

## Minimal scheme for certifying three-outcome qubit measurements in the prepare-and-measure scenario

Jonathan Steinberg \* and H. Chau Nguyen*Naturwissenschaftlich-Technische Fakultät, Universität Siegen, 57068 Siegen, Germany*

Matthias Kleinmann

*Naturwissenschaftlich-Technische Fakultät, Universität Siegen, 57068 Siegen, Germany  
and Faculty of Physics, University of Duisburg–Essen, 47048 Duisburg, Germany*

(Received 27 May 2021; accepted 7 December 2021; published 17 December 2021)

The number of outcomes is a defining property of a quantum measurement, in particular, if the measurement cannot be decomposed into simpler measurements with fewer outcomes. Importantly, the number of outcomes of a quantum measurement can be irreducibly higher than the dimension of the system. The certification of this property is possible in a semi-device-independent way either based on a Bell-like scenario or by utilizing the simpler prepare-and-measure scenario. Here we show that in the latter scenario the minimal scheme for certifying an irreducible three-outcome qubit measurement requires three state preparations and only two measurements and we provide experimentally feasible examples for this minimal certification scheme. We also discuss the dimension assumption characteristic of the semi-device-independent approach and to what extent it can be mitigated.

DOI: [10.1103/PhysRevA.104.062431](https://doi.org/10.1103/PhysRevA.104.062431)

### I. INTRODUCTION

The most general description of a quantum measurement is given by a positive-operator-valued measure (POVM) which is any collection of positive operators summing up to identity. The set of measurements described in this way contains instances which are neither projective nor obtainable by combining projective measurements. This class of genuinely nonprojective measurements can be utilized, for example, in quantum computing [1,2], quantum cryptography [3,4], randomness certification [5], and quantum tomography [6]. However, since genuinely nonprojective measurements cannot be combined from projective measurements, their experimental implementation is difficult and typically requires control over additional degrees of freedoms [7]. It is hence of interest to verify whether an experiment had successfully implemented a nonprojective measurement. Recently, semi-device-independent certification schemes have become the focus of theoretical investigation [8–10] as well as experimental implementation [11–13]. The employed certification schemes can be divided into two classes, those based on Bell-like scenarios [11] and those using a prepare-and-measure scenario [8–10,12].

In the former case, an entangled state is distributed to two spatially separated measurement stations and the correlations between the different measurements at each station can then certify the presence of a genuinely nonprojective measurement. In the latter case, the certification consists of several preparation procedures, possibly intermediate transforma-

tions, and subsequent measurements on the same system. Both scenarios are device independent because only very rudimentary assumptions need to be made about the implementation details of the state preparation and measurement devices. However, because nonprojective measurements can always be implemented on a system with enlarged Hilbert space, it is common to both scenarios that they add an extra assumption, namely, an upper limit on the dimension of the system, rendering the scenarios semi-device-independent.

It should be noted that the certification of a genuinely nonprojective measurements is usually performed in a slightly different context where the number of outcomes of a measurement plays a key role [8–12]. If a measurement cannot be implemented by using measurements with a lower number of outcomes, then the number of outcomes is irreducible. Since projective measurements cannot have more outcomes than the dimension of the system, it is hence sufficient to certify excess outcomes in order to certify that a measurement is also genuinely nonprojective. For the purposes of this paper, where we consider three-outcome measurements on a qubit, this distinction is not relevant, because for qubits, the number of outcomes of a measurement is irreducibly 3 if and only if the measurement is genuinely nonprojective.

In the prepare-and-measure scenario, so far, the certification schemes for three-outcome and four-outcome qubit measurements are based on linear inequalities satisfied for all correlations requiring fewer outcomes. The schemes use at least two additional measurement settings [12] to achieve a certification. In this paper we do not restrict ourselves to linear inequalities and we show that the minimal scheme consists of three different state preparations and only one auxiliary measurement. Our analysis is complemented with examples

\*steinberg@physik.uni-siegen.de

which allow a simple certification of a three-outcome qubit measurement.

Our paper is organized as follows. In Sec. II we revisit the concept of prepare-and-measure scenarios and fix our notation and terminology. In Sec. III we define the operational setup in which a three-outcome qubit measurement can be certified. We prove the necessity of one auxiliary measurement and three preparation procedures and give corresponding examples. We proceed in Sec. IV by discussing a possible mitigation of the assumption of an upper bound on the dimension before we conclude with a discussion in Sec. V.

## II. PREPARE-AND-MEASURE SCENARIO

A prepare-and-measure scenario can be understood as a setup that is composed solely of a preparation device and a measurement device. An experimenter can choose among  $s$  different preparation procedures labeled by  $x \in \{1, \dots, s\}$  and  $m$  different measurements labeled by  $y \in \{1, \dots, m\}$ . After choosing one particular pair  $(x, y)$  the experimenter produces the state  $x$  on which the measurement  $y$  is performed. Consequently, this experiment can be described by the experimentally accessible correlations  $p(a|x, y)$  which give the probability of obtaining outcome  $a$  when performing measurement  $y$  on preparation  $x$ . It is important to note that the preparation as well as the measurement apparatuses are considered as a black box, that is, no assumptions are made about the state  $\rho_x$  and the measurement description  $M_y$ . This is with the exception that we assume that the dimension of the underlying Hilbert space is fixed. In this sense, properties of the system that can be deduced from the experimental data alone are semi-device-independent.

### A. Structure of the measurements

Any  $n$ -outcome measurement  $M$  on a  $d$ -dimensional Hilbert space can be written as a POVM, that is, as positive-semidefinite operators  $M = (M_1, \dots, M_n)$  satisfying  $\sum_a M_a = \mathbb{1}$ . The operators  $M_a$  are the effects of  $M$  and are associated with the outcomes  $a \in \{1, \dots, n\}$  of  $M$ . The set of all POVMs is convex, that is, it is closed with respect to taking probabilistic mixtures and efficient algorithms are known in order to decompose a given POVM into extremal POVMs [14]. By virtue of the Born rule, the probability  $p(a)$  of obtaining outcome  $a$  for a quantum state  $\rho$  is given by  $p(a) = \text{tr}(\rho M_a)$ .

It is interesting to notice that a similar notion of POVMs can also be introduced in classical probability theory. Here a general  $n$ -outcome measurement on a  $d$ -dimensional classical system is given by  $n$  vectors from the  $d$ -dimensional unit cube  $C_d = \{\mathbf{x} \in \mathbb{R}^d \mid 0 \leq x_i \leq 1\}$  such that they sum up to  $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^d$ . Therefore, the set of all classical  $n$ -outcome measurements is equivalent to the set of all right stochastic  $d \times n$  matrices. We mention that this classical case is identical to the quantum case when one restricts all effects and states to be diagonal in some fixed basis.

When a measurement has only two nonzero outcomes the measurement is dichotomic, for three nonzero outcomes it is trichotomic, and it is  $n$ -chotomic in the case of  $n$  nonzero outcomes. An  $n$ -outcome POVM  $M$  can be simulated [15]

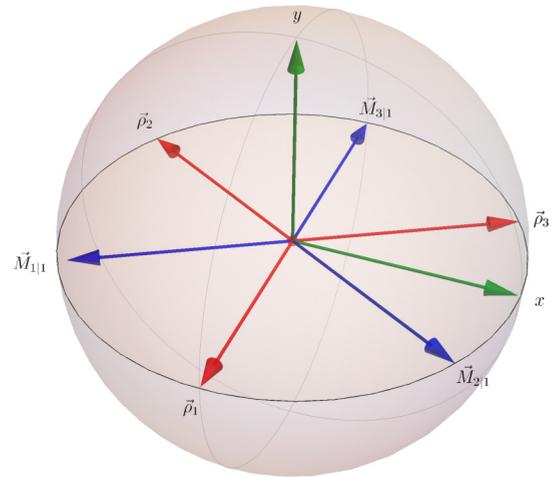


FIG. 1. Representation of the state and effect configuration for generating trichotomic correlations with the trine POVM. The states  $\rho_x$  are obtained from the vectors labeled by  $\vec{\rho}_x$  via  $\rho_x = \frac{1}{2}(\mathbb{1} + \sum_k [\vec{\rho}_x]_k \sigma_k)$  for  $x \in \{1, 2, 3\}$  and the effects  $M_{a|1}$  are obtained from the vectors labeled by  $\vec{M}_{a|1}$  via  $M_{a|1} = \frac{1}{3}(\mathbb{1} + \sum_k [\vec{M}_{a|1}]_k \sigma_k)$  for  $a \in \{1, 2, 3\}$ .

with  $n'$ -chotomic POVMs  $(N_\ell)_\ell$  if there exists a probability distribution  $(p_\ell)_\ell$  such that  $M = \sum_\ell p_\ell N_\ell$ . Otherwise the measurement is irreducibly  $n$ -chotomic. For  $n = 3$  and  $n' = 2$  the simulation reduces to the randomization of three dichotomic measurements, that is,

$$(M_1, M_2, M_3) = p_1(N_{1|1}, N_{2|1}, 0) + p_2(0, N_{2|2}, N_{3|2}) + p_3(N_{1|3}, 0, N_{3|3}), \quad (1)$$

where we write  $N_{a|\ell}$  for the outcome  $a$  of the measurement  $N_\ell$ . These reducible three-outcome measurements form a convex subset of the set of all measurements. While in the  $d$ -dimensional classical probability theory all measurements are reducible to  $d$ -outcome measurements, this is not the case in quantum theory [16]. An archetypal counterexample is the trine POVM  $S$  which is composed of three qubit effects  $S_a = \frac{2}{3}|S_a\rangle\langle S_a|$ , where the  $|S_a\rangle$  for  $a = 1, 2, 3$  are located in a plane of the Bloch sphere and are rotated by an angle of  $\frac{2}{3}\pi$  against each other (see Fig. 1).

### B. Unambiguous state discrimination

Unambiguous state discrimination [17] (USD) is a special instance of quantum state estimation. While the power of quantum state tomography relies on the access to a sufficiently high number of identically prepared quantum states, here only a single copy of the input state is available. In what follows we formulate the task of USD for the special case where the system subject to discrimination is prepared in one of two pure states. Then one party (Alice) randomly but with equal probability chooses one of two states  $|\psi_1\rangle$  and  $|\psi_2\rangle$  which are known to both parties and sends it to a receiver (Bob). By measuring the incoming state, Bob must either correctly identify the state or declare that he does not know the answer, yielding an inconclusive result. Naturally, as soon as Alice's two states are not perfectly distinguishable,  $\langle\psi_1|\psi_2\rangle \neq 0$ , Bob

cannot archive a unit success rate. It is well known [17] that for qubits Bob's best measurement is given by the irreducibly trichotomic POVM  $\mathbf{M} = (M_1, M_2, M_3)$  with

$$M_1 = \frac{\mathbb{1} - |\psi_2\rangle\langle\psi_2|}{1 + |\langle\psi_1|\psi_2\rangle|}, \quad M_2 = \frac{\mathbb{1} - |\psi_1\rangle\langle\psi_1|}{1 + |\langle\psi_1|\psi_2\rangle|}, \quad (2)$$

and  $M_3 = \mathbb{1} - M_1 - M_2$ . With this construction one can see directly that if the measurement yields the outcome  $a = 1$  or  $a = 2$ , then one can conclude that the received state was  $|\psi_a\rangle$ , while outcome  $a = 3$  does not allow a definite statement.

Motivated by the above considerations, we define now a family of correlations that is particularly useful for our later analysis. We consider a prepare-and-measure scenario with three preparations  $\rho_1, \rho_2$ , and  $\rho_3$  and two measurements  $\mathbf{M}_1 = (M_{1|1}, M_{2|1}, M_{3|1})$  and  $\mathbf{M}_2 = (M_{1|2}, M_{2|2})$ . The states and the measurements are chosen in such a way that if  $\mathbf{M}_2$  yields the outcome 1 this implies that the received state was not  $\rho_1$  and if the obtained state is  $\rho_2$ , then  $\mathbf{M}_2$  produces with certainty outcome 2. For  $\mathbf{M}_1$  we impose that from outcome 1 (2) it follows that the state was not  $\rho_3$  ( $\rho_2$ ). Since any POVM has to obey the normalization condition, the last effect can always be calculated from the previous ones. Hence we arrange the correlations  $p(a|x, y) = \text{tr}(\rho_x M_{a|y})$  in a  $3 \times 3$  matrix  $\mathcal{P}$ , where the rows correspond to the states and the columns to the effects  $M_{1|1}, M_{2|1}, M_{1|2}$ , that is,

$$\mathcal{P} = \begin{pmatrix} p(1|1, 1) & p(2|1, 1) & 0 \\ p(1|2, 1) & 0 & 1 \\ 0 & p(2|3, 1) & p(1|3, 2) \end{pmatrix}. \quad (3)$$

Since correlations of this form are motivated by USD, we refer to them as USD correlations. In particular, if the dimension of the system is known to be  $d$ , we write  $\text{USD}_d$  for the set of all USD correlation achievable under this constraint. Obviously the sets  $\text{USD}_d$  obey the inclusion  $\text{USD}_d \subset \text{USD}_{d+1}$  and furthermore, as we point out in Sec. III A, dimension 3 is already sufficient to achieve all USD correlations,  $\text{USD}_d = \text{USD}_3$  for all  $d > 3$ . We are therefore particularly interested in the qubit case, for which we have the following characterization.

*Lemma 1.* The set  $\text{USD}_2$  consists exactly of all correlations of the form

$$\mathcal{P}(p, q, \xi) = \begin{pmatrix} p\xi & q & 0 \\ p(1 - \xi) & 0 & 1 \\ 0 & q(1 - \xi) & \xi \end{pmatrix}, \quad (4)$$

with  $p, q, \xi \in [0, 1]$  such that  $(1 - p)(1 - q) \geq p q \xi$ . In addition, for a given correlation matrix  $\mathcal{P}$ , the states and measurements realizing  $\mathcal{P}$  are unique, up to a global unitary transformation.

The proof is given in Appendix A.

### III. CERTIFICATION OF TRICHOTOMIC MEASUREMENTS

We approach now the question whether the difference between reducible and irreducible  $n$ -outcome measurements can be observed using only the correlations, that is, whether there exists an irreducible  $n$ -outcome POVM  $\mathbf{M} = (M_1, \dots, M_n)$  together with auxiliary measurements  $\mathbf{M}_2, \dots, \mathbf{M}_m$  and states

$\rho_1, \dots, \rho_s$  such the correlations  $p(a|x, y) = \text{tr}(\rho_x M_{a|y})$  cannot stem from a reducible measurement. Correlations of this type are genuinely  $n$ -chotomic; otherwise they are simulable  $n$ -chotomic. Clearly, if such correlations exist, they enable us to certify that the measurement  $\mathbf{M}_1$  is indeed irreducible. To point out the difference between both questions consider the following example. Suppose that one implements the trine POVM  $\mathbf{S}$  on a qubit system, which is an irreducible three-outcome measurement. We show in Sec. III A that this measurement alone can never yield correlations which cannot be explained by a reducible measurement. Therefore, an irreducible three-outcome measurement does not necessarily define genuine trichotomic correlations, while genuine trichotomic correlations always involve an irreducible measurement.

#### A. Minimal scenario

Suppose we want to certify that a given measurement apparatus implements an irreducible  $n$ -outcome POVM  $\mathbf{M}_1$ . What is the minimal scenario in which one can conclude from the output statistics alone that the POVM is irreducible? In other words, what is the minimal number of state preparations  $s$  and auxiliary measurements  $m - 1$ ?

Clearly, if one has access to  $s$  preparations and  $m - 1$  auxiliary measurements, the set of all possible correlations  $p(a|x, y)$  that can be obtained in a scenario without any constraint on the dimension of the system yields a convex set. As it turns out, this convex set is a polytope, whose extremal points correspond to deterministic correlations [18], that is, where all  $p(a|x, y)$  are either 0 or 1. If the dimension  $d$  of the system is at least  $s$  then these extremal points can be obtained from a fixed choice of  $s$  orthogonal states  $\rho_x = |\psi_x\rangle\langle\psi_x|$  and at most  $s$ -chotomic measurements with effects  $M_{a|y} = \sum_x p(a|x, y)\rho_x$ . Hence, all correlations can be written as convex combination of deterministic strategies. Since all extreme points use the same states, the convex coefficients can be absorbed into the effects yielding at most  $s$ -chotomic POVMs. For our case of an irreducible three-outcome measurement on a qubit,  $n = 3$  and  $d = 2$ , this implies that at least  $s = 3$  different states are required. It also follows that  $\text{USD}_d \subset \text{USD}_3$  since only three different states are used in the USD correlations.

Regarding the number of auxiliary measurements  $m - 1$ , we consider the case where no auxiliary measurements are used. In this scenario, we write  $\mathcal{C}_d(s, n)$  for the convex hull of all correlations on a  $d$ -dimensional classical system and  $\mathcal{Q}_d(s, n)$  for the convex hull of all correlations on a  $d$ -dimensional quantum system. Note that these sets are only of importance for determining the minimal scenario and are not used subsequently. In Theorem 3 in Ref. [19] it was established that both sets are equal,  $\mathcal{Q}_d(s, n) = \mathcal{C}_d(s, n)$ . Since all correlations in  $\mathcal{C}_d(s, n)$  can be obtained from  $d$ -chotomic measurements, this property also follows for all correlations in  $\mathcal{Q}_d(s, n)$ . Therefore, the smallest scenario which allows us to certify an irreducible three-outcome measurement on a qubit includes at least two measurements  $\mathbf{M}_1 = (M_{1|1}, M_{2|1}, M_{3|1})$  and the auxiliary measurement  $\mathbf{M}_2 = (M_{1|2}, M_{2|2})$  (see also Fig. 2). The set of all correlations achievable in the minimal

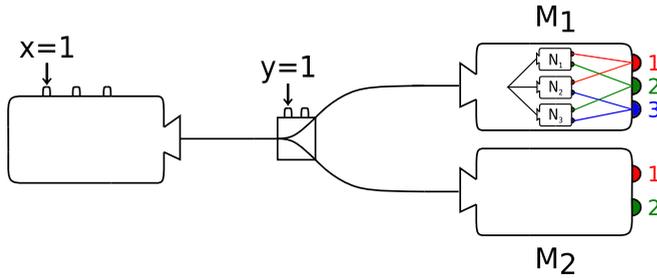


FIG. 2. Prepare-and-measure setup of the minimal scenario producing simulable trichotomic correlations. On the left-hand side one can choose between three different preparation procedures  $x \in \{1, 2, 3\}$  and the corresponding qubit state is sent to one of two measurement devices (right-hand side). The measurement device is chosen by the experimenter and this choice is denoted by  $y \in \{1, 2\}$ . For  $y = 1$  the measurement can yield three different outcomes, but governed by a specific inner mechanism. It consists of three two-outcome measurements, one of which is chosen at random. Depending on the outcome of this measurement, an outcome is assigned to the overall three-outcome measurement. If  $y = 2$  the measurement device is a simple two-outcome measurement.

scenario with fixed dimension  $d$  is subsequently denoted by  $COR_d$ . Furthermore, we write  $SIM_d$  for the set of simulable trichotomic correlations within  $COR_d$ . All sets that are relevant for our discussion can be concisely described as follows:  $USD_d$  is the set of correlations of the form (3) that can be obtained from a  $d$ -level quantum system,  $SIM_d$  is the set of simulable trichotomic correlations that can be obtained from a  $d$ -level quantum system in the minimal scenario, and  $COR_d$  is the set of all correlations achievable in the minimal scenario with fixed dimension  $d$ . Clearly, the correlations  $USD_d$  are a subset of  $COR_d$  and  $SIM_d \cap USD_d$  are the simulable trichotomic correlations within  $USD_d$ .

### B. Geometry of the trichotomic correlations

In this section we investigate how the sets  $SIM_2$  and  $COR_2$  are related. In contrast to Bell-like scenarios, where the convex hull of the correlations is taken [11,20], here we restrict the study to the bare sets  $COR_2$  and  $SIM_2$ . These sets are not convex, as can be seen by considering the subset  $USD_2$ , characterized by Lemma 1. In fact, it is evident that the correlation matrix of  $D_1 = \mathcal{P}(1, 0, 1)$  and  $D_2 = \mathcal{P}(1, 1, 0)$  can be realized with dichotomic measurements, that is,  $D_1, D_2 \in USD_2 \cap SIM_2$ . However, no convex combination  $D_\lambda = \lambda D_1 + (1 - \lambda)D_2$  with  $0 < \lambda < 1$  can be written in the form  $\mathcal{P}(p, q, \xi)$  as given by Eq. (4). Here it is important to note that the correlations  $D_\lambda$  are still USD correlations. Hence  $D_\lambda \in COR_2$  already implies  $D_\lambda \in USD_2$ . The relation of the points  $D_1$  and  $D_2$  and their convex mixture  $D_\lambda$  to the sets  $USD_2$  and  $SIM_2$  is illustrated in Fig. 3. Since  $USD_2$  is an affine section of  $COR_2$ , it follows that neither  $COR_2$  nor  $SIM_2$  is convex.

Our next step is to establish that not all qubit USD correlations are simulable trichotomic. We have the following theorem, which we prove in Appendix B.

*Theorem 1.* There exist correlations in  $USD_2$  that are not contained in the convex hull of  $USD_2 \cap SIM_2$ .

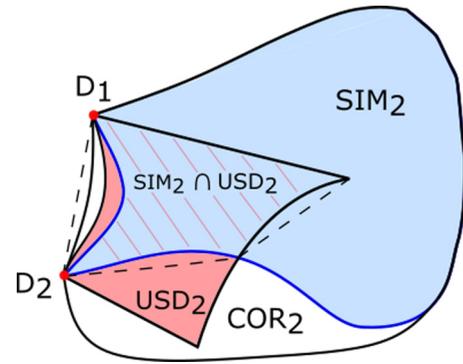


FIG. 3. Schematic illustration of the relations among the sets  $USD_2$ ,  $SIM_2$ , and  $COR_2$ . The set  $COR_2$  of all correlations that can be obtained with a qubit in the minimal scenario is obviously a superset of  $SIM_2$ , the set of all simulable correlations in the minimal scenario as well as  $USD_2$ , the set of all correlations that can be obtained with a qubit and are of the form (3). The points  $D_1$  and  $D_2$  represent deterministic correlations contained in  $USD_2 \cap SIM_2$ , hence located at the boundary of any of the three sets. The dashed line represents the convex hull of the set  $USD_2 \cap SIM_2$ .

Hence, even the convex hull of the simulable trichotomic qubit USD correlations does not cover all qubit USD correlations. This might raise the expectation that there exists a linear inequality separating  $COR_2$  and  $SIM_2$ , despite the nonconvexity properties discussed above. However, note that the statement of Theorem 1 only concerns the subset of USD correlations and one cannot conclude that the convex hull of  $SIM_2$  is a proper subset of  $COR_2$ .

### C. Genuine trichotomic correlations

For an experimental certification of an irreducible three-outcome measurement it is essential to find correlations  $\mathcal{P} \in COR_2$  such that closest simulable correlations  $\mathcal{P}' \in SIM_2$  have a distance  $r$  of reasonable size. We measure this distance either in terms of the supremum norm, yielding  $r_\infty$ , or in terms of the Euclidean norm, yielding  $r_2$ , that is,

$$r_\infty = \max_{i,j} |\mathcal{P}_{i,j} - \mathcal{P}'_{i,j}|, \quad (5)$$

$$r_2 = \left[ \sum_{i,j} (\mathcal{P}_{i,j} - \mathcal{P}'_{i,j})^2 \right]^{1/2}. \quad (6)$$

According to Theorem 1, we can preliminarily focus on the family of USD correlations, since Theorem 1 guarantees that there exist USD correlations  $\mathcal{P} \in USD_2$  such that  $r > 0$ .

In order to compute  $r_2$  and  $r_\infty$  we rely on numerical optimization over the set  $SIM_2$ . The optimization is non-linear and involves three (possibly mixed) states as well as four dichotomic POVMs. It is important to note that if the states or the effects are fixed, the problem can be rephrased as a semidefinite program (SDP) and becomes thereby easy to solve numerically. We use this fact to implement an alternating optimization (seesaw algorithm [21]); a detailed description is provided in Appendix C. We mention that, technically, this optimization algorithm only yields guaranteed upper bounds on the distances, because it is based on

finding the correlations in  $\text{SIM}_2$  closest to  $\mathcal{P}$ . While this fact is not of practical concern, we discuss in Appendix D also how strict, but rather rough, lower bounds can be obtained. The largest distance we found is realized by the choice of  $\mathcal{P} = \mathcal{P}(0.577, 0.726, 0.276)$ , where the seesaw algorithm yields  $r_2 \approx 0.0391$  and  $r_\infty \approx 0.0177$ . Using the algorithm described in Appendix D, the computation of the lower bound yields  $r_\infty \approx 0.0022$ , hence one order of magnitude smaller. However, the lower bound should only emphasize that the distance is indeed positive, hence consistent with Theorem 1. Larger distances can be achieved using correlations which are not confined to  $\text{USD}_2$ . In particular, choosing an arrangement involving the trine POVM (see Fig. 1 and Appendix E), we find  $r_2 \approx 0.0686$  and  $r_\infty \approx 0.0342$ . Hence we obtain only moderate experimental requirements for a certification of an irreducible three-outcome measurement.

#### IV. ESTIMATING THE STATE SPACE DIMENSION

As discussed in the Introduction, in order to certify an irreducible  $n$ -outcome measurement, knowledge of the dimension of the prepared system is necessary. While it is possible that the dimension can be convincingly deduced from the experimental setup, for a fully device-independent procedure, the dimension of the system has to be determined from correlation data alone. However, the dimension of a system is a physically ill-defined object, in the sense that any description of a system can always be embedded into a higher-dimensional system. For example, we can treat a qubit as a restricted theory of a qutrit. However, from an operational point of view, one can still assess the dimension by determining the effective dimension, that is, the minimal dimension which explains the experimental data. A dimension witness [22] might seem to be the appropriate tool for this purposes, since it gives a procedure to determine a lower bound on the dimension. However, this is not sufficient for our purposes, because it does not exclude that the effective dimension can be higher than the dimension witnessed.

In general, we assume that the procedure to determine the effective dimension of a system consists of  $s$  different state preparations and  $m$  measurements. The correlations  $p(a|x, y)$  form a matrix  $(\mathcal{A}_{x,\ell})_{x,\ell}$ , where  $x$  labels the states,  $a$  labels the outcome of measurement  $y$ , and  $\ell$  enumerates all outcomes of all measurements. For a  $d$ -level system, the rank of this matrix can be at most the affine dimension of the state space, that is,  $d^2 - 1$ . Hence determining the rank of the matrix  $\mathcal{A}$  can give an estimate of the effective dimension of the system. In practice, one would choose a large number of preparation procedures and a large number of measurement procedures with the expectation that an estimate of the rank of  $\mathcal{A}$  produces a reliable estimate of the affine dimension of the state space. We mention that for consistency reasons, the preparation-and-measurement procedures should include those required to certify the irreducibility of the  $n$ -outcome measurement.

While this approach can work in principle, it has to be considered with care. For an implementation of an irreducible three-outcome measurement on a qubit, it is typically necessary to dilate the three-outcome measurement to a projective measurement on a higher-dimensional system [7]. Despite this, it still makes sense to speak about an irreducible three-

outcome measurement if the additional dimensions used for the dilated measurement are not accessible due to a physical mechanism that reduces the dimension before entering the measurement station. In a setup using the polarization degree of freedom of a photon, this may be achieved, for example, by means of a single mode fiber. Mathematically, such a mechanism corresponds to a completely positive map  $\Phi$  so that the correlations are obtained through  $p(a|x, y) = \text{tr}[\Phi(\rho_x)M_{a|y}]$ . However, then the rank of the matrix  $\mathcal{A}$  alone is insufficient to establish the effective dimension of the system, as can be seen by considering the dephasing qutrit-qutrit channel  $\Phi : \rho \mapsto \sum_k |k\rangle\langle k| \rho |k\rangle\langle k|$ . Using this channel, the matrix  $\mathcal{A}$  will have rank 3, which would suggest an effective dimension of  $d = 2$ , while the actual effective dimension is  $d = 3$ . This can be overcome by certifying that the shape of the state and effect space correspond to a qubit. For methods to implement such a certification refer to Ref. [23].

#### V. DISCUSSION

We studied the structure of the correlations produced by irreducible three-outcome qubit measurements in the prepare-and-measure scenario. Our goal was a minimal scenario in terms of the number of experimental devices required. Using only one auxiliary measurement, we found that the genuine trichotomic correlations and the simulable trichotomic correlations can be separated in the Euclidean norm by  $r_2 = 0.0686$ . Consequently, all nine entries in the correlation matrix must be determined with an absolute error of roughly  $r_2/\sqrt{9} \approx 0.0229$ . We mention the great similarity of this setup to the one used in Ref. [12], namely, the states and the measurements are the same, but in our setup we omit one of the auxiliary measurements. While this scheme is motivated by symmetry considerations, we also used the USD correlations to systematically describe a subset of the trichotomic correlations. However, within these correlations the largest distance we obtained is smaller,  $r_2 \approx 0.0347$ . We mention that the results in Ref. [9] are also based on USD, but with the different goal to certify any of several trichotomic measurement.

Our results are not based on a linear inequality separating genuine and simulable trichotomic correlations and such an inequality is not necessary for the purpose of an experimental certification. However, we also established in Theorem 1 that such an inequality exists when the analysis is constrained to the USD correlations. It is now an interesting open question whether in our minimal scenario this also holds when considering all correlations, since this would imply that also in the minimal scenario the convex hull of the simulable correlations does not include all trichotomic correlations. A further interesting question is if it is possible to find a systemic approach to construct families of distributions in a minimal scenario that are able to determine whether an implemented  $n$ -chotomic measurement is irreducible. Such an approach would unify recent results that actually appear in separate contexts.

#### ACKNOWLEDGMENTS

We thank Rene Schwonnek, Xiao-Dong Yu, and Eric Chitambar for discussions and the University of Siegen for enabling our computations through the HoRUS and the OMNI

cluster. This work was supported by the Deutsche Forschungsgemeinschaft (German Research Foundation, Projects No. 447948357 and No. 440958198), the Sino-German Center for Research Promotion (Project No. M-0294), and the ERC (Consolidator Grant No. 683107/TempoQ). J.S. acknowledges support from the House of Young Talents of the University of Siegen.

### APPENDIX A: PROOF OF LEMMA 1

Suppose that  $\mathcal{P} \in \text{USD}_2$ . This has consequences for the measurements  $\mathbf{M}_1 = (M_{1|1}, M_{2|1}, M_{3|1})$  and  $\mathbf{M}_2 = (M_{1|2}, M_{2|2})$  and the states  $\rho_1$ ,  $\rho_2$ , and  $\rho_3$  that can realize the correlations. In particular,  $\mathcal{P}_{2,3} = \text{tr}[\rho_2 M_{1|2}] = 1$  and  $\mathcal{P}_{1,3} = 0$  imply  $\rho_2 = M_{1|2} = |\varphi\rangle\langle\varphi|$  and  $\rho_1 = |\varphi^\perp\rangle\langle\varphi^\perp|$ , where  $|\varphi\rangle$  and  $|\varphi^\perp\rangle$  are two orthonormal vectors. With a similar argument we obtain  $M_{2|1} = q|\varphi^\perp\rangle\langle\varphi^\perp|$  with  $0 \leq q \leq 1$ . It remains to consider the consequences of  $\mathcal{P}_{3,1} = 0$  for  $\rho_3$  and  $M_{1|1}$ . This requires  $M_{1|1} = p|\eta^\perp\rangle\langle\eta^\perp|$  and  $\rho_3 = |\eta\rangle\langle\eta|$  for some orthonormal vectors  $|\eta\rangle$  and  $|\eta^\perp\rangle$  and  $0 \leq p \leq 1$ .

Without loss of generality, we can assume

$$|\eta\rangle = \sqrt{\xi}|\varphi\rangle + \sqrt{1-\xi}e^{i\phi}|\varphi^\perp\rangle, \quad (\text{A1})$$

$$|\eta^\perp\rangle = \sqrt{1-\xi}|\varphi\rangle - \xi e^{i\phi}|\varphi^\perp\rangle, \quad (\text{A2})$$

where  $0 \leq \xi \leq 1$  and  $\phi \in \mathbb{R}$ . This yields immediately Eq. (4) together with the conditions  $p, q, \xi \in [0, 1]$ . For  $\mathbf{M}_1$  to form a POVM it remains to verify that  $M_{3|1} = \mathbb{1} - M_{1|1} - M_{2|1}$  is positive semidefinite. This reduces here to  $\text{tr}(M_{3|1}) \geq 0$  and  $\det(M_{3|1}) \geq 0$  and can be equivalently expressed as the single condition  $(1-p)(1-q) \geq pq\xi$ .

From the above construction it is also immediately clear that, conversely, any choice of  $p, q, \xi$  satisfying the constraints in Lemma 1 is in  $\text{USD}_2$ . Finally, given  $\mathcal{P}$ , all effects and states are fixed by the above considerations, except for the choice of the orthonormal basis  $\{|\varphi\rangle, e^{i\phi}|\varphi^\perp\rangle\}$ , proving the claim of a unique representation up to a unitary transformation.

### APPENDIX B: PROOF OF THEOREM 1

We first parametrize the correlations  $\mathcal{D} \in \text{SIM}_2 \cap \text{USD}_2$  with  $\mathcal{D}_{3,3} \neq 0$ . By virtue of Eq. (1) and using the proof of Lemma 1 above, the effects of the simulated trichotomic POVM  $\mathbf{M}_1$  can be written as

$$M_{1|1} = p|\eta^\perp\rangle\langle\eta^\perp| = \kappa_1 F_1 + \kappa_3(\mathbb{1} - F_3), \quad (\text{B1})$$

$$M_{2|1} = q|\varphi^\perp\rangle\langle\varphi^\perp| = \kappa_1(\mathbb{1} - F_1) + \kappa_2 F_2, \quad (\text{B2})$$

where, compared to Eq. (1), we write  $\kappa_j$  in place of  $p_j$  and  $F_j$  in place of  $N_{j|j}$ . For  $\kappa_1 \neq 0$  it follows that  $F_1 \propto |\eta^\perp\rangle\langle\eta^\perp|$  and  $\mathbb{1} - F_1 \propto |\varphi^\perp\rangle\langle\varphi^\perp|$ , yielding  $F_1 = |\eta^\perp\rangle\langle\eta^\perp| = |\varphi\rangle\langle\varphi|$ . This is in contradiction to the assumption  $\mathcal{D}_{3,3} \neq 0$  by virtue of  $\mathcal{D}_{3,3} = \text{tr}(M_{1|2}\rho_3)$  with  $\rho_3 = |\eta\rangle\langle\eta|$  and  $M_{1|2} = |\varphi\rangle\langle\varphi|$ . Hence  $\kappa_1 = 0$ . Writing  $\kappa \equiv \kappa_2 = 1 - \kappa_3$ , Eq. (B2) implies  $F_2 = f_2|\varphi^\perp\rangle\langle\varphi^\perp|$  with  $0 \leq f_2 \leq 1$  and  $\kappa f_2 = q$ . Similarly, from Eq. (B1) one obtains  $\mathbb{1} - F_3 = f_3|\eta^\perp\rangle\langle\eta^\perp|$  with  $(1-\kappa)f_3 = p$ . Therefore,  $\mathcal{D} \in \text{USD}_2 \cap \text{SIM}_2$  with  $\mathcal{D}_{3,3} \neq 0$

if and only if

$$\mathcal{D} = \begin{pmatrix} f_3(1-\kappa)\xi & f_2\kappa & 0 \\ f_3(1-\kappa)(1-\xi) & 0 & 1 \\ 0 & f_2\kappa(1-\xi) & \xi \end{pmatrix}, \quad (\text{B3})$$

with  $f_2, f_3, \kappa, \xi \in [0, 1]$  and  $\xi \neq 0$ .

Next we show that for any  $0 < \xi < 1$  there exist correlations  $\mathcal{P} \in \text{USD}_2 \setminus \text{SIM}_2$ . For this we consider the linear map

$$W : \mathcal{P} \mapsto -\mathcal{P}_{1,1} - \mathcal{P}_{1,2} + \mathcal{P}_{3,2} + \mathcal{P}_{3,3}. \quad (\text{B4})$$

For  $\mathcal{P}(p, q, \xi)$  as in Eq. (4) and the choice  $p = \frac{1}{2}$  and  $q = 1/(\xi + 1)$ , we verify that for any  $0 < \xi < 1$  the constraint  $(1-p)(1-q) \geq pq\xi$  is satisfied and in addition  $W(\mathcal{P}) < 0$  holds. However, for any  $\mathcal{D}$  as in Eq. (B3), we have  $W(\mathcal{D}) = \xi[1 - f_2\kappa - f_3(1-\kappa)] \geq 0$  and in addition for any  $\mathcal{T} \in \text{USD}_2$  with  $\mathcal{T}_{3,3} = 0$  we have immediately  $W(\mathcal{T}) = 0$ . Consequently,  $W(\mathcal{S} - \mathcal{P}) > 0$  for our choice of  $\mathcal{P}$  and any  $\mathcal{S} \in \text{SIM}_2 \cap \text{USD}_2$ , that is,  $\mathcal{S} \neq \mathcal{P}$ .

From this observation, it readily follows that the convex hull of  $\text{USD}_2 \cap \text{SIM}_2$  does not contain all of  $\text{USD}_2$ . This is the case because  $W$  is a linear map and thus its minimum over the convex hull of  $\text{USD}_2 \cap \text{SIM}_2$  is attained already for some  $\mathcal{S} \in \text{USD}_2 \cap \text{SIM}_2$ . However, for this set we just proved that  $W(\mathcal{S}) > W(\mathcal{P})$  for certain  $\mathcal{P} \in \text{USD}_2$ .

### APPENDIX C: UPPER BOUND ON THE MINIMAL DISTANCE

We compute the maximal radius  $r$  of a ball  $B_r(\mathcal{P})$  around correlations  $\mathcal{P} \in \text{COR}_2 \setminus \text{SIM}_2$  such that  $B_r(\mathcal{P}) \cap \text{SIM}_2$  is empty. As we mention in the main text, we consider this maximal ball with respect to the supremum norm  $r = r_\infty$  and the Euclidean norm  $r = r_2$  [see Eqs. (5) and (6), respectively]. Correspondingly, computing  $r$  can be formulated as the optimization problem to minimize a real parameter  $t$  over all  $\mathcal{Q} \in \text{SIM}_2$  such that

$$-t \leq \mathcal{P}_{i,j} - \mathcal{Q}_{i,j} \leq t \quad \forall i, j \quad (\text{C1})$$

in the case of  $r_\infty$  and

$$\sum_{i,j} (\mathcal{P}_{i,j} - \mathcal{Q}_{i,j})^2 \leq t^2 \quad (\text{C2})$$

in the case of  $r_2$ .

We write  $F_1 = M_{1|1}$ ,  $F_2 = M_{2|1}$ , and  $F_3 = M_{1|2}$  so that  $\mathcal{Q}_{i,j} = \text{tr}(\rho_i F_j)$ . If we keep the effects fixed, then the optimization is a semidefinite program of the following type: Minimize  $t$  under the constraints  $\rho_i \geq 0$  and  $\text{tr}(\rho_i) = 1$  for  $i = 1, 2, 3$  and either the linear constraint (C1) or the quadratic-convex constraint (C2). Similarly, if we keep the states fixed, then the optimization is again a semidefinite program, however now with the constraints on the states replaced by constraints on the effects, namely,

$$F_1 = F'_1 + F'_0, \quad F_2 = F'_2 + f_0\mathbb{1} - F'_0, \quad (\text{C3})$$

$$0 \leq F'_0 \leq f_0\mathbb{1}, \quad 0 \leq F'_1 \leq f_1\mathbb{1}, \quad 0 \leq F'_2 \leq f_2\mathbb{1}, \quad (\text{C4})$$

$$0 \leq F_3 \leq \mathbb{1}, \quad f_0 + f_1 + f_2 = 1. \quad (\text{C5})$$

Since these small semidefinite programs can be solved very fast numerically, this invites a seesaw optimization [21] where one alternates between the two optimizations until  $t$  converges.

We implement this seesaw algorithm using the PYTHON library PICOS with the CVXOPT back end. As the criterion for convergence we take  $t_{n-1} - t_n < 10^{-6}$ , where  $t_n$  is the value after  $n$  seesaw iterations. This convergence happens after at most 300 iterations. We repeat the optimization 4500 times, each time with different starting values for  $\rho_1$ ,  $\rho_2$ , and  $\rho_3$ , where we take pure states chosen randomly according to the Haar measure and then decrease the purity  $\text{tr}(\rho^2)$  to be uniformly in the interval  $[\frac{1}{2}, 1]$ . The same optimal value is always reached independently of the start values for the Euclidean norm, while it occurs only for about 1% of the start values in the case of the supremum norm.

#### APPENDIX D: LOWER BOUND ON THE MINIMAL DISTANCE

In this Appendix we describe an algorithm that yields a lower bound on the distance between a given correlation and the set  $\text{SIM}_2$ . In the following we will write  $B(P, \epsilon)$  for the ball at  $P \in \text{COR}_2$  of radius  $\epsilon$  in the metric induced by the  $\infty$ -norm. Considering a given correlation  $P \in \text{USD}_2$  of the form in Eq. (3), we are interested in the condition on the states and the measurements under which we have  $Q \in B(P, \epsilon)$  for  $\epsilon > 0$ . Clearly, any  $Q \in \text{COR}_2$  is of the form

$$Q = \begin{pmatrix} \text{tr}(\rho_1 M_{111}) & \text{tr}(\rho_1 M_{211}) & \text{tr}(\rho_1 M_{112}) \\ \text{tr}(\rho_2 M_{111}) & \text{tr}(\rho_2 M_{211}) & \text{tr}(\rho_2 M_{112}) \\ \text{tr}(\rho_3 M_{111}) & \text{tr}(\rho_3 M_{211}) & \text{tr}(\rho_3 M_{112}) \end{pmatrix}. \quad (\text{D1})$$

However, it turns out to be useful to group the parameters in (D1) into two different classes. Define  $a = (\rho_1, \rho_2, M_{112})$  and  $b = (\rho_3, M_{111}, M_{211})$ , which together form the full set of parameters of the distribution  $Q = Q(a, b)$ . The question of how large can  $\epsilon$  be chosen such that there is no  $Q(a, b) \in B(P, \epsilon)$  if the trichotomic measurement  $M_1 = (M_{111}, M_{211}, M_{311})$  is limited to be simulable by dichotomic measurements. Observe that if  $Q(a, b)$  are USD correlations, then the elements  $Q_{13}$ ,  $Q_{22}$ ,  $Q_{31}$ , and  $Q_{23}$  all are 0 or 1 and this imposes strong constraints on the parameters  $a$ . In fact, if  $Q(a, b) = P$ , then the value of  $a$  is fixed up to a unitary transformation, as we have seen in Appendix B. We call this fixed value  $a^*$ . In the following, we will extend the argument in Appendix B to show that  $Q(a, b) \in B(P, \epsilon)$  implies that  $a \in B(a^*, \mathcal{O}(\sqrt{\epsilon}))$ . Then, using the continuity of  $Q$ , we can show that  $a \in B(a^*, \mathcal{O}(\sqrt{\epsilon}))$  implies  $Q(a, b) \in B(Q(a^*, b), \mathcal{O}(\sqrt{\epsilon}))$ . Since  $Q(a, b) \in B(P, \epsilon)$  and  $Q(a, b) \in B(Q(a^*, b), \mathcal{O}(\sqrt{\epsilon}))$ , it follows that  $Q(a^*, b) \in B(P, \epsilon + \mathcal{O}(\sqrt{\epsilon}))$  by the triangular inequality. We have thus reduced the problem of asking for the existence of  $a$  and  $b$  such that  $Q(a, b) \in B(P, \epsilon)$  to simply asking for the existence of  $b$  such that  $Q(a^*, b) \in B(P, \epsilon + \mathcal{O}(\sqrt{\epsilon}))$ . The latter means that, for a fixed value of  $\epsilon$  and fixed values of  $\rho_1$ ,  $\rho_2$ , and  $M_{112}$ , we ask for the feasibility of  $\rho_3$ ,  $M_{111}$ , and  $M_{211}$  such that  $Q \in B(P, \epsilon + \mathcal{O}(\sqrt{\epsilon}))$ . This is not yet an SDP; however, it can be decomposed into a finite number of SDPs by scan-

ning the values of  $\rho_3$  and bounding the error in the finite scanning.

The detailed calculation of the above program is straightforward yet cumbersome. We will first introduce some notation. For convenience we parametrize the effects and the states by the Bloch coordinates. More precisely, for an effect  $E$ , we write  $(x_0, \vec{x})$ , which means  $E = \frac{1}{2} \sum_{i=0}^3 x_i \sigma_i$ , where  $\sigma_0 := \mathbb{1}$  and  $\sigma_i$  are the Pauli matrices. Similarly, for a state  $\rho$  we write  $\vec{y}$  with  $\rho = \frac{1}{2} (\mathbb{1} + \sum_{i=1}^3 y_i \sigma_i)$ . In particular, we write  $\rho_i = (1, \vec{r}_i)$ ,  $M_{111} = (x_{01}, \vec{x}_1)$ ,  $M_{211} = (x_{02}, \vec{x}_2)$ , and  $M_{112} = (y_0, \vec{y})$ . Because of the unitary freedom, we can always assume  $\vec{x}_1$  and  $\vec{x}_2$  to have no  $z$  component and at the same time  $\vec{r}_2$  to have no  $y$  component. We also write  $Q \in P \pm \epsilon$  as a synonym of  $Q \in B(P, \epsilon)$ .

First, we show that one can bound the trace of the effect from below, given a lower bound on the probability for that effect.

*Lemma 2.* Let  $\rho = (1, \vec{r})$  be a state and  $F = (x_0, \vec{x})$  be an effect. We have (i) if  $\text{tr}[F\rho] \geq a$  then  $x_0 \geq a$  and (ii) if  $\text{tr}[F\rho] \leq b$  then  $x_0 \leq 1 + b$ .

*Proof.* We first show (i). We have  $\text{tr}[F\rho] = \frac{1}{2}(x_0 + \vec{x}\vec{r}) \geq a$ . Because  $\vec{x}\vec{r} \leq |\vec{x}||\vec{r}| \leq x_0|\vec{r}| \leq x_0$ , we have  $x_0 + x_0 \geq 2a$  or  $x_0 \geq a$ . For (ii) we have similarly  $\text{tr}[F\rho] = \frac{1}{2}(x_0 + \vec{x}\vec{r}) \leq b$ ; then  $\frac{1}{2}(2 - x_0 - \vec{x}\vec{r}) \geq 1 - b$ . Then applying (i), we find  $2 - x_0 \geq 1 - b$  or  $x_0 \leq 1 + b$ .

The following corollary is an application of the lemma to the condition  $Q \in P \pm \epsilon$ .

*Corollary 1 (estimation of the traces of effects).* For  $Q \in P \pm \epsilon$  we need (i)  $1 - \epsilon \leq y_0 \leq 1 + \epsilon$ , (ii)  $c_1 - \epsilon \leq x_{01} \leq 1 + \epsilon$  with  $c_1 = \max\{P_{12}, P_{22}\}$ , and (iii)  $c_2 - \epsilon \leq x_{02} \leq 1 + \epsilon$  with  $c_2 = \max\{P_{13}, P_{33}\}$ .

*Proof.* (i) is the direct consequence of  $0 \leq Q_{13} \leq \epsilon$  and  $1 \geq Q_{23} \geq 1 - \epsilon$ . (ii) is the direct consequence of  $Q_{11} \in P_{11} \pm \epsilon$  and  $Q_{21} \in P_{21} \pm \epsilon$ . Likewise, (iii) is the direct consequence of  $Q_{12} \in P_{12} \pm \epsilon$  and  $Q_{32} \in P_{32} \pm \epsilon$ .

*Corollary 2 (estimation of  $\vec{r}_1$  and  $\vec{r}_2$  by  $\vec{y}$ ).* (i)  $Q_{23} \geq 1 - \epsilon$  implies  $|\vec{y} - \vec{r}_2| \leq \sqrt{4\epsilon + \epsilon^2}$  and (ii)  $Q_{13} \leq \epsilon$  implies  $|\vec{y} + \vec{r}_1| \leq \sqrt{4\epsilon + \epsilon^2}$ .

*Proof.* (i) We have  $Q_{23} = \frac{1}{2}(y_0 + \vec{y}\vec{r}_2) \geq 1 - \epsilon$ ; then  $\vec{y}\vec{r}_2 \geq 2(1 - \epsilon) - y_0$ . Further,  $\vec{y}\vec{r}_2 \geq 2(1 - \epsilon) - y_0$  leads to  $\vec{y}^2 + \vec{r}_2^2 - (\vec{y} - \vec{r}_2)^2 \geq 4(1 - \epsilon) - 2y_0$ . Because  $|\vec{y}| \leq \min\{y_0, 2 - y_0\}$  and  $\vec{r}_2^2 \leq 1$ , we obtain  $(\min\{y_0, 2 - y_0\})^2 + 2y_0 - 3 + 4\epsilon \geq (\vec{y} - \vec{r}_2)^2$ . The left-hand side is maximized at  $y_0 = 1 + \epsilon$ , which leads to  $4\epsilon + \epsilon^2 \geq (\vec{y} - \vec{r}_2)^2$ . The proof of (ii) is analogous to (i).

Now we consider the constraint of the form  $\text{tr}[\rho F] \leq \epsilon$ . We show that if the trace of  $F$  is bounded from below, we can bound the purity of  $\rho$ . This also extends to the consideration of each Bloch components of the Bloch vectors.

*Lemma 3.* Let  $\rho = (1, \vec{r})$  be a state and  $F = (x_0, \vec{x})$  be an effect with  $x_0 \geq c$  and  $\vec{x}$  having no  $z$  component. Then  $\text{tr}[\rho F] \leq \epsilon$  implies  $|\vec{r}| \geq |\vec{r}_{xy}| \geq 1 - \frac{2\epsilon}{c}$  and  $|r_z| \leq \sqrt{1 - (1 - 2\epsilon/c)^2}$ . As a consequence, if we let  $\vec{n}$  be the unit vector of the direction of  $\vec{r}_{xy}$ , then  $|\vec{r}_{xy} - \vec{n}| \leq 2\epsilon/c$  and  $|\vec{r} - \vec{n}| \leq 2\epsilon/c + \sqrt{1 - (1 - 2\epsilon/c)^2}$ . Here  $\vec{r}_{xy}$  denotes the projection of  $\vec{r}$  onto the  $xy$  plane and  $\vec{r}_z$  onto the  $z$  axis.

*Proof.* Observe that  $\text{tr}[\rho F] = \frac{1}{2}(x_0 + \vec{r}\vec{x}) = \frac{1}{2}(x_0 + \vec{r}_{xy}\vec{x}) \leq \epsilon$ . This leads to  $-\vec{r}_{xy}\vec{x} \geq x_0 - 2\epsilon$ . Then we have  $|\vec{r}_{xy}|x_0 \geq x_0 - 2\epsilon$  and so  $|\vec{r}| \geq |\vec{r}_{xy}| \geq 1 - 2\epsilon/x_0 \geq$

$1 - 2\epsilon/c$ . Also  $1 \geq \bar{r}_z^2 + \bar{r}_{xy}^2 \geq \bar{r}_z^2 + (1 - 2\epsilon/c)^2$ , so  $|\bar{r}_z| \leq \sqrt{1 - (1 - 2\epsilon/c)^2}$ . The latter part of the statement is obvious.

*Corollary 3 (estimation of  $\bar{r}_2$  and  $\bar{r}_3$  by the unit vectors of their projection onto the  $xy$  plane).* Let  $\bar{r}_{2xy}$  denote the  $xy$  component and  $\bar{r}_{2z}$  denote the  $z$  component of  $\bar{r}_2$ . Further, let  $\bar{n}$  be the unit vector in the direction of  $\bar{r}_{2xy}$  and  $\bar{t}$  be the unit vector in the direction of  $\bar{r}_{3xy}$ . Then we obtain the following as a direct consequence of Lemma 3: (i)  $Q_{31} \leq \epsilon$  implies  $|\bar{r}_{3xy} - \bar{t}| \leq$

$2\epsilon/(c_1 - \epsilon) = \epsilon_{3xy}$  and  $|\bar{r}_{3z}| \leq \sqrt{1 - [1 - 2\epsilon/(c_1 - \epsilon)]^2} = \epsilon_{3z}$  and (ii)  $Q_{22} \leq \epsilon$  implies  $|\bar{r}_{2xy} - \bar{n}| \leq 2\epsilon/(c_2 - \epsilon) = \epsilon_{2xy}$  and  $|\bar{r}_{2z}| \leq \sqrt{1 - [1 - 2\epsilon/(c_2 - \epsilon)]^2} = \epsilon_{2z}$ .

Now let us consider the cost of pinning the values of  $\bar{r}_1$ ,  $\bar{r}_2$ , and  $\bar{r}_3$ . It is best to start with  $\bar{r}_2$ . We consider the estimation of  $\bar{r}_2$  by  $\bar{n}$ . This leads to a new estimation of the matrix elements involving  $\rho_1$  and  $\rho_2$ . In particular, we have

$$\begin{aligned} Q_{11} &= \frac{1}{2}(x_{01} + \bar{r}_1 \bar{x}_1) \\ &= \frac{1}{2}[x_{01} + (-\bar{n})\bar{x}_1 + (\bar{n} - \bar{r}_{2xy})\bar{x}_1 + (\bar{r}_2 - \bar{y})\bar{x}_1 + (\bar{y} + \bar{r}_1)\bar{x}_1] \\ &\in \frac{1}{2}[x_{01} + (-\bar{n})\bar{x}_1 \pm (\epsilon_{2xy} + 2\sqrt{4\epsilon + \epsilon^2})], \end{aligned} \quad (D2)$$

$$\begin{aligned} Q_{12} &= \frac{1}{2}(x_{02} + \bar{r}_1 \bar{x}_2) \\ &= \frac{1}{2}[x_{02} + (-\bar{n})\bar{x}_2 + (\bar{n} - \bar{r}_{2xy})\bar{x}_2 + (\bar{r}_2 - \bar{y})\bar{x}_2 + (\bar{y} + \bar{r}_1)\bar{x}_2] \\ &\in \frac{1}{2}[x_{02} + (-\bar{n})\bar{x}_2 \pm (\epsilon_{2xy} + 2\sqrt{4\epsilon + \epsilon^2})], \end{aligned} \quad (D3)$$

$$Q_{21} = \frac{1}{2}(x_{01} + \bar{r}_2 \bar{x}_1) = \frac{1}{2}[x_{01} + \bar{n}\bar{x}_1 + (\bar{r}_{2xy} - \bar{n})\bar{x}_1] \in \frac{1}{2}[x_{01} + \bar{n}\bar{x}_1 \pm \epsilon_{2xy}], \quad (D4)$$

$$Q_{22} = \frac{1}{2}(x_{02} + \bar{r}_2 \bar{x}_2) = \frac{1}{2}[x_{02} + \bar{n}\bar{x}_2 + (\bar{r}_{2xy} - \bar{n})\bar{x}_2] \in \frac{1}{2}[x_{02} + \bar{n}\bar{x}_2 \pm \epsilon_{2xy}]. \quad (D5)$$

Let us consider the error one introduces by setting  $\bar{r}_{3xy}$  equal to  $\bar{t}$ . This only affects the last row of the correlation table

$$\begin{aligned} Q_{33} &= \frac{1}{2}(y_0 + \bar{y}\bar{r}_3) = \frac{1}{2}[y_0 + \bar{r}_2\bar{r}_3 + (\bar{y} - \bar{r}_2)\bar{r}_3] \in \frac{1}{2}[y_0 + \bar{r}_2\bar{r}_3 \pm \sqrt{4\epsilon + \epsilon^2}] \\ &\in \frac{1}{2}[y_0 + \bar{r}_{2xy}\bar{r}_{3xy} + \bar{r}_{2z}\bar{r}_{3z} \pm (\sqrt{4\epsilon + \epsilon^2})] \\ &\in \frac{1}{2}[y_0 + \bar{n}\bar{t} + \bar{n}(\bar{r}_{3xy} - \bar{t}) + (\bar{r}_{2xy} - \bar{n})\bar{t} + (\bar{r}_{2xy} - \bar{n})(\bar{r}_{3xy} - \bar{t}) \pm (\epsilon_{2z}\epsilon_{3z} + \sqrt{4\epsilon + \epsilon^2})] \\ &\in \frac{1}{2}[y_0 + \bar{n}\bar{t} \pm (\epsilon_{2xy} + \epsilon_{3xy} + \epsilon_{2xy}\epsilon_{3xy} + \epsilon_{2z}\epsilon_{3z} + \sqrt{4\epsilon + \epsilon^2})], \end{aligned} \quad (D6)$$

$$\begin{aligned} Q_{31} &= \frac{1}{2}(x_{01} + \bar{x}_1\bar{r}_3) = \frac{1}{2}(x_{01} + \bar{x}_1\bar{r}_{3xy}) = \frac{1}{2}[x_{01} + \bar{x}_1\bar{t} + \bar{x}_1(\bar{r}_{3xy} - \bar{t})] \\ &\in \frac{1}{2}[x_{01} + \bar{x}_1\bar{t} \pm \epsilon_{3xy}], \end{aligned} \quad (D7)$$

$$\begin{aligned} Q_{32} &= \frac{1}{2}(x_{02} + \bar{x}_2\bar{r}_3) = \frac{1}{2}(x_{02} + \bar{x}_2\bar{r}_{3xy}) = \frac{1}{2}[x_{02} + \bar{x}_2\bar{t} + \bar{x}_2(\bar{r}_{3xy} - \bar{t})] \\ &\in \frac{1}{2}[x_{02} + \bar{x}_2\bar{t} \pm \epsilon_{3xy}]. \end{aligned} \quad (D8)$$

To sum up, as long as  $Q \in P \pm \epsilon$ , Eqs. (D2)–(D8) should be satisfied. This further implies the constraints on the values on the right-hand sides of Eqs. (D2)–(D8). For example, Eq. (D2) together with  $Q_{11} \in P_{11} \pm \epsilon$  implies  $\frac{1}{2}(x_{01} + \bar{n}\bar{x}_1) \in P_{11} \pm [\epsilon + (\epsilon_{2xy} + 2\sqrt{4\epsilon + \epsilon^2})]$ , and so on. Therefore, given  $\epsilon > 0$  arbitrary, we have to ask whether there exist feasible  $\bar{n}$  and  $\bar{t}$  and  $(x_{10}, \bar{x}_1)$  and  $(x_{20}, \bar{x}_2)$  which are simulable by dichotomic measurements such that all these constraints are satisfied.

Notice that, due to the unitary freedom mentioned at the beginning of this Appendix, one can always set  $\bar{n} = (1, 0, 0)$ . If we further set  $\bar{t} = (\cos(\varphi), \sin(\varphi), 0)$ , asking for the existence of  $(x_{10}, \bar{x}_1)$  and  $(x_{20}, \bar{x}_2)$  which are simulable by dichotomic measurements is an SDP.

To conclude, we have the following algorithm. Scanning over  $\bar{t} = (\cos(\varphi), \sin(\varphi), 0)$  with certain finite step in  $\varphi$ , for any value of  $\bar{t}$ , set  $\bar{r}_2 = \bar{n}$ ,  $\bar{r}_1 = -\bar{n}$ ,  $\bar{y} = \bar{n}$ , and  $\bar{r}_3 = \bar{t}$ . Set a value of  $\epsilon$  and test the SDP of finding reducible  $M_{1|1}$  and  $M_{2|1}$  such that all the above-mentioned constraints are satisfied. At  $\epsilon = 0$ , the SDP is infeasible. We can implement a bisection method to find the exact transition point where the SDP is infeasible. We obtain the critical error tolerance  $\epsilon_c(\varphi)$ . By scanning all values of  $\varphi$ , we find  $\epsilon^* = \min \epsilon_c(\varphi)$ .

For practical purposes, the above-described procedure is sufficient. In principle, one may still object that the number  $\epsilon^*$  may not be reliable because of the finite step scanning over the values of  $\varphi$ . This objection can also be addressed. The idea is that the error introduced by the finite steps can in fact be

bounded by bounding the variation of the function  $\epsilon_c(\varphi)$ . That is, we can find a number  $C$  such that  $|\epsilon_c(\varphi + x) - \epsilon_c(\varphi)| \leq C\delta$  for all  $\varphi \in [0, 2\pi]$  and  $|x| \leq \delta$ . Note that the variation of  $\varphi$  only affects  $Q_{3k}$ . A variation of  $\delta$  in  $\varphi$  gives rise to a variation of  $\delta\vec{r}$  with  $|\delta\vec{r}| \leq \delta$ . One then sees that  $\delta Q_{3k} \leq \delta/2$ . Therefore,  $|\epsilon_c(\varphi + x) - \epsilon_c(\varphi)| \leq \delta/2$  for any value of  $\varphi$  and  $|x| \leq \delta$ . If one selects a step in  $\varphi$  with size  $\delta$ , the error in the global minimum  $\epsilon^* = \min \epsilon_c(\varphi)$  is bounded by  $\delta/4$  (since the maximum distance from any point to a computed point is  $\delta/2$ ). Taking  $\delta = 2\pi \times 10^{-5}$  is sufficient to bounding the error by  $\pi/2 \times 10^{-5}$ . An adaptive scheme of varying the step sizes over different regimes of  $\varphi$  can be utilized to speed up the computation.

#### APPENDIX E: CORRELATIONS FROM THE TRINE POVM

At the end of Sec. III C and in Fig. 1 we used an arrangement of states and effects involving the trine POVM. This arrangement produces the genuine trichotomic

correlations

$$\mathcal{P}_{\text{trine}} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{3}{4} \\ 0 & \frac{1}{2} & \frac{3}{4} \end{pmatrix} \quad (\text{E1})$$

with a Euclidean distance of  $r_2 \approx 0.0686$  to the simulable trichotomic correlations. The correlations  $\mathcal{P}_{\text{trine}}$  are obtained with the states

$$\begin{aligned} \rho_1 &= \frac{1}{2}(\mathbb{1} + \sigma_z), \\ \rho_2 &= \frac{1}{2}\left(\mathbb{1} - \frac{\sqrt{3}}{2}\sigma_x - \frac{1}{2}\sigma_z\right), \\ \rho_3 &= \frac{1}{2}\left(\mathbb{1} + \frac{\sqrt{3}}{2}\sigma_x - \frac{1}{2}\sigma_z\right) \end{aligned} \quad (\text{E2})$$

and the measurement effects  $M_{1|1} = \frac{2}{3}(\mathbb{1} - \rho_3)$ ,  $M_{2|1} = \frac{2}{3}(\mathbb{1} - \rho_2)$ ,  $M_{3|1} = \frac{2}{3}(\mathbb{1} - \rho_1)$ ,  $M_{1|2} = \mathbb{1} - \rho_1$ , and  $M_{2|2} = \rho_1$ .

- 
- [1] D. Bacon, A. M. Childs, and W. van Dam, Optimal measurements for the dihedral hidden subgroup problem, *Chicago J. Theor. Comput. Sci.* **2006**, 2 (2006).
- [2] D. Bacon, A. M. Childs, and W. van Dam, in *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science, Pittsburgh, 2005* (IEEE, Piscataway, 2005), pp. 469–478.
- [3] J. M. Renes, Spherical-code key-distribution protocols for qubits, *Phys. Rev. A* **70**, 052314 (2004).
- [4] C. H. Bennett, Quantum Cryptography Using Any Two Nonorthogonal States, *Phys. Rev. Lett.* **68**, 3121 (1992).
- [5] A. Acín, S. Pironio, T. Vértesi, and P. Wittek, Optimal randomness certification from one entangled bit, *Phys. Rev. A* **93**, 040102(R) (2016).
- [6] A. Bisio, G. Chiribella, G. M. D’Ariano, S. Facchini, and P. Perinotti, Optimal quantum tomography, *IEEE J. Sel. Top. Quantum Electron.* **15**, 1646 (2009).
- [7] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer, Dordrecht, 1993).
- [8] P. Mironowicz and M. Pawłowski, Experimentally feasible semi-device-independent certification of four-outcome positive-operator-valued measurements, *Phys. Rev. A* **100**, 030301(R) (2019).
- [9] A. Tavakoli, Semi-Device-Independent Certification of Independent Quantum State and Measurement Devices, *Phys. Rev. Lett.* **125**, 150503 (2020).
- [10] N. Miklin, J. J. Borkala, and M. Pawłowski, Semi-device-independent self-testing of unsharp measurements, *Phys. Rev. Res.* **2**, 033014 (2020).
- [11] E. S. Gómez, S. Gómez, P. González, G. Cañas, J. F. Barra, A. Delgado, G. B. Xavier, A. Cabello, M. Kleinmann, T. Vértesi, and G. Lima, Device-Independent Certification of a Non-projective Qubit Measurement, *Phys. Rev. Lett.* **117**, 260401 (2016).
- [12] A. Tavakoli, M. Smania, T. Vértesi, N. Brunner, and M. Bourennane, Self-testing nonprojective quantum measurements in prepare-and-measure experiments, *Sci. Adv.* **6**, eaaw6664 (2020).
- [13] M. Smania, P. Mironowicz, M. Nawareg, M. Pawłowski, A. Cabello, and M. Bourennane, Experimental certification of an informationally complete quantum measurement in a device-independent protocol, *Optica* **7**, 123 (2020).
- [14] G. Sentís, B. Gendra, S. D. Bartlett, and A. C. Doherty, Decomposition of any quantum measurement into extremals, *J. Phys. A: Math. Theor.* **46**, 375302 (2013).
- [15] M. Oszmaniec, L. Guerini, P. Wittek, and A. Acín, Simulating Positive-Operator-Valued Measures with Projective Measurements, *Phys. Rev. Lett.* **119**, 190501 (2017).
- [16] P. Busch, M. Grabowski, and P. J. Lahti, *Measurements, Operational Quantum Physics* (Springer 2009), Pt. 1, Chap. 4, pp. 95–115.
- [17] I. D. Ivanovic, How to differentiate between non-orthogonal states, *Phys. Lett. A* **123**, 257 (1987).
- [18] J. Hoffmann, C. Spee, O. Gühne, and C. Budroni, Structure of temporal correlations of a qubit, *New J. Phys.* **20**, 102001 (2018).
- [19] P. E. Frenkel and M. Weiner, Classical information storage in an  $n$ -level quantum system, *Commun. Math. Phys.* **340**, 563 (2015).
- [20] M. Kleinmann and A. Cabello, Quantum Correlations are Stronger Than All Nonsignaling Correlations Produced by  $n$ -Outcome Measurements, *Phys. Rev. Lett.* **117**, 150401 (2016).
- [21] K. F. Pál and T. Vértesi, Maximal violation of a bipartite three-setting, two-outcome Bell inequality using infinite-dimensional quantum systems, *Phys. Rev. A* **82**, 022116 (2010).
- [22] N. Brunner, M. Navascués, and T. Vértesi, Dimension Witnesses and Quantum State Discrimination, *Phys. Rev. Lett.* **110**, 150501 (2013).
- [23] M. D. Mazurek, M. F. Pusey, K. J. Resch, and R. W. Spekkens, Experimentally bounding deviations from quantum theory in the landscape of generalized probabilistic theories, *PRX Quantum* **2**, 020302 (2021).