Calibration-robust entanglement detection beyond Bell inequalities

Tobias Moroder[1] and Oleg Gittsovich[2]

[1]*Institut für Quantenoptik und Quanteninformation, Österreichische Akademie der Wissenschaften, Technikerstraße 21A, A-6020 Innsbruck, Austria*

[2]*Department of Physics & Astronomy, Institute for Quantum Computing University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1*

(Received 19 December 2011; published 5 March 2012)

Most entanglement verification is examined in either the completely characterized or the totally device-independent scenario. The assumptions imposed by these extreme cases are often either too weak or too strong for real experiments. Here we investigate this detection task in the intermediate regime where there is partial knowledge of the measured observables, considering cases like orthogonal, sharp, or only dimension-bounded measurements. We show that for all these assumptions it is not necessary to violate a corresponding Bell inequality in order to detect entanglement. We derive strong detection criteria that can be directly evaluated for experimental data and which are robust against large classes of calibration errors. The conditions are even capable of detecting bound entanglement under the sole assumption of dimension-bounded measurements.

DOI: 10.1103/PhysRevA.85.032301 PACS number(s): 03.67.Mn, 03.65.Ud, 03.65.Ta

I. INTRODUCTION

Entanglement is the most striking phenomenon in the quantum world. It provides the resource for fascinating new applications such as quantum computing, teleportation, or unconditional secure communiction. These possibilities sparked great interest in this resource, and many current experiments strive to realize strong and robust entanglement. Consequently many different detection methods have been developed in recent years; for reviews, see Refs. [1,2].

Reliable entanglement verification must fulfill certain criteria [3]. Most importantly it should not depend on the preparation procedure, and the only available information about the presence of entanglement should be obtained via measurements of the underlying system. However there is still one open choice left: Usually each classical outcome is associated with an operator describing the measurement apparatus, say, outcome k corresponds to the measurement operator M_k. Equipped with this quantum-mechanical meaning, one can employ for instance the tool of entanglement witnesses [4–6] to decide on the presence of entanglement. This standard scenario might be too optimistic for applications because it crucially relies on the correctness of the employed operator description of the measuring device, and clearly if the true measurement apparatus functions are different then anything can go wrong. For mere entanglement detection these deviations might be called systematic errors, but for applications where the presence of entanglement is essential, secure communication being a prominent example, these deviations are undesired pitfalls [7,8]. Thus in contrast to the completely characterized scenario there is also the other extreme where one does not need any specific quantum-mechanical model at all. Although it is surprising at first, even in this most pessimistic, device-independent case it is possible to infer entanglement for good enough data, for example, by use of Bell inequalities [9–11].

Detection of entanglement in a completely device-independent manner has recently attracted a lot of interest; in particular verification in multipartite settings [12–14] since it was realized that this task differs from the corresponding nonlocality one [12]. This was surprising because device-independent entanglement detection and exclusion of a local-hidden-variable model are equivalent problems in the bipartite case [15,16]. Steering inequalities are entanglement detection methods in hybrid scenarios, i.e., one party is complete characterized, the other totally uncharacterized [17]. Only a few results and techniques have addressed the detection of entanglement in a partially characterized setting so far. The authors of Refs. [18,19] considered the case of sharp, orthogonal qubit measurements and showed that much more entangled states can in fact be detected than with the corresponding Bell inequality. For instance it was proven that the bound appearing in the famous Bell correlation term of the Clauser-Horne-Shimony-Holt (CHSH) inequality [10] can actually be reduced from 2 to $\sqrt{2}$ when the measurements satisfy this extra constraint. These results were extended in order to provide even quantitative bounds on the amount of entanglement in Ref. [20]. In order to devise more robust entanglement verification methods that avoid fake entanglement detection under misspecification of the employed observables, a technique called the squash model is very useful [21,22]. Applied to entanglement verification, this notion, usually common in quantum key distribution, can even be extended [23]. Finally, in Ref. [24] types of Bell inequalities are introduced that can be applied if the commutator of different measurement settings is known.

The purpose of the current paper is to investigate the verification task in the intermediate regime where one possesses some partial knowledge of the employed measurement devices. Despite its elegance the completely device-independent setting suffers from the drawback that it is hard to address experimentally. In fact no Bell experiments so far were fully device independent due to certain loopholes, but these are often irrelevant if certain other assumptions hold. However, such assumptions, like the fair sampling condition, effectively can be regarded as partial information, e.g., that the "inconclusive" measurement outcome is the same for both settings for the fair sampling case. But clearly, if these additional assumptions are not satisfied, the corresponding implementation of the Bell test also does not provide a positive entanglement check [25,26]. Besides, the completely device-independent scenario is often also too pessimistic. Of course, for applications like quantum key distribution this very pessimistic viewpoint is legitimate but for the mere verification of entanglement in an experiment this seems like breaking a butterfly on a wheel. Although we do not address the question of how such partial knowledge can be obtained or justified, we nevertheless believe that certain deviations in a measurement description are more serious than others. For example, the assumption that the measurement of the electronic state of an ion in a trap is very well described by a qubit measurement seems much easier to assure than the assumption that the performed measurements are really orthogonal or that they are true projectors. Independent of this discussion of which scenario is now more reasonable for which situation, our investigations shed some light on the question of which assumptions are more crucial than others in order to verify entanglement. In addition, the derived entanglement criteria are more robust against calibration errors while they still keep a large detection strength, in particular when compared to Bell inequalities.

In the following we first focus on the scenario investigated in Ref. [19] and analyze entanglement verification in the simplest possible setting of two different dichotomic measurement settings per side. We direct our attention to different assumptions about qubit measurements and distinguish three different classes: sharp, orthogonal, and completely uncharacterized qubit measurements. We provide a solution in terms of the singular values of a corresponding data matrix and find that one detects a much larger fraction of states than with the completely device-independent setting. This already provides examples where one detects entanglement although the corresponding Bell inequalities, the CHSH inequalities [10] in this case, are not violated, with the sole extra assumption that the dimension of the underlying quantum system is fixed. Additionally we consider specific observations where the additional knowledge of sharpness or orthogonality is irrelevant for the detection strength and already the dimension restriction suffices to verify exactly the same amount of entangled states as with completely characterized, i.e., sharp and orthogonal, measurements. After considering these various scenarios for two qubits we focus on the extensions to more dichotomic measurements with completely unspecified measurements restricted only by the underlying dimension. We derive a criterion that is applicable for any of these settings and show that it is capable of detecting entanglement in data originating from bound entangled states. Moreover the criterion even shows that with uncharacterized qutrit measurements one can verify more entanglement than with the corresponding CHSH inequalities. Note that dimensions $d \geqslant 4$ are not relevant, because then one effectively equals the detection strength of the Bell inequalities [15,16].

The outline is as follows: In Sec. II we precisely define entanglement verification under partial information on the performed measurements. Section III starts with a discussion about different assumptions on the measurements. In addition we provide some further notation and background knowledge about the entanglement criterion that we employ for our purpose. Section III E finally contains the above-mentioned results for the two-qubit case, whereas Sec. IV is devoted to the general scenario of n uncharacterized dichotomic qudit measurements. Finally we conclude and comment on possible further extensions and directions in Sec. V. Some technical details of the proofs can be found in the Appendix.

II. PROBLEM DEFINITION

Suppose that Alice and Bob observe an outcome probability distribution denoted as $P(x,y|a,b)$, where a labels different measurement choices with corresponding outcomes x for Alice, and similarly for Bob. These observed data have a quantum-mechanical representation if there exists a quantum state ρ_{AB} and corresponding measurements, i.e., sets of positive operator-valued measures (POVM) for Alice $\{M_x^a\}$ and Bob $\{M_y^b\}$ such that

$$P(x,y|a,b) = \big(\rho_{\text{AB}} M_x^a \otimes M_y^b\big), \quad \forall \quad x,y,a,b. \tag{1}$$

The observed data are said to verify entanglement if and only if all states ρ_{AB} that satisfy this relation are entangled. Note that the measurement description is crucial here because it ties a quantum-mechanical meaning to the classical outcomes.

In the following we consider the alternative that only partial knowledge is possessed about the measurement description, meaning that the POVMs describing the measurement are not known completely. Hence each local measurement characterization is only assured to lie within a certain class. This set of possible POVMs will be denoted by $\mathcal{M}_{\text{A}}$ for Alice and $\mathcal{M}_{\text{B}}$ for Bob. In this case, successful entanglement detection implies that for all measurement descriptions only entangled states give rise to the observed data. More precisely, if $\mathcal{S}$ denotes the set of states having a quantum representation in accordance with the assumed measurement description,

$$\mathcal{S} = \big\{\rho_{\text{AB}} |\ \exists \big\{M_x^a\big\} \in \mathcal{M}_{\text{A}}, \big\{M_y^b\big\} \in \mathcal{M}_{\text{B}} : \\ P(x,y|a,b) = \text{tr}\big(\rho_{\text{AB}} M_x^a \otimes M_y^b\big),\ \forall\ x,y,a,b\big\}, \tag{2}$$

then the observed data $P(x,y|a,b)$ certify entanglement if and only if all states of this set $\mathcal{S}$ are entangled.

Let us comment on the two extreme cases: If the measurements are completely specified each set $\mathcal{M}$ consists of only one possible element. In this case the question of whether given observations verify entanglement is completely answered, for example, with the help of entanglement witnesses, even if the measurements do not provide full tomography [27]. In the other extreme that the measurements are completely unspecified the sets $\mathcal{M}$ consist of all possible POVMs in all possible dimensions, and the data correspond exclusively to entangled states if and only if a Bell inequality with the specified number of settings and outcomes is violated [15].

Finally let us stress one more technical point: We do not assume any "convexification" of the problem as, for example, employed in Ref. [28] for dimension witnesses. Convexification would mean that if two different observations P_1 and P_2 have a separable quantum representation then so also does their convex combination $\lambda P_1 + (1-\lambda)P_2$ for all $\lambda \in [0,1]$. However, the problem is that the quantum representations might need different measurements, so that directly taking the convex combination on the level of quantum states does not work anymore. In Sec. III E we provide an explicit example of this nonconvexity. If one manually convexifies the problem, e.g., by allowing that states and measurements can be conditioned on an extra random variable (or shared randomness in the language of Ref. [28]), one misses the entanglement of several data sets which could be detected otherwise. Thus the effectiveness of this intermediate approach lies, to some extent, in explicitly paying attention to the nonconvex structure.

III. QUBIT CASE

This section concentrates on the two-qubit scenario. We start with the definition of different measurement assumptions followed by an explicit parametrization. Afterward we introduce the notation of a data matrix in order to express our results more compactly and also state the entanglement criterion that is employed to prove the main results in the last part of this section.

A. Different measurement assumptions

First, let us specify more closely the different measurement properties which were abstractly described by the set $\mathcal{M}$ in the previous section. We consider the simplest nontrivial case: Each party has two different measurement settings each of which has two different outcomes. Any such dichotomic measurement is more compactly determined by the difference of two POVM elements, e.g., with the first setting for Alice we associate the operator

$$A = M^a_{+1} - M^a_{-1} \quad \Leftrightarrow \quad M^a_{\pm 1} = \tfrac{1}{2}(\mathbb{1} \pm A). \tag{3}$$

Here $x = \pm 1$ labels the two different outcomes, while the resolution of the POVM elements $M^a_{\pm 1}$ follows because of normalization. In order that this operator A describes a valid POVM it must satisfy the conditions

$$A - \mathbb{1} \geqslant 0, \quad \mathbb{1} - A \geqslant 0. \tag{4}$$

Let us remark that this condition is still independent of any dimension restriction and it will reappear in a later section for the more general case of uncharacterized qudit measurements. The operator for the second choice of Alice is denoted by A', while Bob's choices are given by B and B', respectively.

The following definition summarizes the different qubit specifications that we consider. Let us point out that the dimension restriction seems to us the first nontrivial assumption that one can make.[1]

Definition 1 (Qubit measurement models). For two dichotomic measurements, characterized by the operators A and A' according to Eq. (3) and satisfying Eq. (4), we distinguish the following cases:

(i) Uncharacterized qubit measurements: Both operators act on *the same* qubit.

(ii) Sharp qubit measurements: The POVM elements are rank-1 projectors on the same qubit, i.e., A and A' have eigenvalues ± 1.

(iii) Orthogonal qubit measurements: The eigenbasis of A and A' are mutually unbiased.[2]

B. Parametrization of POVM elements

In the following we introduce a parametrization of the POVM elements corresponding to different measurement scenarios. This parametrization will be convenient later for the technical proofs. Additionally it should further clarify the different measurement properties.

1. Sharp qubit measurements

In this case the operators A and A' can be written as follows:

$$A = \cos(\theta)\sigma_i + \sin(\theta)\sigma_j, \tag{5}$$

$$A' = \cos(\theta)\sigma_i - \sin(\theta)\sigma_j, \tag{6}$$

where σ_i and σ_j are two different, possibly rotated, Pauli operators, i.e., they can be written as $\sigma_i = \hat{u}_i \cdot \vec{\sigma}$ and $\sigma_j = \hat{u}_j \cdot \vec{\sigma}$ with two unit vectors $\hat{u}_i, \hat{u}_j \in \mathbb{R}^3$ satisfying $\hat{u}_i \cdot \hat{u}_j = 0$. The parameter θ characterizes the tilt between the measurement directions. Note that this relation can also be reversed, i.e., to express the Pauli operator in terms of the considered measurement operators. Formally, the relation between orthogonal and nonorthogonal observables is described by

$$\begin{bmatrix} \mathbb{1} \\ \sigma_i \\ \sigma_j \end{bmatrix} = R(\theta) \begin{bmatrix} \mathbb{1} \\ A \\ A' \end{bmatrix} \tag{7}$$

with[3]

$$R(\theta) = \begin{bmatrix} 1 & & \\ & \frac{1}{2\cos(\theta)} & \frac{1}{2\cos(\theta)} \\ & \frac{1}{2\sin(\theta)} & -\frac{1}{2\sin(\theta)} \end{bmatrix}. \tag{8}$$

If we refer only to the 2×2 submatrix, formed by the second and third columns and rows, we use the label $R_2(\theta)$.

2. Orthogonal qubit measurements

When the measurements are orthogonal but not necessarily sharp we directly employ the reverse parametrization

$$\sigma_i = x_1 \mathbb{1} + x_2 A, \tag{9}$$

$$\sigma_j = x_3 \mathbb{1} + x_4 A', \tag{10}$$

with $x_i \in \mathbb{R}$. Note that in order for A and A' to correspond to the physical observables given by Eq. (4) these parameters must satisfy $x_2 \geqslant 1 + |x_1|$ and $x_4 \geqslant 1 + |x_3|$. Here we can choose without loss of generality $x_2, x_4 > 0$ to be positive,

[1] The only other alternative would be to provide a distance measure for the set of POVMs, i.e., $\delta(\{M^a_{x,\mathrm{ideal}}\},\{M^a_{x,\mathrm{true}}\}) \leqslant \epsilon$ which quantifies the difference between the true and the ideal measurement descriptions. If one wants to make this bound independent of the dimension, this norm must be independent of the dimension as well. However, we cannot think of any reasonable distance here.

[2] Note that this does not imply the orthogonality of A and A' with respect to the Hilbert-Schmidt norm.

[3] For notational clarity if matrix entries are blank they are equal to zero; if they can be arbitrary they are symbolized by $*$.

by selecting σ_i or $-\sigma_i$ appropriately. Formally the sharp and nonsharp observables can be related by

$$\begin{bmatrix} \mathbb{1} \\ \sigma_i \\ \sigma_j \end{bmatrix} = S(\vec{x}) \begin{bmatrix} \mathbb{1} \\ A \\ A' \end{bmatrix} \tag{11}$$

with

$$S(\vec{x}) = \begin{bmatrix} 1 & & \\ x_1 & x_2 & \\ x_3 & & x_4 \end{bmatrix}. \tag{12}$$

3. Uncharacterized qubit measurement

The remaining case of totally uncharacterized qubit measurements can be considered as a combination of the above two cases. The overall transformation is given by

$$\begin{bmatrix} \mathbb{1} \\ \sigma_i \\ \sigma_j \end{bmatrix} = R(\theta) S(\vec{x}) \begin{bmatrix} \mathbb{1} \\ A \\ A' \end{bmatrix}. \tag{13}$$

The first operation S turns the operators A and A' into sharp, but not necessarily orthogonal, measurements, which are considered afterward by applying the transformation R.

C. Data matrix

In order to express our results let us define some further notation. The observed data $P(x,y|a,b)$ are compactly expressed in terms of a data matrix D_3, given by the matrix of expectation values

$$D_3 = \begin{bmatrix} \langle \mathbb{1} \rangle & \langle B \rangle & \langle B' \rangle \\ \langle A \rangle & \langle AB \rangle & \langle AB' \rangle \\ \langle A' \rangle & \langle A'B \rangle & \langle A'B' \rangle \end{bmatrix}. \tag{14}$$

For convenience we often abbreviate the 2×2 submatrix containing only the full correlations, i.e., built up by the second and third rows and columns, as D_2. Our criteria are typically given in terms of the singular values of this submatrix, denoted as $\lambda_{1/2} \geqslant 0$.

D. Employed entanglement criterion

For entanglement detection we employ a criterion which is a direct corollary of the computable cross-norm or realignment (CCNR) criterion [29,30]. The corollary is formulated in terms of the singular values of the correlation matrix T_3 given by

$$T_3 = \begin{bmatrix} \langle \mathbb{1} \rangle & \langle \sigma_i^B \rangle & \langle \sigma_j^B \rangle \\ \langle \sigma_i^A \rangle & \langle \sigma_i^A \sigma_i^B \rangle & \langle \sigma_i^A \sigma_j^B \rangle \\ \langle \sigma_j^A \rangle & \langle \sigma_j^A \sigma_i^B \rangle & \langle \sigma_j^A \sigma_j^B \rangle \end{bmatrix}. \tag{15}$$

Note that this correlation matrix T_3 represents a special data matrix D_3 for which the employed measurement operators are sharp and orthogonal. Because of those similarities we employ a similar label T_2 in order to refer to the full correlation 2×2 submatrix.

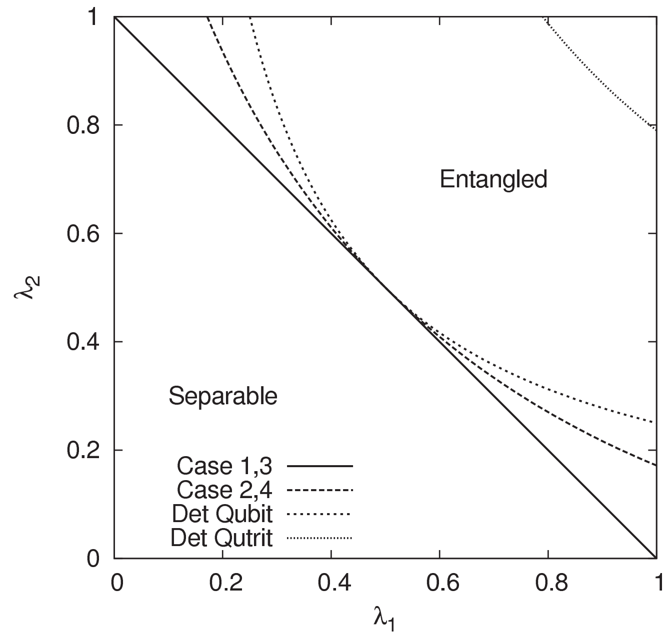

FIG. 1. Different detection regions for a data matrix with singular values λ_1 and λ_2. The solid and long-dashed lines correspond to the case of a data matrix with vanishing marginals as discussed in Proposition 2. The remaining two lines correspond to the determinant detection rule $\det(D) = \lambda_1 \lambda_2$ given in Proposition 4 for qubits (short-dashed line) and qutrits (dotted line).

Proposition 1 (Corollary of the CCNR criterion). Given the correlation matrix T_3 with ordered singular values $\lambda_0 \geqslant \lambda_1 \geqslant \lambda_2 \geqslant 0$. Then the CCNR criterion implies that any separable state necessarily satisfies

$$\|T_3\|_1 = \lambda_0 + \lambda_1 + \lambda_2 \leqslant 2. \tag{16}$$

If the correlation matrix has vanishing marginals, i.e., $\langle \sigma_k^A \rangle = \langle \sigma_k^B \rangle = 0$ for all $k \in \{i,j\}$, and $\lambda_0 = 1$ then this condition is also sufficient.

For completeness we provide a proof of this proposition in Appendix A. With this stage set we will state in the next section our main results on entanglement verification in different two-qubit scenarios.

E. Main results

In the following we state and prove our main results for qubits. We first consider the special case that the observed data matrix D_3 has vanishing marginals. We obtain a complete solution for different scenarios if we use only the knowledge of the singular values that appear, i.e., the criteria are minimized over all data matrices with fixed singular values. Using additional structure of the observation improves the detection strength as we will see later in Proposition 3. For comparison, the different detection regions are visualized in Fig. 1.

Proposition 2 (Data matrix with zero marginals). Given a data matrix D_3 of the form

$$D_3 = \begin{bmatrix} 1 & \\ & D_2 \end{bmatrix}, \tag{17}$$

where the full correlation data matrix D_2 is characterized by the singular values $\lambda_1 \geqslant \lambda_2 \geqslant 0$. These data verify entanglement under the assumption that the qubit measurements of both Alice and Bob are

(1) sharp and orthogonal: $\lambda_1 + \lambda_2 > 1$;
(2) sharp: $\sqrt{\lambda_1} + \sqrt{\lambda_2} > \sqrt{2}$;
(3) orthogonal: $\lambda_1 + \lambda_2 > 1$;
(4) uncharaterized: $\sqrt{\lambda_1} + \sqrt{\lambda_2} > \sqrt{2}$.

The definitions of these properties are given in Definition 1 and these bounds are tight for the considered scenario.

Remark 1. Note that we assume that D_3 actually originates from a quantum state under the considered measurement scenario, which can be assured, for example, if the singular values satisfy $1 \geqslant \lambda_1 \geqslant \lambda_2 \geqslant 0$.

Proof. Case (1) of sharp and orthogonal measurements has already been discussed in Ref. [19]. Alternatively it is a direct application of Proposition 1.

All other scenarios are proven along the following lines: Given the data matrix D_3 one first reconstructs the corresponding correlation matrix T_3 by the appropriate transformations $S(\vec{x})$ and $R(\theta)$ as given in Sec. III B. In order to certify entanglement one employs Proposition 1. However, since T_3 depends on the transformation parameters, e.g., $\theta,\vec{x},\ldots$, one needs to optimize over all such choices. This will in general result in lower bounds on the singular values of the data matrix. If these bounds are tight then the provided condition is necessary and sufficient in order to detect entanglement with the provided data.

Case (2): First let us concentrate on the sharp but not necessarily orthogonal case, in which the correlation matrix is given by $T_3 = R(\alpha)D_3R(\beta)^T$. For the block-diagonal data matrix the resulting correlation matrix is of similar block structure, i.e., $T_3 = \mathrm{diag}[1,T_2]$ with

$$T_2 = R_2(\alpha)D_2R_2^T(\beta). \tag{18}$$

Hence, if the ordered singular values of T_2 are denoted as $t_1 \geqslant t_2 \geqslant 0$, then Proposition 1 states that the state is entangled if and only if $t_1 + t_2 > 1$ holds for all values of α and β. In order to minimize the sum $t_1 + t_2$ over the angles, we first lower-bound this quantity by an expression containing only the singular values of the appearing transformations since this is more easily optimized in the end.

The lower bound is derived using the inverse relation of Eq. (18),

$$D_2 = R_2(\alpha)^{-1}T_2R_2^T(\beta)^{-1} \tag{19}$$

with

$$R_2(\alpha)^{-1} = \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ \cos(\alpha) & -\sin(\alpha) \end{bmatrix}. \tag{20}$$

In the following discussion we employ the abbreviations $a_1 \geqslant a_2 \geqslant 0$ and $b_1 \geqslant b_2 \geqslant 0$ for the ordered singular values of $R_2^T(\alpha)^{-1}$ and $R_2^T(\beta)^{-1}$, respectively. Furthermore, recall that D_2 is characterized by its two singular values $\lambda_1 \geqslant \lambda_2 \geqslant 0$. For the matrices on the left- and right-hand sides of Eq. (19) the following relations hold:

$$t_1t_2 = \frac{\lambda_1\lambda_2}{a_1a_2b_1b_2}, \tag{21}$$

$$t_1^2 + t_2^2 \geqslant \frac{(\lambda_1+\lambda_2)^2}{a_1^2b_1^2 + a_2^2b_2^2}. \tag{22}$$

The proof of these two relations involves some technical details and is given in Appendix B 1. Employment of these two identities provides

$$(t_1+t_2)^2 \geqslant \min_{\alpha,\beta}\left[\frac{(\lambda_1+\lambda_2)^2}{a_1^2b_1^2 + a_2^2b_2^2} + 2\frac{\lambda_1\lambda_2}{a_1a_2b_1b_2}\right] \tag{23a}$$

$$\geqslant \frac{1}{4}(\sqrt{\lambda_1}+\sqrt{\lambda_2})^4. \tag{23b}$$

The last inequality arises if one employs the true singular values $a_1(\alpha),\ldots$ and performs the minimization; for an explicit proof of this optimization see Lemma 1 in Appendix A. Equation (23b) confirms that the state is entangled if and only if $\sqrt{\lambda_1} + \sqrt{\lambda_1} > \sqrt{2}$, where the sufficiency follows from the fact that all appearing inequalities can also be achieved with equality. This finalizes the proof of Case (2).

Case (3): Next consider the orthogonal case. The correlation matrix $T_3 = S(\vec{x})D_3S(\vec{y})^T$, with appropriate "sharpener" transformations $S(\vec{x})$ given by Eq. (12), has the following block structure:

$$T_3 = \begin{bmatrix} 1 & \\ x & S_x \end{bmatrix}\begin{bmatrix} 1 & \\ & D_2 \end{bmatrix}\begin{bmatrix} 1 & y^T \\ & S_y^T \end{bmatrix} \tag{24a}$$

$$= \begin{bmatrix} 1 & y^T \\ x & xy^T + S_xD_2S_y^T \end{bmatrix}. \tag{24b}$$

Here we used an analog block decomposition for $S(\vec{x})$ with column vector $x = [x_1,x_3]^T$ and diagonal submatrix $S_x = \mathrm{diag}[x_2,x_4]$ and corresponding abbreviations for the transformation of Bob. In order to prove the statement we will use the following three inequalities for the ordered singular values of T_3:

$$t_0 \geqslant 1, \tag{25}$$

$$t_0t_1 \geqslant \lambda_1, \tag{26}$$

$$t_0t_1t_2 \geqslant \lambda_1\lambda_2. \tag{27}$$

The proof of these inequalities is given in Appendix B 2. Further, as shown in Lemma 2 of Appendix C these conditions, together with the ordering condition $1 \geqslant \lambda_1 \geqslant \lambda_2 \geqslant 0$, ensure that

$$t_0 + t_1 + t_2 \geqslant 1 + \lambda_1 + \lambda_2. \tag{28}$$

If this ordering is not valid, i.e., $\lambda_1 > 1$, entanglement directly follows because

$$t_0 + t_1 + t_2 \geqslant t_0 + t_1 \geqslant 2\sqrt{t_0t_1} > 2 \tag{29}$$

via the inequality of arithmetic and geometric means. Thus in total $\lambda_1 + \lambda_2 > 1$ is necessary and sufficient for entanglement; the sufficiency is because one detects the same result as in the more restrictive case of sharp, orthogonal measurements.

This finishes the proof for the orthogonal case of qubit measurements.

Case (4): For the remaining scenario of fully uncharacterized qubit measurements we can largely employ the previous results. In this scenario the correlation matrix $T_3 = R(\alpha)S(\vec{x})D_3S(\vec{y})^T R(\beta)^T$ is given by

$$T_3 = \begin{bmatrix} 1 & \\ & R_2(\alpha) \end{bmatrix} \begin{bmatrix} 1 & y^T \\ x & xy^T + S_x D_2 S_y^T \end{bmatrix} \begin{bmatrix} 1 & \\ & R_2(\beta)^T \end{bmatrix} \quad (30a)$$

$$= \begin{bmatrix} 1 & \tilde{y}^T \\ \tilde{x} & \tilde{x}\tilde{y}^T + R_2(\alpha) S_x D_2 S_y^T R_2(\beta)^T \end{bmatrix} \quad (30b)$$

with $\tilde{x} = R_2(\alpha)x, \tilde{y} = R_2(\beta)y$. In this case the important submatrix is

$$\bar{T}_2 = R_2(\alpha)S_x D_2 S_y^T R_2(\beta)^T = R_2(\alpha)\bar{D}_2 R_2(\beta)^T, \quad (31)$$

which can be considered as the central submatrix of the sharp case, Eq. (18), but where the transformation is applied to $\bar{D}_2$ instead of the true data matrix D_2 itself. From the sharp case we know that the singular values of this matrix $\bar{T}_2$, denoted as $\bar{t}_1 \geqslant \bar{t}_2$, satisfy

$$\bar{t}_1 + \bar{t}_2 \geqslant \tfrac{1}{2}(\sqrt{\bar{\lambda}_1} + \sqrt{\bar{\lambda}_2})^2 \geqslant \tfrac{1}{2}(\sqrt{\lambda_1} + \sqrt{\lambda_2})^2, \quad (32)$$

where $\bar{\lambda}_i \geqslant \lambda_i$ are the singular values of $\bar{D}_2$. Next, using similar arguments as already presented in the unsharp case but orthogonal case we can derive the following set of inequalities for the singular values of T_3:

$$t_0 \geqslant 1, \quad (33)$$

$$t_0 t_1 \geqslant \bar{t}_1, \quad (34)$$

$$t_0 t_1 t_2 \geqslant \bar{t}_1 \bar{t}_2. \quad (35)$$

Use of Lemma 2 and Eq. (32) once more provides

$$t_0 + t_1 + t_2 \geqslant 1 + \bar{t}_1 + \bar{t}_2 \geqslant 1 + \tfrac{1}{2}(\sqrt{\lambda_1} + \sqrt{\lambda_2})^2 \quad (36)$$

if one has the ordering $1 \geqslant \bar{t}_1 \geqslant \bar{t}_2 \geqslant 0$. If $\bar{t}_1 > 1$ one verifies entanglement again by the inequality of the arithmetic and geometric means. This finally shows that the state is entangled if and only if $\sqrt{\lambda_1} + \sqrt{\lambda_2} > \sqrt{2}$, which proves the claim for uncharacterized qubit measurements. ■

Next let us provide an important numerical example. First it demonstrates that the orthogonal case is indeed different from the completely characterized case. Together with Proposition 2 it shows that the sharp case and the orthogonal case are indeed inequivalent to each other, i.e., there are observations which are exclusively detected by one of these two scenarios. Additionally, this example proves that the entanglement verification in the unsharp case is in fact a nonconvex problem. This means that one must be very careful in applying Proposition 2; it is, for example, not possible to use it on a "depolarized" version of the observed data matrix D_3, i.e., the one that one obtains by setting the marginals equal to zero.

Example 1. The data matrix

$$D_3 = \begin{bmatrix} 1 & 1-\sqrt{3} & \\ 1-\sqrt{3} & (15-8\sqrt{3})/2 & \\ & & 1/2 \end{bmatrix} \quad (37a)$$

$$\approx \begin{bmatrix} 1 & -0.73 & \\ -0.73 & 0.57 & \\ & & 0.5 \end{bmatrix} \quad (37b)$$

can originate from a separable state in the case of orthogonal qubit measurements, but verifies entanglement for sharp qubit measurements.

Moreover, it shows that the unsharp scenarios are indeed nonconvex problems, i.e., a convex combination of two separable data matrices might no longer be separable.

Proof. First let us give the separable state and its corresponding measurements that are consistent with the given data matrix. This also discloses the generation of this example. The data matrix given by Eq. (37a) is obtained by measuring the separable state

$$\rho_{\text{sep}} = \tfrac{1}{4}\left[\mathbb{1} \otimes \mathbb{1} + \tfrac{1}{2}(\sigma_x \otimes \sigma_x + \sigma_z \otimes \sigma_z)\right] \quad (38)$$

with unsharp measurements A parametrized according to Eq. (9) with $\sigma_i = \sigma_x$, $x_1 = 1 + \sqrt{3}$, and $x_2 = 1 + x_1$, while $A' = \sigma_z$, and we have the same settings for Bob's side. The reason for the choice of x_1 becomes clear later.

Verifying entanglement under the premise of sharp qubit observables goes as follows: Note that if the measurements are even orthogonal application of the CCNR criterion of Proposition 1 immediately shows entanglement. In the nonorthogonal case one utilizes again the transformations $R(\alpha)$ and $R(\beta)$ in order to generate T_3 as it was done in Proposition 2. Even with nonvanishing marginals the important submatrix is given by $T_2 = R_2(\alpha)D_2R_2(\beta)^T$ with D_2 being the submatrix containing the full correlations. Although it is not directly stated as the CCNR criterion, a state is already entangled if the singular values of T_2 satisfy $t_1 + t_2 > 1$.[4] As shown in the proof of Proposition 2, this relation is assured if the singular values of D_2 fulfill $\sqrt{\lambda_1} + \sqrt{\lambda_2} > \sqrt{2}$. Since the data matrix given by Eq. (37a) satisfies this condition, this proves that D_3 cannot be compatible with a separable state under sharp qubit measurements.

Finally one needs to verify using sharp measurements that the data matrix given by Eq. (37a) is at all consistent with a valid quantum state. This is necessary in order to assure that the set $\mathcal{S}$ defined as in Eq. (2) is indeed nonempty. However, the operator

$$\rho_{\text{ent}} = \frac{1}{4}\left(y\sigma_y \otimes \sigma_y + \sum_{i,j \in \{0,x,y\}} [D_3]_{ij} \sigma_i \otimes \sigma_j \right) \quad (39)$$

[4] If the correlation matrix T_3 corresponds to a separable state, then so also does $\bar{T}_3$ where the marginals have been inverted. Since correlation matrices of separable states form a convex structure, this assures that the depolarized version $\tilde{T}_3 = (T_3 + \bar{T}_3)/2 = \text{diag}[1, T_2]$ is also separable. Application of the CCNR criterion to $\tilde{T}_3$ assures entanglement if $t_1 + t_2 > 1$ is fulfilled.

with $y = 4\sqrt{3} - 7$ represents a valid state compatible with the data matrix if one employs the sharp measurements $A = B = \sigma_x, A' = B' = \sigma_z$. Let us point out that this physical condition determined the parameter x_1: We optimized the detection condition $\sqrt{\lambda_1} + \sqrt{\lambda_2}$ while still keeping the data compatible with a valid state.

Furthermore, this example provides an explicit instance for the failure of convexity.[5] If D_3 corresponds to a separable state then so also does its marginal inverted version $\bar{D}_3$ because it effectively represents only a classical outcome interchange $+1 \leftrightarrow -1$ on both sides. Thus the data matrix given by Eq. (37a) with marginals $-(1-\sqrt{3})$ is also separable. However, taking an equal mixture leads to the depolarized data matrix $\tilde{D}_3 = \mathrm{diag}[1, D_2]$, which would verify entanglement according to Proposition 2. ■

The following proposition demonstrates that the entanglement properties are not fully determined by the singular values of the observed data matrix. Moreover, for this special data structure it is interesting to observe that the extra knowledge of sharpness and orthogonality is irrelevant for the detection strength and mere information about the dimension of the measurements suffices to verify the same fraction of entanglement. Furthermore, since these observations satisfy all CHSH inequalities [10], complete device independent detection is not possible.

Proposition 3 (Diagonal data matrix). Observations of a diagonal data matrix

$$D_3 = \begin{bmatrix} 1 & & \\ & \lambda_1 & \\ & & \lambda_2 \end{bmatrix} \tag{40}$$

with $1 \geqslant \lambda_{1/2} \geqslant 0$ verify entanglement under the assumption of qubit measurements if and only if $\lambda_1 + \lambda_2 > 1$. Hence one verifies the same fraction as with sharp, orthogonal qubit measurements. In contrast, completely device-independent entanglement verification fails.

Proof. The proof runs analogous to that of the qubit measurement scenario of Proposition 2. Note that the full correlation matrix T_3 is given by Eq. (30b) and that one detects entanglement if and only if the singular values of $\bar{T}_2 = R_2(\alpha)\bar{D}_2 R_2(\beta)^T$ given by Eq. (31), satisfy $\bar{t}_1 + \bar{t}_2 > 1$. However, in contrast to Proposition 2 it is now possible to derive a tighter lower bound by exploiting the diagonal structure of

$$\bar{D}_2 = \begin{bmatrix} x_2 y_2 \lambda_1 & \\ & y_2 y_4 \lambda_2 \end{bmatrix} = \begin{bmatrix} \bar{\lambda}_1 & \\ & \bar{\lambda}_2 \end{bmatrix}. \tag{41}$$

The singular values of $\bar{D}_2$ are given by the diagonal entries, which fulfill $\bar{\lambda}_i \geqslant \lambda_i$, due to the constraints on the parameters x_i and y_i. In order to finish the proof we employ as usual certain inequalities for singular values of the matrices $\bar{T}_2$ and D_2:

$$\bar{t}_1 \bar{t}_2 \geqslant \lambda_1 \lambda_2, \tag{42}$$

$$\bar{t}_1^2 + \bar{t}_2^2 \geqslant \lambda_1^2 + \lambda_2^2. \tag{43}$$

These inequalities imply $\bar{t}_1 + \bar{t}_2 \geqslant \lambda_1 + \lambda_2$ and therefore prove the claim of the proposition.

In order to prove the inequality (42) one applies the determinant multiplication rule together with the property $|\det[R_2(\cdot)]| \geqslant 1$ that can be checked directly from the definition given by Eq. (8). The second inequality (43) is verified by

$$\bar{t}_1^2 + \bar{t}_2^2 = \mathrm{tr}\big(\bar{T}_2 \bar{T}_2^T\big) \tag{44a}$$

$$= \tfrac{1}{16}\{(\bar{\lambda}_1 + \bar{\lambda}_2)^2[\csc(\alpha)^2\csc(\beta)^2 + \sec(\alpha)^2\sec(\beta)^2] + (\bar{\lambda}_1 - \bar{\lambda}_2)^2[\csc(\alpha)^2\sec(\beta)^2 + \sec(\alpha)^2\csc(\beta)^2]\} \tag{44b}$$

$$\geqslant \tfrac{1}{2}[(\bar{\lambda}_1 + \bar{\lambda}_2)^2 + (\bar{\lambda}_1 - \bar{\lambda}_2)^2] \geqslant \lambda_1^2 + \lambda_2^2, \tag{44c}$$

where the first inequality is obtained by minimizing each term within the square brackets separately.

Completely device-independent entanglement verification fails because all possible Bell inequalities are satisfied, which is equivalent to a separable quantum representation in the bipartite case [15,16]. ■

IV. QUTRITS AND BEYOND

As shown in the previous section, just having the knowledge that one is measuring a qubit is enough to detect entanglement even if the corresponding Bell inequalities, the set of inequivalent CHSH inequalities, are not violated. Nevertheless, if all these Bell inequalities are satisfied then the observed data can be reproduced by appropriate measurements onto a higher-dimensional separable state [15,16]; for the considered case this would be in dimensions $4 \otimes 4$. Thus the only other nontrivial case is the instance of qutrits. In the following we prove that even the qutrit assumption alone suffices to detect more entanglement than with the CHSH Bell inequalities.

Rather than defining different notions of sharpness or orthogonality for higher-dimensional measurements or more settings, we focus on the completely uncharacterized case of n dichotomic measurements on a d-dimensional system. Each dichotomic measurement is uniquely determined by the operator given by the difference of two POVM elements and is denoted as A_i with $i = 1, \ldots, n$ in the following. The only defining inequality for all these operators A_i, besides that they are all acting on the same d-dimensional Hilbert space $\mathbb{C}^d$, is the condition of Eq. (4) which ensures that they correspond to valid quantum measurements. Similar conditions are imposed for the measurements for Bob labeled as B_i. We employ once more the notion of a data matrix D, each entry defined as $[D]_{ij} = \langle A_i \otimes B_j \rangle$ for $i, j = 0, \ldots, n$, with $A_0 = B_0 = \mathbb{1}$, such that it also contains the observed marginals. Obviously we also have $[D]_{00} = 1$, which we always assume to be fulfilled if we speak about a data matrix. This is again employed to provide a more compact solution, which is stated in the following proposition. In addition, from the above-mentioned qutrit example in the CHSH case, it has a few more consequences which are commented on afterward. The conditions for qubits and qutrits are plotted in Fig. 1.

[5]Note that this is not shown via Proposition 2 because one employs only some restricted information from the observations, solely knowledge of the singular values.

Proposition 4 (Data matrix for n dichotomic measurements on two qudits). If the data matrix D corresponding to n dichotomic measurements satisfies

$$|\det(D)| > \left(\frac{d}{n+1}\right)^{n+1} \tag{45}$$

one verifies entanglement under the assumption of d-dimensional measurements.

If one has at least as many settings as dimensions, i.e., $n \geqslant d$, already the condition

$$|\det(D)| > \left(\frac{d-1}{n}\right)^{n} \tag{46}$$

ensures entanglement. For the special case of two settings $n = 2$ and qutrits $d = 3$ the condition can be improved to

$$|\det(D)| > \tfrac{64}{81}. \tag{47}$$

Proof. We follow a similar proof technique as in the previous section. The data matrix is transformed with appropriate corrections G_a and G_b in order to form a kind of correlation matrix $C = G_a D G_b^T$ for which one employs a known entanglement criterion.

This correlation matrix C is very similar to the previously employed matrix T_3. Here it is defined as $C[\rho]_{ij} = \mathrm{tr}(\rho K_i^{\mathrm{A}} \otimes K_j^{\mathrm{B}})$ where each local set K_i consists of orthonormal (with respect to the Hilbert-Schmidt inner product) observables, i.e., $\mathrm{tr}(K_i K_j) = \delta_{ij}$. Using the inclusion principle in a similar fashion as in the proof of Proposition 1 one finds that the state is entangled if the singular values of C, denoted as c_i, fulfill

$$\|C\|_1 = \sum_{i=0}^{n} c_i > 1. \tag{48}$$

In the following we bound this trace norm by the determinant of the correlation matrix $|\det(C)| = \prod_i c_i$ and the extra knowledge that the largest singular value satisfies $c_0 \geqslant 1/d$ which is implied by the choice $K_0 = \mathbb{1}/\sqrt{d}$ and the inclusion principle. This provides the following estimate:

$$\sum_{i=0}^{n} c_i \geqslant \min_{c_0 \geqslant 1/d} c_0 + \sum_{i=1}^{n} c_i \geqslant \min_{c_0 \geqslant 1/d} c_0 + n\left(\prod_{i=1}^{n} c_i\right)^{1/n} \tag{49a}$$

$$\geqslant \min_{c_0 \geqslant 1/d} c_0 + n\left(\frac{|\det(C)|}{c_0}\right)^{1/n} \tag{49b}$$

$$= \begin{cases} (n+1)|\det(C)|^{1/(n+1)} & \text{if} \quad |\det(C)|^{1/(n+1)} \geqslant \frac{1}{d}, \\ \frac{1}{d} + n\,[d|\det(C)|]^{1/n} & \text{else.} \end{cases} \tag{49c}$$

Here we employed the inequality of arithmetic and geometric means in the second step, while the optimization given by Eq. (49b) is performed using standard analysis. Note that the first solution in Eq. (49c) is the unconstrained optimum that could have been inferred directly by applying the inequality of arithmetic and geometric mean to all terms. However, under the constraint on the largest singular value this solution is only reached if the determinant of the correlation matrix satisfies the stated extra condition. This distinction is necessary for the improved condition in the case $n \geqslant d$.

The remaining strategy is to lower-bound each solution of Eq. (49c) by an expression involving the data matrix. Afterward one investigates which conditions assure that this lower bound actually exceeds 1, such that, in the spirit of Eq. (48), it would signal entanglement. The more stringent of these two cases will be the final entanglement criterion. Here we need the following inequality that relates the determinant of correlation and the data matrix:

$$|\det(C)| = |\det(D)|\,|\det(G_a)|\,\left|\det\left(G_b^T\right)\right| \tag{50a}$$

$$\geqslant |\det(D)| d^{-(n+1)} \tag{50b}$$

which follows from the bound $|\det(G)| \geqslant d^{-(n+1)/2}$ for each of the above-mentioned transformations G_a and G_b, proven in Appendix D.

Let us start with the second solution and employ Eq. (50b), resulting in

$$\frac{1}{d} + n\,[d|\det(C)|]^{1/n} \geqslant \frac{1}{d}\left(1 + n|\det(D)|^{1/n}\right), \tag{51}$$

which is larger than 1 if and only if the condition given by Eq. (46) holds. For the first solution one obtains

$$(n+1)|\det(C)|^{1/(n+1)} \geqslant \frac{n+1}{d}\max\left[1, |\det(D)|^{1/(n+1)}\right]. \tag{52}$$

The first part in the maximum follows from the region constraint on $|\det(C)|$, while the second part is obtained using Eq. (50b). The maximum appears because both bounds are valid. The right-hand side of Eq. (52) is larger than 1 if already one of the terms is. For the general case of this proposition one chooses the second part of this maximum, which is larger than 1 if and only if the determinant of the data matrix satisfies Eq. (45). Because this condition is weaker than the previous condition from Eq. (51), this is the entanglement criterion for the general case. For the special configuration of $n \geqslant d$ we employ the first part of the maximum since it always exceeds 1. Hence only the condition from Eq. (51) is relevant for this case, which proves the proposition of the case $n \geqslant d$.

The improved condition for the qutrit case follows from a sharper lower bound on the transformations G_a and G_b given by $|\det(G)| \geqslant \sqrt{3}/8$, which is proven in Appendix D. ■

It is worth stressing that one can employ Proposition 4 not only with the determinant of the full data matrix D, but also for each subdeterminant. This describes the case that certain measurement settings are left out, which is useful when two or more settings coincide or are linearly dependent, in which case the determinant of the whole data matrix vanishes.

It is interesting that Proposition 4 detects bound entanglement. In what follows we provide an explicit example of a positive partial transpose (PPT) bound entangled state that is detected via the criterion given by Proposition 4, i.e., solely by having knowledge about the underlying dimension. The state

$$\rho_{\mathrm{BFP}} = \tfrac{1}{6}(\Phi^+_{\mathrm{AB}}\Psi^-_{\mathrm{A'B'}} + \Psi^+_{\mathrm{AB}}\Psi^+_{\mathrm{A'B'}} + \Psi^-_{\mathrm{AB}}\Phi^-_{\mathrm{A'B'}} + \Phi^-_{\mathrm{AB}}\Psi^+_{\mathrm{A'B'}} + \Phi^-_{\mathrm{AB}}\Psi^-_{\mathrm{A'B'}} + \Phi^-_{\mathrm{AB}}\Phi^-_{\mathrm{A'B'}}), \tag{53}$$

with $\Phi^+, \ldots, \Psi^-$ denoting the projectors onto the standard two-qubit Bell states, has been shown to be a $4 \otimes 4$ bipartite PPT bound entangled state [31] under the splitting $AA'|BB'$. Assume that one performs "good" enough measurements described by all traceless operators $A_{kl} = \sigma_k^{A} \otimes \sigma_l^{A'}$ built up by tensor products of the identity and the Pauli operators, and the same measurements for Bob. Mixed with white noise $\rho(p) = (1-p)\rho_{BFP} + p\mathbb{1}/16$, the criterion given by Eq. (46) becomes

$$|\det(D)| = \left[\frac{(1-p)}{3}\right]^{15} > \left(\frac{3}{15}\right)^{15}, \qquad (54)$$

and thus verifies entanglement as long as $p < 2/5 = 0.4$. This value seems to coincide with the point where entanglement disappears, i.e., for $p \geqslant 0.41$ the state is separable using the method of Ref. [32]. This detection capability represents a clear advantage over Bell inequalities, where it is still unknown if measurements on a bipartite bound entangled state can violate a Bell inequality at all, though recent results seem to falsify this belief [33].

To conclude this section we consider the two-qubit Werner states $\rho_W(p) = (1-p)\Psi^- + p\mathbb{1}/4$ and three orthogonal standard measurements, i.e., $\sigma_x, \sigma_y, \sigma_z$ for both sides. Then the above-stated conditions verify entanglement as long as the white-noise parameter satisfies $p < 2/3 \approx 0.67$, which coincides with the point where entanglement vanishes. This represents another important improvement with respect to Bell inequalities, since, first, there is hardly any good Bell inequality known that detects entanglement if $p > 1 - 1/\sqrt{2} \approx 0.29$ and, second, above $p \geqslant 7/12 \approx 0.58$ it is known that no measurement would violate a Bell inequality [34].

V. CONCLUSION AND OUTLOOK

We have investigated the task of entanglement detection for cases where only some partial information about the performed measurements is known or assumed. The considered scenarios included properties like sharpness and orthogonality, and cases where only the dimension of the underlying measurements is fixed. Via this extra information one verifies more data as resulting only from entangled states than in the totally device-independent setting while still keeping a good detection strength in comparison to the fully characterized case.

There are many further research lines connecting from here: A thorough investigation of higher-dimensional states and more measurement settings is clearly interesting in order to clarify the power but also the limitations of this intermediate approach. For that one should be aware that the current methods represent only the first steps toward these directions. Here alternative tools might be necessary; even an efficient numerical approach would be of great help. Another approach would be the investigation of detection methods for the multipartite case in a similar intermediate setting. Since our criteria rest on an entanglement criterion based on a correlation matrix, this extension might be possible utilizing recent detection methods for genuine multipartite entanglement using the correlation tensor [35,36]. Regarding explicit tasks, since our results show that one verifies entanglement even if the underlying Bell inequality is not violated, this means that the lower bound on the concurrence given in Ref. [37] could be improved. This partially characterized scenario might also be useful in order to obtain steering inequalities which are more robust against calibration errors, in a similar spirit as in Ref. [38]. Finally it is tempting to apply these result also to quantum key distribution operated in a similar intermediate setting as described here, which has already been started in Ref. [39]. For that in particular Proposition 3 is interesting, because it describes exactly the kind of observations that one expects in an entanglement-based Bennett-Brassard 1984 (BB84) protocol. Since it states that one verifies the same fraction of entanglement as with totally characterized measurements, this hints that a very strong "semi-device-independent" key rate could be obtained if one just possesses the knowledge that one measures a qubit. This would allow a much larger freedom in finding appropriate squash models for quantum key distribution since the measurements no longer have to be fixed [21].

ACKNOWLEDGMENTS

We thank J. I. de Vicente, O. Gühne, B. Jungnitsch, and N. Lütkenhaus for stimulating discussions, in particular J. I. de Vicente for much help regarding the more general case. This work has been supported by the FWF (START prize No. Y376-N16) and the EU (Marie Curie Grant No. CIG 293993/ENFOQI). O.G. is grateful for support from the Industry Canada and NSERC Strategic Project Grant (SPG) FREQUENCY.

APPENDIX A: PROOF OF PROPOSITION 1

In this Appendix we provide the proof of the CCNR criterion adapted to our partial information setting.

Proof. It is only necessary to consider the case of partial information, since the CCNR criterion is typically formulated in terms of the full density operator. Following Ref. [40] the CCNR criterion can be expressed as follows.

Let T_4 denote the complete correlation matrix of two qubits, which results from T_3 by addition of the remaining Pauli operator for each local side. Suppose that its corresponding ordered singular values are denoted as $t_0 \geqslant t_1 \geqslant t_2 \geqslant t_3 \geqslant 0$. Then the CCNR criterion states that for any separable state these singular values fulfill $\sum_{i=0}^{3} t_i \leqslant 2$.

According to the inclusion principle (cf. Corollary 3.1.3 of Ref. [41]), the ordered singular values are lower bounded by the singular values of any submatrix. Thus one obtains $t_i \geqslant \lambda_i$ for $i = 0,1,2$, such that one arrives at

$$\lambda_0 + \lambda_1 + \lambda_2 \leqslant \sum_{i=0}^{3} t_i \leqslant 2. \qquad (A1)$$

Whenever this condition is violated the state must necessarily be entangled.

In the case of vanishing marginals with $\lambda_0 = 1$ the condition transforms into $\lambda_1 + \lambda_2 \leqslant 1$. Note that the parameters $\lambda_{1/2} \leqslant 1$ are also the singular values of the submatrix T_2. Using similar techniques as presented in Ref. [42], it is possible to fit a rotated Bell-diagonal separable state to these data. This is achieved along the following lines: First, consider another correlation matrix $T[\rho]_{\alpha\beta} = \mathrm{tr}(\rho\sigma_\alpha \otimes \sigma_\beta)$ which is built up by the standard Pauli operators $\{\sigma_x,\sigma_y,\sigma_z\}$ for each local side. Assume that the given submatrix T_2 is precisely the upper-two block of such a correlation matrix, i.e., $T = \text{diag}[T_2,0]$ filled with additional zero entries. The singular value decomposition of this correlation matrix is given by $T = O_a \Lambda O_b$ with singular values $\Lambda = \text{diag}[\lambda_1,\lambda_2,0] \geqslant 0$ and O_a and O_b being special orthogonal matrices of similar block-diagonal form, i.e., $O_a = \text{diag}[\bar{O}_a, \pm 1]$ and an analogous form of O_b. Here, note that $\bar{O}_a$ and $\bar{O}_b$ are the (not necessarily special) orthogonal matrices from the singular value decomposition of $T_2 = \bar{O}_a \text{diag}[\lambda_1,\lambda_2]\bar{O}_b^T$.

Next let us discuss the special case of a diagonal correlation matrix, i.e., $T[\rho] = \text{diag}[\lambda_1,\lambda_2,0]$: These data correspond to a Bell-diagonal state, abstractly expressed as $\rho_{\text{bs}} = \sum_i p_i |bs_i\rangle\langle bs_i|$ with standard Bell states $|bs_i\rangle$ and appropriate weights equal to $p_i = (1 \pm \lambda_1 \pm \lambda_2)/4$ having all four combinations. The above-stated condition ensures that all probabilities are indeed non-negative and are upper bounded by $1/2$, which ensure separability in this case [42].

For the general case one employs the relation that any special orthogonal transformed correlation matrix corresponds to some special unitary transformation on the level of quantum states [42],

$$O_a T[\rho] O_b^T = T[U_a \otimes U_b \rho U_a^\dagger \otimes U_b^\dagger]. \tag{A2}$$

Since local unitary transformations do not change the entanglement properties, the appropriately transformed state $U_a \otimes U_b \rho_{\text{bs}} U_a^\dagger \otimes U_b^\dagger$ is the actual separable state for the general correlation matrix T. Finally, this rotation property is exploited once more to support the assumption that T_2 is the submatrix of the first two rows of T, since any two orthogonal vectors can be rotated such that it matches these axis. This finally proves the claim. ■

APPENDIX B: DETAILS FOR TWO-QUBIT CASE

1. Relations for sharp measurements, Eqs. (21) and (22)

Equation (21) is a direct consequence of the determinant multiplication rule, i.e., $\det(AB) = \det(A)\det(B)$, and of the fact that the absolute value of the determinant is equal to the product of its singular values. In order to derive the second inequality one employs the singular value identities

$$\sum_{i=1}^{k} \sigma_i(AB) \leqslant \sum_{i=1}^{k} \sigma_i(A)\sigma_i(B), \tag{B1}$$

$$\prod_{i=1}^{k} \sigma_i(AB) \leqslant \prod_{i=1}^{k} \sigma_i(A)\sigma_i(B), \tag{B2}$$

where $\sigma_i(\cdot)$ denotes the decreasing ordered singular values; cf. Ref. [41]. Application of these identities to Eq. (19) leads to

$$\begin{aligned}
\lambda_1 + \lambda_2 &\leqslant b_1\sigma_1\{[R_2(\alpha)]^{-1}T_2\} + b_2\sigma_2\{[R_2(\alpha)]^{-1}T_2\} && \text{(B3a)}\\
&\leqslant b_1\sigma_1\{[R_2(\alpha)]^{-1}T_2\} \\
&\quad + b_2\{a_1t_1 + a_2t_2 - \sigma_1[R_2(\alpha)^{-1}T_2]\} && \text{(B3b)}\\
&\leqslant (a_1b_1)t_1 + (a_2b_2)t_2 && \text{(B3c)}\\
&\leqslant \sqrt{\left(a_1^2b_1^2 + a_2^2b_2^2\right)\left(t_1^2 + t_2^2\right)}, && \text{(B3d)}
\end{aligned}$$

where the Cauchy-Schwarz inequality is applied in the last step.

2. Relations for orthogonal measurements, Eqs. (25)–(27)

The first condition given by Eq. (25) follows from the inclusion principle [41] using the first entry as a submatrix. The last inequality Eq. (27) holds because of the determinant multiplication rule,

$$\begin{aligned}
t_0t_1t_2 &= |\det[S(\vec{x})]||\det(D_3)||\det[S(\vec{y})]| && \text{(B4a)}\\
&= \left|\det\left(S_x D_2 S_y^T\right)\right| = (x_2x_4)(\lambda_1\lambda_2)(y_2y_4) && \text{(B4b)}\\
&\geqslant \lambda_1\lambda_2, && \text{(B4c)}
\end{aligned}$$

and the bounds on the appearing parameters, e.g., $x_2 \geqslant 1 + |x_1| \geqslant 1$.

In order to prove Eq. (26) we apply the inclusion principle to a particular chosen 2×2 submatrix. For this argument the matrix $\bar{D}_2 = S_x D_2 S_y^T$ attains special importance. First, note that the singular values of $\bar{D}_2$, denoted as $\bar{\lambda}_1 \geqslant \bar{\lambda}_2$, satisfy $\bar{\lambda}_i \geqslant \lambda_i$ because the transformations satisfy $S_x,S_y - \mathbb{1} \geqslant 0$. Next, employ the singular value decomposition $\bar{D}_2 = U\Sigma V^T$ with $\Sigma = \text{diag}[\bar{\lambda}_1,\bar{\lambda}_2]$. Since the singular values of T_3 remain invariant under orthogonal transformations, we apply appropriate orthogonal matrices to diagonalize $\bar{D}_2$, which leads to

$$\begin{bmatrix} 1 & \\ & U^T \end{bmatrix} T_3 \begin{bmatrix} 1 & \\ & V \end{bmatrix} = \begin{bmatrix} 1 & \bar{y}^T \\ \bar{x}^T & \bar{x}\bar{y}^T + \Sigma \end{bmatrix}, \tag{B5}$$

with $\bar{x} = U^T x, \bar{y} = V^T y$. Using the submatrix formed by the first two rows and columns one obtains

$$\begin{bmatrix} 1 & \bar{y}_1 \\ \bar{x}_1 & \bar{x}_1\bar{y}_1 + \bar{\lambda}_1 \end{bmatrix}, \tag{B6}$$

which has determinant $\bar{\lambda}_1 \geqslant \lambda_1$. Then the inclusion principle directly states Eq. (26).

APPENDIX C: OPTIMIZATION PROBLEMS

In this Appendix we prove two lemmas concerning optimization problems appearing in the proof of Proposition 2. They are mainly given for completeness of the paper.

Lemma 1 Suppose $\lambda_1,\lambda_2 \geqslant 0$. Then the solution of

$$\min_{\alpha,\beta}\left[\frac{(\lambda_1+\lambda_2)^2}{a_1^2b_1^2 + a_2^2b_2^2} + 2\frac{\lambda_1\lambda_2}{a_1a_2b_1b_2}\right] \tag{C1}$$

with $a_1 = \sqrt{2\cos(\alpha)^2} \geqslant a_2 = \sqrt{2\sin(\alpha)^2}$ and similarly for b_i with another angle β is

$$\tfrac{1}{4}(\sqrt{\lambda_1} + \sqrt{\lambda_2})^4. \tag{C2}$$

Proof. First note that the ordering of the singular values does not modify the solution if the optimization is performed over the full period of each angle, since wrong ordering only leads to larger function values. Use of the parametrization $\alpha = \gamma + \delta$, $\beta = \gamma - \delta$ simplifies the function to

$$\frac{(\lambda_1+\lambda_2)^2}{2 + \cos(4\gamma) + \cos(4\delta)} + \frac{4\lambda_1\lambda_2}{|\cos(4\gamma) - \cos(4\delta)|}. \tag{C3}$$

Hence it effectively depends only on $u = \cos(4\gamma)$ and $v = \cos(4\delta)$, which are both in the interval $[-1,1]$. Note that the boundary of this feasible set is characterized by either u or v taking on the extreme values of this interval. In the following we prove that the minimum lies at this boundary.

Use of the linear variable transformation $x = u + v$, $y = u - v$ changes the function to

$$\frac{(\lambda_1 + \lambda_2)^2}{2 + x} + \frac{4\lambda_1\lambda_2}{|y|}. \tag{C4}$$

Now, taking partial derivatives, one finds that this function is decreasing with respect to x in the interval $(-2,2]$ and depending on the sign of y also decreasing in the $+y$ or $-y$ direction (the exceptional cases $y = 0$ or $x = -2$ can be excluded since they have no minima). Since the variables x and y are bounded this shows that the minimum is attained at the boundary. Going back to the form given by Eq. (C3) means that either $\cos(4\gamma) = \pm 1$ or $\cos(4\delta) = \pm 1$. Here it does not matter which cosine is put to its extreme values since one needs to consider only one of them. Only one of these values ± 1 leads to the solution, for the other one can directly verify that its solution is larger than given by Eq. (C2). Concluding, only the following function needs to be optimized over the angle ψ:

$$\frac{(\lambda_1 + \lambda_2)^2}{3 + \cos(\psi)} + \frac{4\lambda_1\lambda_2}{1 - \cos(\psi)}. \tag{C5}$$

This optimization can be performed directly by looking for the vanishing derivatives and using only the real solutions. This finally leads to the solution given by Eq. (C2). ■

Lemma 2. Suppose $\lambda_0 \geqslant \lambda_1 \geqslant \lambda_2 \geqslant 0$. Then the solution of

$$\begin{aligned} \min\ & \mu_0 + \mu_1 + \mu_2 \\ \text{s. t.}\ & \mu_0 \geqslant \lambda_0, \quad \mu_0\mu_1 \geqslant \lambda_0\lambda_1, \\ & \mu_0\mu_1\mu_2 \geqslant \lambda_0\lambda_1\lambda_2, \\ & \mu_0 \geqslant \mu_1 \geqslant \mu_2 \geqslant 0, \end{aligned} \tag{C6}$$

is $\lambda_0 + \lambda_1 + \lambda_2$.

Proof. In order to prove the lemma let us first state the solution of the following subproblem:

$$\min_{x \geqslant x_{\min} \geqslant 0} x + \frac{\lambda}{x} = \begin{cases} 2\sqrt{\lambda} & \text{if } \sqrt{\lambda} > x_{\min}, \\ x_{\min} + \frac{\lambda}{x_{\min}} & \text{else,} \end{cases} \tag{C7}$$

with $\lambda > 0$ which follows using standard analysis. Note that $\sqrt{\lambda}$ is the argument of the only positive constrained minimum.

Now we turn to the intended problem given by Eq. (C6). First consider the case that the parameter μ_0 is fixed and that we bound the sum of the remaining two parameters from below, which provides

$$\min_{\mu_1,\mu_2} \mu_1 + \mu_2 \geqslant \min_{\mu_1 \geqslant \bar{\mu}_1} \mu_1 + \left(\frac{\lambda_0\lambda_1\lambda_2}{\mu_0}\right)\frac{1}{\mu_1} \tag{C8a}$$

$$\geqslant \begin{cases} 2\sqrt{\frac{\lambda_0\lambda_1\lambda_2}{\mu_0}} & \text{if} \quad \mu_0 \geqslant \bar{\mu}_0, \\ \frac{\lambda_0\lambda_1}{\mu_0} + \lambda_2 & \text{if} \quad \lambda_0 \leqslant \mu_0 \leqslant \bar{\mu}_0, \end{cases} \tag{C8b}$$

using the abbreviations $\bar{\mu}_1 = \lambda_0\lambda_1/\mu_0$ and $\bar{\mu}_0 = \lambda_0\lambda_1/\lambda_2$. In the first inequality we employ the lower bound on μ_2 from the problem formulation. The second inequality is an application of the subproblem in which the conditions are reexpressed in terms of the parameter μ_0. In the following we use these lower bounds and consider variations of μ_0 within the corresponding valid region. For simplicity let us assume strict inequality $\lambda_0 > \lambda_1 > \lambda_2$; the statement with equality of some or all parameters follows then by continuity.

Let us start with the case $\lambda_0 \leqslant \mu_0 \leqslant \bar{\mu}_0$. Use of the derived lower bound gives

$$\min_{\substack{\mu_1,\mu_2 \\ \lambda_0 \leqslant \mu_0 \leqslant \bar{\mu}_0}} \mu_0 + \mu_1 + \mu_2 \geqslant \min_{\mu_0 \geqslant \lambda_0} \mu_0 + \frac{\lambda_0\lambda_1}{\mu_0} + \lambda_2 \tag{C9a}$$

$$\geqslant \lambda_0 + \lambda_1 + \lambda_2, \tag{C9b}$$

via another application of the given subproblem. Note that the unconstrained minimum is not achieved, i.e., $\sqrt{\lambda_0\lambda_1} \not\geqslant \lambda_0$ because of the strict ordering.

Next consider the case $\mu_0 \geqslant \bar{\mu}_0$, for which we have to employ the other lower bound in Eq. (C8b). This leads to

$$\min_{\substack{\mu_1,\mu_2 \\ \mu_0 \geqslant \bar{\mu}_0}} \mu_0 + \mu_1 + \mu_2 \geqslant \min_{\mu_0 \geqslant \bar{\mu}_0} \mu_0 + 2\sqrt{\frac{\lambda_0\lambda_1\lambda_2}{\mu_0}} \tag{C10a}$$

$$\geqslant \frac{\lambda_0\lambda_1}{\lambda_2} + 2\lambda_2 > \lambda_0 + \lambda_1 + \lambda_2. \tag{C10b}$$

The optimization problem appearing in Eq. (C10a) is very similar to our subproblem, in particular it is convex again for positive μ_0. Its only positive constrained minimum is at the argument $\sqrt[3]{\lambda_0\lambda_1\lambda_2}$ which, however, is outside the allowed region, i.e., $\bar{\mu}_0 > \sqrt[3]{\lambda_0\lambda_1\lambda_2}$ due to the strict ordering of the λ's. Thus the optimum is attained at the boundary $\mu_0 = \bar{\mu}_0$. The last inequality represents another consequence of the ordering property since it is equivalent to $\lambda_0(\lambda_1 - \lambda_2) > \lambda_2(\lambda_1 - \lambda_2)$. ■

APPENDIX D: TRANSFORMATION DETERMINANTS

In this Appendix we provide a proof for the bounds on the determinant of the transformations G used in the proof of Proposition 4. It is a direct consequence of the Gram-Schmidt procedure.

Lemma 3. The linear transformation G that maps the operators from the data matrix $\{A_i\}$ all acting on $\mathbb{C}^d$ with $K_0 = \mathbb{1}$ and $-\mathbb{1} \leqslant A_i \leqslant \mathbb{1}$ for all $i = 1, \ldots, n$ into an orthonormal operator set $\{K_i\}$, i.e., $\text{tr}(K_i K_j) = \delta_{ij}$, fulfills

$$|\det(G)| \geqslant d^{-(n+1)/2}. \tag{D1}$$

For the case $n = 2$, $d = 3$ the bound can be improved to

$$|\det(G)| \geqslant \frac{\sqrt{3}}{8}. \tag{D2}$$

Proof. The linear operation G can be obtained, for instance, using the Gram-Schmidt process. Without loss of generality this procedure can be decomposed into two operations $G = ON$. The first transformation N should map each operator A_i to its normalized form $\tilde{A}_i = A_i/\sqrt{\text{tr}(A_i^2)}$, such that the second transformation O only needs to orthogonalize them. This second linear operation always stretches the "vectors" $\tilde{A}_i$ such that the volume spanned by this set always increases; therefore $|\det(O)| \geqslant 1$. The first transformation N is a diagonal matrix with entries given by $1/\sqrt{\mathrm{tr}(A_i^2)}$. Since each operator A_i is bounded by the identity due to the restriction that it describes a valid measurement, one obtains $\mathrm{tr}A_i^2 \leqslant \mathrm{tr}\mathbb{1} \leqslant d$ or finally

$$|\det(G)| = |\det(O)|\,|\det(N)| \geqslant d^{-(n+1)/2}. \quad \text{(D3)}$$

For the special case of $n = 2$ and $d = 3$ we explicitly carry out the Gram-Schmidt process and minimize the determinant under the given constraints. We consider the case that G maps the operators to the orthonormal set $\{K_0 = \mathbb{1}/\sqrt{3}, K_1, K_2\}$. This resulting operation G is of triangular form, i.e.,

$$G_a = \begin{bmatrix} \frac{1}{\sqrt{3}} & & \\ * & N_1 & \\ * & * & N_2 \end{bmatrix}, \quad \text{(D4)}$$

which has a determinant $N_1 N_2/\sqrt{3}$. In order to obtain a lower bound one needs to minimize $N_1 N_2$ or maximize $1/(N_1 N_2)^2$, which results in

$$\left|\left[\mathrm{tr}(A^2) - \frac{[\mathrm{tr}(A)]^2}{3}\right]\left[\mathrm{tr}(A'^2) - \frac{[\mathrm{tr}(A')]^2}{3}\right]\right| - \left[\mathrm{tr}(AA') - \frac{\mathrm{tr}(A)\mathrm{tr}(A')}{3}\right]^2 \quad \text{(D5a)}$$

$$\leqslant \left|\left[\mathrm{tr}(A^2) - \frac{[\mathrm{tr}(A)]^2}{3}\right]\left[\mathrm{tr}(A'^2) - \frac{[\mathrm{tr}(A')]^2}{3}\right]\right| \leqslant \left(\frac{8}{3}\right)^2. \quad \text{(D5b)}$$

The last inequality originates from the bound

$$\mathrm{tr}(A^2) - \frac{[\mathrm{tr}(A)]^2}{3} \quad \text{(D6)}$$

$$= \frac{2}{3}\left(l_1^2 + l_2^2 + l_3^2 - l_1 l_2 - l_1 l_3 - l_2 l_3\right) \leqslant \frac{8}{3}, \quad \text{(D7)}$$

where l_i are the eigenvalues of A that satisfy the box constraint $|l_i| \leqslant 1$ due to the measurement condition of Eq. (4). In this expression at most two of the last three terms can be positive but one must necessarily be negative. Suppose that this term is $-l_2 l_3 < 0$; then $l_2^2 + l_3^2 - l_2 l_3 \leqslant 1$ holds given the stated box constraints. ■

[1] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009).
[2] O. Gühne and G. Tóth, Phys. Rep. **474**, 1 (2009).
[3] S. J. van Enk, N. Lütkenhaus, and H. J. Kimble, Phys. Rev. A **75**, 052318 (2007).
[4] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1 (1996).
[5] B. Terhal, Phys. Lett. A **271**, 319 (2000).
[6] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri, Phys. Rev. A **69**, 022308 (2004).
[7] B. Qi, C. H. F. Fung, H. K. Lo, and X. Ma, Quantum Inf. Comput. **7**, 73 (2007).
[8] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov, Nat. Photonics **4**, 686 (2010).
[9] J. S. Bell, Physics (Long Island City, NY) **1**, 195 (1964).
[10] J. F. Clauser and M. A. Horne, Phys. Rev. D **10**, 526 (1974).
[11] A. Peres, Found. Phys. **29**, 589 (1999).
[12] J. D. Bancal, N. Gisin, Y.-C. Liang, and S. Pironio, Phys. Rev. Lett. **106**, 250404 (2011).
[13] K. F. Pál and T. Vértesi, Phys. Rev. A **83**, 062123 (2011).
[14] N. Brunner, J. Sharam, and T. Vértesi, e-print arXiv:1110.5512 [Phys. Rev. Lett. (to be published)].
[15] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).
[16] A. Acín, N. Gisin, and L. Masanes, Phys. Rev. Lett. **97**, 120405 (2006).
[17] H. M. Wiseman, S. J. Jones, and A. C. Doherty, Phys. Rev. Lett. **98**, 140402 (2007).
[18] S. M. Roy, Phys. Rev. Lett. **94**, 010402 (2005).
[19] J. Uffink and M. Seevinck, Phys. Lett. A **372**, 1205 (2008).
[20] P. Lougovski and S. J. van Enk, Phys. Rev. A **80**, 034302 (2009).
[21] N. J. Beaudry, T. Moroder, and N. Lütkenhaus, Phys. Rev. Lett. **101**, 093601 (2008).
[22] T. Tsurumaru and K. Tamaki, Phys. Rev. A **78**, 032302 (2008).
[23] T. Moroder, O. Gühne, N. J. Beaudry, M. Piani, and N. Lütkenhaus, Phys. Rev. A **81**, 052342 (2010).
[24] E. Shchukin and W. Vogel, e-print arXiv:0902.3962.
[25] A. A. Semenov and W. Vogel, Phys. Rev. A **83**, 032119 (2011).
[26] I. Gerhardt, Q. Liu, A. Lamas-Linares, J. Skaar, V. Scarani, V. Makarov, and C. Kurtsiefer, Phys. Rev. Lett. **107**, 170404 (2011).
[27] M. Curty, M. Lewenstein, and N. Lütkenhaus, Phys. Rev. Lett. **92**, 217903 (2004).
[28] N. Brunner, S. Pironio, A. Acín, N. Gisin, A. A. Méthot, and V. Scarani, Phys. Rev. Lett. **100**, 210503 (2008).
[29] O. Rudolph, Quantum Inf. Process. **4**, 219 (2005).
[30] K. Chen and L.-A. Wu, Quantum Inf. Comput. **3**, 193 (2003).
[31] F. Benatti, R. Floreanini, and M. Piani, Open Syst. Inf. Dyn. **11**, 325 (2004).
[32] J. T. Barreiro, P. Schindler, O. Gühne, T. Monz, M. Chwalla, C. F. Roos, M. Hennrich, and R. Blatt, Nat. Phys. **6**, 943 (2010).
[33] T. Vértesi and N. Brunner, Phys. Rev. Lett. **108**, 030403 (2012).
[34] J. Barrett, Phys. Rev. A **65**, 042302 (2002).
[35] J. I. de Vicente and M. Huber, Phys. Rev. A **84**, 062306 (2011).
[36] W. Laskowski, M. Markiewicz, T. Paterek, and M. Zukowski, Phys. Rev. A **84**, 062305 (2011).
[37] Y.-C. Liang, T. Vértesi, and N. Brunner, Phys. Rev. A **83**, 022108 (2011).

[38] D. H. Smith, G. Gillett, M. de Almeida, C. Branciard, A. Fedrizzi, T. J. Weinhold, A. Lita, B. Calkins, T. Gerrits, S. W. Nam, and A. G. White, Nat. Commun. **3**, 625 (2012).

[39] M. Pawlowski and N. Brunner, Phys. Rev. A **84**, 010302(R) (2011).

[40] O. Gühne, M. Mechler, G. Tóth, and P. Adam, Phys. Rev. A **74**, 010301(R) (2006).

[41] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis* (Cambridge University Press, Cambridge, 1991).

[42] R. Horodecki and M. Horodecki, Phys. Rev. A **54**, 1838 (1996).