Characterizing finite-dimensional quantum behavior

Miguel Navascués,[1] Adrien Feix,[2,3] Mateus Araújo,[2,3] and Tamás Vértesi[4]
[1]*Department of Physics, Bilkent University, Ankara 06800, Turkey*
[2]*Faculty of Physics, University of Vienna, Boltzmanngasse 5, 1090 Vienna, Austria*
[3]*Institute for Quantum Optics and Quantum Information (IQOQI), Boltzmanngasse 3, 1090 Vienna, Austria*
[4]*Institute for Nuclear Research, Hungarian Academy of Sciences, P.O. Box 51, H-4001 Debrecen, Hungary*
(Received 19 August 2015; published 21 October 2015)

We study and extend the semidefinite programming (SDP) hierarchies introduced in Navascués and Vértesi [Phys. Rev. Lett. **115**, 020501 (2015)] for the characterization of the statistical correlations arising from finite-dimensional quantum systems. First, we introduce the dimension-constrained noncommutative polynomial optimization (NPO) paradigm, where a number of polynomial inequalities are defined and optimization is conducted over all feasible operator representations of bounded dimensionality. Important problems in device-independent and semi-device-independent quantum information science can be formulated (or almost formulated) in this framework. We present effective SDP hierarchies to attack the general dimension-constrained NPO problem (and related ones) and prove their asymptotic convergence. To illustrate the power of these relaxations, we use them to derive a number of dimension witnesses for temporal and Bell-type correlation scenarios, and also to bound the probability of success of quantum random access codes.

DOI: 10.1103/PhysRevA.92.042117 PACS number(s): 03.65.Ud, 03.67.Hk, 03.67.Mn

I. INTRODUCTION

Many problems in quantum information theory can be formulated as optimizations over operator algebras of a given dimensionality. Let us quickly review some of them.

In one-way quantum communication complexity [1,2], two separate parties, call them Alice and Bob, are respectively handed the bit strings $x,y \in \{0,1\}^n$. Bob's task consists in guessing the value of the Boolean function $f(x,y) \in \{0,1\}$, and, to this aim, we allow Alice to send him a D-dimensional quantum system. Under these conditions, computing the maximum probability that Bob's guess is correct amounts to optimizing over all possible D-dimensional quantum states prepared by Alice and over all possible measurements conducted by Bob on such states.

In Bell scenarios [3], two or more distant parties conduct measurements over an unknown quantum state. It has been observed that, even if we do not assume any knowledge whatsoever about the mechanisms of the measurement devices, it is sometimes possible to lower bound the dimensionality D of the quantum systems accessible to each party by virtue of the correlations between the measurement results alone [4–7]. In this regard, deriving *dimension witnesses*, i.e., statistical inequalities satisfied by the correlations achievable through multipartite quantum systems of local dimension D, can be understood as an optimization over entangled states and measurement operators.

Entanglement distillation [8], or the capacity to prepare states close to a pure singlet given a number of mixed states through local operations and classical communication (LOCC), is one of the most conventional problems in quantum information science. More generally, determining whether the state transformation $\rho \to \sigma$ can be effected via LOCC can be interpreted as a feasibility problem, where the free variable is the corresponding LOCC map. If we restrict to local protocols or one-way LOCC, the set of relevant maps admits a simple characterization in terms of tensor products of Kraus operators satisfying certain quadratic constraints.

The above problems involve optimizations over a tuple of noncommuting variables $X_1,\dots,X_n$ satisfying a number of polynomial constraints, such as $X_i^2 = X_i = X_i^\dagger$ (for projectors) or $X_iX_i^\dagger = X_i^\dagger X_i = \mathbb{I}$ (for unitaries). The number of total constraints is typically so low that, even fixing the dimensionality D of the spaces where these operators act, we find a continuum of inequivalent representations.

Analogous problems emerge in the black-box approach to quantum information theory [3,9–11], where the only constraints considered are essentially commutation relations between projection operators implemented by distant parties. The characterization of quantum nonlocality has boosted the field of noncommutative polynomial optimization (NPO) theory [12–14], where the goal is, precisely, to conduct optimizations over all tuples of operators satisfying a number of polynomial inequalities. NPO theory achieves this via hierarchies of semidefinite programming (SDP) [15] relaxations whose first levels approximate quite well the space of feasible solutions.

Unfortunately, NPO theory does not offer any means to bound or fix the dimension of the Hilbert spaces where such operators act. Since the aforementioned problems in quantum information theory become senseless or trivial in the high dimensionality limit, one would not expect NPO to be of any use.

This view changed with the publication of Ref. [16], where a systematic way to devise hierarchies of SDP relaxations for a wide class of NPO problems under dimension constraints was introduced. Such relaxations, which seem to work quite well in practice, were used to derive a number of new results in quantum nonlocality and quantum communication complexity. Important theoretical aspects, such as the completeness of the hierarchies, or the explicit nature of the dimension constraints, were nonetheless left out. Actually, from a reading of Ref. [16], it is not even clear which problems can be attacked with the new tools.

In this paper, we generalize the SDP schemes proposed in Ref. [16] to cover *all* NPO problems where the dimensionality of the relevant Hilbert spaces is bounded. We prove the convergence of the resulting SDP hierarchies and discuss their efficient implementation. Finally, we use them to derive a number of results in quantum information theory: from new bounds for quantum random access codes (QRACs) for both real and complex quantum systems to semi-device-independent positive operator valued measure (POVM) detection [17,18] and from the characterization of temporal correlations [19] under dimension constraints to the exploration of tripartite Bell scenarios where the dimensionality of just one of the parties is limited.

The structure of this paper is as follows. First, we define the generic problem of NPO under polynomial constraints. Then, in Sec. III, we present a hierarchy of SDP relaxations to tackle it. In Sec. IV we prove the convergence of this hierarchy. As it turns out, a straightforward implementation of the hierarchy would converge too slowly to be of much use, given reasonable computational resources. Hence, in Sec. V we give some hints to boost the speed of convergence of the method—to make it practical—and exemplify its application by solving a specific problem on temporal correlations [19]. In Sec. VI we explore the performance of related SDP hierarchies to characterize quantum nonlocality under dimension constraints and quantum communication complexity. In Sec. VII we offer some advice on how to code the corresponding programs. Finally, we summarize our conclusions.

II. NONCOMMUTATIVE POLYNOMIAL OPTIMIZATION UNDER DIMENSION CONSTRAINTS

Consider the set S of all n-tuples of self-adjoint operators $(X_1,\dots,X_n)$ satisfying the relations $\mathcal{R}=\{q_i(X)\geqslant 0: i=1,\dots,m\}$. Here $q_i(X)$ denotes a Hermitian polynomial of the variables $X_1,\dots,X_n$, while the notation $A\geqslant 0$ signifies that operator A is positive semidefinite. We call each feasible tuple $(X_1,\dots,X_n)$ a *representation* of the polynomial relations $\mathcal{R}$. Given a Hermitian polynomial $p(X)$ and a natural number D, the problem we want to address is how to maximize the maximum eigenvalue of $p(X)$ over all representations of $\mathcal{R}$ of dimension D or smaller. In other words, we want to solve the problem

$$p^\star=\max_{\mathcal{H},X,\psi}\langle\psi|p(X)|\psi\rangle$$

$$\text{such that}\qquad \dim(\mathcal{H})\leqslant D,\quad q_i(X)\geqslant 0,\quad \text{for } i=1,\dots,m, \tag{1}$$

where the maximization is supposed to take place over all Hilbert spaces $\mathcal{H}$ with $\dim(\mathcal{H})\leqslant D$, all tuples of operators $(X_1,\dots,X_n)\subset B(\mathcal{H})$, and all normalized states $|\psi\rangle\in\mathcal{H}$. Note that, if it were not for the dimension restriction D, the above would be a regular NPO problem [14].

We say that the relations $\mathcal{R}$ satisfy the *Archimedean condition* if there exist polynomials $f_j(X),g_{ij}(X)$ such that

$$C-\sum_i X_i^2=\sum_j f_j(X)^\dagger f_j(X)+\sum_{i,j}g_{ij}(X)^\dagger q_i(X)g_{ij}(X). \tag{2}$$

In the following, we provide a hierarchy of SDP relaxations for this problem. Such a hierarchy provides a decreasing sequence of values $p_1\geqslant p_2\geqslant\cdots$ such that $p^k\geqslant p^\star,\forall k$. Moreover, if the Archimedean condition holds,[1] then the hierarchy can be shown complete, i.e., $\lim_{k\to\infty}p^k=p^\star$.

III. THE METHOD

Let $y=(y_w)_{|w|\leqslant 2k}$ be a sequence of complex numbers labeled by monomials w of the variables $X_1,\dots,X_n$ of degree $|w|$ smaller than or equal to $2k$. Such a sequence is called a $2k$th-*order moment vector*. Given y, the kth-order *moment matrix* $M_k(y)$ is an array whose rows and columns are labeled by monomials of $X_1,\dots,X_n$ of degree at most k, and such that

$$M_k(y)_{u,v}=y_{u^\dagger v}. \tag{3}$$

Given $y=(y_w)_{|w|\leqslant 2k}$ and a Hermitian polynomial $q(X)=\sum_w q_w w(X)$, where the w in the summation ranges over all monomials of $X_1,\dots,X_n$ of degree at most $\deg(q)$, the corresponding kth-order *localizing matrix* is defined as

$$M_k(qy)_{u,v}=\sum_w q_w y_{u^\dagger wv}, \tag{4}$$

with $|u|,|v|\leqslant k-\lfloor\frac{\deg(q)}{2}\rfloor$.

A sequence $y=(y_w)_{|w|\leqslant 2k}$ admits a *quantum representation* if there exists a representation $(X_1,\dots,X_n)\subset B(\mathcal{H})$ of relations $\{q_i(X)\geqslant 0\}_i$, with $\dim(\mathcal{H})\leqslant D$, and a normalized vector $|\psi\rangle\in\mathcal{H}$ such that $y_w=\langle\psi|w(X)|\psi\rangle$. It is a standard result in NPO theory that, if (yw) admits a moment representation (of whatever dimensionality), then $M_k(y)$ and $M_k(q_i y)$ must be positive semidefinite matrices for all orders k [14].

The above positive semidefinite constraints are not dimension dependent and are actually obeyed by momenta emerging from representations of $\{q_i(X)\geqslant 0\}_i$ of arbitrary (even infinite) dimensionality. The key to introduce dimension constraints is to acknowledge that moment vectors (y_w) admitting a quantum representation satisfy a number of extra linear restrictions depending on the value of D.

Some of such restrictions arise due to matrix polynomial identities (MPIs) [20]: These are polynomials $s(X)$ of the variables $X_1,\dots,X_n$, which are identically zero when evaluated on matrices of dimensionality D or smaller. For $D=1$, all MPIs reduce to commutators, i.e., $[A_i,A_j]=0$, if $A_i,A_j\in B(\mathbb{C})$. Identifying $A_i=X_i$, this implies that sequences $y=(y_w)_{|w|\leqslant 2k}$ admitting a one-dimensional moment representation must satisfy $y_{X_1X_2}-y_{X_2X_1}=0$. Actually, for any value of D there exist MPIs from which nontrivial linear constraints on y can be derived. For $D=2$, all MPIs are generated by composition of the identities $[[A_1,A_2]^2,A_3]=0$ and $\sum_{\pi\in S_4}\operatorname{sgn}(\pi)A_{\pi(1)}A_{\pi(2)}A_{\pi(3)}A_{\pi(4)}=0$, where S_4 denotes the set of all permutations of four elements. The latter identity is a particular case of the family of polynomial identities I_d, with

$$\sum_{\pi\in S_d}\operatorname{sgn}(\pi)A_{\pi(1)}\cdots A_{\pi(d)}=0. \tag{5}$$

It can be proven that all $D\times D$ matrices satisfy I_{2D} [20], also called the *standard identity*. The problem of determining

[1] Actually, as we will see, a weaker condition suffices.

the generators of *all* MPIs for dimensions D greater than two is, however, open.

A nontrivial relaxation of problem (1) is thus

$$p^k = \max_y \sum_w p_w y_w,$$
$$\text{such that} \qquad (6)$$
$$y \in S_D^k, \quad y_{\mathbb{I}} = 1, \quad M_k(y) \geqslant 0,$$
$$M(q_i y) \geqslant 0, \quad \text{for} \quad i = 1,\dots,m,$$

where S_D^k denotes the span of the set of feasible sequences $y = (y_w)_{|w|\leqslant 2k}$. This is a semidefinite program, and, as such, can be solved efficiently for moment matrices of moderate size (around 200×200) using a normal desktop PC [15].

Equivalently, we can reexpress the positivity conditions as $\hat{M}_k \equiv M_k(y) \oplus \bigoplus_{i=1}^m M_k(q_i y) \geqslant 0$ and rewrite the objective function as a linear combination of the entries of the first diagonal block of $\hat{M}$. That way, we can regard the block-diagonal matrix $\hat{M}$ (and not y) as our free variable, hence arriving at the program

$$p^k = \max_{\hat{M}} \sum_w p_w \hat{M}_{w,\mathbb{I}},$$
$$\text{such that} \qquad (7)$$
$$\hat{M} \in \mathcal{M}_D^k, \quad \hat{M}_{\mathbb{I},\mathbb{I}} = 1, \quad \hat{M} \geqslant 0,$$

where $\mathcal{M}_d^k$ denotes the span of the set of feasible extended moment matrices. This reformulation of problem (6), although conceptually more cumbersome, leads to simpler computer codes.

The key to implementing either program is, of course, to identify the subspaces $S_D^k, \mathcal{M}_D^k$. We now provide two methods to do so. Both have advantages and disadvantages. In [21] we provide yet a third method, which, although more complicated than the other two, requires considerably less memory and time resources, making it suitable for high-order relaxations.

A. The randomized method

We sequentially generate n-tuples of random Hermitian $D \times D$ complex matrices $X^j \equiv (X_1^j,\dots,X_n^j)$ and normalized random vectors $|\psi^j\rangle \in \mathbb{C}^D$, which we use to build moment and localizing matrices $M_{u,v}^j = \langle\psi^j|u(X^j)^\dagger v(X^j)|\psi^j\rangle$, $M(q_i)_{u,v}^j = \langle\psi^j|u(X^j)^\dagger q_i(X) v(X^j)|\psi^j\rangle$, respectively. Their direct sum will constitute an extended moment matrix $\hat{M}^j$. Adopting the Hilbert-Schmidt scalar product $\langle A,B\rangle = \mathrm{tr}(A^\dagger B)$, one can apply the Gram-Schmidt process[2] to the resulting sequence of feasible extended moment matrices in order to obtain an orthogonal basis $\tilde{M}^1, \tilde{M}^2, \dots$ for the space spanned by such matrices. We notice that, for some number N, $\tilde{M}^{N+1} = 0$, up to numerical precision. This is the point at which to terminate the Gram-Schmidt process and define the normalized matrices $\{\Gamma^i \equiv \frac{\tilde{M}^j}{\sqrt{\mathrm{tr}(\tilde{M}^j)^2}} : j = 1,\dots,N\}$. It is easy to see that, even though the matrix basis $\{\Gamma^j\}_{j=1}^N$ was obtained randomly, the space it represents is always the same, namely, $\mathcal{M}_D^k$.

Indeed, let $N = \dim(\mathcal{M}_D^k)$, and suppose that $\tilde{M}^1,\dots,\tilde{M}^{j-1}$ are nonzero, with $j \leqslant N$. Then the entries of the matrix $\tilde{M}^j$ will be polynomials of the components of X^j, ψ^j. Since $j \leqslant N$, there exists a choice of $\vec{z}^j$ such that $\tilde{M}^j(X^j,\psi^j) \neq 0$. The probability that a nonzero polynomial vanishes when evaluated randomly is zero, and so we conclude that $\tilde{M}^j$ will be nonzero with probability 1. On the other hand, $\tilde{M}^1 = \hat{M} \neq 0$, so by induction we have that N randomly chosen moment matrices will span $\mathcal{M}_D^k$ with certainty. Consequently, $\tilde{M}^{N+1} = 0$ indicates when to stop the procedure.

Remark 1. A cautionary note is in order. For high-order k, it is expected that program (7), as written, will not admit strongly feasible points. That is, the subspace $\mathcal{M}_D^k$ will not contain any positive definite matrix. This can be problematic, as many SDP solvers need strong feasibility to operate. The solution is to add up the random extended moment matrices $\hat{M}^j$ as we produce them, i.e., to compute the operator $T = \frac{1}{N}\sum_{j=1}^N \hat{M}^j$. Since $\{\hat{M}^j\}$ were randomly generated, it can be argued that the support of any matrix $\hat{M} \in \mathcal{M}_D^k$ is contained in the support of T. Let V be any matrix mapping supp(T) to $\mathbb{C}^{\dim[\mathrm{supp}(T)]}$ isometrically. We just need to replace the positivity condition $\hat{M} \geqslant 0$ in (7) with $V\hat{M}V^\dagger \geqslant 0$, which, by definition, admits a strictly feasible point.

This method has the advantage that it is extremely easy to program, more so when the constraints $\{q_i(X) \geqslant 0\}$ reduce to polynomial identities, as we will see. One disadvantage is that, in practice, the decision to stop the protocol amounts to verifying that the entries $\tilde{M}^{N+1}$ are zero up to ϵ precision. Choosing the right value for the threshold ϵ is a delicate matter: If too small, the algorithm will not halt; if too large, the algorithm will stop before it finds a complete basis for $\mathcal{M}_D^k$. The second problem is that, due to rounding errors, it is possible that the algorithm will not identify the right subspace, but only an approximation to it.

B. The deterministic method

We choose a simple distribution $f(X,\psi)dXd\psi$, say, a Gaussian, for the entries of each of the matrices $X_1,\dots,X_n$ and the components of the un-normalized vector ψ. Then we define the components of the $2k$th moment vector $y(X,\psi)$ via the relation $y(X,\psi)_w \equiv \langle\psi|w(X)|\psi\rangle$. Then we compute analytically the matrix

$$S \equiv \int f(X,\psi)dXd\psi\, y(X,\psi)y(X,\psi)^\dagger. \qquad (8)$$

Clearly, the space S_D^k corresponds to the support of S. Diagonalizing S and keeping just the eigenvectors with nonzero eigenvalue we hence obtain an orthonormal basis for S_D^k.

The disadvantage of this method is that it involves symbolic computations, and hence, depending on the platform used, it is either more difficult to code or results in slower programs.

IV. CONVERGENCE OF THE HIERARCHY

Let p^k denote the result of the kth-order relaxation (6) of problem (1), and let y^k be the corresponding minimizer $2k$th-order moment vector.

[2]One could even do better using a numerically stable variant, such as the modified Gram-Schmidt method [22].

Now let us assume that the Archimedean condition (2) is met and call r the degree of the polynomial on the right-hand side of Eq. (2). Expressing the polynomials f_i, g_{ij} as $f_i(X) = \sum_v f_i^v v(X)$, $g_{ij}(X) = \sum_v g_{ij}^v v(X)$, it follows that

$$C y^k_{u^\dagger u} - \sum_{l=1}^n y^k_{u^\dagger X_l^2 u} = \sum_i \sum_{v,w} (f_i^v)^* f_i^w y^k_{u^\dagger v^\dagger w u} + \sum_{i,j} \sum_{v,w,s} (g_{ij}^v)^* g_{ij}^w q_i^s y^k_{u^\dagger v^\dagger s w u} \quad (9)$$

for $|u| \leqslant k - \lceil \frac{r}{2} \rceil$. Due to positive semidefiniteness of the moment and localizing matrices, the right-hand side of the above equation is non-negative, implying that $C y^k_{u^\dagger u} \geqslant y^k_{u^\dagger X_l^2 u}$ for all X_l. By induction, it follows that

$$C^{|u|} \geqslant y^k_{u^\dagger u} \quad (10)$$

for all sequences $|u| \leqslant k - \lceil \frac{r}{2} \rceil$. Such moments correspond to the diagonal entries of the moment matrix $M_k(y^k)$. Since $M_k(y^k) \geqslant 0$, it follows that $|y_{|w|}| \leqslant C^{|w|/2}$ for all monomials $|w| \leqslant 2k - 2\lceil \frac{r}{2} \rceil$.

Now, for each vector y^k, replace with zeros all entries y^k_w, with $|w| > 2k - 2\lceil \frac{r}{2} \rceil$, and complete the resulting $2k$th-order moment vector to an ∞-order moment vector by placing even more zeros. We arrive at an infinite sequence $\hat{y}^s, \hat{y}^{s+1}, \ldots$ of vectors, with $|\hat{y}^k_w| \leqslant C^{|w|}$ for all k,w. By the Banach-Alaoglu theorem [23], this sequence has a converging subsequence,[3] and we call $\hat{y}$ the corresponding limit.

$\hat{y}$ satisfies $\hat{y}_{\mathbb{I}} = 1$, and $M_k(\hat{y}), M_k(q_i\hat{y}) \geqslant 0$ for all k,i. By successive Cholesky decompositions of $M_k(\hat{y})$ for $k = s, s+1, \ldots$, we find a sequence of complex vectors $(|u\rangle)_u$ with the property $\hat{y}_{u^\dagger v} = \langle u|v\rangle$ for all monomials u,v. Call $H \equiv \text{span}\{|u\rangle : u\}$. We define the action of the operator $\tilde{X}_i$ on this (nonorthogonal) basis by

$$\tilde{X}_i|u\rangle = |X_i u\rangle \quad (11)$$

and extend its definition to $\text{span}\{|u\rangle : u\}$ by linearity. To prove that this definition is consistent, we need to show that, if $\sum_u c_u|u\rangle = \sum_u d_u|u\rangle$ for two different linear combinations $(c_u)_u, (d_u)_u$, then $\sum_u c_u|X_i u\rangle = \sum_u d_u|X_i u\rangle$. Indeed, note that, for any vector $|w\rangle$,

$$\begin{aligned}\langle w|\sum_u c_u|X_i u\rangle &= \sum_u c_u\langle w|X_i u\rangle = \sum_u c_u\langle X_i w|u\rangle \\ &= \langle X_i w|\sum_u c_u|u\rangle = \langle X_i w|\sum_u d_u|u\rangle \\ &= \langle w|\sum_u d_u|X_i u\rangle, \end{aligned} \quad (12)$$

where the second and fifth equalities follow from $\langle w|X_i u\rangle = y_{w^\dagger X_i u} = y_{(X_i w)^\dagger u} = \langle X_i w|u\rangle$. This relation holds for arbitrary $|w\rangle$, so the vectors $\sum_u c_u|X_i u\rangle, \sum_u d_u|X_i u\rangle$ must be identical. Similarly, it can be verified immediately that $\tilde{X}_i$ is a symmetric operator, since $\langle u|\tilde{X}_i|v\rangle = y_{u^\dagger X_i v} = y^*_{v^\dagger X_i u} = \langle u|\tilde{X}_i|v\rangle$.

From the positive semidefiniteness of the localizing matrices $M(q_i\hat{y})$, it can be shown that $\langle\phi|q_i(\tilde{X})|\phi\rangle \geqslant 0$ for all $|\phi\rangle \in H$ and $i = 1, \ldots, m$. The Archimedean condition implies, moreover, that $\langle\phi|\tilde{X}_i^2|\phi\rangle \leqslant C\langle\phi|\phi\rangle$. From this observation it is trivial to extend the action of $\tilde{X}_i$ to $\tilde{\mathcal{H}}$, the closure of H, and hence we arrive at a Hilbert space $\tilde{\mathcal{H}}$ and a set of operators $\tilde{X}_1, \ldots, \tilde{X}_n$ such that $q_i(\tilde{X}) \geqslant 0$ for $i = 1, \ldots, m$ and $\hat{y}_u = \langle\tilde{\psi}|u(\tilde{X})|\tilde{\psi}\rangle$ for $|\tilde{\psi}\rangle \equiv |\mathbb{I}\rangle$.

Note as well that, by construction, these operators satisfy all MPIs for dimension D. Now, call $\mathcal{A}$ the von Neumann algebra generated by $\tilde{X}_1, \ldots, \tilde{X}_n$. By von Neumann's 1949 result [24], such an algebra must decompose as a direct integral of types I, II, and III factors [25]. That is,

$$\mathcal{A} = \int^\oplus d\mu^{\text{I}}(y)\mathcal{A}^{\text{I}}_y \oplus \int^\oplus d\mu^{\text{II}}(y)\mathcal{A}^{\text{II}}_y \oplus \int^\oplus d\mu^{\text{III}}(y)\mathcal{A}^{\text{III}}_y. \quad (13)$$

Type I factors are isomorphic to $B(\mathcal{H})$ for Hilbert spaces $\mathcal{H}$ of finite or infinite dimensionality [25]. Since $\mathcal{A}$ must satisfy the MPIs for dimension D, that excludes Hilbert spaces of dimension $d > D$ from the first term of the right-hand side of (13). Moreover, in the Appendix it is proven that types II and III factors violate the standard identity (5) for all values of d.

It follows that we can write our operators $\tilde{X}_1, \ldots, \tilde{X}_n$ as

$$\tilde{X}_i = \int^\oplus d\mu^{\text{I}}(y)\tilde{X}_{i,y}, \quad (14)$$

where each $\tilde{X}_{i,y}$ acts on a Hilbert space $\mathcal{H}_y$ with $\dim(\mathcal{H}_y) \leqslant D$. From $q_i(\tilde{X}) \geqslant 0$, it follows that $q_i(\tilde{X}_y) \geqslant 0$ for $i = 1, \ldots, m$. Hence, $\hat{p}$ is a convex combination of feasible values of $\langle p(X)\rangle$ and so $\hat{p} \leqslant p^\star$. On the other hand, $p^k \geqslant p^\star$ for $k \geqslant s$. Thus, $\hat{p} = \lim_{k\to\infty} p^k \geqslant p^\star$, proving the convergence of the hierarchy.

Remark 2. Note that we just invoked the Archimedean condition (2) to establish the existence of $\hat{y}$ and, later, the boundedness of the operators $\tilde{X}_1, \ldots, \tilde{X}_n$. Both results also follow from the weaker *D-dimensional Archimedean condition*,

$$C - \sum_i X_i^2 = \sum_j f_j(X)^\dagger f_j(X) + \sum_{i,j} g_{ij}(X)^\dagger q_i(X) g_{ij}(X) + h_D(X), \quad (15)$$

where $h_D(X)$ is an MPI for dimension D.

Remark 3. If we take $D = 1$, then the MPIs will force all operators $X_1, \ldots, X_n$ to commute with each other. In that case, the SDP hierarchy reduces to the Lasserre-Parrilo hierarchy for polynomial minimization [26,27].

Remark 4. So far, we have been assuming that the variables $X_1, \ldots, X_n$ are Hermitian. If a subset of them is not, one can still define a converging SDP hierarchy, by considering all possible monomials of the variables $X_1, \ldots, X_n$ and their adjoints $X_1^\dagger, \ldots, X_n^\dagger$ in the definition of moment vectors and moment matrices. In that case, the left-hand side of the D-dimensional Archimedean condition (15) must be replaced with $C - \sum_i X_i X_i^\dagger - X_i^\dagger X_i$.

[3] Technically, the Banach-Alaoglu theorem must be applied to the sequence $\hat{z}^s, \hat{z}^{s+1}, \ldots$, where $z_u = \frac{y_u}{C^{|u|/2}}$.

V. EXPLOITING POLYNOMIAL CONSTRAINTS

It is a basic result in operator algebras that MPIs for $D \times D$ matrices must have degree at least $2D$ [20]. This implies that we would need to implement the Dth relaxation of (6) or (7) in order to obtain nontrivial D-dimensional constraints. Even for problems involving a small number of noncommuting variables, this becomes impractical already for $D = 5$. Hence, if we wish to conduct optimizations over matrices of dimensions greater than 2 or 3, we must rely on linear restrictions other than those derived from MPIs.

Most NPO problems relevant in quantum information science involve polynomial *identities* rather than polynomial *inequalities*. That is, constraints of the form $q(X) \geqslant 0$ are complemented with $-q(X) \geqslant 0$, and so $q(X) = 0$ must hold for all representations of $\{q_i(X) \geqslant 0\}$. The strategy we follow to solve these kinds of problems is to divide the representations of $X_1, \ldots, X_n$ into different classes r in such a way that any two representations belonging to the same class r can be connected by a continuous trajectory of feasible representations in r. As we will see, each of these classes will satisfy nontrivial low degree polynomial identities, which we translate into linear constraints at the level of moment matrices (vectors). By carrying out a relaxation of the form (6) for each possible class r and taking the greatest result, we hence obtain an upper bound on the solution of the general problem (1).

For instance, suppose that $\{q_i(X) \geqslant 0\}$ contains relations of the form $X_i^2 = 1$ for $i = 1,2,3$. For $D = 2$, there are two possibilities:

(1) $X_i = \pm\mathbb{I}$ for some $i \in \{1,2,3\}$;

(2) $X_i \neq \pm\mathbb{I}$ for all i, in which case it can be shown that the operators satisfy the identities

$$[X_1,\{X_2,X_3\}_+] = [X_2,\{X_1,X_3\}_+] = [X_3,\{X_1,X_2\}_+] = 0. \tag{16}$$

In either case, the noncommuting variables satisfy nontrivial polynomial constraints of degree smaller than 4, the smallest possible degree of an MPI for $D = 2$. A way to attack this problem is therefore to define an SDP relaxation of the form (7) for each case, enforcing the corresponding extra linear constraints on the moment matrix (or moment vector).

Note also that, if we further assume that the matrices $X_1, \ldots, X_n$ are real, then we can add constraints of the sort

$$\{X_1,[X_2,X_3]\}_+ = 0 \tag{17}$$

in the second case. This approach hence allows us (in principle) to distinguish between real and complex matrix algebras.

The objective, again, is to identify all possible linear restrictions on $\hat{M}_k$ or y within a given class r. Fortunately, most NPO problems in quantum information science have the peculiarity that random representations of a given class can be generated efficiently.

Continuing with the previous example, suppose that we wish to optimize over six dichotomic operators; i.e., the polynomial constraints are, precisely, $X_i^2 = \mathbb{I}$ for $i = 1,2,\ldots,6$. For simplicity, let us denote the first four operators as $X_{00}, X_{01}, X_{10}, X_{11}$ and the last two as Y_1, Y_2. We want to maximize the average value of the operator

$$p(X) \equiv \sum_{j=1,2} \sum_{c_1,c_2=0,1} (-1)^{c_j} X_{c_1c_2} Y_j X_{c_1,c_2}. \tag{18}$$

Note that we can write each dichotomic operator as $X_i = (-1)^a \frac{2E_a^i - \mathbb{I}}{2}$, where $\{E_a^i\}$ are projection operators satisfying $E_0^i + E_1^i = \mathbb{I}$. Substituting in (18), we have that the objective function $\langle\psi|p(X)|\psi\rangle$ is equal to

$$4 \sum_{c_1,c_2,s=0,1} P\big(s,c_1\big|X_{c_1c_2},Y_1\big) + P\big(s,c_2\big|X_{c_1c_2},Y_2\big) - 16, \tag{19}$$

where $P(a_1,a_2|x_1,x_2) = \langle\psi|E_{a_1}^{x_1} E_{a_2}^{x_2} E_{a_1}^{x_1}|\psi\rangle$. This corresponds to the *temporal correlations scenario* defined in [19], where sequential dichotomic projective measurements are conducted over a quantum system and a record of the measurements $x_1,x_2,\ldots$ implemented, as well as the measurement outcomes $a_1,a_2,\ldots$, is kept. The goal is to limit the statistics $P(a_1,\ldots,a_n|x_1,\ldots,x_n)$ obtained after several repetitions of the experiment.

As noted in [14,19], when the dimensionality of the quantum system is unrestricted, the set of all feasible distributions $P(a_1,\ldots,a_n|x_1,\ldots,x_n)$, and hence the optimal value of (18), can be characterized by a single SDP. Using the SDP solver MOSEK [28], we find that, for $D = \infty$, $p^\star = 8$ up to seven decimal places.

Suppose, however, that we have the promise that the system has dimension $D = 2$. The problem we want to solve is therefore

$$p^\star = \max_{\mathcal{H},X,\psi} \langle\psi|p(X)|\psi\rangle$$

such that

$$\dim(\mathcal{H}) \leqslant 2, \quad \mathbb{I} - X_i^2 = 0, \quad \text{for} \quad i = 1,\ldots,6. \tag{20}$$

We start by dividing the representations of two-dimensional dichotomic operators into classes. For any dichotomic operator, the rank of the projector $E \equiv \frac{X+\mathbb{I}}{2}$ can be 0, 1, or 2. For $r = 0,2$, the corresponding operator is $X = -\mathbb{I}$ or $X = \mathbb{I}$, respectively. For $r = 1$, a random dichotomic operator X can be generated as $X = 2\frac{|v\rangle\langle v|}{\langle v|v\rangle} - \mathbb{I}$, where $v \in \mathbb{C}^2$ is a random complex vector. Since we are dealing with six noncommuting variables, there are $3^6 = 729$ classes, labeled by the vector $\vec{r} \in \{0,1,2\}^6$, with $\text{rank}(X_i + \mathbb{I}) = r_i$.

For a fixed value of $\vec{r}$, we sequentially generate random 6-tuples of dichotomic operators $X^j \equiv (X_1^j,\ldots,X_6^j)$, with the required rank constraints, as well as a sequence of random normalized vectors $|\psi^j\rangle \in \mathbb{C}^2$. As before, we use each pair $(X^j,|\psi^j\rangle)$ to generate a random feasible moment matrix M_k^j. Note that, since the conditions $X_i^2 = \mathbb{I}$ are implicit in each moment matrix, it is not necessary to include localizing matrices in our description (they would amount to zero diagonal blocks in the extended moment matrix). Notice as well that, given a feasible moment matrix M_k, its complex conjugate M_k^* is also feasible. Since $p(X)$ is a real linear combination of Hermitian monomials, the objective function will have the same value for both M_k and M_k^* [and thus for the real feasible moment matrix $\text{Re}(M_k) = \frac{M_k}{2} + \frac{M_k^*}{2}$]. This implies that, in order to define an SDP relaxation for (20), it suffices to consider the sequence of real matrices $\text{Re}(M_k^1), \text{Re}(M_k^2), \ldots$.

Applying the modified Gram-Schmidt method to that sequence until we find linear dependence, we obtain an orthonormal basis for $\mathcal{M}_{D,\vec{r}}^k$, the space of all real feasible moment matrices for representations of the class $\vec{r}$. This time, the fact that this randomizing method works with probability one is a consequence that the projection of the randomly generated moment matrix $\mathrm{Re}(M_k^{j+1})$ onto the orthogonal complement of the space spanned by $\mathrm{Re}(M_k^1),\ldots,\mathrm{Re}(M_k^j)$ is a matrix whose entries are rational functions of the randomly generated vectors used to build X^{j+1} and $|\psi^{j+1}\rangle$. If $\mathrm{Re}(M_k^1),\ldots,\mathrm{Re}(M_k^j)$ do not span $\mathcal{M}^k_{D,\vec{r}}$, then there exists a choice for those vectors such that the projected matrix is nonzero; i.e., at least one of such rational functions is nonzero. It is a well-known fact that the probability that a randomly evaluated nonzero rational function vanishes is zero.

Alternatively, we can identify $\mathcal{S}^k_{D,\vec{r}}$, the space of feasible $2k$th-order moment vectors for representations in the class $\vec{r}$ by parametrizing each normalized vector needed to build X or ψ by two angles ϕ,φ, and constructing the corresponding moment vector $y(\vec{\phi},\vec{\varphi})$. Then the entries of the matrix

$$S \equiv \int d\vec{\phi}d\vec{\varphi}\mathrm{Re}(y(\vec{\phi},\vec{\varphi}))\mathrm{Re}(y(\vec{\phi},\vec{\varphi})^\dagger) \tag{21}$$

can be computed analytically. Its support will coincide with $\mathcal{S}^k_{D,\vec{r}}$.

One way or another, we must solve the program

$$p^k = \max_{\hat{M}} \sum_w p_w \hat{M}_{w,\mathbb{I}},$$

$$\text{such that}$$

$$\hat{M} \in \mathcal{M}^k_{D\vec{r}}, \quad \hat{M}_{\mathbb{I},\mathbb{I}} = 1, \quad \hat{M} \geqslant 0, \tag{22}$$

for all possible classes $\vec{r}$. For $k=2$, again using MOSEK [28], we find $p^2 \approx 5.656\,854$, definitely smaller than the free limit.

VI. SIMILARLY INSPIRED SDP HIERARCHIES

In the following we introduce two problems in quantum information science which, while not exactly fitting in the class of problems (1), can be similarly reduced to SDP hierarchies.

A. Quantum nonlocality under dimension constraints

The scenario is as follows. Two distant parties, call them Alice and Bob, conduct measurements on a bipartite quantum system. We denote Alice's (Bob's) measurement setting by x (y) and her (his) measurement outcome by a (b). We wish to bound a linear functional of the statistics $P(a,b|x,y)$ they will observe, under the assumption that Alice's and Bob's spaces are, at most, D dimensional. If Alice and Bob's outcomes are binary, i.e., $a,b \in \{0,1\}$, the problem can be shown to be equivalent to

$$\max \sum_{x,y,a,b} B^{x,y}_{a,b} P(a,b|x,y),$$

$$\text{such that} \tag{23}$$

$$P(a,b|x,y) = \langle\psi|E^x_a \otimes F^y_b|\psi\rangle,$$

where $\{E^x_a,F^y_b\}$ are projection operators acting on $\mathbb{C}^D$, with $\sum_a E^x_a = \sum_b F^y_b = \mathbb{I}_D$ and $|\psi\rangle \in \mathbb{C}^D \otimes \mathbb{C}^D$.

Following the last section, we divide the representations of the operators E^x_a,F^y_b into different classes labeled by the vectors $\vec{r},\vec{t}$, with $\mathrm{rank}(E^x_a) = r^x_a$, $\mathrm{rank}(F^y_b) = t^y_b$. For each representation class $\vec{r},\vec{t}$, we try to characterize the span $\mathcal{M}^k_{D,\vec{r}\vec{t}}$ of feasible kth-order moment matrices. To do so, we sequentially generate random normalized states $|\psi^j\rangle \in \mathbb{C}^D \otimes \mathbb{C}^D$ and projectors $E^{x,j}_a, F^{y,j}_b \in B(\mathbb{C}^D)$, satisfying the rank conditions $\mathrm{rank}(E^{x,j}_a) = r^x_a$, $\mathrm{rank}(F^{y,j}_b) = t^y_b$. Given $E^{x,j}_a,F^{y,j}_b$, we define the projectors $\bar{E}^{x,j}_a \equiv E^{x,j}_a \otimes \mathbb{I}_D$ and $\bar{F}^{y,j}_b \equiv \mathbb{I}_D \otimes F^{y,j}_b$, which we use to generate a feasible kth-order moment matrix M^j_k. By subjecting the resulting sequence of moment matrices to the modified Gram-Schmidt orthogonalization, we obtain a basis for $\mathcal{M}^k_{D,\vec{r}\vec{t}}$. The SDP to solve is hence

$$B^k \equiv \max \sum_{x,y,a,b} B^{x,y}_{a,b} (M_k)_{\bar{E}^x_a,\bar{F}^y_b}$$

$$\text{such that} \tag{24}$$

$$(M_k)_{\mathbb{I},\mathbb{I}} = 1, \quad M_k \geqslant 0, \quad M_k \in \mathcal{T}^k_{D,\vec{r},\vec{t}}.$$

We again advise the reader to check that $T \equiv \frac{1}{N}\sum_{j=1}^N M^j_k$ is positive definite: Otherwise, a projection of M_k onto the support of T is necessary to guarantee the strict feasibility of the associated SDP; see Remark 1.

The completeness of the above SDP hierarchy can be established easily: Following the same lines as in Sec. IV, we prove the existence of a (in general, infinite-dimensional) representation $\tilde{E}^x_a,\tilde{F}^y_b \subset B(\mathcal{H})$, with $[\tilde{E}^x_a,\tilde{F}^y_b] = 0$ for all x,y,a,b, and a state $\tilde{\psi}$, such that $\sum_{x,y,a,b} B^{x,y}_{a,b}\langle\tilde{\psi}|\tilde{E}^x_a\tilde{F}^y_b|\tilde{\psi}\rangle$ coincides with the asymptotic limit $\hat{B} \equiv \lim_{k\to\infty} B^k$. The center of the algebra $\mathcal{A}$ generated by $\{\tilde{E}^x_a : x,a\}$ decomposes $\mathcal{H}$ into a direct integral of sectors $\mathcal{H}_z$. By construction, in each sector z, $\mathcal{A}$ boils down to a type I factor $\mathcal{A}_z$ of dimension smaller than or equal to D. Being a type I factor, we can write $\mathcal{H}_z = \mathcal{H}^A_z \otimes \mathcal{H}^B_z$; then $\mathcal{A}_z \sim B(\mathcal{H}^A_z) \otimes \mathbb{I}$ and $\mathcal{A}'_z \sim \mathbb{I} \otimes B(\mathcal{H}^B_z)$, where $\mathcal{A}'_z$ denotes the commutant of $\mathcal{A}_z$. It follows that

$$\tilde{E}^x_a = \int^\oplus d\mu(z)\tilde{E}^x_{a,z} \otimes \mathbb{I}_{B,z},$$

$$\tilde{F}^y_b = \int^\oplus d\mu(z)\mathbb{I}_{A,z} \otimes \tilde{F}^y_{b,z}, \tag{25}$$

where $\dim(\mathcal{H}_{A,Y}) \leqslant D$. Likewise, it can also be shown that the algebra generated by $\tilde{F}^y_{b,z}$ decomposes as a direct integral of finite-dimensional algebras with dimension smaller than or equal to D. $\hat{B}$ is thus a convex combination of feasible points and, as such, it represents a lower bound for the original problem (23).

1. *Examples*

Here we give examples of maximizing the violation of bipartite Bell inequalities with binary outcomes for different dimensionality of the component spaces. Some of the examples have already appeared in Ref. [16]. First we discuss the I_{3322} inequality, the only tight three-setting, two-outcome Bell inequality and its modified version. Then we move on to one more setting per party. For all the subsequent computations we used the solvers MOSEK [28] and SEDUMI [29] through the interface YALMIP [30], which we ran on a memory-enhanced desktop PC (with 128 GB RAM).

a. I3322. First we considered the I_{3322} inequality [31], which is the member of the I_{NN22} family $N \geqslant 2$. Recently, it has been proven that qubit systems are not enough to attain the overall quantum maximum 0.2509. Rather, the best value in $\mathbb{C}^2 \times \mathbb{C}^2$ systems is 0.25 [32,33]. Using SDP, we reproduced the maximum value of 0.25 in dimensions $\mathbb{C}^3 \times \mathbb{C}^3$ as well up to eight significant digits [16]. The size of the moment matrix was 76, involving 1240 linear constraints. The computations took about 5 min for a fixed rank combination of measurements. Note that the hierarchy of Moroder *et al.* [33], by limiting the negativity [34] of the bipartite quantum state, also gives a (not necessarily tight) upper bound on the Bell violation for a fixed dimension of the quantum state. Indeed, this method works for $\mathbb{C}^2 \times \mathbb{C}^2$ systems by returning a violation of 0.25 of the I_{3322} inequality [33]. However, for $\mathbb{C}^3 \times \mathbb{C}^3$ systems it does not seem to converge (see Fig. 1 of [33]).

The SDP method also allows the user to upper bound the maximum quantum violation of a Bell inequality using a fixed two-qudit state. Let us choose the three-dimensional maximally entangled state, $|\psi\rangle = (|00\rangle + |11\rangle + |22\rangle)/\sqrt{3}$. We find the value of 0.229 771, which is saturated by see-saw computation; hence, the presented upper bound is tight. The computation involved a 116-dimensional moment matrix (on a partial four-level relaxation) along with 1060 linear constraints. The program took 2 min to complete for a fixed rank combination of projective measurements. We mention a related problem, where the maximal violation of I_{3322} has been computed for the maximally entangled state (of unrestricted dimensionality). This has been solved both analytically [35] and using a relaxation method [36] by returning the value of 0.25.

b. Modified I3322. Though I_{3322} inequality likely requires infinite dimensions to achieve the maximal violation 0.2509 [13], $D = 12$ seems to be the smallest local dimension surpassing the qubit bound 0.25 [37]. It is an open question whether there exists a Bell inequality for which the maximum quantum violation in a given dimension is a strictly monotonic function of the dimension. Below we give such a candidate. To this end, we modify the I_{3322} inequality by introducing a parameter $c \geqslant 1$,

$$\begin{aligned} I_{3322}(c) = {} & E_1^A + E_2^A + E_1^B + E_2^B \\ & - (E_{1,1} + E_{1,2} + E_{2,1} + E_{2,2}) \\ & + c(E_{1,3} + E_{3,1} - E_{2,3} - E_{3,2}) \leqslant 4c, \end{aligned} \tag{26}$$

where the correlator $E_{x,y}$ between measurement x by Alice and measurement y by Bob is defined as $E_{x,y} = P(a = b|x,y) - P(a \neq b|x,y)$, $a,b \in \{0,1\}$, and E_x^A denotes the marginal of Alice's measurement setting x (and E_y^B is similarly defined for Bob). This inequality is symmetric for exchange of Alice and Bob and returns the original I_{3322} inequality (written in terms of correlators) for parameter $c = 1$.

Setting $c = 2$, we used the see-saw variational technique [38,39] to find a lower bound on the maximal violation for any dimension $2 \leqslant d \leqslant 15$, which we observe to be gradually increasing with dimension. We conjecture that the bounds are tight. Table I shows results up to $D = 6$ concerning both the lower (see-saw) and the upper bounds (SDP). Accordingly, the bounds for $D = 2,3$ are indeed tight. Computationally, the most challenging case was obtaining the upper bound in $D = 4$. It involved 3514 constraints; the dimension of the moment matrix is 184 and it took roughly 40 min for MOSEK to complete the task for a given rank combination of the measurements. The quantum maximum in dimensions 5 and 6 coincide with the NPA bound on level 3 up to the shown digits. We pose it as a challenge to prove tightness of the see-saw bound for $D = 4$ (or possibly higher dimensions) by exploiting the symmetric structure of the inequality (26) using techniques such as in Refs. [33,40].

TABLE I. Quantum bounds for different local dimensions on the violation of the $I_{3322}(2)$ inequality computed using see-saw search/SDP computation. Bounds for $D = 2,3$ are tight since the see-saw and SDP bounds match. As an overall upper bound, the Navascués-Pironio-Acín (NPA) [12,13] hierarchy on level 3 gives 8.075 937.

D	Lower bound	Upper bound
2	8.013 177	8.013 177
3	8.024 050	8.024 050
4	8.032 766	8.071 722
5	8.039 579	8.075 937
6	8.056 714	8.075 937

c. I4422 family. A one-parameter family of four-setting inequalities is given in Ref. [32]. These inequalities are not tight but they have a quite simple structure. They look as follows for $c \geqslant 0$:

$$\begin{aligned} I_{4422}(c) = {} & cE_1^A + (E_{1,1} + E_{1,2} + E_{2,1} - E_{2,2}) \\ & + (E_{3,3} + E_{3,4} + E_{4,3} - E_{4,4}) \leqslant 4 + c. \end{aligned} \tag{27}$$

When $c = 0$, it is a direct sum of two Clauser-Horne-Shimony-Holt (CHSH) [9] inequalities; hence, maximum violation is attained with qubit systems. However, by setting $c > 0$, it may serve as a dimension witness. In particular, for $c = 1$ (the value used in Eq. (19) of [32]), its maximal violation in $\mathbb{C}^2 \times \mathbb{C}^2$ systems is upper bounded by the value of 5.8515 [32]. However, using our SDP tool, this upper bound turns out to be not tight: We certify a smaller value of 5.8310, which is matched by the see-saw method. Further, by raising the dimension to $\mathbb{C}^3 \times \mathbb{C}^3$, we get the same amount of violation. The SDP computation returning 5.8310 in $\mathbb{C}^3 \times \mathbb{C}^3$ was quite demanding: It required a 130-dimensional moment matrix and took about 2 h of computational time. The value 5.8310 must be compared to the maximum value of $2\sqrt{2} + \sqrt{10} \approx 5.9907$, achievable in $\mathbb{C}^4 \times \mathbb{C}^4$ systems. In contrast to our certified value 5.8310, the corresponding $\mathbb{C}^3 \times \mathbb{C}^3$ value arising from Moroder *et al.* hierarchy [33] (on their level 2) is a higher value of 5.9045.

d. I4722 inequality. It is also worth mentioning a situation (actually, this is the only case we are aware of) for which a previous SDP method introduced in Ref. [32] outperforms our present SDP method. We tested the method in case of asymmetric Bell inequalities, that is, when the number of settings on the two sides are not the same. For the sake of comparison, we have chosen a correlation-type Bell inequality from [41], already analyzed in [32],

$$\begin{aligned} I_{4722} = {} & E_{11} + E_{21} + E_{31} + E_{41} \\ & +(E_{12} - E_{22}) + (E_{31} - E_{33}) + (E_{41} - E_{44}) \\ & +(E_{25} - E_{35}) + (E_{26} - E_{46}) + (E_{37} - E_{47}) \\ & \leqslant 8, \end{aligned} \tag{28}$$

which consists of four and seven binary-outcome settings on Alice and Bob's respective sides. In Ref. [32] a method is presented in Sec. III B, which is particularly suited to asymmetric Bell setups. This way, the best upper bound obtained for qubit systems is 10.5102, whereas the best lower bound value of 10.4995 is due to see-saw search. The quantum maximum, attainable with two ququarts, is 10.5830. Using our present SDP technique and a desktop PC, unfortunately, we did not manage to go below the global maximum 10.5830.

Suppose now that Alice and Bob are conducting nonbinary measurements, that is, measurements with more than just two outcomes. Then, in order to consider the most general measurements they could perform, we must model their measurement devices via POVMs, rather than projective measurements. There are two ways to accomplish this.

(1) We can replace constraints of the sort $(E_a^x)^2 = E_a^x$ in (23) with the positive semidefinite constraints $E_a^x \geqslant 0$ at the cost of having to add the corresponding localizing matrices to (24). Although converging, this method does not seem to behave well in our numerical experiments.

(2) Alternatively, we can exploit the fact that any d-outcome POVM $\{E_a \geqslant 0\} \subset B(\mathbb{C}^D)$ can be realized in an extended Hilbert space $\mathbb{C}^d \otimes \mathbb{C}^D$ via a projective measurement of the form $M_a = U(|a\rangle\langle a| \otimes \mathbb{I}_D)U^\dagger$, where $U \in B(\mathbb{C}^d \otimes \mathbb{C}^D)$ is a unitary matrix [42]. Indeed, taking the state to be $\rho = |0\rangle\langle 0| \otimes |\psi\rangle\langle\psi|$ and choosing U appropriately, it can be verified that

$$\mathrm{tr}(\rho M_a) = \mathrm{tr}(E_a|\psi\rangle\langle\psi|), \tag{29}$$

for $a = 0,\dots,d-1$ and all states $|\psi\rangle$.

In the hierarchy to implement, random states of the form $|0\rangle\langle 0|_{A'} \otimes |\psi\rangle\langle\psi|_{AB} \otimes |0\rangle\langle 0|_{B'}$ are generated. For each random state, we construct a moment matrix containing the operators $\bar{E}_a^x = U^x(|a\rangle\langle a| \otimes \mathbb{I}_D)(U^x)^\dagger \otimes \mathbb{I}_D \otimes \mathbb{I}_d$, $\bar{F}_b^y = \mathbb{I}_d \otimes \mathbb{I}_D \otimes V^y(\mathbb{I}_D \otimes |b\rangle\langle b|)(V^y)^\dagger$ and the projectors $P_A = |0\rangle\langle 0| \otimes \mathbb{I}_D^{\otimes 2} \otimes \mathbb{I}_d$, $P_B = \mathbb{I}_d \otimes \mathbb{I}_D^{\otimes 2} \otimes |0\rangle\langle 0|$.

The convergence of this hierarchy follows from the fact that the algebras generated by $\{P_A\bar{E}_a^x P_A\}$, $\{P_B\bar{F}_b^y P_B\}$ cannot violate D-dimensional MPIs.

2. *Examples*

We now apply our method to place nontrivial upper bounds on the quantum violation of Bell inequalities using genuine POVM measurements for some of the settings. Note that for binary-outcome settings general POVM measurements are not relevant; hence, we have to consider Bell inequalities with at least one nonbinary setting. To this end, we consider the simplest tight Bell inequality due to Pironio beyond genuine two-outcome inequalities [4,43]. In this inequality, Alice has three binary-outcome measurements, and Bob has two settings: The first one has binary outcomes and the second one has ternary outcomes. If we allow Bob to use general POVM measurements on his second setting, the two-qubit quantum maximum $(\sqrt{2}-1)/2 \approx 0.2071$ is recovered up to computer precision on level 3 of the SDP hierarchy. Hence, in this particular Bell inequality the use of general measurements does not provide any advantage over projective ones. Let us note that the quantum maximum without dimension constraints is a larger value, 0.2532, which can be obtained using a two-qutrit system and projective measurements [4]. We also applied the above method to the Collins-Gisin-Linden-Massar-Popescu (CGLMP) inequality [44] in order to prove the conjecture that the qubit bound 0.2071 using projective measurements is optimal (i.e., general POVM measurements do not help to improve the bound). However, in that case, we were unable to go below the known overall quantum maximum given in Refs. [13,45].

The previous approach can be easily extended to characterize the statistics of multipartite scenarios where the local dimensionality of all parties is bounded from above. More interestingly, it can also be adapted to deal with multipartite Bell scenarios where only a subset of the parties has limited dimensionality.

Consider, for instance, a tripartite scenario where Alice and Bob's measurement devices are unconstrained, but the dimensionality of the third system (say, Charlie's) is bounded by D. We want to generate a basis for the corresponding space of truncated moment matrices, with rows and columns labeled by strings of operators of the form $u(AB)v(C)$, where $u(AB)$ $[v(C)]$ denotes a string of Alice and Bob's (Charlie's) operators of length at most k_{AB} (k_C).

The key is to realize that, in a multipartite (complex) Hilbert space, the space of feasible moment matrices is spanned by moment matrices corresponding to separable states. Hence, in order to attack this problem, we start by generating a sequence of *complex* D-dimensional moment matrices for Charlie's system alone. After applying Gram-Schmidt to these complex matrices, we obtain the basis of Hermitian matrices $\{M_j\}_{j=1}^N$. Next we generate a basis for Alice and Bob's moment matrices. Since their dimension is unconstrained, such matrices are expressed as

$$\Gamma_{k_{AB}} = \sum_{|u|\leqslant 2k_{AB}} c_u N_u + c_u^* N_{u^\dagger}, \tag{30}$$

where N_u is a matrix defined by

$$(N_u)_{v,w} = 1, \quad \text{if } v^\dagger w = u; \\ 0, \quad \text{otherwise.} \tag{31}$$

The overall moment matrix for the whole system can then be expressed as $M = \sum_{u,j} M_j \otimes (c_{u,j}N_u + c_{u,j}^* N_{u^\dagger})$.

Since we are just interested in optimizing a real linear combination of real entries of M—corresponding to the measured probabilities $P(a,b,c|x,y,z)$—we can take the real part of the above matrix, and so we end up with the relaxation

$$\max \sum_{x,y,z,a,b,c} B_{a,b,c}^{x,y,z} M_{E_a^x,F_b^y G_c^z},$$
$$\text{such that}$$
$$M_{\mathbb{I},\mathbb{I}} = 1, \quad M \geqslant 0, \tag{32}$$
$$M = \sum_{u,j} c_{u,j}^{\mathcal{R}} \mathrm{Re}(M_j) \otimes (N_u + N_{u^\dagger}) \\ - c_{u,j}^{\mathbb{I}} \mathrm{Im}(M_j) \otimes (N_u - N_{u^\dagger}),$$

where $c_{u,j}^{\mathcal{R}}$ $(c_{u,j}^{\mathbb{I}})$ denotes the real (imaginary) part of $c_{u,j}$. This is an SDP with real variables.

3. Examples

We now show applications of the above SDP method tailored to multipartite systems. As a first example, a three-party system is considered for which Alice possesses a qubit and the other two parties (Bob and Charlie) have no restriction on the dimensionality of the Hilbert spaces. We are able to fully reproduce the bounds obtained in Ref. [32]. In the next example, we extend Alice's Hilbert space to a qutrit, thereby certifying genuine four-dimensional entanglement. Then we move to a four-party (translationally invariant) Bell scenario and certify that a Bell value above a certain threshold cannot be obtained with symmetric measurements (that is, when each four parties measure the same observables in the first and second respective settings).

a. I333 inequality. We consider the following three-party three-setting permutationally invariant Bell inequality [32]

$$\begin{aligned} I_{333} = \mathrm{sym}\{&-P(A_1) - 2P(A_3) + P(A_1,B_1) \\ &- P(A_1,B_2) + P(A_1,B_3) - 2P(A_2,B_2) \\ &+ 2P(A_2,B_3) - 2P(A_3,B_3)\} \leqslant 0. \end{aligned} \tag{33}$$

Here we used the short-hand notation $P(A_x,B_y) = p(0,0|x,y)$, $P(A_x) = p(0|x)$ and similarly for the other parties. Notice that the Bell expression above consists of only two-body correlators and single-party marginal terms, which usually provide an advantage in experiments. Such Bell inequalities have been proposed in Ref. [46] to detect nonlocality in multipartite quantum systems for any number of parties (however, those inequalities involve only two settings per party; hence, they can be maximally violated with qubit systems, unlike the present example). In Eq. (33), $\mathrm{sym}\{X\}$ means that every term occurring in X should be symmetrized with respect to all possible permutations of the parties, e.g., $\mathrm{sym}\{P(A_1,B_1)\} = P(A_1,B_1) + P(B_1,C_1) + P(A_1,C_1)$.

We next compute upper bounds on the quantum violations assuming different dimensionality of the Hilbert spaces. Lower bound values, on the other hand, are obtained from see-saw iteration in a prior work [32]. Table II summarizes the results. Values with an asterisk (*) have been established in the present work. Notation $(D_1 D_2 D_3)$ refers to the dimensionalities of Alice, Bob, and Charlie's Hilbert spaces, respectively. Notice that due to symmetry of the Bell inequality (33), the same bounds apply to any permutations of $(D_1 D_2 D_3)$. Establishing upper bound on case (222) with nondegenerate measurements was the most time consuming task, the corresponding SDP problem involved 4894 constraints and took 3 h to be solved using MOSEK; still the lower bound value has not been saturated. Computing the upper bound for the case $(2\infty\infty)$ required to run the hybrid method (Alice was given level 2 of the qubit hierarchy, whereas Bob and Charlie's system was computed on NPA level $1 + AB$). In that case, we managed to close the gap between the lower and the upper bound values, thereby reproducing the result of Ref. [32]. We also computed $(3\infty\infty)$ upper bound and recovered the global maximum of 0.196 285 2 certified by the NPA hierarchy (Alice was given the level 3 of the qutrit hierarchy and took 24 h for MOSEK to solve the resulting SDP). Accordingly, any Bell violation of I_{333} bigger than 0.178 689 7 cannot be attained with dimensionalities $(2\infty\infty)$ (plus the two other permutations), implying that the underlying three-party state ρ_{ABC} has at least Schmidt number vector (3,3,3) (see, e.g., Refs. [47,48]). Moreover, any pure state decomposition of ρ_{ABC} contains at least one state $\sigma_{ABC} = |\psi\rangle\langle\psi|$ such that the rank of each single-party marginal σ_A, σ_B, and σ_C is greater than 2. In short, a Bell violation of I_{333} bigger than 0.178 689 7 detects in a device-independent way that the three-party state is genuinely three-dimensional entangled.

b. I444 inequality. We construct a three-party Bell inequality which cannot be violated maximally in state spaces $\mathbb{C}^3 \times \mathbb{C}^D \times \mathbb{C}^D$ (and arbitrary permutations thereof) for any dimension D. This extends the previous example to the case when Alice's state space is restricted to a qutrit (instead of a qubit). In particular, the maximal violation is attained in $\mathbb{C}^4 \times \mathbb{C}^4 \times \mathbb{C}^4$. Hence, this certifies that the underlying three-party quantum state is genuinely four-dimensional entangled.

Let us consider the following three-party, four-setting Bell inequality [49],

$$I_{444} = \mathrm{CHSH}_{AB} + \mathrm{CHSH}_{A'C} + \mathrm{CHSH}_{B'C'} \leqslant 6, \tag{34}$$

where A and A' denote different sets of measurements for party A, and we use similar notation for parties B and C. In Ref. [49] it has been proved that the maximum quantum violation attainable with biseparable states is $S = 4 + 2\sqrt{2} \approx 6.8284$. Hence, if the above bound is exceeded in a Bell experiment, we can conclude that the state is genuinely tripartite entangled [50,51]. The same bound can be derived by using the SDP techniques of Ref. [52] based on the NPA hierarchy. Below we extend this result to the realm of genuine higher-dimensional entanglement.

To this end, we replace $\mathrm{CHSH}_{B'C'}$ with the Tsirelson bound $2\sqrt{2}$ [53]. This places an upper bound on I_{444} in (34).

TABLE II. Lower bounds (LB) and upper bounds (UB) on the violation of the I_{333} inequality in various local dimensions. Values with an asterisk (*) have been established in the present work. The notation $(D_1 D_2 D_3)$ refers to the dimensionalities of Alice, Bob, and Charlie's Hilbert spaces, respectively. The sign ∞ denotes no restriction on dimension of the respective party. Abbreviation Deg/No-deg refers to the situation when Alice has at least one degenerate measurement/all measurements are nondegenerate (i.e., rank 1 projectors). The qutrit value (333) is the overall quantum maximum certified by the NPA hierarchy [12]. The upper bound value for $(2\infty\infty)$ in the degenerate case was obtained using the NPA hierarchy as well.

	LB (222)	UB (222)	LB $(2\infty\infty)$	UB $(2\infty\infty)$	LB (333)
No-deg	0.044 348 4	0.054 136 2*	0.178 394 6	0.178 394 6*	0.196 285 2
Deg	0.178 394 6	0.178 394 6*	0.178 689 7	0.178 689 7	

TABLE III. Maximum quantum bounds on Bell inequality I_{444} in (34) for different local dimensions of Alice. The first column (labeled by Bisep) stands for the case when Alice has a classical system and the other two parties have unrestricted dimensionalities. All bounds are tight as they are matched with lower bounds arising from see-saw iteration.

Bisep	$(2\infty\infty)$	$(3\infty\infty)$	(444)
6.828 427	7.656 854	7.971 284	8.485 281

Therefore, we are left with optimizing $\text{CHSH}_{AB} + \text{CHSH}_{A'C} + 2\sqrt{2}$ for $\mathbb{C}^3 \times \mathbb{C}^D \times \mathbb{C}^D$ systems, where D denotes arbitrary dimension. To do so, we classify Alice's four observables according to their traces ($\pm 1, \pm 3$) and in each case we can solve the problem with SDP for the hybrid multipartite case. Notice, however, that Bob and Charlie have only two binary-outcome measurements; hence, Jordan's Lemma applies and we can assume that Bob and Charlie have traceless qubit observables [54]. Then the problem goes back to upper bounding $\text{CHSH}_{AB} + \text{CHSH}_{A'C} + 2\sqrt{2}$ in $\mathbb{C}^3 \times \mathbb{C}^2 \times \mathbb{C}^2$, which can be straightforwardly done using our SDP tools. By running the SDP, the maximum turns out to be $36/7 + 2\sqrt{2}$ up to the numerical precision of the solver MOSEK. We also solved the problem assuming that Alice has a qubit yielding the upper bound $2 + 4\sqrt{2}$ up to computer precision. Results are summarized in Table III. All bounds are tight as they are saturated using see-saw search.

Consider now the so-called fully connected Bell state, that is, a three-party state for which any two parties share a two-qubit Bell pair,

$$|\psi_{444}\rangle = |\varphi^+\rangle_{AB} \otimes |\varphi^+\rangle_{A'C} \otimes |\varphi^+\rangle_{B'C'}. \tag{35}$$

With this particular $\mathbb{C}^4 \times \mathbb{C}^4 \times \mathbb{C}^4$ state and measurement settings optimal for CHSH violation, we get the overall quantum maximum of $6\sqrt{2} \approx 8.485\,281$ for the Bell inequality (34). By adding a certain amount of white noise to the state (35):

$$\rho_{\text{noisy}} = p|\psi_{444}\rangle\langle\psi_{444}| + \frac{1-p}{4^3}\mathbb{I}_{4\times4\times4}, \tag{36}$$

we get the critical visibility $p_{\text{crit}} = (36/7 + 2\sqrt{2})/(6\sqrt{2}) \approx 0.939\,425$, above which we can detect the state (36) to be genuinely four-dimensional entangled. We believe this threshold is low enough to be interesting from an experimental point of view as well.

c. I2222 with symmetric measurements. In Ref. [55], a search has been conducted for all three- and four-partite binary-outcome Bell inequalities involving two-body correlators that obey translationally symmetry. Any translationally invariant Bell inequality is provably maximally violated by a translationally invariant state when all parties measure the same set of observables (of unlimited dimensionality). Numerical investigations in Ref. [55] suggest that it is not true anymore if we restrict the local Hilbert space dimension of the parties. Let us pick #64 inequality from Table II in Ref. [55]. Due to the fact that these Bell inequalities involve two dichotomic measurements per site, Jordan's lemma applies and the maximum violation is given by $\beta_Q = 6 + 2\sqrt{2}$ in qubit systems. Due to numerics, this value is achieved with different pairs of qubit observables. Indeed, running our SDP program by building up the bases from random symmetric measurements, we certify that applying the same settings at all sites does not allow us to violate the Bell inequality #64 in Table II of Ref. [55] (i.e., $\beta_Q^{TI} = \beta_c$ in the notation of the corresponding reference).

B. One-way quantum communication complexity

Consider the following communication scenario. Alice and Bob are given inputs x, y with probability $p(x,y)$ and have the task to compute the Boolean function $f(x,y)$. To do so, we allow Alice to transmit a D-dimensional quantum system to Bob, who, upon receiving it, must make a guess b on $f(x,y)$. We wish to find the strategy which will allow Alice and Bob to maximize the probability that Bob's guess is correct, i.e., $b = f(x,y)$. For example, in a QRAC [56], the inputs $\vec{x}, y$ can take values in $\{0,1\}^k$ and $\{1,\dots,k\}$, respectively, and the function to compute is $f(\vec{x},y) = x_y$.

This scenario can be modeled by assuming that Alice prepares a pure quantum state $\rho_x \equiv |\psi_x\rangle\langle\psi_x| \in B(\mathbb{C}^D)$ depending on her input x. Bob will conduct a two-outcome projective measurement labeled by y and defined by the projection operators $\{F_b^y : b = 0,1\}$, whose outcome will be Bob's guess. In sum, we need to solve the problem

$$\begin{gathered} \max \sum_{x,y} p(x,y)\mathrm{tr}\big(\rho_x F^y_{f(x,y)}\big), \\ \text{such that} \\ \mathrm{tr}(\rho_x) = 1, \quad \rho_x^2 = \rho_x, \quad \big(F_b^y\big)^2 = F_b^y, \\ \rho_x, \quad F_b^y \in B(\mathbb{C}^D). \end{gathered} \tag{37}$$

This problem can be reformulated by assuming that the initial state of Alice's system corresponds to $\rho_{x=0}$ and for any other input x she sends the state $V_x\rho_0 V_x^\dagger$, where V_x is a unitary operator that can be chosen self-adjoint, i.e., $V_x^2 = \mathbb{I}$. The resulting problem belongs to the class (1), and hence there is a converging SDP hierarchy to attack it. We observed that, in practice, such an SDP hierarchy gave good predictions for $D = 2$ at $k = 2$. For $D = 3$, a third-order relaxation did not suffice to reach the optimal probability of success in $2 \to 1$ QRAC [56].

We believe that the main reason for such a slow convergence rate is that the above proposal relies solely on MPIs to enhance dimension constraints. In order to devise a practical SDP hierarchy for problem (37), one needs to find a reformulation of problem (37) where the space $\mathcal{M}^k_{D,\vec{r}}$ is dimension-dependent even for low values of k. One such reformulation is immediate: Regard ρ_x as rank 1 projectors and assume that the state of the system is the (not normalized) tracial state $\mathbb{I}_D$.

The resulting hierarchy of SDPs should be easy to guess. First, we divide the representations of problem (37) into different classes r depending on the rank of the projectors $\{F_0^y\}$. Second, we generate random states ρ_x and projectors $\{F_0^y\}$ within the class r, and, taking the state of the system to be $\mathbb{I}_D$, we use them to build random feasible moment matrices. Those allow us to characterize the space $\mathcal{M}^k_{D,\vec{r}}$. Note that dimension constraints on $\mathcal{M}^k_{D,\vec{r}}$ are present for all D even for $k = 1$. For example, for any feasible first-order moment matrix M, $M_{\mathbb{I},\mathbb{I}} = D \times M_{\mathbb{I},\rho_x}$.

The above SDP hierarchy gives good results in practice, but we were not able to prove its convergence. Following Sec. IV, it can be shown that, for any class r, one can define a representation $\tilde{F}_b^y,\tilde{\rho}_x$ and a tracial state $D|\tilde{\psi}\rangle\langle\tilde{\psi}|$ which recover the limiting value of the hierarchy of SDPs. Furthermore, the operator algebra decomposes into a direct integral of representations z with the property that $\text{rank}(\tilde{F}_{b,z}^y) \leqslant r_b^y$, $\text{rank}(\mathbb{I}-\rho_{x,z}) \leqslant D-1$, $\text{rank}(\rho_{x,z}) \leqslant 1$. If the dimensionality of $\mathcal{H}_z$ is D, that defines a feasible point of problem (37). However, if the dimensionality of $\mathcal{H}_z$ is strictly smaller, we run into trouble: In such representations, $\rho_{x,z}$ can vanish for some values of x. Constraints such as $\text{tr}(\rho_z)=1$ can be accounted for by other representations t, where $\rho_{x,t}$ is a rank 1 projector, since $D\langle\tilde{\psi}_t|\rho_{x,t}|\tilde{\psi}_t\rangle = \frac{D}{\dim(\mathcal{H}_t)} > 1$.

One possibility to suppress the effects of lower finite-dimensional representations is to add "noncommuting constants," i.e., certain extra operators whose operator relations cannot be realized in dimensions lower than D. For instance, in order to guarantee that all representations Y have dimension $D=2$, we could include the Pauli matrices σ_z,σ_x as operators in the moment matrix M_k. With these extra variables, proving convergence can be done by appealing to the convergence of the Lasserre-Parrilo hierarchy [26,27]. However, we did not find a single situation in our numerical experiments where adding noncommuting constants to fix the dimension was of any advantage.

1. *Examples*

We explore how the relaxation of the above communication problem performs in practice. To do so, we establish (usually tight) upper bounds in QRAC for various values of k and dimension D. We also recompute quantum bounds for the witnesses I_N of Gallego *et al.* [57]. We further distinguish between real and complex Hilbert spaces and detect general POVM measurements assuming that Alice communicates Bob a quantum system of fixed dimension $D=2$. Note that Ref. [58] investigates a generalized QRAC problem where Alice's inputs $\vec{x}$ take values from a string of dits (instead of bit-strings). In that case, our SDP method also showed good performance [58].

a. QRAC. We suppose the QRAC has independently and uniformly distributed inputs and Alice is allowed to transmit Bob a D-level quantum system. We use the notation of Ref. [59] and we denote the average success probability of the optimal $k \to \log_2(D)$ QRAC by $P_{\max}[k \to \log_2(D)]$.

It was known previously from Ref. [56] that $P_{\max}(2\to 1) = 1/2+\sqrt{2}/4$. This is actually the value given by our SDP code at order 2, up to numerical precision. Likewise, when Alice is allowed to transmit a qutrit (case $D=3$), our relaxation based on tracial states at the same order 2 gives $P_{\max} \leqslant 0.904\,508\,50$, which matches with high numerical precision the lower bound value obtained via see-saw technique. Another method from the literature to attack this problem is the Mironowicz-Li-Pawłowski (MLP) SDP hierarchy [60], whose second-order relaxation gives us the (nontight) upper bound of 0.926 835 5.

One can prove that the MLP hierarchy does not converge in general. To do that, first notice that any QRAC can be rewritten as a full-correlation Bell inequality by defining Alice's observables as $A_x = 2\rho_x - \mathbb{I}$. Then, the only constraint that the MLP hierarchy adds to the NPA hierarchy is that $\langle A_x\rangle = 2-D$. Taking $D=2$, we see that the problem of calculating $P_{\max}(k\to 1)$ reduces to maximizing the violation of a full-correlation Bell inequality constraining (some of) its marginals to be uniform. By Tsirelson's theorem, the maximum is anyway attained when the marginals are uniform, so the constraint is automatically satisfied [53,61]. This means we can simply solve Tsirelson's SDP to find out this maximum and the minimal dimension necessary to attain it [61]. For $k=4$, we see that $D=4$ is necessary to reach the maximum, so the hierarchy did not converge to the maximum for $D=2$.

By increasing the dimension D and the parameter k, the second-order relaxation of the hierarchy based on the tracial states also performs well. The entries in the first three rows of Table IV are from Ref. [16], which shows lower and upper bounds on the average success probability for QRAC $k \to \log_2(D)$ for $k=3$ and for different values of D. The upper bounds (UB) are computed via our SDP using a normal desktop PC, and took less than 1 h for any of the D values (assuming a given rank-combination of measurements) using the solver SEDUMI [29]. The upper bounds (UB$'$) are resulting from the second-order relaxation of the MLP method [60]. We also show upper bounds (UB$''$) derived from the Moroder *et al.* [33] hierarchy by fixing negativity $(D-1)/2$ and adding the constraint $P(a|x)=1/D$ on Alice's marginal distributions. As Table IV shows, except for $D=2,4$, where the outputs of all methods coincide, the new tool gives predictions $\sim 10^{-2}$ more accurate than the MLP method and the method based on Moroder *et al.* hierarchy.

Let us pick $P_{\max}(3\to 1) = 0.788\,675$ from Table IV. This is precisely the value given by the construction of Chuang [56] proving optimality of the *complex* qubit value. However, we can apply in this case the same ideas to characterize the properties of *real* qubit systems as well. By generating the basis from randomly chosen real-valued qubit states ρ_x and projectors $\{F_0^y\}$, we get the (tight) upper bound 0.769 672 3. Hence, this simple example allows us to distinguish between real and complex two-level systems.

b. I_N family. As another example, we used the SDP program based on tracial states to recompute the maximal quantum value of the prepare-and-measure dimension witnesses I_N defined in Ref. [57], Table I. The second relaxation of the

TABLE IV. Lower (LB) and various upper bounds (UB, UB$'$, UB$''$) on $P_{\max}[3 \to \log_2(D)]$ detailed in the text.

D	2	3	4	5	6	7
LB	0.788 675	0.832 273	0.908 248	0.924 431	0.951 184	0.969 841
UB	0.788 675	0.832 273	0.908 248	0.924 445	0.954 123	0.969 841
UB$'$	0.788 675	0.853 553	0.908 248	0.934 264	0.957 785	0.979 567
UB$''$	0.788 675	0.852 156	0.908 248	0.931 201	0.954 140	0.977 072

TABLE V. Lower and upper bounds on I_7 witness for dimensions $D = 2,\ldots,6$. LB stands for the data in Ref. [63], Table I, whereas LB$'$ is due to our see-saw technique and UB is resulting from our SDP method.

D	2	3	4	5	6
LB	17.3976	20.7143	23.2167	24.8978	26.1017
LB$'$	17.3976	20.7085	23.2167	24.8987	26.1017
UB	17.3976	20.7718	23.2180	24.8991	26.1019

SDP hierarchy turns out to produce upper bounds for cases $N = 3,4$ and $D = 2,3$ which match the lower bounds obtained with see-saw method. Let us note that the conclusions of the experimental paper [62] relied on the conjecture that the inequality $I_4 \leqslant 7.9689$ cannot be violated by quantum systems of dimension $D = 3$.

In a recent experimental paper [63], the I_N dimension witness has been investigated for $N = 7$. Using a heuristic search, lower bounds are provided for dimensions $D = 2,\ldots,6$, which were conjectured to be optimal. The first row (LB) in our Table V shows lower bound results of Ref. [63], which match our lower bounds (LB$'$ in second row), except the case $D = 3$, where we got a slightly higher lower bound value. This also justifies the need to relaxation methods providing certified upper bound values. Remarkably, our SDP upper bound values (UB in third row) are close to the LB$'$ values differing in the worst case $D = 3$ in the second digit. For instance, Ref. [63] provides the experimental value of $I_7 = 25.44 \pm 0.02$ for the preparation of six-dimensional states. Therefore, our UB value of 24.8991 for $D = 5$ certifies the generation of at least six-dimensional quantum states.

c. POVM witnesses. Gisin [64] asked if there exists a Bell inequality which requires POVMs for optimal violation on some quantum state. This question has been answered affirmatively in case of two-qubit states (see, e.g., Refs. [17,18,65,66]). However, the question is still open for Bell inequalities defining facet of the local polytope (though, numerical study suggests the existence of such cases for three parties and high-dimensional states [67]). We pose a similar question in the prepare-and-measure communication scenario: By fixing dimension (say, Alice is allowed to send Bob a two-level system), are there witnesses which allow higher violations when Bob performs general POVM measurements instead of standard projective measurements? Our SDP tools allow one to certify the existence of such POVM witnesses.

To this end, we pick the V_4 witness of Ref. [68] and consider dimension $D = 2$. This witness consists of four preparations ($x = 1,\ldots,4$) on Alice's side and six measurements ($y = 1,\ldots,6$) on Bob's side. For $D = 2$, the maximum value of V_4 equals $2\sqrt{6}$, which can be attained if Alice prepares four states pointing toward the vertices of the regular tetrahedron (corresponding to SIC POVM elements). We add a four-outcome measurement to Bob ($y = 7$) with four outcomes $b = 1,2,3,4$ to the original V_4 witness and define the modified V_4 witness as follows:

$$V_4' = V_4 - \sum_{i=1}^{4} P(b = i|x = i, y = 7) \leqslant 2\sqrt{6} \simeq 4.8990. \tag{38}$$

We remark that a similar modification was used in the context of Bell nonlocality in Ref. [65]. As the last term in the inequality cannot be positive, the qubit bound $2\sqrt{6} \simeq 4.8990$ using POVMs follows from the bound on V_4. Indeed, by using the known optimal qubit settings for V_4, the only way of getting the maximal violation of $2\sqrt{6}$ is when Bob's POVM elements in setting $y = 7$ are antialigned with the four tetrahedron states prepared by Alice, so that all probabilities $P(b = i|x = i, y = 7)$ become zero. By assuming projective qubit measurements for Bob and running SDP in the case of $D = 2$, we obtain $2(\sqrt{2} + 1) \simeq 4.8284$ up to numerical precision on the witness V_4' in Eq. (38). Hence, any value bigger than 4.8284 for V_4' certifies in a semi-device-independent way that Bob's measurement $y = 7$ was, in fact, a general POVM measurement.

VII. SOME TIPS ON IMPLEMENTATION

In this section, we offer some tips on implementing the programs defined above. As we will see, those tricks will make the hierarchy for NPO under dimension constraints much easier to code and modify than its dimension-free counterpart [14]. For simplicity, we assume that our program does not involve localizing matrices, i.e., that all polynomial restrictions appear as identities. We also presume that operator representations can be divided into classes $\vec{r}$ and that generating a random instance of each class can be done efficiently.

First, we need a subroutine [X,rho] = genSamp(D,n,r) that generates a cell-array X of random operators $X_0, X_1, \ldots, X_n$, with $X_0 = \mathbb{I}_D$ and $X_1, \ldots, X_n \in B(\mathbb{C}^D)$ satisfying the appropriate class constraints, determined by the vector r. GENSAMP must also return a quantum state $\rho \in B(\mathbb{C}^D)$ (in our examples, either a pure random state $|\psi\rangle\langle\psi|$, the maximally entangled state, or the un-normalized maximally mixed state).

Second, we need a subroutine G = buildG(X,rho,k) that, given the cell array X and an index k, generates a kth-order moment matrix of the form

$$G_{\vec{i},\vec{j}} = \mathrm{tr}\big(\rho X_{i_k}^\dagger \cdots X_{i_1}^\dagger X_{j_1} \cdots X_{j_k}\big), \tag{39}$$

where $\vec{i},\vec{j} \in \{0,\ldots,n\}^k$ if the variables $X_1,\ldots,X_n$ are Hermitian or $\vec{i},\vec{j} \in \{0,\ldots,2n\}^k$ if they are not. In either case, monomials of G can be accessed via $\mathrm{tr}(G|\vec{i}\rangle\langle\vec{j}|)$, choosing $\vec{i},\vec{j}$ appropriately. Note that, in this representation, different columns of G correspond to the same operator, e.g., 01 and 10. That does not matter, because, at the end of the day, redundant columns will be suppressed by the matrix V mapping the space where G is defined to the support of span$\{G\}$; see Remark 1.

If the reader is an OCTAVE or MATLAB user, the following considerations will lead to very fast code.

Any $D \times D$ matrix $A = \sum_{i,j=1}^{D} A_{i,j}|i\rangle\langle j|$ can be represented in vector form as $|A\rangle = \sum_{i,j} A_{ij}|i\rangle|j\rangle$: To go from A to $|A\rangle$, one can invoke the in-built function RESHAPE. It can be verified that $(\mathbb{I}_D \otimes \langle\psi^+| \otimes \mathbb{I}_D)|A\rangle|B\rangle = |AB\rangle$, where $|\psi^+\rangle$ is the non-normalized maximally entangled state $|\psi^+\rangle = \sum_{i=1}^{D} |i,i\rangle$. Similarly, $(\mathbb{I}_D \otimes \langle\phi|)|A\rangle = A|\phi^*\rangle$, and, hence, $\langle A|(\mathbb{I}_D \otimes \rho^*)|B\rangle = \mathrm{tr}(A^\dagger B\rho)$ for any Hermitian matrix $\rho \in B(C^D)$.

Now, for random operators $X_1,\ldots,X_n$ in the class r, define $\Lambda=\sum_{i=0}^{n}|X_i\rangle\langle i|$. From the above, it follows that

$$(\mathbb{I}_D\otimes\langle\psi^+|\otimes\mathbb{I}_D)\left(\Lambda\otimes\sum_{j_2,\ldots,j_l}|X_{j_2}\cdots X_{j_l}\rangle\langle j_2,\ldots,j_l|\right)$$
$$=\sum_{j_1,\ldots,j_l}|X_{j_1}\cdots X_{j_l}\rangle\langle j_1,\ldots,i_l|. \quad (40)$$

By tensoring Λ sequentially and projecting on the maximally entangled state, we thus (quickly) obtain the operator

$$C\equiv\sum_{j_1,\ldots,j_k}|X_{j_1}\cdots X_{j_k}\rangle\langle j_1,\ldots,i_k|. \quad (41)$$

Then it can be verified that

$$G=C^{\dagger}(\mathbb{I}_D\otimes\rho^*)C. \quad (42)$$

Finally, we must code a third subroutine [basisMat,V]= buildBasis(D,n,r,k) that, by repeatedly calling GENSAMP and BUILDG, derives an orthonormal matrix basis for span$\{G\}$ and the isometry V described in Remark 1. From there, it is straightforward to implement program (7).

Notice that all operator relations are determined by GENSAMP only. This means, for example, that if we wish to optimize over projection (unitary operators) all we need to do is program GENSAMP to generate random tuples of projection (unitary) operators $P_1,\ldots,P_n$ $(U_1,\ldots,U_n,U_1^{\dagger},\ldots,U_n^{\dagger})$. Switching from one type of polynomial constraints to another can thus be done straightforwardly with the above implementation.

VIII. CONCLUSION

In this paper we have extended the notion of NPO to scenarios where the dimensionality of the spaces where the noncommuting variables act is bounded from above. We have presented a complete hierarchy of SDP relaxations to solve such problems and we have explored its performance by applying it to solve a number of open problems in quantum information theory.

Our research raises several questions which deserve further study. The first one is whether the hierarchy of relaxations proposed to study one-way quantum complexity is complete. As we showed, the main obstacle to prove convergence lies in interference from low-dimensional operator representations with “dark states.” For $D=2$, the extreme points of such additional one-dimensional representations are finite, and thus the problem of determining whether the hierarchy converges in a given prepare-and-measure scenario amounts to proving that all such points can be reproduced by two-level quantum systems. Similarly, the convergence of the sequence of relaxations to the maximum value of a specific functional for arbitrary D can be verified by establishing an upper bound for the dark-state value smaller than or equal to the value of a concrete D-dimensional realization. It would be more satisfactory, though, to have a general convergence result.

Another open problem is whether the relaxations proposed for optimizations over finite-dimensional *real* operator algebras actually converge. Here we are faced with the problem that certain operator representations, irreducible in the real space, can be expressed as a direct sum of nontrivial representations if we allow complex unitary transformations. Hence, if we wished to follow the proof for complex algebras, we would encounter problems at the step of applying von Neumann’s direct-integral theorem [24].

Finally, note that in this work we have not exploited the symmetry of the functionals to optimize. If our aim is to bound quantum nonlocality under dimension constraints, that leaves us with a method that, in the best scenario, would allow us to conduct optimizations for $D=2,3$ with a normal computer. Which dimensionalities could become accessible if we chose to play with symmetries is an intriguing question.

ACKNOWLEDGMENTS

M.N. acknowledges support by Spanish MINECO Project No. FIS2013-40627-P. T.V. acknowledges support from the Hungarian National Research Fund OTKA (Grant No. K111734) and a János Bolyai Grant of the Hungarian Academy of Sciences. M.A. and A.F. acknowledge support from the European Commission project RAQUEL (Grant No. 323970) and the Austrian Science Fund (FWF) through the Special Research Programme FoQuS, Individual Project (Grant No. 2462) and the Doctoral Programme CoQuS.

APPENDIX: TYPE II AND III FACTORS VIOLATE THE STANDARD MPI

The purpose of this appendix is to prove its title. To do so, we require the following lemma.

Lemma 1. Let $\mathcal{A}$ be a type II or III factor. Then for all n there exists nontrivial projectors $\{P_i\}_{i=1}^n\subset\mathcal{A}$ such that

(1) $\sum_{i=1}^n P_i=\mathbb{I}$;

(2) $P_1\mathcal{A}P_2\mathcal{A}\cdots P_n\mathcal{A}\neq 0$.

Proof. Suppose that the statement is not true for general n and let N be the greatest number such that it holds. Note that $N>1$. Indeed, if $P_1\mathcal{A}P_2=0$ for nontrivial P_1,P_2, then for all $x\in\mathcal{A}$,

$$[x,P_1]=(P_1+P_2)[x,P_1](P_1+P_2)$$
$$=P_1xP_1+P_2xP_1-P_1xP_1-P_1xP_2=0. \quad (A1)$$

That is impossible, because factors, by definition, are central; i.e., the only elements of $\mathcal{A}$ commuting with $\mathcal{A}$ are multiples of the identity.

Now suppose that $\{P_i\}_{i=1}^N$ satisfy the conditions of the lemma. Then we can always write $P_N=P_N'+P_{N+1}'$, where P_N',P_{N+1}' are nonzero projectors such that $P_1'\mathcal{A}P_2'\mathcal{A}\cdots P_N'\mathcal{A}\neq 0$. Call $\mathcal{B}$ the algebra generated by $\mathcal{A}P_1'\mathcal{A}P_2'\mathcal{A}\cdots P_N'\mathcal{A}$, and let Π denote its identity, i.e., a projector $\Pi\in\mathcal{B}$ such that $\Pi y=y$ for all $y\in\mathcal{B}$. Note that, due to the definition of $\mathcal{B}$, $\mathcal{A}\mathcal{B}=\mathcal{B}\mathcal{A}=\mathcal{B}$, and hence $x\Pi,\Pi x\in\mathcal{B}$ for all $x\in\mathcal{A}$. It follows that $\Pi x=\Pi x\Pi=x\Pi$; that is, $[x,\Pi]=0$.

On the other hand, N is the greatest number such that the conditions of the lemma hold, and so $\mathcal{A}P_1'\cdots\mathcal{A}P_N'\mathcal{A}P_{N+1}'=P_{N+1}'\mathcal{A}P_1'\cdots\mathcal{A}P_N'\mathcal{A}=0$. In other words, $\mathcal{B}P_{N+1}'=0$, and, consequently, $\Pi P_{N+1}'=0$. We conclude that Π, which is neither 0 nor the identity, must commute with all $x\in\mathcal{A}$, contradicting the centrality of $\mathcal{A}$. ■

Now, given n and a type II or III factor $\mathcal{A}$, choose orthogonal projectors $\{P_i\}_{i=1}^n \subset \mathcal{A}$ and operators $x_1,\ldots,x_n \in \mathcal{A}$ such that $P_1x_1P_2x_2\cdots P_n \neq 0$. Then define the operators $E_i \equiv P_ix_iP_{i+1}$ for $i=1,\ldots,n-1$ and compute the fundamental polynomial $I_{n-1}(E)$. Due to the orthogonality of $\{P_i\}$, the only nonvanishing product is $E_1E_2\cdots E_{n-1} = P_1x_1P_2x_2\cdots P_n \neq 0$. $\mathcal{A}$, hence, violates $I_{n-1}=0$.

[1] A. C.-C. Yao, Quantum circuit complexity, *Proceedings of the 34th IEEE FOCS* (IEEE Computer Society Press, Los Alamitos, CA, 1993), p. 352.
[2] H. Buhrman, R. Cleve, S. Massar, and R. de Wolf, Nonlocality and communication complexity, Rev. Mod. Phys. **82**, 665 (2010).
[3] John S. Bell, On the einstein podolsky rosen paradox, Physics **1**, 195 (1964).
[4] N. Brunner, S. Pironio, A. Acín, N. Gisin, A. A. Méthot, and V. Scarani, Testing the Dimension of Hilbert Spaces, Phys. Rev. Lett. **100**, 210503 (2008).
[5] K. F. Pál and T. Vértesi, Efficiency of higher dimensional Hilbert spaces for the violation of Bell inequalities, Phys. Rev. A **77**, 042105 (2008).
[6] S. Wehner, M. Christandl, and A. C. Doherty, Lower bound on the dimension of a quantum system given measured data, Phys. Rev. A **78**, 062112 (2008).
[7] J. Briët, H. Buhrman, and B. Toner, A generalized grothendieck inequality and nonlocal correlations that require high entanglement, Commun. Math. Phys. **305**, 827 (2011).
[8] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wooters, Purification of Noisy Entanglement and Faithful Teleportation Via Noisy Channels, Phys. Rev. Lett. **76**, 722 (1996).
[9] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, Phys. Rev. Lett. **23**, 880 (1969).
[10] L. A. Khalfin and B. S. Tsirelson, Quantum and quasi-classical analogues of Bell inequalities, *Symposium on the Foundations of Modern Physics*, edited by P. Lahti *et al.* (World Scientific Publishing, Singapore, 1985), pp. 441–460.
[11] Sandu Popescu and Daniel Rohrlich, Quantum nonlocality as an axiom, Found. Phys. **24**, 379 (1994).
[12] M. Navascués, S. Pironio, and A. Acín, Bounding the Set of Quantum Correlations, Phys. Rev. Lett. **98**, 010401 (2007).
[13] M. Navascués, S. Pironio, and A. Acín, A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations, New J. Phys. **10**, 073013 (2008).
[14] M. Navascués, S. Pironio and A. Acín, Convergent relaxations of polynomial optimization problems with noncommuting variables, SIAM J. Optim. **20**, 2157 (2010).
[15] L. Vandenberghe and S. Boyd, Semidefinite programming, SIAM Rev. **38**, 49 (1996).
[16] M. Navascués and T. Vértesi, Bounding the Set of Finite Dimensional Quantum Correlations, Phys. Rev. Lett. **115**, 020501 (2015).
[17] T. Vértesi and E. Bene, Two-qubit bell inequality for which positive operator-valued measurements are relevant, Phys. Rev. A **82**, 062115 (2010).
[18] J. F. Barra, E. S. Gómez, G. Cañas, W. A. T. Nogueira, L. Neves, and G. Lima, Higher quantum bound for the vértesi-bene-bell inequality and the role of positive operator-valued measures regarding its threshold detection efficiency, Phys. Rev. A **86**, 042114 (2012).
[19] C. Budroni, T. Moroder, M. Kleinmann, and O. Gühne, Bounding Temporal Correlations, Phys. Rev. Lett. **111**, 020403 (2013).
[20] C. Procesi, *Rings with Polynomial Identities* (Dekker, New York, 1973).
[21] M. Navascués and T. Vértesi, The structure of matrix product states, arXiv:1509.04507.
[22] Å. Björck and C. C. Paige, Loss and recapture of orthogonality in the modified gramschmidt algorithm, SIAM. J. Matrix Anal. Appl. **13**, 176 (1992).
[23] M. Reed and B. Simon, *Functional Analysis* (Academic Press, New York, 1980).
[24] J. von Neumann, On rings of operators. Reduction theory, Ann. Math., 2nd Ser. **50**, 401 (1949).
[25] M. Takesaki, *Theory of Operator Algebras I* (Springer, Berlin, 1979).
[26] J. B. Lasserre, Global optimization with polynomials and the problem of moments, SIAM J. Optim. **11**, 796 (2001).
[27] P. Parrilo, Structured semidefinite programs and semialgebraic geometry methods in robstness and optimization, Ph.D. thesis, California Institute of Technology, Pasadena, CA, 2000.
[28] The Mosek optimization toolbox for MATLAB manual, Version 7.0 (Revision 140) (MOSEK ApS, Denmark), http://docs.mosek.com/7.0/toolbox/index.html.
[29] J. F. Sturm, Using sedumi 1.02, a matlab toolbox for optimization over symmetric cones, Optim. Methods Software **11**, 625 (1999).
[30] J. Löfberg, YALMIP: A toolbox for modeling and optimization in MATLAB, in *Proceedings of the CACSD Conference, Taipei, Taiwan* (IEEE, 2004), pp. 284–289.
[31] D. Collins and N. Gisin, A relevant two qubit Bell inequality inequivalent to the CHSH inequality, J. Phys. A: Math. Gen. **37**, 1775 (2004).
[32] M. Navascués, G. de la Torre, and T. Vértesi, Characterization of Quantum Correlations with Local Dimension Constraints and its Device-Independent Applications, Phys. Rev. X **4**, 011011 (2014).
[33] T. Moroder, J.-D. Bancal, Y.-C. Liang, M. Hofmann, and O. Gühne, Device-Independent Entanglement Quantification and Related Applications, Phys. Rev. Lett. **111**, 030501 (2013).
[34] G. Vidal and R. F. Werner, Computable measure of entanglement, Phys. Rev. A **65**, 032314 (2002).
[35] T. Vidick and S. Wehner, More nonlocality with less entanglement, Phys. Rev. A **83**, 052310 (2011).
[36] B. Lang, T. Vértesi, and M Navascués, Closed sets of correlations: Answers from the zoo, J. Phys. A: Math. Theor. **47**, 424029 (2014).
[37] K. F. Pál and T. Vértesi, Quantum bounds on bell inequalities, Phys. Rev. A **79**, 022120 (2009).

[38] R. F. Werner and M. M. Wolf, All-multipartite bell-correlation inequalities for two dichotomic observables per site, Phys. Rev. A **64**, 032112 (2001).

[39] K. F. Pál and T. Vértesi, Maximal violation of the i3322 inequality using infinite dimensional quantum systems, Phys. Rev. A **82**, 022116 (2010).

[40] D. Rosset, J.-D. Bancal, and N. Gisin, Classifying 50 years of Bell inequalities, J. Phys. A: Math. Theor. **47**, 424022 (2014).

[41] T. Vértesi and K. F. Pál, Bounding the dimension of bipartite quantum systems, Phys. Rev. A **79**, 042106 (2009).

[42] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2000).

[43] S. Pironio, All Clauser–Home–Shimony–Holt polytopes, J. Phys. A: Math. Theor. **47**, 424020 (2014).

[44] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu, Bell Inequalities for Arbitrarily High-Dimensional Systems, Phys. Rev. Lett. **88**, 040404 (2002).

[45] A. Acín, T. Durt, N. Gisin, and J. I. Latorre, Quantum nonlocality in two three-level systems, Phys. Rev. A **65**, 052325 (2002).

[46] J. Tura, R. Augusiak, A. B. Sainz, T. Vértesi, M. Lewenstein, and A. Acín, Detecting nonlocality in many-body quantum states, Science **344**, 1256 (2014).

[47] M. Huber, F. Mintert, A. Gabriel, and B. C. Hiesmayr, Detection of High-Dimensional Genuine Multipartite Entanglement of Mixed States, Phys. Rev. Lett. **104**, 210501 (2010).

[48] C. Spengler, M. Huber, A. Gabriel, and B. C. Hiesmayr, Examining the dimensionality of genuine multipartite entanglement, Quantum Inf. Process. **12**, 269 (2013).

[49] N. Brunner and T. Vértesi, Persistency of entanglement and nonlocality in multipartite quantum systems, Phys. Rev. A **86**, 042113 (2012).

[50] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, Rev. Mod. Phys. **81**, 865 (2009).

[51] O. Gühne and G. Tóth, Entanglement detection, Phys. Rep. **474**, 1 (2009).

[52] J.-D. Bancal, N. Gisin, Y.-C. Liang, and S. Pironio, Device-Independent Witnesses of Genuine Multipartite Entanglement, Phys. Rev. Lett. **106**, 250404 (2011).

[53] B. S. Cirel'son, Quantum generalizations of Bell's inequality, Lett. Math. Phys. **4**, 93 (1980).

[54] Ll. Masanes, Extremal quantum correlations for N parties with two dichotomic observables per site, arXiv:quant-ph/0512100.

[55] J. Tura, A. B. Sainz, T. Vértesi, A. Acín, M. Lewenstein, and R. Augusiak, Translationally invariant multipartite Bell inequalities involving only two-body correlators, J. Phys. A: Math. Theor. **47**, 424024 (2014).

[56] A. Ambainis, A. Nayak, A. Ta-Shma, and U. Vazirani, Dense quantum coding and quantum finite automata, J. ACM **49**, 496 (2002).

[57] R. Gallego, N. Brunner, C. Hadley, and A. Acín, Device-Independent Tests of Classical and Quantum Dimensions, Phys. Rev. Lett. **105**, 230501 (2010).

[58] A. Tavakoli, A. Hameedi, B. Marques, and M. Bourennane, Quantum Random Access Codes Using Single d-level systems, Phys. Rev. Lett. **114**, 170502 (2015).

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

[60] P. Mironowicz, H.-W. Li, and M. Pawłowski, Properties of dimension witnesses and their semidefinite programming relaxations, Phys. Rev. A **90**, 022322 (2014).

[61] D. Avis, S. Moriyama, and M. Owari, From Bell inequalities to Tsirelson's theorem, IEICE Trans. Fundam. **92**, 1254 (2009).

[62] M. Hendrych, R. Gallego, M. Mičuda, N. Brunner, A. Acín, and J. P. Torres, Experimental estimation of the dimension of classical and quantum systems, Nat. Phys. **8**, 588 (2012).

[63] V. D'Ambrosio, F. Bisesto, F. Sciarrino, J. F. Barra, G. Lima, and A. Cabello, Device-Independent Certification of High-Dimensional Quantum Systems, Phys. Rev. Lett. **112**, 140503 (2014).

[64] N. Gisin, Bell inequalities: many questions, a few answers, in *Quantum Reality, Relativistic Causality, and Closing the Epistemic Circle* (Springer, Berlin, 2009), p. 125.

[65] A. Acín, S. Pironio, T. Vértesi, and P. Wittek, Optimal randomness certification from one entangled bit, arXiv:1505.03837.

[66] M. Kleinmann and A. Cabello, Quantum correlations cannot arise from measurements with limited number of outcomes, arXiv:1505.04179.

[67] B. Grandjean, Y.-C. Liang, J.-D. Bancal, N. Brunner, and N. Gisin, Bell inequalities for three systems and arbitrarily many measurement outcomes, Phys. Rev. A **85**, 052113 (2012).

[68] N. Brunner, M. Navascués, and T. Vértesi, Dimension Witnesses and Quantum State Discrimination, Phys. Rev. Lett. **110**, 150501 (2013).