Robust and Versatile Black-Box Certification of Quantum Devices

Tzyh Haur Yang,[1] Tamás Vértesi,[2] Jean-Daniel Bancal,[1] Valerio Scarani,[1,3] and Miguel Navascués[4]
[1]*Centre for Quantum Technologies, National University of Singapore, 3 Science drive 2, Singapore 117543, Singapore*
[2]*Institute for Nuclear Research, Hungarian Academy of Sciences, P.O. Box 51, H-4001 Debrecen, Hungary*
[3]*Department of Physics, National University of Singapore, 2 Science drive 3, Singapore 117542, Singapore*
[4]*School of Physics, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, United Kingdom*
(Received 10 January 2014; published 22 July 2014)

Self-testing refers to the fact that, in some quantum devices, both states and measurements can be assessed in a black-box scenario, on the sole basis of the observed statistics, i.e., without reference to any prior device calibration. Only a few examples of self-testing are known, and they just provide nontrivial assessment for devices performing unrealistically close to the ideal case. We overcome these difficulties by approaching self-testing with the semidefinite programing hierarchy for the characterization of quantum correlations. This allows us to improve dramatically the robustness of previous self-testing schemes; e.g., we show that a Clauser-Horne-Shimony-Holt violation larger than 2.57 certifies a singlet fidelity of more than 70%. In addition, the versatility of the tool brings about self-testing of hitherto impossible cases, such as the robust self-testing of nonmaximally entangled two-qutrit states in the Collins-Gisin-Linden-Massar-Popescu scenario.

DOI: 10.1103/PhysRevLett.113.040401 PACS numbers: 03.65.Ta, 03.67.Mn

Introduction.—The validation and certification of sources and measurement apparatuses constitute a fundamental step of science and technology. One does not buy the elements to set up an experiment without first assessing their quality, and one should not make claims about the final results of an experiment without several checks. Usually, a variety of assumptions goes into these procedures. For instance, the certification of a device often depends on the fact that other devices are properly calibrated [1]. In the last few years, it has been noticed that tasks like quantum key distribution [2] and random number generation [3,4] can be validated based only on minimal assumptions and on the statistics observed *a posteriori*. The idea consists in looking for statistics that violate Bell inequalities [5]; the minimal assumptions that go into this so-called device-independent assessment are essentially no signaling (which could, in principle, be guaranteed by putting a sufficient distance between the devices) and measurement independence (i.e., the possibility of performing different measurements on the same setup, a cornerstone of the scientific method) [6,7].

Rather than certifying that some device can accomplish a task, one may want to certify the device itself, which in turn would provide certification for any possible further task one may want to perform with it. For instance, if the device is a source, this would amount to performing a "blind tomography" where measurement devices are treated as black boxes. It has long been known that this is possible in some specific and ideal cases. Famously, if the Clauser-Horne-Shimony-Holt (CHSH) inequality [8] is violated at its maximal value $2\sqrt{2}$, the devices are certified to be performing complementary measurements on two effective qubits in the maximally entangled state [9–11]. Another criterion that certifies the same state and measurements was put forward by Mayers and Yao, who called the whole task self-testing of quantum apparatuses [12].

In addition to being tailored for a two-qubit singlet, these pioneering works are unapplicable to real-world devices because they only discuss the statistics of the ideal case. A first step towards the resolution of this issue was taken when several self-testing schemes were shown to be "robust" (or "rigid") [13–16]; the most advanced of these results applies to a multiple-copy scenario and certifies the state as a resource for universal quantum computation [17]. Despite the name, however, these results tolerate only tiny deviations from the ideal case. Take again the certification of the two-qubit singlet based on the CHSH inequality: even for the largest reported experimental violation, which is 2.827 ± 0.0017 [18], i.e., only 0.1% away from the ideal value, none of the robust self-testing approaches quoted above provide a nontrivial bound on the singlet fidelity.

One may surmise that this could be an intrinsic limitation on the ambitious task of self-testing. Here, we show that this is not the case: we demonstrate that a CHSH violation of 2.827 is only compatible with a singlet fidelity larger than 99.83%. This real-life robustness is only one of the benefits of the method that we introduce. Indeed, our approach formalizes the idea of swapping black boxes with trusted systems [12] with the semidefinite characterization of quantum correlations [19], which makes it especially versatile. We demonstrate this explicitly with several examples, all of which are robust. Notably, we describe the self-testing of qutrit states with ternary-outcome measurements, which would not be possible with previous techniques.

In most self-testing works, the assumption is made that the tested devices behave independently and in an identical way (i.i.d.) over the runs. This assumption may sound problematic, as it may fail in real situations (e.g., if a source is drifting). Fortunately, tools have been developed to deal with the general case of Bell-based tests where each realization of the box can be different from the previous one and may even depend on all previous operations effected on the system [20–22]. With these tools, the results obtained with i.i.d. hold true in the general case, in the asymptotic limit of infinitely many runs. In this Letter, we work only in that limit, so we take i.i.d. for granted in the rest of the Letter.

For clarity of presentation, we now introduce our method with the basic example of two-qubit singlet state certification via the CHSH inequality. A few other applications are discussed in the remainder of the Letter, and many more are left for future work.

Bound on the singlet fidelity from CHSH.—Let us consider a bipartite experiment with binary inputs $x, y \in \{0, 1\}$ and binary outputs $a, b \in \{0, 1\}$. After querying the boxes a large number of times, one can reconstruct the measurement statistics $P(ab|xy)$; the CHSH inequality is violated if $\mathcal{B}_{\mathrm{CHSH}} = \sum_{abxy}(-1)^{a+b+xy}P(ab|xy) > 2$ [23]. If a violation is observed, the measured state must be entangled, and it must even be a maximally entangled singlet state $|\psi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$ if the violation is maximal. Our goal is to quantify how far from the singlet the state can be, in terms of fidelity, when the violation is not maximal. Since nothing guarantees that the state in the boxes is a two-qubit state, one must clarify what the fidelity with the singlet means at all. The idea of self-testing consists in swapping part of the content of the black boxes into a trusted system (in this case, two qubits) initially prepared in a suitable dummy state. The singlet fidelity of the final two-qubit state is then well defined.

Specifically, let the trusted auxiliary qubits A' and B' be prepared in the state $|0\rangle$. Then, some local unitaries $\mathcal{S}_{AA'}$ and $\mathcal{S}_{BB'}$ are applied between these trusted systems and their respective boxes, as shown in Fig. 1. Such hypothetical operations leave the trusted systems in the state

$$\rho_{\mathrm{swap}} = \mathrm{tr}_{AB}[\mathcal{S}\rho_{AB} \otimes |00\rangle\langle 00|_{A'B'}\mathcal{S}^\dagger], \tag{1}$$

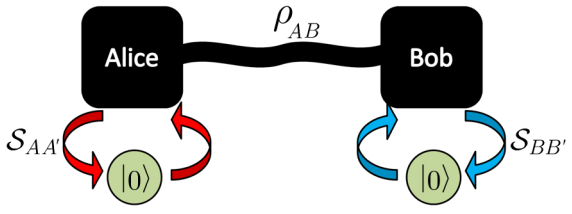

FIG. 1 (color online). The swap concept: Characteristics of black boxes are assessed by considering the effect of swap operations between these black boxes and trusted systems (initialized in the state $|0\rangle$ here).

where $\mathcal{S} = \mathcal{S}_{AA'} \otimes \mathcal{S}_{BB'}$. This operation is a local isometry from the black box to the trusted space, as usually considered in self-testing. One wants to choose $\mathcal{S}$ such that $F = \langle\psi^-|\rho_{\mathrm{swap}}|\psi^-\rangle$ is large, possibly maximal.

It is crucial to stress that this isometry is the virtual procedure that allows one to define a figure of merit, not a procedure that must be implemented in the lab for the certification to be possible. All that needs to be done in the lab is to collect the data that lead to reconstructing $P(ab|xy)$. Therefore, the alleged swap operation $\mathcal{S}$ itself must be defined, and its performance evaluated, from the observed statistics and the belief that whatever happens can be described within the framework of quantum theory. The latter tells us that, to any input x of Alice, there corresponds in the box one Hermitian operator E_a^x for each outcome a, which can be taken as a projector since the dimension of the system being measured is not restricted. The same holds for Bob. Based on these existing projectors, it is convenient to define the Hermitian and unitary operators $A_x = \mathrm{E}_0^x - \mathrm{E}_1^x$ and $B_y = \mathrm{F}_0^y - \mathrm{F}_1^y$. Also, we describe the ideal state as

$$|\bar{\psi}\rangle = \cos\left(\frac{\pi}{8}\right)|\phi^+\rangle + \sin\left(\frac{\pi}{8}\right)|\psi^+\rangle, \tag{2}$$

which is maximally entangled and therefore equivalent to $|\psi^-\rangle$ up to local unitaries. This is chosen for convenience of notation since this state achieves $\mathcal{B}_{\mathrm{CHSH}} = 2\sqrt{2}$ for the operators

$$\bar{A}_0 = \bar{B}_0 = \sigma_z, \qquad \bar{A}_1 = \bar{B}_1 = \sigma_x. \tag{3}$$

All the framework is set. In order to guess a good construction for $\mathcal{S}$, we get inspiration from the ideal case. If the system in each box were indeed a qubit, the swap operations could be realized by combining three CNOT gates [24]. Further, using Eq. (3), the CNOT that has A as target and A' as control can be written as $\bar{U}_{AA'} = \mathbb{1} \otimes |0\rangle\langle 0| + \bar{A}_1 \otimes |1\rangle\langle 1|$; the CNOT with reversed roles can be written as $\bar{V}_{AA'} = (\mathbb{1} + \bar{A}_0)/2 \otimes \mathbb{1} + ((\mathbb{1} - \bar{A}_0)/2) \otimes \sigma_x$. Having noticed this, for the untrusted case, we can tentatively define

$$\mathcal{S}_{AA'} = U_{AA'}V_{AA'}U_{AA'} \tag{4}$$

with

$$U_{AA'} = \mathbb{1} \otimes |0\rangle\langle 0| + A_1 \otimes |1\rangle\langle 1|,$$
$$V_{AA'} = \frac{\mathbb{1} + A_0}{2} \otimes \mathbb{1} + \frac{\mathbb{1} - A_0}{2} \otimes \sigma_x, \tag{5}$$

and similarly for Bob. These operations are unitary for all A_0 and A_1 unitary and Hermitian. Obviously, their actual actions may differ from perfect swaps. For instance, suppose that the states and measurements in the boxes are equivalent to Eqs. (2) and (3) up to local unitaries: the swapped state is

always found to be $\rho_{\text{swap}} = |\bar{\psi}\rangle\langle\bar{\psi}|$ rather than its unitary equivalent. In other words, on maximally entangled two-qubit states and complementary measurements, this $\mathcal{S}$ acts as a "clever swap" that compensates for local unitaries to always produce the desired output state.

Now that $\mathcal{S}$ is given explicitly in terms of A_0, A_1, B_0, and B_1, the partial trace (1) can be formally computed [25]: the entries of ρ_{swap} are given by linear combinations of correlation terms from the set $c = \{c_{\mathbb{1}} = \text{tr}(\rho_{AB}\mathbb{1}), c_{A_0} = \text{tr}(\rho_{AB}A_0), \ldots, c_{A_0A_1B_0} = \text{tr}(\rho_{AB}A_0A_1B_0), \ldots\}$. The fidelity $\bar{F} = \langle\bar{\psi}|\rho_{\text{swap}}|\bar{\psi}\rangle$ is hence a linear combination of these moments, and so is the CHSH expression. This allows one to relate the observed CHSH violation to the overlap. Since any such moments that proceed from a quantum realization satisfy some semidefinite constraints [19,26], a lower bound on the fidelity of the swapped state is obtained by solving the following semidefinite program (SDP):

$$\begin{aligned} f = \min\ & \langle\bar{\psi}|\rho_{\text{swap}}|\bar{\psi}\rangle \\ \text{s. t.}\quad & c \in \mathcal{Q}_n, \\ & c_{A_0B_0} + c_{A_1B_0} + c_{A_0B_1} - c_{A_1B_1} = \mathcal{B}_{\text{CHSH}}, \end{aligned} \tag{6}$$

where $\mathcal{Q}_n$ is a relaxation of the quantum set. We run the SDP for various values of $\mathcal{B}_{\text{CHSH}}$. The result is the lowest curve of Fig. 2. It is now simple to add constraints: for instance, the actual statistics may correspond to isotropic boxes, i.e., $c_{A_0B_0} = c_{A_1B_0} = c_{A_0B_1} = -c_{A_1B_1}$ and $c_{A_x} = c_{B_y} = 0$, and these conditions can be added to the SDP.

Remarks on the method.—The crucial element of our method is the swap operator $\mathcal{S}$. Once expressed from the expected behavior of the boxes, and guaranteed to be unitary, the fidelity becomes a linear combination of moments c, which allow its optimization by SDP. The observed statistics enter this SDP as constraints. The outcome of the SDP is a lower bound on the desired value for two reasons: first, because one finds the minimum fidelity within $\mathcal{Q}_n$, so the fidelity within the quantum set can only be larger, and second, because the choice of $\mathcal{S}$ may not be optimal. For a given choice of $\mathcal{S}$, one may be able to prove that the SDP bound is tight by exhibiting an explicit quantum strategy which reaches the bound. At the moment of writing, we do not know how to estimate how far from optimal a choice of $\mathcal{S}$ can be, but the examples shown in this Letter demonstrate that intuitive constructions of the swap based on the expected realization of the boxes lead already to much better bounds than the previously reported ones.

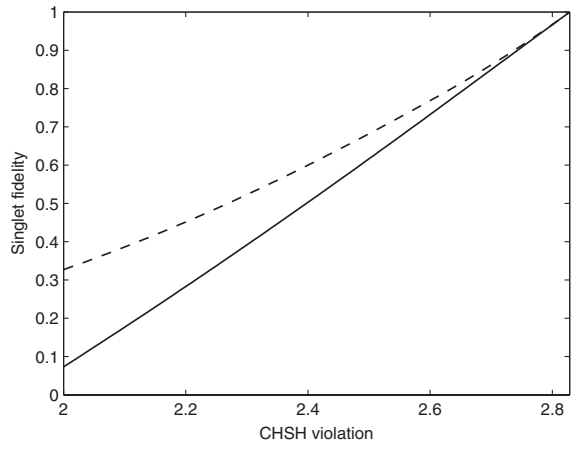

FIG. 2. Minimal singlet fidelity as a function of CHSH violation. The solid line denotes a lower bound on the fidelity for generic boxes, the dashed one a lower bound for isotropic boxes. Improved bounds are presented in Ref. [27] using optimized swap operators.

The versatility of the method is therefore evident. Having shown that it provides very robust bounds on the most studied example of self-testing, we move to apply it to a case for which no method was previously known: the self-testing of a partially entangled qutrit state through ternary-outcome statistics. Later, we shall also present an example of self-testing of measurements; several other examples are left for a forthcoming paper [27].

Partially entangled qutrits.—Self-testing of qutrits with ternary measurements, and more generally of box scenarios with more than two outputs per box, was not possible to analyze with Jordan's lemma [28], as used in Refs. [16,17]. With our method, we can achieve it by simply transposing the analysis of the CHSH inequality to the Collins-Gisin-Linden-Massar-Popescu (CGLMP) inequality $\mathcal{B}_{\text{CGLMP}} \geq 1$ [29].

The maximum quantum violation of this inequality in the case of three outcomes was conjectured to be $\mathcal{B}_{\text{CGLMP}}(p) = (12 - \sqrt{33})/9 \approx 0.6950$ [30]; this was later verified with SDP, up to numerical precision [19]. Moreover, it is believed that the maximal quantum violation can only be achieved with the nonmaximally entangled state

$$|\bar{\psi}\rangle = \frac{1}{\sqrt{2+\gamma^2}}(|00\rangle + \gamma|11\rangle + |22\rangle), \tag{7}$$

where $\gamma = (\sqrt{11} - \sqrt{3})/2$. This conjecture will be proved as a corollary of our self-testing.

The only technical step consists in finding a suitable $\mathcal{S}$ for this situation. CNOT operators for qutrit states take a different form than Eq. (5). However, they can still be expressed in terms of the measurement operators $(\bar{E}_a^x, \bar{F}_b^y)$ that yield the maximal CGLMP violation following the technique presented in the Supplemental Material [31] (more details are given in Ref. [27]). Once this is done, again, we obtain the formal expression of the two-qutrit swapped state ρ_{swap}; then, we run the SDP to obtain a lower bound on its fidelity with the reference state $|\bar{\psi}\rangle$ as a function of the CGLMP violation. The result is shown in Fig. 3. In particular, the fact that $\langle\bar{\psi}|\rho_{\text{swap}}|\bar{\psi}\rangle = 1$ when the violation is maximal shows that any quantum system maximally violating the CGLMP inequality is indeed isometrically equivalent to $|\bar{\psi}\rangle$.

Measurement estimation.—As the last application of our method in this Letter, we consider certifying measurements

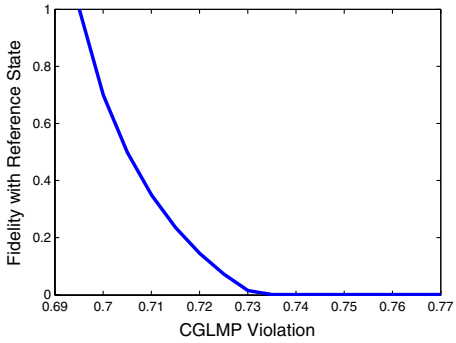

FIG. 3 (color online). Minimum fidelity of the swapped state with the reference state (7) as a function of the three-outcome Bell inequality $\mathcal{B}_{\mathrm{CGLMP}}$.

rather than states. Suppose that, rather than verifying that $|\psi\rangle$ is close to $|\bar{\psi}\rangle$, we are interested in learning to which degree the actual measurements $\{F_b^y\}$ that Bob's box is performing are well described by some matrices $\{\bar{F}_b^y\}$. The virtual procedure is again based on the intuition of the swap and thus demonstrates another use of the swap operator $\mathcal{S}$ introduced earlier: this time, consider the task of swapping into the box an arbitrary trusted state, then probe the box with different measurements y. The figure of merit should quantify how close to the ideal case the boxes perform.

For definiteness, let us practice this intuition in the CHSH case (left-hand side of Fig. 4). We conjecture that Bob's observables are close to $\bar{B}_0 = \sigma_z$ and $\bar{B}_1 = \sigma_x$. To quantify this hypothesis, we define the figure of merit as

$$\tau \equiv \frac{1}{2}\{P(0|0,|0\rangle) + P(1|0,|1\rangle) + P(0|1,|+\rangle) + P(1|1,|-\rangle)\} - 1, \tag{8}$$

where $P(b|y,|\varphi\rangle)$ denotes the probability of obtaining result b when the trusted qubit was prepared in state $|\varphi\rangle$ and one presses button y after applying the full swap [Eq. (4)] to Bob's box. τ is a number ranging from -1 to $+1$, and $\tau = 1$ is achievable only in the ideal case. As before, each $P(b|y,|\varphi\rangle)$ (and hence τ) is a linear expression in the moments c; so, a lower bound can be found with the SDP. The result is shown in the right-hand side of Fig. 4 for the case of isotropic boxes. This confirms that Bob's measurements are essentially σ_z and σ_x when CHSH takes a value close to $2\sqrt{2}$.

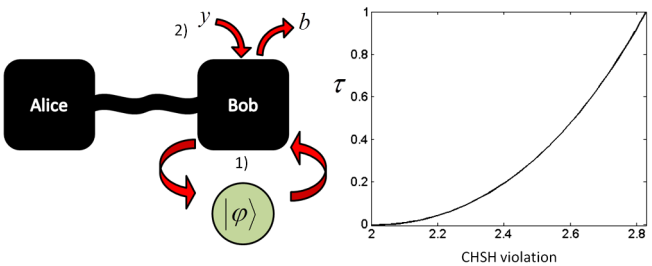

FIG. 4 (color online). Estimation of Bob's measurements. The protocol works in two steps: (1) We implement a full SWAP of Bob's box and his trusted qubit that we prepare in state $|\varphi\rangle$. (2) We implement measurement B_y and study the resulting statistics.

Conclusion.—We have described an approach to self-testing that provides much more robust bounds than previously reported and is at the same time very versatile: once the swap operator is constructed, the details of the scenario (ideal cases, figure of merit to be used) enter as parameters. The construction of unitaries $\mathcal{S}$ that provide optimal bounds remains a challenge, but one that can be met with an intuitive understanding of the problem at hand. We have illustrated the power of the method with a few paradigmatic results: the first bound on the singlet fidelity based on CHSH that is robust for real experiments (Fig. 2), the first report of self-testing of qutrits using ternary measurements (which also solves a standing conjecture about the kind of states required to violate the CGLMP inequality maximally), and an example of certification of measurements.

This work is funded by the Singapore Ministry of Education (partly through the Academic Research Fund Tier 3 MOE2012-T3-1-009) and by the National Research Foundation of Singapore. M. N. acknowledges support from the John Templeton Foundation, the European Commission (EC) STREP "RAQUEL", and the MINECO Project No. FIS2008-01236, with the support of FEDER funds. T. V. acknowledges financial support from a János Bolyai Grant of the Hungarian Academy of Sciences, the Hungarian National Research Fund OTKA (PD101461), and the TÁMOP-4.2.2.C-11/1/KONV-2012-0001 Project.

[1] J. S. Lundeen, A. Feito, H. Coldenstrodt-Ronge, K. L. Pregnell, Ch. Silberhorn, T. C. Ralph, J. Eisert, M. B. Plenio, and I. A. Walmsley, Nat. Phys. **5**, 27 (2009).
[2] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Phys. Rev. Lett. **98**, 230501 (2007).
[3] R. Colbeck and A. Kent, J. Phys. A **44**, 095305 (2011).
[4] S. Pironio, A. Acín, S. Massar, A. B. de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe, Nature (London) **464**, 1021 (2010).
[5] J. S. Bell, Physics **1**, 195 (1964).
[6] V. Scarani, Acta Phys. Slovaca **62**, 347 (2012).
[7] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Rev. Mod. Phys. **86**, 419 (2014).
[8] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).
[9] S. J. Summers and R. F. Werner, Commun. Math. Phys. **110**, 247 (1987); refer to Theorem 2.3.
[10] S. Popescu and D. Rohrlich, Phys. Lett. A **169**, 411 (1992).
[11] B. S. Tsirelson, Hadronic J. Suppl. **8**, 329 (1993).

[12] D. Mayers and A. Yao, Quantum Inf. Comput. **4**, 273 (2004).
[13] The first instance in which deviations from the ideal case were studied is by S. J. Summers and R. F. Werner, Ann. Inst. Henri Poincaré **49**, 215 (1988). But, application to experiments was not in view in that series of works, and no effort of optimizing the robustness of the estimates was made.
[14] F. Magniez, D. Mayers, M. Mosca, and H. Ollivier, arXiv:quant-ph/0512111.
[15] M. McKague, T. H. Yang, and V. Scarani, J. Phys. A **45**, 455304 (2012).
[16] C. A. Miller and Y. Shi, arXiv:1207.1819.
[17] B. W. Reichardt, F. Unger, and U. Vazirani, Nature (London) **496**, 456 (2013).
[18] B. G. Christensen *et al.*, Phys. Rev. Lett. **111**, 130406 (2013).
[19] M. Navascués, S. Pironio, and A. Acín, Phys. Rev. Lett. **98**, 010401 (2007).
[20] J. Barrett, D. Collins, L. Hardy, A. Kent, and S. Popescu, Phys. Rev. A **66**, 042111 (2002).
[21] R. D. Gill, in *Proceedings of the Foundations of Probability and Physics—2*, Series of Mathematical Modelling in Physics, Engineering, and Cognitive Science (Växjö University Press, Växjö, 2003), Vol. 5, p. 179.
[22] Y. Zhang, S. Glancy, and E. Knill, Phys. Rev. A **84**, 062118 (2011).
[23] Here and throughout the whole text, we consider only asymptotic statements; i.e., we assume that the measurement statistics are perfectly reconstructed. Also, it is understood that the observed Bell violation is not fake: the detection loophole must be closed, and no signaling must be assumed or guaranteed.
[24] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).
[25] Having chosen the ancillas A' and B' in the $|0\rangle$ state, we could have used $\mathcal{S}_{AA'} = U_{AA'}V_{AA'}$ and the same for Bob because the first U operator acts trivially. This is effectively equivalent to the isometry used in Ref. [15]: therefore, the enormous difference in robustness between that result and ours shows the power of the SDP tool.
[26] L. Vandenberghe and S. Boyd, SIAM Rev. **38**, 49 (1996).
[27] J.-D. Bancal, M. Navascués, V. Scarani, T. Vértesi, and Y. H. Yang, arXiv:1307.7053.
[28] C. Jordan, Bull. Soc. Math. Fr. **3**, 103 (1875).
[29] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu, Phys. Rev. Lett. **88**, 040404 (2002).
[30] A. Acín, T. Durt, N. Gisin, and J. I. Latorre, Phys. Rev. A **65**, 052325 (2002).
[31] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.113.040401 for more details concerning the swap method for the CGLMP inequality.