More nonlocality with less entanglement

Thomas Vidick and Stephanie Wehner

Phys. Rev. A **83**, 052310 — Published 13 May 2011

DOI: 10.1103/PhysRevA.83.052310

More non-locality with less entanglement

Thomas Vidick[1] and Stephanie Wehner[2]

[1] *Computer Science division, UC Berkeley, 94720 Berkeley, USA*[*]

[2] *Center for Quantum Technologies, National University of Singapore, 2 Science Drive 3, 117543 Singapore*[†]

(Dated: March 30, 2011)

Recent numerical investigations (Pál and Vértesi, PRA (2010)) suggest that the I3322 inequality, arguably the simplest extremal Bell inequality after the CHSH inequality, has a very rich structure in terms of the entangled states and measurements that maximally violate it. Here, we show that for this inequality the maximally entangled state of any dimension achieves the same violation than just a single EPR pair. In contrast, stronger violations can be achieved using higher dimensional states which are *less* entangled. This shows that the maximally entangled state is not the most nonlocal resource, even when one restricts attention to the most simple extremal Bell inequalities.

I. INTRODUCTION

Entanglement is a powerful resource, facilitating computation, communication, or more generally any nonlocal task. Like all resources it is useful to be able to measure it, so that entangled states could be ranked according to their usefulness for a given task. A very natural measure for the entanglement of any bipartite state $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ is the entropy of entanglement $E(\Psi) = S(\rho_A)$ [5], where $S(\rho_A) = -\mathrm{Tr}(\rho_A \log \rho_A)$ is the von Neumann entropy and $\rho_A = \mathsf{tr}_B(|\Psi\rangle\langle\Psi|)$ is the reduced density operator of $|\Psi\rangle$ on one of the two subsystems. In any dimension d this measure is maximized by the maximally entangled state

$$|\Psi_d\rangle = \frac{1}{\sqrt{d}} \sum_{j=1}^{d} |j\rangle|j\rangle \ . \tag{1}$$

Since $|\Psi_d\rangle$ exhibits the largest amount of entanglement, it would be natural to guess that it would indeed be the most useful state for any nonlocal task. This belief is reinforced by the fact that this state has proven extremely useful for many quantum information problems (e.g. [6, 8, 17]), and is by itself a sufficient resource for the creation of *any* other nonlocal state as soon as one allows local operations and classical communication (LOCC) [6]. Moreover, it is known that any shared pure state $|\Psi\rangle$ violates a Bell inequality if and only if it is entangled [19, 36], suggesting that the amount of entanglement may play a central role in quantifying the strength of nonlocal correlations.

For a long time it was implicitly assumed that $|\Psi_d\rangle$ is the most useful state with respect to violation of Bell inequalities [4]. The first doubts cast on this conjecture stem from a result by Eberhard [16] who showed that when it comes to closing the detection efficiency loophole less entangled states can be more useful. More recently, such doubts were confounded by the surprising fact that, at least in small dimensions in which numerical experiments can be conducted, there are inequalities for which the maximally entangled state does not give the maximum violation. More specifically, for every dimension d there is a Bell inequality (such as the CGLMP inequality [14]) which in that dimension is maximally violated by a state different from the maximally entangled state — a state with *lower* entanglement [1, 2, 39]. Conversely, it is sometimes necessary to use a larger amount of certain maximally non-local resources in order to simulate all possible correlations coming from some less entangled state, compared to what is necessary to simulate those coming from the maximally entangled state [11]. This has prompted the realization that *nonlocality* might be a resource of a different nature than entanglement, and many other examples have been discovered in the realm of Bell inequalities and quantum cryptography (see [29] for a survey), as well as quantum information theory [7, 9]. In a recent breakthrough, Junge and Palazuelos [22] showed using tools employed in the study of operator algebras and a probabilistic argument that there *exists* a family of Bell inequalities for which the maximally entangled state of any dimension can only lead to arbitrarily weaker violations than optimal. However, these Bell inequalities are very large and non-explicit.

A. The I_{3322} inequality

The only extremal [40] Bell inequality with two settings and two outcomes per site is the CHSH inequality, for which it is known that achieving violations close to optimal requires the use of a state arbitrarily close to an EPR pair [21]; optimal measurements are also well-understood [28].

In general, the nonlocal properties of Bell inequalities with two settings and two outcomes per site are reasonably well-understood. In that case it is known that we may without loss of generality restrict our attention to entangled states with local dimension 2 only [20, 27], as they are sufficient to reproduce all possible correlations. As a consequence, those inequalities lend themselves to extensive numerical and analytical investigations.

In contrast, as soon as one considers inequalities with more than two settings per site, the minimal local dimension required to achieve optimal violation is not known. In fact, recent extensive numerical investigations [33] suggest that the simplest extremal inequality after CHSH, the I_{3322} inequality

[*] Electronic address: `vidick@eecs.berkeley.edu`
[†] Electronic address: `wehner@nus.edu.sg`

(first introduced in [18], its name refers to the fact it has three settings and two outcomes per site), allows for a surprisingly complex structure of the maximally violating states.

We will use $\{A_j\}_{j\in\{1,2,3\}}$ and $\{B_k\}_{k\in\{1,2,3\}}$ to denote the measurement operators for the first of the two possible outcomes for Alice and Bob respectively. Using the common shorthands

$$\langle A_j B_k\rangle := \langle\Psi|A_j\otimes B_k|\Psi\rangle\ , \tag{2}$$
$$\langle A_j\rangle := \langle\Psi|A_j\otimes \mathsf{id}|\Psi\rangle\ , \tag{3}$$
$$\langle B_k\rangle := \langle\Psi|\mathsf{id}\otimes B_k|\Psi\rangle\ , \tag{4}$$

we define

$$\begin{aligned}\langle I_{3322}\rangle := &-\langle A_2\rangle - \langle B_1\rangle - 2\langle B_2\rangle + \langle A_1B_1\rangle\\ &+\langle A_1B_2\rangle + \langle A_2B_1\rangle + \langle A_2B_2\rangle - \langle A_1B_3\rangle\\ &+\langle A_2B_3\rangle - \langle A_3B_1\rangle + \langle A_3B_2\rangle\end{aligned} \tag{5}$$

While for classical correlations we have

$$\langle I_{3322}\rangle \le 0\ , \tag{6}$$

there exist measurements [13] such that using just one EPR pair (i.e. $|\Psi\rangle = |\Psi_2\rangle$) one can get

$$\langle I_{3322}\rangle = \frac{1}{4}\ . \tag{7}$$

Yet, the precise maximum of $\langle I_{3322}\rangle$ over all quantum states and measurements remains unknown. Numerical upper-bounds were obtained using a SDP hierarchy in [15]. This was followed by recent exhaustive numerical investigations by Pál and Vértesi [33], who report very interesting results. Their experiments suggest that the optimum violation of (6), even though it only involves a constant number of settings and outcomes, might only be reached in infinite dimension. Indeed, they find strategies obtaining a value of at least $0.25084...$ (matching the upper bound up to precision 10^{-7} in dimension ≈ 100), and moreover in their experiments this value keeps increasing as the dimension of the strategies is allowed to increase. Moreover, even though the observables which achieve the maximum violation in a given dimension have a rather simple and systematic form, the corresponding state has an interesting distribution of Schmidt coefficients, and it is quite far from the maximally entangled state. Their results, however, provide no indication of whether similar violations might be reached (perhaps at the price of increased dimension) with much simpler entangled states, such as the maximally entangled state.

B. Result

Our main result is that indeed the maximally entangled state does not lead to optimal violation of the I_{3322} inequality. In fact, a maximally entangled state of dimension d is no more useful than one of dimension 2, that is, a single EPR pair. More precisely, we show the following

Theorem 1. *For all dimensions $d \ge 0$, and any observables, using the maximally entangled state $|\Psi\rangle = |\Psi_d\rangle$ can lead to a violation of at most*

$$\langle I_{3322}\rangle \le \frac{1}{4}\ . \tag{8}$$

Note that in contrast with previous work, (7) tells us that a value of $1/4$ can be attained using just one EPR pair, and hence the maximally entangled state in any dimension is no more powerful than the maximally entangled state for $d = 2$.

Our result gives a strong demonstration that maximally entangled states are not the most nonlocal. Posterior to our work, a result appeared [25] showing, among other things, that there also exists Bell inequalities with two outcomes and two settings for which the maximally entangled state is not optimal. Similar results also follow from previous work [16, 20, 27]. However, these inequalities are somewhat artificial (the motivation for the work [25] is in a different context in which such inequalities are indeed interesting); in particular it is known that they are not extremal and one already knows [20, 27] that an optimal violation can always be reached with local dimension 2. Our result contributes to the understanding of more complex Bell inequalities, by showing that a similar phenomenon arises naturally and in a setting where using states of arbitrarily large dimension *can* actually be helpful — but, as we show, maximally entangled states themselves are not.

C. Generic states

Before embarking on our proof, it is worth pointing out that there does in fact exist a generic family of states that always allow us to obtain the maximum violation for *any* Bell inequality. These states, however, exhibit less entanglement than the maximally entangled state of same dimension. This “universal” family of states are known as *embezzlement states* [38]. They previously played an important role in more involved tasks in quantum information theory, namely the so-called quantum reverse Shannon theorem [7, 9], which provided another example where the maximally entangled state is not sufficient to achieve the corresponding channel simulation result, but the universal embezzlement states are. The key property of the d-dimensional embezzlement state $|\Phi_d\rangle$ that is used is that, for *any* pure state $|\Psi\rangle$, there exists d and d' such that $|\Phi_d\rangle \approx |\Phi_{d'}\rangle\otimes|\Psi\rangle$, where the equivalence only requires the application of local unitaries on each system; no communication is needed [38]. Since an embezzlement state can be used to obtain any other pure state by local unitary operations, it immediately follows that any Bell inequality can be maximally violated by an embezzlement state (of possibly higher dimension), as pointed out recently in [32]. This demonstrates that, even though in small dimensions it might seem like every inequality has its own specialized maximizing state, if one allows the dimension to grow larger, then a simple class of states is sufficient to obtain maximal violations.

II. USING THE MAXIMALLY ENTANGLED STATE

We now give a detailed overview of the proof of our main result (Theorem 1), relegating technical details to the appendix. Throughout we will refer to a particular choice of measurements applied to the maximally entangled state as a *strategy*. Since our game is binary, it is known that we may assume without loss of generality (and without affecting the underlying state) that the operators used by Alice and Bob are projectors [12, Proposition 2], and we will denote them by $\{A_j, \mathsf{id} - A_j\}$ for Alice and $\{B_k, \mathsf{id} - B_k\}$ for Bob. We will also refer to

$$\omega := \langle I_{3322} \rangle \tag{9}$$

as the *value* of a particular strategy. Our goal is to show that ω is at most $1/4$, irrespective of the dimension d. We first introduce an important tool in our analysis, the CS decomposition of a pair of projectors. This decomposition was also at the heart of the results in [20, 27], where it was used to handle the case of only *two* observables per site.

The CS decomposition. Given a pair of d-dimensional projectors P and Q, there exists an orthonormal basis in which the two projectors are jointly block-diagonal (see for instance [10]). Moreover, the blocks can be either 1-dimensional, in which case P and Q either have a 0 or a 1 in that block, or 2-dimensional, in which case they can be written in the form

$$P = \frac{1}{2}\begin{pmatrix} 1-c & -s \\ -s & 1+c \end{pmatrix}\,, \tag{10}$$

$$Q = \frac{1}{2}\begin{pmatrix} 1-c & s \\ s & 1+c \end{pmatrix}\,, \tag{11}$$

for some coefficients $c \in (-1,1)$ and $s = \sqrt{1-c^2}$. The angles θ such that $c = \cos\theta$ are called the *principal angles* between the subspaces on which P and Q project.

Our proof proceeds in two steps. Step 1 is to show that we can greatly simplify the form of Alice's and Bob's measurement operators. The main idea is to show using the CS decomposition that for any strategy maximizing (5) there exists a basis in which all measurements are tridiagonal [41]. This lets us greatly reduce the number of parameters and give a relatively simple analytic expression for the value ω of the strategy. Step 2 consists in upper-bounding this simple expression using standard analytic techniques.

A. Step 1: A simple joint normal form

This is arguably the most crucial step in our proof, as it lets us show that a completely arbitrary strategy given by projectors $\{A_j, B_k\}_{j,k=1,\ldots,3}$ can be put into a much simpler form without decreasing its value. As we mentioned previously, the key idea is to apply the CS decomposition twice, once to the pair (A_1, A_2), and once to the pair (B_1, B_2). This results in two orthonormal bases $\mathcal{B}_A$ and $\mathcal{B}_B$ such that the matrices of (A_1, A_2) in $\mathcal{B}_A$ are block-diagonal, with blocks of the form (10) for A_1 and (11) for A_2, and similarly for (B_1, B_2) in $\mathcal{B}_B$. We number the blocks of (A_1, A_2) using even indices $2, \ldots, d$ and call the corresponding coefficients c_{2i}, s_{2i}; the blocks of (B_1, B_2) are numbered using odd indices $1, \ldots, d+1$ and corresponding coefficients c_{2i+1}, s_{2i+1}.

In general the bases $\mathcal{B}_A$ and $\mathcal{B}_B$ are unrelated, but we argue that, under the condition that the strategy maximizes (5), they must in fact be permutations of one another. To see this, note that (5) can be re-written as

$$\begin{aligned}&\langle I_{3322}\rangle = \\ &\langle A_1 + A_2, B_1 + B_2\rangle + \langle A_2 - A_1, B_3\rangle + \langle A_3, B_2 - B_1\rangle \\ &- \langle A_2, \mathsf{id}\rangle - \langle \mathsf{id}, B_1\rangle - 2\langle \mathsf{id}, B_2\rangle\end{aligned} \tag{12}$$

where

$$\langle A, B\rangle = \frac{1}{d}\mathrm{Tr}(A^T B) \tag{13}$$

and we used that if $|\Psi\rangle$ is the maximally entangled state then

$$\langle \Psi | A \otimes B | \Psi\rangle = \langle A, B\rangle\ . \tag{14}$$

Note that since the A_j operators always appear on the left of the tensor product (Alice's side), we will henceforth argue about A_j^T rather than A_j, omitting the transpose sign for simplicity of notation. For the moment, let's ignore the contribution of the last three terms in (12). Observe that A_3 (resp. B_3) only appears in the term $\langle A_3, B_2 - B_1\rangle$ (resp. $\langle A_2 - A_1, B_3\rangle$). When maximizing over A_3 it is thus clear that the optimal choice is to make A_3 the projector onto the positive eigenspace of $B_2 - B_1$ (resp. B_3 to project on the positive eigenspace of $A_2 - A_1$). This in particular implies that the value of those two terms is *independent* of the choice of $\mathcal{B}_B$ (resp. $\mathcal{B}_A$). Hence the choice of the bases $\mathcal{B}_A$, $\mathcal{B}_B$ only bears influence on the value of the first term in (12).

Let us now examine the first term. Note that the precise form (10), (11) in which we wrote the CS decomposition ensures that $A_1 + A_2$ is diagonal in $\mathcal{B}_A$ (resp. $B_1 + B_2$ in $\mathcal{B}_B$). It is well known (see Claim 6 in the appendix) that $\langle A_1 + A_2, B_1 + B_2\rangle$ is maximized whenever the vectors in $\mathcal{B}_B$ are a permutation of those in $\mathcal{B}_A$. It follows that for the optimal choice of bases $A_1 + A_2$ and $B_1 + B_2$ will necessarily be simultaneously diagonal.

However, this does not necessarily imply that the blocks of (A_1, A_2) are aligned with those of (B_1, B_2), as corresponding pairs of basis vectors need not match — in fact, if they did, then it is not hard to see that the strategy would be reduced to a convex combination of 2-dimensional strategies, which would conclude our proof. Nevertheless, by a simple argument we can show that without loss of generality the blocks are simply "shifted": there exists an ordering of $\mathcal{B}_A = \{e_1, \ldots, e_d\}$ such that if the blocks of (A_1, A_2) correspond to pairs $(e_1, e_2), (e_3, e_4), \ldots$ then those of (B_1, B_2) can be seen to correspond to pairs $(e_d, e_1), (e_2, e_3), \ldots$.

The exact form we obtain for the strategies is given in Definition 4 in the Appendix, and gaps in the argument above are filled in the proof of Lemma 5, which can informally be summarized as follows.

Lemma 2 (Lemma 5, informal)**.** *There exists a basis* $(e_1, \ldots, e_d)$ *in which*

- (A_1, A_2, B_3) *(resp.* (B_1, B_2, A_3)*) are jointly block-diagonal.*
- *The blocks corresponding to each of these decompositions are shifted: blocks of* (A_1, A_2, B_3) *correspond to pairs* (e_{2i-1}, e_{2i})*, while blocks of* (A_1, A_2, B_3) *correspond to pairs* (e_{2i}, e_{2i+1})*.*
- *The blocks of* (A_1, A_2) *are of the form* (10), (11) *with coefficients* (c_{2i}, s_{2i})*,* $i = 1, \ldots, d/2$*, while those of* (B_1, B_2) *are of the same form with corresponding coefficients* (c_{2i+1}, s_{2i+1})*,* $i = 0, \ldots, d/2 - 1$*.*

B. Step 2: The value of a strategy in joint normal form

Once we have found a nice basis in which to express all observables appearing in the strategy, it should appear as no surprise that the value of (5) should be easily expressible as a function of the coefficients $(c_i)_{i=1,\ldots,d}$, since these are the only free parameters left in our choice of strategy. In fact, fixing coefficients c_i where i is even, it is not hard to determine the optimal choice of coefficients c_i for odd i. This reduces the size of our problem to the $d/2$ parameters $c_2, \ldots, c_d$. One can then show that the strategy has the following value (cf. Lemma 8 for a more precise statement):

$$\omega = \frac{1}{d}\sum_{i=1}^{d/2} f(c_{2i-1}, c_{2i+1}) + \frac{c_1 - c_{d+1}}{2d} \, , \tag{15}$$

where

$$f(x, y) = \sqrt{(x+y)^2 + 1} + \frac{1}{2}\sqrt{1 - x^2} + \frac{1}{2}\sqrt{1 - y^2} - 2 \, .$$

We have thus rephrased the problem of maximizing $\langle I_{3322} \rangle$ over all strategies in terms of maximizing ω over all admissible coefficients $(c_{2i-1})_{i=1,\ldots,d/2+1}$. To prove our claim, it only remains to prove an upper bound on ω, which can be done using standard analytical techniques provided in the appendix.

Lemma 3. *Let* $c_{2i-1} \in [-1, 1]$*, for* $i = 1, \ldots, d/2 + 1$*. Then the expression* $\omega = \omega(c_i)$ *in* (15) *is upper-bounded by* $\frac{1}{4}$*.*

III. CONCLUSION AND OPEN QUESTIONS

We have provided a concrete example of a simple inequality for which it can be shown that the maximally entangled state of any dimension is not the most nonlocal state. An interesting question, already asked in [33], is whether one can show that optimal violation of the I_{3322} inequality requires a state of infinite dimension. This is strongly suggested by the strategies found numerically by Pál and Vértesi, which, even though they are based on an entangled state which is very far from the maximally entangled state, have a matrix form which is quite similar to the one in Def. 4. Extending our argument to show that Alice and Bob's measurements always have this form, even when they do not use the maximally entangled state, would be a big step towards proving that no finite-dimensional strategy is optimal [3]. This would not only have very interesting consequences for our understanding of Bell inequalities, but also for the optimization of polynomials with non-commutative variables. In particular, it would imply that the SDP hierarchies suggested in [15, 30, 31] only converge in the limit of infinitely many levels, which is an open problem even outside the realm of quantum information.

Acknowledgments

We thank Salman Beigi for interesting discussions, and Oded Regev for helpful comments. TV was supported by ARO Grant W911NF-09-1-0440 and NSF Grant CCF-0905626. SW was supported by the National Research Foundation (Singapore), and the Ministry of Education (Singapore). TV is grateful to CQT, Singapore, for hosting him while part of this work was done.

[1] A. Acin, T. Durt, N. Gisin, and J. I. Latorre. Quantum nonlocality in two three-level systems. *Phys. Rev. A*, 65:052325, 2002.
[2] A. Acin, R. Gill, and N. Gisin. Optimal bell tests do not require maximally entangled states. *Phys. Rev. Lett.*, 95:210402, 2005.
[3] S. Beigi. Personal communication, 2010.
[4] J. S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1:195–200, 1965.
[5] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher. Concentrating partial entanglement by local operations. *Phys. Rev. A*, 53:2046, 1996.
[6] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett*, 70:1895–1899, 1993.
[7] C. H. Bennett, I. Devetak, A. W. Harrow, P. W. Shor, and A. Winter. Quantum Reverse Shannon Theorem. arXiv:0912.5537, December 2009.
[8] C. H. Bennett and S. Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Phys. Rev. Lett.*, 69:2881–2884, 1992.
[9] M. Berta, M. Christandl, and R. Renner. A conceptually simple proof of the quantum reverse shannon theorem. arXiv:0912.3805, 2009.
[10] R Bhatia. *Matrix Analysis*. Graduate Texts in Mathematics. Springer, New York, 1997.
[11] N. Brunner, N. Gisin, and V. Scarani. Entanglement and nonlocality are different resources. *New J. of Physics*, 7:88, 2005.
[12] R. Cleve, P. Høyer, B. Toner, and J. Watrous. Consequences and limits of nonlocal strategies. *Proceedings of the 19th IEEE Con-*

ference on Computational Complexity, pages 236–249, 2004.
[13] D. Collins and N. Gisin. A relevant two qubit Bell inequality inequivalent to the CHSH inequality. *J. Phys. A: Math. Gen.*, 37:1775–1787, 2004.
[14] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu. Bell Inequalities for Arbitrarily High-Dimensional Systems. *Phys. Rev. Lett.*, 88(4):040404, January 2002.
[15] A. C. Doherty, Y. Liang, B. Toner, and S. Wehner. The quantum moment problem and bounds on entangled multi-prover games. In *Proc. IEEE Conference on Computational Complexity*, pages 199–210, 2008.
[16] P. Eberhard. Background level and counter efficiencies requires for a loophole free Einstein-Podolsky-Rosen experiment. *Phys. Rev. A (R)*, 47:747–750, 1993.
[17] A. Ekert. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett*, 67:661–663, 1991.
[18] M. Froissart. Constructive generalization of Bells inequalities. *Il Nuovo Cimento B*, 64(2):241–251, August 1981.
[19] N. Gisin. Bell's inequality holds for all non-product states. *Phys. Lett. A*, 154:201–202, 1991.
[20] P.L. Halmos. Two subspaces. *Trans. Amer. Math. Soc*, 144:381, 1969.
[21] R Horodecki. Violating Bell inequality by mixed ? states: necessary and sufficient condition. *Physics Letters A*, 200(5):340–344, May 1995.
[22] M. Junge and C. Palazuelos. Large violation of bell inequalities with low entanglement. arXiv:1007.3043, 2010.
[23] R. König and S. Wehner. A strong converse for classical channel coding using entangled inputs. *Phys. Rev. Lett.*, 103:070504, 2009.
[24] M. Laurent. Sums of squares, moment matrices and optimization over polynomials. *Emerging Applications of Algebraic Geometry*, 149:157–270, 2009.
[25] Yeong-Cherng Liang, Tamas Vertesi, and Nicolas Brunner. Device-independent bounds on entanglement. arXiv:1012.1513, 2010.
[26] J. Löfberg. Yalmip : A toolbox for modeling and optimization in MATLAB. In *Proc. CACSD Conference*, 2004.
[27] Ll. Masanes. Extremal quantum correlations for N parties with two dichotomic observables per site. quant-ph/0512100, 2005.
[28] D Mayers and A Yao. Quantum Cryptography with Imperfect Apparatus. In *Proceedings of 39th IEEE FOCS*, 1998.
[29] A. A. Methot and V. Scarani. An anomaly of non-locality. *QIC*, 7:157–170, January 2007.
[30] M. Navascues, S. Pironio, and A. Acin. Bounding the set of quantum correlations. *Phys. Rev. Lett.*, 98:010401, 2007.
[31] M. Navascues, S. Pironio, and A. Acin. A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. *New J. of Physics*, 10:073013, 2008.
[32] M. Oliveira. Embezzlement States are Universal for Non-Local Strategies. arXiv:1009.0771, September 2010.
[33] Károly Pál and Tamás Vértesi. Maximal violation of a bipartite three-setting, two-outcome Bell inequality using infinite-dimensional quantum systems. *Phys. Rev. A*, 82(2):022116, August 2010.
[34] P. Parrilo. *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*. PhD thesis, California Institute of Technology, 2000.
[35] P. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Math. Prog. Ser. B*, 96(2):293–320, 2003.
[36] S. Popescu and D. Röhrlich. Generic quantum nonlocality. *Phys. Lett. A*, 166:293–297, 1992.
[37] J. F. Sturm. Using sedumi 1.02, a matlab toolbox for optimization over symmetric cones, 1998.
[38] W. van Dam and P. Hayden. Universal entanglement transformations without communication. *Phys. Rev. A (R)*, 67(6):060302, 2003.
[39] S. Zohren and R. D. Gill. Maximal violation of the collins-gisin-linden-massar-popescu inequality for infinite dimensional states. *Phys. Rev. Lett*, 100:120406, 2008.
[40] The CHSH inequality is extremal in the sense that violation (by a certain state) of any two-setting inequality implies the same state also violates CHSH.
[41] A matrix is tridiagonal if its only non-zero entries are on the main diagonal and the two diagonals right above and under it.

Appendix A: A joint normal form for strategies using the maximally entangled state

The goal of this section is to prove Lemma 5, which shows that any optimal strategy must have a certain simple joint normal form. Before we define it precisely, note that in order for the strategy $\{A_j, B_k\}_{j,k=1,\ldots,3}$ to be optimal, for a fixed choice of $\{B_k\}$ it is necessary that the operators $\{A_j\}$ be chosen so as to maximize

$$\langle\Psi|A_1 \otimes (B_1 + B_2 - B_3)|\Psi\rangle \tag{A1}$$
$$\langle\Psi|A_2 \otimes (B_1 + B_2 + B_3 - \mathsf{id})|\Psi\rangle \tag{A2}$$
$$\langle\Psi|A_3 \otimes (B_2 - B_1)|\Psi\rangle \tag{A3}$$

while for fixed $\{A_j\}$, the $\{B_k\}$ should maximize

$$\langle\Psi|B_1 \otimes (A_1 + A_2 - A_3 - \mathsf{id})|\Psi\rangle \tag{A4}$$
$$\langle\Psi|B_2 \otimes (A_1 + A_2 + A_3 - 2\mathsf{id})|\Psi\rangle \tag{A5}$$
$$\langle\Psi|B_3 \otimes (A_2 - A_1)|\Psi\rangle \tag{A6}$$

Since $|\Psi\rangle$ is the maximally entangled state, for any A and B we have $\langle\Psi|A \otimes B|\Psi\rangle = \frac{1}{d}\mathrm{Tr}(AB^T) =: \langle A, B\rangle$, where $\langle\cdot,\cdot\rangle$ denotes the real Hilbert-Schmidt matrix inner product. To simplify notation, and since the A_j operators always appear on the left of the tensor product (Alice's side), we will argue about A_j^T rather than A_j, omitting the transpose sign. Hence given for instance B_1, B_2 and B_3, the A_1 maximizing (A1) is simply the projector on the positive eigenspace of $B_1 + B_2 - B_3$. In particular, if B_1, B_2 and B_3 have a joint block-diagonalization this will be reflected in $B_1 + B_2 - B_3$ and hence in A_1. This observation, combined with the CS decomposition for a pair of projectors, will let us find a simple joint form for all the A_j and B_k, as explicited in the following definition.

Definition 4. *For any $c \in [-1, 1]$, let $s = \sqrt{1 - c^2}$ and define the 2-dimensional projectors*

$$P_1(c) := \frac{1}{2}\begin{pmatrix} 1-c & -s \\ -s & 1+c \end{pmatrix}, \tag{A7}$$

$$P_2(c) := \frac{1}{2}\begin{pmatrix} 1-c & s \\ s & 1+c \end{pmatrix}, \tag{A8}$$

$$P_3 := \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \tag{A9}$$

We say that d-dimensional projectors $\{A_j, B_k\}$ are in joint normal form *if there exists a basis of $\mathbb{C}^d$ such that either*

- *For even dimensions d, there exist reals $c_i \in [-1,1]$, $i = 1,\dots,d+1$ such that:*
  - *A_1 (resp. A_2) is block-diagonal with blocks $L_1^{2i} = P_1(c_{2i})$ (resp. $L_2^{2i} = P_2(c_{2i})$), $i = 1\dots d/2$*
  - *B_3 is block-diagonal with blocks all identical to P_3.*
  - *B_1 (resp. B_2) is block-diagonal, with the first block R_1^1 (resp. R_2^1) one-dimensional equal to $\left(\frac{1-c_1}{2}\right)$, the following $d/2-1$ blocks $R_1^{2i+1} = P_1(-c_{2i+1})$ (resp. $R_2^{2i+1} = P_2(-c_{2i+1})$), $i = 1\dots d/2-1$, and the last block $R_1^{d+1} = \left(\frac{1-c_{d+1}}{2}\right)$ (resp. $R_2^{d+1} = \left(\frac{1-c_{d+1}}{2}\right)$).*
  - *A_3 is block-diagonal with its first block one-dimensional equal to (1), the following blocks all identical to P_3, and the last block one-dimensional equal to (1).*
- *For odd dimensions d, there exist reals $c_i \in [-1,1]$, $i = 1,\dots,d+1$ such that:*
  - *A_1 (resp. A_2) is block-diagonal with $(d-1)/2$ 2-dimensional blocks $L_1^{2i} = P_1(c_{2i})$ (resp. $L_2^{2i} = P_2(c_{2i})$), $i = 1\dots(d-1)/2$, and a final 1-dimensional block $L_1^{d+1} = \left(\frac{1-c_{d+1}}{2}\right)$ (resp. $L_2^{d+1} = \left(\frac{1-c_{d+1}}{2}\right)$),*
  - *B_3 is block-diagonal with the first $(d-1)/2$ blocks all identical to P_3, and the last one 1-dimensional equal to (1).*
  - *B_1 (resp. B_2) is block-diagonal with an initial one-dimensional block $R_1^1 = \left(\frac{1-c_1}{2}\right)$ (resp. $R_2^1 = \left(\frac{1-c_1}{2}\right)$) and the following $(d-1)/2$ blocks $R_1^{2i+1} = P_1(-c_{2i+1})$ (resp. $R_2^{2i+1} = P_2(-c_{2i+1})$, $i = 1\dots(d-1)/2$.*
  - *A_3 is block-diagonal, with the first 1-dimensional block equal to (1), and all following blocks identical to P_3.*

Or the same as above, but with the roles of $\{A_1, A_2, B_3\}$ and $\{B_1, B_2, A_3\}$ exchanged.

The main lemma of this section is the following:

Lemma 5. *Suppose A_1, A_2, A_3 and B_1, B_2, B_3 are six d-dimensional projectors achieving the maximum of* (5) *over all d-dimensional strategies using the maximally entangled state $|\Psi\rangle$. Then there is a $d' \leq d$, and a d'-dimensional strategy in joint normal form which achieves a value at least as large as that of $\{A_j, B_k\}$.*

Proof. Apply the CS decomposition to A_1 and A_2, resulting in a joint block-diagonalization basis $\{|e_i\rangle\}_i$, and to B_1 and B_2, resulting in $\{|f_i\rangle\}_i$. We first show that we may take $\{|e_i\rangle\} = \{|f_i\rangle\}$ without lowering the value of the strategy.

As we already noted, the optimal choice for A_3 (resp. B_3) is the projector on the positive eigenspace of $B_2 - B_1$ (resp. $A_2 - A_1$). This implies that the value of (A3) does not depend on the choice of basis $\{|e_i\rangle\}$, but only on the eigenvalues of $B_2 - B_1$. Hence of all the terms in (5), the only ones whose value depends on the choice of the bases $\{|e_i\rangle\}$ and $\{|f_i\rangle\}$ can be grouped together as $\langle\Psi|(A_1 + A_2)\otimes(B_1 + B_2)|\Psi\rangle$.

Claim 6. *Let $|\Psi\rangle = \frac{1}{\sqrt{d}}\sum_i |i\rangle|i\rangle$, and $A = \sum_i \alpha_i|u_i\rangle\langle u_i|$ and $B = \sum_i \beta_i|v_i\rangle\langle v_i|$ positive. Then the expression $\langle\Psi|A\otimes B|\Psi\rangle$ is maximized when the $|u_i\rangle, |v_i\rangle$ are a permutation of the Schmidt basis of $|\Psi\rangle$.*

Proof. For any two matrices A, B we have $\langle\Psi|A\otimes B|\Psi\rangle = \frac{1}{d}\mathrm{Tr}(A^T B)$. Note that A^T has the same eigenvalues as A. We then have by [23, Lemma IV.11] that there exists a permutation $\pi \in S_d$ such that

$$\frac{1}{d}\mathrm{tr}(A^T B) \leq \sum_{j=1}^{d} \lambda^A_{\pi(j)}\lambda^B_j \,, \tag{A10}$$

where $\lambda_1^A, \dots, \lambda_d^A$ and $\lambda_1^B, \dots, \lambda_d^B$ are the eigenvalues of A and B respectively. □

Given our specific choice of basis for the block-diagonalization, we have that $A_1 + A_2$ (resp. $B_1 + B_2$) is diagonal in the basis $\{|e_i\rangle\}$ (resp. $\{|f_i\rangle\}$), hence Claim 6 shows that these two bases may be taken equal (up to permutation) without lowering the value of the strategy.

We call a strategy given by projectors $\{A_j, B_k\}_{j,k}$ *irreducible* if it cannot be decomposed as a direct sum of lower-dimensional strategies. We show that any irreducible strategy has the form described in Definition 4.

Claim 7. *Suppose $\{A_j, B_j\}$ is irreducible. If d is even, then either all blocks of the joint decomposition of $\{A_1, A_2, B_3\}$ and $\{B_1, B_2, A_3\}$ are two-dimensional, or $\{A_1, A_2, B_3\}$ have exactly two 1-dimensional blocks and $\{B_1, B_2, A_3\}$ none (or vice-versa). If d is odd, then each of $\{A_1, A_2, B_3\}$ and $\{B_1, B_2, A_3\}$ have exactly one common 1-dimensional block.*

Proof. We treat the case of even dimension, the odd-dimensional case being analogous. Reason by contradiction and first assume e.g. that $\{A_1, A_2, B_3\}$ each have more than two 1-dimensional blocks in their joint block-diagonalization. We show that there is a non-trivial subspace stabilized by all operators $\{A_j, B_k\}$, contradicting the strategy's irreducibility.

Let $|e_1\rangle$ be the vector corresponding to a one-dimensional block of $\{A_1, A_2, B_3\}$. Since the $\{|f_i\rangle\}$ are a permutation of $\{|e_i\rangle\}$, there exists an i_1 such that $|f_{i_1}\rangle = |e_1\rangle$. There are two possibilities for $|f_{i_1}\rangle$: either it is a joint eigenvector of B_1, B_2 and A_3 (i.e. it corresponds to a one-dimensional block in their joint block-diagonalization), or there exists an index i_2 such that $\mathrm{Span}\{|f_{i_1}\rangle, |f_{i_2}\rangle\}$ is left invariant by the action of B_1, B_2 and A_3 (i.e. it corresponds to a two-dimensional block). In the first case we have already found a strict subspace $\mathrm{Span}\{|e_1\rangle\}$ stabilized by all $\{A_j, B_k\}$. In the second case we can iterate this procedure, assuming without loss of generality that $|e_2\rangle = |f_{i_2}\rangle$. There are again two cases: either $|e_2\rangle$ corresponds to a 1-dimensional block of $\{A_1, A_2, B_3\}$, in which case $\mathrm{Span}\{|e_1\rangle, |e_2\rangle\}$ is a non-trivial stable subspace, or there is a vector $|e_3\rangle$ such that $(|e_2\rangle, |e_3\rangle)$ corresponds to a 2-dimensional block of $\{A_1, A_2, B_3\}$. We will then find an i_3 such that $|f_{i_3}\rangle = |e_3\rangle$, and so on.

In all cases, the process must end as soon as one of the vectors $|e_k\rangle$ encountered corresponds to a 1-dimensional block of $\{A_1, A_2, B_3\}$. Given our assumption that there were three or more such blocks, we have found a strict subspace stabilized by all $\{A_j, B_k\}$, contradicting the irreducibility assumption. □

As a consequence of Claim 7, we can block-diagonalize the pair of projectors (A_1, A_2) with blocks

$$L_1^{2i} = \frac{1}{2}\begin{pmatrix} 1-c_{2i} & -s_{2i} \\ -s_{2i} & 1+c_{2i} \end{pmatrix} , \tag{A11}$$

$$L_2^{2i} = \frac{1}{2}\begin{pmatrix} 1-c_{2i} & s_{2i} \\ s_{2i} & 1+c_{2i} \end{pmatrix} , \tag{A12}$$

where $c_{2i} \in (-1, 1)$ and $s_{2i} = \sqrt{1-c_{2i}^2}$, together possibly with an initial and final 1-dimensional blocks, depending on the parity of the dimension.

In the definition of a normal form we also require the one-dimensional blocks to have the same coefficients for both A_1 and A_2, which is is easily seen to hold without loss of generality from the optimality of the strategy $\{A_j, B_k\}$. Indeed, let i be the index of such a block, corresponding to vector $|e_i\rangle$; A_1 and A_2 are necessarily chosen so as to maximize the value of (A1) and (A2) respectively, and the coefficient in front of $(A_1)_{i,i}$ and $(A_2)_{i,i}$ will be the same in both equations, so that the optimal choice is the same. Similarly, the matrices (B_1, B_2) can be block-diagonalized with blocks:

$$R_1^{2i+1} = \frac{1}{2}\begin{pmatrix} 1+c_{2i+1} & -s_{2i+1} \\ -s_{2i+1} & 1-c_{2i+1} \end{pmatrix} , \tag{A13}$$

$$R_2^{2i+1} = \frac{1}{2}\begin{pmatrix} 1+c_{2i+1} & s_{2i+1} \\ s_{2i+1} & 1-c_{2i+1} \end{pmatrix} . \tag{A14}$$

Finally, it is easy to infer from (A3) (resp. (A6)) the necessary form of A_3 (resp. B_3): indeed, it is simply the projector on the positive eigenspace of $B_2 - B_1$ (resp. $A_2 - A_1$), which is a block P_3 whenever B_1, B_2 (resp. A_1, A_2) have a common 2-dimensional block, and a block (1) whenever B_1, B_2 (resp. A_1, A_2) have a common one-dimensional block. □

Appendix B: The value of a strategy in joint normal form

In this section we derive an expression for the value obtained in (5) for any strategy in joint normal form (Lemma 8), and then show how it can be upper-bounded by analytical techniques (Lemma 10).

Lemma 8. *Suppose $\{A_j, B_k\}$ is a strategy in joint normal form, described by a certain block structure and corresponding sequence of coefficients c_i. Then the value of* (5) *for this strategy for even dimensions d is given by*

$$\omega = \frac{1}{d}\sum_{i=1}^{d/2} f(c_{2i-1}, c_{2i+1}) + \frac{c_1 - c_{d+1}}{2d} , \tag{B1}$$

and for odd dimension d by

$$\omega = \frac{1}{d}\sum_{i=1}^{(d-1)/2} f(c_{2i-1}, c_{2i+1}) + \frac{1}{d}\Big(c_d\, c_{d+1} + \frac{c_1 - c_{d+1}}{2} - 1 + \frac{1}{2}\sqrt{1-c_d^2}\Big) , \tag{B2}$$

where

$$f(x,y) = \sqrt{(x+y)^2+1} + \frac{1}{2}\sqrt{1-x^2} + \frac{1}{2}\sqrt{1-y^2} - 2 . \tag{B3}$$

Proof. We treat the cases of even and odd dimension separately.

a. d even. In that case we know that the block-diagonalization of either $\{A_1, A_2, B_3\}$ or $\{B_1, B_2, A_3\}$ contains exactly two 1-dimensional blocks, while the other contains none. We assume that $\{B_1, B_2, A_3\}$ has no 1-dimensional blocks; the other case is treated symmetrically. In this case we can write

$$A_2 = \frac{1}{2}\begin{pmatrix} 1-c_2 & s_2 & 0 & 0 & \cdots \\ s_2 & 1+c_2 & 0 & 0 & \cdots \\ 0 & 0 & 1-c_4 & s_4 & \cdots \\ 0 & 0 & s_4 & 1+c_4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} , \tag{B4}$$

$$B_2 = \frac{1}{2}\begin{pmatrix} 1-c_1 & 0 & 0 & 0 & \cdots \\ 0 & 1+c_3 & s_3 & 0 & \cdots \\ 0 & s_3 & 1-c_3 & 0 & \cdots \\ 0 & 0 & 0 & 1+c_5 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} , \tag{B5}$$

where A_1 and B_1 are identical to A_2 and B_2 respectively but have their off-diagonal elements negated, and $c_1, c_{d+1} \in \{-1, 1\}$.

Fixing the coefficients of B_1 and B_2, we can derive constraints on those of A_1 and A_2 from the constraint that they should be chosen so as to maximize (A1) and (A2). The two equations are similar; let's look at (A2). Its value can be calculated as

$$\begin{aligned}
&\frac{1}{d}\sum_{i,j}(A_2)_{i,j}\big((B_1)_{i,j} + (B_2)_{i,j} + (B_3)_{i,j} - \delta_{i,j}\big) \\
&= \frac{1}{d}\sum_{i=1}^{d/2}\Big(\frac{1}{2}\big(1-c_{2i}\big)\big((1-c_{2i-1}) + \frac{1}{2} - 1\big) \\
&\quad + \frac{1}{2}\big(1+c_{2i}\big)\big((1+c_{2i+1}) + \frac{1}{2} - 1\big)\Big) \\
&\quad + \frac{1}{d}\sum_{i=1}^{d/2}\frac{s_{2i}}{2} \\
&= \frac{1}{d}\sum_{i=1}^{d/2}\Big(\frac{1-c_{2i}}{2}\Big(\frac{1}{2} - c_{2i-1}\Big) \\
&\quad + \frac{1+c_{2i}}{2}\Big(\frac{1}{2} + c_{2i+1}\Big) + \frac{s_{2i}}{2}\Big)
\end{aligned} \tag{B6}$$

Setting $\tau_{2i} = (c_{2i-1} + c_{2i+1})/2$, for a fixed c_{2i-1}, c_{2i+1} the choice of c_{2i} which maximizes (B6) is $c_{2i} = 2\tau_{2i}\,(4\tau_{2i}^2 + 1)^{-1/2}$, which gives a value of $\frac{1}{d}\sum_{i=1}^{d/2}\sqrt{4\tau_{2i}^2+1}/2+1/2+(c_{2i+1}-c_{2i-1})/2$ for (A2). (A1) is maximized for the same choice of coefficients, and has exactly the same value. Concerning (A3), we find that its value is simply

$$\frac{1}{d}\sum_{i,j}(A_3)_{i,j}\big((B_2)_{i,j}-(B_1)_{i,j}\big) = \frac{1}{d}\sum_{i=1}^{d/2-1} s_{2i+1}\,. \tag{B7}$$

Combining (A1),(A2) and (A3), and subtracting $(1/d)(\mathrm{Tr}(B_1) + 2\mathrm{Tr}(B_2))$, we obtain the value of (5), which is thus

$$\begin{aligned}\omega = \frac{1}{d}\sum_{i=1}^{d/2}&\left(\sqrt{(c_{2i-1}+c_{2i+1})^2+1}+1\right)\\ &+\frac{1}{d}(c_{d+1}-c_1)+\frac{1}{d}\sum_{i=1}^{d/2-1}\sqrt{1-c_{2i+1}^2}\\ &-3\left(\frac{1}{2}+\frac{c_{d+1}-c_1}{2d}\right)\,,\end{aligned} \tag{B8}$$

where we replaced $s_{2i+1} = \sqrt{1-c_{2i+1}^2}$. Using the definition of f, this can be re-written as

$$\omega = \frac{1}{d}\sum_{i=1}^{d/2} f(c_{2i-1},c_{2i+1}) + \frac{c_1 - c_{d+1}}{2d}$$

b. d odd. In that case, each of $\{A_1, A_2, B_3\}$ and $\{B_1, B_2, A_3\}$ must have a 1-dimensional-block in their joint block-diagonalization; say that the one for $\{A_1, A_2, B_3\}$ is the last block while the one for $\{B_1, B_2, A_3\}$ is the first block. We can proceed exactly as above to evaluate the value of this strategy, under the condition that it is optimal and hence maximizes (A1)-(A3), which lets us express the even coefficients c_{2i} as a function of the odd ones c_{2i+1}. Omitting a few calculations very similar to the ones we performed in the even-dimensional case, we obtain that the value of this solution is

$$\begin{aligned}\omega &= \frac{1}{d}\sum_{i=1}^{(d-1)/2}\left(\sqrt{(c_{2i-1}+c_{2i+1})^2+1}+1\right)+\frac{1}{d}\big(c_d - c_1\big)\\ &\quad+\frac{1}{d}\big(1-c_{d+1}\big)\big(\frac{1}{2}-c_d\big)\\ &\quad+\frac{1}{d}\sum_{i=1}^{(d-1)/2}\sqrt{1-c_{2i+1}^2}-3\left(\frac{1}{2}-\frac{c_1}{2d}\right) \qquad\qquad (\mathrm{B9})\\ &=\frac{1}{d}\sum_{i=1}^{(d-1)/2}(a_i-2) \qquad\qquad (\mathrm{B10})\\ &\quad+\frac{1}{d}\Big(c_d\,c_{d+1}+\frac{c_1-c_{d+1}}{2}-1+\frac{1}{2}\sqrt{1-c_d^2}\Big)\,.\end{aligned}$$

□

It now remains to bound ω. The following claim, proven in Section C, will be useful.

Claim 9. *Let $f(x,y) = \sqrt{(x+y)^2+1} + \sqrt{1-x^2}/2 + \sqrt{1-y^2}/2 - 2$ be defined on $[-1,1]^2$. Then*

1. *The maximum of $f(a,b) + f(b,c)$ over all $a,b,c \in [-1,1]^2$ such that $a+b \geq 0$ and $b+c \leq 0$ is less than $.244$.*
2. *The maximum of $f(1,b)+f(b,c)$ over all $b,c \in [-1,1]^2$ such that $1+b \geq 0$ and $b+c \leq 0$ is less than $.103$.*
3. *The maximum of $f(a,1)$ over all $a \in [-1,1]$ is less than $.368$.*

Lemma 10. *Let $c_i \in [-1,1]$, for $i = 1\ldots d+1$. Then the expression $\omega = \omega(c_i)$ in both* (B1) *and* (B2) *is upper-bounded by $\frac{1}{4}$.*

Proof. First note that the maximum value of the expression $c_d\,c_{d+1} + \frac{c_1-c_{d+1}}{2} - 1 + \frac{1}{2}\sqrt{1-c_d^2}$ over all $c_1, c_{d+1} \in \{-1,1\}$ and $c_d \in [-1,1]$ is less than $1/4$, hence (B2) is always lower than (B1). Hence it is sufficient to show that $\omega = \frac{1}{d}\sum_{i=1}^{d/2} f(c_{2i-1},c_{2i+1}) + \frac{c_1-c_{d+1}}{2d}$ is upper-bounded by $1/4$, for any even d and $(c_2,\ldots,c_d) \in [-1,1]^{d-1}$ and $c_1, c_{d+1} \in \{-1,1\}$.

It is easy to verify that $f(x,y) \leq 1/2$ on the square $(x,y) \in [-1,1]^2$. Unfortunately, the extra term $\frac{c_1-c_{d+1}}{2d}$ potentially induces an additive $1/d$, so that it is not so immediate to bound ω. Note that we can assume that $c_1 = 1$ and $c_{d+1} = -1$, since otherwise the bound follows trivially from the upper-bound on $f(x,y) \leq 1/2$.

Given the value of c_1 and c_{d+1}, there must exist an i such that $c_{2i-1} + c_{2i+1} \geq 0$ and $c_{2i+1} + c_{2i+3} \leq 0$; let i_0 be the first such i. We distinguish four cases, depending on the value of i_0.

- If $d = 4$, one gets that $f(1,c_3) + f(c_3,-1) < 0$. Adding $(c_1 - c_{d+1})/8$, one can see that $\omega < 1/4$. We assume $d > 4$ for the remaining cases.
- If $i_0 = 1$, we can use the second bound in Claim 9 to bound $f(c_1,c_3) + f(c_3,c_5)$ by $.103$, since $c_1 = 1$. In this case the value of $f(c_1,c_3) + f(c_3,c_5) + f(c_{d-1},c_{d+1})$ is at most $.103 + .368 < .5$. Adding $1 = (c_1 - c_{d+1})/2$ and dividing by d, we see that $\omega < 1/4$ irrespective of the value of the other c_i (recall that $f(x,y) \leq 1/4$ for all (x,y)).
- If $i_0 = d/2-1$, the same bound can be obtained by symmetry.
- Otherwise $1 < i_0 < d/2-1$, in which case by using the first and last bounds from Claim 9 we see that the value of $f(c_1,c_3)+f(c_{2i-1},c_{2i+1})+f(c_{2i+1},c_{2i+3})+f(c_{d-1},c_{d+1})$ is at most $.244 + 2\cdot .368 < 1$. Again adding $1 = (c_1 - c_{d+1})/2$ and dividing by d, one sees that $\omega < 1/4$ irrespective of the value of the other c_i.

□

Appendix C: Details of Claim 9

We now provide the details of Claim 9. To find the claimed upper bounds we use a well-established optimization technique based on a hierarchy of semidefinite programs (SDPs) backed by the real Positivstellensatz [34, 35]. More specifically, if t denotes a claimed upper bound, our goal will be to show that for any variables a, b and c satisfying the constraints we have $t - h(a,b,c) \geq 0$, where $h(a,b,c)$ denotes the function we wish to optimize in case 1, 2 or 3. To this end, we will first rewrite any terms involving $\sqrt{\cdot}$ in the function $h(a,b,c)$ in terms of additional variables. Second, we will use polynomial optimization techniques from [34, 35] to obtain the bound t. This is exactly analogous to the techniques established in quantum information to obtain bounds on quantum violation of Bell inequalities [15, 30, 31].

We would like to emphasize that whereas semidefinite programming, as for example performed in Matlab, is a numerical technique, if a bound t_ℓ is obtained at level ℓ of the SDP hierarchy then it is in principle possible to extract an *analytical* proof that t_ℓ is an upper-bound on the corresponding expression h from the numerics. That is, we do not rely on any heuristic optimization methods that are not guaranteed to provide a rigorous bound.

1. Case 3

For completeness, we provide a brief informal sketch of this method for case 3; details can be found in [34, 35], or in the dual view of the SDP, as explained in this survey [24]. First of all, substituting

$$x^2 := (a+1)^2 + 1 = a^2 + 2a + 2 \,, \tag{C1}$$
$$z^2 := 1 - a^2 \,, \tag{C2}$$

our goal of showing that $t = 0.368$ is an upper bound to $f(a,1)$ can be restated as showing that

$$\begin{array}{ll} \text{we have} & t \geq x + \frac{1}{2}z - 2 \\ \text{whenever} & x^2 = a^2 + 2a + 1 \\ & z^2 = 1 - a^2 \\ & -1 \leq a \leq 1 \,. \end{array}$$

For simplicity, we will without loss of generality ignore the last constraint. Now note that if we were able to find polynomials t_1 and t_2 in variables x, z, and a such that

$$\begin{aligned} p := t - \left(x + \frac{1}{2}z - 2\right) - t_1(a^2 + 2a + 2 - x^2) \\ - t_2(1 - a^2 - z^2) = s_0 \,, \end{aligned} \tag{C3}$$

where s_0 is a polynomial in x, z and a which is a sum of squares, then for any variables satisfying the desired constraints $t - (x + \frac{1}{2}z - 2) \geq 0$ since s_0 is always positive. Our goal can thus be rephrased as searching for suitable polynomials t_1 and t_2 such that we can rewrite the resulting polynomial as a sum of squares. Very intuitively, level ℓ of the SDP hierarchy searches for such polynomials up to degree 2ℓ by searching for a matrix Q_ℓ such that $Q_\ell \geq 0$ and for $v_\ell = (1, a, x, z, \ldots)$ being the vector of all possible monomials up to degree ℓ where we have $v_\ell^\dagger Q_\ell v_\ell = p$. To convince ourselves, note that this means we search for $Q_\ell \geq 0$ such that

$$\begin{aligned} t - \left(x + \frac{1}{2}z - 1\right) = v_\ell^\dagger Q_\ell v_\ell \\ + t_1(a^2 + 2a + 2 - x^2) + t_2(1 - a^2 - z^2) \end{aligned} \tag{C4}$$

and thus for variables satisfying the constraints

$$t - \left(x + \frac{1}{2}z - 1\right) = v_\ell^\dagger Q_\ell v_\ell \,, \tag{C5}$$

which is clearly positive. The actual sums of squares polynomials s_0 can be obtained from Q by diagonalizing $Q = U^\dagger D U$ where U is unitary and D is a diagonal matrix. Since D only has positive entries ($Q \geq 0$), we obtain that $s_0 = \sum_j d_j (Uv)_j^\dagger (Uv)_j$ is indeed a sum of squares.

It turns out that for case 3, we can already find such a matrix Q at level $\ell = 0$ of the SDP, that is, $t_1, t_2 \in \mathbb{R}$ are simply scalars. To see how this works explicitly, let us first rewrite the polynomials above in terms of matrices. Let

$$M_0 := \begin{pmatrix} -2 & \frac{1}{2} & \frac{1}{4} & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{4} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \,, \tag{C6}$$

$$M_1 := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \,, \tag{C7}$$

$$M_2 := \begin{pmatrix} 2 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \,, \tag{C8}$$

$$T := \begin{pmatrix} t & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \,. \tag{C9}$$

Clearly, for

$$v := (1\ x\ z\ a)^T \,, \tag{C10}$$

we have

$$v^\dagger M_0 v = x + \frac{1}{2}z - 2 \,, \tag{C11}$$
$$v^\dagger M_1 v = 1 - a^2 - z^2 \,, \tag{C12}$$
$$v^\dagger M_2 v = a^2 + 2a + 2 - x^2 \,. \tag{C13}$$

From the numerical solutions obtained by Matlab with SeDuMi [37] and YALMIP [26], we can guess an analytical solution given by

$$t_1 = 0.51 \tag{C14}$$
$$t_2 = 0.24 \tag{C15}$$
$$t = 0.368 \tag{C16}$$

for which we can easily verify that

$$Q_0 := S - M_0 - t_1 M_1 - t_2 M_2 \geq 0 \;, \tag{C17}$$

which concludes our claim.

2. Cases 1 and 2

The bounds for cases 1 and 2 are obtained analogously. The only difference is that we have to deal with more variables. Again, we first introduce auxiliary variables to eliminate terms containing $\sqrt{\cdot}$. We then search for suitable polynomials like t_1 and t_2 above. Unlike for the simple case 3, the desired bounds are not obtained at level $\ell = 0$ of the hierarchy. However, they are already found at level $\ell = 1$, and an analytical solution can again be extracted. Yet, since at level $\ell = 1$ we observe polynomials of degree up to 2 in both the original and the auxiliary variables (in total 6 for case 2, and 8 for case 1) the resulting problem is already rather large (involving matrices of size 82×82 for case 1). We do not include these matrices here, but the Matlab scripts that can be used to extract the analytical values are available upon request.