Universal bound on the cardinality of local hidden variables in networks

Rosset, Denis; Gisin, Nicolas; Wolfe, Elie

Quantum Information and Computation, Vol. 18, No. 11&12 (2018) 0910–0926

UNIVERSAL BOUND ON THE CARDINALITY OF LOCAL HIDDEN VARIABLES IN NETWORKS

DENIS ROSSET[1,2,3a] NICOLAS GISIN[1] *and* ELIE WOLFE[3]
[1] *Group of Applied Physics, Université de Genève, 1211 Genève, Switzerland*
[2] *Department of Physics, National Cheng Kung University, Tainan 701, Taiwan*
[3] *Perimeter Institute for Theoretical Physics, 31 Caroline St. N, Waterloo, Ontario, Canada, N2L 2Y5*

Received October 26, 2017
Revised August 12, 2018

We present an algebraic description of the sets of local correlations in arbitrary networks, when the parties have finite inputs and outputs. We consider networks generalizing the usual Bell scenarios by the presence of multiple uncorrelated sources. We prove a finite upper bound on the cardinality of the value sets of the local hidden variables. Consequently, we find that the sets of local correlations are connected, closed and semialgebraic, and bounded by tight polynomial Bell-like inequalities.

Keywords: Nonlocality, Quantum Networks, Causal Structures

Communicated by: S Braunstein & R Blatt

1 Introduction

Bell's theorem opened a new perspective for the study of quantum systems, as it predicted that quantum systems exhibit a wider range of correlations than systems restricted to classical information. The first results concerned two observers sharing a single resource modeled using a local hidden variable: the corresponding set of correlations is a polytope [1.2]. This mathematical structure enables straightforward checking of the locality of correlations by linear programming [3], and the facets of the polytope provide ready-to-use linear inequalities. In the multipartite version of Bell's locality [2,4], the networks describe several observers sharing a single resource; there, the set of local correlations is still a polytope. In all these studies, it is customary to identify the local hidden variable with a list of deterministic strategies implemented by the parties. The (finite) number of those strategies provides an upper bound on the cardinality of the local hidden variable, that is the number of different values it has to take to reproduce all local correlations [5].

Later, the description of networks of uncorrelated sources led to local models containing independent local hidden variables [6,7], extending the idea of "local beables" originated by John Bell [8]. There, the correlation sets are no longer polytopes, being not even convex. However, it is still possible, in some cases, to identify local hidden variables with deterministic strategies and to provide a bound on their cardinality [7]. The local sets of several networks have been characterized, at least partially, in the probability space [6,7,9–14] or in the entropy

[a]physics@denisrosset.com

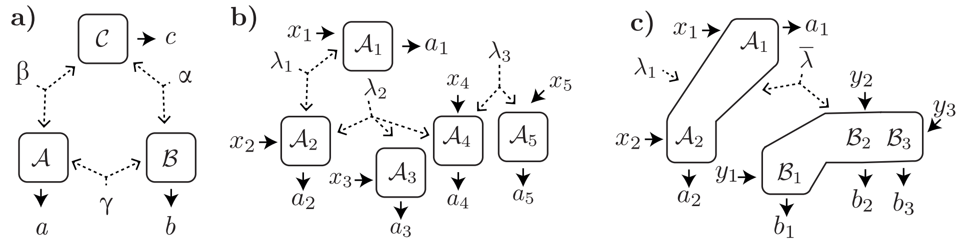

Fig. 1. In a), the triangle network Δ, where three parties share three bipartite local hidden variables α, β, γ and produce binary outputs, without receiving any input. In b), an example of complex network. In c), the renaming/grouping of parties used in Proposition 3 for the same network.

space [15–17], but there is no general method providing a list of inequalities in contrast to the case of Bell locality. However, the inflation technique [14] provides a hierarchy that converges [18] to the local set.

As the description using marginal entropies loses information, we focus on the probability space in the present work. As done usually for the characterization of local sets, we assume that inputs and outputs are taken from finite sets — however, we do not assume this restriction on the local hidden variables *a priori*. In any network, the characterization of the local correlations is tractable algebraically as long as one condition is satisfied: that *all* local hidden variables take a finite number of values. When this condition holds, the set of local correlations is described by a system of polynomial inequalities [19], reminiscent of the linear Bell inequalities bounding the local set in Bell scenarios. Correlations can also be tested for nonlocality by various algorithms. Thus, an important open question is whether local hidden variables can be restricted to finite sets without loss of generality [7,19].

Consider, as a motivating example, the triangle scenario (abbreviated Δ) shown in Figure 1a, introduced independently in [7] and [20]. The parties A, B, C share three independent local hidden variables with values α, β, γ, according to the connections $A \leftarrow (\beta, \gamma)$, $B \leftarrow (\gamma, \alpha)$ and $C \leftarrow (\alpha, \beta)$ — each party is connected to two variables, and outputs a single bit, written respectively $a, b, c = 0, 1$. The joint probability distribution $P(abc)$ describes the behavior of the network.[b] We now consider a particular local behavior. The local hidden variables are all uniformly distributed between 0 and 1. Each party outputs the Boolean result of the comparison "$\lambda_1 \geq \lambda_2$", where (λ_1, λ_2) corresponds to the pair of variables connected to it: for example, $a = 1$ if and only if $\beta \geq \gamma$. As the underlying Δ-local model has a cyclic symmetry, the resulting correlations are symmetric under cyclic permutation of parties. We compute easily $P(abc) = 0$ if $a = b = c$ and $P(abc) = 1/6$ otherwise. Surprisingly, this behavior, which we write $\vec{P}_{\neq} = (0, 1, 1, 1, 1, 1, 1, 0)/6$ by enumerating the indices (a, b, c) in the lexicographic order, does not have a symmetric *finite* Δ-local model, that is a symmetric model where $\alpha \in \Omega_\alpha$ with Ω_α finite, as we now prove.

Without loss of generality, a finite local model is written using local hidden variables $\alpha, \beta, \gamma \in \{1, \ldots, u\}$ with respective distributions P_α, P_β, P_γ and local response functions P_{A},

[b] Writing $P(abc)$ is an abuse of notation as it does not distinguish between the random variable a itself and the *value* a taken by the random variable in particular cases. When necessary, we write $P(0)$ explicitly as $P(a = 0)$ or $P_a(0)$.

P_{B}, P_{C}:

$$P(abc) = \sum_{\alpha,\beta,\gamma=1}^{u} P_\alpha(\alpha) P_\beta(\beta) P_\gamma(\gamma) P_{\mathrm{A}}(a|\beta\gamma) P_{\mathrm{B}}(b|\gamma\alpha) P_{\mathrm{C}}(c|\alpha\beta). \tag{1}$$

When the model is symmetric, we have $P_\alpha = P_\beta = P_\gamma$ and $P_{\mathrm{A}} = P_{\mathrm{B}} = P_{\mathrm{C}}$. To reproduce the correlations $\vec{P}_{\neq}$, we have $u = P_\alpha(\alpha = i) > 0$ and $v = P_{\mathrm{A}}(a = j|\beta\gamma = ii) > 0$ for some values i, j. Then $P(abc = jjj) \geq u^3 v^3 > 0$ is a contradiction. We could think of this as a hint that finite models do not exist for the correlations $\vec{P}_{\neq}$. This intuition would be incorrect, as finite-valued models exist at the price of breaking symmetry (see Appendix B). Moreover, we show in this work that finite-valued models are actually universal, as long as the sets of inputs and outputs employed by the parties are themselves finite.

The manuscript is structured as follows. We first establish the notation and our main claim in Section 2. The main claim is proved twice in Section 3, first using a simplified construction, and then by a refinement that achieves a better upper bound on the cardinality; we also briefly study the impact of our result on the dimensionality of quantum states.

2 Definitions and main claim

The study of nonlocality can be generalized to arbitrary networks of sources and parties. Let us consider a network of m parties sharing n sources as in Figure 1b. The parties are written $\mathcal{A}_1, \ldots, \mathcal{A}_m$, and have inputs $x_1, \ldots, x_m$ and outputs $a_1, \ldots, a_m$ taken from finite sets such that the observations are described by joint probability distribution:

$$P(a_1 \ldots a_m | x_1 \ldots x_m) = P\left(\overline{a}|\overline{x}\right) \tag{2}$$

collecting $\overline{a} = (a_1, \ldots, a_m)$ and $\overline{x} = (x_1, \ldots, x_m)$.

By numbering the input and output values, we identify $x_i = 1, \ldots, X_i$ and $a_i = 1, \ldots, A_i$, and enumerate the coefficients of $P(a_1 \ldots a_m | x_1 \ldots x_m)$ in a vector $\vec{P} \in \mathbb{R}^d$ where

$$d = X_1 \ldots X_m A_1 \ldots A_m\ . \tag{3}$$

We describe the connections in the network by the incidence matrix $I \in \{0,1\}^{m \times n}$ where $I_{ij} = 1$ when the i-th party is connected to the j-th source. A network $\mathcal{N}$ is then described by the size of the input/output sets $\overline{X} = (X_1, \ldots, X_m)$, $\overline{A} = (A_1, \ldots, A_m)$ and the incidence matrix I.

In the network $\mathcal{N}$, we turn to the description of local models. Each source $\mathcal{S}_j$ produces a local hidden variable $\lambda_j \in \Omega_j$, taken from the value set Ω_j with probability measure [c]ρ_j. Each party $\mathcal{A}_i$ receives an input x_i along with the local hidden variables $\lambda_{[i]} = \{\lambda_j | I_{ij} = 1\}$, and processes them according to the response function $P_i(a_i|x_i\lambda_{[i]})$. Then:

$$P\left(\overline{a}|\overline{x}\right) = \prod_{j=1}^{n} \int_{\Omega_j} \mathrm{d}\rho_j(\lambda_j) \prod_{i=1}^{m} P_{\mathcal{A}_i}(a_i|x_i\lambda_{[i]}). \tag{4}$$

The local model $\mathcal{M}$ is fully described by the value sets Ω_j, the probability densities ρ_j and the response functions $P_i(a_i|x_i\lambda_{[i]})$. The set of network-local (or $\mathcal{N}$-local) correlations $\mathcal{L}$ is the set of all $P\left(\overline{a}|\overline{x}\right)$ reproduced by some model $\mathcal{M}$ according to (4). Our main objective is to prove that all $\mathcal{N}$-local correlations can be produced by a simple model.

[c]The usual notation used in the study of Bell locality uses a probability *density* $\rho_j(\lambda_j)$ while the existence of a proper measure for λ_j is implicitly assumed. The notation used here is equivalent and slightly shorter.

Proposition 1 *Any $\vec{P} \in \mathcal{L} \subset \mathbb{R}^d$ can be reproduced with a generic model of the form:*

$$P\left(\overline{a}|\overline{x}\right) = \prod_{j=1}^{n} \sum_{\lambda_j=1}^{d+1} P_{\lambda_j}(\lambda_j) \prod_{i=1}^{m} P_{\mathcal{A}_i}(a_i|x_i\lambda_{[i]}), \tag{5}$$

where the local hidden variables λ_j are integers in the set $\{1, \ldots, d+1\}$.

The proof will directly follow from Proposition 2 below, which bounds the cardinality of network-local models.

3 Cardinality of network-local models

An estimation of the complexity of the model $\mathcal{M}$ is given by the cardinality of the input sets Ω_j. When Ω_j is finite, we write $c_j = |\Omega_j|$ and otherwise $c_j = \infty$. The overall complexity is given by the tuple $\overline{c} = (c_1, \ldots, c_n)$. We compare the power of given cardinalities by the componentwise partial order: $\overline{c}\overline{c}'$ if $c_j c'_j$ for all j.

We now look for an upper bound $\overline{c}_{\text{ub}}$ in any network such that all local $\vec{P} \in \mathcal{L}$ can be realized with models of cardinality $\overline{c}\overline{c}_{\text{ub}}$. Our bound is not minimal but scales linearly with the dimension of $\vec{P}$. Note that the cardinalities $\overline{c}$ are not totally ordered and thus there could well be several minimal upper bounds $\overline{c}$ in a given network.

Our construction rests on the existence of individual bounds on the cardinality of each source. We show that there exists a finite upper bound u_1 on the cardinality c_1 of the source $\mathcal{S}_1$: any model that reproduces a local behavior $\vec{P}$ with cardinality $c_1 > u_1$ can be modified into a model with $c'_1 u_1$ without changing $c_2, \ldots, c_n$. The same proposition holds for each source $\mathcal{S}_j$ with a corresponding upper bound u_j. Thus, the cardinality $\overline{c}_{\text{ub}} = (u_1, \ldots, u_n)$ is sufficient to reproduce all local behaviors.

We first give a simplified version of our construction, and then improve the obtained upper bound in a refinement.

3.1 Finite upper bound: simplified version

Without loss of generality, we consider the cardinality of the source $\mathcal{S}_1$.

Proposition 2 *In a given network $\mathcal{N}$, let $\mathcal{M}$ be a model for the local behavior $\vec{P} \in \mathbb{R}^d$, with cardinality $\overline{c} = (c_1, c_2, \ldots, c_n)$. Then there is a model $\mathcal{M}'$ of cardinality $\overline{c}' = (d+1, c_2, \ldots, c_n)$ that reproduces $\vec{P}$.*

Proof. Let us perform the following experiment with the model $\mathcal{M}$. We sample randomly a value $\lambda_1 = \mu \in \Omega_1$ from the source $\mathcal{S}_1$, and then replace $\mathcal{S}_1$ by a deterministic source that always outputs $\lambda_1 = \mu$. The behavior of the resulting model is:

$$P_\mu\left(\overline{a}|\overline{x}\right) = \left[\prod_{j=2}^{n} \int_{\Omega_j} \mathrm{d}\rho_j(\lambda_j) \prod_{i=1}^{m} P_{\mathcal{A}_i}(a_i|x_i\lambda_{[i]})\right]_{\lambda_1=\mu}. \tag{6}$$

Now, the vector $\vec{P}_\mu$ is a random variable that depends on μ with average value:

$$\langle\vec{P}_\mu\rangle = \int_{\Omega_1} \mathrm{d}\rho_1(\mu)\vec{P}_\mu, \tag{7}$$

and this average value is equal to the original behavior $\vec{P}$. Let $U = \{\vec{P}_\mu | \mu \in \Omega_1\} \subset \mathbb{R}^d$ be the set of possible values of $\vec{P}_\mu$. By construction, $\langle \vec{P}_\mu \rangle$ is a convex mixture of points in U. Using Carathéodory's theorem (see Appendix 1), we can write $\langle \vec{P}_\mu \rangle$ as a convex combination of at most $d+1$ points of U with weights $\{w_k\}$:

$$\langle \vec{P}_\mu \rangle = \sum_{k=1}^{d+1} w_k \vec{P}_{\mu_k} \tag{8}$$

as each point of U can be realized with a $\mu \in \Omega_1$ (not necessarily unique). Then:

$$\vec{P} = \sum_{k=1}^{d+1} w_k \left[\prod_{j=2}^{n} \int_{\Omega_j} \mathrm{d}\rho_j(\lambda_j) \prod_{i=1}^{m} P_{\mathcal{A}_i}(a_i | x_i \lambda_{[i]}) \right]_{\lambda_1 = \mu_k} . \tag{9}$$

Thus, we can replace ρ_1 by a probability distribution ρ_1' on a discrete set $\Omega_1' = \{\mu_1, \ldots, \mu_{d+1}\}$ with weights w_k, and obtain a model $\mathcal{M}'$ that reproduces the behavior $\vec{P}$ with cardinality $c_1 d + 1$, while other elements of $\mathcal{M}$ are left unchanged. □

The proof of Proposition 1 follows directly. Once all sets Ω_j have been replaced by sets of finite cardinality, we simplify the model structure by replacing all local variables by integers $\lambda_1'', \ldots, \lambda_n'' \in \{1, \ldots, d+1\}$ indexing the elements in the finite sets $\Omega_1', \ldots, \Omega_n'$. In the end, any $\mathcal{N}$-local behavior can be reproduced by the generic model (5).

This result simplifies the study of local models in arbitrary networks. When all the involved sets of values are finite, the set of correlations is parameterized by the generic model (5), a polynomial system involving a finite number of equations and unknowns (as already noted by various authors [19,21]. This mathematical structure enables the generalization of several concepts used in the study of Bell locality. As detailed in Appendix D, the set of network-local correlations is a closed semialgebraic set bounded by system of polynomial inequalities of the form $f(\vec{P}) \geq 0$ (for Bell, it was a polytope bounded by a finite number of linear inequalities). Thus, any nonlocal behavior $\vec{P}$ violates such an inequality by a nonzero amount, a fact that can be tested experimentally. The membership problem (is $\vec{P}$ local?) is a polynomial feasibility problem, solved for example by sum-of-squares relaxations that provide the relevant inequality (for Bell, a linear program).

Our result also simplifies the machinery of proofs, as it removes the conceptual difficulties of continuous models such as non-empty sets of measure zero. An example is given in Appendix C, where we provide an elementary proof that the behavior $P_{abc}(000) = P_{abc}(111) = 1/2$ is non-Δ-local.

Our construction (and its refinements below) rely on Carathéodory's theorem, and is not directly constructive. However, as it provides an upper bound on all c_j, we can always find a successful realization of $\vec{P}$ using a model of bounded cardinality. We provide an example of such an exhaustive search in Appendix B.

3.2 Finite upper bound: refined version

We refine our bound by observing two properties of the set U used in the proof of Proposition 2. Firstly, the bound provided by Carathéodory's theorem depends on the affine dimension of U, which is always less than d. Secondly, we replace Carathéodory's theorem by a variant due to Fenchel, and use the fact that U can always be taken as connected.

Let $\mathcal{P}$ be the set of arbitrary nonsignaling probability distributions $P(\overline{a}|\overline{x})$ in a network $\mathcal{N}$. Its affine dimension [22] is given by:

$$\mathrm{affdim}(\mathcal{P}) = \prod_{i=1}^{m} [X_i(A_i - 1) + 1] - 1, \tag{10}$$

where (A_i, X_i) is the number of (outputs, inputs) of the i-th party. This affine dimension is made explicit, for example, by the Collins-Gisin parameterization of $\mathcal{P}$ [23].

As before, we consider the local hidden variable λ_1 without loss of generality. We write $\overline{\lambda} = (\lambda_2, \ldots, \lambda_n)$ with corresponding probability measure $\overline{\rho}$ over a set $\overline{\Omega}$. We rename the parties and local hidden variables as follows. As drawn in Figure 1c, we write $\mathcal{A}_1, \ldots, \mathcal{A}_N$ the parties connected to the local hidden variable $\lambda_1 \in \Omega_1$, with inputs $x_1, \ldots, x_N$ and outputs $a_1, \ldots, a_N$, collected in $\overline{a} = (a_1, \ldots, a_N)$, $\overline{x} = (x_1, \ldots, x_N)$. We write $\mathcal{B}_1, \ldots, \mathcal{B}_{n-N}$ be the remaining parties, with inputs $y_1, \ldots, y_{n-N}$ and outputs $b_1, \ldots, b_{n-N}$, which we also collected in $\overline{b}$ and $\overline{y}$. The behavior $\vec{P}$ is written:

$$\begin{aligned} P(a_1 \ldots a_N b_1 \ldots b_{n-N} | x_1 \ldots x_N y_1 \ldots y_{n-N}) &= P\left(\overline{a}\overline{b}\middle|\overline{xy}\right) \\ &= \int_{\Omega_1} \mathrm{d}\rho_1(\lambda_1) \int_{\overline{\Omega}} \mathrm{d}\overline{\rho}\left(\overline{\lambda}\right) P_{\overline{\mathrm{A}}}\left(\overline{a}\middle|\overline{x}\lambda_1\overline{\lambda}\right) P_{\overline{\mathrm{B}}}\left(\overline{b}\middle|\overline{y\lambda}\right), \end{aligned} \tag{11}$$

where $P_{\overline{\mathrm{A}}}\left(\overline{a}\middle|\overline{x}\lambda_1\overline{\lambda}\right)$ collects the response functions of $A_1 \ldots A_N$ and $P_{\overline{\mathrm{B}}}\left(\overline{b}\middle|\overline{y\lambda}\right)$ the response functions of $B_1 \ldots B_{n-N}$.

Proposition 3 *Let $\mathcal{M}$ be a model for the local behavior $\vec{P} \in \mathbb{R}^d$, of cardinality $\overline{c} = (c_1, c_2, \ldots, c_n)$. Then there is a model $\mathcal{M}'$ of cardinality $\overline{c}' = (u_1, c_2, \ldots, c_n)$ that reproduces $\vec{P}$ with $u_1 = \mathrm{affdim}(\mathcal{P}_{\mathrm{AB}}) - \mathrm{affdim}(\mathcal{P}_B)$, where $\mathcal{P}_{AB}$ is the set of nonsignaling $P\left(\overline{a}\overline{b}\middle|\overline{xy}\right)$ and $\mathcal{P}_B$ the set of nonsignaling $P\left(\overline{b}\middle|\overline{y}\right)$.*

Proof. As before, for a fixed value $\lambda_1 = \mu$, we write:

$$P_\mu\left(\overline{a}\overline{b}\middle|\overline{xy}\right) = \left[\int_{\overline{\Omega}} \mathrm{d}\overline{\rho}\left(\overline{\lambda}\right) P_{\overline{\mathrm{A}}}\left(\overline{a}\middle|\overline{x}\lambda_1\overline{\lambda}\right) P_{\overline{\mathrm{B}}}\left(\overline{b}\middle|\overline{y\lambda}\right)\right]_{\lambda_1=\mu}, \tag{12}$$

with $U = \{\vec{P}_\mu | \mu \in \Omega_1\}$. Firstly, we remark that the marginal distribution $P_\mu\left(\overline{b}\middle|\overline{y}\right)$ does not depend on μ:

$$P_\mu\left(\overline{b}\middle|\overline{y}\right) = \int_{\overline{\Omega}} \mathrm{d}\overline{\rho}\left(\overline{\lambda}\right) P_{\overline{\mathrm{B}}}\left(\overline{b}\middle|\overline{y\lambda}\right) = P\left(\overline{b}\middle|\overline{y}\right). \tag{13}$$

Secondly, the set U can always be made connected by modifying the model as follows. We replace λ_1 by $\lambda_1' = (\lambda_1, \nu)$, where $\nu \in [0,1]$ represents noise strength. We modify the response function of $\mathcal{A}_1$ such that:

$$P'_{\mathcal{A}_1}(a_1 | x_1 \lambda_{[1]} \nu) = \nu P_1(a_1|x_1) + (1-\nu) P_{A_1}(a_1 | x_1 \lambda_{[1]}), \tag{14}$$

where P_1 is the uniformly random distribution, and so on for all $\mathcal{A}_i$. The new model still reproduces the behavior $\vec{P}$ provided we fix always $\nu = 0$. When we fix $\nu = 1$, we obtain:

$$P_{\nu=1}\left(\overline{a}\overline{b}\middle|\overline{xy}\right) = P_1\left(\overline{a}|\overline{x}\right) P\left(\overline{b}\middle|\overline{y}\right), \tag{15}$$

which does not depend on μ, and the path obtained by varying ν between 0 and 1 is continuous. Then, any point in U can be brought to $\vec{P}_{\nu=1}$, and U is connected. The connectedness of U is a special case of Prop. 3.15 in Ref. [19].

An upper bound on the affine dimension of U is given by the affine dimension of the set of nonsignaling $P\left(\overline{a}\overline{b}\middle|\overline{x}\overline{y}\right)$, after removing the degrees of freedom fixed by the constant marginal $P\left(\overline{b}\middle|\overline{y}\right)$, see Eq. (13). Thus:

$$\mathrm{affdim}(U) u_1 \equiv \mathrm{affdim}(\mathcal{P}_{\mathrm{AB}}) - \mathrm{affdim}\left(\mathcal{P}_{\mathrm{B}}\right). \tag{16}$$

As the set U is connected, we apply Fenchel's variant of Carathéodory's theorem and the number of values in the convex decomposition is upper bounded by u_1. □

3.3 Examples

In the Δ network presented in the introduction, we consider, without loss of generality, the cardinality of the variable α. The affine dimension of $P(abc)$ is 7 while the affine dimension of $P(a)$ is 1; thus $u_1 = 6$ and any Δ-local distribution can be reproduced with $\Omega_\alpha, \Omega_\beta, \Omega_\gamma$ containing at most 6 values.

The bilocal scenario considered considered in [7] has three parties, $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$, with binary inputs and outputs $a, b, c, x, y, z = 0, 1$ connected by local hidden variables $\lambda_{\mathcal{AB}}$ and $\lambda_{\mathcal{BC}}$. The affine dimension of $P(abc|xyz)$ is 26, the affine dimension of $P(bc|yz)$ is 8; thus the cardinality of $\lambda_{\mathcal{AB}}$ is upper bounded by 18. However, by enumerating the deterministic strategies corresponding to $\mathcal{A}$, the cardinality of $\lambda_{\mathcal{AB}}$ is maximum 4.

This shows that the upper bound presented in this paper is not always optimal. Indeed, consider a Bell scenario where $\mathcal{A}$ and $\mathcal{B}$ have binary outputs $a, b = 0, 1$ but no input, and are connected to the variable λ. The affine dimension of $P(ab)$ is 3, and thus we obtain the upper bound 3 on the cardinality of λ. However, we can also identify $\lambda = a$ and write $P(ab) = P(a)P(b|a)$ to provide a model with cardinality 2.

3.4 Comparison to quantum resources

What happens when we try to extend our result to other types of resources, for example quantum states? Consider, for example, the bipartite scenario where a quantum source produces the state $\rho_{\mathrm{AB}} = \sum_i p_i |\varphi_i\rangle\langle\varphi_i|$. Then, we observe:

$$P(ab|xy) = \mathrm{tr}\left[\rho_{\mathrm{AB}}\left(\Pi^{\mathrm{A}}_{a|x} \otimes \Pi^{\mathrm{B}}_{b|y}\right)\right] = \sum_i p_i \left\langle \varphi_i \middle| \Pi^{\mathrm{A}}_{a|x} \otimes \Pi^{\mathrm{B}}_{b|y} \middle| \varphi_i \right\rangle. \tag{17}$$

By convexity, an argument similar to Proposition 2 selects at most $d+1$ elements in the decomposition of ρ_{AB}. Thus, the *rank* of ρ_{AB} can be always be upper-bounded, while the *dimension* of the quantum state is unchanged (or decreased, depending on whether the dimension is the Hilbert space dimension or the Schmidt number). However, the defining feature of the size of a quantum state is its dimension, rather than its rank: we know that mixed states can always be replaced by pure states of higher dimension. The recent work of Slofstra [24,25] proves that some quantum correlations require infinite dimensional systems, even in scenarios when the cardinality of inputs and outputs is finite. This highlights a qualitative difference between classical and quantum systems.

4 Conclusion

We proved that finite local hidden variables suffice to reproduce local correlations in any network. In consequence, we showed that local sets of correlations are semialgebraic, tightly bounded by a finite number of polynomial inequalities. We also outlined algebraic methods to check the nonlocality of given correlations. Our construction rests on the dimension of the output space and does not exploit much of the network structure. It is thus likely that the bound we propose can be improved. For example, in the case of the triangle network Δ, Proposition 3 proposes an upper bound of 6, while the behavior $P_{\neq}$ has a local model using only bits. We leave as an open question the closing of the gap between these bounds.

However, for some purposes, a tight upper bound does not matter as only the *existence* of a finite model simplifies the calculations. Indeed, finite probability spaces are conceptually much simpler, as the machinery of probability measures/densities can replaced by discrete distributions. Finally, we exhibited a generalization of our method to quantum states, and outlined the distinct notions of rank and dimension, which are conflated in the classical case.

Acknowledgements

We thank many colleagues for discussions in the last three years; among them Jean-Daniel Bancal, Nicolas Brunner, Rafael Chaves, Flavien Hirsch, Yeong-Cherng Liang, Marc-Olivier Renou and Rob Spekkens. This work is supported by the Swiss National Science Foundation via the Mobility Fellowship P2GEP2_162060 and the NCCR-QSIT; also by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science and Economic Development and by the Province of Ontario through the Ministry of Research and Innovation. This publication was made possible through the support of a grant from the John Templeton Foundation. We also acknowledge support by the Ministry of Education, Taiwan, R.O.C., through Aiming for the Top University Project granted to the National Cheng Kung University (NCKU), and by the Ministry of Science and Technology, Taiwan (Grants No. 104-2112-M-006-021-MY3).

References

pt!

1. I. Pitowsky, "Correlation polytopes: Their geometry and complexity," Mathematical Programming **50**, 395 (1991).
2. N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, "Bell nonlocality," Reviews of Modern Physics **86**, 419 (2014).
3. D. Kaszlikowski, P. Gnacinski, M. Zukowski, W. Miklaszewski, and A. Zeilinger, "Violations of Local Realism by Two Entangled N-Dimensional Systems Are Stronger than for Two Qubits," Physical Review Letters **85**, 4418 (2000).
4. N. D. Mermin, "Extreme quantum entanglement in a superposition of macroscopically distinct states," Physical Review Letters **65**, 1838 (1990).
5. J. M. Donohue and E. Wolfe, "Identifying nonconvexity in the sets of limited-dimension quantum correlations," Physical Review A **92**, 062120 (2015).
6. C. Branciard, N. Gisin, and S. Pironio, "Characterizing the Nonlocal Correlations Created via Entanglement Swapping," Physical Review Letters **104**, 170401 (2010).
7. C. Branciard, D. Rosset, N. Gisin, and S. Pironio, "Bilocal versus nonbilocal correlations in entanglement-swapping experiments," Physical Review A **85**, 032119 (2012).
8. J. Bell, "On the Einstein-Podolsky-Rosen Paradox," Physics **1**, 195 (1964).

9. A. Tavakoli, P. Skrzypczyk, D. Cavalcanti, and A. Acn, “Nonlocal correlations in the star-network conguration,” Physical Review A **90**, 062109 (2014).
10. D. Rosset, C. Branciard, T. J. Barnea, G. Putz, N. Brunner, and N. Gisin, Nonlinear Bell Inequalities Tailored for Quantum Networks,” Physical Review Letters **116**, 010403 (2016).
11. A. Tavakoli, “Quantum correlations in connected multipartite Bell experiments,” Journal of Physics A **49**, 145304 (2016).
12. A. Tavakoli, Bell-type inequalities for arbitrary noncyclic networks,” Physical Review A **93**, 030101 (2016).
13. R. Chaves, “Polynomial Bell Inequalities,” Physical Review Letters **116**, 010402 (2016).
14. E. Wolfe, R. W. Spekkens, and T. Fritz, “The Ination Technique for Causal Inference with Latent Variables,” arXiv:1609.00672 (2016).
15. R. Chaves and T. Fritz, “Entropic approach to local realism and noncontextuality,” Physical Review A **85**, 032113 (2012).
16. R. Chaves, L. Luft, and D. Gross, “Causal structures from entropic information: Geometry and novel scenarios,” New Journal of Physics **16**, 043001 (2014).
17. J. Henson, R. Lal, and M. F. Pusey, “Theory-independent limits on correlations from generalized Bayesian networks,” New Journal of Physics **16**, 113043 (2014).
18. M. Navascues and E. Wolfe, “The ination technique solves completely the classical inference problem,” arXiv:1707.06476 (2017).
19. T. Fritz, “Beyond Bells theorem: Correlation scenarios,” New Journal of Physics **14**, 103001 (2012).
20. B. Steudel and N. Ay, “Information-theoretic inference of common ancestors,” Entropy **17**, 2304 (2015).
21. D. Geiger, D. Heckerman, H. King, and C. Meek, “Stratied exponential families: Graphical models and model selection,” Ann. Statist. **29**, 505 (2001).
22. S. Boyd and L. Vandenberghe, Convex Optimization (Cambridge University Press, 2004).
23. D. Collins and N. Gisin, “A relevant two qubit Bell inequality inequivalent to the CHSH inequality,” Journal of Physics A **37**, 1775 (2004).
24. W. Slofstra, “Tsirelsons problem and an embedding theorem for groups arising from non-local games,” (2016), arXiv:1606.03140.
25. W. Slofstra, “The set of quantum correlations is not closed,” (2017), arXiv:1703.08618.
26. I. Barany and R. Karasev, Notes about the Caratheodory number,” Discrete Computational Geometry **48**, 783 (2012).
27. J. Lofberg, “YALMIP : A Toolbox for Modeling and Optimization in MATLAB,” in In Proceedings of the CACSD Conference (Taipei, Taiwan, 2004).
28. H. P. Williams, “Fouriers Method of Linear Programming and Its Dual,” The American Mathematical Monthly **93**, 681 (1986).
29. J. B. Lasserre, “Global Optimization with Polynomials and the Problem of Moments,” SIAM Journal on Optimization **11**, 796 (2001).
30. P. A. Parrilo, “Semidenite programming relaxations for semialgebraic problems,” Mathematical Programming **96**, 293 (2003).
31. J. Bochnak, M. Coste, and M.-F. Roy, Real Algebraic Geometry (Springer Berlin Heidelberg, 1998).
32. J. Rotman, An Introduction to Algebraic Topology (Springer Science Business Media, 1998).
33. S. Basu, R. Pollack, and M.-F. Coste-Roy, Algorithms in Real Algebraic Geometry (Springer Berlin Heidelberg, 2006).
34. C. M. Lee and R. W. Spekkens, “Causal inference via algebraic geometry: feasibility tests for functional causal structures with two binary observed variables,” Journal of Causal Inference (2015).
35. A. Papachristodoulou, J. Anderson, G. Valmorbida, S. Prajna, P. Seiler, and P. A. Parrilo, SOSTOOLS: Sum of Squares Optimization Toolbox for MATLAB (arXiv:1310.4716, 2013) available from eng.ox.ac.uk/control/sostools, cds.caltech.edu/sostools and mit.edu/parrilo/sostools.

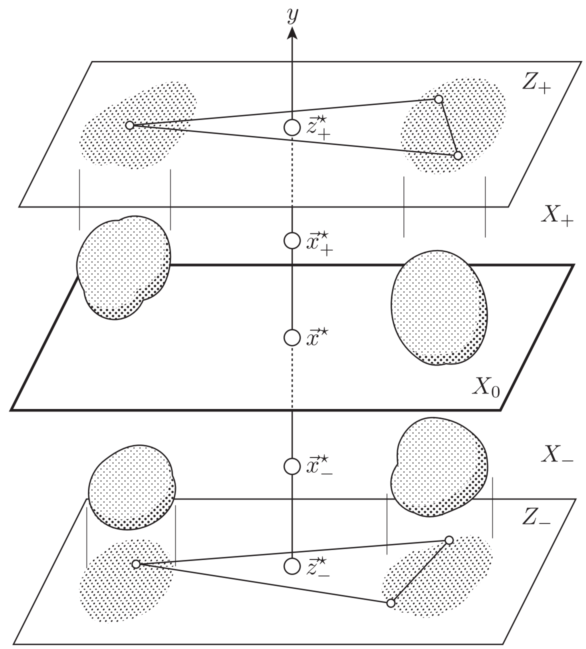

Fig. A.1. For $d = 3$, the decomposition of X, represented by dotted blobs, into sets $X_\pm$ depending on the sign of y. The top and bottom planes represent the projection $Z_\pm$, with the triangles showing that $\vec{z}^\star_\pm$ lies in the convex hull of at most d points.

Appendix A Variants of Carathéodory's theorem

Let $X \subset \mathbb{R}^d$ a bounded subset of $\mathbb{R}^d$ and ρ a probability measure on (the Borel subsets of) $\mathbb{R}^d$. We assume that X has full affine dimension $a = d$. Otherwise, when $a < d$, we use the fact that affine maps preserve convex decompositions, and study the image of X under an injective affine map $f : \mathbb{R}^d \to \mathbb{R}^a$. As noted in several steps of the proof below, we make liberal use of affine maps to bring points at convenient locations to simplify the notation.

Let $\vec{x}^*$ be the center of mass of ρ:

$$\vec{x}^* = \int_X \mathrm{d}\rho(\vec{x})\ \vec{x}. \tag{A.1}$$

When X is compact (or simply finite), the standard formulation [26] of Carathéodory's theorem states that there exists at most $d+1$ elements $\vec{x}_i \in X$ such that $\vec{x}^*$ is a convex mixture of those:

$$\vec{x}^* = \sum_{i=1}^{d+1} w_i \vec{x}_i, \quad \sum_i w_i = 1, \quad \forall i, w_i \geq 0. \tag{A.2}$$

Moreover, when X is also connected, Fenchel's variant [26] of the theorem reduces the upper bound to d instead of $d+1$.

When the set X is not closed, we first reduce to the finite case before applying the original theorem. Without loss of generality, we shift X such that $\vec{x}^*$ is the origin. We rewrite $\vec{x} = (y, \vec{z})$ with $y \in \mathbb{R}$ and $\vec{z} \in \mathbb{R}^{d-1}$ and split X into three sets: X_+, X_0 and X_- depending on the sign of y (see Figure A.1). We restrict and normalize ρ to each of these three sets to obtain ρ_+, ρ_0 and ρ_-, and compute:

$$\vec{x}^*_0 = \int_{X_0} \mathrm{d}\rho_0(\vec{x})\ \vec{x}, \qquad \vec{x}^*_\pm = \int_{X_\pm} \mathrm{d}\rho_\pm(\vec{x})\ \vec{x}, \tag{A.3}$$

such that:

$$\vec{x}^* = w_+\vec{x}^*_+ + w_0\vec{x}^*_0 + w_-\vec{x}^*_- \tag{A.4}$$

for nonnegative, normalized weights $w_0, w_\pm$. When any of $X_\pm$ or X_0 is empty (or of measure zero), we simply remove it from this decomposition. Note that $\vec{x}^*_\pm$, $\vec{x}^*_0$ are convex mixtures on their own and are not necessarily elements of the sets $X_\pm$, X_0. However, $\vec{x}^*_\pm$, $\vec{x}^*_0$ can be obtained by a mixture of a finite number of points of those sets as we show below.

We proceed by induction. When $d = 1$, the solution is trivial when $X_0 \neq \emptyset$. Otherwise, both X_+ and X_- are not empty and we can replace $\vec{x}^*_\pm$ by any point in $X_\pm$, adjusting the weights $w_\pm$ as required.

For $d > 1$, we first look at $\vec{x}^*_0$. If X_0 is not empty, we find a finite decomposition for $\vec{x}^*_0$ by applying the proposition for $(d-1)$. We then write $\vec{x}^*$ as a mixture of $\vec{x}^*_0$ and the point $\vec{x}^\star = (w_+\vec{x}^*_+ + w_-\vec{x}^*_-)/(w_+ + w_-)$. It remains to show that the point $\vec{x}^\star$ has a finite decomposition (and now, we identify $\vec{x}^\star = \vec{x}^*$ when X_0 has measure zero). Without loss of generality, we assume again that $\vec{x}^\star = \vec{0}$. We write:

$$\vec{x}^\star = \begin{pmatrix} 0 \\ \vec{0} \end{pmatrix} = \frac{1}{w_+ + w_-}\left[w_+\begin{pmatrix} y^\star_+ \\ \vec{z}^\star_+ \end{pmatrix} + w_-\begin{pmatrix} y^\star_- \\ \vec{z}^\star_- \end{pmatrix}\right]. \tag{A.5}$$

Using an affine transform, we can always ensure that $\vec{z}^\star_\pm = \vec{0}$ without changing the y coordinates (shear mapping). We first consider the projection of X_+ on the last $d-1$ coordinates:

$$Z_+ = \{\vec{z}_+ \text{ s.t. } \exists y_+, (y_+, \vec{z}_+) \in X_+\}. \tag{A.6}$$

By using the proposition for $d-1$, we find a convex decomposition of $\vec{z}^\star_+$ using a finite number of points:

$$\vec{0} = \vec{z}^\star_+ = \sum_i p_i\vec{z}_i, \quad \vec{z}_i \in Z_+, \quad p_i \geq 0, \quad \sum_i p_i = 1. \tag{A.7}$$

By picking for each $\vec{z}_i$ a corresponding point $\vec{x}_i = (y_i, \vec{z}_i) \in X_+$, we obtain a convex decomposition:

$$\vec{x}^\circ_+ = \begin{pmatrix} y^\circ_+ \\ \vec{0} \end{pmatrix} = \sum_i p_i \begin{pmatrix} y_i \\ \vec{z}_i \end{pmatrix}, \tag{A.8}$$

where $y^\circ_+ > 0$ by construction. The same argument for X_- provides a convex decomposition of $\vec{x}^\circ_- = (y^\circ_-, \vec{0})$ with $y^\circ_- < 0$. Thus, $\vec{x}^\star = \vec{0}$ can be written as a convex mixture of the two points $\vec{x}^\circ_\pm$, which can in turn be written as a combination of a finite number of points.

Note that in Proposition 3, we use the fact that X is closed *and* connected, which can always be achieved by first reducing to the finite case, then adding the connectivity using the trick mentioned in the proof.

Appendix B Asymmetric models for $P_{\neq}$

We provide in this Appendix two asymmetric finite Δ-local models for the behavior $\vec{P}_{\neq}$ presented in the introduction. The first model has cardinality $\bar{c} = (|\Omega_\alpha|, |\Omega_\beta|, |\Omega_\gamma|) = (3, 2, 6)$, while the second model has $\bar{c} = (2, 2, 2)$.

B.1 By application of our proposition

We take the original model and apply the construction of Proposition 2 to each variable in turn, modifying their distributions such that the model always reproduces $\vec{P}_{\neq}$ after each step. We start with β, and $|\Omega'_\beta| > 1$, otherwise $P_{\neq}$ would factorize as $P_{\neq}(ac) = P_{\neq}(a)P_{\neq}(c)$. We try with $\beta_+, \beta_- \in \Omega_\beta$ and a weight $v \in [0, 1]$, such that β_+ is distributed with probability v and β_- with probability $1 - v$.

We obtain the solution $\beta_\pm = \left(3 \pm \sqrt{3}/6\right)$, with $v = 1/2$. We turn to α, and first try to use three $\alpha_1, \alpha_2, \alpha_3 \in \Omega_\alpha$ distributed with the respective nonnegative weights $u_1 + u_2 + u_3 = 1$. We obtain a model with $u_1 = u_3 = \left(3 - \sqrt{3}\right)/6$, $\alpha_1 = \left(3 - \sqrt{3}\right)/12$, $\alpha_2 = 1/2$ and $\alpha_3 = 1 - \alpha_1$. We turn finally to γ, and observe that by the definition of the response functions "$\lambda_1 \geq \lambda_2$?", the most general set of values for γ is given by $\Omega'_\gamma = \{\gamma_1, \ldots, \gamma_6\}$ where the γ_k are contained in the six strict intervals between the seven numbers $\{0, \alpha_1, \beta_1, \alpha_2, \beta_2, \alpha_3, 1\}$ and the exact values of γ_k do not matter. We consider a distribution of γ_k with weights w_k, and obtain $w_1 = w_2 = w_5 = w_6 = \left(3 - \sqrt{3}\right)/12$ and $w_3 = w_4 = 1/\left(2\sqrt{3}\right)$. Our transformed model has $|\Omega'_\alpha| = 3$, $|\Omega'_\beta| = 2$ and $|\Omega'_\gamma| = 6$.

B.2 Using a combinatorial approach

We found a minimal model for the correlations $\vec{P}_{\neq}$ with $\bar{c} = (2, 2, 2)$ using an hybrid combinatorial-algebraic search. We start with $\alpha, \beta, \gamma = 0, 1$ and all coefficients $P_\alpha(\alpha), P_\beta(\beta), P_\gamma(\gamma)$ *strictly* between 0 and 1 (as any $c_i = 1$ is impossible). We first solve the following relaxed problem. Instead of computing the exact value of $P(a|\beta\gamma)$, we only tracked the possible values of a for the four (β, γ) pairs: whether $P(a = 0|\beta\gamma) = 0$, $P(a = 0|\beta\gamma) = 1$ or $P(a = 0|\beta\gamma) \in]0, 1[$; the same for $P(b|\gamma\alpha)$ and $P(c|\alpha\beta)$. For example, it is impossible to have $P(a = 0|\beta\gamma) \in]0, 1[$ for all (β, γ) pairs, as it renders the condition $P(000) = 0$ impossible to satisfy.

In total, there are half a million (3^{12}) possibilities to check. After removing the obvious symmetries (permutation of parties, bit flips of α, β, γ), we are left with 4 cases corresponding to polynomial feasibility problems of degree 4 involving between 4 and 5 unknowns. We checked those cases manually using a branch-and-bound solver (BMIBNB [26]) and found the following model. The local hidden variables have distributions $P_\alpha(0) = P_\beta(0) = 1/3$ and $P_\gamma(0) = 1/4$. The response functions of A and B are deterministic such that $a = \beta\gamma$ and $b = 1 \oplus \gamma\alpha$, while Charlie uses an additional uniformly random bit ν so that $c = \nu$ when $\alpha = \beta$, and $c = \alpha$ otherwise.

Without solving the combinatorial problem first, the algebraic problem involves monomials of degree 6 in 15 variables, and is thus out of reach of branch-and-bound or sum-of-squares relaxation methods.

Appendix C Example of an elementary proof using finite models

In the Δ network, consider the correlations $\vec{P}_{=} = (1/2, 0, 0, 0, 0, 0, 0, 1/2)$ where the outputs are always correlated ($a = b = c$). These correlations were proven non-Δ-local [19] using entropic inequalities, or in [14] using the inflation technique. We here provide an elementary proof based on exploiting the generality of finite Δ-local models[d]

Using Proposition 1, we assume that the local hidden variable sets Ω_α, Ω_β and Ω_γ are finite. By embedding any local source of randomness in a local hidden variable, we can assume the local response functions to be deterministic [2]. Because $P_{=}(000) > 0$, there exists a triplet of values $(\alpha_0, \beta_0, \gamma_0)$ resulting in[e] $a = b = c = 0$, with the triplet probability $P_{\alpha\beta\gamma}(\alpha_0\beta_0\gamma_0) > 0$. A similar argument exhibits $(\alpha_1, \beta_1, \gamma_1)$ such that $a = b = c = 1$ with $P_{\alpha\beta\gamma}(\alpha_1\beta_1\gamma_1) > 0$. Now, consider $(\alpha, \beta, \gamma) = (\alpha_0, \beta_0, \gamma_1)$; surely, $P_{\alpha\beta\gamma}(\alpha_0\beta_0\gamma_1) > 0$. Because $c = 0$, we must have $a = 0$ whenever $(\beta, \gamma) = (\beta_0, \gamma_1)$. Now, when $(\alpha, \beta, \gamma) = (\alpha_1, \beta_0, \gamma_1)$ we have $b = 1$, and thus we must have $a = 1$ for $(\beta, \gamma) = (\beta_0, \gamma_1)$, which is a contradiction.

The proof does not hold when $\Omega_\alpha, \Omega_\beta, \Omega_\gamma$ are infinite: while the probability density $\rho_\alpha(\alpha_0)\rho_\beta(\beta_0)\rho_\gamma(\gamma_1)$ could be nonzero, the event $(\alpha, \beta, \gamma) = (\alpha_0, \beta_0, \gamma_1)$ could be happen *almost never*, and we would not achieve the desired contradiction.

Appendix D Tools for the study of network-local sets

We review first the tools used to study the sets of Bell-local correlations, before going to the general case, where we will obtain the following results, valid in any network:

- The membership problem (is $\vec{P} \in \mathcal{L}$?) can be solved, in principle, with tools of real algebraic geometry.

[d]The authors wish to acknowledge that App. C appears to be an instance of a more general proof technique developed earlier and independently by T. Fritz. Fritz's proof was communicated to the authors in private communication; his single-author unpublished proof technique is otherwise largely superceded by Ref. [14].

[e]That is $P_{\mathrm{A}}(0|\beta_0\gamma_0) = P_{\mathrm{B}}(0|\gamma_0\alpha_0) = P_{\mathrm{C}}(0|\alpha_0\beta_0) = 1$.

– The network-local set of correlations $\mathcal{L}$ is closed, connected, and described by a finite number of nonstrict polynomial inequalities.

Computational requirements, however, preclude the use of such tools except in the simplest cases.

D.1 Bell-local sets

All Bell-local correlations can be written as a convex mixture of deterministic behaviors. We follow the notation of the review [2] :

$$\vec{P} = \sum_\lambda q_\lambda \vec{d}_\lambda = D\vec{q}, \qquad \vec{q} \geq 0, \qquad \sum_\lambda q_\lambda = 1, \tag{D.1}$$

where the inequality $\vec{q} \geq 0$ is understood componentwise. We also collected the deterministic behaviors column-wise in the matrix D. We write $\mathcal{L}$ the set of all Bell-local $\vec{P}$. Note that the construction of the Bell-local set from deterministic strategies preserves a large part of the mathematical structure, as the labels λ in the sum above contains the deterministic outputs of each input of each party.

D.2 Solving the membership problem

We consider the following expression for vectors $\vec{\xi} \in \mathbb{R}^{\dim(\vec{P})}$, $0\vec{\zeta} \in \mathbb{R}^{\dim(\vec{q})}$:

$$I = \vec{\xi}^\top (\vec{P} - D\vec{q}) + \vec{\zeta}^\top \vec{q} \geq 0 \tag{D.2}$$

which is nonnegative by construction (we put no term for the normalization of $\vec{q}$ as it follows from the normalization of $\vec{P}$). Whenever $\vec{\zeta}^\top - \vec{\xi}^\top D = 0$, the inequality reduces to an inequality valid for all Bell-local correlations: $I = \vec{\xi}^\top \vec{P} \geq 0$. Thus, when faced with a normalized $\vec{P}$, we can search for a certificate that proves $\vec{P} \notin \mathcal{L}$:

$$\begin{array}{rcl} \nu = \min_{\vec{\zeta},\vec{\xi}} \vec{\xi}^\top \vec{P} & & \text{such that} \\ \vec{\zeta} & \geq & 0 \\ \vec{\zeta}^\top - \vec{\xi}^\top D & = & 0. \end{array} \tag{D.3}$$

which is a linear program. When the optimal solution has $\nu^* < 0$, the behavior $\vec{P}$ is not local. Following the theory of polytope duality, all $\vec{P} \notin \mathcal{L}$ can be detected.

D.3 Computing a complete description

We just showed how to solve the membership problem for Bell-local sets. However, if we want a complete description of $\mathcal{L}$, we can obtain it by variable elimination. Consider the set X:

$$X = \left\{ \vec{x} = (\vec{P}, \vec{q}) \text{ s.t. } \vec{P} = D\vec{q}, \quad \vec{q} \geq 0, \quad \sum_\lambda q_\lambda = 1 \right\}. \tag{D.4}$$

This set is a polytope in $\mathbb{R}^{\dim(\vec{P})+\dim(\vec{q})}$ described by linear equalities and inequalities. We now consider the projection of X on the dimensions corresponding to $\vec{P}$:

$$X_{|\vec{P}} = \left\{ \vec{P} \text{ such that } \exists \vec{q}, (\vec{P}, \vec{q}) \in X \right\}. \tag{D.5}$$

The resulting polytope, equivalent to the local set $\mathcal{L} = X_{|\vec{P}}$ can be computed using Fourier-Motzkin elimination [28] . The process is however quite demanding and has only been completed for simple scenarios.

D.4 General networks

In more general networks, the network-local correlation set $\mathcal{L}$ is no longer a polytope. The reduction to deterministic strategies still works, as any source of randomness can be embedded in a local hidden variable [2] . However, the parties no longer have access to the same local hidden variable to perform a convex mixture of those strategies.

In some networks, such as those of bilocal scenarios [7] , the reduction to finite models is done while preserving much of the structure. As we will see in Appendix E, this structure can be helpful to reduce the complexity of the local model. Such constructions also work in other networks without loops using a construction similar to the one in [10].

In networks involving loops, such as the triangle network of Figure 1a, we have to resort to Proposition 3. After computing the cardinality of all value sets, we parameterize the model using discrete probability distributions. For example, for the triangle scenario, we need $3 \cdot (6-1)$ coefficients for the distributions of α, β, γ and $3 \cdot (2-1) \cdot 6^2$ coefficients to parameterize the local response functions (having removed the degree of freedom of the normalization), as in Eq. (1). We collect all these coefficients in a vector $\vec{q}$, along with linear inequalities that enforce their nonnegativity:

$$g_j = \vec{a}_j^\top \vec{q} - b_j \geq 0, \tag{D.6}$$

where $\vec{a}_j \in \mathbb{R}^{\dim(\vec{q})}, b_j \in \mathbb{R}$. The fact that $\vec{q}$ reproduces the behavior $\vec{P}$ is given by the relation (5), which we abbreviate

$$f_i = \vec{c}_i^\top \vec{P} - h_i = 0, \tag{D.7}$$

where $\vec{c}_i \in \mathbb{R}^{\dim(\vec{P})}$ and $h_i(\vec{q})$ is a polynomial in the coefficients of $\vec{q}$.

D.5 Solving the membership problem

For all polynomials $F_i(\vec{P}, \vec{q})$, $G_{jk}(\vec{P}, \vec{q})$ and $L_l(\vec{P}, \vec{q})$, the following expression is nonnegative:

$$I = \sum_i F_i f_i + \sum_{jk} (G_{jk})^2 g_j + \sum_l (L_l)^2 \geq 0. \tag{D.8}$$

In comparison with (D.2), we now have polynomial coefficients, and we employed squared polynomials to enforce nonnegativity.[f] When all the monomials in $\vec{q}$ cancel, we are left with a polynomial inequality in $\vec{P}$ valid for all network-local behaviors. As with the linear program (D.3), we can detect points outside $\mathcal{L}$ by the minimization:

$$\begin{aligned} \nu = \min_{\{F_i\},\{G_{jk}\},\{L_l\}} \quad & I(\vec{P}) \qquad \text{such that} \\ & I(\vec{P}, \vec{q}) = I(\vec{P}) \end{aligned} \tag{D.9}$$

which can be formulated as a semidefinite program by upper bounding the degree of all $\{F_i\}$, $\{G_{jk}\}$, $\{L_l\}$. By increasing this bound step by step, we get a semidefinite hierarchy, proven to converge [29,30] .

In the triangle network, what are the requirements of this hierarchy for the bound of $|\Omega_\alpha| = |\Omega_\beta| = |\Omega_\gamma| 6$ proven in Proposition 3 ? First, observe that Eq. (1) can be written:

$$P(abc) = \sum_{\alpha,\beta,\gamma=1}^{6} P(a\beta|\gamma) P(b\gamma|\alpha) P(c\alpha|\beta), \tag{D.10}$$

with $P(a\beta|\gamma) = P(\beta)P(a|\beta\gamma), \ldots$ and the constraints suitably modified. The number of degrees of freedom of $P(a\beta|\gamma)$ is at least $r^2 + r - 1$, where $r = 6$ is the rank of the local hidden variables, with the total number of degrees of freedom $\mathcal{D} = 3(r^2 + r - 1) = 123$. The highest degree of involved polynomials is 3 in Eq. (D.10). Thus the relaxation should be of degree at least 2, and involve semidefinite matrices of row and column size $\mathcal{D}(\mathcal{D}+1)/2 = 7626$, which is out of our reach. However, the complexity decreases rapidly with r: for $r = 5$, we obtain the size 3828×3828, while $r = 4$ has size 1653×1653. Thus, better upper bounds are not only interesting in theory; they render the membership problem tractable in practice.

D.6 Computing a complete description

We can also characterize the set $\mathcal{L}$ by considering:

$$X = \left\{ \vec{x} = (\vec{P}, \vec{q}) \text{ s.t. } f_i(\vec{P}, \vec{q}) = 0 \text{ and } g_j(\vec{q}) \geq 0 \right\}, \tag{D.11}$$

[f]Being a sum-of-squares is a sufficient condition for a polynomial to be nonnegative.

$=\langle ..\rangle$	1	C_0	C_1
1		$1-\eta$	$1-\eta$
A_0	$\eta-1$	$-(\eta-1)^2$	$-(\eta-1)^2$
A_1	$1-\eta$	$(\eta-1)^2$	$(\eta-1)^2$

$=\langle .B_0.\rangle$	1	C_0	C_1
1	0	0	0
A_0	0	$\eta^2/2$	$-\eta^2/2$
A_1	0	$-\eta^2/2$	$\eta^2/2$

$=\langle .B_1.\rangle$	1	C_0	C_1
1	0	0	0
A_0	0	$\eta^2/2$	$\eta^2/2$
A_1	0	$\eta^2/2$	$\eta^2/2$

$=\langle .B_0B_1.\rangle$	1	C_0	C_1
1	0	0	0
A_0	0	0	0
A_1	0	0	0

Table E.1. All correlators involved in a 2-locality test involving two singlet states and inefficient detectors for Alice and Charlie, with efficiency η.

$=\langle ..\rangle$	1	C_0	C_1	C_0C_1
1		$1-\eta$	$1-\eta$	ξ
A_0	$\eta-1$	$-(\eta-1)^2$	$-(\eta-1)^2$	$(\eta-1)\xi$
A_1	$1-\eta$	$(\eta-1)^2$	$(\eta-1)^2$	$(1-\eta)\xi$
A_0A_1	ζ	$(1-\eta)\zeta$	$(1-\eta)\zeta$	$\xi\zeta$

$=\langle .B_0.\rangle$	1	C_0	C_1	C_0C_1
1	0	0	0	0
A_0	0	$\eta^2/2$	$-\eta^2/2$	0
A_1	0	$-\eta^2/2$	$\eta^2/2$	0
A_0A_1	0	f_2	$-f_2$	0

$=\langle .B_1.\rangle$	1	C_0	C_1	C_0C_1
1	0	0	0	0
A_0	0	$\eta^2/2$	$\eta^2/2$	f_1
A_1	0	$\eta^2/2$	$\eta^2/2$	f_1
A_0A_1	0	0	0	0

$=\langle .B_0B_1.\rangle$	1	C_0	C_1	C_0C_1
1	0	0	0	0
A_0	0	0	0	0
A_1	0	0	0	0
A_0A_1	0	0	0	0

Table E.2. Bilocal model for the correlations after symmetrization. The number of variables has been greatly reduced.

which a subset of $\mathbb{R}^{\dim(\vec{P})+\dim(\vec{q})}$ characterized by polynomial (in)equalities: thus a semialgebraic set [31] . Its projection $X_{|\vec{P}}$ on the variables $\vec{P}$ is also a semialgebraic set. As X is a compact set, by the Tube Lemma [32] , $X_{|\vec{P}}$ is bounded and closed as well. Thus $\mathcal{L}$ is a semialgebraic closed set. Using the Finiteness Theorem [31, Thm 2.7.1] , $\mathcal{L}$ can be written as a finite union:

$$\mathcal{L} = X_{|\vec{P}} = \bigcup_i \mathcal{L}_i \ , \qquad \mathcal{L}_i = \left\{ \vec{P} \in \mathbb{R}^{\dim(\vec{P})} \text{ s.t. } f_{ij}(\vec{P}) \geq 0, \forall j \in J_i \right\} \ , \tag{D.12}$$

and $|J_i| A(A+1)/2$, where A is the affine dimension of the nonsignaling space of behaviors $\vec{P}$, and the bound on the number of inequalities is given by the Bröcker-Scheiderer Theorem [31, Thm 10.4.8] . We also know that $\mathcal{L}$ is connected: by including a noise parameter in each local hidden variable, as done in the proof of Proposition 3, it is possible to connect all behaviors to the uniformly random distribution.

In principle, $X_{|\vec{P}}$ can be computed using the cylindrical algebraic decomposition algorithm [33] , such as implemented by the `Reduce` function of Mathematica. In practice, the process is extremely demanding and can only be completed in very simple cases [34] ; moreover the output seldom matches the nice form (D.12).

In the next Appendix, we demonstrate the usefulness of sum-of-squares in the context of networks, by proving algebraically a numerical result provided in [7] .

Appendix E Example of a sum-of-squares proof

We give below an example of sum-of-squares relaxations applied to the characterization of network-local sets, taken from the section III.C.2 of [7] . There, the authors studied the correlations coming from an entanglement-swapping experiment and its resistance to detector inefficiencies. However, the result $\eta_{\text{biloc}} = 2/3$ was obtained numerically. We provide an algebraic proof below using a sum-of-squares decomposition.

E.1 2-local model

First, we summarize the 2-local model considered in [7] . We consider an entanglement swapping experiment, where the source $\mathcal{S}_1$ is connected to Alice and Bob, while the source $\mathcal{S}_2$ is connected to Bob and Charlie. Bob does not receive an input, but outputs two bits, $B_0 = \pm 1$ and $B_1 = \pm 1$. Alice and Charlie have binary inputs $x, y = 0, 1$ with corresponding outputs $A = \pm 1$ and $C = \pm 1$. From the bilocality condition, we write:

$$P(AB_0B_1C|xz) = \int P(A|x\lambda_1)P(B_0B_1|\lambda_1\lambda_2)P(C|z\lambda_2)\mathrm{d}\pi_1(\lambda_1)\mathrm{d}\pi_2(\lambda_2). \tag{E.1}$$

In [7] , this was shown to be equivalent to the existence of an underlying distribution of A_0, A_1, B_0, B_1, C_0, C_1, where A_x and C_z are the deterministic values of A and C for all inputs. As these are ± 1-valued variables, we can compute the averages $\langle A_0^{i_0} A_1^{i_1} B_0^{j_0} B_1^{j_1} C_0^{k_0} C_1^{k_1} \rangle$ for $i_0, i_1, j_0, j_1, k_0, k_1 = 0, 1$. As A_0, A_1 and C_0, C_1 can all be observed at the same time, the following distribution exists and is nonnegative:

$$\begin{aligned} &P(A_0{=}\alpha_0, A_1{=}\alpha_1, B_0{=}\beta_0, B_1{=}\beta_1, C_0{=}\gamma_0, C_1{=}\gamma_1) \\ &\qquad = \frac{1}{64}\Big\langle (1+\alpha_0 A_0)(1+\alpha_1 A_1)(1+\beta_0 B_0)(1+\beta_1 B_1)(1+\gamma_0 C_0)(1+\gamma_1 C_1) \Big\rangle \geq 0, \end{aligned} \tag{E.2}$$

and the independence of Λ_1, Λ_2 reduces to:

$$\langle A_0^{i_0} A_1^{i_1} C_0^{k_0} C_1 \rangle = \langle A_0^{i_0} A_1^{i_1} \rangle \langle C_0^{k_0} C_1^{k_1} \rangle \tag{E.3}$$

for all $i_0, i_1, k_0, k_1 = 0, 1$. The $\langle \ldots \rangle$ containing both A_0A_1 or both C_0C_1 cannot be obtained from the distribution $P(AB_0B_1C|xz)$ and are unknown parameters of the model.

E.2 Quantum correlations to test

The quantum correlations are obtained from the state $\rho_{\text{ABC}} = \rho_{\text{AB}} \otimes \rho_{\text{BC}} = |\Psi^-\rangle\langle\Psi^-| \otimes |\Psi^-\rangle\langle\Psi^-|$ distributed by the sources $\mathcal{S}_1$ and $\mathcal{S}_2$. Alice and Charlie have both inefficient detectors with detection efficiency η. When $\eta = 1$, Alice and Charlie use the projective measurements with ± 1-valued outcomes:

$$\overline{A}_0 = \overline{C}_0 = \frac{\sigma_z + \sigma_x}{\sqrt{2}}, \qquad \overline{A}_1 = \overline{C}_1 = \frac{\sigma_z - \sigma_x}{\sqrt{2}}, \tag{E.4}$$

but in case of nondetection, Alice outputs $A = (-1)^{x+1}$ and Charlie always outputs $C = +1$. Thus, their effective measurement operators are:

$$A_0 = \eta\overline{A}_0 - (1-\eta)1, \qquad A_1 = \eta\overline{A}_1 + (1-\eta)1, \qquad C_z = \eta\overline{C}_z + (1-\eta)1\ . \tag{E.5}$$

Bob performs a Bell state measurement $\{\mathcal{B}_{B_0B_1}\}$ with four outcomes, described by two bits $B_0, B_1 = \pm 1$:

$$\mathcal{B}_{++} = |\Phi^+\rangle\langle\Phi^+|, \quad \mathcal{B}_{-+} = |\Phi^-\rangle\langle\Phi^-|, \quad \mathcal{B}_{+-} = |\Psi^+\rangle\langle\Psi^+|, \quad \mathcal{B}_{--} = |\Psi^-\rangle\langle\Psi^-|, \tag{E.6}$$

so that the value of the bits B_0, B_1 and their parity B_0B_1, are given by:

$$B_0 = \mathcal{B}_{++} + \mathcal{B}_{+-} - \mathcal{B}_{-+} - \mathcal{B}_{--}, \quad B_1 = \mathcal{B}_{++} - \mathcal{B}_{+-} + \mathcal{B}_{-+} - \mathcal{B}_{--}, \quad B_0B_1 = \mathcal{B}_{++} - \mathcal{B}_{+-} - \mathcal{B}_{-+} + \mathcal{B}_{--}. \tag{E.7}$$

The corresponding correlations are given in Table E.1.

E.3 Simplified 2-local model

We find that the correlations are invariant under a symmetry group G of order 8, with generators g_{AB}, g_{BC} and g_{ABC}:

$$g_{\mathrm{AB}} = \begin{cases} A_0 \to & -A_1 \\ A_1 \to & -A_0 \\ B_1 \to & -B_1 \end{cases}, \quad g_{\mathrm{BC}} = \begin{cases} C_0 \to & C_1 \\ C_1 \to & C_0 \\ B_0 \to & -B_0 \end{cases}, \quad g_{\mathrm{ABC}} = \begin{cases} A_0 \to & -C_0 \\ A_1 \to & C_1 \\ C_0 \to & -A_0 \\ C_1 \to & A_1 \\ B_0 \to & B_1 \\ B_1 \to & B_0 \end{cases}. \tag{E.8}$$

Both g_{AB} and g_{BC} can be applied on any bilocal model by embedding a flag in $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively, while g_{ABC} is not compatible with the model. We thus apply the symmetries g_{AB} and g_{BC} on our bilocal model filled with the values of Table E.1 to obtain the correlations in Table E.2, written using four unknowns ξ, ζ, f_1, f_2 and the detection efficiency η. We thus need to find the range of η such that values for the unknown variables exist and Eq. (E.2) is satisfied (the other constraints are already taken care of).

The solution ($\eta \in [0, 2/3]$) was found using a sum-of-squares relaxation with the solver SOSTOOLS [35] ; we then worked the analytic solution below from the numerical output of the solver.

E.4 Sum-of-squares solution

Among the inequalities coming from (E.2), the following inequalities are relevant, writing $\bar{\xi} = 1 - \xi$ and $\bar{\zeta} = 1 + \zeta$:

$$\begin{array}{lllll} 0 g_1 = & 64P(++++-+) & = & \bar{\xi}\bar{\zeta} - 2(f_1 + f_2), \\ 0 g_2 = & 64P(-+-+-+) & = & 4\bar{\xi} - \xi\zeta - 2\eta^2 - 2\bar{\xi}\eta - 2f_2, \\ 0 g_3 = & 64P(+-++--) & = & \bar{\xi}\bar{\zeta} - 2\eta\bar{\zeta} - 2\bar{\xi}\eta + 4\eta^2, \\ 0 g_4 = & 64P(-+++--) & = & \bar{\xi}\bar{\zeta} + \bar{\xi}(2\eta - 4) - \bar{\zeta}2\eta + 8\eta - 2\eta^2, \\ 0 g_5 = & 64P(+-++-+) & = & 2\bar{\xi}\eta - \bar{\xi}\bar{\zeta} - 2\eta^2 + 2f_2, \\ 0 g_6 = & 64P(++++--) & = & 2\bar{\zeta}\eta - \bar{\xi}\bar{\zeta} - 2\eta^2 + 2f_1. \end{array} \tag{E.9}$$

We form the conic combinations:

$$\begin{array}{llll} 0 & F_+ & = g_1 + 2g_3 + g_5 + g_6 & = \left(2\eta - \bar{\zeta}\right)\left(2\eta - \bar{\xi}\right), \\ 0 & F_- & = 2g_1 + g_3 + 2g_5 + 2g_6 & = -\left(2\eta - \bar{\zeta}\right)\left(2\eta - \bar{\xi}\right), \\ 0 & I & = (2g_1 + g_2 + 3g_5 + 2g_6) & = \bar{\xi} + \eta\bar{\xi} + \eta\bar{\zeta} - \bar{\xi}\bar{\zeta} - 3\eta^2, \\ 0 & J & = (3g_3 + g_4)/4 & = \left(\eta - \bar{\zeta} + 1\right)\left(2\eta - \bar{\xi}\right), \end{array}$$

From $F_\pm$, we deduce that either $\left(2\eta - \bar{\zeta}\right) = 0$ or $\left(2\eta - \bar{\xi}\right) = 0$. We examine first $\bar{\zeta} = 2\eta$, and substitute:

$$I = \bar{\xi}(1 - \eta) - \eta^2 \geq 0, \quad J = (1 - \eta)\left(2\eta - \bar{\xi}\right) \geq 0. \tag{E.10}$$

Definition E.1 *A polynomial* p *to be a* sum-of-squares polynomial *(SOS) if there is a decomposition:*

$$p = \sum_k (m_k)^2, \tag{E.11}$$

where m_k *are polynomials in* ξ, η.

For SOS polynomials p_1, p_2 and p_3, we have $K = p_1 + p_2 I + p_3 J \geq 0$. To find a bound on the detection efficiency η_{bound}, we write:

$$K = \eta_{\mathrm{bound}} - \eta = p_1 + p_2 I + p_3 J \geq 0, \tag{E.12}$$

which will prove that bilocal models satisfy $\eta\eta_{\mathrm{bound}}$. The bound $\eta_{\mathrm{bound}} = 2/3$ is verified using:

$$p_1 = \frac{(2 - 3\eta)^2}{6}, \quad p_2 = p_3 = \frac{1}{2}. \tag{E.13}$$

The procedure for $\bar{\xi} = 2\eta$ is similar and also gives $\eta_{\mathrm{bound}} = 2/3$.