi_{y_2}}\right), \tag{A2}$$

where the special case of orthogonal state is accounted for by defining $\tau = 1$ if and only if $\langle\psi_{y_1}|\psi_{y_2}\rangle = 0$ and otherwise $\tau = 0$. For any given set of preparations, we can now evaluate the best performance in randomised USD for stochastic projective measurements to be

$$\begin{aligned}\mathcal{S}_t(\{\psi\}) &= \frac{1}{2^{2t}}\sum_{y_1<y_2}\left(1+|\langle\psi_{y_1}|\psi_{y_2}\rangle|^2-\tau_{\psi_{y_1},\psi_{y_2}}\right)^{2t}\\ &= \frac{1}{2^{2t}}\sum_{y_1<y_2}\sum_{n=0}^{2t}\binom{2t}{n}|\langle\psi_{y_1}|\psi_{y_2}\rangle|^{2n}\left(1-\tau_{\psi_y,\psi_{y'}}\right)^{2t-n}\\ &= \frac{1}{2^{2t+1}}\sum_{n=1}^{2t}\binom{2t}{n}(V_n-N)+\frac{1}{2^{2t}}\sum_{y_1<y_2}\left(1-\tau_{\psi_y,\psi_{y'}}\right).\end{aligned} \tag{A3}$$

In order to obtain a bound valid for all projective measurements and all state ensembles, we must find a lower bound on the above expression valid for all states. However, this appears not to be straightforward due to the right-most term.

However, by focusing on the most relevant case of qubits we can solve the problem. Consider the example of $(N,d,t)=(4,2,2)$. Since we only have four states, the number of possible pairwise orthogonalities is small. One can exhaustively consider the different orthogonality configurations that influence the right-most term in (A3). This straightforwardly leads to the finding that the optimal configuration features no orthogonalities among the four states. We can then obtain a lower bound on (A3) via the global lower bound on the frame potential (it gives $\mathcal{S}_2 \geq 11/10$). However, this bound is sub-optimal since four qubit states cannot be used to form the 3- or 4-design that appear in the final expression in (A3). A better bound can be obtained by directly exploiting the Bloch sphere parameterisation to reliably minimise the final expression in (A3). This leads to an optimal configuration being four Bloch vectors pointing to the vertices of a tetrahedron (quantum 2-design). Since this means $|\langle\psi_{y_1}|\psi_{y_2}\rangle|^2 = 1/3$ for $y_1 \neq y_2$. Inserted into Eq (A3), we obtain that projective measurements must obey $\mathcal{S}_2 \geq 32/27$.

The sizable gap between $\mathcal{S}_2 \geq 32/27$ and the best quantum result at $\mathcal{S}_2 = 2/3$ allows for the certification of non-projectiveness to be robust to errors. Consider for instance that Bob's detectors succeed with probability η. Due to the unambiguity of the discimniation, failed events must be mapped to $b = \perp$. We therefore have that $p^y_{\text{usd}} = \eta\left(1-|\langle\psi_{y_1}|\psi_{y_2}\rangle|\right) = \eta\frac{\sqrt{3}-1}{\sqrt{3}}$. We therefore have

$$\mathcal{S}_t = 6\times\left[1-\eta\frac{\sqrt{3}-1}{\sqrt{3}}\right]^4. \tag{A4}$$

The critical value of η for certifying the implementation of non-projective measurements is obtained from solving $\mathcal{S}_t =$

$32/27$. The critical detection efficiency becomes

$$\eta = \frac{3+\sqrt{3}}{6} \approx 78.9\%. \tag{A5}$$

Bilocal Bell inequalities violated by the quantum Elegant Joint Measurement

Armin Tavakoli,[1] Nicolas Gisin,[1] and Cyril Branciard[2]

[1]*Département de Physique Appliquée, Université de Genève, CH-1211 Genève, Switzerland*
[2]*Université Grenoble Alpes, CNRS, Grenoble INP, Institut Néel, 38000 Grenoble, France*

Network Bell experiments give rise to a form of quantum nonlocality that conceptually goes beyond Bell's theorem. We investigate here the simplest network, known as the bilocality scenario. We depart from the typical use of the Bell State Measurement in the network central node and instead introduce a family of symmetric iso-entangled measurement bases that generalise the so-called Elegant Joint Measurement. This leads us to report noise-tolerant quantum correlations that elude bilocal variable models. Inspired by these quantum correlations, we introduce network Bell inequalities for the bilocality scenario and show that they admit noise-tolerant quantum violations. In contrast to many previous studies of network Bell inequalities, neither our inequalities nor their quantum violations are based on standard Bell inequalities and standard quantum nonlocality. Moreover, we pave the way for an experimental realisation by presenting a simple two-qubit quantum circuit for the implementation of the Elegant Joint Measurement and our generalisation.

Introduction.— The violation of Bell inequalities is a hallmark property of quantum theory. It asserts that the predictions of quantum theory cannot be accounted for by any physical model based only on local variables [1]. Such violations, referred to as quantum nonlocality, do not only provide insights in the foundations of quantum theory, but they also constitute a powerhouse for a broad scope of applications in quantum information science [2].

A standard Bell experiment features a source that emits a pair of particles shared between two space-like separated observers who perform local and independent measurements. In quantum theory the particles can be entangled, thus enabling global randomness [3]. In contrast, in local variable models aiming to simulate the quantum predictions, the particles are endowed with classically correlated stochastic properties that locally determine the outcome of a given measurement. Many decades of research on Bell inequalities have brought a relatively deep understanding of quantum nonlocality and have established standard methods for characterising correlations in both quantum models and local variable models [2].

The last decade witnessed a significant conceptual advance: much attention was directed at going beyond correlations in standard Bell experiments in favour of investigating correlations in networks featuring many observers and several independent sources of particles [4, 5]. While a standard Bell experiment may be viewed as a trivial network (with a single source), the introduction of multiple independent sources is conceptually interesting since it brings into play new physical ingredients and corresponds to the topology of future quantum networks. In contrast to standard Bell experiments, network Bell experiments feature some observers who hold independent particles (from different sources) and therefore *a priori* share no correlations. Moreover, entanglement can be distributed in the network, in particular to initially independent observers, through the process of entanglement swapping [6]. Recent years have seen much attention being directed at characterising classical, quantum and post-quantum correlations in networks, many times through the construction of network Bell inequalities and the exploration of their quantum violations [7–26]. In general, this is challenging due to the fact that the introduction of multiple independent sources makes the set of local variable correlations non-convex [4].

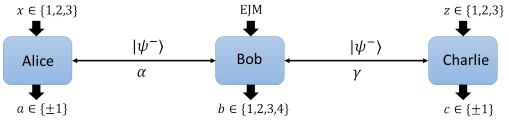

FIG. 1: Bilocality scenario: Bob independently shares a "state" with Alice and Charlie, respectively. In a quantum experiment, these are independent, typically entangled quantum states ($|\psi^-\rangle$), while in a bilocal model these are associated to independent local variables (α and γ).

Here, we focus on the simplest nontrivial network Bell experiment, known as the *bilocality scenario*. It features two independent sources that each produce a pair of particles. The first pair is shared between observers Alice and Bob while the second pair is shared between Bob and another observer, Charlie (see Figure 1). Interestingly, there are known Bell inequalities for the bilocality scenario (bilocal inequalities), i.e. inequalities for the observed correlations that are satisfied by all local variable models respecting the independence of the two sources. Importantly, these inequalities are also known to admit quantum violations. The quantum violations typically arise from Bob implementing a Bell State Measurement (BSM, encountered in quantum teleportation [27] and entanglement swapping [6]). Conspicuously, both the inequalities and their reported violations strongly resemble those encountered in the standard Bell experiments (see e.g. [7, 18, 19, 28]). For instance the standard bilocal inequality, first presented in Ref. [7], is essentially built on the Clauser-Horne-Shimony-Holt (CHSH) inequality [29] and its quantum violations through the BSM turn out to effectively correspond to Bob in a coordinated manner separately testing the CHSH inequality with Alice and Charlie respectively. Indeed, the BSM measurement amounts to performing simultaneously the two commuting measurements of $\sigma_1 \otimes \sigma_1$ and $\sigma_3 \otimes \sigma_3$ (where $(\sigma_1, \sigma_2, \sigma_3)$ are the three Pauli observables) on Bob's two independent qubits, and ample numerical evidence shows that the optimal measurements settings for Alice and Charlie are at ± 45 degrees on the Bloch sphere, i.e. ex-

actly those settings tailored for the CHSH inequality. Given this close resemblance to the CHSH inequality, it is perhaps unsurprising that the critical singlet visibility, required for two identical noisy singlet states to enable a violation, is the same as that encountered in the CHSH inequality, namely $\frac{1}{\sqrt{2}}$ for each state.

Here we investigate quantum nonlocality in the bilocality scenario that is not based on the BSM and does not directly trace back to standard quantum nonlocality as in the previous cases. To this end, we present a family of two-qubit entangled measurements generalising the so-called Elegant Joint Measurement (EJM) [30]. These allow Bob to effectively distribute (in an entanglement swapping scenario) different entangled states to Alice and Charlie from those obtained through a BSM. We investigate bilocal models for the resulting correlations, show explicit quantum violations of bilocality and obtain the critical visibility per singlet for a quantum violation. Subsequently, we introduce new bilocal inequalities tailored to our quantum correlations and show that they can detect quantum nonlocality in the network at reasonable singlet visibilities. Furthermore, towards experimental demonstrations of quantum violations of network Bell inequalities, that are not based on standard Bell inequalities, we explore the implementation of our generalised EJM. We prove that it cannot be implemented in linear optical schemes without auxiliary photons but that it can be implemented with a simple two-qubit quantum circuit.

Entangled measurements with tetrahedral symmetry.— We consider symmetric entangled measurements on two qubits that, most naturally, have four outcomes. Specifically, we present a family of bases $\{|\Phi_b^\theta\rangle\}_{b=1}^4$ of the two-qubit Hilbert space, parametrised by $\theta \in [0, \frac{\pi}{2}]$, such that all elements are equally entangled and, moreover, the four local states, corresponding to either qubit being traced out, form a shrunk regular tetrahedron inside the Bloch sphere.

To construct such bases, let us first introduce the pure qubit states $|\vec{m}_b\rangle$ that point (on the Bloch sphere) towards the four vertices

$$\begin{aligned} \vec{m}_1 &= (+1,+1,+1)\,, \quad \vec{m}_2 = (+1,-1,-1)\,, \\ \vec{m}_3 &= (-1,+1,-1)\,, \quad \vec{m}_4 = (-1,-1,+1) \end{aligned} \tag{1}$$

of a regular tetrahedron, as well as the states $|-\vec{m}_b\rangle$ with the antipodal direction. Specifically, we write these tetrahedron vertices in cylindrical coordinates as $\vec{m}_b = \sqrt{3}\left(\sqrt{1-\eta_b^2}\cos\varphi_b, \sqrt{1-\eta_b^2}\sin\varphi_b, \eta_b\right)$ and define

$$|\pm\vec{m}_b\rangle = \sqrt{\frac{1\pm\eta_b}{2}}e^{-i\varphi_b/2}|0\rangle \pm \sqrt{\frac{1\mp\eta_b}{2}}e^{i\varphi_b/2}|1\rangle. \tag{2}$$

Our family of generalised EJM bases, with the above properties, is then given by

$$|\Phi_b^\theta\rangle = \frac{\sqrt{3}+e^{i\theta}}{2\sqrt{2}}|\vec{m}_b, -\vec{m}_b\rangle + \frac{\sqrt{3}-e^{i\theta}}{2\sqrt{2}}|-\vec{m}_b, \vec{m}_b\rangle \tag{3}$$

Notice that for $\theta = 0$, we obtain the EJM introduced in Ref. [30] (the largest local tetrahedron in our family, of radius $\frac{\sqrt{3}}{2}$), while for $\theta = \frac{\pi}{2}$, we obtain the BSM (the smallest local tetrahedron, of radius zero) up to local unitaries (which can for instance be chosen as $U_1 \otimes U_2 = \mathbb{1} \otimes e^{\frac{2\pi i}{3}\frac{\sigma_1+\sigma_2+\sigma_3}{\sqrt{3}}}$ to recover the standard BSM). By varying θ, we thus continuously interpolate between the EJM and the BSM.

Quantum correlations.— We consider a specific quantum implementation of the bilocality experiment illustrated in Figure 1. We let Bob apply the generalised EJM and consider that both sources emit pairs of qubits corresponding to noisy singlets (so-called Werner states [31])

$$\rho_i = V_i|\psi^-\rangle\langle\psi^-| + \frac{1-V_i}{4}\mathbb{1}, \tag{4}$$

for $i \in \{1,2\}$ where $V_i \in [0,1]$ denotes the visibility of each singlet $|\psi^-\rangle = \frac{1}{\sqrt{2}}(|0,1\rangle - |1,0\rangle)$. By applying his measurement onto distributed (pure) singlets, Bob effectively prepares Alice's and Charlie's joint state in an entangled state similar to that of Eq. (3), up to a change in signs for $\vec{m}_b$ and θ. Due to the tetrahedral structure of the distributed states we expect to find strong correlations between Alice and Charlie when they perform measurements of the three Pauli observables [32]. We therefore let each of them have three possible measurement settings $x, z \in \{1,2,3\}$ (corresponding to the observables (σ_x, σ_z)), with binary outcomes denoted $a, c \in \{+1,-1\}$.

To reflect the symmetry of our scenario, it is convenient to identify Bob's outcome b with the corresponding vector $\vec{m}_b$ from Eq. (1), i.e., to write b as ±1-valued 3-vector $b = (b^1, b^2, b^3)$. The conditional probability distribution $p(a,b,c|x,z)$ obtained in the experiment can then be characterised in terms of the single-, two- and three-party correlators $\langle A_x\rangle$, $\langle B^y\rangle$, $\langle C_z\rangle$, $\langle A_xB^y\rangle$, $\langle B^yC_z\rangle$, $\langle A_xC_z\rangle$ $(= \langle A_x\rangle\langle C_z\rangle$ in the bilocality scenario) and $\langle A_xB^yC_z\rangle$ for all $x,y,z \in \{1,2,3\}$, with e.g. $\langle A_xB^yC_z\rangle = \sum_{a,b^1,b^2,b^3,c=\pm1} a\, b^y\, c\, p(a,b,c|x,z)$ and similarly for the other correlators [33]. For the quantum correlation p_{Q}^θ obtained from our above choice of states and measurements, these correlators become

$$\begin{aligned} &\langle A_x\rangle = \langle B^y\rangle = \langle C_z\rangle = \langle A_xC_z\rangle = 0, \\ &\langle A_xB^y\rangle = -\tfrac{V_1}{2}\cos\theta\,\delta_{x,y}, \quad \langle B^yC_z\rangle = \tfrac{V_2}{2}\cos\theta\,\delta_{y,z}, \\ &\langle A_xB^yC_z\rangle = \begin{cases} -\frac{V_1V_2}{2}(1+\sin\theta) & \text{if } xyz \in \{123,231,312\} \\ -\frac{V_1V_2}{2}(1-\sin\theta) & \text{if } xyz \in \{132,213,321\} \\ 0 & \text{otherwise} \end{cases}, \end{aligned} \tag{5}$$

where δ is the Kronecker symbol.

Simulating quantum correlations in bilocal models.— Let us first investigate whether the quantum probability distribution p_{Q}^θ admits a bilocal model. In such a model, each pair of particles is associated to a local variable denoted α and γ respectively (see Figure 1). Alice's (Charlie's) outcome is determined by her (his) setting and α (γ). Since they each have three possible settings, we can without loss of generality represent the local variables as triples $\alpha = (\alpha_1, \alpha_2, \alpha_3)$ and $\gamma = (\gamma_1, \gamma_2, \gamma_3)$ with entries ±1, with each α_x, γ_z denoting Alice's or Charlie's deterministic outcome for the setting x or

z. A bilocal model can thus be written as

$$p_{\text{biloc}}(a,b,c|x,z) = \sum_{\alpha,\gamma} q_\alpha^{(1)} q_\gamma^{(2)} \delta_{a,\alpha_x} \delta_{c,\gamma_z} p(b|\alpha,\gamma), \quad (6)$$

where $\{q_\alpha^{(1)}\}_\alpha$ and $\{q_\gamma^{(2)}\}_\gamma$ are probability distributions representing the stochastic nature of the local variables α and γ respectively, and $p(b|\alpha,\gamma)$ are probability distributions representing the stochastic response of Bob upon receiving (α,γ).

The central question is whether the quantum correlations characterised by Eq. (5) can be simulated in a bilocal model. We investigate the matter with three different approaches. Firstly, we set $V \equiv V_1 = V_2$ (equal noise on both sources), and $\theta = 0$ (as in the original EJM [30]). By employing semidefinite relaxations of the set of bilocal correlations, one can obtain a necessary condition for the existence of a bilocal model [34]. An evaluation of the relevant semidefinite program guarantees that a violation of bilocality is obtained whenever $V \gtrsim 83\,\%$ [35]. However, this bound is not expected to be tight due to the non-convex nature of the set of quantum correlations with independent sources.

Secondly, we provide a better characterisation of the power of bilocal models by explicitly considering their ability to simulate the quantum correlations. To this end, we have used an efficient search method which exploits that the numerical difficulties associated with the bilocality assumption are significantly reduced if the bilocal model first undergoes a Fourier transformation [7]. For the case of $V \equiv V_1 = V_2$ and $\theta = 0$ considered above, we look for the largest V for which $p_{\text{Q}}^{\theta=0}$ admits a bilocal model, and find the critical visibility

$$V_{\text{crit}} \approx 79.1\,\%. \quad (7)$$

To further explore different values of (V_1, V_2), we then also consider, for a given V_1, the largest V_2 for which a bilocal model exists. Figure 2 shows the region in the (V_1, V_2)-plane for which we find a bilocal simulation of p_{Q}^{θ} (still for $\theta = 0$ here; the analysis for $\theta > 0$ is presented in Appendix A). It also displays the product V_1V_2 associated to the boundary of the bilocal region (the critical pairs). Interestingly, the product of the critical visibilities is not constant. This is in stark contrast with previously studied quantum correlations that arise from the BSM [7] for which the product of visibilities determines the existence of a bilocal model. Notably, also the violations of many bilocal inequalities that are based on coordinated tests of standard Bell inequalities [7, 9, 18, 19, 28] are determined by such products of visibilities.

Thirdly, we employ an intuitive ansatz for analytically constructing bilocal models that mimic the symmetry of $p_{\text{Q}}^{\theta=0}$. Namely, we impose that the (unobserved) probability distribution $p_{\text{biloc}}(\alpha,b,\gamma) = q_\alpha^{(1)} q_\gamma^{(2)} p(b|\alpha,\gamma)$ of the bilocal model should have the same tetrahedral symmetry: for every permutation π of the tetrahedron vertices $\{\vec{m}_b\}$ in Eq. (1), extended to the opposite vertices via $\pi(-\vec{m}_b) = -\pi(\vec{m}_b)$, and applied to the 3-vector variables α, b, γ, one should have $p_{\text{biloc}}(\pi(\alpha),\pi(b),\pi(\gamma)) = p_{\text{biloc}}(\alpha,b,\gamma)$. Under this symmetry ansatz, we are able to analytically construct efficient bilocal simulations of $p_{\text{Q}}^{\theta=0}$. Interestingly, along the entire boundary of the bilocal region, the obtained results match those presented in Figure 2 up to the fifth decimal digit. This shows that simple and highly symmetric bilocal models are very nearly optimal for simulating $p_{\text{Q}}^{\theta=0}$. These bilocal models and the critical visibilities are detailed in Appendix B.

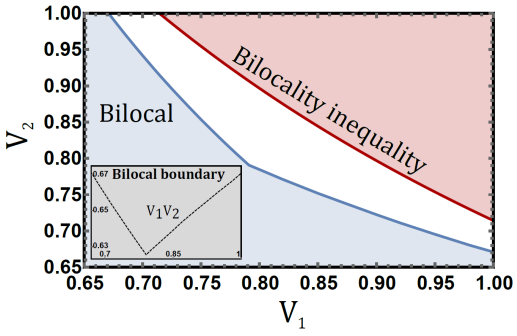

FIG. 2: The blue region represents the set of bilocal quantum correlations $p_{\text{Q}}^{\theta=0}$ in the plane of visibilities (V_1, V_2), with the dashed line in the inset figure showing the product of the visibilites on the boundary of this bilocal region. The red area is the part of the quantum region that can be detected as non-bilocal through the violation of our bilocal inequality (9).

Bilocal Bell inequalities.— We now draw inspiration from the structure of the nonbilocal quantum corelations obtained from the EJM to construct a bilocal inequality. Hence, in contrast to several previous bilocal inequalities, the present one is neither based on, nor apparently resembles, a standard Bell inequality. Also, naturally, such an inequality applies to detecting the non-bilocality of general probability distributions, not only p_{Q}^{θ}. To build the Bell expression, we introduce the following quantities

$$S = \sum_{y=z} \langle B^y C_z \rangle - \sum_{x=y} \langle A_x B^y \rangle,$$

$$T = \sum_{x\neq y\neq z\neq x} \langle A_x B^y C_z \rangle, \quad Z = \max\left(\mathcal{C}_{\text{other}}\right), \quad (8)$$

where $\mathcal{C}_{\text{other}} = \{|\langle A_x\rangle|, |\langle A_x B^y\rangle|, \ldots, |\langle A_x B^y C_z\rangle|\}$ is the set of the absolute values of all one-, two- and three-party correlators other than those appearing in the expressions of S and T. This leads us to the following bilocal inequality:

$$\mathcal{B} \equiv \frac{S}{3} - T \overset{\text{biloc}}{\leq} 3 + 5Z. \quad (9)$$

Notice that the Z quantity makes this general inequality nonlinear. The most interesting case is however when $Z = 0$, as satisfied by the quantum correlation of Eq. (5). For this case, we have proved the bilocal bound under the previously considered symmetry ansatz (which in fact enforces $Z = 0$, see Appendix C). Then, we have also confirmed the bilocal bound using two different numerical methods applied to general bilocal models [36]. We find that the bilocal inequality above, for $Z = 0$, is tight in the sense that it constitutes one of the facets of the projection of the "$Z = 0$ slice" of the bilocal set of correlations onto the (S,T)-plane. Remarkably,

this projection of the $Z = 0$ slice is delimited by linear inequalities, as further described in Appendix D; this stands in contrast to previous bilocal inequalities which use nonlinear Bell expressions. Finally, for $Z > 0$, we have again applied the same numerical search methods to justify the correction term $5Z$ in the bilocal bound of $\mathcal{B}$. Notably, more accurate corrections are also possible (see Appendix E).

For our quantum correlation of Eq. (5), we straightforwardly obtain $(S, T, Z) = \left(3\frac{V_1+V_2}{2}\cos\theta, -3V_1V_2, 0\right)$, and $\mathcal{B} = 3V_1V_2 + \frac{V_1+V_2}{2}\cos\theta$. In the noiseless case ($V_1 = V_2 = 1$), we thus get $\mathcal{B} = 3 + \cos\theta$, which gives a violation of our bilocal inequality (9) for our whole family of generalised EJM (i.e., the whole range of θ), except for the special case of a BSM ($\theta = \frac{\pi}{2}$, for which our quantum correlation turns out to be bilocal: see Appendix A). In contrast, when white noise is present and both sources are equally noisy ($V \equiv V_1 = V_2$), we get a violation of our inequality whenever $3V^2 + V\cos\theta > 3$. For $\theta = 0$, the critical visibility per singlet required for a violation is

$$V_{\text{crit}} = \frac{\sqrt{37}-1}{6} \approx 84.7\,\%. \tag{10}$$

This shows that the quantum violation is robust to white noise on the singlet states, but not optimally robust as no violation is found here for $V \in [0.791, 0.847]$. More generally, the bilocal inequality enables the detection of quantum correlations in a sizable segment of the (V_1, V_2)-plane (see Figure 2).

Finally, we note that several different bilocal inequalities can be constructed based on the correlations from the EJM. As another example, in Appendix F we consider the following Bell expression

$$\mathcal{B}' \equiv \sum_{x,b}\sqrt{p(b)\left(1 - b^x E_b^{\mathrm{A}}(x)\right)} + \sum_{z,b}\sqrt{p(b)\left(1 + b^z E_b^{\mathrm{C}}(z)\right)} + \sum_{x\neq z,b}\sqrt{p(b)\left(1 - b^x b^z E_b^{\mathrm{AC}}(x,z)\right)}, \tag{11}$$

where $E_b^{\mathrm{A}}(x)$, $E_b^{\mathrm{C}}(z)$ and $E_b^{\mathrm{AC}}(x,z)$ denote one- and two-party expectation values for Alice and Charlie, conditioned on Bob's output $b = (b^1, b^2, b^3)$ (see Appendix F). Numerical methods similar to the previous ones are employed to evidence that $\mathcal{B}' \leq 12\sqrt{3} + 2\sqrt{15}$ holds for bilocal models. In Appendix F we prove that there are quantum distributions whose non-bilocality is detected with this bilocal inequality but not with the inequality (9). Furthermore, we also prove that if Bob has uniformly distributed outcomes ($p(b) = \frac{1}{4}$), then $\mathcal{B}' \lesssim 30.70$ is respected by all quantum models with independent sources and hence it constitutes a quantum Bell inequality for the network [35].

Implementation of the Elegant Joint Measurement.— It is both interesting and practically relevant to address the question of how one may implement experimentally the EJM and its generalisation. In general, the implementation of joint (two-qubit) measurements requires the interaction of different signals. Optical implementations are of particular interest since they are common and convenient for Bell-type experiments. However, many such measurements, including the BSM, cannot be implemented with the basic tools applied in linear optics schemes (phase-shifters and beam splitters) when no auxiliary photons are present [37]. It turns out that our family of generalised EJM as defined by Eq. (3) can also not be implemented with two-photon linear optics, as can be shown by evaluating the criterion provided in Ref. [38]. More sophisticated tools are therefore required.

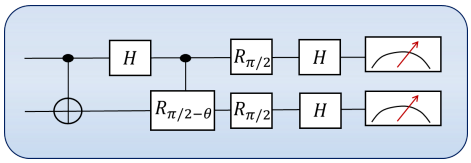

FIG. 3: Quantum circuit for implementing our family of generalised Elegant Joint Measurements parameterised by θ. A controlled-NOT gate is followed by a Hadamard rotation ($H = (\sigma_1 + \sigma_3)/\sqrt{2}$) on the control qubit, a controlled phase shift gate $R_{\pi/2-\theta}$, and a separate rotation of each qubit composed of $R_{\pi/2}$ and H. Finally, a measurement is performed in the basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$.

Our measurement family can in fact be implemented by the two-qubit circuit presented in Figure 3. This circuit maps the four measurement basis states $\{|\Phi_b^\theta\rangle\}_b$ onto the computational basis product states $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ (up to global phases). The proposed implementation involves (in addition to single-qubit gates) two different controlled unitary operations, namely a standard controlled-NOT gate and a controlled implementation of the phase shift gate

$$R_\phi = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{pmatrix}. \tag{12}$$

We remark that this controlled phase gate itself can be implemented using two controlled-NOT gates and unitaries acting on the target qubit as described in Ref. [39]. Finally, notice that when $\theta = \frac{\pi}{2}$, we have $R_{\pi/2-\theta} = \mathbb{1}$ and thus the circuit only involves a single two-qubit gate, just like the standard scheme for a BSM [40].

Discussion and open questions.— We have investigated quantum violations of bilocality based on the Elegant Joint Measurement and a new generalisation thereof. In contrast to several previous works in which quantum correlations were generated through a Bell State Measurement, our setup does not effectively reduce to separate implementations of the standard CHSH scenario. We nevertheless constructed new bilocal inequalities, and exhibited violations that we could not directly trace back to violations of a standard Bell inequality. Finally, we paved the way towards a bilocality experiment based on the EJM by constructing a quantum circuit for its implementation.

Several intriguing questions are left open. 1) What is the largest possible quantum violation of the bilocal inequalities? 2) Can the inequalities be proven in full generality? We note that the semidefinite relaxation methods of [34] can be exploited to place a bilocal bound on $\mathcal{B}$, albeit perhaps not tight. 3) How can one formalise the intuitive idea that some bilocal inequalities may or may not trace back to standard Bell in-

equalities? 4) Can our EJM family be further generalised for two higher-dimensional systems or for more than two qubits such that it preserves its elegant properties? 5) Are there any other correlations obtained using our EJM family that would be of particular interest to study (in the bilocality scenario or beyond), and more generally, could the introduced family of measurements have other interesting applications in quantum information science?

Acknowledgments

We thank Alejandro Pozas-Kerstjens for sharing with us both his codes and results for semidefinite programs based on Ref. [34], and Norbert Lütkenhaus for directing us into Ref. [38]. This work was supported by the Swiss National Science Foundation via the NCCR-SwissMap.

[1] J. S. Bell, On the Einstein Podolsky Rosen Paradox, Physics **1**, 195 (1964).
[2] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).
[3] N. Gisin, *Quantum Chance, nonlocality, teleportation and other quantum marvels*, Springer, 2014.
[4] C. Branciard, N. Gisin, and S. Pironio, Characterizing the Nonlocal Correlations Created via Entanglement Swapping, Phys. Rev. Lett. **104**, 170401 (2010).
[5] T. Fritz, Beyond Bell's theorem: correlation scenarios, New J. Phys. **14** 103001 (2012).
[6] M. Żukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert, "Event-ready-detectors" Bell experiment via entanglement swapping, Phys. Rev. Lett. **71**, 4287 (1993).
[7] C. Branciard, D. Rosset, N. Gisin, and S. Pironio, Bilocal versus nonbilocal correlations in entanglement-swapping experiments, Phys. Rev. A **85**, 032119 (2012).
[8] R. Chaves and T. Fritz, Entropic approach to local realism and noncontextuality, Phys. Rev. A **85**, 032113 (2012).
[9] A. Tavakoli, P. Skrzypczyk, D. Cavalcanti, and A. Acín, Nonlocal correlations in the star-network configuration, Phys. Rev. A **90**, 062109 (2014).
[10] J. Henson, R. Lal and M. F. Pusey, Theory-independent limits on correlations from generalised Bayesian networks, New J. Phys. **16**, 113043 (2014).
[11] C. J. Wood and R. W. Spekkens, The lesson of causal discovery algorithms for quantum correlations: causal explanations of Bell-inequality violations require fine-tuning, New J. Phys. **17** 033002 (2015).
[12] R. Chaves, R. Kueng, J. B. Brask, and D. Gross, Unifying Framework for Relaxations of the Causal Assumptions in Bell's Theorem, Phys. Rev. Lett. **114**, 140403 (2015).
[13] A. Tavakoli, Quantum Correlations in Connected Multipartite Bell Experiments, J. Phys. A: Math and Theor **49**, 145304 (2016).
[14] T. Fritz, Beyond Bell's Theorem II: Scenarios with arbitrary causal structure, Comm. Math. Phys. **341**, 391-434 (2016).
[15] D. Rosset, C. Branciard, T. J. Barnea, G. Pütz, N. Brunner, and N. Gisin, Nonlinear Bell Inequalities Tailored for Quantum Networks, Phys. Rev. Lett. **116**, 010403 (2016).
[16] R. Chaves, Polynomial Bell Inequalities, Phys. Rev. Lett. **116**, 010402 (2016).
[17] A. Tavakoli, Bell-type inequalities for arbitrary noncyclic networks, Phys. Rev. A **93**, 030101(R) (2016).
[18] A. Tavakoli, M-O. Renou, N. Gisin, and N. Brunner, Correlations in star networks: from Bell inequalities to network inequalities, New J. Phys. **19**, 073003 (2017).
[19] F. Andreoli, G. Carvacho, L. Santodonato, R. Chaves and F. Sciarrino, Maximal violation of n-locality inequalities in a star-shaped quantum network, New J. Phys. **19**, 113020 (2017).
[20] T. C. Fraser and E. Wolfe, Causal compatibility inequalities admitting quantum violations in the triangle structure, Phys. Rev. A **98**, 022113 (2018).
[21] M-X. Luo, Computationally Efficient Nonlinear Bell Inequalities for Quantum Networks, Phys. Rev. Lett. **120**, 140402 (2018).
[22] E. Wolfe, R. W. Spekkens, and T. Fritz, The Inflation Technique for Causal Inference with Latent Variables, J. Causal Inference **7**, 2 (2019).
[23] E. Wolfe, A. Pozas-Kerstjens, M. Grinberg, D. Rosset, A. Acín, and M. Navascues, Quantum Inflation: A General Approach to Quantum Causal Compatibility, arXiv:1909.10519
[24] M-O. Renou, E. Bäumer, S. Boreiri, N. Brunner, N. Gisin, and S. Beigi, Genuine Quantum Nonlocality in the Triangle Network, Phys. Rev. Lett. **123**, 140401 (2019).
[25] M-O. Renou, Y. Wang, S. Boreiri, S. Beigi, N. Gisin, and N. Brunner, Limits on Correlations in Networks for Quantum and No-Signaling Resources, Phys. Rev. Lett. 123, 070403 (2019).
[26] N. Gisin, J-D. Bancal, Y. Cai, A. Tavakoli, E. Z. Cruzeiro, S. Popescu, and N. Brunner, Constraints on nonlocality in networks from no-signaling and independence, Nat Commun **11**, 2378 (2020).
[27] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels, Phys. Rev. Lett. **70**, 1895 (1993).
[28] N. Gisin, Q. Mei, A. Tavakoli, M-O. Renou, and N. Brunner, All entangled pure quantum states violate the bilocality inequality, Phys. Rev. A **96**, 020304(R) (2017).
[29] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, Phys. Rev. Lett. **23**, 880 (1969).
[30] N. Gisin, Entanglement 25 Years after Quantum Teleportation: Testing Joint Measurements in Quantum Networks, Entropy **21**, 325 (2019).
[31] R. F. Werner, Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model, Phys. Rev. A **40**, 4277 (1989).
[32] Another natural scenario is to let Alice and Charlie perform four measurements, with Bloch vectors pointing to the vertices of a tetrahedron. It turns out, however, that the resulting correlations are less robust to noise than those obtained from measuring the three Pauli observables.
[33] Specifically, one has $p(a,b^1,b^2,b^3,c|x,z) = \frac{1}{16}\big[1 + a\langle A_x\rangle + \sum_y b^y\langle B^y\rangle + c\langle C_z\rangle + \sum_y ab^y\langle A_xB^y\rangle + \sum_y b^yc\langle B^yC_z\rangle + ac\langle A_xC_z\rangle + \sum_y ab^yc\langle A_xB^yC_z\rangle\big]$. Notice the (convenient) redundancy in our encoding of Bob's outcome, as the product of its three ±1-valued bits $b^1b^2b^3$ is always $+1$. As in Ref. [7], we write y as superscripts in b^y, B^y to distinguish the case where the outputs (b^y) are all observed together, from the case of outputs obtained for different inputs (as in A_x, C_z).
[34] A. Pozas-Kerstjens, R. Rabelo, L. Rudnicki, R. Chaves, D. Cav-

alcanti, M. Navascués, and A. Acín, Bounding the Sets of Classical and Quantum Correlations in Networks, Phys. Rev. Lett. **123**, 140503 (2019).
[35] Private communication with A. Pozas-Kerstjens.
[36] We have used the search method based on the Fourier transform of p_{biloc} and standard brute-force search using Matlab's fmincon module to confirm our bilocal inequalities.
[37] N. Lütkenhaus, J. Calsamiglia, and K-A. Suominen, On Bell measurements for teleportation, Phys. Rev. A **59**, 3295 (1999).
[38] P. van Loock and N. Lütkenhaus, Simple criteria for the implementation of projective measurements with linear optics, Phys. Rev. A **69**, 012302 (2004).
[39] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter, Elementary gates for quantum computation, Phys. Rev. A **52**, 3457 (1995).
[40] M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information (10th Anniversary edition) Cambridge University Press, 2010.

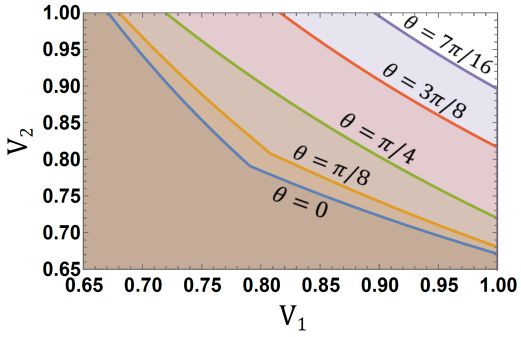

FIG. 4: Bilocal regions of the quantum correlations p_{Q}^{θ} in the (V_1, V_2)-plane, for different values of θ.

Appendix A: Bilocal simulation for intermediate measurements

Here, we explore the possibility of a bilocal simulation of quantum correlations based on the measurement family intermediate between the EJM and the BSM. Firstly, we consider the case in which Alice and Charlie perform the measurements $(\sigma_1, \sigma_2, \sigma_3)$ and Bob performs the intermediate measurement corresponding to a fixed θ. The resulting correlators are given in Eq. (5) of the main text.

In order to search for the region in the (V_1, V_2)-plane for which a bilocal simulation is possible, we have used the numerical method mentioned in the main text (where we presented the analysis for $\theta = 0$). Specifically, for a given θ, we search for a brute-force solution to $p_{\text{Q}}^{\theta} = p_{\text{biloc}}$ where we first apply a Fourier transform to the problem. This transforms probabilities into correlators. Some of these correlators are fixed immediately by the constraint $p_{\text{Q}}^{\theta} = p_{\text{biloc}}$. Those that are not fixed correspond to non-observable correlators (say e.g. $\langle A_1 A_2 C_1 \rangle$) and represent the internal degrees of freedom of the bilocal model, which we optimise over (under the constraint that they define valid probabilities). The main benefit of this method is that source-independence, appearing on the level of the free correlators, translates into simple conditions that are either linear or quadratic. This makes the numerical search more efficient and accurate; see Ref. [7] for a more detailed description. In Figure 4, we display the boundary of the bilocal region found through this method for several different values of θ. We find that as we depart further from the EJM, i.e as we increase θ, the region that admits a bilocal simulation grows larger. For $\theta = \frac{\pi}{2}$, the quantum correlation p_{Q}^{θ} is found to be bilocal[1] for all visibilities V_1, V_2.

It is interesting to note that while the non-bilocal region

[1] An explicit bilocal model for $p_{\text{Q}}^{\theta=\frac{\pi}{2}}$ is obtained by letting α be any of the 4 vectors $-\vec{m}_b$ and γ be any of the 4 vectors $\vec{m}_b$ of Eq. (1), with equal probabilities, and by defining $p(b|\alpha,\gamma) = \frac{1+3V_1V_2}{4}$ if $-\alpha = b = \gamma$ or $\det(-\alpha, b, \gamma) > 0$, and $p(b|\alpha,\gamma) = \frac{1-V_1V_2}{4}$ otherwise.

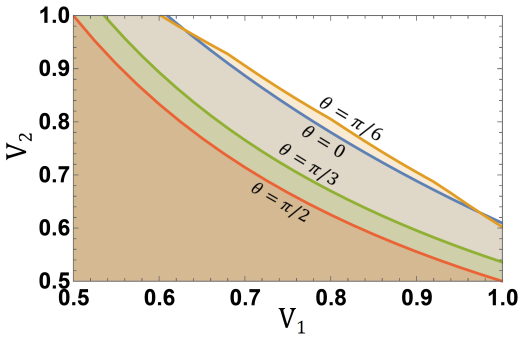

FIG. 5: Bilocal regions of the quantum correlations when Alice and Charlie perform the measurements $\left(\frac{\sigma_3+\sigma_1}{\sqrt{2}}, \sigma_2, \frac{\sigma_3-\sigma_1}{\sqrt{2}}\right)$ for different values of θ.

appears to be vanishing in Figure 4 as Bob's measurement approaches the BSM, the standard bilocal inequality [7], which is based on the BSM, admits a robust quantum violation. This suggests that as θ grows larger, and the measurement becomes less similar to the EJM and more similar to the BSM, Alice and Charlie would benefit from changing the orientation of their local measurements. We illustrate this by letting Alice and Charlie measure in the bases $\left(\frac{\sigma_3+\sigma_1}{\sqrt{2}}, \sigma_2, \frac{\sigma_3-\sigma_1}{\sqrt{2}}\right)$. For several values of θ, we plot the region in the (V_1, V_2)-plane for which we find a bilocal simulation of the quantum correlations thus obtained (still considering measurements on noisy singlet states): see Figure 5. In Figure 5 we see that the trend observed in Figure 4 is reversed; for larger values of θ, the bilocal region is shrinking. In particular, for the BSM ($\theta = \frac{\pi}{2}$), the boundary of the bilocal region is characterised by $V_1 V_2 = \frac{1}{2}$ which is the same as that obtained in the standard bilocal inequality [7]. However, the bilocal region in Figure 5 is not monotonic in θ: the bilocal region for $\theta = \frac{\pi}{6}$ is typically larger than that of the EJM ($\theta = 0$). Typically, for small values of θ, the re-oriented local measurements of Alice and Charlie do not constitute an improvement over the previous $(\sigma_1, \sigma_2, \sigma_3)$ measurements.

All this illustrates the fact that the choice of Alice and Charlie's measurements have a nontrivial implication on the existence or non-existence of a bilocal model for the quantum correlations under investigation. Although the choice of measurements $(\sigma_1, \sigma_2, \sigma_3)$ for Alice and Charlie that we considered in the main text looks appropriate when Bob performs the EJM, it is seen to be nonoptimal when Bob uses the generalised EJM family, for general values of $\theta > 0$. Finding the optimal measurements to unveil quantum nonbilocality in a given scenario is certainly not a trivial task.

Appendix B: Bilocal models with tetrahedral symmetry

We detail here a simple and analytical family of bilocal models exhibiting the tetrahedral symmetry outlined in the main text. These models provide bilocal simulations of the quantum correlation $p_{\rm Q}^{\theta=0}$ for visibilities (V_1, V_2) very close to the critical ones, above which $p_{\rm Q}^{\theta=0}$ becomes nonbilocal. Along the boundary of the bilocal set in the (V_1, V_2) plane (shown in Figure 2 of the main text), the difference between the critical pairs obtained by numerical search without our symmetry assumption and those obtained for the model below is of the order of 10^{-5} only; e.g., for the symmetric noise case ($V_1 = V_2$), numerical optimisation over all bilocal models gave us a critical visibility of $V_{\rm crit} \approx 0.790896$, while the symmetric model below gives $V_{\rm crit} \approx 0.790871$.

For convenience in the presentation below, let us denote by $\mathcal{T}_+ = \{\vec{m}_b\}_{b=1,\ldots,4}$ the set of tetrahedron vertices $\vec{m}_b$ of Eq. (1) and by $\mathcal{T}_- = \{-\vec{m}_b\}_{b=1,\ldots,4}$ the set of opposite vectors. For any $\alpha = (\alpha_1, \alpha_2, \alpha_3) \in \mathcal{T}_\pm$ (for any $\gamma = (\gamma_1, \gamma_2, \gamma_3) \in \mathcal{T}_\pm$, respectively), let us define $\tilde{\alpha} = \pm\alpha \in \mathcal{T}_+$ ($\tilde{\gamma} = \pm\gamma \in \mathcal{T}_+$) to be the vector in $\mathcal{T}_+$ along the same direction as α (γ) and possibly with the opposite sign, if $\alpha \in \mathcal{T}_-$ (if $\gamma \in \mathcal{T}_-$).

In general, $p_{\rm Q}^\theta = p_{\rm biloc}$ leads to a large system of equations. However, the symmetry ansatz greatly simplifies the problem. Note first that the requirement that $p_{\rm biloc}(\pi(\alpha), \pi(b), \pi(\gamma)) = p_{\rm biloc}(\alpha, b, \gamma)$ for all permutations π of the tetrahedron (as defined in the main text) imposes that the probabilities $q_\alpha^{(1)}$ ($q_\gamma^{(2)}$, resp.) are the same for all four values of α (γ) in $\mathcal{T}_+$, and the same for all four values of α (γ) in $\mathcal{T}_-$. Defining $q_\pm^{(1)} = \sum_{\alpha\in\mathcal{T}_\pm} q_\alpha^{(1)} \in [0,1]$ and $q_\pm^{(2)} = \sum_{\gamma\in\mathcal{T}_\pm} q_\gamma^{(2)} \in [0,1]$, the weights $q_\alpha^{(1)}$ ($q_\gamma^{(2)}$) are then all either equal to $\frac{1}{4}q_+^{(1)}$ ($\frac{1}{4}q_+^{(2)}$) or to $\frac{1}{4}q_-^{(1)}$ ($\frac{1}{4}q_-^{(2)}$), depending on whether α (γ) is in $\mathcal{T}_+$ or $\mathcal{T}_-$.

In turn, it also follows that Bob's response functions conditioned on the local variables α, γ have the symmetry $p(\pi(b)|\pi(\alpha), \pi(\gamma)) = p(b|\alpha, \gamma)$ for all permutations π of the tetrahedron. With this symmetry (and noting that b, just like $\tilde{\alpha}$ and $\tilde{\gamma}$, is always in $\mathcal{T}_+$), Bob's response functions can be defined by only specifying for instance, for each of the four cases where $\alpha \in \mathcal{T}_\pm$ and $\gamma \in \mathcal{T}_\pm$: *(i)* the probabilities that $b = \tilde{\alpha} = \tilde{\gamma}$ when $\tilde{\alpha} = \tilde{\gamma}$, which we denote by $q_{b=\tilde{\alpha}=\tilde{\gamma}|\tilde{\alpha}=\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma}$ (with the superscripts $\tau_\alpha, \tau_\gamma = \pm$ referring to $\alpha \in \mathcal{T}_{\tau_\alpha}$ and $\gamma \in \mathcal{T}_{\tau_\gamma}$, and such that the probabilities that b takes any of the three values other than $\tilde{\alpha} = \tilde{\gamma}$ is, by symmetry, $(1 - q_{b=\tilde{\alpha}=\tilde{\gamma}|\tilde{\alpha}=\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma})/3$); *(ii)* the probabilities that $b = \tilde{\alpha}$ and *(iii)* the probabilities that $b = \tilde{\gamma}$ when $\tilde{\alpha} \neq \tilde{\gamma}$, which we denote by $q_{b=\tilde{\alpha}|\tilde{\alpha}\neq\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma}$ and $q_{b=\tilde{\gamma}|\tilde{\alpha}\neq\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma}$, resp. (such that the probabilities that b takes any of the two values other than $\tilde{\alpha}$ and $\tilde{\gamma}$, when these are different, is $(1 - q_{b=\tilde{\alpha}|\tilde{\alpha}\neq\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma} - q_{b=\tilde{\gamma}|\tilde{\alpha}\neq\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma})/2$). All in all (and noting that $q_-^{(i)} = 1 - q_+^{(i)}$ for $i = 1, 2$), any bilocal model with the tetrahedral symmetry considered here can thus be defined by just the 14 parameters $q_+^{(1)}, q_+^{(2)}, q_{b=\tilde{\alpha}=\tilde{\gamma}|\tilde{\alpha}=\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma}, q_{b=\tilde{\alpha}|\tilde{\alpha}\neq\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma}, q_{b=\tilde{\gamma}|\tilde{\alpha}\neq\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma}$ (for each of the four combinations of τ_α, τ_γ).[2]

[2] Note that $p_{\rm Q}^\theta$ does not have the ("full") tetrahedral symmetry considered here when $\theta > 0$, as the correlators $\langle A_x B^y C_z\rangle$ in Eq. (5) are different for

To find the critical visibilities (V_1, V_2) for which such symmetric models can reproduce the quantum correlation $p_{\text{Q}}^{\theta=0}$, we let V_1 take different fixed values, and optimise over the 14 weights above, together with V_2, so as to find the largest possible V_2 allowing for $p_{\text{Q}}^{\theta=0}$ to be reproduced. Numerically we found, for large enough V_1 (namely, $V_1 \gtrsim 0.791$), that the optimal strategies were to take

$$\begin{aligned}
&q_+^{(1)} \approx q_+^{(2)}, \\
&q_{b=\tilde{\alpha}=\tilde{\gamma}|\tilde{\alpha}=\tilde{\gamma}}^{+,+} \approx 0, \quad q_{b=\tilde{\alpha}|\tilde{\alpha}\neq\tilde{\gamma}}^{+,+} \approx 0, \quad q_{b=\tilde{\gamma}|\tilde{\alpha}\neq\tilde{\gamma}}^{+,+} \approx 1, \\
&q_{b=\tilde{\alpha}=\tilde{\gamma}|\tilde{\alpha}=\tilde{\gamma}}^{-,+} \approx 1, \quad q_{b=\tilde{\alpha}|\tilde{\alpha}\neq\tilde{\gamma}}^{-,+} \approx 0, \quad q_{b=\tilde{\gamma}|\tilde{\alpha}\neq\tilde{\gamma}}^{-,+} = q_0, \\
&q_{b=\tilde{\alpha}=\tilde{\gamma}|\tilde{\alpha}=\tilde{\gamma}}^{+,-} \approx 1, \quad q_{b=\tilde{\alpha}|\tilde{\alpha}\neq\tilde{\gamma}}^{+,-} \approx 0, \quad q_{b=\tilde{\gamma}|\tilde{\alpha}\neq\tilde{\gamma}}^{+,-} \approx 0, \\
&q_{b=\tilde{\alpha}=\tilde{\gamma}|\tilde{\alpha}=\tilde{\gamma}}^{-,-} \approx 0, \quad q_{b=\tilde{\alpha}|\tilde{\alpha}\neq\tilde{\gamma}}^{-,-} \approx 1, \quad q_{b=\tilde{\gamma}|\tilde{\alpha}\neq\tilde{\gamma}}^{-,-} \approx 0, \qquad \text{(B1)}
\end{aligned}$$

for some value $q_0 \in [0,1]$ (that depends on V_1). E.g., when $\alpha, \gamma \in \mathcal{T}_+$ (in which case $\tilde{\alpha} = \alpha, \tilde{\gamma} = \gamma$), then the model should return any of the three values $b \neq \alpha, \gamma$ (with equal probabilities) if $\alpha = \gamma$, or should return $b = \gamma$ if $\alpha \neq \gamma$; when $\alpha \in \mathcal{T}_-, \gamma \in \mathcal{T}_+$ (in which case $\tilde{\alpha} = -\alpha, \tilde{\gamma} = \gamma$), then the model should return $b = -\alpha = \gamma$ if $-\alpha = \gamma$, or should return $b = \gamma$ with probability q_0 or any of the two values $b \neq -\alpha, \gamma$ with equal probabilities $(1 - q_0)/2$ if $-\alpha \neq \gamma$; etc.

By imposing that the 14 parameters of the model satisfy Eq. (B1) with strict equalities, it becomes possible to construct the model analytically. To reproduce the correlation $p_{\text{Q}}^{\theta=0}$, for a given value of V_1, the remaining free parameters need to take the values

$$\begin{aligned}
V_2 &= \frac{58 + 9V_1 - 12\sqrt{2V_1 - 8/9}}{27(1 + 2V_1)}, \\
q_+^{(1)} &= q_+^{(2)} = \frac{2}{3} - \frac{\sqrt{2V_1 - 8/9}}{2}, \\
q_0 &= \frac{6\sqrt{2V_1 - 8/9} + 9V_2 - 9V_1 - 2}{3\sqrt{2V_1 - 8/9} + 8 - 9V_1}, \qquad \text{(B2)}
\end{aligned}$$

which completes the full specification of our family of bilocal models for $p_{\text{Q}}^{\theta=0}$, and for some very close-to-optimal visibilities (V_1, V_2).

Note that our models here work for visibilities $V_1 \geq V_2$; for $V_2 \geq V_1$ similar models can be found, with the roles of V_1 and V_2 exchanged in the construction above. For $V_1 = V_2$, the critical visibility $V_{\text{crit}} \approx 0.791$ is obtained as the unique solution to the first line of Eq. (B2), after imposing $V_1 = V_2 = V_{\text{crit}}$. Note also that $V_1 \geq V_2$ ensures in particular that $V_1 \geq V_{\text{crit}} > 4/9$, so that the square roots in Eq. (B2) take real values.

even and odd permutations of $\{x, y, z\}$. One may also define bilocal models with a "relaxed" tetrahedral symmetry matching the symmetry of $p_{\text{Q}}^{\theta>0}$, by allowing for different probabilities for the two values of b other than $\tilde{\alpha}$ and $\tilde{\gamma}$ when these are different (just depending on the sign of $\det(\tilde{\alpha}, b, \tilde{\gamma})$ to preserve some symmetry). This would add four parameters to the model (one for each combination of τ_α, τ_γ). As an example, the explicit bilocal model given in Footnote A for $p_{\text{Q}}^{\theta=\frac{\pi}{2}}$ has this relaxed tetrahedral symmetry. In the remaining part of these appendices however, by tetrahedral symmetry we will refer to the "full" tetrahedral symmetry.

Appendix C: Proof of the first bilocal inequality under tetrahedral symmetry

Here we prove the bilocal inequality (9) for models that satisfy our tetrahedral symetry ansatz.

For such models, as defined in Appendix B in terms of the 14 parameters $q_+^{(1)}, q_+^{(2)}, q_{b=\tilde{\alpha}=\tilde{\gamma}|\tilde{\alpha}=\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma}, q_{b=\tilde{\alpha}|\tilde{\alpha}\neq\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma}, q_{b=\tilde{\gamma}|\tilde{\alpha}\neq\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma}$, the one-, two- and three-party correlators are found to be

$$\begin{aligned}
&\langle A_x \rangle = \langle B^y \rangle = \langle C_z \rangle = \langle A_x C_z \rangle = 0, \\
&\langle A_x B^y \rangle = \delta_{x,y} \sum_{\tau_\alpha,\tau_\gamma=\pm1} q_{\tau_\alpha}^{(1)} q_{\tau_\gamma}^{(2)} \, \tau_\alpha \Big(q_{b=\tilde{\alpha}|\tilde{\alpha}\neq\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma} - \tfrac{1 - q_{b=\tilde{\alpha}=\tilde{\gamma}|\tilde{\alpha}=\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma}}{3} \Big), \\
&\langle B^y C_z \rangle = \delta_{y,z} \sum_{\tau_\alpha,\tau_\gamma=\pm1} q_{\tau_\alpha}^{(1)} q_{\tau_\gamma}^{(2)} \, \tau_\gamma \Big(q_{b=\tilde{\gamma}|\tilde{\alpha}\neq\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma} - \tfrac{1 - q_{b=\tilde{\alpha}=\tilde{\gamma}|\tilde{\alpha}=\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma}}{3} \Big), \\
&\langle A_x B^y C_z \rangle = \delta_{x\neq y\neq z} \sum_{\tau_\alpha,\tau_\gamma=\pm1} q_{\tau_\alpha}^{(1)} q_{\tau_\gamma}^{(2)} \, \tau_\alpha \tau_\gamma \Big(\tfrac{1}{2} - \tfrac{1 - q_{b=\tilde{\alpha}=\tilde{\gamma}|\tilde{\alpha}=\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma}}{3} \\
&\qquad\qquad - \tfrac{q_{b=\tilde{\alpha}|\tilde{\alpha}\neq\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma} + q_{b=\tilde{\gamma}|\tilde{\alpha}\neq\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma}}{2} \Big) \qquad \text{(C1)}
\end{aligned}$$

(with $\delta_{x\neq y\neq z} = 1$ if x, y, z are all distinct, $\delta_{x\neq y\neq z} = 0$ otherwise). From these we get[3]

$$\begin{aligned}
&\frac{S}{3} - T \\
&= q_+^{(1)} q_+^{(2)} \Big[2q_{b=\tilde{\alpha}|\tilde{\alpha}\neq\tilde{\gamma}}^{+,+} + 4q_{b=\tilde{\gamma}|\tilde{\alpha}\neq\tilde{\gamma}}^{+,+} - 1 - 2q_{b=\tilde{\alpha}=\tilde{\gamma}|\tilde{\alpha}=\tilde{\gamma}}^{+,+} \Big] \\
&+ q_+^{(1)} q_-^{(2)} \Big[3 - 4\tfrac{1 - q_{b=\tilde{\alpha}=\tilde{\gamma}|\tilde{\alpha}=\tilde{\gamma}}^{+,-}}{3} - 4q_{b=\tilde{\alpha}|\tilde{\alpha}\neq\tilde{\gamma}}^{+,-} - 4q_{b=\tilde{\gamma}|\tilde{\alpha}\neq\tilde{\gamma}}^{+,-} \Big] \\
&+ q_-^{(1)} q_+^{(2)} \Big[3 - 8\tfrac{1 - q_{b=\tilde{\alpha}=\tilde{\gamma}|\tilde{\alpha}=\tilde{\gamma}}^{-,+}}{3} - 2q_{b=\tilde{\alpha}|\tilde{\alpha}\neq\tilde{\gamma}}^{-,+} - 2q_{b=\tilde{\gamma}|\tilde{\alpha}\neq\tilde{\gamma}}^{-,+} \Big] \\
&+ q_-^{(1)} q_-^{(2)} \Big[4q_{b=\tilde{\alpha}|\tilde{\alpha}\neq\tilde{\gamma}}^{-,-} + 2q_{b=\tilde{\gamma}|\tilde{\alpha}\neq\tilde{\gamma}}^{-,-} - 1 - 2q_{b=\tilde{\alpha}=\tilde{\gamma}|\tilde{\alpha}=\tilde{\gamma}}^{-,-} \Big] \qquad \text{(C2)}
\end{aligned}$$

and $Z = 0$.

Recalling that all parameters $q_{(\cdots)}^{\tau_\alpha,\tau_\gamma}$ of the symmetric model are between 0 and 1, and that they further satisfy $q_{b=\tilde{\alpha}|\tilde{\alpha}\neq\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma} + q_{b=\tilde{\gamma}|\tilde{\alpha}\neq\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma} \leq 1$, one can easily see that each term in square brackets above is upper-bounded by 3. As $S/3 - T$ is obtained as a convex combination of these four terms (with the weights $q_\pm^{(1)} q_\pm^{(2)}$), then it is also upper-bounded by 3—which indeed proves our inequality (9) for bilocal models satisfying the tetrahedral symmetry assumption (for which $Z = 0$).

We believe it should be possible to prove that any general bilocal model for correlations satisfying $Z = 0$ can be "symmetrised" into a bilocal model with the tetrahedral symmetry

[3] Note that bilocal models with the "relaxed" tetrahedral symmetry as described in Footnote B give the same values of S and T, so that the proof here also applies to such models.

considered here, that would have the same values of S and T. This would give a general proof of our bilocal inequality (9), for the $Z = 0$ case. However, the details here remain to be worked out properly, so that we rely for now on (trustworthy) numerical optimisations.

Appendix D: "$Z = 0$ slice" of the bilocal set in the (S, T)-plane

It clearly appears that a case of particular interest in our study is when $Z = 0$—as satisfied in particular by the quantum correlation p_{Q}^{θ} we investigate, and by any bilocal model with the tetrahedral symmetry considered previously. The choice to define and look at the quantities S and T, as defined in Eq. (8), is then rather naturally dictated by the specific forms of the correlators, Eq. (5) for p_{Q}^{θ}.[4]

To get some idea of what the set of bilocal correlations looks like, it is instructive to look at the *projection* onto the (S, T) plane of its *slice* where $Z = 0$. This projection, obtained through numerical optimisation to check the (non)bilocality of various points (S, T), is shown on Figure 6. Quite remarkably, and contrarily to all (nontrivial, multidimensional) bilocal sets previously studied in the literature, it appears that the bilocal set in this projected slice is delimited by linear inequalities, namely:

$$\pm\frac{S}{3} - T \overset{\substack{\mathrm{biloc}\\ Z=0}}{\leq} 3, \quad \pm S \overset{\substack{\mathrm{biloc}\\ Z=0}}{\leq} 3, \quad \pm S + T \overset{\substack{\mathrm{biloc}\\ Z=0}}{\leq} 3. \tag{D1}$$

We also verified these six inequalities via numerical optimisations, as we did for our other bilocal inequalities presented in this paper. These can also be proven for bilocal models with the tetrahedral symmetry in the same way as in the previous appendix. The first of these inequalities, with a $+$ sign, corresponds precisely to our bilocal inequality (9) for the $Z = 0$ case. As we see, it thus appears to be "tight", in the sense of defining a facet of the bilocal set in the projected $Z = 0$ slice.

To complete the picture, one may also look at the set of local correlations. This forms a convex polytope in the full correlation space, so it is expected to also be delimited by linear inequalities in the projected $Z = 0$ slice. We find here that its facets are defined by

$$\pm S \overset{\substack{\mathrm{loc}\\ Z=0}}{\leq} 3, \quad \pm T \overset{\substack{\mathrm{loc}\\ Z=0}}{\leq} 4, \quad \pm S + \frac{T}{2} \overset{\mathrm{loc}}{\leq} 3 \tag{D2}$$

(with the last pair of inequalities holding in fact for general local models, without the $Z = 0$ restriction); see Figure 6. We note that the correlations p_{Q}^{θ} satisfy the above inequalities and, more generally, they admit a local model.

[4] One may also naturally refine the analysis by defining and considering $S^{AB} = \sum_{x=y} \langle A_x B^y \rangle$, $S^{BC} = \sum_{y=z} \langle B^y C_z \rangle$, $R^{+} = \sum_{xyz \in \{123,231,312\}} \langle A_x B^y C_z \rangle$ and $R^{-} = \sum_{xyz \in \{132,213,321\}} \langle A_x B^y C_z \rangle$ (such that $S = S^{BC} - S^{AB}$ and $R = R^{+} + R^{-}$).

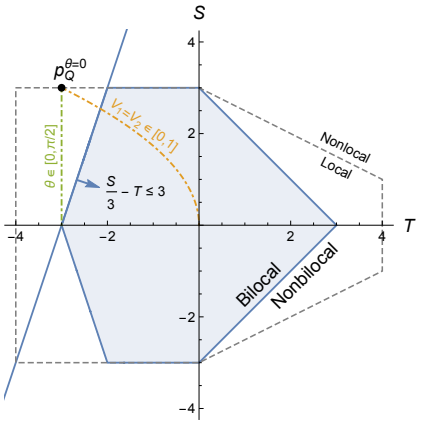

FIG. 6: Projection of the "$Z = 0$ slice" of the correlation space onto the (S, T)-plane. The blue region represents the projection of the bilocal set, delimited by the inequalities in Eq. (D1). The gray dashed lines delimit the projection of the local set, according to Eq. (D2). The black point shows the projection of the quantum correlation $p_{\mathrm{Q}}^{\theta=0}$ in the noiseless case ($V_1 = V_2 = 0$). The orange and green dash-dotted curves show the projections of $p_{\mathrm{Q}}^{\theta=0}$ for symmetric noise $V_1 = V_2 \in [0, 1]$ and the projections of p_{Q}^{θ} for all $\theta \in [0, \frac{\pi}{2}]$ in the noiseless case, respectively, with the former entering the bilocal set for visibilities $V_1 = V_2 = V_{\mathrm{crit}}$ given by Eq. (10), and the latter remaining nonbilocal as long as $\theta < \frac{\pi}{2}$.

Appendix E: Z-correction of the bilocal inequality

As we just saw, when restricting to the case where $Z = 0$, the bilocal inequality presented in Eq. (9) of the main text is tight in the (S, T) plane. However, when Z is perturbed away from zero (e.g. due to small experimental errors), then Eq. (9) is not tight anymore.

We have numerically computed the largest values of $\mathcal{B} = \frac{S}{3} - T$ attainable for a given value of Z. This can be efficiently incorporated into the optimisation by placing the linear constraint $-Z \leq \langle \cdot \rangle \leq Z$ on all the one-, two- and three-party correlators that do not appear in S and T. The results of the optimisation are displayed in Figure 7. The simplest correction term that can be added to the bilocal bound for $Z = 0$ in order to account for the case where $Z > 0$ is a linear correction of $5Z$, as illustrated and as we considered in Eq. (9). However, it is clear that more precise correction terms to the bilocal bound are also possible.

Appendix F: A second bilocal inequality

We detail the second bilocal inequality mentioned in the main text. Alike the first bilocal inequality, it is inspired by the quantum correlations based on the EJM. In order to detect non-bilocal correlations without imposing additional constraints on the correlations, it is typically necessary to employ

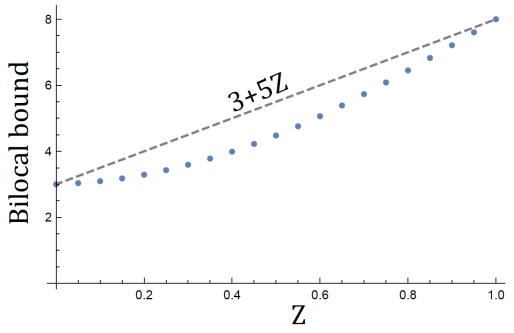

FIG. 7: Results for the numerical optimisation of the bilocal bound of $\mathcal{B}$ for various values of Z (blue dots), and a linear correction $(5Z)$ to the bilocal bound of 3 associated to the case of $Z = 0$.

nonlinear expressions to capture the non-convexity of the set of bilocal correlations. We thus consider here the Bell expression (Eq. (11) of the main text)

$$\mathcal{B}' \equiv \sum_{x,b} \sqrt{p(b)\left(1 - b^x E_b^{\mathrm{A}}(x)\right)} + \sum_{z,b} \sqrt{p(b)\left(1 + b^z E_b^{\mathrm{C}}(z)\right)} + \sum_{x\neq z,b} \sqrt{p(b)\left(1 - b^x b^z E_b^{\mathrm{AC}}(x,z)\right)}, \tag{F1}$$

where we have defined the conditional one-party correlators $E_b^{\mathrm{A}}(x) \equiv \sum_{a,c} a\, p(a,c|b,x,z)$ and $E_b^{\mathrm{C}}(z) \equiv \sum_{a,c} c\, p(a,c|b,x,z)$ and the conditional two-party correlators $E_b^{\mathrm{AC}}(x,z) \equiv \sum_{a,c} ac\, p(a,c|b,x,z)$, and where as in the main text Bob's output b is written as $b = (b^1, b^2, b^3)$, with each $b^y = \pm 1$. In terms of the (non-conditional) correlators considered previously, one has $p(b) = \frac{1}{4}(1 + \sum_y b^y \langle B^y\rangle)$, $p(b)E_b^{\mathrm{A}}(x) = \frac{1}{4}(\langle A_x\rangle + \sum_y b^y \langle A_x B^y\rangle)$, $p(b)E_b^{\mathrm{C}}(z) = \frac{1}{4}(\langle C_z\rangle + \sum_y b^y \langle B^y C_z\rangle)$, and $p(b)E_b^{\mathrm{AC}}(x,z) = \frac{1}{4}(\langle A_x C_z\rangle + \sum_y b^y \langle A_x B^y C_z\rangle)$.

1. Bilocal bound

The bilocal bound on $\mathcal{B}'$ was obtained numerically, by optimising general models using two different numerical search methods [36]. We found in particular that the same upper bound was obtained (up to machine precision) by bilocal models with the tetrahedral symmetry considered before; let us thus consider such models to obtain the analytical expression for the bilocal bound.

Recall that for these models, $Z = 0$—i.e., $\langle A^x\rangle = \langle B^y\rangle = \langle C^z\rangle = \langle A^x C^z\rangle = 0$; the bipartite correlators $\langle A_x B^y\rangle$ and $\langle B^y C_z\rangle$ are also 0 whenever $x \neq y$ and $y \neq z$, resp.; and the tripartite correlators $\langle A_x B^y C_z\rangle$ are also 0 whenever x, y, z are not all different. It follows that $p(b) = \frac{1}{4}$, $E_b^{\mathrm{A}}(x) = b^x \langle A_x B^x\rangle$, $E_b^{\mathrm{C}}(z) = b^z \langle B^z C_z\rangle$, and $E_b^{\mathrm{AC}}(x,z) = \delta_{x\neq z} b^x b^z \langle A_x B^{y\neq x,z} C_z\rangle$ (where we used the fact that $b^1 b^2 b^3 = 1$, and where the superscript $y \neq x, z$ denotes the unique value of y different from both x and z when $x \neq z$), so that $\mathcal{B}'$ can be written as

$$\mathcal{B}' = 2\sum_x \sqrt{1 - \langle A_x B^x\rangle} + 2\sum_z \sqrt{1 + \langle B^z C_z\rangle} + 2\sum_{x\neq z} \sqrt{1 - \langle A_x B^{y\neq x,z} C_z\rangle}. \tag{F2}$$

Using Eq. (C1) we obtain more specifically, in terms of the 14 parameters $q_+^{(1)}, q_+^{(2)}, q_{b=\tilde{\alpha}=\tilde{\gamma}|\tilde{\alpha}=\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma}, q_{b=\tilde{\alpha}|\tilde{\alpha}\neq\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma}, q_{b=\tilde{\gamma}|\tilde{\alpha}\neq\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma}$ defining a symmetric bilocal model (see Appendix B):

$$\mathcal{B}' = 6\sqrt{\sum_{\tau_\alpha,\tau_\gamma} q_{\tau_\alpha}^{(1)} q_{\tau_\gamma}^{(2)} \left(1 - \tau_\alpha\, q_{b=\tilde{\alpha}|\tilde{\alpha}\neq\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma} + \tau_\alpha \frac{1 - q_{b=\tilde{\alpha}=\tilde{\gamma}|\tilde{\alpha}=\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma}}{3}\right)} + 6\sqrt{\sum_{\tau_\alpha,\tau_\gamma} q_{\tau_\alpha}^{(1)} q_{\tau_\gamma}^{(2)} \left(1 + \tau_\gamma\, q_{b=\tilde{\gamma}|\tilde{\alpha}\neq\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma} - \tau_\gamma \frac{1 - q_{b=\tilde{\alpha}=\tilde{\gamma}|\tilde{\alpha}=\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma}}{3}\right)} + 12\sqrt{\sum_{\tau_\alpha,\tau_\gamma} q_{\tau_\alpha}^{(1)} q_{\tau_\gamma}^{(2)} \left(\begin{array}{c} 1 - \tau_\alpha\tau_\gamma \frac{1}{2} + \tau_\alpha\tau_\gamma \frac{1 - q_{b=\tilde{\alpha}=\tilde{\gamma}|\tilde{\alpha}=\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma}}{3} \\ + \tau_\alpha\tau_\gamma \frac{q_{b=\tilde{\alpha}|\tilde{\alpha}\neq\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma} + q_{b=\tilde{\gamma}|\tilde{\alpha}\neq\tilde{\gamma}}^{\tau_\alpha,\tau_\gamma}}{2} \end{array}\right)}. \tag{F3}$$

One can then use the trivial (and all saturable) bounds $-q_{b=\tilde{\alpha}|\tilde{\alpha}\neq\tilde{\gamma}}^{+,+} \leq 0$, $-q_{b=\tilde{\alpha}|\tilde{\alpha}\neq\tilde{\gamma}}^{+,-} \leq 0$, $-\frac{1 - q_{b=\tilde{\alpha}=\tilde{\gamma}|\tilde{\alpha}=\tilde{\gamma}}^{-,+}}{3} \leq 0$ and $q_{b=\tilde{\alpha}|\tilde{\alpha}\neq\tilde{\gamma}}^{-,-} \leq 1$ under the first square root, $q_{b=\tilde{\gamma}|\tilde{\alpha}\neq\tilde{\gamma}}^{+,+} \leq 1$, $-q_{b=\tilde{\gamma}|\tilde{\alpha}\neq\tilde{\gamma}}^{+,-} \leq 0$, $-\frac{1 - q_{b=\tilde{\alpha}=\tilde{\gamma}|\tilde{\alpha}=\tilde{\gamma}}^{-,+}}{3} \leq 0$ and $-q_{b=\tilde{\gamma}|\tilde{\alpha}\neq\tilde{\gamma}}^{-,-} \leq 0$ under the second square root, and $q_{b=\tilde{\alpha}|\tilde{\alpha}\neq\tilde{\gamma}}^{+,+} + q_{b=\tilde{\gamma}|\tilde{\alpha}\neq\tilde{\gamma}}^{+,+} \leq 1$, $-q_{b=\tilde{\alpha}|\tilde{\alpha}\neq\tilde{\gamma}}^{+,-} - q_{b=\tilde{\gamma}|\tilde{\alpha}\neq\tilde{\gamma}}^{+,-} \leq 0$, $-\frac{1 - q_{b=\tilde{\alpha}=\tilde{\gamma}|\tilde{\alpha}=\tilde{\gamma}}^{-,+}}{3} \leq 0$ and $q_{b=\tilde{\alpha}|\tilde{\alpha}\neq\tilde{\gamma}}^{-,-} + q_{b=\tilde{\gamma}|\tilde{\alpha}\neq\tilde{\gamma}}^{-,-} \leq 1$ under the third square root, to upper-bound $\mathcal{B}'$ above by a (saturable) expression that does no longer contain the 7 different parameters involved here. This leaves us with only 7 (out of the initial 14) free parameters to optimise for the symmetric models, at which point we resort to numerical means. We find in particular that the maximum of the $\mathcal{B}'$ expression is obtained by choosing $q_+^{(1)} = q_+^{(2)} = 1$ or $q_+^{(1)} = q_+^{(2)} = 0$. In the first case, we thus obtain

$$\mathcal{B}' \leq 18\sqrt{1 + \frac{1 - q_{b=\tilde{\alpha}=\tilde{\gamma}|\tilde{\alpha}=\tilde{\gamma}}^{+,+}}{3}} + 6\sqrt{2 - \frac{1 - q_{b=\tilde{\alpha}=\tilde{\gamma}|\tilde{\alpha}=\tilde{\gamma}}^{+,+}}{3}}, \tag{F4}$$

which reaches its maximum for $q_{b=\tilde{\alpha}=\tilde{\gamma}|\tilde{\alpha}=\tilde{\gamma}}^{+,+} = 0$. Thus, we find the bilocal bound to be

$$\mathcal{B}' \overset{\text{biloc}}{\leq} 12\sqrt{3} + 2\sqrt{15} \approx 28.53. \tag{F5}$$

We reiterate that, although obtained explicitly here for bilocal models with the tetrahedral symmetry, this bound was verified numerically to hold for general bilocal models.

2. Quantum violations

The bilocal bound is violated by the quantum correlations p_{Q}^{θ}, based on the family of measurements generalising the

EJM. Evaluating the conditional one- and two-party correlators, we obtain

$$E_b^{\mathrm{A}}(x) = -b^x \frac{V_1}{2}\cos\theta, \quad E_b^{\mathrm{C}}(z) = b^z \frac{V_2}{2}\cos\theta,$$
$$E_b^{\mathrm{AC}}(x,z) = \begin{cases} -b^x b^z \frac{V_1V_2}{2}(1+\sin\theta) & \text{if } xz \in \{13,21,32\} \\ -b^x b^z \frac{V_1V_2}{2}(1-\sin\theta) & \text{if } xz \in \{12,23,31\} \\ 0 & \text{otherwise} \end{cases}. \tag{F6}$$

Together with $p_{\mathrm{Q}}^{\theta}(b) = \frac{1}{4}$ for all b, this gives

$$\mathcal{B}' = 6\sqrt{1+\tfrac{V_1}{2}\cos\theta} + 6\sqrt{1+\tfrac{V_2}{2}\cos\theta} + 6\sqrt{1+\tfrac{V_1V_2}{2}(1+\sin\theta)} + 6\sqrt{1+\tfrac{V_1V_2}{2}(1-\sin\theta)}. \tag{F7}$$

In the noiseless case ($V_1 = V_2 = 1$), this gives a violation of the bilocal inequality (F5) for all θ in the range $0 \leq \theta \lesssim 0.254\,\pi$, with a maximal value of

$$\mathcal{B}' = 12\sqrt{6} \approx 29.39, \tag{F8}$$

obtained for $\theta = 0$. For $\theta = 0$ precisely, allowing now for symmetric noise, we find a violation for visibilities $V_1 = V_2$ larger than the critical visibility $V_{\text{crit}} \approx 88.0\,\%$.

3. Comparison between our two bilocal inequalities

From the results above it seems that our second bilocal inequality, Eq. (F5), is less powerful than our first one, Eq. (9), at detecting the non-bilocality of the quantum correlation p_{Q}^{θ}: indeed it detects it only for a restricted range of θ, and for larger visibilities. More generally, we find that any choice of parameters (V_1, V_2, θ) for which p_{Q}^{θ} violates our second inequality already violates our first inequality.

However, looking beyond the specific quantum correlation p_{Q}^{θ}, one can find non-bilocal correlations that violate Eq. (F5) but not Eq. (9). An example is for instance given by the (local) correlation defined by

$$\langle A_x\rangle = \langle B^y\rangle = \langle C_z\rangle = \langle A_x C_z\rangle = 0,$$
$$\langle A_x B^y\rangle = -\frac{1}{2}\delta_{x,y}, \quad \langle B^y C_z\rangle = \frac{1}{2}\delta_{y,z},$$
$$\langle A_x B^y C_z\rangle = -\frac{1}{3}\delta_{x\neq y\neq z}, \tag{F9}$$

which gives $(S,T,Z) = (3,-2,0)$ and thus satisfies Eq. (9) (and in fact, satisfies all inequalities (D1) that bound the (S,T)-projection of the bilocal set when $Z = 0$, see Appendix D), while $\mathcal{B}' = 6\sqrt{6} + 8\sqrt{3} \approx 28.55 > 12\sqrt{3} + 2\sqrt{15} \approx 28.53$ violates Eq. (F5).

It is clear that many different bilocal inequalities could be obtained by considering various types of nonlinear functions of the correlations, as we did here with $\mathcal{B}'$. Some may be found to be better-suited for certain correlations of interest, other than p_{Q}^{θ}.

4. Stronger-than-quantum nonlocality

We have also used the Bell expression (F1) to detect stronger-than-quantum network nonlocality. This is achieved by deriving a quantum Bell inequality for the network, i.e. a non-trivial bound on $\mathcal{B}'$ satisfied by all quantum models with two independent sources. The bound is established under the mild restriction that Bob has uniform outcomes, i.e. $p(b) = \frac{1}{4}$ for all b. To this end, we consider the use of a simple concavity inequality to linearize $\mathcal{B}'$: for any $a_1, \ldots, a_n \geq 0$, it holds that

$$\sum_{i=1}^{n}\sqrt{a_i} \leq \sqrt{n\sum_{i=1}^{n} a_i}, \tag{F10}$$

with equality if and only if all a_i are equal. Since the Bell expression $\mathcal{B}'$ is a sum of square-root expressions, using the concavity inequality above allows us to bound it with an expression that is a square-root of the corresponding sums. One thus finds

$$\mathcal{B}' \leq \sqrt{48\,\mathcal{B}'_{\text{lin.}}}, \tag{F11}$$

where we have defined the linear expression

$$\mathcal{B}'_{\text{lin.}} \equiv \sum_{x,b}\frac{1}{4}\left(1 - b^x E_b^{\mathrm{A}}(x)\right) + \sum_{z,b}\frac{1}{4}\left(1 + b^z E_b^{\mathrm{C}}(z)\right) + \sum_{x\neq z,b}\frac{1}{4}\left(1 - b^x b^z E_b^{\mathrm{AC}}(x,z)\right). \tag{F12}$$

We can now bound $\mathcal{B}'_{\text{lin.}}$ for quantum models, with independent sources, by using the semidefinite relaxations of Ref. [34]. Thanks to codes provided by A. Pozas-Kerstjens, we have been able to evaluate the third level SDP relaxation described in [34] and obtain $\mathcal{B}'_{\text{lin.}} \lesssim 19.64$. This corresponds to $\mathcal{B}' \leq \sqrt{48\,\mathcal{B}'_{\text{lin.}}} \lesssim 30.70$.

Compounds of symmetric informationally complete measurements and their application in quantum key distribution

Armin Tavakoli,[1] Ingemar Bengtsson,[2] Nicolas Gisin,[1] and Joseph M. Renes[3]

[1]*Département de Physique Appliquée, Université de Genève, CH-1211 Genève, Switzerland*
[2]*Stockholms Universitet, AlbaNova Fysikum SE-106 91 Stockholm, Sweden*
[3]*Institute for Theoretical Physics, ETH Zürich, Switzerland*

Symmetric informationally complete measurements (SICs) are elegant, celebrated and broadly useful discrete structures in Hilbert space. We introduce a more sophisticated discrete structure compounded by several SICs. A SIC-compound is defined to be a collection of d^3 vectors in d-dimensional Hilbert space that can be partitioned in two different ways: into d SICs and into d^2 orthonormal bases. While a priori their existence may appear unlikely when $d > 2$, we surprisingly answer it in the positive through an explicit construction for $d = 4$. Remarkably this SIC-compound admits a close relation to mutually unbiased bases, as is revealed through quantum state discrimination. Going beyond fundamental considerations, we leverage these exotic properties to construct a protocol for quantum key distribution and analyze its security under general eavesdropping attacks. We show that SIC-compounds enable secure key generation in the presence of errors that are large enough to prevent the success of the generalisation of the six-state protocol.

Introduction.— Quantum information theory has established a permanent link between the foundations of quantum theory and quantum information technologies. This has reinvigorated interest in understanding the ultimate limitations of quantum states and measurements as discrete structures in Hilbert space. Quantum states and measurements have a rich geometry that has no counterpart in classical models. Therefore, it is unsurprising that the most elegant and sophisticated discrete structures that can be found in Hilbert space frequently also are the most celebrated and useful resources for the processing of quantum information.

An outstanding example is known as a *symmetric informationally complete* set of pure quantum states (SIC). A SIC is a maximal set (size d^2) of d-dimensional states, $\{|\phi_k\rangle\}_{k=1}^{d^2}$, with the property that the overlap between any pair of states has the same magnitude:

$$|\langle\phi_k|\phi_l\rangle|^2 = \frac{d\delta_{k,l}+1}{d+1}, \tag{1}$$

where the constant on right-hand-side is fixed by normalisation. Interestingly, a SIC can both be interpreted as a set of states (as above) and as a generalised quantum measurement (positive operator-valued measure, POVM) with d^2 possible outcomes. The measurement operators in such a SIC-POVM are merely the subnormalised projectors of a SIC, namely $\{\frac{1}{d}|\phi_k\rangle\langle\phi_k|\}_{k=1}^{d^2}$.

SICs have been investigated for a long time in many different contexts [1–5]. Their relevance in pure mathematics is remarkably diverse [6–8] and they even have technological applications in high-resolution radar [9] and speech recognition [10]. However, their interest in physics stems from their prominent role in quantum information theory [5]. SIC-POVMs are key tools for quantum state tomography [11–13], which has motivated their experimental realisation in high-dimensional Hilbert spaces [14–16]. Generally, SICs and SIC-POVMs are used in a range of protocols: quantum key distribution (QKD) [17–19], entanglement detection [20–22], device-independent random number generation [23, 24], dimension witnessing [25] and characterisation of quantum devices [26–30]. Moreover, SICs have been studied in the context of quantum nonlocality [24, 31–33] and they have an interesting foundational role in QBism [34]. All this has triggered much interest in addressing the existence of SICs in general Hilbert space dimensions. Presently, existence has been proven numerically at least up to $d = 151$ [5, 35–37] and is conjectured for any d (see [37] for a review).

In this work, we introduce a natural discrete Hilbert space structure that is compounded of many separate SICs. The resulting *SIC-compound* is a set of d^3 pure d-dimensional quantum states, denoted $\{|\psi_{jk}\rangle\}_{jk}$ for $j \in [d^2]$ and $k \in [d]$ (where $[s] = \{1, \ldots, s\}$) with the following two properties:

I For every k, the states $\{|\psi_{jk}\rangle\}_j$ form a SIC.

II For every j, the states $\{|\psi_{jk}\rangle\}_k$ form an orthonormal (ON) basis of Hilbert space.

In a handy terminology, we say that a SIC-compound is composed of d "orthogonal SICs", in the sense that elements numbered j in the d SICs are orthogonal to each other. Indeed, given that the existence of SICs is a longstanding open problem [38], deciding the existence of a SIC-compound for a given d is expected to be even more challenging. A priori, it may seem unlikely that SIC-compounds exist at all when $d > 2$ (it turns out that $d = 2$ is exceptional). We address the existence of SIC-compounds for $d = 3, \ldots, 8$. For $d = 3$ we prove that no SIC-compound exists and for $d = 5, 6, 7, 8$ we give evidence in support of the same conclusion. Remarkably, however, for $d = 4$ we are able to analytically construct a SIC-compound, thus proving that they, in fact, can exist in higher-dimensional Hilbert spaces. The many symmetries of the SIC-compound, which go beyond its defining properties, allow it to be represented as a Latin square. Moreover, we find that the SIC-compound admits a strong connection to mutually unbiased bases (MUBs) which is revealed through quantum state discrimination. Equipped with the fundamental understanding of the SIC-compound, we consider its practical application for quantum information processing. Specifically, we place the SIC-compound at the heart of protocols for QKD, analyze their security under coherent attacks and show their improved

robustness as compared to the four-dimensional counterpart of the six-state protocol [39] (which extends the celebrated BB84 protocol [40]).

Qubit SIC-compound.— It is instructive to first consider the simple example of a qubit SIC-compound. In terms of the Bloch sphere representation, a SIC corresponds to four unit Bloch vectors such that any pair has equal magnitude overlap. Hence, the four vectors point to the vertices of a regular tetrahedron. For each vector, the unique orthogonal state is represented by the antipodal Bloch vector, and therefore the four antipodal Bloch vectors also form a regular tetrahedron. By construction, the two SICs together form a SIC-compound. Their convex hull is a cube inscribed in the Bloch sphere.

Generating SICs.— When $d > 2$, the existence of a SIC-compound is far less clear. In order to address the matter, one benefits much from the established knowledge of SICs which heavily exploits the Weyl-Heisenberg (WH) group. This group has two generators, X and Z, which are required to satisfy the relations $X^d = Z^d = \mathbb{1}$ and $ZX = \omega XZ$, where $\omega = e^{\frac{2\pi i}{d}}$. Every known SIC (with a single exception in dimension 8 [3]) has been obtained by applying the WH group in the following ansatz,

$$|\phi_j\rangle = X^{j_1} Z^{j_2} |\varphi\rangle, \tag{2}$$

for $j \equiv (j_1, j_2) \in [d]^2$ and for a suitably chosen so-called fiducial state $|\varphi\rangle$. The group generators can conveniently be chosen as the so-called shift and clock operators

$$X = \sum_{k=0}^{d-1} |k+1\rangle\langle k| \qquad Z = \sum_{k=0}^{d-1} \omega^k |k\rangle\langle k|. \tag{3}$$

For $d = 2, 3$ all SICs are obtained via this ansatz [41, 42] and the same is true for any prime d provided that the SIC admits some group structure [43]. Moreover, there is numerical evidence supporting that all SICs for $d = 4, 5, 6, 7$ can be obtained via the WH group [44].

No qutrit SIC-compound.— Consider the case of qutrits ($d = 3$). In view of the above, by showing that no SIC-compound can be obtained via the WH group, we disprove their existence in full generality. Note that the problem is substantially simplified due to the fact that Eq (2) generates SICs by unitarily acting on a fiducial state. Therefore, in order to construct orthogonal SICs, we must only find orthogonal fiducial states. However, for qutrit systems there are uncountably many relevant fiducial states [4, 5] (for a fixed representation of the WH group). Fortunately, using the representation in Eq (3), they all admit a simple parameterisation which allows us efficiently investigate their orthogonalities. In Appendix A, we detail the analysis for $d = 3$ and show that no more than two orthogonal SICs can be constructed. An example of two orthogonal SICs is straightforwardly obtained from choosing the two fiducial vectors $|\varphi_1\rangle = \frac{1}{\sqrt{2}}(1, 1, 0)^{\mathrm{T}}$ and $|\varphi_2\rangle = \frac{1}{\sqrt{2}}(1, -1, 0)^{\mathrm{T}}$.

Ququart SIC-compound.— For the case of ququarts ($d = 4$), in contrast to qutrits, there are only 256 fiducial states [45] that yield SICs under the ansatz (2) (for a fixed representation). Within these, one can find a SIC-compound with a simple analytical form. To present it, we change the representation of the WH-group so that the generators are written as [46]

$$X = e^{\frac{i\pi}{4}} \begin{pmatrix} 0 & i & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & i & 0 \end{pmatrix}, \quad Z = e^{\frac{i\pi}{4}} \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{pmatrix}. \tag{4}$$

Note that the global phase factors only serve to ensure the correct sign of X^d and Z^d. Consider also the unitary operators

$$U = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & i \\ 1 & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{pmatrix}, \qquad V = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix}, \tag{5}$$

which generate a projective representation of the Klein four-group $\mathbb{Z}_2 \times \mathbb{Z}_2$. Application of $\mathbb{1}$, U, V, and UV on the vector $|\varphi_1\rangle = (t, i, i, i)^{\mathrm{T}}/n$ produces an orthonormal basis, where $t = \sqrt{2+\sqrt{5}}$ and $n = \sqrt{5+\sqrt{5}}$. Call these states $\{|\varphi_k\rangle\}_{k\in[4]}$. Then it can be easily verified that the states $|\psi_{jk}\rangle = X^{j_1} Z^{j_2} |\varphi_k\rangle$ form a SIC for each value of k, where $j = (j_1, j_2)$. By construction, the states $\{|\psi_{jk}\rangle\}_{k\in[4]}$ form an ON-basis for each of the 16 values of j. We remark that if the computational basis is chosen as separable, all 64 states are iso-entangled [47]; the entanglement negativity is $\frac{1}{n^2}\sqrt{1+t^2}$. This constitutes an interesting parallel to the concept of iso-entangled MUBs [48] (which upholds the same degree of entanglement per state as the SIC-compound [49]).

By definition, the ququart SIC-compound contains four SICs and 16 ON-bases of $\mathbb{C}^4$. Interestingly, it turns out that it upholds two additional symmetries (that have no counterpart in the qubit SIC-compound). Firstly, a careful examination of $\{|\psi_{jk}\rangle\}_{j,k}$ shows that every state is not a member of precisely one ON-basis, but in fact of two different ON-bases. Therefore, the SIC-compound houses an additional 16 ON-bases. Secondly, one finds that every state $|\psi_{jk}\rangle$ upholds the defining (SIC-like) overlap property (1) with 27 other states in the SIC-compound, instead of the expected 15. The additional 12 SIC-like overlaps originate from an additional SIC which shares four states with the defining SIC in the compound. Thus, every state is a member of two distinct SICs (see Ref [47] and Appendix C) that have four elements in common.

Since we are now faced with a total of 8 SICs and 32 ON-bases present in the compound, one benefits from nicely organising the elements. A useful observation is that for each of the four defining SICs, one can find four sets of four states such that each is an orbit under the WH subgroup $\{\mathbb{1}, X^2, Z^2, X^2Z^2\}$ (again a projective Klein four-group). By suitably permuting the label $j \in [16]$ in $\{|\psi_{jk}\rangle\}_{jk}$, so that j_1 indexes the subgroup and j_2 indexes the application of $\mathbb{1}$, X, Z, and XZ, we can group these orbits together and represent the SIC-compound as a Latin square (see Figure 1).

On existence in $d = 5, 6, 7, 8$.— For dimensions $d = 5, 6, 7, 8$, using the representation (3), there are only finitely many relevant fiducial states to be considered [45, 50]. The number of states that yield SICs when the WH group, in the representation (3), is applied to them can be regarded as

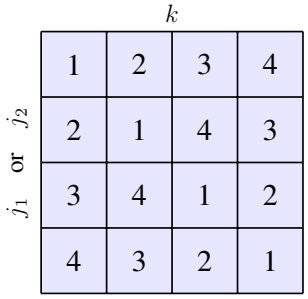

FIG. 1: Schematic of the 64 states in the ququart SIC-compound. First, let us index the columns by $k \in [4]$ and the rows by $j_2 \in [4]$ and let each block contain the four states $\{|\psi_{jk}\rangle\}_{j_1=1}^{4}$ (recall $j = (j_1, j_2)$). Then, each column corresponds to one of the defining SICs of the compound. The collection of elements in the identically labelled ('1','2','3' and '4') blocks constitute the four additional SICs present in the compound. Secondly, let us view the Latin square as an illustration of the 16 individual states in each row of the previous interpretation. The block with coordinates (j_2, k) corresponds to the state $|\psi_{jk}\rangle$ (for any chosen row index j_1). Each row (of four states) then corresponds to a defining ON-basis. For $j_1 = 1, 2, 3, 4$, the collection of elements with identical labels ('1','2','3' and '4') constitute the total of 16 additional ON-bases present in the compound.

known if we combine the high quality numerical results of Ref [35] with the group theoretical analysis of Ref [45]. We have enumerated all of them and exhaustively checked the number of orthogonal SICs that can be constructed using these states. We find that the number of orthogonal SICs varies (2, 4, 2 and 5 respectively) and that no SIC-compound can be constructed. Reminding ourselves of the strong numerical evidence in support of there not existing any other SICs than those that we have explicitly constructed for $d = 5, 6, 7$, our results render the existence of a SIC-compound for $d = 5, 6, 7$ very unlikely. However, as previously mentioned, dimension 8 also houses SICs that are not based on the WH group [3, 51, 52]. Whether a SIC-compound can be formed from these exceptional SICs is left as an open question.

Furthermore, in Appendix B we present a method for certifying [53] a SIC-compound (if it exists) or falsifying their existence (if it does not exist) under the sole assumption of dimension d.

Discriminating the SIC-compound with MUBs.—The ququart SIC-compound admits a simple operational relation to a set of four MUBs. Consider that for fixed j_2 and k, we try to discriminate between the four (equiprobable) states $\{|\psi_{j,k}\rangle\}_{j_1}$. Since these states are linearly independent, we can use the "pretty good measurement" [54] which is the ON-basis obtained from $|\xi_{j,k}\rangle = T_{j_2,k}^{-1/2}|\psi_{j,k}\rangle$ by varying j_1, where $T_{j_2,k} = \sum_{j_1} |\psi_{j,k}\rangle\langle\psi_{j,k}|$. Measuring in this basis is in fact optimal for minimising the error probability of the discrimination, which follows from [55]. Moreover, the resulting bases for given j_2 but different k are identical, while the bases for different j_2 are mutually unbiased. Thus, the four rows of the Latin square correspond to four MUBs which, interestingly, are iso-entangled with the largest possible entanglement negativity (each basis element has an entanglement negativity of $\frac{1}{\sqrt{8}}$). In Appendix C, we show that the relation between the SIC-compound and the four MUBs is not a coincidence but traces back to the fact that the Clifford group contains a copy of the bipartite WH group. Finally, we note that also the fifth MUB (the computational basis) emerges from state discrimination in the SIC-compound: a state is randomly sampled from a given column of the Latin square and we are asked to determine which row it belongs to. The optimal measurement is the computational basis.

Application in QKD.— Let us now consider the usefulness of the $d = 4$ SIC-compound in QKD. Consider a prepare and measure QKD scheme in which Alice transmits a random state $|\psi_{jk}\rangle$ and Bob randomly measures in one of the 16 defining ON bases of the SIC-compound. A variety of specific QKD protocols can be constructed from this starting point, depending on how Alice and Bob transform or "sift" their resulting data into the "raw key". Here we focus on just two sifting protocols. As in the original BB84 protocol, Alice and Bob can use the bases in the compound, taking their k values as the sifted key when their j_1 and j_2 values both match. Call this Sifting B. Another possibility, which we denote Sifting A, is that j_1 is taken as the sifted key value when their j_2 values agree, but their k values *disagree*. (This turns out to be slightly more favorable than when the k values match.) Both protocols finish with the standard steps of parameter estimation, information reconciliation, and privacy amplification to output a secure key. Since both protocols use the same prepare and measure setup but differ only in the classical postprocessing, we will see that Alice and Bob can first perform parameter estimation on their data and then decide which sifting strategy to employ.

We establish the security of both protocols against arbitrary attacks by adapting the methods of [56–58] to ensure security against collective attacks and then invoking [59] to ensure security against arbitrary attacks. The analysis proceeds in the entanglement-based scenario of the protocol. Here Eve supplies Alice and Bob with many copies of an arbitrary bipartite state ρ_{AB}, to which she retains the purification in system E, and Alice and Bob each randomly measure the bases associated with the compound on their respective subsystems. The resulting statistics of their classical measurement choices j and results k, as well as the possible collective attacks, are precisely the same as the prepare-and-measure scenario.

Crucially, the symmetries of the SIC-compound translate into symmetries of both sifting protocols, and this simplifies the form of ρ_{AB}. As we show in Appendix D, for both Sifting A and B we can assume without loss of generality that $\rho_{AB} = (1-p-q)\Phi_{AB} + q\pi_{AB} + p\kappa_{AB}$ for some positive parameters q, p with $q + p \leq 1$, where Φ_{AB} is the maximally-entangled state, π_{AB} is the maximally-mixed state, and κ_{AB} is the diagonal state of perfect uniform correlation. In other words, the joint state is a partially depolarized and dephased maximally-entangled state.

Alice and Bob can determine both p and q in the parameter estimation phase as follows. It turns out that the probability of sifting success for Sifting A increases with increasing q, while

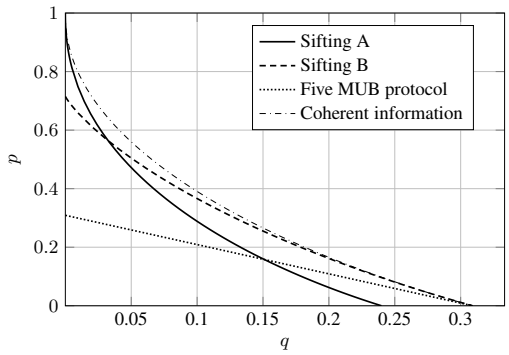

FIG. 2: Regions of positive key rate for various protocols. For each q, the curves show the value of p such that the key rate is zero. Sifting B outperforms the analog of the qubit six-state protocol using a full set of five MUBs. Sifting A can tolerate $p \to 1$ as $q \to 0$. Together, Sifting A and B nearly replicate the region of positive coherent information $-H(A|B)_\rho$ from the state ρ_{AB}.

the probability of error in the raw key depends on both p and q. Therefore, before they commit to either sifting procedure, Alice and Bob can use their data to determine both parameters and only then decide which sifting procedure is more appropriate. Knowing the state ρ_{AB}, it is then a simple matter to apply known bounds on the rate of key extraction using information reconciliation and privacy amplification.

Fig. 2 depicts the values of q and p which lead to positive key rates. It also displays the region of positive key for the generalisation of the six-state protocol to $d = 4$ (using a full set of five MUBs). To enable a fair comparison, the latter protocol also discards sifting information [61]. Its symmetries ensure that it treats all states delivered by Eve as depolarized maximally-entangled states, so that when the actual joint state is of the form ρ_{AB} above, it sees a depolarization rate of $1 - p - q$. Therefore, the region of positive rate for the five MUBs protocol is symmetric under interchange of p and q. Using the rate expression derived in [62], we find the threshold for $p = 0$ to be $q \approx 0.309$. This is also the threshold of the Sifting B protocol.

Conclusions.— We have introduced SIC-compounds as an elegant and sophisticated discrete structure in Hilbert space. Against initial intuition, we found that SIC-compounds can exist beyond qubit systems and explicitly constructed a four-dimensional SIC-compound. We found that it upholds many unexpected symmetries as well as an operational connection to mutually unbiased bases. Then, through our example of SIC-compounds, we illustrated that foundational understanding of discrete structures of quantum systems not only are interesting in themselves but that they also serve as new, powerful, tools for quantum information processing. We applied SIC-compounds towards quantum key distribution and showed that they can produce secure key in relevant situations in which the generalisation of the six-state protocol no longer is useful.

Lastly, we ask whether four-dimensional SIC-compounds can be used to construct interesting entangled measurements of two (or more) four-dimensional systems; generalising the measurements of [65, 66] .

Acknowledgments

This work was supported by the Swiss National Science Foundation via the NCCR-SwissMap.

[1] O. Hesse, Über die Wendepuncte der Curven dritter Ordnung, J. Reine Angew. Math. **28**, 97 (1844).
[2] P. Delsarte, J. M. Goethels and J.J. Seidel, Bounds for systems of lines and Jacobi polynomials. Philips Res. Rep. **30**, 91 (1975); reprinted in Geometry and Combinatorics, D.G. Corneil and R. Mathon, Eds.; Academic Press, San Diego, USA, 1991.
[3] S.G. Hoggar, Two quaternionic 4-polytopes, In The Geometric Vein: The Coxeter Festschrift; C. Davis, B. Grünbaum, F. A. Sherk, Eds.; Springer: New York, NY, USA, 1981.
[4] G. Zauner: Quantendesigns. Grundzuge einer nichtkommutativen Designtheorie, PhD thesis, Universität Wien, 1999; also published as Quantum designs: Foundations of a noncommutative design theory, Int. J. Quant. Inf. **9**, 445 (2011).
[5] J. M. Renes, R. Blume-Kohout, A. J. Scott, C. M. Caves, Symmetric informationally complete quantum measurements, J. Math. Phys. **45**, 2171 (2004).
[6] M. Appleby, H. Yadsan-Appleby and G. Zauner, Galois automorphisms of a symmetric measurement. Quantum Inf. Comput. **13**, 672 (2013).
[7] M. Appleby, S. Flammia, G. McConnell and J. Yard, SICs and Algebraic Number Theory Found. Phys. **47**, 1042 (2017).
[8] D. M. Appleby, C. A. Fuchs, and H. Zhu, Group theoretic, Lie algebraic and Jordan algebraic formulations of the sic existence problem, Quantum Inf. Comput. **15**, 61 (2015).
[9] S. D. Howard, A. R. Calderbank and W. Moran, The finite Heisenberg-Weyl groups in radar and communications, EURASIP J. Appl. Signal Process. 1 (2006).
[10] R. Balan, B.G. Bodmann, P. G. Casazza and D. Edidin, Painless reconstruction from magnitudes of frame coefficients, J. Fourier Anal. Appl. **15**, 488 (2009).
[11] A. J. Scott, Tight informationally complete quantum measurements, J. Phys. A: Math. Gen. **39** 13507 (2006).
[12] H. Zhu and B-G. Englert, Quantum state tomography with fully symmetric measurements and product measurements, Phys. Rev. A **84**, 022327 (2011).
[13] D. Petz and L. Ruppert, Efficient quantum tomography needs complementary and symmetric measurements, Rep. Math. Phys. **69**, 161 (2012).
[14] Z. E. D. Medendorp, F. A. Torres-Ruiz, L. K. Shalm, G. N. M. Tabia, C. A. Fuchs, and A. M. Steinberg, Experimental characterization of qutrits using symmetric informationally complete positive operator-valued measurements, Phys. Rev. A **83**, 051801(R) (2011).
[15] W. M. Pimenta, B. Marques, T. O. Maciel, R. O. Vianna, A.

Delgado, C. Saavedra, and S. Pádua, Minimum tomography of two entangled qutrits using local measurements of one-qutrit symmetric informationally complete positive operator-valued measure, Phys. Rev. A **88**, 012112 (2013).
[16] N. Bent, H. Qassim, A. A. Tahir, D. Sych, G. Leuchs, L. L. Sánchez-Soto, E. Karimi, and R. W. Boyd, Experimental realization of quantum tomography of photonic qudits via symmetric informationally complete positive operator-valued measures, Phys. Rev. X **5**, 041006 (2015).
[17] J. M. Renes, Equiangular spherical codes in quantum cryptography, Quant. Inf. Comput. **5**, 080 (2005).
[18] B-G. Englert, D. Kaszlikowski, H. K. Ng, W. K. Chua, J. Řeháček, and Janet Anders, Efficient and robust quantum key distribution with minimal state tomography, arXiv:quant-ph/0412075v4
[19] F. Bouchard, K. Heshami, D. England, R. Fickler, R. W. Boyd, B-G. Englert, L. L. Sánchez-Soto, and E. Karimi, Experimental investigation of high-dimensional quantum key distribution protocols with twisted photons, Quantum **2**, 111 (2018).
[20] A. Kalev and J. Bae, Optimal approximate transpose map via quantum designs and its applications to entanglement detection, Phys. Rev. A **87**, 062314 (2013).
[21] J. Shang, A. Asadian, H. Zhu, and O. Gühne, Enhanced entanglement criterion via symmetric informationally complete measurements, Phys. Rev. A **98**, 022309 (2018).
[22] J. Bae, B. C. Hiesmayr, and D. McNulty, Linking entanglement detection and state tomography via quantum 2-designs, New J. Phys. **21** 013012 (2019).
[23] A. Acín, S. Pironio, T. Vértesi, and P. Wittek, Optimal randomness certification from one entangled bit, Phys. Rev. A **93**, 040102(R) (2016).
[24] A. Tavakoli, M. Farkas, D. Rosset, J-D. Bancal, and J. Kaniewski, Mutually unbiased bases and symmetric informationally complete measurements in Bell experiments: Bell inequalities, device-independent certification and applications, arXiv:1912.03225
[25] N. Brunner, M. Navascués and T. Vértesi Dimension Witnesses and Quantum State Discrimination, Phys. Rev. Lett. **110**, 150501 (2013).
[26] A. Tavakoli, D. Rosset, and M-O. Renou, Enabling computation of correlation bounds for finite-dimensional quantum systems via symmetrization, Phys. Rev. Lett. **122**, 070501 (2019).
[27] A. Tavakoli, M. Smania, T. Vértesi, N. Brunner, and M. Bourennane, Self-testing non-projective quantum measurements in prepare-and-measure experiments, Science Advances **6**, 16 (2020).
[28] M. Smania, P. Mironowicz, M. Nawareg, M. Pawlowski, A. Cabello, and M. Bourennane, Experimental device-independent certification of a symmetric, informationally complete, positive operator-valued measure, Optica **7**, 123 (2020).
[29] P. Mironowicz and M. Pawłowski, Experimentally feasible semi-device-independent certification of four-outcome positive-operator-valued measurements, Phys. Rev. A **100**, 030301(R) (2019).
[30] A. Tavakoli, Semi-device-independent certification of independent quantum state and measurement devices, arXiv:2003.03859
[31] T. Vértesi and E. Bene, Two-qubit Bell inequality for which positive operator-valued measurements are relevant, Phys. Rev. A **82**, 062115 (2010).
[32] A. Tavakoli and N Gisin, The Platonic solids and fundamental tests of quantum mechanics, arXiv:2001.00188
[33] N. Gisin, Entanglement 25 Years after Quantum Teleportation: Testing Joint Measurements in Quantum Networks, Entropy **21**, 325 (2019).
[34] C. A. Fuchs and R. Schack, Quantum-Bayesian coherence, Rev. Mod. Phys. **85**, 1693 (2013).
[35] A. J. Scott and M. Grassl, SIC-POVMs: A new computer study, J. Math. Phys. **51**, 042203 (2010)
[36] A. J. Scott, SICs: Extending the list of solutions, arXiv:1703.03993
[37] C. A. Fuchs, M. C. Hoang, and B. C. Stacey, The SIC Question: History and State of Play, Axioms **21**, 6 (2017).
[38] P. Horodecki, L. Rudnicki, and K. Życzkowski, Five open problems in quantum information, arXiv:2002.03233
[39] Dagmar Bruß, Optimal Eavesdropping in Quantum Cryptography with Six States Phys. Rev. Lett. **81**, 3018 (1998).
[40] C. H. Bennett and G. Brassard, Quantum cryptography: Public key distribution and coin tossing, Proceedings of IEEE International Conference on Computers, Systems and Signal Processing, volume 175, page 8, New York 1984; reprinted in Theoretical Computer Science **560**, 7 (2014).
[41] F. Szöllősi, All complex equiangular tight frames in dimension 3, arXiv:1402.6429
[42] L. P. Hughston and S. M. Salamon, Surveying points in the complex projective plane, Adv. in Mathematics **286**, 1017 (2016).
[43] H. Zhu, SIC POVMs and Clifford groups in prime dimensions, J. Phys. A: Math. Theor. **43** 305305 (2010)
[44] H. Zhu, Quantum State Estimation and Symmetric Informationally Complete POMs, Doctoral thesis, National University of Singapore (2012).
[45] D. M. Appleby, Symmetric informationally complete-positive operator valued measures and the extended Clifford group, Journal of Mathematical Physics **46**, 052107 (2005).
[46] D. M. Appleby, I. Bengtsson, S. Brierley, M. Grassl, D. Gross, and J.-Å. Larsson, The monomial representations of the Clifford group, Quantum Inf. Comp. **12**, 0404 (2012).
[47] H. Zhu, Y. S. Teo, and B-G. Englert, Structure of Two-qubit Symmetric Informationally Complete POVMs, Phys. Rev. A **82**, 042308 (2010).
[48] J. Czartowski, D. Goyeneche, M. Grassl, and Karol Życzkowski, Iso-entangled mutually unbiased bases, symmetric quantum measurements and mixed-state designs, Phys. Rev. Lett. **124**, 090503 (2020).
[49] Note that the negativity, $\mathcal{N}$, is one-to-one with the purity, T, of a subsystem; $T = 1 - 2\mathcal{N}^2$. The average purity in both SICs and complete sets of MUBs is determined by the 2-design property — which explains their equal degree of entanglement.
[50] Specifically 2000, 3456, 16464 and 24576 for $d = 5, 6, 7, 8$ respectively. In dimensions 7 and 8 there are two inequivalent sets of fiducial states - both of which must be taken into account ($5488 + 10976 = 16464$ and $8192 + 16384 = 24576$). See [45]. In dimension 8, the are four orthogonal SICs in the smaller orbit which grows to five if also the second orbit is included.
[51] A. Szymusiak and W. Słomczyński, Informational power of the Hoggar SIC-POVM, Phys. Rev. A **94**, 012122 (2016).
[52] B. C. Stacey, Sporadic SICs and Exceptional Lie Algebras arXiv:1911.05809
[53] A. Tavakoli, J. Kaniewski, T. Vértesi, D. Rosset and N. Brunner, Self-testing quantum states and measurements in the prepare-and-measure scenario, Phys. Rev. A **98**, 062307 (2018).
[54] P. Hausladen and W. K. Wootters, A 'Pretty Good' Measurement for Distinguishing Quantum States, Journal of Modern Optics, **41**, 12 (1994).
[55] S. M. Barnett, Minimum-error discrimination between multiply symmetric states Phys. Rev. A **64**, 030303 (2001).
[56] B. Kraus, N. Gisin, and R. Renner, Lower and Upper Bounds

on the Secret-Key Rate for Quantum Key Distribution Protocols Using One-Way Classical Communication, Phys. Rev. Lett. **95**, 080501 (2005).
[57] R. Renner, N. Gisin, and B. Kraus, Information-theoretic security proof for quantum-key-distribution protocols, Phys. Rev. A **72**, 012332 (2005).
[58] J. M. Renes and M. Grassl, Generalized decoding, effective channels, and simplified security proofs in quantum key distribution, Phys. Rev. A **74**, 022317 (2006).
[59] M. Christandl, R. König, and R. Renner, Postselection Technique for Quantum Channels with Applications to Quantum Cryptography, Phys. Rev. Lett. **102**, 020504 (2009).
[60] I. Devetak and A. Winter, Distillation of secret key and entanglement from quantum states, Proc. Roy. Soc. A **461**, 207 (2005).
[61] We could consider tomographic versions of all three protocols, as the SIC-compound is tomographically complete, but the resulting protocols are more complicated as they must estimate a large number of attack parameters.
[62] L. Sheridan and V. Scarani, Security proof for quantum key distribution using qudit systems, Phys. Rev. A **82**, 030301 (2010).
[63] H. Zhu, Twin Heisenberg-Weyl groups and the Clifford hierarchy, unpublished manuscript.
[64] K. Blanchfield, Orbits of mutually unbiased bases, J. Phys. A: Math. Theor. **47** 135303 (2014).
[65] N. Gisin, Entanglement 25 Years after Quantum Teleportation: Testing Joint Measurements in Quantum Networks, Entropy **21**, 325 (2019).
[66] A. Tavakoli, N. Gisin and C. Branciard, Bilocal Bell inequalities violated by the quantum Elegant Joint Measurement, arXiv:2006.16694
[67] M. Navascués and T. Vértesi, Bounding the Set of Finite Dimensional Quantum Correlations, Phys. Rev. Lett. **115**, 020501 (2015).

Appendix A: No SIC-compound for $d = 3$

We fix the representation of the WH group to $X = \sum_{k=0}^{d-1} |k+1\rangle\langle k|$ and $Z = \sum_{k=0}^{d-1} \omega^k |k\rangle\langle k|$. For this fixed representation, we prove that no SIC-compound exists for $d = 3$. It is known that there are infinitely many fiducial states in $d = 3$ [4, 5]. They can be parameterised using a complete set of mutually unibased bases, which can be written (without normalisation) as follows:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{bmatrix}, \quad \begin{bmatrix} 1 & \omega & \omega \\ \omega & 1 & \omega \\ \omega & \omega & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & \omega^2 & \omega^2 \\ \omega^2 & 1 & \omega^2 \\ \omega^2 & \omega^2 & 1 \end{bmatrix}. \tag{A1}$$

All [41, 42] fiducial states can be obtained via the following [1] procedure. Choose any one of the four bases. Then, choose any pair of elements within the basis. Denote the first element by $|e_1\rangle$ and the second element by $|e_2\rangle$. The vector $|\phi\rangle = \left(|e_1\rangle - e^{i\theta}|e_2\rangle\right)/\sqrt{2}$, for any $\theta \in [0, 2\pi]$, is a valid fiducial state. Repeating this procedure for all twelve relevant pairs appearing in Eq. (A1), one obtains the complete set of fiducial states.

The task of showing that no three fiducial states can form an ON-basis is significantly simplified by the fact that the problem is invariant in such a way that we can without loss of generality choose the first fiducial vector correspoding to the two first elements of the first basis in Eq. (A1), namely $|\phi_1\rangle = \left(|0\rangle - e^{i\theta_1}|1\rangle\right)/\sqrt{2}$. Moreover, since every basis in Eq. (A1) can be transformed into every other basis in Eq. (A1), it is sufficient to search for an ON-basis with respect to all fiducial states associated to, for instance, the second basis. We name the three elements of the second basis (represented in Eq. (A1) by the Fourier matrix) $\{|f_1\rangle, |f_2\rangle, |f_3\rangle\}$. Writing $|\phi_2\rangle = \left(|f_1\rangle - e^{i\theta_2}|f_2\rangle\right)/\sqrt{2}$, we straightforwardly obtain that

$$0 \stackrel{!}{=} \langle\phi_1|\phi_2\rangle \Leftrightarrow \begin{cases} \cos\theta_1 - \cos(\theta_1+\theta_2) - \cos(\theta_2+\frac{\pi}{3}) = 1 \\ \sin\theta_1 - \sin(\theta_1+\theta_2) - \sin(\theta_2+\frac{\pi}{3}) = 0. \end{cases} \tag{A2}$$

The solutions are found at $(\theta_1, \theta_2) = (2\pi/3, \pi/3)$ and $(\theta_1, \theta_2) = (5\pi/3, 4\pi/3)$. To show that no third orthogonal fiducial state exists, we also consider the cases of $|\phi_2'\rangle = \left(|f_1\rangle - e^{i\theta_3}|f_3\rangle\right)/\sqrt{2}$ and $|\phi_2{''}\rangle = \left(|f_2\rangle - e^{i\theta_2}|f_3\rangle\right)/\sqrt{2}$. These give equations analogous to Eq. (A2), each with two solutions. Inspecting these few cases, one easily finds that no orthogonalities exist among these solutions. Thus, we conclude that no qutrit SIC-compound exists. However, as is clear from the above, it is possible to construct two orthogonal qutrit SICs. The perhaps easiest example corresponds to the two orthogonal fiducial states

$$|\phi_1\rangle = \frac{(1,1,0)^{\mathrm{T}}}{\sqrt{2}}, \qquad |\phi_2\rangle = \frac{(1,-1,0)^{\mathrm{T}}}{\sqrt{2}}. \tag{A3}$$

Appendix B: Certification and falsification of SIC-compounds

We show that SIC-compounds can be certified in a semi-device-independent manner [53] (provided that they exist) and that existence can be disproved using hierarchies of increasingly precise necessary conditions that each can be evaluated as a semidefinite program.

Consider a prepare-and-measure scenario in which Alice has a random input $x \in [d^2]$ and Bob has an input (y, y') which labels all pairs of elements in $[d^2]$. For convention, we take $y < y'$. Each measurement of Bob has binary outcomes $b \in [2]$. Alice's states are of dimension no greater than d. In Refs [25, 26], it was shown that the quantum maximum of the following functional

$$S' = \sum_{(y,y')} p(b=1|y,(y,y')) + p(b=2|y',(y,y')) \tag{B1}$$

is uniquely achieved in by Alice's states forming a SIC. Thus, it semi-device-independently certifies SIC preparations. Moreover, one can add another (single) setting to Bob, $z \in [1]$, which has $o \in [d^2]$ possible outcomes, such that the modified functional

$$S = S' + \sum_{x=1}^{d^2} p(o=x|x,z=1) \tag{B2}$$

achieves its quantum maximum when both S' and the above sum individually are maximal. The optimal quantum value obeys [26]

$$\max_Q S \leq \frac{1}{2}\sqrt{d^5(d-1)^2(d+1)} + \binom{d^2}{2} + d, \qquad \text{(B3)}$$

which can be saturated if and only if Alice prepares a SIC (provided it exists) and the setting z corresponds to the aligned SIC-POVM (obtained from Alice's sub-normalised preparations).

We will use this already known communication game for SICs as a building block to construct a communication game for SIC-compounds. Let Alice have inputs $x \in [d^2]$ and $i \in [d]$. Bob takes inputs (y, y') and $j \in [d]$ and returns a binary outcome. Moreover, Bob additionally has d settings labelled $z \in [d]$ which have d^2 possible outcomes. We are only interested in cases in which $r \equiv i = j$. Let Alice and Bob play the above game (for SICs) d times in parallel: each implementation (indexed by r) uses the preparations $\{(x, i = r)\}_x$ and the measurements $\{(y, y', j = r) \cup (z = r)\}_{y,y'}$. We label the score in the r'th game by S_r. Naturally, these scores are so far independent since they each correspond to independent sets of preparations and measurements. If all S_r are maximal, it thus certifies that Alice and Bob have implemented d independent pairs of SIC preparations and SIC-POVMs. In order to certify a SIC-compound, we need to enforce the orthogonality of the d SICs.

To that end, we add a penalty term. If Alice's preparation is (x, i) and Bob implements one of his additional settings with $z \neq i$, then the outcome $o = x$ must never occur. If this holds true for every $(x, i, z \neq i)$, it is equivalent to a SIC-compound given that we already know that Alice must prepare SICs. Therefore, we choose our final correlation functional as

$$H = \frac{1}{d}\sum_{r=1}^{d} S_r - \sum_{\substack{x \\ i \neq z}} p(o = x|(x, i), z). \qquad \text{(B4)}$$

Using (B3) it follows that

$$\max_Q H \leq \max_Q S, \qquad \text{and that} \qquad \text{(B5)}$$

$$H = \max_Q S \Leftrightarrow \text{Alice prepares a SIC-compound.} \qquad \text{(B6)}$$

Thus, we have constructed a quantum communication game in which the optimal correlations are uniquely attained by SIC-compounds.

This has two notable consequences. Firstly, we may numerically search for SIC-compounds by attempting to maximise H (which can be efficiently done through alternating convex searches). Secondly, if one can prove that H cannot attain the value (B3) in a quantum model, one falsifies the existence of any SIC-compound in the given dimension. To enable such a proof, one can use the hierarchy of semidefinite relaxations of the set of dimensionally restricted quantum correlations [67]. However, the computational requirements are significant due to the large number of preparations and measurements. Nevertheless, semidefinite relaxations can be evaluated by employing the symmetrisation techniques of Ref [26]. For instance, we consider the (trivial) case of deciding the existence of three orthogonal SICs for $d = 2$. The existence of a SIC-compound would enable $H \approx 12.899$ while our semidefinite relaxation proves that no larger value is possible in quantum theory than $H \approx 12.728$. We could also evaluate the case of three orthogonal SICs in dimension three, but were unable to obtain a bound on H smaller than that achieved by a SIC-compound (our SDP matrix is of size 3915). The falsification (which we have already shown analytically) could require a higher-level relaxation.

Appendix C: SIC-compounds and MUBs in dimension four

Standard lore has it that SICs and MUBs are unrelated in four dimensions. SICs appear as orbits of the Weyl–Heisenberg group, and the SIC-compound is an orbit under a subgroup of the normalizer of the Weyl–Heisenberg group. MUBs on the other hand are obtained from the bipartite Heisenberg group. Since the two groups are different, one does not expect a connection between SICs and MUBs. Nevertheless we found a connection, and it is interesting to see how this arises.

To see this we first recapitulate the analysis by Zhu et al. [47, 63], which shows that in this dimension the Clifford group contains two normal copies of the Weyl–Heisenberg group. The Clifford group contains the symplectic group $SL(2)$ with matrix elements chosen to be integers modulo 8. Its representation is fixed once the representation of the Weyl–Heisenberg group is fixed [45]. The subgroup of $SL(2)$ that transforms a given compound to itself is generated by the order 4 symplectic matrices

$$G_1 = \begin{pmatrix} 3 & 0 \\ 6 & 3 \end{pmatrix}, \qquad G_2 = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}, \qquad \text{(C1)}$$

together with an order 3 Zauner matrix [45] which plays no role in this Appendix. The corresponding unitaries are denoted U_{G_1} and U_{G_2}. The generators of the twin Weyl–Heisenberg group are then represented by [47, 63].

$$\tilde{X} = e^{\frac{i\pi}{4}} U_{G_2} X Z = e^{\frac{i\pi}{4}} \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}, \qquad \text{(C2)}$$

$$\tilde{Z} = U_{G_1} Z = e^{\frac{i\pi}{4}} \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & i \\ i & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \qquad \text{(C3)}$$

The presence of this 'extra' Weyl–Heisenberg group explains why the $4 \cdot 16$ vectors in the compound can be regrouped in such a way that $4 + 4$ SICs appear [63].

But the bipartite Heisenberg group is lurking here as well. A straightforward calculation verifies that

$$X^2 = \sigma_z \otimes \mathbb{1} \,, \qquad Z^2 = \mathbb{1} \otimes \sigma_z \,,$$
$$-iX\tilde{Z} = \sigma_y \otimes \sigma_y \,, \qquad Z\tilde{X} = \mathbb{1} \otimes \sigma_x \,, \tag{C4}$$

where $\sigma_x, \sigma_y, \sigma_z$ are the usual Pauli matrices. These local operators generate the bipartite Heisenberg group, they leave a given SIC compound invariant, and they can be used to create the MUBs mentioned in the main text.

The usual construction of five MUBs proceeds by dividing the bipartite Heisenberg group into maximal abelian subgroups. In the main text we obtained 4 MUBs, all of them unbiased relative to the computational basis, as an orbit under the bipartite Heisenberg group. This is the Alltop construction of MUBs. The fact that this construction works in dimension 4 is already known [64], but the relation to the Weyl-Heisenberg Clifford group is new.

Appendix D: QKD security proof details

Following [56–58], we can treat the sifting operation as a quantum operation as follows. Since the SIC-compound forms a single POVM, measurement can be described by the isometry $|\phi\rangle \mapsto \frac{1}{4}\sum_{jk} |j\rangle|k\rangle\langle\psi_{jk}|\phi\rangle$, followed by usual projective meaurement of the $|j\rangle$ and $|k\rangle$ registers. Sifting can then be regarded as projective measurement of the appropriate registers, either (j_2, k) or (j_1, j_2), followed by postselection based on comparing the results using public communication. Thus, each (j_2, k) combination in Sifting A, for instance, gives rise to a Kraus operator $S_{j_2,k}$ which maps the AB system to the raw keys $K_A K_B$ according to $S_{j_2,k} : |\phi\rangle_A \otimes |\psi\rangle_B \mapsto N\sum_{i,i'} |i\rangle_{K_A}|i'\rangle_{K_B}\langle\psi_{i,j_2,k}|_A\langle\psi^*_{i',j_2,k}|_B$, where N is a normalization factor. (Recall that the conversion requires Bob to use the complex conjugate states $|\psi^*_{jk}\rangle$.) The case of Sifting B is entirely similar.

In this formalism it is now easy to confirm that the sifting procedure is covariant under the automorphism G of the SIC-compound, which is generated by X, Z, U, V, and one further unitary operator, W, which cyclically permutes the last three vector components and leaves the first fixed. Then, in the case of Sifting A, for any element $Y \in G$ and combination (j_2, k), the operator $S_{j_2,k}Y \otimes Y^* = S_{j'_2,k'}$ for some j'_2 and k' (up to a phase), because the automorphism generators each preserve the individual rows and columns of the Latin square. Importantly, in both sifting procedures under consideration, the protocol discards the information besides the sifted key, e.g. the (j_2, k) values in Sifting A and the (j_1, j_2) values in Sifting B. Therefore we may average the input state ρ_{AB} over G, since the protocol will effectively only see the state $\bar{\rho}_{AB} = \sum_{Y\in G} Y \otimes Y^* \rho_{AB} Y^\dagger \otimes Y^T$. Straightforward calculation shows that $\bar{\rho}_{AB} = (1-p-q)\Phi_{AB} + q\pi_{AB} + p\kappa_{AB}$ for some positive parameters q, p with $q + p \leq 1$, where Φ_{AB} is the maximally-entangled state, π_{AB} is the maximally-mixed state, and κ_{AB} is the diagonal state of perfect uniform correlation.

The protocol proceeds to distill secret key from the raw key using information reconciliation and privacy amplification. Given a post-sifted state $\sigma_{K_A K_B E}$, we can appeal to the rate formula of [60], $r \geq H(K_A|E)_\sigma - H(K_A|K_B)_\sigma$, where $H(K_A|E)_\sigma$ is the conditional entropy. The post-sifted state will be of the form $\sigma_{K_A K_B E} = \mathcal{M}(S_{1,1}\bar{\rho}_{ABE}S^\dagger_{1,1})$, where $\mathcal{M}$ denotes the measurement of the K_A and K_B systems, each in the standard basis. This is a slight departure from and improvement on [56–58], which for simplicity uses only the Bell-diagonal part of $S_{1,1}\bar{\rho}_{ABE}S^\dagger_{1,1}$. This lowers the key rate and is unnecessary here as the state $\bar{\rho}_{ABE}$ itself is of a very simple form.