Polynomial Bell Inequalities

Rafael Chaves
Institute for Physics & FDM, University of Freiburg, 79104 Freiburg, Germany
and Institute for Theoretical Physics, University of Cologne, 50937 Cologne, Germany
(Received 17 June 2015; published 7 January 2016)

It is a recent realization that many of the concepts and tools of causal discovery in machine learning are highly relevant to problems in quantum information, in particular quantum nonlocality. The crucial ingredient in the connection between both fields is the mathematical theory of causality, allowing for the representation of arbitrary causal structures and providing a rigorous tool to reason about probabilistic causation. Indeed, Bell's theorem concerns a very particular kind of causal structure and Bell inequalities are a special case of linear constraints following from such models. It is thus natural to look for generalizations involving more complex Bell scenarios. The problem, however, relies on the fact that such generalized scenarios are characterized by polynomial Bell inequalities and no current method is available to derive them beyond very simple cases. In this work, we make a significant step in that direction, providing a new, general, and conceptually clear method for the derivation of polynomial Bell inequalities in a wide class of scenarios. We also show how our construction can be used to allow for relaxations of causal constraints and naturally gives rise to a notion of nonsignaling in generalized Bell networks.

DOI: 10.1103/PhysRevLett.116.010402

Bell's theorem [1] demonstrates that our classical conceptions of causal relations must be taken with care, as they fail to commit with the results obtained in quantum experiments performed by distant parties, the phenomenon of quantum nonlocality. Remarkably, in order to prove the emergence of nonlocal correlations, it is sufficient to consider a very simple causal structure, where due to the distance between the parties it is natural to assume, at least classically, that the correlations between them are mediated via causal influences originating in a common local hidden variable (LHV). Yet, quantum theory predicts that the correlations obtained via local measurements performed on distant entangled particles are incompatible with any classical theory fulfilling such a natural causal description.

Recently, it has been realized that one can significantly expand the notion of quantum nonlocality by considering more complex causal structures going beyond the usual LHV models [2–10]. At the basis of this new research program lies the mathematical theory of causality [11,12], which provides a rigorous and systematic way to reason about causal relations and causal structures. This realization has already lead to new insights about the tension between quantum mechanics and causality [6,13–16] and reveals a much richer structure of quantum correlations than the one we could naively presume from Bell's paradigmatic causal structure alone [2–9].

As an illustration of such generalized scenarios and its applications, consider an entanglement swapping experiment [17]. Starting with two independent pairs of entangled particles and jointly measuring one particle from each pair, we can generate entanglement and nonlocal correlations between the two remaining particles, even though they have never interacted. In this case, to contrast the quantum and classical descriptions we have to consider a finer structure for the underlying causal model that should now be described by uncorrelated LHVs and thus introduces additional structure to the set of allowed correlations [3]. Generally, since models with many independent sources are more restrictive to classical explanations, they offer a novel new route to decrease the requirements on experimental implementations of Bell tests [3,4,18]. In fact, as shown in Refs. [19,20], arranging multiple copies of an entangled (but local at the single-copy level) state into complex networks may reveal its nonlocality. Furthermore, scenarios with many independent sources of quantum states are ubiquitous in quantum information, e.g., quantum networks [21,22] and quantum repeaters [23]. Thus, understanding generalized Bell scenarios is not only of fundamental interest but also of high practical relevance.

Within that context, the basic question to be solved is how to derive Bell inequalities for general Bell scenarios, whose classical causal description will be named here as generalized local hidden variables (GLHV) models. Bell inequalities play a fundamental role in the study of nonlocality, since it is via their violation that we can witness the nonlocal character of experimental data. Unfortunately, opposed to usual Bell scenarios, GLHV models imply a nonconvex region—characterized by polynomial Bell inequalities—of correlations that are compatible with it. Given this difficulty, only sparse results have been obtained for GLHV models, either using coarse-grained information [6,8,24–26] or considering particular scenarios and techniques of limited application [3,4,27,28]. In causal inference, techniques from algebraic geometry (AG) were

 © 2016 American Physical Society

shown to provide, in principle, a general solution to this problem [29]. Unfortunately, given their double exponential computational complexity [30–33], their application to Bell scenarios is intractable already for the simplest possible models [34,35]. Even the most computationally amenable tools from AG [34,36,37] are unable to characterize structures beyond 5 binary variables [34]. It is clear that in spite of the existence of general purpose methods, they are far from delivering a relevant and practical tool for the study of the emergence of nonlocality in complex causal structures.

In this Letter we propose a new and general method for deriving polynomial inequalities in generalized Bell scenarios. As opposed to other available methods requiring a high level of expertise in algebraic geometry [29,34], our approach involves basic concepts from convex optimization and thus provides a more accessible tool for the analytic derivation of inequalities in a variety of scenarios. Moreover, in spite of the intrinsic high computational complexity of the problem, our approach is computationally more accessible than previous attempts. Finally, our construction allows for relaxations of causal constraints and, as shown in the Supplemental Material [38], naturally introduces a notion of nonsignaling [45] in generalized scenarios.

Bell inequalities, causal structures, and algebraic geometry.—Bell scenarios beyond LHV models can be represented via the graphical notation of directed acyclic graphs (DAG), where nodes stand for variables and directed arrows represent causal relations [11,12]. LHV models correspond to DAGs with a single hidden variable. For instance, the DAG in Fig. 1(a) represents the usual causal structure from a Bell experiment, where a common source produces particles emitted to two observers that at each

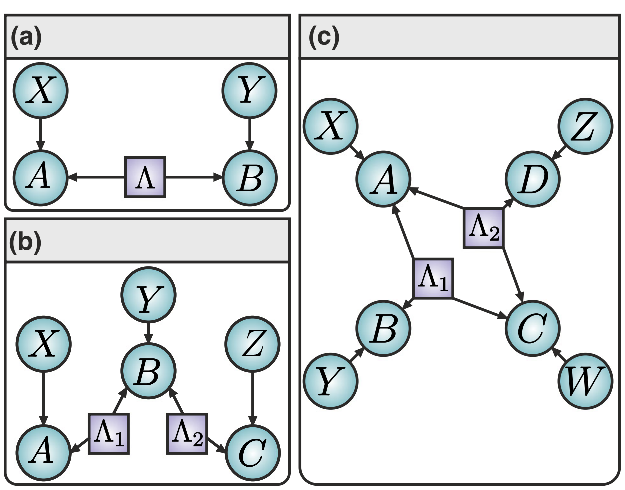

FIG. 1. DAG representation of causal structures. (a) Bipartite LHV model. (b) Tripartite GLHV model with 2 independent hidden variables representing an entanglement swapping experiment [3,17]. (c) Four-partite GLHV model with 2 hidden variables.

round of the experiment measure a given observable, labeled by X and Y, respectively, obtaining outcomes A and B. GLHV models have a similar physical intuition, the difference being that they are represented by DAGs with $n \geq 2$ independent hidden variables [Figs. 1(b)–1(c)]. The causal relations implied by DAGs are captured by the (conditional) independencies (CI) implied by the graph [11]. For instance, for the LHV model in Fig. 1(a) it follows that $p(x,y,\lambda) = p(x)p(y)p(\lambda)$ and $p(a|x,y,\lambda) = p(a|x,\lambda)$ (similarly to b).

To contrast the difference in the geometry of correlations of LHV and GLHV models, consider the models in Figs. 1(a)–1(b). From the CIs implied by the DAG in Fig. 1(a) it follows that any observable data $p(a,b|x,y)$ compatible with a LHV model can be decomposed as

$$p(a,b|x,y) = \sum_{\lambda} p(a|x,\lambda)p(b|y,\lambda)p(\lambda). \qquad (1)$$

That is, any LHV distribution lies inside the convex set defined by Eq. (1), the correlation polytope $\mathbb{C}$ [46,47]. In this geometric picture, (linear) Bell inequalities are nothing other than facets of $\mathbb{C}$. Given the list of the extremal points of $\mathbb{C}$, finding its facets amounts to dualize the description of the polytope, the facet enumeration problem.

In turn, from the DAG in Fig. 1(b), it follows that any GLHV model compatible with it can be written as

$$p(a,b,c|x,y,z) = \sum_{\lambda_1,\lambda_2} p(a|x,\lambda_1)p(b|y,\lambda_1,\lambda_2)p(c|z,\lambda_2)p(\lambda_1)p(\lambda_2). \qquad (2)$$

Because of the independence of the underlying hidden sources $[p(\lambda_1,\lambda_2) = p(\lambda_1)p(\lambda_2)]$, Eq. (2) implies a nonconvex region. Therefore, the techniques developed for LHV models can no longer be applied.

In this case, one can, in principle, resort to the AG approach [29], where the constraints implied by a DAG are encoded in a semialgebraic set, a list of polynomial equalities and inequalities in all variables composing the DAG. Given that some of the variables are not observable they need to be eliminated from our description. Formally, the problem is equivalent to quantifier elimination: the projection of a semialgebraic set onto a subspace of it, that by the Tarski-Seidenberg theorem is again a semialgebraic set [29,30]. Via quantifier elimination we obtain a full description, in terms of polynomial inequalities, of the marginal scenario of interest associated with any DAG. The problem with the usual methods [29,34,36] resides in its complexity, that not only is it double exponential, but also depends on the domain size of all model variables, including hidden ones. Even for the simplest possible Bell scenario, the bipartite LHV model in Fig. 1(a) with dichotomic inputs and outputs $(x,y,a,b = 0,\ 1)$, the hidden variable is 16-dimensional, implying a total 256-dimensional object, far beyond computational reach

[34,35]. It is clear that AG methods, while allowing for the characterization of arbitrary DAGs, are not a viable option in the study of Bell scenarios.

New method for deriving polynomial Bell inequalities.— To circumvent this problem, our method relies on the realization that Bell inequalities can be seen as constraints arising from a marginal problem [48]: given some marginal distributions of n variables is it possible to find a joint distribution of all variables, such that this distribution marginalizes to the given ones? To see that Bell's theorem is indeed a particular marginal problem, notice that the LHV description (1) is equivalent to the existence of a joint distribution $\mathbf{p} = p(a_0, \ldots, a_{m_x}, b_0, \ldots, b_{m_y})$ (represented as a vector $\mathbf{p}$) describing the probability for outcomes of all possible measurements, where a_i labels the outcome a given that $x = i = \{0, \ldots, m_x\}$ and similarly for b. Since $\mathbf{p}$ defines a valid probability, it is constrained by a set of linear inequalities $L\mathbf{p} \geq 0$ given by $p_i \geq 0$ (positivity) and $\sum_i p_i = 1$ (normalization) defining the simplex polytope $\mathbb{P}$ [49]. Moreover, given that at each round of the experiment only one a_i and one b_j can be measured simultaneously, $\mathbf{p}$ defines a nonobservable quantity. However, the constraints on $\mathbf{p}$ will also imply constraints on the level of the observable distributions $p(a_i, b_j)$. These are exactly Bell inequalities, that in this picture can be understood as necessary and sufficient conditions for the marginal problem to have a positive answer.

Similarly, a GLHV model also implies the existence of a joint distribution $\mathbf{p}$ characterized by linear inequalities $L\mathbf{p} \geq 0$. The difference being that GLHV models also imply a set of polynomial inequalities $W(\mathbf{p})\mathbf{p} \geq 0$ as described by the CIs implied by the associated DAG. This is the case in the DAG of Fig. 1(b) implying the independence relation $p(a, c) = p(a)p(c)$ and in many other relevant scenarios in quantum information [3,4,6–10,17,19–23]. In these cases, the GLHV model is characterized by the intersection of $\mathbb{P}$ with $W(\mathbf{p})\mathbf{p} \geq 0$, again a semialgebraic set but one that now does not depend on hidden variables of the DAG anymore. Clearly, this enormously reduces the number of variables requiring double exponential complexity algorithms to be eliminated. On the negative side, not all DAGs display CIs on the level of $\mathbf{p}$; in this case, AG methods are likely to be the only possible route.

Our approach consists of performing a quantifier elimination (over unobservable terms) in the joint system of inequalities $L\mathbf{p} \geq 0$ and $W(\mathbf{p})\mathbf{p} \geq 0$, to which one could use standard techniques such as cylindrical algebraic decomposition [50]. Instead, we propose a new quantifier elimination procedure that combines facet enumeration and the Fourier-Motzkin (FM) elimination [51], basic tools from convex optimization—therefore avoiding the conceptual and technical challenges from an AG description.

The method proceed as follows. Given the scenario of interest, we need to define $\mathbf{p}_{\mathrm{O}}$ and $\mathbf{p}_{\mathrm{NO}}$ standing, respectively, to the set of components p_i that we want to keep or not in our description. We also need to define $\mathbf{p}_{\mathrm{LNO}}$ and $\mathbf{p}_{\mathrm{WNO}}$. The first corresponds to components in $\mathbf{p}_{\mathrm{NO}}$ appearing only in $L\mathbf{p} \geq 0$, the latter describes components in $\mathbf{p}_{\mathrm{NO}}$ also appearing in $W(\mathbf{p})\mathbf{p} \geq 0$. Terms $\mathbf{p}_{\mathrm{LNO}}$ can be eliminated via FM elimination over $L\mathbf{p} \geq 0$, obtaining a new set of linear relations $L'\mathbf{p} \geq 0$. Equivalently, to obtain $L'\mathbf{p} \geq 0$, we can instead resort to the usual facet enumeration procedure.

The terms in $\mathbf{p}_{\mathrm{WNO}}$ have to be eliminated considering $L'\mathbf{p} \geq 0$ and $W(\mathbf{p})\mathbf{p} \geq 0$ jointly. To that aim, notice that this system of inequalities can be linearized by considering some of the variables as free parameters of the problem. Given $W(\mathbf{p})\mathbf{p} \geq 0$ there is going to be a minimum set of nonobservable variables $\mathbf{p}_{\mathrm{FWNO}}$ that need to be set to free parameters in order to linearize the problem. Chosen $\mathbf{p}_{\mathrm{FWNO}}$ we can perform a FM elimination over the remaining terms in $\mathbf{p}_{\mathrm{WNO}}$, obtaining a new system of inequalities $W'(\mathbf{p})\mathbf{p} \geq 0$ depending nonlinearly on terms $\mathbf{p}_{\mathrm{FWNO}}$. To obtain inequalities depending on the observable data only we have to further eliminate the terms $\mathbf{p}_{\mathrm{FWNO}}$. In this last step, the algorithm relies on the usual quantifier elimination methods. We highlight that this quantifier elimination will be performed on a much smaller number of variables given by $|\mathbf{p}_{\mathrm{FWNO}}|$. In fact, for the bilocality scenario [Fig. 1(b)] this last step will involve the elimination of a single variable only, clearly illustrating the huge computational benefit of this approach.

Examples.—Our method provides a computationally more accessible route as compared to usual approaches from AG. However, given the intrinsic high complexity of the problem, its computational treatment is still bounded to cases with few variables. This is similar to what happens in the usual LHV models, since the full characterization of even the seemingly simple Bell causal structure in Fig. 1(a) is in general computationally intractable [46]. To be able to obtain Bell inequalities for more complex LHV models [52–55], the conceptually simple and clear geometric understanding of Bell polytopes [46] has become an indispensable tool. Our approach provides the equivalent conceptually simple description of GLHV models and as shown next (and detailed in Ref. [38]) allows for the straightforward derivation of polynomial inequalities in a variety of scenarios.

For simplicity, in the following we focus on dichotomic outcomes (e.g., $a_i = 0, 1$). It is then convenient to consider the equivalent description of the problem in terms of the correlation vector $\mathbf{E}$ with components given by expectation values, e.g., $\langle A_i B_j \rangle = \sum_{a_i, b_j} (-1)^{a_i + b_j} p(a_i, b_j)$. The vectors $\mathbf{E}$ and $\mathbf{p}$ are linearly related as $\mathbf{E} = T^{-1}\mathbf{p}$ implying that $\mathbf{E}$ must obey linear inequalities $T\mathbf{E} \geq 0$.

Consider, for instance, the following question: given a certain Bell inequality valid for LHV models, how is it modified in the presence of further causal constraints? For example, how is the LHV inequality for a tripartite system

$$|I| + |J| \leq 4, \tag{3}$$

where $I = \sum_{x,z=0,1} \langle A_x B_0 C_z \rangle$ and $J = \sum_{x,z=0,1} (-1)^{x+z} \times \langle A_x B_1 C_z \rangle$, modified if we assume the bilocality constraint [3] following from the DAG in Fig. 1(b)? To prove that Eq. (3) is a valid inequality it is sufficient to consider the inequalities following from $T\mathbf{E} \geq 0$:

$$\pm I - \langle A_0 A_1 \rangle - \langle C_0 C_1 \rangle - \langle A_0 A_1 C_0 C_1 \rangle \leq 1, \tag{4}$$

$$\pm J + \langle A_0 A_1 \rangle + \langle C_0 C_1 \rangle - \langle A_0 A_1 C_0 C_1 \rangle \leq 1. \tag{5}$$

Summing these inequalities and using $\langle A_0 A_1 C_0 C_1 \rangle \leq 1$ we obtain Eq. (3). Instead, considering the independence constraint following from the DAG in Fig. 1(b),

$$\langle A_0 A_1 C_0 C_1 \rangle = \langle A_0 A_1 \rangle \langle C_0 C_1 \rangle, \tag{6}$$

and substituting it in Eqs. (4) and (5), after some manipulations we can combine them into the polynomial inequality

$$2\langle A_0 A_1 \rangle^2 + (\pm J \mp I)\langle A_0 A_1 \rangle - (\pm I \pm J + 2) \leq 0. \tag{7}$$

As discussed in the general algorithm, we arrive at an inequality having a nonlinear dependence on nonobservable terms, in this case $\langle A_0 A_1 \rangle$. Here, the quantifier elimination of the unobservable term $\langle A_0 A_1 \rangle$ corresponds simply to solve the quadratic equation (7). The minimum of the polynomial in Eq. (7) is achieved at $\langle A_0 A_1 \rangle = (\pm I \mp J)/4$, implying an inequality in terms of observable data only:

$$-(1/8)(\pm I - \mp J)^2 - (\pm I \pm J + 2) \leq 0, \tag{8}$$

equivalent to the bilocality inequality derived in Refs. [3,4,27], using much more involved and less general techniques. Consider the correlation $I = J = 2$ that can be achieved quantum mechanically with two copies of Bell states shared between the parties [27]. This correlation admits a LHV model but violates Eq. (8), illustrating the fact that GLHV models are more restrictive to classical explanations and therefore can witness a larger class of nonlocal correlations.

Another nice feature of our construction is the fact that independencies are not required to hold exactly: we can quantify how much a given constraint must be relaxed in order to explain some experimental data [15,56]. In the bilocality scenario, allowing for correlations $\mathcal{C}_{AC} \geq |\langle A_0 A_1 C_0 C_1 \rangle - \langle A_0 A_1 \rangle \langle C_0 C_1 \rangle|$ between A and C, it follows that

$$-(1/8)(\pm I - \mp J)^2 - (\pm I \pm J + 2) \leq 2\mathcal{C}_{AC}; \tag{9}$$

that is, the violation of Eq. (8) quantifies the degree of correlation required to classically reproduce some non-bilocal correlation. Considering again $I = J = 2$, we see that it requires a maximum correlation $\mathcal{C}_{AC} = 1$ to be reproduced.

To further illustrate the practicality of our method we also analytically derive in Ref. [38] new polynomial inequalities for scenarios with more measurement settings or parties. The relevance of such scenarios stems from the fact that they typically allow for reductions in the experimental requirements (e.g., detection efficiency) for the observation of nonlocality [55,57–59]. Using our method to derive new inequalities and analyze how they may reduce experimental requirements is an interesting topic for future research.

For instance, considering the four-partite scenario in Fig. 1(c), it follows that a similar inequality to Eq. (8) also holds. The DAG in Fig. 1(c) involves 10 nodes, being completely intractable with computational tools. In turn, considering the bilocality scenario in Fig. 1(b) with 3 measurement settings, it follows that

$$-(1/8)(I - J + 16)^2 + 8I \leq 0 \tag{10}$$

holds, with $I = \sum_{x,z=0,1,2} \langle A_x B_0 C_z \rangle$ and $J = \sum_{x,z=0,1,2} (-1)^{x+z} \langle A_x B_1 C_z \rangle$. Choosing $I = J = 9v$ (achievable in quantum mechanics for $v \leq 1/2$) it follows that for $v \leq 5/9$ the correlation is local. However, Eq. (10) is violated for $v > 4/9$, illustrating the gap between the local and bilocal sets in this case.

Discussion.—Complex causal structures beyond the usual LHV models offer an almost unexplored territory for generalizations of Bell's theorem. The basic question to be solved in this quest is how to derive polynomial Bell inequalities bounding the classical correlations compatible with them. In this work we made an important step in that direction. We proposed a new and general method that can be readily applied to a wide range of scenarios, considering its application in few GLHV models and deriving polynomial Bell inequalities characterizing them. Our approach not only provides a more accessible computational route but also, given its conceptual clarity, allows for analytical derivations of polynomial Bell inequalities. Furthermore, it allows for relaxations of causal constraints and naturally leads to a notion of nonsignaling correlations in GLHV models [38].

Given the fundamental role that Bell inequalities play in the study and practical applications of nonlocality, we believe that our results will motivate and set a basic tool for future research. The natural next step is to put the machinery to use in a variety of scenarios, understanding how known inequalities are modified in the presence of extra underlying causal constraints and deriving new inequalities well suited, for example, to decrease the requirements on experimental implementations. It would be interesting to investigate the role of polynomial Bell inequalities in practical applications of nonlocality, such as quantum cryptography [60], randomness generation [61,62], or distributed computing [63]. For instance, the amount of violation of usual Bell inequalities can be directly associated with the probability of success in communication complexity problems [63,64]. Are there any communication problems associated with polynomial Bell inequalities? Another possibility is to find Tsirelson's bounds [65,66] associated with generalized inequalities, that is, their maximum violation achievable with quantum correlations. Related to that and inspired by results such as information causality [67], it would also be relevant to derive information-theoretical principles for these more complex causal structures [9].

We acknowledge financial support from the Excellence Initiative of the German Federal and State Governments (Grants No ZUK 43 and 81), the U.S. Army Research Office under Contracts No. W911NF-14-1-0098 and No. W911NF-14-1-0133 (Quantum Characterization, Verification, and Validation), and the DFG (GRO 4334 and SPP 1798).

Note added.—After completion of this work, a related and complementary work discussing polynomial Bell inequalities has appeared [68]. Also, algebraic geometry methods have been recently proposed in Ref. [37]. Combining these and the present results opens interesting lines for future research.

[1] J. S. Bell, Physics **1**, 195 (1964).
[2] S. Popescu, Phys. Rev. Lett. **74**, 2619 (1995).
[3] C. Branciard, N. Gisin, and S. Pironio, Phys. Rev. Lett. **104**, 170401 (2010).
[4] C. Branciard, D. Rosset, N. Gisin, and S. Pironio, Phys. Rev. A **85**, 032119 (2012).
[5] R. Gallego, L. E. Würflinger, R. Chaves, A. Acín, and M. Navascués, New J. Phys. **16**, 033037 (2014).
[6] T. Fritz, New J. Phys. **14**, 103001 (2012).
[7] T. Fritz, arXiv:1404.4812 [Commun. Math. Phys. (to be published)].
[8] J. Henson, R. Lal, and M. F. Pusey, New J. Phys. **16**, 113043 (2014).
[9] R. Chaves, C. Majenz, and D. Gross, Nat. Commun. **6**, 5766 (2015).
[10] R. Chaves, L. Luft, T. O. Maciel, D. Gross, D. Janzing, and B. Schölkopf, *Proceedings of the 30th Conference on Uncertainty in Artificial Intelligence (2014)*, p. 112.
[11] J. Pearl, *Causality* (Cambridge University Press, Cambridge, England, 2009).
[12] P. Spirtes, N. Glymour, and R. Scheienes, *Causation, Prediction, and Search*, 2nd ed. (MIT Press, Cambridge, MA, 2001).
[13] E. G. Cavalcanti and R. Lal, J. Phys. A **47**, 424018 (2014).
[14] C. J. Wood and R. W. Spekkens, New J. Phys. **17**, 033002 (2015).
[15] R. Chaves, R. Kueng, J. B. Brask, and D. Gross, Phys. Rev. Lett. **114**, 140403 (2015).
[16] K. Ried, M. Agnew, L. Vermeyden, D. Janzing, R. W. Spekkens, and K. J. Resch, Nat. Phys. **11**, 414 (2015).
[17] M. Zukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert, Phys. Rev. Lett. **71**, 4287 (1993).
[18] P. H. Eberhard, Phys. Rev. A **47**, R747 (1993).
[19] D. Cavalcanti, M. L. Almeida, V. Scarani, and A. Acin, Nat. Commun. **2**, 184 (2011).
[20] A. Sen(De), U. Sen, C. Brukner, V. Buzek, and M. Zukowski, Phys. Rev. A **72**, 042310 (2005).
[21] A. Acín, J. I. Cirac, and M. Lewenstein, Nat. Phys. **3**, 256 (2007).
[22] H. J. Kimble, Nature (London) **453**, 1023 (2008).
[23] N. Sangouard, C. Simon, H. de Riedmatten, and N. Gisin, Rev. Mod. Phys. **83**, 33 (2011).
[24] B. Steudel and N. Ay, Entropy **17**, 2304 (2015).
[25] R. Chaves and T. Fritz, Phys. Rev. A **85**, 032113 (2012).
[26] R. Chaves, L. Luft, and D. Gross, New J. Phys. **16**, 043001 (2014).
[27] A. Tavakoli, P. Skrzypczyk, D. Cavalcanti, and A. Acín, Phys. Rev. A **90**, 062109 (2014).
[28] K. Mukherjee, B. Paul, and D. Sarkar, arXiv:1411.4188.
[29] D. Geiger and C. Meek, *Proceedings of the 15th Conference on Uncertainty in Artificial Intelligence* (Morgan Kaufmann Publishers Inc., Burlington, Massachusetts, 1999), p. 226.
[30] A. Tarski, *A Decision Method for Elementary Algebra and Geometry* (Springer, New York, 1998).
[31] V. Weispfenning, J. Symb. Comput. **5**, 3 (1988).
[32] C. W. Brown and J. H. Davenport, *Proceedings of the 2007 International Symposium on Symbolic and Algebraic Computation* (ACM, New York, 2007), p. 54.
[33] J. H. Davenport and J. Heintz, J. Symb. Comput. **5**, 29 (1988).
[34] L. D. Garcia, M. Stillman, and B. Sturmfels, J. Symb. Comput. **39**, 331 (2005).
[35] G. Ver Steeg and A. Galstyan, *Proceedings of the 27th Conference on Uncertainty in Artificial Intelligence* (AUAI Press, 2011).
[36] C. Kang and J. Tian, arXiv:1206.5275.
[37] C. M. Lee and R. W. Spekkens, arXiv:1506.03880.
[38] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.116.010402, which includes Refs. [39–44], for a detailed discussion about computational complexity issues, nonsignaling correlations in generalized Bell scenarios, and the derivation of the polynomial Bell inequalities.
[39] B. Chazelle, Discrete Computational Geometry **10**, 377 (1993).
[40] C. Budroni and A. Cabello, J. Phys. A **45**, 385304 (2012).
[41] T. Fritz and R. Chaves, IEEE Trans. Inf. Theory **59**, 803 (2013).
[42] P. Kurzyńskiand D. Kaszlikowski, Phys. Rev. A **89**, 012103 (2014).
[43] B. Bonet, *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence* (Morgan Kaufmann Publishers Inc., Burlington, Massachusetts, 2001), p. 48.
[44] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts, Phys. Rev. A **71**, 022101 (2005).
[45] S. Popescu and D. Rohrlich, Found. Phys. **24**, 379 (1994).
[46] I. Pitowsky, *Quantum Probability-Quantum Logic*, Lecture Notes in Physics (Springer-Verlag, New York, 1989).
[47] I. Pitowsky, Math. Program. **50**, 395 (1991).
[48] A. Fine, Phys. Rev. Lett. **48**, 291 (1982).
[49] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, England, 2009).

[50] G. E. Collins, in *Automata Theory and Formal Languages 2nd GI Conference Kaiserslautern, May 20–23, 1975* (Springer, New York, 1975), p. 134.
[51] H. P. Williams, Am. Math. Mon. **93**, 681 (1986).
[52] R. F. Werner and M. M. Wolf, Phys. Rev. A **64**, 032112 (2001).
[53] M. Zukowski and C. Brukner, Phys. Rev. Lett. **88**, 210401 (2002).
[54] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu, Phys. Rev. Lett. **88**, 040404 (2002).
[55] D. Collins and N. Gisin, J. Phys. A **37**, 1775 (2004).
[56] M. J. W. Hall, Phys. Rev. Lett. **105**, 250404 (2010).
[57] A. Cabello and J.-A. Larsson, Phys. Rev. Lett. **98**, 220402 (2007).
[58] N. Brunner, N. Gisin, V. Scarani, and C. Simon, Phys. Rev. Lett. **98**, 220403 (2007).
[59] J. B. Brask and R. Chaves, Phys. Rev. A **86**, 010103 (2012).
[60] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Phys. Rev. Lett. **98**, 230501 (2007).
[61] R. Colbeck, Ph.D. thesis, University of Cambridge, 2007.
[62] S. Pironio *et al.*, Nature (London) **464**, 1021 (2010).
[63] C. Brukner, M. Zukowski, J.-W. Pan, and A. Zeilinger, Phys. Rev. Lett. **92**, 127901 (2004).
[64] H. Buhrman, R. Cleve, S. Massar, and R. de Wolf, Rev. Mod. Phys. **82**, 665 (2010).
[65] B. Cirel'son, Lett. Math. Phys. **4**, 93 (1980).
[66] M. Navascués, S. Pironio, and A. Acín, Phys. Rev. Lett. **98**, 010401 (2007).
[67] M. Pawlowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Zukowski, Nature (London) **461**, 1101 (2009).
[68] D. Rosset, C. Branciard, T. J. Barnea, G. Pütz, N. Brunner, and N. Gisin, following letter, Phys. Rev. Lett. **116**, 010403 (2016).