EPR steering and its application to fundamental questions in bell nonlocality

Bowles, Joseph

Université de Genève
Département de Physique Théorique

Faculté des Sciences
Professeur N. Brunner

EPR Steering and its Application to Fundamental Questions in Bell Nonlocality

THÈSE

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

par

Joseph Bowles

de

Harrow (Royaume-Uni)

Thèse No 4953

Genève
Atelier Repromail, Université de Genève
2016

UNIVERSITÉ DE GENÈVE

FACULTÉ DES SCIENCES

Doctorat ès sciences
Mention physique

Thèse de *Monsieur Joseph BOWLES*

intitulée :

"EPR Steering and its Application to Fundamental Questions in Bell Nonlocality"

La Faculté des sciences, sur le préavis de Monsieur N. BRUNNER, professeur assistant et directeur de thèse (Département de physique théorique), Monsieur N. GISIN, professeur ordinaire (Groupe de physique appliquée), Monsieur J. BRASK, docteur (Groupe de physique appliquée), Monsieur S. PIRONIO, (Service de physique théorique, Université libre de Bruxelles, Belgique) et Monsieur A. ACIN, professeur (Quantum information theory group, Institute of Phonotic Sciences, Barcelona, Spain) , autorise l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 12 juillet 2016

Thèse - 4953 -

Le Doyen

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

This thesis is dedicated to the study of foundational concepts in quantum theory from a perspective of quantum information theory. We consider the measurement statistics obtained from multipartite quantum systems within the frameworks of Bell nonlocality and EPR steering (Part I), and the communication and processing of single quantum systems in networks of devices (Part II).

One of the most striking aspects of quantum theory is its departure from classical notions of locality. A typical state in quantum theory cannot be written as the product of its local subsystems, but instead must be described by a global state of which the local subsystems are said to be entangled. This inseparability at the mathematical level of the theory is known to persist even at the observational level, where certain entangled quantum states, via the violation of Bell inequalities, display correlations that defy any explanation satisfying classical notions of local causality.

Interestingly however, it is known that not all entangled quantum states lead to Bell inequality violation, showing that entanglement and Bell nonlocality are in general different resources. Whereas much effort has been devoted to studying the set of states that violate a Bell inequality, relatively little effort has been given to the converse question of characterising those states which do not. The central focus of Part I is thus an investigation into the sets of entangled states in quantum theory that do not allow for any Bell inequality violation. Our motivation is in part fundamental, since a better understanding of such sets allows for a better understanding of the relationship between the fields of entanglement and Bell nonlocality, and in part practical, since Bell nonlocality serves as the resource for so called device-independent protocols in quantum information processing.

After introducing Bell nonlocality (Chapter 2), our starting point will be the recently developed field of EPR steering (Chapter 3). We first present

two results relating specifically to this field, namely demonstration of asymmetric EPR steering (Chapter 4) and a criterion which serves as a necessary condition for EPR steering for two-qubit states (Chapter 5). We then exploit a connection to Bell nonlocality, allowing us to use the tools of EPR steering to answer fundamental questions relating to Bell nonlocality. This proves to be particularly fruitful, leading to a number of applications (Chapter 4) and allowing one to prove an in-equivalence between entanglement and nonlocality for quantum systems of arbitrary numbers of parties (Chapter 6). We then explore the classical resources required to simulate entanglement (Chapter 7), our main result being the first example of an entangled quantum state that can be simulated using classical resources of finite dimension.

Part II switches the focus to single quantum systems and the processing of these systems in communication networks. Here, quantum theory is known to provide significant advantages over classical information processing, for example through algorithms of quantum computing, communication complexity, cryptography and random number generation.

An important resource of interest for such tasks is dimension, i.e. Hilbert space dimension of the quantum systems used in the networks. We present methods to certify the dimension of quantum systems processed by networks of devices, tailored to the case where the devices share classical correlations (Chapter 7), and the case in which they are independent (Chapter 8). We also use these tests to compare the power of quantum vs classical communication, and show that quantum systems outperform classical systems of the same dimension under significant experimental noise. This leads to a quantum random number certification protocol that is extremely tolerant to noise (Chapter 8).

While the thesis should serve as a concise overview of all main results, the reader is directed towards the published papers for some additional results and detailed proofs.

Résumé

Cette thèse est consacrée à l'étude des concepts fondamentaux de la physique quantique d'un point de vue de l'information quantique. Les statistiques de mesures obtenues via des systèmes quantiques multipartites sont étudiées dans le cadre de la nonlocalité de Bell et du steering EPR (Partie I). Par ailleurs, nous discutons des concepts de dimension et d'aléa dans des réseaux quantiques (Partie II).

Un des aspects les plus étonnants de la mécanique quantique est son désaccord avec la notion classique de localité. Un état quantique typique ne peut pas être décrit comme le produit de sous-systèmes locaux, mais par un état global dont les sous-systèmes sont intriqués. Cette inséparabilité au niveau mathématique persiste même au niveau de l'observation. En effet, certains états intriqués, via la violation des inégalités de Bell, demontrent des fortes corrélations qui défient toute explication par une théorie locale.

Il existe pourtant des états quantiques intriqués qui ne violent aucune inégalité de Bell, ce qui montre que l'intrication et la nonlocalité sont en général des ressources différentes. Tandis que beaucoup d'effort a été dédié à l'étude des états qui violent des inégalités de Bell, relativement peu est connu à propos des états intriqués, dits locaux, qui ne violent aucune inégalité de Bell. Le focus de la première partie de la thèse est l'étude de ces états. Notre motivation est en partie fondamentale, car cet étude permet une meilleure compréhension de la relation entre les concepts de l'intrication et de la nonlocalité de Bell, et en partie pratique puisque la nonlocalité de Bell constitue une ressource pour les protocoles de l'information quantiques dits "device-independent".

Après avoir rappelé les concepts de la nonlocalité de Bell (Chapitre 2), notre point de départ sera le domaine récement développé du steering EPR (Chapitre 3). Nous présentons d'abord deux résultats relatifs à ce sujet. Le premier démontre que le steering EPR est fondamentalement asymétrique (Chapitre 4). Le second est un critère nécessaire pour montrer qu'un état de deux qubits est

une ressource pour le steering EPR (Chapitre 5). Ensuite, nous exploitons une connexion avec la nonlocalité de Bell, ce qui nous permet d'utiliser les outils du steering EPR pour répondre à des questions fondamentales de la nonlocalité de Bell. Ceci s'avère particulièrement fructueux, permettant plusièurs applications (Chapter 5) et une démonstration de l'inéquivalence entre l'intrication et la nonlocalité pour les systèmes quantiques d'un nombre arbitraire de sous-systèmes (Chapter 6). La Partie I est conclut en explorant les ressources classiques nécessaires pour la simulation de l'intrication (Chapitre 7). Notre résultat central est la démonstration de l'existence d'états intriqués admettant une simulation classique avec des ressources de dimension finie.

La second partie de la thèse se concentre sur les systèmes quantiques uniques dans les réseaux de communication. Dans ce scènario, les états quantiques offrent des avantages considérables sur leurs contreparties classiques, par exemple dans les domaines du calcul, de la cryptographie et de la géneration de nombres aléatoires.

Une ressource importante pour de telles tâches est la dimension, i.e. la dimension de l'espace de Hilbert des systèmes quantiques utilisés dans ces réseaux. Nous présentons des méthodes pour certifier la dimension d'états quantiques traités par des réseaux de communication. Ces méthodes sont adaptées au cas où les appareils partagent des correlations classiques (Chapitre 7), et au cas ou ils sont indépendants (Chapitre 8). On compare la puissance de la communication quantique avec la communication classique, démontrant ainsi que la performance des systèmes quantiques surpasse celle des systèmes classique de même dimension, même en présence d'un fort bruit experimental. Ces résultats mènent également un protocole pour la certification de nombres aléatoires quantiques qui est extrêmement robuste au bruit experimental (Chapitre 8).

Contents

Chapter 1

Overview

Part I: Outline

The early 20^{th} century was an exciting time for physics. Whereas Einstein was creating a revolution in our understanding of space, time and gravity, an independent quantum revolution was under way. The year 1900 saw the publication of Planck's work on the quantum nature of energy exchange [1], followed in 1905 by Einstein's proposal of the quantisation of radiation as an explanation of the photoelectric effect [2]. The work eventually resulted in the Heisenberg and Schrödinger formalisms of quantum mechanics [3, 4] marking a profound paradigm shift in physics. Quantum theory has since been unprecedented in its predictive power, providing breakthroughs spanning all areas of physics.

Despite its undeniable success at observational predictions, questions were being raised about the foundations of the theory. In their famous 1935 paper, Einstein, Podolsky and Rosen (EPR) [5] put forward their belief as to why quantum theory cannot be considered a complete theory of nature. Underpinning their argument was the notion of local causality, that operations performed in one local laboratory should not influence the state of a system at another distant laboratory. Under this assumption, the nonlocal nature of wavefunction collapse suggests that the theory is incomplete, in the sense that new additional variables should be added to render it consistent.

In the same year, Schrödinger [6] expanded on the ideas of EPR and showed that such a phenomenon arises from a property of certain quantum states known as entanglement: such states cannot be understood as the sum of their individual subsystems but rather by a global state which encompasses all subsystems. With the continuing success of quantum mechanics as a predictive tool however, the conclusion of EPR was mostly considered a philosophical

footnote of the theory and went largely ignored. It was not until 1964 that John Bell offered a conclusive reply [7], leading to the field of Bell nonlocality to which Part I of the thesis is dedicated and which we introduce in Chapter 2. Bell proved that *no* theory satisfying the notion of local causality can reproduce the correlations of separated measurements on certain quantum states. Consequently, orthodox pictures of the nature of physics in space and time as favoured by EPR had to be abandoned. Moreover, such claims soon became experimentally testable via the violation of so-called Bell inequalities [8,9] and have been recently demonstrated beyond all reasonable doubt via loophole free experiments [10–12].

The phenomenon of Bell nonlocality can be seen as a consequence of the existence of entangled states in quantum theory as introduced by Schrödinger [6], and for a time it was presumed that entanglement and nonlocality were two sides of the same coin. It was not until 1989 that the the precise relationship between the state dependent notion of entanglement and the possibility of nonlocality (i.e. Bell inequality violation) was investigated by Werner [13]. There, he showed the existence of entangled quantum states which do not lead to Bell inequality violation. Hence, although entanglement is necessary for nonlocality, in some cases it is not sufficient. 30 years on from Werner's result, a precise understanding of the relationship between entanglement and nonlocality is still missing. In Part I of the thesis we present progress in this direction, as outlined below.

Local hidden variable models
In order to prove that a particular quantum state does not violate any Bell inequality, one must construct a local hidden variable (LHV) model, which guarantees that all possible correlations from the state satisfy the notion of local causality. This is typically a difficult task since the LHV model must be guaranteed to work for all choices of measurements, and consequently relatively few states admitting LHV models are known. Much of Part I of this thesis is dedicated to studying the set of states admitting LHV models, and presenting new methods for their construction. This then allows us to answer fundamental questions about the relationship between entanglement and nonlocality.

EPR Steering as a tool for the study of Bell nonlocality
The phenomenon discussed by EPR and expanded on by Schrödinger has recently been formalised in quantum information theory under the name of EPR steering [14]. The previous years have seen a large volume of work focused on this subject, which we introduce in Chapter 3. Importantly for us, EPR

steering can be linked to Bell nonlocality. Fundamental questions about Bell nonlocality can then be tackled with the tools of EPR steering, which we exploit in later chapters.

In Chapter 4 we demonstrate asymmetric steering, that is, a quantum state for which EPR steering is possible in only one direction. This settles a question that was open since the beginning of the field [14]. Chapter 5 presents a sufficient criterion which guarantees that a two-qubit state does not lead to EPR steering. This in turn implies that such states do not exhibit Bell nonlocality and provides general classes of entangled quantum states admitting LHV models. We then present some applications of this result to the fields of Bell nonlocality and measurement incompatibility. Some further applications of this result are then presented. In Chapter 6 we consider multipartite systems. Here, we present a family of genuinely multipartite entangled states for any number of parties which admit LHV models, proving an in-equivalence of entanglement and nonlocality for any number of parties. Our construction is based upon ideas of EPR steering and uses states obtained from the criterion of Chapter 5.

LHV models using finite resources
Although it is known that certain entangled states admit LHV models, a common feature of all previously known models is that they require shared classical resources of infinite dimension, despite the corresponding simulated quantum state having finite Hilbert space dimension. In Chapter 8 we investigate whether this is necessarily the case of all LHV models. We prove the existence of entangled quantum states admitting LHV models that can be constructed using shared classical resources of finite dimension, and discuss the case of simulating nonlocal quantum states using classical communication.

Part II: Outline

The second half of the thesis is focused on the certification of dimension and randomness from quantum systems.

Certification of dimension
The concept of dimension (i.e. the number of independent degrees of freedom of a system) plays a vital role in many areas of physics and information theory. Given the fragile nature of quantum systems however, creating and controlling systems of high dimension becomes increasingly difficult, and experimentalists are constantly pushing to manipulate quantum systems of ever higher dimension. This has its motivation, for example, in the practical implementation of

quantum computing or communication protocols, where quantum systems of a given dimension are known to dramatically outperform their classical counterparts.

It is thus desirable to develop methods to certify the dimension of physical systems produced in the laboratory. Chapter 8 and Chapter 9 are focused on this problem. We work in the "device-independent" framework, whereby one does not assume precise experimental control over the devices and works only with the statistics of the experimental data. This is of practical significance since experimental implementations are typically subject to sources of error.

In Chapter 8 we outline the general framework, building on that of [15]. We consider networks of preparation, transformation and measurement devices which model the physical devices used in a laboratory experiment and develop tests of dimension for these networks. We also investigate the power of quantum vs classical systems in such networks and show that quantum systems can significantly outperform classical systems of the same dimension at certain tasks. In Chapter 9 we consider a related problem featuring a simple network of two devices, under the additional assumption that the devices act independently. We show that here, quantum systems can outperform classical systems, even under arbitrary large sources of background noise.

Certification of randomness

The generation of high quality random numbers is an important problem that finds application in cryptography, simulation and gambling. However, practically all commercial methods used to generate random numbers use classical pseudo-random functions, which are fundamentally deterministic, or noise from complex physical processes (such as thermal/electronic noise), which is impossible to fully characterise. This makes for a reliable estimate on the quality of such sources particularly troublesome. The fact that the outcomes of measurements on quantum systems are inherently probabilistic has sparked a field of research focused on exploiting quantum systems for the generation and certification of random numbers [16–20]. In this way, reliable bounds of the quality of a source of random numbers can be guaranteed.

In the final part of Chapter 9 we present a protocol that allows one to certify the quality of random numbers produced when qubit quantum systems are prepared and subsequently measured. The protocol is based on the tests of dimension presented in the same chapter. We work in the "semi-device-independent" scenario, where total experimental control over the quantum systems is not required. As with our tests of dimension, true randomness can be certified under large experimental noise and low detection efficiency, which has made possible a corresponding experimental realisation.

Part I

Bell nonlocality and the classical simulation of entangled quantum states

Chapter 2

Bell nonlocality

2.1 The Bell scenario

Consider two parties (which we will call Alice and Bob from here on in) at separate locations in space. A source is located between the two which provides them with some shared physical system (see Fig. 2.2). Alice receives a uniformly chosen classical variable $x \in \{0, \cdots, n-1\}$ (called her input) and must return another classical variable $a = \{0, \cdots, m-1\}$ (called her output). Similarly, Bob receives his input $y = \{0, \cdots, n-1\}$, outputting $b = \{0, \cdots, m-1\}$. The labels of the inputs/outputs do not carry any particular significance but simply refer abstractly to a number of possible measurements that could be performed on some shared system held by Alice and Bob, and their corresponding outcomes (for example polarisation measurements on a system of two photons). We make the additional assumption that Alice and Bob cannot communicate their given inputs to each other. This is somewhat natural given that they are separated spatially, but can be enforced, for example, by ensuring that the space-time event where Alice receives her input is space-like separated from the event where Bob gives his output and vice-versa.

This process is then repeated many times; in each round Alice and Bob receive their inputs and must give their outputs. In the limit of infinitely many rounds[1], we may thus define the conditional probability distributions $p(ab|xy)$, i.e. the probability that in a given round Alice and Bob give the outputs a, b upon receiving the inputs x, y. Throughout this thesis, we will generally refer to the set of these distributions as the *correlations* produced in a Bell scenario. As we will see, a striking feature of quantum theory is that it

[1]Note that we are making an assumption of independent, identically distributed rounds. For a discussion of memory effects in Bell tests, see for example [21].

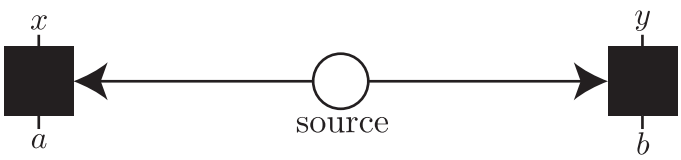

Figure 2.1: The Bell scenario. Two separated parties (called Alice and Bob) represented by the black boxes share a physical system provided by a source, then receive inputs x and y and provide outcomes a and b.

can produce correlations strictly stronger than those allowed by any classical theory.

2.1.1 Classical correlations

Our aim here is to characterise the most general correlations that can result from the laws of classical physics in a Bell scenario. Following Bell [7], this can be formalised as follows. Given her input x, Alice outputs a according to a probability distribution $p(a|x)$. Similarly, Bob outputs according to $p(b|y)$. Note that at this stage we have assumed a notion of *locality*: since Alice does not know Bob's input y, her output can depend only on her input x and so $p(a|xy) = p(a|x)$ (and likewise for Bob $p(b|xy) = p(b|y)$). We further assume that these distributions are deterministic, i.e. for a given input x, Alice will return some output a with probability 1. At this stage we therefore have

$$p(ab|xy) = p(a|x)p(b|y), \tag{2.1}$$

which will also be a deterministic function since it is a product of deterministic functions.

In a typical experiment, the observed correlations will neither be of the above form (i.e. product form) nor deterministic since there may be some unobserved or unknown classical variables leading to correlations and/or noise. This can be taken care of by introducing a shared classical random variable $\lambda \in \Lambda$ (which may take a potentially continuous range of values), which Alice and Bob can use to probabilistically mix deterministic strategies. Denoting the probability density of this random variable by q_λ so that $\int_\Lambda q_\lambda \mathrm{d}\lambda = 1$, we thus have

$$p(ab|xy) = \int_\Lambda q_\lambda p_\lambda(a|x)p_\lambda(b|y)\mathrm{d}\lambda, \tag{2.2}$$

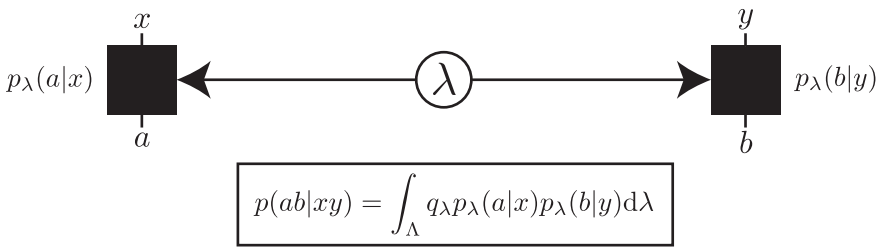

Figure 2.2: Representation of a LHV model which defines the set of classical correlations in a Bell scenario.

where each $p_\lambda(a|x)$, $p_\lambda(b|y)$ is a deterministic function[2]. If we have a finite number of inputs and outputs, we have a finite number of possible deterministic functions $p_\lambda(a|x)$, $p_\lambda(b|y)$ and so one can replace the above integral by a finite sum without loss of generality:

$$p(ab|xy) = \sum_\lambda q_\lambda p_\lambda(a|x) p_\lambda(b|y). \tag{2.3}$$

Here, λ is summed over all possible combinations of deterministic functions for Alice and Bob. For historical reasons, the shared variable λ is called a *local hidden variable* and (2.2) a *local hidden variable model* for the correlations (see Fig. 2.2 for a graphical representation of the above).

With a little thought, one can convince oneself that any correlations arising from the laws of classical physics (classical mechanics and special relativity, for example) must give rise to correlations of this form. More generally, any theory satisfying the above principle of locality will produce such correlations. Such correlations are hence termed *local*, and throughout this thesis we will use the words classical and local synonymously to describe correlations of the form (2.2).

2.1.2 Quantum correlations

We now wish to provide the same analysis for quantum theory (see Fig. 2.3). Alice and Bob's local measurements will be governed by measurement operators $M_{a|x} \geq 0$, $M_{b|y} \geq 0$ acting on the local Hilbert spaces of Alice and Bob and such that $\sum_a M_{a|x} = \sum_b M_{b|y} = \mathbb{1}$ (where the identity acts on the relevant

[2]Note that we could allow these functions to be non-deterministic. However, we can always hide any indeterminism by adding some classical random variables to λ that simulate this indeterminism (since any finite distribution $p(a|x)$ can be written as a convex mixture over the deterministic functions $p_\lambda(a|x)$)

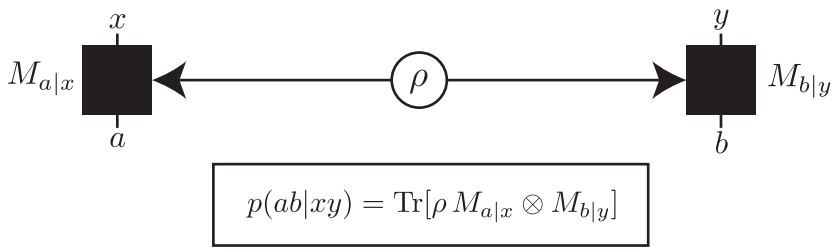

Figure 2.3: Bipartite correlations attainable in a Bell test using quantum theory.

Hilbert space). Since we have the full power of quantum mechanics, we replace the classical hidden variable λ by any valid bipartite quantum state ρ acting on the joint Hilbert space of Alice and Bob. This state, can, for example, be entangled between the Hilbert spaces of Alice and Bob. The correlations are then given via the Born rule as

$$p(ab|xy) = \mathrm{Tr}\left[M_{a|x} \otimes M_{b|y}\, \rho\right]. \tag{2.4}$$

It is a simple exercise to show[3] that any correlations of the form (2.2) can be written in the form (2.4). Whether the converse question is true, i.e. whether quantum correlations admit a local hidden variable model explanation and if so when, will be the general question of interest of this part of the thesis.

2.2 The CHSH inequality

We start by rewriting (2.3) as

$$p(ab|xy) = \sum_{\lambda} q_\lambda D_\lambda(ab|xy), \tag{2.5}$$

where $D_\lambda(ab|xy) = p_\lambda(a|x)p_\lambda(b|y)$ and $\sum_\lambda q_\lambda = 1$. Equation (2.5) can thus be seen as the convex hull of a finite number of possible deterministic distributions indexed by the variable λ. This allows for a rather elegant description of the set of classical correlations in terms of convex polytopes. As an illustrative example, we take the simplest nontrivial scenario where both Alice and Bob receive one of two possible inputs, giving one of two possible outcomes (that is, $n = m = 2$).

[3]for example, the classical hidden variable can be modelled as a diagonal (hence separable) quantum state of sufficiently high dimension $\rho = \sum_\lambda q_\lambda |\lambda\rangle\langle\lambda|$.

One possible deterministic strategy for Alice is to output $a = 0$ for $x = 0, 1$, that is $[p(0|0), p(0|1)] = [1, 1]$. Another is given by the case where she outputs $a = 1$ for $x = 0, 1$, so $[p(0|0), p(0|1)] = [0, 0]$. Finally we have the cases where she outputs $a = 0$ for one input and $a = 1$ for the other, corresponding to $[0, 1]$ and $[1, 0]$. Hence, we have 4 possible deterministic strategies for Alice's local function $p(a|x)$. For Bob we have the same and so we have a total of 4×4 deterministic functions $D_\lambda(ab|xy)$, $\lambda = 1, \cdots, 16$.

We now define the "correlation vector"

$$\mathbf{p} = [p(00|00), p(01|00), p(10|00), \cdots, p(11|11)] \tag{2.6}$$

whose elements are the values of the conditional probability distribution $p(ab|xy)$. Since we have $m = n = 2$ this vector lives in a space of dimension 16 and every possible set of correlations $p(ab|xy)$ corresponds to a point in this space. In fact, we can reduce the size of this space by exploiting some known linear dependencies between the $p(ab|xy)$. The first four of these are given by the normalisation of the probabilities, that is

$$\sum_{ab} p(ab|xy) = 1 \quad \forall x, y. \tag{2.7}$$

The second set of dependencies are given by *no-signalling*, reflecting the fact that Alice cannot communicate her input to Bob (and vice-versa). Mathematically, this means that Bob's (Alice's) marginal distribution should be independent of Alice's (Bob's) choice of input, giving us

$$\sum_{a} p(ab|xy) = p(b|y) \quad \forall y, b \tag{2.8}$$

$$\sum_{b} p(ab|xy) = p(a|x) \quad \forall x, a. \tag{2.9}$$

This gives us another 4 dependencies (we can consider e.g. $a = 0, b = 0$ due to normalisation). We may therefore reduce the size of our space by 8 so that $\mathbf{p}$ now lives in a space of dimension 8.

The set of classical correlations is then the convex hull of the 16 deterministic vectors $\mathbf{p}_\lambda$, which geometrically forms a convex polytope in dimension 8 (see Fig.
2.4 for a graphical representation of this). This polytope is completely characterised by a finite set of inequalities that are linear in the probabilities, which define the facets of the polytope. For our simple case, one finds a number of facets corresponding to the normalisation of probabilities and a single facet

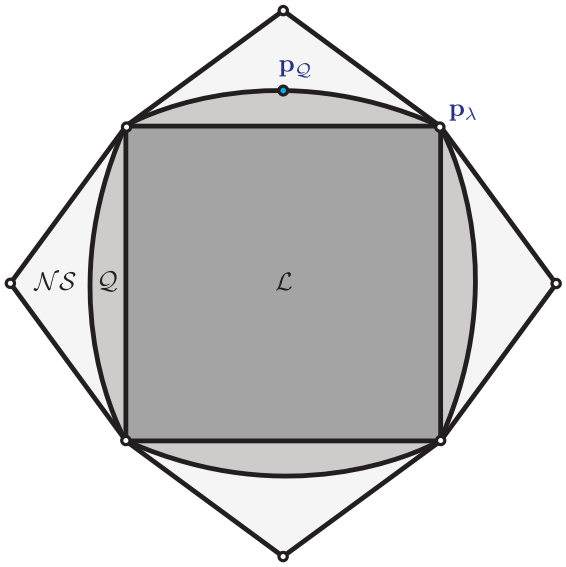

Figure 2.4: Graphical (two dimensional) representation of the set of Bell local, quantum and non-signalling correlations $\mathbf{p}$ for the two inputs, two outputs. The vertices of the central square represent the deterministic classical correlations. The set of Bell local correlations, $\mathcal{L}$, is the convex hull of these vertices (dark grey square). Note that the full space is actually 8-dimensional and we have more deterministic vertices than shown. The facets of this polytope are called Bell inequalities. Outside of this lies the set of correlations achievable using quantum systems, $\mathcal{Q}$. Finally, the set of non-signalling correlations $\mathcal{NS}$ (those which respect (2.8)) is also a polytope lying outside of the quantum set. The point $\mathbf{p}_\mathcal{Q}$ corresponds to the maximal quantum violation of the CHSH inequality.

inequality (up to relabelling of input/outputs) given by

$$\mathcal{I}[p(ab|xy)] = E_{00} + E_{01} + E_{10} - E_{11} \leq 2, \tag{2.10}$$

where $E_{ij} = \sum_{ab}(-1)^{a+b}p(ab|xy)$. This is known as the CHSH Bell inequality [8], and is hence satisfied by all correlations of the form (2.2). Since (2.10) is linear in the probabilities, one may alternatively write

$$\mathbf{p} \cdot \boldsymbol{\mathcal{I}} \leq 2 \tag{2.11}$$

for some real vector $\boldsymbol{\mathcal{I}}$ of the same dimension of $\mathbf{p}$.

For the case with n inputs and m outputs the procedure is much the same, and we arrive at a local polytope $\mathbb{P}$ (of much higher dimension) whose non-

normalisation facets we call *Bell inequalities*[4]. A set of distributions $p(ab|xy)$ given by a correlation vector $\mathbf{p}$ is then Bell local iff it belongs to the polytope $\mathbb{P}$ and hence satisfies all facet Bell inequalities $\boldsymbol{\mathcal{I}}^{(n)}$ (indexed by n):

$$\mathbf{p} \in \mathbb{P} \iff \mathbf{p} \cdot \boldsymbol{\mathcal{I}}^{(n)} \leq C^{(n)} \quad \forall n. \tag{2.12}$$

Computational methods [22, 23] can be used for finding such facets, however quickly become infeasible so that only simple scenarios can be solved (in general the problem is known to be NP-hard in the number of inputs [24]).

2.3 Quantum nonlocality

We have now done all the necessary work to lead us to Bell's famous theorem [7]:

Theorem 2.3.1 (Bell's Theorem)**.** There exist correlations $p(ab|xy)$ arising from local measurements on bipartite quantum states that cannot be written in the form (2.2).

To see this, consider the (maximally entangled) pure state

$$\rho = |\psi^-\rangle\langle\psi^-|, \tag{2.13}$$

where

$$|\psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle). \tag{2.14}$$

Parameterise measurements using the Bloch vector notation:

$$M_{0|j} = \frac{\mathbb{1} + \vec{m}_j \cdot \vec{\sigma}}{2}\ ; \quad M_{1|j} = \mathbb{1} - M_{0|j}, \tag{2.15}$$

where $\vec{m}_j$ is a normalised Bloch vector and $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ is the vector of Pauli matricies. The state (2.14) has the interesting property that $E_{xy} = -\vec{a}_x \cdot \vec{b}_y$ where $\vec{a}_x$, $\vec{b}_y$ are the Bloch vectors of Alice and Bob's measurements. If we chose Alice's measurement Bloch vectors to be in the z and x directions, $a_1 = \hat{e}_z$, $\vec{a}_2 = \hat{e}_x$, and Bob's measurement Bloch vectors to be $\vec{b}_y = (\hat{e}_z \pm \hat{e}_x)/\sqrt{2}$, we find

$$\mathcal{I}[p(ab|xy)] = 2\sqrt{2}, \tag{2.16}$$

[4] Facets of the local polytope $\mathbb{P}$ are generally called tight Bell inequalities. More generally, any function $f(p(ab|xy))$ such that $f(p(ab|xy) > L$ for some L implies that the correlations lie outside of $\mathbb{P}$ (and are hence nonlocal) is called a Bell inequality

hence violating the CHSH Bell inequality. For obvious reasons, such correlations are termed *nonlocal*. It hence follows that no classical theory can reproduce the correlations of quantum mechanics.

Experimental observation of the violation of the CHSH Bell inequality was made soon after its discovery [25] using entangled photons pairs, and eventually enforcing space like separation between the measurement events [26]. More recently, a number of experimental groups have achieved a so called "loophole free" violation of the CHSH inequality [10–12], confirming that nature is indeed nonlocal.

2.4 Not all entangled states are nonlocal: Werner's model

Since separable states always give rise to local correlations, the violation of a Bell inequality can be seen as a certification of entanglement. A relevant question is then: does entanglement *always* lead to nonlocality? That is, given any entangled state, can one always find local measurements such that (2.2) is not satisfied?

It turns out that the answer to this question is no, as first shown by Werner [13] in 1989. Consider the one parameter family of two qubit mixed states (now called Werner states):

$$\rho_\alpha = \alpha|\psi^-\rangle\langle\psi^-| + (1-\alpha)\mathbb{1}/4. \tag{2.17}$$

These states are entangled for $\alpha > 1/3$ (as verified, for example, via the PPT criterion [27]). For $\alpha > 1\sqrt{2}$ the state violates the CHSH Bell inequality. Werner showed that for $\alpha \leq 1/2$, all the correlations obtained from ρ_α using projective measurements satisfy (2.2) and thus never lead to Bell inequality violation.

In order to prove this, we need to exhibit an explicit local hidden variable (LHV) model for the state, i.e. some hidden variable λ with distribution q_λ and local distributions $p_\lambda(a|\vec{x})$, $p_\lambda(b|\vec{y})$ for the measurements $M_{a|\vec{x}}$, $M_{b|\vec{y}}$ of Alice/Bob with Bloch vectors $\vec{x}$, $\vec{y}$ such that

$$\mathrm{Tr}\left[\rho_{\frac{1}{2}} M_{a|\vec{x}} \otimes M_{b|\vec{y}}\right] = \int_\Lambda q_\lambda \, p_\lambda(a|\vec{x}) p_\lambda(b|\vec{y}) \mathrm{d}\lambda \tag{2.18}$$

(the case $\alpha < 1/2$ is straightforward since it corresponds to adding white noise). This can be achieved by choosing λ to be a unit vector on the 3-sphere, which we write as $\lambda = \vec{\lambda}$, distributed uniformly $q_{\vec{\lambda}} = \frac{1}{4\pi}$. Intuitively,

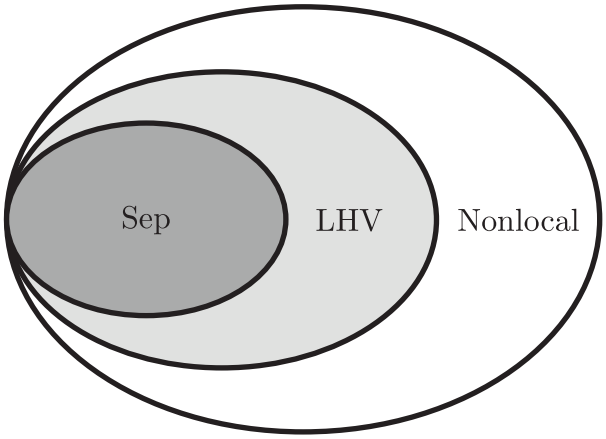

Figure 2.5: Representation of the sets of separable, local (states admitting LHV models), and nonlocal quantum states.

$\vec{\lambda}$ can be thought to represent a qubit pure state with Bloch vector $\vec{\lambda}$. The distributions $p_{\vec{\lambda}}(a|\vec{x})$ and $p_{\vec{\lambda}}(b|\vec{y})$ are given by

$$p_{\vec{\lambda}}(a|\vec{x}) = \frac{1-(-1)^a \operatorname{sgn}[\vec{x}\cdot\vec{\lambda}]}{2}\ ; \quad p_{\vec{\lambda}}(b|\vec{y}) = \frac{1+(-1)^b\, \vec{y}\cdot\vec{\lambda}}{2}. \tag{2.19}$$

Hence, the state ρ_α does not violate any Bell inequality for $\alpha \leq \frac{1}{2}$ if arbitrary projective measurements are performed. One may further wonder whether (2.2) may be violated if one is allowed to perform general measurements (positive-operator valued measures, or POVMs) on the same state. Barrett [28] proved that if $\alpha \leq 5/12$ this is not the case, and we hence have a class of entangled state that admits a LHV model even for general measurements.

Since there exist entangled quantum states admitting LHV models, the set of states admitting a LHV model is a superset the set of separable states (see Fig. 2.5). However, relatively few such states are known [29–32]. Much of this half of the thesis will be dedicated to the study of the LHV set, i.e. finding entangled quantum states admitting LHV models, with an aim of better understanding the relationship between entanglement and nonlocality. In order to do this we will use tools from the newly developed field of EPR steering, which we introduce in the following chapter.

Chapter 3

EPR steering

At the heart of the EPR argument [5] lies the fact that one party, by making measurements on one half of an entangled state, can non-locally influence the reduced state of the other party's subsystem. This concept was proposed by Schrödinger [6] and first explored in continuous variable systems [33, 34]. Recently, the phenomenon has been formalised in the quantum information setting [14] (we focus on a more recent, although essentially equivalent approach, see [35] for a review), which we briefly outline here.

Again we have two parties, Alice and Bob, that share some bipartite quantum state ρ. Alice performs a measurement on her half of the state given by the measurement operators $M_{a|x}$. We then define the resulting subnormalised reduced state on Bob's subsystem

$$\sigma_{a|x} \equiv \mathrm{Tr}_A \left[M_{a|x} \otimes \mathbb{1} \, \rho \right], \tag{3.1}$$

where we have $p(a|x) = \mathrm{Tr}\, \sigma_{a|x}$. The set of matrices $\{\sigma_{a|x}\}_a$ for fixed x is called a *measurement ensemble* and the set of all measurement ensembles an *assemblage*. Notice that since $\sum_a M_{a|x} = \mathbb{1}$, $\sum_a \sigma_{a|x} = \mathrm{Tr}_A[\rho] = \rho_B$ for all x. Formally, we say that an assemblage demonstrates EPR steering from Alice to Bob (shortened to steering from Alice to Bob from here on in) if it does not admit a decomposition of the form

$$\sigma_{a|x} = \int_\Lambda q_\lambda p_\lambda(a|x) \sigma_\lambda \mathrm{d}\lambda, \quad \forall a, x \tag{3.2}$$

where q_λ is again a normalised distribution over $\lambda \in \Lambda$, $p_\lambda(a|x)$ is an arbitrary probability distribution, and σ_λ is a quantum state (called a local hidden state) acting on the Hilbert space of Bob. The right hand side of the above can be understood as follows (see Fig. 3.1). A source between Alice and Bob chooses

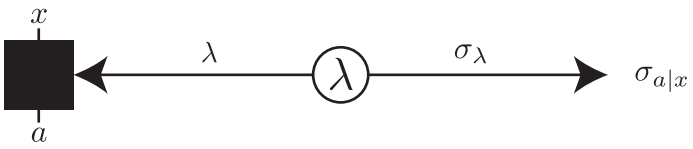

Figure 3.1: Representation of (3.2). The source prepares the classical variable λ, and sends this to Alice. To Bob, the local hidden state σ_λ is sent. If Alice outputs with probability $p_\lambda(a|x)$ Bob receives the state $\sigma_{a|x}/\operatorname{Tr}\sigma_{a|x}$ (after averaging over λ) with probability $\operatorname{Tr}\sigma_{a|x}$.

the local hidden variable λ with probability density q_λ, and sends to Bob the local hidden state σ_λ and to Alice the variable λ. Alice, upon receiving her input x, outputs a with probability $p_\lambda(a|x)$. The corresponding state held by Bob is then $\sigma_{a|x}/\operatorname{Tr}[\sigma_{a|x}]$ occurring probability $p(a|x) = \operatorname{Tr}\sigma_{a|x}$. To simplify things slightly, one can further absorb the q_λ into the σ_λ, defining $\tilde{\sigma}_\lambda = q_\lambda\sigma_\lambda$. Equivalently to (3.2), one then has

$$\sigma_{a|x} = \int_\Lambda p_\lambda(a|x)\tilde{\sigma}_\lambda \mathrm{d}\lambda \quad \forall a, x \ , \tag{3.3}$$

with $\tilde{\sigma}_\lambda \geq 0$ and $\int \operatorname{Tr}\tilde{\sigma}_\lambda = 1$.

Since all separable state lead to assemblages of the form (3.3), it follows that if the assemblage $\{\sigma_{a|x}\}$ demonstrates steering then the state ρ must be entangled. One also sees that Alice and Bob play asymmetric roles. Generally, one refers to Alice as the untrusted party and Bob the trusted party. The reason for this is that we do not assume anything about the distributions $p_\lambda(a|x)$, but we assume we know the reduced states $\sigma_{a|x}/\operatorname{Tr}\sigma_{a|x}$ of Bob via tomography. Contrasting this with Bell nonlocality, for example, both parties are untrusted. In the following chapter we will explore this asymmetry in more detail.

3.1 EPR steering and LHV models

If all assemblages that can be produced from a state ρ admit a decomposition of the form (3.2) (i.e. for all possible measurements $M_{a|x}$) then the state ρ is called unsteerable from Alice to Bob. Every state which is unsteerable admits a LHV model, which can be seen as follows. Define $p_\lambda(b|y) = \operatorname{Tr}[\sigma_\lambda M_{b|y}]$. This

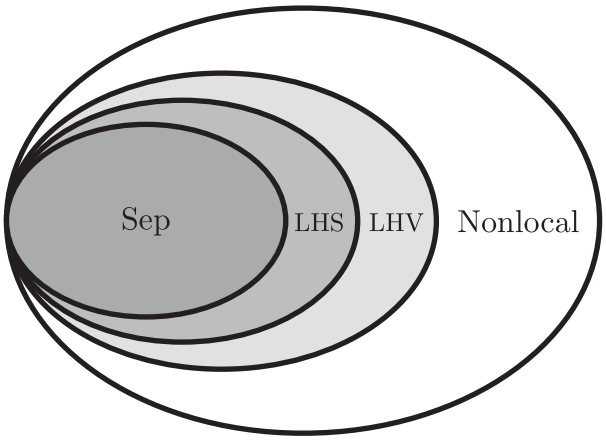

Figure 3.2: Representation of the sets of separable, unsteerable (LHS), Bell local (LHV) and nonlocal states. The set of unsteerable states is a subset of the set of local states.

then defines a LHV model

$$\begin{aligned} p(ab|xy) &= \int_\Lambda q_\lambda p_\lambda(a|x) \operatorname{Tr}\left[\sigma_\lambda M_{b|y}\right] \mathrm{d}\lambda \\ &= \operatorname{Tr}\left[\left(\int_\Lambda q_\lambda p_\lambda(a|x)\sigma_\lambda \mathrm{d}\lambda\right) M_{b|y}\right] \\ &= \operatorname{Tr}\left[\sigma_{a|x} M_{b|y}\right] = \operatorname{Tr}\left[\operatorname{Tr}_A\left[M_{a|x}\otimes \mathbb{1}\,\rho\right] M_{b|y}\right] \\ &= \operatorname{Tr}\left[M_{a|x}\otimes M_{b|y}\,\rho\right], \end{aligned} \tag{3.4}$$

hence reproducing the correlations of the state ρ. Consequently, unsteeable states occupy a middle ground between separable and nonlocal states (see figure (3.2)). This allows one to study fundamental questions related to Bell nonlocality whilst working in the steering framework. As we will see in this and later chapters, this can be a particularly fruitful approach.

The first line of (3.4):

$$p(ab|xy) = \int_\Lambda q_\lambda p_\lambda(a|x) \operatorname{Tr}\left[\sigma_\lambda M_{b|y}\right] \mathrm{d}\lambda \tag{3.5}$$

is called a *local hidden state* (LHS) model, since it is a special case of a LHV model where Bob's response function takes the form $p_\lambda(b|y) = \operatorname{Tr}[\sigma_\lambda M_{b|y}]$, with σ_λ called the local hidden state (in comparison to a local hidden variable). The above can be seen as the definition of steering for correlations: a set of correlations $p(ab|xy)$ demonstrates steering from Alice to Bob iff they do not admit a decomposition of the form (3.5).

3.2 EPR steering and one-sided maps

The fact that one of the parties in the steering scenario is trusted allows one to define a useful lemma (see also [36, 37]):

Lemma 3.2.1. Let Ω be a positive linear map, and let $\{\sigma_{a|x}\}$ be an assemblage that does not demonstrate steering from Alice to Bob (i.e. that admits a decomposition (3.3)). Then the assemblage $\{\sigma^{\Omega}_{a|x}\}$ given by

$$\sigma^{\Omega}_{a|x} = \frac{\Omega[\sigma_{a|x}]}{\mathrm{Tr}\left[\Omega[\rho_B]\right]} \tag{3.6}$$

where $\rho_B = \sum_a \sigma_{a|x}$ does also not demonstrate steering from Alice to Bob. Furthermore, if Ω is invertible and its inverse map positive, then $\{\sigma_{a|x}\}$ demonstrates steering form Alice to Bob if and only if $\{\sigma^{\Omega}_{a|x}\}$ demonstrates steering from Alice to Bob.

Proof. Using (3.3) we find

$$\frac{\Omega[\sigma_{a|x}]}{\mathrm{Tr}\left[\Omega[\rho_B]\right]} = \int_\Lambda p_\lambda(a|x)\frac{\Omega[\tilde{\sigma}_\lambda]}{\mathrm{Tr}\left[\Omega[\rho_B]\right]}\mathrm{d}\lambda = \int_\Lambda p_\lambda(a|x)\tilde{\sigma}^{\Omega}_\lambda \mathrm{d}\lambda. \tag{3.7}$$

Note that $\tilde{\sigma}^{\Omega}_\lambda \geq 0$ and

$$\int_\Lambda \mathrm{Tr}\left[\tilde{\sigma}^{\Omega}_\lambda\right]\mathrm{d}\lambda = \frac{\mathrm{Tr}\left[\Omega[\int_\Lambda \tilde{\sigma}_\lambda \mathrm{d}\lambda]\right]}{\mathrm{Tr}\left[\Omega[\rho_B]\right]} = 1 \tag{3.8}$$

since by summing (3.3) over a we have

$$\int_\Lambda \tilde{\sigma}_\lambda \mathrm{d}\lambda = \rho_B. \tag{3.9}$$

If the map Ω is invertible then one has

$$\sigma_{a|x} = \frac{\Omega^{-1}\left[\sigma^{\Omega}_{a|x}\right]}{\mathrm{Tr}\left[\Omega^{-1}\left[\sigma^{\Omega}_{a|x}\right]\right]}. \tag{3.10}$$

Hence, if Ω^{-1} positive then $\{\sigma_{a|x}\}$ demonstrates steering if and only if $\{\sigma^{\Omega}_{a|x}\}$ demonstrates steering. □

From this we can prove the following Theorem, which will be of use throughout this part of the thesis.

Theorem 3.2.1. Let Ω be a positive linear map. If the state ρ admits a LHS model from Alice to Bob, then the state

$$\rho_\Omega = \frac{\mathbb{1}_A \otimes \Omega[\rho]}{\mathrm{Tr}\left[\Omega[\rho_B]\right]} \tag{3.11}$$

admits a LHS model from Alice to Bob. Furthermore, if Ω is invertible and its inverse map positive then the statement is if and only if.

Proof. The proof of this follows from the fact that

$$\mathrm{Tr}_A\left[M_{a|x} \otimes \mathbb{1}\ \rho_\Omega\right] = \frac{\Omega\left[\sigma_{a|x}\right]}{\mathrm{Tr}\left[\Omega[\rho_B]\right]}. \tag{3.12}$$

which by Lemma 3.2.1 does not demomstrate steering for all $M_{a|x}$ and hence admits a LHS model. To prove the if and only if condition, one uses the fact that

$$\rho = \frac{\mathbb{1}_A \otimes \Omega^{-1}\left[\rho_\Omega\right]}{\mathrm{Tr}\left[\Omega^{-1}\left[\rho_B^\Omega\right]\right]} \tag{3.13}$$

with $\rho_B^\Omega = \mathrm{Tr}_A\, \rho_\Omega$. □

3.3 Steering as a semi-definite program

Many problems in steering can be tackled using the tools of semi-definite programming (SDP) optimisation [38]. For example, for a finite number of inputs and outputs, deciding if (3.3) is satisfied can be cast as a semi-definite program feasibility problem [39, 40]. This follows from that fact that for a fixed number of inputs and outputs, one may take the functions $p_\lambda(a|x)$ to be the set of all deterministic functions (as was the case in Bell nonlocality). We must then check

$$\sigma_{a|x} = \sum_\lambda p_\lambda(a|x)\tilde{\sigma}_\lambda \mathrm{d}\lambda \ \ \forall a, x\ ; \quad \sum_\lambda \mathrm{Tr}[\tilde{\sigma}_\lambda] = 1\ ; \quad \tilde{\sigma}_\lambda \geq 0\ \ \forall \lambda. \tag{3.14}$$

The first two of these conditions are linear in the optimisation variables $\tilde{\sigma}_\lambda$ and the final a semi-definite constraint, hence the above can be solved with a semi-definite program. SDP methods for steering go far beyond this simple example - see [35] for a review.

Chapter 4

One-way EPR steering

Alice and Bob play asymmetric roles in the steering scenario. A natural question is whether there exist states that are steerable uniquely in one direction [14]. That is, does there exist a state ρ that admits a LHS model from Bob to Alice, however does not admit a LHS model from Alice to Bob. This has fundamental interest, since it would show an inherent asymmetry of quantum steering for quantum states, but also of potential practical interest, since there are many quantum information tasks in which one party is trusted (yourself, for example), but not the other. The existence of one-way steering was proven for the set of gaussian measurements on a family of continuous variable states [41,42] (with corresponding experimental observation [43]). Considering non-gaussian measurements however, the result was not known to hold, and indeed for Bell nonlocality it is known that for certain states non-guassian measurements can be necessary to violate a Bell inequality [44]. Considering more general measurements on quantum states the question thus remained open, mainly due to the difficulty in constructing LHS models for all measurements for (necessarily asymmetric) quantum states.

We answer the question of one-way EPR steering **[Paper A]**, considering arbitrary projective measurements on the two qubit state given by:

$$\rho(\alpha) = \alpha|\psi^-\rangle\langle\psi^-| + \frac{1-\alpha}{5}\left(2\,|0\rangle\langle 0| \otimes \frac{\mathbb{1}}{2} + 3\frac{\mathbb{1}}{2} \otimes |1\rangle\langle 1|\right). \tag{4.1}$$

This state is entangled in the range $\alpha > 1/19(-6+5\sqrt{6}) \approx 0.3288$ (as verified using the positive partial transpose criterion [27]). For $\alpha \leq 1/2$, $\rho(\alpha)$ we show that $\rho(\alpha)$ admits a LHS model from Bob to Alice, however for $\alpha \gtrsim 0.4983$ no model can exist from Alice to Bob (i.e. the state is steerable from $A \to B$). Hence in the range $0.4983 \leq \alpha \leq 1/2$ the state is one-way steerable from Alice to Bob. Below we show this in two parts, first focusing on the LHS model

from Bob to Alice, then the demonstration of steering from Alice to Bob. In the following chapter (Section 5.3.1), we present the simplest possible example of one-way steering based on a different class of two-qubit states.

4.1 No steering from Bob to Alice

We wish to provide a LHS model for $\rho(\alpha)$ for $\alpha = 1/2$. We therefore need to define the hidden states σ_λ, distribution q_λ and response functions $p_\lambda(b|\vec{x})$ such that

$$p(ab|xy) = \mathrm{Tr}\left[\rho(1/2) M_{a|\vec{x}} \otimes M_{b|\vec{y}}\right] \tag{4.2}$$

$$= \int_\Lambda q_\lambda \,\mathrm{Tr}\left[\sigma_\lambda M_{a|\vec{x}}\right] p_\lambda(b|y)\mathrm{d}\lambda, \tag{4.3}$$

(see Eq. 3.5) where we take $a = \pm 1$, $b = \pm 1$. For projective measurements with Bloch vectors $\vec{x}$, $\vec{y}$, the correlations of this state are completely defined by

$$\langle ab\rangle = -\frac{1}{2}\vec{x}\cdot\vec{y}\,; \qquad \langle a\rangle = \frac{x_3}{5}\,; \qquad \langle b\rangle = -\frac{3y_3}{10}\,, \tag{4.4}$$

where $\langle ab\rangle = \sum_{ab} ab\, p(ab|\vec{x}\vec{y})$ is the expectation value of the product of outcomes ab (and similarly for $\langle a\rangle$, $\langle b\rangle$), and x_3, y_3 are the z-components of the Bloch vectors $\vec{x}$, $\vec{y}$.

To simulate these statistics, we chose the LHS model

$$\sigma_{\vec{\lambda}} = |\vec{\lambda}\rangle\langle\vec{\lambda}|\,; \quad q_{\vec{\lambda}} = \frac{1+\lambda_3}{4\pi}\,; \quad b = -\lambda_0 \mathrm{sgn}(\vec{y}\cdot\vec{\lambda}), \tag{4.5}$$

where $\vec{\lambda} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ is the Bloch vector for the hidden state and $\lambda_0 = \pm 1$ is a local variable for Bob. For now we assume $\lambda = 1$ always. Unlike before, the shared variable is not uniformly distributed on the sphere, but is weighted towards to top of the sphere due to the term λ_3. Working in spherical co-ordinates one has $\lambda_3 = \cos\theta$ and

$$\langle ab\rangle = -\int_\phi\int_\theta \frac{1+\cos\theta}{4\pi}\,\mathrm{Tr}\left[A_{\vec{x}}|\vec{\lambda}\rangle\langle\vec{\lambda}|\right]\mathrm{sgn}[\vec{y}\cdot\vec{\lambda}]\sin\theta\mathrm{d}\theta\mathrm{d}\phi \tag{4.6}$$

$$= -\int_\phi\int_\theta \frac{1+\cos\theta}{4\pi}\,\vec{x}\cdot\vec{\lambda}\,\mathrm{sgn}[\vec{y}\cdot\vec{\lambda}]\sin\theta\mathrm{d}\theta\mathrm{d}\phi \tag{4.7}$$

$$= -\int_\phi\int_\theta \frac{1}{4\pi}\,\vec{x}\cdot\vec{\lambda}\,\mathrm{sgn}[\vec{y}\cdot\vec{\lambda}]\sin\theta\mathrm{d}\theta\mathrm{d}\phi$$

$$- \int_\phi\int_\theta \frac{\cos\theta}{4\pi}\,\vec{x}\cdot\vec{\lambda}\,\mathrm{sgn}[\vec{y}\cdot\vec{\lambda}]\sin\theta\mathrm{d}\theta\mathrm{d}\phi. \tag{4.8}$$

Here $A_{\vec{x}} = M_{+1|\vec{x}} - M_{-1|\vec{x}}$ is Alice's observable. The second of these two terms is equal to zero. This can easily be seen since if the vector $\vec{\lambda}$ has angle θ, the flipped vector $-\vec{\lambda}$ has angle $\pi - \theta$. Under this transformation $\theta \to \pi - \theta$ one has $\cos\theta \to -\cos\theta$, $\sin\theta \to \sin\theta$, $\vec{x}\cdot\vec{\lambda} \to -\vec{x}\cdot\vec{\lambda}$ and $\mathrm{sgn}[\vec{y}\cdot\vec{\lambda}] \to -\mathrm{sgn}[\vec{y}\cdot\vec{\lambda}]$. It follows that the contribution to the second term of $\langle ab\rangle$ coming from $\vec{\lambda}$ is precisely equal and opposite than the contribution from $\vec{\lambda}$, and the integral vanishes. We are thus left with

$$\langle ab\rangle = -\int_\phi \int_\theta \frac{1}{4\pi}\, \vec{x}\cdot\vec{\lambda}\,\mathrm{sgn}[\vec{y}\cdot\vec{\lambda}]\sin\theta \mathrm{d}\theta\mathrm{d}\phi \tag{4.9}$$

$$= -\frac{1}{2}\vec{x}\cdot\vec{y}, \tag{4.10}$$

which follows from the fact that the above is precisely the same correlation predicted by Werner's model (see (2.19)) with the uniform distribution $q_{\vec{\lambda}} = 1/4\pi$.

The marginal term $\langle a\rangle$ for Alice is a straightforward integration of trigonometric functions

$$\langle a\rangle = \int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} \sin\theta \mathrm{d}\theta\, \frac{1+\cos\theta}{4\pi}\, \vec{x}\cdot\vec{\lambda} \tag{4.11}$$

$$= \frac{x_3}{3}. \tag{4.12}$$

To calculate Bob's marginal $\langle b\rangle$ is a little more involved. Since the function $\mathrm{sgn}(\vec{y}\cdot\vec{\lambda})$ is symmetric about $\vec{\lambda} = \vec{y}$, it is useful to rotate into the frame $(\tilde{\theta}, \tilde{\phi})$ in which the $\tilde{\theta} = 0$ direction is aligned with $\vec{y}$. This will involve rotating the original frame by an angle θ_B, where θ_B is the angle of $\vec{y}$ to the (original) z direction. In this frame, the distribution $q_{\vec{\lambda}}$ is given by

$$q(\tilde{\theta}, \tilde{\phi}) = \frac{1 - \sin\theta_B \sin\phi \sin\tilde{\theta} + \cos\theta_B \cos\tilde{\theta}}{4\pi}. \tag{4.13}$$

This frame has the advantage that $\mathrm{sgn}(\vec{y}\cdot\vec{\lambda}) = +1$ for $0 \le \tilde{\theta} \le \pi/2$ and $\mathrm{sgn}(\vec{y}\cdot\vec{\lambda}) = -1$ for $\pi/2 < \tilde{\theta} \le \pi$. We may thus split the integral in two so that

$$\langle b\rangle = -\int_0^{2\pi} \mathrm{d}\tilde{\phi} \left[\int_0^{\pi/2} q(\tilde{\theta},\tilde{\phi})\sin\tilde{\theta}\mathrm{d}\tilde{\theta} - \int_{\pi/2}^{\pi} q(\tilde{\theta},\tilde{\phi})\sin\tilde{\theta}\mathrm{d}\tilde{\theta}\right]. \tag{4.14}$$

This involves only integration over trigonometric functions and gives

$$\langle b\rangle = -\frac{y_3}{2}. \tag{4.15}$$

Hence the above model produces the correct joint correlations $\langle ab\rangle$ but not the marginal statistics $\langle a\rangle$. This can be remedied by using the variable λ_0 (which we previously set to 1). This is due to the fact that the marginal statistics predicted by the model are in fact too strong. If we chose a distribution $p(\lambda_0 = -1) = f$ then one finds

$$\langle a\rangle = (1-2f)\frac{x_3}{3}\ ; \quad \langle b\rangle = (1-2f)\frac{y_3}{2}. \tag{4.16}$$

Thus, if we chose $f = \frac{1}{5}$ we obtain the correct statistics for the state. Note that the joint statistics $\langle ab\rangle$ do not depend on on f.

4.2 Steering from Alice to Bob

The next step involves showing that no LHS model can exist from Alice to Bob. This can be achieved using semi-definite programming (SDP) techniques. We perform a fixed number m of measurements given by the observables M_i for Alice, and the three Pauli observables σ_j, $j = 1, 2, 3$ for Bob. We further define $\sigma_0 = \mathbb{1}$. We then have

$$\langle ab\rangle = \mathrm{Tr}\left[\rho(\alpha) M_i \otimes \sigma_j\right] \quad j > 0\ ; \quad \langle b\rangle = \mathrm{Tr}\left[\rho(\alpha)\mathbb{1} \otimes \sigma_j\right]. \tag{4.17}$$

One has $\langle a\rangle_i = \langle ab\rangle_{i0}$. We may then use use the following SDP [39] (which is essentially the same SDP as Section 3.3) for deciding if a given assemblage is steerable or not.

$$\begin{aligned} &\alpha^* = \max \alpha \\ &\text{s.t.} \sum_\lambda E_\lambda(i)\,\mathrm{Tr}\left[\rho_\lambda \sigma_j\right] = \langle ab\rangle_{ij}, \quad \sum_\lambda \mathrm{Tr}\left[\rho_\lambda \sigma_j\right] = \langle b\rangle_j \qquad (4.18) \\ &\quad \sum_\lambda \mathrm{Tr}\left[\sigma_\lambda\right] = 1\ ; \quad \sigma_\lambda \geq 0 \quad \forall \sigma_\lambda. \qquad (4.19) \end{aligned}$$

Here $E_\lambda(i) = \pm 1$ represent the deterministic response functions of Alice (given now as correlators), and the summation over λ is over all such deterministic strategies for her inputs. The value of α^* gives upper bounds to the critical value at which the state demonstrates steering for a given set of measurements. In table 4.2, upper bounds for α^* obtained via SDP optimisation are presented. For $m > 14$ one sees that the state is steerable for $\alpha \geq 0.4983$. This therefore proves that in the range $0.4983 \leq \alpha < 0.5$, the state is one-way steerable. This result can be made analytic via the construction of a steering inequality which guarantees that the assemblage demonstrates steering (See Paper A Appendix).

m	2	3	4	5	6	7	8
α^*	0.6951	0.5661	0.5424	0.5302	0.5156	0.5120	0.5088
m	9	10	11	12	13	14	
α^*	0.5037	0.5030	0.5014	0.5005	0.4993	0.4983	

Table 4.1: Threshold values α^* for which the state $\rho(\alpha)$ is steerable from Alice to Bob. The optimisation is conducted over all possible steering tests where Alice performs $m = 2, \ldots, 14$ projective measurements.

4.2.1 A stronger example

One can in fact obtain a stronger example of one-way steering using essentially the same LHS model as above (although this was no realised at the time of publication). If one fixes $\lambda_0 = 1$, the correlations (4.9), (4.11), (4.15) simulate the correlations of the state (written in the Pauli basis)

$$\rho = \frac{1}{4}\left[\mathbb{1} + \frac{1}{3}\sigma_z \otimes \mathbb{1} - \frac{1}{2}\mathbb{1} \otimes \sigma_z - \frac{1}{2}\sum_i \sigma_i \otimes \sigma_i\right]. \tag{4.20}$$

This state has stronger asymmetry than (4.1) and steering from Alice to Bob can be more easily demonstrated using the same SDP techniques (six measurements for Alice are enough here).

4.3 Outlook

This work has inspired a number of subsequent works. Notably, one way steering has been extended to the scenario in which the LHS model for the unsteerable direction is valid even if general measurements (POVMs) are performed [36]. There have also been experimental realisations of one-way steering for finite dimensional quantum states [45, 46].

Given the asymmetric nature of one-way steering, one could imagine applications in asymetric quantum information processing tasks. This is at the moment somewhat missing and would be an interesting avenue of research. Finally, EPR steering has recently been formalised in the multipartite setting [47]. A natural extension to this work would be to explore assymetric steering in multipartite states. For example, does there exist a tripartite state (with parties ABC), such that A can steer to B, B to C, C to A, but not in the reverse direction?

In Section 5.3.1 of the following chapter, we present the simplest possible

demonstration of one-way EPR steering, based on a different set of unsteerable states.

Chapter 5

Sufficient condition for unsteerability

Whereas it is relatively easy to verify if a state is steerable (e.g. via SDP methods for a fixed number of measurements [39]), it is generally much harder to guarantee that a state is unsteerable. This is due to the fact that one must verify that (3.3) is satisfied for all possible measurements. For this reason, relatively few states admitting LHS [13, 29, 48], or more general LHV models are known [31, 49, 50]. In **[Paper B]**, we present a simple sufficient criterion for a state of two-qubits to be unsteerable for arbitrary projective measurements, hence implying that ρ admits a LHS model for projective measurements. This provides one with the possibility of finding new families of states admitting LHS (and hence LHV) models. One important concept needed to prove this criterion is that of canonical states, as described in the following section. We then apply our criterion to a family of two-qubit states, which leads to a number of applications in both Bell nonlocality and measurement incompatibility.

5.1 Canonical states for steering

Consider a mixed entangled state ρ of two qubits, with reduced state for Bob $\rho_B = \mathrm{Tr}_A[\rho]$. Define the (state dependent) positive linear map

$$\Omega_\rho[\rho] = \rho_B^{-\frac{1}{2}}\ \rho\ \rho_B^{-\frac{1}{2}}. \tag{5.1}$$

Note that Ω is invertible since ρ_B is full rank. It follows from Theorem 3.2.1 that the state

$$\rho' = \frac{\mathbb{1} \otimes \rho_B^{-\frac{1}{2}} \, \rho \, \mathbb{1} \otimes \rho_B^{-\frac{1}{2}}}{\mathrm{Tr}\left[\mathbb{1} \otimes \rho_B^{-\frac{1}{2}} \, \rho \, \mathbb{1} \otimes \rho_B^{-\frac{1}{2}}\right]} \tag{5.2}$$

is unsteerable from Alice to Bob if and only if ρ is unsteerable from Alice to Bob. Hence, if we aim to prove the unsteerability of a two-qubit mixed state[1] ρ, we may focus our attention on the corresponding canonical state ρ' defined above. Since by construction we have $\mathrm{Tr}_A[\rho'] = \rho'_B = \mathbb{1}/2$ it follows that

$$\rho' = \frac{1}{4}\left(\mathbb{1} + \vec{a} \cdot \vec{\sigma} \otimes \mathbb{1} + \sum_i T_i \sigma_i \otimes \sigma_i\right) \tag{5.3}$$

where $\vec{a}$ is Alice's local Bloch vector and $T = \mathrm{diag}(T_1, T_2, T_3)$ is the diagonal correlation matrix of the state (which can always be made diagonal by local unitaries [51]).

5.2 Criterion for unsteerability

The criterion is as follows:

Theorem 5.2.1. Let ρ be a two-qubit state with corresponding canonical form ρ' given by (5.3). If

$$\max_{\vec{x}} \left[\, (\vec{a} \cdot \vec{x})^2 + 2 \, ||T\vec{x}|| \, \right] \leq 1, \tag{5.4}$$

where $\vec{x}$ is a normalized vector and $||\cdot||$ the euclidean vector norm, then ρ is unsteerable from Alice to Bob, considering arbitrary projective measurements.

Proof. We first characterize the assemblage resulting from projective measurements on a state in the canonical form ρ'. Alice's measurement is given by a Bloch vector $\vec{x}$ and output $a = \pm 1$, corresponding to operators $M_{\pm|\vec{x}} = (\mathbb{1} + \vec{x} \cdot \vec{\sigma})/2$. For $a = +1$, the steered state in Bloch vector notation is (see for example [48])

$$\sigma_{+|\vec{x}} = \mathrm{Tr}_A(M_{+|\vec{x}} \otimes \mathbb{1}\rho') = \frac{1}{4}[(1 + \vec{a} \cdot \vec{x})\mathbb{1} + T\vec{x} \cdot \vec{\sigma}]. \tag{5.5}$$

[1] The case of pure states is irrelevant here since all entangled pure states are steerable [39]

Notice that the above state is diagonal in the basis $\{|\vec{s}\rangle, |-\vec{s}\rangle\}$, with Bloch vector $\vec{s} = \frac{T\vec{x}}{||T\vec{x}||}$; we omit the $\vec{x}$ dependence to ease notation. The eigenvalues of $\sigma_{+|\vec{x}}$ are

$$\alpha(\vec{x}) = \frac{1}{4}(1 + \vec{a}\cdot\vec{x} + ||T\vec{x}||),\ \beta(\vec{x}) = \frac{1}{4}(1 + \vec{a}\cdot\vec{x} - ||T\vec{x}||). \tag{5.6}$$

Note that by construction $\alpha(\vec{x}) \geq \beta(\vec{x})$. One can then find $\sigma_{-1|\vec{x}}$ from $\sigma_{+|\vec{x}}$ by using $\sigma_{-1|\vec{x}} = \rho_B - \sigma_{+1|\vec{x}}$.

To prove unsteerability of ρ we may focus on the unsteerability of the canonical state ρ'. We therefore construct a LHS for the assemblage given by (5.5), which can be seen as a generalisation of Werner's model. Notice that Werner's model (Eq. 2.19) can be seen as a local hidden state model with local hidden states given by the pure states

$$\sigma_{\vec{\lambda}} = |\vec{\lambda}\rangle\langle\vec{\lambda}| \tag{5.7}$$

with a uniform distribution

$$q_{\vec{\lambda}} = \frac{1}{4\pi}. \tag{5.8}$$

This ensures that

$$\int q_{\vec{\lambda}}\, \sigma_{\vec{\lambda}}\, \mathrm{d}^2\vec{\lambda} = \frac{\mathbb{1}}{2} = \rho_B \tag{5.9}$$

as required by summing (3.2) over a. Our canonical states ρ' are defined precisely so that $\mathrm{Tr}_A[\rho] = \rho'_B = \mathbb{1}/2$. We also therefore take the the same hidden states and (uniform) distribution as Werner. However, we modify the response function $p_{\vec{\lambda}}(a = \pm|\vec{x})$ such that

$$p_{\vec{\lambda}}(\pm|\vec{x}) = \frac{1 \pm \mathrm{sgn}(\vec{s}(\vec{x})\cdot\vec{\lambda} - c(\vec{x}))}{2}, \tag{5.10}$$

where $\vec{s}$ given by $\vec{s} = T\vec{x}/||T\vec{x}||$ defines the direction of the steered state for a given $\vec{x}$ and $c(\vec{x}) \in [-1, 1]$ (to be chosen). The intuition (see also Fig. 5.1) is that with this choice of $\vec{s}$, the steered state in our LHS model (after averaging over $\vec{\lambda}$) will have Bloch vector in the direction $\vec{s}$, as required from (5.5). The number $c(\vec{x})$ can be used to make the marginal distribution $p(a|\vec{x})$ for Alice non-uniform without modifying the direction of the Bloch vector of the steered state (e.g. for $c(\vec{x}) = -1$ we have $p(a|\vec{x}) = 1$). For each measurement $\vec{x}$, we thus need to find $\vec{s}(\vec{x})$ and $c(\vec{x})$ such that the LHS model reproduces the steered state (5.5).

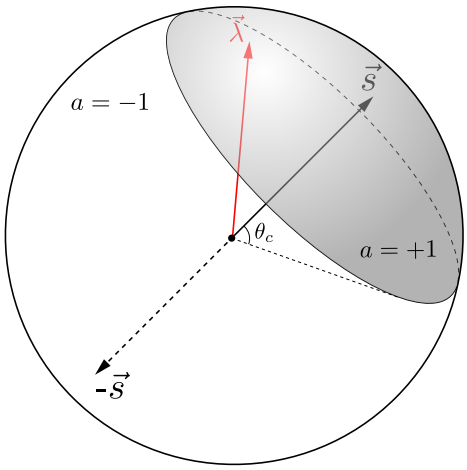

Figure 5.1: Illustration of Alice's response function (5.10) in our LHS model. If $\mathrm{sgn}(\vec{s}(\vec{x}) \cdot \vec{\lambda} - c(\vec{x})) \geq 0$ then $a = +1$ (shaded spherical cap, with angle $\theta_c = \arccos[c]$), otherwise $a = -1$. The steered state then corresponds to the average (sub-normalized) density matrix obtained by integrating pure qubit states $|\vec{\lambda}\rangle$ over the shaded region.

To do this, note that we need only concentrate on the case $a = +1$; the case $a = -1$ is automatically satisfied from $\sigma_{+1|\vec{x}} + \sigma_{-1|\vec{x}} = \rho_B$ and (5.9). We now calculate the assemblage predicted by this model, given by

$$\sigma^{\mathrm{LHS}}_{+1|\vec{x}} = \int \sigma_{\vec{\lambda}}\, p_{\vec{\lambda}}(+|\vec{x}) d\vec{\lambda} = \frac{1}{4\pi} \int |\vec{\lambda}\rangle\langle\vec{\lambda}| p_{\vec{\lambda}}(+|\vec{x}) d\vec{\lambda}. \tag{5.11}$$

We parameterise the state $|\vec{\lambda}\rangle$ using the Bloch decomposition in the basis $\{|\vec{s}\rangle, |-\vec{s}\rangle\}$:

$$|\vec{\lambda}\rangle = |\vec{\lambda}(\theta, \phi)\rangle = \cos\frac{\theta}{2}|\vec{s}\rangle + \sin\frac{\theta}{2} e^{i\phi}|-\vec{s}\rangle. \tag{5.12}$$

Working in this basis and integrating over the spherical cap of Fig. 5.1 for which $a = +1$, we have

$$\sigma^{\mathrm{LHS}}_{+1|\vec{x}} = \int_0^{2\pi} \int_0^{\theta_c} \begin{pmatrix} \cos^2\frac{\theta}{2} & \cos\frac{\theta}{2}\sin\frac{\theta}{2} e^{-i\phi} \\ \cos\frac{\theta}{2}\sin\frac{\theta}{2} e^{i\phi} & \sin^2\frac{\theta}{2} \end{pmatrix} \frac{\sin\theta \mathrm{d}\phi \mathrm{d}\theta}{4\pi},$$

where $\theta_c = \arccos[c(\vec{x})]$ is the angle of the spherical cap. Since $\int_0^{2\pi} e^{i\phi} \mathrm{d}\phi = 0$, the off-diagonal components will be zero, and $\sigma^{\mathrm{LHS}}_{+1|\vec{x}}$ is therefore diagonal in

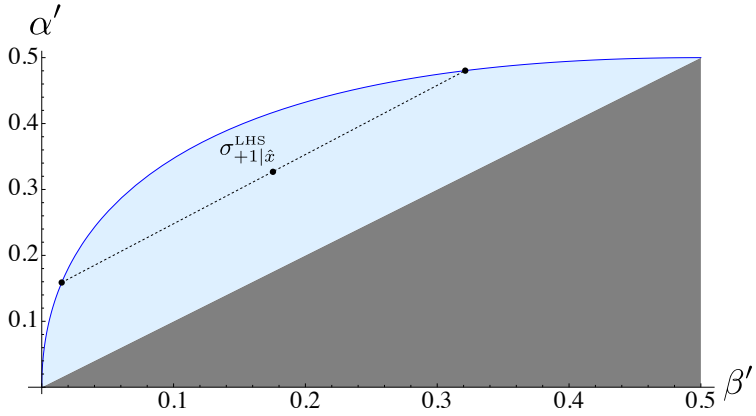

Figure 5.2: Plot of the achievable range of eigenvalues (α', β') in our LHS model (for a fixed direction $\vec{s}$). The upper blue curve corresponds to the condition $\alpha' = \sqrt{2\beta'} - \beta'$ and is achieved by the response functions (5.10); any point in the light blue area below may be achieved by taking a suitable convex combination of these functions (e.g. dashed line). Since we have $\alpha' \geq \beta'$, the grey area is not of interest.

the $\{|\vec{s}\rangle, |-\vec{s}\rangle\}$ basis, as desired. From this, the eigenvalues of $\sigma^{\rm LHS}_{+1|\vec{x}}$, $\alpha'(\vec{x})$ and $\beta'(\vec{x})$, are given by

$$\alpha'(\vec{x}) + \beta'(\vec{x}) = \frac{1}{2}\int_0^{\theta_c} \sin\theta \mathrm{d}\theta = \frac{1-\cos\theta_c}{2}; \tag{5.13}$$

$$\alpha'(\vec{x}) - \beta'(\vec{x}) = \frac{1}{2}\int_0^{\theta_c} \cos\theta \sin\theta \mathrm{d}\theta = \frac{1-\cos^2\theta_c}{4}. \tag{5.14}$$

Upon using $\theta_c = \arccos[c(\vec{x})]$ one then finds

$$\alpha'(\vec{x}) + \beta'(\vec{x}) = \frac{1}{2}(1 - c(\vec{x})); \tag{5.15}$$

$$\alpha'(\vec{x}) - \beta'(\vec{x}) = \frac{1}{4}(1 - c^2(\vec{x})), \tag{5.16}$$

from which we get the eigenvalues as a function of $c(\vec{x})$ as

$$\alpha'(\vec{x}) = \sqrt{2\beta'(\vec{x})} - \beta'(\vec{x})\,;\ \beta'(\vec{x}) = \frac{1}{8}(1 - c(\vec{x}))^2, \tag{5.17}$$

corresponding to the curve of Fig. 5.2. Since this curve is concave, by fixing $\vec{s}$ and taking convex combinations of the response functions (5.10) with different

$c(\vec{x})$, we may prepare any steered states corresponding to (α', β') below this curve, leading finally to

$$\beta'(\vec{x}) \leq \alpha'(\vec{x}) \leq \sqrt{2\beta'(\vec{x})} - \beta'(\vec{x}). \tag{5.18}$$

This corresponds to the blue area in Fig. 5.2. We thus conclude that the model reproduces the assemblage of any canonical state ρ, as long as its eigenvalues satisfy the above relation, i.e. $\alpha(\vec{x}) \leq \sqrt{2\beta(\vec{x})} - \beta(\vec{x})$, for any measurement vector $\vec{x}$, or equivalently

$$\max_{\vec{x}} \left[(\alpha(\vec{x}) + \beta(\vec{x}))^2 - 2\beta(\vec{x}) \right] \leq 0. \tag{5.19}$$

Using (5.6) to convert this maximisation into Bloch vector notation we arrive at Theorem 5.2.1.

□

One natural question is whether our criterion is both necessary and sufficient, i.e. whether it completely characterises the set of unsteerable two-qubit states. Using the criterion alone, we know this to not be the case. For example take the classically correlated state $\rho = \frac{1}{2}[|00\rangle\langle 00| + |11\rangle\langle 11|]$ and choose $\vec{x} = (0, 0, 1)$. One easily sees that our criterion is violated, however the state is separable and thus clearly unsteerable. However, it may be possible to incorporate other ideas in order to make the condition necessary and sufficient. For example, one could consider the set of states given by

$$\rho = p\rho_{\text{US}} + (1 - p)\rho_{\text{SEP}}, \tag{5.20}$$

where ρ_{US} is a state satisfying our criterion, ρ_{SEP} a separable state and $0 \leq p \leq 1$. Since mixing an unsteerable state with a separable state does not change its steerability, this also defines a (larger) set of unsteerable states which could potentially include all unsteerable two-qubit states. Whether this is the case is still unknown.

5.2.1 A useful example

The criterion (5.4) can be used to prove analytically the unsteerability of families of two qubit states for projective measurements. Here, we give one such example which will prove to be useful in later chapters of the thesis. Our states of interest are the two parameter family of "generalised Werner states":

$$\rho(p, \chi) = p|\psi_\chi\rangle\langle\psi_\chi| + (1 - p)\rho_\chi^A \otimes \frac{\mathbb{1}}{2}, \tag{5.21}$$

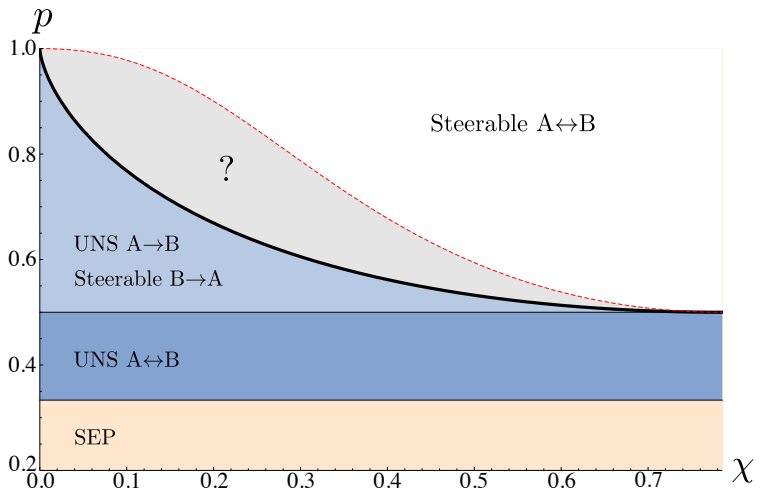

Figure 5.3: Characterisation of entanglement and steering for states $\rho(p, \chi)$. The solid black curve corresponds to (5.22), obtained from our unsteerability criterion. The state is separable in the light orange region, unsteerable (in both directions) in the dark blue region, unsteerable only from Alice to Bob (hence one-way steerable) in the light blue region, and two-way steerable in the white region (obtained from equation 19 of [48]). What happens in the grey region is an interesting open question.

with $p \in [0, 1]$ and $\chi \in [0, \pi/4]$. Note that for $\chi = \pi/4$ we recover the two-qubit Werner states. Using our criterion, we are able to prove unsteerability form Alice to Bob in the range

$$\cos^2(2\chi) \geq \frac{2p-1}{(1-p)p^3}, \tag{5.22}$$

see the thick black curve of Fig. 5.22 for a graph of this curve. For a proof of the above see Appendix A.

5.3 Applications

The state (5.21) has a number of interesting applications. This is due in part to the fact that for very small χ, the state is highly asymmetric yet remains entangled. We now present a number of these applications.

5.3.1 Simplest one-way steering

The states (5.21) in fact provide another example of one-way steering for two-qubit states. From these we can construct the simplest example possible of one-way steering, that is, a two-qubit state that is one-way steerable, where only two measurement settings are needed to prove steering in one direction. Note that by applying the filter $F_\chi = \mathrm{diag}(1/\cos\chi, 1/\sin\chi)$ we obtain

$$\frac{1}{2}F_\chi \otimes \mathbb{1}\rho(p,\chi)F_\chi \otimes \mathbb{1} = \rho(p,\pi/4), \tag{5.23}$$

which is a Werner state with visibility p. Since this state is steerable (in both directions) for $p > 1/2$ [14], it follows from Theorem 3.2.1 and the fact that F_χ is invertible that the state $\rho(p,\chi)$ is steerable from Bob to Alice for $p > 1/2$ for $\chi > 0$. Hence, for the parameter range $\chi > 0$, $1/2 < p < p^*(\chi)$, where $p^*(\chi)$ is determined by (5.4), the states $\rho(p,\chi)$ are one-way steerable (for projective measurements) from Bob to Alice only (see Fig. 5.22).

Furthermore, the Werner state violates the CHSH Bell inequality for $p > 1/\sqrt{2}$. Since Bell nonlocality is strictly stronger than EPR steering, any Bell nonlocal state is steerable. The state $\rho(p,\pi/4)$ is therefore steerable (in both directions) for only two measurement settings for $p > 1/\sqrt{2}$. Again by using F_χ one has from Theorem 3.2.1 that $\rho(p,\chi)$ is steerable for two measurement settings from Bob to Alice for $\chi > 0$. This can also be extended to the case where the LHS model for the state is valid for POVM measurements by considering the states

$$\rho_{\mathrm{POVM}}(p,\chi) = \frac{1}{2}\rho(p,\chi) + \frac{1}{2}|0\rangle\langle 0| \otimes \rho_B \tag{5.24}$$

where one can show (see Paper B Section IVA) one-way steering for two measurements settings in the range $p > 0.83353$ and corresponding χ given by (5.22).

5.3.2 Sufficient condition for joint measurability

The above criterion also finds application in quantum measurement theory. This follows from the direct connection existing between steering and the notion of joint measurability of a set of quantum measurements [52,53], which has already found applications, see e.g. [54]. A collection of projective measurements are said to be compatible if their corresponding observables commute (i.e. they share the same eigenvectors). For POVM measurements, this can be generalised to a notion of joint measurability. A set of measurements $\{M_{a|x}\}$

is said to be jointly measurable [55] if there exists a joint POVM $\{G_\lambda\}$ with outcomes λ and probability distributions $p(a|x,\lambda)$, from which the statistics of any of the measurements $\{M_{a|x}\}$ can be recovered by a suitable post processing, that is

$$M_{a|x} = \int G_\lambda p(a|x,\lambda)\mathrm{d}\lambda \qquad \forall\, a, x. \tag{5.25}$$

It has been shown that a set of POVMs is non-jointly measurable if and only if it can be used to demonstrate steering on some bipartite entangled state [52, 53]. In this sense steering and measurement incompatible are very closely related.

From this, one can convert our criterion for unsteerability into a criterion for joint measurability of a (potentially continuous) set of POVM measurements. Let $\{M_{\pm|x}\}$ be a set of dichotomic qubit POVMs with

$$M_{+|x} = \frac{1}{2}\left(k_x \mathbb{1} + \vec{m}_x \cdot \vec{\sigma}\right) \tag{5.26}$$

with $||\vec{m}_x|| \leq k_x \leq 2 - ||\vec{m}_x||$, and $M_{-|x} = \mathbb{1} - M_{+|x}$. Then the set $\{M_{\pm|x}\}$ is jointly measurable if

$$k_x(k_x - 2) + 2||\vec{m}_x|| \leq 0 \tag{5.27}$$

for all x. This can be seen as follows. A set of measurements $\{M_{\pm|x}\}$ is jointly measurable if and only if the assemblage given by $\sigma_{\pm|x} = \rho^{\frac{1}{2}} M_{\pm|x} \rho^{\frac{1}{2}}$, where ρ is a full-rank quantum state [56], is unsteerable. Choosing $\rho = \mathbb{1}/2$ we get the corresponding assemblage $\sigma_{\pm|x} = \rho^{1/2}\, M_{\pm|x}\, \rho^{1/2} = \frac{1}{2} M_{\pm|x}$. Following Theorem 1, condition (5.27) ensures the unsteerability of $\sigma_{\pm|x}$, and consequently the joint measurability of $\{M_{\pm|x}\}$. As with our criterion for unsteerability, it is not known if the above leads to a necessary and sufficient condition for joint measurability.

5.3.3 LHV model for incompatible measurements

In the bipartite scenario, if one party has a set of jointly measurable POVMs, it follows that the statistics must be local [53]. Like entanglement, then, non-joint measurability for both parties is a necessary requirement for nonlocality. Using our criterion, one can show that there exist non-jointly measurable sets of measurements for Alice that do not lead to nonlocal correlations, considering arbitrary dihcotomic measurements for Bob and an arbitrary shared entangled

state. Hence, for arbitrary dihcotomic measurements for Bob, non-joint measurability of measurements is not sufficient for nonlocality. This includes the case, for example, where the parties share an arbitrary two-qubit entangled state and Bob performs arbitrary projective measurements. This can be contrasted with the existence of entangled quantum states that admit LHV models and thus do not lead to nonlocality. Specifically, we consider the (continuous) set of noisy projective measurements

$$M^{\eta}_{\pm|\vec{y}} = \frac{1}{2}\left(1 + \eta\, \vec{y}\cdot\vec{\sigma}\right). \tag{5.28}$$

This set of measurements is jointly measurable iff $\eta \leq 1/2$ [54]. Note that

$$\mathrm{Tr}\left[M_{\pm|\vec{x}} \otimes M^{\eta}_{\pm|\vec{y}}\, |\psi_\chi\rangle\langle\psi_\chi|\right] = \mathrm{Tr}\left[M_{\pm|\vec{x}} \otimes M_{\pm|\vec{y}}\, \rho(\eta,\chi)\right], \tag{5.29}$$

where $\rho(\eta,\chi)$ is given by (5.21). One can thus recast the problem of the LHV simulation of the set of noisy projective measurements (5.28) and arbitrary projective measurements for Alice on any bipartite state as finding a LHV model for the state (5.21) for all χ for a given $p = \eta$. This eventually leads to a proof [**Paper C**] that the set of measurements with $\eta < 0.503$ (or $\eta < 0.515$ if SDP methods are used) admits a LHV model for any state where Bob performs arbitrary dihcotomic measurements. Extending this to POVM measurements for Bob is currently a work in progress and would lead to a proof that non-joint measurability does not imply Bell inequality violation in general.

5.3.4 No hidden nonlocality for qubits

A bipartite state is said to have hidden nonlocality if ρ admits a LHV model however the state

$$\rho_F = \frac{F_A \otimes F_B \rho F_A^\dagger \otimes F_B^\dagger}{\mathrm{Tr}[F_A \otimes F_B \rho F_A^\dagger \otimes F_B^\dagger]} \tag{5.30}$$

does not admit a LHV model, hence violates some Bell inequality. The matrices F_A, F_B are called local filters. The filtering operation, for example, corresponds to making the local measurements $\{F_A, \mathbb{1} - F_A\}$, $\{F_B, \mathbb{1} - F_B\}$ on the state ρ and postselecting the first outcomes. In the context of sequential measurement scenarios, hidden nonlocality gives a sufficient condition for the state not to admit a so-called sequential LHV model (see [57]). Hidden nonlocality was first shown to exist by Popescu [58], where the state before the filtering admits a LHV model for projective measurements. We then extended

this result to the case in which the LHV model before filtering is valid for general POVM measurements [**Paper D**]. One may wonder whether all entangled states could have their nonlocality revealed in such a way, i.e. for any entangled state, can one always find local filters that reveal the nonlocality of the state? For the case of projective measurements, the criterion (5.4) can be used to show that this is not the case. Specifically, consider the two-qubit Werner states

$$\rho_p = p|\psi^-\rangle\langle\psi^-| + (1-p)\frac{\mathbb{1}}{4}. \tag{5.31}$$

We first consider a filter on Alice's subsystem:

$$\rho_{F_A} = \frac{F_A \otimes \mathbb{1}\, \rho_p\, F_A^\dagger \otimes \mathbb{1}}{\mathrm{Tr}\left[F_A \otimes \mathbb{1}\, \rho_p\, F_A^\dagger \otimes \mathbb{1}\right]}. \tag{5.32}$$

Due to the $U \otimes U$ symmetry of the Werner state, it is sufficient to restrict to diagonal filters $F_A = \mathrm{diag}(\cos\chi, \sin\chi)$ for $\chi \in [0, \pi/4]$. One then finds

$$\rho_{F_A} = p|\psi_\chi\rangle\langle\psi_\chi| + (1-p)\rho_\chi^A \otimes \frac{\mathbb{1}}{2} = \rho(p,\chi). \tag{5.33}$$

Hence, this state is precisely the state (5.21). From the condition (5.22), one sees that this state admits a LHS model from Alice to Bob for all χ if $p \leq 1/2$. From Theorem (3.2.1) we may therefore apply any filter to Bob's subsystem and still have a LHS model for the filtered state. From this it follows that the filtered state

$$\rho_F = \frac{F_A \otimes F_B\, \rho_p\, F_A^\dagger \otimes F_B^\dagger}{\mathrm{Tr}\left[F_A \otimes F_B\, \rho_p\, F_A^\dagger \otimes F_B^\dagger\right]} \tag{5.34}$$

admits a LHS model from Alice to Bob for projective measurements (and thus a LHV model) for all possible filters for $p \leq 1/2$. This state therefore does not violate any Bell inequality for arbitrary projective measurements. This result can thus be seen as a generalisation of Werner's original result, extended to include hidden nonlocality. The case where one demands that the filtered state remains local even for POVM measurements can be tackled using a combination of the above and SDP techniques (work in progress).

5.4 Outlook

Related work has also discussed the LHS simulation of two-qubit states, notably for Bell diagonal states [48] and more recently general two-quibt states

[59]. One common property of all these works (including this chapter) is that only steering under projective measurements is considered. Naturally, one would therefore like to extend our criterion and others to include POVM measurements. This seems somewhat challenging given the added complexity of the set of POVM measurements. An alternative possibility to proving a similar result would be to look at whether POVM measurements are useful for steering. Evidence suggests that POVMs are generally not useful for steering, that is, a given state is steerable if and only if it is steerable using projective measurements. Proving this would therefore immediately generalise any result holding for projective measurements. Similarly, one would like to find criteria which extend beyond the qubit-qubit case. It is unlikely that our criterion can be straightforwardly extended to higher dimensional systems since it uses as a central tool the Bloch sphere representation which is applicable only to qubit systems.

Chapter 6

LHV simulation of multipartite entangled quantum states

So far our discussion of nonlocality has been limited to bipartite states. Considering multipartite states, the relationship between entanglement and nonlocality is far less explored. One important question here is whether there exist multipartite entangled states that admit LHV models, as we have seen is the case in the bipartite scenario. To make the question nontrivial it is necessary to consider states that have genuine multipartite entanglement, the strongest form of multipartite entanglement. This question has been explored for the case of tripartite states, where it was shown that such states exist [60], and another work has explored the link between genuine multipartite entanglement and genuine multipartite nonlocality [61]. We present progress in this direction [**Paper E**] by using tools from EPR steering to show the existence of a family of genuinely multipartite entangled states that admit LHV models, for any number of parties. Thus, the strongest form of multipartite entanglement does not lead to any nonlocality if one considers non-sequential local measurements. Considering sequential measurements however, we will see that our construction allows us to prove the existence of multipartite hidden nonlocality. This therefore highlights the potential importance of performing sequential measurements to reveal the nonlocality of entangled quantum states.

6.1 Multipartite LHV models

Consider a state ρ of N parties, where party i can make measurements labelled x_i obtaining outcomes a_i, specified by the measurement operators $M_{a_i|x_i}$, with $M_{a_i|x_i} \geq 0$ and $\sum_{a_i} M_{a_i|x_i} = \mathbb{1}$. The probability to see the outputs $\mathbf{a} =$

$(a_1, \cdots, a_N)$ given the inputs $\mathbf{x} = (x_1, \cdots, x_N)$ is given by

$$p(\mathbf{a}|\,\mathbf{x}) = \mathrm{Tr}\left[\rho\,(\otimes_{i=1}^{N} M_{a_i|x_i})\right]. \tag{6.1}$$

The state ρ is called (fully) local if, for all possible measurement operators $M_{a_i|x_i}$, the statistics $p(\mathbf{a}|\,\mathbf{x})$ can be reproduced by a local hidden variable (LHV) model:

$$p(\mathbf{a}|\mathbf{x}) = \int_{\Lambda} q_\lambda p_\lambda(a_1|x_1) p_\lambda(a_2|x_2) \cdots p_\lambda(a_N|x_N) \mathrm{d}\lambda, \tag{6.2}$$

where q_λ is a probability density over the shared variable $\lambda \in \Lambda$ and the $p_\lambda(a_i|x_i)$'s are probability distributions, called local response functions. Likewise, if the above cannot be satisfied then the state is said to be nonlocal, as witnessed by the violation of (some) Bell inequality.

One may also consider a weaker notion of locality, whereby the correlations are not demanded to be local with respect to all parties (as in (6.2)), but instead to be (mixtures of) correlations that are each local across some bipartition. Denoting by $(b, \bar{b}) \in \mathcal{B}$ a bipartition of the parties, these correlations take the form

$$p(\mathbf{a}|\mathbf{x}) = \sum_{(b,\bar{b})\in\mathcal{B}} p_b \int_{\Lambda_b} q_\lambda^b p_\lambda(\mathbf{a}_b|\mathbf{x}_b) p_\lambda(\mathbf{a}_{\bar{b}}|\mathbf{x}_{\bar{b}}) \mathrm{d}\lambda, \tag{6.3}$$

where $\mathbf{a}_b$, $\mathbf{x}_b$ denote the inputs and outputs for the bipartition b, p_b is a probability distribution and q_λ^b is a probability density over $\lambda \in \Lambda_b$ for each b. Note that (6.2) implies (6.3) but not necessarily the converse. Correlations which cannot be written in the above form are called *genuinely multipartite nonlocal* (GMNL) and represent the strongest form of multipartite nonlocality [62]. Here, for simplicity, we put no restrictions on the probability distributions $p_\lambda(\mathbf{a}_b|\mathbf{x}_b)$, $p_\lambda(\mathbf{a}_{\bar{b}}|\mathbf{x}_{\bar{b}})$ other than positivity and normalisation (for example they may be signalling); note that more sophisticated definitions of GMNL were proposed [63,64]. The N-party Greenberger-Horne-Zeilinger (GHZ) state $|\mathrm{GHZ}\rangle = (|0\rangle^{\otimes N} + |1\rangle^{\otimes N})/\sqrt{2}$ is known to produce correlations which are GMNL, as proven by the violation of the Svetlichny inequalities [62, 65, 66].

Recall that we would like to find multipartite entangled states admitting LHV models, i.e. satisfying (6.2) for all local measurements. Trivially, one can achieve this by taking a bipartite state $\rho_{A_1A_2}$ admitting a LHV model and constructing the state

$$\rho_{A_1\cdots A_N} = \rho_{A_1A_2} \otimes \rho_{A_3} \otimes \rho_{A_3} \otimes \cdots \otimes \rho_{A_N}. \tag{6.4}$$

Since $\rho_{A_1A_2}$ is entangled, so is $\rho_{A_1\cdots A_N}$, and the LHV model for $\rho_{A_1A_2}$ can be easily extended to the whole state as the remaining subsystems are in a separable state. To make the question relevant, we therefore require a stronger notion of multipartite entanglement, called *genuine multipartite entanglement*.

6.2 Genuine multipartite entanglement

Consider N parties sharing a multipartite quantum state ρ acting on $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_N$, where $\mathcal{H}_i$ is the local Hilbert space of party i. Denote again by $(b, \bar{b}) \in \mathcal{B}$ a bipartition of the N parties. If ρ can be decomposed as a mixture of states that are each separable on some bipartition of the Hilbert space then we have

$$\rho = \sum_{(b,\bar{b})\in\mathcal{B}} p_b \left(\sum_j q_j^b |\Phi_j\rangle\langle\Phi_j|_b \otimes |\Phi_j\rangle\langle\Phi_j|_{\bar{b}} \right), \tag{6.5}$$

with $\sum_b p_b = \sum_j q_j^b = 1$ and $|\Phi_j\rangle\langle\Phi_j|_b$ acts on the Hilbert space specified by the partition b (and similarly for $|\Phi_j\rangle\langle\Phi_j|_{\bar{b}}$). If ρ does not admit such a decomposition then it is genuinely multipartite entangled (GME). Such states can thus not be created via local operations and communication using only biseparable states.

Determining whether a given state is GME is challenging, as one must search over all possible decompositions (6.5). However, there are sufficient conditions for an N-qubit state to be GME [67–69] (see also [70]).

One such condition, which can be seen as a generalisation of concurrence for multi-qubit systems, is given be the following [67–69]. Write the state ρ in the canonical basis $|0,0,\cdots,0\rangle, |0,0,\cdots,1\rangle, \cdots, |1,1,\cdots,1\rangle$ as

$$\rho = \begin{pmatrix} c_1 & & & & & & z_1 \\ & c_2 & & & & z_2 & \\ & & \ddots & & \iddots & & \\ & & & c_n & z_n & & \\ & & & z_n^* & d_n & & \\ & & \iddots & & \ddots & & \\ & z_2^* & & & & d_2 & \\ z_1^* & & & & & & d_1 \end{pmatrix} \tag{6.6}$$

(we only write the elements of interest), where $n = 2^{N-1}$. Then ρ is GME if

$$C(\rho) = 2 \max_i \{|z_i| - w_i\} > 0, \tag{6.7}$$

where $w_i = \sum_{j \neq i}^{n} \sqrt{c_j d_j}$. For the case of qubit X-matrices (of the above form), the above condition is both necessary and sufficient. Below we will use this condition to ensure that a state is GME.

6.3 Method

Our construction will make use of tools from EPR steering and the following generalisation of Theorem 3.2.1 1 for multiple copies of a bipartite state. The basic idea is that by taking many copies of an unsteerable state in a star-network configuration (see Fig. 6.1), one may make a joint measurement on the central parties, swapping the entanglement to the remaining parties such that they are projected on to a state that is GME. Importantly, if the original state is unsteerable, the swapped state admits a LHV model. Concretely, we have the following lemma:

Lemma 6.3.1. Let ρ be a quantum state acting on $\mathcal{H}_{A_1} \otimes \mathcal{H}_{B_1}$. The state $\rho^{\otimes N}$ therefore acts on $\mathcal{H}_{A_1} \otimes \cdots \otimes \mathcal{H}_{A_N} \otimes \mathcal{H}_{B_1} \otimes \cdots \otimes \mathcal{H}_{B_N} = \mathcal{H}_A \otimes \mathcal{H}_B$. Furthermore let Ω_B be a completely positive linear map acting on $\mathcal{H}_B$. If ρ is unsteerable from A_1 to B_1, then the N-party state

$$\rho_{A_1 \cdots A_N} = \frac{\mathrm{Tr}_B \left[\mathbb{1}_A \otimes \Omega_B(\rho^{\otimes N}) \right]}{\mathrm{Tr}[\mathbb{1}_A \otimes \Omega_B(\rho^{\otimes N})]} \tag{6.8}$$

admits a local hidden variable model, of the form (6.2), on the N-partition $A_1/A_2/\cdots/A_{N-1}/A_N$.

The set of measurements for which the multipartite LHV model will hold will depend on the original LHS model for ρ, e.g. if the LHS model for ρ is valid for projective/POVM measurements, then the LHV model for $\rho_{A_1 \cdots A_N}$ will be valid for projective/POVM measurements.

The LHV model for $\rho_{A_1 \cdots A_N}$ is constructed as follows. Denote the i^{th} state of $\rho^{\otimes N}$ by ρ_i. Since ρ_i is unsteerable, it admits a LHS model with local hidden states σ_{λ_i} distributed according to q_{λ_i}, and local response functions $p_{\lambda_i}(a_i|x_i)$. The LHV model for $\rho_{A_1 \cdots A_N}$ then has a local hidden variable $\vec{\lambda} = (\lambda_1, \lambda_2, \cdots, \lambda_N)$ distributed according to

$$Q(\vec{\lambda}) = \frac{\prod_i q_i}{\mathcal{N}} \mathrm{Tr}[\Omega_B(\otimes_i \sigma_{\lambda_i})], \tag{6.9}$$

where $\mathcal{N}$ ensures the normalisation of Q. One sees that the initially uncorrelated variables λ_i of the shared variable $\vec{\lambda}$ in the multipartite LHV model are correlated using the map Ω_B and the local hidden states. This can be seen as a classical analogue of entanglement swapping whereby a local CP map on the central parties is used to quantumly correlate a collection of initially uncorrelated quantum systems. For a full proof of the above lemma see the Appendix of [Paper E].

6.4 GME states with fully local models

6.4.1 Projective measurements

We now use Lemma 6.3.1 to construct N-qubit GME states which admit fully local models for projective measurements for all N. Specifically, consider the class of two-qubit states (defined previously in (5.21))

$$\rho(p,\chi) = p|\psi_\chi\rangle\langle\psi_\chi| + (1-p)\rho_A^\chi \otimes \frac{\mathbb{1}}{2}, \tag{6.10}$$

where $0 \leq p \leq 1$, $0 \leq \chi \leq \pi/4$, $|\psi_\chi\rangle = \cos\chi|00\rangle + \sin\chi|11\rangle$, and $\rho_A^\chi = \mathrm{Tr}_B\,|\psi_\chi\rangle\langle\psi_\chi|$. These states are entangled for all $\chi \in]0,\pi/4]$ if $p > 1/3$. As we have seen (Eq. 5.22) these states admit a LHS model for projective measurements from Alice to Bob if

$$\cos^2(2\chi) \geq \frac{2p-1}{(2-p)p^3}. \tag{6.11}$$

Hence, for any $0 \leq p < 1$ one may find a corresponding $\chi > 0$ such that $\rho(p,\chi)$ is unsteerable. We define the completely positive linear map

$$\Omega_B(\sigma) = F_B \sigma F_B^\dagger\,,\ F_B = |0\rangle\left[\langle 0,0,\cdots,0| + \langle 1,1,\cdots,1|\right],$$

which projects the systems of $B_1\cdots B_N$ onto a N-qubit GHZ state. We may now define the N-party state $\rho_{A_1\cdots A_N}$ by using $\rho(p,\chi)$ and Ω_B in (6.8). One can show (see [Paper E] Appendix B) that the concurrence of this state for a fixed N, p, χ is given by

$$C(\rho_{A_1\cdots A_N}) = \frac{2\sin^N(2\chi)\left(p^N + \left[\frac{1+p}{2}\right]^N + \left[\frac{1-p}{2}\right]^N - 1\right)}{\left[1+p\cos 2\chi\right]^N + \left[1-p\cos 2\chi\right]^N}. \tag{6.12}$$

It follows that for any N, one can find parameters p,χ such that (i) condition (6.11) is satisfied (ensuring that $\rho(p,\chi)$ has a LHS model), and (ii)

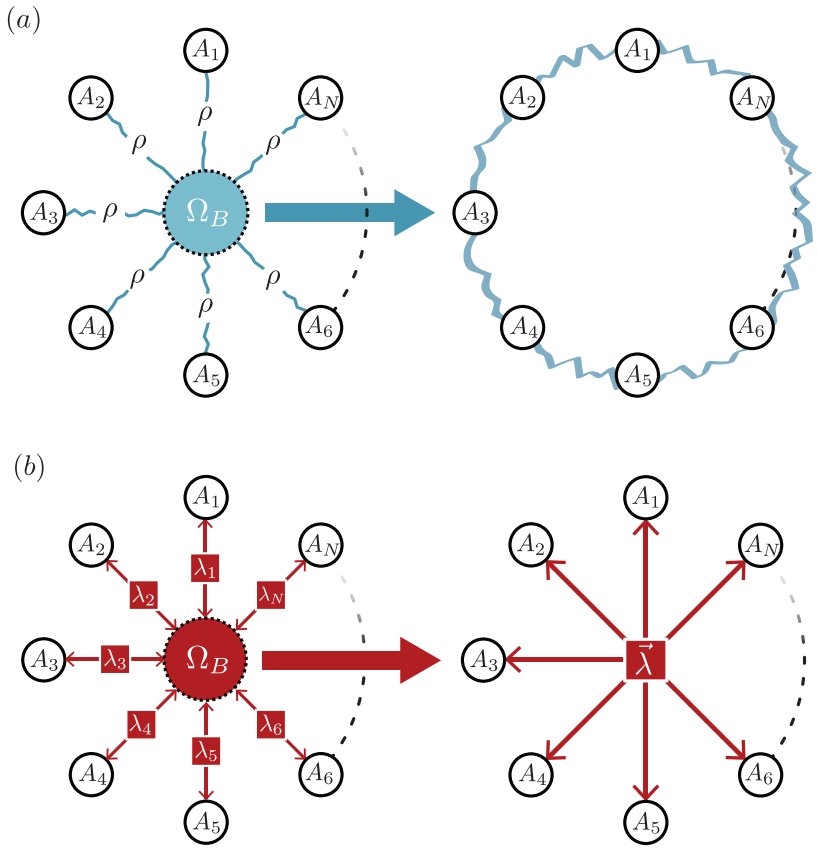

Figure 6.1: Construction of multipartite states admitting a fully local model. (a) Construction of the state. First, place N copies of a bipartite state ρ in a star-shaped network. Then, apply a map Ω_B at the central node (i.e. on parties $B_1 \cdots B_N$), and trace out these parties. We thus obtain an N-partite state, $\rho_{A_1 \cdots A_N}$ (represented by the blue wiggly line), shared by parties $A_1 \cdots A_N$. (b) LHV model. If ρ admits a LHS model, one can simulate the correlations of the star-shaped network for $\rho^{\otimes N}$, whereby the central node receives the hidden states σ_{λ_i} independently from each source and the parties A_i receive hidden variables λ_i. One may now correlate the individual λ_i's by having the map Ω_B act on the hidden states, i.e. we can define a new distribution over $\vec{\lambda} = (\lambda_1, \cdots, \lambda_N)$ that depends on $\mathrm{Tr}[\Omega_B(\otimes_i \sigma_{\lambda_i})]$. If each party A_i uses the same response function as in the LHS model for ρ, then the resulting statistics on parties $A_1 \cdots A_N$ simulate exactly the state $\rho_{A_1 \cdots A_N}$.

$C(\rho_{A_1\cdots A_N}) > 0$, proving that $\rho_{A_1\cdots A_N}$ is GME. To give a specific example, take $p = 1 - 1/N^2$ and $\chi > 0$ such that (6.11) is saturated. One sees that the denominator of (6.12) and $\sin^N 2\chi$ are both positive. We therefore need

$$p^N + \left[\frac{1+p}{2}\right]^N + \left[\frac{1-p}{2}\right]^N > 1 \tag{6.13}$$

to be positive for all $N \geq 2$. For the case $N = 2$ one has $p = 3/4$ and this is easily verified. For $N > 2$, upon substituting $p = 1 - 1/N^2$ the left hand side becomes

$$\begin{aligned}\left[1 - \frac{1}{N^2}\right]^N + \left[1 - \frac{1}{2N^2}\right]^N + \left[\frac{1}{2N^2}\right]^N \\ > 2\left[1 - \frac{1}{N^2}\right]^N > 2\left[1 - \frac{1}{N}\right] > 1\end{aligned} \tag{6.14}$$

where for the first inequality use the fact that $[1-1/N^2]^N < [1-1/2N^2]^N$ and $[1/2N^2]^N > 0$, and the second inequality follows from Bernoulli's inequality. Thus, these states are GME and admit a LHV model for projective measurements.

6.4.2 POVM measurements

Since the LHS model for $\rho(p, \chi)$ is valid only for projective measurements, the previous model is only valid for local projective measurements $\rho_{A_1\cdots A_N}$. One would therefore like to extend this to include POVM measurements. While the states $\rho(p, \chi)$ are not known to admit a LHS model for POVMs, we can nevertheless proceed differently. Starting from $\rho_{A_1\cdots A_N}$, we can construct another state, ρ_{GME}, which is both GME and local for POVM measurements.

Specifically, define $\rho_{A_1\cdots A_k} = \text{Tr}_{A_{k+1}\cdots A_N}[\rho_{A_1\cdots A_N}]$ and denote by $\circlearrowleft[\rho]$ the unnormalised and symmetrised version of ρ. Then the state

$$\rho_{\text{GME}} = \frac{1}{2^N}\left[\rho_{A_1\cdots A_N} + \sum_{j=0}^{N-1} \circlearrowleft\left[\rho_{A_1\cdots A_j} \otimes |2\rangle\langle 2|^{\otimes N-j}\right]\right] \tag{6.15}$$

admits a fully local model, for arbitrary local POVMs. Note that $|2\rangle\langle 2|$ denotes the projector onto a subspace orthogonal to the qubit subpace. The above follows from a straightforward extension of Protocol 2 of Ref. [49] to the case of N parties.

Finally, we have to show that the state is GME. Note that if each party makes a local projection on the qubit subspace $|0\rangle\langle 0| + |1\rangle\langle 1|$ then the resulting (renormalised) state is $\rho_{A_1\cdots A_N}$, which is GME. Since one cannot create GME using stochastic local operations, it follows that ρ_{GME} is GME.

6.5 Genuine multipartite hidden nonlocality

We have shown that GME states can admit a fully local LHV models for arbitrary non-sequential measurements. A natural question now is whether these states have hidden nonlocality [49, 57, 58]. This is possible if one can transform the original state via local stochastic operations, i.e. local filters, to another state that violates some Bell inequality. Below we will see that the states ρ_{GME} have genuine multipartite hidden nonlocality. Furthermore, the activation of nonlocality is maximal, in the sense that the filtered state exhibits GMNL, despite the initial state being fully local.

Consider N parties sharing ρ_{GME}. Let each party perform a local filtering operation given by

$$G_\epsilon = \epsilon|0\rangle\langle 0| + |1\rangle\langle 1|, \tag{6.16}$$

hence transforming ρ_{GME} to the state

$$\rho_\epsilon = \frac{G_\epsilon^{\otimes N}\rho_{\text{GME}}\, G_\epsilon^{\otimes N}}{\text{Tr}[G_\epsilon^{\otimes N}\rho_{\text{GME}}\, G_\epsilon^{\otimes N}]}. \tag{6.17}$$

For $\epsilon = \tan\chi$ (where χ is the parameter in (6.10)), the filtered states is essentially a pure N-party GHZ state $[|0\rangle^{\otimes N} + |1\rangle^{\otimes N}]/\sqrt{2}$. Specifically, the fidelity between the two states is given by

$$\begin{aligned}\mathcal{F}(\rho_\epsilon, |\text{GHZ}\rangle\langle\text{GHZ}|) &= \langle\text{GHZ}|\rho_\epsilon|\text{GHZ}\rangle \\ &= \frac{1}{2}\left[p^N + \left(\frac{1+p}{2}\right)^N + \left(\frac{1-p}{2}\right)^N\right].\end{aligned} \tag{6.18}$$

which tends to 1 when p is sufficiently close to 1 (p is the parameter appearing in the state (6.10) used to construct ρ_{GME}). Since the GHZ state is known to exhibit GMNL for any N, in particular via violation of the Svetlichny inequalities [65, 66] (which are robust to noise), it follows that ρ_ϵ can also be made GMNL.

6.6 Outlook

The study of LHV models for multipartite entangled states is almost unexplored. The tools in this chapter (i.e. Lemma 6.3.1) are likely to be useful in future studies. Even in the simplest case where one starts with two unsteerable states and obtains a bipartite state after the map Ω_B, we have explored very

little the set of possible states one could obtain. Our methods could thus find new LHV models for entangled quantum states even in the bipartite scenario, as well as the multipartite scenario. The concept of steering has also recently been extended to the multipartite setting [47] and our tools may also prove to be useful here.

Finally, we note that even though the states presented here admit LHV models for single measurements, this fails dramatically when sequential measurements are made, since the state becomes GMNL. It is thus desirable to look for multipartite entangled states (if such states exist) that remain local even if sequences of measurements are performed. Note that this is still open even in the bipartite scenario - for progress here see the previous section 5.3.4.

Chapter 7

Simulation of entangled quantum states with finite resources

By this point it is well established that there exist entangled quantum states that admit LHV models and thus do not violate any Bell inequality. In this sense, these states can be seen as classical since they can be simulated using purely classical resources (namely classical shared randomness). However, all examples presented thus far (and in all other works) share one property in common: they all require a shared classical variable of infinite dimension. What precisely we mean by dimension is as follows. Consider a LHV model which makes use of a shared variable $\lambda \in \Lambda$. Define the dimension of the shared variable to be the cardinality of the set Λ, $D = |\Lambda|$, that is, the number of labels that we need to describe a general λ. Since all models presented thus far use a continuous shared variable $\vec{\lambda} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ requiring two real numbers θ, ϕ, for these we have $|\Lambda| = \infty$.

One could thus argue that such models are in some sense unphysical, since they require the transmission of an infinite amount of classical information. This disparity is particularly evident when one compares to the case of quantum systems: to produce the correlations of a (Bell local) two-qubit Werner state we require an entangled quantum system of total Hilbert space dimension 4, or, using Werner's LHV model, classical systems of infinite dimension. The central motivation of this work is to investigate the classical cost of simulating correlations of entangled states: is it possible to construct LHV models for entangled quantum states with $D < \infty$? And if so, what is the minimum D required to simulate these states? We prove the existence of entangled quantum states that can be simulated by LHV models with classical shared

resources of finite dimension [**Paper F**]. We also consider a related problem of simulating the correlations of nonlocal quantum states using finite classical communication and finite classical shared resources.

7.1 Simulating separable states

Any separable state is trivially local and can decomposed as [71, 72]

$$\rho = \sum_{\lambda=1}^{D} p_\lambda \rho_A^\lambda \otimes \rho_B^\lambda \tag{7.1}$$

where $D = d^4$ is the local Hilbert space dimension of ρ. This follows from Carathéodory's theorem [73] since the set of $d \times d$ states lives in a convex space of dimension $d^2 - 1$. Hence we have a shared variable λ, distributed according to p_λ. Choosing $p_\lambda(a|x) = \mathrm{Tr}[\rho_A^\lambda M_{a|x}]$ and $p_\lambda(b|y) = \mathrm{Tr}[\rho_B^\lambda M_{b|y}]$ for Bob, we simulate the state ρ. Hence, for separable states we have $D \leq d^4$ (note that for two-qubit separable states $D = 4$ is enough [74]).

7.2 Simulating entangled Werner states

If we wish to simulate an entangled state using a λ of finite dimension, we clearly cannot use the same method as for separable states, since by definition they do not admit a separable decomposition. We nevertheless show that this can be done, first by considering the simulation of entangled two-qubit Werner states of a finite number of measurements for one party, then extending this to the set of all measurements on a more noisy version of the state. After this we present a general construction. For more details see Appendix [Paper F].

7.2.1 Fixed measurements

Consider measurements for Alice given by $M_{a|\vec{x}_i} = \frac{1}{2}[1 + a\vec{x}_i \cdot \vec{\sigma}]$ with $a = \pm 1$. We restrict ourselves to a finite number of measurements given by the Bloch vectors $\{\vec{x}_i\}$, $i = 1, \cdots, 12$ which are the vertices of a icosahedron inscribed inside the Bloch sphere (see Fig. 7.1). For now, Alice will be restricted to making only measurements from this set and no others. For Bob, we allow arbitrary projective measurements $M_{b|\vec{y}} = \frac{1}{2}[1 + b\vec{y} \cdot \vec{\sigma}]$, ± 1 given by Bloch vectors $\vec{y}$. Our aim is to reproduce the correlations of the Werner state

$$\rho_\alpha = \alpha \, |\psi^-\rangle\langle\psi^-| + (1 - \alpha)\frac{\mathbb{1}}{4} \tag{7.2}$$

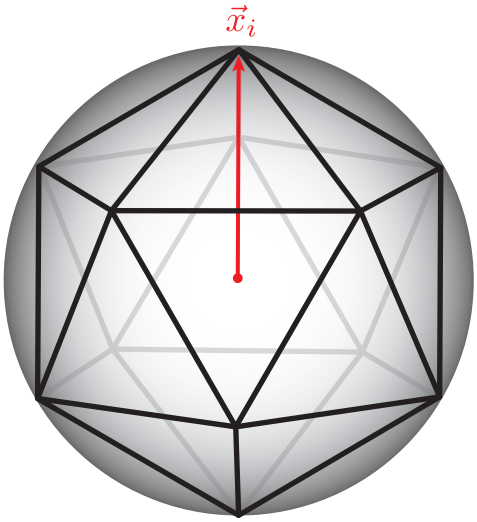

Figure 7.1: Vertices of the icosahedron inscribed in the Bloch sphere used as finite shared randomness to construct our LHV model for the Werner state.

whose correlations for local projective measurements given by Bloch vectors $\vec{x}$, $\vec{y}$ are defined by:

$$\langle ab\rangle = -\alpha\, \vec{x}\cdot\vec{y}\,; \quad \langle a\rangle = \langle b\rangle = 0. \tag{7.3}$$

Here, $\langle ab\rangle = \sum_{ab} ab\, p(ab|\vec{x}\vec{y})$ is the expectation value of the product of outcomes $a, b = \pm 1$ (and similarly for $\langle a\rangle, \langle b\rangle$).

We now describe the LHV model for these measurements. We introduce the local hidden variable $\lambda = 1, \cdots, 12$ distributed uniformly, $q_\lambda = \frac{1}{12}$. Since $D = 12$ we are using a finite amount of shared randomness. Each value of λ corresponds to each of the $\vec{x}_\lambda$ from Alice's measurement set. Hence, the shared randomness can be thought of as a uniform distribution over the vertices of the same icosahedron. We then define the response functions

$$p_\lambda(a|\vec{x_i}) = \frac{1 - a\,\mathrm{sgn}[\vec{x}_i\cdot\vec{x}_\lambda]}{2}\,; \quad p_\lambda(b|\vec{y}) = \frac{1 + b\,\vec{y}\cdot\vec{x}_\lambda}{2}. \tag{7.4}$$

Notice that the icosahedron is such that

$$\forall\, \vec{x}_\lambda\, \exists\, \vec{x}_{\lambda'}\ \text{s.t.}\ \vec{x}_\lambda = -\vec{x}_{\lambda'}\,; \quad \sum_{\lambda\,\text{s.t}\,\vec{x}_\lambda\cdot\vec{x}_j\geq 0} \vec{x}_\lambda = \gamma\,\vec{x}_j \quad \forall\, j. \tag{7.5}$$

The second condition says that if one sums all the vectors in one half of the sphere defined by $\vec{x}_j$, one recovers a vector proportional to $\vec{x}_j$ given by the

factor γ. For the icosahedron, one has

$$\gamma = 1 + \sqrt{5}. \tag{7.6}$$

With these properties it is then easy to see that

$$\langle a \rangle = \frac{1}{12} \sum_{\lambda} -\mathrm{sgn}[\vec{x}_i \cdot \vec{x}_\lambda] = 0 \; ; \quad \langle b \rangle = \frac{1}{12} \sum_{\lambda} \vec{y} \cdot \vec{x}_\lambda = 0. \tag{7.7}$$

and that

$$\begin{aligned} \langle ab \rangle &= \frac{1}{12} \sum_{\lambda} -\mathrm{sgn}[\vec{x}_i \cdot \vec{x}_\lambda] \, \vec{y} \cdot \vec{x}_\lambda \\ &= \frac{1}{12} \left[- \sum_{\lambda \,\mathrm{s.t}\, \vec{x}_\lambda \cdot \vec{x}_i \geq 0} \vec{x}_\lambda \cdot \vec{y} + \sum_{\lambda \,\mathrm{s.t}\, \vec{x}_\lambda \cdot \vec{x}_i < 0} \vec{x}_\lambda \cdot \vec{y} \right] \\ &= -\frac{1}{6} \sum_{\lambda \,\mathrm{s.t}\, \vec{x}_\lambda \cdot \vec{x}_i \geq 0} \vec{x}_\lambda \cdot \vec{y} = -\frac{\gamma}{6} \, \vec{x}_i \cdot \vec{y}. \end{aligned} \tag{7.8}$$

For these measurements we therefore simulate the Werner state with $\alpha = \gamma/6 \approx 0.54$. Note that we could have chosen other polyhedra satisfying the properties (7.5) (such as all platonic solids except the tetrahedron), resulting in different visibilities (see Table I of Appendix [Paper F]).

7.2.2 Shrinking the sphere

We now explain how this may be extended to all projective measurements. Our first step is to notice that the noisy projective measurement

$$M^{\ell}_{a|\vec{x}} = \frac{\mathbb{1} + a \, \ell \, \vec{x} \cdot \vec{\sigma}}{2} \tag{7.9}$$

can be simulated if $\ell \vec{x}$ lies inside the convex hull of $\{\vec{x}_i\}$, that is, if one can find a decomposition

$$\ell \, \vec{x} = \sum_i \omega_i \vec{x}_i \; ; \quad \sum_i \omega_i = 1 \; ; \quad \omega_i \geq 0. \tag{7.10}$$

This follows from the linearity of quantum mechanics since using the above one has

$$\mathrm{Tr}\left[M^{\ell}_{a|\vec{x}} \otimes M_{b|\vec{y}} \, \rho \right] = \sum_i \omega_i \mathrm{Tr}\left[M_{a|\vec{x}_i} \otimes M_{b|\vec{y}} \, \rho \right] \tag{7.11}$$

$$= q_\lambda \sum_\lambda \left[\sum_i \omega_i \, p_\lambda(a|\vec{x}_i) \right] p_\lambda(b|\vec{y}), \tag{7.12}$$

where in the last line we have used the fact that the state admits a model for the fixed measurements $\{\vec{x}_i\}$. Hence, if upon receiving λ Alice chooses to simulate the measurement of $\vec{x}_i$ with probability ω_i, we simulate $M^{\ell}_{a|\vec{x}}$.

For any polyhedron there will be a maximum ℓ such that all vectors $\vec{x}$ admit a decomposition (7.10). We call this ℓ the *shrinking factor* of the polyhedron. Geometrically this corresponds to the radius of the largest sphere centred on the origin that one can inscribe inside the polyhedron. For the icosahedron we find

$$\ell = \sqrt{(5 + 2\sqrt{5})/15} \approx 0.795. \tag{7.13}$$

It follows that given a model for the fixed measurements $\{\vec{x}_i\}$ for Alice, one has a model for all noisy measurements $M^{\ell}_{a|\vec{x}}$, where ℓ is the shrinking factor of the polyhedron.

The final ingredient is to realise that this noise in the measurements can be passed onto the state, since one has

$$\mathrm{Tr}\left[M^{\ell}_{a|\vec{x}} \otimes M_{b|\vec{y}}\,\rho\right] = \mathrm{Tr}\left[M_{a|\vec{x}} \otimes M_{b|\vec{y}}\,(\ell\rho + (1-\ell)\frac{\mathbb{1}}{2} \otimes \rho_B)\right]. \tag{7.14}$$

The statistics are therefore equivalent to the projective measurements $M_{a|\vec{x}}$ on the state $\ell\rho + (1-\ell)\mathbb{1} \otimes \rho_B$. Since we have the Werner state, which has $\rho_B = \mathbb{1}/2$ this noise changes the state ρ_α to

$$\rho_\alpha \to \ell\,\rho_\alpha + (1-\ell)\frac{\mathbb{1}}{4} = \rho_{\ell\,\alpha}, \tag{7.15}$$

i.e. we reduce the visibility by a factor ℓ. For the fixed icosahedron one therefore finds a final visibility

$$\frac{\ell\gamma}{6} \approx 0.43, \tag{7.16}$$

which is larger than $\frac{1}{3}$ hence entangled. Compactly, we have the full protocol:

Protocol 1. Alice and Bob share $\lambda \in \{1, \ldots, 12\}$, uniformly distributed. Upon receiving setting $\vec{x}$, Alice calculates the subnormalised vector $\vec{x}' = \ell\vec{x}$. This ensures that $\vec{x}'$ lies inside the convex hull of V and so Alice can find a convex decomposition $\vec{x}' = \sum_i \omega_i \vec{x}_i$ with $\sum_i \omega_i = 1$ and $\omega_i \geq 0$. Then, with probability ω_i, she outputs $a = \pm 1$ with probability $(1 \mp \mathrm{sgn}[\vec{x}_\lambda \cdot \vec{x}_i])/2$. Bob, upon receiving $\vec{y}$, outputs $b = \pm 1$ with probability $(1 \pm \vec{y} \cdot \vec{v}_\lambda)/2$.

7.2.3 General construction for Werner states

The above model will not work for general polyhedra, since they will not necessarily satisfy the properties (7.5). One can, however, relax these conditions and work with polyhedra satisfying only the first of the two conditions. We thus now have polyhedra, defined by a set of normalised vertices $\{v_\lambda\}$ such that

$$\forall\, \vec{v}_\lambda \,\exists\, \vec{v}_{\lambda'} \text{ s.t. } \vec{v}_\lambda = -\vec{v}_{\lambda'}. \tag{7.17}$$

Again, we will first be interested in a finite set of measurements $\{\vec{x}_i\}$ for Alice and arbitrary projective measurements for Bob, however this time the set $\{x_i\}$ is defined by

$$\sum_{\lambda \text{ s.t } \vec{v}_\lambda \cdot \vec{v}_i \geq 0} \vec{v}_\lambda = \gamma_i\, \vec{x}_i. \tag{7.18}$$

Hence, for each $\vec{v}_i$ we have a corresponding $\vec{x}_i$ and γ_i. We further define $\gamma_{\min} = \min_i \gamma_i$. Following a similar reasoning to before, we see that if we take the uniform distribution $q_\lambda = 1/D$ and Alice outputs with

$$p_\lambda(a|\vec{x}_i) = \frac{1 - a\ \mathrm{sgn}[\vec{x}_i \cdot \vec{x}_\lambda]}{2} \tag{7.19}$$

with probability $\gamma_{\min}/\gamma_i$ and uniformly $p_\lambda(\pm|\vec{x}_i) = 1/2$ otherwise, we find (see [Paper F])

$$\langle ab \rangle = -\frac{\gamma_{\min}}{D} \vec{x}_i \cdot \vec{y}. \tag{7.20}$$

Hence, we simulate a Werner state with $\alpha = \gamma_{\min}/D$. Again the property (7.17) ensures that $\langle a \rangle = \langle b \rangle = 0$. Finally we may extend this to all projective measurements for Alice at the cost of reducing the overall visibility by considering the convex hull M of the set of fixed measurements $\vec{x}_i$ with corresponding shrinking factor ℓ. This leads to a final visibility

$$\alpha = \frac{2\ell}{D} \gamma_{\min}. \tag{7.21}$$

The full protocol is as follows:

Protocol 2. Alice and Bob share $\lambda \in \{1, \ldots, D\}$ uniformly distributed. Upon receiving setting $\vec{x}$, Alice calculates the subnormalised vector $\vec{x}' = \ell \vec{x}$ where ℓ is the radius of the largest sphere fitting inside M and centred on the

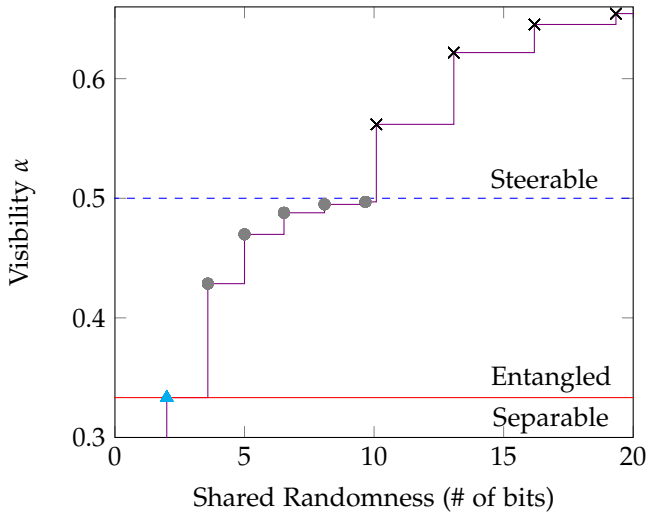

Figure 7.2: Visibility of the Werner state simulated using finite shared randomness as a function of the number of bits needed to encode λ. The rounded dots correspond to the model described here. Higher visibilities (crosses) can be obtained by generalising this idea and using linear programming techniques (see Section 7.3).

origin. This ensures that $\vec{x}'$ lies inside the convex hull of M and Alice can therefore find a convex decomposition $\vec{x}' = \sum_{i=1}^{D} \omega_i \vec{x}_i$. Then, with probability $p_i = \omega_i \gamma_{\min}/\gamma_i$ she outputs $a = -\mathrm{sgn}(\vec{v}_i \cdot \vec{v}_\lambda)$, and with probability $(1 - \sum_i p_i)$ she outputs a random bit. Bob, upon receiving $\vec{b}$, outputs $b = \pm 1$ with probability $(1 \pm \vec{b} \cdot \vec{v}_\lambda)/2$.

One can now define a sequence of polyhedra tending to a sphere such that

$$\frac{\gamma_{\min}}{D} \to \frac{1}{4}\,; \qquad \gamma \to 1 \tag{7.22}$$

as $D \to \infty$. In this limit we thus recover Werner's model $\alpha = 1/2$ and for all other polyhedra we have $\alpha < \frac{1}{2}$. An example of such a sequence is as follows (for example, starting from the icosahedron). Define a polyhedron that is the union of the icosahedron and its normalised dual (which is the dodecahedron). This defines the next polyhedron in the sequence. This is then repeated, at each step taking the union of the current polyhedron and its normalised dual. In Fig. 7.2 (round dots) the visibility versus the number of bits required to encode the shared randomness is plotted for the first 5 members of this

sequence.

7.3 General method

The above ideas can be generalised to find LHV simulations of other entangled quantum states. Specifically we have the following

Theorem 7.3.1. Consider a state ρ (of dimension $d \times d$) admitting a LHV model for all projective measurements. Then, a LHV model using only finite shared randomness can simulate all projective measurements on the state

$$\rho(\eta) = \eta^2 \rho + \eta(1-\eta)\left(\frac{\mathbb{1}}{d} \otimes \rho_B + \rho_A \otimes \frac{\mathbb{1}}{d}\right) + (1-\eta)^2 \frac{\mathbb{1} \otimes \mathbb{1}}{d^2}$$

for any $0 \leq \eta < 1$. Here $\rho_{A,B} = \mathrm{Tr}_{B,A}(\rho)$.

Proof. Notice that the ability to pass noise in the measurements on to noise on the state works for general quantum states. Consider the set of all projective measurements $\{A_a\}$ for Alice and $\{B_b\}$ for Bob. Define the sets of noisy projective measurements $\{A_a(\eta)\}$ and $\{B_b(\eta)\}$ given by

$$A_a(\eta) = \eta A_a + (1-\eta)\frac{\mathbb{1}}{d}\,; \quad B_b(\eta) = \eta B_b + (1-\eta)\frac{\mathbb{1}}{d}. \tag{7.23}$$

Then one has

$$\mathrm{Tr}[A_a \otimes B_b \rho(\eta)] = \mathrm{Tr}[A_a(\eta) \otimes B_b(\eta)\rho], \tag{7.24}$$

i.e. one can pass noise in the measurement on to the state, resulting in a new state $\rho(\eta)$. Now, for any η one can find a finite set of projective measurements such that the set of noisy projective measurements lies inside their convex hull, that is

$$\sum_i \omega_i \, M_a^i = M_a(\eta)\,; \quad \sum_j \nu_j \, M_b^j = M_b(\eta). \tag{7.25}$$

One can thus simulate the noisy measurements using a LHV for the fixed set of measurements since via the linearity of quantum mechanics

$$\mathrm{Tr}\left[M_a(\eta) \otimes M_b(\eta)\ \rho\right] = \sum_{ij} \omega_i \nu_j \, \mathrm{Tr}\left[M_a^i \otimes M_b^j \rho\right]. \tag{7.26}$$

Hence to simulate $M_a(\eta)$ Alice should used the LHV simulation for the fixed measurement M_a^i with probability given by ω_i (and similarly for Bob). Finally, if the probability distribution $p(ab|xy)$ resulting from these fixed sets of measurements admits a LHV model, then it can be decomposed as the convex combination of a finite set of local deterministic functions (since the set of local distributions forms a polytope). Hence, one can simulate these measurements (and therefore the set of noisy measurements) using a finite amount of shared randomness. Passing this noise on to the state completes the proof. □

Using tools of linear programming and the above method, we are then able to extend the range of α for which we simulate the Werner state past $\alpha = 1/2$. This is presented in Fig.7.2.

7.3.1 Extension to POVM measurements

So far we have been considering the simulation of the set of projective measurements only. Using the tools of [Paper D] (protocol 2) one can extend our results to POVM measurements. Specifically, one has that if the state ρ admits a LHV with finite shared randomness then the state

$$\rho' = \frac{1}{(d+1)^2}\left(\rho + d(\rho_A \otimes F + F \otimes \rho_B) + d^2 F \otimes F\right) \tag{7.27}$$

admits a local model for POVMs using also k bits of shared randomness. Here $F = |d+1\rangle\langle d+1|$ denotes a projector onto a subspace orthogonal to the support of ρ, hence ρ' is entangled iff ρ is entangled and of local dimension $d+1$.

7.4 Classical communication cost of simulating nonlocal correlations

We now discuss the case of simulating nonlocal (i.e. Bell inequality violating quantum states). Here, we will need classical communication since no LHV model exists. Previous protocols required either infinite communication or infinite shared randomness (see e.g. [75–77]). Here we construct protocols using finite classical resources.

7.4.1 Communication protocols without shared randomness

We first note that any full rank quantum state can be simulated using finite (although potentially large) classical communication and no shared randomness. This is achieved as follows. Any full rank state can be written in the form $\rho = \alpha|\Psi\rangle\langle\Psi| + (1-\alpha)\mathbb{1}/d^2$, where $|\Psi\rangle$ is an arbitrary entangled state of dimension $d \times d$, and $\alpha < 1$. Upon receiving her measurement setting $A = \{A_a\}$, Alice outputs a according to the distribution $p(a) = \mathrm{Tr}(\rho_A A_a)$ where ρ_A is Alice's reduced state. For output a, Bob should hold the (normalized) state $\rho_B^a = \mathrm{Tr}_A(A_a \otimes \mathbb{1}\rho)/p(a)$. Since ρ_B^a is full-rank (by construction), then for any $\alpha < 1$, there exists a polyhedron V (with D vertices, each representing a pure quantum state $|\phi_i\rangle\langle\phi_i| \in V$ of dimension d) such that Alice can decompose ρ_B^a as a convex combination

$$\rho_B^a = \sum_i^D \omega_i |\phi_i\rangle\langle\phi_i| \tag{7.28}$$

of the vertices of V. With probability ω_i (the coefficient of vertex i in the decomposition) Alice sends label i to Bob, who can then locally reconstruct the corresponding pure state (knowing V). The model thus reproduces the statistics of ρ using $\log_2(D)$ bits of communication, however this will diverge when α is large.

7.4.2 Communication protocols with shared randomness

We now consider the case where we allow for a finite amount of shared randomness. This allows us to reduce the dimension of the classical communication needed to simulate entanglement. Specifically, we construct communication models for Werner states using finite communication and finite shared randomness. Consider a polyhedron V with D vertices satisfying (7.17), with corresponding $\gamma_{\min}$ and shrinking factor ℓ. Our model uses $n \log_2(D)$ bits of shared randomness and $\log_2 n$ bits of communication (in the worst case), and simulates $\rho_W(\alpha)$ for

$$\alpha = \frac{\gamma_{\min}}{\gamma_{\max}} \left(1 - [1 - \frac{2\gamma_{\min}}{D}]^n\right) \ell^2 \tag{7.29}$$

where $\gamma_{\max} = \max_i(\gamma_i)$. Note that by choosing a symmetric enough polyhedron with $\ell \approx 1$ and $2\gamma_{\min}/D \approx 1/2$ we can simulate a $\rho_W(\alpha)$ for $\alpha \to 3/4$ with

$n = 2$ (when $D \to \infty$). Hence, using finite shared randomness and a single bit of communication suffices to simulate a nonlocal quantum state. See [Paper F] Appendix D for details.

7.5 Outlook

This work has inspired subsequent work on algorithmic methods for the construction of LHV models of entangled quantum states [78, 79]. Here, the ideas of above are generalised so that linear programming and SDP methods can be used to computationally find LHV and LHS models for entangled states of (in principle) any dimension. This has led to the first LHV simulation of non-full rank and bound entangled states [78], as well as thousands of new examples of bipartite and genuinely multipartite entangled states admitting LHV models [79]. However fundamental questions still remain. For example, what is the absolute minimum required number of bits needed to simulate entangled states of dimension d? Is it strictly larger than the $2 \log_2 d$ bits needed to simulate a separable state? Similarly, do there exist entangled states that necessarily require an infinite amount of shared randomness to simulate?

Part II

Certification of dimension and randomness from quantum systems

Chapter 8

Certification of dimension in network scenarios

The concept of dimension plays a important role in the areas of computation, communication, complexity, to the physics of thermodynamics/statistical mechanics. From a quantum information processing perspective, dimension (that is, Hilbert space dimension) is a valuable resource and it is therefore desirable to develop methods that can be used to certify the dimension of a given physical system under consideration, which is precisely the aim of this chapter.

Here we will work in the device-independent framework, which means that we may treat the devices used in the certification protocol as black boxes and work only with the statistics of measurement outcomes. Interestingly, one can stil certify the dimension of physical systems in this scenario. This problem was originally posed in the Bell scenario [80–82], where the amount of violation of certain Bell inequalities can be used as a certification of the dimension of bipartite entangled quantum states. These ideas were then later formalised in a prepare-and-measure scenario, where devices prepare, communicate and subsequently measure physical systems [15]. Here, we develop this latter scenario, extending the prepare-and-measure scenario to include networks of devices that can communicate and process physical systems [**Paper G**]. We construct new tests which certify the dimension of the systems used in several simple networks involving three devices. We also study the relative power of quantum and classical systems, and show that quantum systems can outperform classical systems of the same dimension. We then show that the advantage offered by quantum systems of a given dimension over their classical counterparts is significantly stronger in our simple networks of devices compared to the original prepare-and-measure scenario.

We will focus on a device independent approach, first introduced in [15],

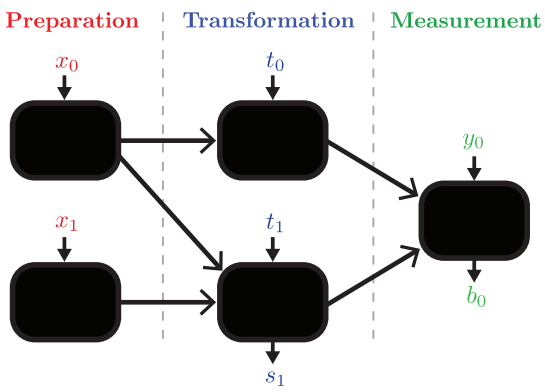

Figure 8.1: A possible network scenario featuring preparation, transformation and measurement. Vertical arrows represent classical inputs and outputs into the devices. Horizontal lines represent communication channels (either quantum or classical) between devices

where one makes very few assumptions on the functioning of the devices used.

8.1 General scenario

The general scenario we wish to consider is a network of devices exchanging and processing information, as represented in Fig. 8.1. Devices are represented by black boxes. An arrow connecting two devices represents a (one-way) communication channel between them.

A network consists of three levels: (i) a number of preparation devices, (ii) a number of transformation devices and (iii) a number of measurement devices. In each round of the experiment, the observer chooses the preparations $\mathbf{x}$, the transformations $\mathbf{t}$ and the measurement settings $\mathbf{y}$. He then obtains measurement outcomes $\mathbf{b}$; note that transformation devices can also provide outcomes, denoted $\mathbf{s}$. More precisely, we have that the choice of preparations is given by $\mathbf{x} = \{x_i\}$, where x_i denotes the input for device i. The choice of transformations is $\mathbf{t} = \{t_j\}$, where t_j denotes the input for device j, and the (possible) outcomes are $\mathbf{s} = \{s_j\}$, where s_j denotes the output of device j. Finally, the choice of measurement settings is $\mathbf{y} = \{y_k\}$, where y_k denotes the input for measurement device k, and gives outcomes $\mathbf{b} = \{b_k\}$, where b_k is the output of measurement device k. The experiment is therefore characterized

by the data

$$p(\mathbf{b}, \mathbf{s}|\mathbf{x}, \mathbf{t}, \mathbf{y}), \tag{8.1}$$

that is, the conditional probabilities of observing outputs $\mathbf{b}, \mathbf{s}$ given inputs $\mathbf{x}, \mathbf{t}, \mathbf{y}$. A general scenario is thus specified by a directed graph representing the network, and the number of inputs and outputs for each of the devices (which we will here consider to be finite).

In this network, the devices exchange information encoded in physical systems. For instance, upon receiving input x_i, each preparation device emits a system, the state of which is adapted depending on x_i. Which physical system is used, and what mechanism is used to encode information in it, is completely unknown to the observer, who has only access to inputs and outputs of the black boxes. That is, we work in a device-independent scenario.

Now the main point is the following. Clearly, the amount of information about x_i which can be encoded in the system will depend on its dimension (i.e. the number of independent degrees of freedom of the system). Therefore, we expect that a restriction on the dimension will in general limit the possible observable data (8.1). Consider for instance the case in which the outputs $\mathbf{b}$ contain all information about the inputs $\mathbf{x}$. This implies that the mediating physical systems had enough dimensions for encoding $\mathbf{x}$ perfectly.

The main question we will discuss in the present work is to understand the limitations on the data, arising from constraints on the dimension of the mediating systems. This will allow us to find lower bounds on the dimension of the systems present in a network for given data (8.1). In particular, we will discuss bounds for both classical and quantum systems. Notably, we will see that for a fixed dimension, quantum systems outperform classical ones.

8.2 Classical networks

For the sake of clarity, we will focus on the network consisting of one preparation device, followed by a single transformation device, and finally a single measurement device (see Fig. 8.2). The data is thus given by the conditional distribution $p(b, s|x, t, y)$; we consider a finite (but otherwise unspecified) number of inputs and outputs. Note that the methods discussed below can be straightforwardly generalised to more general networks.

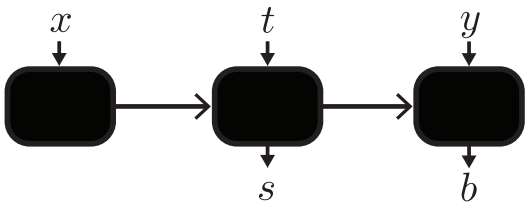

Figure 8.2: A simple network consisting of a preparation, a transformation and a measurement device. The set of possible distributions of inputs and outputs, $p(bs|xty)$, will depend on the dimension of the communication allowed between the devices and whether the communication is classical or quantum.

8.2.1 Basics

We start our analysis by considering classical communication between the devices. Denote by c_0 the communication sent from the preparation device to the transformation device, and c_1 the communication sent from the transformation device to the measurement device. We consider communication of bounded dimension d, that is

$$c_0, c_1 \in \{1, \cdots, d\}. \tag{8.2}$$

Upon receiving input x, the preparation device sends communication c_0, with probability $p(c_0|x)$. In turn, upon receiving input t and communication c_0 (from the preparation device), the transformation device outputs s and sends communication c_1 to the measurement device with probability $p(s, c_1|t, c_0)$. Finally, upon receiving measurement setting y and communication c_1, the measurement device outputs b with probability $p(b|y, c_1)$. We thus have that

$$p(b, s|x, t, y) = \sum_{c_0, c_1 = 1}^{d} p(c_0|x) p(s, c_1|t, c_0) p(b|y, c_1). \tag{8.3}$$

We first consider the case in which all devices act deterministically. That is, each of the previously mentioned probabilities are either 0 or 1. It follows that each probability $p(b, s|x, t, y)$ also takes only values 0 or 1. We refer to these sets of data as 'deterministic strategies'.

In general, we also want to include the possibility that the devices in the network output probabilistically, and moreover that they follow a common strategy. That is, the behaviour of the devices might be correlated, due to some (common) internal variable $\lambda \in \Lambda$ (referred to as shared randomness). The set

of possible distributions now becomes all convex combinations of deterministic strategies:

$$p(b,s|x,t,y) = \int_\Lambda q_\lambda \mathrm{d}\lambda \sum_{c_0,c_1=1}^{d} p_\lambda(c_0|x) p_\lambda(s,c_1|t,c_0) p_\lambda(b|y,c_1),$$

where q_λ is a normalized probability density over λ and $p_\lambda(c_0|x)$ denotes the probability for the preparation device to send c_0, given input x and internal variable λ, and so on.

Any set of data that cannot be decomposed in the form (8.4) therefore requires the use of communication (c_0 and/or c_1) of dimension strictly greater than d. In the next sections we will see how to test whether a given set of data can be decomposed in the above form or not. This will provide the 'dimension witnesses' we are looking for.

8.2.2 Geometrical interpretation

The above ideas admit an elegant description in geometrical terms. The idea is essentially the same as for Bell nonlocality as described in chapter 1 adapted here to the prepare-and measure-scenario [15].

The goal is to characterize the set of distributions (8.4) in geometrical terms. Consider first one particular set of data $p(b,s|x,t,y)$. This distribution can be viewed as a vector $\mathbf{p}$ where each component of the vector corresponds to one of the probabilities $p(b,s|x,t,y)$ appearing in the data. Hence $\mathbf{p} \in \mathbb{R}^D$, where

$$D = |b|\,|s|\,|x|\,|t|\,|y| \tag{8.4}$$

with $|b|$ denoting the alphabet size of b, that is the number of possible outcomes b, and similarly for other symbols.

Next, consider the entire set of distributions admitting a decomposition of the form (8.4), that is, all sets of data that can be obtained by using communication c_0 and c_1 of dimension d. This set, denoted $\mathbb{P}_d$, thus forms a subspace of $\mathbb{R}^D$. In fact, $\mathbb{P}_d$ forms a convex polytope. Its extremal points (or vertices) correspond to the deterministic strategies, that is, the set of distributions of the form (8.3), for which $p(b,s|x,t,y) \in \{0,1\}$ for all b,s,x,t,y. Alternatively, the polytope $\mathbb{P}_d$ can also be characterized by its facets (of which there is a finite number, since the number of vertices is finite). Formally, facets are given by linear inequalities

$$\mathbf{p} \cdot \boldsymbol{\mathcal{I}} = \sum_{b,s,x,t,y} \alpha^{b,s}_{x,t,y}\, p(b,s|x,t,y) \leq C_d \tag{8.5}$$

where $\alpha^{b,s}_{x,t,y}$ and C_d are real numbers (usually integers). $\boldsymbol{\mathcal{I}}$ is the D-dimensional vector, with components $\alpha^{b,s}_{x,t,y}$, associated to the facet, i.e. orthogonal to the hyperplane given by the facet. Therefore we have that

$$\mathbf{p} \in \mathbb{P}_d \iff \mathbf{p} \cdot \boldsymbol{\mathcal{I}}^n_d \leq C^n_d \tag{8.6}$$

where the $\boldsymbol{\mathcal{I}}^n_d$'s represent all the facet inequalities (labelled by n). Moreover, we have that $\mathbb{P}_d \subseteq \mathbb{P}_{d+1}$, since all strategies involving d-dimensional communication can always be realised using communication of dimension $d+1$. Note also the similarity to (2.12), since the set of local distributions in a Bell scenario also forms a convex polytope.

In practice, the polytope $\mathbb{P}_d$ can be constructed for simple networks, i.e. few devices and small alphabets for the inputs and outputs. Specifically, one starts by listing the deterministic strategies, i.e. the vertices of the polytope. Then, appropriate software (see e.g. [22, 23]) allows one to find the facets of the polytope. As with nonlocality, the problem however becomes intractable beyond simple scenarios on standard computers.

Finally, note that one can slightly reduce the complexity of the problem by taking into account certain constraints on the data $p(b,s|x,t,y)$. This allows one to discard certain (redundant) components of $\mathbf{p}$. In particular, we have here the normalization conditions

$$\sum_{b,s} p(b,s|x,t,y) = 1 \quad \forall x,t,y \tag{8.7}$$

and the condition that

$$\sum_{b} p(b,s|x,t,y) = p(s|x,t) \quad \forall s,x,t,y. \tag{8.8}$$

That is, the output s of the transformation device does not depend on the choice of input y for the measuring device. This follows from the fact that y can in principle be chosen after the output s is obtained. For more general networks, it is important to take all such 'no-signaling' conditions into account in order to reduce the complexity of the problem.

8.2.3 Classical dimension witnesses

Our main goal is to develop methods for testing whether a given set of data $p(b,s|x,t,y)$ is compatible with a particular network sending communication of bounded dimension. To address this question, we will now discuss the concept of 'dimension witnesses', hence generalising the ideas of Ref. [15] to networks.

Consider linear combinations of the form:

$$W = \mathbf{w} \cdot \mathbf{p} = \sum_{b,s,x,t,y} \omega_{xty}^{bs} p(b,s|x,t,y) \leq C_d, \tag{8.9}$$

where $\mathbf{w}$ is a D-dimensional vector, with real components ω_{xty}^{bs}, and C_d is a real number. We say that an inequality of the above form is a *linear classical dimension witness of dimension* d, if (i) the inequality holds for any distribution $p(b,s|x,t,y)$ realisable with classical communication of dimension d, and (ii) there exists at least one distribution $p(b,s|x,t,y)$ (involving systems of dimension at least $d+1$) for which the inequality is violated.

The geometrical ideas discussed in the previous subsection are relevant here, as they will allow us to construct dimension witnesses. Take one facet inequality of the polytope $\mathbb{P}_d$: property (i) above will immediately be satisfied. In general, there will also exist a vector $\mathbf{p} \in \mathbb{P}_{d'}$ with $d < d'$ that will violate the facet inequality, and hence (ii) is also satisfied. Such facet inequalities will be called 'tight dimension witnesses'. In fact, the complete list of the facets of $\mathbb{P}_d$ will provide a complete list of dimension witnesses, which allow one to find the minimal dimension of the communication necessary to reproduce a given set of data.

In the section 8.5, we will present several examples of dimension witnesses.

8.3 Quantum networks

We now move to the case of quantum communication networks. Here, the classical channels are replaced by quantum channels. Our goal is thus to characterize the sets of data compatible with sending quantum communication of bounded Hilbert space dimension in the network. For the sake of clarity, we will also focus on the simple network of Fig. 8.2.

8.3.1 Basics

Consider again the network consisting of one preparation device, followed by a transformation device, and finally by a measurement device. The devices can now produce, process, and measure quantum systems. The constraint we consider is that the quantum systems transmitting information between the devices are of Hilbert space dimension bounded by d.

Let us first consider the preparation device. Upon receiving input x, the device prepares a d-dimensional quantum system in state ρ_x, which is sent to the transformation device. In turn, the transformation device receives input t,

as well as the quantum communication ρ_x, produces an outcome s, and sends a d-dimensional quantum system to the measurement device. The action of the transformation device can thus be represented by a set of completely positive (CP) maps $\{\Phi_{s|t}\}$ (acting on $\mathbb{C}^d$), such that $\sum_s \Phi_{s|t}$ is completely positive and trace preserving (CPTP): this ensures that $\sum_s p(s|x,t) = 1$ for all x, t. Note that, since we impose that all communication is of bounded dimension d, we restrict to CP maps which do not increase the Hilbert space dimension. With probability $\mathrm{Tr}[\Phi_{s|t}(\rho_x)]$ the transformation device outputs s, and sends the quantum state

$$\Phi_{s|t}(\rho_x)/\,\mathrm{Tr}[\Phi_{s|t}(\rho_x)] \tag{8.10}$$

to the measuring device. Finally, upon receiving this quantum communication and the input y, the measuring device provides an output b. This is represented by a set of measurement operators $M_{b|y}$ (acting on $\mathbb{C}^d$), such that $M_{b|y} \geq 0$ and $\sum_b M_{b|y} = \mathbb{1}$.

Putting all this together we obtain that

$$p(b,s|x,t,y) = \mathrm{Tr}\left(\Phi_{s|t}(\rho_x) M_{b|y}\right). \tag{8.11}$$

Any set of data admitting a decomposition of this form is thus realisable with quantum communication of dimension d. On the contrary, if such a decomposition cannot be found, then higher dimensional quantum systems must have been used.

As in the case of classical networks, it is also relevant to allow for the devices to act according to a common strategy λ. In this case, the set of compatible distributions is therefore the convex hull of those of the form (8.11):

$$p(b,s|x,t,y) = \int_\Lambda \mathrm{Tr}\left(\Phi^\lambda_{s|t}(\rho^\lambda_x) M^\lambda_{b|y}\right) \pi(\lambda)\mathrm{d}\lambda, \tag{8.12}$$

where now the states, transformations and measurements are written with λ dependence. Finally, note that one could also consider the case in which the devices share quantum correlations, i.e. initial entanglement (see Section 8.5.4 for an example).

8.3.2 Quantum dimension witnesses

The problem is now to test whether a given set of data $p(b,s|x,t,y)$ is compatible with a particular network sending quantum communication of bounded Hilbert space dimension. Similarly to the classical case discussed above, we now define 'quantum dimension witnesses'.

Consider again linear inequalities of the form

$$W = \mathbf{w} \cdot \mathbf{p} = \sum_{b,s,x,t,y} \omega_{xty}^{bs}\, p(b,s|x,t,y) \leq Q_d, \tag{8.13}$$

with $\mathbf{w}$ a D-dimensional vector, with real components ω_{xty}^{bs}, and Q_d a real number. In analogy to the classical case, W is a *linear quantum dimension witness of dimension d* if (i) the above inequality is satisfied by all sets of data $p(b,s|x,t,y)$ realisable with quantum communication of dimension d, and (ii) using quantum communication of dimension greater than d allows one to violate the inequality. Note that the set of distributions $\mathbf{p}$ admitting a quantum d-dimensional realisation is generally not a polytope. Therefore a more efficient characterisation of these sets involve witnesses that are nonlinear in the observed probabilities, see for example [83].

Finding quantum dimension witnesses is generally a harder task than in the classical case. To the best of our knowledge, there are no known efficient computational methods for this problem; see however Refs [84] for recent progress.

8.4 Testing non-classicality

An interesting development related to dimension tests is the possibility of certifying non-classicality of communication in a device-independent way, assuming an upper-bound on the dimension. This aspect was discussed in Ref. [15] for simple prepare-and-measure scenarios. Here we consider this problem in the context of more general networks.

Before moving on, it is important to understand why an assumption on the dimension is necessary in order to make the problem non-trivial. Consider for instance the network of Fig. 2. If the dimension is not limited, then the input settings of the preparation and transformation devices, x and t, can be perfectly transmitted to the final measurement device. Since the transformation device has all information about x and t, and the measuring device has all information about x, t, y, it follows that any possible statistics $p(b,s|x,t,y)$ can be reproduced. This implies that nontrivial bounds can only be placed if $|c_0| < |x|$ and/or $|c_1| < |x||t|$.

8.4.1 Non-classicality tests based on dimension witnesses

Considering systems of a fixed dimension, quantum communication can outperform classical communication. This advantage can be revealed by using dimension witnesses. Specifically, by using a well-chosen quantum strategy

involving states of Hilbert space dimension d, it is possible to violate certain classical dimension witnesses of dimension d. More formally, we say that a dimension witness with the following property

$$W = \mathbf{w} \cdot \mathbf{p} \leq C_d < Q_d \tag{8.14}$$

can be used as non-classicality tests for systems of dimension d. Consider a set of data $\mathbf{p}_Q$ such that $W = \mathbf{w} \cdot \mathbf{p}_Q > C_d$. This implies the use of genuinely quantum systems for reproducing $\mathbf{p}_Q$, under the assumption that the experiment involves systems of dimension d. In Section 8.5, we will discuss several examples.

8.4.2 Quantifying quantum advantage

It is useful to quantify the advantage offered by quantum resources over classical ones. In the present context, several figures of merit can be considered. First, the amount of violation of a given dimension witness could be used, however this will generally depend on how the witness is expressed, and will not allow one to compare different witnesses. Hence, here we use the notion of noise tolerance, which has a more physical interpretation, and will allow us to compare various witnesses.

Consider a quantum experiment (with systems of dimension d) and its corresponding set of data $\mathbf{p}_Q$, which is found to violate a classical dimension witness, i.e. $W = \mathbf{w} \cdot \mathbf{p}_Q > C_d$. The noise tolerance of the quantum point $\mathbf{p}_Q$ for this dimension witness is defined as the minimal fraction of white noise, η, such that the distribution

$$\mathbf{p}_0 = (1 - \eta)\mathbf{p}_Q + \eta\mathbf{p}_{\mathbb{1}} \tag{8.15}$$

does not violate the witness, i.e. $W = \mathbf{w} \cdot \mathbf{p}_0 = C_d$. Here $\mathbf{p}_{\mathbb{1}}$ denotes white noise, i.e. $p_{\mathbb{1}}(b, s|x, t, y) = \frac{1}{|b|\,|s|}$ is the uniform distribution for all x, t, y. In a practical context, considering noisy distributions of the form (8.15) is quite natural, due to unavoidable technical imperfections, e.g. losses or misalignment of the preparations.

8.4.3 Bounded noise tolerance in prepare-and-measure scenarios involving qubits

It turns out that the noise tolerance of qubit strategies is bounded for any dimension witness in the prepare-and-measure scenario. More precisely, any set

of data obtained from qubits and projective measurements can be reproduced using one classical bit if the noise level η satisfies

$$\eta \geq \eta^* = 1 - \frac{1}{k_3} \approx 0.34, \tag{8.16}$$

where k_3 is the Grothendieck constant [85] of order three[1]. Hence, in the prepare-and-measure scenario, no dimension witness for classical bits and projective measurements can be violated for $\eta \geq \eta^*$. This can be proven using a known LHV model for projective measurements for the Werner state of dimension 2 for weight $\frac{1}{k_3}$ [31]. For a proof of the above see Paper G Section V.

As mentioned, the above result holds only if the measurement device performs a projective measurement. Since any two outcome qubit measurement can be written as a convex mixture of projective measurements, the result can be extended to all two outcome scenarios. One can extend further to general positive operator-valued measurements at the cost of a larger η^* by using Werner's model [13] for the state (2.17) with $\alpha = \frac{1}{2}$, leading to $\eta^* = \frac{1}{2}$.

8.5 Case studies

We now present several case studies, illustrating the relevance of the concepts and tools discussed above. We first discuss two examples of networks where preparation, transformation, and measurement devices are 'in a line'. We then discuss two examples based on a different network, featuring two separate preparations devices and one measurement device. Note that such a network has been considered in different contexts. Notably, this was studied in communication complexity, in the so-called simultaneous message passing model [86], e.g. quantum fingerprinting [87], but also for the black-box certification of entangled measurements [88–90], and the Pusey-Barrett-Rudolph theorem [91].

In all cases quantum systems are shown to provide significant advantage over classical systems of the same dimension. Moreover, in all examples (except for the third one), this quantum advantage is stronger compared to the simple prepare-and-measure scenario, in terms of noise tolerance. This suggests that the simulation of quantum strategies becomes significantly harder in the case of networks, even if they feature only few devices.

[1]Note that only upper and lower bounds are known for k_3; see e.g. T. Vértesi, Phys. Rev. A **78**, 032112 (2008).

8.5.1 Three devices in a line: simple case

We start with the network of Fig. 8.2, considering one of the simplest (non-trivial) configurations in terms of the number of inputs and outputs. Specifically, we have $|x| = 3$ and $|t| = |y| = |b| = 2$. Note that the transformation device does not give any outcome (i.e. $|s| = 1$). We label the inputs and outputs: $x \in \{0, 1, 2\}$ and $t, y, b \in \{0, 1\}$. Hence a set of data is characterised by $D = 24$ probabilities $p(b|x, t, y)$. However, considering normalisation conditions, this number is reduced to 12; specifically, the probabilities $p(1|x, t, y) = 1 - p(0|x, t, y)$ are redundant and can thus be omitted.

Applying the method described in Section 8.2.2 we have fully characterised the polytope $\mathbb{P}_2$, that is, the set of distributions achievable for $c_0, c_1 \in \{0, 1\}$. Using the software *PORTA*, we could find the complete list of facets of $\mathbb{P}_2$, which can be grouped (under relabeling of inputs and outputs) into 1870 inequivalent classes of dimension witnesses[2].

Here, we present one class of tight dimension witnesses, a member of which can be written in simple form:

$$W_J = p_{011} + p_{101} + p_{110} + p_{200} - p_{000} - p_{001} - p_{010} - p_{211} \leq 2, \tag{8.17}$$

where we write $p_{xty} = p(b = 0|x, t, y)$. Using qubits we can significantly outperform this bound. Consider general pure qubit preparations:

$$|\psi(\theta, \phi)\rangle = \cos(\frac{\theta}{2})|0\rangle + \sin(\frac{\theta}{2}) \exp(i\phi)|1\rangle. \tag{8.18}$$

Specifically, for preparations $x = 0, 1, 2$ take $|\psi(\frac{\pi}{2}, 0)\rangle$, $|\psi(\frac{\pi}{2}, \frac{3\pi}{4})\rangle$ and $|\psi(\frac{\pi}{2}, \frac{-3\pi}{4})\rangle$ respectively. Next consider the transformation device, parametrized by

$$\Phi_{t=0} = \mathbb{1}_2 \quad , \quad \Phi_{t=1} = \exp(-i\frac{\pi}{4}\sigma_z), \tag{8.19}$$

where $\sigma_z = \text{diag}(1, -1)$ is the Pauli z matrix. Finally, for the measuring device, we have the measurement operators

$$M_{0|0} = |\psi(\frac{\pi}{2}, \frac{-3\pi}{4})\rangle\langle\psi(\frac{\pi}{2}, \frac{-3\pi}{4})| \; ; \quad M_{0|1} = |\psi(\frac{\pi}{2}, \frac{3\pi}{4})\rangle\langle\psi(\frac{\pi}{2}, \frac{3\pi}{4})|. \tag{8.20}$$

Calculating the resulting probabilities, via Eq. (8.11), and inserting them into (8.17), we obtain

$$W_J = 2 + \sqrt{2} \approx 3.41. \tag{8.21}$$

[2]For the full list of inequalities, contact *joseph.bowles@unige.ch*

The above qubit strategy thus clearly violates the witness (8.17), and can therefore not be reproduced with classical bits; classical trits must be used. Numerical optimization strongly suggests that this qubit strategy is optimal.

The noise tolerance of the above qubit strategy is

$$\eta = \sqrt{2} - 1 \approx 0.41. \tag{8.22}$$

Notably, this value exceeds the bound $\eta^* \approx 0.34$ (see Section 8.4.2) for any prepare-and-measure scenario. Hence the advantage offered by qubits compared to classical bits is stronger compared to any witness in the prepare-and-measure scenario.

8.5.2 Distributed $3 \to 1$ random access code

As a second example, we consider a task inspired from the information-theoretic task of a random access code (RAC) [92]. Specifically, we consider a distributed version of the $3 \to 1$ RAC featuring three devices in a line. Consider 3 bits a_0, a_1, a_2 randomly taken from a uniform distribution. These bits will determine the inputs of the preparation and transformation devices, namely: $x = (a_0, a_1)$ and $t = a_0 \oplus a_2$. Again, the transformation device has no output. The measuring devices has a ternary input $y = 0, 1, 2$. Similarly to a RAC, the goal is to have the output $b = a_y$. Hence we can define the following witness (for the scenario $|x| = 4$, $|t| = |b| = 2$, $|y| = 3$, and $|s| = 1$) which is the average success probability:

$$W_{\text{D-RAC}} = \frac{1}{24} \sum_{\substack{a_0 a_1 \\ a_2 y}} p(b = a_y | x = (a_0, a_1), t = (a_0 \oplus a_2), y) \leq C_d. \tag{8.23}$$

We first discuss the case of classical communication. For bits we obtain the bound $C_2 = \frac{2}{3}$. For the case of classical trits, $c_0, c_1 \in \{0, 1, 2\}$, we get $C_3 = 19/24$. In order to achieve success with probability one, i.e. $W_{\text{D-RAC}} = 1$, eight-dimensional systems are required. Using qubits, we can achieve up to

$$W_{\text{D-RAC}} = Q_2 = \frac{1}{2}(1 + \frac{1}{\sqrt{3}}) \approx 0.79. \tag{8.24}$$

For the quantum strategy which achieves this see Paper E. The noise tolerance of this strategy is given by

$$\eta = 1 - \frac{1}{\sqrt{3}} \approx 0.43 \tag{8.25}$$

which again exceeds the bound for the prepare-and-measure scenario, $\eta^* \approx 0.34$.

8.5.3 Two preparation devices, one measurement device: simple case

Finally we consider a scenario with two preparation devices sending communication to a measurement device (see Fig. 8.3 (a)). A simple non-trival scenario here is one in which both preparation devices receive a ternary input. We denote the input of the first device $x_0 \in \{0, 1, 2\}$, and the input of the second $x_1 \in \{0, 1, 2\}$. The measurement device has no input (i.e. a fixed measurement) and provides a binary output $b = \{0, 1\}$. That is, we have $|x_0| = |x_1| = 3$, $|y| = 1$ and $|b| = 2$.

We consider the case in which the channels carry classical bits, i.e. $c_0, c_1 \in \{0, 1\}$. In this case we have fully characterised the polytope $\mathbb{P}_2$: it features 13 non-trivial classes of facets (see Paper E Appendix). Here we focus on one particular class (witness 1 in Appendix), represented by the following witness:

$$W_K = -p_{00} + p_{01} + p_{02} - p_{10} - p_{12} + p_{20} + p_{21} - p_{22} \leq 2.$$

With qubit strategies one can achieve a value of $\frac{5}{2}$ with a corresponding noise tolerance $\eta = 0.2$ (See Paper E).

One can derive an upper bound on W_K for separable measurement operators of the form $M_b = \sum_i M^i_{b,1} \otimes M^i_{b,2}$ where $M^i_{b,k}$ is a positive operator acting on the system sent by preparation device k. Numerical tests suggest that the optimal value is $W_K \approx 2.337$. This suggests that W_K may also be used as a test of the non-separability of a set of measurement operators.

8.5.4 Nonlocal dense coding

As the last example, we present a dimension witness for a task which can be viewed as a nonlocal version of dense coding [93]. As in the previous example, we consider the case of two preparation devices and one measuring device.

Here each preparation device receives two input bits: $x_0 = (u_0, u_1)$ for the first and $x_1 = (v_0, v_1)$ for the second. The measurement device receives $y = 0, 1$ as input, and provides two output bits $\mathbf{b} = (b_0, b_1)$. The rules of the game are the following (see Fig. 8.3(b)). On the one hand, for $y = 0$, the outputs should satisfy $(b_0, b_1) = (u_0 \oplus v_0, u_1 \oplus v_1)$. On the other hand, for $y = 1$, the output bits should satisfy $(b_0, b_1) = (u_0 \oplus v_1, u_1 \oplus v_0)$. Furthermore, there is a penalty if both b_0 and b_1 are guessed incorrectly. This corresponds to the witness

$$\begin{aligned} W_D =& \langle (b_0, b_1) = (u_0 \oplus v_0\bar{y} \oplus v_1 y, u_1 \oplus v_1\bar{y} \oplus v_0 y)\rangle \\ & - \langle (\bar{b}_0, \bar{b}_1) = (u_0 \oplus v_0\bar{y} \oplus v_1 y, u_1 \oplus v_1\bar{y} \oplus v_0 y)\rangle \leq C_d, \end{aligned} \tag{8.26}$$

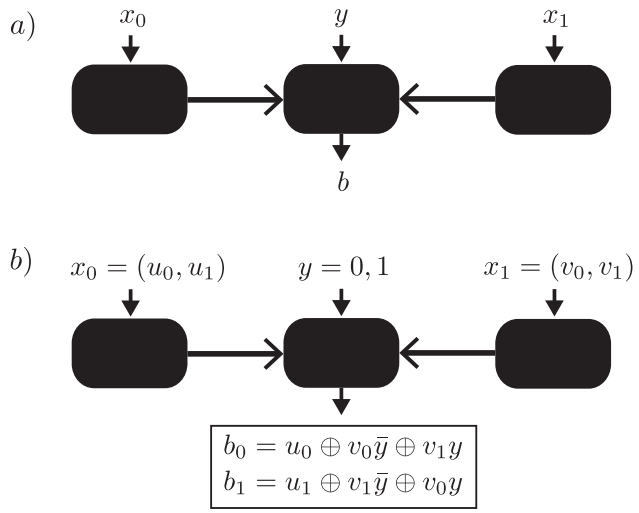

Figure 8.3: (a) A simple network involving two preparation devices (left and right) and a single measurement device (center). (b) A dimension witness for this network, referred to as nonlocal dense coding.

where $\bar{y} = y \oplus 1$, and the average $\langle\cdot\rangle$ is taken over all inputs:

$$\langle(b_0, b_1)\rangle = \frac{1}{32} \sum_{\substack{u_0,u_1 \\ v_0,v_1,y}} p(b_0, b_1|u_0, u_1, v_0, v_1, y). \tag{8.27}$$

For classical bits, we have $C_2 = \frac{1}{4}$. Using classical trits, we get $C_3 = \frac{9}{16}$. Sending four dimensional systems achieves success probability one. Considering qubit strategies (see Paper E Appendix) we can achieve

$$W_D = Q_2 = \frac{1}{2} \tag{8.28}$$

which appears optimal from numerical tests. This corresponds to a noise tolerance of $\eta = \frac{1}{2}$, which represents a considerable improvement over the simple prepare-and-measure scenario. Additionally, one may also wish to consider the possibility that the devices share quantum correlations (i.e. initial entanglement). Allowing for this considerably enhances the success probability (still using qubit communication), which becomes maximal, that is $W_D = 1$. The strategy is the following. The preparation devices now share a singlet state. Upon receiving the inputs $x_0 = (u_0, u_1)$ and $x_1 = (v_0, v_1)$, the preparation devices locally rotate the singlet state to

$$(\sigma_x^{u_1}\sigma_z^{u_0}) \otimes (\sigma_x^{v_1}\sigma_z^{v_0})|\psi^-\rangle. \tag{8.29}$$

The measurement device then performs a projective measurement onto the entangled basis

$$M_{b_0 b_1|y} = \sigma_x^{b_1} \sigma_z^{b_0} \otimes H^y |\psi^-\rangle, \tag{8.30}$$

where $|\psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$ is the singlet state and $H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ is the Hadamard matrix. The noise tolerance for this strategy is $\eta = \frac{3}{4}$.

8.6 Outlook

Most of the work on dimension witnesses (and related work on communication complexity) has so far focused on the prepare-and-measure scenario [83,84,94–97]. One notable exception is that of [98], which looks are a sequence of N devices in a line and shows that in the limit $N \to \infty$ one requires classical systems of infinite dimension in order to simulate the communication of a qubit. Given that we have seen a significant improvement over the prepare-and-measure scenario by adding a third device, it would be interesting to investigate whether qubit communication requires infinite classical dimension in a network featuring a finite number of devices, with continuous inputs. For example, how much communication is needed to simulate the preparation of an arbitrary qubit, arbitrary unitary, followed by arbitrary measurement in the line network of Section 8.5.1?

Going to more complex networks, the computational techniques used here to find dimension witnesses will no longer be tractable. It is therefore desirable to develop methods to construct dimension witnesses for networks

The techniques of dimension witnesses can also be used to construct quantum randomness certification [99, 100] as well as quantum key distribution protocols [101], where one makes an assumption of an upper bound of the dimension of the communicated systems (see also the following chapter). It is likely that the witnesses presented in this section could also be of use here, and would likely have higher tolerances to experimental errors.

Chapter 9

Certification of dimension and randomness using independent devices

In the previous chapter we presented methods to certify device independently the dimension of physical systems placed in a network of devices. Here, we consider a similar problem in which the devices are independent, meaning that they do not have access to shared classical randomness. This renders the problem somewhat difficult mathematically, since it introduces non-convexity into the sets of probability distributions obtainable using a given dimension. Nevertheless, we find families of (nonlinear) dimension witnesses that characterise these sets [**Paper H**] (see also [102]). The performance of classical vs quantum systems of the same dimension was also discussed in the previous chapter. Here, we will see that under the assumption of independence, quantum systems vastly outperform classical systems of the same dimension and allow for arbitrary tolerance to noise.

To tools of dimension certification have also recently been applied to the certification of random numbers from quantum systems [99, 100]. We present random number certification protocols based on our tests of dimension under the assumption that we have independent preparation and measurement devices [**Paper J**]. Unlike previous protocols, the additional assumption of independence allows for extremely high tolerance to experimental noise, which has made possible a corresponding experimental implementation [**Paper J**].

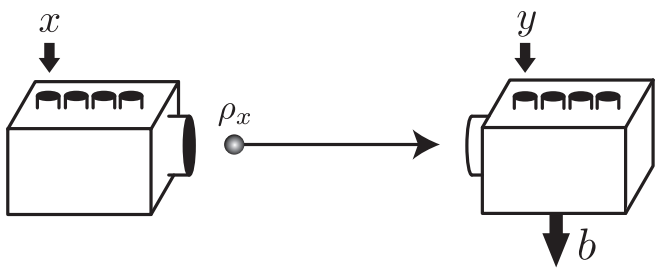

Figure 9.1: The prepare-and-measure scenario. Alice receives a classical input x and sends the d-dimensional state ρ_x to Bob. Bob, upon receiving his classical input y and the sent state, outputs b. The experiment is then characterised by the probability distribution $p(b|xy)$.

9.1 The prepare-and-measure scenario

We consider the simplest possible scenario of a single preparation and measurement device (see Fig. 9.1), which we call a prepare-and-measure scenario. Alice receives an input x and sends a system of dimension d to Bob. Bob then receives his input y and outputs b by performing a measurement on the system sent by Alice. We hence obtain a set of distributions $p(b|xy)$ from which we would like to conclude something about the dimension of sent system.

9.1.1 Classical systems

We first consider deterministic strategies. Alice receives x, sends a classical message $m = 1, \cdots d$ to Bob, who outputs b given y and m. Summing over the possible messages we have

$$p(b|xy) = \sum_{m=1}^{d} p(m|x)p(b|ym) \tag{9.1}$$

where $p(m|x), p(b|ym) = 0, 1$ are deterministic functions. We can then introduce local variables $\lambda \in \Lambda$ and $\mu \in M$ for Alice/Bob with joint distribution $Q_{\lambda,\mu}$ to allow for mixed strategies:

$$p(b|xy) = \int_{\Lambda}\int_{M} Q_{\lambda,\mu} \sum_{m=1}^{d} p_\lambda(m|x)p_\mu(b|ym)\mathrm{d}\mu\mathrm{d}\lambda. \tag{9.2}$$

Here, we will be interested in independent devices. This means that the variables λ, μ should be uncorrelated $Q_{\lambda,\mu} = q_\lambda r_\mu$. Integrating the above over λ,

μ one then obtains

$$p(b|xy) = \sum_{m=1}^{d} \int_{\Lambda} q_\lambda p_\lambda(m|x) \mathrm{d}\lambda \int_M r_\mu p_\mu(b|ym) \mathrm{d}\mu \tag{9.3}$$

$$= \sum_{m=1}^{d} p(m|x) p(b|ym). \tag{9.4}$$

That is, we have the same form as (9.1), however $p(m|x)$, $p(b|ym)$ can now be arbitrary probability distributions. Note that since in general one has $p\, q_{\lambda_1} r_{\mu_1} + (1-p) q_{\lambda_2} r_{\mu_2} \neq q_{\lambda_3} r_{\mu_3}$ the set of distributions given by (9.3) is non-convex.

9.1.2 Quantum systems

In the case of independent quantum systems Alice prepares a state ρ_x and Bob measures according to measurement operators $M_{b|y}$. We therefore have

$$p(b|xy) = \mathrm{Tr}\left[\rho_x M_{b|y}\right], \tag{9.5}$$

where ρ and $M_{b|y}$ act on $\mathbb{C}^d$.

9.2 Classical dimension witnesses for dimension 2

9.2.1 BB84 witness

We first focus of the case of dimension 2 with four preparations $x = 0, ..., 3$ and two measurements $y = 0, 1$. Consider the following matrix

$$\mathbf{W}_2 = \begin{pmatrix} p(0,0) - p(1,0) & p(2,0) - p(3,0) \\ p(0,1) - p(1,1) & p(2,1) - p(3,1) \end{pmatrix} \tag{9.6}$$

where we write $p(x,y) = p(b=0|x,y)$ for simplicity. For any strategy involving a classical bit (i.e. its statistics admits a decomposition of the form (9.3) with $d = 2$), one has that

$$W_2 = |\det(\mathbf{W}_2)| = 0. \tag{9.7}$$

The proof is straightforward. Note that for any statistics of the form (9.3) with $d = 2$, we have that using $p(1|x) = 1 - p(0|x)$

$$p(x,y) = p(0|x)[p(0|0,y) - p(0|1,y)] + p(0|1,y). \tag{9.8}$$

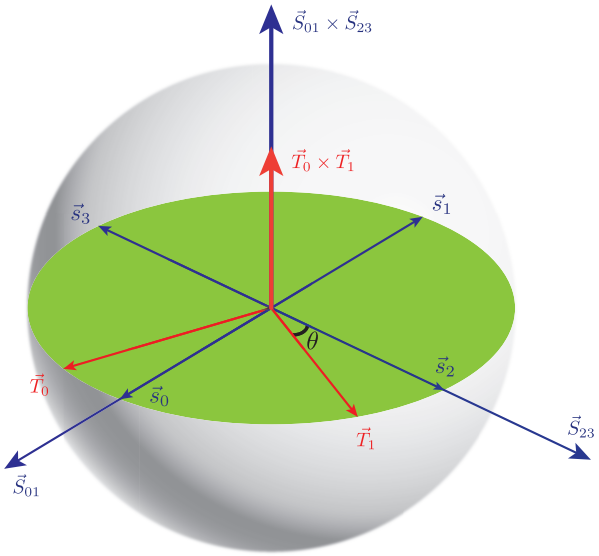

Figure 9.2: Qubit strategy that achieves $|\det \mathbf{W}_2| = 1$. Note that due to the rotational invariance (in the plane) of the vector cross product, the angle θ can be chosen freely.

Hence we write

$$p(x,y) - p(x',y) = [s(0|x) - s(0|x')][t(0|0,y) - t(0|1,y)] = S_{xx'}T_y \tag{9.9}$$

from which it follows that

$$W_2 = \begin{vmatrix} S_{01}T_0 & S_{23}T_0 \\ S_{01}T_1 & S_{23}T_1 \end{vmatrix} = 0. \tag{9.10}$$

Note that due to the non-convexity of the set of classical distributions using a classical bit, the witness is nonlinear (quadratic) in the observed probabilities $p(b|xy)$. In fact, this witness turns out to characterise fully the set of experiments involving a classical bit. Specifically, for any experiment achieving $W_2 = 0$ (for all relabelings of the preparation x), there exists a decomposition of the form (9.3) with $d = 2$ (see Paper F Appendix A).

Next we consider the performance of qubit strategies, i.e. statistics of the form (9.5) with $d = 2$. States are given by density matrices $\rho_x = (\mathbb{1}_2 + \vec{s}_x \cdot \vec{\sigma})/2$ and measurement operators by $M_{0|y} = c_y \mathbb{1}_2 + \vec{T}_y \cdot \vec{\sigma}/2$, where $\vec{s}_x$ and $\vec{T}_y$ are Bloch vectors and $|\vec{T}_y| \leq |c_y| \leq 1 - |\vec{T}_y|/2$. Similarly to above, we write

$$p(x,y) - p(x',y) = \mathrm{Tr}[(\rho_x - \rho_{x'})M_{0|y}] = \vec{S}_{xx'} \cdot \vec{T}_y \tag{9.11}$$

where $\vec{S}_{xx'} = (\vec{s}_x - \vec{s}_{x'})/2$. Finally, we get

$$W_2 = \begin{vmatrix} \vec{S}_{01} \cdot \vec{T}_0 & \vec{S}_{23} \cdot \vec{T}_0 \\ \vec{S}_{01} \cdot \vec{T}_1 & \vec{S}_{23} \cdot \vec{T}_1 \end{vmatrix} = (\vec{S}_{01} \times \vec{S}_{23}) \cdot (\vec{T}_0 \times \vec{T}_1) \leq 1 \tag{9.12}$$

since $|\vec{S}_{01} \times \vec{S}_{23}| \leq 1$ and $|\vec{T}_0 \times \vec{T}_1| \leq 1$. This bound for qubit strategies is tight, and can be reached as follows: choose the preparations to be the pure qubit states given by $\vec{s}_0 = -\vec{s}_1 = \hat{z}$, $\vec{s}_2 = -\vec{s}_3 = \hat{x}$, and the measurements by the vectors $\vec{T}_0 = \cos\theta\hat{z} + \sin\theta\hat{x}$ and $\vec{T}_1 = \sin\theta\hat{z} - \cos\theta\hat{x}$. Notice that we are free to choose any angle θ here, due to the rotational invariance of the cross product in the plane. For $\theta = 0$ we get the usual BB84 states and measurements. The value of W_2 for qubit strategies has a simple geometrical interpretation presented in Fig 9.2.

9.2.2 Tetrahedral witness

We now consider a second witness which appears to be maximally violated by choosing preparations that form a tetrahedron in the Bloch sphere and measurements of $\sigma_x, \sigma_y, \sigma_z$. We therefore have 4 preparations $x = 0, 1, 2, 3$ and 3 measurements $y = 0, 1, 2$. Define the matrix

$$\mathbf{W}_{\text{tet}} = \begin{pmatrix} p(0,0) - p(1,0) & p(1,0) - p(2,0) & p(2,0) - p(3,0) \\ p(0,1) - p(1,1) & p(1,1) - p(2,1) & p(2,1) - p(3,1) \\ p(0,2) - p(1,2) & p(1,2) - p(2,2) & p(2,2) - p(3,2) \end{pmatrix}. \tag{9.13}$$

The witness is then again determined by taking the determinant $W_{\text{tet}} = |\det \mathbf{W}_{\text{tet}}|$. Converting this to Bloch vector notation we have

$$W_{\text{tet}} = \begin{vmatrix} \vec{S}_{01} \cdot \vec{T}_0 & \vec{S}_{12} \cdot \vec{T}_0 & \vec{S}_{23} \cdot \vec{T}_0 \\ \vec{S}_{01} \cdot \vec{T}_1 & \vec{S}_{12} \cdot \vec{T}_1 & \vec{S}_{23} \cdot \vec{T}_1 \\ \vec{S}_{01} \cdot \vec{T}_2 & \vec{S}_{12} \cdot \vec{T}_2 & \vec{S}_{23} \cdot \vec{T}_2 \end{vmatrix} \tag{9.14}$$

This can be expressed using the generalised cross product defined as follows. The cross product $\vec{S}_0 \times \vec{S}_1 \times \cdots \times \vec{S}_{k-1}$ of k vectors in $\mathbb{R}^{k+1}$ is defined as the unique vector $\vec{u} \in \mathbb{R}^{k+1}$ such that $\vec{V} \cdot \vec{u} = \det(V, \vec{S}_0, \vec{S}_1, \cdots, \vec{S}_{k-1})$ for all $\vec{V} \in \mathbb{R}^{k+1}$ (see e.g. [103]). We may therefore rewrite the above as

$$W_{\text{tet}} = (\vec{S}_{01} \times \vec{S}_{12} \times \vec{S}_{23}) \cdot (\vec{T}_0 \times \vec{T}_1 \times \vec{T}_2). \tag{9.15}$$

Here, one must extend all vectors so that they are defined in $\mathbb{R}^4$ so that the cross product is properly defined. Numerical evidence strongly suggests that this is maximised by taking preparations $\vec{s}_x$ that form a tetrahedron in the Bloch sphere and measurements in the directions x, y, z. This strategy scores $W_{\text{tet}} = \frac{2}{3\sqrt{3}}$.

9.3 Tolerance to noise

In the previous section ((8.4.3)), we saw that assuming the devices have access to shared classical randomness, the maximum tolerance to white noise in the prepare-and-measure scenario is $\eta^* = 0.34$ for projective measurements and $\eta^* = 0.5$ for POVM measurements (although smaller bounds may be achievable). Note that if one assumes the devices are independent as above, essentially any qubit strategy scores $|W_2| > 0$, suggesting a high tolerance to noise of qubits compared to classical bits. Indeed, one finds that if one starts with a distribution $p_Q(x, y)$ which achieves $|W_2| = Q$, then the noisy strategy

$$p(x, y) = (1 - \eta)p_Q + \eta\, p_N(y) \tag{9.16}$$

achieves $|W_2| = (1 - \eta)^2 Q$. Here, p_N is any noise distribution which depends upon y only (the same result also holds for the witness W_{tet} for the same reasons). Hence, one may tolerate an arbitrary amount of noise, such as low detector efficiencies or background noise. This is in stark contrast to the case where the devices share classical correlations where typically high efficiencies are required [102].

The above witnesses also display interesting tolerance to misalignment errors in the preparations/measurements. Note that due to the rotational invariance (in the plane) of the cross product, the value of W_2 is unaffected by applying any unitary rotation (in the relevant plane) to either the preparations of measurements. For example, this corresponds to the freedom of choosing θ in Fig. 9.2. For the witnesss W_{tet} this is even more extreme. Since both cross products appearing in (9.15) are by definition orthogonal to the subspaces defined by $\{\vec{s}_x\}$ and $\{\vec{m}_y\}$ (i.e. they both point into the 4^{th} dimension), one may apply an arbitrary unitary to all preparations and/or measurements without affecting the value of the witness. This is of practical interest since there is therefore no need to align the preparation and measurement devices.

9.4 Dimension witnesses for all dimensions

We now generalise the above witness for testing classical and quantum systems of arbitrary dimension (we again consider binary outcomes). Consider a scenario with $2k$ preparations and k binary measurements. Construct the $k \times k$ matrix

$$\mathbf{W}_k(i, j) = p(2j, i) - p(2j + 1, i) \tag{9.17}$$

with $0 \leq i, j \leq k-1$. As above, the witness is given by $W_k = |\det(\mathbf{W}_k)|$. For classical systems of dimension d, one has that

$$W_k = 0 \quad \text{for } d \leq k, \tag{9.18}$$

while one can have $W_k \geq 1$ for $d > k$. For quantum systems of dimension d, we get

$$W_k = 0 \quad \text{for } d \leq \sqrt{k}, \tag{9.19}$$

while $W_k > 0$ is possible whenever $d > \sqrt{k}$. Hence we obtain a quadratic separation between classical and quantum dimensions, using a number of preparations and measurements that grows only linearly with the dimension. The origin of the quadratic separation arises from the fact that quantum systems require quadratically many more real parameters ($d^2 - 1$ for Hilbert space dimension d) to parameterise them than their classical counterparts (which require only $d - 1$).

9.5 Semi-device independent randomness certification

The generation of high quality random numbers is an important problem that finds application in cryptography, simulation and gambling. However, essentially all methods used commercially to generate random numbers use classical processes (for example complex mathematical functions or electronic/atmospheric noise) and are hence fundamentally deterministic. This makes for a reliable estimate on the quality of such sources particularly troublesome.

The fact that the outcomes of measurements on quantum systems are inherently probabilistic has sparked a field of research focused on exploiting quantum systems for the generation and certification of random numbers [16–20]. In this section we apply the techniques of dimension witnesses to the certification of randomness in prepare-and-measure scenarios (see [Paper I]).

9.5.1 The semi-device independent scenario

One may easily generate perfect random numbers if one assumes perfect control and knowledge of the experimental setup: simply prepare for example the qubit state $|0\rangle$ and measure in the σ_x basis. The Born rule then tells us that the outcome of the measurement of σ_x

$$p(\pm|\sigma_x) = |\langle \pm|0\rangle|^2 = \frac{1}{2} \tag{9.20}$$

is uniform and random (here $|\pm\rangle$ denote the $\pm$ eigenstates of σ_x). In practice, however, one cannot prepare perfect eigenstates and the certification of randomness becomes more difficult. For example, if the preparation device malfunctions and instead prepares the state $|+\rangle$ with probability $\frac{1}{2}$ and $|-\rangle$ with probability $\frac{1}{2}$ (say due to some classical source of noise in the preparation device) then one will also find $p(\pm|\sigma_x) = \frac{1}{2}$. Hence, upon observing $p(\pm|\sigma_x) = \frac{1}{2}$ one cannot be sure if the corresponding bits produced by the experiment come from the intrinsic randomness present in quantum mechanics or from some classical source of noise present in the devices.

The aim of the semi-device independent scenario is to overcome this difficulty by relaxing some of the experimental requirements, so that randomness can still be certified if the experimenters do not have perfect control over the devices. Specifically we work in the prepare-and-measure scenario and assume the following:

- Settings x, y may be chosen independently of the devices. i.e. the devices do not know x, y prior to inputting.
- Independent, identically distributed rounds (i.i.d).
- Independent preparation and measurement devices (i.e. the devices do not share classical correlations).
- Qubit channel capacity: The information about the choice of preparation x retrieved by the measurement device (via a measurement on the sent particle) is contained in a 2-dimensional quantum subspace (a qubit).

Hence, one does not assume perfect knowledge of the functioning of the devices, but simply the above list of assumptions which must be justified by a particular experimental setup (see [Paper I] for an in depth discussion of this). The aim is now to certify quantum randomness under these assumptions, which we do with the aid of the witness W_2.

9.5.2 Randomness certification based on W_2

Defining again λ and μ to be local classical variables for the preparation and measurement devices, a general experiment satisfying the above can be written

$$p(b|xy) = \sum_{\lambda,\mu} q_\lambda r_\mu p_{\lambda\mu}(b|xy) \tag{9.21}$$

where $p_{\lambda\mu}(b|xy) = \mathrm{Tr}[\rho_x^\lambda M_{b|y}^\mu]$ and ρ_x, $M_{b|y}^\mu$ act on $\mathbb{C}^2$. We would like to ensure that the randomness produced in the experiment does not come from the

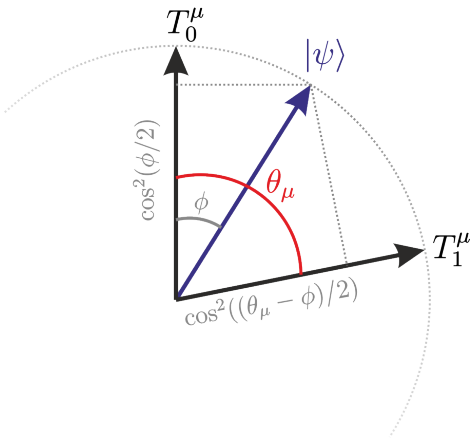

Figure 9.3: Measurement vectors are given by the vectors $\vec{T}_b^\mu$ and their corresponding probabilities given a state $|\psi\rangle$. In order to maximise the second line of (9.24) one should chose a state which lies in the middle of these two vectors. This gives guessing probability $(1+\cos(\theta_\mu/2))/2$.

classical random variables λ, μ. We therefore define the guessing probability for a given strategy and inputs x, y

$$p_g^{xy} = \sum_{\lambda,\mu} q_\lambda r_\mu \max_b p_{\lambda\mu}(b|xy). \tag{9.22}$$

If the probabilities $p_{\lambda\mu}$ are deterministic then one has $\max_b p_{\lambda\mu}(b|xy) = 1$ and any indeterminism in the statistics can be explained by the classical random variable λ, μ. The corresponding guessing probability is therefore 1. Any $p_g^{xy} < 1$ therefore indicates genuine quantum randomness.

Given some experimental data $p(b|xy)$, in order to certify randomness, we need to maximise this guessing probability over all possible strategies. Furthermore, we take the average over the uniformly chosen inputs $x = 0, \cdots, 3$, $y = 0, 1$. The maximal average guessing probability associated to the data $p(b|xy)$ is therefore

$$p_g^*[p(b|xy)] = \max_{q_\lambda, r_\mu, \rho_x^{\lambda,\mu} M_{b|y}^\mu} \frac{1}{8} \sum_{xy} p_g^{xy}, \tag{9.23}$$

such that (9.21) is satisfied. One may now use the witness W_2 given in (9.12) to certify randomness. We first fix a particular choice of the variables λ, μ.

We then have

$$p_g = \frac{1}{8}\sum_{x,y}\max_b p_{\lambda\mu}(b|x,y) \tag{9.24}$$

$$\leq \frac{1}{2}\max_x \sum_y \max_b p_{\lambda\mu}(b|x,y) \tag{9.25}$$

$$\leq \frac{1+\cos(\theta_\mu/2)}{2} \tag{9.26}$$

where θ_μ denotes the angle between Bob's two measurement. The reasoning of the derivation is as follows. The best guessing probability averaged over inputs of Alice is bounded by the maximum over her inputs. This gives the first inequality and allows us to focus on the best possible state that Alice can send. Next, Bob has two measurements described by Bloch vectors $\vec{T}^\mu_{0,1}$, and θ_μ is the angle between them. The best guessing probability averaged over his inputs is obtained by sending a state which lies in the middle between his measurements on the Bloch sphere (see Fig. 9.3). For such a state, the outcome probabilities for the two values of b are $\cos^2(\theta_\mu/4)$, and $\sin^2(\theta_\mu/4)$. Choosing the larger value and using the double-angle formula, one arrives at the second inequality.

Next, we note that from (9.12) for fixed λ, μ one has

$$W_2 \leq |\vec{T}^\mu_0 \times \vec{T}^\mu_1| \leq \sin\theta_\mu. \tag{9.27}$$

After some basic trigonometry, combining this with the above expression for p_g gives

$$p_g \leq \frac{1}{2}\left[1+\sqrt{\frac{1+\sqrt{1-W_2^2}}{2}}\right]. \tag{9.28}$$

Finally, to remove the assumption of fixed λ, μ, one can use concavity arguments to show that this bound holds in general, allowing for arbitrary distributions q_λ, r_μ of the classical random variables. We thus have

$$p_g^* \leq \frac{1}{2}\left[1+\sqrt{\frac{1+\sqrt{1-W_2^2}}{2}}\right]. \tag{9.29}$$

Finally, one must also take into account a small correction due to the effect of finite statistics (see Appendix [Paper I]). The amount of extractable randomness (in bits) is then given by the min entropy $-\log_2 p_g$, plotted in figure 9.4.

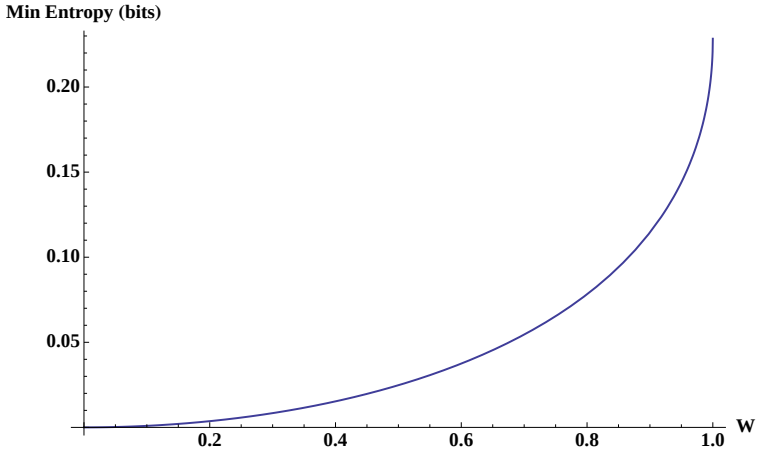

Figure 9.4: Min entropy $-\log_2(p_g)$ as a function of W_2, where p_g is given by (9.29). The actual min entropy of the source will be slightly smaller after finite size effects are taken into account.

The data can then be fed into a randomness extractor to produce a uniformly random seed given a short random seed.

As expected, the ability to extract randomness from W_2 is very robust to experimental noise. In [Paper I] an all optical experiment implementing the above randomness certification protocol is presented, generating 23 bits/sec of certified randomness.

9.6 Outlook

Other works have also investigated the problem of testing dimension using independent devices. Notably, in [102], the tolerance to detector inefficiencies in semi-device independent protocols based on dimension was discussed, where it was also found that extremely high tolerances are possible. Other protocols for randomness certification based on the assumption of independent devices have been developed in the semi-device-independent scenario [104–106]. Notably, [105] work in the same scenario as we do, however their methods have the disadvantage that they are based on numerics. Ref. [104] considers a scenario in which the measurement device is assumed to perform measurements on to a qubit subspace, which effectively fixes the preparation device to prepare qubit states. Their assumptions are thus stronger than those presented here, leading to higher rates. Further to this, quantum key distribution has also been considered, assuming independent devices in a prepare-and-measure

scenario [107].

There is still much work to be done in characterising the set of distributions obtainable from finite dimensional quantum systems. In the case where the devices are correlated, the problem is known to admit semi-definite programming relaxations [84], however this is unlikely to be the case for independent devices due to the non-convexity of the sets. It would be interesting to know, for example, if the witnesses $|W_k|$ give a complete characterisation, as was the case for W_2 (although this will probably need dimension witnesses featuring more than only binary outputs). Also, can one find nonlinear dimension witness which characterise the set of quantum, rather than classical distributions?

From a semi-device independent perspective, one could argue that the assumptions of independent devices and i.i.d. rounds are too strong. It would therefore be interesting to either remove the assumption of i.i.d. or develop sets of different assumptions which can be more easily justified in experimental setups.

Finally, our methods may be of use in other problems with non-convex geometry. One possibility is that of the nonlocality in networks of independent sources [108–111], where new techniques are much needed.

Chapter 10

Conclusion and Future Directions

The results presented in this thesis shed new light on fundamental questions about quantum correlations. From the perspective of Bell nonlocality, our work, as well as settling certain questions, naturally leads to new unanswered ones, many of which can be found in the outlook sections following each chapter. For example, questions regarding the precise relationship between entanglement and Bell nonlocality remain. Perhaps the holy grail in this respect would be a definitive answer as to whether entanglement always leads to correlations that are in disagreement with any local hidden variable description, and if so, what is the simplest scenario in which such a demonstration is possible? Whereas we have seen many examples in this thesis that this is not the case in the standard Bell scenario of non-sequential measurements on a single copy of the state, considering more general scenarios relatively little is known. If one extends the standard scenario to include hidden nonlocality (i.e. preprocessing of the state by local filters), then we have given considerable evidence that entanglement does not always lead to nonlocality (see Section 5.3.4). Whether this is also the case if one considers the complete set of probability distributions resulting from an arbitrary number of sequential measurements is still unknown. Moving to more complex scenarios, one could also consider taking multiple copies of the state and performing joint measurements, combining this with the filtering operations or sequential measurements or consider networks of states subject to independence conditions as is the case with N-locality [110, 112]. Constructing LHV models (if they exist) for such scenarios is understandably a formidable task, however the tools of EPR steering may again prove useful here.

We have also seen that it is possible to simulate entangled quantum states

with finite classical resources (Chapter 7). However, in our general construction, when one approaches the surface of the set of local states, an infinite amount of shared randomness is needed. It is therefore of interest to know whether this will always be the case, or on the contrary, if finite dimensional local quantum states can always be simulated with finite classical resources. Similarly, it would be interesting to get lower bounds on the dimension of the shared randomness needed to simulate a d-dimensional entangled states, and to know if this is strictly larger than that needed for separable states of the same dimension. Answers to either of these questions would provide insight into the connection between entanglement and nonlocality.

Moving to the certification of dimension in networks, much work is still to be done. The computational methods of Chapter 8 quickly become infeasible, and so developing more powerful tests in more complex scenarios will require new ideas. The simple examples presented here may provide the intuition from which these ideas may come. This is also the case for the scenario of independent devices of Chapter 9, where it is desirable to extend our witnesses to handle a larger number of outputs, as this is likely to be more relevant when considering higher dimensional systems than qubits. Another possible avenue of research here is the possibility of using such witnesses for the self-testing of certain ensembles of states and measurements. For example, the witnesses of Section 9.2 could potentially be used to self-test BB84 type setups and the preparation of the "tetrahedral" qubit states discussed therein. Progress here would also likely be useful for constructing more powerful protocols for random number certification.

Appendix A

Proof of the unsteerability of $\rho(p, \chi)$

Here we show that for the class of states (6.10), Theorem 1 implies that the $\rho(p, \chi)$ is unsteerable if

$$\cos^2 2\chi \geq \frac{2p-1}{(2-p)p^3}. \tag{A.1}$$

To do this, we first consider states in canonical form (5.3), which satisfy $\vec{a} = (0, 0, a_z)$ and $|T_x| = |T_y|$. In order to perform the maximisation of Theorem 1, we parameterize $\hat{x}$ using spherical co-ordinates $\hat{x} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. Our criterion may now be written as

$$\begin{aligned} &\max_{\theta,\phi} F(\theta, \phi) \leq 1\,, \qquad (A.2)\\ &F(\theta, \phi) = (\vec{a}\cdot\hat{x})^2 + 2||T\hat{x}|| \\ &\qquad = \cos^2\theta\ a_z^2 + 2\sqrt{T_x^2 + \cos^2\theta\ (T_z^2 - T_x^2)}. \end{aligned}$$

Unsurprisingly, F depends only on θ since the problem is symmetric with respect to the x and y directions and we may ignore the maximisation over ϕ. Note that if $|T_z| = |T_x|$ then the maximisation occurs at $\theta = 0$ and our condition for unsteerability becomes

$$a_z^2 + 2|T_z| \leq 1. \tag{A.3}$$

In the case $|T_z| \neq |T_x|$, one should find the extremal points of $F(\theta)$ and prove that they do not exceed 1. To find these extrema we solve

$$\frac{\mathrm{d}F}{\mathrm{d}\theta} = -\sin 2\theta \left(a_z^2 + \frac{T_z^2 - T_x^2}{\sqrt{T_x^2 + \cos^2\theta(T_z^2 - T_x^2)}} \right) = 0. \tag{A.4}$$

From $\sin 2\theta = 0$ we have solutions $\theta = 0, \pi/2, \pi$, and possibly other solutions given by

$$a_z^2 + \frac{T_z^2 - T_x^2}{\sqrt{T_x^2 + \cos^2\theta(T_z^2 - T_x^2)}} = 0. \tag{A.5}$$

We now derive conditions such that (A.5) has no solution. After rearranging (A.5) we have

$$\cos^2\theta = \frac{T_x^2}{T_x^2 - T_z^2} - \frac{T_x^2 - T_z^2}{a_z^4}. \tag{A.6}$$

This has no solution if the RHS is greater than 1 or less than 0. Hence we have two conditions

$$\frac{T_x^2}{T_x^2 - T_z^2} < \frac{T_x^2 - T_z^2}{a_z^4} \quad \text{or} \quad \frac{T_z^2}{T_x^2 - T_z^2} > \frac{T_x^2 - T_z^2}{a_z^4}. \tag{A.7}$$

If one of the above conditions is fulfilled we therefore have extrema for $\theta = 0, \pi/2, \pi$ only. In this case, and since $F(0) = F(\pi)$, our condition for unsteerability becomes

$$\max_\theta F(\theta) = \max\{\ a_z^2 + 2|T_z|\ ,\ 2|T_x|\ \} \leq 1. \tag{A.8}$$

We now move to the explicit case of $\rho(p, \chi)$. We find a canonical state with $|T_x| = |T_y|$, $\vec{a} = (0, 0, a_z)$ and

$$a_z = \frac{(1-p^2)\cos 2\chi}{1 - p^2\cos^2 2\chi}; \quad T_z = \frac{p(1-\cos^2 2\chi)}{1 - p^2\cos^2 2\chi}; \quad T_x = \sqrt{\frac{p^2(1-\cos^2 2\chi)}{1 - p^2\cos^2 2\chi}}. \tag{A.9}$$

We now introduce the ansatz (for $p \geq \frac{1}{2}$)

$$\cos^2 2\chi = \frac{2p-1}{(2-p)p^3}. \tag{A.10}$$

Eliminating the variable χ we find

$$a_z^2 = \frac{(2-p)(2p-1)}{p}; \quad T_z = \frac{(1-p)^2}{p}; \quad T_x = 1 - p. \tag{A.11}$$

For the case $p = \frac{1}{2}$ we have $|T_z| = |T_x|$ and one finds that (A.3) is satisfied. For $p > \frac{1}{2}$ we show that the second condition of (A.7) holds. To this end, we calculate

$$\frac{T_z^2}{T_x^2 - T_z^2} - \frac{T_x^2 - T_z^2}{a_z^4} = \frac{(3-p)(1-p)^3}{(p-2)^2(2p-1)}. \tag{A.12}$$

This is easily seen to be positive for $p \in]\frac{1}{2}, 1]$, and so $F(\theta)$ has extrema at $\theta = 0, \pi, \pi/2$ only. It therefore remains to prove (A.8). We find

$$a_z^2 + 2|T_z| = 1\,, \quad 2|T_x| = 2(1-p). \tag{A.13}$$

and so (A.8) is satisfied for $p > \frac{1}{2}$. This proves that the state $\rho(p,\chi)$ is unsteerable if $p \geq \frac{1}{2}$ and p and χ satisfy (A.10), which corresponds to the black curve of Fig. 5.22 in the main text. Finally, we note that for a fixed χ, lowering p amounts to putting more weight on the separable part of the state. Since a convex combination of an unsteerable state with a separable state is also unsteerable, all points below the curve of Fig. 5.22 are also unsteerable. Hence, we arrive at (5.22).

Bibliography

[1] M. Planck, "Zur Theorie des Gesetzes der Energieverteilung im Normalspektrum," *Verhandlungen der Deutschen Physikalischen Gesellschaft*, vol. 2, p. 237, 1900.

[2] A. Einstein, "Über einen die erzeugung und verwandlung des lichtes betreffenden heuristischen gesichtspunkt," *Annalen der Physik*, vol. 322, p. 132, Jul 1905.

[3] H. A. Kramers and W. Heisenberg, "Über die Streuung von Strahlung durch Atome, Zeitschrift für Physik," *Zeitschrift für Physik*, vol. 31, no. 1, pp. 681–708, 1925.

[4] E. Schrödinger, "An undulatory theory of the mechanics of atoms and molecules," *Physical Review*, vol. 28, no. 6, pp. 1049–1070, 1926.

[5] A. Einstein, B. Podolsky, and N. Rosen, "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?," *Physical Review*, vol. 47, no. 10, pp. 777–780, 1935.

[6] E. Schrödinger *Proc. Camb. Phil. Soc.*, vol. 31, p. 555, 1935.

[7] J. Bell, "On the Einstein Podolsky Rosen paradox," *Physics*, vol. 1, p. 195, 1964.

[8] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, "Proposed experiment to test local hidden-variable theories," *Phys. Rev. Lett.*, vol. 23, pp. 880–884, Oct 1969.

[9] A. Aspect, P. Grangier, and G. Roger, "Experimental realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A new violation of Bell's inequalities," *Physical Review Letters*, vol. 49, no. 2, pp. 91–94, 1982.

[10] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. ABellán, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson, "Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres," *Nature*, vol. 526, no. 7575, pp. 682–686, 2015.

[11] M. Giustina, M. A. M. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J. Å. Larsson, C. ABellán, W. Amaya, V. Pruneri, M. W. Mitchell, J. Beyer, T. Gerrits, A. E. Lita, L. K. Shalm, S. W. Nam, T. Scheidl, R. Ursin, B. Wittmann, and A. Zeilinger, "Significant-Loophole-Free Test of Bell's Theorem with Entangled Photons," *Physical Review Letters*, vol. 115, no. 25, 2015.

[12] L. K. Shalm, E. Meyer-Scott, B. G. Christensen, P. Bierhorst, M. A. Wayne, M. J. Stevens, T. Gerrits, S. Glancy, D. R. Hamel, M. S. Allman, K. J. Coakley, S. D. Dyer, C. Hodge, A. E. Lita, V. B. Verma, C. Lambrocco, E. Tortorici, A. L. Migdall, Y. Zhang, D. R. Kumor, W. H. Farr, F. Marsili, M. D. Shaw, J. A. Stern, C. ABellán, W. Amaya, V. Pruneri, T. Jennewein, M. W. Mitchell, P. G. Kwiat, J. C. Bienfang, R. P. Mirin, E. Knill, and S. W. Nam, "Strong Loophole-Free Test of Local Realism," *Physical Review Letters*, vol. 115, no. 25, 2015.

[13] R. F. Werner, "Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model," *Physical Review A*, vol. 40, no. 8, pp. 4277–4281, 1989.

[14] H. M. Wiseman, S. J. Jones, and A. C. Doherty, "Steering, entanglement, nonlocality, and the Einstein-Podolsky-Rosen paradox," *Physical Review Letters*, vol. 98, no. 14, 2007.

[15] R. Gallego, N. Brunner, C. Hadley, and A. Acin, "Device-independent tests of classical and quantum dimensions.," *Physical Review Letters*, vol. 105, no. 23, p. 230501, 2010.

[16] T. Jennewein, U. Achleitner, G. Weihs, H. Weinfurter, and A. Zeilinger, "A Fast and Compact Quantum Random Number Generator," *Review of Scientific Instruments*, vol. 71, no. 4, pp. 1675–1680, 2000.

[17] A. Stefanov, N. Gisin, O. Guinnard, L. Guinnard, and H. Zbinden, "Optical quantum random number generator," *Journal of Modern Optics*, vol. 47, no. 4, pp. 595–598, 2000.

[18] J. Rarity, P. Owens, and P. Tapster, "Quantum Random-number Generation and Key Sharing," *Journal of Modern Optics*, vol. 41, no. 12, pp. 2435–2444, 1994.

[19] R. Colbeck, "Quantum and relativistic protocols for secure multi-party computation," *arXiv:0911.3814*, 2009.

[20] S. Pironio, a. Acín, S. Massar, a. B. 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*, vol. 464, no. 7291, pp. 1021–1024, 2010.

[21] J. Barrett, D. Collins, L. Hardy, A. Kent, and S. Popescu, "Quantum nonlocality, Bell inequalities, and the memory loophole," *Phys. Rev. A*, vol. 66, p. 042111, Oct 2002.

[22] "Porta/panda." `http://www.iwr.uni-heidelberg.de/groups/comopt/`.

[23] "Lrs." `http://www.cgm.cs.mcgill.ca/avis/C/lrs.html`.

[24] I. Pitowsky, *Quantum Probability - Quantum Logic*, vol. 1 of *321*. Springer-Verlag Berlin Heidelberg, 1989.

[25] S. J. Freedman and J. F. Clauser, "Experimental test of local hidden-variable theories," *Phys. Rev. Lett.*, vol. 28, pp. 938–941, Apr 1972.

[26] A. Aspect, P. Grangier, and G. Roger, "Experimental tests of realistic local theories via Bell's theorem," *Phys. Rev. Lett.*, vol. 47, pp. 460–463, Aug 1981.

[27] M. Horodecki, P. Horodecki, and R. Horodecki, "Separability of mixed states: necessary and sufficient conditions," *Physics Letters A*, vol. 223, no. 1-2, pp. 1–8, 1996.

[28] J. Barrett, "Nonsequential positive-operator-valued measurements on entangled mixed states do not always violate a Bell inequality," *Phys. Rev. A*, vol. 65, p. 042302, Mar 2002.

[29] J. Bowles, T. Vértesi, M. T. Quintino, and N. Brunner, "One-way Einstein-Podolsky-Rosen steering," *Physical Review Letters*, vol. 112, no. 20, 2014.

[30] R. Augusiak, M. Demianowicz, and A. Acín, "Local hidden variable models for entangled quantum states," *Journal of Physics A: Mathematical and Theoretical*, vol. 47, no. 42, p. 424002, 2014.

[31] A. Acín, N. Gisin, and B. Toner, "Grothendieck's constant and local models for noisy entangled quantum states," *Physical Review A*, vol. 73, no. 6, 2006.

[32] M. L. Almeida, S. Pironio, J. Barrett, G. Tóth, and A. Acín, "Noise robustness of the nonlocality of entangled quantum states," *Physical Review Letters*, vol. 99, no. 4, 2007.

[33] M. D. Reid, "Demonstration of the Einstein-Podolsky-Rosen paradox using nondegenerate parametric amplification," *Physical Review A*, vol. 40, no. 2, pp. 913–923, 1989.

[34] M. D. Reid, P. D. Drummond, W. P. Bowen, E. G. Cavalcanti, P. K. Lam, H. A. Bachor, U. L. Andersen, and G. Leuchs, "Colloquium: The Einstein-Podolsky-Rosen paradox: From concepts to applications," *Reviews of Modern Physics*, vol. 81, no. 4, pp. 1727–1751, 2009.

[35] D. Cavalcanti and P. Skrzypczyk, "Quantum steering: a short review with focus on semidefinite programming," *arXiv:1604.00501*, 2016.

[36] M. T. Quintino, T. Vértesi, D. Cavalcanti, R. Augusiak, M. Demianowicz, A. Acín, and N. Brunner, "Inequivalence of entanglement, steering, and Bell nonlocality for general measurements," *Physical Review A - Atomic, Molecular, and Optical Physics*, vol. 92, no. 3, 2015.

[37] J. Bowles, F. Hirsch, M. T. Quintino, and N. Brunner, "Sufficient criterion for guaranteeing that a two-qubit state is unsteerable," *Phys. Rev. A*, vol. 93, p. 022121, Feb 2016.

[38] S. P. B. L. Vandenberghe, *Convex Optimization*, vol. 1. Cambridge University Press, 2004.

[39] P. Skrzypczyk, M. Navascues, and D. Cavalcanti, "Quantifying Einstein-Podolsky-Rosen steering," *Physical Review Letters*, vol. 112, no. 18, 2014.

[40] M. F. Pusey, "Negativity and steering: A stronger Peres conjecture," *Physical Review A - Atomic, Molecular, and Optical Physics*, vol. 88, no. 3, 2013.

[41] S. L. W. Midgley, A. J. Ferris, and M. K. Olsen, "Asymmetric gaussian steering: When Alice and Bob disagree," *Phys. Rev. A*, vol. 81, p. 022101, Feb 2010.

[42] M. K. Olsen, "Asymmetric gaussian harmonic steering in second-harmonic generation," *Phys. Rev. A*, vol. 88, p. 051802, Nov 2013.

[43] V. Händchen, T. Eberle, S. Steinlechner, A. Samblowski, T. Franz, R. F. Werner, and R. Schnabel, "Observation of one-way Einstein-Podolsky-Rosen steering," *Nature Photonics*, vol. 6, no. 9, pp. 598–601, 2012.

[44] K. Banaszek and K. Wódkiewicz, "Testing quantum nonlocality in phase space," *Phys. Rev. Lett.*, vol. 82, pp. 2009–2013, Mar 1999.

[45] S. Wollmann, N. Walk, A. J. Bennet, H. M. Wiseman, and G. J. Pryde, "Observation of genuine one-way Einstein-Podolsky-Rosen steering," *Phys. Rev. Lett.*, vol. 116, p. 160403, Apr 2016.

[46] K. Sun, X.-J. Ye, J.-S. Xu, X.-Y. Xu, J.-S. Tang, Y.-C. Wu, J.-L. Chen, C.-F. Li, and G.-C. Guo, "Experimental quantification of asymmetric Einstein-Podolsky-Rosen steering," *Phys. Rev. Lett.*, vol. 116, p. 160404, Apr 2016.

[47] D. Cavalcanti, P. Skrzypczyk, G. H. Aguilar, R. V. Nery, P. H. S. Ribeiro, and S. P. Walborn, "Detection of entanglement in asymmetric quantum networks and multipartite quantum steering," *Nat Commun*, vol. 6, 08 2015.

[48] S. Jevtic, M. J. W. Hall, M. R. Anderson, M. Zwierz, and H. M. Wiseman, "Einstein–Podolsky–Rosen steering and the steering ellipsoid," *J. Opt. Soc. Am. B*, vol. 32, pp. A40–A49, Apr 2015.

[49] F. Hirsch, M. T. Quintino, J. Bowles, and N. Brunner, "Genuine hidden quantum nonlocality," *Physical Review Letters*, vol. 111, no. 16, 2013.

[50] 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.*, vol. 116, p. 130401, Mar 2016.

[51] R.Horodecki and M. Horodecki, "Information-theoretic aspects of inseparability of mixed states," *Physical Review A*, vol. 54, no. 3, pp. 1838–1843, 1996.

[52] R. Uola, T. Moroder, and O. Gühne, "Joint Measurability of Generalized Measurements Implies Classicality," *Physical Review Letters*, vol. 113, no. 16, 2014.

[53] M. T. Quintino, T. Vértesi, and N. Brunner, "Joint measurability, Einstein-Podolsky-Rosen steering, and Bell nonlocality," *Physical Review Letters*, vol. 113, no. 16, 2014.

[54] T. Heinosaari, J. Kiukas, and D. Reitzner, "Noise robustness of the incompatibility of quantum measurements," *Physical Review A - Atomic, Molecular, and Optical Physics*, vol. 92, no. 2, 2015.

[55] P. L. P. Busch and P. Mittelstaedt, *The Quantum Theory of Measurement, Lecture Notes in Physics Monographs*, vol. 2. Springer-Verlag Berlin Heidelberg, 1996.

[56] R. Uola, C. Budroni, O. Gühne, and J. P. Pellonpää, "One-to-One Mapping between Steering and Joint Measurability Problems," *Physical Review Letters*, vol. 115, no. 23, 2015.

[57] R. Gallego, L. E. Wúrflinger, R. Chaves, A. Acín, and M. Navascués, "Nonlocality in sequential correlation scenarios," *New Journal of Physics*, vol. 16, 2014.

[58] S. Popescu, "Bell's inequalities and density matrices: Revealing "hidden" nonlocality," *Physical Review Letters*, vol. 74, no. 14, pp. 2619–2622, 1995.

[59] T. V. H. Chau Nguyen, "Necessary and sufficient condition for steerability of two-qubit states by the geometry of steering outcomes," *arXiv:1604.03815*, 2016.

[60] G. Tóth and A. Acín, "Genuine tripartite entangled states with a local hidden-variable model," *Physical Review A*, vol. 74, no. 3, 2006.

[61] R. Augusiak, M. Demianowicz, J. Tura, and A. Acín, "Entanglement and Nonlocality are Inequivalent for Any Number of Parties," *Physical Review Letters*, vol. 115, no. 3, 2015.

[62] G. Svetlichny, "Distinguishing three-body from two-body nonseparability by a Bell-type inequality," 1987.

[63] J. D. Bancal, J. Barrett, N. Gisin, and S. Pironio, "Definitions of multipartite nonlocality," *Physical Review A - Atomic, Molecular, and Optical Physics*, vol. 88, no. 1, 2013.

[64] R. Gallego, L. E. Würflinger, A. Acín, and M. Navascués, "Operational framework for nonlocality," *Phys. Rev. Lett.*, vol. 109, p. 070401, Aug 2012.

[65] M. Seevinck and G. Svetlichny, "Bell-type inequalities for partial separability in N-particle systems and quantum mechanical violations.," *Physical review letters*, vol. 89, no. 1, p. 060401, 2002.

[66] D. Collins, N. Gisin, S. Popescu, D. Roberts, and V. Scarani, "Bell-Type Inequalities to Detect True n-Body Nonseparability," *Physical Review Letters*, vol. 88, p. 170405, 2002.

[67] Z. H. Ma, Z. H. Chen, J. L. Chen, C. Spengler, A. Gabriel, and M. Huber, "Measure of genuine multipartite entanglement with computable lower bounds," *Physical Review A - Atomic, Molecular, and Optical Physics*, vol. 83, no. 6, 2011.

[68] M. Huber, F. Mintert, A. Gabriel, and B. C. Hiesmayr, "Detection of high-dimensional genuine multipartite entanglement of mixed states," *Physical Review Letters*, vol. 104, no. 21, 2010.

[69] O. Gühne and M. Seevinck, "Separability criteria for genuine multiparticle entanglement," *New Journal of Physics*, vol. 12, 2010.

[70] C. Eltschka and J. Siewert, "Quantifying entanglement resources," *Journal of Physics A: Mathematical and Theoretical*, vol. 47, p. 424005, 2014.

[71] K. Z. I. Bengtsson, *Geometry of Quantum States: An Introduction to Quantum Entanglement*. Cambridge University Press, 2006.

[72] P. Horodecki, "Separability criterion and inseparable mixed states with positive partial transposition," *Physics Letters A*, vol. 232, no. 5, pp. 333 – 339, 1997.

[73] C. Carathéodory, "Über den variabilitätsbereich der fourier'schen konstanten von positiven harmonischen funktionen," *Rendiconti del Circolo Matematico di Palermo (1884-1940)*, vol. 32, no. 1, pp. 193–217, 1911.

[74] W. K. Wootters, "Entanglement of formation of an arbitrary state of two qubits," *Phys. Rev. Lett.*, vol. 80, pp. 2245–2248, Mar 1998.

[75] C. Branciard and N. Gisin, "Quantifying the nonlocality of greenberger-horne-zeilinger quantum correlations by a bounded communication simulation protocol," *Physical Review Letters*, vol. 107, no. 2, 2011.

[76] B. F. Toner and D. Bacon, "Communication cost of simulating Bell correlations.," *Physical review letters*, vol. 91, no. 18, p. 187904, 2003.

[77] O. Regev and B. Toner, "Simulating quantum correlations with finite communication," in *Proceedings - Annual IEEE Symposium on Foundations of Computer Science, FOCS*, pp. 384–394, 2007.

[78] F. Hirsch, M. T. Quintino, T. Vértesi, M. F. Pusey, and N. Brunner, "Algorithmic construction of local hidden variable models for entangled quantum states," *arXiv:1512.00262*, 2016.

[79] D. Cavalcanti, L. Guerini, R. Rabelo, and P. Skrzypczyk, "General method for constructing local-hidden-variable models for multiqubit entangled states," *arXiv:1512.00277*, 2016.

[80] N. Brunner, S. Pironio, A. Acin, N. Gisin, A. A. Méthot, and V. Scarani, "Testing the dimension of hilbert spaces," *Phys. Rev. Lett.*, vol. 100, p. 210503, May 2008.

[81] T. Vértesi and K. F. Pál, "Bounding the dimension of bipartite quantum systems," *Phys. Rev. A*, vol. 79, p. 042106, Apr 2009.

[82] J. Briët, H. Buhrman, and B. Toner, "A generalized grothendieck inequality and nonlocal correlations that require high entanglement," *Communications in Mathematical Physics*, vol. 305, no. 3, pp. 827–843, 2011.

[83] N. Brunner, M. Navascués, and T. Vértesi, "Dimension witnesses and quantum state discrimination," *Phys. Rev. Lett.*, vol. 110, p. 150501, Apr 2013.

[84] M. Navascués and T. Vértesi, "Bounding the set of finite dimensional quantum correlations," *Phys. Rev. Lett.*, vol. 115, p. 020501, Jul 2015.

[85] A. Grothendieck, "Résumé de la théorie métrique des produits tensoriels topologiques," *Bol. Soc. Mat. Sao Paulo*, no. 8, pp. 1–79, 1953.

[86] H. Buhrman, R. Cleve, S. Massar, and R. de Wolf, "Nonlocality and communication complexity," *Rev. Mod. Phys.*, vol. 82, pp. 665–698, Mar 2010.

[87] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf, "Quantum fingerprinting," *Phys. Rev. Lett.*, vol. 87, p. 167902, Sep 2001.

[88] T. Vértesi and M. Navascués, "Certifying entangled measurements in known hilbert spaces," *Phys. Rev. A*, vol. 83, p. 062112, Jun 2011.

[89] R. Rabelo, M. Ho, D. Cavalcanti, N. Brunner, and V. Scarani, "Device-independent certification of entangled measurements," *Phys. Rev. Lett.*, vol. 107, p. 050502, Jul 2011.

[90] A. Bennet, T. Vértesi, D. J. Saunders, N. Brunner, and G. J. Pryde, "Experimental semi-device-independent certification of entangled measurements," *Phys. Rev. Lett.*, vol. 113, p. 080405, Aug 2014.

[91] M. F. Pusey, J. Barrett, and T. Rudolph, "On the reality of the quantum state," *Nature Physics*, vol. 8, no. 6, pp. 476–479, 2012.

[92] A. Ambainis, A. Nayak, A. Ta-Shma, and U. Vazirani, "Dense quantum coding and quantum finite automata," *Journal of the ACM (JACM)*, vol. 49, no. 4, pp. 496–511, 2002.

[93] S. Wiesner, *Conjugate coding*, vol. 15. 1983.

[94] P. Mironowicz, H.-W. Li, and M. Pawłowski, "Properties of dimension witnesses and their semidefinite programming relaxations," *Phys. Rev. A*, vol. 90, p. 022322, Aug 2014.

[95] A. Tavakoli, A. Hameedi, B. Marques, and M. Bourennane, "Quantum random access codes using single d-level systems," *Phys. Rev. Lett.*, vol. 114, p. 170502, Apr 2015.

[96] S. Wehner, M. Christandl, and A. C. Doherty, "Lower bound on the dimension of a quantum system given measured data," *Phys. Rev. A*, vol. 78, p. 062112, Dec 2008.

[97] R. Raz, "Exponential separation of quantum and classical communication complexity," *Proceedings of the 31st Annual ACM Symposium on Theory of Computing*, pp. 358–367, 1999.

[98] E. F. Galvão, "Feasible quantum communication complexity protocol," *Phys. Rev. A*, vol. 65, p. 012318, Dec 2001.

[99] H.-W. Li, M. Pawłowski, Z.-Q. Yin, G.-C. Guo, and Z.-F. Han, "Semi-device-independent randomness certification using $n \to 1$ quantum random access codes," *Phys. Rev. A*, vol. 85, p. 052308, May 2012.

[100] 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*, vol. 84, p. 034301, Sep 2011.

[101] M. Pawłowski and N. Brunner, "Semi-device-independent security of one-way quantum key distribution," *Phys. Rev. A*, vol. 84, p. 010302, Jul 2011.

[102] M. Dall'Arno, E. Passaro, R. Gallego, M. Pawłowski, and A. Acín, "Detection loophole attacks on semi-device-independent quantum and classical protocols," *Quantum Information and Computation*, vol. 15, no. 1-2, pp. 37–49, 2014.

[103] A. Dittmer, "Cross product identities in arbitrary dimension," *American Math. Monthly*, vol. 101, p. 887, 1994.

[104] Z. Cao, H. Zhou, X. Yuan, and X. Ma, "Source-independent quantum random number generation," *Phys. Rev. X*, vol. 6, p. 011020, Feb 2016.

[105] Y.-G. Han, Z.-Q. Yin, H.-W. Li, W. Chen, S. Wang, G.-C. Guo, and Z.-F. Han, "More randomness from a prepare-and-measure scenario with independent devices," *Phys. Rev. A*, vol. 93, p. 032332, Mar 2016.

[106] G. Cañas, J. Cariñe, E. S. Gómez, J. F. Barra, A. CaBello, G. B. Xavier, G. Lima, and M. Pawłowski, "Experimental quantum randomness generation invulnerable to the detection loophole," *arXiv:1410.3443*, 2016.

[107] H.-W. Li, Z.-Q. Yin, W. Chen, S. Wang, G.-C. Guo, and Z.-F. Han, "Quantum key distribution based on quantum dimension and independent devices," *Phys. Rev. A*, vol. 89, p. 032302, Mar 2014.

[108] R. Chaves, R. Kueng, J. B. Brask, and D. Gross, "Unifying framework for relaxations of the causal assumptions in Bell's theorem," *Physical Review Letters*, vol. 114, no. 14, 2015.

[109] A. Tavakoli, P. Skrzypczyk, D. Cavalcanti, and A. Acín, "Nonlocal correlations in the star-network configuration," *Physical Review A - Atomic, Molecular, and Optical Physics*, vol. 90, no. 6, 2014.

[110] C. Branciard, D. Rosset, N. Gisin, and S. Pironio, "Bilocal versus nonbilocal correlations in entanglement-swapping experiments," *Physical Review A - Atomic, Molecular, and Optical Physics*, vol. 85, no. 3, 2012.

[111] D. Rosset, C. Branciard, T. J. Barnea, G. Pütz, N. Brunner, and N. Gisin, "Nonlinear Bell inequalities tailored for quantum networks," *Phys. Rev. Lett.*, vol. 116, p. 010403, Jan 2016.

[112] C. Branciard, N. Gisin, and S. Pironio, "Characterizing the nonlocal correlations created via entanglement swapping," *Physical Review Letters*, vol. 104, no. 17, 2010.

Acknowledgements

First and foremost I thank my supervisor Prof. Nicolas Brunner. His endless energy, enthusiasm and ideas inspired me to study for this thesis and have motivated most of the presented work. I am particularly grateful for the considerable amount of time to which he has dedicated discussions, explanations, and paper writing. They undoubtedly have and will continue to be of enormous benefit.

I would also like to thank all members of the groups in Geneva, past and present. I thank Marco Túlio Quintino and Flavien Hirsch for the many (sometimes useful but always interesting) debates and discussions both in and out of work, and for all the ideas, motivation and late nights which have contributed to a large part of this thesis. I would equally like to thank Jonatan Brask from whom I have learned a great deal. I thank Prof. Nicolas Gisin for inspiration and discussions. In the same respect I thank Prof. Hugo Zbinden, Prof. Yeong Cherng Liang, Denis Rosset, Tomer Barnea, Gilles Pütz and Charles Ci Wen Lim.

Outside of Geneva I am grateful to Prof. Támas Vértesi, Prof. Antonio Acìn, Prof. Stefano Pironio, Marcus Huber, Rafael Chaves, Paul Skrzypczyk and Daniel Cavalcanti for very useful discussions and ideas. I thank Prof. Sandu Popescu, whose fascinating lectures on quantum information were are large part of my inspiration to start this thesis. In the same respect I would like to take this opportunity thank Prof. Dan Browne, with whom I undertook my first serious project in quantum information and who provided much of the inspiration to study in this field.

I thank my family and friends for support, encouragement and love.

Finally I thank you, Alexia. Your support (professional and personal) is invaluable.

Part III

Published Papers

Paper A

One-way Einstein-Podolsky-Rosen Steering

Physical Review Letters **112**, 200402 (2013)

Joseph Bowles, Támas Vértesi, Marco Túlio Quintino, and Nicolas Brunner

One-way Einstein-Podolsky-Rosen Steering

Joseph Bowles,[1] Tamás Vértesi,[2] Marco Túlio Quintino,[1] and Nicolas Brunner[1,3]
[1]*Département de Physique Théorique, Université de Genève, 1211 Genève, Switzerland*
[2]*Institute for Nuclear Research, Hungarian Academy of Sciences, P.O. Box 51, H-4001 Debrecen, Hungary*
[3]*H.H. Wills Physics Laboratory, University of Bristol, Bristol BS8 1TL, United Kingdom*
(Received 24 February 2014; published 22 May 2014)

Einstein-Podolsky-Rosen steering is a form of quantum nonlocality exhibiting an inherent asymmetry between the observers, Alice and Bob. A natural question is then whether there exist entangled states which are one-way steerable, that is, Alice can steer Bob's state, but it is impossible for Bob to steer the state of Alice. So far, such a phenomenon has been demonstrated for continuous variable systems, but with a strong restriction on allowed measurements, namely, considering only Gaussian measurements. Here we present a simple class of entangled two-qubit states which are one-way steerable, considering arbitrary projective measurements. This shows that the nonlocal properties of entangled states can be fundamentally asymmetrical.

DOI: 10.1103/PhysRevLett.112.200402 PACS numbers: 03.65.Ud

The nonlocality of entangled quantum states, first pointed out by Einstein, Podolsky, and Rosen [1], was later proven by Bell [2] to be an inherent feature of the theory. Nowadays quantum nonlocality is considered as a fundamental aspect of the theory and plays a central role in quantum information processing [3,4].

The concept of steering (or Einstein-Podolsky-Rosen steering), proposed by Schrödinger [5], brings an alternative approach to this phenomenon. Consider two remote observers, Alice and Bob, who share a pair of entangled particles. By performing a measurement on her system, Alice can steer the state of Bob's system. Importantly, it is the intrinsic randomness of quantum theory that prevents this effect from leading to instantaneous signaling. First explored in the context of continuous variable systems [6,7], quantum steering was recently formalized as an information-theoretic task by Wiseman *et al.* [8]. Steering finds applications in quantum information processing, e.g., for cryptography [9] and randomness generation [10]. Experimental investigations have been reported [11] with, notably, a recent loophole-free experiment [12]. Steering has also been discussed for detecting entanglement in Bose-Einstein condensates [13] and atomic ensembles [14].

A characteristic trait of steering—distinguishing it from both entanglement and Bell nonlocality—is an asymmetry between the observers. As formalized in [8], a steering test can be viewed as the distribution of entanglement from an untrusted party. Hence, in this protocol, Alice and Bob play different roles which are not interchangeable. Specifically, Alice tries to convince Bob that they share an entangled state. However, Bob does not trust Alice, and thus asks her to remotely steer the state of his particle according to a different measurement basis. Bob can then verify Alice's claim by checking a steering inequality [15], as the violation of such an inequality implies the presence of entanglement. Conversely, if the inequality is satisfied, Bob will not be convinced that entanglement is present, since a local state strategy can, in principle, reproduce the observed data. Interestingly, steering turns out to be a form of quantum nonlocality that is intermediate between entanglement and Bell nonlocality, in the sense that not all entangled states lead to steering, and not all steerable states violate a Bell inequality [8,11].

A natural question, already raised in Ref. [8], is then whether there exists one-way quantum steering. That is, are there entangled states such that steering can occur from Alice to Bob, but not from Bob to Alice? The properties of such states would thus be fundamentally different depending on the role of the observers. On the one hand Alice can convince Bob that the state they share is entangled. On the other hand, it is impossible for Bob to convince Alice that the state is entangled since the observed behavior can be reproduced by a local state model. Note that such a phenomenon cannot occur for pure entangled states, which can always be brought to a symmetric form via local basis change (using the so-called Schmidt decomposition). Hence, one-way quantum steering requires mixed entangled states. So far, it was shown theoretically [16,17] and experimentally [18] that such phenomena can occur in continuous variable systems. However, these results hold only for a restricted class of measurements, namely, Gaussian measurements, and there is no evidence that this asymmetry will persist for more general measurements. In fact, it is known that non-Gaussian measurements are useful in this context, as they are necessary to reveal the nonlocality of certain entangled states [19].

Here we present a simple class of two-qubit entangled states with one-way steering for arbitrary projective measurements. First we show that steering is not possible from Bob to Alice by constructing an explicit local hidden state

model. Then we show that the state is nevertheless steerable when the roles of the parties are interchanged. Making use of techniques recently introduced in Skrzypczyk *et al.* [20], we construct a steering test for demonstrating steering from Alice to Bob. The present work thus demonstrates a fundamental asymmetry in the nonlocal properties of certain entangled states.

We start by introducing the scenario and fixing notations. Consider two remote parties, Alice and Bob, sharing an entangled quantum state ρ_{AB}. By performing a local measurement on her particle, Alice can prepare the state of Bob's particle in different ways. In this work we will focus on the case of two-qubit states ρ_{AB} and local qubit projective measurements. Consider that Alice measures the observable $\vec{x}\cdot\vec{\sigma}$ and obtains outcome $a=\pm 1$; here $\vec{x}$ denotes a vector on the Bloch sphere and $\vec{\sigma}=(\sigma_1,\sigma_2,\sigma_3)$ is the vector of Pauli matrices. Then, Bob's particle is left in the (unnormalized) state

$$\rho_{a|\vec{x}}=\mathrm{tr}_A(\rho_{AB}M_{a|\vec{x}}\otimes\mathbb{I}), \tag{1}$$

where $M_{a|\vec{x}}=(\mathbb{I}+a\vec{x}\cdot\vec{\sigma})/2$ is the projector corresponding to outcome a. The set of unnormalized states $\{\rho_{a|\vec{x}}\}$, referred to as an assemblage, thus characterizes the experiment [8,20,21]. The assemblage characterizes both the conditional probability of Alice's outcome, $p(a|\vec{x})=\mathrm{tr}(\rho_{a|\vec{x}})$, and the (normalized) conditional state prepared for Bob $\hat{\rho}_{a|\vec{x}}=\rho_{a|\vec{x}}/p(a|\vec{x})$. Note that all assemblages satisfy $\sum_a\rho_{a|\vec{x}}=\sum_a\rho_{a|\vec{x}'}$ for all measurement directions $\vec{x}$ and $\vec{x}'$, ensuring that Alice cannot signal to Bob.

In a steering test [8], Alice wants to convince Bob that she can steer his state. Bob, who does not fully trust Alice, wants to verify her claim. In order to do so, he asks Alice to make a measurement in a given direction $\vec{x}$ (chosen from a given set of measurements) and then to announce her result a. By repeating this procedure a sufficient number of times, Bob can estimate the assemblage $\{\rho_{a|\vec{x}}\}$, e.g., via quantum state tomography. Bob's goal is now to find out whether (i) Alice did indeed steer his state by making a measurement on an entangled state ρ_{AB}, or whether (ii) she cheated by using a local hidden state (LHS) strategy, in which no entanglement is involved. In this second case, Alice would prepare a single qubit state ρ_λ and send it to Bob; here λ represents a classical variable known to Alice, with an arbitrary distribution $\omega(\lambda)$. Upon receiving a measurement direction $\vec{x}$ from Bob, Alice announces an outcome a according to a predetermined strategy $p_\lambda(a|\vec{x})$. Hence Bob holds the state

$$\rho_{a|\vec{x}}=\sum_\lambda\omega(\lambda)p_\lambda(a|\vec{x})\rho_\lambda. \tag{2}$$

Therefore, the problem for Bob is to determine whether the states in the assemblage $\{\rho_{a|\vec{x}}\}$ admit a decomposition of the form of equation (2). If this is the case, then Bob will not be convinced that Alice can steer his state. On the other hand, if no decomposition of the form of equation (2) is possible, then Bob will be convinced that Alice did steer his state. More generally, we say that a state ρ_{AB} is unsteerable from Alice to Bob, if the assemblage $\{\rho_{a|\vec{x}}\}$ admits a decomposition of the form of equation (2) for all possible measurement directions $\vec{x}$. On the other hand, if there exists a set of measurement directions such that the corresponding assemblage $\{\rho_{a|\vec{x}}\}$ does not admit a decomposition of the form of equation (2), we say that ρ_{AB} is steerable from Alice to Bob.

A steering test is thus clearly asymmetrical, as the roles played by Alice and Bob are different. Hence it is natural to ask whether there exist entangled states ρ_{AB} that can be steered only in one direction, say from Alice to Bob but not from Bob to Alice. Here we show that such a phenomenon of *one-way steering* occurs for simple two-qubit entangled states, considering arbitrary projective measurements.

Specifically, we consider states of the form

$$\rho_{AB}(\alpha)=\alpha\Psi_-+\frac{1-\alpha}{5}\left(2|0\rangle\langle 0|\otimes\frac{\mathbb{I}}{2}+3\frac{\mathbb{I}}{2}\otimes|1\rangle\langle 1|\right), \tag{3}$$

where $\Psi_-=|\psi^-\rangle\langle\psi^-|$ denotes the projector on the singlet state $|\psi^-\rangle=(|0,1\rangle-|1,0\rangle)/\sqrt{2}$ and $0\le\alpha\le 1$. The state $\rho_{AB}(\alpha)$ is entangled for $\alpha>1/19(-6+5\sqrt{6})\simeq 0.3288$, as can be checked via partial transposition [4]. We will see that in the range $0.4983\lesssim\alpha\le 1/2$, the state $\rho_{AB}(\alpha)$ is one-way steerable. The proof is divided into two parts. First, we show that the state is unsteerable from Bob to Alice by constructing a LHS model for $\rho_{AB}(1/2)$. Second, we show that steering can nevertheless occur from Alice to Bob by showing that the assemblage resulting from 14 well-chosen projective measurements on the state $\rho_{AB}(\alpha)$ with $\alpha\gtrsim 0.4983$ does not admit a decomposition of the form of equation (2).

No steering from B to A.—We construct a LHS model, from Bob to Alice, for arbitrary local projective measurements on $\rho_{AB}(1/2)$. The model works as follows. Bob first sends to Alice a pure qubit state of the form

$$\rho_\lambda=(\mathbb{I}+\lambda_0\vec{\lambda}\cdot\vec{\sigma})/2, \tag{4}$$

where $\lambda_0=\pm 1$, and $\vec{\lambda}=(\lambda_1,\lambda_2,\lambda_3)=(\cos\phi\sin\theta,\sin\phi\sin\theta,\cos\theta)$ is a vector on the Bloch sphere distributed according to the density

$$\omega(\theta,\phi)=\frac{1}{2\pi}\cos^2(\theta/2). \tag{5}$$

That is, the probability of using a given vector $\vec{\lambda}$ depends on its height on the Bloch sphere. Note that λ_0 and $\vec{\lambda}$ represent here the classical variables available to Bob. Upon receiving an arbitrary measurement direction $\vec{y}=(y_1,y_2,y_3)$

from Alice, Bob then announces outcome $b = -\lambda_0 \mathrm{sgn}(\vec{y}\cdot\vec{\lambda})$. Finally, Alice characterizes her state. For convenience, we consider here the case where she performs an arbitrary projective measurement along direction $\vec{x} = (x_1, x_2, x_3)$ with outcome a.

Now we compute the statistics of the above model, focusing first on the case $\lambda_0 = 1$. Because of the form of the state, Eq. (3), we can take $\vec{y} = (0, \sin\theta_B, \cos\theta_B)$ without loss of generality. Moreover, it will be convenient to use a new reference frame such that the $\hat{e}_3 = (0,0,1)$ axis is aligned on Bob's vector $\vec{y}$. Angles and axes in the new frame are denoted with a tilde. First we evaluate the distribution of $\vec{\lambda}$ in the new frame. That is, we compute $\omega(\tilde{\theta},\tilde{\phi})$, with $\vec{\lambda} = (\tilde{\lambda}_1, \tilde{\lambda}_2, \tilde{\lambda}_3) = (\cos\tilde{\phi}\sin\tilde{\theta}, \sin\tilde{\phi}\sin\tilde{\theta}, \cos\tilde{\theta})$. Since the new frame is obtained by performing a rotation of $-\theta_B$ around the $\hat{e}_1 = (1,0,0)$ axis, we have that $\lambda_3 = -\sin\theta_B\tilde{\lambda}_2 + \cos\theta_B\tilde{\lambda}_3$. Moreover, since $\theta = \arccos(\lambda_3)$, we have that

$$\omega(\theta,\phi) = \frac{1}{2\pi}\cos^2\left(\frac{\arccos(\lambda_3)}{2}\right) = \frac{1+\lambda_3}{4\pi}. \tag{6}$$

Hence we get that

$$\omega(\tilde{\theta},\tilde{\phi}) = (1 - \sin\theta_B \sin\tilde{\phi}\sin\tilde{\theta} + \cos\theta_B\cos\tilde{\theta})/4\pi.$$

Next, we write Alice's vector in the new frame, $\vec{x} = (\cos\tilde{\phi}_A\sin\tilde{\theta}_A, \sin\tilde{\phi}_A\sin\tilde{\theta}_A, \cos\tilde{\theta}_A)$. Using the fact that $\mathrm{tr}(\vec{x}\cdot\vec{\sigma}\rho_\lambda) = \vec{x}\cdot\vec{\lambda}$, we obtain the correlation

$$\begin{aligned}\langle ab\rangle &= -\int_0^{2\pi} d\tilde{\phi}\int_0^{\pi}\sin\tilde{\theta}d\tilde{\theta}\omega(\tilde{\theta},\tilde{\phi})(\vec{x}\cdot\vec{\lambda})\mathrm{sgn}(\vec{y}\cdot\vec{\lambda})\\ &= \int_0^{2\pi} d\tilde{\phi}\left(\int_{\pi/2}^{\pi}\sin\tilde{\theta}d\tilde{\theta}\omega(\tilde{\theta},\tilde{\phi})(\vec{x}\cdot\vec{\lambda})\right.\\ &\quad\left. -\int_0^{\pi/2}\sin\tilde{\theta}d\tilde{\theta}\omega(\tilde{\theta},\tilde{\phi})(\vec{x}\cdot\vec{\lambda})\right).\end{aligned} \tag{7}$$

Since $\vec{x}\cdot\vec{\lambda} = \sin\tilde{\theta}\sin\tilde{\theta}_A\cos(\tilde{\phi}-\tilde{\phi}_A) + \cos\tilde{\theta}\cos\tilde{\theta}_A$, we find after a lengthy but straightforward calculation

$$\langle ab\rangle = -\frac{\cos\tilde{\theta}_A}{2} = -\frac{\vec{x}\cdot\vec{y}}{2}. \tag{8}$$

Note that $\tilde{\theta}_A$ is the angle between vectors $\vec{x}$ and $\vec{y}$.

Finally, we calculate the marginals, i.e., the local expectation values for Bob

$$\begin{aligned}\langle b\rangle &= -\int_0^{2\pi} d\tilde{\phi}\int_0^{\pi}\sin\tilde{\theta}d\tilde{\theta}\omega(\tilde{\theta},\tilde{\phi})\mathrm{sgn}(\vec{y}\cdot\vec{\lambda})\\ &= -\frac{\cos\theta_B}{2} = -\frac{y_3}{2}\end{aligned} \tag{9}$$

and for Alice

$$\begin{aligned}\langle a\rangle &= \int_0^{2\pi} d\phi\int_0^{\pi}\sin\theta d\theta\omega(\theta,\phi)(\vec{x}\cdot\vec{\lambda})\\ &= \frac{\cos\theta_A}{3} = \frac{x_3}{3}.\end{aligned} \tag{10}$$

Note that for computing Alice's marginal, it is more convenient to use the original reference frame.

At this point, it is useful to note that the correlations, Eq. (8), correspond exactly to those of the state $\rho_{AB}(1/2)$. Moreover, the marginals, Eqs. (9) and (10), have the right form, but are in fact slightly stronger than those of $\rho_{AB}(1/2)$. In order to weaken the marginals, while keeping the correlation unchanged, we now use the variable λ_0. Specifically, consider the distribution $p(\lambda_0 = -1) = f$. Hence, the marginals are decreased to $\langle a\rangle = (1-2f)x_3/3$ and $\langle b\rangle = (1-2f)y_3/2$. Choosing a flipping probability of $f = 1/5$, we finally get

$$\langle a\rangle = \frac{x_3}{5}, \qquad \langle b\rangle = \frac{3y_3}{10}, \qquad \langle ab\rangle = -\frac{\vec{x}\cdot\vec{y}}{2}. \tag{11}$$

Hence, the model simulates exactly the statistics of local projective measurements on the state $\rho_{AB}(1/2)$. The assemblage $\{\rho_{b|\vec{y}}\}$ observed by Alice is thus identical to the assemblage expected for the state $\rho_{AB}(1/2)$, that is, $\rho_{b|\vec{y}} = \mathrm{tr}_B(\rho_{AB}(1/2)\mathbb{I}\otimes M_{b|\vec{y}})$, where $M_{b|\vec{y}} = (\mathbb{I} + b\vec{y}\cdot\vec{\sigma})/2$. Therefore, the state $\rho_{AB}(1/2)$ is unsteerable from Bob to Alice. The extension to the case $\alpha < 1/2$ is straightforward.

Finally, note that the above model can also be understood as a local hidden variable model; thus, the statistics of local projective measurements on $\rho_{AB}(\alpha)$ with $\alpha \le 1/2$ cannot violate any Bell inequality [22]. This complements a series of works describing entangled states admitting a local hidden variable model [23–26].

Steering from A to B.—We will see now that the situation is completely different when the roles of Alice and Bob are interchanged. Specifically, the state $\rho_{AB}(\alpha)$ with $\alpha \gtrsim 0.4983$ is steerable from Alice to Bob. In order to prove this, we will show that, for a well-chosen set of m projective measurements for Alice, the resulting assemblage $\{\rho_{a|\vec{x}}\}$ obtained on Bob's side cannot be reproduced by any LHS model.

The observables measured by Alice are denoted $A_i = \vec{x}_i\cdot\vec{\sigma}$ with $i = 1,\ldots,m$ and outcome $a = \pm 1$. Bob characterizes the state $\rho_{a|\vec{x}_i}$ by tomography, making measurements represented by the Pauli matrices σ_j, with $j = 1,2,3$, outcome $b = \pm 1$, and $\sigma_0 = \mathbb{I}$. The observed statistics are then given by

$$\begin{aligned}\langle ab\rangle_{ij} &= \mathrm{tr}(\rho_{AB}(\alpha)A_i\otimes\sigma_j),\\ \langle b\rangle_j &= \mathrm{tr}(\rho_{AB}(\alpha)\mathbb{I}\otimes\sigma_j).\end{aligned} \tag{12}$$

Alice's marginals are given by $\langle a\rangle_i = \langle ab\rangle_{i0}$.

Considering a given number of measurements m, we now aim at finding the largest value of α, denoted α^*, for

which the state $\rho_{AB}(\alpha)$ is unsteerable from Alice to Bob. That is, we want to determine the largest α such that the statistics (12) can be reproduced by a LHS model, i.e.,

$$\langle ab\rangle_{ij} = \sum_\lambda E_\lambda(i)\mathrm{tr}(\rho_\lambda\sigma_j), \qquad \langle b\rangle_j = \sum_\lambda \mathrm{tr}(\rho_\lambda\sigma_j), \quad (13)$$

where $E_\lambda(i) = p_\lambda(a=1|i) - p_\lambda(a=-1|i)$ is the expectation value of Alice's outcome a for a given λ and measurement i. Note that here the local states ρ_λ are not normalized, and one has that $\sum_\lambda \mathrm{tr}(\rho_\lambda) = 1$.

To solve this problem we make use of a semi-definite programming (SDP) technique recently developed in [20], for deciding whether a given assemblage $\{\rho_{a|\vec{x}}\}$ belongs to the set of "unsteerable assemblages," that is, whether $\{\rho_{a|\vec{x}}\}$ admits a decomposition of the form equation (2). Our present problem can be solved by the following SDP:

$$\begin{aligned} &\alpha^* \equiv \max \alpha \qquad \text{s.t.} \sum_\lambda E_\lambda(i)\mathrm{tr}(\rho_\lambda\sigma_j) = \langle ab\rangle_{ij}, \\ &\sum_\lambda \mathrm{tr}(\rho_\lambda\sigma_j) = \langle b\rangle_j, \mathrm{tr}\sum_\lambda \rho_\lambda = 1, \qquad \rho_\lambda \geq 0 \quad \forall\ \lambda, \end{aligned} \quad (14)$$

where the optimization variables are α and ρ_λ, and the quantities $\langle ab\rangle_{ij}$ and $\langle b\rangle_j$ are computed as in Eq. (12). Note that we can focus here on LHS strategies for which Alice provides a deterministic outcome a given λ and i [20], that is $E_\lambda(i) = \pm 1$ for $i = 1, \ldots, m$. Hence, we have altogether 2^m possible strategies for Alice to consider. The above SDP is then implemented for each strategy.

Using the above SDP, we can thus estimate, for a particular choice of m measurement directions $\vec{x}_i$ (with $i = 1, \ldots, m$), the threshold value α^* for which the state $\rho_{AB}(\alpha)$ is steerable from Alice to Bob. For fixed m, we then minimize α^* over all possible choices of measurement operators for Alice, using a hill-climbing heuristic algorithm. Results for m up to 14 are presented in Table I. Notably, for $m = 14$ we get $\alpha^* \simeq 0.4983$, thus implying that the state $\rho_{AB}(\alpha)$ with $\alpha \gtrsim 0.4983$ is steerable from Alice to Bob.

Finally, from the result of the above optimization procedure, it is in fact possible to extract an explicit steering inequality. Once the optimal measurement directions $\vec{x}_i$ $(i = 1, \ldots, m)$ have been found via the hill-climbing algorithm, the dual of the SDP problem (14) allows us to extract a linear steering inequality [20] of the form

$$\sum_{i=1}^m \sum_{j=1}^3 s_{ij}\langle ab\rangle_{ij} + \sum_{i=1}^m s_i^A\langle a\rangle_i + \sum_{j=1}^3 s_j^B\langle b\rangle_j \leq L. \quad (15)$$

Such an inequality is characterized by (i) a set of real coefficients: s_{ij}, s_i^A, and s_j^B, and (ii) a bound L which holds for any LHS strategy. In the Supplemental Material [27], we follow the above method to give explicitly a steering inequality featuring $m = 13$ measurements, which is violated by performing appropriate measurements (which we give as well) on the state $\rho_{AB}(1/2)$.

Discussion.—We have shown the existence of entangled states which are one-way steerable when considering arbitrary projective measurements. That is, the nonlocal properties of such states depend on the role played by the parties: while Alice can steer the state of Bob, it is impossible for Bob to steer Alice's state. This shows that quantum nonlocality can be fundamentally asymmetrical. An interesting open question is whether the present result can be extended to the most general measurements, i.e., positive operator valued measures. Moreover, it would be interesting to find an application, e.g., in quantum information processing, of the phenomenon of one-way steering.

TABLE I. Threshold values α^* for which the state $\rho_{AB}(\alpha)$ is steerable from Alice to Bob. The optimization is conducted over all possible steering tests where Alice performs $m = 2, \ldots, 14$ projective measurements.

m	2	3	4	5	6	7	8
α^*	0.6951	0.5661	0.5424	0.5302	0.5156	0.5120	0.5088
m	9	10	11	12	13	14	
α^*	0.5037	0.5030	0.5014	0.5005	0.4993	0.4983	

We thank Jonatan Brask for discussions and acknowledge financial support from the Swiss National Science Foundation (Grants No. PP00P2_138917 and No. QSIT), SEFRI (COST action MP1006) and the EU DIQIP. T. V. thanks support from the János Bolyai Programme, the OTKA (Grant No. PD101461), and the TÁMOP-4.2.2.C-11/1/KONV-2012-0001 project.

[1] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
[2] J. Bell, Physics (Long Island City, N.Y.) **1**, 195 (1964).
[3] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Rev. Mod. Phys. **86**, 419 (2014).
[4] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009).
[5] E. Schrödinger, Proc. Cambridge Philos. Soc. **31**, 555 (1935).
[6] M. D. Reid, Phys. Rev. A **40**, 913 (1989).
[7] M. D. Reid, P. D. Drummond, E. G. Cavalcanti, P. K. Lam, H. A. Bachor, U. L. Andersen, and G. Leuchs, Rev. Mod. Phys. **81**, 1727 (2009).
[8] H. M. Wiseman, S. J. Jones, and A. C. Doherty, Phys. Rev. Lett. **98**, 140402 (2007).
[9] C. Branciard, E. G. Cavalcanti, S. P. Walborn, V. Scarani, and H. M. Wiseman, Phys. Rev. A **85**, 010301(R) (2012).
[10] Y. Z. Law, J.-D. Bancal, and V. Scarani, arXiv:1401.4243.

[11] D. J. Saunders, S. J. Jones, H. M. Wiseman, and G. J. Pryde, Nat. Phys. **6**, 845 (2010).
[12] B. Wittmann, S. Ramelow, F. Steinlechner, N. K. Langford, N. Brunner, H. M. Wiseman, R. Ursin, and A. Zeilinger, New J. Phys. **14**, 053030 (2012); see also D. H. Smith, G. Gillett, M. P. de Almeida, C. Branciard, A. Fedrizzi, T. J. Weinhold, A. Lita, B. Calkins, T. Gerrits, H. M. Wiseman, S. W. Nam, and A. G. White, Nat. Commun. **3**, 625 (2012); A. J. Bennett, D. A. Evans, D. J. Saunders, C. Branciard, E. G. Cavalcanti, H. M. Wiseman, and G. J. Pryde, Phys. Rev. X **2**, 031003 (2012).
[13] Q. Y. He, M. D. Reid, T. G. Vaughan, C. Gross, M. Oberthaler, and P. D. Drummond, Phys. Rev. Lett. **106**, 120405 (2011).
[14] Q. Y. He and M. D. Reid, New J. Phys. **15**, 063027 (2013).
[15] E. G. Cavalcanti, S. J. Jones, H. M. Wiseman, and M. D. Reid, Phys. Rev. A **80**, 032112 (2009).
[16] S. L. W. Midgley, A. J. Ferris, and M. K. Olsen, Phys. Rev. A **81**, 022101 (2010).
[17] M. K. Olsen, Phys. Rev. A **88**, 051802 (2013).
[18] V. Händchen, T. Eberle, S. Steinlechner, A. Samblowski, T. Franz, R. F. Werner, and R. Schnabel, Nat. Photonics **6**, 598 (2012).
[19] K. Banaszek and K. Wódkiewicz, Phys. Rev. Lett. **82**, 2009 (1999).
[20] P. Skrzypczyk, M. Navascués, and D. Cavalcanti, Phys. Rev. Lett. **112**, 180404 (2014).
[21] M. F. Pusey, Phys. Rev. A **88**, 032313 (2013).
[22] Note that our local model is related to that of Ref. [23]; the central difference is that we consider here a nonuniform distribution of vectors $\vec{\lambda}$.
[23] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).
[24] J. Barrett, Phys. Rev. A **65**, 042302 (2002).
[25] M. L. Almeida, S. Pironio, J. Barrett, G. Tóth, and A. Acín, Phys. Rev. Lett. **99**, 040403 (2007).
[26] F. Hirsch, M. T. Quintino, J. Bowles, and N. Brunner, Phys. Rev. Lett. **111**, 160402 (2013).
[27] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.112.200402 for derivation of the steering inequality.

One-way Einstein-Podolsky-Rosen steering / Supplementary material

Joseph Bowles,[1] Tamás Vértesi,[2] Marco Túlio Quintino,[1] and Nicolas Brunner[1,3]

[1]*Département de Physique Théorique, Université de Genève, 1211 Genève, Switzerland*
[2]*Institute for Nuclear Research, Hungarian Academy of Sciences, H-4001 Debrecen, P.O. Box 51, Hungary*
[3]*H.H. Wills Physics Laboratory, University of Bristol, Bristol, BS8 1TL, United Kingdom*

Here we describe explicitly the steering test witnessing the fact that the state $\rho_{AB}(\alpha)$ (see eq. (3) of main text) with $\alpha > 1/2$ (more precisely the proof works for $\alpha > (2268/2269)(1/2) \simeq 0.4998$) is steerable from Alice to Bob.

Here we consider the case of $m = 13$ measurement settings for Alice, characterized by operators of the form $A_i = \vec{x}_i \cdot \vec{\sigma}$ with $i = 1, ..., m$ with outcome $a = \pm 1$. Bob performs tomography, making measurements in the Pauli basis σ_j with $j = 1, 2, 3$, outcome $b = \pm 1$. The observed data is then given by

$$\langle ab \rangle_{ij} = \mathrm{tr}(\rho_{AB}(\alpha) A_i \otimes \sigma_j) \quad (1)$$
$$\langle b \rangle_j = \mathrm{tr}(\rho_{AB}(\alpha) \mathbb{I} \otimes \sigma_j) \quad (2)$$
$$\langle a \rangle_i = \mathrm{tr}(\rho_{AB}(\alpha) A_i \otimes \mathbb{I}). \quad (3)$$

We now construct a linear steering inequality of the form

$$\sum_{i=1}^{m} \sum_{j=1}^{3} s_{ij} \langle ab \rangle_{ij} + \sum_{i=1}^{m} s_i^A \langle a \rangle_i + \sum_{j=1}^{3} s_j^B \langle b \rangle_j \leq L. \quad (4)$$

The inequality is characterized by the matrix $\mathbf{S}$, with real coefficients s_{ij}, and the vectors $\mathbf{S^A}$ and $\mathbf{S^B}$, with real elements s_i^A and s_j^B, respectively. Specifically, we have that

$$\mathbf{S} = \begin{pmatrix} \frac{1}{12} & \frac{6}{175} & \frac{-12}{27} \\ \frac{-9}{79} & \frac{-1}{38} & \frac{-7}{94} \\ \frac{-1}{162} & \frac{18}{133} & \frac{-1}{18} \\ \frac{17}{157} & \frac{-6}{143} & \frac{-10}{141} \\ \frac{5}{62} & \frac{-10}{97} & \frac{-1}{62} \\ 0 & \frac{2}{103} & \frac{-9}{76} \\ \frac{-16}{105} & \frac{1}{89} & 0 \\ \frac{5}{104} & \frac{-6}{79} & \frac{-11}{72} \\ \frac{-4}{73} & \frac{-6}{109} & \frac{-54}{433} \\ \frac{-3}{26} & \frac{-3}{20} & \frac{-2}{83} \\ \frac{10}{179} & \frac{9}{103} & \frac{-13}{121} \\ \frac{1}{132} & \frac{-4}{33} & \frac{-2}{49} \\ \frac{-11}{107} & \frac{14}{139} & \frac{-20}{161} \end{pmatrix}, \quad \mathbf{S^A} = -\begin{pmatrix} \frac{1}{71} \\ \frac{1}{53} \\ \frac{1}{71} \\ \frac{2}{111} \\ \frac{1}{244} \\ \frac{3}{100} \\ 0 \\ \frac{4}{103} \\ \frac{2}{63} \\ \frac{1}{163} \\ \frac{3}{110} \\ \frac{1}{96} \\ \frac{3}{95} \end{pmatrix}, \quad \mathbf{S^B} = \begin{pmatrix} 0 \\ 0 \\ \frac{-15}{59} \end{pmatrix}. \quad (5)$$

The local bound of the above inequality, which holds for any possible LHS model, is $L = 1$. This can be verified using e.g. the techniques of Refs [1, 2].

Now we give the measurement operators for Alice, characterized by Bloch vectors $\vec{x}_i$ with $i = 1, ..., 13$. We have that

$$\mathbf{V} = \begin{pmatrix} \frac{-31}{38} & \frac{-17}{54} & z_1 \\ \frac{69}{82} & \frac{6}{35} & z_2 \\ \frac{5}{111} & \frac{-103}{110} & z_3 \\ \frac{-9}{11} & \frac{7}{23} & z_4 \\ \frac{-52}{83} & \frac{53}{69} & z_5 \\ \frac{-1}{673} & \frac{-8}{49} & z_6 \\ \frac{456}{457} & \frac{-5}{83} & z_7 \\ \frac{-128}{427} & \frac{57}{124} & z_8 \\ \frac{37}{90} & \frac{35}{88} & z_9 \\ \frac{47}{77} & \frac{183}{233} & z_{10} \\ \frac{-37}{94} & \frac{-53}{86} & z_{11} \\ \frac{-3}{58} & \frac{116}{121} & z_{12} \\ \frac{13}{23} & \frac{-76}{137} & z_{13} \end{pmatrix}, \quad (6)$$

where the k-th row of the above matrix is understood to be $\vec{x}_k$. By normalization of the vectors, we have that $z_k^2 = 1 - v_{k1}^2 - v_{k2}^2$ where v_{ij} denote the elements of matrix $\mathbf{V}$, and z_k is chosen to be positive. With this set of measurements performed on the state $\rho_{AB}(1/2)$, we can evaluate the right hand side of (4) giving us the value $S_q > \frac{2269}{2268} > 1 = L$ hence demonstrating steering from Alice to Bob.

[1] D.J. Saunders, S.J. Jones, H.M. Wiseman, and G.J. Pryde, Nat. Phys. **6**, 845 (2010).
[2] M.F. Pusey, Phys. Rev. A **88**, 032313 (2013).

Paper B

Sufficient criterion for guaranteeing that a two-qubit state is unsteerable

Physical Review A **93**, 022121 (2016)

Joseph Bowles, Flavien Hirsch, Marco Túlio Quintino, and Nicolas Brunner

Sufficient criterion for guaranteeing that a two-qubit state is unsteerable

Joseph Bowles, Flavien Hirsch, Marco Túlio Quintino, and Nicolas Brunner
Département de Physique Théorique, Université de Genève, 1211 Genève, Switzerland
(Received 26 October 2015; published 26 February 2016)

Quantum steering can be detected via the violation of steering inequalities, which provide sufficient conditions for the steerability of quantum states. Here we discuss the converse problem, namely, ensuring that an entangled state is unsteerable and hence Bell local. We present a simple criterion, applicable to any two-qubit state, that guarantees that the state admits a local hidden state model for arbitrary projective measurements. Specifically, we construct local hidden state models for a large class of entangled states, which thus cannot violate any steering or Bell inequality. In turn, this leads to sufficient conditions for a state to be only one-way steerable and provides the simplest possible example of one-way steering. Finally, by exploiting the connection between steering and measurement incompatibility, we give a sufficient criterion for a continuous set of qubit measurements to be jointly measurable.

DOI: 10.1103/PhysRevA.93.022121

I. INTRODUCTION

Entanglement lies at the heart of quantum physics. Notably, correlations arising from local measurements performed on separated entangled systems can exhibit nonlocal correlations [1,2]. Specifically, the observed statistics cannot be reproduced using a local hidden variable model, as witnessed by violation of a Bell inequality.

Recently, the effect of Einstein-Podolsky-Rosen steering has brought novel insight into quantum nonlocality. Originally discussed by Schrödinger [3] and used in quantum optics [4], the effect was recently formalized in a quantum information-theoretic setting [5]. Considering two distant observers sharing an entangled state, steering captures the fact that one observer, by performing a local measurement on one's subsystem, can nonlocally steer the state of the other subsystem. Steering can be understood as a form of quantum nonlocality intermediate between entanglement and Bell nonlocality [5,6] and is useful to explore the relation between these concepts. It was demonstrated experimentally (see, e.g., [7]) and finds application in quantum information processing [8–10].

Steering can be detected via steering inequalities (analogous to Bell inequalities) [11], the violation of which provides a sufficient condition for a given quantum state to be steerable. Derived for both discrete and continuous variable quantum systems [11–13], such inequalities can be obtained using semidefinite programming [14–17].

Interestingly, whereas the effect of steering implies the presence of entanglement, the converse does not hold [5]. Specifically, there exist entangled states that provably cannot give rise to steering (and hence are referred to as unsteerable) [5,18], even when general measurements are considered [6]. The correlations of such states can in fact be reproduced without entanglement, using a so-called local hidden state (LHS) model [5], and therefore can never violate any steering inequality. Since a LHS model is a particular case of a local hidden variable model, any unsteerable state is Bell local.

Determining which entangled states are steerable and which ones are not is a challenging problem in general. This is mainly due to the fact that, when constructing a LHS model, one must ensure that the model reproduces the desired quantum correlations for any possible measurements. Local hidden state models have been constructed for entangled states featuring a high level of symmetry [18–22] (see [23] for a review). For more general states very little is known, even for the simplest case of two-qubit states. Based on the concept of the steering ellipsoid [24], Ref. [25] derived a condition guaranteeing unsteerability of Bell diagonal two-qubit state. This method, however, is not applicable to general two-qubit states, for which unsteerability conditions are still missing.

Here, via the construction of a class of LHS models, we derive a simple criterion sufficient for guaranteeing that a two-qubit state is unsteerable, considering arbitrary projective measurements. In turn, this criterion can also be used to guarantee one-way steerability [22,26], a weak form of steering where only one of the observers can steer the state of the other. We illustrate the relevance of the criterion with examples, providing in particular the simplest possible example of one-way steering. Finally, by exploiting the strong connection between steering and measurement incompatibility [27,28], we provide a sufficient condition for a continuous set of dichotomic qubit positive-operator-valued measures (POVMs) to be jointly measurable.

II. PRELIMINARIES

Consider two distant parties, Alice and Bob, sharing an entangled quantum state ρ. On her subsystem, Alice makes measurements, described by operators $\{M_{a|x}\}$, with $M_{a|x} \geqslant 0$ and $\sum_a M_{a|x} = \mathbb{1}$, where x denotes the measurement setting and a its outcome. The possible states of Bob's subsystem, conditioned on Alice's measurement x and her output a, are characterized by a collection of (subnormalized) density matrices $\{\sigma_{a|x}\}_{a,x}$, called an assemblage, with

$$\sigma_{a|x} = \mathrm{Tr}_A(M_{a|x} \otimes \mathbb{1}\rho). \tag{1}$$

Note that it also includes Alice's marginal statistics $p(a|x) = \mathrm{Tr}\,\sigma_{a|x}$. The assemblage $\{\sigma_{a|x}\}$ is called unsteerable if it can be reproduced by a LHS model, i.e., it admits a decomposition

$$\sigma_{a|x} = \sigma^{\mathrm{LHS}}_{a|x} = \int \sigma_\lambda\, p(a|x\lambda) d\lambda \;\forall a,x, \tag{2}$$

where $\{\sigma_\lambda\}$ is a set of positive matrices such that $\int \mathrm{Tr}\,\sigma_\lambda d\lambda = 1$ and the $p(a|x,\lambda)$'s are probability distributions. The

right-hand side of (2) can be understood as follows. Alice sends the quantum state $\sigma_\lambda/\operatorname{Tr}\sigma_\lambda$ to Bob with probability density $\operatorname{Tr}\sigma_\lambda$. Given her measurement input x, she then outputs a with probability $p(a|x,\lambda)$. In this way, Alice can prepare the same assemblage for Bob as if the state ρ had been used, without the need for entanglement. Bob will thus be unable to distinguish whether he and Alice share the entangled state ρ or if the above LHS strategy were used. In contrast, if a decomposition of the form (2) does not exist, which can be certified, for example, via violation of a steering inequality [11], the use of an entangled state is certified. In this case, ρ is termed steerable from Alice to Bob.

Interestingly, not all entangled states are steerable. That is, there exist entangled states, called unsteerable, which admit a LHS model for all projective measurements [5,18] and even considering general POVMs [6]. A natural question is thus to determine which entangled states are steerable and which ones are not. This is a challenging problem, mainly due to the difficulty of constructing LHS models for a continuous set of measurements.

III. SUFFICIENT CRITERION FOR UNSTEERABILITY

Our main result is a simple criterion, sufficient for a two-qubit state to admit a LHS model for arbitrary projective measurements. Consider a general two-qubit state, expressed in the local Pauli basis

$$\rho_0 = \frac{1}{4}\left(\mathbb{1} + \vec{a}_0\cdot\vec{\sigma}\otimes\mathbb{1} + \mathbb{1}\otimes\vec{b}_0\cdot\vec{\sigma} + \sum_{i,j=x,y,z} T^0_{ij}\sigma_i\otimes\sigma_j\right), \tag{3}$$

where $\vec{a}_0$ and $\vec{b}_0$ are Alice and Bob's local Bloch vectors and $\vec{\sigma} = (\sigma_x,\sigma_y,\sigma_z)$ is the vector of Pauli matrices. Our criterion for unsteerability is simply based on the local Bloch vectors and the correlation matrix T^0_{ij}.

The first step consists in converting the state ρ_0 into a canonical form. For this, based on previous work [6,29], we make the following observation.

Lemma 1. Let Λ be a positive linear map on the set of quantum states and

$$\rho_\Lambda = \mathbb{1}\otimes\Lambda(\rho)/\operatorname{Tr}[\mathbb{1}\otimes\Lambda(\rho)] \tag{4}$$

be a valid bipartite quantum state. If ρ is unsteerable from Alice to Bob, then ρ_Λ is also unsteerable from Alice to Bob. Furthermore, if Λ is invertible and its inverse map positive, then ρ is unsteerable from Alice to Bob if and only if ρ_Λ is unsteerable from Alice to Bob.

Note that Λ does not have to be completely positive and may therefore correspond to a nonquantum operation. For a proof see, e.g., Lemma 2 of [6], where the condition of complete positivity can simply be relaxed.

Let us now consider the positive linear map

$$\mathbb{1}\otimes\Lambda(\rho_0) = \mathbb{1}\otimes\rho_B^{-1/2}\rho_0\mathbb{1}\otimes\rho_B^{-1/2}, \tag{5}$$

where $\rho_B = \operatorname{Tr}_A[\rho_0]$. This map is invertible as long as ρ_B is mixed with the inverse (positive) map given by

$$\mathbb{1}\otimes\Lambda^{-1}(\rho_0) = \mathbb{1}\otimes\rho_B^{1/2}\rho_0\mathbb{1}\otimes\rho_B^{1/2}. \tag{6}$$

The interesting property of this map is that when applied to an arbitrary state ρ_0, the resulting state has $\vec{b} = 0$, i.e., Bob's reduced state is maximally mixed [30]. Given the above lemma, the application of the map preserves the steerability (or unsteerability) of ρ_0.

Finally, we apply local unitaries (which also cannot change the steerability of the state) so that our state has a diagonal T matrix, giving us the canonical form

$$\rho = \frac{1}{4}\left(\mathbb{1} + \vec{a}\cdot\vec{\sigma}\otimes\mathbb{1} + \sum_{i=x,y,z} T_i\sigma_i\otimes\sigma_i\right), \tag{7}$$

where $\vec{a}$ and T_i are in general different from the original $\vec{a}_0$ and T^0_{ii}. Below we give a sufficient criterion for the unsteerability of any state ρ expressed in the canonical form. In turn this provides a sufficient criterion for unsteerability of any two-qubit state.

Theorem 1. Let ρ_0 be a two-qubit state with corresponding canonical form ρ as given in Eq. (7). If

$$\max_{\hat{x}}[(\vec{a}\cdot\hat{x})^2 + 2\|T\hat{x}\|] \leqslant 1, \tag{8}$$

where $\hat{x}$ is a normalized vector and $\|\cdot\|$ the Euclidean vector norm, then ρ is unsteerable from Alice to Bob, considering arbitrary projective measurements.

Proof. We first characterize the assemblage resulting from projective measurements on a state in the canonical form ρ. Alice's measurement is given by a Bloch vector $\hat{x}$ and output $a = \pm 1$, corresponding to operators $M_{\pm|\hat{x}} = (\mathbb{1}\pm\hat{x}\cdot\vec{\sigma})/2$. For $a = +1$, the steered state is (see, for example, [25])

$$\sigma_{+|\hat{x}} = \operatorname{Tr}_A(M_{+|\hat{x}}\otimes\mathbb{1}\rho) = \tfrac{1}{4}[(1+\vec{a}\cdot\hat{x})\mathbb{1} + T\hat{x}\cdot\vec{\sigma}]. \tag{9}$$

Notice that the above state is diagonal in the basis $\{|\hat{s}\rangle, |-\hat{s}\rangle\}$, with Bloch vector $\hat{s} = \frac{T\hat{x}}{\|T\hat{x}\|}$; we omit the $\hat{x}$ dependence to simplify notation. The corresponding eigenvalues are

$$\alpha(\hat{x}) = \tfrac{1}{4}(1+\vec{a}\cdot\hat{x}+\|T\hat{x}\|), \quad \beta(\hat{x}) = \tfrac{1}{4}(1+\vec{a}\cdot\hat{x}-\|T\hat{x}\|). \tag{10}$$

Note that by construction $\alpha(\hat{x}) \geqslant \beta(\hat{x})$.

Our goal is now to construct a LHS model for this assemblage. First, the local hidden states σ_λ are taken to be pure qubit states and hence are represented by unit Bloch vectors $\hat{\lambda}$ and uniformly distributed over the sphere

$$\sigma_{\hat{\lambda}} = \frac{|\hat{\lambda}\rangle\langle\hat{\lambda}|}{4\pi}. \tag{11}$$

Normalization is ensured as $\int \operatorname{Tr}[\sigma_{\hat{\lambda}}]d\hat{\lambda} = 1$. This ensures that we obtain the correct reduced state for Bob:

$$\frac{1}{4\pi}\int |\hat{\lambda}\rangle\langle\hat{\lambda}|\, d\hat{\lambda} = \frac{\mathbb{1}}{2} = \rho_B. \tag{12}$$

Next we define Alice's response function to be given by the distribution

$$p(\pm|\hat{x},\hat{\lambda}) = \frac{1\pm\operatorname{sgn}[\hat{s}\cdot\hat{\lambda} - c(\hat{x})]}{2}, \tag{13}$$

parametrized by the function $-1 \leqslant c(\hat{x}) \leqslant 1$, with $\hat{s}$ the Bloch vector of the eigenvector of $\sigma_{+|\hat{x}}$ with the largest eigenvalue. The function (13) can be understood as follows (see Fig. 1). If $\hat{\lambda}$ is in the spherical cap centered on $\hat{s}$ such that

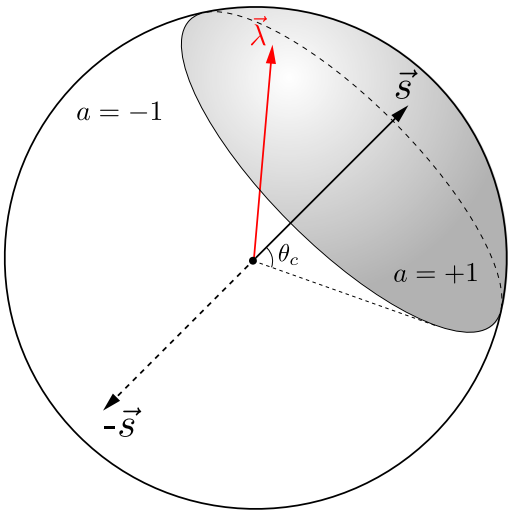

FIG. 1. Illustration of Alice's response function (13) in our LHS model. If $\mathrm{sgn}[\hat{s}\cdot\hat{\lambda} - c(\hat{x})] \geqslant 0$ then $a = +1$ (shaded spherical cap, with angle $\theta_c = \arccos[c]$); otherwise $a = -1$. The assemblage (14) then corresponds to the average (subnormalized) density matrix obtained by integrating pure qubit states $|\hat{\lambda}\rangle$ over the shaded region.

$\hat{\lambda}\cdot\hat{s} \geqslant c(\hat{x})$, the output is $a = +1$; otherwise $a = -1$. Note that we need only concentrate on the case $a = +1$; the case $a = -1$ is automatically satisfied from $\sigma_{+1|\hat{x}} + \sigma_{-1|\hat{x}} = \rho_B$ and (12). We now calculate the assemblage predicted by this model, given by

$$\sigma^{\mathrm{LHS}}_{+1|\hat{x}} = \int \sigma_{\hat{\lambda}}\, p(+|\hat{x},\hat{\lambda})d\hat{\lambda} = \frac{1}{4\pi}\int |\hat{\lambda}\rangle\langle\hat{\lambda}|\, p(+|\hat{x},\hat{\lambda})d\hat{\lambda}.$$

We parametrize the state $|\hat{\lambda}\rangle$ using the Bloch decomposition in the basis $\{|\hat{s}\rangle, |-\hat{s}\rangle\}$:

$$|\hat{\lambda}\rangle = |\hat{\lambda}(\theta,\phi)\rangle = \cos\frac{\theta}{2}|\hat{s}\rangle + \sin\frac{\theta}{2}e^{i\phi}|-\hat{s}\rangle\,. \tag{14}$$

Working in this basis and integrating over the spherical cap for which $a = +1$ (see Fig. 1), (14) becomes

$$\int_0^{2\pi}\int_0^{\theta_c}\begin{pmatrix}\cos^2\frac{\theta}{2} & \cos\frac{\theta}{2}\sin\frac{\theta}{2}e^{-i\phi}\\ \cos\frac{\theta}{2}\sin\frac{\theta}{2}e^{i\phi} & \sin^2\frac{\theta}{2}\end{pmatrix}\frac{\sin\theta\, d\phi\, d\theta}{4\pi},$$

where $\theta_c = \arccos[c(\hat{x})]$ is the angle of the spherical cap. Since $\int_0^{2\pi} e^{i\phi}d\phi = 0$, the off-diagonal components will be zero and $\sigma^{\mathrm{LHS}}_{+1|\hat{x}}$ is therefore diagonal in the $\{|\hat{s}\rangle, |-\hat{s}\rangle\}$ basis, as desired. From this, the eigenvalues of $\sigma^{\mathrm{LHS}}_{+1|\hat{x}}$, i.e., $\alpha'(\hat{x})$ and $\beta'(\hat{x})$, are given by

$$\alpha'(\hat{x}) + \beta'(\hat{x}) = \frac{1}{2}\int_0^{\theta_c}\sin\theta\, d\theta = \frac{1-\cos\theta_c}{2}, \tag{15}$$

$$\alpha'(\hat{x}) - \beta'(\hat{x}) = \frac{1}{2}\int_0^{\theta_c}\cos\theta\sin\theta\, d\theta = \frac{1-\cos^2\theta_c}{4}. \tag{16}$$

Upon using $\theta_c = \arccos[c(\hat{x})]$ one then finds

$$\alpha'(\hat{x}) + \beta'(\hat{x}) = \tfrac{1}{2}[1 - c(\hat{x})], \tag{17}$$

$$\alpha'(\hat{x}) - \beta'(\hat{x}) = \tfrac{1}{4}[1 - c^2(\hat{x})], \tag{18}$$

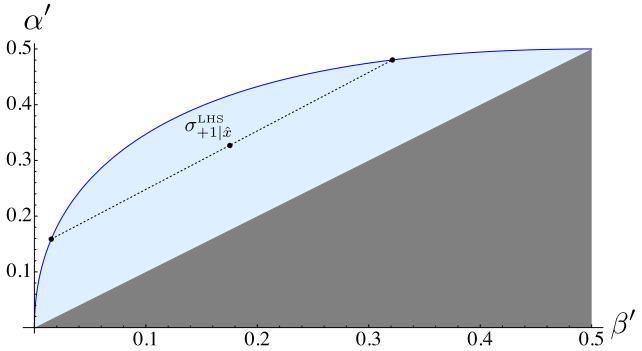

FIG. 2. Plot of the achievable range of eigenvalues (α',β') in our LHS model (for a fixed direction $\hat{s}$). The upper blue curve corresponds to the condition $\alpha' = \sqrt{2\beta'} - \beta'$ and is achieved by the response functions (13); any point in the light blue area below may be achieved by taking a suitable convex combination of these functions (e.g., the dashed line). Since we have $\alpha' \geqslant \beta'$, the gray area is not of interest.

from which we get the eigenvalues as a function of $c(\hat{x})$ as

$$\alpha'(\hat{x}) = \sqrt{2\beta'(\hat{x})} - \beta'(\hat{x}), \quad \beta'(\hat{x}) = \tfrac{1}{8}[1 - c(\hat{x})]^2, \tag{19}$$

corresponding to the curve of Fig. 2. Since this curve is concave, by fixing $\hat{s}$ and taking convex combinations of the response functions (13) with different $c(\hat{x})$, we may prepare any steered states corresponding to (α',β') below this curve, leading finally to

$$\beta'(\hat{x}) \leqslant \alpha'(\hat{x}) \leqslant \sqrt{2\beta'(\hat{x})} - \beta'(\hat{x}). \tag{20}$$

This corresponds to the blue area in Fig. 2. We thus conclude that the model reproduces the assemblage of any canonical state ρ, as long as its eigenvalues satisfy the above relation, i.e., $\alpha(\hat{x}) \leqslant \sqrt{2\beta(\hat{x})} - \beta(\hat{x})$, for any measurement vector $\hat{x}$, or equivalently

$$\max_{\vec{x}}\{[\alpha(\vec{x}) + \beta(\vec{x})]^2 - 2\beta(\vec{x})\} \leqslant 0. \tag{21}$$

Using (10) to convert this maximization into Bloch vector notation, we arrive at (8). ■

A natural question is whether condition (8) is also necessary for unsteerability. Unfortunately, this is not the case. Consider the state $\rho_c = \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|)$, which does not satisfy (8) [choose, e.g., $\hat{x} = (0,0,1)$], but is separable and hence clearly unsteerable. Note however that condition (8) can in fact be strengthened by considering convex combinations with separable states (see Appendix A). An interesting open question is then whether there exist unsteerable states, which cannot be written as convex combinations of unsteerable states satisfying condition (8) and separable states. Nevertheless, condition (8) turns out to be useful for proving the unsteerability of interesting classes of states, as we illustrate below.

IV. APPLICATIONS

We now illustrate the relevance of the above result with some applications. We consider the class of states

$$\rho(p,\chi) = p\,|\psi_\chi\rangle\langle\psi_\chi| + (1-p)\rho^A_\chi \otimes \mathbb{1}/2, \tag{22}$$

where $|\psi_\chi\rangle = \cos\chi\,|00\rangle + \sin\chi\,|11\rangle$ is a partially entangled two-qubit state, $\rho^A_\chi = \mathrm{Tr}_B\,|\psi_\chi\rangle\langle\psi_\chi|$, $p \in [0,1]$, and

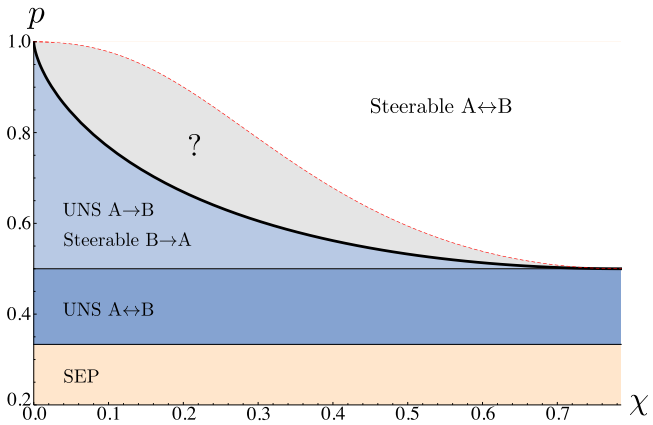

FIG. 3. Characterization of entanglement and steering for states $\rho(p,\chi)$. The solid black curve corresponds to (23), obtained from our unsteerability criterion. The state is separable in the light orange region, unsteerable (in both directions) in the dark blue region, unsteerable only from Alice to Bob (hence one-way steerable) in the light blue region, and two-way steerable in the white region [obtained from Eq. (19) of Ref. [25]]. What happens in the gray region is an interesting open question.

$\chi \in]0,\pi/4]$. The state is entangled for $p > 1/3$. From Theorem 1 it follows that $\rho(p,\chi)$ is unsteerable from Alice to Bob if

$$\cos^2(2\chi) \geqslant \frac{2p-1}{(2-p)p^3}, \tag{23}$$

as we show in Appendix B. This result is illustrated in Fig. 3 (black solid line). Note that our result recovers the case of a two-qubit Werner state $\rho(1/2,\pi/4)$, which admits a LHS model [18] (in both directions).

A. One-way steering

Alice and Bob play different roles in the steering scenario. Hence steerability in one direction (say from Alice to Bob) does not necessarily imply steerability in the other direction (from Bob to Alice). This effect of one-way steering was first observed in the context of continuous variable systems and Gaussian measurements [26]. More recently, an example of a two-qubit one-way steerable state was presented considering arbitrary projective measurements [22]. That is, while Alice can steer Bob using a finite number of measurements, it would be impossible for Bob to steer Alice as the state admits a LHS model (Bob to Alice). Moreover, a qutrit-qubit state was shown to be one-way steerable considering POVMs [6].

Clearly, our results are also useful for capturing one-way steering. Consider a given state ρ, the canonical form of which is found to satisfy condition (8). From Theorem 1 it follows that ρ is unsteerable from Alice to Bob. Moreover, if steerability from Bob to Alice can be verified using standard methods, e.g., via violation of a steering inequality or using semidefinite programming methods [14,15], one-way steerability of ρ is proven.

We present different examples of one-way steering. Our states of interest will be the states $\rho(p,\chi)$ defined above. This state is unsteerable from Alice to Bob for projective measurements when (23) is satisfied, corresponding to the area below the thick black line of Fig. 3. The steerability from Bob to Alice of the above state was discussed in previous works. In particular it was shown that $\rho(p,\chi)$ is unsteerable if $p \leqslant 1/2$ for all χ [20]. However, for $p > 1/2$, the state becomes steerable from Bob to Alice for all χ. This can be seen as follows. By applying on Alice's side the filter $F_\chi = \mathrm{diag}(1/\cos\chi, 1/\sin\chi)$, we obtain the state

$$\tfrac{1}{2} F_\chi \otimes \mathbb{1} \rho(p,\chi) F_\chi \otimes \mathbb{1} = \rho(p,\pi/4), \tag{24}$$

which is simply a Werner state with visibility p. Since this state is steerable for $p > 1/2$ [5], it follows from Lemma 1 that all states $\rho(p,\chi)$ with $p > 1/2$ and satisfying (23) are one-way steerable from Bob to Alice for projective measurements.

Simplest one-way steering

A relevant question to ask is how many measurements are needed in order to demonstrate one-way steering. So far, the only known examples for a two-qubit state required as many as 13 measurements [22] and considered only projective measurements and similarly for the qutrit-qubit example of Ref. [6]. Here we present the simplest possible example of one-way steering, that is, a two-qubit state such that Alice cannot steer Bob even with POVMs, although Bob can steer Alice using only two measurement settings.

We start with the case of projective measurements. We show that the states $\rho(p,\chi)$ with $p > 1/\sqrt{2}$ and satisfying (23) are one-way steerable and only two measurements are required for demonstrating steering from Bob to Alice. To prove this we proceed as follows. First, from Lemma 1 it is sufficient to consider the state $\rho(p,\pi/4)$, i.e., a Werner state [see Eq. (24)]. Since this state violates the Clauser-Horne-Shimony-Holt (CHSH) Bell inequality for $p > 1/\sqrt{2}$ [2], it is nonlocal and thus steerable from Bob to Alice with two measurements.

Next we move to the case of POVMs, building on the above example. Following protocol 2 of [31], we construct the state

$$\rho_{\mathrm{POVM}}(p,\chi) = \tfrac{1}{2}\rho(p,\chi) + \tfrac{1}{2}|0\rangle\langle 0| \otimes \rho_B, \tag{25}$$

where $\rho_B = \mathrm{Tr}_A\, \rho(p,\chi)$, which is now unsteerable from Alice to Bob for POVMs, for p and χ satisfying (23). We now show that steering from Bob to Alice is possible using only two measurements. From Lemma 1 we can focus our analysis on the state

$$\begin{aligned}\rho_F &= \frac{F_\chi \otimes \mathbb{1} \rho_{\mathrm{POVM}} F_\chi \otimes \mathbb{1}}{\mathrm{Tr}\,(F_\chi \otimes \mathbb{1} \rho_{\mathrm{POVM}} F_\chi \otimes \mathbb{1})} \\ &= \frac{\cos^2\chi\,(p|\phi^+\rangle\langle\phi^+| + (1-p)\mathbb{1}/4) + \frac{1}{2}|0\rangle\langle 0| \otimes \rho_B}{\cos^2\chi + 1/2}.\end{aligned} \tag{26}$$

Using the CHSH violation criterion [32], one can find the range of parameters such that ρ_F violates the CHSH inequality and is thus steerable form Bob to Alice with two measurements. We find a parameter range $p > 0.833\,53$ and corresponding χ given by condition (23).

B. Sufficient condition for joint measurability

Theorem 1 also finds application in quantum measurement theory. This follows from the direct connection existing

between steering and the notion of joint measurability of a set of quantum measurements [27,28], which has already found applications (see, e.g., [33]). This allows us to convert our sufficient condition for unsteerability into a sufficient condition for joint measurability of a set of qubit dichotomic POVMs. Notably, this condition is applicable to continuous sets of POVMs.

A set of measurements $\{M_{a|x}\}$ is said to be jointly measurable [34] if there exists a joint POVM $\{G_\lambda\}$ with outcomes λ and probability distributions $p(a|x,\lambda)$, from which the statistics of any of the measurements $\{M_{a|x}\}$ can be recovered by a suitable postprocessing, that is,

$$M_{a|x} = \int G_\lambda\, p(a|x,\lambda) d\lambda \;\forall a,x. \tag{27}$$

Let $\{M_{\pm|x}\}$ be a set of dichotomic qubit POVMs

$$M_{+|x} = \tfrac{1}{2}(k_x \mathbb{1} + \vec{m}_x \cdot \vec{\sigma}), \tag{28}$$

with $\|\vec{m}_x\| \leqslant k_x \leqslant 2 - \|\vec{m}_x\|$ and $M_{-|x} = \mathbb{1} - M_{+|x}$. Then the set $\{M_{\pm|x}\}$ is jointly measurable if

$$k_x(k_x - 2) + 2\|\vec{m}_x\| \leqslant 0 \tag{29}$$

for all x. This can be seen as follows. A set of measurements $\{M_{\pm|x}\}$ is jointly measurable if and only the assemblage given by $\sigma_{\pm|x} = \rho^{1/2} M_{\pm|x} \rho^{1/2}$, where ρ is a full-rank quantum state [35], is unsteerable. Choosing $\rho = \mathbb{1}/2$ we get the corresponding assemblage $\sigma_{\pm|x} = \rho^{1/2} M_{\pm|x} \rho^{1/2} = \frac{1}{2} M_{\pm|x}$. Following Theorem 1, condition (29) ensures the unsteerability of $\sigma_{\pm|x}$ and consequently the joint measurability of $\{M_{\pm|x}\}$.

V. CONCLUSION

We have presented a simple criterion sufficient for a qubit assemblage to admit a LHS model. Notably, our method can guarantee the unsteerability of a general two-qubit state and should thus find applications. We have shown that the criterion allows one to detect entangled states that are only one-way steerable and provides the simplest such examples. Moreover, the criterion is relevant to quantum measurement theory, as it provides a sufficient condition for a continuous set of dichotomic qubit POVMs to be jointly measurable. Further to this, our criterion has also found applications in the connection between measurement incompatibility and nonlocality [36] and multipartite nonlocality [37].

It would be interesting to extend this criterion in several directions. First, can the criterion be strengthened, e.g., by considering convex combinations, in order to become necessary and sufficient? Also, while we focused here on projective measurements, generalizing the method to POVMs would be useful.[1] Whether the present ideas can be adapted to the case of higher-dimensional systems (beyond qubits) is also a natural question. In particular, a natural case to consider is that of entangled states of dimension $d \times 2$, where our method should be directly applicable. Applications to multipartite steering [38,39] would also be interesting.

[1]Note that, starting from our result, one can construct new entangled states admitting a LHS model for all POVMs, using Protocol 2 of Ref. [31].

ACKNOWLEDGMENTS

We thank Charles Xu for discussions and acknowledge financial support from the Swiss National Science Foundation (Grant No. PP00P2_138917 and Starting Grant DIAQ).

APPENDIX A: CONVEX COMBINATIONS OF UNSTEERABLE STATES

Since our criterion (8) is not linear and since it does not detect all separable states, it can be useful to consider convex combination of states. Specifically, consider an entangled unsteerable state of the form

$$\rho = p\sigma + (1-p)\rho_{\text{SEP}}, \tag{A1}$$

where ρ_{SEP} is a separable (hence unsteerable) state and σ is an unspecified state. If σ is unsteerable, then it follows that ρ is unsteerable. However, it could be that, while ρ violates condition (8), σ does not. In this case, the unsteerability of ρ can be shown by finding suitable p and ρ_{SEP} such that

$$\sigma = \frac{\rho - (1-p)\rho_{\text{SEP}}}{p} \tag{A2}$$

satisfies condition (8).

As a simple example, consider the state

$$\rho = \tfrac{1}{2}\sigma + \tfrac{1}{2}\big(\tfrac{1}{2}\,|00\rangle\langle 00| + \tfrac{1}{2}\,|11\rangle\langle 11|\big), \tag{A3}$$

where σ is the two-qubit isotropic state $\sigma = (|\phi^+\rangle\langle\phi^+| + \mathbb{1}/4)/2$. Hence ρ is an equal mixture of σ and the separable classically correlated state. One finds that for ρ, $T_z = 3/4$ and so ρ violates (8) for $\hat{x} = (0,0,1)$. However, the state σ has $T = \mathbb{1}/2$ and $\vec{a} = \vec{0}$ and therefore satisfies (8), hence proving the unsteerability of ρ.

APPENDIX B: PROOF OF UNSTEERABILITY OF $\rho(p,\chi)$

Here we show that for the class of states (22), Theorem 1 implies that the $\rho(p,\chi)$ is unsteerable if

$$\cos^2 2\chi \geqslant \frac{2p-1}{(2-p)p^3}. \tag{B1}$$

To do this, we first consider states in canonical form (7), which satisfy $\vec{a} = (0,0,a_z)$ and $|T_x| = |T_y|$. In order to perform the maximization of Theorem 1, we parametrize $\hat{x}$ using spherical coordinates $\hat{x} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. Our criterion (8) may now be written as

$$\begin{aligned} &\max_{\theta,\phi} F(\theta,\phi) \leqslant 1, \\ &F(\theta,\phi) = (\vec{a}\cdot\hat{x})^2 + 2\|T\hat{x}\| \\ &\qquad = \cos^2\theta\, a_z^2 + 2\sqrt{T_x^2 + \cos^2\theta\,\left(T_z^2 - T_x^2\right)}. \end{aligned} \tag{B2}$$

Unsurprisingly, F depends only on θ since the problem is symmetric with respect to the x and y directions and we may ignore the maximization over ϕ. Note that if $|T_z| = |T_x|$ then the maximization occurs at $\theta = 0$ and our condition for unsteerability becomes

$$a_z^2 + 2|T_z| \leqslant 1. \tag{B3}$$

In the case $|T_z| \neq |T_x|$, one should find the extremal points of $F(\theta)$ and prove that they do not exceed 1. To find these

extrema we solve

$$\frac{dF}{d\theta} = -\sin 2\theta \left(a_z^2 + \frac{T_z^2 - T_x^2}{\sqrt{T_x^2 + \cos^2\theta\left(T_z^2 - T_x^2\right)}} \right) = 0. \tag{B4}$$

From $\sin 2\theta = 0$ we have solutions $\theta = 0, \pi/2, \pi$ and possibly other solutions given by

$$a_z^2 + \frac{T_z^2 - T_x^2}{\sqrt{T_x^2 + \cos^2\theta\left(T_z^2 - T_x^2\right)}} = 0. \tag{B5}$$

We now derive conditions such that (B5) has no solution. After rearranging (B5) we have

$$\cos^2\theta = \frac{T_x^2}{T_x^2 - T_z^2} - \frac{T_x^2 - T_z^2}{a_z^4}. \tag{B6}$$

This has no solution if the right-hand side is greater than 1 or less than 0. Hence we have two conditions

$$\frac{T_x^2}{T_x^2 - T_z^2} < \frac{T_x^2 - T_z^2}{a_z^4} \quad \text{or} \quad \frac{T_z^2}{T_x^2 - T_z^2} > \frac{T_x^2 - T_z^2}{a_z^4}. \tag{B7}$$

If one of the above conditions is fulfilled we therefore have extrema for $\theta = 0, \pi/2, \pi$ only. In this case, since $F(0) = F(\pi)$, our condition for unsteerability becomes

$$\max_\theta F(\theta) = \max\left\{a_z^2 + 2|T_z|, 2|T_x|\right\} \leqslant 1. \tag{B8}$$

We now move to the explicit case of $\rho(p,\chi)$. We find a canonical state with $|T_x| = |T_y|$, $\vec{a} = (0,0,a_z)$, and

$$\begin{aligned} a_z &= \frac{(1-p^2)\cos 2\chi}{1 - p^2\cos^2 2\chi}, \\ T_z &= \frac{p(1-\cos^2 2\chi)}{1 - p^2\cos^2 2\chi}, \\ T_x &= \sqrt{\frac{p^2(1-\cos^2 2\chi)}{1 - p^2\cos^2 2\chi}}. \end{aligned} \tag{B9}$$

We now introduce the ansatz (for $p \geqslant \frac{1}{2}$)

$$\cos^2 2\chi = \frac{2p-1}{(2-p)p^3}. \tag{B10}$$

Eliminating the variable χ we find

$$\begin{aligned} a_z^2 &= \frac{(2-p)(2p-1)}{p}, \\ T_z &= \frac{(1-p)^2}{p}, \\ T_x &= 1 - p. \end{aligned} \tag{B11}$$

For the case $p = \frac{1}{2}$ we have $|T_z| = |T_x|$ and we find that (B3) is satisfied. For $p > \frac{1}{2}$ we show that the second condition of (B7) holds. To this end, we calculate

$$\frac{T_z^2}{T_x^2 - T_z^2} - \frac{T_x^2 - T_z^2}{a_z^4} = \frac{(3-p)(1-p)^3}{(p-2)^2(2p-1)}. \tag{B12}$$

This is easily seen to be positive for $p \in]\frac{1}{2}, 1]$ and so $F(\theta)$ has extrema at $\theta = 0, \pi, \pi/2$ only. It therefore remains to prove (B8). We find

$$a_z^2 + 2|T_z| = 1, \quad 2|T_x| = 2(1-p) \tag{B13}$$

and so (B8) is satisfied for $p > \frac{1}{2}$. This proves that the state $\rho(p,\chi)$ is unsteerable if $p \geqslant \frac{1}{2}$ and p and χ satisfy (B10), which corresponds to the black curve of Fig. 3. Finally, we note that for a fixed χ, lowering p amounts to putting more weight on the separable part of the state. Since a convex combination of an unsteerable state with a separable state is also unsteerable, all points below the curve of Fig. 3 are also unsteerable. Hence, we arrive at (23).

[1] J. S. Bell, Physics (NY) **1**, 195 (1964).
[2] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Rev. Mod. Phys. **86**, 419 (2014).
[3] E. Schrödinger, Math. Proc. Cambridge Philos. Soc. **32**, 446 (1936).
[4] M. D. Reid, P. D. Drummond, W. P. Bowen, E. G. Cavalcanti, P. K. Lam, H. A. Bachor, U. L. Andersen, and G. Leuchs, Rev. Mod. Phys. **81**, 1727 (2009).
[5] H. M. Wiseman, S. J. Jones, and A. C. Doherty, Phys. Rev. Lett. **98**, 140402 (2007).
[6] M. T. Quintino, T. Vértesi, D. Cavalcanti, R. Augusiak, M. Demianowicz, A. Acín, and N. Brunner, Phys. Rev. A **92**, 032107 (2015).
[7] B. Wittmann *et al.*, New J. Phys. **14**, 053030 (2012); D. H. Smith *et al.*, Nat. Commun. **3**, 625 (2012); A. J. Bennet, D. A. Evans, D. J. Saunders, C. Branciard, E. G. Cavalcanti, H. M. Wiseman, and G. J. Pryde, Phys. Rev. X **2**, 031003 (2012).
[8] C. Branciard, E. G. Cavalcanti, S. P. Walborn, V. Scarani, and H. M. Wiseman, Phys. Rev. A **85**, 010301(R) (2012).
[9] M. Piani and J. Watrous, Phys. Rev. Lett. **114**, 060404 (2015).
[10] Y. Z. Law, L. P. Thinh, J.-D. Bancal, and V. Scarani, J. Phys. A: Math. Theor. **47**, 424028 (2014).
[11] E. G. Cavalcanti, S. J Jones, H. M. Wiseman, and M. D. Reid, Phys. Rev. A **80**, 032112 (2009).
[12] S. P. Walborn, A. Salles, R. M. Gomes, F. Toscano, and P. H. Souto Ribeiro, Phys. Rev. Lett. **106**, 130402 (2011).
[13] E. G. Cavalcanti, C. J. Foster, M. Fuwa, and H. M. Wiseman, J. Opt. Soc. Am. B **32**, A74 (2015).
[14] M. F. Pusey, Phys. Rev. A **88**, 032313 (2013).
[15] P. Skrzypczyk, M. Navascues, and D. Cavalcanti, Phys. Rev. Lett. **112**, 180404 (2014).
[16] I. Kogias, P. Skrzypczyk, D. Cavalcanti, A. Acín, and G. Adesso, Phys. Rev. Lett. **115**, 210401 (2015).
[17] H. Zhu, M. Hayashi, and L. Chen, Phys. Rev. Lett. **116**, 070403 (2016).
[18] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).
[19] J. Barrett, Phys. Rev. A **65**, 042302 (2002).
[20] M. L. Almeida, S. Pironio, J. Barrett, G. Tóth, and A. Acín, Phys. Rev. Lett. **99**, 040403 (2007).

[21] J. Bowles, F. Hirsch, M. T. Quintino, and N. Brunner, Phys. Rev. Lett. **114**, 120401 (2015).
[22] J. Bowles, T. Vértesi, M. T. Quintino, and N. Brunner, Phys. Rev. Lett. **112**, 200402 (2014).
[23] R. Augusiak, M. Demianowicz, and A. Acín, J. Phys. A: Math. Theor. **47**, 424002 (2014).
[24] S. Jevtic, M. Pusey, D. Jennings, and T. Rudolph, Phys. Rev. Lett. **113**, 020402 (2014).
[25] S. Jevtic, M. J. W. Hall, M. R. Anderson, M. Zwierz, and H. M. Wiseman, J. Opt. Soc. Am. B **32**, A40 (2015).
[26] S. L. W. Midgley, A. J. Ferris, and M. K. Olsen, Phys. Rev. A **81**, 022101 (2010).
[27] M. T. Quintino, T. Vértesi, and N. Brunner, Phys. Rev. Lett. **113**, 160402 (2014).
[28] R. Uola, T. Moroder, and O. Gühne, Phys. Rev. Lett. **113**, 160403 (2014).
[29] R. Gallego and L. Aolita, Phys. Rev. X **5**, 041008 (2015).
[30] A. Milne, D. Jennings, S. Jevtic, and T. Rudolph, Phys. Rev. A **90**, 024302 (2014).
[31] F. Hirsch, M. T. Quintino, J. Bowles, and N. Brunner, Phys. Rev. Lett. **111**, 160402 (2013).
[32] R. Horodecki, P. Horodecki, and M. Horodecki, Phys. Lett. A **200**, 340 (1995).
[33] T. Heinosaari, J. Kiukas, and D. Reitzner, Phys. Rev. A **92**, 022115 (2015).
[34] P. Busch, P. Lahti, and P. Mittelstaedt, *The Quantum Theory of Measurement*, Lecture Notes in Physics Monographs Vol. 2 (Springer, Berlin, 1996), pp. 25–90.
[35] R. Uola, C. Budroni, O. Gühne, and J. P. Pellonpää, Phys. Rev. Lett. **115**, 230402 (2015).
[36] M. Quintino, J. Bowles, F. Hirsch, and N. Brunner, arXiv:1510.06722.
[37] J. Bowles, J. Francfort, M. Fillettaz, F. Hirsch, and N. Brunner, arXiv:1511.08401.
[38] D. Cavalcanti, P. Skrzypczyk, G. H. Aguilar, R. V. Nery, P. H. Souto Ribeiro, and S. P. Walborn, Nat. Commun. **6**, 7941 (2015).
[39] Q. Y. He and M. D. Reid, Phys. Rev. Lett. **111**, 250403 (2013).

Paper C

Incompatible quantum measurements admitting a local-hidden-variable model

Physical Review A **93**, 052115 (2016)

Marco Tlio Quintino, Joseph Bowles, Flavien Hirsch, and Nicolas Brunner

PHYSICAL REVIEW A **93**, 052115 (2016)

Incompatible quantum measurements admitting a local-hidden-variable model

Marco Túlio Quintino, Joseph Bowles, Flavien Hirsch, and Nicolas Brunner
Département de Physique Théorique, Université de Genève, 1211 Genève, Switzerland
(Received 28 October 2015; published 23 May 2016)

The observation of quantum nonlocality, i.e., quantum correlations violating a Bell inequality, implies the use of incompatible local quantum measurements. Here we consider the converse question. That is, can any set of incompatible measurements be used in order to demonstrate Bell inequality violation? Our main result is to construct a local hidden variable model for an incompatible set of qubit measurements. Specifically, we show that if Alice uses this set of measurements, then for any possible shared entangled state and any possible dichotomic measurements performed by Bob, the resulting statistics are local. This represents significant progress towards proving that measurement incompatibility does not imply Bell nonlocality in general.

DOI: 10.1103/PhysRevA.93.052115

I. INTRODUCTION

A key aspect of quantum theory is that certain observables cannot be jointly measured, in strong contrast with classical physics. This leads to many prominent quantum features, such as the uncertainty principle and information gain vs disturbance tradeoff, and plays a central role in quantum information processing [1]. The incompatibility of quantum observables is usually captured via the notion of commutativity: incompatible observables do not commute. However, quantum theory allows for more general measurements, so-called positive-operator valued measures (POVM), the incompatibility of which cannot be properly captured using commutativity [2]. Here a natural concept is that of joint measurability [3]. A set of POVMs is said to be jointly measurable if each one of them can be derived from coarse graining of one common POVM. Conversely, if such a joint POVM does not exist, the set is considered incompatible. The concept of joint measurability thus arguably provides a natural separation between classical and nonclassical sets of measurements.

A longstanding question is to understand the relation between the incompatibility of quantum measurements and quantum nonlocality [4,5], another key feature of quantum theory. When performing a set of well-chosen measurements on a shared entangled state, two distant observers can observe nonlocal correlations, i.e., which cannot be explained by a local (i.e., classical) model. The question is then how the nonclassicality of quantum measurements (i.e., their incompatibility) relates to the nonclassicality of quantum correlations detected via violation of a Bell inequality. While the observation of nonlocality implies the use of incompatible measurements (for both observers), the converse is not known. Specifically, the question is the following. For any possible set of incompatible measurements performed by one observer, can we always find a shared entangled state and a set of measurements for the second observer such that the resulting statistics will lead to Bell inequality violation?

In the case of projective measurements, the answer is positive, as proven many years ago [6]. For the case of POVMs, however, the question is much more difficult. In the simplest case of two dichotomic POVMs, Wolf *et al.* [7] proved that incompatibility is equivalent to violation of the Clauser-Horne-Shimony-Holt [8] inequality, confirming previous evidence [9,10]. However, their proof cannot be extended to the general case, as the joint measurability problem cannot be reduced to a pair of POVMs only [2]. For instance, it is possible to have a set of three POVMs which is incompatible, although any pair (among the three) is jointly measurable [11,12]. Recently, a strong connection between joint measurability and Einstein-Podolsky-Rosen (EPR) steering [13], a form of quantum nonlocality strictly weaker than Bell nonlocality [14], has been demonstrated [15–17], leading to interesting results in both areas [18]. More generally, the connection between measurement uncertainty and nonlocality in no-signaling theories has been discussed [19–21].

In the present work we show that a set of incompatible quantum measurements can admit a local-hidden-variable (LHV) model. Specifically, we consider a bipartite Bell test in which Alice performs a given nonjointly measurable set of qubit POVMs. We then show that the statistics of such an experiment, considering an arbitrary shared entangled state and any possible dichotomic measurements performed by Bob, can be exactly reproduced using only classical shared resources. In other words, this set of incompatible measurements, despite having some nonclassical feature, can never lead to nonlocal correlations (considering dichotomic measurements for Bob). A parallel can be drawn to the study, initiated by Werner [22], of quantum states which are entangled (hence nonclassical) but nevertheless admit a LHV model;

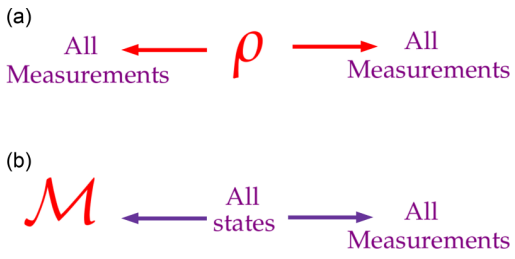

FIG. 1. The problem of classically simulating quantum correlations has two facets. (a) Constructing a LHV model for a given entangled quantum state ρ, considering arbitrary local measurements for Alice and Bob. (b) Constructing a LHV model for a given set of incompatible measurements $\mathcal{M}$ (performed by Alice), considering arbitrary entangled states and arbitrary local measurements Bob. While question (a) has been extensively studied, much less is known about question (b), which is the focus of this work.

see, e.g., [23–28] and [29] for a recent review. In contrast, we show that a set of nonclassical measurements admits a LHV model (see Fig. 1). Finally, we discuss the perspective of extending our result to the most general Bell test, which would thus demonstrate that incompatibility does not imply Bell nonlocality in general.

II. PRELIMINARIES

We start by introducing concepts and notations. Consider a set of N POVMs, given by operators $M_{a|x}$ satisfying $\sum_a M_{a|x} = \mathbb{1}$, $M_{a|x} \geqslant 0$ for $x \in \{1,\ldots,N\}$. This set is said to be jointly measurable if there exists one common POVM, $M_{\vec{a}}$, with outcomes $\vec{a} = [a_{x=1}, a_{x=2}, \ldots, a_{x=N}]$, where a_x gives the outcome of measurement x, that is,

$$M_{\vec{a}} \geqslant 0, \quad \sum_{\vec{a}} M_{\vec{a}} = \mathbb{1}, \quad \sum_{\vec{a}\setminus a_x} M_{\vec{a}} = M_{a|x}\ , \tag{1}$$

where $\vec{a} \setminus a_x$ stands for the elements of $\vec{a}$ except for a_x. Hence, all POVM elements $M_{a|x}$ are recovered as marginals of the *joint observable* $M_{\vec{a}}$. Notably, joint measurability of a set of POVMs does not imply that they commute [30]. Moreover, partial joint measurability does not imply full joint measurability in general [2], contrary to commutation. More generally, any partial compatibility configuration can be realized in quantum theory [31].

The focus of this work is to connect the incompatibility of a set of measurements to quantum nonlocality. We thus consider a Bell scenario featuring two observers, Alice and Bob, sharing an entangled state ρ. Alice and Bob perform local measurements, represented by operators $M_{a|x}$ and $M_{b|y}$. Here x and y denote the choice of measurement settings, while a and b denote the outcomes. The resulting probability distribution is thus given by $p(ab|xy) = \mathrm{tr}(\rho M_{a|x} \otimes M_{b|y})$. This distribution is local (in the sense of Bell) if it admits a decomposition of the form

$$p(ab|xy) = \int d\lambda q(\lambda) p_A(a|x,\lambda) p_B(b|y,\lambda). \tag{2}$$

Here the local model consists of a classical (hidden) variable λ, distributed according to density $q(\lambda)$, and Alice's and Bob's local response functions represented by the probability distributions $p_A(a|x,\lambda)$ and $p_B(b|y,\lambda)$. On the contrary, if a decomposition of the form (2) cannot be found, the distribution $p(ab|xy)$ is termed nonlocal and violates (at least) one Bell inequality [4,5].

It is straightforward to show that if the set of Alice's measurements, $\mathcal{M}_A = \{M_{a|x}\}$, is jointly measurable, the resulting distribution $p(ab|xy)$ is local, for any possible entangled state ρ and arbitrary measurements of Bob; see, e.g., [16]. Indeed, if the set $\mathcal{M}_A$ is compatible, then Alice can recover all statistics from one joint observable. Clearly, no Bell inequality violation can be obtained if Alice always performs the same measurement.

The main goal of this work is to discuss the converse problem. Specifically, given that the set $\mathcal{M}_A$ is incompatible, what can we say about the locality of the distribution $p(ab|xy)$? Previous work [7] demonstrated a striking connection in the simplest case, when $\mathcal{M}_A$ consists of two dichotomic POVMs. Any set $\mathcal{M}_A$ that is not jointly measurable can be used to demonstrate nonlocality. Whether this connection holds for more general sets of POVMs has been an open question since then. Here we show that for certain incompatible sets of POVMs, the resulting distribution $p(ab|xy)$ is always local, considering arbitrary entangled states ρ and arbitrary dichotomic measurements on Bob's side [32].

III. MAIN RESULT

We consider the continuous set of dichotomic qubit POVMs, $\mathcal{M}_A^\eta = \{M_{\pm|\hat{x}}^\eta\}$, with elements

$$M_{\pm|\vec{x}}^\eta = \tfrac{1}{2}(\mathbb{1} \pm \eta\, \hat{x}\cdot\vec{\sigma}) \tag{3}$$

with binary outcome $a = \pm 1$. Here $\hat{x}$ is any vector on the Bloch sphere denoting the measurement direction, and $\vec{\sigma} = (\sigma_1,\sigma_2,\sigma_3)$ is the vector of Pauli matrices. Note that the set $\mathcal{M}_A^\eta$ features a parameter $0 \leqslant \eta \leqslant 1$, representing basically the purity of the POVM elements. For $\eta = 1$, all POVM elements are projectors,

$$\Pi_{\pm|\hat{x}} = \tfrac{1}{2}(\mathbb{1} \pm \hat{x}\cdot\vec{\sigma}). \tag{4}$$

The set $\mathcal{M}_A^{\eta=1}$ is simply the set of all qubit projective measurements and is thus clearly incompatible. For $\eta = 0$, the set contains only the identity (thus clearly compatible). In general the set $\mathcal{M}_A^\eta$ contains noisy measurements, with elements simply given by $M_{\pm|\hat{x}}^\eta = \eta\Pi_{\pm|\hat{x}} + (1-\eta)\mathbb{1}/2$. In fact, the set $\mathcal{M}_A^\eta$ is jointly measurable if and only if $\eta \leqslant 1/2$ [15,16].

Below we will show that there is $\eta^* > 1/2$ such that the set $\mathcal{M}_A^{\eta^*}$ is local in any Bell test, considering arbitrary states ρ and arbitrary dichotomic measurements for Bob. Since $\mathcal{M}_A^{\eta^*}$ is not jointly measurable, this shows that incompatibility is not sufficient for Bell inequality violation in this case. Below we give a full proof of the result, proceeding in several steps.

The first step consists in exploiting the symmetries of the problem in order to find the minimal set of states ρ we need to consider. By linearity of the problem—the probabilities $p(ab|xy)$ are linear in ρ, and the set of local correlations is convex, see, e.g., [5]—we can safely focus on pure states. Indeed, if there was a mixed state ρ leading to Bell inequality violation using measurements $\mathcal{M}_A^{\eta^*}$, there would also be a pure state doing so.

Next, given that $\mathcal{M}_A^{\eta^*}$ consists only of qubit measurements, Alice's subsystem can be considered to be a qubit. Moreover, since we are free to choose convenient local reference frames (i.e., we can apply any local unitaries on Alice and Bob's systems), the shared state ρ (of dimension $2 \times d$) can therefore be expressed in the Schmidt form [1], i.e., $\rho = |\phi_\theta\rangle\langle\phi_\theta|$ with

$$|\phi_\theta\rangle = \cos\theta|00\rangle + \sin\theta|11\rangle \tag{5}$$

and $\theta \in [0,\pi/4]$.

Now we introduce the measurements on Bob's side. Since Bob's system is of rank 2, we can focus here on dichotomic qubit measurements. As any such POVM can be viewed as a projective qubit measurement followed by classical

postprocessing [33], it is sufficient to discuss projective qubit measurements $\Pi_{b|\hat{y}} = (\mathbb{1} + b\,\hat{y}\cdot\vec{\sigma})/2$, where $\hat{y}$ is any vector on the Bloch sphere and $b = \pm 1$.

Our goal is thus to show that there exists $\eta^* > 1/2$ such that the distribution

$$p(ab|xy) = \mathrm{tr}\left(|\phi_\theta\rangle\langle\phi_\theta| M^{\eta^*}_{a|\hat{x}} \otimes \Pi_{b|\hat{y}}\right) \tag{6}$$

is local for any measurement directions $\hat{x}$ and $\hat{y}$, and any state $|\phi_\theta\rangle$. In other words we would like to construct a LHV model for the incompatible set of measurements $\mathcal{M}^{\eta^*}_A$. In order to do so, we start by reformulating the problem by making use of the following relation:

$$\mathrm{tr}\left(|\phi_\theta\rangle\langle\phi_\theta| M^{\eta}_{a|\hat{x}} \otimes \Pi_{b|\hat{y}}\right) = \mathrm{tr}\left(\rho^\eta_\theta \Pi_{a|\hat{x}} \otimes \Pi_{b|\hat{y}}\right) \tag{7}$$

where

$$\rho^\eta_\theta = \eta|\phi_\theta\rangle\langle\phi_\theta| + (1-\eta)\tfrac{\mathbb{1}}{2} \otimes \rho_B \tag{8}$$

and $\rho_B = \mathrm{tr}_A(|\phi_\theta\rangle\langle\phi_\theta|)$. Thus, the problem of constructing a LHV model for $\mathcal{M}^{\eta^*}_A$ (considering dichotomic measurements for Bob) is equivalent to the problem of constructing a LHV model for the class of states $\rho^{\eta^*}_\theta$ (for all $\theta \in [0,\pi/4]$) with arbitrary projective measurements for Alice and Bob. Importantly, it must be shown that $\rho^{\eta^*}_\theta$ admits a LHV model for all $\theta \in [0,\pi/4]$ and for a fixed $\eta^* > 1/2$ (independent of θ).

The locality of the states $\rho^{\eta^*}_\theta$ must be discussed in two steps for different ranges of the parameter θ. First consider the range $\theta \in [0,\pi/4-\epsilon]$ with $\epsilon > 0$. Recently, we presented a sufficient condition for a two-qubit state to admit a LHV model for projective measurements [28]. For states of the form ρ^η_θ, a LHV model was shown to exist given that

$$\cos^2(2\theta) \geqslant \frac{2\eta - 1}{(2-\eta)\eta^3}. \tag{9}$$

Hence for any θ, we get a corresponding value of η for which the state is provably local; see Fig. 2. This clearly guarantees that for $\theta \in [0,\pi/4-\epsilon]$, with $\epsilon > 0$ fixed, we can find

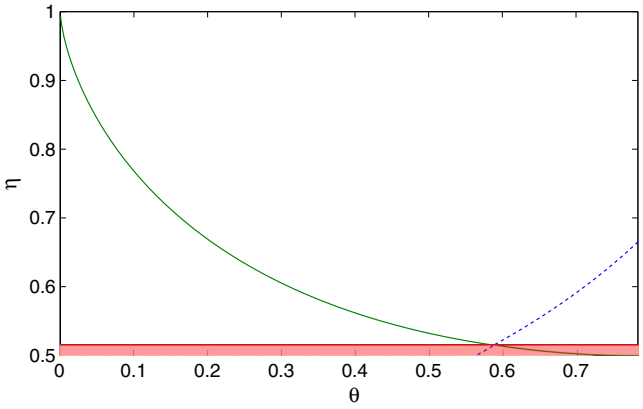

FIG. 2. Parameter region for which the state ρ^η_θ admits a LHV model: first, below the green curve, as given by Eq. (9), and second, below the blue dashed curve, as found via the SDP (14). The two curves cross at $\eta^* \simeq 0.515$. It follows that the state ρ^η_θ is local for $\eta \leqslant \eta^*$ and for all θ, i.e., in the shaded region, below the red horizontal line.

$\eta^* > 1/2$ such that $\rho^{\eta^*}_\theta$ is local. However, when θ gets closer to $\pi/4$, this approach will not work. Indeed, there is no fixed value $\eta^* > 1/2$ for which locality can be guaranteed for any $\theta \in [0,\pi/4]$, as can be seen by continuity of Eq. (9) or from Fig. 2. We thus need to find a different approach for this regime.

We proceed as follows. First note that for the case $\theta = \pi/4$, the state ρ^η_θ is simply a two-qubit Werner state,

$$\rho^\mu_W = \mu|\phi_+\rangle\langle\phi_+| + (1-\mu)\tfrac{\mathbb{1}}{4}, \tag{10}$$

with $|\phi_+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$. Coincidentally, such states admit a LHV model for $\mu \leqslant \mu_{LHV} \simeq 0.66$, considering arbitrary projective measurements [24]. The case $\theta = \pi/4$ is thus covered. Let us next discuss the case of θ in the neighborhood of $\pi/4$. To do so we consider the problem of decomposing the target state ρ^η_θ as a mixture of states admitting a LHV model. Specifically, we demand for which values of θ and η, we can find a convex combination of the form

$$\rho^\eta_\theta = \alpha\rho^{\mu_{LHV}}_W + (1-\alpha)\sigma \tag{11}$$

with $0 \leqslant \alpha \leqslant 1$. Here σ is an unspecified two-qubit state, which we are free to choose. As long as σ admits a LHV model, this implies that ρ^η_θ is local. In order to do so, we simply ensure that

$$\sigma = \frac{\rho^\eta_\theta - \alpha\rho^{\mu_{LHV}}_W}{1-\alpha} \tag{12}$$

is a valid separable state. By setting $\alpha = \frac{1}{\mu_{LHV}}\eta\sin(2\theta)$, we obtain a diagonal matrix σ (for all η and θ). It is straightforward to check that the eigenvalues of σ are positive when

$$\eta \leqslant \frac{\mu_{LHV}}{(1+\mu_{LHV})\cot\theta - \mu_{LHV}}. \tag{13}$$

By combining condition (9) and the above result, it follows that the state ρ^η_θ admits a LHV model for any θ and for $\eta \leqslant \eta^* \simeq 0.503$. Note that a better bound can be obtained using numerical methods. Consider again the problem of finding a decomposition of the form (11) with σ a separable state. For fixed θ, the optimal decomposition can be found via semidefinite programming (SDP):

$$\begin{aligned} &\max\ \eta \\ &\text{s.t. } \rho^\eta_\theta = \alpha\rho^{\mu_{LHV}}_W + \sigma \\ &\sigma \geqslant 0, \quad \sigma^{PT} \geqslant 0, \\ &\mathrm{Tr}\,\sigma + \alpha = 1, \quad \alpha \geqslant 0. \end{aligned} \tag{14}$$

Here σ^{PT} denotes the partial transpose [34] of σ. Verifying that σ^{PT} is positive ensures here that σ is separable [35]. The result of this optimization procedure is shown in Fig. 2. Combining again with condition (9) we get that ρ^η_θ admits a LHV model for $\eta \leqslant \eta^* \simeq 0.515$ (for any θ), for all projective measurements for Alice and Bob.

We therefore conclude that in the range $1/2 < \eta^* \lesssim 0.515$, the set of measurements $\mathcal{M}^{\eta^*}_A$ is incompatible and admits a LHV model. Specifically, $\mathcal{M}^{\eta^*}_A$ can never lead to Bell inequality violation, considering arbitrary shared entangled

states and arbitrary dichotomic measurements performed by the second observers.

Finally, it is worth mentioning that this result can be straightforwardly extended to the case of a set containing only a finite number of incompatible measurements. For instance, we have checked that a set of 12 well-chosen POVMs in $\mathcal{M}_A^\eta$ (chosen rather uniformly on the Bloch sphere) is incompatible for $\eta > 0.512$ via standard SDP techniques [7]. However, this set clearly admits a LHV model for $\eta \lesssim 0.515$.

It would be interesting to see if the result also holds in the simplest case of a set of only three POVMs. Consider, for instance, the three Pauli operators: σ_x, σ_y, and σ_z. Adding noise as in Eq. (3), the resulting POVMs are pairwise jointly measurable, but still not fully jointly measurable, in the range $1/\sqrt{3} < \eta \leqslant 1/\sqrt{2}$ [11,12]. Could such a set of three POVMs admit a LHV model?

IV. DISCUSSION

We discussed the relation between measurement incompatibility and Bell nonlocality. Specifically, we showed that a given set of incompatible qubit measurements can never lead to Bell inequality violation, as it admits a LHV model. Our construction covers the case of any possible shared entangled state and all possible dichotomic measurements performed by the second observer.

The main open question now is whether our result can be extended to nondichotomic measurements on Bob's side. If possible, this would then prove that measurement incompatibility does not imply Bell nonlocality in general [36].

We believe that the prospects for extending our LHV model for the set of measurements $\mathcal{M}_A^\eta$ to general measurements on Bob's side is promising. More precisely, following our approach, this amounts to show that the states ρ_θ^η of Eq. (8) (for a fixed $\eta > 1/2$ and all θ) admit a LHV model, considering arbitrary projective measurements for Alice and arbitrary POVMs for Bob [37]. We conjecture that this is the case, which is also supported by the fact that, so far, there is no example of an entangled state admitting a LHV model for projective measurements but not for POVMs.

ACKNOWLEDGMENTS

We thank Matt Pusey and Tamás Vértesi for discussions. We acknowledge financial support from the Swiss National Science Foundation (Grant No. PP00P2_138917 and Starting Grant DIAQ).

[1] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, Cambridge Series on Information and the Natural Sciences (Cambridge University Press, Cambridge, 2000).

[2] K. Kraus, A. Bohm, J. Dollard, and W. Wootters, *States, Effects, and Operations: Fundamental Notions of Quantum Theory*, Lecture Notes in Physics (Springer-Verlag, Berlin, 1983).

[3] P. Busch, P. Lahti, and P. Mittelstaedt, in *The Quantum Theory of Measurement*, Environmental Engineering Vol. 2 (Springer, Berlin, 1996).

[4] J. S. Bell, On the Einstein-Poldolsky-Rosen paradox, Physics **1**, 195 (1964).

[5] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).

[6] L. A. Khalfin and B. S. Tsirelson, Quantum and quasi-classical analogs of Bell inequalities, in *Symposium on the Foundations of Modern Physics*, edited by Lahti *et al.* (World Scientific, Singapore, 1985), pp. 441–460.

[7] 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).

[8] 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).

[9] E. Andersson, S. M. Barnett, and A. Aspect, Joint measurements of spin, operational locality, and uncertainty, Phys. Rev. A **72**, 042104 (2005).

[10] W. Son, E. Andersson, S. M. Barnett, and M. S. Kim, Joint measurements and Bell inequalities, Phys. Rev. A **72**, 052116 (2005).

[11] T. Heinosaari, D. Reitzner, and P. Stano, Notes on joint measurability of quantum observables, Found. Phys. **38**, 1133 (2008).

[12] Y.-C. Liang, R. W. Spekkens, and H. M. Wiseman, Specker's parable of the overprotective seer: A road to contextuality, nonlocality, and complementarity, Phys. Rep. **506**, 1 (2011).

[13] 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).

[14] M. T. Quintino, T. Vértesi, D. Cavalcanti, R. Augusiak, M. Demianowicz, A. Acín, and N. Brunner, Inequivalence of entanglement, steering, and Bell nonlocality for general measurements, Phys. Rev. A **92**, 032107 (2015).

[15] R. Uola, T. Moroder, and O. Gühne, Joint Measurability of Generalized Measurements Implies Classicality, Phys. Rev. Lett. **113**, 160403 (2014).

[16] M. T. Quintino, T. Vértesi, and N. Brunner, Joint Measurability, Einstein-Podolsky-Rosen Steering, and Bell Nonlocality, Phys. Rev. Lett. **113**, 160402 (2014).

[17] R. Uola, C. Budroni, O. Gühne, and J.-P. Pellonpää, A One-to-One Mapping Between Steering and Joint Measurability Problems, Phys. Rev. Lett. **115**, 230402 (2015).

[18] T. Heinosaari, J. Kiukas, D. Reitzner, and J. Schultz, Incompatibility breaking quantum channels, J. Phys. A: Math. Theor. **48**, 435301 (2015).

[19] J. Oppenheim and S. Wehner, The uncertainty principle determines the nonlocality of quantum mechanics, Science **330**, 1072 (2010).

[20] C. Pfister and S. Wehner, An information-theoretic principle implies that any discrete physical theory is classical, Nat. Commun. **4**, 1851 (2013).

[21] M. Banik, Measurement incompatibility and Schrödinger-Einstein-Podolsky-Rosen steering in a class of probabilistic theories, J. Math. Phys. **56**, 052101 (2015).

[22] R. F. Werner, Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model, Phys. Rev. A **40**, 4277 (1989).

[23] J. Barrett, Nonsequential positive-operator-valued measurements on entangled mixed states do not always violate a Bell inequality, Phys. Rev. A **65**, 042302 (2002).

[24] A. Acín, N. Gisin, and B. Toner, Grothendieck's constant and local models for noisy entangled quantum states, Phys. Rev. A **73**, 062105 (2006).

[25] 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).

[26] F. Hirsch, M. T. Quintino, J. Bowles, and N. Brunner, Genuine Hidden Quantum Nonlocality, Phys. Rev. Lett. **111**, 160402 (2013).

[27] J. Bowles, F. Hirsch, M. T. Quintino, and N. Brunner, Local Hidden Variable Models for Entangled Quantum States Using Finite Shared Randomness, Phys. Rev. Lett. **114**, 120401 (2015).

[28] J. Bowles, F. Hirsch, M. T. Quintino, and N. Brunner, Sufficient criterion for guaranteeing that a two-qubit state is unsteerable, Phys. Rev. A **93**, 022121 (2016).

[29] R. Augusiak, M. Demianowicz, and A. Acín, Local hidden-variable models for entangled quantum states, J. Phys. A: Math. Theor. **47**, 424002 (2014).

[30] P. Kruszyski and W. de Muynck, Compatibility of observables represented by positive operator-valued measures, J. Math. Phys. **28**, 1761 (1987).

[31] R. Kunjwal, C. Heunen, and T. Fritz, All joint measurability structures are quantum realizable, Phys. Rev. A **89**, 052126 (2014).

[32] One could also consider a simpler question, where both Alice and Bob use the same set of measurements. In this case, however, the problem is much easier, and examples of incompatible sets not leading to Bell violation can be found in Ref. [16].

[33] G. Mauro D'Ariano, P. Lo Presti, and P. Perinotti, Classical randomness in quantum measurements, J. Phys. A: Math. Gen. **38**, 5979 (2005).

[34] A. Peres, Separability Criterion for Density Matrices, Phys. Rev. Lett. **77**, 1413 (1996).

[35] M. Horodecki, P. Horodecki, and R. Horodecki, Separability of mixed states: Necessary and sufficient conditions, Phys. Lett. A **223**, 1 (1996).

[36] One may also wonder about multipartite Bell tests. Consider, for instance, a tripartite Bell test, where A performs measurements $\mathcal{M}_A^\eta$. Clearly, nonlocality can be obtained if B and C simply ignore A and make appropriate measurements on a shared singlet. Hence the relevant question here is whether nonlocality can be obtained on the partition $A|BC$, which brings us back to a bipartite Bell test (with arbitrary outputs on Bob's side).

[37] Note that it is sufficient to consider the cases of three- and four-output qubit POVMs, as extremal qubit POVMs have (at most) four outcomes [33].

Paper D

Genuine Hidden Quantum Nonlocality

Physical Review Letters **111**, 160402 (2013)

Flavien Hirsch, Marco Tlio Quintino, Joseph Bowles, and Nicolas Brunner

Genuine Hidden Quantum Nonlocality

Flavien Hirsch,[1] Marco Túlio Quintino,[1] Joseph Bowles,[1] and Nicolas Brunner[1,2]
[1]*Département de Physique Théorique, Université de Genève, 1211 Genève, Switzerland*
[2]*H.H. Wills Physics Laboratory, University of Bristol, Bristol BS8 1TL, United Kingdom*
(Received 19 July 2013; published 16 October 2013)

The nonlocality of certain quantum states can be revealed by using local filters before performing a standard Bell test. This phenomenon, known as hidden nonlocality, has been so far demonstrated only for a restricted class of measurements, namely, projective measurements. Here, we prove the existence of genuine hidden nonlocality. Specifically, we present a class of two-qubit entangled states, for which we construct a local model for the most general local measurements, and show that the states violate a Bell inequality after local filtering. Hence, there exist entangled states, the nonlocality of which can be revealed only by using a sequence of measurements. Finally, we show that genuine hidden nonlocality can be maximal. There exist entangled states for which a sequence of measurements can lead to maximal violation of a Bell inequality, while the statistics of nonsequential measurements is always local.

DOI: 10.1103/PhysRevLett.111.160402 PACS numbers: 03.65.Ud

Performing local measurements on separated entangled particles can lead to nonlocal correlations, as witnessed by the violation of a Bell inequality [1]. This phenomenon, termed quantum nonlocality, has received strong experimental confirmation. Moreover, entanglement and nonlocality are now viewed as fundamental aspects of quantum theory and play a prominent role in quantum information [2,3].

However, 50 years after the discovery of Bell's theorem, we still do not fully understand the relation between entanglement and nonlocality, although significant progress was made [3]. In particular, the most natural question, of which entangled states can lead to nonlocal correlations and which ones cannot, is still open. While it is known that nonlocality is a generic feature for pure entangled states [4,5], the situation for mixed states turns out to be much more complex. First, Werner [6] showed that there exist mixed entangled states (so-called Werner states) that admit a local model for projective measurements. However, it could still be the case that such states violate a Bell inequality when more general measurements, i.e. positive operator value measures (POVMs), are considered. Motivated by this question, Barrett [7] showed that certain noisy Werner states (but nevertheless entangled) admit a local model even when POVMs are considered (see also Ref. [8]).

Another twist to this question was given in Refs. [9,10], proposing Bell tests where observers perform a sequence of measurements—rather than a single measurement. Notably, Popescu [9] showed that Werner states of local dimension $d \geq 5$ can violate a Bell inequality when judicious local filters are applied to the state before performing a standard Bell test. Hence, the local filters reveal the hidden nonlocality of the quantum state. Importantly, the use of local filters does not open any loophole, since the choice of local measurement settings (for the second measurement) can be performed after applying the filters [9,11,12]. While this result shows that sequential measurements can be beneficial in Bell tests, it raises the question of whether they are necessary. Indeed, the crucial point here is that hidden nonlocality has been so far demonstrated only for a restricted class of measurements, namely, projective measurements. Specifically, the Werner states considered by Popescu admit a local model for projective measurements but could in principle violate a Bell inequality when POVMs are considered. Indeed, POVMs are proven to be relevant in Bell tests, as they can increase Bell violation compared to projective measurements [13]. Hence, this raises the question of whether there exists *genuine hidden nonlocality*. That is, do there exist entangled states, the nonlocality of which can be observed only if sequential measurements are used?

Here, we prove the existence of genuine hidden nonlocality. Specifically, we start by presenting a simple class of two-qubit entangled states, for which we construct a local model for POVMs, i.e., arbitrary nonsequential measurements. Next, we show that these states violate the Clauser-Horne-Shimony-Holt (CHSH) [14] Bell inequality when a judiciously chosen sequence of measurements is performed. Hence, this shows that sequential measurements outperform nonsequential ones, and that the nonlocality of certain entangled states can be revealed only through a sequence of measurements. Moreover, our construction provides the simplest example of hidden nonlocality known so far. A central tool for deriving our result is a technique which allows us, starting from a local model for simulating dichotomic projective measurements on a given state, to construct a local model for simulating POVMs on a related (but in general different) state. Finally, we demonstrate that genuine hidden nonlocality can be maximal. Specifically, we present a simple class of qutrit-qutrit entangled states which admit a local model for

POVMs but violate maximally the CHSH inequality when a sequence of measurements is used. Hence, such states are useful resources for information-theoretic tasks based on nonlocality [2,3], although they seem useless at first sight. These results highlight novel aspects of the subtle relation between entanglement and nonlocality.

We start by introducing the scenario and notations. Consider a bipartite Bell scenario in which distant parties, Alice and Bob, perform local measurements on an entangled state ρ of local Hilbert space dimension d. The choice of measurement setting is denoted by x for Alice (y for Bob), and the measurement outcome by a (b for Bob). Each setting is represented by a collection of positive operators acting on $\mathbb{C}^d$, denoted here as $M_{a|x}$ and $M_{b|y}$, satisfying the relations $\sum_a M_{a|x} = \mathbb{1}$ and $\sum_b M_{b|y} = \mathbb{1}$, where $\mathbb{1}$ denotes the identity operator in dimension d. The experiment is then characterized by the joint probability distribution

$$p(ab|xy) = \mathrm{Tr}(M_{a|x} \otimes M_{b|y}\rho). \tag{1}$$

If the distribution $p(ab|xy)$ violates (at least) one Bell inequality, the state ρ is said to be nonlocal. If, on the other hand, the distribution admits a decomposition

$$p(ab|xy) = \int d\lambda\, \omega(\lambda) p(a|x\lambda) p(b|y\lambda) \tag{2}$$

for all possible measurements, the state ρ admits a local model and cannot violate any Bell inequality. Here, λ represents the local hidden variable, distributed according to the density $\omega(\lambda)$. We will consider two separate cases. First, when a decomposition of the form (2) can be found for all projective measurements (i.e., $M_{a|x}^2 = M_{a|x}$ and $M_{b|y}^2 = M_{b|y}$), we say that ρ is local for projective measurements. Second, if a decomposition of the form (2) can be found for all POVMs (arbitrary nonsequential measurements), we say that ρ is local for POVMs.

So far, we have considered a Bell scenario in which each party performs a single measurement on its particle. One can, however, consider a more general measurement scenario, in which each party performs a sequence of measurements [9,10]. For instance, upon receiving their particle, the parties apply a local filtering. In the case that the filtering succeeds on both sides, the parties now hold the "filtered" state

$$\tilde{\rho} = \frac{1}{N}[(F_A \otimes F_B)\rho(F_A^\dagger \otimes F_B^\dagger)], \tag{3}$$

where $N = \mathrm{Tr}[(F_A \otimes F_B)\rho(F_A^\dagger \otimes F_B^\dagger)]$ is a normalization factor, and F_A and F_B are positive operators acting on $\mathbb{C}^d$ representing the local filtering of Alice and Bob. Finally, the parties perform local measurements on $\tilde{\rho}$ and can test a Bell inequality. Here, we will see that such a sequence of measurements is necessary in certain cases. More precisely, there exist entangled quantum states, the nonlocality of which can only be revealed by performing sequential measurements. Thus, such states exhibit genuine hidden nonlocality.

To demonstrate our main result, we proceed in several steps. First, we consider a simple class of entangled two-qubit states, of the form

$$\rho = q\Psi_- + (1-q)|0\rangle\langle 0| \otimes \frac{\mathbb{1}}{2}, \tag{4}$$

where $\Psi_- = |\psi_-\rangle\langle\psi_-|$ denotes the projector on the singlet state $|\psi_-\rangle = (|0,1\rangle - |1,0\rangle)/\sqrt{2}$, and $0 \le q \le 1$. Building upon the models discussed in Refs [15,16], we will see now that state (4) admits a local model for projective measurements when $q \le 1/2$, although it is entangled for all $q > 0$. Specifically, Alice and Bob receive as input a vector $\vec{x}$ and $\vec{y}$ and should simulate the statistics of measuring qubit observables $\vec{x}\cdot\vec{\sigma}$ and $\vec{y}\cdot\vec{\sigma}$ on ρ; here, $\vec{\sigma}$ denotes the vector of Pauli matrices; hence, the measurement outcomes are ± 1.

Protocol 1.—Alice and Bob share a three-dimensional unit vector $\vec{\lambda}$, uniformly distributed on the sphere. Upon receiving $\vec{x}$, Alice tests the shared vector $\vec{\lambda}$. With probability $|\vec{x}\cdot\vec{\lambda}|$, she "accepts" $\vec{\lambda}$ and outputs $a = -\mathrm{sgn}(\vec{x}\cdot\vec{\lambda})$; otherwise, she outputs $a = \pm 1$ with probability $(1 \pm \langle 0|\vec{x}\cdot\vec{\sigma}|0\rangle)/2$. Bob simply outputs $b = \mathrm{sgn}(\vec{y}\cdot\vec{\lambda})$.

The protocol consists of two parts. First, when Alice accepts $\vec{\lambda}$, which occurs on average with probability $1/2$ (independently of $\vec{x}$), $\vec{\lambda}$ is distributed according to the density $\omega(\vec{\lambda}) = |\vec{x}\cdot\vec{\lambda}|/2\pi$ [15,16]. In this case, the correlation between Alice's and Bob's outcomes is

$$\langle ab\rangle = -\frac{1}{2\pi}\int d\vec{\lambda}\,|\vec{x}\cdot\vec{\lambda}|\mathrm{sgn}(\vec{x}\cdot\vec{\lambda})\mathrm{sgn}(\vec{y}\cdot\vec{\lambda}) = -\vec{x}\cdot\vec{y}, \tag{5}$$

where the integral is taken over the sphere. As the marginals are uniform, i.e., $\langle a\rangle = \langle b\rangle = 0$, we recover the singlet correlations. Second, when Alice rejects $\vec{\lambda}$, she simulates locally the statistics of the state $|0\rangle$, while Bob's outcome is uncorrelated. Hence, the model reproduces exactly the statistics of the state (4) for $q = 1/2$, i.e., $\langle ab\rangle = (-\vec{x}\cdot\vec{y})/2$, $\langle a\rangle = x_z/2$, and $\langle b\rangle = 0$. The case $q < 1/2$ is a trivial extension.

At this point, it is relevant to note that after local filtering, the state (4) violates the CHSH inequality $|S| \le 2$ [14], where $S = E_{1,1} + E_{1,2} + E_{2,1} - E_{2,2}$ and $E_{x,y} = \sum_{a,b=\pm 1}(ab)p(ab|xy)$. Specifically, applying filters of the form

$$F_A = \epsilon|0\rangle\langle 0| + |1\rangle\langle 1|, \qquad F_B = \delta|0\rangle\langle 0| + |1\rangle\langle 1|, \tag{6}$$

with $\delta = \epsilon/\sqrt{q}$ to state (4), we obtain the filtered state

$$\tilde{\rho} \simeq \sqrt{q}\Psi_- + (1-\sqrt{q})\frac{|0,1\rangle\langle 0,1| + |1,0\rangle\langle 1,0|}{2} + \mathcal{O}(\epsilon^2),$$

which violates CHSH up to $S = 2\sqrt{1+q}$ (for $\epsilon \to 0$) according the Horodecki criterion [17]. Note that filters (6) are optimal here [18]. Hence, the state (4) exhibits

hidden nonlocality for projective measurements. This shows that hidden nonlocality exists for two-qubit states—the previous example [9] considered Werner states of local dimension $d \geq 5$. However, at this point, we cannot ensure that the state (4) is local for all nonsequential measurements, since Bell violation could in principle be obtained using POVMs. Nevertheless, we will now build upon the above construction to present a state featuring genuine hidden nonlocality.

Our main tool is a protocol for constructing a state which admits a local model for POVMs. Specifically, starting from a state ρ_0 of local dimension d which is local for dichotomic projective measurements, we construct the state

$$\rho' = \frac{1}{d^2}[\rho_0 + (d-1)(\rho_A \otimes \sigma_B + \sigma_A \otimes \rho_B) + (d-1)^2 \sigma_A \otimes \sigma_B], \tag{7}$$

which is local for POVMs. Here, $\sigma_{A,B}$ are arbitrary d-dimensional states, and $\rho_{A,B} = \mathrm{Tr}_{B,A}(\rho_0)$.

Alice receives as input a POVM $\{M_a\}$ (from now on, we omit the subscript x). Without loss of generality, each POVM element M_a can be taken to be proportional to a rank-one projector P_a (see, e.g., Ref. [7]), i.e., $M_a = \alpha_a P_a$ with $\alpha_a \geq 0$ and $\sum_a \alpha_a = d$ by normalization of the POVM. Bob receives POVM $\{M_b\}$ (with $M_b = \beta_b P_b$). The protocol is explained below for Alice; Bob follows the same procedure.

Protocol 2.—(i) Alice chooses projector P_a with probability α_a/d (note that $\sum_a \alpha_a/d = 1$). (ii) She simulates the dichotomic projective measurement $\{P_a, \mathbb{1} - P_a\}$ on state ρ_0. (iii) If the output of the simulation corresponds to P_a, she outputs a. (iv) Otherwise, she outputs (any) a with probability $\mathrm{Tr}(M_a \sigma_A)$.

Let us now show that the protocol simulates ρ'. Note first that the probability that Alice outputs in step (iii) is given by $\sum_a \alpha_a/d \mathrm{Tr}(P_a \rho_A) = 1/d$. We will now evaluate the probability that the parties output given values a and b in the protocol. Four cases are possible: 1. Both Alice and Bob output in step (iii), which occurs with probability $(\alpha_a/d)(\beta_b/d)\mathrm{Tr}(P_a \otimes P_b \rho_0) = (1/d^2)\mathrm{Tr}(M_a \otimes M_b \rho_0)$. 2. Alice outputs in step (iii) and Bob in step (iv), which occurs with probability

$$\sum_k \frac{\alpha_a}{d}\frac{\beta_k}{d}\mathrm{Tr}[P_a(\mathbb{1} - P_k)\rho_0]\mathrm{Tr}(M_b \sigma_B) = \frac{d-1}{d^2}\mathrm{Tr}(M_a \rho_A)\mathrm{Tr}(M_b \sigma_B). \tag{8}$$

3. Alice outputs in step (iv), and Bob in step (iii) has probability $(d - 1/d^2)\mathrm{Tr}(M_a \sigma_A)\mathrm{Tr}(M_b \rho_B)$. 4. Both Alice and Bob output in step (iv), which occurs with probability $[(d-1)^2/d^2]\mathrm{Tr}(M_a \sigma_A)\mathrm{Tr}(M_b \sigma_B)$. Altogether, we have that $p(ab) = \mathrm{Tr}(M_a \otimes M_b \rho')$. Hence, the model reproduces the statistics of arbitrary POVMs on the state ρ'.

We are now ready to show our main result. We use protocol 2 with ρ_0 given by the state of Eq. (4), which is local for projective measurements for $q \leq 1/2$, and choosing $\sigma_{A,B} = |0\rangle\langle 0|$, we obtain a state of the form

$$\rho_G = \frac{1}{4}\Big[q\Psi_- + (2-q)|0\rangle\langle 0| \otimes \frac{\mathbb{1}}{2} + q\frac{\mathbb{1}}{2} \otimes |0\rangle\langle 0| + (2-q)|0,0\rangle\langle 0,0|\Big] \tag{9}$$

which is local for POVMs by construction for $q \leq 1/2$. Nevertheless, ρ_G is nonlocal for any $q > 0$ when an appropriate sequence of measurements is used. In particular, applying filters of the form (6) with $\delta = \epsilon/\sqrt{q}$ to state ρ_G, we obtain

$$\tilde{\rho}_G \simeq \frac{\sqrt{q}}{2}\Psi_- + \left(1 - \frac{\sqrt{q}}{2}\right)\frac{|0,1\rangle\langle 0,1| + |1,0\rangle\langle 1,0|}{2} + \mathcal{O}(\epsilon^2),$$

which violates CHSH up to $S = 2\sqrt{1 + q/4}$ (for $\epsilon \to 0$), according the criteria of Ref. [17]. Hence, sequential measurements are necessary to reveal the nonlocality of the state (9), which therefore exhibits genuine hidden nonlocality.

Finally, we present a stronger version of this phenomenon, showing that there exist quantum states with genuine and maximal hidden nonlocality. That is, although the state admits a local model for POVMs, it violates maximally the CHSH inequality when sequential measurements are used, as the state after filtering is a pure singlet state.

We start here by considering the qutrit-qubit state

$$\rho_E = q\Psi_- + (1-q)|2\rangle\langle 2| \otimes \frac{\mathbb{1}_2}{2}, \tag{10}$$

where $\mathbb{1}_2$ denotes the identity in the $|0\rangle, |1\rangle$ qubit subspace. This state is usually referred to as the "erasure state," as it can be obtained by sending half of a singlet state Ψ_- through an erasure channel; with probability q, the singlet state remains intact, and with probability $(1-q)$, Alice's qubit is lost and replaced by the state $|2\rangle\langle 2|$ (orthogonal to the qubit subspace).

The state (10) is local for dichotomic projective measurements when $q \leq 1/2$. Consider Alice receiving an observable with eigenvalues ± 1, which can always be written as an operator of the form $c_0 \vec{x} \cdot \vec{\sigma} + c_1 \mathbb{1}_2 + R$, where $c_0, c_1 \in [0,1]$, operators $\vec{x} \cdot \vec{\sigma}$ and $\mathbb{1}_2$ act on the $|0\rangle, |1\rangle$ qubit subspace, and operator R has no support in the qubit subspace. The protocol is similar to protocol 1. Alice and Bob share a vector $\vec{\lambda}$. Alice accepts $\vec{\lambda}$ with probability $|\vec{x} \cdot \vec{\lambda}|$, in which case she outputs $a = -\mathrm{sgn}(\vec{x} \cdot \vec{\lambda})$ with probability c_0, and a random bit otherwise. If she rejects $\vec{\lambda}$, she outputs ± 1 with probability $[1 \pm (c_1 + \mathrm{Tr}R)/2]/2$. Bob receives observable $\vec{y} \cdot \vec{\sigma}$ and outputs $b = \mathrm{sgn}(\vec{y} \cdot \vec{\lambda})$.

Noting that Alice accepts $\vec{\lambda}$ with probability $1/2$ on average, we obtain $\langle ab \rangle = -c_0(\vec{x} \cdot \vec{y})/2$, $\langle a \rangle = (c_1 + \mathrm{Tr}R)/2$, and $\langle b \rangle = 0$, which is the statistics of dichotomic projective measurements on state ρ_E for $q = 1/2$. Next, we apply protocol 2 to ρ_E, taking $\sigma_{A,B} = |2\rangle\langle 2|$. Hence, the state

$$\rho_{GM} = \frac{1}{9}[q\Psi_- + (3-q)|2\rangle\langle 2| \otimes \frac{\mathbb{1}_2}{2} + 2q\frac{\mathbb{1}_2}{2} \otimes |2\rangle\langle 2| + (6-2q)|2,2\rangle\langle 2,2|] \quad (11)$$

is local for POVMs for $q \leq 1/2$. To reveal the nonlocality of the above state, we apply filters of the form $F_A = F_B = |0\rangle\langle 0| + |1\rangle\langle 1|$. Hence, after successful filtering, we obtain a pure singlet state, i.e., $\tilde{\rho}_{GM} = \Psi_-$. By performing suitable measurements on $\tilde{\rho}_{GM}$, Alice and Bob can now get maximal violation of the CHSH inequality, i.e., $S = 2\sqrt{2}$ [19]. Therefore, the state (11) has genuine and maximal hidden nonlocality.

Note also that applying the above filters to the erasure state (10) gives a pure singlet state for any $q > 0$. Thus, the erasure state with $0 < q \leq 1/2$ has hidden nonlocality for dichotomic measurements. Moreover, for $q \leq 1/6$, the erasure state admits a local model for projective measurements, as can be shown by using protocol 2 [20]. Hence, such states feature hidden nonlocality for projective measurements.

To summarize, we have shown the existence of genuine hidden nonlocality. That is, there exist entangled quantum states the nonlocality of which can be revealed only via sequential measurements. In certain cases, this nonlocality can even be maximal.

In the present Letter, we have focused on Bell tests in which a single copy of an entangled state is measured in each run of the experiment. It is, however, also relevant to consider the case in which several copies of the state can be measured jointly in each run [21–24]. Notably, it has been shown recently that nonlocality can be superactivated in this scenario [25]. That is, by performing judicious joint measurements on sufficiently many copies of a state ρ, it becomes possible to violate a Bell inequality (with nonsequential measurements), although the state ρ admits a local model for POVMs. More generally, this phenomenon occurs for any entangled state ρ that is useful for teleportation [26]. It is thus interesting to ask whether the nonlocality of the states considered here could also be revealed by allowing for many copies to be measured jointly. However, the current results on superactivation of quantum nonlocality do not detect the states presented here [27], thus leaving the question open. Another point worth mentioning is activation of nonlocality in quantum networks. It would also be relevant to see whether the nonlocality of the states presented here can be activated by placing several copies of them in a quantum network [28]. Concerning the erasure state, Ref. [29] shows that it is a nonlocal resource when placed in a tripartite network; hence, the local model constructed here confirms that activation of nonlocality does indeed occur.

Finally, an interesting open question is whether there exist entangled states for which nonlocality cannot be observed, even considering sequential measurements on an arbitrary number of copies of the state.

We thank Y.-C. Liang for discussions. We also thank A. Acin, F. Brandão, D. Cavalcanti, N. Gisin, and T. Maciel. We acknowledge financial support from the Swiss National Science Foundation (Grant No. PP00P2_138917) and the EU DIQIP.

[1] J. S. Bell, Physics **1**, 195 (1964).
[2] H. Buhrman, R. Cleve, S. Massar, and R. de Wolf, Rev. Mod. Phys. **82**, 665 (2010).
[3] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, arXiv:1303.2849.
[4] N. Gisin, Phys. Lett. A **154**, 201 (1991).
[5] S. Popescu and D. Rohrlich,Phys. Lett. A **166**, 293 (1992).
[6] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).
[7] J. Barrett, Phys. Rev. A **65**, 042302 (2002).
[8] M. L. Almeida, S. Pironio, J. Barrett, G. Toth, and A. Acin, Phys. Rev. Lett. **99**, 040403 (2007).
[9] S. Popescu, Phys. Rev. Lett. **74**, 2619 (1995).
[10] N. Gisin, Phys. Lett. A **210**, 151 (1996).
[11] M. Zukowski, R. Horodecki, M. Horodecki, and P. Horodecki, Phys. Rev. A **58**, 1694 (1998).
[12] S. Teufel, K. Berndl, D. Dürr, S. Goldstein, and N. Zanghi, Phys. Rev. A **56**, 1217 (1997).
[13] T. Vértesi and E. Bene, Phys. Rev. A **82**, 062115 (2010).
[14] J. F. Clauser, M. A. Horne, A. Shimony, and R. Holt, Phys. Rev. Lett. **23**, 880 (1969).
[15] N. Gisin and B. Gisin, Phys. Lett. A **260**, 323 (1999).
[16] J. Degorre, S. Laplante, and J. Roland, Phys. Rev. A **72**, 062314 (2005).
[17] R. Horodecki, P. Horodecki, and M. Horodecki, Phys. Lett. A **200**, 340 (1995).
[18] F. Verstraete and M. M. Wolf, Phys. Rev. Lett. **89**, 170401 (2002).
[19] B. S. Tsirelson, Lett. Math. Phys. **4**, 93 (1980).
[20] Apply protocol 2 to ρ_E with $q = 1/2$ and take $\sigma_A = |2\rangle\langle 2|$. Note that the protocol must be applied solely on Alice's side. Since Bob holds a qubit, the simulation model for dichotomic measurements includes already all projective measurements.
[21] A. Peres, Phys. Rev. A **54**, 2685 (1996).
[22] Y.-C. Liang and A. C. Doherty, Phys. Rev. A **73**, 052116 (2006).
[23] L. Masanes, Y.-C. Liang, and A. C. Doherty, Phys. Rev. Lett. **100**, 090403 (2008).
[24] Y.-C. Liang, L. Masanes, and D. Rosset, Phys. Rev. A **86**, 052115 (2012).
[25] C. Palazuelos, Phys. Rev. Lett. **109**, 190401 (2012).
[26] D. Cavalcanti, A. Acin, N. Brunner, and T. Vértesi, Phys. Rev. A **87**, 042104 (2013).
[27] Currently, the best result shows that superactivation is possible when the state has a singlet fidelity greater than $1/d$ [26], which is not the case for states (9) and (11) (when $q \leq 1/2$) and state (4) (for $q \leq 1/3$).
[28] D. Cavalcanti, M. L. Almeida, V. Scarani, and A. Acin, Nat. Commun. **2**, 184 (2011).
[29] D. Cavalcanti, R. Rabelo, and V. Scarani, Phys. Rev. Lett. **108**, 040402 (2012).

Paper E

Genuinely Multipartite Entangled Quantum States with Fully Local Hidden Variable Models and Hidden Multipartite Nonlocality

Physical Review Letters **116**, 130401 (2016)

Joseph Bowles, Jérémie Francfort, Mathieu Fillettaz, Flavien Hirsch, and Nicolas Brunner

Genuinely Multipartite Entangled Quantum States with Fully Local Hidden Variable Models and Hidden Multipartite Nonlocality

Joseph Bowles, Jérémie Francfort, Mathieu Fillettaz, Flavien Hirsch, and Nicolas Brunner
Département de Physique Théorique, Université de Genève, 1211 Genève, Switzerland
(Received 2 December 2015; published 30 March 2016)

The relation between entanglement and nonlocality is discussed in the case of multipartite quantum systems. We show that, for any number of parties, there exist genuinely multipartite entangled states that admit a fully local hidden variable model, i.e., where all parties are separated. Hence, although these states exhibit the strongest form of multipartite entanglement, they cannot lead to Bell inequality violation considering general nonsequential local measurements. Then, we show that the nonlocality of these states can nevertheless be activated using sequences of local measurements, thus revealing genuine multipartite hidden nonlocality.

DOI: 10.1103/PhysRevLett.116.130401

The relation between quantum entanglement and nonlocality has been studied extensively in recent years; see, e.g., Refs. [1,2]. While both notions turn out to be equivalent for pure states [3,4], the case of a mixed state is still not understood. This is nevertheless desirable given the importance of entanglement and nonlocality from the point of view of the foundations of quantum theory and for quantum information processing [1].

This research was initiated by Werner [5], who presented a class of bipartite entangled states admitting a local hidden variable (LHV) model. This proved that the correlations obtained by performing arbitrary local projective measurements on such states can be perfectly simulated by a LHV model, hence using only classical resources. This was later extended to general nonsequential measurements, i.e., positive operator valued measures (POVMs) [6]. Since such states cannot lead to Bell inequality violation [7], they are referred to as "local" entangled states [8].

It turns out, however, that certain local entangled states can nevertheless lead to nonlocality when a sequence of local measurements is performed [9]. That is, the use of local filters can help to reveal (or activate) the nonlocality of the entangled state. This phenomenon, termed "hidden nonlocality," occurs even for entangled states admitting a LHV model for POVMs [10]. Other works showed that the nonlocality of local entangled states can be activated by performing joint measurements on several copies of the state [11–13], or by placing many copies of the state in a quantum network [14,15].

Whereas the above questions have been intensively discussed for bipartite states, the relation between entanglement and nonlocality for multipartite systems is almost unexplored thus far. Here, one should nevertheless expect interesting and novel phenomenona, due to the rich structure of multipartite entanglement. In particular, there is a hierarchy of different forms of entanglement in multipartite systems, the strongest of which is genuine multipartite entanglement (GME). Similarly, the notion of genuine multipartite nonlocality (GMNL) has been discussed [16–18], which represents the strongest form of nonlocality for multipartite systems. A first natural question is then whether there exist GME states, the correlations of which can be simulated by a LHV model. This was first discussed by Tóth and Acín [19], who presented a GME state of 3 qubits admitting a LHV model, but could not extend their construction to more parties. More recently, Augusiak *et al.* [20] showed the existence of GME states of any number of parties that cannot lead to GMNL. Specifically, the authors discussed a class of GME states of N parties, and constructed a LHV model in which the parties are separated into two groups. However, this model is essentially bipartite, as the N parties cannot be completely separated. Beyond these few exploratory works, nothing is known. to the best of our knowledge.

Here we report progress in understanding the relation between GME and nonlocality. First, we present a general technique for constructing multipartite entangled states admitting a fully LHV model, i.e., where all parties are separated. This allows us to show that there exist GME states of an arbitrary number of systems, which admit a fully LHV model for arbitrary POVM measurements. Moreover, we show that the nonlocality of these states can be activated using sequential measurements. Notably, the use of local filters allows us to obtain GMNL. To summarize, there exist multipartite states, entangled in the strongest possible sense, that do not exhibit even the weakest form of nonlocality when considering nonsequential measurements. However, when using sequences of measurements, the strongest form of multipartite nonlocality can be obtained. We conclude with a series of open questions.

Genuine multipartite entanglement.—Consider N parties sharing a multipartite quantum state ρ acting on $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_N$, where $\mathcal{H}_i$ is the local Hilbert space of party i. Denote by $(b, \bar{b}) \in \mathcal{B}$ a bipartition of the N parties.

If ρ can be decomposed as a mixture of states that are each separable on some bipartition of the Hilbert space, then we have

$$\rho = \sum_{(b,\bar{b})\in\mathcal{B}} p_b \left(\sum_j q_j^b |\Phi_j\rangle\langle\Phi_j|_b \otimes |\Phi_j\rangle\langle\Phi_j|_{\bar{b}} \right), \quad (1)$$

with $\sum_b p_b = \sum_j q_j^b = 1$, and $|\Phi_j\rangle\langle\Phi_j|_b$ acts on the Hilbert space specified by the partition b (and similarly for $|\Phi_j\rangle\langle\Phi_j|_{\bar{b}}$). If ρ does not admit such a decomposition, then it is GME. Such states can thus not be created via local operations and classical communication (LOCC) using only biseparable states.

Determining whether a given state is GME is challenging, as one must search over all possible decompositions [Eq. (1)]. However, there are sufficient conditions for an N-qubit state to be GME [21–23] (see also Ref. [24]). Write the state ρ in the canonical basis $|0,0,\ldots,0\rangle, |0,0,\ldots,1\rangle, \ldots, |1,1,\ldots,1\rangle$ as

$$\rho = \begin{pmatrix} c_1 & & & & & & z_1 \\ & c_2 & & & & z_2 & \\ & & \ddots & & \iddots & & \\ & & & c_n & z_n & & \\ & & & z_n^* & d_n & & \\ & & \iddots & & & \ddots & \\ & z_2^* & & & & d_2 & \\ z_1^* & & & & & & d_1 \end{pmatrix} \quad (2)$$

(we only write the elements of interest), where $n = 2^{N-1}$. Then ρ is GME if

$$C(\rho) = 2\max_i\{|z_i| - w_i\} > 0, \quad (3)$$

where $w_i = \sum_{j\neq i}^n \sqrt{c_j d_j}$. Below, we will use this condition to ensure that a state is GME. Note that the value of $C(\rho)$ can also be used to quantify GME [25], an aspect that, however, will not be discussed here.

Nonlocality.—Consider again the state ρ, where now each party can make measurements labeled x_i obtaining outcomes a_i, specified by the measurement operators $M_{a_i|x_i}$, with $M_{a_i|x_i} \geq 0$ and $\sum_{a_i} M_{a_i|x_i} = \mathbb{1}$. The probability to see the outputs $\mathbf{a} = (a_1, \ldots, a_N)$ given the inputs $\mathbf{x} = (x_1, \ldots, x_N)$ is given by

$$p(\mathbf{a}|\mathbf{x}) = \mathrm{Tr}[\rho(\otimes_{i=1}^N M_{a_i|x_i})]. \quad (4)$$

The state ρ is called (fully) local if, for all possible measurement operators $M_{a_i|x_i}$, the statistics $p(\mathbf{a}|\mathbf{x})$ can be reproduced by a LHV model:

$$p(\mathbf{a}|\mathbf{x}) = \int_\lambda q_\lambda p_\lambda(a_1|x_1) p_\lambda(a_2|x_2) \cdots p_\lambda(a_N|x_N) d\lambda, \quad (5)$$

where q_λ is a probability density over the shared variable λ and $p_\lambda(a_i|x_i)$ are probability distributions, called local response functions. Likewise, if Eq. (5) cannot be satisfied, then the state is said to be nonlocal, as witnessed by the violation of (some) Bell inequality.

One may also consider a weaker notion of locality, whereby the correlations are not demanded to be local with respect to all parties [as in Eq. (5)], but instead to be (mixtures of) correlations that are each local across some bipartition. Again denoting by $(b,\bar{b}) \in \mathcal{B}$ a bipartition of the parties, these correlations take the form

$$p(\mathbf{a}|\mathbf{x}) = \sum_{(b,\bar{b})\in\mathcal{B}} p_b \int_\lambda q_\lambda^b p_\lambda(\mathbf{a}_b|\mathbf{x}_b) p_\lambda(\mathbf{a}_{\bar{b}}|\mathbf{x}_{\bar{b}}) d\lambda, \quad (6)$$

where $\mathbf{a}_b$, $\mathbf{x}_b$ denote the inputs and outputs for the bipartition b. Note that Eq. (5) implies Eq. (6), but not necessarily the converse. Correlations that cannot be written in the above form are called genuinely multipartite nonlocal and represent the strongest form of multipartite nonlocality [16]. Here, for simplicity, we put no restrictions on the probability distributions $p_\lambda(\mathbf{a}_b|\mathbf{x}_b)$, $p_\lambda(\mathbf{a}_{\bar{b}}|\mathbf{x}_{\bar{b}})$ other than positivity and normalization (for example, they may be signaling); note that more sophisticated definitions of GMNL were proposed [17,18]. The N-party Greenberger–Horne–Zeilinger (GHZ) state, $|\mathrm{GHZ}\rangle = (|0\rangle^{\otimes N} + |1\rangle^{\otimes N})/\sqrt{2}$, is known to produce correlations that are GMNL, as proven by the violation of the Svetlichny inequalities [16,27,28].

GME and nonlocality.—The link between GME and nonlocality is almost unexplored thus far. For $N = 3$, Tóth and Acín constructed a genuine tripartite entangled state admitting a fully LHV model [i.e., of the form Eq. (5)] for arbitrary local projective measurements [19]. Recently, Augusiak *et al.* [20] presented GME states of N qubits which cannot lead to GMNL. More precisely, they constructed a LHV model for some bipartition of N qubits, i.e., of the form Eq. (6). However, it is still unknown if there exist GME states that admit LHV models that are fully local, i.e., that satisfy Eq. (5), for any possible measurements. This is what we show in the next section.

Method.—Our main tool is a simple method to construct entangled N-party states which admit a LHV model. Specifically, we start by considering a bipartite entangled state ρ which is "unsteerable," that is, which cannot be used to demonstrate steering. Formally, this means that ρ admits as so-called local hidden state (LHS) model [29]; hence, its correlations can be decomposed as

$$\begin{aligned} p(ab|xy) &= \mathrm{Tr}[\rho M_{a|x} \otimes M_{b|y}] \\ &= \int q_\lambda p_\lambda(a|x) \mathrm{Tr}[\sigma_\lambda M_{b|y}] d\lambda, \end{aligned} \quad (7)$$

where σ_λ is the local hidden state, distributed with density q_λ, and $B_{b|y}$ denotes Bob's measurement operator. Clearly, an unsteerable state is local (with $p(b|y,\lambda) = \mathrm{Tr}[\sigma_\lambda M_{b|y}]$), while the opposite may not hold in general.

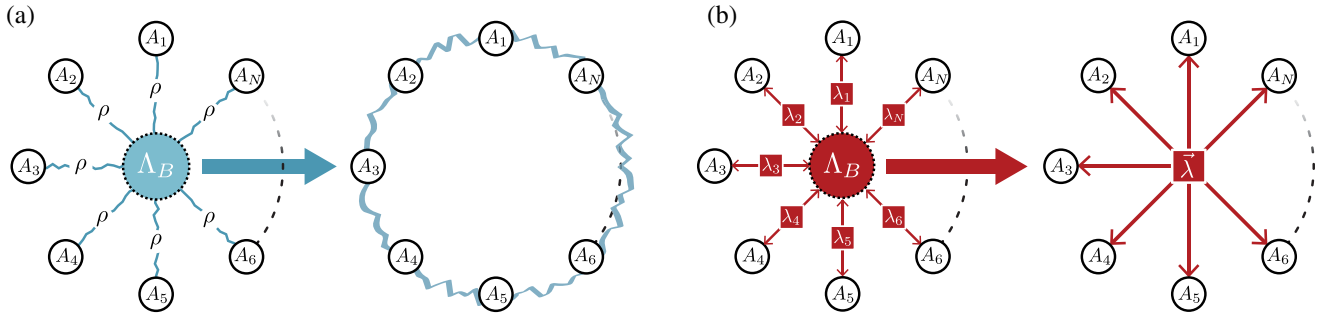

FIG. 1. Construction of multipartite states admitting a fully local model. (a) Construction of the state. First, place N copies of a bipartite state ρ in a star-shaped network. Then, apply a map Λ_B at the central node (i.e., on parties $B_1 \ldots B_N$), and trace out these parties. We thus obtain an N-partite state, $\rho_{A_1 \cdots A_N}$ (represented by the blue wiggly line), shared by parties $A_1 \ldots A_N$. (b) LHV model. If ρ admits a LHS model, one can simulate the correlations of the star-shaped network for $\rho^{\otimes N}$, whereby the central node receives the hidden states σ_{λ_i} independently from each source and the parties A_i receive hidden variables λ_i. One may now correlate the individual λ_i's by having the map Λ_B act on the hidden states; i.e., we can define a new distribution over $\vec{\lambda} = (\lambda_1, \ldots, \lambda_N)$ that depends on $\mathrm{Tr}[\Lambda_B(\otimes_i \sigma_{\lambda_i})]$. If each party A_i uses the same response function as in the LHS model for ρ, then the resulting statistics on parties $A_1 \ldots A_N$ simulate exactly the state $\rho_{A_1 \cdots A_N}$.

Next, we combine several copies of ρ in a star-shaped network (see Fig. 1). This allows one to construct a multipartite entangled state admitting a fully local model. Specifically, we have the following.

Lemma 1.—Let ρ be a quantum state acting on $\mathcal{H}_{A_1} \otimes \mathcal{H}_{B_1}$. The state $\rho^{\otimes N}$ therefore acts on $\mathcal{H}_{A_1} \otimes \cdots \otimes \mathcal{H}_{A_N} \otimes \mathcal{H}_{B_1} \otimes \cdots \otimes \mathcal{H}_{B_N} = \mathcal{H}_A \otimes \mathcal{H}_B$. Furthermore, let Λ_B be a completely positive linear map acting on $\mathcal{H}_B$. If ρ is unsteerable from A_1 to B_1, i.e., admits a decomposition [Eq. (7)], then the N-party state,

$$\rho_{A_1 \cdots A_N} = \frac{\mathrm{Tr}_B[\mathbb{1}_A \otimes \Lambda_B(\rho^{\otimes N})]}{\mathrm{Tr}[\mathbb{1}_A \otimes \Lambda_B(\rho^{\otimes N})]}, \tag{8}$$

admits a local hidden variable model, of the form Eq. (5), on the N-partition $A_1/A_2/\cdots/A_{N-1}/A_N$.

The intuition behind the above lemma is given in Fig. 1. A complete proof is given in Appendix A in Supplemental Material [30].

Note that we have not specified the class of local measurements for which the LHV model is valid in the above lemma. If ρ has a LHS model for projective measurements, then $\rho_{A_1 \cdots A_N}$ will have a LHV model for projective measurements, and similarly for POVMs. Note also that one can generalize slightly the result of Lemma 1 (see Appendix A in Supplemental Material [30]). Specifically, one can use different unsteerable states in each arm of the star-shaped network rather than the same state N times, and one can choose not to perform the trace over B and keep the center party.

GME states with fully local model.—We now use Lemma 1 to construct N-qubit states which admit a fully local model. We then prove these states to be GME for all N. Specifically, consider the class of two-qubit states,

$$\rho_{\alpha,\theta} = \alpha |\psi_\theta\rangle\langle\psi_\theta| + (1-\alpha)\rho_A^\theta \otimes \frac{\mathbb{1}}{2}, \tag{9}$$

where $0 \le \alpha \le 1$, $0 \le \theta \le \pi/4$, $|\psi_\theta\rangle = \cos\theta|00\rangle + \sin\theta|11\rangle$, and $\rho_A^\theta = \mathrm{Tr}_B|\psi_\theta\rangle\langle\psi_\theta|$. These states are entangled for all $\theta \in]0, \pi/4]$, if $\alpha > 1/3$. Furthermore, they are unsteerable from Alice to Bob for arbitrary projective measurements if the relation

$$\cos^2(2\theta) \ge \frac{2\alpha - 1}{(2-\alpha)\alpha^3} \tag{10}$$

holds [31]. Hence, for any $0 \le \alpha < 1$, one may find a corresponding $\theta > 0$ such that $\rho_{\alpha,\theta}$ is unsteerable. We now define the completely positive linear map,

$$\Lambda_B(\sigma) = F_B \sigma F_B^\dagger, \qquad F_B = |0\rangle[\langle 0,0,\ldots,0| + \langle 1,1,\ldots,1|],$$

which projects the systems of $B_1 \ldots B_N$ onto an N-qubit GHZ state. We may now define the N-party state $\rho_{A_1 \cdots A_N}$ by using $\rho_{\alpha,\theta}$ and Λ_B in Eq. (8). In Appendix B in Supplemental Material [30] we show that the concurrence of this state for a fixed N, α, θ is given by

$$C(\rho_{A_1 \cdots A_N}) = \frac{2\sin^N(2\theta)\left(\alpha^N + \left[\frac{1+\alpha}{2}\right]^N + \left[\frac{1-\alpha}{2}\right]^N - 1\right)}{[1+\alpha\cos 2\theta]^N + [1-\alpha\cos 2\theta]^N}. \tag{11}$$

It follows that for any N, one can find parameters α, θ such that (i) condition (10) is satisfied (ensuring that $\rho_{\alpha,\theta}$ has a LHS model) and (ii) $C(\rho_{A_1 \cdots A_N}) > 0$, proving that $\rho_{A_1 \cdots A_N}$ is GME. To give a specific example, take $\alpha = 1 - 1/N^2$ and $\theta > 0$, such that Eq. (10) is saturated. One sees that the denominator of Eq. (11) and $\sin^N 2\theta$ are both positive. We therefore need

$$\alpha^N + \left[\frac{1+\alpha}{2}\right]^N + \left[\frac{1-\alpha}{2}\right]^N > 1 \tag{12}$$

to be positive for all $N \geq 2$. For the case $N = 2$, one has $\alpha = 3/4$ and we find $43/32 > 1$. For $N > 2$, upon substituting $\alpha = 1 - 1/N^2$ the left-hand side becomes

$$\left[1 - \frac{1}{N^2}\right]^N + \left[1 - \frac{1}{2N^2}\right]^N + \left[\frac{1}{2N^2}\right]^N > 2\left[1 - \frac{1}{N^2}\right]^N > 2\left[1 - \frac{1}{N}\right] > 1, \tag{13}$$

where for the first inequality we use the fact that $[1 - 1/N^2]^N < [1 - 1/2N^2]^N$ and $[1/2N^2]^N > 0$, and the second inequality follows from Bernoulli's inequality.

Extension to general measurements.—A natural question is now to find a GME state with a fully local model, considering general POVMs. While the states $\rho_{\alpha,\theta}$ are not known to admit a LHS model for POVMs, we can nevertheless proceed differently. Starting from $\rho_{A_1\cdots A_N}$, we can in fact construct another state, ρ_{GME}, which is both GME and local for POVM measurements.

Specifically, define $\rho_{A_1\cdots A_k} = \mathrm{Tr}_{A_{k+1}\cdots A_N}[\rho_{A_1\cdots A_N}]$ and denote by $\circlearrowleft[\rho]$ the unnormalized and symmetrized version of ρ. Then the state

$$\rho_{\mathrm{GME}} = \frac{1}{2^N}\left[\rho_{A_1\cdots A_N} + \sum_{j=0}^{N-1} \circlearrowleft[\rho_{A_1\cdots A_j} \otimes |2\rangle\langle 2|^{\otimes N-j}]\right] \tag{14}$$

admits a fully local model, for arbitrary local POVMs. Note that $|2\rangle\langle 2|$ denotes the projector onto a subspace orthogonal to the qubit subpace. The above follows from a straightforward extension of Protocol 2 of Ref. [10] to the case of N parties.

To conclude, we have to show that the state is GME. Note that if each party makes a local projection on the qubit subspace $|0\rangle\langle 0| + |1\rangle\langle 1|$, then the resulting (renormalized) state is $\rho_{A_1\cdots A_N}$, which is GME. Since one cannot create GME using stochastic local operations, it follows that ρ_{GME} is GME.

Hidden genuine multipartite nonlocality.—We showed that GME states can admit a fully LHV model for arbitrary nonsequential measurements. A natural question now is whether these states have hidden nonlocality [9], that is, whether nonlocality could be revealed via sequences of measurements. A sufficient condition for the existence of hidden nonlocality is the possibility of transforming the initial state using local stochastic operations, i.e., local filters, to another state that violates some Bell inequality (see, e.g., Ref. [32]). Below, we will see that the states ρ_{GME} have genuine multipartite hidden nonlocality. Furthermore, the activation of nonlocality is maximal, in the sense that the filtered state exhibits GMNL, despite the initial state being fully local.

Consider N parties sharing ρ_{GME}. Let each party perform a local filtering operation given by

$$G_\epsilon = \epsilon|0\rangle\langle 0| + |1\rangle\langle 1|, \tag{15}$$

hence transforming ρ_{GME} to the state

$$\rho_\epsilon = \frac{G_\epsilon^{\otimes N}\rho_{\mathrm{GME}}G_\epsilon^{\otimes N}}{\mathrm{Tr}[G_\epsilon^{\otimes N}\rho_{\mathrm{GME}}G_\epsilon^{\otimes N}]}. \tag{16}$$

In Appendix C of Supplemental Material [30] we prove that for $\epsilon = \tan\theta$ [where θ is the parameter in Eq. (9)] the filtered state is essentially a pure N-party GHZ state $[|0\rangle^{\otimes N} + |1\rangle^{\otimes N}]/\sqrt{2}$. Specifically, the fidelity between the two states is given by

$$\mathcal{F}(\rho_\epsilon, |\mathrm{GHZ}\rangle\langle\mathrm{GHZ}|) = \langle\mathrm{GHZ}|\rho_\epsilon|\mathrm{GHZ}\rangle = \frac{1}{2}\left[\alpha^N + \left(\frac{1+\alpha}{2}\right)^N + \left(\frac{1-\alpha}{2}\right)^N\right], \tag{17}$$

which tends to 1 when α is sufficiently close to 1. Since the GHZ state is known to exhibit GMNL for any N, in particular, via violation of the Svetlichny inequalities [27,28] (which are robust to noise), it follows that ρ_ϵ can also be made GMNL.

Conclusion.—We showed that GME states can admit a fully LHV model, for any number of parties. Thus, while exhibiting the strongest form of multipartite entanglement (GME), these states can never lead to any Bell inequality violation, considering general nonsequential measurements. This can be viewed as a maximal inequivalence between multipartite entanglement and nonlocality. Interestingly, this gap can disappear when sequential measurements are considered, and the strongest form of nonlocality can be activated, thus highlighting the relevance of sequential measurements in multipartite nonlocality.

In the future, it would be interesting to investigate the above questions in quantitative terms. For instance, could one find examples of highly entangled GME states admitting a LHV model? In order to do so, one should choose a specific measure of GME [24] (as there exist no unique measure).

Also, the method we presented for constructing multipartite local entangled states could be further explored. Firstly, one could start from different bipartite unsteerable states; see, e.g., Refs. [33,34]. Secondly, by keeping the central node in the network, one can construct multipartite LHS models where one of the parties has a quantum response function, and hence may prove useful in the study of multipartite steering [35].

Finally, one could ask if there exist GME states admitting LHV models for sequential measurements, although this question is in fact still open even in the bipartite case.

We thank Marco Túlio Quintino for discussions. We acknowledge financial support from the Swiss National Science Foundation (Grant No. PP00P2_138917 and Starting Grant DIAQ) and EU SIQS.

[1] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Rev. Mod. Phys. **86**, 419 (2014).
[2] A. A. Methot and V. Scarani, Quantum Inf. Comput. **7**, 157 (2007).
[3] N. Gisin, Phys. Lett. A **154**, 201 (1991).
[4] S. Popescu and D. Rohrlich, Phys. Lett. A **166**, 293 (1992).
[5] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).
[6] J. Barrett, Phys. Rev. A **65**, 042302 (2002).
[7] J. S. Bell, Physics (Long Island City, N.Y.) **1**, 195 (1964).
[8] R. Augusiak, M. Demianowicz, and A. Acín, J. Phys. A **47**, 424002 (2014).
[9] S. Popescu, Phys. Rev. Lett. **74**, 2619 (1995).
[10] F. Hirsch, M. T. Quintino, J. Bowles, and N. Brunner, Phys. Rev. Lett. **111**, 160402 (2013).
[11] C. Palazuelos, Phys. Rev. Lett. **109**, 190401 (2012).
[12] D. Cavalcanti, A. Acín, N. Brunner, and T. Vértesi, Phys. Rev. A **87**, 042104 (2013).
[13] L. Masanes, Y.-C. Liang, and A. C. Doherty, Phys. Rev. Lett. **100**, 090403 (2008).
[14] D. Cavalcanti, M. L. Almeida, V. Scarani, and A. Acín, Nat. Commun. **2**, 184 (2011).
[15] A. Sen(De), U. Sen, C. Brukner, V. Bužek, and M. Żukowski, Phys. Rev. A **72**, 042310 (2005).
[16] G. Svetlichny, Phys. Rev. D **35**, 3066 (1987).
[17] R. Gallego, L. E. Würflinger, A. Acin, and M. Navascués, Phys. Rev. Lett. **109**, 070401 (2012).
[18] J.-D. Bancal, J. Barrett, N. Gisin, and S. Pironio, Phys. Rev. A **88**, 014102 (2013).
[19] G. Tóth and A. Acín, Phys. Rev. A **74**, 030306 (2006).
[20] R. Augusiak, M. Demianowicz, J. Tura, and A. Acín, Phys. Rev. Lett. **115**, 030404 (2015).
[21] O. Gühne and M. Seevinck, New J. Phys. **12**, 053002 (2010).
[22] M. Huber, F. Mintert, A. Gabriel, and B. C. Hiesmayr, Phys. Rev. Lett. **104**, 210501 (2010).
[23] Z.-H. Ma, Z.-H. Chen, J.-L. Chen, C. Spengler, A. Gabriel, and M. Huber, Phys. Rev. A **83**, 062325 (2011).
[24] C. Eltschka and J. Siewert, J. Phys. A **47**, 424005 (2014).
[25] The quantity $C(\rho)$ is a lower bound for one possible multipartite generalization of the concurrence [21,22], and becomes exact for the case of qubit X matrices [26].
[26] S. M. Hashemi Rafsanjani, M. Huber, C. J. Broadbent, and J. H. Eberly, Phys. Rev. A **86**, 062303 (2012).
[27] D. Collins, N. Gisin, S. Popescu, D. Roberts, and V. Scarani, Phys. Rev. Lett. **88**, 170405 (2002).
[28] M. Seevinck and G. Svetlichny, Phys. Rev. Lett. **89**, 060401 (2002).
[29] H. M. Wiseman, S. J. Jones, and A. C. Doherty, Phys. Rev. Lett. **98**, 140402 (2007).
[30] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.116.130401 for technical proofs of main results. (A) Proof of lemma 1. (B) Calculation of concurrence of GME state. (C) Proof of multipartite hidden nonlocality.
[31] J. Bowles, F. Hirsch, M. T. Quintino, and N. Brunner, Phys. Rev. A **93**, 022121 (2015).
[32] R. Gallego, L. E. Würflinger, R. Chaves, A. Acin, and M. Navascués, New J. Phys. **16**, 033037 (2014).
[33] J. Bowles, T. Vértesi, M. T. Quintino, and N. Brunner, Phys. Rev. Lett. **112**, 200402 (2014).
[34] S. Jevtic, M. J. W. Hall, M. R. Anderson, M. Zwierz, and H. M. Wiseman, J. Opt. Soc. Am. B **32**, A40 (2015).
[35] D. Cavalcanti, P. Skrzypczyk, G. H. Aguilar, R. V. Nery, P. H. Souto Ribeiro, and S. P. Walborn, Nat. Commun. **6**, 7941 (2015).

Genuinely multipartite entangled quantum states with fully local hidden variable models and hidden multipartite nonlocality: Supplementary material

Appendix A: General method proof

Here we give a complete proof of Lemma 1. We define the state ρ_Λ as

$$\rho_\Lambda = \frac{\mathbb{1}_A \otimes \Lambda_B(\rho^{\otimes N})}{\mathcal{N}}, \tag{A1}$$

where

$$\mathcal{N} = \mathrm{Tr}[\mathbb{1}_A \otimes \Lambda_B(\rho^{\otimes N})] = \mathrm{Tr}[\Lambda_B(\rho_B^{\otimes N})] \tag{A2}$$

with $\rho_B = \mathrm{Tr}_A[\rho]$. Note that $\rho_{A_1\cdots A_N} = \mathrm{Tr}_B[\rho_\Lambda]$.

We first show that the $N+1$ party distribution

$$p(\mathbf{a}b|\mathbf{x}y) = \mathrm{Tr}\left[(\otimes_{i=1}^N M_{a_i|x_i}) \otimes M_{b|y}\ \rho_\Lambda\right] \tag{A3}$$

admits a LHV model on the $N+1$ partition $A_1/A_2/\cdots/A_N/B$. Note that since the parties $B_1\cdots B_N = B$ now form a single party, the operator $M_{b|y}$ acts on the Hilbert space $\mathcal{H}_{B_1}\otimes\cdots\otimes\mathcal{H}_{B_N}$ and may be entangled across this space. Since a LHV model for a state clearly implies a LHV model for any subsystems of that state, proving a LHV model for ρ_Λ then implies a LHV model for $\mathrm{Tr}_B[\rho_\Lambda] = \rho_{A_1\cdots A_N}$, proving Lemma 1. To this end, we show the existence of a shared variable $\vec{\lambda}$ with corresponding normalised probability density $Q(\vec{\lambda})$ and response functions for the $N+1$ parties such that the corresponding LHV model reproduces the statistics (A3).

Replacing ρ_Λ by (A1) and denoting the dual map of Λ_B by Λ_B^* we have

$$\begin{aligned} p(\mathbf{a}b|\mathbf{x}y) &= \frac{1}{\mathcal{N}}\mathrm{Tr}\left[(\otimes_i M_{a_i|x_i})\otimes M_{b|y}\ \mathbb{1}_A\otimes\Lambda_B(\rho^{\otimes N})\right] = \frac{1}{\mathcal{N}}\mathrm{Tr}\left[(\otimes_i M_{a_i|x_i})\otimes\Lambda_B^*(M_{b|y})\rho^{\otimes N}\right] \\ &= \frac{1}{\mathcal{N}}\mathrm{Tr}\left[\mathrm{Tr}_A\left[(\otimes_i M_{a_i|x_i})\rho^{\otimes N}\right]\Lambda_B^*(M_{b|y})\right] = \frac{1}{\mathcal{N}}\mathrm{Tr}\left[\left(\otimes_i \mathrm{Tr}_{A_i}\left[M_{a_i|x_i}\otimes\mathbb{1}\,\rho\right]\right)\Lambda_B^*(M_{b|y})\right]. \end{aligned} \tag{A4}$$

Since we assume the state ρ to be unsteerable, it follows that (for examples see [1])

$$\mathrm{Tr}_{A_i}\left[M_{a_i|x_i}\otimes\mathbb{1}\,\rho\right] = \int q_{\lambda_i}p_{\lambda_i}(a_i|x_i)\sigma_{\lambda_i}\mathrm{d}\lambda_i. \tag{A5}$$

Combining this with the above we have

$$\begin{aligned} p(\mathbf{a}b|\mathbf{x}y) &= \frac{1}{\mathcal{N}}\mathrm{Tr}\left[\left(\otimes_i\int_{\lambda_i} q_{\lambda_i}p_{\lambda_i}(a_i|x_i)\sigma_{\lambda_i}\mathrm{d}\lambda_i\right)\Lambda_B^*(M_{b|y})\right] \\ &= \int_{\lambda_1}\cdots\int_{\lambda_N}\frac{q_{\lambda_1}\cdots q_{\lambda_N}}{\mathcal{N}}p_{\lambda_1}(a_1|x_1)\cdots p_{\lambda_N}(a_N|x_n)\,\mathrm{Tr}\left[\left(\otimes_i\sigma_{\lambda_i}\right)\Lambda_B^*(M_{b|y})\right]\mathrm{d}\lambda_1\cdots\mathrm{d}\lambda_N \\ &= \int_{\vec{\lambda}} Q(\vec{\lambda})\,p_{\lambda_1}(a_1|x_1)\cdots p_{\lambda_N}(a_N|x_n)\,\mathrm{Tr}\left[\sigma_{\vec{\lambda}}M_{b|y}\right]\mathrm{d}\vec{\lambda}, \end{aligned} \tag{A6}$$

where $\vec{\lambda} = (\lambda_1,\cdots,\lambda_N)$ and we have

$$\sigma_{\vec{\lambda}} = \frac{\Lambda_B(\otimes_i\sigma_{\lambda_i})}{\mathrm{Tr}[\Lambda_B(\otimes_i\sigma_{\lambda_i})]} \qquad Q(\vec{\lambda}) = \frac{\prod_i q_i}{\mathcal{N}}\mathrm{Tr}[\Lambda_B(\otimes_i\sigma_{\lambda_i})]. \tag{A7}$$

Equation (A6) is now in the precise form of a LHV model. The shared variable consists of the vector $\vec{\lambda} = (\lambda_1,\cdots,\lambda_N)$ which is distributed to the N parties with probability density $Q(\vec{\lambda})$. Conditioned on $\vec{\lambda}$, the response functions for parties $A_1\cdots A_N$ remain unchanged whereas party B outputs according to $p(b|y,\vec{\lambda}) = \mathrm{Tr}[\sigma_{\vec{\lambda}}M_{b|y}]$, which is a valid probability distribution since $\sigma_{\vec{\lambda}}$ is a normalised quantum state. Furthermore since Λ_B is positive we have $Q(\vec{\lambda}) > 0$ and

$$\int_{\vec{\lambda}} Q(\vec{\lambda})\mathrm{d}\vec{\lambda} = \int\frac{\prod_i q_i}{\mathcal{N}}\mathrm{Tr}\left[\Lambda_B(\otimes_i\sigma_{\lambda_i})\right]\mathrm{d}\vec{\lambda} = \frac{1}{\mathcal{N}}\mathrm{Tr}\left[\Lambda_B\left(\otimes_i\int_{\lambda_i} q_{\lambda_i}\sigma_{\lambda_i}\mathrm{d}\lambda_i\right)\right] = \frac{1}{\mathcal{N}}\mathrm{Tr}\left[\Lambda_B\left(\rho_B^{\otimes N}\right)\right] = 1, \tag{A8}$$

where the third line follows from (A5) by setting say $A_{1|x_i} = \mathbb{1}$ and consequently $p(1|x_1,\lambda) = 1$. Hence $Q(\vec{\lambda})$ is indeed a probability density.

Appendix B: Calculation of $\mathbf{C}(\rho_{\mathbf{A_1}\cdots\mathbf{A_N}})$

Here we give a detailed derivation of (11). We first write the state (9) as

$$\rho_{\alpha,\theta} = \left[\frac{1+\alpha}{2}\right]\left(c^2|00\rangle\langle 00| + s^2|11\rangle\langle 11|\right) + \left[\frac{1-\alpha}{2}\right]\left(c^2|01\rangle\langle 01| + s^2|10\rangle\langle 10|\right) + \alpha cs\left(|00\rangle\langle 11| + |11\rangle\langle 00|\right), \tag{B1}$$

where c, s denote $\cos\theta$ and $\sin\theta$ respectively. To begin, we consider the unormalised state

$$\rho_F = \mathrm{Tr}_B\left[\left[\mathbb{1}_A \otimes F_B\right]\rho_{\alpha,\theta}^{\otimes N}\left[\mathbb{1}_A \otimes F_B\right]\right], \tag{B2}$$

where

$$F_B = |0\rangle\langle 00\cdots 0| + |0\rangle\langle 11\cdots 1| \tag{B3}$$

acts on $\mathcal{H}_B$. Notice that $\rho_{A_1\cdots A_N} = \rho_F/\,\mathrm{Tr}[\rho_F]$ and so $C(\rho_{A_1\cdots A_N}) = C(\rho_F)/\,\mathrm{Tr}[\rho_F]$. After performing the partial trace of (B2) we obtain

$$\begin{aligned}\rho_F = \quad & \mathbb{1}_A \otimes \langle 00\cdots 0|_B\,\rho_{\alpha,\theta}^{\otimes N}\,\mathbb{1}_A \otimes |00\cdots 0\rangle_B + \mathbb{1}_A \otimes \langle 11\cdots 1|_B\,\rho_{\alpha,\theta}^{\otimes N}\,\mathbb{1}_A \otimes |11\cdots 1\rangle_B \\ & + \mathbb{1}_A \otimes \langle 00\cdots 0|_B\,\rho_{\alpha,\theta}^{\otimes N}\,\mathbb{1}_A \otimes |11\cdots 1\rangle_B + \mathbb{1}_A \otimes \langle 11\cdots 1|_B\,\rho_{\alpha,\theta}^{\otimes N}\,\mathbb{1}_A \otimes |00\cdots 0\rangle_A.\end{aligned} \tag{B4}$$

We consider each of these four terms separately. For the first term, the only non-zero contributions coming from $\rho_{\alpha,\theta}^{\otimes N}$ will correspond the N-fold tensor product of combinations of the projectors $|00\rangle\langle 00|$ and $|10\rangle\langle 10|$ with their corresponding weights. Hence, this will contribute diagonal terms to ρ_F. For example, the diagonal term corresponding to

$$|01\cdots 0\rangle\langle 01\cdots 0| \tag{B5}$$

where the projector contains m 1's and $N-m$ 0's, will have a corresponding weight

$$c^{2(N-m)}s^{2m}\left[\frac{1+\alpha}{2}\right]^{N-m}\left[\frac{1-\alpha}{2}\right]^{m}. \tag{B6}$$

For the second term of (B4) we will have a similar situation, this time contributing

$$c^{2(N-m)}s^{2m}\left[\frac{1+\alpha}{2}\right]^{m}\left[\frac{1-\alpha}{2}\right]^{N-m} \tag{B7}$$

to the same diagonal element. Adding these two contributions, each diagonal entry of ρ_F containing m 1's and $N-m$ 0's will have weight

$$\gamma(m) = c^{2(N-m)}s^{2m}\left(\left[\frac{1+\alpha}{2}\right]^{N-m}\left[\frac{1-\alpha}{2}\right]^{m} + \left[\frac{1+\alpha}{2}\right]^{m}\left[\frac{1-\alpha}{2}\right]^{N-m}\right). \tag{B8}$$

Turning to the third and fourth terms of (B4) we see that the only nonzero contributions from ρ_0 correspond to the N-fold tensor products of $|00\rangle\langle 11|$ and $|11\rangle\langle 00|$ respectively. These contribute to ρ_F the two off-diagonal terms

$$(\alpha cs)^N|00\cdots 0\rangle\langle 11\cdots 1|, \qquad (\alpha cs)^N|11\cdots 1\rangle\langle 00\cdots 0|. \tag{B9}$$

Hence, we have a ρ_F of the form

$$\rho_F = \begin{pmatrix} \gamma(0) & & & & (\alpha cs)^N \\ & \gamma(1) & & & \\ & & \ddots & & \\ & & & \gamma(N-1) & \\ (\alpha cs)^N & & & & \gamma(N) \end{pmatrix}.$$

We now define the quantity

$$w_0 = \sum_{j=1}^{n}\sqrt{c_j d_j}, \tag{B10}$$

where c_j d_j correspond to entries in (3). Calculating this for ρ_F we obtain

$$\begin{aligned} w_0 &= \sum_{j=1}^{n} \sqrt{c_j d_j} = \frac{1}{2} \sum_{m=0}^{N} \binom{N}{m} \sqrt{\gamma(m)\gamma(N-m)} \\ &= \frac{1}{2} c^N s^N \left(\sum_{m=0}^{N} \binom{N}{m} \left[\frac{1+\alpha}{2}\right]^{N-m} \left[\frac{1-\alpha}{2}\right]^{m} + \sum_{m=0}^{N} \binom{N}{m} \left[\frac{1+\alpha}{2}\right]^{m} \left[\frac{1-\alpha}{2}\right]^{N-m} \right) \\ &= c^N s^N \left(\frac{1+\alpha}{2} + \frac{1-\alpha}{2} \right)^N = c^N s^N. \end{aligned} \tag{B11}$$

Due to the form of ρ_F, we see that $|z_i| - w_i$ can only be positive for $i = 1$. We have

$$\begin{aligned} w_1 &= w_0 - c_1 d_1 = c^N s^N - \sqrt{\gamma(0)\gamma(N)} \\ &= c^N s^N \left(1 - \left[\frac{1+\alpha}{2}\right]^N + \left[\frac{1-\alpha}{2}\right]^N \right). \end{aligned} \tag{B12}$$

We may now calculate

$$|z_1| - w_1 = c^N s^N \left[\alpha^N + \left[\frac{1+\alpha}{2}\right]^N + \left[\frac{1-\alpha}{2}\right]^N - 1 \right]. \tag{B13}$$

Finally, to calculate $C(\rho_{A_1 \cdots A_N}) = C(\rho_F)/\operatorname{Tr}[\rho_F]$ we need to calculate the normalisation $\operatorname{Tr}[\rho_F]$. This is given by

$$\begin{aligned} \operatorname{Tr}[\rho_F] &= \sum_{m=0}^{N} \binom{N}{m} \gamma(m) = \left[c^2 \frac{1+\alpha}{2} + s^2 \frac{1-\alpha}{2} \right]^N + \left[s^2 \frac{1+\alpha}{2} + c^2 \frac{1-\alpha}{2} \right]^N \\ &= \left[\frac{1}{2}(1 + \alpha \cos 2\theta) \right]^N + \left[\frac{1}{2}(1 - \alpha \cos 2\theta) \right]^N. \end{aligned} \tag{B14}$$

Combining this with (B13) and using $\cos\theta \sin\theta = \frac{1}{2} \sin 2\theta$, we arrive at (11). For $\rho_{A_1 \cdots A_N}$ we thus have

$$C(\rho_{A_1 \cdots A_N}) = \frac{2 \sin^N 2\theta \left(\alpha^N + \left[\frac{1+\alpha}{2}\right]^N + \left[\frac{1-\alpha}{2}\right]^N - 1 \right)}{[1 + \alpha \cos 2\theta]^N + [1 - \alpha \cos 2\theta]^N}.$$

Appendix C: Genuine multipartite hidden nonlocality

Here we calculate the fidelity between $\rho_{\tan\theta}$ and the N-party GHZ state. The state ρ_ϵ is given by

$$\rho_\epsilon = \frac{G_\epsilon^{\otimes N} \rho_{\mathrm{GME}} \, G_\epsilon^{\otimes N}}{\operatorname{Tr}[G_\epsilon^{\otimes N} \rho_{\mathrm{GME}} \, G_\epsilon^{\otimes N}]}. \tag{C1}$$

with

$$G_\epsilon = \epsilon |0\rangle\langle 0| + |1\rangle\langle 1|. \tag{C2}$$

Note that since G_ϵ has no support on the $|2\rangle\langle 2|$ subspace, only the first term of (14) will survive the filter. We may therefore replace ρ_{GME} in (C1) by $\rho_{A_1 \cdots A_N}$. To make calculations easier, we begin by working with the unormalised state

$$\tilde{\rho}_\epsilon = G_\epsilon^{\otimes N} \rho_{A_1 \cdots A_N} \, G_\epsilon^{\otimes N}. \tag{C3}$$

Since the filter (C2) is diagonal, $\tilde{\rho}_\epsilon$ will have the same structure as $\rho_{A_1 \cdots A_N}$. It is easy to see that after the filter, a diagonal element which contains m 1's and $N - m$ 0's picks up a factor of $\epsilon^{2(N-m)}$ whereas the two off-diagonal elements pick up each a factor of ϵ^N. We now use the ansatz $\epsilon = \tan\theta$. With this we have

$$\tilde{\rho}_{\tan\theta} = s^{2N} \begin{pmatrix} \gamma'(0) & & & & & \alpha^N \\ & \gamma'(1) & & & & \\ & & \ddots & & & \\ & & & \gamma'(N-1) & & \\ \alpha^N & & & & & \gamma'(N) \end{pmatrix},$$

where

$$\gamma'(m) = \left[\frac{1+\alpha}{2}\right]^{N-m}\left[\frac{1-\alpha}{2}\right]^{m} + \left[\frac{1+\alpha}{2}\right]^{m}\left[\frac{1-\alpha}{2}\right]^{N-m}. \tag{C4}$$

For this state we have $\mathrm{Tr}[\tilde{\rho}_{\tan\theta}] = 2s^{2N}$ and so after renormalising we obtain

$$\rho_{\tan\theta} = \frac{1}{2}\begin{pmatrix} \left[\frac{1+\alpha}{2}\right]^N + \left[\frac{1-\alpha}{2}\right]^N & & & & \alpha^N \\ & \ddots & & & \\ & & \gamma'(m) & & \\ & & & \ddots & \\ \alpha^N & & & & \left[\frac{1+\alpha}{2}\right]^N + \left[\frac{1-\alpha}{2}\right]^N \end{pmatrix}.$$

One can now easily see how this state can be made arbitrarily close to the GHZ state. Taking α close to zero forces the extreme diagonal and off diagonal elements to $\frac{1}{2}$ while forcing all others to zero. Making this quantitative, we compute the fidelity between $\rho_{\tan\theta}$ and the pure GHZ state $|\mathrm{GHZ}\rangle = (|0\rangle^{\otimes N} + |1\rangle^{\otimes N})/\sqrt{2}$:

$$\mathcal{F}(\rho_{\tan\theta}, |\mathrm{GHZ}\rangle\langle\mathrm{GHZ}|) = \langle\mathrm{GHZ}|\rho_{\tan\theta}|\mathrm{GHZ}\rangle = \frac{1}{2}\left[\alpha^N + \left(\frac{1+\alpha}{2}\right)^N + \left(\frac{1-\alpha}{2}\right)^N\right] \tag{C5}$$

which tends to 1 when α tends to 1.

---

[1] H. M. Wiseman, S. J. Jones, and A. C. Doherty, Phys. Rev. Lett. **98**, 140402 (2007).

Paper F

Local Hidden Variable Models for Entangled Quantum States Using Finite Shared Randomness

Physical Review Letters **112**, 200402 (2013)

Joseph Bowles, Támas Vértesi, Marco Túlio Quintino, and Nicolas Brunner

Local Hidden Variable Models for Entangled Quantum States Using Finite Shared Randomness

Joseph Bowles, Flavien Hirsch, Marco Túlio Quintino, and Nicolas Brunner
Département de Physique Théorique, Université de Genève, 1211 Genève, Switzerland
(Received 11 December 2014; published 24 March 2015)

The statistics of local measurements performed on certain entangled states can be reproduced using a local hidden variable (LHV) model. While all known models make use of an infinite amount of shared randomness, we show that essentially all entangled states admitting a LHV model can be simulated with finite shared randomness. Our most economical model simulates noisy two-qubit Werner states using only $\log_2(12) \simeq 3.58$ bits of shared randomness. We also discuss the case of positive operator valued measures, and the simulation of nonlocal states with finite shared randomness and finite communication. Our work represents a first step towards quantifying the cost of LHV models for entangled quantum states.

DOI: 10.1103/PhysRevLett.114.120401 PACS numbers: 03.65.Ud

Introduction.—Quantum systems exhibit a wide range of nonclassical and counterintuitive phenomena, such as quantum entanglement [1] and Bell nonlocality [2,3]. In recent years, considerable effort has been devoted to understanding the relation between entanglement and nonlocality; see [3]. While entanglement is necessary to demonstrate nonlocality (i.e., violation of a Bell inequality), it is not yet clear whether all entangled states can lead to nonlocality when considering the most general scenario [4,5]. Nevertheless, entanglement and nonlocality are proven to be different in the simplest scenario in which local (nonsequential) measurements are performed on a single copy of an entangled state. As discovered by Werner [6], there exist entangled states that can provably not violate any Bell inequality, since the state admits a local hidden variable (LHV) model. While Werner focused on projective measurements, Barrett [7] showed that the result holds for the most general nonsequential measurements, so-called positive operator valued measures (POVMs).

Following these early results, plenty of works have investigated these ideas; see [8] for a recent review. LHV models were reported for entangled states with less symmetry than Werner states [9–13]. Multipartite states were discussed as well [14,15]. Interestingly, it was shown that in certain cases, the nonlocality of local entangled states can be activated, e.g., by considering sequential measurements [11,16]. More recently, interest was devoted to a special class of LHV models, referred to as local hidden state (LHS) models, which naturally arise in the context of Einstein-Podolsky-Rosen (EPR) steering [17,18], and essentially require that the local variable represents a quantum state; see [17] for details, and [6,7,9,12,19] for examples of LHS models.

Here we discuss novel types of questions in this context, namely that of quantifying LHV models. Specifically, given a local entangled state, we ask what resources are required to construct a LHV model; i.e., what is the cost of classically simulating the correlations of the state? As a figure of merit, we consider the minimal dimension of the shared local (hidden) variable that is needed; that is, how much classical information (how many bits) is necessary to encode the local variable? Note that all LHV models constructed so far are maximally costly according to our measure as they make use of shared variables which are continuous. Hence, such models would require a communication channel of infinite capacity, the physical relevance of which is questionable. For instance, in Werner's model, the local variables are unit vectors $\vec{\lambda}$ (e.g., vectors on the Bloch sphere). Importantly, although these vectors are of a given dimension, the model requires an infinite number of them, as vectors $\vec{\lambda}$ are taken from the uniform distribution over the sphere.

Hence, a natural question is whether it would be in fact possible to simulate the correlations of an entangled state using shared variables of finite dimension (i.e., a finite number of shared random bits). Here we show that essentially any entangled state admitting a LHV model can be simulated with finite shared randomness, considering arbitrary local projective measurements. We discuss in detail the case of Werner states of two qubits. We also show that the simulation of arbitrary POVMs on certain entangled states is possible using finite shared randomness. Finally, we consider the simulation of nonlocal entangled states (i.e., which can violate a Bell inequality), in which case communication between the parties is necessary. In particular, we show that the simulation of any full rank entangled state can be achieved using only finite communication.

Our work provides a perspective on understanding how the correlations of local entangled states differ from those of fully separable states. On the one hand, it shows that there is no fundamental difference between the two cases, in the sense that finite shared randomness is enough for both (at least for certain entangled states). Recall that the correlations of separable states can always be simulated

using $4\log_2(d)$ bits [20], where d denotes the local Hilbert space dimension of the state. On the other hand, our results suggest that the simulation of entangled states is in general more costly compared to that of separable states—despite the fact that both classes of states can never lead to Bell inequality violation.

Preliminaries.—We consider a bipartite Bell scenario. Two distant observers, Alice and Bob, share a quantum state ρ (of Hilbert space dimension $d \times d$) and perform local measurements $A = \{A_a\}$ and $B = \{B_b\}$, respectively. The observed statistics are local (in the sense of Bell), if they can be decomposed as follows [2,3]:

$$\mathrm{Tr}(A_a \otimes B_b \rho) = \int \pi(\lambda) p_A(a|A,\lambda) p_B(b|B,\lambda) d\lambda, \quad (1)$$

where λ represents a shared (hidden) variable, distributed according to density $\pi(\lambda)$. If a decomposition of the form (1) exists for all possible local measurements, we say that the state ρ is local as it will never violate any Bell inequality. The LHV model is then characterized by the distributions $\pi(\lambda)$, and $p_A(a|A,\lambda)$, $p_B(b|B,\lambda)$ which are Alice's and Bob's local response functions.

Trivially, any state ρ that is separable is local. Indeed, one can write $\rho = \sum_{\lambda=1}^{d^4} p_\lambda \rho_A^\lambda \otimes \rho_B^\lambda$ [20,21] with $\sum_\lambda p_\lambda = 1$ and $p_\lambda \geq 0$ (note that for two-qubit states, a more economical decomposition exists, involving only four product states [22]). Here the local variable λ is distributed according to p_λ, and the local response functions are simply $p_A(a|A,\lambda) = \mathrm{Tr}(A_a \rho_A^\lambda)$ for Alice and similarly for Bob. Note that the shared variable takes only d^4 different values here, and can thus be encoded in $4\log_2(d)$ bits (for two-qubit states 2 bits are enough). More interestingly, there exist entangled states ρ which are local. The most famous example is the Werner state, which for the case $d = 2$ takes the form

$$\rho_W(\alpha) = \alpha |\psi^-\rangle\langle\psi^-| + (1-\alpha)\mathbb{I}/4, \quad (2)$$

where $|\psi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$ is the singlet state and $\mathbb{I}/4$ is the maximally mixed two-qubit state. After showing that the state $\rho_W(\alpha)$ is entangled for $\alpha > 1/3$, Werner [6] constructed a local model for arbitrary projective measurements for $\alpha \leq 1/2$; later another local model was constructed for $\alpha \lesssim 0.66$ [10]. Considering the most general nonsequential measurements, i.e., POVMs, a local model was presented for $\alpha \leq 5/12$ [7].

A common feature of these local models (and to the best of our knowledge, of all known LHV models) is the fact that the shared variable λ takes an infinite number of different values; typically, λ denotes a (unit) vector, which is taken randomly from a uniform distribution over the sphere. Hence λ requires an infinite number of bits to be encoded, in stark contrast with the case of separable states, where $4\log_2(d)$ bits are enough. Therefore, it is rather natural to ask if this represents a fundamental difference between local entangled states and separable ones. Below we will show that this is not the case, by exhibiting LHV models for entangled states requiring only finite resources, i.e., where λ can be encoded with a finite number of bits.

Simulating Werner states with finite shared randomness.—We present local models using a finite amount of shared randomness, simulating the correlations of Werner states $\rho_W(\alpha)$ for $\alpha < 0.5$ for all projective measurements; extensions to $\alpha \lesssim 0.66$ are given in the next section. Alice and Bob receive here Bloch vectors $\vec{a}$ and $\vec{b}$ (representing observables $A = \vec{a}\cdot\vec{\sigma}$ and similarly for Bob) and should provide outcomes $a, b = \pm 1$ such that

$$\langle a \rangle = \langle b \rangle = 0, \qquad \langle ab \rangle = -\alpha \vec{a}\cdot\vec{b}. \quad (3)$$

For clarity, we start by presenting a simple model using only $\log_2(12)$ bits of shared randomness, which works for $\alpha \lesssim 0.43$. Our model uses the icosahedron, one of the 5 platonic solids in dimension 3. The icosahedron has 12 vertices represented by the normalized vectors $\vec{v}_\lambda \in V$, which satisfy the following properties:

$$\forall \vec{v}_\lambda \, \exists \, \vec{v}_j \text{ such that (s.t.) } \vec{v}_\lambda = -\vec{v}_j \quad (4)$$

$$\sum_{j \text{ s.t. } \vec{v}_j\cdot\vec{v}_\lambda \geq 0} \vec{v}_j = \gamma \vec{v}_\lambda \qquad \forall \lambda, \quad (5)$$

with $\gamma = 1 + \sqrt{5}$. Note that the radius of a sphere inscribed inside the icosahedron is given by $\ell = \sqrt{(5+2\sqrt{5})/15}$. In our model the shared variable $\lambda \in \{1, \ldots, 12\}$ is distributed uniformly and represents one of the 12 vertices of the icosahedron. That is, when Alice and Bob receive λ, they will use vector $\vec{v}_\lambda$.

Protocol 1.—Alice and Bob share $\lambda \in \{1, \ldots, 12\}$, uniformly distributed. Upon receiving setting $\vec{a}$, Alice calculates the subnormalized vector $\vec{a}' = \ell\vec{a}$. This ensures that $\vec{a}'$ lies inside the convex hull of V; hence, Alice can find a convex decomposition $\vec{a}' = \sum_i \omega_i \vec{v}_i$ with $\sum_i \omega_i = 1$ and $\omega_i \geq 0$ (note that any convex decomposition can be chosen). Then, with probability ω_i, she outputs $a = \pm 1$ with probability $(1 \pm \mathrm{sgn}[\vec{v}_\lambda \cdot \vec{v}_i])/2$. Bob, upon receiving $\vec{b}$, outputs $b = \pm 1$ with probability $(1 \mp \vec{b}\cdot\vec{v}_\lambda)/2$.

We now show that the protocol reproduces the desired statistics. We start with the correlator:

$$\langle ab \rangle = -\frac{1}{12} \sum_\lambda \sum_i \omega_i \mathrm{sgn}(\vec{v}_i \cdot \vec{v}_\lambda) \vec{v}_\lambda \cdot \vec{b}. \quad (6)$$

Interchanging the sums, we first calculate

$$\sum_\lambda \mathrm{sgn}(\vec{v}_i \cdot \vec{v}_\lambda) \vec{v}_\lambda \cdot \vec{b} = 2\gamma \vec{v}_i \cdot \vec{b}, \quad (7)$$

which follows from (4) and (5); see details in the Supplemental Material [23]. Inserting the last expression in (6), we get

$$\langle ab\rangle = -\frac{\gamma}{6}\sum_i \omega_i \vec{v}_i\cdot\vec{b} = -\frac{\ell\gamma}{6}\vec{a}\cdot\vec{b} \simeq -0.43\ \vec{a}\cdot\vec{b}. \quad (8)$$

Finally, we compute Alice's marginal

$$\langle a\rangle = -\frac{1}{12}\sum_\lambda \sum_i \omega_i \,\mathrm{sgn}(\vec{v}_i\cdot\vec{v}_\lambda) = 0, \quad (9)$$

which can be seen from (4). Similarly, we get that $\langle b\rangle = 0$. Therefore, the model simulates $\rho_W(\alpha)$ for $\alpha \simeq 0.43$. Extension to smaller values of α is straightforward.

The above protocol can be adapted to any polyhedron satisfying conditions (4) and (5). Natural candidates are the Platonic solids, except for the tetrahedron which does not satisfy (4). Among these, the icosahedron turns out to be optimal here; see Supplemental Material [23]. Hence, in order to simulate Werner states which are more entangled, i.e., going beyond $\alpha \simeq 0.43$, we need another method.

We now present a protocol, which will allow us to relax condition (5). Specifically, we consider again a three-dimensional polyhedron V with D vertices $\vec{v}_i$, but only demand that is satisfy condition (4) (which can always be achieved at the expense of doubling the number of vertices of a given polyhedron). As before, the shared variable $\lambda \in \{1,\ldots,D\}$ encodes the choice of vertex, and is uniformly distributed. Having abandoned condition (5), we have for each vertex $\vec{v}_\lambda$:

$$\sum_{j\,\mathrm{s.t.}\,\vec{v}_j\cdot\vec{v}_\lambda\geq 0} \vec{v}_j = \gamma_\lambda \vec{m}_\lambda, \quad (10)$$

where $\vec{m}_\lambda$ is a normalized vector and generally $\vec{m}_\lambda \neq \vec{v}_\lambda$. Let us define $\gamma_{\min} = \min_\lambda(\gamma_\lambda)$. Note that there are now two polyhedra of interest: (i) V, that is defined by the vertices $\vec{v}_\lambda$ and (ii) M, defined by the vertices $\vec{m}_\lambda$, which are in one-to-one correspondence with the $\vec{v}_\lambda$. Consider the following protocol.

Protocol 2.—Alice and Bob share $\lambda \in \{1,\ldots,D\}$ uniformly distributed. Upon receiving setting $\vec{a}$, Alice calculates the subnormalized vector $\vec{a}' = \ell\vec{a}$ where ℓ is the radius of the largest sphere fitting inside M and centered on the origin. This ensures that $\vec{a}'$ lies inside the convex hull of M and Alice can therefore find a convex decomposition $\vec{a}' = \sum_{i=1}^D \omega_i \vec{m}_i$. Then, with probability $p_i = \omega_i\gamma_{\min}/\gamma_i$ she outputs $a = \mathrm{sgn}(\vec{v}_i\cdot\vec{v}_\lambda)$, and with probability $(1-\sum_i p_i)$ she outputs a random bit. Bob, upon receiving $\vec{b}$, outputs $b = \pm 1$ with probability $(1 \mp \vec{b}\cdot\vec{v}_\lambda)/2$.

The resulting correlations are given by

$$\begin{aligned}\langle ab\rangle &= -\frac{1}{D}\sum_\lambda\sum_i \omega_i \frac{\gamma_{\min}}{\gamma_i}\mathrm{sgn}(\vec{v}_i\cdot\vec{v}_\lambda)\vec{b}\cdot\vec{v}_\lambda \\ &= -\frac{2\gamma_{\min}}{D}\sum_i \frac{\omega_i}{\gamma_i}\sum_{\lambda\,\mathrm{s.t.}\,\vec{v}_\lambda\cdot\vec{v}_i\geq 0}\vec{v}_\lambda\cdot\vec{b} \\ &= -\frac{2\ell}{D}\gamma_{\min}\vec{a}\cdot\vec{b}, \end{aligned} \quad (11)$$

where we have used Eq. (10) in the last step; see Supplemental Material [23] for details. As for protocol 1, using Eq. (4) we get that the marginals $\langle a\rangle = \langle b\rangle = 0$. Hence, the model reproduces the statistics of $\rho_W(\alpha)$ for $\alpha = (2\ell/D)\gamma_{\min}$.

Starting from a sufficiently regular polyhedron with a large number D of vertices $\vec{v}_\lambda$, we can approximate the unit sphere and the factor ℓ can become arbitrary close to one. In the limit $D\to\infty$ we expect to recover the uniform distribution over the sphere and our model therefore becomes equivalent to Werner's model for $\rho_W(1/2)$ [6]. In Fig. 1 we plot upper bounds on the required shared randomness to simulate $\rho(\alpha)$ as a function of α obtained via protocol 2. We use a family of polyhedra, generated iteratively and starting from the icosahedron. To generate the second polyhedron, we take the union of the icosahedron and its normalized dual (which is the dodecahedron), and so on. One can verify that these polyhedra respect condition (4).

Note that the above protocols are LHS models. Hence the above results can be straightforwardly extended to the simulation of entangled states which are obtained via local filtering on the Werner state, e.g., Ref. [9] (see Supplemental Material [23]). Also, it would be interesting to see if more economical models (i.e., using less shared randomness) exist, and if local entangled states require

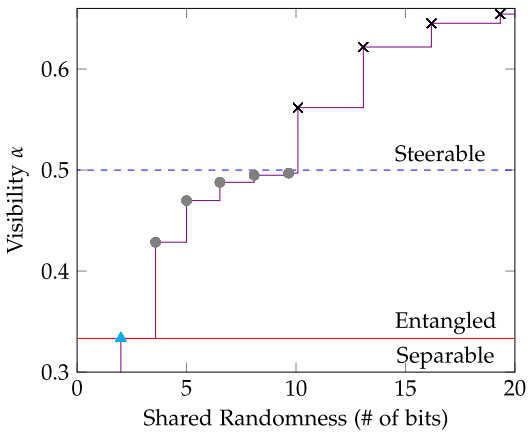

FIG. 1 (color online). Simulation of two-qubit Werner states $\rho_W(\alpha)$ with finite shared randomness. The graph shows the relation between the visibility α (essentially the degree of entanglement) and the amount of shared randomness, quantified in bits. For $\alpha \leq 1/3$ (below the solid line) the state is separable; hence 2 bits of shared randomness suffice (triangle). For $1/3 < \alpha \lesssim 0.43$, the state can be simulated with $\log_2(12)$ bits of shared randomness using protocol 1. For $0.43 \lesssim \alpha \lesssim 0.66$, $\rho_W(\alpha)$ can be simulated with a larger (but nevertheless finite) amount of shared randomness. For $0.43 \lesssim \alpha < 0.5$, we have a LHS model (using protocol 2). For $0.5 < \alpha \lesssim 0.66$ the state becomes steerable but can nevertheless be simulated by a LHV model using finite shared randomness, by applying Result 1 to the model of Ref. [10] (see main text).

more shared randomness compared to separable states. For Werner states, this translates to whether we expect to see a discontinuity at the separable-entangled boundary for $\alpha = 1/3$ (see Fig. 1). We give two partial answers in this direction: (i) for LHS models, the maximum α one can simulate with $D = 4$ is the separable state $\alpha = 1/3$ (see Supplemental Material [23]); (ii) Restricting to equatorial measurements one can achieve $\alpha = 1/2$ with only $D = 4$ (see Supplemental Material [23]).

General results.—In the above, we have focused on a class of highly symmetric states, namely Werner states, and considered only projective measurements. Here we show how local models with finite shared randomness can be constructed for essentially any state that admits a LHV model. We also discuss the case of general measurements, i.e., POVMs.

Result 1: Consider a state ρ (of dimension $d \times d$) admitting a LHV model for all projective measurements. Then, a LHV model using only finite shared randomness can simulate all projective measurements on the state

$$\rho(\eta) = \eta^2\rho + \eta(1-\eta)\left(\frac{\mathbb{I}}{d} \otimes \rho_B + \rho_A \otimes \frac{\mathbb{I}}{d}\right) + (1-\eta)^2 \frac{\mathbb{I} \otimes \mathbb{I}}{d^2}$$

for any $0 \leq \eta < 1$. Here $\rho_{A,B} = \mathrm{Tr}_{B,A}(\rho)$.

Proof.—First, note that it follows from the relation

$$\mathrm{tr}[A_a \otimes B_b \rho(\eta)] = \mathrm{tr}[A_a(\eta) \otimes B_b(\eta)\rho] \qquad (12)$$

that the simulation of projective measurements (given by operators A_a and B_b) on $\rho(\eta)$ is equivalent to the simulation of noisy measurements, given by operators $A_a(\eta) = \eta A_a + (1-\eta)(\mathbb{I}/d)$ and $B_b(\eta) = \eta B_b + (1-\eta)(\mathbb{I}/d)$ on the state ρ. Next, since $A_a(\eta)$ and $B_b(\eta)$ are full rank for any $\eta < 1$, they are not on the border on the set of measurements [24], and can thus be decomposed as convex mixtures over a single set of finitely many projective measurements (more details in the Supplemental Material [23]). Finally, note that the simulation of a finite number of projective measurements on ρ requires only finite shared randomness. This follows from the fact (i) the resulting distribution is local (as ρ admits a LHV model), and (ii) the set of local distributions forms a polytope [3].

Note that the amount of shared randomness needed will depend on the value of η and diverges as $\eta \to 1$.

Result 2: Let us now discuss more general measurements, i.e., POVMs. Consider an entangled state ρ (of dimension $d \times d$) admitting a local model with k bits of shared randomness for projective measurements. We can then construct the state

$$\rho' = \frac{1}{(d+1)^2}[\rho + d(\rho_A \otimes F + F \otimes \rho_B) + d^2 F \otimes F],$$

which admits a local model with k bits of shared randomness for POVMs. Here $F = |d+1\rangle\langle d+1|$ denotes a projector onto a subspace orthogonal to the support of ρ; hence, ρ' is entangled by construction and of local dimension $d+1$. This result follows straightforwardly from Protocol 2 of Ref. [11], since the local model obtained for ρ' makes use of the same shared randomness as the one for ρ.

Finally, we present two examples illustrating the above results. First, applying Result 1 the local model of Ref. [25] allows us to extend our result for two-qubit Werner states. Specifically, we show that $\rho_W(\alpha)$ can be simulated with finite shared randomness for $\alpha \lesssim 0.66$. Upper bounds on the amount of shared randomness are given in Fig. 1 (using again an iterative procedure based on the icosahedron). Notably, this shows that certain states useful for EPR steering can be simulated with finite shared randomness. Second, applying Result 2 to the state [26] $\rho_W(0.43)$, we obtain that the state $\rho = \frac{1}{3}[\rho_W(0.43) + 2|2\rangle\langle 2| \otimes (\mathbb{I}/2)]$ can be simulated for arbitrary POVMs using $\log_2(12)$ bits of shared randomness.

Simulating nonlocal states with finite resources.—Finally, we discuss the simulation of entangled states which are nonlocal. In this case, classical communication from (say) Alice to Bob is required. This communication is sent after Alice has received her input. Two cases can be considered: (i) Alice and Bob are initially uncorrelated (i.e., have no shared randomness), and Alice sends classical information to Bob, (ii) Alice and Bob have access to shared randomness, and Alice sends classical information to Bob. Reference [27] presents a model using no shared randomness and finite expected communication. Other known protocols (see, e.g., [25,28,29]) require, for case (ii), finite communication assisted with infinite shared randomness—hence infinite communication for case (i). Here we present protocols using only finite resources, even in the worst-case scenario.

Considering case (i), we first show that the statistics of any bipartite entangled state ρ of dimension $d \times d$ that is full rank can be simulated with finite communication. Note that a state of full rank does not lie on the border of the set of quantum states [20]. Upon receiving her measurement setting $A = \{A_a\}$, Alice outputs a according to the distribution $p(a) = \mathrm{Tr}(\rho_A A_a)$, where ρ_A is Alice's reduced state. For output a, Bob should hold the (normalized) state $\rho_B^a = \mathrm{Tr}_A(A_a \otimes \mathbb{I}\rho)/p(a)$. Since ρ_B^a is full rank (by construction) there exists a polyhedron V (with D vertices, each representing a pure quantum state of dimension d) such that Alice can decompose ρ_B^a as a convex combination of the vertices of V. With probability ω_i (the coefficient of vertex i in the decomposition) Alice sends label i to Bob, who can then locally reconstruct the corresponding pure state (knowing V). The model thus reproduces the statistics of ρ using $\log_2(D)$ bits of communication.

For case (ii), we show that any state $\rho_W(\alpha)$ [see Eq. (2)], with $\alpha < 1$, can be simulated with finite shared randomness

and finite communication (worst case). In particular, for $\alpha \leq 3/4$ a single bit suffices. To construct such a model, we combine the ideas of Protocol 1 and the simulation model (using 1 bit of communication) for the singlet state of Ref. [28]. See Supplemental Material [23] for details.

Conclusion.—We have shown that the correlations of essentially all entangled states that admit a LHV model can be simulated with finite shared randomness for the case of projective measurements. This shows that the requirement of infinite shared randomness (hence channels with infinite capacity) used in previous models can in fact be dispensed with. Whether this result can be extended to the case of POVMs is a relevant issue.

An interesting open question is to find the minimal amount of shared randomness required to simulate a local entangled state. For a state of local dimension d, are more than $4\log_2(d)$ bits of shared randomness always required, that is, is the simulation of local entangled states strictly more costly than that of separable states? We presented a model using only 2 bits for Werner states of two qubits, but our model works only for equatorial measurements; hence, the question remains open.

We thank Antonio Acín and Marcin Pawlowski for discussions. We acknowledge financial support from the Swiss National Science Foundation (Grant No. PP00P2_138917 and QSIT), and SEFRI (COST action MP1006). J. B. and F. H. contributed equally to this work.

[1] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, Rev. Mod. Phys. **81**, 865 (2009).
[2] J. S. Bell, On the Einstein-Poldolsky-Rosen paradox, Physics (Long Island City, N.Y.) **1**, 195 (1964).
[3] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).
[4] L. Masanes, Y.-C. Liang, and A. C. Doherty, All Bipartite Entangled States Display Some Hidden Nonlocality, Phys. Rev. Lett. **100**, 090403 (2008).
[5] T. Vértesi and N. Brunner, Disproving the Peres conjecture: Bell nonlocality from bipartite bound entanglement, Nat. Commun. **5**, 5297 (2014).
[6] R. F. Werner, Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model, Phys. Rev. A **40**, 4277 (1989).
[7] J. Barrett, Nonsequential positive-operator-valued measurements on entangled mixed states do not always violate a Bell inequality, Phys. Rev. A **65**, 042302 (2002).
[8] R. Augusiak, M. Demianowicz, and A. Acín, Local hidden-variable models for entangled quantum states, J. Phys. A **47**, 424002 (2014).
[9] 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).
[10] A. Acín, N. Gisin, and B. Toner, Grothendieck's constant and local models for noisy entangled quantum states, Phys. Rev. A **73**, 062105 (2006).
[11] F. Hirsch, M. T. Quintino, J. Bowles, and N. Brunner, Genuine Hidden Quantum Nonlocality, Phys. Rev. Lett. **111**, 160402 (2013).
[12] J. Bowles, T. Vértesi, M. T. Quintino, and N. Brunner, One-Way Einstein-Podolsky-Rosen Steering, Phys. Rev. Lett. **112**, 200402 (2014).
[13] M. Cong Tran, W. Laskowski, and T. Paterek, The Werner gap in the presence of simple coloured noise, J. Phys. A **47**, 424025 (2014).
[14] G. Tóth and A. Acín, Genuine tripartite entangled states with a local hidden-variable model, Phys. Rev. A **74**, 030306(R) (2006).
[15] R. Augusiak, M. Demianowicz, J. Tura, and A. Acín, Entanglement and nonlocality are inequivalent for any number of particles, arXiv:1407.3114.
[16] S. Popescu, Bell's Inequalities and Density Matrices: Revealing Hidden Nonlocality, Phys. Rev. Lett. **74**, 2619 (1995).
[17] 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).
[18] M. D. Reid, P. D. Drummond, W. P. Bowen, E. G. Cavalcanti, P. K. Lam, H. A. Bachor, U. L. Andersen, and G. Leuchs, Colloquium: The Einstein-Podolsky-Rosen paradox: From concepts to applications, Rev. Mod. Phys. **81**, 1727 (2009).
[19] S. Jevtic, M. J. W. Hall, M. R. Anderson, M. Zwierz, and H. M. Wiseman, Einstein-Podolsky-Rosen steering and the steering ellipsoid, arXiv:1411.1517.
[20] I. Bengtsson and K. Zyczkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, Cambridge, 2006).
[21] P. Horodecki, Separability criterion and inseparable mixed states with positive partial transposition, Phys. Lett. A **232**, 333 (1997).
[22] W. K. Wootters, Entanglement of Formation of an Arbitrary State of Two Qubits, Phys. Rev. Lett. **80**, 2245 (1998).
[23] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.114.120401 for details of the simulation protocols.
[24] G. Mauro D'Ariano, P. Lo Presti, and P. Perinotti, Classical randomness in quantum measurements, J. Phys. A **38**, 5979 (2005).
[25] O. Regev and B. Toner, Simulating quantum correlations with finite communication, arXiv:0708.0827.
[26] Note that we start here from a LHS, hence the model works directly for POVMs on Bob's side. It is thus sufficient to apply the extension procedure of Result 2 on Alice's side only.
[27] G. Brassard, L. Devroye, and C. Gravel, Exact simulation of the GHZ distribution, arXiv:1303.5942.
[28] B. F. Toner and D. Bacon, Communication Cost of Simulating Bell Correlations, Phys. Rev. Lett. **91**, 187904 (2003).
[29] C. Branciard and N. Gisin, Quantifying the Nonlocality of Greenberger-Horne-Zeilinger Quantum Correlations by a Bounded Communication Simulation Protocol, Phys. Rev. Lett. **107**, 020401 (2011).

Local hidden variable models for entangled quantum states using finite shared randomness
Supplementary Materials

Joseph Bowles,[1,*] Flavien Hirsch,[1,*] Marco Túlio Quintino,[1] and Nicolas Brunner[1]
[1] *Département de Physique Théorique, Université de Genève, 1211 Genève, Switzerland*
(Dated: 25th February 2015)

I. LHS MODELS FOR WERNER STATES USING FINITE SHARED RANDOMNESS

Here, we describe in detail the protocols 1 and 2 of the main text for the simulation of Werner states $\rho_W(\alpha)$ with $\alpha < 1/2$.

Protocol 1. Consider V to be any of the platonic solids (except for the tetrahedron) with D vertices $\vec{v}_i$, satisfying conditions:

$$\forall\ \vec{v}_\lambda\ \exists\ \vec{v}_j \text{ s.t. } \vec{v}_\lambda = -\vec{v}_j \tag{1}$$

$$\sum_{j\ s.t.\ \vec{v}_j\cdot\vec{v}_\lambda\geq 0} \vec{v}_j = \gamma\vec{v}_\lambda \qquad \forall\lambda\ . \tag{2}$$

The shared variable is given by $\lambda \in \{1,\cdots,D\}$, uniformly distributed. Upon receiving λ, Alice and Bob use vector $\vec{v}_\lambda$, and output according to the following response functions:

$$p_A(a|\lambda,\vec{a}) = \frac{1+a\text{sgn}(\vec{v}_\lambda\cdot\vec{a})}{2}, \tag{3}$$

$$p_B(b|\lambda,\vec{b}) = \frac{1-b(\vec{v}_\lambda\cdot\vec{b})}{2}. \tag{4}$$

To begin with, consider the case where Alice's measurements corresponds to one of the vertices of V, *i.e.* $\vec{a}=\vec{v}_i$. Bob's measurement $\vec{b}$ is arbitrary. We obtain the correlator:

$$\begin{aligned}\langle ab\rangle &= -\frac{1}{D}\sum_\lambda \text{sgn}(\vec{v}_i\cdot\vec{v}_\lambda)\vec{b}\cdot\vec{v}_\lambda \\ &= -\frac{1}{D}\Big(\sum_{\lambda|\vec{v}_i\cdot\vec{v}_\lambda\geq 0}\vec{v}_\lambda\cdot\vec{b} - \sum_{\lambda|\vec{v}_i\cdot\vec{v}_\lambda<0}\vec{v}_\lambda\cdot\vec{b}\Big).\end{aligned} \tag{5}$$

From equation (1) we have that

$$\sum_{\lambda|\vec{v}_i\cdot\vec{v}_\lambda\geq 0}\vec{v}_\lambda = -\sum_{\lambda|\vec{v}_i\cdot\vec{v}_\lambda<0}\vec{v}_\lambda \tag{6}$$

hence implying that

$$\langle ab\rangle = -\frac{2}{D}\sum_{\lambda|\vec{v}_i\cdot\vec{v}_\lambda\geq 0}\vec{v}_\lambda\cdot\vec{b} = -\frac{2}{D}\gamma\ \vec{v}_i\cdot\vec{b}$$

where we used equation (2) in the last step. Next we compute the marginals:

$$\langle a\rangle = -\frac{1}{D}\sum_\lambda \text{sgn}(\vec{v}_i\cdot\vec{v}_\lambda) = 0 \tag{7}$$

since for each $\vec{v}_j$ there is an opposite vector $\vec{v}_k = -\vec{v}_j$. Similarly for Bob:

$$\langle b\rangle = -\frac{1}{D}\sum_\lambda(\vec{b}\cdot\vec{v}_\lambda) = 0. \tag{8}$$

Hence the model simulates a Werner state for $\alpha = \frac{2\gamma}{D}$, for the case in which Alice's measurement is one of the vertices of V.

Next we extend the model to an arbitrary projective measurement for Alice, represented by vector $\vec{a}$. Note that for any $\vec{a}$ one can find a set $\{\omega_i\}_{i=1,\ldots,D}$, with $\omega_i \geq 0$ and $\sum_{i=1}^D \omega_i = 1$ such that

$$\sum_{i=1}^D \omega_i\vec{v}_i = \ell\vec{a}, \tag{9}$$

with $\ell < 1$. That is, for each $\vec{a}$ one can find a vector lying in the convex hull of V that lies parallel to $\vec{a}$ and has length ℓ. Hence, ℓ is precisely the radius of the sphere (centered in the origin) inscribed inside V. If upon receiving $\vec{a}$ Alice uses local randomness to simulate the measurement of $\vec{v}_i$ with probability ω_i the overall correlator is given by

$$\langle ab\rangle = \sum_{i=1}^D \omega_i\vec{v}_i\cdot\vec{b} = \ell\vec{a}\cdot\vec{b}. \tag{10}$$

The marginal remains unchanged, *i.e.* $\langle a\rangle = 0$. Hence, the model now simulates a Werner state $\rho_W(\alpha)$ with visibility $\alpha = \frac{2\gamma\ell}{D}$. Indeed the 'shrinking factor' ℓ depends on the choice of polyhedron V.

For each platonic solid, we give the visibility α of the Werner state that is simulated, and the required amount of shared randomness (see Table I). For the dodecahedron and the icosahedron, the model simulates the correlations of an entangled state. Note that the visibility α depends on the ratio of various parameters, hence using a polyhedron with more vertices may result in a lower visibility.

Protocol 2. Protocol 1 can be extended to any polyhedron V (with D vertices) satisfying (1). Hence we now

* These authors contributed equally to this work.

	Shared randomness (bits)	α	Sep/Ent
Octahedron	2.58	0.19	separable
Cube	3	0.29	separable
Dodecahedron	4.32	0.41	entangled
Icosahedron	3.58	0.43	entangled

Table I. For each platonic solid, we give the visibility α of the simulated Werner state. The amount of required shared randomness is given by the number of vertices of the polyhedron.

relax condition (2), and have the relation

$$\sum_{j \ s.t. \ \vec{v}_j \cdot \vec{v}_\lambda \geq 0} \vec{v}_j = \gamma_\lambda \vec{m}_\lambda \tag{11}$$

where $\vec{m}_\lambda$ is a normalized vector and generally $\vec{m}_\lambda \neq \vec{v}_\lambda$. Define $\gamma_{\min} = \min_\lambda(\gamma_\lambda)$. Hence we obtain a second polyhedron M, defined by the vertices $\vec{m}_\lambda$, which are in one-to-one correspondence with the $\vec{v}_\lambda$.

Upon receiving $\lambda \in \{1, ..., D\}$, Alice and Bob use vector $\vec{v}_\lambda$. Similarly to above, let us start with the case where Alice's measurement corresponds to one of the vectors of M, $\vec{a} = \vec{m}_i$. Here we will slightly modify Alice's response function compared to protocol 1. Specifically, Alice now outputs according to

$$p_A(a|\lambda, \vec{m}_i) = \frac{1 + a\mathrm{sgn}(\vec{v}_\lambda \cdot \vec{v}_i)}{2}, \tag{12}$$

with probability $\gamma_{\min}/\gamma_i$, and outputs randomly otherwise. Bob receives an arbitrary projective measurement $\vec{b}$ and outputs as in protocol 1. The correlator is thus given by

$$\begin{aligned} \langle ab \rangle &= -\frac{1}{D} \sum_\lambda \frac{\gamma_{\min}}{\gamma_i} \mathrm{sgn}(\vec{v}_i \cdot \vec{v}_\lambda) \vec{b} \cdot \vec{v}_\lambda \\ &= -\frac{1}{D} \frac{\gamma_{\min}}{\gamma_i} \Big(\sum_{\lambda | \vec{v}_i \cdot \vec{v}_\lambda \geq 0} \vec{v}_\lambda \cdot \vec{b} - \sum_{\lambda | \vec{v}_i \cdot \vec{v}_\lambda < 0} \vec{v}_\lambda \cdot \vec{b} \Big) \\ &= -\frac{2\gamma_{\min}}{D\gamma_i} \sum_{\lambda | \vec{v}_i \cdot \vec{v}_\lambda \geq 0} \vec{v}_\lambda \cdot \vec{b} \\ &= -\frac{2}{D} \gamma_{\min} \ \vec{m}_i \cdot \vec{b}. \end{aligned} \tag{13}$$

Note that condition (1) again ensures that the marginals are uniform. Hence, the model simulates the correlations of a Werner state $\rho_W(\alpha)$ with visibility $\alpha = \frac{2}{D}\gamma_{\min}$ when Alice's measurement corresponds to one of the vertices of M. Following the same reasoning as above (for protocol 1), we can extend the simulation model to the case of an arbitrary projective measurement for Alice. Similarly to above, the resulting visibility is found to be $\alpha = 2\gamma_{\min}\ell/D$, where ℓ is the radius of the sphere inscribed inside M centered on the origin.

Extensions. Protocols 1 and 2 are LHS models. Hence, they can be extended to the simulation of other entangled states, which can be obtained from Werner state via a filtering operation on Bob's side (the trusted party). For instance, Ref. [1] discussed states of the form

$$\rho_{\alpha,\theta} = \alpha \left|\psi_\theta\right\rangle\!\left\langle\psi_\theta\right| + (1-\alpha)\frac{\mathbb{I}}{2} \otimes \rho_B \tag{14}$$

where $\left|\psi_\theta\right\rangle = \cos(\theta)\left|00\right\rangle + \sin(\theta)\left|11\right\rangle$ and $\rho_B = \mathrm{Tr}_A(\left|\psi_\theta\right\rangle\!\left\langle\psi_\theta\right|)$. Our model can be straightforwardly adapted to the above class of states. For a given amount of shared randomness, the model will simulate $\rho_{\alpha,\theta}$ with the same visibility α as for the Werner state.

II. LHS MODELS FOR ENTANGLED WERNER STATES REQUIRE MORE THAN TWO BITS OF SHARED RANDOMNESS

Consider LHS models. Bob's response function is quantum mechanical, given by the trace rule. In order to simulate a Werner state $\rho_W(\alpha)$ with such a model using only two bits of shared randomness, we must have that

$$\langle ab \rangle = \sum_{\lambda=1}^{4} p_\lambda A_\lambda(\vec{a}) \ (\vec{v}_\lambda \cdot \vec{b}) = -\alpha \ \vec{a} \cdot \vec{b}. \tag{15}$$

where $\sum_\lambda p_\lambda = 1$ and $p_\lambda \geq 0$. Here $A_\lambda(\vec{a})$ denotes an arbitrary response function for Alice. For the particular case of $\vec{b} = \vec{a}$, this implies

$$\sum_{k=1}^{4} p_k A_\lambda(\vec{a}) \ (\vec{v}_k \cdot \vec{a}) = -\alpha \tag{16}$$

hence we obtain

$$\alpha \leq \max_\lambda(|\vec{v}_\lambda \cdot \vec{a}|). \tag{17}$$

As this holds for all $\vec{a}$, we have that

$$\alpha \leq \min_{\vec{a}}[\max_\lambda(|\vec{v}_\lambda \cdot \vec{a}|)]. \tag{18}$$

Here the best strategy consists in using the tetrahedron, leading to $\alpha \leq 1/3$. Consequently, any LHS model reproducing the correlations of an entangled Werner state, i.e. $\rho_W(\alpha)$ with $\alpha > 1/3$, requires more than two bits of shared randomness.

III. SIMULATING EQUATORIAL MEASUREMENTS ON WERNER STATES WITH TWO BITS OF SHARED RANDOMNESS

Here we present a model to simulate the statistics of the state $\rho_W(\alpha)$ for $\alpha \leq 1/2$, where all measurement Bloch vectors lie in a plane (taken here to be the $x-y$ plane). Surprisingly, the model only uses two bits of shared randomness. We parametrize Alice's and Bob's measurement vectors on the Bloch equator as $\vec{a} = (\cos(\theta_a), \sin(\theta_a))$ and $\vec{b} = (\cos(\theta_b), \sin(\theta_b))$. Imagine the following model which uses a single bit of shared randomness $\lambda = 0, 1$ with equal

probability. For $\lambda = 0$, Alice outputs according to the probability distribution

$$p_A(a|\lambda = 0, \vec{a}) = \frac{1}{2}(1 + a\cos(\theta_a)) \tag{19}$$

whereas for $\lambda = 1$ she outputs according to

$$p_A(a|\lambda = 1, \vec{a}) = \frac{1}{2}(1 + a\sin(\theta_a)). \tag{20}$$

Bob does exactly the same up to a flip of his output:

$$p_B(b|\lambda = 0, \vec{b}) = \frac{1}{2}(1 - b\cos(\theta_b)) \tag{21}$$

$$p_B(b|\lambda = 1, \vec{b}) = \frac{1}{2}(1 - b\sin(\theta_b)). \tag{22}$$

A short calculation shows that this gives the correlator

$$\begin{aligned}\langle ab\rangle &= -\frac{1}{2}(\cos(\theta_a)\cos(\theta_b) + \sin(\theta_a)\sin(\theta_b)) \\ &= -\frac{1}{2}\vec{a}\cdot\vec{b}.\end{aligned} \tag{23}$$

In order to ensure that we have the correct marginals, we add an additional bit of shared randomness to the model $\mu = 0, 1$ (again uniform). If we have $\mu = 1$ then Alice and Bob should both flip their output, *i.e.*

$$p_A(a|\lambda = 0, \mu, \vec{a}) = \frac{1}{2}(1 + a(-1)^\mu\cos(\theta_a)); \tag{24}$$

$$p_B(a|\lambda = 1, \mu, \vec{a}) = \frac{1}{2}(1 + a(-1)^\mu\sin(\theta_a)) \tag{25}$$

and equivalently for Bob. This then gives uniform marginals $\langle a\rangle = \langle b\rangle = 0$ while keeping the correlator unchanged. Hence, we simulate exactly the statistics of projective equatorial measurements on the state $\rho_W(1/2)$.

It would be interesting to see whether this model can be extended to the whole sphere. Using the techniques of Ref. [2], we did not manage to solve the problem, as the visibility α is reduced in the procedure.

IV. DETAILS OF THE PROOF OF RESULT 1

Here we show that noisy measurements can always be decomposed using a fixed set of finitely many extremal measurements.

Let $\mathcal{A}$ be a projective measurement (acting on $\mathbb{C}^d$) with projectors A_a with $a = 1, .., d$. Let now $\mathcal{A}(\eta)$ be a d-outcome POVM with elements

$$A_a(\eta) = \eta A_a + (1 - \eta)\frac{\mathbb{I}}{d} \tag{26}$$

where $\eta < 1$. This POVM can be decomposed as

$$\mathcal{A}(\eta) = \eta\mathcal{A} + (1 - \eta)\mathcal{I} \tag{27}$$

where $\mathcal{I} = \{\mathbb{I}/d, ..., \mathbb{I}/d\}$ is the identity measurement. It is proven in [3] that a POVM is extremal (i.e. in the border of the set of POVMs) if and only if one of its elements is not full-rank. Hence POVM $\mathcal{A}(\eta)$ is not extremal, since $\eta < 1$. Considering now all possible projective measurements $\mathcal{A}$, and a fixed $\eta < 1$, the set of POVMs $\mathcal{A}(\eta)$ forms a hyper-surface which is strictly contained in the set of all d-outcome POVMs acting on $\mathbb{C}^d$ (i.e. this hyper-surface does not touch the border of the set). Therefore it is possible to construct a polyhedron V made of finitely many points on the border of the set of POVMs which contains this hyper-surface. Hence any point on the hyper-surface (or inside) belongs to the convex hull of V and can then be decomposed as convex combinations of the vertices of V.

V. SIMULATING NONLOCAL WERNER STATES WITH FINITE COMMUNICATION AND FINITE SHARED RANDOMNESS

We now discuss the simulation of a two-qubit Werner state $\rho_W(\alpha)$ for all $\alpha < 1$ with finite communication and finite shared randomness. Consider a polyhedron V with D vertices satisfying (1), with corresponding $\gamma_{\min}$ and shrinking factor ℓ. Our model uses $n\log_2(D)$ bits of shared randomness and $\log_2 n$ bits of communication (in the worst case), and simulates $\rho_W(\alpha)$ for

$$\alpha = \frac{\gamma_{\min}}{\gamma_{\max}}\left(1 - [1 - \frac{2\gamma_{\min}}{D}]^n\right)\ell^2 \tag{28}$$

where $\gamma_{\max} = \max_i(\gamma_i)$. Note that by choosing a symmetric enough polyhedron with $\ell \approx 1$ and $2\gamma_{\min}/D \approx 1/2$ we can simulate a $\rho_W(\alpha)$ for $\alpha \to 3/4$ with $n = 2$ (when $D \to \infty$). Hence, using finite shared randomness and a single bit of communication suffices to simulate a nonlocal quantum state.

Again we consider a polyhedron V with D vertices satisfying (1) from which we can define a second polyhedron M via equation:

$$\sum_{j\ s.t.\ \vec{v}_j\cdot\vec{v}_\lambda \geq 0} \vec{v}_j = \gamma_\lambda \vec{m}_\lambda\ . \tag{29}$$

Let us first discuss the case in which both Alice and Bob's measurement Bloch vectors correspond to one of the vertices of M, *i.e.* Alice gets vector $\vec{a} = \vec{m}_l$ and Bob $\vec{b} = \vec{m}_k$. The protocol is then as follows:

Protocol 4. In each round Alice and Bob receive n numbers $\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$, where each λ_i is uniformly distributed with $\lambda_i \in \{1, \cdots, D\}$. Either Alice will select one of the λ_i or she will reject all of them. Consider a variable T to denote Alice's selection, with $T = 0$ corresponding to rejection. In the first step, Alice concentrates on λ_1 and does one of the following: (i) with probability $|\vec{m}_l \cdot \vec{v}_{\lambda_1}|\gamma_{\min}/\gamma_{\max}$ she selects λ_1 and sets $T = 1$ and moves to the final step, (ii) with probability $1 - |\vec{m}_l \cdot \vec{v}_{\lambda_1}|\gamma_{\min}/\gamma_l$ she discards λ_1 and moves to the second step (concentrating now on λ_2), (iii) she rejects,

sets $T = 0$ and moves to the final step. Hence, at step j (if it is reached), Alice concentrates on λ_j. In the final step, Alice may have selected λ_T or she may have rejected. In the case of rejection ($T = 0$), Alice sends $c = 1$ to Bob and outputs randomly. Otherwise, she sends $c = T$ to Bob and outputs $a = \mathrm{sgn}[\vec{m}_l \cdot v_{\lambda_T}]$. Bob then outputs $b = -\mathrm{sgn}[\vec{v}_k \cdot \vec{v}_{\lambda_c}]$ with probability γ_{min}/γ_k, and otherwise outputs randomly.

For the correlator we have

$$\begin{aligned}\langle ab\rangle = & \qquad (30)\\ &-\frac{\gamma_{\min}}{D^n\gamma_k}\sum_{\{\lambda_i\}}\sum_{t=1}^{n} p(T=t|\{\lambda_i\},\vec{m}_l)\mathrm{sgn}(\vec{m}_l\cdot\vec{v}_{\lambda_t})\mathrm{sgn}(\vec{v}_k\cdot\vec{v}_{\lambda_t})\\ =&-\frac{\gamma_{\min}}{D^n\gamma_k}\sum_{t=1}^{n}\sum_{\lambda_t}^{D}\mathrm{sgn}(\vec{m}_l\cdot\vec{v}_{\lambda_t})\mathrm{sgn}(\vec{v}_k\cdot\vec{v}_{\lambda_t})\\ &\qquad\times\sum_{\{\lambda_{i\neq t}\}} p(T=t|\{\lambda_i\},\vec{m}_l).\end{aligned}$$

From the protocol we have that

$$p(T=t|\{\lambda_i\},\vec{m}_j)=\frac{\gamma_{\min}}{\gamma_{\max}}|\vec{m}_j\cdot\vec{v}_{\lambda_t}|\prod_{j<t}(1-\frac{\gamma_{\min}}{\gamma_j}|\vec{m}_j\cdot\vec{v}_{\lambda_j}|). \qquad (31)$$

From (13) it follows that

$$\begin{aligned}&\frac{\gamma_{\min}}{\gamma_k}\sum_{\lambda_t}|\vec{m}_l\cdot\vec{v}_{\lambda_t}|\mathrm{sgn}(\vec{m}_l\cdot\vec{v}_{\lambda_t})\mathrm{sgn}(\vec{v}_k\cdot\vec{v}_{\lambda_t}) \qquad (32)\\ =&\frac{\gamma_{\min}}{\gamma_k}\sum_{\lambda_t}\vec{m}_l\cdot\vec{v}_{\lambda_t}\mathrm{sgn}(\vec{v}_k\cdot\vec{v}_{\lambda_t})\\ =&2\gamma_{\min}\vec{m}_l\cdot\vec{m}_k.\end{aligned}$$

We then have

$$\langle ab\rangle=\frac{-2\gamma_{\min}^2}{D^n\gamma_{\max}}\vec{m}_l\cdot\vec{m}_k\sum_{t=1}^{n}\sum_{\{\lambda_{i\neq t}\}}\prod_{j<t}(1-\frac{\gamma_{\min}}{\gamma_l}|\vec{m}_l\cdot\vec{v}_{\lambda_j}|) \qquad (33)$$

and so we simulate a Werner state with α given by

$$\alpha=\frac{2\gamma_{\min}^2}{D^n\gamma_{\max}}\sum_{t=1}^{n}\sum_{\{\lambda_{i\neq t}\}}\prod_{j<t}(1-\frac{\gamma_{\min}}{\gamma_l}|\vec{m}_l\cdot\vec{v}_{\lambda_j}|). \qquad (34)$$

We now proceed to simplify the above expression for α and show that it is independent of $\vec{m}_l$. Since each term in the product depends only on a single λ_j we have:

$$\alpha=\frac{2\gamma_{\min}^2}{\gamma_{\max}}\sum_{t=1}^{n}\frac{1}{D^{n-t+1}}\sum_{\{\lambda_{i>t}\}}\prod_{j<t}\sum_{\lambda=1}^{D}\frac{1}{D}(1-\frac{\gamma_{\min}}{\gamma_l}|\vec{m}_l\cdot\vec{v}_{\lambda_j}|).$$

From the definition of γ_l, it follows that $\sum_{\lambda=1}^{D}\frac{1}{D}\frac{\gamma_{\min}}{\gamma_l}|\vec{m}_l\cdot\vec{v}_{\lambda_j}| = 2\gamma_{\min}/D$, and so summing over the $\{\lambda_{i>t}\}$ as well

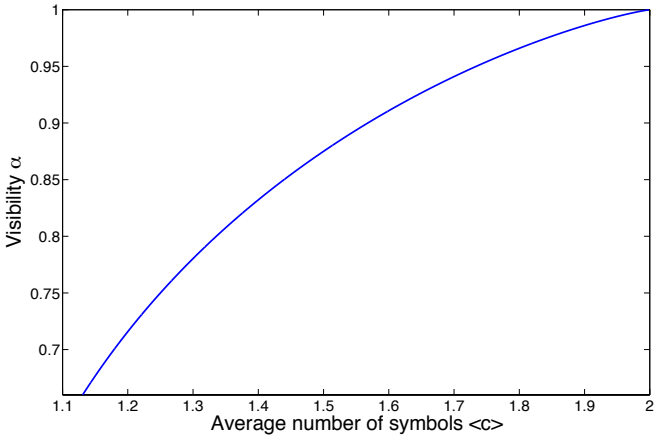

Figure 1. Visibility α of the Werner state $\rho_W(\alpha)$ as a function of the average number of communication $\langle c\rangle$.

we get

$$\begin{aligned}\alpha&=\frac{2\gamma_{\min}^2}{D\gamma_{\max}}\sum_{t=1}^{n}\prod_{j<t}(1-2\gamma_{\min}/D) \qquad (35)\\ &=\frac{2\gamma_{\min}^2}{D\gamma_{\max}}\sum_{t=1}^{n}(1-2\gamma_{\min}/D)^{t-1}\\ &=\frac{\gamma_{\min}}{\gamma_{\max}}(1-(1-2\gamma_{\min}/D)^n),\end{aligned}$$

where in the last line we have used the fact that $\sum_{i=1}^{n}(1-x)^{(i-1)} = (1-(1-x)^n)/x$. Using similar reasoning it is lengthy but straightforward to check that both Alice and Bob's marginals are uniform, *i.e.* $\langle a\rangle = \langle b\rangle = 0$. Finally, we note that we can extend this model to a model for all projective measurements in the same way as previously if Alice and Bob decompose their measurement vectors as convex combinations of vertices of the polyhedron M. This will add a factor ℓ^2 giving the final visibility

$$\alpha=\frac{\gamma_{\min}}{\gamma_{\max}}\left(1-[1-\frac{2\gamma_{\min}}{D}]^n\right)\ell^2. \qquad (36)$$

Average communication. Although the above protocol requires $\log_2(n)$ bits of communication in the worst case, the average amount of communication is typically much smaller, as each λ_i is decreasingly less likely to be selected by Alice. To quantify this, we calculate the average label that is sent by Alice. *i.e.* the average value of the communication c:

$$\begin{aligned}\langle c\rangle&=\frac{1}{D^n}\sum_{\{\lambda_i\}}\sum_{j=1}^{n} j\,p(c=j|\{\lambda_i\}\vec{a}) \qquad (37)\\ &=1+(1-x)x\frac{d}{dx}g_n(x)\end{aligned}$$

with $x = 1-2\gamma_{\min}/D$ and $g_n(x)=\frac{1-x^n}{1-x}$. In the limit of large shared randomness, *i.e.* $D\to\infty$, we get

$$\langle c\rangle = 2-\frac{1+n}{2^n}. \qquad (38)$$

Hence, in this regime, the average value of the communication c remains smaller than 2. Figure 2 shows the visibility α of the simulated state as a function of $\langle c \rangle$. Thus we expect that the model requires only a small amount of average bits of communication although the worst case communication is $\log_2 n$ bits.

[1] 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).

[2] G. Brassard, R. Cleve, and A. Tapp, "Cost of Exactly Simulating Quantum Entanglement with Classical Communication," *Physical Review Letters* **83**, 1874–1877 (1999), quant-ph/9901035.

[3] G. Mauro D'Ariano, P. Lo Presti, and P. Perinotti, "Classical randomness in quantum measurements," *Journal of Physics A Mathematical General* **38**, 5979–5991 (2005).

Paper G

Testing Dimension and Non-classicality in Communication networks

Physical Review A **92**, 022351 (2015)

Joseph Bowles, Nicolas Brunner, and Marcin Pawłowski

PHYSICAL REVIEW A **92**, 022351 (2015)

Testing dimension and nonclassicality in communication networks

Joseph Bowles,[1] Nicolas Brunner,[1] and Marcin Pawłowski[2]
[1]*Département de Physique Théorique, Université de Genève, 1211 Genève, Switzerland*
[2]*Institute of Theoretical Physics and Astrophysics, University of Gdańsk, 80-952 Gdańsk, Poland*
(Received 13 May 2015; published 24 August 2015)

We consider networks featuring preparation, transformation, and measurement devices, in which devices exchange communication via mediating physical systems. We investigate the problem of testing the dimension of the mediating systems in the device-independent scenario, that is, based on observable data alone. A general framework for tackling this problem is presented, considering both classical and quantum systems. These methods can then also be used to certify the nonclassicality of the mediating systems, given an upper bound on their dimension. Several case studies are reported, which illustrate the relevance of the framework. These examples also show that, for fixed dimension, quantum systems largely outperform classical ones. Moreover, the use of a transformation device considerably improves noise tolerance when compared to simple prepare-and-measure networks. These results suggest that the classical simulation of quantum systems becomes costly in terms of dimension, even for simple networks.

DOI: 10.1103/PhysRevA.92.022351 PACS number(s): 03.67.Hk, 03.65.−w

I. INTRODUCTION

The problem of estimating the dimension of an unknown physical system has attracted attention recently. Following early works discussing the problem in the context of Bell inequalities [1–3], a framework was presented for the simplest case of a prepare-and-measure scenario [4]. Such a setup features two devices. First a preparation device, which allows the observers to prepare a physical system in various ways. Second, a measurement device, which allows the observer to perform a measurement on the prepared physical system. It is then possible to find the minimal dimension of the physical system that is compatible with the data. The method is device independent (DI), in the sense that dimension can be certified from the data alone. Techniques tailored for classical [4,5] and quantum [6–8] systems were reported, as well as for the case in which the devices are assumed to be independent [9,10]. The practical relevance of these ideas was recently illustrated [12,13]. Also, the notion of dimension was discussed in more general models beyond quantum theory [11].

A closely related problem is that of testing the nonclassicality of communication. More specifically, considering again the prepare-and-measure setup, it is possible to guarantee the use of quantum communication, under the assumption that the dimension of the system is upper bounded [4]. From a conceptual point of view, this approach aims at quantifying how much classical communication is required to simulate quantum communication [14,15], a relevant problem in the foundations of quantum theory and in communication complexity [16]. Moreover, these ideas are relevant for "semi-device-independent" quantum information processing [17]. Here the correct implementation of a protocol can be guaranteed in a device-independent way, with an additional assumption on the Hilbert space dimension. Protocols for semi-DI quantum key distribution [17–19], randomness certification [20,21], and the characterization of quantum systems [22,23] were discussed, with experimental implementations recently reported [24–26].

More generally, it is natural to consider the problem of testing dimension and nonclassicality in general communication networks, in which black-box devices exchange and process information. To model such a situation, we consider a network composed of preparation devices, transformation devices, and measurement devices (see Fig. 1). First, the preparation devices send out information encoded in physical systems of certain dimension. In turn, these physical systems (and the information they carry) are processed in transformation devices. Finally, the systems are measured (i.e., the information is extracted) using measurement devices. Since we work in the device-independent picture, all devices are represented by black boxes. We therefore have access only to measurement data, that is, the probabilities of obtaining certain measurement results, given the choices of preparations, transformations, and measurements made by the observer. From this data, our goal is then to infer a lower bound on the dimension of the physical systems mediating the information. We will here consider both the case of classical and quantum systems. Moreover, we discuss testing the nonclassicality of communication under the assumption that the dimension is upper bounded. Note that the definition of dimension that we employ here is related to the number of perfectly distinguishable states, i.e., that there should be precisely d perfectly distinguishable states in dimension d. For classical and quantum systems this will coincide with the classical alphabet size and Hilbert space dimension, respectively.

We start by describing the general scenario we consider in Sec. II. Next, we discuss a general framework for addressing this problem for the case of classical systems (Sec. III) and quantum systems (Sec. IV). For the sake of clarity, we present the framework in detail for a simple network, featuring one preparation, one transformation, and one measurement device. We show that the idea of dimension witnesses [4] can be generalized to arbitrary networks, and present methods for deriving optimal witnesses. In Sec. V, we show how dimension witnesses can be used to certify and measure nonclassicality of communication. In order to illustrate the relevance of these methods, we discuss several case studies in Sec. VI, deriving and characterizing dimension witnesses for simple networks. An interesting feature shared by most of these examples is the fact that quantum systems strongly outperform classical

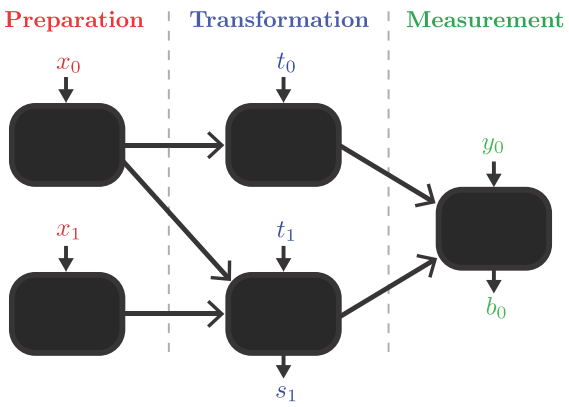

FIG. 1. (Color online) We consider networks featuring preparation, transformation, and measurement devices. All devices receive classical inputs. Transformation and measurement devices provide classical outputs. The arrows between the devices represent communication channels, either quantum or classical.

systems of the same dimension. In fact, we observe a significant enhancement of the advantage offered by quantum systems over classical ones compared to the usual prepare-and-measure scenario. This suggests interesting possibilities for quantum information protocols, and for addressing questions in the foundations of quantum theory. These issues are discussed at the end of the paper, in Sec. VII.

II. GENERAL SCENARIO

The general scenario we wish to consider is a network of devices exchanging and processing information, as represented in Fig. 1. Devices are represented by black boxes. An arrow connecting two devices represents a (one-way) communication channel between them [27].

A network consists of three levels: (i) a number of preparation devices, (ii) a number of transformation devices, and (iii) a number of measurement devices. In each round of the experiment, the observer chooses the preparations $\mathbf{x}$, the transformations $\mathbf{t}$, and the measurement settings $\mathbf{y}$. He then obtains measurement outcomes $\mathbf{b}$; note that transformation devices can also provide outcomes, denoted $\mathbf{s}$. More precisely, we have that the choice of preparations is given by $\mathbf{x} = \{x_i\}$, where x_i denotes the input for device i. The choice of transformations is $\mathbf{t} = \{t_j\}$, where t_j denotes the input for device j, and the (possible) outcomes are $\mathbf{s} = \{s_j\}$, where s_j denotes the output of device j. Finally, the choice of measurement settings is $\mathbf{y} = \{y_k\}$, where y_k denotes the input for measurement device k, and gives outcomes $\mathbf{b} = \{b_k\}$, where b_k is the output of measurement device k. The experiment is therefore characterized by the data

$$p(\mathbf{b},\mathbf{s}|\mathbf{x},\mathbf{t},\mathbf{y}), \tag{1}$$

that is, the conditional probabilities of observing outputs $\mathbf{b},\mathbf{s}$ given inputs $\mathbf{x},\mathbf{t},\mathbf{y}$. A general scenario is thus specified by a directed graph representing the network, and the number of inputs and outputs for each of the devices (which we will here consider to be finite).

In this network, the devices exchange information encoded in physical systems. For instance, upon receiving input x_i, each preparation device emits a system, the state of which is adapted depending on x_i. Which physical system is used, and what mechanism is used to encode information in it, is completely unknown to the observer, who has only access to inputs and outputs of the black boxes. That is, we work in a device-independent scenario.

Now the main point is the following. Clearly, the amount of information about x_i which can be encoded in the system will depend on its dimension (i.e., the number of independent degrees of freedom of the system). Therefore, we expect that a restriction on the dimension will in general limit the possible observable data (1). Consider, for instance, the case in which the outputs $\mathbf{b}$ contain all information about the inputs $\mathbf{x}$. This implies that the mediating physical systems had enough dimensions for encoding $\mathbf{x}$ perfectly.

The main question we will discuss in the present work is to understand the limitations on the data, arising from constraints on the dimension of the mediating systems. This will allow us to find lower bounds on the dimension of the systems present in a network for given data (1). In particular, we will discuss bounds for both classical and quantum systems. Notably, we will see that for a fixed dimension, quantum systems outperform classical ones.

III. CLASSICAL NETWORKS

For the sake of clarity, we will focus on the network consisting of one preparation device, followed by a single transformation device, and finally a single measurement device (see Fig. 2). The data is thus given by the conditional distribution $p(b,s|x,t,y)$; we consider a finite (but otherwise unspecified) number of inputs and outputs. Note that the methods discussed below can be straightforwardly generalized to more general networks.

A. Basics

We start our analysis by considering classical communication between the devices. Denote by c_0 the communication sent from the preparation device to the transformation device, and c_1 the communication sent from the transformation device to the measurement device. We consider communication of bounded dimension d, that is,

$$c_0,c_1 \in \{1,\ldots,d\}. \tag{2}$$

Upon receiving input x, the preparation device sends communication c_0, with probability $p(c_0|x)$. In turn, upon receiving

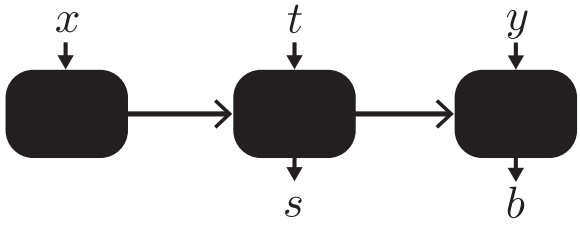

FIG. 2. A simple network consisting of a preparation, a transformation, and a measurement device. The set of possible distributions of inputs and outputs, $p(bs|xty)$, will depend on the dimension of the communication allowed between the devices, and whether the communication is classical or quantum.

input t and communication c_0 (from the preparation device), the transformation device outputs s and sends communication c_1 to the measurement device with probability $p(s,c_1|t,c_0)$. Finally, upon receiving measurement setting y and communication c_1, the measurement device outputs b with probability $p(b|y,c_1)$. We thus have that

$$p(b,s|x,t,y) = \sum_{c_0,c_1=1}^{d} p(c_0|x)p(s,c_1|t,c_0)p(b|y,c_1). \quad (3)$$

We first consider the case in which all devices act deterministically. That is, each of the previously mentioned probabilities are either 0 or 1. It follows that each probability $p(b,s|x,t,y)$ also takes only values 0 or 1. We refer to these sets of data as "deterministic strategies."

In general, we also want to include the possibility that the devices in the network output probabilistically, and moreover that they follow a common strategy. That is, the behavior of the devices might be correlated, due to some (common) internal variable λ (referred to as shared randomness). The set of possible distributions now becomes all convex combinations of deterministic strategies:

$$p(b,s|x,t,y) = \int_\lambda \pi(\lambda)d\lambda \sum_{c_0,c_1=1}^{d} p_\lambda(c_0|x)p_\lambda(s,c_1|t,c_0)p_\lambda(b|y,c_1), \quad (4)$$

where $\pi(\lambda)$ is a normalized probability density over λ and $p_\lambda(c_0|x)$ denotes the probability for the preparation device to send c_0, given input x and internal variable λ, and so on.

Any set of data that cannot be decomposed in the form (4) therefore requires the use of communication (c_0 and/or c_1) of dimension strictly greater than d. In the next sections we will see how to test whether a given set of data can be decomposed in the above form or not. This will provide the "dimension witnesses" we are looking for.

B. Geometrical interpretation

The above ideas admit an elegant description in geometrical terms. Initially developed in the context of Bell nonlocality [28], these ideas were also adapted to the prepare-and-measure scenario [4].

The goal here is to characterize the set of distributions (4) in geometrical terms. Consider first one particular set of data $p(b,s|x,t,y)$. This distribution can be viewed as a vector $\mathbf{p}$ where each component of the vector corresponds to one of the probabilities $p(b,s|x,t,y)$ appearing in the data. Hence $\mathbf{p} \in \mathbb{R}^D$, where

$$D = |b|\,|s|\,|x|\,|t|\,|y| \quad (5)$$

with $|b|$ denoting the alphabet size of b, that is, the number of possible outcomes b, and similarly for other symbols.

Next, consider the entire set of distributions admitting a decomposition of the form (4), that is, all sets of data that can be obtained by using communication c_0 and c_1 of dimension d. This set, denoted $\mathbb{P}_d$, thus forms a subspace of $\mathbb{R}^D$. In fact, $\mathbb{P}_d$ forms a convex polytope. Its extremal points (or vertices) correspond to the deterministic strategies, that is, the set of distributions of the form (3), for which $p(b,s|x,t,y) \in \{0,1\}$ for all b,s,x,t,y. Alternatively, the polytope $\mathbb{P}_d$ can also be characterized by its facets (of which there is a finite number, since the number of vertices is finite). Formally, facets are given by linear inequalities

$$\mathbf{p}\cdot\mathbf{A} = \sum_{b,s,x,t,y} \alpha^{b,s}_{x,t,y}\, p(b,s|x,t,y) \leqslant C_d, \quad (6)$$

where $\alpha^{b,s}_{x,t,y}$ and C_d are real numbers (usually integers). $\mathbf{A}$ is the D-dimensional vector, with components $\alpha^{b,s}_{x,t,y}$, associated to the facet, i.e., orthogonal to the hyperplane given by the facet. Therefore we have that

$$\mathbf{p} \in \mathbb{P}_d \iff \mathbf{p}\cdot\mathbf{A} \leqslant C_d, \quad (7)$$

where the right-hand side means that all facet inequalities are satisfied. Moreover, we have that $\mathbb{P}_d \subseteq \mathbb{P}_{d+1}$, since all strategies involving d-dimensional communication can always be realized using communication of dimension $d+1$.

In practice, the polytope $\mathbb{P}_d$ can be constructed for simple networks, i.e., few devices and small alphabets for the inputs and outputs. Specifically, one starts by listing the deterministic strategies, i.e., the vertices of the polytope. Then, appropriate software (see, e.g., [29,30]) allows one to find the facets of the polytope. Beyond simple cases, however, the problem becomes intractable on standard computers.

Finally, note that one can slightly reduce the complexity of the problem by taking into account certain constraints on the data $p(b,s|x,t,y)$. This allows one to discard certain (redundant) components of $\vec{p}$. In particular, we have here the normalization conditions

$$\sum_{b,s} p(b,s|x,t,y) = 1 \quad \forall x,t,y \quad (8)$$

and the condition that

$$\sum_{b} p(b,s|x,t,y) = p(s|x,t) \quad \forall s,x,t,y. \quad (9)$$

That is, the output s of the transformation device does not depend on the choice of input y for the measuring device. This follows from the fact that y can in principle be chosen after the output s is obtained. For more general networks, it is important to take all such "no-signaling" conditions into account in order to reduce the complexity of the problem.

C. Classical dimension witnesses

Our main goal is to develop methods for testing whether a given set of data $p(b,s|x,t,y)$ is compatible with a particular network sending communication of bounded dimension. To address this question, we will now discuss the concept of "dimension witnesses," hence generalizing the ideas of Ref. [4] to networks.

Consider linear combinations of the form

$$W = \mathbf{w}\cdot\mathbf{p} = \sum_{b,s,x,t,y} \omega^{bs}_{xty}\, p(b,s|x,t,y) \leqslant C_d, \quad (10)$$

where $\mathbf{w}$ is a D-dimensional vector, with real components ω^{bs}_{xty}, and C_d is a real number. We say that an inequality of the above form is a *linear classical dimension witness of dimension* d, if (i) the inequality holds for any distribution

$p(b,s|x,t,y)$ realizable with communication of dimension d, and (ii) there exists at least one distribution $p(b,s|x,t,y)$ (involving systems of dimension at least $d+1$) for which the inequality is violated.

The geometrical ideas discussed in the previous section are relevant here, as they will allow us to construct dimension witnesses. Take one facet inequality of the polytope $\mathbb{P}_d$: property (i) above will immediately be satisfied. In general, there will also exist a vector $\mathbf{p} \in \mathbb{P}_{d'}$ with $d < d'$ that will violate the facet inequality, and hence (ii) is also satisfied. Such facet inequalities will be called "tight dimension witnesses." In fact, the complete list of the facets of $\mathbb{P}_d$ will provide a complete list of dimension witnesses, which allow one to find the minimal dimension of the communication necessary to reproduce a given set of data.

In Sec. VI, we will present several examples of dimension witnesses.

IV. QUANTUM NETWORKS

We now move to the case of quantum communication networks. Here, the classical channels are replaced by quantum channels. Our goal is thus to characterize the sets of data compatible with sending quantum communication of bounded Hilbert space dimension in the network. For the sake of clarity, we will also focus on the simple network of Fig. 2.

A. Basics

Consider again the network consisting of one preparation device, followed by a transformation device, and finally by a measurement device. The devices can now produce, process, and measure quantum systems. The constraint we consider is that the quantum systems transmitting information between the devices are of Hilbert space dimension bounded by d.

Let us first consider the preparation device. Upon receiving input x, the device prepares a d-dimensional quantum system in state ρ_x, which is sent to the transformation device. In turn, the transformation device receives input t, as well as the quantum communication ρ_x, produces an outcome s, and sends a d-dimensional quantum system to the measurement device. The action of the transformation device can thus be represented by a set of completely positive (CP) maps $\{\Phi_{s|t}\}$ (acting on $\mathbb{C}^d$), such that $\sum_s \Phi_{s|t}$ is completely positive and trace preserving: this ensures that $\sum_s p(s|x,t) = 1$ for all x,t. Note that, since we impose that all communication is of bounded dimension d, we restrict to CP maps which do not increase the Hilbert space dimension [31]. With probability $\mathrm{Tr}[\Phi_{s|t}(\rho_x)]$ the transformation device outputs s, and sends the quantum state

$$\Phi_{s|t}(\rho_x)/\,\mathrm{Tr}[\Phi_{s|t}(\rho_x)] \tag{11}$$

to the measuring device. Finally, upon receiving this quantum communication and the input y, the measuring device provides an output b. This is represented by a set of measurement operators $M_{b|y}$ (acting on $\mathbb{C}^d$), such that $M_{b|y} \geqslant 0$ and $\sum_b M_{b|y} = \mathbb{I}$.

Putting all this together we obtain that

$$p(b,s|x,t,y) = \mathrm{Tr}(\Phi_{s|t}(\rho_x) M_{b|y}). \tag{12}$$

Any set of data admitting a decomposition of this form is thus realizable with quantum communication of dimension d. On the contrary, if such a decomposition cannot be found, then higher dimensional quantum systems must have been used.

As in the case of classical networks, it is also relevant to allow for the devices to act according to a common strategy λ. In this case, the set of compatible distributions is therefore the convex hull of those of the form (12):

$$p(b,s|x,t,y) = \int_\lambda \mathrm{Tr}\left(\Phi^\lambda_{s|t}\left(\rho^\lambda_x\right) M^\lambda_{b|y}\right)\pi(\lambda)d\lambda, \tag{13}$$

where now the states, transformations, and measurements are written with λ dependence. Finally, note that one could also consider the case in which the devices share quantum correlations, i.e., initial entanglement (see Sec. VI D for an example).

B. Quantum dimension witnesses

The problem is now to test whether a given set of data $p(b,s|x,t,y)$ is compatible with a particular network sending quantum communication of bounded Hilbert space dimension. Similarly to the classical case discussed above, we now define "quantum dimension witnesses."

Consider again linear inequalities of the form

$$W = \mathbf{w}\cdot\mathbf{p} = \sum_{b,s,x,t,y} \omega^{bs}_{xty}\, p(b,s|x,t,y) \leqslant Q_d, \tag{14}$$

with $\mathbf{w}$ a D-dimensional vector, with real components ω^{bs}_{xty}, and Q_d a real number. In analogy to the classical case, W is a *linear quantum dimension witness of dimension d* if (i) the above inequality is satisfied by all sets of data $p(b,s|x,t,y)$ realizable with quantum communication of dimension d, and (ii) using quantum communication of dimension greater than d allows one to violate the inequality.

Finding quantum dimension witnesses is generally a harder task than in the classical case. To the best of our knowledge, there are no known efficient computational methods for this problem; see, however, Ref. [8] for recent progress.

V. TESTING NONCLASSICALITY

An interesting development related to dimension tests is the possibility of certifying nonclassicality of communication in a device-independent way, assuming an upper bound on the dimension. This aspect was discussed in Ref. [4] for simple prepare-and-measure scenarios. Here we consider this problem in the context of more general networks.

Before moving on, it is important to understand why an assumption on the dimension is necessary in order to make the problem nontrivial. Consider, for instance, the network of Fig. 2. If the dimension is not limited, then the input settings of the preparation and transformation devices, x and t, can be perfectly transmitted to the final measurement device. Since the transformation device has all information about x and t, and the measuring device has all information about x,t,y, it follows that any possible statistics $p(b,s|x,t,y)$ can be reproduced. This implies that nontrivial bounds can only be placed if $|c_0| < |x|$ and/or $|c_1| < |x||t|$.

A. Nonclassicality tests based on dimension witnesses

Considering systems of a fixed dimension, quantum communication can outperform classical communication. This advantage can be revealed by using dimension witnesses. Specifically, by using a well-chosen quantum strategy involving states of Hilbert space dimension d, it is possible to violate certain classical dimension witnesses of dimension d. More formally, we say that a dimension witness with the property

$$W = \mathbf{w}\cdot\mathbf{p} \leqslant C_d < Q_d \tag{15}$$

can be used as nonclassicality tests for systems of dimension d. Consider a set of data $\mathbf{p}_Q$ such that $W = \mathbf{w}\cdot\mathbf{p}_Q > C_d$. This implies the use of genuinely quantum systems for reproducing $\mathbf{p}_Q$, under the assumption that the experiment involves systems of dimension d. In Sec. VI, we will discuss several examples.

B. Quantifying quantum advantage

It is useful to quantify the advantage offered by quantum resources over classical ones. In the present context, several figures of merit can be considered. First, the amount of violation of a given dimension witness could be used, however this will generally depend on how the witness is expressed, and will not allow one to compare different witnesses. Hence, here we use the notion of noise tolerance, which has a more physical interpretation, and will allow us to compare various witnesses.

Consider a quantum experiment (with systems of dimension d) and its corresponding set of data $\mathbf{p}_Q$, which is found to violate a classical dimension witness, i.e., $W = \mathbf{w}\cdot\mathbf{p}_Q > C_d$. The noise tolerance of the quantum point $\mathbf{p}_Q$ for this dimension witness is defined as the minimal fraction of white noise, η, such that the distribution

$$\mathbf{p}_0 = (1-\eta)\mathbf{p}_Q + \eta\mathbf{p}_{\mathbb{I}} \tag{16}$$

does not violate the witness, i.e., $W = \mathbf{w}\cdot\mathbf{p}_0 = C_d$. Here $\mathbf{p}_{\mathbb{I}}$ denotes white noise, i.e., $p_{\mathbb{I}}(b,s|x,t,y) = \frac{1}{|b|\,|s|}$ is the uniform distribution for all x,t,y.

In a practical context, considering noisy distributions of the form (16) is quite natural, due to unavoidable technical imperfections, e.g., losses or misalignment of the preparations.

C. Bounded noise tolerance in prepare-and-measure scenarios involving qubits

It turns out that the noise tolerance of qubit strategies is bounded for any dimension witness in the prepare-and-measure scenario. More precisely, any set of data obtained from qubits and projective measurements can be reproduced using one classical bit if the noise level η satisfies

$$\eta \geqslant \eta^* = 1 - \frac{1}{k_3} \approx 0.34, \tag{17}$$

where k_3 is the Grothendieck constant [32] of order three [33]. Hence, in the prepare-and-measure scenario, no dimension witness for classical bits and projective measurements can be violated for $\eta \geqslant \eta^*$.

We give a proof of the above statement. Consider that the choice of preparation is specified by a vector $\vec{x} \in \mathbb{R}^3$, which represents the Bloch vector of the desired qubit state. Similarly the measurement is specified by a Bloch vector $\vec{y}$, representing the observable $M_{\vec{y}} = \vec{y}\cdot\vec{\sigma}$ (with outcomes $b = \pm 1$), where $\vec{\sigma} = (\sigma_x,\sigma_y,\sigma_z)$ denotes the vector of Pauli matrices. The expected data is therefore

$$p(b|\vec{x},\vec{y}) = \frac{1 + b\,\vec{x}\cdot\vec{y}}{2}. \tag{18}$$

Any such data can be reproduced classically by sending two bits [34]. In order to see this, consider that the preparation and measurement devices share a singlet state $|\psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$. In order to prepare a qubit state corresponding to vector $\vec{x}$, measure the observable $\vec{x}\cdot\vec{\sigma}$ on (half of) the singlet. The result of this measurement is $a = \pm 1$. Then, the state of the other half of the singlet (held by the measuring device) is given by the Bloch vector $-a\vec{x}$. By performing a measurement of the observable $-a\vec{y}\cdot\vec{\sigma}$ on this half of the state, we recover the data (18). The protocol thus requires one bit of communication (to send a), and one singlet state. Using only classical resources, the protocol requires two bits of communication (as the simulation of the singlet state can be done with one bit of communication [34]).

Now, let us see what one can do using only a single bit of communication. The main point is that the simulation of a sufficiently noisy singlet state can be done without communication. That is, there exists a local hidden variable model (for projective measurements) for the state

$$\rho = w\,|\psi^-\rangle\langle\psi^-| + (1-w)\mathbb{I}/4 \tag{19}$$

for $w \leqslant 1/k_3 \leqslant 0.66$ [35]. Considering such a noisy singlet state in the above protocol, we see that it is possible to simulate the data (18) with probability w; with probability $1-w$ we obtain the distribution $\mathbf{p}_{\mathbb{I}}$. Hence, with a noise level $\eta \geqslant \eta^* = 1 - \frac{1}{k_3}$, any qubit strategy can be simulated with one classical bit (and shared randomness).

As mentioned, the above result holds only if the measurement device performs a projective measurement. Since any two outcome qubit measurement can be written as a convex mixture of projective measurements, the result can be extended to all two outcome scenarios. One can extend further to general positive operator-valued measurements at the cost of a larger η^* by using Werner's model [36] for the state (19) with $w = \frac{1}{2}$, leading to $\eta^* = \frac{1}{2}$. This follows from the fact that Werner's model can be seen as a local hidden state model [37], hence the model is valid if general measurements are performed on one side (the trusted party).

VI. CASE STUDIES

We now present several case studies, illustrating the relevance of the concepts and tools discussed above. We first discuss two examples of networks of the form Fig. 2, where preparation, transformation, and measurement devices are "in a line." We then discuss two examples based on a different network, featuring two separate preparation devices and one measurement device. Note that such a network has been considered in different contexts. Notably, this was studied in communication complexity, in the so-called simultaneous message passing model [16], e.g., quantum fingerprinting [38], but also for the black-box certification of entangled measurements [23,26,39], and the Pusey-Barrett-Rudolph theorem [40].

In all cases quantum systems are shown to provide significant advantage over classical systems of the same dimension. Moreover, in all examples (except for the third one), this quantum advantage is stronger compared to the simple prepare-and-measure scenario, in terms of noise tolerance. This suggests that the simulation of quantum strategies becomes significantly harder in the case of networks, even if they feature only few devices.

A. Three devices in a line: Simple case

We start with the network of Fig. 2, considering one of the simplest (nontrivial) configurations in terms of the number of inputs and outputs. Specifically, we have $|x| = 3$ and $|t| = |y| = |b| = 2$. Note that the transformation device does not give any outcome (i.e., $|s| = 1$). We label the inputs and outputs: $x \in \{0,1,2\}$ and $t,y,b \in \{0,1\}$. Hence a set of data is characterized by $D = 24$ probabilities $p(b|x,t,y)$. However, considering normalization conditions, this number is reduced to 12; specifically, the probabilities $p(1|x,t,y) = 1 - p(0|x,t,y)$ are redundant and can thus be omitted.

Applying the method described in Sec. III B we have fully characterized the polytope $\mathbb{P}_2$, that is, the set of distributions achievable for $c_0,c_1 \in \{0,1\}$. Using the software PORTA, we could find the complete list of facets of $\mathbb{P}_2$, which can be grouped (under relabeling of inputs and outputs) into 1870 inequivalent classes of dimension witnesses [41].

Here, we present one class of tight dimension witnesses, a member of which can be written in simple form:

$$W_J = p_{011} + p_{101} + p_{110} + p_{200} - p_{000} - p_{001} - p_{010} - p_{211} \leqslant 2, \quad (20)$$

where we write $p_{xty} = p(b = 0|x,t,y)$. A simple strategy using $c_0,c_1 \in \{0,1\}$ that reaches $W_J = 2$ is as follows. The preparation device sends $c_0 = 0$ for inputs $x = 0,2$, but sends $c_0 = 1$ if $x = 1$. Upon receiving c_0 and input t, the transformation device sends $c_1 = c_0 \oplus t$ to the measurement device (where $\oplus$ denotes addition modulo 2). Finally, the measurement device outputs $b = c_1 \oplus y$. Note also that using classical trits, $c_0,c_1 \in \{0,1,2\}$, we can achieve $W_J = 4$, the maximal possible value.

Using qubits we can significantly outperform classical bits. Consider general pure qubit preparations:

$$|\psi(\theta,\phi)\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + \sin\left(\frac{\theta}{2}\right)\exp(i\phi)\,|1\rangle\,. \quad (21)$$

Specifically, for preparations $x = 0,1,2$ take $|\psi(\frac{\pi}{2},0)\rangle$, $|\psi(\frac{\pi}{2},\frac{3\pi}{4})\rangle$, and $|\psi(\frac{\pi}{2},\frac{-3\pi}{4})\rangle$, respectively. Next consider the transformation device, parametrized by

$$\Phi_{t=0} = \mathbb{I}_2, \quad \Phi_{t=1} = \exp\left(-i\frac{\pi}{4}\sigma_z\right), \quad (22)$$

where $\sigma_z = \mathrm{diag}(1,-1)$ is the Pauli z matrix. Finally, for the measuring device, we have the measurement operators

$$M_{0|0} = \left|\psi\left(\tfrac{\pi}{2},\tfrac{-3\pi}{4}\right)\right\rangle\!\left\langle\psi\left(\tfrac{\pi}{2},\tfrac{-3\pi}{4}\right)\right|, \quad (23)$$

$$M_{0|1} = \left|\psi\left(\tfrac{\pi}{2},\tfrac{3\pi}{4}\right)\right\rangle\!\left\langle\psi\left(\tfrac{\pi}{2},\tfrac{3\pi}{4}\right)\right|. \quad (24)$$

Calculating the resulting probabilities, via Eq. (12), and inserting them into Eq. (20), we obtain

$$W_J = 2 + \sqrt{2} \approx 3.41. \quad (25)$$

The above qubit strategy thus clearly violates the witness (20), and can therefore not be reproduced with classical bits; classical trits must be used. Numerical optimization strongly suggests that this qubit strategy is optimal.

The noise tolerance of the above qubit strategy is

$$\eta = \sqrt{2} - 1 \approx 0.41. \quad (26)$$

Notably, this value exceeds the bound $\eta^* \approx 0.34$ (see Sec. V B) for any prepare-and-measure scenario. Hence the advantage offered by qubits compared to classical bits is stronger compared to what is possible in the prepare-and-measure scenario.

B. Distributed 3 → 1 random access code

As a second example, we consider a task inspired from the information-theoretic task of a random-access code (RAC) [42].

Specifically, we consider a distributed version of the 3 → 1 RAC featuring three devices in a line [see Fig. 3(a)]. Consider three bits a_0,a_1,a_2 randomly taken from a uniform distribution. These bits will determine the inputs of the preparation and transformation devices, namely, $x = (a_0,a_1)$ and $t = a_0 \oplus a_2$. Again, the transformation device has no output. The measuring device has a ternary input $y = 0,1,2$. Similarly to a RAC, the goal is to have the output $b = a_y$. Hence we can define the following witness (for the scenario $|x| = 4$, $|t| = |b| = 2$, $|y| = 3$, and $|s| = 1$) which is the average success probability:

$$W_{\text{D-RAC}} = \tfrac{1}{24}\sum_{\substack{a_0a_1\\a_2y}} p(b = a_y|x = (a_0,a_1),t = (a_0 \oplus a_2),y) \leqslant C_d.$$

We first discuss the case of classical communication. For bits we obtain the bound $C_2 = \frac{2}{3}$, which can be achieved as follows. The preparation device sends $c_0 = a_0$ to the transformation device, who in turn sends $c_1 = c_0$ to the measurement device for both inputs $t = 0,1$. The measurement device outputs $b = c_1 = a_0$. Hence, for $y = 0$ we always have $b = a_y$. However, for $y = 1,2$, success is only achieved with probability $1/2$. Overall, this leads to $C_2 = \frac{2}{3}$. For the case of classical trits, $c_0,c_1 \in \{0,1,2\}$, we get $C_3 = 19/24$. In order to achieve success with probability one, i.e., $W_{\text{D-RAC}} = 1$, eight-dimensional systems are required.

Next, we discuss quantum strategies. Using qubits, we can achieve up to

$$W_{\text{D-RAC}} = Q_2 = \frac{1}{2}\left(1 + \frac{1}{\sqrt{3}}\right) \approx 0.79. \quad (27)$$

The optimal strategy is the following. For input $x = (a_0,a_1)$, choose preparations

$$\left|\psi\left((-1)^{a_1}\arccos\left(\frac{1}{\sqrt{3}}\right) + \pi a_1, \frac{\pi}{4} + \pi a_0\right)\right\rangle, \quad (28)$$

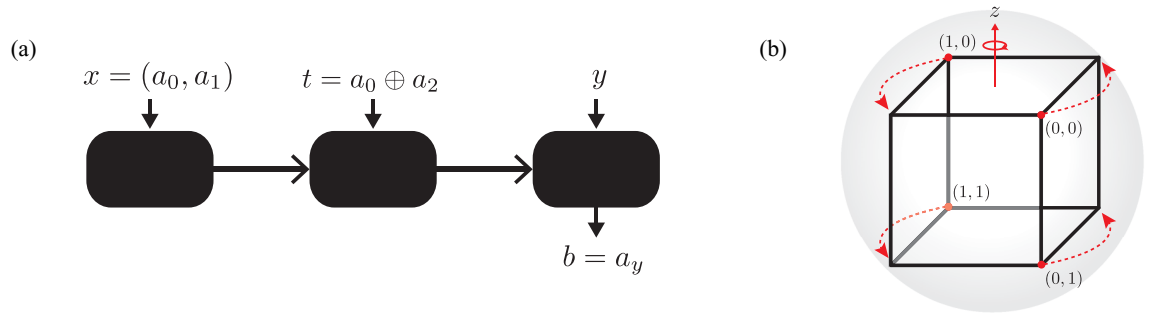

FIG. 3. (Color online) (a) The network corresponding to the distributed $3 \to 1$ random-access code (case study B). Three random bits a_0,a_1,a_2 are used to generate the inputs x and t. Upon receiving input $y = 0,1,2$, the measurement device should output $b = a_y$. The dimension witness $W_{\text{D-RAC}}$ [see Eq. (28)] quantifies the average success probability. (b) Optimal qubit strategy. The four qubit preparations [red dots, corresponding to (a_0,a_1)] are given by the vertices of a cube inscribed inside the Bloch sphere. Upon receiving input $t = a_0 \oplus a_2 = 1$, the transformation device performs a rotation of $\pi/2$ around the z axis if $t = 1$, and the identity otherwise. Finally, by performing a measurement in the x,y,z directions, maximal information about a_0,a_1,a_2 (respectively) is obtained.

which lie at four of the vertices of the cube inscribed inside the Bloch sphere [see Fig. 3(b)]. The transformations are given by

$$\Phi_{t=0} = \mathbb{I}_2, \quad \Phi_{t=1} = \exp\left(i\frac{\pi}{4}\sigma_z\right). \tag{29}$$

Finally, the measuring device performs a measurement in one of three mutually unbiased bases:

$$\begin{aligned} M_{0|0} &= \left|\psi\left(\tfrac{\pi}{2},0\right)\right\rangle\!\left\langle\psi\left(\tfrac{\pi}{2},0\right)\right| = |+x\rangle\langle+x|, \\ M_{0|1} &= |\psi(0,0)\rangle\langle\psi(0,0)| = |+z\rangle\langle+z|, \\ M_{0|2} &= \left|\psi\left(\tfrac{\pi}{2},\tfrac{\pi}{2}\right)\right\rangle\!\left\langle\psi\left(\tfrac{\pi}{2},\tfrac{\pi}{2}\right)\right| = |+y\rangle\langle+y|. \end{aligned} \tag{30}$$

The noise tolerance of this strategy is given by

$$\eta = 1 - \frac{1}{\sqrt{3}} \approx 0.43, \tag{31}$$

which again exceeds the bound for the prepare-and-measure scenario, $\eta^* \approx 0.34$.

Finally, let us comment on the relation of the above game and the standard (prepare-and-measure) $3 \to 1$ RAC. We first note that the optimal qubit strategies for $W_{\text{D-RAC}}$ and the standard RAC are in fact essentially the same [43]. Specifically, the qubit states arriving at the measuring device are identical in both cases [given inputs (a_0,a_1,a_2)]. Hence, this qubit is unaffected by the fact that the inputs are now distributed between the preparation and transformation devices. Indeed, the ability of implementing unitary transformations is central here.

Interestingly, the situation is very different for the case of classical bits. While the average probability of success is 3/4 in the standard RAC, the fact that the inputs are now distributed decreases the average score to 2/3. The reason for this is that the optimal strategy in the standard RAC is to send $c = \text{maj}(a_0,a_1,a_2)$, where $\text{maj}(\cdot)$ denotes the majority function. However, using this strategy requires access to all the input bits a_0,a_1,a_2, which none of the devices in distributed RAC has. The consequence of this is that the noise tolerance of qubit strategies is enhanced in the distributed version of the game, as we showed above.

C. Two preparation devices, one measurement device: Simple case

We now consider a scenario with two preparation devices sending communication to a measurement device [see Fig. 4(a)]. A simple nontrival scenario here is one in which both preparation devices receive a ternary input. We denote the input of the first device $x_0 \in \{0,1,2\}$, and the input of the second $x_1 \in \{0,1,2\}$. The measurement device has no input (i.e., a fixed measurement) and provides a binary output $b = \{0,1\}$. That is, we have $|x_0| = |x_1| = 3$, $|y| = 1$, and $|b| = 2$.

We consider the case in which the channels carry classical bits, i.e., $c_0,c_1 \in \{0,1\}$. In this case we have fully characterized the polytope $\mathbb{P}_2$: it features 13 nontrivial classes of facets which we present in Appendix A. Here we focus on one particular class (witness 1 in Appendix A), represented by the following witness:

$$\begin{aligned} W_K = &-p_{00} + p_{01} + p_{02} - p_{10} \\ &- p_{12} + p_{20} + p_{21} - p_{22} \leqslant 2, \end{aligned}$$

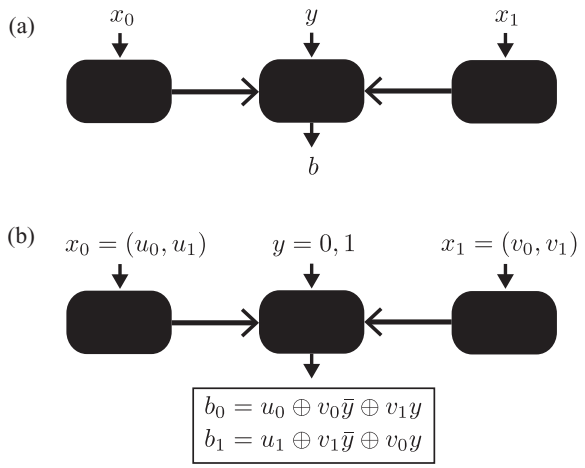

FIG. 4. (a) A simple network involving two preparation devices (left and right) and a single measurement device (center). (b) In case study D, we discuss a dimension witness for this network, referred to as nonlocal dense coding.

where $p_{x_0x_1} = p(b = 0|x_0,x_1)$. An optimal classical bit strategy is as follows. The first preparation device sends $c_0 = 0$ for $x_0 = 0,1$ and $c_0 = 1$ for $x_0 = 2$. The second preparation device sends $c_1 = 1$ for $x_1 = 0,1$ and $c_1 = 0$ for $x_1 = 2$. The measurement device then outputs $b = c_0c_1 \oplus 1$. Clearly, sending classical trits achieves the maximum $W_K = 4$.

Let us now discuss strategies involving qubits. Via numerical optimization we expect a maximal quantum violation of

$$W_K = Q_2 = \tfrac{5}{2}. \tag{32}$$

This can be achieved using the following strategy. The two preparation devices prepare the same states, i.e., we have $\rho_{x_1} = \rho_{x_2}$ for $x_1 = x_2$. For inputs $x_1 = x_2 = 0,1,2$, the preparations are

$$|\psi(-\alpha,0)\rangle\,, \quad |0\rangle\,, \quad |\psi(\alpha,0)\rangle\,, \tag{33}$$

respectively, with $\alpha = 2\arccos\sqrt{\frac{3}{8}}$. The measurement operator for outcome $b = 0$ is a projection onto the entangled subspace:

$$M_0 = |\phi^-\rangle\langle\phi^-| + |\xi(\gamma)\rangle\langle\xi(\gamma)|\,, \tag{34}$$

with $\gamma = \arccos\sqrt{\frac{1}{10}}$ and where

$$|\xi(\gamma)\rangle = \cos\gamma\,|01\rangle - \sin\gamma\,|10\rangle\,. \tag{35}$$

The corresponding noise tolerance is $\eta = 0.2$.

It is relevant to consider a situation in which one channel sends a qubit, while the other one sends a classical bit. Performing numerical optimization, we find a maximal value of $W_K \approx 2.337$ for this case.

Finally, one may also ask if this witness could be used to detect entangled measurements, similarly to Ref. [23]. Specifically, one can derive an upper bound on W_K for separable measurement operators of the form $M_b = \sum_i M^i_{b,1} \otimes M^i_{b,2}$ where $M^i_{b,k}$ is a positive operator acting on the system sent by preparation device k. Numerical tests suggest that the optimal value is $W_K \approx 2.337$. Hence we find the same value as for the above case of hybrid qubit/bit channels. Therefore, we expect that a value $W_K > 2.337$ certifies that (i) both channels send qubits and (ii) the measurement is nonseparable, i.e., has (at least) one entangled eigenstate. Note that the witness (33) has been discussed before in [44] in a similar context, where upper bounds of $W_K \approx 2.506$ and $W_K \approx 2.377$ were found for the case of general and unentangled measurements, supporting our findings.

D. Nonlocal dense coding

As the last example, we present a dimension witness for a task which can be viewed as a nonlocal version of dense coding [45]. As in the previous example, we consider the case of two preparation devices and one measuring device.

Here each preparation device receives two input bits: $x_0 = (u_0,u_1)$ for the first and $x_1 = (v_0,v_1)$ for the second. The measurement device receives $y = 0,1$ as input, and provides two output bits $\mathbf{b} = (b_0,b_1)$. The rules of the game are the following [see Fig. 4(b)]. On the one hand, for $y = 0$, the outputs should satisfy $(b_0,b_1) = (u_0 \oplus v_0, u_1 \oplus v_1)$. On the other hand, for $y = 1$, the output bits should satisfy $(b_0,b_1) = (u_0 \oplus v_1, u_1 \oplus v_0)$. Furthermore, there is a penalty if both b_0 and b_1 are guessed incorrectly. This corresponds to the witness

$$W_D = \langle(b_0,b_1) = (u_0 \oplus v_0\bar{y} \oplus v_1y, u_1 \oplus v_1\bar{y} \oplus v_0y)\rangle - \langle(\bar{b}_0,\bar{b}_1) = (u_0 \oplus v_0\bar{y} \oplus v_1y, u_1 \oplus v_1\bar{y} \oplus v_0y)\rangle \leqslant C_d, \tag{36}$$

where $\bar{y} = y \oplus 1$, and the average $\langle\cdot\rangle$ is taken over all inputs:

$$\langle(b_0,b_1)\rangle = \frac{1}{32}\sum_{\substack{u_0,u_1\\v_0,v_1,y}} p(b_0,b_1|u_0,u_1,v_0,v_1,y). \tag{37}$$

Let us discuss the case of classical communication. For bits, we have $C_2 = \frac{1}{4}$ which can be achieved as follows. The first preparation device sends communication $c_0 = u_0u_1$. Similarly, the second device sends $c_1 = v_0v_1$. The measurement device then outputs $(b_0,b_1) = (c_0 \oplus c_1, c_0 \oplus c_1)$. Using classical trits, we get $C_3 = \frac{9}{16}$. Indeed, sending four-dimensional systems achieves success probability 1.

Next, consider qubit strategies (see Appendix B for more details). Here we can achieve

$$W_D = Q_2 = \tfrac{1}{2}, \tag{38}$$

which appears optimal from numerical tests. This corresponds to a noise tolerance of $\eta = \frac{1}{2}$, which represents a considerable improvement over the simple prepare-and-measure scenario. The strategy is the following. The preparation devices send qubit states

$$\sigma_x^{u_1}\sigma_z^{u_0}\big|\psi\big(\tfrac{\pi}{4},0\big)\big\rangle, \tag{39}$$

$$\sigma_x^{v_1}\sigma_z^{v_0}\big|\psi\big(-\tfrac{3\pi}{4},0\big)\big\rangle \tag{40}$$

for the first and second preparation devices, respectively. The measurement device then performs a projective measurement onto the entangled basis

$$M_{b_0b_1|y} = |\phi(b_0,b_1,y)\rangle\langle\phi(b_0,b_1,y)|\,, \tag{41}$$

where

$$|\phi(b_0,b_1,y)\rangle = \sigma_x^{b_1}\sigma_z^{b_0} \otimes H^y\,|\psi^-\rangle\,, \tag{42}$$

$|\psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$ is the singlet state, and $H = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix}1 & 1\\ 1 & -1\end{smallmatrix}\right)$ is the Hadamard matrix. Note that by using qutrits, one can reach $Q_3 \approx 0.598$ according to numerical optimization. Hence we obtain the following relations $C_2 < Q_2 < C_3 < Q_3$.

Additionally, one may also wish to consider the possibility that the devices share quantum correlations (i.e., initial entanglement). Allowing for this considerably enhances the success probability (still using qubit communication), which becomes maximal, that is $W_D = 1$. The strategy is the following. The preparation devices now share a singlet state. Upon receiving the inputs $x_0 = (u_0,u_1)$ and $x_1 = (v_0,v_1)$, the preparation devices locally rotate the singlet state to

$$\left(\sigma_x^{u_1}\sigma_z^{u_0}\right) \otimes \left(\sigma_x^{v_1}\sigma_z^{v_0}\right)|\psi^-\rangle\,. \tag{43}$$

The measurement device performs the same measurement as above [see (41)]. The noise tolerance for this strategy is $\eta = \frac{3}{4}$.

VII. DISCUSSION

We have discussed the problem of testing the dimension and nonclassicality in communication networks. We have presented methods for addressing these problems, generalizing the concept of dimension witnesses to networks, and discussed several illustrative examples.

We believe our results raise several natural questions. First, it would be interesting to investigate the separation between classical and quantum dimension in more general networks. In particular, what is the classical communication cost (i.e., how many classical dimensions are required) for simulating qubit networks? A potential direction for tackling this problem would be to find a family of dimension witnesses for a scenario featuring one preparation device and one measurement device, but any number of transformation devices in between (here we gave examples for the case of a single transformation device). Notably, Galvão and Hardy [14] proved that, in the case of an infinite number of transformation devices, classical systems of infinite dimension are required for simulating a single qubit. The game discussed in [14] can be recast as a dimension witness. Proving a similar result for the case of a finite number of transformation devices would be relevant. Going beyond qubits is also interesting. In fact, for quantum systems of dimension $d \geqslant 3$, it is not known whether an exact simulation is possible with classical systems of finite dimension, even in the simplest prepare-and-measure scenario.

From a more applied perspective, the ideas discussed could find applications in quantum information processing. Recent works discussed protocols for which the security is based on dimension witnesses, so-called semi-device-independent protocols [17,19–22]. For instance, quantum key distribution and randomness expansion can be achieved, assuming only that the devices prepare and measure qubit systems. Moving to more general networks may allow for more robust and efficient protocols, and other information-theoretic tasks.

ACKNOWLEDGMENTS

This work is supported by FNP programme TEAM and NCN through Grant No. 2014/14/E/ST2/00020, the Swiss National Science Foundation (Grant No. PP00P2_138917 and Starting Grant DIAQ), and SEFRI (COST action MP1006).

APPENDIX A: ALL DIMENSION WITNESSES FOR A SIMPLE NETWORK

Here we present all dimension witnesses for the scenario of Fig. 4(a) with $|x_0| = |x_1| = 3$, $|y| = 1$, and $|b| = 2$. In this scenario, considering classical communication $c_0, c_1 \in \{0,1\}$, there exist 13 nontrivial facets (i.e., facets that do not correspond to the normalization of probabilities). We present the witnesses in tabular form, using the notation

$$\begin{pmatrix} w_{00} & w_{01} & w_{02} \\ w_{10} & w_{11} & w_{12} \\ w_{20} & w_{21} & w_{22} \end{pmatrix} \leqslant C_2 \tag{A1}$$

to describe the witness

$$\sum_{x_0=0}^{2}\sum_{x_1=0}^{2} w_{x_0x_1}\, p(0|x_0,x_1) \leqslant C_2. \tag{A2}$$

The 13 witnesses are

$$(1)\begin{pmatrix} -1 & 1 & 1 \\ -1 & 0 & -1 \\ 1 & 1 & -1 \end{pmatrix} \leqslant 2, \quad (2)\begin{pmatrix} 2 & -1 & 1 \\ 2 & 0 & -2 \\ 0 & -1 & 1 \end{pmatrix} \leqslant 4,$$

$$(3)\begin{pmatrix} 1 & -1 & 1 \\ 1 & -2 & -3 \\ 0 & 2 & -2 \end{pmatrix} \leqslant 2, \quad (4)\begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \leqslant 3,$$

$$(5)\begin{pmatrix} 2 & 2 & 0 \\ 1 & -2 & 0 \\ -1 & 1 & -1 \end{pmatrix} \leqslant 4, \quad (6)\begin{pmatrix} 1 & -2 & 3 \\ 2 & 0 & -2 \\ -1 & 2 & 1 \end{pmatrix} \leqslant 6,$$

$$(7)\begin{pmatrix} 1 & -1 & 2 \\ 1 & 0 & -1 \\ -1 & 1 & 1 \end{pmatrix} \leqslant 4, \quad (8)\begin{pmatrix} 1 & -1 & 2 \\ 2 & 0 & -2 \\ -1 & 1 & 0 \end{pmatrix} \leqslant 4,$$

$$(9)\begin{pmatrix} 2 & -2 & 4 \\ 4 & -1 & -5 \\ -2 & 1 & -1 \end{pmatrix} \leqslant 6, \quad (10)\begin{pmatrix} 1 & -1 & 2 \\ 2 & -3 & -5 \\ -1 & 3 & -3 \end{pmatrix} \leqslant 3,$$

$$(11)\begin{pmatrix} 2 & 2 & -1 \\ 1 & -1 & 0 \\ -1 & 1 & 0 \end{pmatrix} \leqslant 4,$$

$$(12)\begin{pmatrix} 1 & -2 & 3 \\ 3 & 1 & -2 \\ -2 & 3 & 1 \end{pmatrix} \leqslant 8, \quad (13)\begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} \leqslant 2.$$

Note that the last witness, (13), is in fact a lifting from the simplest prepare-and-measure scenario featuring three preparations and two binary measurements. This can be seen by imagining that the first preparation device in our scenario simply acts as a classical input for the measurement device, i.e., x_1 takes the role of y in the prepare-and-measure scenario. Since the channel supports bits, then we must have $y = 0,1$. In the final witness we see that $x_1 = 0,1$ corresponds to $y = 0,1$ and $x_1 = 2$ is never used (since we have all zeros on the bottom row of the witness). Upon interpreting x_1 as y in a prepare-and-measure scenario, the final witness then corresponds to Eq. (6) of [4].

APPENDIX B: QUANTUM VIOLATION IN NONLOCAL DENSE CODING

Here we calculate explicitly the values of (36) for strategies using qubits. We first consider the case where the devices do not share initial entanglement. To ease notation we define

$$|h_+\rangle = \left|\psi\left(\tfrac{\pi}{4},0\right)\right\rangle; \quad |h_-\rangle = \left|\psi\left(-\tfrac{3\pi}{4},0\right)\right\rangle. \tag{B1}$$

Following the preparations and measurements given in the main text, we have

$$\begin{aligned} &p(b_0,b_1|u_0,u_1,v_0,v_1,y) \\ &\quad = |\langle\psi^-|\,\sigma_x^{b_1\oplus u_1}\sigma_z^{b_0\oplus u_0}\otimes H^y\sigma_x^{v_1}\sigma_z^{v_0}\,|h_+\rangle\,|h_-\rangle\,|^2 \\ &\quad = |\langle\psi^-|\,\sigma_x^{b_1\oplus u_1\oplus v_1\bar{y}\oplus v_0y}\sigma_z^{b_0\oplus u_0\oplus v_0\bar{y}\oplus v_1y}\otimes\mathbb{I}\,|h_+\rangle\,|h_-\rangle\,|^2, \end{aligned} \tag{B2}$$

where in the last line we have used

$$H\sigma_x^{v_1}\sigma_z^{v_0} = \sigma_z^{v_0}\sigma_x^{v_1}H \tag{B3}$$

and

$$H\,|h_\pm\rangle = \pm\,|h_\pm\rangle\,. \tag{B4}$$

By writing $|\psi^-\rangle = \frac{1}{\sqrt{2}}(|h_+\rangle\,|h_-\rangle - |h_-\rangle\,|h_+\rangle)$ we see that the probability that $(b_0,b_1) = (u_0 \oplus v_0\bar{y} \oplus v_1 y, u_1 \oplus v_1\bar{y} \oplus v_0 y)$ is given by

$$|\langle\psi^-||h_+\rangle|h_-\rangle|^2 = \tfrac{1}{2}. \tag{B5}$$

The probability that both bits are guessed incorrectly, i.e., $(\bar{b}_0,\bar{b}_1) = (u_0 \oplus v_0\bar{y} \oplus v_1 y, u_1 \oplus v_1\bar{y} \oplus v_0 y)$ is

$$|\langle\psi^-|\sigma_x\sigma_z \otimes \mathbb{I}|h_+\rangle|h_-\rangle|^2 = 0. \tag{B6}$$

Hence, we achieve $W_D = \frac{1}{2}$. In order to treat the case in which the preparation devices share entanglement, we need to replace the state $|h_+\rangle\,|h_-\rangle$ by the singlet state $|\psi^-\rangle$. Hence the probability that $(b_0,b_1) = (u_0 \oplus v_0\bar{y} \oplus v_1 y, u_1 \oplus v_1\bar{y} \oplus v_0 y)$ becomes

$$|\,\langle\psi^-|\psi^-\rangle\,|^2 = 1 \tag{B7}$$

and the game is won perfectly.

[1] N. Brunner, S. Pironio, A. Acín, N. Gisin, A. A. Methot, and V. Scarani, Phys. Rev. Lett. **100**, 210503 (2008).
[2] T. Vértesi and K. F. Pál, Phys. Rev. A **77**, 042106 (2008).
[3] D. Pérez-García, M. M. Wolf, C. Palazuelos, I. Villanueva, and M. Junge, Commun. Math. Phys. **279**, 455 (2008).
[4] R. Gallego, N. Brunner, C. Hadley, and A. Acín, Phys. Rev. Lett. **105**, 230501 (2010).
[5] M. Dall'Arno, E. Passaro, R. Gallego, and A. Acín, Phys. Rev. A **86**, 042312 (2012).
[6] S. Wehner, M. Christandl, and A. C. Doherty, Phys. Rev. A **78**, 062112 (2008).
[7] N. Brunner, M. Navascués, and T. Vértesi, Phys. Rev. Lett. **110**, 150501 (2013).
[8] M. Navascués and T. Vértesi, Phys. Rev. Lett. **115**, 020501 (2015).
[9] J. Bowles, M. T. Quintino, and N. Brunner, Phys. Rev. Lett. **112**, 140407 (2014).
[10] M. Dall'Arno, E. Passaro, R. Gallego, M. Pawłowski, A. Acín, M. Pusey, J. Barrett, and T. Rudolph, Quantum Inf. Comput. **15**, 0037 (2015).
[11] N. Brunner, M. Kaplan, A. Leverrier, and P. Skrzypczyk, New J. Phys. **16**, 123050 (2014).
[12] M. Hendrych, R. Gallego, M. Micuda, N. Brunner, A. Acín, and J. Torres, Nat. Phys. **8**, 588 (2012).
[13] J. Ahrens, P. Badziąg, A. Cabello, and M. Bourennane, Nat. Phys. **8**, 592 (2012).
[14] E. F. Galvão and L. Hardy, Phys. Rev. Lett. **90**, 087902 (2003).
[15] N. Harrigan, T. Rudolph, and S. Aaronson, arXiv:0709.1149.
[16] H. Buhrman, R. Cleve, S. Massar, and R. de Wolf, Rev. Mod. Phys. **82**, 665 (2010).
[17] M. Pawłowski and N. Brunner, Phys. Rev. A **84**, 010302 (2011).
[18] E. Woodhead, Phys. Rev. A **88**, 012331 (2013).
[19] E. Woodhead, C. W. Lim, and S. Pironio, Lect. Notes Comput. Sci. **7582**, 107 (2013).
[20] 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).
[21] H.-W. Li, M. Pawłowski, Z.-Q. Yin, G.-C. Guo, and Z.-F. Han, Phys. Rev. A **85**, 052308 (2012).
[22] Y. C. Liang, T. Vértesi, and N. Brunner, Phys. Rev. A **83**, 022108 (2011).
[23] T. Vértesi and M. Navascués, Phys. Rev. A **83**, 062112 (2011).
[24] T. Lunghi, J. B. Brask, Charles Ci Wen Lim, Q. Lavigne, J. Bowles, A. Martin, H. Zbinden, and N. Brunner, Phys. Rev. Lett. **114**, 150501 (2015).
[25] G. Cañas, J. Carine, E. S. Gómez, J. F. Barra, A. Cabello, G. B. Xavier, G. Lima, and M. Pawłowski, arXiv:1410.3443.
[26] A. Bennet, T. Vértesi, D. J. Saunders, N. Brunner, and G. J. Pryde, Phys. Rev. Lett. **113**, 080405 (2014).
[27] The case of two-way communication could also be considered, but we will not discuss it here.
[28] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Rev. Mod. Phys. **86**, 419 (2014).
[29] See http://www.iwr.uni-heidelberg.de/groups/comopt/.
[30] See http://cgm.cs.mcgill.ca/~avis/C/lrs.html.
[31] Indeed, more general transformations, which increase the Hilbert space dimension, could be considered.
[32] J. L. Krivine, Adv. Math. **31**, 16 (1979).
[33] Note that only upper and lower bounds are known for k_3; see, e.g., T. Vértesi, Phys. Rev. A **78**, 032112 (2008).
[34] B. F. Toner and D. Bacon, Phys. Rev. Lett. **91**, 187904 (2003).
[35] A. Acín, N. Gisin, and B. Toner, Phys. Rev. A **73**, 062105 (2006).
[36] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).
[37] H. M. Wiseman, S. J. Jones, and A. C. Doherty, Phys. Rev. Lett. **98**, 140402 (2007).
[38] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf, Phys. Rev. Lett. **87**, 167902 (2001).
[39] R. Rabelo, M. Ho, D. Cavalcanti, N. Brunner, and V. Scarani, Phys. Rev. Lett. **107**, 050502 (2011).
[40] M. F. Pusey, J. Barrett, and T. Rudolph, Nat. Phys. **8**, 475 (2012).
[41] For the full list of inequalities, contact joseph.bowles@unige.ch.
[42] A. Ambainis, A. Nayak, A. Ta-Shma, and U. Vazirani, J. ACM **49**, 496 (2002).
[43] A. Ambainis, D. Leung, L. Mancinska, and M. Ozols, arXiv:0810.2937.
[44] M. Navascués, G. de la Torre, and T. Vértesi, Phys. Rev. X **4**, 011011 (2014).
[45] S. Wiesner, SIGACT News **15**, 78 (1983).

Paper H

Certifying the Dimension of Classical and Quantum Systems in a Prepare-and-Measure Scenario with Independent Devices

Physical Review Letters **112**, 140407 (2014)

Joseph Bowles, Marco Túlio Quintino, and Nicolas Brunner

Certifying the Dimension of Classical and Quantum Systems in a Prepare-and-Measure Scenario with Independent Devices

Joseph Bowles,[1] Marco Túlio Quintino,[1] and Nicolas Brunner[1,2]
[1]*Département de Physique Théorique, Université de Genève, 1211 Genève, Switzerland*
[2]*H.H. Wills Physics Laboratory, University of Bristol, Bristol BS8 1TL, United Kingdom*
(Received 11 November 2013; published 11 April 2014)

We consider the problem of testing the dimension of uncharacterized classical and quantum systems in a prepare-and-measure setup. Here we assume the preparation and measurement devices to be independent, thereby making the problem nonconvex. We present a simple method for generating nonlinear dimension witnesses for systems of arbitrary dimension. The simplest of our witnesses is highly robust to technical imperfections, and can certify the use of qubits in the presence of arbitrary noise and arbitrarily low detection efficiency. Finally, we show that this witness can be used to certify the presence of randomness, suggesting applications in quantum information processing.

DOI: 10.1103/PhysRevLett.112.140407 PACS numbers: 03.65.Ta, 03.67.-a

The problem of estimating the dimension of uncharacterized physical systems has recently attracted attention. From a fundamental point of view, this problem is well motivated, as it shows that dimension—the number of (relevant) degrees of freedom—of an unknown system can be determined in a device-independent way. That is, dimension can be tested from measurement data alone, in a scenario in which all devices used in the experiment, including the measurement device, are uncharacterized; i.e., no assumption about the internal working of the devices is needed. Beyond the fundamental interest, this problem is also relevant in the context of quantum information, where the dimension of quantum systems—i.e., the Hilbert space dimension—represents a resource for performing information-theoretic tasks. Specifically, higher dimensional quantum systems can increase the performance of certain protocols, and/or simplify their implementation.

First approaches to this problem considered Bell inequality tests [1–6], random access codes [7], and monitoring of an observable of a dynamic system [8]. More recently, a general formalism was developed to estimate the dimension of classical and quantum systems in a prepare-and-measure setup [9], the simplest but also the most general scenario. Consider two uncharacterized devices, hence described as black boxes (see Fig. 1). The first device prepares upon request a physical system in an unknown state ρ_x. A second device then performs a measurement on the system. The observer tests the devices, by choosing a preparation x and a measurement y, then receiving measurement outcome b. Repeating the experiment many times, the observer obtains the probability distribution $p(b|x,y)$, called here the data. The goal for the observer is then to give a lower bound on the dimension of the unknown set of states $\{\rho_x\}$ from the data alone. This can be achieved using "dimension witnesses" [9–11] (see also Refs. [12,13] for different approaches). These ideas were shown to be relevant experimentally [14,15], and for quantum information processing [16,17].

Here we discuss this problem assuming the preparation and measurement devices to be independent. This assumption is rather natural in a device-independent estimation scenario, where devices are uncharacterized but do not conspire maliciously against the observer. The main difficulty of this problem is that it is nonconvex, a feature that makes generic problems with independent variables hard to tackle. Note that previous works on dimension witnesses allowed the devices to be correlated via shared randomness (hence relaxing the independence assumption), making the problem convex. Although these techniques can in principle be applied in our case, they are far from optimal, as we shall see below.

It is therefore desirable to develop novel methods, which is the goal of this work. Specifically, we present a simple technique for deriving nonlinear dimension witnesses, tailored for device-independent tests of dimension assuming independent devices. We derive witnesses for systems of arbitrary dimension, obtaining a quadratic gap between classical and quantum dimensions. The simplest witness is discussed in detail. We show that it is extremely robust to technical imperfections, and can be used to certify the presence of randomness.

Scenario.—We consider the setup of Fig. 1. The experiment is characterized by the set of conditional probabilities

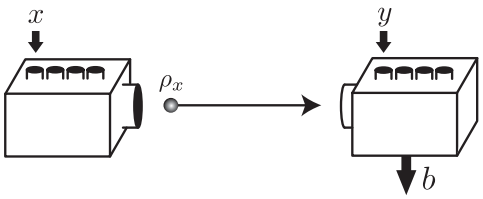

FIG. 1. Prepare-and-measure setup.

$p(b|x,y)$ (i.e., the data) which gives the probability of obtaining outcome b when performing measurement y on preparation x.

Consider first the case of quantum systems. We say that the experiment admits a d-dimensional quantum representation when there exist states ρ_x and measurement operators $M_{b|y}$ both acting on $\mathbb{C}^d$, such that

$$p(b|x,y) = \mathrm{Tr}(\rho_x M_{b|y}). \tag{1}$$

Next consider the situation of classical systems of dimension d. Given the choice of preparation x, the first device sends a classical message $m = 0, \ldots, d-1$. Note that the device may have an internal source of randomness (represented by a random variable λ_1). Hence, which message m is sent depends on both x and λ_1. The measurement device, upon receiving message m, and input y from the observer, delivers an outcome b. As it also features a source of randomness (random variable λ_2), the output b depends on m, y, and λ_2. The behavior observed in the experiment is then given by

$$p(b|x,y) = \int d\lambda_1 d\lambda_2 \rho(\lambda_1,\lambda_2) \sum_{m=0}^{d-1} p(m|x,\lambda_1) p(b|m,y,\lambda_2).$$

The main point now is to consider the joint distribution of random variables $\lambda_{1,2}$. If $\rho(\lambda_1,\lambda_2) \neq \rho_1(\lambda_1)\rho_2(\lambda_2)$, the variables are correlated; hence, the devices may follow a (preestablished) correlated strategy. Previous works focused on this situation. As the set of behaviors of the above form is convex, it can be fully characterized with linear dimension witnesses [9].

Here we consider the situation in which the devices are independent, i.e., $\rho(\lambda_1,\lambda_2) = \rho_1(\lambda_1)\rho_2(\lambda_2)$. That is, although each device features an internal source of randomness, the devices have no shared randomness. In this case, the observed statistics can be written as

$$p(b|x,y) = \sum_{m=0}^{d-1} s(m|x) t(b|m,y) \tag{2}$$

where $s(m|x) = \int d\lambda_1 \rho_1(\lambda_1) p(m|x,\lambda_1)$ is the distribution of possible messages m for each preparation x, and $t(b|m,y) = \int d\lambda_2 \rho_2(\lambda_2) p(b|m,y,\lambda_2)$ is the distribution of outcomes b for measurement y when receiving message m. Below we will see how to characterize the set of behaviors of the form Eq. (2). This will require nonlinear dimension witnesses as the set is nonconvex.

Determinant witness.—In this work we focus on experiments with binary outcomes, denoted $b = 0, 1$. We will construct nonlinear witnesses based on the determinant of a matrix. We first discuss the simplest case, with four preparations $x = 0, \ldots, 3$ and two measurements $y = 0, 1$. Consider the following matrix

$$\mathbf{W}_2 = \begin{pmatrix} p(0,0) - p(1,0) & p(2,0) - p(3,0) \\ p(0,1) - p(1,1) & p(2,1) - p(3,1) \end{pmatrix} \tag{3}$$

where we write $p(x,y) = p(b=0|x,y)$ for simplicity. For any strategy involving a classical bit [i.e., its statistics admits a decomposition of the form Eq. (2) with $d = 2$], one has that

$$W_2 = \det(\mathbf{W}_2) = 0. \tag{4}$$

The proof is straightforward. Note that for any statistics of the form Eq. (2) with $d = 2$, we have that $p(x,y) = s(0|x)[t(0|0,y) - t(0|1,y)] + t(0|1,y)$. Hence we write

$$\begin{aligned} p(x,y) - p(x',y) &= [s(0|x) - s(0|x')][t(0|0,y) - t(0|1,y)] \\ &= S_{xx'} T_y \end{aligned} \tag{5}$$

from which it follows that

$$W_2 = \begin{vmatrix} S_{01}T_0 & S_{23}T_0 \\ S_{01}T_1 & S_{23}T_1 \end{vmatrix} = 0. \tag{6}$$

An interesting feature of the above witness is that it is given by an equality, whereas linear witnesses are given by inequalities [9]. Moreover, our witness turns out to characterize fully the set of experiments involving a classical bit. Specifically, for any experiment achieving $W_2 = 0$ (for all relabelings of the preparation x), there exists a decomposition of the form Eq. (2) with $d = 2$ (see Supplemental Material [18]). Note that if the preparation and measurement devices are correlated, then classical bit strategies can reach $W_2 = 1$. Consider for instance the equal mixture of the two following deterministic strategies: (i) $s(0|x) = 1$ iff $x = 0, 3$ and $t(0|m,y) = m + y \bmod 2$, (ii) $s(0|x) = 1$ iff $x = 0, 2$ and $t(0|m,y) = m$. Hence we get $\mathbf{W}_2 = \mathbb{I}_2$ and $W_2 = 1$. This shows that our witness is tailored for the case in which the devices are independent.

Next we investigate the performance of qubit strategies, i.e., statistics of the form Eq. (1) with $d = 2$. States are given by density matrices $\rho_x = (\mathbb{I}_2 + \vec{s}_x \cdot \vec{\sigma})/2$ and measurement operators by $M_{0|y} = c_y \mathbb{I}_2 + \vec{T}_y \cdot \vec{\sigma}/2$, where $\vec{s}_x$ and $\vec{T}_y$ are Bloch vectors and $|c_y| \leq 1$ [19]. Similarly to above, we write

$$p(x,y) - p(x',y) = \mathrm{Tr}[(\rho_x - \rho_{x'}) M_{0|y}] = \vec{S}_{xx'} \cdot \vec{T}_y \tag{7}$$

where $\vec{S}_{xx'} = (\vec{s}_x - \vec{s}_{x'})/2$. Finally, we get

$$W_2 = \begin{vmatrix} \vec{S}_{01} \cdot \vec{T}_0 & \vec{S}_{23} \cdot \vec{T}_0 \\ \vec{S}_{01} \cdot \vec{T}_1 & \vec{S}_{23} \cdot \vec{T}_1 \end{vmatrix} = (\vec{S}_{01} \times \vec{S}_{23}) \cdot (\vec{T}_0 \times \vec{T}_1) \leq 1 \tag{8}$$

since $|\vec{S}_{01} \times \vec{S}_{23}| \leq 1$ and $|\vec{T}_0 \times \vec{T}_1| \leq 1$. This bound for qubit strategies is tight, and can be reached as follows:

choose the preparations to be the pure qubit states given by $\vec{s}_0 = -\vec{s}_1 = \hat{z}$, $\vec{s}_2 = -\vec{s}_3 = \hat{x}$, and the measurements by the vectors $\vec{T}_0 = \cos\theta\hat{z} + \sin\theta\hat{x}$ and $\vec{T}_1 = \sin\theta\hat{z} - \cos\theta\hat{x}$. Notice that we are free to choose any angle θ here, due to the rotational invariance of the cross product in the plane. For $\theta = 0$ we get the usual BB84 states and measurements.

It is relevant to note that essentially any qubit strategy achieves $|W_2| > 0$. Only very specific alignments of the qubit preparations and measurements (a set of measure zero) achieve $W_2 = 0$. Therefore, a generic qubit strategy always outperforms the most general strategy involving a bit.

This suggests that our witness is well suited for distinguishing data involving classical bits and qubits. To illustrate the robustness of our witness, we investigate the effect of technical imperfections, such as background noise and limited detection efficiency (of the detector inside the measurement device), on a generic qubit strategy given by the data $p_Q(x, y)$ achieving $|W_2| = Q > 0$. Say that an error occurs with probability $1 - \eta$, for instance the emitted particle is lost. Hence the observed statistics is given by

$$p(x, y) = \eta p_Q(x, y) + (1 - \eta)p_N(y), \tag{9}$$

where we consider a noise model of the form $p_N(x, y) = p_N(y)$; i.e., the noise is independent of the choice of preparation x. The difference in probabilities entering the witness is then independent of the noise term: $p(x, y) - p(x', y) = \eta[p_Q(x, y) - p_Q(x', y)]$, and thus the observed value of the witness is $W_2 = \eta^2 Q$, which is strictly positive whenever $Q > 0$. Hence, for an arbitrary amount of background noise and/or an arbitrarily low efficiency, a generic qubit strategy will outperform any classical bit strategy; see Ref. [20] for a related result. This is indeed in stark contrast with previous witnesses, which can only tolerate a finite amount of noise and require a high efficiency [11].

Finally, we comment on strategies involving higher dimensional systems. Using a classical trit one achieves $|W_2| \leq 1$ [21], while numerical analysis shows that $|W_2| \leq 1.299$ for qutrit strategies. This shows that the value of W_2 is useful to assess dimension. To reach the algebraic maximum of $W_2 = 2$, systems of dimension (at least) $d = 4$ (either classical or quantum) are required.

Determinant witness for all dimensions.—We now generalize the above witness for testing classical and quantum systems of arbitrary dimension. Consider a scenario with $2k$ preparations and k binary measurements. Construct the $k \times k$ matrix

$$\mathbf{W}_k(i, j) = p(2j, i) - p(2j + 1, i) \tag{10}$$

with $0 \leq i, j \leq k - 1$. As above, the witness is given by $W_k = |\det(\mathbf{W}_k)|$. We will see that, for classical systems of dimension d, one has that

$$W_k = 0 \quad \text{for } d \leq k, \tag{11}$$

while one can have $W_k \geq 1$ for $d > k$. For quantum systems of dimension d, we get

$$W_k = 0 \quad \text{for } d \leq \sqrt{k}, \tag{12}$$

while $W_k > 0$ is possible whenever $d > \sqrt{k}$. Hence we obtain a quadratic separation between classical and quantum dimensions, using a number of preparations and measurements that grows only linearly.

To prove the above claims, it is enough to focus on quantum strategies. Consider matrices of the form

$$\rho_x = \frac{1}{d}(\mathbb{I}_d + \phi_d \vec{s}_x \cdot \vec{\lambda}), \tag{13}$$

with $\vec{s}_x \in \mathbb{R}^{d^2-1}$, $|\vec{s}_x| \leq 1$, $\vec{\lambda}$ the vector of the $d^2 - 1$ Gell-Mann matrices (generalized Pauli matrices, satisfying $\mathrm{tr}(\lambda_i) = 0$ and $\mathrm{tr}(\lambda_i\lambda_j) = 2\delta_{ij}$) and $\phi_d = \sqrt{[d(d-1)]/2}$. While all matrices of the above form are valid quantum density matrices for $|\vec{s}_x| \leq 2/d$ [22], this is not the case in general (although this will not affect our argument). Similarly we write measurement operators as $M_{0|y} = c_y\mathbb{I}_d + \phi_d\vec{T}_y \cdot \vec{\lambda}/d$ with $\vec{T}_y \in \mathbb{R}^{d^2-1}$, $|\vec{T}_y|, |c_y| \leq 1$[19], and get that

$$\mathbf{W}_k(i, j) = \mathrm{Tr}[(\rho_{2j} - \rho_{2j+1})M_{0|i}] = \vec{S}_j \cdot \vec{T}_i \tag{14}$$

with $\vec{S}_j = (1 - (1/d))(\vec{s}_{2j} - \vec{s}_{2j+1})$. Thus, as before, the entries of the matrix $\mathbf{W}_k$ are given by scalar products of vectors. Similarly to the qubit construction of Eq. (8), the witness W_k can be expressed using cross products, generalized here to arbitrary dimensions.

Specifically, the cross product $\vec{S}_0 \times \vec{S}_1 \times \cdots \times \vec{S}_{k-1}$ of k vectors in $\mathbb{R}^{k+1}$ is defined as the unique vector $\vec{u} \in \mathbb{R}^{k+1}$ such that $\vec{V} \cdot \vec{u} = \det(\vec{S}_0, \vec{S}_1, \cdots, \vec{S}_{k-1})$ for all $\vec{V} \in \mathbb{R}^{k+1}$ (see, e.g., Ref. [23]). It follows that $\vec{S}_0 \times \cdots \times \vec{S}_{k-1} = 0$ iff $\vec{S}_0, \cdots, \vec{S}_{k-1}$ are linearly dependent. Furthermore, similarly to Eq. (8), we have that

$$\begin{aligned} W_k &= |\det(\mathbf{W}_k)| \\ &= |(\vec{S}_0 \times \cdots \times \vec{S}_{k-1}) \cdot (\vec{T}_0 \times \cdots \times \vec{T}_{k-1})|. \end{aligned}$$

To conclude, we relate the dimension of the quantum systems to the linear (in)dependence of the set of vectors $\vec{S}_j$ and $\vec{T}_i$. Note that we must ensure here that the vectors $\vec{S}_j$, $\vec{T}_i$ are in $\mathbb{R}^{k+1}$, via an embedding or by using only a restricted set of parameters. As d-dimensional quantum systems have $d^2 - 1$ parameters, we see that the vectors $\vec{S}_j$ (and similarly for $\vec{T}_i$) can span a subspace of dimension at most $d^2 - 1$.

Hence, if $d \leq \sqrt{k}$, the vectors $\vec{S}_j$ cannot be linearly independent, and we get $W_k = 0$. On the contrary if $d > \sqrt{k}$, the vectors $\vec{S}_j$ and $\vec{T}_i$ can be chosen to be linearly independent, and we have $W_k > 0$. Take for instance $\vec{S}_j$ to be parallel to $\vec{T}_i$, and $|\vec{s}_j|$, $|\vec{T}_i| \leq 2/d$ ensuring that all preparations and measurements are represented by valid operators. Note, however, that this construction is suboptimal in general, as one can obtain $W_k = 1$ with quantum states of dimension $d > \sqrt{k}$ (with d an integer prime power), using a mutually unbiased basis (see Supplemental Material [18]).

The proof for classical systems can be derived by noting that any classical strategy using d-dimensional states can be recast as a quantum strategy using diagonal density matrices acting on $\mathbb{C}^d$. Since we have only $d-1$ parameters in this case, it follows from the above that $W_k = 0$ when $d \leq k$. For $d > k$, one can get $W_k \geq 1$. The lower bound is obtained by considering the following strategy: if x is even, then send $m = x/2$, else send $m = d$; for the measurement device, output $b = 0$ iff $y = m$. Note that for this strategy, we get $\mathbf{W}_k = \mathbb{I}_k$, hence $W_k = 1$. An interesting question is to find the algebraic maximum of W_k, and the minimal dimension for classical and quantum systems required to attain it. Note that this problem is related to that of finding the determinant of a Hadamard matrix. Hence we get the bound $W_k \leq k^{k/2}$, which is tight iff there exists a Hadamard matrix of size $k \times k$.

Certifying randomness.—The fact that the determinant witness can distinguish between classical and quantum systems (given a bound on the dimension) suggests applications in randomness certification. Here we investigate the connection between the amount of violation of the witness W_2 and the intrinsic randomness of the of the underlying statistics, assuming that the preparation device emits qubit states.

Consider the quantity

$$\bar{p} = \frac{1}{4} \sum_{x,y=0,1} \max_b p(b|x,y), \tag{15}$$

i.e., the average guessing probability of the outcome b for preparations $x = 0, 1$. Randomness can be quantified by the min-entropy of $\bar{p}$, i.e., $H_{\min}(\bar{p}) = -\log_2(\bar{p})$, which gives the number of random bits extractable from the experiment (per run). Now for a given amount of violation of the witness $W_2 = Q > 0$, we want to find out the maximal value of $\bar{p}$ over all qubit strategies which are compatible with the value $W_2 = Q > 0$. In other words, what is the minimal amount of randomness compatible with a certain violation of the witness? To answer this question, we solve numerically the following problem. We maximize $\bar{p}$ subject to the constraints: $W_2 = Q$, $p(b|x,y) = \mathrm{Tr}(\rho_x M_{b|y})$ where ρ_x, $M_{b|y}$ are arbitrary qubit states and measurement operators.

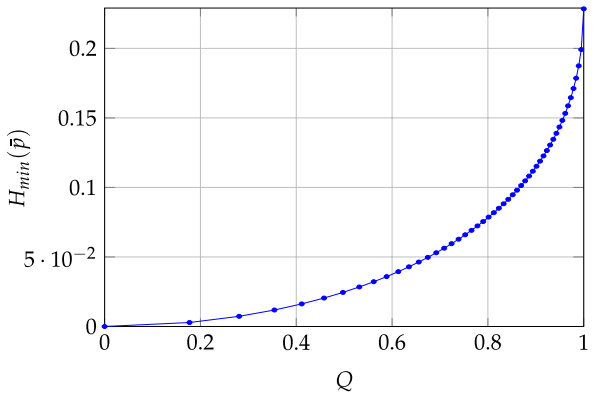

FIG. 2 (color online). Average certifiable randomness $H_{\min}(\bar{p})$ using the witness W_2. For any amount of violation of the witness $W_2 = Q > 0$, randomness can be certified.

In Figure 2, we plot the amount of randomness $H_{\min}(\bar{p})$ as a function of the value Q of the witness W_2. We see that for any amount of violation, randomness can be certified. In other words, from the sole knowledge of the value of W_2, one can upper bound the probability of correctly guessing the output b, for any observer knowing the detailed qubit strategy that is being used. Importantly, the quantity $H_{\min}(\bar{p})$ captures here the intrinsic quantum randomness of the experiment, but is independent of any randomness generated locally in the devices (used, e.g., to create mixed state preparations). These issues will be discussed in detail in a forthcoming work [24], where a protocol for randomness certification will be presented.

Discussion.—We have presented a method for testing the dimension of classical and quantum systems of arbitrary dimension. Moreover, the simplest of our witnesses is highly robust to noise and can be used to certify randomness without the need of high visibilities and efficiencies. Hence we believe these ideas are relevant in practice. In this perspective, it will be necessary to make a statistical analysis in the spirit of Refs. [25] for taking finite size effects into account [24]. Finally, from a more abstract point of view, the ideas presented here could be useful in other nonconvex problems involving independent variables, such as Bell tests with independent sources [26,27], and more general marginal problems [28].

We thank J. B. Brask, Ci Wen Lim, M. Pawłowski, S. Pironio, E. Woodhead, and H. Zbinden for discussions, and acknowledge financial support from the Swiss National Science Foundation (Grant No. PP00P2_138917) and the EU DIQIP.

[1] N. Brunner, S. Pironio, A. Acín, N. Gisin, A. A. Méthot, and V. Scarani, Phys. Rev. Lett. **100**, 210503 (2008).
[2] K. F. Pál and T. Vértesi, Phys. Rev. A **77**, 042105 (2008).

Certifying the dimension of classical and quantum systems in a prepare-and-measure scenario with independent devices — Supplementary material

Joseph Bowles,[1] Marco Túlio Quintino,[1] and Nicolas Brunner[1,2]

[1]*Département de Physique Théorique, Université de Genève, 1211 Genève, Switzerland*
[2]*H.H. Wills Physics Laboratory, University of Bristol, Bristol, BS8 1TL, United Kingdom*

A. Full characterisation of the set of classical bit strategies with the witness W_2— In the main text we showed that $W_2 = 0$ for strategies involving a classical bit. Here we will see that the converse holds. That is, any statistics achieving $W_2 = 0$ for all possible relabelings of the preparation label x, can be realised with a classical bit strategy.

Consider the matrix $\mathbf{W}_2$, here rewritten as

$$\mathbf{W}_2 = \begin{pmatrix} a_1 - a_2 & b_1 - b_2 \\ c_1 - c_2 & d_1 - d_2 \end{pmatrix} \tag{1}$$

with $a_1 = p(0,0)$, $a_2 = p(1,0)$ and so on. Without loss of generality, we take $d_2 \geq d_1 \geq c_2 \geq c_1$ which can be achieved via a relabelling of the preparations x. The conditions $W_2 = 0$, considering all relabelings of x, are then given by

$$\begin{aligned} (a_1 - a_2)(d_1 - d_2) &= (b_1 - b_2)(c_1 - c_2) \\ (a_1 - b_2)(d_1 - c_2) &= (b_1 - a_2)(c_1 - d_2) \\ (a_1 - b_1)(c_2 - d_2) &= (a_2 - b_2)(c_1 - d_1) \end{aligned} \tag{2}$$

where the second and third equations correspond to relabelling x according to $1 \leftrightarrow 3$ and $1 \leftrightarrow 2$ respectively. To show that there exists a decomposition of the form (**??**) with $d = 2$, we solve for $\{s(m|x)\}, \{t(0|m,y\}$ (with $m = 0,1$, $x = 0,...,3$, and $y = 0,1$) the set of equations given by the conditions (2) and the 8 conditions given by $p(x,y) = \sum_{m=0,1} s(m|x)t(0|m,y)$. A solution is given by $t(0|0,0) = b_1$, $t(0|0,1) = d_1$, $t(0|1,0) = b_2$, $t(0|1,1) = d_2$, $s(0|0) = (d_2 - c_1)/(d_2 - d_1)$, $s(0|1) = (d_2 - c_2)/(d_2 - d_1)$, $s(0|2) = 1$ and $s(0|3) = 0$. Hence any matrix (1) satisfying conditions (2) admits a decomposition of the form (**??**) with $d = 2$. Thus the determinant witness characterises fully the set of distributions obtained from strategies involving a classical bit.

B. Quantum strategy using mutually unbiased bases— Consider a quantum system of dimension d, for which we have $n \leq d + 1$ mutually unbiased bases (MUBs) denoted by $\mathcal{M}_\alpha = \{|\psi_{i|\alpha}\rangle\}$, where $\alpha = 0, \cdots, n-1$ and $i = 0, \cdots, d-1$. Due to the properties of MUBs the projectors $\pi_{i|\alpha} = |\psi_{i|\alpha}\rangle\langle\psi_{i|\alpha}|$ satisfy $\mathrm{tr}(\pi_{i|\alpha}\pi_{j|\alpha}) = \delta_{ij}$ and $\mathrm{tr}(\pi_{i|\alpha}\pi_{j|\beta}) = 1/d$ for $\alpha \neq \beta$. The main idea now will be to construct a quantum strategy for which we get $\mathbf{W}_k = \mathbb{I}_k$ and so $W_k = 1$.

Consider first the upper left block of $\mathbf{W}_k$ of size $d-1 \times d-1$. Concentrating on the first basis $\mathcal{M}_0$, we choose the preparations as $\rho_{2j} = \pi_{j|0}$ and $\rho_{2j+1} = \pi_{d-1|0}$, (with $j = 0, \cdots, d-2$) and measurement projectors as $M_{0|i} = \pi_{i|0}$, (where $i = 0, \cdots, d-2$). Hence for this block we have that

$$p(2j,i) - p(2j+1,i) = \mathrm{tr}([\pi_{j|0} - \pi_{d-1|0}]\pi_{i|0}) = \delta_{ij} \tag{3}$$

since $\mathcal{M}_0$ is an orthonormal basis, and so the first $d-1 \times d-1$ block of $\mathbf{W}_k$ is the identity matrix $\mathbb{I}_{d-1}$.

We then move on to the next $d-1$ preparations and measurements, keeping the same pattern but using the next basis $\mathcal{M}_1$. That is, we choose $\rho_{2j+2(d-1)} = \pi_{j|1}$, $\rho_{2j+1+2(d-1)} = \pi_{d-1|1}$ and $M_{0|i+d-1} = \pi_{i|1}$ $(i,j = 0, \cdots, d-2)$. We continue this pattern until we have used up all n MUBs. This will give us a $(d-1)n \times (d-1)n$ matrix. Via the same argument as above, each $d-1 \times d-1$ block on the diagonal of $\mathbf{W}_k$ will be equal to $\mathbb{I}_{d-1}$. All off-diagonal blocks are the zero matrix since they contain preparations and measurements that belong to different MUBs. Indeed we have that $\mathrm{tr}([\pi_{i|\alpha} - \pi_{j|\alpha}]\pi_{k|\beta}) = 0$ when $\alpha \neq \beta$.

Hence when $(d-1)n \geq k$, we get that $\mathbf{W}_k = \mathbb{I}_k$, hence $W_k = 1$. If the Hilbert space dimension d is an integer power of a prime, then there exist $d+1$ MUBs [W. Wooters, B. Fields, Annals of Phys. **191**, 363 (1989)]. In this case, one has that $W_k = 1$ for $d \geq \sqrt{k+1}$, or equivalently $d > \sqrt{k}$.

Paper I

Self-Testing Quantum Random Number Generator

Physical Review Letters **114**, 150501 (2015)

Tommaso Lunghi, Jonatan Bohr Brask, Charles Ci Wen Lim, Quentin Lavigne, Joseph Bowles, Anthony Martin, Hugo Zbinden, and Nicolas Brunner

Self-Testing Quantum Random Number Generator

Tommaso Lunghi,[1] Jonatan Bohr Brask,[2] Charles Ci Wen Lim,[1] Quentin Lavigne,[1] Joseph Bowles,[2] Anthony Martin,[1] Hugo Zbinden,[1] and Nicolas Brunner[2]
[1]*Group of Applied Physics, Université de Genève, 1211 Genève, Switzerland*
[2]*Département de Physique Théorique, Université de Genève, 1211 Genève, Switzerland*
(Received 14 November 2014; published 15 April 2015)

The generation of random numbers is a task of paramount importance in modern science. A central problem for both classical and quantum randomness generation is to estimate the entropy of the data generated by a given device. Here we present a protocol for self-testing quantum random number generation, in which the user can monitor the entropy in real time. Based on a few general assumptions, our protocol guarantees continuous generation of high quality randomness, without the need for a detailed characterization of the devices. Using a fully optical setup, we implement our protocol and illustrate its self-testing capacity. Our work thus provides a practical approach to quantum randomness generation in a scenario of trusted but error-prone devices.

DOI: 10.1103/PhysRevLett.114.150501 PACS numbers: 03.67.Ac, 42.50.Ex

Given the importance of randomness in modern science and beyond, e.g., for simulation algorithms and for cryptography, an intense research effort has been devoted to the problem of extracting randomness from quantum systems. Devices for quantum random number generation (QRNG) are now commercially available. All of these schemes work essentially according to the same principle, exploiting the randomness of quantum measurements. A simple realization consists in sending a single photon on a 50/50 beam splitter and detecting the output path [1–3]. Other designs were developed, based on measuring the arrival time of single photons [4–7], the phase noise of a laser [8–10], vacuum fluctuations [11,12], and even mobile phone cameras [13].

A central issue in randomness generation is the problem of estimating the entropy of the bits that are generated by a device, i.e., how random is the raw output data. When a good estimate is available, appropriate postprocessing can be applied to extract true random bits from the raw data (via a classical procedure termed randomness extractor [14]). However, poor entropy estimation is one of the main weaknesses of classical RNG [15], and can have important consequences. In the context of QRNG, entropy estimates for specific setups were recently provided using sophisticated theoretical models [16,17]. Nevertheless, this approach has several drawbacks. First, these techniques are relatively cumbersome, requiring estimates for numerous experimental parameters which may be difficult to precisely assess in practice. Second, each study applies to a specific experimental setup, and cannot be used for other implementations. Finally, it offers no real-time monitoring of the quality of the RNG process, hence no protection against unnoticed misalignment (or even failures) of the experimental setup.

It is therefore highly desirable to design QRNG techniques which can provide a real-time estimate of the output entropy. An elegant solution is provided by the concept of device-independent QRNG [18,19], where randomness can be certified and quantified without relying on a detailed knowledge of the functioning of the devices used in the protocol. Nevertheless, the practical implementation of such protocols is extremely challenging as it requires the genuine violation of Bell's inequality [19,20]. Alternative approaches were proposed [21] but their experimental implementation suffers from loopholes [22]. More recently, an approach based on the uncertainty principle was proposed but requires a fully characterized measurement device [23].

Here, we present a simple and practical protocol for self-testing QRNG. Based on a prepare-and-measure setup, our protocol provides a continuous estimate of the output entropy. Our approach requires only a few general assumptions about the devices (such as quantum systems of bounded dimension) without relying on a detailed model of their functioning. This setting is relevant to real-world implementations of randomness generation, and is well adapted to a scenario of trusted but error-prone providers, i.e., a setting where the devices used in the protocol are not actively designed to fool the user, but where implementation may be imperfect. The key idea behind our protocol is to certify randomness from a pair of incompatible quantum measurements. As the incompatibility of the measurements can be directly quantified from experimental data, our protocol is self-testing. That is, the amount genuine quantum randomness can be quantified directly from the data, and can be separated from other sources of randomness such as fluctuations due to technical imperfections. We implemented this scheme with standard technology, using a single photon source and fibered telecommunication components. We implement the complete QRNG protocol, achieving a rate 23 certified random bits per second, with 99% confidence.

Protocol.—Our protocol, sketched in Fig. 1, uses two devices which, respectively, prepare and measure an

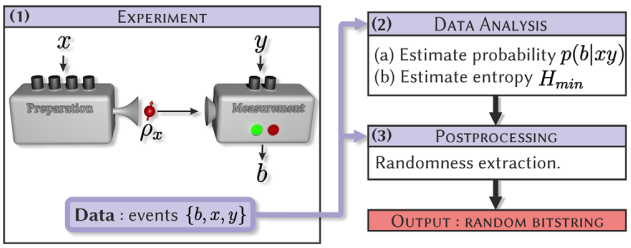

FIG. 1 (color online). Sketch of the protocol. The self-testing QRNG protocol consists of three distinct steps. (1) First, an experiment is performed where, in each round, the user chooses a preparation x and a measurement y, and obtains an outcome b. (2) From the raw data, the distribution $p(b|x, y)$ can be estimated leading to an estimate for the value of the witness W, from which the entropy of the raw data can be quantified. (3) Based on the entropy bound, appropriate postprocessing of the raw data is performed, in order to extract the final random bit string.

uncharacterized qubit system. In each round of the protocol, the observer chooses settings among four possible preparations, $x = 0, 1, 2, 3$, and two measurements $y = 0, 1$, resulting in a binary outcome $b = \pm 1$. To model imperfections, we represent the internal state of each device by a random variable—λ for the preparation device and μ for the measurement device—which are unknown to the observer. As we work in a scenario where the devices are not maliciously conspiring against the user, we assume the devices to be independent, i.e., $p(\lambda, \mu) = q(\lambda)r(\mu)$, where $\int d\lambda q(\lambda) = \int d\mu r(\mu) = 1$.

In each round of the experiment, the preparation device emits a qubit state ρ_x^λ which depends on the setting x and on the internal state λ. Similarly, the measurement device performs a measurement M_y^μ. Thus the distributions of λ and μ determine the distributions of the prepared states and the measurements. As the observer has no access to the variables λ and μ, he will observe

$$p(b|x,y) = \int d\lambda q(\lambda) \int d\mu r(\mu) p(b|x,y,\lambda,\mu) = \mathrm{Tr}\left(\rho_x \frac{\mathbb{1} + bM_y}{2}\right) = \frac{1}{2}(1 + b\vec{S}_x \cdot \vec{T}_y), \quad (1)$$

where

$$\rho_x = \int d\lambda q(\lambda) \rho_x^\lambda = \frac{1}{2}(\mathbb{1} + \vec{S}_x \cdot \vec{\sigma}), \quad (2)$$

$$M_y = \int d\mu r(\mu) M_y^\mu = \vec{T}_y \cdot \vec{\sigma}. \quad (3)$$

Here, $\vec{S}_x$ and $\vec{T}_y$ denote the Bloch vectors of the (average) states and measurements, and $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ is the vector of Pauli matrices.

The task of the observer is to estimate the amount of genuine quantum randomness generated in this setup, based only on the observed distribution $p(b|x, y)$. This is a nontrivial task as the apparent randomness of the distribution $[0 < p(b|x, y) < 1]$ can have different origins. On the one hand, it could be genuine quantum randomness. That is, if in a given round of the experiment, the state ρ_x^λ is not an eigenstate of the measurement operator M_y^μ, then the outcome b cannot be predicted with certainty, even if the internal states λ and μ are known, i.e., $0 < p(b|x, y, \lambda, \mu) < 1$. On the other hand, the apparent randomness may be due to technical imperfections, that is, to fluctuations of the internal states λ and μ. Consider the following example: The preparation device emits the states $\rho_x^{\lambda=0} = |0\rangle\langle 0|$ and $\rho_x^{\lambda=1} = |1\rangle\langle 1|$ with $q(\lambda = 0, 1) = 1/2$. For a measurement of the observable $M_y = \hat{z} \cdot \vec{\sigma}$, one obtains that $p(b|x, y) = 1/2$. However, these data clearly contain no quantum randomness, since the outcome b can be perfectly guessed if the internal state λ is known.

Our protocol allows the observer to separate quantum randomness from the randomness due to technical noise. The key technical tool of our protocol is a function recently presented in [24], which works as a "dimension witness." Given data $p(b|x, y)$, the quantity

$$W = \begin{vmatrix} p(1|0,0) - p(1|1,0) & p(1|2,0) - p(1|3,0) \\ p(1|0,1) - p(1|1,1) & p(1|2,1) - p(1|3,1) \end{vmatrix} \quad (4)$$

captures the quantumness of the preparation and measurements. Specifically, if the preparations are classical (i.e., there exists a basis in which all states ρ_x^λ are diagonal), one has that $W = 0$, while a generic qubit strategy achieves $0 \leq W \leq 1$ [24]. $W > 0$ guarantees that the measurements performed by Bob are incompatible (see [25]) and since it is then impossible to simultaneously assign deterministic outcomes to them, this enables us to bound the guessing probability and certify randomness. Given x, y, and knowledge of the internal states λ, μ, the best guess for b is given by $\max_b p(b|x, y, \lambda, \mu)$. Assuming uniformly distributed x and y, the average probability of guessing b fulfils the following inequality (see [25]):

$$p_{\text{guess}} = \frac{1}{8} \sum_{x,y,\lambda,\mu} q_\lambda r_\mu \max_b p(b|x,y,\lambda,\mu) \leq \frac{1}{2}\left(1 + \sqrt{\frac{1 + \sqrt{1 - W^2}}{2}}\right). \quad (5)$$

Therefore, the guessing probability can be upper bounded by a function of W, which can be determined directly from the data $p(b|x, y)$. Finally, to extract random bits from the raw data, we use a randomness extraction procedure. The number of random bits that can be extracted per experimental run is given by the min-entropy $H_{\min} = -\log_2 p_{\text{guess}}$ [27]. Hence $H_{\min}$ is the relevant parameter for determining how the raw data must be postprocessed. Note that randomness can be extracted for any $W > 0$, since $p_{\text{guess}} < 1$ in this case.

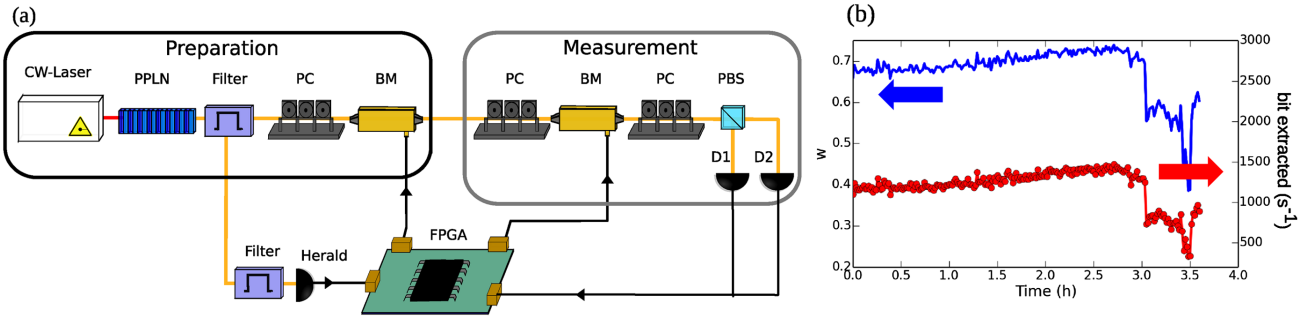

FIG. 2 (color online). Implementing the self-testing QRNG. (a) Experimental setup. (b) Real-time evolution of the witness value W (blue) and randomness generation rate (bits extracted per second; red). After 3 h, the air conditioning in the laboratory is switched off, which leads to misalignment of the optical components. In turn, this leads to a significant drop of the witness value W and corresponding entropy.

The maximal value of $W = 1$ can be reached using the set of preparations and measurements: $\vec{S}_0 = -\vec{S}_1 = \vec{T}_0 = \hat{z}$ and $\vec{S}_2 = -\vec{S}_3 = \vec{T}_1 = \hat{x}$, which correspond to the BB84 QKD protocol [28]. In this case, we can certify randomness with min-entropy $H_{\min} \simeq 0.2284$. Using other preparations and measurements, e.g., if the system is noisy or becomes misaligned, one will typically obtain $0 < W < 1$. Nevertheless, for any value $W > 0$, randomness can be certified, and the corresponding min-entropy can be estimated using Eq. (5). Our protocol is therefore self-testing, since the evaluation of W allows quantifying the amount of randomness in the data. In turn, this allows one to perform adapted postprocessing in order to finally extract random bits.

To conclude this section, we discuss the assumptions which are required in our protocol: (i) *Choice and distribution of settings.*—The devices make no use of any prior information about the choice of settings x and y. (ii) *Internal states of the devices are independent and identically distributed (i.i.d).*—The distributions $q(\lambda)$ and $r(\mu)$ do not vary between experimental rounds. (iii) *Independent devices.*—The preparation and measurement devices are independent, in the sense that $p(\lambda, \mu) = q(\lambda)r(\mu)$. (iv) *Qubit channel capacity.*—The information about the choice of preparation x retrieved by the measurement device (via a measurement on the mediating particle) is contained in a two-dimensional quantum subspace (a qubit).

Assumptions (i) and (iii) are arguably rather natural in a setting where the devices are produced without malicious intent. They concern the independence of devices used in the protocol, namely the preparation and measurement devices, and the choice of settings. When these are produced by trusted (or simply different) providers, it is reasonable to assume that there are no (built-in) preestablished correlations between the devices and that the settings x, y can be generated independently, e.g., using a pseudo RNG. Assumptions (ii) and (iv) are stronger, and will have to be justified for the particular implementation at hand. The content of assumption (ii) is essentially that the devices are memoryless (internal states do not depend on previous events). We believe this assumption can likely be weakened, since randomness can in fact be guaranteed in the presence of certain memory effects, in particular, the experimentally relevant afterpulsing effect (see [25]). Finally, note that assumption (iv) restricts the amount of information about x that is retrieved by the measuring device (via a measurement on the mediating particle), but not the information about x contained in the mediating particle itself. In other words, it might be the case that information about x leaks out from the preparation device via side channels, but we assume that these side channels are not maliciously exploited by the measurement device.

Experiment.—We implemented the above protocol using a fully guided optical setup [see Fig. 2(a)]. The qubit preparations are encoded in the polarization state of single photons, generated via a heralded single-photon source based on a continuous wave (CW) spontaneous parametric down conversion (SPDC) process in a periodically poled lithium niobate (PPLN) waveguide [29]. The idler photon is detected with a ID220 free-running InGaAs/InP single-photon detector (SPD) (herald) with 20% detection efficiency and 20 μs dead time. The polarization is rotated using a polarization controller (PC) and an electro-optical birefringence modulator (BM) based on a lithium niobate waveguide phase modulator. The preparations $x = \{0, 1, 2, 3\}$ correspond, respectively, to the diagonal (D), antidiagonal (A), circular right (R), and circular left (L) polarization states. For the measurement device, polarization measurements are done using a BM and a PC followed by a polarization beam splitter and two ID210 InGaAs/InP SPDs (with a 1.5 ns gate and 25% detection efficiency) triggered by a detection at the heralding detector. The measurements $y = \{0, 1\}$ correspond, respectively, to the $\{D, A\}$ basis and the $\{R, L\}$ basis. The number of photon pairs generated by the SPDC source is set to obtain a count rate at the heralding detector of about 30 kHz, which corresponds to a probability of single photon emission of $p_1 = 6.5 \times 10^{-4}$ per gate, and a two photon emission

$p_2 = p_1^2/2 = 2.1 \times 10^{-7}$ per gate. A field-programmable-gate-array board (FPGA) continuously generates sequences of three pseudorandom bits. Upon successful heralding, these three bits are used to choose (x, y). Finally, the FPGA records the outcome b (whether each ID210 detector has clicked or not).

We briefly discuss to which extent the assumptions of the protocol fit to our implementation. First, the choice of preparation and measurement, x and y, are made by the FPGA using a linear-feedback shift register pseudo RNG [30]. This RNG provides a deterministic cyclic function sampled by the heralding detector. Since the sampling is asynchronous with respect to the RNG rate, the output is uniform and (i) is fulfilled. The BMs are separated spatially by 1 m, their temperature is controlled independently, and the voltages are applied with independent electronic circuits. Any cross talk between them, e.g., due to stray electric fields, can be safely neglected; hence, (iii) is also satisfied. Concerning assumption (ii), we evaluate the distribution $p(b|x, y)$ after every minute of acquisition. Therefore, we need to consider memory effects with time characteristics shorter than one minute. Two main effects should be considered: charge accumulation in the birefringence modulator, and afterpulsing in the detectors, which is a common issue in standard QRNG approaches [4,16]. Importantly, our protocol is robust to afterpulsing (see [25]). Charge effects in the modulator are relevant only for modulation slower than 1 Hz [31]. Finally, the qubit assumption (iv) is arguably the most delicate one. As the choice of preparation x is encoded in the polarization of a single photon, (iv) seems justified. However, a small fraction of heralded events corresponds to multiphoton pulses, in which (iv) is not valid. To take these events into account, we extend our theoretical analysis (see [25]). We show that quantum randomness can still be guaranteed even when (iv) is not fulfilled in all experimental events, provided that the fraction of events violating (iv) can be bounded and is small enough compared to the total number of successful events. To verify this assumption, the probability of single and multiphoton pulses must be properly calibrated. For our single-photon source, the ratio of multiphoton events vs heralds is given by $\sim p_1/2 = 3.25 \times 10^{-4}$, and our method can be applied.

We ran the experiment estimating W for the data accumulated each minute. As discussed in [25], the estimation of W considers finite-size effects and the size of the randomness extractor is determined based on the value of W [16,32]. In the best conditions, our setup generates about 402 bits/s of raw data (before the extractor). The witness corresponds to a value of $W = 0.76$. After extraction, we get final random bits at a rate of 23 bits/s with a confidence of 99%. Note that the confidence level is set when accounting for finite size effects; a higher confidence can be chosen at the expense of a lower rate. Note also that this rate is limited by the slow repetition rate of the experiment (limited by the dead time of the heralding detector) and by the losses in the optical implementation (channel transmission is ~8%; total efficiency ~2%). Figure 2(b) shows the estimated value of W over 3.5 h and the rate at which the final random bits are generated. To demonstrate the self-testing capacity of our protocol, we switched off the air conditioning in the room after 3 h. This impacts the alignment of the setup. As can be seen from Fig. 2(b), the witness value W drops, reflecting the fact that the distributions of internal states [$q(\lambda)$ and $r(\mu)$] changed. In turn, this forces us to perform more postprocessing, resulting in a lower randomness generation rate. Nevertheless, the quality of the final random bits is still guaranteed. This shows that our setup can warrant the generation of high quality randomness, without active stabilization or precise modeling of the impact of the temperature increase.

The quality of the generated randomness can be assessed by checking for patterns and correlations in the extracted bits. We performed standard statistical tests, as defined by NIST, and although not all tests could be performed due to the small size of the sample, all performed tests were successful (see [25]). We stress that these tests do not constitute a proof of randomness (which is impossible); however, failure to pass any of them would indicate the presence of correlations among the output bits.

Finally, we comment on the influence of losses. In the above analysis, we discarded inconclusive events in which the photon was not detected at the measuring device, although the emission of a single photon was heralded by the source. Therefore, our analysis is subject to an additional assumption, namely, that of fair sampling, which we believe is rather natural in the case of nonmalicious devices. Note, however, that this is not necessary strictly speaking, as our protocol is in principle robust to arbitrarily low detection efficiency [24]. Performing the data analysis without the fair-sampling assumption (in which case the inconclusive events are attributed the outcome -1) we obtain witness values of $W \sim 1.5 \times 10^{-4}$, corresponding to $H_{\min} \sim 2.0 \times 10^{-9}$. In this case, the rate for generating random bits drops considerably to 6×10^{-5} bits/s, but importantly does not vanish. Hence, our setup can be used to certify randomness without requiring the fair-sampling assumption. We note that even a small increase in efficiency would lead to a large improvement in rate. E.g., an increase from our current 2% to 10% would already give ~0.04 bits/s while an overall efficiency of 50% would be enough to reach 23 bits/s without postselection, equal to our current postselected rate.

Conclusion.—We have presented a protocol for self-testing QRNG, which allows for real-time monitoring of the entropy of the raw data. This allows adapting the randomness extraction procedure in order to continuously generate high quality random bits. Using a fully optical guided implementation, we have demonstrated that our protocol is practical and efficient, and illustrated its

self-testing capacity. Our work thus provides an approach to QRNG, which can be viewed as intermediate between the standard (device-dependent) approach and the device-independent one.

Compared to the device-dependent approach, our protocol delivers a stronger form of security requiring less characterization of the physical implementation, at the price of a reduced rate compared to commercial QRNGs such as ID Quantique QUANTIS which reaches 4 Mbits/s. A fully device-independent approach [18,19], on the other hand, offers even stronger security [in particular assumptions (ii)–(iv) can be relaxed, hence offering robustness to side channels and memory effects], but its practical implementation is extremely challenging. Proof-of-principle experiments require state-of-the-art setups but could achieve only very low rates [19,20]. Our approach arguably offers a weaker form of security, but can be implemented with standard technology. Our work considers a scenario of trusted but error-prone devices, which we believe to be relevant in practice.

We thank Antonio Acin, Stefano Pironio, Valerio Scarani, and Eric Woodhead for discussions; Raphael Houlmann and Claudio Barreiro for technical support; Batelle and ID Quantique for providing the PPLN waveguide. We acknowledge financial support from the Swiss National Science Foundation (Grant No. PP00P2_138917, Starting Grant DIAQ, and QSIT), SEFRI (COST action MP1006), and the EU project SIQS.

T. L. and J. B. B. contributed equally to this work.

Note added.—After submission of this work, several related works have appeared [33–35].

[1] J. Rarity, P. Owens, and P. Tapster, J. Mod. Opt. **41**, 2435 (1994).
[2] A. Stefanov, N. Gisin, O. Guinnard, L. Guinnard, and H. Zbinden, J. Mod. Opt. **47**, 595 (2000).
[3] T. Jennewein, U. Achleitner, G. Weihs, H. Weinfurter, and A. Zeilinger, Rev. Sci. Instrum. **71**, 1675 (2000).
[4] J. F. Dynes, Z. L. Yuan, A. W. Sharpe, and A. J. Shields, Appl. Phys. Lett. **93**, 031109 (2008).
[5] M. Wahl, M. Leifgen, M. Berlin, T. Rhlicke, H.-J. Rahn, and O. Benson, Appl. Phys. Lett. **98**, 171105 (2011).
[6] Y.-Q. Nie, H.-F. Zhang, Z. Zhang, J. Wang, X. Ma, J. Zhang, and J.-W. Pan, Appl. Phys. Lett. **104**, 051110 (2014).
[7] M. Stipčević and B. M. Rogina, Rev. Sci. Instrum. **78**, 045104 (2007).
[8] B. Qi, Y.-M. Chi, H.-K. Lo, and L. Qian, Opt. Lett. **35**, 312 (2010).
[9] A. Uchida, K. Amano, M. Inoue, K. Hirano, S. Naito, H. Someya, OowadaIsao, T. Kurashige, M. Shiki, S. Yoshimori, K. Yoshimura, and P. Davis, Nat. Photonics **2**, 728 (2008).
[10] C. Abellán, W. Amaya, M. Jofre, M. Curty, A. Acín, J. Capmany, V. Pruneri, and M. W. Mitchell, Opt. Express **22**, 1645 (2014).
[11] C. Gabriel, C. Wittmann, D. Sych, R. Dong, W. Mauerer, U. L. Andersen, C. Marquardt, and G. Leuchs, Nat. Photonics **4**, 711 (2010).
[12] T. Symul, S. M. Assad, and P. K. Lam, Appl. Phys. Lett. **98**, 231103 (2011).
[13] B. Sanguinetti, A. Martin, H. Zbinden, and N. Gisin, Phys. Rev. X **4**, 031056 (2014).
[14] N. Nisan and A. Ta-Shma, J. Comput. Syst. Sci. **58**, 148 (1999).
[15] Y. Dodis, D. Pointcheval, S. Ruhault, D. Vergniaud, and D. Wichs, in *Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security CCS '13* (ACM, New York, 2013), p. 647.
[16] D. Frauchiger, R. Renner, and M. Troyer, arXiv:1311.4547.
[17] X. Ma, F. Xu, H. Xu, X. Tan, B. Qi, and H.-K. Lo, Phys. Rev. A **87**, 062327 (2013).
[18] R. Colbeck, Ph.D. thesis, Trinity College, University of Cambridge [arXiv:0911.3814].
[19] S. Pironio, A. Acín, S. Massar, A. B. de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe, Nature (London) **464**, 1021 (2010).
[20] B. G. Christensen, K. T. McCusker, J. B. Altepeter, B. Calkins, T. Gerrits, A. E. Lita, A. Miller, L. K. Shalm, Y. Zhang, S. W. Nam, N. Brunner, C. C. W. Lim, N. Gisin, and P. G. Kwiat, Phys. Rev. Lett. **111**, 130406 (2013).
[21] 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); H.-W. Li, M. Pawłowski, Z.-Q. Yin, G.-C. Guo, and Z.-F. Han, Phys. Rev. A **85**, 052308 (2012).
[22] M. Dall'Arno, E. Passaro, R. Gallego, M. Pawlowski, and A. Acin, Quantum Inf. Comput. **15**, 0037 (2015).
[23] G. Vallone, D. G. Marangon, M. Tomasin, and P. Villoresi, Phys. Rev. A **90**, 052327 (2014).
[24] J. Bowles, M. T. Quintino, and N. Brunner, Phys. Rev. Lett. **112**, 140407 (2014).
[25] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.114.150501, which includes Ref. [26], for details of the proof of randomness and discussion of assumptions, afterpulsing, multiphoton events, and statistical tests.
[26] W. Hoeffding, J. Am. Stat. Assoc. **58**, 13 (1963).
[27] R. Koenig, R. Renner, and C. Schaffner, IEEE Trans. Inf. Theory **55**, 4337 (2009).
[28] C. H. Bennett and G. Brassard, in *Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing, Bangalore, India* (IEEE, New York, 1984), p. 27.
[29] S. Tanzilli, A. Martin, F. Kaiser, M. De Micheli, O. Alibart, and D. Ostrowsky, Laser Photonics Rev. **6**, 115 (2012).
[30] P. Alfke, *Efficient Shift Registers, LFSR Counters, and Long Pseudo-Random Sequence Generators* (Xilinx Inc., San Jose, CA, 1996).
[31] E. Wooten, K. Kissa, A. Yi-Yan, E. Murphy, D. Lafaw, P. Hallemeier, D. Maack, D. Attanasio, D. Fritz, G. McBrien, and D. Bossi, IEEE J. Sel. Top. Quantum Electron. **6**, 69 (2000).
[32] M. Troyer and R. Renner, A Randomness Extractor for the Quantis Device, http://www.idquantique.com/images/stories/PDF/quantis-random-generator/quantis-rndextract-techpaper.pdf.
[33] M. W. Mitchell, C. Abellan, and W. Amaya, Phys. Rev. A **91**, 012314 (2015).
[34] G. Cañas *et al.*, arXiv:1410.3443.
[35] J. Y. Haw *et al.*, arXiv:1411.4512.

A self-testing quantum random number generator : Supplementary material

Tommaso Lunghi,[1,*] Jonatan Bohr Brask,[2,*] Charles Ci Wen Lim,[1] Quentin Lavigne,[1] Joseph Bowles,[2] Anthony Martin,[1] Hugo Zbinden,[1] and Nicolas Brunner[2]

[1] *Group of Applied Physics, Université de Genève, 1211 Genève, Switzerland*
[2] *Département de Physique Théorique, Université de Genève, 1211 Genève, Switzerland*

In this Supplementary material we provide a proof of randomness for our protocol along with the required assumptions in Sec. I. We show that our protocol is robust to detector afterpulsing in Sec. II. We show how to account for multi-photon events in Sec. III, and we account for finite-size effects in Sec. IV. Finally, we discuss statistical tests applied to the output data.

I. PROOF OF RANDOMNESS

Here we provide a lower bound on the randomness in the observed output using the dimension witness of Ref. [1]. The devices are assumed to be independent, but each device features an internal source of randomness, represented by the variable λ for Alice, and variable μ for Bob. Our goal is to upper bound the probability of guessing the output b that one would have if λ and μ were known, averaged over all inputs and values of the local random variables. Before proceeding with the proof, we first establish the setting in which we will work and state the assumptions made.

A. Setting and assumptions

A priori, the probability of observing a certain output in a given round of the experiment could depend on everything that happened before, and later events could be correlated with the observation of a certain output. However, we will introduce several assumptions which ensure that we can speak about output probabilities without referring to specific rounds as well as the independence of the devices. Let us associate random variables B_i, X_i, Y_i, Λ_i, M_i with the output, the inputs, and the internal variables in round i, and let us write $\vec{B}_i$ for the set of variables $B_1, ..., B_i$ etc. Also, let us denote the probabilities for the random variables to take on specific values by lower case symbols, e.g. $p(x_i) = P(X_i = x_i)$ and $p(\vec{b}_i|\vec{x}_i, \vec{y}_i) = P(\vec{B}_i = \vec{b}_i|\vec{X}_i = \vec{x}_i, \vec{Y}_i = \vec{y}_i)$.

Our first assumption is that *all inputs are independent of each other and the devices*. Formally, X_i is independent of X_j for any $j \neq i$ and of $\vec{Y}_{i-1}$, $\vec{\Lambda}_{i-1}$, $\vec{M}_{i-1}$, and similarly for Y_i. Our second assumption is that *the output in a given round depends only on the inputs in that round and the current state of the devices*. Formally, B_i is conditionally independent of $\vec{B}_{i-1}$, $\vec{X}_{i-1}$, $\vec{Y}_{i-1}$, $\vec{\Lambda}_{i-1}$, and $\vec{M}_{i-1}$ given X_i, Y_i, Λ_i, and M_i. Our third assumption is that *the devices do not record the outputs*. Formally, Λ_i and M_i are independent of $\vec{B}_{i-1}$. Under these assumptions, the probability for a certain string of outputs to occur factorises

$$p(\vec{b}_n|\vec{x}_n, \vec{y}_n, \vec{\lambda}_n, \vec{\mu}_n) = \prod_{i=1}^{n} p(b_i|x_i, y_i, \lambda_i, \mu_i). \quad (1)$$

This can be seen by repeated application of Bayes' rule. The probability to correctly guess the output string $\vec{b}_n$ knowing all the inputs and internal variables in an experiment with n rounds is

$$\begin{aligned} p^g_{\vec{x}_n\vec{y}_n\vec{\lambda}_n\vec{\mu}_n} &= \max_{\vec{b}_n} p(\vec{b}_n|\vec{x}_n, \vec{y}_n, \vec{\lambda}_n, \vec{\mu}_n) \\ &= \prod_{i=1}^{n} \max_b p(b|x_i, y_i, \lambda_i, \mu_i), \end{aligned} \quad (2)$$

and it follows that

$$\begin{aligned} \log(p^g_{\vec{x}_n\vec{y}_n\vec{\lambda}_n\vec{\mu}_n}) &= \sum_{i=1}^{n} \log(\max_b p(b|x_i, y_i, \lambda_i, \mu_i)) \\ &\leq n \log(\frac{1}{n}\sum_{i=1}^{n} \max_b p(b|x_i, y_i, \lambda_i, \mu_i)). \end{aligned} \quad (3)$$

We now assume that *the distribution of the internal randomness is fixed for the duration of the experiment*. Formally, the $\vec{\Lambda}_n$ are identically distributed, and the $\vec{M}_n$ as well. With this assumption, for $n \to \infty$ the sum in the last line above is equivalent to averaging over the inputs and internal variables, that is, it equals

$$\sum_{x,y}\sum_{\lambda,\mu} \max_b p(b|x, y, \lambda, \mu)p(x, y)p(\lambda, \mu). \quad (4)$$

With the final assumption that *the devices are independent*, formally that the $\vec{\Lambda}_n$ are independent of the $\vec{M}_n$, it follows from our proof below that this quantity is bounded by a function of the observed witness value $f(W)$. This implies that in the limit of large n

$$p^g_{\vec{x}_n\vec{y}_n\vec{\lambda}_n\vec{\mu}_n} \leq f(W)^n, \quad (5)$$

and hence the entropy per bit in the output string is bounded by

$$H = -\frac{1}{n}\log_2(p^g_{\vec{x}_n\vec{y}_n\vec{\lambda}_n\vec{\mu}_n}) \geq -\log_2(f(W)). \quad (6)$$

* These authors contributed equally to this work.

We have assumed that the internal random variables are identically distributed in every round. On the physical level, the corresponding requirement is that any external parameters which influence the distributions q_λ, r_μ, such as e.g. temperature, vary slowly on the time-scale of one experimental run, i.e. the time required to gather enough data to estimate the witness value W. In our experimental implementation this time-scale is about one minute. Between different experimental runs there is no requirement for q_λ, r_μ to stay unchanged. We have also assumed that the internal variables are independent of the outputs. Note however that we believe that these assumptions can be relaxed. For example, detector afterpulsing breaks the second assumption, but randomness can nevertheless be certified in our protocol as demonstrated in Sec. II.

B. Proof

Having established the above assumptions, we can now go ahead with our randomness proof without reference to any specific round of the experiment, i.e. we can work just with the distribution $p(b|x,y,\lambda,\mu)$. For given inputs and λ, μ, the guessing probability for this distribution is

$$p^g_{xy\lambda\mu} = \max_b p(b|x,y,\lambda,\mu). \tag{7}$$

The average guessing probability p^g is the average of $p^g_{xy\lambda\mu}$ over the distribution of inputs and local randomness. To proceed, however, we will first derive an upper bound on $p^g_{\lambda\mu}$, defined to be the average over the inputs only.

We consider the witness W of the main text. We thus have four preparations, $x = 0,1,2,3$ and two measurements $y = 0,1$. Consider choices of preparations and measurements which are uniformly random (as explained in the main text, pseudorandomness is sufficient here), i.e. each combination x, y occurs with probability $1/8$. We have that

$$\begin{aligned} p^g_{\lambda\mu} &= \frac{1}{8}\sum_{x,y}\max_b p(b|x,y,\lambda,\mu) \\ &\leq \frac{1}{2}\max_x \sum_y \max_b p(b|x,y,\lambda,\mu) \\ &\leq \frac{1+\cos(\theta_\mu/2)}{2} \end{aligned} \tag{8}$$

where θ_μ denotes the angle between Bob's two measurement. The reasoning of the derivation is as follows. The best guessing probability averaged over inputs of Alice is bounded by the maximum over her inputs. This gives the first inequality and allows us to focus on the best possible state that Alice can send. Next, Bob has two measurements described by Bloch vectors $\vec{T}^\mu_{0,1}$, and θ_μ is the angle between them. The best guessing probability averaged over his inputs is obtained by sending a state which lies in the middle between his measurements on the Bloch sphere (see Fig. 1). For such a state, the outcome probabilities for the two values of b are $\cos^2(\theta_\mu/4)$, and $\sin^2(\theta_\mu/4)$. Choosing the larger value and using the double-angle formula, one arrives at the second inequality.

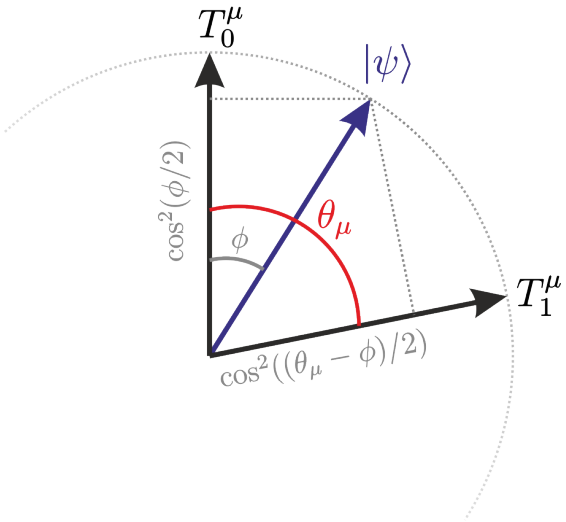

FIG. 1. Cut through the Bloch sphere showing the measurements of Bob, and a state $|\psi\rangle$ lying in the same plane. The probabilities of outcome, say, $b = 1$ are given by the projections of $|\psi\rangle$ onto $T^\mu_{0,1}$. The probabilities when $|\psi\rangle$ makes an angle ϕ with T^μ_0 are indicated. To maximise the average of these, one must choose $\phi = \theta_\mu/2$. Note that choosing a state out of the plane of the measurements can only decrease the guessing probability.

Now we use the fact that a bound on the angle θ_μ can be derived from the witness value for fixed local randomness $W_{\lambda,\mu}$. One has that (see [1])

$$W_{\lambda,\mu} \leq |\vec{T}^\mu_0 \times \vec{T}^\mu_1| \leq \sin\theta_\mu \tag{9}$$

For maximally anti-commuting measurements, we get $W_{\lambda,\mu} = 1$. Combining (8) and (9), we get

$$p^g_{\lambda,\mu} \leq \frac{1}{2}\left(1+\sqrt{\frac{1+\sqrt{1-W^2_{\lambda,\mu}}}{2}}\right) \equiv f(W_{\lambda,\mu}). \tag{10}$$

We note that the function f is concave and decreasing.

Next, we establish the following convexity property of the witness (in a slight abuse of notation, W denotes the observed value of the witness when λ, μ are not known)

$$W \leq \sum_{\lambda,\mu} q_\lambda r_\mu W_{\lambda,\mu}. \tag{11}$$

To see that this holds, consider the entries of the matrix defining W. They are of the form $p(1|x,y) - p(1|x',y)$.

When the devices have internal randomness, we can write

$$\begin{aligned} p(1|x,y) - p(1|x',y) &= \sum_{\lambda,\mu} q_\lambda r_\mu \left(\mathrm{Tr}[\rho_x^\lambda \Pi_{1|y}^\mu] - \mathrm{Tr}[\rho_{x'}^\lambda \Pi_{1|y}^\mu] \right) \\ &= \sum_{\lambda,\mu} q_\lambda r_\mu \vec{S}_{xx'}^\lambda \cdot \vec{T}_y^\mu \\ &= \left(\sum_\lambda q_\lambda \vec{S}_{xx'}^\lambda \right) \cdot \left(\sum_\mu r_\mu \vec{T}_y^\mu \right) \\ &\equiv \vec{S}_{xx'} \cdot \vec{T}_y, \end{aligned} \tag{12}$$

where ρ_x^λ are the states produced by Alice's box, and $\Pi_{1|y}^\mu = (\mathbb{1} + M_y^\mu)/2$ are the projection operators of Bob corresponding to outcome 1, $\vec{T}_y^\mu$ is the Bloch vector corresponding to M_y^μ and $S_{xx'}^\lambda$ is the difference of the Bloch vectors for ρ_x^λ and $\rho_{x'}^\lambda$ (see [1]). Now, from [1] it follows that

$$W = (S_{01} \times S_{23}) \cdot (T_0 \times T_1) \tag{13}$$

$$= \sum_{\lambda,\lambda',\mu,\mu'} q_\lambda q_{\lambda'} r_\mu r_{\mu'} (S_{01}^\lambda \times S_{23}^{\lambda'}) \cdot (T_0^\mu \times T_1^{\mu'}) \tag{14}$$

$$= \sum_{\lambda,\lambda',\mu,\mu'} q_\lambda q_{\lambda'} r_\mu r_{\mu'} |S_{01}^\lambda \times S_{23}^{\lambda'}| |T_0^\mu \times T_1^{\mu'}| \cos\phi_{\lambda,\lambda',\mu,\mu'} \tag{15}$$

where $\phi_{\lambda,\lambda',\mu,\mu'}$ denotes the angle between the vectors $(S_{01}^\lambda \times S_{23}^{\lambda'})$ and $(T_0^\mu \times T_1^{\mu'})$. Next we notice that, for fixed λ, μ, μ', there will be a value of λ' such that $|S_{01}^\lambda \times S_{23}^{\lambda'}| \cos\phi_{\lambda,\lambda',\mu,\mu'}$ is maximal. If we label this value λ and set $q_{\lambda'} = 1$ when $\lambda' = \lambda$ this can only increase the expression. We thus obtain:

$$W \leq \sum_{\lambda,\mu,\mu'} q_\lambda r_\mu r_{\mu'} |S_{01}^\lambda \times S_{23}^{\lambda}| |T_0^\mu \times T_1^{\mu'}| \cos\phi_{\lambda,\mu,\mu'} \tag{16}$$

Using a similar argument, we can eliminate μ':

$$W \leq \sum_{\lambda,\mu} q_\lambda r_\mu |S_{01}^\lambda \times S_{23}^{\lambda}| |T_0^\mu \times T_1^{\mu}| \cos\phi_{\lambda,\mu} \tag{17}$$

$$= \sum_{\lambda,\mu} q_\lambda r_\mu W_{\lambda,\mu}. \tag{18}$$

We are now ready to bound the guessing probability p^g. Using the definition of p^g, (10), and (11) we have

$$p^g = \sum_{\lambda,\mu} q_\lambda r_\mu p_{\lambda,\mu}^g \tag{19}$$

$$\leq \sum_{\lambda,\mu} q_\lambda r_\mu f(W_{\lambda,\mu}) \tag{20}$$

$$\leq f(\sum_{\lambda,\mu} q_\lambda r_\mu W_{\lambda,\mu}) \tag{21}$$

$$\leq f(W) \tag{22}$$

where in the third line we have used Jensen's inequality and concavity of f, and in the last line we have used that f is decreasing. Hence, we finally get

$$p^g \leq \frac{1}{2} \left(1 + \sqrt{\frac{1 + \sqrt{1 - W^2}}{2}} \right) \tag{23}$$

which gives the desired upper bound on the guessing probability as a function of the observed value of the witness W. This bound is tight when maximal violation of the witness is achieved, i.e. $W = 1$. In Sec. IV, we provide the calculation for the maximum number of extractable random bits.

Finally, we provide a proof of the relation between W and the commutativity of the measurements. We write $M_y^\mu = \vec{T}_y^\mu \cdot \vec{\sigma}$, and we have

$$\begin{aligned} \int d\mu r(\mu) || \left[M_0^\mu, M_1^\mu \right] || &= \int d\mu r(\mu) || \left[\vec{T}_0^\mu \cdot \vec{\sigma}, \vec{T}_1^\mu \cdot \vec{\sigma} \right] || \\ &= \int d\mu r(\mu) || 2i (\vec{T}_0^\mu \times \vec{T}_1^\mu) \cdot \vec{\sigma} || \\ &= 2 \int d\mu r(\mu) |\vec{T}_0^\mu \times \vec{T}_1^\mu| \\ &\geq 2 \int d\lambda d\mu q(\lambda) r(\mu) W_{\lambda,\mu} \\ &\geq 2W, \end{aligned} \tag{24}$$

where we have used (9) and (11).

II. CERTIFYING RANDOMNESS IN THE PRESENCE OF AFTERPULSING

In the following we show that although afterpulsing a priory violates the i.i.d. assumption *(iii)*, the self-testing nature of our protocol captures the effect. When afterpulsing is present, the witness value is reduced correspondingly and randomness can still be certified.

To see this, we first consider a hypothetical experiment in which the outputs are generated as follows: in a fraction η of events, the experiment follows and ideal quantum qubit implementation while for the remaining events an outcome is generated at random by the measurement device, determined only by some internal random variable μ independent of the inputs. Let us denote the witness value computed from the whole dataset W, and the value which would be obtained from only the quantum events $\tilde{W}$. To an observer who does not know μ, the non-quantum events look just like uniform noise and the witness values fulfil $W = \eta^2 \tilde{W}$ [1]. At the same time, this scenario meets all of the assumptions in the proof of randomness of Sec. I. Therefore, for an observer with perfect knowledge of μ, who can hence perfectly predict the output for the non-quantum events, the guessing probability on the whole dataset is bounded by

$$p^g \leq f(W) = f(\eta^2 \tilde{W}). \tag{25}$$

We now show that the witness value is reduced in a similar way for afterpulsing, and hence even if the outputs from afterpulsing events can be perfectly predicted, our bound on the randomness still holds.

Consider an experiment generating a set $S = \{(b_1, x_1, y_1), \ldots, (b_N, x_N, y_N)\}$ of N events. The first thing to notice is that afterpulsing is probabilistic: in any given event either there is an afterpulse or there is not. We can therefore think of S as consisting of a set $\tilde{S}$ of $\tilde{N}$ events with no afterpulse and $N - \tilde{N}$ additional afterpulsing events. Let N_{bxy} denote the number of events in S with outcome b and inputs x, y, and $\tilde{N}_{bxy}$ the events in $\tilde{S}$, and define N_{xy}, $\tilde{N}_{xy}$ similarly. For simplicity let us consider the limit of large N such that finite size effects can be neglected. Since the inputs are chosen uniformly $N_{xy} = N/8$. We note that the probability for an afterpulse to occur in a given round i of the experiment does not depend on the inputs x_i, y_i in that round. The number of afterpulses is therefore the same for all combinations of x,y, and $\tilde{N}_{xy} = \eta N/8$ with $\eta = \tilde{N}/N$. In any afterpulsing event, the outcome b_i is also uncorrelated to the inputs x_i, y_i in that round (since $b_i = b_{i-1}$). This means that the effect of afterpulsing when counting events can be written

$$N_{b,x,y} = \tilde{N}_{b,x,y} + c_b, \tag{26}$$

where, importantly, c_b is independent of x (also of y and indeed it may be independent of b, but this is not important in the following).

The witness value on the dataset S is computed from the frequencies $\nu_{b|xy} = N_{b,x,y}/N_{x,y}$. Using the above, we can write

$$\nu_{b|xy} = \frac{\tilde{N}_{b,x,y} + c_b}{N/8} = \frac{\eta \tilde{N}_{b,x,y}}{\eta N/8} + \frac{8c_b}{N} = \eta \tilde{\nu}_{b|xy} + \frac{8c_b}{N}, \tag{27}$$

where $\tilde{\nu}_{b|xy} = \tilde{N}_{b,x,y}/\tilde{N}_{x,y}$ is the frequency one would have obtained considering only the set $\tilde{S}$. Now, since the last term above is independent of x and since the witness is computed solely from terms of the form $\nu_{1|xy} - \nu_{1|x'y}$, we have that

$$W = \eta^2 \tilde{W}, \tag{28}$$

where $\tilde{W}$ is the witness value which one would obtain from the events $\tilde{S}$ without afterpulsing. Since the reduction in W when afterpulses are added is exactly the same as in the scenario above where events with perfectly predictable outputs were added, it follows that even if afterpulse events would be perfectly predictable, the bound (25) on the guessing probability still holds.

III. ACCOUNTING FOR MULTI-PHOTON EVENTS

For real-world sources it is challenging to guarantee that they are of qubit nature. In particular, single-photon sources based on spontaneous parametric down conversion process or weak coherent sources have non-zero probability of emitting more than one photon, violating the qubit assumption.

Given an imperfect source which does not always satisfy the qubit assumption, we would like to say something about the witness violation corresponding to events that do satisfy the assumption. In particular, we would like a lower bound on this violation in terms of the observed, experimental probability distribution and some guarantee on the fraction of non-qubit events. Even without a detailed model of the source, it is possible to determine this fraction e.g. using knowledge of the photon statistics.

A. Bounding the violation for given qubit fraction

To derive a bound on the quantum violation, we will assume that each experimental round either satisfies the qubit assumption, or not. That is, the conditional probability distribution for the experiment can be modeled as

$$p(b|xz) = \alpha p_{qa}(b|xz) + (1-\alpha) p_{\bar{q}a}(b|xz), \tag{29}$$

where α is the fraction of qubit events, p_{qa} is the distribution corresponding to the qubit events, and $p_{\bar{q}a}$ is an unrestricted distribution. The witness value is given in terms of the probabilities by $|W|$, where

$$W = \begin{vmatrix} p(1|0,0) - p(1|1,0) & p(1|2,0) - p(1|3,0) \\ p(1|0,1) - p(1|1,1) & p(1|2,1) - p(1|3,1) \end{vmatrix}. \tag{30}$$

From the model (29), it follows that the expected witness value must satisfy

$$W = |\alpha^2 W_{qa} + (1-\alpha)^2 W_{\bar{q}a} + \alpha(1-\alpha)(G + G')|, \tag{31}$$

where W_{qa}, $W_{\bar{q}a}$ are the determinants corresponding to distributions p_{qa} and $p_{\bar{q}a}$ respectively, and

$$G = \begin{vmatrix} p_{qa}(1|0,0) - p_{qa}(1|1,0) & p_{qa}(1|2,0) - p_{qa}(1|3,0) \\ p_{\bar{q}a}(1|0,1) - p_{\bar{q}a}(1|1,1) & p_{\bar{q}a}(1|2,1) - p_{\bar{q}a}(1|3,1) \end{vmatrix}$$

$$G' = \begin{vmatrix} p_{\bar{q}a}(1|0,0) - p_{\bar{q}a}(1|1,0) & p_{\bar{q}a}(1|2,0) - p_{\bar{q}a}(1|3,0) \\ p_{qa}(1|0,1) - p_{qa}(1|1,1) & p_{qa}(1|2,1) - p_{qa}(1|3,1) \end{vmatrix}$$

To bound the qubit violation for a given expected observed violation we should minimise $|W_{qa}|$ subject to the constraint (31). However, if a certain value W can be attained for a fixed value of $|W_{qa}|$, then attaining all smaller values requires even less qubit violation. We may therefore just as well look for the maximal W for fixed $|W_{qa}|$. Any value above this maximum guarantees a qubit violation of at least $|W_{qa}|$. The maximum has a simple form. It is given by

$$\max W = \max \begin{Bmatrix} 4\alpha(1-\alpha) + \alpha(2\alpha - 1) W_{qa} \\ 2(1-\alpha) + \alpha W_{qa} \end{Bmatrix}. \tag{32}$$

The first thing we notice is that when $\max W$ in (32) is less than 1, it is always given by the first line. This is the relevant case for certifying randomness in practice. Solving for the qubit violation, given an observed violation less than unity we have the bound

$$W_{qa} \geq \frac{1}{\alpha(2\alpha-1)}[W - 4\alpha(1-\alpha)]. \tag{33}$$

Second, we note that for $\alpha > 1/2$ the maximum (32) is always larger than 1. This means that to be able to certify randomness in practice, we need a minimal fraction of events satisfying the qubit assumption of

$$\alpha > \frac{1}{2}. \tag{34}$$

Third, for a given value of α there is a minimal observed violation below which the bound (33) becomes trivial and no randomness can be certified. We must have

$$W > 4\alpha(1-\alpha). \tag{35}$$

B. Estimating the qubit fraction

For an implementation with a particular source, we need an estimate or a lower bound on the fraction of qubit events α. Source and detector inefficiency, and transmission losses lead to inconclusive events, and our estimate of α should be consistent with how these events are dealt with.

In the scenario of non-malicious, error-prone devices considered here, it is rather natural to discard inconclusive events (e.g. assuming fair-sampling) and then compute W from the remaining data. To be able to evaluate (33) in this case, one needs to estimate α when inconclusive events are discarded. It is also natural to assume that all events with at most one photon emitted obey the qubit assumption.

With these assumptions, let q denote the probability for the source to emit at most one photon and consider an experiment with N events and M conclusive events. Before post-selection, asymptotically the fraction of events that obey the qubit assumption is then $\alpha = q$. For a finite number of events, we can put a conservative estimate, i.e., a lower bound, on the number of events N_α that satisfy the qubit assumption, within a given confidence. In particular, under the assumption that we know q, the behaviour of the source is modelled by a family of N Bernoulli trials parameterized by q, and thus the estimation problem can be solved by using the Chernoff-Hoeffding tail inequality. More formally, let $\nu > 0$ be the failure probability of the estimation process and $t > 0$ be the margin parameter, then

$$P(N_\alpha \leq qN - t) \leq \exp(-2Nt^2) = \nu, \tag{36}$$

which implies that $N_\alpha > qN - t$ is true with probability at least $1-\nu$. Equivalently, the fraction of qubit events without post-selection is $\alpha > q - t/N$ with probability at least $1-\nu$. The margin parameter t can be expressed in terms of N and ν as $t = \sqrt{1/(2N)\log(1/\nu)}$.

To account for post-selection, we conservatively assume that all multi-photon events are conclusive. Asymptotically, the fraction of non-qubit events will be $(1-q)N/M$, so $\alpha = 1 - (1-q)N/M$. For finite N we have that after post-selection

$$\alpha \geq 1 - (1-q)\frac{N}{M} - \frac{t}{M} \tag{37}$$

with probability at least $1-\nu$, with ν and t given by (36).

IV. SECURITY ANALYSIS

In this section, we show that with the observed experimental statistics, it is possible to provide a bound on the number of random bits that can be extracted from the raw data set, Z, which takes values from a set of all binary strings, $\mathcal{Z}$ of length m. Our approach essentially uses the (quantum) leftover hash lemma, which states that the amount of private randomness is approximately equal to the min-entropy characterization of the raw data Z. More specifically, it says that the number of extractable random bits (that is independent of variables X, Y, L) is roughly given by $\mathrm{H}_{\min}(Z|XYL)$. Here, we recall that variables X and Y are the inputs of Alice and Bob, respectively, and L is the classical register capturing all information about the local variables λ and μ. The min-entropy of Z given XYL has a clear operational meaning when casted in terms of the guessing probability, i.e., $\mathrm{H}_{\min}(Z|XYL) = -m\log_2 p_{\mathrm{guess}}$: it measures the probability of correctly guessing Z when given access to classical side-information XYL.

On a more concrete level, the leftover hash lemma employs a family of *universal hash functions* to convert Z into an output string S (of size ℓ) that is close to a uniform string conditioned on side-information XYL. In particular, we say that the output string S is Δ-*close* to uniform conditioned on XYL, if

$$\frac{1}{2}\sum_{s,x,y,l} |P_{SXYL} - U_S P_{XYL}| \leq \Delta, \tag{38}$$

where U_S is the uniform distribution of S. The quality of the output string is directly related to the number of extractable random bits, i.e.,

$$\ell = \left\lfloor \mathrm{H}_{\min}(Z|XYL) - 2\log_2\frac{1}{2\Delta} \right\rfloor. \tag{39}$$

Therefore, to bound ℓ, we only need to fix a security level $\epsilon_{\mathrm{sec}} \geq \Delta$ and find a lower bound on the min-entropy term. Using the definition of conditional min-entropy and the assumption that Z is generated from an iid process, we

have

$$\ell = \left\lfloor m - m\log_2\left(1+\sqrt{\frac{1+\sqrt{1-W^2}}{2}}\right) - 2\log_2\frac{1}{2\Delta}\right\rfloor. \tag{40}$$

Accordingly, the rate of extraction is ℓ/m, and it converges to the min-entropy rate when $m\to\infty$ (therefore $\Delta\to 0$). At the moment, our bound on ℓ is written in terms of the expected value of W, which is not directly accessible in the experiment. In order to relate the W to the set of experimental statistics $\mathcal{E} := \{n^+_{x,y}/n_{x,y}\}_{x,y}$, we first use the Chernoff-Hoeffding tail inequality [2], which provides an upper bound on the probability that the sum of random variables deviates from its expected value. We get

$$p(1|x,y) - t(\epsilon_{\rm pe}, n_{x,y}) \overset{\epsilon_{\rm pe}}{\leq} \frac{n^+_{x,y}}{n_{x,y}} \overset{\epsilon_{\rm pe}}{\leq} p(1|xy) + t(\epsilon_{\rm pe}, n_{x,y}), \tag{41}$$

where $t(\epsilon_{\rm pe}, n_{x,y}) := \sqrt{\log(1/\epsilon_{\rm pe})/(2n_{x,y})}$. Here, relations with oversetting $\epsilon_{\rm pe}$ means that the relations are probabilistically true, i.e., the relations hold except with probability $\epsilon_{\rm pe}$. For our purposes later, we denote $p^{\pm}_{x,y} := p(1|x,y) \pm t(\epsilon_{\rm pe}, n_{xy})$. In the following, we introduce an estimate of the expected W, i.e.,

$$W \overset{\epsilon'}{\geq} W_{\min} := \min_{q_{x,y}\in(p^-_{x,y},p^+_{x,y})} \left|W(\{q_{x,y}\})\right|, \tag{42}$$

where $\epsilon' = 8\epsilon_{\rm pe}$ and

$$W(\{q_{x,y}\}) := \det\begin{bmatrix} q_{0,0}-q_{1,0} & q_{2,0}-q_{3,0} \\ q_{0,1}-q_{1,1} & q_{2,1}-q_{3,1}\end{bmatrix}. \tag{43}$$

Next, we need to bound the maximum fraction of non qubit events, $1-\alpha$. Following the discussion in Sec. III, with post-selection we expect α to be $1-\frac{p_2}{p_1+p_2}$ (p_2 and p_1 are the probabilities of the SPDC to emit, respectively, a double pair or a single-photon pair). In the scenario where N preparations are made, by using the Chernoff-Hoedffing tail inequality, we have that

$$\alpha \overset{\epsilon''}{\geq} \hat{\alpha} := 1 - \left[\frac{p_2}{p_1+p_2} + t(\epsilon'', N)\right]. \tag{44}$$

Plugging this into Eq. (C5), we get

$$W_{\rm qa} \overset{\epsilon'+\epsilon''}{\geq} \frac{W_{\min} - 4\hat{\alpha}(1-\hat{\alpha})}{\hat{\alpha}(2\hat{\alpha}-1)}. \tag{45}$$

Therefore, the effective violation is

$$\hat{W}_{\rm eff} := \frac{W_{\min} - 4\hat{\alpha}(1-\hat{\alpha})}{2\hat{\alpha}-1}. \tag{46}$$

Note that the effective violation is obtained by fixing the violation due to non qubit contribution to be zero. In other words, the effective violation measures the amount of randomness in Z. That is, we have

$$\ell = \left\lfloor m - m\log_2\left(1+\sqrt{\frac{1+\sqrt{1-\hat{W}^2_{\rm eff}}}{2}}\right) - 2\log_2\frac{1}{2\Delta}\right\rfloor.$$

Finally, by choosing $\Delta=\epsilon$ and fixing $\epsilon_{\rm pe}=\epsilon''=\epsilon$, the output string S is 10ϵ-close to uniform conditioned on XYL. In the actual implementation we chose $\epsilon = 10^{-3}$.

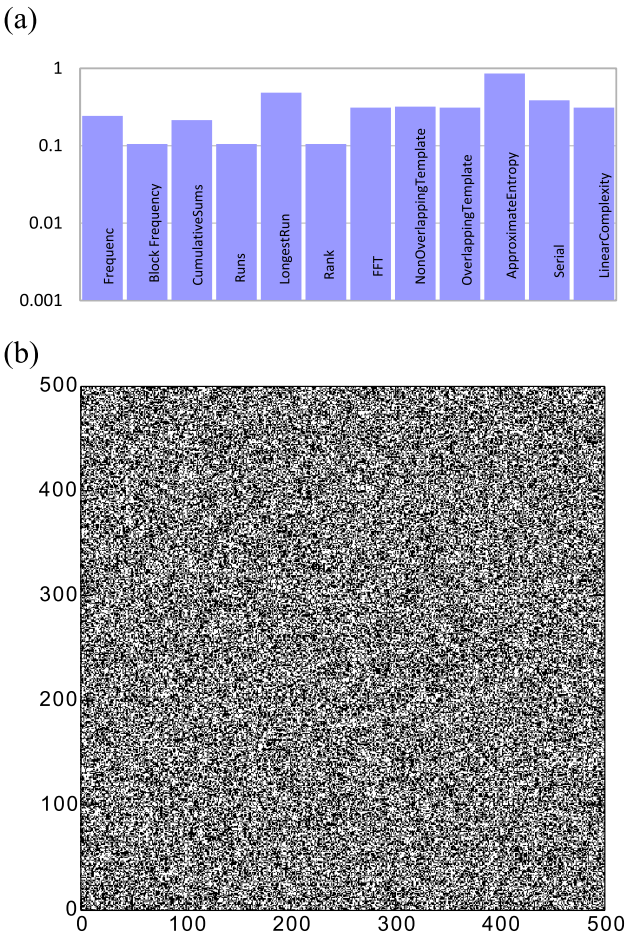

FIG. 2. (a) NIST tests of the data at the output of the extractor. (b) Binary image (500×500) of the extracted random bits.

V. OUTPUT DATA ANALYSIS

We performed tests for assessing the quality of the generated randomness, looking for patterns and correlations in the output data. We performed standard statistical test, as defined by NIST. For each test, the p-value is the result of the Kolmogorov-Smirnov test, and must satisfy $0.01 \leq p \leq 0.99$ to be considered successful. Although

		Measurement			
		D/A		R/L	
		D_1	D_2	D_1	D_2
Preparation	D	5903	97	3515	2485
	A	172	5828	2950	3050
	R	2825	3175	5914	86
	L	3565	2435	199	5801

TABLE I. Sample of raw data taken during one minute under good alignment conditions.

not all tests could be performed due to the small size of the sample, all performed tests were successful (see FIG. 2-(a)). A more visual approach to detecting patterns is illustrated in FIG. 2-(b), where we display 250000 bits in a 500×500 matrix as a black-and-white image. Any repeated pattern or regular structure in the image would indicate correlations among the bits. No pattern appears.

VI. EXAMPLE OF RAW DATA

Here, for completeness, we present an extract of the raw data from our experiment, see Tab. I. The data corresponds to one minute of integration, under good alignment conditions. We give the detector counts observed for each detector (D_1 and D_2), for each measurement setting y and preparation setting x. As mentioned in the main text, the preparations $x = \{0, 1, 2, 3\}$ correspond respectively to the diagonal (D), anti-diagonal (A), circular right (R) and circular left (L) polarization states. The measurements $y = \{0, 1\}$ correspond respectively to the {D,A} basis and the {R,L} basis. In other words, we use the preparations and measurements of the BB84 protocol.

Based on the raw data, we evaluate the asymptotic probability distribution $p(b|x, y)$ using the method presented in Section IV, and then evaluate the witness value W. While perfect BB84 preparations and measurements would give $W = 1$ in the asymptotic limit, the observed value is reduced. This is partly due to alignment errors, but especially to finite-size effects. To illustrate, we compute the W value corresponding to the data in Tab. I with and without accounting for finite-size effects. We find $W = 0.92$ and $W = 0.79$ respectively. These values correspond to visibilities of $V = \sqrt{W} \approx 0.96$ and $V \approx 0.89$ respectively, with respect to the ideal BB84 preparations and measurements mixed with white noise. Note that $W = 0.79$ is not far from the average $W = 0.76$ observed under good conditions (see the main text).

[1] J. Bowles, M. T. Quintino, and N. Brunner, Phys. Rev. Lett. **112**, 140407 (2014).

[2] W. Hoeffding, Journal of the American Statistical Association **58**, 13 (1963).

Paper J

Single-electron entanglement and nonlocality

New Journal of Physics **18**, 043036 (2016)

David Dasenbrook, Joseph Bowles, Jonatan Bohr Brask, Patrick Hofer, Christian Flindt and Nicolas Brunner

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
11 December 2015

REVISED
2 March 2016

ACCEPTED FOR PUBLICATION
8 April 2016

PUBLISHED
26 April 2016

PAPER

Single-electron entanglement and nonlocality

David Dasenbrook[1,3], Joseph Bowles[1], Jonatan Bohr Brask[1], Patrick P Hofer[1], Christian Flindt[2] and Nicolas Brunner[1]

[1] Département de Physique Théorique, Université de Genève, 1211 Genève, Switzerland
[2] Department of Applied Physics, Aalto University, FI-00076 Aalto, Finland
[3] Author to whom any correspondence should be addressed.

E-mail: david.dasenbrook@unige.ch

Keywords: entanglement, quantum transport, single-electron source, electronic interferometer, nonlocality

Abstract

Motivated by recent progress in electron quantum optics, we revisit the question of single-electron entanglement, specifically whether the state of a single electron in a superposition of two separate spatial modes should be considered entangled. We first discuss a gedanken experiment with single-electron sources and detectors, and demonstrate deterministic (i. e. without post-selection) Bell inequality violation. This implies that the single-electron state is indeed entangled and, furthermore, nonlocal. We then present an experimental scheme where single-electron entanglement can be observed via measurements of the average currents and zero-frequency current cross-correlators in an electronic Hanbury Brown–Twiss interferometer driven by Lorentzian voltage pulses. We show that single-electron entanglement is detectable under realistic operating conditions. Our work settles the question of single-electron entanglement and opens promising perspectives for future experiments.

1. Introduction

The field of electron quantum optics has witnessed strong experimental advances over a short period of time [1]. Electronic analogues of the Mach–Zehnder [2], Hanbury Brown–Twiss [3] and Hong-Ou-Mandel interferometers [4] can now be implemented with edge channels of the integer quantum hall effect functioning as wave guides for electrons. At the same time, the recent realization of coherent single-electron emitters is opening up avenues for the controlled manipulation of few-particle electronic states [5–8]. In parallel to these developments, a number of theoretical proposals have been put forward to entangle electrons, e.g. in edge channels [9–13], using either the electron spin or the orbital degrees of freedom. The entanglement is detected by violating a Bell inequality [14, 15] formulated in terms of zero-frequency current cross-correlations [16–18]. While early proposals focus on electron sources driven by static voltages, more recent works investigate the on-demand generation of entangled states using dynamic single-electron emitters [19–23].

For spin or orbital entanglement, several particles are involved and the particles are entangled in the spin or the orbital degrees of freedom, respectively. A conceptually different notion of entanglement is provided by entangled states of different occupation numbers. In this case, the entanglement is between different modes, and the relevant degree of freedom is the particle number in each mode. It is a question that has been much debated whether a state of a single particle in a superposition of two spatially separate modes should be considered entangled [24–30]. For photons (and other Bosons) it is by now well established that the answer is yes, and that the entanglement is in fact useful in quantum communication applications [31, 32]. For electrons (and other Fermions), the situation is different because of charge and parity superselection rules, and the question still causes controversy [33–37].

Here, we revisit this question motivated by the recent development of dynamic single-particle sources in electron quantum optics. We demonstrate rigorously that the answer for electrons is affirmative based on the situation sketched in figure 1(a): two independent sources each produce a single electron which is delocalized with one part transmitted to location A and the other to B. Using only local operations (LOs) and measurements at each location, a Bell inequality between A and B is violated deterministically, i.e. without post-selection. This

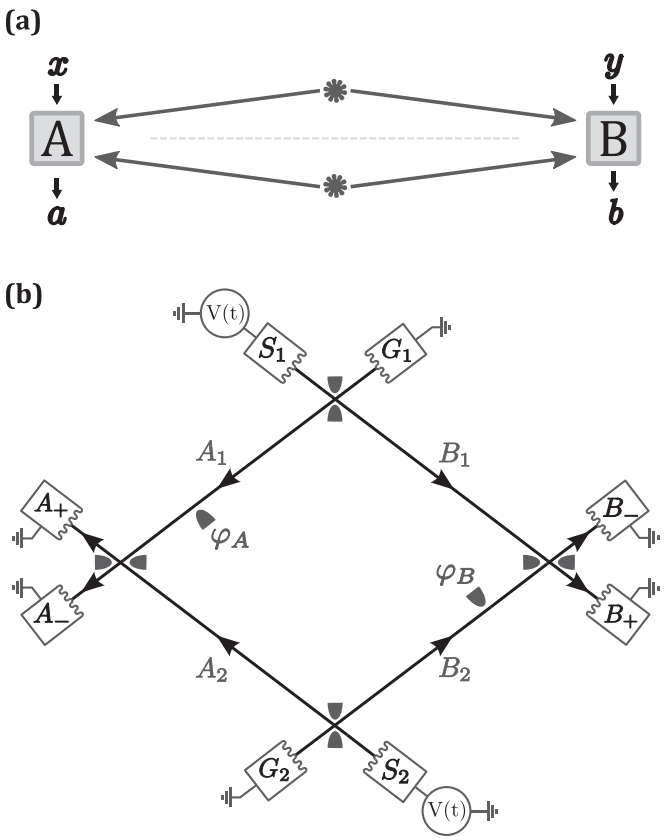

Figure 1. Schematic setup. (**a**) Two independent single-electron sources emit delocalized electrons towards the locations A and B. A Bell test is performed using local operations and measurements at A and B. If the resulting data $p(ab|xy)$ violates a Bell inequality, A and B necessarily share entanglement. Hence, the sources must emit entangled states. (**b**) Electronic Hanbury Brown–Twiss interferometer realizing the idea in (a) for an experimental demonstration of single-electron entanglement. Single-electron excitations are generated at the source contacts S_1 and S_2 and travel to the outputs $A_\pm$ and $B_\pm$. The contacts G_1 and G_2 are grounded.

necessarily implies that there is entanglement between A and B. Since the sources are independent this in turn implies that the state emitted by a single source is entangled between regions A and B. Specifically, we show that such a situation can be realized in an electronic Hanbury Brown–Twiss interferometer driven by Lorentzian voltage pulses as illustrated in figure 1(b). Notably, the single-electron entanglement can be observed from current cross-correlation measurements at the outputs of the interferometer.

2. Single-particle entanglement

We start with a brief introduction to single-particle entanglement. A single particle in a superposition of two different locations can be described by the state

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle_A\,|1\rangle_B + |1\rangle_A\,|0\rangle_B), \tag{1}$$

where the numbers in the kets indicate the particle numbers in the spatially separated modes. The basic question is whether such a state is entangled. One can ask the question both for Bosons and for Fermions, in particular for photons and electrons. To answer affirmatively, the entanglement must be experimentally detectable.

Entanglement should be verified directly from measurements on each spatial mode in equation (1), e.g. by testing the observations against a Bell inequality [14, 15]. If arbitrary measurements were possible, equation (1) should indeed be considered entangled since it for example violates the Clauser–Horne–Shimony–Holt (CHSH) Bell inequality [38]. However, the possible measurements may be limited because the state equation (1) is a single-particle state. Violating the CHSH inequality requires measurements which are not diagonal in the occupation number basis, i.e. they should contain projections onto superpositions of states with different particle numbers such as $(|0\rangle + |1\rangle)/\sqrt{2}$. One may therefore expect a fundamental difference between photons and electrons because global charge conservation and parity superselection [39, 40] forbids such superpositions for electrons [30, 41].

For photons it is by now established that the state given in equation (1) is entangled and in fact useful for applications in quantum communication [32, 42]. Experimental demonstrations of single-photon entanglement have been reported using homodyne [43, 44] and weak displacement measurements [45, 46]. These measurements require the use of coherent states of light (laser light), which introduces additional particles. These particles provide a reference frame between the observers [30, 47]. Alternatively, single-photon entanglement can be converted into entanglement between two atoms [31]. In equation (1), the numbers 0, 1 then represent internal atomic states and entanglement can be verified straightforwardly. Importantly, since the conversion process involves only LOs, one concludes that the original single-photon state given in equation (1) must have been entangled. These procedures, however, cannot be straightforwardly applied to Fermions (for example, there is no equivalent of coherent states for Fermions). Hence, a more careful analysis is necessary as we show in the following.

3. Single-electron entanglement and nonlocality

We consider the experiment pictured in figure 1(b) and now argue that single-electron entanglement is observable. To keep the analysis simple, we work at zero temperature and assume that the sources create single electronic excitations above the Fermi sea which can be detected one by one. These assumptions do not contradict any fundamental principle such as charge conservation. We consider the possibility of an experimental implementation with current technology later on.

Single electrons are excited above the Fermi sea at the sources S_1 and S_2, and are coherently split and interferred on electronic beamsplitters—quantum point contacts (QPCs) tuned to half transmission. Tunable phases φ_A and φ_B can be applied in one arm on either side of the interferometer. The phases can be tuned using side gates or by changing the magnetic flux Φ through the device. In the latter case, we have $2\pi\Phi/\Phi_0 = \varphi_A + \varphi_B$, where $\Phi_0 = h/e$ is the magnetic flux quantum.

Labelling the modes as indicated in the figure, in second quantized notation the top beam splitter implements the transformation $a^\dagger_{S_1} \to (a^\dagger_{A_1} + a^\dagger_{B_1})/\sqrt{2}$, $a^\dagger_{G_1} \to (a^\dagger_{A_1} - a^\dagger_{B_1})/\sqrt{2}$ and similarly for the others. Here, we have introduced the Fermionic creation and annihilation operators $a^\dagger_\alpha$ and a_α for electrons above the Fermi sea in mode α. Considering just the top source (S_1), the state created after the beam splitter is thus

$$\frac{1}{\sqrt{2}}(a^\dagger_{A_1} + a^\dagger_{B_1})|0\rangle, \tag{2}$$

where the state $|0\rangle$ represents the undisturbed Fermi sea. This is the electronic version of equation (1), and we use the interferometer to demonstrate that the state indeed is entangled between the regions A and B.

The joint initial state of the two sources is $a^\dagger_{S_1} a^\dagger_{S_2} |0\rangle$, and the state evolution up to the output of the interferometer is then

$$\begin{aligned} a^\dagger_{S_1} a^\dagger_{S_2} |0\rangle &\to \frac{1}{2}(a^\dagger_{A_1} e^{i\varphi_A} + a^\dagger_{B_1})(a^\dagger_{A_2} + a^\dagger_{B_2} e^{i\varphi_B})|0\rangle \\ &\to \frac{1}{4}[a^\dagger_{A_+} a^\dagger_{B_+}(e^{i\varphi} - 1) + a^\dagger_{A_+} a^\dagger_{B_-}(e^{i\varphi} + 1) \\ &\quad + a^\dagger_{A_-} a^\dagger_{B_+}(e^{i\varphi} + 1) + a^\dagger_{A_-} a^\dagger_{B_-}(e^{i\varphi} - 1) \\ &\quad - 2e^{i\varphi_A} a^\dagger_{A_+} a^\dagger_{A_-} + 2e^{i\varphi_B} a^\dagger_{B_+} a^\dagger_{B_-}]|0\rangle, \end{aligned} \tag{3}$$

where $\varphi = \varphi_A + \varphi_B$ and we have used the Fermionic anti-commutation relations $\{a^\dagger_i, a_j\} = \delta_{ij}$ and $\{a^\dagger_i, a^\dagger_j\} = \{a_i, a_j\} = 0$. We omit terms where two electrons go to the same output since these are ruled out by the Pauli exclusion principle[4].

Assuming that single-electron detection is possible, the state given in equation (3) can be seen to violate the CHSH inequality using the following strategy: the phases φ^x_A and φ^y_B are determined by the inputs $x, y = 0, 1$, and the binary outputs $a, b = \pm 1$ are determined by outputting ± 1 when one click is observed in detector $A_\pm$ (similarly for B). In cases where both or none of the detectors click, the outputs are defined to be $+1$ and -1 respectively. We denote the probability for outputs a, b given inputs x, y by $P(ab|xy)$. The correlator defined as

$$E_{xy} = \sum_{a,b} abP(ab|xy) \tag{4}$$

[4] Such terms vanish due to the Fermionic anti-commutation relations, e.g. $2(a^\dagger_{A_1})^2 = \{a^\dagger_{A_1}, a^\dagger_{A_1}\} = 0$.

is then given by

$$E_{xy} = -\frac{1 + \cos(\varphi_A^x + \varphi_B^y)}{2}. \tag{5}$$

If the experiment can be explained by a local hidden variable model, then the CHSH inequality holds [38]

$$S = |E_{00} + E_{01} + E_{10} - E_{11}| \leqslant 2. \tag{6}$$

Now, with the choice $\varphi_A^0 = 0$, $\varphi_A^1 = \pi/2$, $\varphi_B^0 = -3\pi/4$, and $\varphi_B^1 = 3\pi/4$, we find

$$S = 1 + \sqrt{2} > 2. \tag{7}$$

Thus, the CHSH inequality is clearly violated. Since the state given in equation (3) violates a Bell inequality between A and B, it must necessarily be entangled. Note that this Bell inequality violation is not subjected to the detection loophole [15], as our scheme does not involve any post-selection. Furthermore, the state given in equation (3) was created by LOs on two copies of the state given in equation (2) coming from two independent sources. Since any product of separable states is separable, it follows that the state given in equation (2) must itself be entangled. We thus conclude that the state of a single electron split between two modes is entangled.

It should be pointed out that the setup in figure 1(b) is similar to the Hanbury Brown–Twiss interferometer for electrons, as theoretically proposed [12] and experimentally realized [3] using edge states of a two-dimensional electron gas in the integer quantum hall regime. However, in these works maximal CHSH inequality violation ($S = 2\sqrt{2}$) is achieved by post-selection on the subspace of one electron on each side of the interferometer (effectively post-selecting a maximally entangled state), which is interpreted as two-electron orbital entanglement. Here, by contrast, our scheme involves no post-selection and we do not achieve maximal CHSH violation, but in turn we can demonstrate single-electron entanglement.

It should also be noted that the possibility of using two copies of a single electron entangled state in order to distill one entangled two-electron state has been discussed in [29, 48]. There, the idea is that each observer performs a nondemolition measurement of the local electron number and then post-selects on the cases where a single electron is detected on each side. Alternatively, the distilled entanglement can be transferred to a pair of additional target particles [49], in which case however single-electron nonlocality cannot be unambiguously concluded. Again, as argued above, our setup involves no post-selection and is thus conceptually different. Moreover, the setup does not require nondemolition measurements.

The scheme described so far is a thought experiment, demonstrating that single-electron entanglement in theory is observable. In principle, nothing prevents its realization. Single-electron sources [5, 7, 8] and electronic beam splitters have been experimentally realized and the first steps towards single-electron detectors [6, 50] have recently been taken. Still, realizing our thought experiment is at present challenging, mainly because of the requirement to detect single electrons. To relax this constraint, we discuss in the next section an experiment which only relies on measurements of the average current and the zero-frequency current-correlators. These are standard measurements which would also demonstrate single-electron entanglement, albeit under slightly stronger assumptions about the experimental implementation.

4. Observing single-electron entanglement

We consider again the setup in figure 1(b), but now discuss a detection scheme which is feasible using existing technology. Specifically, we consider measurements of zero-frequency currents and current correlators as an alternative to single-electron detection. We give a detailed description of the single-electron sources and the interferometer based on Floquet scattering theory [51–54]. This allows us to investigate realistic operating conditions such as finite electronic temperatures and dephasing. As we will see, it is possible to demonstrate single-electron entanglement under one additional assumption, namely that the measurement of the mean current and the zero-frequency current correlators amounts to taking ensemble averages over the state in each period of the driving. This is a reasonable assumption if the period of the driving is so long that only one electron from each source is traversing the interferometer at any given time.

For the single-electron sources, we consider the application of Lorentzian-shaped voltage pulses to the contacts [7, 8, 55–58]. A driven mesoscopic capacitor [5] can be used instead. Electrons leaving a contact pick up a time-dependent phase

$$\varphi(t) = -\frac{e}{\hbar}\int_{-\infty}^{t} V(t')\,\mathrm{d}t', \tag{8}$$

where the voltage applied to the contact has the form

$$eV(t) = \sum_{j=-\infty}^{\infty} \frac{2\hbar\Gamma}{(t - n\mathcal{T})^2 + \Gamma^2}. \tag{9}$$

At zero temperature, this results in the excitation of exactly one electron out of the Fermi sea (and one hole going into the contact) without any additional electron-hole pairs. This quasiparticle is called a leviton [7, 8]. In equation (9), the temporal width of the pulse is denoted as Γ and $\mathcal{T}$ is the period of the driving.

Floquet scattering theory provides us with a convenient theoretical framework to describe the periodically driven interferometer [51–54]. By Fourier transforming equation (8), we obtain the Floquet scattering matrix of the driven contacts as

$$\mathcal{S}_l(n) = \begin{cases} -2e^{-n\Omega\Gamma}\sinh(\Omega\Gamma) & n > 0 \\ e^{-\Omega\Gamma} & n = 0 \\ 0 & n < 0. \end{cases} \tag{10}$$

These are the amplitudes for an electron at energy E to leave the contact at energy $E_n = E + n\hbar\Omega$, having absorbed ($n > 0$) or emitted ($n < 0$) $|n|$ energy quanta of size $\hbar\Omega$, where $\Omega = 2\pi/\mathcal{T}$ is the frequency of the driving.

The scattering matrix of the interferometer can be found as follows. Since there are eight terminals in total (four inputs and four outputs), the scattering matrix of the interferometer is an 8×8 matrix. However, due to the chirality of the edge states, electrons leaving an input contact can only travel to an output. This allows us to work with an effective 4×4 scattering matrix connecting every possible input to every possible output. Including the phases φ_A and φ_B, that the particles pick up when travelling from input 1 to location A or from input 2 to B, the scattering matrix reads

$$\mathcal{S} = \begin{pmatrix} r_1 r_A e^{i\varphi_A} & r_1 t_A e^{i\varphi_A} & t_1 t_B & t_1 r_B \\ t_1 r_A e^{i\varphi_A} & t_1 t_A e^{i\varphi_A} & -r_1 t_B & -r_1 r_B \\ t_2 t_A & -t_2 r_A & -r_2 r_B e^{i\varphi_B} & r_2 t_B e^{i\varphi_B} \\ -r_2 t_A & r_2 r_A & -t_2 r_B e^{i\varphi_B} & t_2 t_B e^{i\varphi_B} \end{pmatrix}. \tag{11}$$

Here, $t_{1(2)}$ refers to the transmission amplitude of the QPCs after source 1 (2) and $t_{A(B)}$ is the amplitude for the QPC located at $A(B)$. The r's are the corresponding reflection amplitudes. The rows number the possible inputs S_1, G_1, S_2 and G_2 (in this order) and the columns the possible outputs $A+$, $A-$, $B+$ and $B-$. We have chosen all amplitudes to be real and inserted factors of -1 for half of the reflection amplitudes to ensure the unitarity of the scattering matrix. Below, we consider only half-transparent beam splitters and thus set all amplitudes to $1/\sqrt{2}$.

To obtain the combined Floquet scattering matrix of the interferometer and the single-electron sources, we multiply every matrix element of the stationary $\mathcal{S}$-matrix corresponding to a voltage-biased input (i. e. the first and third rows) by $\mathcal{S}_l(n)$ and every element corresponding to a grounded input (i. e. the second and fourth rows) by δ_{n0}. In doing so, we assume that the two electron sources are perfectly synchronized and all arms of the interferometer have the same length. The resulting Floquet scattering matrix $\mathcal{S}_F(E_n, E) \equiv \mathcal{S}_F(n)$ is the basis of all calculations below.

The current operator in output α is given by [59]

$$I_\alpha = \frac{e}{h}\int_{-\infty}^{\infty} \{c_\alpha^\dagger(E)c_\alpha(E) - b_\alpha^\dagger(E)b_\alpha(E)\}\,dE, \tag{12}$$

where the operators $c_\alpha(E)$ ($b_\alpha(E)$) annihilate an incoming (outgoing) electron in lead α at energy E. Outgoing electrons from the leads are distributed according to the Fermi–Dirac distribution function

$$\langle b_\alpha^\dagger(E)b_\beta(E')\rangle = \delta_{\alpha\beta}\delta(E - E')\frac{1}{e^{E/(k_BT)} + 1}, \tag{13}$$

where T is the electronic temperature and we have set the Fermi level in all reservoirs to $E_F = 0$. The scattered electrons are not in thermal equilibrium. We find their distribution by relating the incoming electrons to the outgoing ones via the Floquet scattering matrix as [53]

$$c_\alpha(E) = \sum_{n=-\infty}^{\infty}\sum_{\beta}[\mathcal{S}_F(E, E_n)]_{\alpha\beta}b_\beta(E_n). \tag{14}$$

4.1. Zero temperature

At zero temperature, the average currents and the zero-frequency current correlators can be calculated analytically using equations (12) and (14). For example, the average current at output $A+$ reads

$$\langle I_{A+}\rangle = \frac{e}{\mathcal{T}}(T_2 T_A + T_1 R_A), \tag{15}$$

where $T_i = |t_i|^2$ and $R_i = |r_i|^2$ ($i = 1, 2, A, B$). The zero-frequency current cross-correlator is defined as

$$P_{\alpha\beta} = \langle I_\alpha I_\beta\rangle - \langle I_\alpha\rangle\langle I_\beta\rangle. \tag{16}$$

For the cross-correlator between the $A+$ and $B+$ outputs we obtain

$$P_{A+B+} = -\frac{e^2}{\mathcal{T}}\,|t_2 t_A r_2 t_B e^{i\varphi_B} + t_1 r_A r_1 r_B e^{-i\varphi_A}|^2. \tag{17}$$

Note that the average currents are insensitive to the phases φ_A and φ_B, whereas the current cross-correlators depend on their sum $\varphi_A + \varphi_B$. This is known as the two-particle Aharonov–Bohm effect [12].

We now formulate the CHSH inequality [38] for our system. The leviton annihilation operator is [58]

$$a_\alpha = \sqrt{2\Gamma}\sum_{E>0} e^{-\Gamma E/\hbar} b_\alpha(E). \tag{18}$$

At zero temperature, we can express the operator of the number of levitons emitted from lead α per period in terms of the current operator as

$$a_\alpha^\dagger a_\alpha = \frac{\mathcal{T}}{e} I_\alpha. \tag{19}$$

This allows us to relate the current operator for a given detector at A or B to an operator on the modes on side A or B before the final beam splitter and the phase shift, see figure 1(b). Taking for instance the detector A_+ and transforming equation (19) through the beam splitter and the phase shift, we get

$$\begin{aligned} a_{A_+}^\dagger a_{A_+} &\to \frac{1}{2}(e^{-i\varphi_A} a_{A_1}^\dagger + a_{A_2}^\dagger)(e^{i\varphi_A} a_{A_1} + a_{A_2}) \\ &= \frac{1}{2}(a_{A_1}^\dagger a_{A_1} + a_{A_2}^\dagger a_{A_2}) + \frac{1}{2}(e^{-i\varphi_A} a_{A_1}^\dagger a_{A_2} + e^{i\varphi_A} a_{A_2}^\dagger a_{A_1}). \end{aligned} \tag{20}$$

To gain an intuitive understanding of this operator, we consider its restriction to the single-electron subspace, i.e. the case where there is exactly one electron on side A of the interferometer. In this case, the first term in equation (20) is just $1/2$. The Hilbert space is two-dimensional and the states $a_{A_1}^\dagger|0\rangle$, $a_{A_2}^\dagger|0\rangle$ form a basis. In this basis, the second term in equation (20) is $(\cos(\varphi_A)\sigma_x + \sin(\varphi_A)\sigma_y)/2$, with σ_x, σ_y, σ_z being the usual Pauli matrices. Thus, in the single-electron subspace we have

$$I_{A_+} = \frac{e}{2\mathcal{T}}(1 + \sigma_{\varphi_A}^A), \tag{21}$$

where $\sigma_{\varphi_A}^A = \cos(\varphi_A)\sigma_x^A + \sin(\varphi_A)\sigma_y^A$ is the rotated Pauli matrix in the x–y plane, acting on side A. From this we see that, in the single-electron subspace, measuring I_{A_+} is equivalent to measuring $\sigma_{\varphi_A}^A$. Similar expressions can be obtained for the currents at the other detectors, and thus, by measuring the currents at the four outputs, we can measure any combination of Pauli operators in the two-qubit subspace with a single electron on each side of the interferometer.

With this in mind, we define the observables

$$X_A^{\varphi_A} = \frac{2\mathcal{T}}{e} I_{A_+}^{\varphi_A} - 1, \qquad X_B^{\varphi_B} = \frac{2\mathcal{T}}{e} I_{B_+}^{\varphi_B} - 1, \tag{22}$$

where the current for a given phase setting φ is denoted as I_α^φ. In the subspace with one electron on each side of the interferometer, these correspond to measuring (rotated) Pauli operators. Events where two or no electrons arrive on the same side will give contributions of $+1$ or -1 respectively, see equation (19), independent of the phase settings, analogously to the output strategy in the previous section. At zero temperature the correlator becomes

$$\langle X_A^{\varphi_A} X_B^{\varphi_B}\rangle = -\frac{1 + \cos(\varphi_A + \varphi_B)}{2}, \tag{23}$$

showing that the joint statistics is the same as in section 3, where single-electron detection was assumed. Here, however, we interpret the current expectation values entering in the correlator, such as $\langle I_{A_+}^{\varphi_A}\rangle$, as the result of time-integrated measurements. We thus assume that a measurement of the time-integrated current and the zero-frequency current correlators amounts to taking ensemble averages over the state in each period of the driving. The statistics obtained from the time-integrated current measurement is then the same as what one would obtain by averaging over several periods of the driving with single-electron detection. Under this assumption, we can again consider the CHSH inequality

$$S = |\langle X_A^{\varphi_A^0} X_B^{\varphi_B^0} + X_A^{\varphi_A^0} X_B^{\varphi_B^1} + X_A^{\varphi_A^1} X_B^{\varphi_B^0} - X_A^{\varphi_A^1} X_B^{\varphi_B^1}\rangle| \leqslant 2. \tag{24}$$

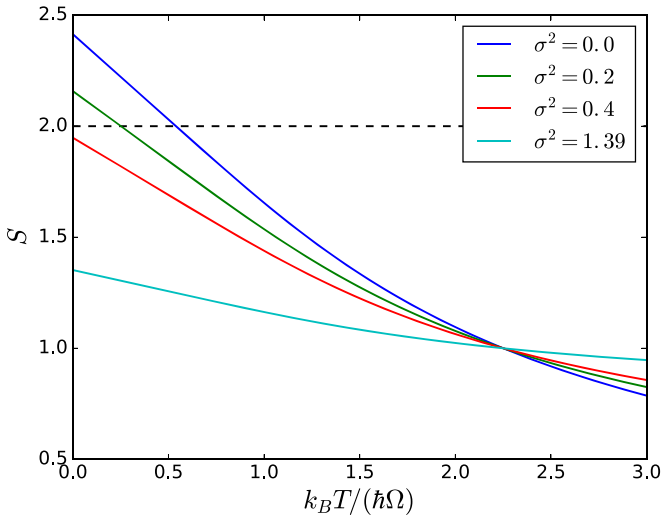

Figure 2. Maximal value of the CHSH parameter as a function of temperature. The Bell angles are $\varphi_A = 0$, $\varphi'_A = \pi/2$, $\varphi_B = -\pi/4$ and $\varphi'_B = 5\pi/4$. The dephasing parameter σ^2 is the variance of the distribution of the sum of the phases $\varphi_A + \varphi_B$. The dashed line indicates the CHSH bound.

It is easy to see that the choice $\varphi^0_A = 0$, $\varphi^1_A = \pi/2$, $\varphi^0_B = -3\pi/4$, $\varphi^1_B = 3\pi/4$ leads to a violation, giving

$$S = 1 + \sqrt{2} > 2. \tag{25}$$

This finally shows us that this scheme makes it possible to observe single-electron entanglement using zero-frequency measurements only.

We note that our results for the current and the zero-frequency noise do not depend on the pulse width Γ. As such, our measurement strategy based on equation (22) would also work with constant voltages as realized in the experiment by Neder *et al* [3], and the CHSH violation of equation (25) would be obtained. However, to unambiguously demonstrate single-electron entanglement, in line with the thought experiment described in section 3, it is important that only one electron from each source is traversing the interferometer at any given time. We therefore need to work with a long period and well-separated pulses, as opposed to constant voltages.

It is instructive to compare our proposal to the previous work of Samuelsson *et al* [12]. Although the two setups are similar, the detection scheme discussed here is different. This significantly changes the interpretation of the observations. The measurement scheme suggested by Samuelsson *et al* is formulated in terms of coincidence rates [12, 60]. The corresponding observable is then sensitive only to the part of the state with a single electron on each side of the interferometer. Thus, the measurement effectively corresponds to performing post-selection, discarding the part of the state where two electrons are on the same side. In this case, the CHSH inequality is maximally violated ($S = 2\sqrt{2}$), as the post-selected state is a maximally entangled two-qubit state. The Bell inequality is then violated because of the two-electron orbital entanglement [12]. By contrast, our measurement strategy is sensitive to the entire state (including terms with two electrons on the same side) and does not imply any effective post-selection. For this reason we reach a lower CHSH violation, $S = 1 + \sqrt{2}$. However, we observe in turn single-electron entanglement.

4.2. Finite temperatures and dephasing

At finite temperatures, additional excitations in terms of electron-hole pairs are expected. Consequently, equation (19) does not hold any longer. The operators in equation (22) are thus not strictly bounded between –1 and +1, although values outside this range should be rare at low temperatures. Since the CHSH parameter S is a monotonically decreasing function of temperature, a violation of the CHSH inequality at finite temperatures indicates that the corresponding zero temperature state is unambiguously entangled. We will thus continue to use equation (24) to detect single-particle entanglement.

At finite temperatures, the average current and the zero-frequency current correlators can be calculated numerically. Figure 2 shows the maximal value of the CHSH parameter (using the same phase settings as above) as a function of the electronic temperature. In the absence of any additional dephasing mechanisms (blue curve), the CHSH inequality can be violated up to a temperature of $k_B T \approx 0.5\hbar\Omega$. For a typical driving frequency of 5 GHz [7, 8], this corresponds to a temperature of about 120 mK, which is well within experimental reach.

Due to interactions with the electrons in the underlying Fermi sea as well as with nearby conductors, the injected single-electron states may experience decoherence and dephasing. Here we do not give a microscopic

model for theses interactions, but instead we introduce a phenomenological dephasing parameter σ^2 which denotes the variance of the total phase $\varphi_A + \varphi_B$ in a model that leads to Gaussian phase averaging. Previous experiments have shown that this is the dominant effect of the interaction of electronic interferometers with their environments [2, 61]. At zero temperature, the correlator in equation (23) then becomes

$$\langle X_A^{\varphi_A} X_B^{\varphi_B} \rangle = -\frac{1 + \mathrm{e}^{-\sigma^2} \cos(\varphi_A + \varphi_B)}{2}, \tag{26}$$

making a Bell violation possible up to $\sigma^2 \lesssim 0.35$. At finite temperatures, an analogous expression can be found [60] and the dephasing has a similar qualitative effect. Figure 2 shows that for small values of the dephasing parameter, a CHSH violation is still possible at low enough temperatures, while for $\sigma^2 \gtrsim 0.35$, the entanglement cannot be detected any longer. We note that the visibility of the current correlators observed in the experiment by Neder *et al* [3] is too low to violate equation (24) in this setup. It corresponds to a dephasing parameter of $\sigma^2 \approx 1.39$ (light blue line in figure 2). Nevertheless, by a careful design of the interferometer the dephasing may be further reduced, bringing the measurement described here within experimental reach.

5. Conclusions

We have revisited the question of single-electron entanglement. Specifically, we have demonstrated theoretically that the state of a single electron in a superposition of two separate spatial modes is entangled. As we have shown, single-electron entanglement can in principle be observed in an electronic Hanbury Brown–Twiss interferometer based on single-electron sources, electronic beam splitters, and single-electron detectors. Unlike earlier proposals for generating entanglement in electronic conductors, our scheme does not rely on any post-selection procedures. Since single-electron detectors are still under development, we have devised an alternative experimental scheme based on existing technology using average current and cross-correlation measurements. With these developments, the experimental perspectives for observing single-electron entanglement seem promising.

Acknowledgments

DD and PPH gratefully acknowledge the hospitality of Aalto University and McGill University, respectively. CF is affiliated with Centre for Quantum Engineering at Aalto University. This work was supported by the Swiss National Science Foundation (grants 200020_150082, PP00P2_138917 and Starting Grant DIAQ) and the Academy of Finland.

References

[1] Bocquillon E *et al* 2014 Electron quantum optics in ballistic chiral conductors *Ann. Phys.* **526** 1
[2] Ji Y, Chung Y, Sprinzak D, Heiblum M, Mahalu D and Shtrikman H 2003 An electronic Mach–Zehnder interferometer *Nature* **422** 415
[3] Neder I, Ofek N, Chung Y, Heiblum M, Mahalu D and Umansky V 2007 Interference between two indistinguishable electrons from independent sources *Nature* **448** 333
[4] Bocquillon E, Freulon V, Berroir J-M, Degiovanni P, Plaçais B, Cavanna A, Jin Y and Fève G 2013 Coherence and indistinguishability of single electrons emitted by independent sources *Science* **339** 1054
[5] Fève G, Mahé A, Berroir J-M, Kontos T, Plaçais B, Glattli D C, Cavanna A, Etienne B and Jin Y 2007 An on-demand coherent single-electron source *Science* **316** 1169
[6] Fletcher J D *et al* 2013 Clock-controlled emission of single-electron wave packets in a solid-state circuit *Phys. Rev. Lett.* **111** 216807
[7] Dubois J, Jullien T, Portier F, Roche P, Cavanna A, Jin Y, Wegscheider W, Roulleau P and Glattli D C 2013 Minimal-excitation states for electron quantum optics using levitons *Nature* **502** 659
[8] Jullien T, Roulleau P, Roche B, Cavanna A, Jin Y and Glattli D C 2014 Quantum tomography of an electron *Nature* **514** 603
[9] Lesovik G B, Martin T and Blatter G 2001 Electronic entanglement in the vicinity of a superconductor *Eur. Phys. J.* B **24** 287
[10] Oliver W D, Yamaguchi F and Yamamoto Y Jan 2002 Electron entanglement via a quantum dot *Phys. Rev. Lett.* **88** 037901
[11] Beenakker C W J, Emary C, Kindermann M and van Velsen J L 2003 Proposal for production and detection of entangled electron-hole pairs in a degenerate electron gas *Phys. Rev. Lett.* **91** 147901
[12] Samuelsson P, Sukhorukov E V and Büttiker M 2004 Two-particle Aharonov–Bohm effect and entanglement in the electronic Hanbury Brown–Twiss setup *Phys. Rev. Lett.* **92** 026805
[13] Scarani V, Gisin N and Popescu S 2004 Proposal for energy-time entanglement of quasiparticles in a solid-state device *Phys. Rev. Lett.* **92** 167901
[14] Bell J 1964 On the Einstein Podolsky–Rosen paradox *Physics* **1** 195
[15] Brunner N, Cavalcanti D, Pironio S, Scarani V and Wehner S 2014 Bell nonlocality *Rev. Mod. Phys.* **86** 419
[16] Kawabata S 2001 Test of Bell's inequality using the spin filter effect in ferromagnetic semiconductor microstructures *J. Phys. Soc. Japan* **70** 1210–3
[17] Samuelsson P, Sukhorukov E V and Büttiker M Oct 2003 Orbital entanglement and violation of Bell inequalities in mesoscopic conductors *Phys. Rev. Lett.* **91** 157002
[18] Samuelsson P, Neder I and Büttiker M 2009 Reduced and projected two-particle entanglement at finite temperatures *Phys. Rev. Lett.* **102** 106804

[19] Splettstoesser J, Moskalets M and Büttiker M 2009 Two-particle nonlocal Aharonov–Bohm effect from two single-particle emitters *Phys. Rev. Lett.* **103** 076804
[20] Sherkunov Y, d'Ambrumenil N, Samuelsson P and Büttiker M Feb 2012 Optimal pumping of orbital entanglement with single-particle emitters *Phys. Rev.* B **85** 081108
[21] Hofer P P and Büttiker M Dec 2013 Emission of time-bin entangled particles into helical edge states *Phys. Rev.* B **88** 241308
[22] Vyshnevyy A A, Lebedev A V, Lesovik G B and Blatter G 2013 Two-particle entanglement in capacitively coupled Mach–Zehnder interferometers *Phys. Rev.* B **87** 165302
[23] Dasenbrook D and Flindt C 2015 Dynamical generation and detection of entanglement in neutral leviton pairs *Phys. Rev.* B **92** 161412
[24] Tan S M, Walls D F and Collett M J 1991 Nonlocality of a single photon *Phys. Rev. Lett.* **66** 252
[25] Hardy L 1994 Nonlocality of a single photon revisited *Phys. Rev. Lett.* **73** 2279
[26] Greenberger D M, Horne M A and Zeilinger A 1995 Nonlocality of a single photon? *Phys. Rev. Lett.* **75** 2064
[27] Vaidman L 1995 Nonlocality of a single photon revisited again *Phys. Rev. Lett.* **75** 2063
[28] Hardy L 1995 Hardy replies *Phys. Rev. Lett.* **75** 2065
[29] Wiseman H M and Vaccaro J A 2003 Entanglement of indistinguishable particles shared between two parties *Phys. Rev. Lett.* **91** 097902
[30] Bartlett S D, Rudolph T and Spekkens R W 2007 Reference frames, superselection rules, and quantum information *Rev. Mod. Phys.* **79** 555
[31] van Enk S J 2005 Single-particle entanglement *Phys. Rev.* A **72** 064306
[32] Sangouard N and Zbinden H 2012 What are single photons good for? *J. Mod. Opt.* **59** 1458
[33] Lebedev A V, Blatter G, Beenakker C W J and Lesovik G B 2004 Entanglement in mesoscopic structures: role of projection *Phys. Rev.* B **69** 235312
[34] Wiseman H M, Bartlett S D and Vaccaro J A 2004 Ferreting out the fluffy bunnies: entanglement constrained by generalized superselection rules *Laser Spectroscopy* vol 1 (Singapore: World Scientific) p 307
[35] Samuelsson P, Sukhorukov E V and Büttiker M 2005 Quasi-particle entanglement: redefinition of the vacuum and reduced density matrix approach *New J. Phys.* **7** 176
[36] Giovannetti V, Frustaglia D, Taddei F and Fazio R 2007 Characterizing electron entanglement in multiterminal mesoscopic conductors *Phys. Rev.* B **75** 241305
[37] Sherkunov Y, Zhang J, d'Ambrumenil N and Muzykantskii B 2009 Optimal electron entangler and single-electron source at low temperatures *Phys. Rev.* B **80** 041313
[38] Clauser J F, Horne M A, Shimony A and Holt R A 1969 Proposed experiment to test local hidden-variable theories *Phys. Rev. Lett.* **23** 880
[39] Friis N 2016 Reasonable Fermionic quantum information theories require relativity *New J. Phys.* **18** 033014
[40] Amosov G G and Filippov S N 2015 Spectral properties of reduced Fermionic density operators and parity superselection rule (arXiv:1512.01828)
[41] Schuch N, Verstraete F and Cirac J I 2004 Nonlocal resources in the presence of superselection rules *Phys. Rev. Lett.* **92** 087904
[42] Sangouard N, Simon C, de Riedmatten H and Gisin N 2011 Quantum repeaters based on atomic ensembles and linear optics *Rev. Mod. Phys.* **83** 33
[43] Babichev S A, Appel J and Lvovsky A I 2004 Homodyne tomography characterization and nonlocality of a dual-mode optical qubit *Phys. Rev. Lett.* **92** 193601
[44] Fuwa M, Takeda S, Zwierz M, Wiseman H M and Furusawa A 2015 Experimental proof of nonlocal wavefunction collapse for a single particle using homodyne measurements *Nat. Commun.* **6** 6665
[45] Hessmo B, Usachev P, Heydari H and Björk G 2004 Experimental demonstration of single photon nonlocality *Phys. Rev. Lett.* **92** 180401
[46] Monteiro F, Caprara Vivoli V, Guerreiro T, Martin A, Bancal J-D, Zbinden H, Thew R T and Sangouard N 2015 Revealing genuine optical-path entanglement *Phys. Rev. Lett.* **114** 170504
[47] Brask J B, Chaves R and Brunner N 2013 Testing nonlocality of a single photon without a shared reference frame *Phys. Rev.* A **88** 012111
[48] Vaccaro J A, Anselmi F and Wiseman H M 2003 Entanglement of identical particles and reference phase uncertainty *Int. J. Quantum Inf.* **1** 427
[49] Ashhab S, Maruyama K and Nori F 2007 Detecting mode entanglement: the role of coherent states, superselection rules, and particle statistics *Phys. Rev.* A **76** 052113
[50] Thalineau R, Wieck A D, Bäuerle C and Meunier T 2014 Using a two-electron spin qubit to detect electrons flying above the Fermi sea (arXiv:1403.7770)
[51] Pedersen M H and Büttiker M 1998 Scattering theory of photon-assisted electron transport *Phys. Rev.* B **58** 12993
[52] Moskalets M and Büttiker M 2002 Floquet scattering theory of quantum pumps *Phys. Rev.* B **66** 205320
[53] Moskalets M 2011 *Scattering Matrix Approach to Non-Stationary Quantum Transport* (London: Imperial College Press)
[54] Dubois J, Jullien T, Grenier C, Degiovanni P, Roulleau P and Glattli D C 2013 Integer and fractional charge Lorentzian voltage pulses analyzed in the framework of photon-assisted shot noise *Phys. Rev.* B **88** 085301
[55] Levitov L S, Lee H and Lesovik G B 1996 Electron counting statistics and coherent states of electric current *J. Math. Phys.* **37** 4845
[56] Ivanov D A, Lee H W and Levitov L S 1997 Coherent states of alternating current *Phys. Rev.* B **56** 6839–50
[57] Lebedev A V, Lesovik G B and Blatter G 2005 Generating spin-entangled electron pairs in normal conductors using voltage pulses *Phys. Rev.* B **72** 245314
[58] Keeling J, Klich I and Levitov L S 2006 Minimal excitation states of electrons in one-dimensional wires *Phys. Rev. Lett.* **97** 116403
[59] Blanter Ya M and Büttiker M 2000 Shot noise in mesoscopic conductors *Phys. Rep.* **336** 1
[60] Samuelsson P, Neder I and Büttiker M 2009 Entanglement at finite temperatures in the electronic two-particle interferometer *Phys. Scr.* **T137** 014023
[61] Roulleau P, Portier F, Glattli D C, Roche P, Cavanna A, Faini G, Gennser U and Mailly D Oct 2007 Finite bias visibility of the electronic Mach–Zehnder interferometer *Phys. Rev.* B **76** 161309

Paper K

Anonymous Quantum Nonlocality

Physical Review Letters **113**, 130401 (2014)

Yeong-Cherng Liang, Florian John Curchod, Joseph Bowles, and Nicolas Gisin

Anonymous Quantum Nonlocality

Yeong-Cherng Liang,[1,*] Florian John Curchod,[2,3,†] Joseph Bowles,[4] and Nicolas Gisin[3]
[1]Institute for Theoretical Physics, ETH Zurich, 8093 Zurich, Switzerland
[2]ICFO–Institut de Ciències Fotòniques, 08860 Castelldefels, Barcelona, Spain
[3]Group of Applied Physics, University of Geneva, CH-1211 Geneva 4, Switzerland
[4]Department of Theoretical Physics, University of Geneva, 1211 Geneva, Switzerland
(Received 19 May 2014; published 23 September 2014)

We investigate the phenomenon of *anonymous quantum nonlocality*, which refers to the existence of multipartite quantum correlations that are not local in the sense of being Bell-inequality-violating but where the nonlocality is—due to its biseparability with respect to *all* bipartitions—seemingly nowhere to be found. Such correlations can be produced by the nonlocal collaboration involving definite subset(s) of parties but to an outsider, the identity of these nonlocally correlated parties is completely anonymous. For all $n \geq 3$, we present an example of an n-partite quantum correlation exhibiting anonymous nonlocality derived from the n-partite Greenberger-Horne-Zeilinger state. An explicit biseparable decomposition of these correlations is provided for any partitioning of the n parties into two groups. Two applications of these anonymous Greenberger-Horne-Zeilinger correlations in the device-independent setting are discussed: multipartite secret sharing between any two groups of parties and bipartite quantum key distribution that is robust against nearly arbitrary leakage of information.

DOI: 10.1103/PhysRevLett.113.130401 PACS numbers: 03.65.Ud, 03.67.Dd, 03.67.Mn

Quantum correlations that violate a Bell inequality [1], a constraint first derived in the studies of local-hidden-variable theories, were initially perceived only as a counterintuitive feature with no classical analog. With the discovery of quantum information science, these intriguing correlations have taken the new role as a resource. For instance, in nonlocal games [2], the presence of a Bell-inequality-violating (hereafter, referred to as nonlocal) correlation signifies the usage of strategies that cannot be achieved using only shared randomness. They are also an indispensable resource in quantum information and communication tasks such as the reduction of communication complexity [3], the distribution of secret keys using untrusted devices [4,5], as well as the certification and expansion of randomness [6], etc. (see [7] for a review).

Thus far, prior studies of quantum nonlocality have focussed predominantly on the bipartite setup. However, as with quantum entanglement [8,9], correlations between measurement outcomes can exhibit a much richer structure in the multipartite setup. Consider a multipartite Bell-type experiment with the ith party's choice of measurement setting (input) denoted by $x_i = 0, 1$ and the corresponding outcome (output) by $a_i = \pm 1$. Already in the tripartite setting [10], quantum mechanics allows for correlations—a collection of joint conditional probability distributions $\vec{P} = \{P(\vec{a}|\vec{x})\} = \{P(a_1a_2a_3|x_1x_2x_3)\}$—that cannot be reproduced even when subsets of the parties are allowed to share some nonlocal resource $\mathcal{R}$ [11,12]. (Throughout, we focus on nonlocal resources $\mathcal{R}$ that respect the no-signaling conditions [13,14] which dictate, e.g., that each marginal distribution of $P_i^{\mathcal{R}}(a_ja_k|x_jx_k)$ can be defined independent of the input of the other party.) Such genuinely tripartite nonlocal correlations are, by definition, those that *cannot* be written in the so-called biseparable form:

$$\begin{aligned} P(\vec{a}|\vec{x}) \neq & \sum_{\nu} p_\nu P_\nu(a_1|x_1) P_\nu^{\mathcal{R}}(a_2a_3|x_2x_3) \\ & + \sum_{\mu} p_\mu P_\mu(a_2|x_2) P_\mu^{\mathcal{R}}(a_1a_3|x_1x_3) \\ & + \sum_{\lambda} p_\lambda P_\lambda(a_3|x_3) P_\lambda^{\mathcal{R}}(a_1a_2|x_1x_2), \end{aligned} \tag{1}$$

where $\sum_{i\in\{\lambda,\mu,\nu\}} p_i = 1$, $p_i \geq 0$ for all $i \in \{\lambda, \mu, \nu\}$ and $P_i^{\mathcal{R}}(a_ja_k|x_jx_k)$ is *any* bipartite correlation allowed by the resource $\mathcal{R}$ [11,12]. In a Bell-type experiment, the presence of genuine multipartite nonlocality [15–19] is a manifestation of genuine multipartite entanglement [8]; it thus facilitates the detection of the latter in a *device-independent* manner, i.e., without relying on any assumption about the measurements being performed nor the dimension of the underlying Hilbert space. (It is also possible to detect genuine multipartite entanglement in a device-independent manner without the detection of genuine multipartite nonlocality; see [18,20]). In contrast, correlations that are biseparable, cf. Eq. (1), receive almost no attention. Apart from being a tool in the derivation of Bell-type inequalities for genuine multipartite nonlocality, is this kind of correlation interesting in its own right? Here, we answer this question affirmatively via the phenomenon of anonymous nonlocality (ANL), an intriguing feature that is only present in biseparable correlations. We will also provide evidence showing that ANL can be a powerful resource, allowing one to design device-independent quantum cryptographic

protocols that can guard against a particular kind of attack by any postquantum, but nonsignaling adversary.

Biseparable correlations and anonymous nonlocality.— To appreciate the peculiarity manifested by ANL, let us start by considering the simplest, tripartite scenario. Clearly, among the subsets of correlations that can be decomposed in the form of the right-hand side of Eq. (1) are those that satisfy

$$P(\vec{a}|\vec{x}) = \sum_{\nu} p_\nu P_\nu(a_1|x_1) P^{\mathcal{R}}_\nu(a_2a_3|x_2x_3), \tag{2a}$$

$$= \sum_{\mu} p_\mu P_\mu(a_2|x_2) P^{\mathcal{R}}_\mu(a_1a_3|x_1x_3), \tag{2b}$$

$$= \sum_{\lambda} p_\lambda P_\lambda(a_3|x_3) P^{\mathcal{R}}_\lambda(a_1a_2|x_1x_2), \tag{2c}$$

where p_ν, p_μ, $p_\lambda \geq 0$ for all ν, μ, and λ, but in contrast with Eq. (1), we now have $\sum_\nu p_\nu = \sum_\mu p_\mu = \sum_\lambda p_\lambda = 1$. Equations (2a)–(2c) imply that the correlation can be produced *without* having any nonlocal collaboration between the isolated party and the remaining two parties (as a group). Naively, one may thus expect that all correlations satisfying these equations must also be local in the sense of being non-Bell-inequality-violating (henceforth abbreviated as local). However, there exist [21] quantum correlations that satisfy Eqs. (2a)–(2c) as well as

$$P(\vec{a}|\vec{x}) \neq \sum_{\theta} p_\theta P_\theta(a_1|x_1) P_\theta(a_2|x_2) P_\theta(a_3|x_3), \tag{2d}$$

for *any* conditional distributions $P_\theta(a_i|x_i)$ and *any* normalized weights p_θ. In other words, $\vec{P}$ satisfying Eq. (2) is nonlocal but this nonlocality is (i) not genuinely tripartite (it is biseparable), (ii) not attributable to any of the two-partite marginals [Eqs. (2a)–(2c) imply that all marginals are local], and (iii) not attributable to any bipartition of the three parties. The nonlocality present in *any* correlations satisfying Eq. (2) is thus in some sense nowhere to be found.

We now provide a very simple example of a correlation satisfying Eq. (2), and more generally the property of being (1) nonlocal and (2) biseparable with respect to all bipartitions in an arbitrary n-partite scenario. Consider the n-partite Greenberger-Horne-Zeilinger (GHZ) state [22] $|\mathrm{GHZ}\rangle_n = 1/\sqrt{2}(|0\rangle^{\otimes n} + |1\rangle^{\otimes n})$ and the local measurement of σ_x and σ_y. The resulting correlation is

$$P(\vec{a}|\vec{x}) = P^n_{\mathrm{GHZ}}(\vec{a}|\vec{x}) = \frac{1}{2^n}\left[1 + \cos\left(\mathbf{x}\frac{\pi}{2}\right)\prod_{i=1}^{n} a_i\right], \tag{3}$$

where $\mathbf{x} = \sum_i x_i$, and we have identified $x_i = 0\,(1)$ as the σ_x (σ_y) measurement [see, e.g., Eq. (23) of [23]]. For *all* $n \geq 3$, we show [24] that n-partite correlations of the form of Eq. (3) admit a biseparable decomposition with respect to *any* partitioning of the n parties into two groups. Specifically, for $n = 3$, this decomposition, cf. Eq. (2a), involves $p_\nu = \frac{1}{4}$ for all ν, $P_\nu(a_1|x_1) = 0, 1$ and $P^{\mathcal{R}}_\nu(a_2a_3|x_2x_3)$ is the correlation associated with the so-called Popescu-Rohrlich (PR) box [13]—a hypothetical, stronger-than-quantum, but nonsignaling resource. [In the tripartite scenario, the biseparability of the GHZ correlation was also discovered independently in [25] (see also [26])]. To see that these correlations are nonlocal, it suffices to note [24] that Eq. (3) violates the Mermin-Ardehali-Belinskiĭ-Klyshko-Bell inequality [27,28] (even maximally [29] for all odd $n \geq 3$).

Consider now an alternative way to understand the nonlocality associated with Eq. (2). Operationally, Eq. (2c) implies that $\vec{P}$ can be produced by, e.g., party 1 signaling classically to party 2, and all parties responding according to the information that they received and some predefined strategy λ. By symmetry of Eqs. (2a)–(2c), the same can be achieved by having only nonlocal collaboration between any two out of the three parties. Thus, while the correlation can be produced by having only a definite subset of parties collaborating nonlocally, the identity of these nonlocally collaborating parties is anonymous to an outsider who only has access to $\vec{P}$. Indeed, even if an outsider is given the promise that a fixed subset of the parties have collaborated nonlocally, it is impossible for him to tell if, say, party 1 and 2 have collaborated nonlocally in generating $\vec{P}$. Importantly, the anonymity present in these correlations differs from the case where a classical mixture of the different bipartitions is necessary, cf. Fig. 1 (see [30,31] for examples of such classical anonymity). In this latter case, it is indeed possible to identify the parties that must have collaborated nonlocally, even though this identification is generally not possible at any single run of the experiment.

As noted above, for all $n \geq 3$, the GHZ correlations of Eq. (3) are nonlocal but can nevertheless be produced by splitting the parties into any two groups, and disallowing any nonlocal collaboration between these groups. Thus, the

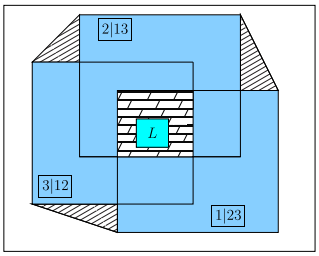

FIG. 1 (color online). Schematic representation of the various sets of tripartite correlations. Correlations biseparable with respect to party i in one group and parties j and k in the other lie in the (light blue) rectangle labeled by "$i|jk$". The convex hull of the three biseparable sets $i|jk$, where $i, j, k \in \{1, 2, 3\}$ is represented by the filled convex region and gives correlations decomposable as the right-hand side of Eq. (1). The blank region between the outermost box and the filled convex region represents correlations that are genuinely tripartite nonlocal. Intersection of the three biseparable subsets $i|jk$ gives correlations satisfying Eqs. (2a)–(2c); its subset featuring ANL is the tiled region while local correlations lie in the (cyan) rectangle L. Hatched regions represent biseparable correlations where classical mixture of different bipartitions is necessary for their production.

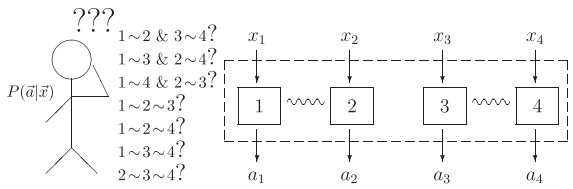

FIG. 2. ANL in the quadripartite scenario. Each participating party is abstractly represented by a box labeled by the party number. The correlations were produced by having parties 1 and 2, as well as 3 and 4 collaborated nonlocally (symbolized by ∼). To an outsider who only has access to $\vec{a}$ and $\vec{x}$, even if one is given the promise that the correlations were produced by the four parties separated into two fixed groups, it is impossible to tell which actual partitioning of the parties generated these correlations.

anonymity present in these correlations is even more striking in the $n > 3$ scenarios: not only are the groups of parties sharing $\mathcal{R}$ unidentifiable in an unambiguous manner, even the size of the groups are also not identifiable (see Fig. 2). For example, when $n = 4$, the correlations satisfy

$$P(\vec{a}|\vec{x}) = \sum_{\lambda_1} q_{\lambda_1} P_{\lambda_1}(a_1|x_1) P^{\mathcal{R}}_{\lambda_1}(a_2 a_3 a_4|x_2 x_3 x_4), \tag{4a}$$

$$= \sum_{\lambda_2} q_{\lambda_2} P_{\lambda_2}(a_2|x_2) P^{\mathcal{R}}_{\lambda_2}(a_1 a_3 a_4|x_1 x_3 x_4), \tag{4b}$$

$$= \cdots, \tag{4c}$$

$$= \sum_{\mu_3} q_{\mu_3} P^{\mathcal{R}}_{\mu_3}(a_1 a_4|x_1 x_4) P^{\mathcal{R}}_{\mu_3}(a_2 a_3|x_2 x_3), \tag{4d}$$

$$P(\vec{a}|\vec{x}) \neq \sum_{\theta} q_\theta \prod_{i=1}^{4} P_\theta(a_i|x_i), \tag{4e}$$

where $\sum_{\lambda_i} p_{\lambda_i} = \sum_{\mu_j} p_{\mu_j} = 1$, $p_i \geq 0$ for all i, j, and "$\cdots$" indicates other possible biseparable decompositions that have been omitted. From Eq. (4), we see that the quadripartite GHZ correlation could have been produced by having any three parties collaborating nonlocally, or any two groups of two parties collaborating nonlocally within each group. From the correlation itself, it is simply impossible to distinguish these possibilities apart (Fig. 2).

Let us now briefly comment on the relationship between ANL and multipartite entanglement. Clearly, one expects that there must also be features analogous to ANL in the studies of multipartite entanglement. Indeed, the first of such examples dates back to the three-qubit bound entangled [32] SHIFT state [33] where its entanglement was dubbed delocalized [34] since it is separable with respect to all bipartitions, yet not fully separable. A more recent example [21] involves a three-qubit bound entangled state which even violates a Bell inequality, thus giving also an example of anonymous quantum correlation. An important difference between their example and the tripartite case of our GHZ example is that their correlation can be produced by a biseparable tripartite entangled state whereas ours necessarily requires a genuinely tripartite entangled state. More generally, for all odd $n \geq 3$, we show [24] that the correlations of Eq. (3) can only be produced by genuinely n-partite entangled state. Our examples thus show that the generation of ANL does not require delocalized entanglement.

Perfect correlations with uniform marginals.—From Eq. (3), we see that whenever an odd number of parties measure in the σ_y basis, the product of outcomes $\prod_i a_i$ gives ± 1 with equal probability; otherwise, it is either perfectly correlated or perfectly anticorrelated. Moreover, it follows from Eq. (3) that all marginal distributions of these correlations are uniformly random. Next, we present two quantum cryptography protocols that exploit these strong but anonymous correlations.

Application I: Multipartite secret sharing (MSS).—Imagine that n parties wanted to share a secret message between any two complementary subgroups as they desire, i.e., between any subgroup of k parties ($k \leq n - 1$) and the subgroup formed by the remaining parties. Suppose, moreover, that the shared secret is to be recovered by these subgroups only when all parties within each group collaborate (so that it is unnecessary to trust all parties within each group). A possibility to achieve this consists of (i) the n parties share (many copies of) $|\mathrm{GHZ}_n\rangle$, (ii) each party randomly measures either the σ_x or the σ_y observable, (iii) the n parties are randomly separated into two groups and all parties assigned to the same group collaborate to compare their inputs and outputs, (iv) both groups announce their sum of inputs, (v) parties in the same group compute the product of their measurement outcome and deduce, using Eq. (3), the shared secret bit upon learning the sum of inputs of the other group, (vi) parties in one group use the shared secret keys to encrypt the message and send it to the other.

In the device-independent setting, security analysis is carried out by treating each physical subsystem together with their measurement device as a black box; conclusions are drawn directly from the measurement statistics. Indeed, the above protocol does not rely on the assumption of a GHZ state nor the particular measurements being performed, but rather the strong correlation present in Eq. (3)—for the right combination of inputs, the product of outputs are perfectly (anti-) correlated. (This happens in half the cases. In the other cases, the correlation is useless for key generation.) Thus, the protocol essentially works by first distributing the correlated data needed to establish the secret keys, and performing the secret sharing [35] between any two complementary subgroups of the n participating parties as they deem fit. Since the product of outcomes for each group is uniformly random, the protocol is secure against cheating by any dishonest parties within the group; no one can retrieve the shared key without collaborating with everyone else within the same group. What about eavesdropping by an external, postquantum but nonsignaling adversary Eve?

Since the GHZ correlations of Eq. (3) are biseparable, a naive attack by Eve may consist in preparing for the n parties the biseparable, nonsignaling boxes that reproduces exactly Eq. (3). For instance, in the tripartite case, in accordance with the biseparable decomposition, she would

prepare with equal probability four different versions of a deterministic box for one of the parties, and, correspondingly, four different versions of a PR box for the remaining two parties. If the decomposition that she chooses matches exactly the way the parties are separated into two groups, then after step (iv), she learns exactly the key and hence the message shared by these parties. In this case, the product of outcomes for each group is a deterministic function (of the sum of inputs) known to Eve. The secret sharing protocol of Hillery, Bužek, and Berthiaume [36] is thus insecure against this kind of attack by a nonsignaling adversary. However, as the grouping is decided only after the measurement phase, she can guess the bipartition correctly only with a chance of $\frac{1}{3}$ in the tripartite case, and more generally $(2^{n-1}-1)^{-1}$ in the n-partite scenario. Evidently, this guessing probability rapidly approaches 0 as n increases, making it extremely difficult for Eve to succeed with this eavesdropping strategy for large n.

Application II: Bipartite leakage-resilient QKD.—Next, let us describe a quantum key distribution (QKD) protocol between two parties, A and B, which is as leakage resilient [37] as one could hope for. The protocol consists of (i) preparation of many copies of $|\mathrm{GHZ}_n\rangle$, (ii) for each of these n-partite systems, a *randomly chosen subset*, say, k of the n subsystems are distributed to A, while the remaining $n-k$ subsystems are distributed to B, (iii) for each of these subsystems, A and B randomly measure σ_x or σ_y, (iv) both parties announce their sum of inputs, (v) for each n-partite system distributed from the source, A and B compute the product of their local measurement outcomes and deduce, using Eq. (3), the shared secret bit upon learning the sum of inputs of the other party.

As with the MSS protocol described above, the secret key is established through the perfect (anti-) correlation present in the product of the outputs. Moreover, the gist of the protocol only relies on the correlation given by Eq. (3), rather than the actual state and measurement giving rise to this correlation, rendering the protocol ideal for device-independent analysis. However, in contrast with usual device-independent cryptography where leakage of information is not allowed, the above protocol is as leakage resilient as one can hope for—the adversary Eve can certainly recover the secret key if all the output bits from either party leak to her, but if she misses merely one output bit from each party, the additional information that she gains from the leakage cannot improve her guess of the secret key. Now, if we assume that Eve has no control over how the subsystems are distributed in step (ii) [38], but otherwise only constrained by the nonsignaling principle, then as with the MSS protocol, for n sufficiently large, her advantage of preparing some biseparable, nonsignaling boxes for A and B is minimal.

Discussion.—Let us now comment on some possible directions for future research. Clearly, we have only provided intuitions on why the protocols proposed above may be secure even in a device-independent setting. For odd $n \geq 3$, since the GHZ correlations violate the Mermin-Bell inequality maximally (see [24]), the result of Franz *et al.* [39] implies that these correlations are necessarily monogamous with respect to any potential quantum eavesdropper. This strongly suggests that if we assume an independent and identically distributed scenario, a formal security proof of these protocols against a quantum adversary may be given even in the case with noisy correlations (because of the noise robustness of the Mermin-Bell violation of $\vec{P}^n_{\mathrm{GHZ}}(\vec{a}|\vec{x})$, the ANL of $\vec{P}^n_{\mathrm{GHZ}}(\vec{a}|\vec{x})$ is also extremely robust to noise), and in a device-independent setting. Evidently, a security proof without this assumption is even more desirable, and a possible path towards this is to prove that the protocols are even secure against an adversary that is only constrained by the nonsignaling principle [13]. Our arguments as to why the protocols are not immediately susceptible to a straightforward attack by such an eavesdropper, despite the fact that the correlations are biseparable, is an evidence pointing in this direction.

For leakage-resilient QKD, one could also imagine, instead of the above protocol, doing an existing QKD protocol many times in parallel and then using the XOR of the secret key bits to generate the final secret key. Although such a protocol requires many more qubits to establish the final secret key, it can clearly offer a high level of leakage resilience. How would such a protocol perform compared with the above protocol based on $|\mathrm{GHZ}_n\rangle$? This certainly deserves some further investigation.

Coming back to ANL itself, let us note that the requirement of (1) nonlocality and (2) biseparability with respect to all bipartitions may arguably not, by themselves, imply that an outsider cannot attribute unambiguously the nonlocality to any definite subset(s) of the n parties. For instance, one may start with the tripartite GHZ correlation $P^3_{\mathrm{GHZ}}(\vec{a}|\vec{x})$, cf. Eq. (3), and trivially construct an example $P' = P^3_{\mathrm{GHZ}}(\vec{a}|\vec{x}) \prod_{i=4}^{n} P(a_i|x_i)$ for arbitrary n parties by introducing parties that are uncorrelated with the first three. While such an n-partite correlation P' indeed satisfies the two requirements stated above, one can unambiguously attribute the nonlocality present only to the three parties that give rise to $P^3_{\mathrm{GHZ}}(\vec{a}|\vec{x})$. Note, however, that such an identification is incomplete since the production of such a biseparable correlation only requires the nonlocal collaboration between two parties, and it is still impossible for an outsider to determine which two parties have collaborated nonlocally in producing the given correlation (Fig. 1 and Fig. 2). A more precise definition of ANL may thus require also a specification of the extent (size) of the nonlocal resource needed in producing the given correlation, a task that shall be pursued elsewhere [31]. For our GHZ examples, except for the cases where n is even with $n/2$ odd, it can be shown [24] using the result of [16] that the correlations of Eq. (3) are not triseparable, i.e., not producible by a partitioning of the parties into three groups (where only parties within the same group are allowed to collaborate nonlocally). Hence, the generation of these correlations indeed requires the nonlocal collaboration of at least $\lceil n/2 \rceil$ parties in one group; an analogous statement for the remaining cases would be desirable.

We are grateful to David Jennings for suggesting the terminology "Anonymous Nonlocality," and to Rotem Arnon-Friedman, Jean-Daniel Bancal, Nicolas Brunner, Stefano Pironio, as well as Renato Renner for stimulating discussions. This work is supported by the Swiss National Science Foundation Grant No. PP00P2 138917, SEFRI (COST action MP1006), Swiss NCCR "Quantum Science and Technology," and the CHIST-ERA DIQIP. F. J. C. acknowledges support from the John Templeton Foundation. Y. C. L. and F. J. C. contributed equally towards this work.

*yliang@phys.ethz.ch
†florian.curchod@icfo.es

[1] J. S. Bell, Physics (NY) **1**, 195 (1964).
[2] R. Cleve, P. Høyer, B. Toner, and J. Watrous, in *Proceedings of the 19th IEEE Annual Conference on Computational Complexity, Amherst, MA, 2004* (IEEE Conference Proceedings, New York, 2004), p. 236.
[3] H. Buhrman, R. Cleve, and W. van Dam, SIAM J. Comput. **30**, 1829 (2001); Č. Brukner, M. Żukowski, J.-W. Pan, and A. Zeilinger, Phys. Rev. Lett. **92**, 127901 (2004).
[4] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991); J. Barrett, L. Hardy, and A. Kent, Phys. Rev. Lett. **95**, 010503 (2005); A. Acín, N. Gisin, and L. Masanes, Phys. Rev. Lett. **97**, 120405 (2006); A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Phys. Rev. Lett. **98**, 230501 (2007).
[5] A. Acín, S. Massar, and S. Pironio, New J. Phys. **8**, 126 (2006).
[6] R. Colbeck, Ph.D. Dissertation, University of Cambridge, 2006, arXiv:0911.3814; 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 (London) **464**, 1021 (2010).
[7] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Rev. Mod. Phys. **86**, 419 (2014).
[8] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009); O. Gühne and G. Tóth, Phys. Rep. **474**, 1 (2009).
[9] M. P. Seevinck and J. Uffink, Phys. Rev. A **78**, 032101 (2008); S. Szalay and Z. Kökényesi, Phys. Rev. A **86**, 032341 (2012).
[10] G. Svetlichny, Phys. Rev. D **35**, 3066 (1987).
[11] R. Gallego, L. E. Würflinger, A. Acín, and M. Navascués, Phys. Rev. Lett. **109**, 070401 (2012).
[12] J.-D. Bancal, J. Barrett, N. Gisin, and S. Pironio, Phys. Rev. A **88**, 014102 (2013).
[13] S. Popescu and D. Rohrlich, Found. Phys. **24**, 379 (1994).
[14] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts, Phys. Rev. A **71**, 022101 (2005).
[15] D. Collins, N. Gisin, S. Popescu, D. Roberts, and V. Scarani, Phys. Rev. Lett. **88**, 170405 (2002); M. Seevinck and G. Svetlichny, Phys. Rev. Lett. **89**, 060401 (2002); N. S. Jones, N. Linden, and S. Massar, Phys. Rev. A **71**, 042329 (2005); J.-D. Bancal, N. Brunner, N. Gisin, and Y.-C. Liang, Phys. Rev. Lett. **106**, 020405 (2011); J.-L. Chen, D.-L. Deng, H.-Y. Su, C. Wu, and C. H. Oh, Phys. Rev. A **83**, 022316 (2011).
[16] J.-D. Bancal, C. Branciard, N. Gisin, and S. Pironio, Phys. Rev. Lett. **103**, 090503 (2009).
[17] L. Aolita, R. Gallego, A. Cabello, and A. Acín, Phys. Rev. Lett. **108**, 100401 (2012).
[18] J.-D. Bancal, C. Branciard, N. Brunner, N. Gisin, and Y.-C. Liang, J. Phys. A **45**, 125301 (2012).
[19] Q. Chen, S. Yu, C. Zhang, C. H. Lai, and C. H. Oh, Phys. Rev. Lett. **112**, 140404 (2014).
[20] J.-D. Bancal, N. Gisin, Y.-C. Liang, and S. Pironio, Phys. Rev. Lett. **106**, 250404 (2011); K. F. Pál and T. Vértesi, Phys. Rev. A **83**, 062123 (2011); T. Moroder, J.-D. Bancal, Y.-C. Liang, M. Hofmann, and O. Gühne, Phys. Rev. Lett. **111**, 030501 (2013).
[21] T. Vértesi and N. Brunner, Phys. Rev. Lett. **108**, 030403 (2012).
[22] D. M. Greenberger, M. A. Horne, and A. Zeilinger, in *Bell's Theorem, Quantum Theory, and Conceptions of the Universe*, edited by M. Kafatos (Kluwer, Dordrecht, 1989), p. 69; N. D. Mermin, Phys. Rev. Lett. **65**, 3373 (1990).
[23] J. J. Wallman, Y.-C. Liang, and S. D. Bartlett, Phys. Rev. A **83**, 022110 (2011).
[24] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.113.130401 for a detailed derivation of all the results presented in this Letter.
[25] S. Pironio, J.-D. Bancal, and V. Scarani, J. Phys. A **44**, 065303 (2011).
[26] J. L. Cereceda, Phys. Rev. A **66**, 024102 (2002); P. Mitchell, S. Popescu, and D. Roberts, Phys. Rev. A **70**, 060101(R) (2004).
[27] N. D. Mermin, Phys. Rev. Lett. **65**, 1838 (1990); M. Ardehali, Phys. Rev. A **46**, 5375 (1992).
[28] S. M. Roy and V. Singh, Phys. Rev. Lett. **67**, 2761 (1991); A. V. Belinskiĭ and D. N. Klyshko, Phys. Usp. **36** 653 (1993); N. Gisin and H. Bechmann-Pasquinucci, Phys. Lett. A **246**, 1 (1998).
[29] R. F. Werner and M. M. Wolf, Phys. Rev. A **61**, 062102 (2000).
[30] F. J. Curchod, Y.-C. Liang, and N. Gisin, arXiv:1310.7598 [J. Phys. A (to be published)].
[31] F. J. Curchod, N. Gisin, and Y.-C. Liang (in preparation).
[32] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **80**, 5239 (1998).
[33] C. H. Bennett, D. P. DiVincenzo, T. Mor, P. W. Shor, J. A. Smolin, and B. M. Terhal, Phys. Rev. Lett. **82**, 5385 (1999).
[34] D. P. DiVincenzo, T. Mor, P. W. Shor, J. A. Smolin, and B. M. Terhal, Commun. Math. Phys. **238**, 379 (2003).
[35] B. Schneier, *Applied Cryptography* (Wiley, New York, 1996), p. 70; J. Gruska, *Foundations of Computing* (Thomson Computer Press, London, 1997), p. 504.
[36] M. Hillery, V. Bužek, and A. Berthiaume, Phys. Rev. A **59**, 1829 (1999).
[37] S. Dziembowski and K. Pietrzak, in *Proceedings of IEEE 49th Annual IEEE Symposium on Foundations of Computer Science*, p. 293; F.-X. Standaert, O. Pereira, Y. Yu, J.-J. Quisquater, M. Yung, and E. Oswald, in *Towards Hardware-Intrinsic Security* (Springer, Berlin, 2010), p. 99; F. G. Lacerda, J. M. Rennes, and R. Renner, arXiv:1404.7516v1.
[38] Instead of this assumption, A and B can employ additional measurement settings, cf. [5], to certify that the overall correlations indeed exhibit genuine multipartite nonlocality and they are then again not susceptible to such an attack.
[39] T. Franz, F. Furrer, and R. F. Werner, Phys. Rev. Lett. **106**, 250502 (2011).

Supplementary Mateirals for "Anonymous Quantum Nonlocality"

Yeong-Cherng Liang,[1] Florian John Curchod,[2,3] Joseph Bowles,[4] and Nicolas Gisin[3]

[1] *Institute for Theoretical Physics, ETH Zurich, 8093 Zurich, Switzerland.*
[2] *ICFO–Institut de Ciències Fotòniques, 08860 Castelldefels (Barcelona), Spain.*
[3] *Group of Applied Physics, University of Geneva, CH-1211 Geneva 4, Switzerland.*
[4] *Department of Theoretical Physics, University of Geneva, 1211 Geneva, Switzerland.*

Appendix A: An explicit biseparable decomposition of the n-partite GHZ correlations

For the n-partite Greenberger-Horne-Zeilinger (GHZ) [1] state and the situation where all parties measure either the 0th-observable σ_x or the 1st observable σ_y, the resulting correlation of Eq. (3) can be rewritten in terms of the *correlator*, i.e., the expectation value of the product of outcomes:[1]

$$E(\vec{x}) = \sum_{a'_1,a'_2,\ldots,a'_n=0,1} (-1)^{\sum_i a'_i} P(\vec{a}'|\vec{x}) = \cos\left(\mathbf{x}\frac{\pi}{2}\right) \quad (3)$$

where for conciseness of subsequent presentation we have used, instead, $a'_i = \frac{a_i+1}{2} = 0, 1$ to denote the output and as before, $\mathbf{x} = \sum_i x_i$ to denote the sum of inputs. Note that all the full n-partite correlators depend only on the parity of $\mathbf{x}$ and $\mathbf{x}/2$ whereas all the marginal correlators vanish.

We now move on to the proof that the correlation (3) is biseparable with respect to *all* bipartitions whenever parties in each group are allowed to share arbitrary post-quantum but non-signaling [3, 4] (NS) resources, while parties in different groups can only be correlated through shared randomness. Note that the biseparability of Eq. (3) under the NS constraint implies that if parties in the same group are allowed to share a stronger resource, such as a Svetlichny resource [5], or some other one-way signaling resource discussed in [6, 7], the correlation must remain biseparable.

For clarity we first give a proof for the 3-partite case. The extension to arbitrary numbers of parties follows by generalizing these ideas. We note that for the three party GHZ state, if we have $\mathbf{x} = 0$ modulo 2, the correlation can be written as:

$$a'_1 + a'_2 + a'_3 = \cos\left(\mathbf{x}\frac{\pi}{2}\right) \overset{\mathbf{x}=0,2}{=} x_1x_2 + x_3, \quad (A1)$$

where the addition is modulo 2. For the other choices of inputs with $\mathbf{x} = 1$ there is zero correlation, and all one and two party marginals are random for all inputs. We note 4 different strategies that are capable of reproducing the correlation (A1) for the case $\mathbf{x} = 0$:

(i) Parties 1 & 2 share a Popescu-Rohrlich (PR) box [3] $a'_1 + a'_2 = x_1x_2$ and the third party owns a deterministic box $a'_3 = x_3$.

(ii) Parties 1 & 2 share an 'anti' PR box $a'_1+a'_2 = x_1x_2+1$ and the third party owns a deterministic box $a'_3 = x_3+1$.
(iii) Parties 1 & 2 share an input flipped PR box $a'_1 + a'_2 = (x_1 + 1)(x_2 + 1)$ and party 3 owns the deterministic box $a'_3 = 1$.
(iv) Parties 1 & 2 share an 'anti' input flipped PR box $a'_1 + a'_2 = (x_1 + 1)(x_2 + 1) + 1$ and party 3 owns the deterministic box $a'_3 = 0$.

From the form of Eq. (A1) it is clear that strategy (i) and (ii) will give the correct full correlation. Strategies (iii) and (iv) follow from the fact that if $\mathbf{x} = 0$ we have $x_1x_2 + x_3 = x_1x_2 + x_1 + x_2 = (x_1 + 1)(x_2 + 1) + 1$. Hence for $\mathbf{x} = 0$ we always simulate the correct correlation. For the cases $\mathbf{x} = 1$, we note that now $x_1x_2 + x_3 = (x_1 + 1)(x_2 + 1)$ and so strategies (iii) and (iv) will give exactly the opposite correlation to strategies (i) and (ii) for these inputs. Mixing equally all four strategies therefore gives no correlation for these inputs. The mixture of the different PR box symmetries and deterministic strategies for party 3 then ensures that all 1 and two party marginals are random. Due to the symmetry of the GHZ correlation, the same construction will therefore work for any bipartition of the 3 parties.

Let us now give a proof for the n-partite case. Consider the four families of n-partite NS boxes, labeled by μ_1, μ_2, μ_3 and μ_4:

$$\begin{aligned}
P^n_{\mu_1}(\vec{a}'|\vec{x}) &= \frac{1}{2^{n-1}}\delta_{\sum_{i=1}^n a'_i - H^n_0(\vec{x}) - H^n_3(\vec{x}) \bmod 2},\\
P^n_{\mu_2}(\vec{a}'|\vec{x}) &= \frac{1}{2^{n-1}}\delta_{\sum_{i=1}^n a'_i - H^n_0(\vec{x}) - H^n_1(\vec{x}) \bmod 2},\\
P^n_{\mu_3}(\vec{a}'|\vec{x}) &= \frac{1}{2^{n-1}}\delta_{\sum_{i=1}^n a'_i - H^n_1(\vec{x}) - H^n_2(\vec{x}) \bmod 2},\\
P^n_{\mu_4}(\vec{a}'|\vec{x}) &= \frac{1}{2^{n-1}}\delta_{\sum_{i=1}^n a'_i - H^n_2(\vec{x}) - H^n_3(\vec{x}) \bmod 2},
\end{aligned} \quad (A2)$$

where $H^n_\ell(\vec{x}) = \sum_{j=0}^{\lfloor \frac{n-\ell}{4} \rfloor} F(4j + \ell, \vec{x})$,

$$F(k, \vec{x}) = \sum_G \prod_{i\in G} x_i \prod_{j\in G'} (x_j + 1) \quad (A3)$$

and the sum $\sum_G$ is over all $G \subseteq [n] = \{1, 2, \ldots, n\}$ with group size $|G| = k$, and G' is the complement of G in $[n]$. Essentially, each term involved in the summand in $F(k, \vec{x})$, and hence $H^n_\ell(\vec{x})$ defines a distinct combination of inputs $\vec{x} = \vec{x}'$ such that $H^n_\ell(\vec{x}') = 1 \bmod 2$, and hence making the outputs anti-correlated. For instance, $F(0, \vec{x})$ only makes a nontrivial combination to $H^n_0(\vec{x})$ if all the inputs x_i are 0.

From Eq. (A2), it is easy to verify that for all $1 \leq k \leq n - 1$, the k-partite marginals of $P^n_j(\vec{a}'|\vec{x})$ are $1/2^k$

[1] To arrive at this n-partite correlator, see, eg., Eq. (23) of [2].

and these correlations indeed define NS probability distributions. Moreover, from Eq. (A2) and these marginal distributions, one can show that these NS boxes give rise to vanishing marginal correlators and the following full n-partite correlators:

$$\begin{aligned} E(\vec{x})_{\mu_1} &= (-1)^{H_0^n(\vec{x})\oplus H_3^n(\vec{x})} = -E(\vec{x})_{\mu_3}, \\ E(\vec{x})_{\mu_2} &= (-1)^{H_0^n(\vec{x})\oplus H_1^n(\vec{x})} = -E(\vec{x})_{\mu_4}, \end{aligned} \tag{A4}$$

where in Eq. (A4), $\oplus$ denotes sum modulo 2 and in arriving at the second equality in each line, we have employed the identity $\sum_{j=0}^{n} F(j) = 1$ that holds for all n-bit strings $\vec{x}$.[2] To gain some intuition on these NS boxes, we note that for $n = 1$, the $\mu_{1/3}$ boxes correspond to the deterministic strategies $a' = x \oplus 1$ and $a' = x$ whereas the $\mu_{2/4}$ boxes correspond to the deterministic strategies $a' = 1$ and $a' = 0$. Similarly, for $n = 2$, the $\mu_{1/3}$ boxes correspond to the PR boxes defined by $a'_1+a'_2 = (x_1+1)(x_2+1)$ and $a'_1+a'_2 = (x_1+1)(x_2+1)\oplus 1$ whereas the $\mu_{2/4}$ boxes correspond to the PR boxes defined by $a'_1 + a'_2 = x_1x_2 \oplus 1$ and $a'_1 + a'_2 = x_1x_2$. For $n = 3$, all these NS boxes correspond to some version of NS box 46 described in [8]. It is conceivable that these boxes are extremal NS distributions for all n.

To reproduce the correlations given in Eq. (3) using biseparable $\mathcal{NS}$ resources with k parties in one group and the remaining $(n-k)$ parties in the other group, it suffices to consider an equal-weight mixture of the following four strategies:

1. The group of k parties share the k-partite version of the μ_1 box and the remaining parties share the $(n-k)$-partite version of the μ_2 box.

2. The group of k parties share the k-partite version of the μ_3 box and the remaining parties share the $(n-k)$-partite version of the μ_4 box.

3. The group of k parties share the k-partite version of the μ_2 box and the remaining parties share the $(n-k)$-partite version of the μ_1 box.

4. The group of k parties share the k-partite version of the μ_4 box and the remaining parties share the $(n-k)$-partite version of the μ_3 box.

For $n = 3$, the above strategy corresponds to a mixture of 4 different versions of the NS box 2 in [8]. In general, to verify that the above strategy indeed gives rise to Eq. (3), we first remark that each of these strategies also reproduces Eq. (3) for the case when $\sum_i x_i$ is even. To see this, we use the fact that NS box μ_1 gives anti-correlation (i.e., expectation value -1) only if either $\sum_i x_i/2$ or $(1 + \sum_i x_i)/2$ is even; NS box μ_2 gives anti-correlation only if $\sum_i x_i/2$ is even or $(1+\sum_i x_i)/2$ is odd; NS box μ_3 gives anti-correlation only if either $\sum_i x_i/2$ or $(1 + \sum_i x_i)/2$ is odd; NS box μ_4 gives anti-correlation only if $\sum_i x_i/2$ is odd or $(1 + \sum_i x_i)/2$ is even. Moreover, since strategy 1 and 3 are such that the correlation produced by parties in the same group are exactly opposite (likewise for strategy 2 and 4), we see that all the less-than-n-partite correlators, as well as the full n-partite correlator when $\sum_{i=1}^{n} x_i$ is odd, indeed vanishes as claimed.

[2] This last sum involves all possible combinations of inputs and thus for all input bit strings $\vec{x}$, there is exactly one term in the expression that does not vanish, therefore giving the identity.

Appendix B: Mermin-Bell violation of the GHZ correlations

Here, we compute the quantum expectation value of the GHZ correlations for the Mermin Bell inequality [9, 10] (here written in the form derived in [2])[3]

$$\left|\mathcal{B}_{\pm}^{n}\right| = 2^{\frac{1-n}{2}}\Big|\sum_{\vec{x}\in\{0,1\}^n}\cos\left\{\frac{\pi}{4}\left[1\pm(n-2\mathbf{x})\right]\right\}E\left(\vec{x}\right)\Big| \leq 1. \tag{B1}$$

The above Bell expression can be rewritten as:

$$\begin{aligned} 2^{\frac{n-1}{2}}\left|\mathcal{B}_{\pm}^{n}\right| =& \Big|\sum_{\vec{x}\in\{0,1\}^n}\cos\left\{\frac{\pi}{4}\left[1\pm(n-2\mathbf{x})\right]\right\}E(\vec{x})\Big|, \\ =& \Big|\sum_{\vec{x}\in\{0,1\}^n}\cos\left[\frac{\pi}{4}(1\pm n)\right]\cos\left(\mathbf{x}\frac{\pi}{2}\right)E(\vec{x}) \\ & \pm\sum_{\vec{x}\in\{0,1\}^n}\sin\left[\frac{\pi}{4}(1\pm n)\right]\sin\left(\mathbf{x}\frac{\pi}{2}\right)E(\vec{x})\Big|. \end{aligned}$$

For the GHZ correlation of Eq. (3), this simplifies to

$$\begin{aligned} \left|\mathcal{B}_{\pm}^{n}\right| =& 2^{\frac{1-n}{2}}\Big|\sum_{\vec{x}\in\{0,1\}^n,\,\mathbf{x}\,\text{even}}\cos\left[\frac{\pi}{4}(1\pm n)\right]\cos^2\left(\mathbf{x}\frac{\pi}{2}\right)\Big|, \\ =& 2^{\frac{n-1}{2}}\left|\cos\left[\frac{\pi}{4}(1\pm n)\right]\right|, \end{aligned}$$

giving

$$\max_{\pm}\left|\mathcal{B}_{\pm}^{n}\right| = \begin{cases} 2^{\frac{n-1}{2}} : n\,\text{odd} \\ 2^{\frac{n-2}{2}} : n\text{ even} \end{cases}, \tag{B2}$$

i.e., achieving maximal [11] possible quantum value of $\left|\mathcal{B}_{\pm}^{n}\right|$ for odd n.

[3] $\mathcal{B}_{+}^{n}$ is the same Bell expression as the usual one obtained through the recursive formula [10]; it can also be obtained by flipping all the inputs in $\mathcal{B}_{-}^{n}$.

Appendix C: Quantum biseparable bound of the n-partite Mermin-Bell expression

For arbitrary odd $n \geq 3$, the Mermin-Bell expression $\mathcal{B}_{+}^{n}$ given on the left-hand-side of Eq. (B1) is equivalent to a special case of a general family of permutationally invariant Bell expression described in Eq. (22) of [12],

$$\Omega_{n,2,2;\delta_{\mathbf{x},0}\cdot r} = 2^{n-2} - 2^{\frac{n-3}{2}}\mathcal{B}_{+}^{n} \tag{C1}$$

From Eq. (23) of [12], it can be shown that the above expression admits the following upper bound on the quantum biseparable bound:

$$\Omega_{n,2,2;\delta_{\mathbf{x},0}\cdot r} \geq 2^{n-3}(2-\sqrt{2}). \tag{C2}$$

Combining these two equations and after some straightforward computations, we get the following upper bound on the quantum biseparable bound for the Mermin-Bell expression:

$$\mathcal{B}_+^n \leq 2^{\frac{n}{2}-1}. \tag{C3}$$

For arbitrary even $n \geq 2$, the Mermin-Bell expression $\mathcal{B}_+^n$ given on the left-hand-side of Eq. (B1) is equivalent to the following Bell expression described in Eq. (1) of [12],

$$\mathcal{I}_{n,2,2} = 2^{n-1} - 2^{\frac{n-2}{2}}\mathcal{B}_+^n \tag{C4}$$

From Eq. (25) of [12], we know that the above expression admits the following upper bound on the quantum biseparable bound:

$$\mathcal{I}_{n,2,2} \geq 2^{n-2}. \tag{C5}$$

Combining these two equations, we arrive, again, at Eq. (C3).

To see that the biseparable bound of Eq. (C3) is tight, it suffices to note that the biseparable quantum state

$$|\psi\rangle = |\mathrm{GHZ}_{n-1}\rangle \otimes |0\rangle \tag{C6}$$

and the local observables

$$\begin{aligned} A_{x_i} &= \cos\alpha_{x_i}\sigma_x + \sin\alpha_{x_i}\sigma_y \quad \forall\ i = 1,\ldots,n-1, \\ A_{x_i} &= \beta_{x_i}\mathbb{1} \quad \text{for} \quad i = n. \end{aligned} \tag{C7}$$

with $\alpha_0 = -\frac{\pi}{4(n-1)}$, $\alpha_1 = -\frac{\pi}{2} - \frac{\pi}{4(n-1)}$, $\beta_0 = -\sqrt{2}\sin\frac{n\pi}{4}$, and $\beta_1 = \sqrt{2}\cos\frac{n\pi}{4}$ indeed give rise to a quantum value of $\mathcal{B}_+^n$ of $2^{\frac{n}{2}-1}$. Since $\mathcal{B}_-^n$ can be obtained from $\mathcal{B}_+^n$ by flipping all the inputs, the same quantum biseparable bound holds for $\mathcal{B}_-^n$.

Since the GHZ correlations of Eq. (3) give Eq. (B2), we see that for odd n, the generation of these correlations necessarily requires a genuinely n-partite entangled state, independent of the underlying Hilbert space dimension.

Appendix D: m-separability and multipartite nonlocality underlying the n-partite GHZ correlations

For odd n, we know from the main theorem of [13] that a quantum violation of $\left|\mathcal{B}_\pm^n\right| = 2^{\frac{n-1}{2}}$ implies that it is impossible to reproduce these GHZ correlations using any 3-separable resource (i.e., a partitioning of the parties into three groups, and where the parties within each group can share even arbitrary nonlocal resource).

For even n, let us evaluate the the quantum value of the following Bell expression [13]:

$$\begin{aligned} |\mathcal{B}_\Sigma^n| =& \frac{1}{\sqrt{2}}\left|\mathcal{B}_+^n + \mathcal{B}_-^n\right|, \\ =& \frac{1}{\sqrt{2}}\Big|\sum_{\vec{x}\in\{0,1\}^n}\sum_{s=0,1}\cos\frac{\pi}{4}[1+(-1)^s(n-2\mathbf{x})]\,E(\vec{x})\Big|, \\ =& \frac{1}{\sqrt{2}}\Big|\sum_{\vec{x}\in\{0,1\}^n} 2\cos\frac{\pi}{4}\cos\Big[\frac{\pi}{4}(n-2\mathbf{x})\Big]E(\vec{x})\Big|, \\ =& \Big|\sum_{\vec{x}\in\{0,1\}^n}\cos\Big[\frac{\pi}{4}(n-2\mathbf{x})\Big]E(\vec{x})\Big|. \end{aligned} \tag{D1}$$

For even n and $E(\vec{x})$ of Eq. (3), this becomes

$$\Big|\sum_{\vec{x}\in\{0,1\}^n,\,\mathbf{x}\,\text{even}}\cos\frac{n\pi}{4}\cos^2\mathbf{x}\frac{\pi}{2}\Big| = 2^{n-1}\left|\cos\frac{n\pi}{4}\right|,$$

giving a value of 2^{n-1} for even $\frac{n}{2}$ and 0 for odd $\frac{n}{2}$.

Again, note from the main theorem of [13] that for even n, any correlation producible by a partition of the n parties into 3 groups (each sharing some Svetlichny resource $\mathcal{S}$ [5–7]) can at most give a value of $\mathcal{B}_\Sigma^n = 2^{n-2}$. This means that, as with odd n, the n-partite GHZ correlation for even n with even $\frac{n}{2}$ is not producible by any partition of the parties into 3 groups, even if parties in each group are allowed to share whatever nonlocal resource.

Together with the biseparable decomposition obtained for these correlations, the above results on m-separability imply that for (1) odd n and (2) even n with even $\frac{n}{2}$, generation of the GHZ correlations of Eq. (3) requires the nonlocal collaboration of at least $\lceil\frac{n}{2}\rceil$ parties in one group.

[1] D. M. Greenberger, M. A. Horne, A. Zeilinger, in "Bell's Theorem, Quantum Theory, and Conceptions of the Universe", edited by M. Kafatos (Kluwer, Dordrecht, 1989), 69-72; N. D. Mermin, Phys. Rev. Lett. **65**, 3373 (1990).

[2] J. J. Wallman, Y.-C. Liang, and S. D. Bartlett, Phys. Rev. A **83**, 022110 (2011).

[3] S. Popescu and D. Rohrlich, Found. Phys. **24**, 379 (1994).
[4] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts, Phys. Rev. A **71**, 022101 (2005).
[5] G. Svetlichny, Phys. Rev. D **35**, 3066 (1987).
[6] R. Gallego, L. E. Würflinger, A. Acín, and M. Navascués, Phys. Rev. Lett. **109**, 070401 (2012).
[7] J.-D. Bancal, J. Barrett, N. Gisin and S. Pironio, Phys. Rev. A **88**, 014102 (2013).
[8] S. Pironio, J.-D. Bancal, and V. Scarani, J. Phys. A: Math. Theor. **44**, 065303 (2011).
[9] N. D. Mermin, Phys. Rev. Lett. **65**, 1838 (1990); M. Ardehali, Phys. Rev. A **46**, 5375 (1992).
[10] S. M. Roy and V. Singh, Phys. Rev. Lett. **67**, 2761 (1991); A. V. Belinskiĭ and D. N. Klyshko, Phys. Usp. **36** 653 (1993); N. Gisin and H. Bechmann-Pasquinucci, Phys. Lett. A **246**, 1 (1998).
[11] R. F. Werner and M. M. Wolf, Phys. Rev. A **61**, 062102 (2000).
[12] J.-D. Bancal, C. Branciard, N. Brunner, N. Gisin, and Y.-C. Liang, J. Phys. A: Math. Theor. **45**, 125301 (2012).
[13] J.-D. Bancal, C. Branciard, N. Gisin and S. Pironio, Phys. Rev. Lett. **103**, 090503 (2009).