Bounds on quantum correlations in Bell-inequality experiments

Yeong-Cherng Liang* and Andrew C. Doherty†
School of Physical Sciences, The University of Queensland, Queensland 4072, Australia

(Received 16 August 2006; published 9 April 2007)

Bell-inequality violation is one of the most widely known manifestations of entanglement in quantum mechanics; indicating that experiments on physically separated quantum mechanical systems cannot be given a local realistic description. However, despite the importance of Bell inequalities, it is not known in general how to determine whether a given entangled state will violate a Bell inequality. This is because one can choose to make many different measurements on a quantum system to test any given Bell inequality and the optimization over measurements is a high-dimensional variational problem. In order to better understand this problem we present algorithms that provide, for a given quantum state and Bell inequality, both a lower bound and an upper bound on the maximal violation of the inequality. In many cases these bounds determine measurements that would demonstrate violation of the Bell inequality or provide a bound that rules out the possibility of a violation. Both bounds apply techniques from convex optimization and the methodology for creating upper bounds allows them to be systematically improved. Examples are given to illustrate how these algorithms can be used to conclude definitively if some quantum states violate a given Bell inequality.

DOI: 10.1103/PhysRevA.75.042103 PACS number(s): 03.65.Ud, 03.67.Mn, 02.60.Pn

I. INTRODUCTION

The nature of entanglement has always been bewildering ever since its first appearance in the literature [1,2]. This is more so after Bell's seminal work in 1964 [3], in which he showed, using what is now known as a Bell inequality, that some experimental statistics of the spin-singlet state are intrinsically incompatible with local realism.

For a long time after that, it seems to have been generally assumed that entanglement and the violation of Bell inequalities are synonymous. The first counterexample to that commonly held intuition was provided by Werner [4], where he showed that there are bipartite mixed two-qudit states, now known as the Werner states, that are entangled and yet do not violate a large class of Bell inequalities.

Soon after that, it was demonstrated by Gisin [5] and later by Gisin and Peres [6] that all bipartite pure entangled states violate the Bell-Clauser-Horne-Shimony-Holt inequality [7]; the generalization to multipartite pure entangled states was also carried out by Popescu and Rohrlich by invoking appropriate postselection [8]. Three years later, Horodecki *et al.* provided the first analytic criterion [9] to determine if a two-qubit state violates the Bell-Clauser-Horne-Shimony-Holt (henceforth abbreviated as Bell-CHSH) inequality.

The Horodecki criterion is, unfortunately, also the only analytic criterion that we have in determining if a broad class of quantum states, namely two-qubit states, can be simulated classically. For specific quantum states, there are examples where explicit local hidden variables (LHV) models have been constructed to reproduce the quantum mechanical predictions, thereby ruling out the possibility that these quantum states may violate a Bell inequality [4,10].

In general, however, to determine if a quantum state violates a Bell inequality is a high-dimensional variational problem, which requires a nontrivial optimization of a Hermitian operator $\mathcal{B}$ (now known as the *Bell* operator [11]) over the various possible measurement settings that each observer may perform. This optimization does not appear to be convex and is possibly *NP*-hard [12].

Except for the simplest scenario where one deals with Bell-CHSH inequality [7], in conjunction with a two-qubit state [9], or a maximally entangled pure state [6,13], and its mixture with the completely mixed state [14], very few analytic results for the optimal measurements are known. As such, for the purpose of characterizing quantum states that are incompatible with local realistic description, efficient algorithms to perform this state-dependent optimization are very desirable.

Bell-inequality violation is also relevant in various aspects of quantum information processing, in particular, quantum teleportation [15], quantum key distribution [16,17], and reduction of communication complexity [18]. Recently, it has even been argued [17] that Bell-inequality violation is necessary to guarantee the security of some entanglement-based quantum key distribution protocols.

On the other hand, state-independent bounds of quantum correlations have also been investigated since the early 1980s. In particular, Tsirelson [19] has demonstrated, using what is now known as Tsirelson's vector construction, that in a Bell-CHSH setup, bipartite quantum systems of arbitrary dimensions cannot exhibit correlations stronger than $2\sqrt{2}$, a value now known as Tsirelson's bound. Recently, analogous bounds for more complicated Bell inequalities have also been investigated by Filipp and Svozil [20], Buhrman and Massar [21], Wehner [22], Toner [23], Avis *et al.* [24], and Navascués *et al.* [25]. On a related note, bounds on quantum correlations for given local measurements, rather than given quantum state, have also been investigated by Cabello [26] and Bovino *et al.* [27].

In this paper, we will present, respectively, in Secs. II B and II C, two algorithms that were developed to provide a lower bound and an upper bound on the maximal expectation

*Electronic address: ycliang@physics.uq.edu.au
†Electronic address: doherty@physics.uq.edu.au

value of a Bell operator for a given quantum state. The second algorithm is another instance where a nonlinear optimization problem is approximated by a hierarchy of semidefinite programs, each giving a better bound of the original optimization problem [25,28–30]. In its simplest form, it provides a bound that is apparently state-independent, and thus provides a (not necessarily tight) bound on the maximum attainable quantum correlations in a Bell inequality experiment.

In Sec. III, we will derive, based on the second algorithm, a necessary condition for a class of two-qudit states to violate the Bell-CHSH inequality. We will also demonstrate how the two algorithms can be used in tandem to determine if some quantum states violate a Bell inequality. Some limitations of these algorithms will then be discussed. We will conclude with a summary of results and some possibilities for future research.

Throughout, boldfaced Latin letters, e.g., $\mathbf{x}$ will be used to denote a column vector whereas $\mathbf{0}$ and $\mathbf{1}$, respectively, represent a zero block matrix and an identity matrix. Moreover, the (i,j) entry of a matrix M will be denoted by $[M]_{ij}$.

II. BOUNDS ON QUANTUM CORRELATIONS

A. Preliminaries

Bell inequalities are inequalities derived from the assumption of a general LHV model. A particular Bell inequality deals with a specific experimental setup, say a source that distributes pairs of particles to two experimenters (hereafter called Alice and Bob), and where each of them can perform, respectively, m_A and m_B alternative measurements that would each generate n_A and n_B distinct outcomes. For each of these setups, a Bell inequality places a bound on the experimental statistics obtained from the corresponding Bell experiments. If there exists a LHV model that saturates the bound then the inequality is said to be tight [32]. In what follows, we will adopt the notation introduced in Ref. [31] and refer to a tight Bell inequality for such an experimental setup as a Bell-$m_Am_Bn_An_B$ inequality.

A Bell inequality for correlations, such as the Bell-CHSH inequality [7] typically involves statistical constraints on some linear combination of *correlation functions*. Similarly, a Bell inequality for probabilities, such as the Bell-Clauser-Horne (henceforth abbreviated as Bell-CH) inequality [33], places bounds on some linear combinations of joint and marginal *probabilities* of experimental outcomes. In either case, a general Bell inequality takes the form

$$\mathcal{S}_{\text{LHV}} \leqslant \beta_{\text{LHV}}, \tag{1}$$

where β_{LHV} is a real number and $\mathcal{S}_{\text{LHV}}$ involves a specific linear combination of correlation functions or joint and marginal probabilities of experimental outcomes.

To compare with predictions given by quantum mechanics, these correlation functions, or probabilities, are calculated using the quantum mechanical rules. The bounds on $\mathcal{S}_{\text{LHV}}$ then translate into corresponding bounds β_{LHV} on the expectation value of some Hermitian observable that describes the Bell-inequality experiment, this observable is known as the *Bell operator* $\mathcal{B}$ [11]. The restriction that the given Bell inequality is satisfied in the experiment is then

$$\mathcal{S}_{\text{QM}}(\rho,\mathcal{B}) = \text{tr}(\rho\mathcal{B}) \leqslant \beta_{\text{LHV}}. \tag{2}$$

The Bell operator depends on the choice of measurements at each of the sites (polarizer angles, for example). These measurements will be described by a set of Hermitian operators $\{O_m\}$. For correlation inequalities these are simply the measured observables at each stage of the Bell measurement, while for general probability inequalities the O_m are elements of the positive-operator-valued measures (POVMs) that describe the measurements at each site. We will denote this expectation value by $\mathcal{S}_{\text{QM}}(\rho,\{O_m\})$ when we want to emphasize its dependence on the choice of local Hermitian observables O_m. Ideally the choice of measurement should give the maximal expectation value of the Bell operator, for which we will give the notation

$$\mathcal{S}_{\text{QM}}(\rho) \equiv \max_{\{O_m\}} \mathcal{S}_{\text{QM}}(\rho,\{O_m\}). \tag{3}$$

It is this implicitly defined function that will give us information about which states violate a given Bell inequality.

As an example, let us recall the Bell-CHSH inequality [7], which is a dichotomic (i.e., two-outcome) Bell correlation inequality that involves two parties, and where the two possible measurement outcomes are assigned the values ±1:

$$\mathcal{S}_{\text{LHV}} = E(A_1,B_1) + E(A_1,B_2) + E(A_2,B_1) - E(A_2,B_2) \leqslant 2. \tag{4}$$

In the above expression, the correlation function $E(A_k,B_l)$ represents the expectation value of Alice's measurement outcome times Bob's measurement outcome, given that she has chosen to measure the observable A_k and he has chosen to measure the observable B_l. In quantum mechanics, these correlation functions are computed using

$$E(A_k,B_l) = \text{tr}(\rho A_k \otimes B_l). \tag{5}$$

Substituting this into Eq. (4) and comparing with Eq. (2), one finds that the corresponding Bell operator reads

$$\mathcal{B}_{\text{CHSH}} = A_1 \otimes (B_1 + B_2) + A_2 \otimes (B_1 - B_2). \tag{6}$$

Determining the maximal Bell-inequality violation for a given ρ, $\mathcal{S}_{\text{QM}}(\rho)$, requires a maximization by varying over all possible choices of $\{O_m\}$, i.e., A_k and B_l in the case of Eq. (6). Whether we are interested in correlation inequalities or in Bell inequalities for probabilities the (bipartite) Bell operator has the general structure

$$\mathcal{B} = \sum_{K,L} b_{KL} A_K \otimes B_L. \tag{7}$$

In the case of a Bell inequality for probabilities the indices K, L are collective indices describing both a particular measurement setting and a particular outcome for each observer. For correlation inequalities they refer simply to the measurement settings as in the Bell-CHSH case described in detail above.

In what follows, we will present two algorithms which we have developed specifically to perform the maximization over choice of measurements. The first, which we will abbreviate as LB, provides a *lower bound* on the maximal expectation value and can be implemented for any Bell inequality. This bound makes use of the fact that the objective function $\mathcal{S}_{\mathrm{QM}}$ is bilinear in the observables O_m, that is it is linear in the A_K for fixed B_L and likewise linear in the B_L for fixed A_K. The second bound, which we will abbreviate as UB, provides an *upper bound* on $\mathcal{S}_{\mathrm{QM}}(\rho)$ by regarding $\mathcal{S}_{\mathrm{QM}}(\rho,\{O_m\})$ as a polynomial function of the variables that define the various O_m and applying general techniques for finding such bounds on polynomials [28,29].

Both of these make use of convex optimization techniques in the form of a semidefinite program (SDP) [34,35]. A semidefinite program is an optimization over Hermitian matrices. The objective function depends linearly on the matrix variable (as expectation values do in quantum mechanics, for example) and the optimization is carried out subjected to the constraint that the matrix variable is positive semidefinite and satisfies various affine constraints. Any semidefinite program may be written in the following *standard form*:

$$\text{maximize} \quad -\mathrm{tr}[F_0 Z] \tag{8a}$$

$$\text{subject to} \quad \mathrm{tr}[F_m Z] = c_m \quad \forall\, m, \tag{8b}$$

$$Z \geqslant 0, \tag{8c}$$

where F_0 and all the F_m's are Hermitian matrices and the c_m are real numbers that together specify the optimization; Z is the Hermitian matrix variable to be optimized.

An SDP also arises naturally in the *inequality form*, which seeks to minimize a linear function of the optimization variables $\mathbf{x} \in \mathbb{R}^n$, subjected to a linear matrix inequality:

$$\text{minimize} \quad \mathbf{x}^T \mathbf{c}' \tag{9a}$$

$$\text{subject to} \quad G_0 + \sum_m \mathbf{x}_m G_m \geqslant 0. \tag{9b}$$

As in the standard form, G_0 and all the G_m's are Hermitian matrices, while $\mathbf{c}'$ is a real vector of length n.

B. Lower bound on $\mathcal{S}_{\mathrm{QM}}(\rho)$

The key idea behind the LB algorithm is to realize that when measurements for all but one party are *fixed*, the optimal measurements for the remaining party can be obtained efficiently using convex optimization techniques, in particular an SDP. Thus we can fix Bob's measurements and find Alice's optimal choice, at least numerically; with this optimized measurements for Alice, we can further find the optimal measurements for Bob (for this choice of Alice's settings), and then Alice again and so on and so forth until $\mathcal{S}_{\mathrm{QM}}(\rho,\{O_m\})$ converges within the desired numerical precision [36].

Back in 2001, Werner and Wolf [37] presented a similar iterative algorithm, by the name of *See-Saw iteration*, to maximize the expectation value of the Bell operator for a correlation inequality involving only dichotomic observables [39]. As a result we will focus here on the (straightforward) generalization to the widest possible class of Bell inequalities. In the work of Werner and Wolf [37] it turned out that once the dichotomic observables for one party are fixed, optimization of the other party's observables can be carried out explicitly. This turns out to be true for any dichotomic Bell inequality and we will return to this question in Sec. II B 3.

1. General settings

First we must develop a more explicit notation for a general Bell inequality for probabilities [38]. Let us consider a Bell-$m_A m_B n_A n_B$ inequality for probabilities. We will denote the POVM element associated with the κth outcome of Alice's kth measurement by A_k^κ while B_l^λ is the POVM element associated with the λth outcome of Bob's lth measurement. Moreover, let d_A and d_B, respectively, be the dimension of the state space that each of the A_k^κ and B_l^λ acts on. Then it follows from Born's rule that

$$p_{AB}^{\kappa\lambda}(k,l) = \mathrm{tr}(\rho A_k^\kappa \otimes B_l^\lambda), \tag{10a}$$

$$p_A^\kappa(k) = \mathrm{tr}(A_k^\kappa \otimes \mathbf{1}_{d_B}), \quad p_B^\lambda(l) = \mathrm{tr}(\rho \mathbf{1}_{d_A} \otimes B_l^\lambda), \tag{10b}$$

where $p_{AB}^{\kappa\lambda}(k,l)$ refers to the joint probability that the κth experimental outcome is observed at Alice's site and the λth outcome at Bob's, given that Alice performs the kth and Bob performs the lth measurement. The marginal probabilities $p_A^\kappa(k)$ and $p_B^\lambda(l)$ are similarly defined. A general Bell operator for probabilities can then be expressed as

$$\mathcal{B} = \sum_{k=1}^{m_A} \sum_{\kappa=1}^{n_A} \sum_{l=1}^{m_B} \sum_{\lambda=1}^{n_B} b_{kl}^{\kappa\lambda} A_k^\kappa \otimes B_l^\lambda, \tag{11}$$

where $b_{kl}^{\kappa\lambda}$ are determined from the given Bell inequality. Note that the sets of POVM elements $\{A_k^\kappa\}_{\kappa=1}^{n_A}$ and $\{B_l^\lambda\}_{\lambda=1}^{n_B}$ satisfy

$$\sum_{\kappa=1}^{n_A} A_k^\kappa = \mathbf{1}_{d_A} \quad \text{and} \quad \sum_{\lambda=1}^{n_B} B_l^\lambda = \mathbf{1}_{d_B} \quad \forall\, k,l, \tag{12a}$$

$$A_k^\kappa \geqslant 0, \quad B_l^\lambda \geqslant 0 \quad \forall\, k,l,\kappa,\lambda. \tag{12b}$$

2. Iterative semidefinite programming algorithm

To see how to develop a lower bound on $\mathcal{S}_{\mathrm{QM}}(\rho)$ by fixing the observables at one site and optimizing the other, we observe that upon substituting Eq. (11) into Eq. (2), the left-hand side of the inequality can be rewritten as

$$\mathcal{S}_{\mathrm{QM}}(\rho, A_k^\kappa, B_l^\lambda) = \sum_{l,\lambda} \mathrm{tr}(\rho_{B_l^\lambda} B_l^\lambda), \tag{13}$$

where $\rho_{B_l^\lambda} \equiv \Sigma_{k,\kappa} b_{kl}^{\kappa\lambda}\, \mathrm{tr}_A[\rho(A_k^\kappa \otimes \mathbf{1}_{d_B})]$. $\mathrm{tr}_A[\cdot]$ is the partial trace over subsystem A.

Notice that if $\rho_{B_l^\lambda}$ are held constant by fixing all of Alice's measurement settings (given by the set of A_k^κ) then $\rho_{B_l^\lambda}$ is a constant matrix independent of the B_l^λ. Thus the objective function is linear in these variables. The constraints that $\{B_l^\lambda\}_{\lambda=1}^{n_B}$ form a POVM for each value of l is a combination of affine and matrix non-negativity constraints. As a result it is fairly clear that the following problem is an SDP in standard form:

$$\text{maximize}_{\{B_l^\lambda\}} \ \mathcal{S}_{\mathrm{QM}}(\rho, A_k^\kappa, B_l^\lambda) \tag{14a}$$

$$\text{subject to} \quad \sum_{\lambda=1}^{n_B} B_l^\lambda = \mathbf{1}_{d_B} \quad \forall\, l, \tag{14b}$$

$$B_l^\lambda \geqslant 0 \quad \forall\, l, \lambda. \tag{14c}$$

Explicit forms for the matrices F_m and values c_m of Eq. (8) can be found in Appendix A.

Exactly the same analysis follows if we fix Bob's measurement settings and optimize over Alice's POVM elements instead. To arrive at a local maximum of $\mathcal{S}_{\mathrm{QM}}(\rho, A_k^\kappa, B_l^\lambda)$, it therefore suffices to start with some random measurement settings for Alice (or Bob), and optimize over the two parties' settings iteratively. A (nontrivial) lower bound on $\mathcal{S}_{\mathrm{QM}}(\rho)$ can then be obtained by optimizing the measurement settings starting from a set of randomly generated initial guesses.

It is worth noting that in any implementation of this algorithm, physical observables $\{A_k^\kappa, B_l^\lambda\}$ achieving the lower bound are constructed when the corresponding SDP is solved. In the event that the lower bound is greater than the classical threshold β_{LHV}, then these observables can, in principle, be measured in the laboratory to demonstrate a Bell-inequality violation of the given quantum state.

We have implemented this algorithm in MATLAB [40] to search for a lower bound on $\mathcal{S}_{\mathrm{QM}}(\rho)$ in the case of Bell-CH [33], I_{3322}, I_{4422}, I_{2233}, and I_{2244} inequality [31], and with the local dimension d up until 32. Typically, with no more than 50 iterations, the algorithm already converges to a point that is different from a local maximum by no more than 10^{-9}. To test against the effectiveness of finding $\mathcal{S}_{\mathrm{QM}}(\rho)$ using LB, we have randomly generated 200 Bell-CH violating two-qubit states and found that on average, it takes about six random initial guesses before the algorithm gives $\mathcal{S}_{\mathrm{QM}}(\rho, \{O_m\})$ that is close to the actual maximum (computed using Horodecki's criterion [9] within 10^{-5}). Specific examples regarding the implementation of this algorithm will be discussed in Sec. III.

Two other remarks concerning this algorithm should now be made. First, the algorithm is readily generalized to multipartite Bell inequalities for probabilities: one again starts with some random measurement settings for all but one party, and optimizes over each party iteratively. Also, it is worth noting that this algorithm is not only useful as a numerical tool, but for specific cases it can also provide useful analytic criterion. In particular, when applied to the Bell-CH inequality [33] for two-qubit states, the analysis for dichotomic observables discussed in the next section allows one to recover Horodecki's criterion [9], i.e., the necessary and sufficient condition for two-qubit states to violate the Bell-CHSH inequality [7,41].

3. Two-outcome Bell experiment

We will show that, just as in the case of correlation inequalities [37], the local optimization can be solved analytically for two-outcome measurements. If we denote by "±" the two outcomes of the experiments, it follows from Eq. (12) that the POVM element B_l^- can be expressed as a function of the complementary POVM element B_l^+, i.e., $B_l^- = \mathbf{1}_{d_B} - B_l^+$, subjected to $0 \leqslant B_l^+ \leqslant \mathbf{1}_{d_B}$. We then have

$$\sum_{\lambda=\pm} \mathrm{tr}(\rho_{B_l^\lambda} B_l^\lambda) = \mathrm{tr}[(\rho_{B_l^+} - \rho_{B_l^-}) B_l^+] + \mathrm{tr}(\rho_{B_l^-}).$$

The above expression can be maximized by setting the positive semidefinite operator B_l^+ to be the projector onto the positive eigenspace of $\rho_{B_l^+} - \rho_{B_l^-}$. In a similar manner, we can also write

$$\sum_{\lambda=\pm} \mathrm{tr}(\rho_{B_l^\lambda} B_l^\lambda) = \mathrm{tr}[(\rho_{B_l^-} - \rho_{B_l^+}) B_l^-] + \mathrm{tr}(\rho_{B_l^+}),$$

which can be maximized by setting B_l^- to be the projector onto the nonpositive eigenspace of $\rho_{B_l^+} - \rho_{B_l^-}$. Notice that this choice is consistent with our earlier choice of B_l^+ for the + outcome POVM element in that they form a valid POVM. Since there can be no difference in these maxima, we may write the maximum as their average, i.e.,

$$\sum_{\lambda=\pm} \mathrm{tr}(\rho_{B_l^\lambda} B_l^\lambda) = \frac{1}{2}\|\rho_{B_l^+} - \rho_{B_l^-}\| + \frac{1}{2}\sum_{\lambda=\pm} \mathrm{tr}(\rho_{B_l^\lambda}),$$

where $\|O\|$ is the trace norm of the Hermitian operator O [42]. Carrying out the optimization for each of the l settings, the optimized $\mathcal{S}_{\mathrm{QM}}(\rho, A_k^\kappa, B_l^\lambda)$, as an implicit function of Alice's POVM $\{A_k^\kappa\}$, is given by

$$\mathcal{S}_{\mathrm{QM}}(\rho, A_k^\kappa) = \frac{1}{2}\sum_l \|\rho_{B_l^+} - \rho_{B_l^-}\| + \frac{1}{2}\sum_l \sum_{\lambda=\pm} \mathrm{tr}(\rho_{B_l^\lambda}). \tag{15}$$

An immediate corollary of the above result is that for the optimization of a two-outcome Bell operator for probabilities, it is unnecessary for any of the two observers to perform generalized measurements described by a POVM; von Neumann projective measurements are sufficient [43]. In practice, this simplifies any analytic treatment of the optimization problem as a generic parametrization of a POVM is a lot more difficult to deal with, thereby supporting the simplification adopted in Ref. [20].

Nevertheless, it may still be advantageous to consider generic POVMs as our initial measurement settings when implementing the algorithm numerically. This is because the local maximum of $\mathcal{S}_{\mathrm{QM}}(\rho, \{O_m\})$ obtained using the iterative procedure is a function of the initial guess. In particular, it was found that the set of local maxima attainable could change significantly if the ranks of the initial measurement projectors are altered. As such, it seems necessary to step through various ranks of the starting projectors to obtain a good lower bound on $\mathcal{S}_{\mathrm{QM}}(\rho)$. Even then, we have also found examples where this does not give a lower bound on $\mathcal{S}_{\mathrm{QM}}(\rho)$ that is as good as when generic POVMs are used as the initial measurement operators.

C. Upper bound on $\mathcal{S}_{\rm QM}(\rho)$

A major drawback of the above algorithm, or the analogous algorithm developed by Werner and Wolf [37] for Bell correlation inequalities is that, except in some special cases, it is generally impossible to tell if the maximal $\mathcal{S}_{\rm QM}(\rho,\{O_m\})$ obtained through this optimization procedure corresponds to the global maximum $\mathcal{S}_{\rm QM}(\rho)$.

Nontrivial upper bounds on $\mathcal{S}_{\rm QM}(\rho)$, nevertheless, can be obtained by considering *relaxations* of the global optimization problem. In a relaxation, a (possibly nonconvex) maximization problem is modified in some way so as to yield a more tractable optimization that bounds the optimization of interest. One example of a variational upper bound that exists for any optimization problem is the *Lagrange dual* optimization that arises in the method of Lagrange multipliers [35].

To see how to apply existing studies in the optimization literature to find upper bounds on $\mathcal{S}_{\rm QM}(\rho)$, let us first remark that the global objective function $\mathcal{S}_{\rm QM}(\rho,\{O_m\})$ can be mapped to a polynomial function in real variables, for instance, by expanding all the local observables $\{O_m\}$ and the density matrix ρ in terms of Hermitian basis operators. In the same manner, matrix equality constraints, such as that given in Eq. (12a), can also be mapped to a set of polynomial equalities by requiring that the matrix equality holds componentwise. Now, it is known that a hierarchy of global bounds of a polynomial function, subjected to polynomial equalities and inequalities, can be achieved by solving suitable SDPs [28,29]. Essentially, this is achieved by approximating the original nonconvex optimization problem by a series of convex ones in the form of a SDP, each giving a better bound of the original polynomial objective function.

At the bottom of this hierarchy is the lowest order relaxation provided by the *Lagrange dual* of the original nonconvex problem. By considering *Lagrange multipliers* that depend on the original optimization variables, higher order relaxations to the original problem can be constructed to give tighter upper bounds on $\mathcal{S}_{\rm QM}(\rho)$ (see Appendix B for more details).

In the following, we will focus our discussion on a general two-outcome Bell (correlation) inequality, where the observables $\{O_m\}$ are only subjected to matrix equalities. In particular, we will show that the global optimization problem for these Bell inequalities is a *quadratically constrained quadratic-program* (QCQP), i.e., one whereby the objective function and the constraints are both *quadratic* in the optimization variables. Then, we will demonstrate explicitly how the Lagrange dual of this QCQP, which is known to be an SDP, can be constructed. The analogous analytic treatment is apparently formidable for higher order relaxations. Nonetheless, there exists a third-party MATLAB [40] toolbox known as the SOSTOOLS which is tailored specifically for this kind of optimization problem [45,46].

Numerically, we have implemented the algorithm for several two-outcome correlation inequalities and will discuss the results in greater detail in Sec. III. For a general Bell inequality where each O_m is also subjected to a linear matrix inequality (LMI) like Eq. (12b), the algorithm can still be implemented, for instance, by requiring that all the principle submatrices of O_m have non-negative determinants [42]. This then translates into a set of polynomial inequalities which fit into the framework of a general polynomial optimization problem (see Appendix B). However, a more effective approach would retain the structure of linear matrix inequalities constraining a polynomial optimization problem; we leave the investigation of these bounds to further work.

1. Global optimization problem

Now, let us consider a dichotomic Bell correlation inequality where Alice an Bob can respectively perform m_A and m_B alternative measurements. A general Bell correlation operator for such an experimental setup can be written as [47]

$$\mathcal{B}=\sum_{k=1}^{m_A}\sum_{l=1}^{m_B} b_{kl} O_k \otimes O_{l+m_A}, \tag{16}$$

where b_{kl} are determined from the given Bell correlation inequality, O_k for $k=1,\ldots,m_A$ refers to the kth Hermitian observable measured by Alice, and O_{l+m_A} for $l=1,\ldots,m_B$ refers to the lth Hermitian observable measured by Bob. Furthermore, by convention, these dichotomic observables can be chosen to have eigenvalues ± 1 [48] and thus

$$O_m^\dagger O_m=(O_m)^2=\mathbf{1}_d \quad \forall\, m, \tag{17}$$

where we have assumed for simplicity that all the local observables O_m act on a state space of dimension d [49].

The global optimization problem derived from a dichotomic Bell correlation inequality thus takes the form of

$$\text{maximize} \quad \text{tr}\,(\rho\mathcal{B}) \tag{18a}$$

$$\text{subject to} \quad O_m^2=\mathbf{1}_d \quad \forall\, m=1,2,\ldots. \tag{18b}$$

For any $m\times n$ complex matrices, we will now define vec(A) to be the $m\cdot n$ dimensional vector obtained by stacking all columns of A on top of one another. By collecting all the vectorized observables together

$$\mathbf{w}^\dagger \equiv [\text{vec}(O_1)^\dagger \text{vec}(O_2)^\dagger \cdots \text{vec}(O_{m_A+m_B})^\dagger],$$

and using the identity

$$\text{tr}(\rho O_k\otimes O_{l+m_A})=\text{vec}(O_k)^\dagger (V\rho)^{T_A}\text{vec}(O_{l+m_A}), \tag{19}$$

with V being the flip operator such that $V|i\rangle|j\rangle=|j\rangle|i\rangle$ and $(\cdot)^{T_A}$ being the partial transposition with respect to subsystem A, we can write the objective function more explicitly as

$$\mathcal{S}_{\rm QM}(\rho,\{O_m\})=\text{tr}(\rho\mathcal{B})=-\,\mathbf{w}^\dagger\Omega_0\mathbf{w}, \tag{20}$$

where

$$\Omega_0\equiv\frac{1}{2}\begin{pmatrix}\mathbf{0} & -b\otimes R\\ -b^T\otimes R^\dagger & \mathbf{0}\end{pmatrix},$$

b is a $m_A\times m_B$ matrix with $[b]_{kl}=b_{kl}$ [cf. Eq. (16)] and $R\equiv(V\rho)^{T_A}$. In this form, it is explicit that the objective function is *quadratic* in vec(O_m). Similarly, by requiring that the matrix equality holds componentwise, we can get a set of equality constraints, which are each *quadratic* in $\text{vec}(O_m)$. The global optimization problem (18) is thus an instance of a QCQP.

On a related note, for any Bell inequality experiments where measurements are restricted to the projective type, the global optimization problem is also a QCQP. To see this, we first note that the global objective function for the general case, as follows from Eqs. (2) and (7), is always quadratic in the local Hermitian observables $\{A_K, B_L\}$. The requirement that these measurement operators are projectors amounts to requiring

$$A_K^2 = A_K, \quad B_L^2 = B_L, \quad \forall\, K,L, \tag{21}$$

which are quadratic constraints on the local Hermitian observables. Since we have shown in Sec. II B 3 that for any two-outcome Bell inequality for probabilities, it suffices to consider projective measurements in optimizing $\mathcal{S}_{\text{QM}}(\rho,\{O_m\})$, it follows that the global optimization problem for these Bell inequalities is always a QCQP.

2. State-independent bound

As mentioned above, the lowest order relaxation to the global optimization problem (18) is simply the Lagrange dual of the original QCQP. This can be obtained via the *Lagrangian* of the global optimization problem (18)

$$\mathcal{L}(\{O_m\},\Lambda_m) = \mathcal{S}_{\text{QM}}(\rho,\{O_m\}) - \sum_{m=1}^{m_A+m_B} \text{tr}[\Lambda_m(O_m^2 - \mathbf{1}_d)], \tag{22}$$

where Λ_m is a matrix of Lagrange multipliers associated with the mth matrix equality constraint. With no loss of generality, we can assume that Λ_m's are Hermitian.

Notice that for all values of $\{O_m\}$ that satisfy the constraints, the Lagrangian $\mathcal{L}(\rho,\{O_m\},\Lambda_m)=\mathcal{S}_{\text{QM}}(\rho,\{O_m\})$. As a result, if we maximize the Lagrangian without regard to the constraints we obtain an upper bound on the maximal expectation value of the Bell operator

$$\max_{\{O_m\}} \mathcal{L}(\rho,\{O_m\},\Lambda_m) \geq \mathcal{S}_{\text{QM}}(\rho). \tag{23}$$

The Lagrange dual optimization simply looks for the best such upper bound.

In order to maximize the Lagrangian we rewrite the Lagrangian with Eq. (20) and the identity

$$\text{tr}(\Lambda_m O_m O_m^\dagger) = \text{vec}(O_m)^\dagger \mathbf{1}_d \otimes \Lambda_m \text{vec}(O_m) \tag{24}$$

to obtain

$$\mathcal{L}(\mathbf{w},\Lambda_m) = -\mathbf{w}^\dagger \Omega \mathbf{w} + \sum_{m=1}^{m_A+m_B} \text{tr}\,\Lambda_m, \tag{25}$$

where $\Omega \equiv \Omega_0 + \oplus_{m=1}^{m_A+m_B} \mathbf{1}_d \otimes \Lambda_m$. Note that each of the diagonal blocks $\mathbf{1}_d \otimes \Lambda_m$ is of the same size as the matrix R.

To obtain the dual optimization problem, we maximize the Lagrangian over $\mathbf{w}$ to obtain the *Lagrange dual function*

$$g(\Lambda_m) \equiv \sup_{\mathbf{w}} \mathcal{L}(\mathbf{w},\Lambda_m). \tag{26}$$

As noted above $g(\Lambda_m) \geq \mathcal{S}_{\text{QM}}(\rho)$ for all choices of Λ_m. Moreover, this supremum over $\mathbf{w}$ is unbounded above unless $\Omega \geq 0$, in which case the supremum is attained by setting $\mathbf{w} = 0$ in Eq. (25). Hence the *Lagrange dual* optimization, which seeks for the best upper bound of Eq. (18) by minimizing Eq. (26) over the Lagrange multipliers, reads

$$\text{minimize} \quad \sum_{i=1}^{m_A+m_B} \text{tr}\,\Lambda_m \quad \text{subject to} \quad \Omega \geq 0. \tag{27}$$

By expanding Λ_m in terms of Hermitian basis operators satisfying Eq. (A1),

$$\Lambda_m = \sum_{n=0}^{d^2-1} \lambda_{mn}\sigma_n, \tag{28}$$

the optimization problem (27) is readily seen to be an SDP in the inequality form (9).

For Bell-CHSH inequality and the correlation equivalent of I_{3322} inequality, it was observed numerically that the upper bound obtained via the SDP (27) is always *state-independent*. In fact, for 1000 randomly generated two-qubit states, and 1000 randomly generated two-qutrit states, the Bell-CHSH upper bound of $\mathcal{S}_{\text{QM}}(\rho)$ obtained through Eq. (27) was never found to differ from the Tsirelson bound [19] by more than 10^{-7}. In fact by finding an explicit feasible solution to the optimization problem dual to Eq. (27), Wehner has shown that the upper bound obtained here can be no better than that obtained by Tsirelson's vector construction for correlation inequalities [50].

In a similar manner, we have also investigated the upper bound of $\mathcal{S}_{\text{QM}}(\rho)$ for some dichotomic Bell probability inequalities using the lowest order relaxation to the corresponding global optimization problem. Interestingly, the numerical upper bounds obtained from the analog of Eq. (27) for these inequalities, namely the Bell-CH inequality, the I_{3322} inequality, and the I_{4422} inequality, are also found to be state-independent and are given by 0.207 106 7, 0.375, and 0.669 346 1, respectively.

3. State-dependent bound

Although the state-independent upper bounds obtained above are interesting in their own right, our main interest here is to find an upper bound on $\mathcal{S}_{\text{QM}}(\rho)$ that does depend on the given quantum state ρ. This can be obtained, with not much extra cost, from the Lagrange dual to a more-refined version of the original optimization problem.

To appreciate that, let us first recall that each dichotomic Hermitian observable O_m can only have eigenvalues ± 1. It follows that their trace

$$z_m \equiv \text{tr}(O_m) \tag{29a}$$

can only take on the following values:

$$z_m = -d, -d+2, \ldots, d-2, d. \tag{29b}$$

In particular, if $z_m = \pm d$ for any m, then $O_m = \pm \mathbf{1}_d$ and it is not difficult to show that the Bell-CHSH inequality cannot be violated for this choice of observable.

Better Lagrange dual bounds arise from taking these additional constraints (29a) and (29b) explicitly into account. We found it most convenient to express the original optimization problem in terms of real variables given by the expansion coefficients of O_m in terms of a basis for Hermitian matrices that includes the (traceless) Gell-Mann matrices and the identity matrix [cf. Eq. (A1)]. For details see Appendix C. The result is a set of SDPs, one for each of the various choices of z_m. The lowest order upper bound on $\mathcal{S}_{\mathrm{QM}}(\rho)$ can then be obtained by stepping through the various choices of z_m given in Eq. (29b), solving each of the corresponding SDP, and taking their maximum. The results of this approach will be discussed later, for now it suffices to note that tighter bounds can be obtained which are explicitly state dependent.

4. Higher order relaxations

The higher order relaxations simply arise from allowing the Lagrange multipliers λ to be polynomial functions of $\{O_m\}$ rather than constants. In this case, it is no longer possible to optimize over the primal variables in the Lagrangian analytically but let us consider the following optimization

$$\begin{aligned} &\text{minimize} \quad \gamma \\ &\text{subject to} \quad \gamma - \mathcal{S}_{\mathrm{QM}}(\rho,\mathbf{x}) = \mu(\mathbf{x}) + \sum_i \lambda_i(\mathbf{x}) f_{\mathrm{eq},i}(\mathbf{x}), \end{aligned} \tag{30}$$

where each of the λ_i's are polynomial functions of $\mathbf{x}$ and $\mu(\mathbf{x})$ is a sum of squares (SOS) polynomial and therefore non-negative. That is $\mu(\mathbf{x}) = \sum_j [h_j(\mathbf{x})]^2 \geqslant 0$ for some set of real polynomials $h_j(\mathbf{x})$. The variables of the optimizations are $\mathbf{x}$ and the coefficients that define the polynomials $\mu(\mathbf{x})$ and λ_i. Notice that we have $\gamma \geqslant \mathcal{S}_{\mathrm{QM}}(\rho,\mathbf{x})$ whenever the constraints are satisfied so that once again we have a global upper bound on $\mathcal{S}_{\mathrm{QM}}(\rho,\mathbf{x})$. This optimization can be implemented numerically by restricting $\mu(\mathbf{x})$ and λ_i to be of some fixed degree. The Lagrange dual optimization (27) arises from choosing the degree of λ_i to be zero. It is known that for any fixed degree this optimization is an SDP [28,29] and we have implemented up to degree four using SOSTOOLS [45,46]. Schmüdgen's theorem guarantees that by increasing the degree of the polynomials in the relaxation we obtain bounds approaching the true maximum $\mathcal{S}_{\mathrm{QM}}(\rho)$. This is a special case of the general procedure described in [28,46,51] which is able to handle inequality constraints. For more details see Appendix B.

III. APPLICATIONS AND LIMITATIONS

In this section, we will look at some concrete examples of how the two algorithms can be used to determine if some quantum states violate a Bell inequality. Specifically, we begin by looking at how the second algorithm can be used to determine, both numerically and analytically, if some bipartite qudit state violates the Bell-CHSH inequality. Then in Sec. III B, we demonstrate how the two algorithms can be used in tandem to determine if a class of two-qubit states violate a Bell-3322 inequality [31]. We will conclude this section by pointing out some limitations of the UB algorithm that we have observed.

A. Bell-CHSH violation for two-qudit states

The Bell-CHSH inequality, as given by Eq. (4), is one that amounts to choosing

$$b = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{31}$$

For low-dimensional quantum systems, an upper bound on $\mathcal{S}_{\mathrm{QM}}(\rho)$ can be efficiently computed in MATLAB following the procedures described in Sec. II C 3. However, for high-dimensional quantum systems, intensive computational resources are required to compute this upper bound, which may render the computation infeasible in practice. Fortunately, for a specific class of two-qudit states, namely those whose coherence vectors [52] vanish, it can be shown that (Appendix D) their $\mathcal{S}_{\mathrm{QM}}(\rho) = \max_{\mathcal{B}_{\mathrm{CHSH}}} \langle \mathcal{B}_{\mathrm{CHSH}} \rangle_\rho$ cannot exceed

$$\max_{z_1,z_2,z_3,z_4} 2\sqrt{2} s_1 d \sqrt{\prod_{i=1}^{2} \frac{2d^2 - z_{2i-1}^2 - z_{2i}^2}{2d^2}} + \sum_{k,l} b_{kl} \frac{z_k z_{l+2}}{d^2}, \tag{32}$$

where s_1 is the largest singular value of the matrix R' defined in Eq. (C2a), and z_m is the trace of the dichotomic observable O_m given in Eqs. (29a) and (29b).

To violate the Bell-CHSH inequality, we must have $\mathcal{S}_{\mathrm{QM}}(\rho) > 2$, hence for this class of quantum states, the Bell-CHSH inequality cannot be violated if

$$\max_{z_1,z_2,z_3,z_4} \sqrt{2} s_1 d \sqrt{\prod_{i=1}^{2} \frac{2d^2 - z_{2i-1}^2 - z_{2i}^2}{2d^2}} + \sum_{k,l} b_{kl} \frac{z_k z_{l+2}}{2d^2} \leqslant 1. \tag{33}$$

Since the Bell-CHSH inequality is tight Eq. (33) guarantees the existence of a LHV model for this experimental setup [54]. Essentially, the semianalytic upper bound Eq. (32) was obtained by considering a particular choice of Lagrange multipliers in the Lagrange dual function (26). Hence it is generally not as tight as the upper bound obtained numerically using the procedures described in Sec. II C 3 (see Appendix D for details).

As an example, consider the d-dimensional isotropic state, i.e., a mixture of the d-dimensional maximally entangled state $|\Psi_d^+\rangle \equiv 1/\sqrt{d}\sum_{j=1}^{d} |j\rangle|j\rangle$ and the completely mixed state:

$$\rho_{I_d} = p|\Psi_d^+\rangle\langle\Psi_d^+| + (1-p)\frac{\mathbf{1}_{d^2}}{d^2}. \tag{34}$$

As can be verified using the positive partial transposition (PPT) criterion [55,56], this state is entangled if and only if

TABLE I. The various threshold values for isotropic states. The first column of the table is the dimension of the local subsystem d. From the second column to the fifth column, we have, respectively, the value of p beyond which the state is entangled p_{Ent}, the value of p below which Eq. (33) is satisfied $p_{\mathrm{UB\text{-}semianalytic}}$, the value of p below which the upper bound obtained from lowest order relaxation is compatible with Bell-CHSH inequality, and the value of p beyond which a Bell-CHSH violation has been observed using the LB algorithm.

d	p_{Ent}	$p_{\mathrm{UB\text{-}semianalytic}}$	$p_{\mathrm{UB\text{-}numerical}}$	p_{LB}
2	0.33333	0.70711	0.70711	0.70711
3	0.25000	0.70711	0.70711	0.76297
4	0.20000	0.65465	0.65465	0.70711
5	0.16667	0.63246	0.63246	0.74340
10	0.09091	0.51450		0.70711
25	0.03846	0.36490		0.71516
50	0.01961	0.26963		0.70711

$p > p_{\mathrm{Ent}} \equiv 1/(d+1)$. Using the procedures outlined in Sec. II C 3, we can numerically compute, up until $d=5$, the threshold value of p below which there can be no violation of the Bell-CHSH inequality; these critical values, denoted by $p_{\mathrm{UB\text{-}numerical}}$ can be found in column 4 of Table I. Similarly, we can numerically compute the corresponding threshold values, denoted by $p_{\mathrm{UB\text{-}semianalytic}}$, using Eq. (33). It is worth noting that these threshold values, as can be seen from columns 3 and 4 of Table I, agree exceptionally well, thereby suggesting that the computationally feasible criterion given by Eq. (33) may be exact for the isotropic states.

B. I_{3322}-violation for two-qubit states

Next, we look at how the two algorithms can be used in tandem to determine if some two-qubit states violate a Bell-3322 inequality, in particular the I_{3322} inequality [31]. This Bell inequality is interesting in that there are quantum states that violate this new inequality but not the Bell-CH nor the Bell-CHSH inequality. The analog of Horodecki's criterion for this inequality is thus very desired for the characterization of quantum states that are inconsistent with local realistic description.

To the best of our knowledge, such an analytic criterion is yet to be found. However, by combining the two algorithms presented above, we can often offer a definitive, yet nontrivial, conclusion about the compatibility of a quantum state description with that given by LHV models. To begin with, we recall that the I_{3322} inequality reads

$$\mathcal{S}_{\mathrm{LHV}} = p_{AB}^{+-}(1,1) + p_{AB}^{+-}(1,2) + p_{AB}^{+-}(1,3) + p_{AB}^{+-}(2,1) + p_{AB}^{+-}(2,2) - p_{AB}^{+-}(2,3) + p_{AB}^{+-}(3,1) - p_{AB}^{+-}(3,2) - p_A^{+}(1) - 2p_B^{-}(1) - p_B^{-}(2) \leqslant 0. \tag{35}$$

Together with Eqs. (2) and (10) one can then obtain the corresponding Bell operator for probabilities:

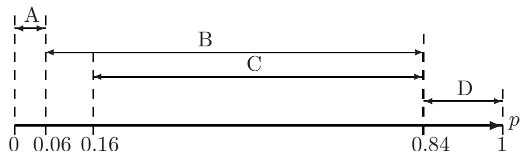

FIG. 1. Domains of p where the compatibility between local realism and quantum mechanical description given by ρ_{CG} was studied via the I_{3322} inequality. From right to left are, respectively, the domain of p whereby ρ_{CG} is (D) found to violate the I_{3322} inequality; (C) found to give a lowest order upper bound that is compatible with the I_{3322} inequality; (B) found to give a higher order upper bound that is compatible with the I_{3322} inequality; and (A) not known if it violates the I_{3322} inequality.

$$\mathcal{B}_{3322} = A_1^{+} \otimes (B_1^{-} - B_2^{+} + B_3^{-}) - A_2^{+} \otimes B_3^{-} - A_2^{-} \otimes (B_1^{-} + B_2^{-}) - A_3^{+} \otimes B_2^{-} - A_3^{-} \otimes B_1^{-}, \tag{36}$$

where we have also used Eq. (12a) to arrive at this form. For convenience, we will adopt the notation that $O_m^{\pm} \equiv A_m^{\pm}$ for $m=1,2,3$ and $O_m^{\pm} \equiv B_{m-3}^{\pm}$ for $m=4,5,6$. In these notations, the global optimization problem for this Bell inequality can be written as

$$\text{maximize} \quad \mathrm{tr}(\rho \mathcal{B}_{3322}) \tag{37a}$$

$$\text{subject to} \quad (O_m^{\pm})^2 = O_m, \tag{37b}$$

which is again a QCQP [57]. The lowest order relaxation to this problem can thus be obtained by following similar procedures as that described in Sec. II C.

To obtain a *state-dependent* upper bound on $\mathcal{S}_{\mathrm{QM}}(\rho)$ for this inequality, we have to impose the analog of Eq. (29b), i.e.,

$$z_m^{\pm} = \mathrm{tr}(O_m^{\pm}) = 0,1,\ldots,d, \tag{38}$$

for each of the POVM elements. For small d, numerical upper bounds on $\mathcal{S}_{\mathrm{QM}}(\rho)$ can then be solved for using SOSTOOLS. As an example, let us now look at how this upper bound, together with the LB algorithm, has enabled us to determine if a class of mixed two-qubit states violates the I_{3322} inequality.

The mixed two-qubit state

$$\rho_{CG} = p|\Psi_{2:1}\rangle\langle\Psi_{2:1}| + (1-p)|0\rangle\langle 0| \otimes |1\rangle\langle 1|, \quad 0 \leqslant p \leqslant 1, \tag{39}$$

can be understood as a mixture of the pure product state $|0\rangle|1\rangle$ and the nonmaximally entangled two-qubit state $|\Psi_{2:1}\rangle = \frac{1}{\sqrt{5}}(2|0\rangle|0\rangle + |1\rangle|1\rangle)$. As can be easily verified using the PPT criterion [55], this state is entangled for $0 < p \leqslant 1$. In particular, the mixture with $p=0.85$ was first presented in Ref. [31] as an example of a two-qubit state that violates the I_{3322} inequality but not the Bell-CH nor the Bell-CHSH inequality.

Given the above observation, a natural question that one can ask is, at what values of p does ρ_{CG} violate the I_{3322} inequality? Using the LB algorithm, we have found that for $p \gtrsim 0.837\,82$ (domain D in Fig. 1), ρ_{CG} violates the I_{3322} inequality. As we have pointed out in Sec. II B, observables that lead to the observed level of I_{3322} violation can be readily read off from the output of the SDP.

On the other hand, through the UB algorithm, we have also found that, with the lowest order relaxation, the states do not violate this three-setting inequality for $0.160\,23 \lesssim p \lesssim 0.836\,25$ (domain C in Fig. 1); with a higher order relaxation, this range expands to $0.062\,91 \lesssim p \lesssim 0.837\,82$ (domain B in Fig. 1); the next order relaxation is, unfortunately, beyond what the software can handle.

The algorithms alone therefore leave a tiny gap at $0 < p \lesssim 0.062\,91$ (domain A in Fig. 1) where we could not conclude if ρ_{CG} violates the I_{3322} inequality. Nevertheless, if we recall that the set of quantum states not violating a given Bell inequality is convex and that $\rho_{CG}(p=0)$, being a pure product state, cannot violate any Bell inequality, we can immediately conclude that ρ_{CG} with $0 \leqslant p \lesssim 0.837\,82$ cannot violate the I_{3322} inequality. As such, together with convexity arguments, the two algorithms allow us to fully characterize the state ρ_{CG} compatible with LHV theories, when each observer is only allowed to perform three different dichotomic measurements.

C. Limitations of the UB algorithm

As can be seen in the above examples, the UB algorithm does not always provide a very good upper bound for $\mathcal{S}_{\rm QM}(\rho)$. In fact, it has been observed that for pure product states, the algorithm with lowest order relaxation always returns a state-independent bound (the Tsirelson bound in the case of Bell-CHSH inequality). As such, for mixed states that can be decomposed as a high-weight mixture of pure product state and some other entangled state, the upper bound given by UB is typically bad. To illustrate this, let us consider the following one-parameter PPT bound entangled state [58,59]:

$$\rho_H = \frac{8p}{8p+1}\rho_{\rm Ent} + \frac{1}{8p+1}|\Psi_p\rangle\langle\Psi_p|, \quad 0<p<1, \tag{40a}$$

where

$$\rho_{\rm Ent} = \frac{3}{8}|\Psi_3^+\rangle\langle\Psi_3^+| + \frac{1}{8}\sum_{i,j=0,i\neq j}^{2}|i\rangle\langle i|\otimes|j\rangle\langle j| - \frac{1}{8}|2\rangle\langle 2|\otimes|0\rangle\langle 0|,$$

$$|\Psi_p\rangle = |2\rangle\left(\sqrt{\frac{1+p}{2}}|0\rangle + \sqrt{\frac{1-p}{2}}|2\rangle\right). \tag{40b}$$

It can be shown, for example, using the range criterion [59] that this state is entangled when $0<p\leqslant 1$. It is, however, well-known that a bipartite PPT state cannot violate the Bell-CH nor the Bell-CHSH inequality [60].

When tested with the UB algorithm using the lowest order relaxation, it turned out that some of these upper bounds are actually above the threshold of Bell-CH violation (see Fig. 2). In fact, the upper bound obtained for the pure product state, $\rho_H(p=0)=|\Psi_p\rangle\langle\Psi_p|$ is actually the maximal achievable Bell-CH violation given by a quantum state [19].

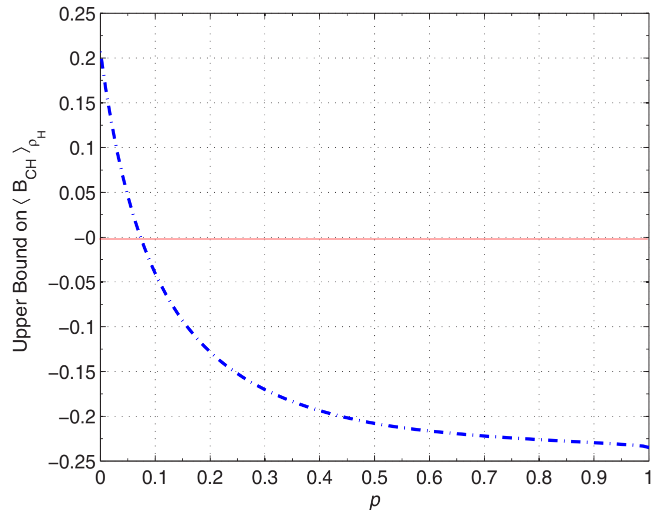

FIG. 2. (Color online) Numerical upper bound on maximal $\langle\mathcal{B}_{\rm CH}\rangle_{\rho_H}$ obtained from the UB algorithm using lowest order relaxation and Eq. (38). The dotted horizontal line is the threshold above which no LHV description is possible.

IV. CONCLUSION

In this paper, we have presented two algorithms which can be used to determine, respectively, a lower bound (LB) and an upper bound (UB) on the maximal expectation value of a Bell operator for a given quantum state, i.e., $\mathcal{S}_{\rm QM}(\rho)$. In particular, we have demonstrated how one can make use of the upper bound to derive a necessary condition for two-qudit states with vanishing coherence vectors to violate the Bell-CHSH inequality.

For low dimensional quantum systems, we have also demonstrated how one can make use of the two algorithms to determine, numerically, if the quantum mechanical prediction is compatible with local realistic description. On a separate note, these algorithms have also been applied to the search of maximal-Bell-inequality-violation in the context of collective measurements without postselection [61].

As with many other numerical optimization algorithms, the algorithm to determine a lower bound (LB) on $\mathcal{S}_{\rm QM}(\rho)$ can only guarantee the convergence to a local maximum. The UB algorithm, on the other hand, provides an (often loose) upper bound on $\mathcal{S}_{\rm QM}(\rho)$. In the event that these bounds agree (up until reasonable numerical precision), we know that optimization of the corresponding Bell operator using LB has been achieved. This ideal scenario, however, is not as common as we would like it to be. In particular, the UB algorithm with lowest order relaxation has been observed to give rather bad bounds for states with a high-weight mixture of pure product states (although it appears that we can often rule out the possibility of a violation in this situation by convexity arguments as in Sec. III B). A possibility to improve these bounds, as suggested by the work of Nie *et al.* [62], is to incorporate the *Karush-Kuhn-Tucker* optimality condition as an additional constraint to the problem. We have done some preliminary studies on this but have not so far found any improvement in the bounds obtained but this deserves further study.

As of now, we have only implemented the UB algorithm to determine upper bounds on $\mathcal{S}_{\rm QM}(\rho)$ for dichotomic Bell inequalities. For Bell inequalities with more outcomes, the local Hermitian observables are generally also subjected to constraints in the form of a LMI. Although the UB algorithm can still be implemented for these Bell inequalities by first mapping the LMI to a series of polynomial inequalities, this approach seems blatantly inefficient. Future work to remedy this difficulty is certainly desirable.

Finally, despite the numerical and analytic evidence at hand, it is still unclear why the lowest order relaxation to the global optimization problem, as described in Sec. II C 2, seems to always give rise to a bound that is state-independent and how generally this is true. Some further investigation on this may be useful, particularly to determine whether the lowest order relaxation is always state independent even for inequalities that are not correlation inequalities. If so this could complement the methods of Refs. [22,24,25] for finding state-independent bounds on Bell inequalities.

ACKNOWLEDGMENTS

This work was supported by the Australian Research Council. We thank Ben Toner and Stephanie Wehner for fruitful discussions. Y.-C. Liang would also like to thank Chris Foster and Shiang Yong Looi for their suggestions on the actual implementation of the algorithms.

APPENDIX A: EXPLICIT FORMS OF SEMIDEFINITE PROGRAMS

Here, we provide an explicit form for the matrices F_m and constants c_m that define the semidefinite program (14). By setting

$$Z = B_1^1 \oplus \cdots \oplus B_1^{n_B} \oplus B_2^1 \oplus \cdots \oplus B_{m_B}^{n_B},$$

$$F_0 = -(\rho_{B_1^1} \oplus \cdots \oplus \rho_{B_1^{n_B}} \oplus \rho_{B_2^1} \oplus \cdots \oplus \rho_{B_{m_B}^{n_B}})$$

in Eq. (8), we see that the inequality constraint (8c) of the SDP entails the positive semidefiniteness of the POVM elements $\{B_l^\lambda\}_{\lambda=1}^{n_B}$, and hence Eq. (14c). On the other hand, the equality constraint (8b), together with the appropriate choice of F_m and c_m, ensures that the normalization condition (14b) is satisfied.

In particular, each F_m is formed from a direct sum of Hermitian basis operators. A convenient choice of such basis operators is given by the traceless Gell-Mann matrices, denoted by $\{\sigma_i\}_{n=1}^{d^2-1}$, supplemented by $\sigma_0 = \mathbf{1}_d/\sqrt{d}$ such that

$$\mathrm{tr}(\sigma_n \sigma_{n'}) = \delta_{nn'} \quad \text{and} \quad \mathrm{tr}(\sigma_n) = \sqrt{d}\,\delta_{n0}, \tag{A1}$$

where $d=d_B$ is the dimension of the state space that each B_l^λ acts on. A typical F_m then consists of n_B diagonal blocks of σ_i at positions corresponding to the n_B POVM elements $\{B_l^\lambda\}_{\lambda=1}^{n_B}$ in Z for a fixed l. For instance, the set of F_m

$$F_m = \begin{pmatrix} \sigma_{m-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma_{m-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad 1 \leqslant m \leqslant d^2,$$

together with $c_m = \sqrt{d}\,\delta_{m1}$ entails the normalization of $\{B_1^\lambda\}_{\lambda=1}^{n_B}$, i.e., $\sum_{\lambda=1}^{n_B} B_1^\lambda = \mathbf{1}_{d_B}$; the remaining $(m_B-1)d^2$ F_m are defined similarly and can be obtained by shifting the nonzero diagonal blocks diagonally downward by appropriate multiples of n_B blocks. The SDP thus consists of solving Eq. (8) for a $m_B n_B d \times m_B n_B d$ Hermitian matrix Z subjected to $d^2 m_B$ affine constraints.

APPENDIX B: SEMIDEFINITE RELAXATION TO THE GLOBAL OPTIMIZATION PROBLEM

The global optimization problem, either in the form of Eq. (18) for a two-outcome Bell correlation inequality, or Eq. (37) for a two-outcome Bell inequality for probabilities, is a QCQP. As was demonstrated in Sec. II C, an upper bound on $\mathcal{S}_{\rm QM}(\rho)$ can then be obtained by considering the corresponding *Lagrange Dual*.

More generally, the global optimization problem can be mapped to a real polynomial optimization problem:

$$\text{maximize} \quad f_{\rm obj}(\mathbf{y}) \tag{B1a}$$

$$\text{subject to} \quad f_{{\rm eq},i}(\mathbf{y}) = 0, \quad i = 1,2,\ldots,N_{eq}, \tag{B1b}$$

$$f_{{\rm ineq},j}(\mathbf{y}) \geqslant 0, \quad j = 1,2,\ldots,N_{ineq}, \tag{B1c}$$

where $\mathbf{y}$ is a vector of *real* variables formed by the expansion coefficients of local observables $\{O_m\}$ in terms of Hermitian basis operators.

Results from semialgebraic geometry dictate that an upper bound for $f_{\rm obj}(\mathbf{y})$ can be computed using Positivstellensatz-based relaxations (see, for example, Ref. [28] and references therein). In particular, γ will be an upper bound on the constrained optimization problem (B1) if there exists a set of sum of squares (SOS) $\mu_i(\mathbf{y})$'s [i.e., non-negative, real polynomials that can be written as $\sum_j [h_j(\mathbf{y})]^2$ with $h_j(\mathbf{y})$ being some real polynomials of $\mathbf{y}$] and a set of real polynomials $\nu_j(\mathbf{y})$ such that [28,46,51]

$$\gamma - f_{\rm obj}(\mathbf{y}) = \mu_0(\mathbf{y}) + \sum_j \nu_j(\mathbf{y}) f_{{\rm eq},j}(\mathbf{y}) + \sum_i \mu_i(\mathbf{y}) f_{{\rm ineq},i}(\mathbf{y}) + \sum_{i_1,i_2} \mu_{i_1,i_2}(\mathbf{y}) f_{{\rm ineq},i_1}(\mathbf{y}) f_{{\rm ineq},i_2}(\mathbf{y}) + \cdots. \tag{B2}$$

The relaxed optimization problem then consists of minimizing γ subjected to the above constraint. Clearly, at values of $\mathbf{y}$ where the constraints are satisfied, γ gives an upper bound on $f_{\rm obj}(\mathbf{y})$. The auxiliary polynomials $\nu_j(\mathbf{y})$ and SOS $\mu_i(\mathbf{y})$ then serve as the Lagrange multipliers in the relaxed optimization problem.

For a fixed degree of the above expression, this relaxed optimization problem can be cast as an SDP in the form of Eq. (9) [28]. For the lowest order relaxation, the auxiliary polynomials $\nu_j(\mathbf{y})$ and SOS $\mu_i(\mathbf{x})$ are chosen such that the degree of the expression in Eq. (B2) is no larger than the maximum degree of the set of polynomials

$$f_{\text{obj}}(\mathbf{y}),f_{\text{eq},1}(\mathbf{y}),\ldots,f_{\text{eq},N_{\text{eq}}}(\mathbf{y}),f_{\text{ineq},1}(\mathbf{y}),\ldots,f_{\text{ineq},N_{\text{ineq}}}(\mathbf{y});$$

for a QCQP with no inequality constraints, this amounts to setting all the $\mu_i(\mathbf{y})$ to zero and all the $\nu_j(\mathbf{y})$ to numbers.

For higher order relaxation, we increase the degree of the expression in Eq. (B2) by increasing the degree of the auxiliary polynomials. At the expense of involving more computational resources, a tighter upper of $f_{\text{obj}}(\mathbf{y})$ can then be obtained by solving the corresponding SDPs.

APPENDIX C: LOWEST ORDER RELAXATION WITH OBSERVABLES OF FIXED TRACE

To obtain a tighter upper bound by using the lowest order relaxation to Eq. (18), we found it most convenient to express Eq. (18) in terms of the *real* optimization variables,

$$y_{mn} \equiv \text{tr}(O_m\sigma_n), \quad n=0,1,\ldots,d^2-1, \tag{C1}$$

which are just the expansion coefficients of each O_m in terms of a set of Hermitian basis operators satisfying Eq. (A1). The constraints can then be taken care of by setting each $y_{m0}=z_m/\sqrt{d}$. It is also expedient to express the density matrix ρ in terms of the same basis of Hermitian operators

$$\rho = \frac{\mathbf{1}_{d^2}}{d^2} + \sum_{i=1}^{d^2-1}([\mathbf{r}_A]_i\sigma_i\otimes\sigma_0 + [\mathbf{r}_B]_i\sigma_0\otimes\sigma_i) + \sum_{i,j=1}^{d^2-1}[R']_{ij}\sigma_i\otimes\sigma_j,$$

where

$$[R']_{ij} = \text{tr}(\rho\sigma_i\otimes\sigma_j), \tag{C2a}$$

$$[\mathbf{r}_A]_i \equiv \text{tr}(\rho\sigma_i\otimes\sigma_0), \quad [\mathbf{r}_B]_j \equiv \text{tr}(\rho\sigma_0\otimes\sigma_j); \tag{C2b}$$

$\mathbf{r}_A$, $\mathbf{r}_B$ are simply the coherence vectors that have been studied in the literature [53].

We can thus incorporate the constraints (29) by expressing the Lagrangian (22) as a function of the reduced set of variables

$$\mathbf{y}'^T \equiv [y_{11}\ y_{12}\ \cdots\ y_{1\ d^2-1}\ y_{21}\ \cdots\ y_{m_A+m_B\ d^2-1}], \tag{C3}$$

while all the $y_{m0}=z_m/\sqrt{d}$ are treated as fixed parameters of the problem. With this change in basis, and after some algebra, the Lagrangian can be rewritten as

$$\mathcal{L}(\mathbf{y}',\lambda_{mn}) = \sum_{m=1}^{m_A+m_B}\lambda_{m0}\left(\sqrt{d}-\frac{z_m^2}{d\sqrt{d}}\right) + \sum_{k,l} b_{kl}\frac{z_k z_{l+m_A}}{d^2} - \frac{1}{\sqrt{d}}(\mathbf{l}-\mathbf{r})^T\mathbf{y}' - \mathbf{y}'^T\Omega'\mathbf{y}', \tag{C4}$$

where λ_{mn} are defined in Eq. (28),

$$\Omega' \equiv \frac{1}{2}\begin{pmatrix} \mathbf{0} & -b\otimes R' \\ -b^T\otimes R'^T & \mathbf{0}\end{pmatrix} + \bigoplus_{m=1}^{m_A+m_B} M_m,$$

$$\mathbf{l} \equiv \text{vec}(L), \quad \mathbf{r} \equiv \begin{pmatrix}\mathbf{t}_A\otimes\mathbf{r}_A \\ \mathbf{t}_B\otimes\mathbf{r}_B\end{pmatrix}, \tag{C5a}$$

and for $i,j=1,2,\ldots,d^2-1$,

$$[L]_{jm} = 2z_m\lambda_{mj}, \quad [\mathbf{t}_A]_k = \sum_l b_{kl}z_{l+m_A}, \quad [\mathbf{t}_B]_l = \sum_k b_{kl}z_k,$$

$$M_m = \sum_{n=0}^{d^2-1}\lambda_{mn}P_n, \quad [P_n]_{ij} = \frac{1}{2}\text{tr}(\sigma_n[\sigma_i,\sigma_j]_+); \tag{C5b}$$

$[\sigma_i,\sigma_j]_+ \equiv \sigma_i\sigma_j+\sigma_j\sigma_i$ is the anticommutator of σ_i and σ_j.

As before, we now maximize the Lagrangian (C4) over $\mathbf{y}'$ to obtain the corresponding Lagrange dual function. The latter, however, is unbounded above unless

$$\begin{pmatrix} -2t & \frac{1}{\sqrt{d}}(\mathbf{l}^T-\mathbf{r}^T) \\ \frac{1}{\sqrt{d}}(\mathbf{l}-\mathbf{r}) & 2\Omega' \end{pmatrix} \geqslant 0, \tag{C6}$$

for some finite t. The convex optimization problem dual to Eq. (18) with fixed trace for each observables is thus

$$\text{minimize} \quad \sum_{m=1}^{m_A+m_B}\lambda_{m0}\left(\sqrt{d}-\frac{z_m^2}{d\sqrt{d}}\right) + \sum_{k,l}b_{kl}\frac{z_k z_{l+m_A}}{d^2} - t$$

$$\text{subject to} \quad \begin{pmatrix} -2t & \frac{1}{\sqrt{d}}(\mathbf{l}^T-\mathbf{r}^T) \\ \frac{1}{\sqrt{d}}(\mathbf{l}-\mathbf{r}) & 2\Omega' \end{pmatrix} \geqslant 0. \tag{C7}$$

APPENDIX D: NONLOCALITY CRITERION

To derive the semianalytic criterion Eq. (33), we now note that any choices of $\{\lambda_{mn}\}_{n=0}^{d^2-1}$ that satisfy constraint (C6) will provide an upper bound on the corresponding $\mathcal{S}_{\text{QM}}(\rho)$. In particular, an upper bound can be obtained by setting

$$\lambda_{mn} = \delta_{n0}[\lambda_A(\delta_{m1}+\delta_{m2}) + \lambda_B(\delta_{m3}+\delta_{m4})] \tag{D1}$$

and solving for λ_A, λ_B that satisfy the constraint (C6). With this choice of the Lagrange multipliers, and for quantum states with vanishing coherence vectors, the constraint (C6) becomes

$$\begin{pmatrix} -2t & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{2\lambda_A}{\sqrt{d}}\mathbf{1}_2 \otimes \mathbf{1}_{d^2-1} & -b \otimes R' \\ \mathbf{0} & -b \otimes R'^T & \frac{2\lambda_B}{\sqrt{d}}\mathbf{1}_2 \otimes \mathbf{1}_{d^2-1} \end{pmatrix} \geq 0, \quad \text{(D2)}$$

where b and R' are defined, respectively, in Eq. (31) and Eq. (C2a). This, in turn is equivalent to

$$-t \geq 0, \quad \text{(D3a)}$$

$$\begin{pmatrix} \frac{2\lambda_A}{\sqrt{d}}\mathbf{1}_2 \otimes \mathbf{1}_{d^2-1} & -b \otimes R' \\ -b \otimes R'^T & \frac{2\lambda_B}{\sqrt{d}}\mathbf{1}_2 \otimes \mathbf{1}_{d^2-1} \end{pmatrix} \geq 0. \quad \text{(D3b)}$$

Using Schur's complement [42], the constraint (D3b) can be explicitly solved to give

$$\lambda_A \lambda_B \geq \frac{1}{2} s_1^2 d,$$

where s_1 is the largest singular value of the matrix R'. Substituting this and Eq. (D3a) into Eq. (C7), and after some algebra, we see that $\langle \mathcal{B}_{\mathrm{CHSH}} \rangle_\rho$ for a quantum state ρ with vanishing coherence vectors cannot be greater than

$$\max_{z_1,z_2,z_3,z_4} 2\sqrt{2} s_1 d \sqrt{\prod_{i=1}^{2} \frac{2d^2 - z_{2i-1}^2 - z_{2i}^2}{2d^2}} + \sum_{k,l} b_{kl} \frac{z_k z_{l+2}}{d^2}.$$

For ρ to violate the Bell-CHSH inequality, we must have this upper bound greater than the classical threshold value of 2 [cf. Eq. (4)]. Hence a sufficient condition for ρ to satisfy the Bell-CHSH inequality is given by Eq. (33).

[1] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
[2] E. Schrödinger, Naturwiss. **23**, 807 (1935).
[3] J. S. Bell, Physics (Long Island City, N.Y.) **1**, 195 (1964).
[4] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).
[5] N. Gisin, Phys. Lett. A **154**, 201 (1991).
[6] N. Gisin and A. Peres, Phys. Lett. A **162**, 15 (1992).
[7] J. F. Clauser, M. A. Horne, A. Shimony, and R. Holt, Phys. Rev. Lett. **23**, 880 (1969); J. S. Bell, in *Foundation of Quantum Mechanics, Proceedings of the International School of Physics "Enrico Fermi," course IL* (Academic, New York, 1971), pp. 171–181.
[8] S. Popescu and D. Rohrlich, Phys. Lett. A **166**, 293 (1992).
[9] R. Horodecki, P. Horodecki, and M. Horodecki, Phys. Lett. A **200**, 340 (1995).
[10] J.-Å. Larsson, Phys. Lett. A **256**, 245 (1999); J. Barrett, Phys. Rev. A **65**, 042302 (2002); B. M. Terhal, A. C. Doherty, and D. Schwab, Phys. Rev. Lett. **90**, 157903 (2003); J. Degorre, S. Laplante, and J. Roland, Phys. Rev. A **72**, 062314 (2005); A. Acín, N. Gisin, and B. Toner, *ibid.* **73**, 062105 (2006).
[11] S. L. Braunstein, A. Mann, and M. Revzen, Phys. Rev. Lett. **68**, 3259 (1992).
[12] M. M. Deza and M. Laurent, *Geometry of Cuts and Metrics*, Vol. 15, in *Algorithms and Combinatorics* (Springer, New York, 1997).
[13] S. Popescu and D. Rohrlich, Phys. Lett. A **169**, 411 (1992).
[14] T. Ito, H. Imai, and D. Avis, Phys. Rev. A **73**, 042109 (2006).
[15] S. Popescu, Phys. Rev. Lett. **72**, 797 (1994); R. Horodecki, M. Horodecki, and P. Horodecki, Phys. Lett. A **222**, 21 (1996).
[16] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991); A. Acín, N. Gisin, Ll. Masanes, and V. Scarani, Int. J. Quantum Inf. **2**, 23 (2004); J. Barrett, L. Hardy, and A. Kent, Phys. Rev. Lett. **95**, 010503 (2005).
[17] A. Acín, N. Gisin, and Ll. Masanes, Phys. Rev. Lett. **97**, 120405 (2006); Ll. Masanes and A. Winter, e-print quant-ph/0606049.
[18] H. Buhrman, R. Cleve, and W. van Dam, SIAM J. Comput. **30**, 1829 (2001); V. Scarani and N. Gisin, Phys. Rev. Lett. **87**, 117901 (2001); Č. Brukner, M. Żukowski, and A. Zeilinger, *ibid.* **89**, 197901 (2002); Č. Brukner, M. Żukowski, J. W. Pan, and A. Zeilinger, *ibid.* **92**, 127901 (2004).
[19] B. S. Cirel'son, Lett. Math. Phys. **4**, 93 (1980).
[20] S. Filipp and K. Svozil, Phys. Rev. Lett. **93**, 130407 (2004); S. Filipp and K. Svozil, Phys. Rev. A **69**, 032101 (2004).
[21] H. Buhrman and S. Massar, Phys. Rev. A **72**, 052103 (2005).
[22] S. Wehner, Phys. Rev. A **73**, 022110 (2006).
[23] B. F. Toner, e-print quant-ph/0601172.
[24] D. Avis, H. Imai, and T. Ito, J. Phys. A **39**, 11283 (2006).
[25] M. Navascués, S. Pironio, and A. Acín, Phys. Rev. Lett. **98**, 010401 (2007).
[26] A. Cabello, Phys. Rev. Lett. **92**, 060403 (2004).
[27] F. A. Bovino, G. Castagnoli, I. P. Degiovanni, and S. Castelletto, Phys. Rev. Lett. **92**, 060404 (2004).
[28] P. A. Parrilo, Math. Program. **96**, 293 (2003); P. A. Parrilo and B. Sturmfels, DIMACS Series on Discrete Math. Theoretical Comput. Sci. **60**, 83 (2003).
[29] J. B. Lasserre, SIAM J. Optim. **11**, 796 (2001).
[30] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri, Phys. Rev. A **69**, 022308 (2004); J. Eisert, P. Hyllus, O. Gühne, and M. Curty, *ibid.* **70**, 062317 (2004).
[31] D. Collins and N. Gisin, J. Phys. A **37**, 1775 (2004).
[32] For a given experimental setup, the set of local realistic correlations defines a convex polytope; any facet of this polytope is a tight Bell inequality for the given experimental setup.
[33] J. F. Clauser and M. A. Horne, Phys. Rev. D **10**, 526 (1974).
[34] L. Vandenberghe and S. Boyd, SIAM Rev. **38**, 49 (1996).
[35] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge, New York, 2004).
[36] That such an iterative algorithm using SDP can lead to a local maximum of $\mathcal{S}_{\mathrm{QM}}(\rho, A_k^{\kappa}, B_l^{\lambda})$ was also discovered independently by Ito *et al.* [14].

[37] R. F. Werner and M. M. Wolf, Quantum Inf. Comput. **1**, 1 (2001).

[38] For a Bell correlation inequality, we can apply LB by first rewriting the corresponding Bell operator (16) in terms of POVM elements that form the measurement outcomes of each O_m.

[39] A dichotomic observable is a Hermitian observable with only two distinct eigenvalues.

[40] MATLAB is a trademark of The Math Works, Inc., Natick, MA.

[41] A quantum state violates the Bell-CH inequality if and only if it violates the Bell-CHSH inequality.

[42] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University, Cambridge, England, 1990); G. Strang, *Linear Algebra and Its Applications* (Harcourt Brace, San Diego, 1988).

[43] As was pointed out in Ref. [14], this sufficiency also follows from Theorem 5.4 in Ref. [44].

[44] R. Cleve, P. Høyer, B. Toner, and J. Watrous, in *Proceedings of the 19th IEEE Annual Conference on Computational Complexity, Amherst, MA , 2004* (IEEE, New York, 2004), pp. 236–249.

[45] S. Prajna, A. Papachristodoulou, P. Seiler, and P. A. Parrilo, SOSTOOLS: Sum of squares optimization toolbox for MATLAB, available from http://www.cds.caltech.edu/sostools and http://www.mit.edu/~parrilo/sostools (2004).

[46] S. Prajna, A. Papachristodoulou, P. Seiler, and P. A. Parrilo, *Lecture Notes in Control and Information Sciences Vol. 312* (Springer, Berlin, 2005), p. 273–292.

[47] Strictly speaking, a general Bell correlation operator may also contain marginal terms like $O_m \otimes \mathbf{1}_{d_B}$ and $\mathbf{1}_{d_A} \otimes O_m$. For brevity, we will not consider such Bell operators in the following discussion. It should, nevertheless, be clear to the readers that the following arguments are readily generalized to include Bell operators of this more general kind.

[48] The choice of eigenvalues, nonetheless, has to be consistent with the values assigned to the two outcomes of the respective dichotomic Bell experiments.

[49] In general, the composite system may consist of subsystems of different dimensions. All the following arguments can be readily generalized to this more general scenario.

[50] S. Wehner (private communication); see also Ref. [22].

[51] K. Schmüdgen, Math. Ann. **289**, 203 (1991).

[52] These are generalization of the Bloch vectors representation for higher dimensional quantum systems; see also Ref. [53].

[53] G. Kimura, Phys. Lett. A **314**, 339 (2003); M. S. Byrd and N. Khaneja, Phys. Rev. A **68**, 062322 (2003).

[54] A. Fine, Phys. Rev. Lett. **48**, 291 (1982).

[55] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996); M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1 (1996).

[56] M. Horodecki, P. Horodecki, and R. Horodecki, in *Quantum Information: An Introduction to Basic Theoretical Concepts and Experiments*, Springer Tracts in Modern Physics (Springer, Berlin, 2001).

[57] As was mentioned in Sec. II C 1, the same remark applies to any other bipartite Bell-inequality for probabilities with only two outcomes per measurement setting.

[58] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **80**, 5239 (1998).

[59] P. Horodecki, Phys. Lett. A **232**, 333 (1997).

[60] R. F. Werner and M. M. Wolf, Phys. Rev. A **61**, 062102 (2000).

[61] Y.-C. Liang and A. C. Doherty, Phys. Rev. A **73**, 052116 (2006).

[62] J. Nie, J. Demmel, and B. Sturmfels, Math. Program. **106**, 587 (2006).