Self-testing of Pauli observables for device-independent entanglement certification

Joseph Bowles,[1] Ivan Šupić,[1] Daniel Cavalcanti,[1] and Antonio Acín[1,2]
[1]*ICFO-Institut de Ciencies Fotoniques, The Barcelona Institute of Science and Technology, 08860 Castelldefels (Barcelona), Spain*
[2]*ICREA - Institució Catalana de Recerca i Estudis Avancats, 08011 Barcelona, Spain*

(Received 27 March 2018; revised manuscript received 26 September 2018; published 29 October 2018)

We present self-testing protocols to certify the presence of tensor products of Pauli measurements on maximally entangled states of local dimension 2^n for $n \in \mathbb{N}$. This provides self-tests of sets of informationally complete measurements in arbitrarily high dimension. We then show that this can be used for the device-independent certification of the entanglement of all bipartite entangled states by exploiting a connection to measurement-device-independent entanglement witnesses and quantum networks. This work extends a more compact parallel work on the same subject [Bowles *et al.*, Phys. Rev. Lett. **121**, 180503 (2018)] and provides all the required technical proofs.

DOI: 10.1103/PhysRevA.98.042336

I. INTRODUCTION

Unjustified or mistaken assumptions about the physics of a quantum information protocol can result in errors that jeopardize the protocol's validity [1,2]. The *device-independent* approach attempts to overcome this problem by keeping assumptions to a minimum; devices in the protocol are treated as black boxes, and the only information available is their input-output statistics. Interestingly, due to the existence of quantum nonlocality [3,4], protocols can still be made to function in this scenario and many quantum information tasks now have device-independent formulations, including protocols for quantum random number certification [5–7], quantum key distribution [8–10], and the characterisation of quantum properties [11,12].

A common device-independent task is that of *entanglement certification*. Here one aims to certify the presence of entanglement in a quantum state from the correlations between local measurement outcomes, and is typically achieved via the violation of a Bell inequality. The central limitation here is that there exist entangled mixed states that admit a so-called local hidden variable model [13–15] and thus do not violate any Bell inequality. Device-independent entanglement certification of such states is therefore impossible via the standard approach. A partial solution to this problem recently came in the form of measurement-device-independent entanglement witnesses (MDIEWs) [16–18]. Here one can achieve entanglement certification of all entangled states by replacing the classical inputs in a Bell test by a set of trusted quantum input states. This approach, however, is only partially device independent since it requires perfect knowledge of the input states.

A closely related task to entanglement certification is that of *self-testing* [19]. In a self-testing protocol, one aims to certify, or self-test, the presence of a target entangled state and/or target set of measurements via the observation of nonlocal correlations. Essentially, this requires finding a Bell inequality whose maximum violation is achieved uniquely by the target state and measurements of interest. A significant literature on self-testing exists [20–26], and it is known for example that all bipartite pure entangled states can be self-tested [27]. The self-testing of quantum measurements is however much less explored, although some results are known [28,29].

In this work (see also [30] for a more compact version) we combine results in the field of self-testing with techniques from MDIEWs to construct device-independent protocols that are capable of certifying the entanglement of all bipartite entangled states. To do this, we move to a scenario involving a network of quantum states that allows us to overcome the limitations of the standard approach. Intuitively, our protocols can be understood as a device-independent extension of MDIEWs, in which the input quantum states are certified device independently via a self-testing protocol. The technical preliminaries to this result include new results concerning the parallel self-testing of Pauli observables and may be of independent interest. In particular, we prove self-testing of tensor products of Pauli observables on maximally entangled states of local dimension 2^n, $n \in \mathbb{N}$, treating a well-known problem that arises when dealing with complex-valued measurements. We note that an analogous result to this was independently proven in Ref. [31] in the context of delegated quantum computation.

The paper is organized as follows. The first two sections focus on the technical ground work in self-testing that are needed for our entanglement certification protocols. In Sec. II we introduce self-testing and revisit the problem that arises with complex-valued measurements. In Sec. II A we focus on the simplest case of two qubits and prove self-testing of the three Pauli observables, before tackling the more involved case of general dimension in Secs. II B and II C and discussing noise-robust versions of these results in Sec. II D. We then move to our protocols for entanglement certification, outlining our network scenario in Sec. III, presenting our entanglement certification protocols in Secs. III A–III D, and finally discussing our results.

II. SELF-TESTING

Suppose two parties, Charlie and Alice,[1] share the quantum state $|\psi\rangle$ and perform local measurements labeled by z and x, obtaining outcomes c and a. From the Born rule, the observed probabilities take the form

$$p(ca|zx) = \mathrm{tr}[|\psi\rangle\langle\psi|\mathsf{M}_{c|z} \otimes \mathsf{M}_{a|x}], \tag{1}$$

where $\mathsf{M}_{c|z}$, $\mathsf{M}_{a|x}$ denote the local measurement operators, and where we have purified states and measurements so that our state is a pure state and our measurements projective. In principle, many different combinations of states and measurements could give rise to the same correlations $p(ca|zx)$. To self-test a target quantum state $|\psi'\rangle$, one must find correlations which are produced uniquely by $|\psi'\rangle$ up to a certain equivalence class, hence certifying the state $|\psi'\rangle$ (up to equivalence) from knowledge of the correlations alone. In the first works on self-testing, this equivalence class is captured by the notion of a local isometry, which takes into account the possibility of unobservable local unitary operations applied to the state and measurements, possible embedding in a Hilbert space of larger dimension and/or the existence of additional degrees of freedom. Note that via the Schmidt decomposition, the freedom of local unitary operations implies that one may assume that the target state $|\psi'\rangle$ can be expressed with real numbers only without loss of generality. The precise definition of self-testing of quantum states is then as follows.

Definition II.1. We say that the correlations $p^*(ca|zx)$ self-test the state $|\psi'\rangle \in \mathcal{H}_{\mathrm{C}'} \otimes \mathcal{H}_{\mathrm{A}'}$ if for all states and all measurement operators satisfying (1) for $p(ca|zx) = p^*(ca|zx)$ there exist Hilbert spaces $\mathcal{H}_{\mathrm{C}}$, $\mathcal{H}_{\mathrm{A}}$ such that $|\psi\rangle \in \mathcal{H}_{\mathrm{C}} \otimes \mathcal{H}_{\mathrm{A}}$, a local auxiliary state $|00\rangle \in \mathcal{H}_{\mathrm{C}'} \otimes \mathcal{H}_{\mathrm{A}'}$ and a local unitary operator U such that

$$U[|\psi\rangle \otimes |00\rangle] = |\xi\rangle \otimes |\psi'\rangle, \tag{2}$$

where $|\xi\rangle \in \mathcal{H}_{\mathrm{C}} \otimes \mathcal{H}_{\mathrm{A}}$ (usually called a junk state) is any state representing possible additional degrees of freedom.

Intuitively, self-testing means proving the existence of local channels (given by the local unitaries and local auxiliary states) which extract the target state $|\psi'\rangle$ from the physical state $|\psi\rangle$ into the $\mathcal{H}_{\mathrm{C}'} \otimes \mathcal{H}_{\mathrm{A}'}$ space.

One may further be interested in certifying that the measurement operators are equivalent to some target measurements $\{\mathsf{M}'_{c|z}\}$, $\{\mathsf{M}'_{a|x}\}$ acting on $|\psi'\rangle$. To begin with, let us assume that the target measurements can be expressed using real numbers alone, i.e., $(\mathsf{M}'_{c|z})^* = \mathsf{M}'_{c|z}$ for all c, z and $(\mathsf{M}'_{a|x})^* = \mathsf{M}'_{a|x}$ for all a, x. We then have the following definition.

Definition II.2. We say that the correlations $p^*(ca|zx)$ self-test the state $|\psi'\rangle$ and real-valued measurements $\{\mathsf{M}'_{c|z}\}$, $\{\mathsf{M}'_{a|x}\}$ if $p^*(ca|zx)$ self-tests the state $|\psi'\rangle$ according to definition II.1 and furthermore

$$U[\mathsf{M}_{c|z} \otimes \mathsf{M}_{a|x}\,|\psi\rangle \otimes |00\rangle] = |\xi\rangle \otimes (\mathsf{M}'_{c|z} \otimes \mathsf{M}'_{a|x}|\psi'\rangle)$$

for each c, a, z, x.

[1]We avoid the usual convention of Alice and Bob for readability with later sections of this paper where our choice will become more natural.

In other words, applying the measurements $\mathsf{M}_{c|z}$, $\mathsf{M}_{a|x}$ to the state $|\psi\rangle$ is equivalent to applying $\mathsf{M}'_{c|z}$, $\mathsf{M}'_{a|x}$ to $|\psi'\rangle$ under the action of the local unitaries.

For measurements that cannot be expressed using real numbers alone an additional complication arises, as noted in the early works on self-testing [32] (see also [28,29]). This is due to the fact that quantum correlations are invariant under transposition (or equivalently, complex conjugation) of the state and measurement operators:

$$\mathrm{tr}[|\psi'\rangle\langle\psi'|\mathsf{M}'_{c|z} \otimes \mathsf{M}'_{a|x}] = \mathrm{tr}\left[|\psi'\rangle\langle\psi'|\mathsf{M}'^{T}_{c|z} \otimes \mathsf{M}'^{T}_{a|x}\right] \tag{3}$$

(where M^T denotes the transposition operation and we assume the state $|\psi'\rangle$ to be real as above). Note that the transposition operation maps valid measurement operators to valid measurement operators, however is not unitary. This means that the measurements $\{\mathsf{M}'_{c|z}\}$, $\{\mathsf{M}'_{a|x}\}$ cannot be self-tested using definition II.2. That is, there always exists an alternative realization using the transposed measurements which cannot be brought to the target measurements using local isometries alone. For such measurements, the most we can hope to certify is that the measurement operators correspond to the target set up to the additional freedom of local transpositions on both subsystems. To deal with this possibility and following the method of [28], we introduce additional local Hilbert spaces $\mathcal{H}_{\mathrm{C}''}$ and $\mathcal{H}_{\mathrm{A}''}$ which act as a control space for possible transposition of the measurement operators. Our precise definition of self-testing is as follows.

Definition II.3. We say that the correlations $p^*(ca|zx)$ self-test the state $|\psi'\rangle \in \mathcal{H}_{\mathrm{C}'} \otimes \mathcal{H}_{\mathrm{A}'}$ and (complex-valued) measurements $\{\mathsf{M}'_{c|z}\}$, $\{\mathsf{M}'_{a|x}\}$ if for all states and all measurement operators satisfying (1) for $p(ca|zx) = p^*(ca|zx)$ there exist Hilbert spaces $\mathcal{H}_{\mathrm{C}}$, $\mathcal{H}_{\mathrm{A}}$ such that $|\psi\rangle \in \mathcal{H}_{\mathrm{C}} \otimes \mathcal{H}_{\mathrm{A}}$, a local auxiliary state $|00\rangle \in [\mathcal{H}_{\mathrm{C}''} \otimes \mathcal{H}_{\mathrm{C}'}] \otimes [\mathcal{H}_{\mathrm{A}''} \otimes \mathcal{H}_{\mathrm{A}'}]$ and a local unitary operator U such that

$$\begin{aligned} &U[\mathsf{M}_{c|z} \otimes \mathsf{M}_{a|x}\,|\psi\rangle \otimes |00\rangle] \\ &\quad = \tilde{\mathsf{M}}_{c|z} \otimes \tilde{\mathsf{M}}_{a|c}[|\xi_0\rangle \otimes |00\rangle + |\xi_1\rangle \otimes |11\rangle] \otimes |\psi'\rangle, \end{aligned} \tag{4}$$

where $|\xi_j\rangle \in \mathcal{H}_{\mathrm{C}} \otimes \mathcal{H}_{\mathrm{A}}$ are some unknown subnormalized junk states such that $\langle\xi_0|\xi_0\rangle + \langle\xi_1|\xi_1\rangle = 1$ and the $\tilde{\mathsf{M}}$ operators are related to the target measurements by

$$\tilde{\mathsf{M}}_{c|z} = \mathbb{1}^{\mathrm{C}} \otimes [\mathsf{M}_0 \otimes \mathsf{M}'_{c|z} + \mathsf{M}_1 \otimes (\mathsf{M}'_{c|z})^T], \tag{5}$$

$$\tilde{\mathsf{M}}_{a|x} = \mathbb{1}^{\mathrm{A}} \otimes [\mathsf{M}_0 \otimes \mathsf{M}'_{a|x} + \mathsf{M}_1 \otimes (\mathsf{M}'_{a|x})^T], \tag{6}$$

with $\mathsf{M}_0 + \mathsf{M}_1 = \mathbb{1}^{\mathrm{C}''}$ and $\langle 0|\mathsf{M}_0|0\rangle = \langle 1|\mathsf{M}_1|1\rangle = 1$.

The above measurements can be understood as "controlled transposition" measurements: one first measures the double primed auxiliary spaces with the measurement $\{\mathsf{M}_0, \mathsf{M}_1\}$; conditioned on this outcome, one then measures the target measurement or its transposition on the target state $|\psi'\rangle$. Due to the form of the measurement operators and state $|\xi_0\rangle \otimes |00\rangle + |\xi_1\rangle \otimes |11\rangle$, one sees that this transposition is correlated between Charlie and Alice, as implied from (3). The probability that this transposition is applied depends on the norm of the vectors $|\xi_j\rangle$, however it is generally unknown since the self-testing data does not allow one to infer the form of these states. Note that one may only wish to self-test a set of

measurements for one of the parties, say Charlie (as will be the case for us); here one would simply replace the measurement operators for Alice by the identity operator in the above.

The central task in self-testing is thus to construct the local unitary U in order to prove statements following the above definitions. In order to do this, one typically considers linear combinations of the probabilities $p(ca|zx)$ (corresponding to some Bell inequality) of the form

$$\mathcal{I}[p(ca|zx)] = \sum_{c,a,z,x} \beta_{ca}^{zx}\, p(ca|zx), \tag{7}$$

for which the maximal value in quantum theory $\mathcal{I} = \mathcal{I}_{\max}$ occurs using the target state and measurements. The observation $\mathcal{I} = \mathcal{I}_{\max}$ then implies relations between the state and measurements performed in the experiment via (7), and one can prove the existence of the local unitary from the measurement operators themselves. A large number of self-testing results are known. For example, if (7) corresponds to the CHSH Bell inequality, maximum violation implies that one can self-test the presence of a maximally entangled state of dimension two $|\Phi^+\rangle = \frac{1}{\sqrt{2}}[|00\rangle + |11\rangle]$, and measurements of σ_x, σ_z for Charlie and $[\sigma_x \pm \sigma_z]/\sqrt{2}$ for Alice [19,23,33]. More generally, one can self-test any pure bipartite entangled two-qubit state $|\psi\rangle = \cos\theta|00\rangle + \sin\theta|11\rangle$ when (7) corresponds to the tilted CHSH Bell inequality [21]. Self-testing of higher dimensional bipartite pure states is given in Refs. [24–27]. Furthermore, a large class of multipartite states can be self-tested by exploiting the methods applied to self-testing of bipartite states [22].

The majority of self-tests mentioned above are useful for the certification of measurements. However, most of these results apply to the self-testing of real-valued measurements due to the added complication definition II.3. The simplest set of measurements which cannot be expressed using real numbers alone is given by the three Pauli observables $\sigma_z, \sigma_x, \sigma_y$. In Sec. II A we prove self-testing statements for these measurements, inspired by the approach of [28] where similar results were obtained. We then extend this to a parallel self-test in Secs. II B and II C in order to prove self-testing statements for n-fold tensor products of the Pauli measurements, which form an informationally complete set in dimension 2^n.

A. Self-testing of Pauli measurements

We begin by proving a self-testing statement for the maximally entangled state of two qubits $|\Phi^+\rangle = \frac{1}{\sqrt{2}}[|00\rangle + |11\rangle]$ and the three Pauli observables for Charlie. Since there does not exist a two-qubit basis in which these observables can be written using real numbers only, our self-testing statement will be of the form of definition II.3. We note that this is not the first proof of such a result; similar results have been obtained in previous works by generalizing the Mayers-Yao self-test [28], by studying the properties of the "elegant" Bell inequality [34,35] and combinations of the CHSH Bell inequality [34] and in a more general approach to the problem [29] focused on commutation relations.

Before proceeding we first clarify some notation. Superscript of an operator denotes the Hilbert space on or in which the operator acts or lives, e.g., X^{C} denotes a linear operator on the space $\mathcal{H}_{\mathrm{C}}$ and $|\psi\rangle^{\mathrm{CA}} \in \mathcal{H}_{\mathrm{C}} \otimes \mathcal{H}_{\mathrm{A}}$. Unless explicitly written, we omit tensor products acting on the remaining Hilbert space, e.g., $\mathsf{X}^{\mathrm{C}}|\psi\rangle^{\mathrm{CA}}$ should be understood as $\mathsf{X}^{\mathrm{C}} \otimes \mathbb{1}^{\mathrm{A}}|\psi\rangle^{\mathrm{CA}}$. This convention then follows for the product of operators, e.g., $\mathsf{X}^{\mathrm{C}}\mathsf{E}^{\mathrm{A}}|\psi\rangle^{\mathrm{CA}}$ should be understood as $\mathsf{X}^{\mathrm{C}} \otimes \mathsf{E}^{\mathrm{A}}|\psi\rangle^{\mathrm{CA}}$.

The scenario we consider for the self-testing is as follows. Charlie and Alice share a bipartite quantum state $|\psi\rangle \in \mathcal{H}_{\mathrm{C}} \otimes \mathcal{H}_{\mathrm{A}}$. Charlie has a choice of three measurements $z = 1, 2, 3$, with outcomes $c = \pm 1$ denoted by the observables X^{C}, Y^{C}, and Z^{C}. Alice has a choice of six ± 1 valued measurements $x = 1, \ldots, 6$, $a = \pm 1$, denoted by the observables $\mathsf{D}^{\mathrm{A}}_{\mathrm{z,x}}, \mathsf{E}^{\mathrm{A}}_{\mathrm{z,x}}, \mathsf{D}^{\mathrm{A}}_{\mathrm{x,y}}, \mathsf{E}^{\mathrm{A}}_{\mathrm{x,y}}, \mathsf{D}^{\mathrm{A}}_{\mathrm{z,y}}, \mathsf{E}^{\mathrm{A}}_{\mathrm{z,y}}$. Note that each of these observables is Hermitian and unitary. We then consider the following Bell operator (introduced in Ref. [34]), which we call the triple CHSH Bell operator

$$\begin{aligned}\mathcal{B} = {} & \mathsf{Z}^{\mathrm{C}}\big(\mathsf{D}^{\mathrm{A}}_{\mathrm{z,x}} + \mathsf{E}^{\mathrm{A}}_{\mathrm{z,x}}\big) + \mathsf{X}^{\mathrm{C}}\big(\mathsf{D}^{\mathrm{A}}_{\mathrm{z,x}} - \mathsf{E}^{\mathrm{A}}_{\mathrm{z,x}}\big) \\ & + \mathsf{Z}^{\mathrm{C}}\big(\mathsf{D}^{\mathrm{A}}_{\mathrm{z,y}} + \mathsf{E}^{\mathrm{A}}_{\mathrm{z,y}}\big) - \mathsf{Y}^{\mathrm{C}}\big(\mathsf{D}^{\mathrm{A}}_{\mathrm{z,y}} - \mathsf{E}^{\mathrm{A}}_{\mathrm{z,y}}\big) \\ & + \mathsf{X}^{\mathrm{C}}\big(\mathsf{D}^{\mathrm{A}}_{\mathrm{x,y}} + \mathsf{E}^{\mathrm{A}}_{\mathrm{x,y}}\big) - \mathsf{Y}^{\mathrm{C}}\big(\mathsf{D}^{\mathrm{A}}_{\mathrm{x,y}} - \mathsf{E}^{\mathrm{A}}_{\mathrm{x,y}}\big).\end{aligned} \tag{8}$$

This Bell operator consists of a sum of three CHSH Bell operators; each line itself is a CHSH Bell operator and each X, Y, and Z observable appears in two of the lines. The correlations that we use for self-testing correspond to those which maximize $\langle\psi|\mathcal{B}|\psi\rangle$, which has maximum value $6\sqrt{2}$ (since each CHSH operator is upper bounded by $2\sqrt{2}$). This can be achieved by taking the following states and observables

$$\begin{aligned} |\psi\rangle &= |\Phi^+\rangle = \frac{1}{\sqrt{2}}[|00\rangle + |11\rangle], \\ \mathsf{Z}^C &= \sigma_z, \quad \mathsf{X}^C = \sigma_x, \quad \mathsf{Y}^C = \sigma_y, \\ \mathsf{D}^{\mathrm{A}}_{i,j} &= \frac{\sigma_i + \sigma_j}{\sqrt{2}}, \quad \mathsf{E}^{\mathrm{A}}_{i,j} = \frac{\sigma_i - \sigma_j}{\sqrt{2}}, \end{aligned} \tag{9}$$

for $(i, j) = (\mathrm{z,x}), (\mathrm{z,y}), (\mathrm{x,y})$. The basic intuition of the self-testing is that since maximal violation of a single CHSH inequality requires anticommuting qubit observables on a maximally entangled state [36], the maximum value of (8) should imply three mutually anticommuting observables on the maximally entangled state, given by the three Pauli observables (or their transpositions). Indeed, we will see that this is the case.

One way to achieve this is to build a sum-of-squares (SOS) decomposition of the shifted Bell operator $6\sqrt{2}\mathbb{1} - \mathcal{B}$ of the form

$$6\sqrt{2}\mathbb{1} - \mathcal{B} = \sum_{\lambda} P_\lambda^\dagger P_\lambda. \tag{10}$$

Such a decomposition is given by

$$\begin{aligned} & 2(6\sqrt{2}\mathbb{1} - \mathcal{B}) \\ & = \left[\mathsf{Z}^{\mathrm{C}} - \frac{1}{\sqrt{2}}\big(\mathsf{D}^{\mathrm{A}}_{\mathrm{z,x}} + \mathsf{E}^{\mathrm{A}}_{\mathrm{z,x}}\big)\right]^2 + \left[\mathsf{X}^{\mathrm{C}} - \frac{1}{\sqrt{2}}\big(\mathsf{D}^{\mathrm{A}}_{\mathrm{z,x}} - \mathsf{E}^{\mathrm{A}}_{\mathrm{z,x}}\big)\right]^2 \\ & + \left[\mathsf{Z}^{\mathrm{C}} - \frac{1}{\sqrt{2}}\big(\mathsf{D}^{\mathrm{A}}_{\mathrm{z,y}} + \mathsf{E}^{\mathrm{A}}_{\mathrm{z,y}}\big)\right]^2 + \left[\mathsf{Y}^{\mathrm{C}} + \frac{1}{\sqrt{2}}\big(\mathsf{D}^{\mathrm{A}}_{\mathrm{z,y}} - \mathsf{E}^{\mathrm{A}}_{\mathrm{z,y}}\big)\right]^2 \\ & + \left[\mathsf{X}^{\mathrm{C}} - \frac{1}{\sqrt{2}}\big(\mathsf{D}^{\mathrm{A}}_{\mathrm{x,y}} + \mathsf{E}^{\mathrm{A}}_{\mathrm{x,y}}\big)\right]^2 + \left[\mathsf{Y}^{\mathrm{C}} + \frac{1}{\sqrt{2}}\big(\mathsf{D}^{\mathrm{A}}_{\mathrm{x,y}} - \mathsf{E}^{\mathrm{A}}_{\mathrm{x,y}}\big)\right]^2. \end{aligned} \tag{11}$$

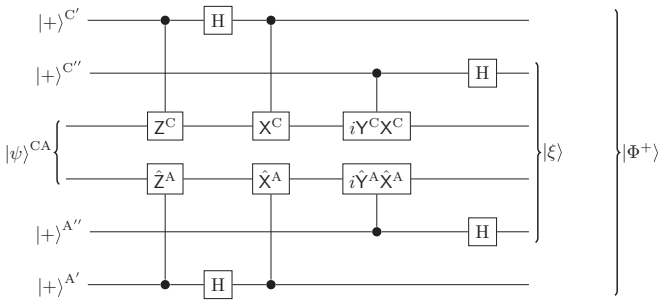

FIG. 1. Self-testing circuit used for the proof of Lemma 1. The unitaries $\hat{\mathsf{Z}}^{\mathrm{A}}$, $\hat{\mathsf{X}}^{\mathrm{A}}$, $\hat{\mathsf{Y}}^{\mathrm{A}}$ can be found in Appendix A.

Here the P_λ's are Hermitian and so $P_\lambda^\dagger P_\lambda = P_\lambda^2$. At maximal value one has $\langle\psi|\mathcal{B}|\psi\rangle = 6\sqrt{2}$ and so

$$\sum_\lambda \langle\psi|P_\lambda^\dagger P_\lambda|\psi\rangle = 0. \tag{12}$$

Since each term in the above is greater or equal to zero we have $P_\lambda|\psi\rangle = 0$ for all λ. Applying this to the SOS decomposition (11) gives

$$\mathsf{Z}^{\mathrm{C}}|\psi\rangle = \frac{1}{\sqrt{2}}\big[\mathsf{D}^{\mathrm{A}}_{\mathrm{z,x}} + \mathsf{E}^{\mathrm{A}}_{\mathrm{z,x}}\big]|\psi\rangle = \frac{1}{\sqrt{2}}\big[\mathsf{D}^{\mathrm{A}}_{\mathrm{z,y}} + \mathsf{E}^{\mathrm{A}}_{\mathrm{z,y}}\big]|\psi\rangle, \tag{13}$$

$$\mathsf{X}^{\mathrm{C}}|\psi\rangle = \frac{1}{\sqrt{2}}\big[\mathsf{D}^{\mathrm{A}}_{\mathrm{z,x}} - \mathsf{E}^{\mathrm{A}}_{\mathrm{z,x}}\big]|\psi\rangle = \frac{1}{\sqrt{2}}\big[\mathsf{D}^{\mathrm{A}}_{\mathrm{x,y}} + \mathsf{E}^{\mathrm{A}}_{\mathrm{x,y}}\big]|\psi\rangle, \tag{14}$$

$$\mathsf{Y}^{\mathrm{C}}|\psi\rangle = \frac{1}{\sqrt{2}}\big[\mathsf{E}^{\mathrm{A}}_{\mathrm{z,y}} - \mathsf{D}^{\mathrm{A}}_{\mathrm{z,y}}\big]|\psi\rangle = \frac{1}{\sqrt{2}}\big[\mathsf{E}^{\mathrm{A}}_{\mathrm{x,y}} - \mathsf{D}^{\mathrm{A}}_{\mathrm{x,y}}\big]|\psi\rangle. \tag{15}$$

Since for any two unitary observables G_1 and G_2, the composite observables $\frac{\mathsf{G}_1+\mathsf{G}_2}{\sqrt{2}}$ and $\frac{\mathsf{G}_1-\mathsf{G}_2}{\sqrt{2}}$ anticommute by construction, from the above three equations it follows that on the support of state $|\psi\rangle$ observables Z^{C}, X^{C}, and Y^{C} mutually anticommute:

$$\{\mathsf{Z}^{\mathrm{C}}, \mathsf{X}^{\mathrm{C}}\}|\psi\rangle = \{\mathsf{Z}^{\mathrm{C}}, \mathsf{Y}^{\mathrm{C}}\}|\psi\rangle = \{\mathsf{X}^{\mathrm{C}}, \mathsf{Y}^{\mathrm{C}}\}|\psi\rangle = 0. \tag{16}$$

The conditions (13)–(15) and (16) allow us to construct a local unitary which will give us our desired self-testing. This unitary can be understood via the circuit of Fig. 1, and is based on the swap gate introduced in Ref. [23] and is the same as the circuit found in Ref. [28]. The unitaries $\hat{\mathsf{Z}}^{\mathrm{A}}$, $\hat{\mathsf{X}}^{\mathrm{A}}$, $\hat{\mathsf{Y}}^{\mathrm{A}}$ are regularized versions of the operators

$$\mathsf{Z}^{\mathrm{A}} = \frac{\mathsf{D}^{\mathrm{A}}_{\mathrm{z,x}} + \mathsf{E}^{\mathrm{A}}_{\mathrm{z,x}}}{\sqrt{2}}, \quad \mathsf{X}^{\mathrm{A}} = \frac{\mathsf{D}^{\mathrm{A}}_{\mathrm{z,x}} - \mathsf{E}^{\mathrm{A}}_{\mathrm{z,x}}}{\sqrt{2}}, \quad \mathsf{Y}^{\mathrm{A}} = \frac{\mathsf{E}^{\mathrm{A}}_{\mathrm{z,y}} - \mathsf{D}^{\mathrm{A}}_{\mathrm{z,y}}}{\sqrt{2}}.$$

For example, $\hat{\mathsf{Z}}^{\mathrm{A}}$ is obtained by setting all zero eigenvalues of Z^{A} to one and then defining $\hat{\mathsf{Z}}^{\mathrm{A}} = \mathsf{Z}^{\mathrm{A}}|\mathsf{Z}^{\mathrm{A}}|^{-1}$. Using standard techniques (see Appendix A), these can be shown to act in the same way as the nonregularized versions. With this we are ready to present the first of our self-testing lemmas.

Lemma 1. Let the state $|\psi\rangle \in \mathcal{H}_{\mathrm{C}} \otimes \mathcal{H}_{\mathrm{A}}$ and ± 1 outcome observables X^{C}, Y^{C}, Z^{C}, $\mathsf{D}^{\mathrm{A}}_{\mathrm{z,x}}$, $\mathsf{E}^{\mathrm{A}}_{\mathrm{z,x}}$, $\mathsf{D}^{\mathrm{A}}_{\mathrm{x,y}}$, $\mathsf{E}^{\mathrm{A}}_{\mathrm{x,y}}$, $\mathsf{D}^{\mathrm{A}}_{\mathrm{z,y}}$, $\mathsf{E}^{\mathrm{A}}_{\mathrm{z,y}}$ satisfy

$$\langle\psi|\mathcal{B}|\psi\rangle = 6\sqrt{2}. \tag{17}$$

Then there exist local auxiliary states $|00\rangle \in [\mathcal{H}_{\mathrm{C}''} \otimes \mathcal{H}_{\mathrm{C}'}] \otimes [\mathcal{H}_{\mathrm{A}''} \otimes \mathcal{H}_{\mathrm{A}'}]$ and a local unitary U such that

$$U[|\psi\rangle \otimes |00\rangle] = |\xi\rangle \otimes |\Phi^+\rangle^{\mathrm{C'A'}}, \tag{18}$$

$$U[\mathsf{X}^{\mathrm{C}}|\psi\rangle \otimes |00\rangle] = |\xi\rangle \otimes \sigma_{\mathrm{x}}^{\mathrm{C'}}|\Phi^+\rangle^{\mathrm{C'A'}}, \tag{19}$$

$$U[\mathsf{Z}^{\mathrm{C}}|\psi\rangle \otimes |00\rangle] = |\xi\rangle \otimes \sigma_{\mathrm{z}}^{\mathrm{C'}}|\Phi^+\rangle^{\mathrm{C'A'}}, \tag{20}$$

$$U[\mathsf{Y}^{\mathrm{C}}|\psi\rangle \otimes |00\rangle] = \sigma_{\mathrm{z}}^{\mathrm{C''}}|\xi\rangle \otimes \sigma_{\mathrm{y}}^{\mathrm{C'}}|\Phi^+\rangle^{\mathrm{C'A'}}, \tag{21}$$

where $|\xi\rangle$ takes the form

$$|\xi\rangle = |\xi_0\rangle^{\mathrm{CA}} \otimes |00\rangle^{\mathrm{C''A''}} + |\xi_1\rangle^{\mathrm{CA}} \otimes |11\rangle^{\mathrm{C''A''}}. \tag{22}$$

Note that the complex observable σ_{y} has an additional σ_{z} measurement on the C″ space, as expected from definition II.3. Hence, the measurement Y can be understood as first measuring σ_{z} on the state $|\xi\rangle$, whose outcome decides whether $\pm\sigma_{\mathrm{y}}$ is performed on the state $|\Phi^+\rangle$. The probability that the observables $\{\sigma_{\mathrm{x}}, \sigma_{\mathrm{y}}, \sigma_{\mathrm{z}}\}$ are used rather than the transposed measurements $\{\sigma_{\mathrm{x}}, -\sigma_{\mathrm{y}}, \sigma_{\mathrm{z}}\}$ is given by the probability to obtain $+1$ for the $\sigma_{\mathrm{z}}^{\mathrm{C''}}$ measurement. As mentioned in Sec. II, this probability remains unknown since one does not know the precise form of $|\xi\rangle$ from the self-testing correlations alone. The proof of Lemma 1 can be found in Appendix A.

B. Parallel self-testing of Pauli observables

The protocol described above can be extended to a parallel self-test. Here our aim is to self-test the n-fold tensor product of the maximally entangled state $|\Phi^+\rangle^{\otimes n}$ (which itself is a maximally entangled state of dimension 2^n) and all combinations of n-fold tensor products of Pauli measurements for Charlie, i.e., $\sigma_{i_1} \otimes \sigma_{i_2} \otimes \cdots \otimes \sigma_{i_n}$ for $i_j = \mathrm{x}, \mathrm{y}, \mathrm{z}$. This is achieved by an n-fold maximal parallel violation of the Bell inequality used in Lemma 1. As a basis we use the techniques of [37], where parallel self-testing of σ_{x} and σ_{z} observables on the maximally entangled state was proven. Besides [37], parallel self-testing of n-fold tensor products of maximally entangled pairs of qubits has been presented in [24,38] and in Ref. [25] for $n = 2$. This section can thus be seen as an extension of these results to all three Pauli observables. Although we use the term "self-testing" here, we will see that simply performing the protocol of Lemma 1 in parallel does not lead to a self-test according to definition II.3. In the following subsection we correct this by adding additional Bell state measurements between local subsystems.

The scenario we consider is as follows. Charlie and Alice share the state $|\psi\rangle \in \mathcal{H}_{\mathrm{C}} \otimes \mathcal{H}_{\mathrm{A}}$. Charlie has a choice of 3^n measurements collected into the vector $\mathbf{z} = (z_1, z_2, \ldots, z_n)$ with $z_i = 1, 2, 3$, and each measurement has 2^n possible outcomes given by $\mathbf{c} = (c_1, c_2, \ldots, c_n)$ with $c_i = \pm 1$. Similarly, Alice has a choice of 6^n measurements given by the vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ with $x_i = 1, 2, 3, 4, 5, 6$, each with 2^n possible outputs given by $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ with $a_i = \pm 1$. Fixing a value of i we thus have three possible settings for Charlie and six for Alice, corresponding to the self-test of the previous section that we now perform in parallel. In order to achieve this we will define an analogous Bell operator to (8) for each value of i.

To this end, we denote Charlie's and Alice's measurement projectors by $\Pi^{\mathrm{C}}_{\mathbf{c}|\mathbf{z}}$ and $\Pi^{\mathrm{A}}_{\mathbf{a}|\mathbf{x}}$, respectively. We then define the following unitary observables for Charlie:

$$\mathsf{O}_{i|\mathbf{z}} = \sum_{\mathbf{c}|c_i=+1} \Pi^{\mathrm{C}}_{\mathbf{c}|\mathbf{z}} - \sum_{\mathbf{c}|c_i=-1} \Pi^{\mathrm{C}}_{\mathbf{c}|\mathbf{z}}. \tag{23}$$

These operators can be understood as ± 1 valued observables that depend on the output c_i only for a particular choice of input $\mathbf{z}$, and are thus analogous to one of the three Pauli measurements (given by the value z_i) acting on the ith subspace of the maximally entangled state. Next we define the operators

$$\mathsf{Z}^{\mathrm{C}}_i = \frac{1}{3^{n-1}} \sum_{\mathbf{z}|z_i=1} \mathsf{O}_{i|\mathbf{z}}, \tag{24}$$

$$\mathsf{X}^{\mathrm{C}}_i = \frac{1}{3^{n-1}} \sum_{\mathbf{z}|z_i=2} \mathsf{O}_{i|\mathbf{z}}, \tag{25}$$

$$\mathsf{Y}^{\mathrm{C}}_i = \frac{1}{3^{n-1}} \sum_{\mathbf{z}|z_i=3} \mathsf{O}_{i|\mathbf{z}}, \tag{26}$$

that is, the average observables compatible with a particular choice of z_i.

Similarly for Alice we define the unitary observables

$$\mathsf{P}_{i|\mathbf{x}} = \sum_{\mathbf{a}|a_i=+1} \Pi^{\mathrm{A}}_{\mathbf{a}|\mathbf{x}} - \sum_{\mathbf{a}|a_i=-1} \Pi^{\mathrm{A}}_{\mathbf{a}|\mathbf{x}} \tag{27}$$

and the six operators

$$\begin{aligned}
\mathsf{D}^{\mathrm{A}}_{\mathrm{zx},i} &= \frac{1}{6^{n-1}} \sum_{\mathbf{x}|x_i=1} \mathsf{P}_{i|\mathbf{x}}, \quad \mathsf{E}^{\mathrm{A}}_{\mathrm{zx},i} = \frac{1}{6^{n-1}} \sum_{\mathbf{x}|x_i=2} \mathsf{P}_{i|\mathbf{x}}, \\
\mathsf{D}^{\mathrm{A}}_{\mathrm{zy},i} &= \frac{1}{6^{n-1}} \sum_{\mathbf{x}|x_i=3} \mathsf{P}_{i|\mathbf{x}}, \quad \mathsf{E}^{\mathrm{A}}_{\mathrm{zy},i} = \frac{1}{6^{n-1}} \sum_{\mathbf{x}|x_i=4} \mathsf{P}_{i|\mathbf{x}}, \\
\mathsf{D}^{\mathrm{A}}_{\mathrm{xy},i} &= \frac{1}{6^{n-1}} \sum_{\mathbf{x}|x_i=5} \mathsf{P}_{i|\mathbf{x}}, \quad \mathsf{E}^{\mathrm{A}}_{\mathrm{xy},i} = \frac{1}{6^{n-1}} \sum_{\mathbf{x}|x_i=6} \mathsf{P}_{i|\mathbf{x}}.
\end{aligned} \tag{28}$$

We now consider Bell operators of the form

$$\begin{aligned}
\mathcal{B}_i = {} & \mathsf{Z}^{\mathrm{C}}_i \big(\mathsf{D}^{\mathrm{A}}_{\mathrm{zx},i} + \mathsf{E}^{\mathrm{A}}_{\mathrm{zx},i}\big) + \mathsf{X}^{\mathrm{C}}_i \big(\mathsf{D}^{\mathrm{A}}_{\mathrm{zx},i} - \mathsf{E}^{\mathrm{A}}_{\mathrm{zx},i}\big) \\
& + \mathsf{Z}^{\mathrm{C}}_i \big(\mathsf{D}^{\mathrm{A}}_{\mathrm{zy},i} + \mathsf{E}^{\mathrm{A}}_{\mathrm{zy},i}\big) - \mathsf{Y}^{\mathrm{C}}_i \big(\mathsf{D}^{\mathrm{A}}_{\mathrm{zy},i} - \mathsf{E}^{\mathrm{A}}_{\mathrm{zy},i}\big) \\
& + \mathsf{X}^{\mathrm{C}}_i \big(\mathsf{D}^{\mathrm{A}}_{\mathrm{xy},i} + \mathsf{E}^{\mathrm{A}}_{\mathrm{xy},i}\big) - \mathsf{Y}^{\mathrm{C}}_i \big(\mathsf{D}^{\mathrm{A}}_{\mathrm{xy},i} - \mathsf{E}^{\mathrm{A}}_{\mathrm{xy},i}\big).
\end{aligned} \tag{29}$$

This is simply the Bell inequality (8), for the inputs z_i and x_i averaged over all compatible $\mathbf{z}$ and $\mathbf{x}$. One can thus obtain $\langle\psi|\mathcal{B}_i|\psi\rangle = 6\sqrt{2}$ for each i by taking n copies of the maximally entangled state of dimension two and adopting the previous measurement strategy (9) independently on each of the copies. From the observation of maximal violation for all i, a self-testing circuit (a parallel version of the circuit of Lemma 1) can be constructed, see Fig. 4 in Appendix C. We then have the following lemma.

Lemma 2. Let the state $|\psi\rangle \in \mathcal{H}_{\mathrm{C}} \otimes \mathcal{H}_{\mathrm{A}}$ and operators $\mathsf{Z}^{\mathrm{C}}_i$, $\mathsf{X}^{\mathrm{C}}_i$, $\mathsf{Y}^{\mathrm{C}}_i$, $\mathsf{D}^{\mathrm{A}}_{\mathrm{zx},i}$, $\mathsf{E}^{\mathrm{A}}_{\mathrm{zx},i}$, $\mathsf{D}^{\mathrm{A}}_{\mathrm{zy},i}$, $\mathsf{E}^{\mathrm{A}}_{\mathrm{zy},i}$, $\mathsf{D}^{\mathrm{A}}_{\mathrm{xy},i}$, $\mathsf{E}^{\mathrm{A}}_{\mathrm{xy},i}$ defined above satisfy

$$\langle\psi|\mathcal{B}_i|\psi\rangle = 6\sqrt{2}, \tag{30}$$

for every $i \in \{1, \ldots n\}$. Then there exists a local unitary U, local registers $|00\rangle \in \otimes_{i=1}^{n} [\mathcal{H}_{\mathrm{C}''_i} \otimes \mathcal{H}_{\mathrm{C}'_i}] \otimes [\mathcal{H}_{\mathrm{A}''_i} \otimes \mathcal{H}_{\mathrm{A}'_i}]$, and a normalized state $|\xi\rangle$ such that

$$\begin{aligned}
U[|\psi\rangle \otimes |00\rangle] &= |\xi\rangle \otimes \big[\otimes_{i=1}^{n} |\Phi^+\rangle^{\mathrm{C}'_i\mathrm{A}'_i}\big], \\
U\big[\mathsf{Z}^{\mathrm{C}}_j |\psi\rangle \otimes |00\rangle\big] &= |\xi\rangle \otimes \big[\sigma_{\mathrm{z}}^{\mathrm{C}'_j} \otimes_{i=1}^{n} |\Phi^+\rangle^{\mathrm{C}'_i\mathrm{A}'_i}\big], \\
U\big[\mathsf{X}^{\mathrm{C}}_j |\psi\rangle \otimes |00\rangle\big] &= |\xi\rangle \otimes \big[\sigma_{\mathrm{x}}^{\mathrm{C}'_j} \otimes_{i=1}^{n} |\Phi^+\rangle^{\mathrm{C}'_i\mathrm{A}'_i}\big], \\
U\big[\mathsf{Y}^{\mathrm{C}}_j |\psi\rangle \otimes |00\rangle\big] &= \sigma_{\mathrm{z}}^{\mathrm{C}''_j} |\xi\rangle \otimes \big[\sigma_{\mathrm{y}}^{\mathrm{C}'_j} \otimes_{i=1}^{n} |\Phi^+\rangle^{\mathrm{C}'_i\mathrm{A}'_i}\big],
\end{aligned}$$

for every $j \in \{1, 2, \ldots n\}$, where $|\xi\rangle$ takes the form

$$|\xi\rangle = \sum_{\bar{q}} |\xi_{\bar{q}}\rangle^{\mathrm{CA}} \otimes |\bar{q}\bar{q}\rangle^{\mathrm{C}''\mathrm{A}''} \tag{31}$$

and the sum is over all bit strings $\bar{q} = (0, 1)^n$.

The proof of the above Lemma can be found in Appendix C. Note that since the self-tested measurements are extremal then the above statement must hold not only for the operators Z_j, X_j, Y_j but for each of the observables $\mathsf{O}_{i|\mathbf{z}}$ appearing in their definition, which implies that the input z_i indeed corresponds to the desired Pauli measurement on the correct subspace. The measurement $\sigma_{\mathrm{z}}^{\mathrm{C}''_j}$ on the state $|\xi\rangle$ again plays the role of deciding whether the measurement $\sigma_{\mathrm{y}}^{\mathrm{C}'_j}$ or $-\sigma_{\mathrm{y}}^{\mathrm{C}'_j}$ is performed on the maximally entangled state. However, note that due to the form of $|\xi\rangle$, this is not guaranteed to be correlated with the other measurements of σ_y on different local subspaces. As a result, one cannot equate this freedom to a local transposition on *all* of Charlie's subsystems, as needed from definition II.3. In the following section we show how to overcome this problem by introducing additional measurement for Alice.

C. Aligning reference frames

As mentioned, Lemma 2 suffers from one drawback, namely that the y direction for each of Charlie's local subsystems need not be aligned. For example, if we take the case $n = 2$, Lemma 2 gives four possibilities for Charlie's effective measurements on the maximally entangled state given by $\{\sigma_{\mathrm{x}}, \pm\sigma_{\mathrm{y}}, \sigma_{\mathrm{z}}\} \otimes \{\sigma_{\mathrm{x}}, \pm\sigma_{\mathrm{y}}, \sigma_{\mathrm{z}}\}$. The probability that each of these strategies is used is unknown and could, for example, be $\frac{1}{4}$ for each. In this case, when the first subsystem measures σ_{y}, the second subsystem has equal probability to measure either σ_{y} or $-\sigma_{\mathrm{y}}$. This lack of alignment is an artifact from performing the protocol of Lemma 1 in parallel without trying to introduce any dependencies between the n individual self-tests. In the following we show that one can further restrict the the state $|\xi\rangle$ to be of the form

$$|\xi\rangle = |\xi_0\rangle \otimes |00\cdots0\rangle^{\mathrm{C}''\mathrm{A}''} + |\xi_1\rangle \otimes |11\cdots1\rangle^{\mathrm{C}''\mathrm{A}''} \tag{32}$$

by introducing additional Bell state measurements between subsystems of Alice. Since $|\xi\rangle$ now has only two terms, the flipping of the σ_{y} measurements is always correlated; either none of the measurements are flipped (each subsystem measures σ_{y}) or all the measurements are flipped (each subsystem measures $-\sigma_{\mathrm{y}}$). We note that an analogous result was independently obtained in Ref. [31] (see Lemma 8 therein) using a similar approach.

To illustrate the basic idea let us again consider the case $n = 2$, and assume we adopt the ideal measurement strategy

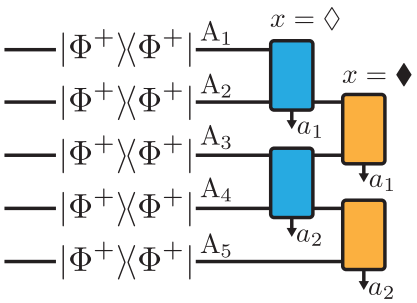

FIG. 2. Graphical representation of the additional measurements performed by Alice for $x = \diamondsuit$ and $x = \blacklozenge$. Boxes between subspaces represent Bell state measurements.

[i.e., the strategy (9) in parallel]. We now add an additional Bell state measurement for Alice which she performs on her two halves of the maximally entangled states. If Alice receives the outcome corresponding to the projector $|\Phi^+\rangle\langle\Phi^+|$, via entanglement swapping Charlie will hold the state $|\Phi^+\rangle$ in his local subsystem (for the other outcomes he will hold a different Bell state). This state has correlations $\langle\Phi^+|\sigma_x \otimes \sigma_x|\Phi^+\rangle = +1$, $\langle\Phi^+|\sigma_y \otimes \sigma_y|\Phi^+\rangle = -1$, $\langle\Phi^+|\sigma_z \otimes \sigma_z|\Phi^+\rangle = +1$. Hence, in order to reproduce these correlations, the direction of Charlie's two measurements of σ_y need to be correlated as otherwise we would not have perfect anticorrelation for the measurement $\sigma_y \otimes \sigma_y$. In the following we formalize this intuition to strengthen Lemma 2 so that $|\xi\rangle$ is of the form (32).

The precise scenario we consider is the following. In addition to the 6^n measurements of Lemma 2 given by the vector $\mathbf{x}$, Alice has two extra measurements denoted by $x = \diamondsuit$ and $x = \blacklozenge$. These measurements have respectively 4^m and $4^{m'}$ outcomes, where $m = \lfloor \frac{n}{2} \rfloor$ and $m' = \lfloor \frac{n-1}{2} \rfloor$, which are grouped into the vectors $\mathbf{a} = (a_1, a_2, \ldots, a_m)$ and $\mathbf{a} = (a_1, a_2, \ldots, a_{m'})$ with $a_i = 0, 1, 2, 3$. We denote by $\Pi_{\mathbf{a},\diamondsuit}$ and $\Pi_{\mathbf{a},\blacklozenge}$ the projectors corresponding to the outcomes of these measurements and define the projectors for $l = 1, \ldots, n$,

$$\mathsf{S}_{l,a^*} = \sum_{\mathbf{a}:a_l=a^*} \Pi_{\mathbf{a}|\diamondsuit}, \quad \mathsf{T}_{l,a^*} = \sum_{\mathbf{a}:a_l=a^*} \Pi_{\mathbf{a}|\blacklozenge}, \tag{33}$$

that is, the projectors onto the the subspace corresponding to $a_l = a^*$ for the two measurements.

To generate our self-testing correlations we use the same strategy as Lemma 2 for the inputs $\mathbf{x}$ and $\mathbf{z}$. The two new measurements for Alice $x = \diamondsuit, \blacklozenge$ correspond to Bell state measurements between successive pairs of qubits of her system, where the Bell state measurements for the input $\blacklozenge$ are shifted with respect to those for $\diamondsuit$ (see Fig. 2). Specifically,

$$\Pi_{\mathbf{a},\diamondsuit} = \bigotimes_{l=1}^{\lfloor \frac{n}{2} \rfloor} |\Psi_{a_i}\rangle\langle\Psi_{a_i}|^{\mathrm{A}_{2l-1}\mathrm{A}_{2l}}, \tag{34}$$

$$\Pi_{\mathbf{a},\blacklozenge} = \bigotimes_{l=1}^{\lfloor \frac{n-1}{2} \rfloor} |\Psi_{a_i}\rangle\langle\Psi_{a_i}|^{\mathrm{A}_{2l}\mathrm{A}_{2l+1}}, \tag{35}$$

where $\{|\Psi_0\rangle, |\Psi_1\rangle, |\Psi_2\rangle, |\Psi_3\rangle\} = \{|\Phi^+\rangle, |\Phi^-\rangle, |\Psi^+\rangle, |\Psi^-\rangle\}$. With this choice, the correlations are given by Table I, which follow from the correlations of the four Bell states. We are now ready for our final self-testing lemma (see Appendix D).

TABLE I. Elements of the table give correlation $\langle\psi|C \otimes R|\psi\rangle$ where C is the operator labeling the column and R the operator labeling the row.

	$\mathbb{1}$	$\mathsf{Z}_{2l-1}\mathsf{Z}_{2l}$	$\mathsf{X}_{2l-1}\mathsf{X}_{2l}$	$\mathsf{Y}_{2l-1}\mathsf{Y}_{2l}$
$\mathsf{S}_{l,0}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$-\frac{1}{4}$
$\mathsf{S}_{l,1}$	$\frac{1}{4}$	$\frac{1}{4}$	$-\frac{1}{4}$	$\frac{1}{4}$
$\mathsf{S}_{l,2}$	$\frac{1}{4}$	$-\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
$\mathsf{S}_{l,3}$	$\frac{1}{4}$	$-\frac{1}{4}$	$-\frac{1}{4}$	$-\frac{1}{4}$
	$\mathbb{1}$	$\mathsf{Z}_{2l}\mathsf{Z}_{2l+1}$	$\mathsf{X}_{2l}\mathsf{X}_{2l+1}$	$\mathsf{Y}_{2l}\mathsf{Y}_{2l+1}$
$\mathsf{T}_{l,0}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$-\frac{1}{4}$
$\mathsf{T}_{l,1}$	$\frac{1}{4}$	$\frac{1}{4}$	$-\frac{1}{4}$	$\frac{1}{4}$
$\mathsf{T}_{l,2}$	$\frac{1}{4}$	$-\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
$\mathsf{T}_{l,3}$	$\frac{1}{4}$	$-\frac{1}{4}$	$-\frac{1}{4}$	$-\frac{1}{4}$

Lemma 3. Let the state $|\psi\rangle \in \mathcal{H}_C \otimes \mathcal{H}_A$ and ± 1 outcome observables $\mathsf{X}^C, \mathsf{Y}^C, \mathsf{Z}^C, \mathsf{D}^A_{zx}, \mathsf{E}^A_{zx}, \mathsf{D}^A_{xy}, \mathsf{E}^A_{xy}, \mathsf{D}^A_{zy}, \mathsf{E}^A_{zy}$ satisfy the conditions of Lemma 2 so that $|\xi\rangle$ has the form (31). Furthermore, let projectors S_{l,a^*} and T_{l,a^*} satisfy the correlations given in Table I for all l. Then $|\xi\rangle$ has the form

$$|\xi\rangle = |\xi_0\rangle \otimes |0\cdots0\rangle + |\xi_1\rangle \otimes |1\cdots1\rangle. \tag{36}$$

Note that $|\xi\rangle$ now has the form of definition II.3 as desired.

D. Noise robustness

It is important to study the noise robustness of Lemmas 1–3 as it is impossible to achieve perfect self-testing correlations in practice. In the same way as related works [25,26,37,38], Lemma 1 and Lemma 2 can be made noise robust. In Appendix B we show how precise robustness bounds can be estimated for Lemma 1. For instance, if we have a nonmaximal value $\langle\psi|\mathcal{B}|\psi\rangle = 6\sqrt{2} - \epsilon$, Eq. (18) from Lemma 1 becomes

$$\|U(|\psi\rangle^{\mathrm{CA}} \otimes |00\rangle) - |\xi\rangle^{\mathrm{CC''AA''}} \otimes |\Phi^+\rangle^{\mathrm{C'A'}}\| \leqslant c\sqrt{\epsilon},$$

where $c = 55 + 36\sqrt{2}$. Similar statements can be derived for Eqs. (19)–(21). For robust statements of Lemma 2, we point the reader to [37] where the same techniques can be applied to our results to obtain polynomial robustness bounds; we do not elaborate further here since such calculations are based on well established methods and are not particularly enlightening. Concerning Lemma 3, we show that given a noise robust Lemma 2, one can extend this to a robust version of Lemma 3 (see Appendix E). These robustness statements will become relevant later in order to make the entanglement certification protocols of Sec. III tolerant to experimental noise.

III. DEVICE-INDEPENDENT ENTANGLEMENT CERTIFICATION

In this section we show how to make use of the preceding self-testing results to construct device-independent entanglement certification protocols for all bipartite entangled quantum states. The precise scenario that we consider is a quantum network featuring three bipartite states: ϱ^{AB} shared between Alice and Bob, and two auxiliary states ϱ^{CA_0} and $\varrho^{\mathrm{B}_0\mathrm{D}}$ shared between Charlie and Alice, and Bob and Daisy, respectively.

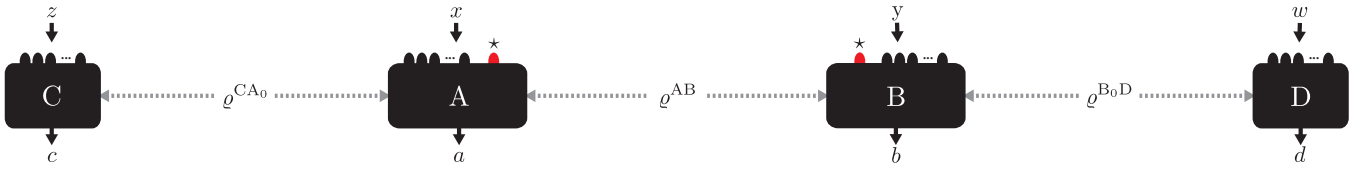

FIG. 3. Device-independent scenario for our entanglement certification protocol. The correlations $p(c, a, b, d|z, x, y, w)$ are checked for (i) maximal violation of a Bell inequality in each of the marginal distributions $p(c, a|z, x)$, $p(b, d|y, w)$ which via self-testing certifies that the states $\varrho^{\mathrm{CA_0}}$, $\varrho^{\mathrm{B_0D}}$ are maximally entangled and that the measurements of Charlie and Daisy are Pauli measurements, and (ii) violation of an additional inequality $\mathcal{I}[p(c, a, b, d|z, x = \star, y = \star, w)]$ where Alice and Bob perform the measurements $x = \star$, $y = \star$, which certifies the entanglement of ϱ^{AB} given (i) is satisfied.

Denoting the set of linear operators on Hilbert space $\mathcal{H}$ by $\mathcal{B}(\mathcal{H})$ we have $\varrho^{\mathrm{AB}} \in \mathcal{B}(\mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}})$, $\varrho^{\mathrm{CA_0}} \in \mathcal{B}(\mathcal{H}_{\mathrm{C}} \otimes \mathcal{H}_{\mathrm{A_0}})$, and $\varrho^{B_0 D} \in \mathcal{B}(\mathcal{H}_{\mathrm{B_0}} \otimes \mathcal{H}_{\mathrm{D}})$. We are interested in certifying the entanglement of the state ϱ^{AB} when placed in a line network (see Fig. 3) featuring the auxiliary states $\varrho^{\mathrm{CA_0}}$ and $\varrho^{\mathrm{B_0D}}$. In such a network, the correlations $\{p(c, a, b, d|z, x, y, w)\}$ are given by

$$p(c, a, b, d|z, x, y, w) = \mathrm{tr}\left[\mathsf{M}^{\mathrm{C}}_{c|z} \otimes \mathsf{M}^{\mathrm{A_0A}}_{a|x} \otimes \mathsf{M}^{\mathrm{BB_0}}_{b|y} \otimes \mathsf{M}^{\mathrm{D}}_{d|w}\, \varrho^{\mathrm{CA_0}} \otimes \varrho^{\mathrm{AB}} \otimes \varrho^{\mathrm{B_0D}}\right], \tag{37}$$

where the $\mathsf{M}_{i|j}$ are the local measurement operators for each party. In the device-independent scenario, one only has access to the observed correlations $p(c, a, b, d|z, x, y, w)$. Hence, a device-independent certification of the entanglement of ϱ^{AB} is possible only if the observed correlations cannot be reproduced by (37) for any separable ϱ^{AB}. That is, one must show

$$p(c, a, b, d|z, x, y, w) \neq \mathrm{tr}\left[\mathsf{M}'^{\mathrm{C}}_{c|z} \otimes \mathsf{M}'^{\mathrm{A_0A}}_{a|x} \otimes \mathsf{M}'^{\mathrm{BB_0}}_{b|y} \otimes \mathsf{M}'^{\mathrm{D}}_{d|w}\, \varrho'^{\mathrm{CA_0}} \otimes \varrho^{\mathrm{AB}}_{\mathrm{SEP}} \otimes \varrho'^{\mathrm{B_0D}}\right] \tag{38}$$

for any choice of separable $\varrho^{\mathrm{AB}}_{\mathrm{SEP}}$, and any local measurement operators $\mathsf{M}'_{i|j}$ and auxiliary states $\varrho'^{\mathrm{CA_0}}$ and $\varrho'^{\mathrm{B_0D}}$. Note that the auxiliary states may be entangled and that since we impose no constraints on the dimension of the auxiliary systems in Eq. (38), we may purify them and take all measurements to be projective without loss of generality.

As we work in the device-independent scenario, all devices are treated as black boxes that process classical information. The precise assumptions we then make about the experiment are as follows.

(1) States and measurements are described by quantum mechanics.

(2) The rounds of the experiment are independent and identically distributed (i.i.d.).

(3) The network of Fig. 3 correctly describes the experimental setup.

The first two of these assumptions are standard in device-independent studies (ideally one would like to drop the second assumption, see [11,39] for some recent progress). The last assumption is required so that we may write our probabilities in the form (37). Physically this assumption means that one is able to prepare the three states independently and that they are trusted to interact in the way described by the network of Fig. 3 (for example the state $\varrho^{\mathrm{CA_0}}$ should only interact with Charlie and Alice and not Bob or Daisy).

A. Certification protocols

We now present our entanglement certification protocols. These can be seen as a device-independent extension of the measurement-device-independent entanglement witnesses (MDIEWs) presented in previous works [16–18]. There, measurement devices are treated as black boxes, however inputs are given as a set of known informationally complete quantum states (in contrast to using classical variables as inputs). Then, an entanglement certification protocol can be built for every entangled state starting from an entanglement witness for the state. However, since this scheme requires a set of trusted input quantum states it is only partially device independent. To see how these protocols can be made fully device independent (i.e., how to remove the trust on the input states) consider that in the network of Fig. 3 the auxiliary states are given by maximally entangled states and that the complete set of projectors for Charlie's (Daisy's) measurements form an informationally complete set. This can in fact be certified device independently using the self-testing protocols of the first part of the paper (see Lemmas 1 and 3). With this, the states that Alice (Bob) receives in the Hilbert space $\mathcal{H}_{\mathrm{A_0}}$ ($\mathcal{H}_{\mathrm{B_0}}$) conditioned on the different inputs and outputs of Charlie (Daisy) also form an informationally complete set. By interpreting these states as the inputs in a MDIEW protocol, one is essentially in the MDIEW scenario and the same techniques can be applied. Here one has to be a bit careful due to the issue of transposition encountered in the self-testing sections, which we deal with in Appendix F.

We now formalize this intuition and move to the main result of this section.

Main result. The entanglement of all bipartite entangled states can be certified device independently in the network of Fig. 3.

In order to show this, we give an explicit family of entanglement certification protocols. The protocols we consider have the same structure for all states and are summarized as follows:

Entanglement certification protocol

(i) *Generation of correlations.* The parties perform local measurements on their subsystems to obtain the correlations $p(c, a, b, d|z, x, y, w)$.

(ii) The following is then verified:

Self-testing. The marginal distributions $p(c, a|z, x)$ and $p(b, d|y, w)$ maximally violate a Bell inequality that certifies that the auxiliary states each contain a maximally entangled

state and that Charlie and Daisy each perform Pauli measurements on their subsystems.

Entanglement certification. The correlations violate an additional inequality $\mathcal{I}(p(c,a,b,d|z,x,y,w) \geqslant 0$ that certifies ϱ^{AB} is entangled.

For now, we have the unrealistic requirement that we have a maximum violation of a Bell inequality in step (ii). This can be weakened to allow for some noise on the statistics, which we discuss in Sec. III D. We now describe in detail the above protocol, starting with the case of two-qubit states.

B. Entanglement certification of all two-qubit entangled states

We start by defining the scenario in which we work. Charlie and Daisy both have a choice of three measurements $z, w = 1, 2, 3$ and Alice and Bob both have a choice of seven inputs $x, y = 1, 2, 3, 4, 5, 6, \star$. All outputs are ± 1 valued.

(i) Generation of correlations. To generate the correlations in step (i) of the protocol, the parties chose $\varrho^{\mathrm{CA}_0} = \varrho^{\mathrm{B}_0\mathrm{D}} = |\Phi^+\rangle\langle\Phi^+|$. Measurements for inputs $z = 1, 2, 3$ and $x = 1, \ldots, 6$ for Charlie and Alice should be chosen so that the conditions of Lemma 1 are satisfied, i.e., given by the qubit observables

$$\sigma_{\mathrm{z}},\ \sigma_{\mathrm{x}},\ \sigma_{\mathrm{y}} \quad z = 1, 2, 3, \tag{39}$$

$$\frac{\sigma_{\mathrm{z}} \pm \sigma_{\mathrm{x}}}{\sqrt{2}},\ \frac{\sigma_{\mathrm{z}} \pm \sigma_{\mathrm{y}}}{\sqrt{2}},\ \frac{\sigma_{\mathrm{x}} \pm \sigma_{\mathrm{y}}}{\sqrt{2}} \quad x = 1, \ldots, 6, \tag{40}$$

acting on the $\mathcal{H}_{\mathrm{C}}$ and $\mathcal{H}_{\mathrm{A}_0}$ spaces, respectively. Measurements for Daisy and Bob are defined analogously. Lastly, the measurement operators for inputs $x = \star$, $y = \star$ are projections onto the maximally entangled state:

$$\mathsf{M}^{\mathrm{AA}_0}_{+|\star} = \mathsf{M}^{\mathrm{B}_0\mathrm{B}}_{+|\star} = |\Phi^+\rangle\langle\Phi^+|. \tag{41}$$

(ii) Self-testing—Our next step is to define the Bell inequality used in step (ii) of the protocol. Here we focus on Charlie and Alice; the Bell inequality used by Daisy and Bob is the same. The inequality we consider is constructed by combining three CHSH Bell inequalities [36]. Define the expectation value for inputs z, x as

$$E_{z,x} = \sum_{c,a=\pm 1} ca\ p(c,a|z,x). \tag{42}$$

We then define the triple CHSH Bell inequality

$$\begin{aligned}\mathcal{J} = {} & E_{1,1} + E_{1,2} + E_{2,1} - E_{2,2} \\ & + E_{1,3} + E_{1,4} - E_{3,3} + E_{3,4} \\ & + E_{2,5} + E_{2,6} - E_{3,5} + E_{3,6}.\end{aligned} \tag{43}$$

Note that each line in the above is a CHSH inequality, and each of Charlie's inputs appears in two of the lines, and that at this stage the inputs $x, y = \star$ remain unused. Using the states and measurements above one finds $\mathcal{J} = 6\sqrt{2}$. Via Lemma 1, this provides a self-test of the auxiliary states and measurements of Charlie and Daisy defined in step (i), up to local transposition.

Entanglement certification. Our next task is to construct the inequality used in the final step of the protocol. The inequality is constructed from an entanglement witness $\mathcal{W}$ for the state ϱ^{AB}. We thus have $\mathrm{tr}[\mathcal{W}\sigma] \geqslant 0$ for all separable states σ and $\mathrm{tr}[\mathcal{W}\varrho^{\mathrm{AB}}] < 0$. Consider the projectors $\pi_{c|z} = \frac{1}{2}[\mathbb{1} + c\,\sigma_z]$ with $c = \pm 1$ and $z = 1, 2, 3$, that is, projectors onto the plus and minus eigenspaces of the Pauli operators. Since these form an (overcomplete) basis of the set of Hermitian matrices, any entanglement witness may be decomposed as

$$\mathcal{W} = \sum_{cdzw} \omega^{zw}_{cd}\, \pi_{c|z} \otimes \pi_{d|w}. \tag{44}$$

To define our inequality, we make use of the additional inputs for both Alice and Bob $x = \star$ and $y = \star$. The inequality is then given by

$$\mathcal{I} = \sum_{cdzw} \omega^{zw}_{cd}\ p(c,+,+,d|z,x=\star,y=\star,w) \geqslant 0 \tag{45}$$

and is satisfied for all separable states but violated using ϱ^{AB}.

We first show that one can achieve $\mathcal{I} < 0$ for entangled ϱ^{AB}. Using the states and measurements defined above one has

$$\begin{aligned} & p(c,+,+,d|z,x=\star,y=\star,w) \\ & \quad = \mathrm{tr}[\pi_{c|z} \otimes |\Phi^+\rangle\langle\Phi^+| \otimes |\Phi^+\rangle\langle\Phi^+| \\ & \qquad \otimes \pi_{d|w}\, |\Phi^+\rangle\langle\Phi^+| \otimes \varrho^{\mathrm{AB}} \otimes |\Phi^+\rangle\langle\Phi^+|] \qquad (46) \\ & \quad = \frac{1}{4}\,\mathrm{tr}\left[|\Phi^+\rangle\langle\Phi^+| \otimes |\Phi^+\rangle\langle\Phi^+|\, \pi^T_{c|z} \otimes \varrho^{\mathrm{AB}} \otimes \pi^T_{d|w}\right] \qquad (47) \\ & \quad = \frac{1}{16}\,\mathrm{tr}[\pi_{c|z} \otimes \pi_{d|w}\, \varrho^{\mathrm{AB}}], \qquad (48) \end{aligned}$$

where we have used $\mathrm{tr}_{\mathrm{A}}[|\Phi^+\rangle\langle\Phi^+|\, \pi^{\mathrm{A}}_{i|j} \otimes \mathbb{1}] = \frac{1}{2}\pi^T_{i|j}$ in the third and fourth line. One thus has

$$\mathcal{I} = \frac{1}{16} \sum_{czdw} \omega^{zw}_{cd}\, \mathrm{tr}[\pi_{c|z} \otimes \pi_{d|w}\, \varrho^{\mathrm{AB}}], \tag{49}$$

$$\mathcal{I} = \frac{1}{16}\,\mathrm{tr}[\mathcal{W}\varrho^{\mathrm{AB}}] < 0, \tag{50}$$

which follows from the fact that $\mathcal{W}$ is an entanglement witness for the state.

We now consider the case in which ϱ_{AB} is separable. In general, if the self-testing part of the protocol is satisfied then one can show that

$$\mathcal{I} = \mathrm{tr}[\mathcal{W}\,\Lambda(\varrho^{\mathrm{AB}})], \tag{51}$$

where Λ is a local, positive map on separable quantum states (see Appendix F 1 for details). Hence $\Lambda(\varrho^{\mathrm{AB}})$ is a separable state and $\mathcal{I} \geqslant 0$. A crucial observation in the proof of the above is that although the measurements for Charlie and Daisy are only certified via self-testing up to a possible transposition, this uncertainty can be mapped to possible local transpositions on the state ϱ^{AB}. Since local transpositions map separable states to separable states, this ensures that a false-positive certification of entanglement does not occur.

C. Entanglement certification of high dimensional states

The previous protocol for two-qubit states can be applied in parallel to construct entanglement certification protocols for bipartite states of any dimension. In the following we construct protocols for states of local dimension 2^n where $n = 2, 3, \ldots$. Since a state of local dimension d can be seen as a particular case of a state of dimension 2^n for some $n \geqslant \log_2 d$ this implies a protocol for any dimension.

The scenario we consider is as follows. Charlie and Daisy each have 3^n inputs, given by the vectors $\mathbf{z} = (z_1, \ldots, z_n)$ and $\mathbf{w} = (w_1, \ldots, w_n)$ with $z_i, w_i = 1, 2, 3$, each with 2^n outcomes given by $\mathbf{c} = (c_1, \ldots, c_n)$ and $\mathbf{d} = (d_1, d_n)$ with $c_i, d_i = \pm 1$. Alice and Bob each have 6^n inputs given by the vectors $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{y} = (y_1, \ldots, y_n)$ with $x_i, y_i = 1, \ldots, 6$, with outcomes $\mathbf{a} = (a_1, \ldots, a_n)$, $\mathbf{b} = (b_1, \ldots, b_n)$ with $a_i, b_i = \pm 1$. Furthermore, Alice and Bob each have two additional inputs $x = \diamondsuit, \blacklozenge$ and $y = \diamondsuit, \blacklozenge$ with $4^{\lfloor \frac{n}{2} \rfloor}$ and $4^{\lfloor \frac{n-1}{2} \rfloor}$ outputs, respectively (as in Lemma 3), and inputs $x = \star$ and $y = \star$ with outputs $a = \pm 1$, $b = \pm 1$ [to be used in step (iii) of the protocol].

(i) Generation of correlations. Since we will perform the previous protocol in parallel, the Hilbert spaces of the auxiliary systems are written as the tensor product of n qubit spaces: $\mathcal{H}_{\mathrm{C}} = \otimes_i \mathcal{H}_{\mathrm{C}_i}$, $\mathcal{H}_{\mathrm{A}_0} = \otimes_i \mathcal{H}_{\mathrm{A}_{0i}}$ (and similarly for Daisy, Bob). The auxiliary states are then n-fold tensors of maximally entangled states on each two-qubit subspace:

$$\varrho^{\mathrm{CA}_0} = \otimes_{i=1}^n |\Phi^+\rangle\langle\Phi^+|^{\mathrm{C}_i\mathrm{A}_{0i}}, \quad \varrho^{\mathrm{B}_0\mathrm{D}} = \otimes_{i=1}^n |\Phi^+\rangle\langle\Phi^+|^{\mathrm{B}_{0i}\mathrm{D}_i}.$$

Measurements are a parallel version of the measurements (39) and (40), i.e., they are given by n-fold tensor products of the measurements (39) and (40), acting on each maximally entangled state. For example $z_i = 1, 2, 3$ corresponds to a measurement of $\sigma_{\mathrm{z}}, \sigma_{\mathrm{x}}, \sigma_{\mathrm{y}}$ on the ith subsystem of Charlie with outcome c_i. As before, the measurements $\mathsf{M}_{+|\star}$ are projections onto the maximally entangled state:

$$\mathsf{M}_{+|\star}^{\mathrm{AA}_0} = \mathsf{M}_{+|\star}^{\mathrm{B}_0\mathrm{B}} = |\Phi^+\rangle\langle\Phi^+|, \tag{52}$$

where here $|\Phi^+\rangle = \frac{1}{\sqrt{2^n}} \sum_i |ii\rangle \in \mathcal{H}_{\mathrm{C}} \otimes \mathcal{H}_{\mathrm{A}_0}$. Finally, the measurements for the inputs $x, y = \diamondsuit, \blacklozenge$ are chosen to be tensor products of Bell state measurements between successive pairs of qubits of the local subsystems of Alice and Bob, and where the Bell state measurements for the input $\diamondsuit$ are shifted with respect to those for $\blacklozenge$ (see Fig. 2 and Sec. II C for more details).

(ii) Self-testing. The Bell inequality is now a parallel version of (43) (again we just describe the inequality for Charlie and Alice). Define the average expectation value for the bits c_i, a_i given $z_i = z$, $x_i = x$ as

$$E_{z,x}^i = \frac{1}{3^{n-1}6^{n-1}} \sum_{\substack{\mathbf{z}|z_i=z \\ \mathbf{x}|x_i=x}} \sum_{\mathbf{c},\mathbf{a}} c_i a_i \; p(\mathbf{c}, \mathbf{a}|\mathbf{z}, \mathbf{x}). \tag{53}$$

For each i, we now have the triple CHSH Bell inequality:

$$\begin{aligned} \mathcal{J}_i = {} & E_{1,1}^i + E_{1,2}^i + E_{2,1}^i - E_{2,2}^i \\ & + E_{1,3}^i + E_{1,4}^i - E_{3,3}^i + E_{3,4}^i \\ & + E_{2,5}^i + E_{2,6}^i - E_{3,5}^i + E_{3,6}^i. \end{aligned} \tag{54}$$

For the entanglement certification protocol we require maximum violation of each of these inequalities, i.e.,

$$\sum_{i=1}^n \mathcal{J}_i = n6\sqrt{2}. \tag{55}$$

We further require that the measurements $x, y = \diamondsuit, \blacklozenge$ correctly reproduce the Bell state measurement correlations given in Table I, which is achieved by our chosen measurement strategy and detailed in Sec. II C. With these conditions met, we may apply Lemma 3 and move on to the entanglement certification of ϱ^{AB}.

(iii) Entanglement certification. Similarly to (44), we may decompose an entanglement witness for $\varrho^{\mathrm{AB}} \in \otimes_i [\mathcal{H}_{\mathrm{A}_i} \otimes \mathcal{H}_{\mathrm{B}_i}]$ using tensor products of Pauli projectors as an (overcomplete) basis:

$$\mathcal{W} = \sum_{\mathbf{c},\mathbf{d},\mathbf{z},\mathbf{w}} \omega_{\mathbf{cd}}^{\mathbf{zw}} \otimes_i \left[\pi_{c_i|z_i}^{\mathrm{A}_i} \otimes \pi_{d_i|w_i}^{\mathrm{B}_i}\right]. \tag{56}$$

The inequality that is used to certify entanglement is then

$$\mathcal{I} = \sum_{\mathbf{c},\mathbf{d},\mathbf{z},\mathbf{w}} \omega_{\mathbf{cd}}^{\mathbf{zw}} \, p(\mathbf{c}, +, +, \mathbf{d}|\mathbf{z}, x = \star, y = \star, \mathbf{w}) \geqslant 0, \tag{57}$$

which for separable states gives

$$\mathcal{I} = \mathrm{tr}[\mathcal{W}\, \Lambda(\varrho^{\mathrm{AB}})] \geqslant 0, \tag{58}$$

where Λ is again a local positive map on separable states (see Appendix F 2 for a full proof). Note here that simply using two-qubit strategy in parallel (i.e., using Lemma 2) without the additional Bell state measurements for inputs $x, y = \diamondsuit, \blacklozenge$ would lead to problems. This is because the measurements for Charlie and Daisy would be certified only up to possible flipping of any number of their n σ_{y} measurements. When mapping this uncertainty to the state ϱ^{AB}, this corresponds to possible local transposition on *part of a local subsystem* of ϱ^{AB}, which may map separable states to unphysical (non-positive) states. Hence, the additional Bell state measurements ensure that either none or all σ_{y} measurements are flipped, corresponding to a transposition of the entire local subsystem of ϱ^{AB} so that the map Λ is positive on separable states.

Finally, we show that $\mathcal{I}$ is violated by ϱ^{AB}. Using the measurement strategy above and that $\mathrm{tr}_{\mathrm{A}}[|\Phi^+\rangle\langle\Phi^+|\, \pi_{i|j}^{\mathrm{A}} \otimes \mathbb{1}] = \frac{1}{d}\pi_{i|j}^T$ for the maximally entangled state of dimension d, it is straightforward to show using the same technique as (46)–(48) that

$$\mathcal{I} = \frac{1}{d^4} \sum_{\mathbf{c},\mathbf{d},\mathbf{z},\mathbf{w}} \omega_{\mathbf{cd}}^{\mathbf{zw}} \, \mathrm{tr}\left[\otimes_i \left(\pi_{c_i|u_i}^{\mathrm{A}_i} \otimes \pi_{d_i|w_i}^{\mathrm{B}_i}\right) \varrho^{\mathrm{AB}}\right] \tag{59}$$

$$= \frac{1}{d^4} \mathrm{tr}[\mathcal{W}\varrho^{\mathrm{AB}}] < 0, \tag{60}$$

thus certifying the entanglement of ϱ^{AB}.

D. Noise robust entanglement certification

A natural question to ask is whether the above certification protocols can be extended to tolerate small amounts of experimental noise. Indeed, this can be achieved using robust versions of Lemmas 1 and 3. The intuitive argument goes as follows. Imagine each of our probabilities differ from the ideal self-testing statistics by some small amount ϵ. Then, the states that Alice and Bob receive from the auxiliary systems conditioned on Charlie's and Daisy's measurement outcomes should be close to eigenstates of products of Pauli operators. This implies that the analogous operator to $\mathcal{W}$ appearing in Eq. (60) is close to the desired witness, which can be used to bound the maximum value of $\mathcal{I}$ for separable states to be

$$\mathcal{I} \geqslant -c(\epsilon) \tag{61}$$

for some positive function $c(\epsilon)$ such that $c(0) = 0$. Unsurprisingly, this means that some weakly entangled states close to the separable set are no longer certified by the method. The amount of noise that can be tolerated by a typical state before it can no longer be certified depends on the optimality of the robustness bounds of the self-testing lemmas; given current techniques the noise tolerance is expected to be small. For a detailed proof and discussion of (61) see Appendix G. For a specific analysis for the class of two-qubit Werner states, see Appendix H.

IV. DISCUSSION AND CONCLUSION

We have shown that all bipartite entangled quantum states are capable of producing correlations that cannot be obtained using separable states by placing them in a larger network of auxiliary states and using tools from self-testing and measurement-device-independent entanglement witnesses. It is desirable to strengthen the self-testing part of our protocol; in particular, improved robustness bounds for self-testing would immediately translate into better noise tolerance of our protocols. One would most likely be able to achieve this using the protocols presented in Ref. [31] where self-testing statements for Pauli observables are presented with a robustness scaling that is independent of n. Furthermore, the choice of measurements used for self-testing could be made much more efficient. In general, one needs d^2 linearly independent projectors to form an informationally complete set, however for local dimension 2^n we make use of an overcomplete basis of 6^n projectors (coming from the tensor product of Pauli projectors), a difference that is exponential in n. Hence, a more efficient self-test of informationally complete sets of measurements would improve the efficiency of the protocol. Furthermore, given a particular state, one typically does not need the full set of projectors in order to write an entanglement witness for the state. It would therefore be interesting to study self-testing protocols that certify only those projectors that appear in a particular decomposition of an entanglement witness.

Although we have focused on the task of entanglement certification, our technique can in principle be applied to other convex sets of quantum states other than the separable set where linear witnesses can also be used. Due to the ambiguity of local unitaries and local transpositions in the self-testing part of our protocol, such sets would need to be closed under local unitary operations and local transpositions (as is the case for the separable set). For example, one could apply the same technique to certify entangled states with negative partial transpose. Finally, it would also be interesting to investigate the possibility of using our general technique for other device-independent tasks, for example using similar ideas to [40–42] to construct device-independent quantum key distribution protocols, or to generalize our protocol for the certification of genuine multipartite entanglement.

ACKNOWLEDGMENTS

The authors are thankful for useful discussions to Paul Skrzypczyk, Nicolas Brunner, Marco Túlio Quintino, Flavien Hirsch, Thomas Vidick, Matteo Lostaglio, Michał Oszmaniec, and Alexia Salavrakos. This work was supported by the Ramón y Cajal fellowship, Spanish MINECO (QIBEQI FIS2016-80773-P and Severo Ochoa SEV-2015-0522), the AXA Chair in Quantum Information Science, Generalitat de Catalunya (CERCA Programme), Fundació Privada Cellex, and ERC CoG QITBOX. COST project CA16218 NANOCOHYBRI and Juan de la Cierva-formación are also acknowledged.

APPENDIX A: PROOF OF LEMMA 1

In this Appendix we prove Lemma 1 from the main text. Define the following operators:

$$\begin{aligned} \mathsf{Z}^{\mathrm{A}}_{\mathrm{z,x}} &= \frac{\mathsf{D}^{\mathrm{A}}_{\mathrm{z,x}} + \mathsf{E}^{\mathrm{A}}_{\mathrm{z,x}}}{\sqrt{2}}, \quad \mathsf{X}^{\mathrm{A}}_{\mathrm{z,x}} = \frac{\mathsf{D}^{\mathrm{A}}_{\mathrm{z,x}} - \mathsf{E}^{\mathrm{A}}_{\mathrm{z,x}}}{\sqrt{2}}, \\ \mathsf{Z}^{\mathrm{A}}_{\mathrm{z,y}} &= \frac{\mathsf{D}^{\mathrm{A}}_{\mathrm{z,y}} + \mathsf{E}^{\mathrm{A}}_{\mathrm{z,y}}}{\sqrt{2}}, \quad \mathsf{Y}^{\mathrm{A}}_{\mathrm{z,y}} = \frac{\mathsf{D}^{\mathrm{A}}_{\mathrm{z,y}} - \mathsf{E}^{\mathrm{A}}_{\mathrm{z,y}}}{\sqrt{2}}, \\ \mathsf{X}^{\mathrm{A}}_{\mathrm{x,y}} &= \frac{\mathsf{D}^{\mathrm{A}}_{\mathrm{x,y}} + \mathsf{E}^{\mathrm{A}}_{\mathrm{x,y}}}{\sqrt{2}}, \quad \mathsf{Y}^{\mathrm{A}}_{\mathrm{x,y}} = \frac{\mathsf{D}^{\mathrm{A}}_{\mathrm{x,y}} - \mathsf{E}^{\mathrm{A}}_{\mathrm{x,y}}}{\sqrt{2}}. \end{aligned} \tag{A1}$$

From (13)–(15) we have

$$\begin{aligned} \mathsf{Z}^{\mathrm{A}}_{\mathrm{z,x}}|\psi\rangle &= \mathsf{Z}^{\mathrm{A}}_{\mathrm{z,y}}|\psi\rangle, \\ \mathsf{X}^{\mathrm{A}}_{\mathrm{z,x}}|\psi\rangle &= \mathsf{X}^{\mathrm{A}}_{\mathrm{x,y}}|\psi\rangle, \\ \mathsf{Y}^{\mathrm{A}}_{\mathrm{z,y}}|\psi\rangle &= \mathsf{Y}^{\mathrm{A}}_{\mathrm{x,y}}|\psi\rangle. \end{aligned} \tag{A2}$$

Hence, defining

$$\mathsf{Z}^{\mathrm{A}} \equiv \mathsf{Z}^{\mathrm{A}}_{\mathrm{z,x}}, \quad \mathsf{X}^{\mathrm{A}} \equiv \mathsf{X}^{\mathrm{A}}_{\mathrm{z,x}}, \quad \mathsf{Y}^{\mathrm{A}} \equiv \mathsf{Y}^{\mathrm{A}}_{\mathrm{z,y}} \tag{A3}$$

we have from (13)–(16) the conditions

$$\mathsf{Z}^{\mathrm{C}}|\psi\rangle = \mathsf{Z}^{\mathrm{A}}|\psi\rangle, \quad \mathsf{X}^{\mathrm{C}}|\psi\rangle = \mathsf{X}^{\mathrm{A}}|\psi\rangle, \quad \mathsf{Y}^{\mathrm{C}}|\psi\rangle = -\mathsf{Y}^{\mathrm{A}}|\psi\rangle, \tag{A4}$$

$$\{\mathsf{Z}^{\mathrm{C}}, \mathsf{X}^{\mathrm{C}}\}|\psi\rangle = 0, \quad \{\mathsf{Z}^{\mathrm{C}}, \mathsf{Y}^{\mathrm{C}}\}|\psi\rangle = 0, \quad \{\mathsf{Y}^{\mathrm{C}}, \mathsf{X}^{\mathrm{C}}\}|\psi\rangle = 0, \tag{A5}$$

$$\{\mathsf{Z}^{\mathrm{A}}, \mathsf{X}^{\mathrm{A}}\}|\psi\rangle = 0, \quad \{\mathsf{Z}^{\mathrm{A}}, \mathsf{Y}^{\mathrm{A}}\}|\psi\rangle = 0, \quad \{\mathsf{Y}^{\mathrm{A}}, \mathsf{X}^{\mathrm{A}}\}|\psi\rangle = 0. \tag{A6}$$

Note that the operators $\mathsf{Z}^{\mathrm{A}}, \mathsf{X}^{\mathrm{A}}, \mathsf{Y}^{\mathrm{A}}$ are not necessarily unitary. We may define the regularized versions of these operators $\hat{\mathsf{Z}}^{\mathrm{A}}, \hat{\mathsf{X}}^{\mathrm{A}}, \hat{\mathsf{Y}}^{\mathrm{A}}$ which are obtained from the original operators by renormalizing all eigenvalues to ± 1 and setting any zero eigenvalues to 1 (without changing the eigenvectors). Using standard techniques (for example see [21,43]) one can show that the regularized operators respect the same conditions, that is,

$$\mathsf{Z}^{\mathrm{C}}|\psi\rangle = \hat{\mathsf{Z}}^{\mathrm{A}}|\psi\rangle, \quad \mathsf{X}^{\mathrm{C}}|\psi\rangle = \hat{\mathsf{X}}^{\mathrm{A}}|\psi\rangle, \quad \mathsf{Y}^{\mathrm{C}}|\psi\rangle = -\hat{\mathsf{Y}}^{\mathrm{A}}|\psi\rangle, \tag{A7}$$

$$\{\mathsf{Z}^{\mathrm{C}}, \mathsf{X}^{\mathrm{C}}\}|\psi\rangle = 0, \quad \{\mathsf{Z}^{\mathrm{C}}, \mathsf{Y}^{\mathrm{C}}\}|\psi\rangle = 0, \quad \{\mathsf{Y}^{\mathrm{C}}, \mathsf{X}^{\mathrm{C}}\}|\psi\rangle = 0, \tag{A8}$$

$$\{\hat{\mathsf{Z}}^{\mathrm{A}}, \hat{\mathsf{X}}^{\mathrm{A}}\}|\psi\rangle = 0, \quad \{\hat{\mathsf{Z}}^{\mathrm{A}}, \hat{\mathsf{Y}}^{\mathrm{A}}\}|\psi\rangle = 0, \quad \{\hat{\mathsf{Y}}^{\mathrm{A}}, \hat{\mathsf{X}}^{\mathrm{A}}\}|\psi\rangle = 0. \tag{A9}$$

Let us prove the first equality from (A7), the other two being analogous. The following chain of equalities is satisfied

$$\begin{aligned}\|(\hat{\mathsf{Z}}^{\mathrm{A}}-\mathsf{Z}^{\mathrm{A}})|\psi\rangle\| &= \|(\mathbb{1}-(\hat{\mathsf{Z}}^{\dagger})^{\mathrm{A}}\mathsf{Z}^{\mathrm{A}})|\psi\rangle\| \\ &= \|(\mathbb{1}-|\mathsf{Z}^{\mathrm{A}}|)|\psi\rangle\| \qquad \text{(A10)} \\ &= \|(\mathbb{1}-|\mathsf{Z}^{\mathrm{C}}\mathsf{Z}^{\mathrm{A}}|)|\psi\rangle\| \\ &\leqslant \|(\mathbb{1}-\mathsf{Z}^{\mathrm{C}}\mathsf{Z}^{\mathrm{A}})|\psi\rangle\| = 0, \qquad \text{(A11)}\end{aligned}$$

where the first equality comes from the fact that $(\hat{\mathsf{Z}}^{\dagger})^{\mathrm{A}}$ is unitary, the second equality just uses the definition of $\hat{\mathsf{Z}}^{\mathrm{A}}$. The third equality is equivalent to $|\mathsf{Z}^{\mathrm{C}}\mathsf{Z}^{\mathrm{A}}| = |\mathsf{Z}^{\mathrm{A}}|$, which is correct because Z^{C} is unitary. The inequality is a consequence of $A \leqslant |A|$, and finally the last equality is the consequence of (A4).

We may now verify equations (18) to (22) of Lemma 1 using the above conditions. The precise isometry that we use is shown in Fig. 1. We first verify that the circuit acts correctly on the state $|\psi\rangle^{\mathrm{CA}}$. Up to and including the second set of controlled gates the circuit is the well known SWAP circuit, and it is well known (see, e.g., [28]) that this extracts the maximally entangled state in to the primed auxiliary systems. At this point our state is thus

$$|{+}{+}\rangle^{\mathrm{C''A''}}\frac{\mathbb{1}+\mathsf{Z}^{\mathrm{C}}}{\sqrt{2}}|\psi\rangle^{\mathrm{CA}}\otimes|\Phi^{+}\rangle^{\mathrm{C'A'}}. \qquad \text{(A12)}$$

Let us denote $|\phi\rangle^{\mathrm{CA}} = \frac{1}{\sqrt{2}}[\mathbb{1}+\mathsf{Z}^{\mathrm{C}}]|\psi\rangle^{\mathrm{CA}}$. The third pair of controlled gates evolves the system to

$$\tfrac{1}{2}[|00\rangle^{\mathrm{C''A''}}|\phi\rangle^{\mathrm{CA}}+|01\rangle^{\mathrm{C''A''}}i\hat{\mathsf{Y}}^{\mathrm{A}}\hat{\mathsf{X}}^{\mathrm{A}}|\phi\rangle^{\mathrm{CA}}+|10\rangle^{\mathrm{C''A''}}i\mathsf{Y}^{\mathrm{C}}\mathsf{X}^{\mathrm{C}}|\phi\rangle^{\mathrm{CA}}-|11\rangle^{\mathrm{C''A''}}\mathsf{Y}^{\mathrm{C}}\mathsf{X}^{\mathrm{C}}\hat{\mathsf{Y}}^{\mathrm{A}}\hat{\mathsf{X}}^{\mathrm{A}}|\phi\rangle^{\mathrm{CA}}]|\Phi^{+}\rangle^{\mathrm{C'A'}}.$$

From (A7)–(A9) it follows that $\hat{\mathsf{Y}}^{\mathrm{A}}\hat{\mathsf{X}}^{\mathrm{A}}|\phi\rangle^{\mathrm{CA}} = \mathsf{Y}^{\mathrm{C}}\mathsf{X}^{\mathrm{C}}|\phi\rangle^{\mathrm{CA}}$ and so

$$\tfrac{1}{2}[|00\rangle^{\mathrm{C''A''}}|\phi\rangle^{\mathrm{CA}}+|01\rangle^{\mathrm{C''A''}}i\mathsf{Y}^{\mathrm{C}}\mathsf{X}^{\mathrm{C}}|\phi\rangle^{\mathrm{CA}}+|10\rangle^{\mathrm{C''A''}}i\mathsf{Y}^{\mathrm{C}}\mathsf{X}^{\mathrm{C}}|\phi\rangle^{\mathrm{CA}}+|11\rangle^{\mathrm{C''A''}}|\phi\rangle^{\mathrm{CA}}]|\Phi^{+}\rangle^{\mathrm{C'A'}}. \qquad \text{(A13)}$$

Finally the last two Hadamards lead to

$$\frac{1}{2\sqrt{2}}[|00\rangle^{\mathrm{C''A''}}(\mathbb{1}+i\mathsf{Y}^{\mathrm{C}}\mathsf{X}^{\mathrm{C}})(\mathbb{1}+\mathsf{Z}^{\mathrm{C}})|\psi\rangle^{\mathrm{CA}}+|11\rangle^{\mathrm{C''A''}}(\mathbb{1}-i\mathsf{Y}^{\mathrm{C}}\mathsf{X}^{\mathrm{C}})(\mathbb{1}+\mathsf{Z}^{\mathrm{C}})|\psi\rangle^{\mathrm{CA}}]|\Phi^{+}\rangle^{\mathrm{C'A'}} = |\xi\rangle^{\mathrm{CC''AA''}}\otimes|\Phi^{+}\rangle^{\mathrm{C'A'}} \qquad \text{(A14)}$$

as claimed. Following the same method and using (A7)–(A9), one easily verifies

$$\begin{aligned}U(\mathsf{X}^{\mathrm{C}}|\psi\rangle^{\mathrm{CA}}\otimes|00\rangle) &= |\xi\rangle^{\mathrm{CC''AA''}}\otimes\sigma_{\mathrm{x}}^{\mathrm{C'}}|\Phi^{+}\rangle^{\mathrm{C'A'}}, \\ U(\mathsf{Z}^{\mathrm{C}}|\psi\rangle^{\mathrm{CA}}\otimes|00\rangle) &= |\xi\rangle^{\mathrm{CC''AA''}}\otimes\sigma_{\mathrm{z}}^{\mathrm{C'}}|\Phi^{+}\rangle^{\mathrm{C'A'}}. \qquad \text{(A15)}\end{aligned}$$

The case $\mathsf{Y}^{\mathrm{C}}|\psi\rangle^{\mathrm{CA}}\otimes|00\rangle$ is a bit more involved. After the second pair of controlled gates the state is transformed to

$$|{+}{+}\rangle^{\mathrm{C''A''}}\frac{1}{\sqrt{2}}i\mathsf{Y}^{\mathrm{C}}\mathsf{X}^{\mathrm{C}}(\mathbb{1}+\mathsf{Z}^{\mathrm{C}})|\psi\rangle^{\mathrm{CA}}\sigma_{\mathrm{y}}^{\mathrm{C'}}|\Phi^{+}\rangle^{\mathrm{C'A'}}.$$

The third pair of controlled gates then transforms the state to

$$\frac{1}{4\sqrt{2}}[|00\rangle^{\mathrm{C''A''}}i\mathsf{Y}^{\mathrm{C}}\mathsf{X}^{\mathrm{C}}|\phi\rangle^{\mathrm{CA}}+|01\rangle^{\mathrm{C''A''}}|\phi\rangle^{\mathrm{CA}}+|10\rangle^{\mathrm{C''A''}}|\phi\rangle^{\mathrm{CA}}+|11\rangle^{\mathrm{C''A''}}i\mathsf{Y}^{\mathrm{C}}\mathsf{X}^{\mathrm{C}}|\phi\rangle^{\mathrm{CA}}]\sigma_{\mathrm{y}}^{\mathrm{C'}}|\Phi^{+}\rangle^{\mathrm{C'A'}},$$

which is simplified by two last Hadamards to

$$\frac{1}{2\sqrt{2}}[|00\rangle^{\mathrm{C''A''}}(\mathbb{1}+i\mathsf{Y}^{\mathrm{C}}\mathsf{X}^{\mathrm{C}})(\mathbb{1}+\mathsf{Z}^{\mathrm{C}})|\psi\rangle^{\mathrm{CA}}-|11\rangle^{\mathrm{C''A''}}(\mathbb{1}-i\mathsf{Y}^{\mathrm{C}}\mathsf{X}^{\mathrm{C}})(\mathbb{1}+\mathsf{Z}^{\mathrm{C}})|\psi\rangle^{\mathrm{CA}}]\sigma_{\mathrm{y}}^{\mathrm{C'}}|\Phi^{+}\rangle^{\mathrm{C'A'}} = \sigma_{\mathrm{z}}^{\mathrm{C''}}|\xi\rangle^{\mathrm{CC''AA''}}\otimes\sigma_{\mathrm{y}}^{\mathrm{C'}}|\Phi^{+}\rangle^{\mathrm{C'A'}} \qquad \text{(A16)}$$

This thus concludes the proof of Lemma 1.

APPENDIX B: ROBUST VERSION OF LEMMA 1

Following the approach from [20,23] we study how Lemma 1 is affected when the achieved Bell inequality (30) violation is $6\sqrt{2}-\epsilon$. Looking at SOS decomposition (11) one can see that each of the terms must be smaller or equal to $\sqrt{\epsilon}$, leading to

$$\begin{aligned}\|(\mathsf{Z}^{\mathrm{C}}-\mathsf{Z}^{\mathrm{A}})|\psi\rangle\| &\leqslant \sqrt{\epsilon}, \\ \|(\mathsf{X}^{\mathrm{C}}-\mathsf{X}^{\mathrm{A}})|\psi\rangle\| &\leqslant \sqrt{\epsilon}, \qquad \text{(B1)} \\ \|(\mathsf{Y}^{\mathrm{C}}+\mathsf{Y}^{\mathrm{A}})|\psi\rangle\| &\leqslant \sqrt{\epsilon},\end{aligned}$$

$$\begin{aligned}\|(\mathsf{Z}^{\mathrm{C}}-\hat{\mathsf{Z}}^{\mathrm{A}})|\psi\rangle\| &\leqslant 2\sqrt{\epsilon}, \\ \|(\mathsf{X}^{\mathrm{C}}-\hat{\mathsf{X}}^{\mathrm{A}})|\psi\rangle\| &\leqslant 2\sqrt{\epsilon}, \qquad \text{(B2)} \\ \|(\mathsf{Y}^{\mathrm{C}}+\hat{\mathsf{Y}}^{\mathrm{A}})|\psi\rangle\| &\leqslant 2\sqrt{\epsilon},\end{aligned}$$

$$\begin{aligned}\|\{\mathsf{Z}^{\mathrm{C}},\mathsf{X}^{\mathrm{C}}\}|\psi\rangle\| &\leqslant (4+4\sqrt{2})\sqrt{\epsilon}, \\ \|\{\mathsf{Z}^{\mathrm{C}},\mathsf{Y}^{\mathrm{C}}\}|\psi\rangle\| &\leqslant (6+6\sqrt{2})\sqrt{\epsilon}, \qquad \text{(B3)} \\ \|\{\mathsf{Y}^{\mathrm{C}},\mathsf{X}^{\mathrm{C}}\}|\psi\rangle\| &\leqslant (8+8\sqrt{2})\sqrt{\epsilon}.\end{aligned}$$

Let us first note that the error coming from the regularizing operators on Alice's side is

$$\|(\hat{\mathsf{Z}}^{\mathrm{A}}-\mathsf{Z}^{\mathrm{A}})|\psi\rangle\| = \|(\mathbb{1}-(\hat{\mathsf{Z}}^{\dagger})^{\mathrm{A}}\mathsf{Z}^{\mathrm{A}})|\psi\rangle\| = \|(\mathbb{1}-|\mathsf{Z}^{\mathrm{A}}|)|\psi\rangle\| = \|(\mathbb{1}-|\mathsf{Z}^{\mathrm{C}}\mathsf{Z}^{\mathrm{A}}|)|\psi\rangle\| \leqslant \|(\mathbb{1}-\mathsf{Z}^{\mathrm{C}}\mathsf{Z}^{\mathrm{A}})|\psi\rangle\| = \sqrt{\epsilon},$$

and similarly for $\hat{\mathsf{X}}^{\mathrm{A}}$ and $\hat{\mathsf{Y}}^{\mathrm{A}}$. Taking this into account inequalities in the second line follow from the corresponding inequalities in the first line and the triangle inequality $\|a+b\| \leqslant \|a\| + \|b\|$. The first inequality in the third line is obtained through the following chain of inequalities:

$$\begin{aligned}
&\|(\mathsf{Z}^{\mathrm{C}}\mathsf{X}^{\mathrm{C}}+\mathsf{X}^{\mathrm{C}}\mathsf{Z}^{\mathrm{C}})|\psi\rangle\| \\
&\leqslant \|\mathsf{Z}^{\mathrm{C}}(\mathsf{X}^{\mathrm{C}}-\mathsf{X}^{\mathrm{A}})|\psi\rangle\| + \|(\mathsf{Z}^{\mathrm{C}}\mathsf{X}^{\mathrm{A}}+\mathsf{X}^{\mathrm{C}}\mathsf{Z}^{\mathrm{A}})|\psi\rangle\| + \|\mathsf{X}^{\mathrm{C}}(\mathsf{Z}^{\mathrm{A}}-\mathsf{Z}^{\mathrm{C}})|\psi\rangle\| \\
&\leqslant \sqrt{\epsilon} + \|\mathsf{X}^{\mathrm{A}}(\mathsf{Z}^{\mathrm{C}}-\mathsf{Z}^{\mathrm{A}})|\psi\rangle\| + \|(\mathsf{Z}^{\mathrm{A}}\mathsf{X}^{\mathrm{A}}+\mathsf{X}^{\mathrm{A}}\mathsf{Z}^{\mathrm{A}})|\psi\rangle\| + \|\mathsf{Z}^{\mathrm{A}}(\mathsf{X}^{\mathrm{C}}-\mathsf{X}^{\mathrm{A}})|\psi\rangle\| + \sqrt{\epsilon} \\
&\leqslant 2\sqrt{\epsilon} + \frac{1}{\sqrt{2}}\left\|\left(\mathsf{D}^{\mathrm{A}}_{\mathrm{z,x}}-\mathsf{E}^{\mathrm{A}}_{\mathrm{z,x}}\right)(\mathsf{Z}^{\mathrm{C}}-\mathsf{Z}^{\mathrm{A}})|\psi\rangle\right\| + \|(\mathsf{Z}^{\mathrm{A}}\mathsf{X}^{\mathrm{A}}+\mathsf{X}^{\mathrm{A}}\mathsf{Z}^{\mathrm{A}})|\psi\rangle\| + \frac{1}{\sqrt{2}}\left\|\left(\mathsf{D}^{\mathrm{A}}_{\mathrm{z,x}}+\mathsf{E}^{\mathrm{A}}_{\mathrm{z,x}}\right)(\mathsf{X}^{\mathrm{C}}-\mathsf{X}^{\mathrm{A}})|\psi\rangle\right\| \\
&\leqslant (2+2\sqrt{2})\sqrt{\epsilon} + \|\mathsf{Z}^{\mathrm{A}}(\mathsf{X}^{\mathrm{A}}-\hat{\mathsf{X}}^{\mathrm{A}})|\psi\rangle\| + \|\hat{X}^{\mathrm{A}}(\mathsf{Z}^{\mathrm{A}}-\hat{\mathsf{Z}}^{\mathrm{A}})|\psi\rangle\| + \|(\hat{\mathsf{Z}}^{\mathrm{A}}\hat{\mathsf{X}}^{\mathrm{A}}+\hat{\mathsf{X}}^{\mathrm{A}}\hat{\mathsf{Z}}^{\mathrm{A}})|\psi\rangle\| \\
&\quad + \|\mathsf{X}^{\mathrm{A}}(\mathsf{Z}^{\mathrm{A}}-\hat{\mathsf{Z}}^{\mathrm{A}})|\psi\rangle\| + \|\hat{\mathsf{Z}}^{\mathrm{A}}(\mathsf{X}^{\mathrm{A}}-\hat{\mathsf{X}}^{\mathrm{A}})|\psi\rangle\| \\
&\leqslant (4+4\sqrt{2})\sqrt{\epsilon}.
\end{aligned}$$

Note that if the violation of Bell inequality is $6\sqrt{2}-\epsilon$ not all terms from the first line of (B1) can simultaneously be equal to $\sqrt{\epsilon}$, but for our purposes a tight multiplicative factor is not of primary interest. The second and the third inequality from the third line of (B1) are derived in an analogous manner as the first one, with the additional factors coming from the convention used in Eq. (A3) which leads to

$$\begin{aligned}
\left\|\left(\mathsf{Z}^{\mathrm{A}}-\mathsf{Z}^{\mathrm{A}}_{\mathrm{z,y}}\right)|\psi\rangle\right\| &\leqslant \sqrt{\epsilon}, \\
\left\|\left(\mathsf{X}^{\mathrm{A}}-\mathsf{X}^{\mathrm{A}}_{\mathrm{x,y}}\right)|\psi\rangle\right\| &\leqslant \sqrt{\epsilon}, \\
\left\|\left(\mathsf{Y}^{\mathrm{A}}-\mathsf{Y}^{\mathrm{A}}_{\mathrm{x,y}}\right)|\psi\rangle\right\| &\leqslant \sqrt{\epsilon}.
\end{aligned}$$

To check the error accumulated when obtaining the final statement from Lemma 1 we will repeatedly use the triangle inequality and bounds from (B1). To get (A12) the first inequality from the second line of (B1) has to be used four times, the second one is used twice, and the anticommuting bound from the third line of (B1) has to be used once. To obtain (A13) the second and the third inequality from the second line and all three inequalities from the third line of (B1) are each used twice. All together these bounds imply

$$\begin{aligned}
&\|U(|\psi\rangle^{\mathrm{CA}}\otimes|00\rangle) - |\xi\rangle^{\mathrm{CC''AA''}}\otimes|\Phi^{+}\rangle^{\mathrm{C'A'}}\| \\
&\quad\leqslant (55+36\sqrt{2})\sqrt{\epsilon}.
\end{aligned}$$

A similar asymptotic bounds can be obtained for the robust versions of Eqs. (19), (21), and (20), the only difference being in the number of times each of the inequalities from (B1) have to be used.

APPENDIX C: PROOF OF LEMMA 2

The proof of Lemma 2 is split into two parts. The first part proves the necessary self-testing relations between the state and measurements needed to construct the self-testing circuit. The second part verifies that the circuit acts as claimed.

1. Self-testing relations

Here we follow closely the proof of [37], adapting it the allow for additional σ_{y} measurements. We first define the following sets of operators:

$$\begin{aligned}
\left\{\mathsf{Z}_i^{(k)}\right\}_k &= \{\mathsf{O}_{i|\mathbf{z}}|z_i=1\}, \\
\left\{\mathsf{X}_i^{(k)}\right\}_k &= \{\mathsf{O}_{i|\mathbf{z}}|z_i=2\}, \\
\left\{\mathsf{Y}_i^{(k)}\right\}_k &= \{\mathsf{O}_{i|\mathbf{z}}|z_i=3\},
\end{aligned} \tag{C1}$$

for $k=1,\ldots,3^{n-1}$ and ordered according to some relation $\mathbf{z}<\mathbf{z}'$. Similarly we define

$$\begin{aligned}
\left\{\mathsf{D}^{(l)}_{\mathrm{zx},i}\right\}_l &= \{\mathsf{P}_{i|\mathbf{x}}|x_i=1\}, \\
\left\{\mathsf{E}^{(l)}_{\mathrm{zx},i}\right\}_l &= \{\mathsf{P}_{i|\mathbf{x}}|x_i=2\}, \\
\left\{\mathsf{D}^{(l)}_{\mathrm{zy},i}\right\}_l &= \{\mathsf{P}_{i|\mathbf{x}}|x_i=3\},
\end{aligned} \tag{C2}$$

$$\begin{aligned}
\left\{\mathsf{E}^{(l)}_{\mathrm{zy},i}\right\}_l &= \{\mathsf{P}_{i|\mathbf{x}}|x_i=4\}, \\
\left\{\mathsf{D}^{(l)}_{\mathrm{xy},i}\right\}_l &= \{\mathsf{P}_{i|\mathbf{x}}|x_i=5\}, \\
\left\{\mathsf{E}^{(l)}_{\mathrm{xy},i}\right\}_l &= \{\mathsf{P}_{i|\mathbf{x}}|x_i=6\}.
\end{aligned} \tag{C3}$$

for $l=1,\ldots,6^{n-1}$ ordered according to some relation $\mathbf{x}<\mathbf{x}'$. Averaging over these sets we thus obtain the operators in Eqs. (24)–(28). We may now write

$$\begin{aligned}
\langle\psi|\mathcal{B}_i|\psi\rangle = \frac{1}{3^{n-1}6^{n-1}}\sum_{k,l}\langle\psi|\big[&\mathsf{Z}_i^{(k)}\left(\mathsf{D}^{(l)}_{\mathrm{zx},i}+\mathsf{E}^{(l)}_{\mathrm{zx},i}\right) \\
&+\mathsf{X}_i^{(k)}\left(\mathsf{D}^{(l)}_{\mathrm{zx},i}-\mathsf{E}^{(l)}_{\mathrm{zx},i}\right)+\mathsf{Z}_i^{(k)}\left(\mathsf{D}^{(l)}_{\mathrm{zy},i}+\mathsf{E}^{(l)}_{\mathrm{zy},i}\right) \\
&-\mathsf{Y}_i^{(k)}\left(\mathsf{D}^{(l)}_{\mathrm{zy},i}-\mathsf{E}^{(l)}_{\mathrm{zy},i}\right)+\mathsf{X}_i^{(k)}\left(\mathsf{D}^{(l)}_{\mathrm{xy},i}+\mathsf{E}^{(l)}_{\mathrm{xy},i}\right) \\
&-\mathsf{Y}_i^{(k)}\left(\mathsf{D}^{(l)}_{\mathrm{xy},i}-\mathsf{E}^{(l)}_{\mathrm{xy},i}\right)\big]|\psi\rangle = 6\sqrt{2}
\end{aligned} \tag{C4}$$

for all $i=1,\ldots,n$. Note that since the maximum value of the triple CHSH inequality is $6\sqrt{2}$ and that the above is a convex mixture of triple CHSH inequalities for different k,l, for each

k, l we have

$$\begin{aligned}\langle\psi|\big[&\mathsf{Z}_i^{(k)}\big(\mathsf{D}_{\mathrm{zx},i}^{(l)}+\mathsf{E}_{\mathrm{zx},i}^{(l)}\big)+\mathsf{X}_i^{(k)}\big(\mathsf{D}_{\mathrm{zx},i}^{(l)}-\mathsf{E}_{\mathrm{zx},i}^{(l)}\big)\\&+\mathsf{Z}_i^{(k)}\big(\mathsf{D}_{\mathrm{zy},i}^{(l)}+\mathsf{E}_{\mathrm{zy},i}^{(l)}\big)\\&-\mathsf{Y}_i^{(k)}\big(\mathsf{D}_{\mathrm{zy},i}^{(l)}-\mathsf{E}_{\mathrm{zy},i}^{(l)}\big)+\mathsf{X}_i^{(k)}\big(\mathsf{D}_{\mathrm{xy},i}^{(l)}+\mathsf{E}_{\mathrm{xy},i}^{(l)}\big)\\&-\mathsf{Y}_i^{(k)}\big(\mathsf{D}_{\mathrm{xy},i}^{(l)}-\mathsf{E}_{\mathrm{xy},i}^{(l)}\big)\big]|\psi\rangle=6\sqrt{2}. \qquad \text{(C5)}\end{aligned}$$

Now, we may again use the SOS decomposition (11) for each i, k, l leading to

$$\mathsf{Z}_i^{(k)}|\psi\rangle=\frac{\mathsf{D}_{\mathrm{zx},i}^{(l)}+\mathsf{E}_{\mathrm{zx},i}^{(l)}}{\sqrt{2}}|\psi\rangle=\frac{\mathsf{D}_{\mathrm{zy},i}^{(l)}+\mathsf{E}_{\mathrm{zy},i}^{(l)}}{\sqrt{2}}|\psi\rangle, \qquad \text{(C6)}$$

$$\mathsf{X}_i^{(k)}|\psi\rangle=\frac{\mathsf{D}_{\mathrm{zx},i}^{(l)}-\mathsf{E}_{\mathrm{zx},i}^{(l)}}{\sqrt{2}}|\psi\rangle=\frac{\mathsf{D}_{\mathrm{xy},i}^{(l)}+\mathsf{E}_{\mathrm{xy},i}^{(l)}}{\sqrt{2}}|\psi\rangle, \qquad \text{(C7)}$$

$$\mathsf{Y}_i^{(k)}|\psi\rangle=\frac{\mathsf{D}_{\mathrm{zy},i}^{(l)}-\mathsf{E}_{\mathrm{zy},i}^{(l)}}{\sqrt{2}}|\psi\rangle=\frac{\mathsf{D}_{\mathrm{xy},i}^{(l)}-\mathsf{E}_{\mathrm{xy},i}^{(l)}}{\sqrt{2}}|\psi\rangle, \qquad \text{(C8)}$$

which we may write as

$$\begin{aligned}\mathsf{Z}_i^{(k)}|\psi\rangle&=\mathsf{Z}_{i+n}^{(l)}|\psi\rangle,\\ \mathsf{X}_i^{(k)}|\psi\rangle&=\mathsf{X}_{i+n}^{(l)}|\psi\rangle,\\ \mathsf{Y}_i^{(k)}|\psi\rangle&=\mathsf{Y}_{i+n}^{(l)}|\psi\rangle,\end{aligned} \qquad \text{(C9)}$$

where

$$\begin{aligned}\mathsf{Z}_{i+n}^{(l)}&=\frac{\mathsf{D}_{\mathrm{zx},i}^{(l)}+\mathsf{E}_{\mathrm{zx},i}^{(l)}}{\sqrt{2}},\\ \mathsf{X}_{i+n}^{(l)}&=\frac{\mathsf{D}_{\mathrm{zx},i}^{(l)}-\mathsf{E}_{\mathrm{zx},i}^{(l)}}{\sqrt{2}},\\ \mathsf{Y}_i^{(k)}|\psi\rangle&=\frac{\mathsf{D}_{\mathrm{zy},i}^{(l)}-\mathsf{E}_{\mathrm{zy},i}^{(l)}}{\sqrt{2}}.\end{aligned} \qquad \text{(C10)}$$

As before, Eqs. (C6)–(C8) imply mutual anticommuntation of Alice's operators:

$$\big\{\mathsf{Z}_i^{(k)},\mathsf{X}_i^{(k)}\big\}=0,\quad \big\{\mathsf{Z}_i^{(k)},\mathsf{Y}_i^{(k)}\big\}=0,\quad \big\{\mathsf{X}_i^{(k)},\mathsf{Y}_i^{(k)}\big\}=0\quad \forall i,k. \qquad \text{(C11)}$$

Defining

$$\begin{aligned}\mathsf{Z}_{i+n}&=\frac{1}{6^{n-1}}\sum_l \mathsf{Z}_{i+n}^{(l)},\\ \mathsf{X}_{i+n}&=\frac{1}{6^{n-1}}\sum_l \mathsf{X}_{i+n}^{(l)},\\ \mathsf{Y}_{i+n}&=-\frac{1}{6^{n-1}}\sum_l \mathsf{Y}_{i+n}^{(l)},\end{aligned} \qquad \text{(C12)}$$

we have from (C9)

$$\begin{aligned}\mathsf{Z}_i^{(k)}|\psi\rangle&=\mathsf{Z}_{i+n}|\psi\rangle,\\ \mathsf{X}_i^{(k)}|\psi\rangle&=\mathsf{X}_{i+n}|\psi\rangle,\\ \mathsf{Y}_i^{(k)}|\psi\rangle&=-\mathsf{Y}_{i+n}|\psi\rangle\end{aligned} \qquad \text{(C13)}$$

for all k. Note that the operators $\mathsf{Z}_{i+n}, \mathsf{X}_{i+n}, \mathsf{Y}_{i+n}$ are not necessarily unitary. We therefore define the regularized versions

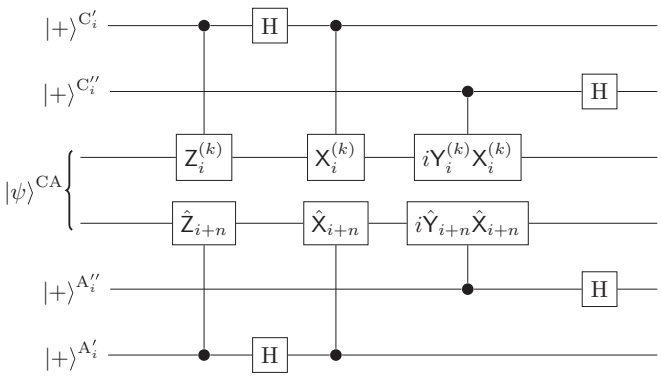

FIG. 4. Circuit diagram representing the local unitary of Lemma 2. The total unitary consists of applying this circuit for each $i = 1, \ldots, n$, and k can be chosen to be any number $k = 1, \ldots, 3^{n-1}$ (for example $k = 1$).

of these operators, denoted by $\hat{\mathsf{Z}}_{i+n}$, $\hat{\mathsf{X}}_{i+n}$, and $\hat{\mathsf{Y}}_{i+n}$, which using standard techniques (see for example [21,43]) can be shown to have the same properties:

$$\begin{aligned}\mathsf{Z}_i^{(k)}|\psi\rangle&=\hat{\mathsf{Z}}_{i+n}|\psi\rangle,\\ \mathsf{X}_i^{(k)}|\psi\rangle&=\hat{\mathsf{X}}_{i+n}|\psi\rangle,\\ \mathsf{Y}_i^{(k)}|\psi\rangle&=-\hat{\mathsf{Y}}_{i+n}|\psi\rangle.\end{aligned} \qquad \text{(C14)}$$

At this point we are nearly ready to construct our self-testing unitary. However, we still need to prove that $P_i^{(k)}$ and $P_j^{(k)}$ for $P \in \{\mathsf{X}, \mathsf{Y}, \mathsf{Z}\}$ commute for $i \neq j$. Here we again use the method of [37] to achieve this, which we restate here. Note that for every $i \neq j$, if we fix $z_i = 1$ and $z_j = 1$, there are 3^{n-2} choices for Charlie's measurement vector $\mathbf{z}$. There are thus 3^{n-2} pairs of indices (k, k') such that operators $\mathsf{Z}_i^{(k)}$ and $\mathsf{Z}_i^{(k')}$ are built from the same set of orthogonal projectors that commute by construction. We thus have 3^{n-2} equations of the form

$$\mathsf{Z}_i^{(k)}\mathsf{Z}_j^{(k')}|\psi\rangle=\mathsf{Z}_j^{(k')}\mathsf{Z}_i^{(k)}|\psi\rangle. \qquad \text{(C15)}$$

Choosing a pair (k, k') and using (C13) and the fact that operators on Charlie and Alice's subsystems commute we then obtain

$$\mathsf{Z}_i^{(k)}\mathsf{Z}_{n+j}|\psi\rangle=\mathsf{Z}_j^{(k')}\mathsf{Z}_{n+i}|\psi\rangle, \qquad \text{(C16)}$$

$$\mathsf{Z}_{n+j}\mathsf{Z}_i^{(k)}|\psi\rangle=\mathsf{Z}_{n+i}\mathsf{Z}_j^{(k')}|\psi\rangle, \qquad \text{(C17)}$$

$$\mathsf{Z}_{n+j}\mathsf{Z}_{n+i}|\psi\rangle=\mathsf{Z}_{n+i}\mathsf{Z}_{n+j}|\psi\rangle. \qquad \text{(C18)}$$

In fact, by working backwards using different values of k, k' and (C13) again, one sees

$$\mathsf{Z}_i^{(k)}\mathsf{Z}_j^{(k')}|\psi\rangle=\mathsf{Z}_j^{(k')}\mathsf{Z}_i^{(k)}|\psi\rangle\quad \forall k,k',i\neq j. \qquad \text{(C19)}$$

In the same fashion, one proves

$$\mathsf{X}_i^{(k)}\mathsf{X}_j^{(k')}|\psi\rangle=\mathsf{X}_j^{(k')}\mathsf{X}_i^{(k)}|\psi\rangle\quad \forall k,k',i\neq j, \qquad \text{(C20)}$$

$$\mathsf{Y}_i^{(k)}\mathsf{Y}_j^{(k')}|\psi\rangle=\mathsf{Y}_j^{(k')}\mathsf{Y}_i^{(k)}|\psi\rangle\quad \forall k,k',i\neq j, \qquad \text{(C21)}$$

$$\mathsf{X}_i^{(k)}\mathsf{Y}_j^{(k')}|\psi\rangle = \mathsf{Y}_j^{(k')}\mathsf{X}_i^{(k)}|\psi\rangle \quad \forall\, k,k',i\neq j, \tag{C22}$$

$$\mathsf{X}_i^{(k)}\mathsf{Z}_j^{(k')}|\psi\rangle = \mathsf{Z}_j^{(k')}\mathsf{X}_i^{(k)}|\psi\rangle \quad \forall\, k,k',i\neq j, \tag{C23}$$

$$\mathsf{Y}_i^{(k)}\mathsf{Z}_j^{(k')}|\psi\rangle = \mathsf{Z}_j^{(k')}\mathsf{Y}_i^{(k)}|\psi\rangle \quad \forall\, k,k',i\neq j. \tag{C24}$$

We have now finished the necessary groundwork to construct the self-testing circuit of Lemma 2.

2. Verification of circuit

The circuit we use (see Fig. 4) is a parallel version of the circuit used in the two qubit case. To prove that it functions correctly, we make repeated use of the properties (C11), (C14), and (C19)–(C24). Before the action of the first controlled gate the system is in state

$$|\psi\rangle^{\mathrm{CA}}\frac{1}{2^{2n}}\sum_{p,q,r,s\in(0,1)^n}|p\rangle^{\mathrm{C}'}|q\rangle^{\mathrm{C}''}|r\rangle^{\mathrm{A}'}|s\rangle^{\mathrm{A}''}, \tag{C25}$$

and after the first controlled gate the state evolves to

$$\frac{1}{2^{2n}}\sum_{p,q,r,s\in(0,1)^n}\left[\otimes_{i=1}^n\left(\mathsf{Z}_i^{(k)}\right)^{p_i}(\hat{\mathsf{Z}}_{i+n})^{r_i}|\psi\rangle^{\mathrm{CA}}\right]|p\rangle^{\mathrm{C}'}|q\rangle^{\mathrm{C}''}|r\rangle^{\mathrm{A}'}|s\rangle^{\mathrm{A}''}, \tag{C26}$$

where $p_i(r_i)$ is the ith element of string $p(r)$. Hadamard gates evolve the state to

$$\frac{1}{2^{3n}}\sum_{p,q,r,s\in(0,1)^n}\left[\otimes_{i=1}^n\left(\mathbb{1}+(-1)^{p_i}\mathsf{Z}_i^{(k)}\right)\left(\mathbb{1}+(-1)^{r_i}\hat{\mathsf{Z}}_{i+n}\right)|\psi\rangle^{\mathrm{CA}}\right]|p\rangle^{\mathrm{C}'}|q\rangle^{\mathrm{C}''}|r\rangle^{\mathrm{A}'}|s\rangle^{\mathrm{A}''}, \tag{C27}$$

and the second controlled gates lead to

$$\frac{1}{2^{3n}}\sum_{p,q,r,s\in(0,1)^n}\left[\otimes_{i=1}^n\left(\mathsf{X}_i^{(k)}\right)^{p_i}\left(\mathbb{1}+(-1)^{p_i}\mathsf{Z}_i^{(k)}\right)(\hat{\mathsf{X}}_{n+i})^{r_i}\left(\mathbb{1}+(-1)^{r_i}\hat{\mathsf{Z}}_{i+n}\right)|\psi\rangle^{\mathrm{CA}}\right]|p\rangle^{\mathrm{C}'}|q\rangle^{\mathrm{C}''}|r\rangle^{\mathrm{A}'}|s\rangle^{\mathrm{A}''}. \tag{C28}$$

Relations (C14) and (C23) allow us to simplify this to

$$\frac{1}{2^{3n}}\sum_{p,q,r,s\in(0,1)^n}\left[\otimes_{i=1}^n\left(\mathsf{X}_i^{(k)}\right)^{p_i}\left(\mathbb{1}+(-1)^{p_i}\mathsf{Z}_i^{(k)}\right)(\hat{\mathsf{X}}_{n+i})^{r_i}\left(\mathbb{1}+(-1)^{r_i}\mathsf{Z}_i^{(k)}\right)|\psi\rangle^{\mathrm{CA}}\right]|p\rangle^{\mathrm{C}'}|q\rangle^{\mathrm{C}''}|r\rangle^{\mathrm{A}'}|s\rangle^{\mathrm{A}''}. \tag{C29}$$

Unitarity and hermiticity of $\mathsf{Z}_i^{(k)}$ implies $(\mathbb{1}+\mathsf{Z}_i^{(k)})(\mathbb{1}-\mathsf{Z}_i^{(k)})|\psi\rangle=0$ and $\frac{1}{4}(\mathbb{1}+\mathsf{Z}_i^{(k)})(\mathbb{1}+\mathsf{Z}_i^{(k)})|\psi\rangle=\frac{1}{2}(\mathbb{1}+\mathsf{Z}_i^{(k)})|\psi\rangle$ so that for every i the state of the system can be further simplified to obtain

$$\frac{1}{2^{2n}}\sum_{p,q,s\in(0,1)^n}\left[\otimes_{i=1}^n\left(\mathsf{X}_i^{(k)}\right)^{p_i}\left(\mathbb{1}+(-1)^{p_i}\mathsf{Z}_i^{(k)}\right)(\hat{\mathsf{X}}_{n+i})^{p_i}|\psi\rangle^{\mathrm{CA}}\right]|p\rangle^{\mathrm{C}'}|q\rangle^{\mathrm{C}''}|p\rangle^{\mathrm{A}'}|s\rangle^{\mathrm{A}''}. \tag{C30}$$

This can be further simplified by using (C11) and (C20):

$$\frac{1}{2^{2n}}\sum_{p,q,s\in(0,1)^n}\left[\otimes_{i=1}^n\left(\mathbb{1}+\mathsf{Z}_i^{(k)}\right)|\psi\rangle^{\mathrm{CA}}\right]|p\rangle^{\mathrm{C}'}|q\rangle^{\mathrm{C}''}|p\rangle^{\mathrm{A}'}|s\rangle^{\mathrm{A}''}=\frac{1}{2^{\frac{3n}{2}}}\sum_{q,s\in(0,1)^n}\left[\otimes_{i=1}^n\left(\mathbb{1}+\mathsf{Z}_i^{(k)}\right)|\psi\rangle^{\mathrm{CA}}\right]\left[\otimes_{i=1}^n|\Phi^+\rangle^{\mathrm{C}_i'\mathrm{A}_i'}\right]|q\rangle^{\mathrm{C}''}|s\rangle^{\mathrm{A}''}. \tag{C31}$$

Already here the state of the primed auxiliaries (extraction auxiliaries in the following text) is n-fold tensor product of maximally entangled pairs of qubits. Since the rest of the circuit does not affect extraction auxiliaries for the sake of simplicity it will be omitted from the following expressions. Following the action of the third pair of controlled gates the system evolves to

$$\frac{1}{2^{\frac{3n}{2}}}\sum_{q,s\in(0,1)^n}\left[\otimes_{i=1}^n\left(i\mathsf{Y}_i^{(k)}\mathsf{X}_i^{(k)}\right)^{q_i}\left(\mathbb{1}+\mathsf{Z}_i^{(k)}\right)(i\hat{\mathsf{Y}}_{n+i}\hat{\mathsf{X}}_{n+i})^{s_i}|\psi\rangle^{\mathrm{CA}}\right]|q\rangle^{\mathrm{C}''}|s\rangle^{\mathrm{A}''}. \tag{C32}$$

By virtue of (C14), (C11), (C24), (C22), and (C23) this simplifies to

$$\frac{1}{2^{\frac{3n}{2}}}\sum_{q,s\in(0,1)^n}\left[\otimes_{i=1}^n\left(i\mathsf{Y}_i^{(k)}\mathsf{X}_i^{(k)}\right)^{q_i+s_i}\left(\mathbb{1}+\mathsf{Z}_i^{(k)}\right)|\psi\rangle^{\mathrm{CA}}\right]|q\rangle^{\mathrm{C}''}|s\rangle^{\mathrm{A}''}. \tag{C33}$$

Finally, at the end of the circuit, after the action of the second pair of Hadamards we have

$$\frac{1}{2^{\frac{5n}{2}}}\sum_{q,s,\bar{q},\bar{s}\in(0,1)^n}\left[\otimes_{i=1}^n(-1)^{\bar{q}_iq_i+\bar{s}_is_i}\left(i\mathsf{Y}_i^{(k)}\mathsf{X}_i^{(k)}\right)^{q_i+s_i}\left(\mathbb{1}+\mathsf{Z}_i^{(k)}\right)|\psi\rangle^{\mathrm{CA}}\right]|\bar{q}\rangle^{\mathrm{C}''}|\bar{s}\rangle^{\mathrm{A}''}. \tag{C34}$$

Note that each term from the sum is characterized by a pair of strings $(\bar{q},\bar{s})$ and a set of pairs of strings Ξ, such that $q''_j+s''_j=q'_j+s'_j$ for every $q'',s'',q',s'\in\Xi$ and every j. We show that the multiplicative factor in front of every term is equal to zero whenever $\bar{q}'\neq\bar{s}'$. Let us assume $\bar{q}'=\bar{s}'$. The multiplicative factor for a term corresponding to a pair of strings q',s' is equal to

$$(-1)^{\sum_{q',s'\in\Xi,j}\bar{q}'_jq'_j+\bar{s}'_js'_j}=(-1)^{\sum_{q',s'\in\Xi,j}\bar{q}'_j(q'_j+s'_j)}=\pm1,$$

i.e., all the terms come with the same sign, since sum is over q',s' which have fixed $q'_j+s'_j$ for every j. Contrarily, in case $\bar{q}'\neq\bar{s}'$ the multiplicative factor for a term corresponding to a pair of strings q',s' is equal to

$$(-1)^{\sum_{q',s'\in\Xi,j}\bar{q}'_jq'_j+\bar{s}'_js'_j}=(-1)^{\sum_{q',s'\in\Xi,j}\bar{q}'_j(q'_j+s'_j)+(\bar{s}'_j-\bar{q}'_j)s'_j}=\begin{cases}\pm1 & \text{when} \quad \sum_j s'_j=0,\\ \mp1 & \text{when} \quad \sum_j s'_j=1,\end{cases}=0.$$

In this case value of s'_j determines the sign of the terms, and for half of the terms it is equal 0 (one sign) and for the half it is equal to 1 (opposite sign). This means that only terms of the sum which survive are those corresponding to $\bar{q}=\bar{s}$:

$$\frac{1}{2^{\frac{5n}{2}}}\sum_{q,s,\bar{q}\in(0,1)^n}\left[\otimes_{i=1}^n(-1)^{\bar{q}_i(q_i+s_i)}\left(i\mathsf{Y}_i^{(k)}\mathsf{X}_i^{(k)}\right)^{q_i+s_i}\left(\mathbb{1}+\mathsf{Z}_i^{(k)}\right)|\psi\rangle^{\mathrm{CA}}\right]|\bar{q}\bar{q}\rangle^{\mathrm{C''A''}}. \tag{C35}$$

The sum has 2^{3n} different contributions (one for each triple $q,s,\bar{q}$), but there are 2^{2n} different terms, meaning that each term has contributions from 2^n different pairs of strings (q,s). This reduces the multiplicative factor in front of the sum to $2^{-\frac{3n}{2}}$. After summing over q,s and making some rearrangements the expression reduces to

$$|\xi\rangle=\frac{1}{2^{\frac{3n}{2}}}\sum_{\bar{q}\in(0,1)^n}\left[\otimes_{i=1}^n\left(\mathbb{1}+(-1)^{\bar{q}_i}i\mathsf{Y}_i^{(k)}\mathsf{X}_i^{(k)}\right)\left(\mathbb{1}+\mathsf{Z}_i^{(k)}\right)|\psi\rangle^{\mathrm{CA}}\right]|\bar{q}\bar{q}\rangle^{\mathrm{C''A''}}. \tag{C36}$$

Finally, by returning the state of extraction auxiliary systems one obtains the statement from Lemma 2:

$$U[|\psi\rangle^{\mathrm{CA}}\otimes|00\rangle]=|\xi\rangle\otimes_{i=1}^n|\Phi^+\rangle^{\mathrm{C'_iA'_i}}. \tag{C37}$$

Before calculating the output of the circuit when the input is $\mathsf{Z}_i^{(k)}|\psi\rangle$ let us acknowledge that $\mathsf{Z}_i^{(k)}|\psi\rangle=\mathsf{Z}_i^{(l)}|\psi\rangle$ for any two l and k, which can be seen from (C14) which is satisfied for any k. The same holds for $\mathsf{X}_i^{(k)}|\psi\rangle$ and $\mathsf{Y}_i^{(k)}|\psi\rangle$. By repeating the same procedure as in the derivation above one can confirm two more statements from Lemma 2 for any k and j:

$$U\left[\mathsf{Z}_j^{(k)}|\psi\rangle^{\mathrm{CA}}\otimes|00\rangle\right]=|\xi\rangle\left[\sigma_z^{\mathrm{C'}_j}\otimes_{i=1}^n|\Phi^+\rangle^{\mathrm{C'_iA'_i}}\right],\quad U\left[\mathsf{X}_j^{(k)}|\psi\rangle^{\mathrm{CA}}\otimes|00\rangle\right]=|\xi\rangle\left[\sigma_x^{\mathrm{C'}_j}\otimes_{i=1}^n|\Phi^+\rangle^{\mathrm{C'_iA'_i}}\right]. \tag{C38}$$

The situation when the input state is $\mathsf{Y}_j^{(k)}|\psi\rangle$ is a bit more complicated so more details of the derivation will be presented. After the second pair of controlled gates the state of the system is

$$\frac{1}{2^{3n}}\sum_{p,q,r,s\in(0,1)^n}\left[\otimes_{i=1}^n\left(\mathsf{X}_i^{(k)}\right)^{p_i}\left(\mathbb{1}+(-1)^{p_i}\mathsf{Z}_i^{(k)}\right)\mathsf{Y}_j^{(k)}(\hat{\mathsf{X}}_{n+i})^{r_i}(\mathbb{1}+(-1)^{r_i}\hat{\mathsf{Z}}_{i+n})|\psi\rangle^{\mathrm{CA}}\right]|p\rangle^{\mathrm{C'}}|q\rangle^{\mathrm{C''}}|r\rangle^{\mathrm{A'}}|s\rangle^{\mathrm{A''}}, \tag{C39}$$

which due to Eqs. (C11) and (C24) simplifies to

$$\frac{1}{2^{3n}}\sum_{p,q,r,s\in(0,1)^n}\left[\otimes_{i=1}^n\left(\mathsf{X}_i^{(k)}\right)^{p_i}\mathsf{Y}_j^{(k)}\left(\mathbb{1}+(-1)^{p_i\oplus\delta_{ij}}\mathsf{Z}_i^{(k)}\right)(\hat{\mathsf{X}}_{n+i})^{r_i}(\mathbb{1}+(-1)^{r_i}\hat{\mathsf{Z}}_{i+n})|\psi\rangle^{\mathrm{CA}}\right]|p\rangle^{\mathrm{C'}}|q\rangle^{\mathrm{C''}}|r\rangle^{\mathrm{A'}}|s\rangle^{\mathrm{A''}}, \tag{C40}$$

By using (C19) and (C11) and the fact that $\frac{\mathbb{1}+\mathsf{Z}_i^{(k)}}{2}$ and $\frac{\mathbb{1}-\mathsf{Z}_i^{(k)}}{2}$ are projectors onto different eigenspaces of $\mathsf{Z}_i^{(k)}$ the above reduces to

$$\frac{1}{2^{2n}}\sum_{q,r,s\in(0,1)^n}\left[\otimes_{i=1}^n(-1)^{r_i\oplus\delta_{ij}}\mathsf{Y}_j^{(k)}\mathsf{X}_j^{(k)}\left(\mathbb{1}+\mathsf{Z}_i^{(k)}\right)|\psi\rangle^{\mathrm{CA}}\right]|r\oplus1_j\rangle^{\mathrm{C'}}|q\rangle^{\mathrm{C''}}|r\rangle^{\mathrm{A'}}|s\rangle^{\mathrm{A''}}, \tag{C41}$$

where 1_j is an n-element string whose jth element is one with all the other elements being zeros. The last expression can be rewritten in the following way:

$$\frac{1}{2^{\frac{3n}{2}}}\sum_{q,s\in(0,1)^n}\left[\otimes_{i=1}^n i\mathsf{Y}_j^{(k)}\mathsf{X}_j^{(k)}\left(\mathbb{1}+\mathsf{Z}_i^{(k)}\right)|\psi\rangle^{\mathrm{CA}}\right]\sigma_y^{\mathrm{C'}_j}\left[\otimes_{i=1}^n|\Phi^+\rangle^{\mathrm{C'_iA'_i}}\right]|q\rangle^{\mathrm{C''}}|s\rangle^{\mathrm{A''}}. \tag{C42}$$

Since the rest of the circuit does not affect the state of extraction auxiliaries we will drop it from the following few equations. After applying the third pair of controlled gates on this state one obtains

$$\frac{1}{2^{\frac{3n}{2}}}\sum_{q,s\in(0,1)^n}\left[\otimes_{i=1}^n\left(i\mathsf{Y}_i^{(k)}\mathsf{X}_i^{(k)}\right)^{q_i+\delta_{ij}}\left(\mathbb{1}+\mathsf{Z}_i^{(k)}\right)(i\hat{\mathsf{Y}}_{i+n}\hat{\mathsf{X}}_{i+n})^{s_i}|\psi\rangle^{\mathrm{CA}}\right]|q\rangle^{\mathrm{C''}}|s\rangle^{\mathrm{A''}}, \tag{C43}$$

which due to (C14) and anticommuting relations (C11) reduces to

$$\frac{1}{2^{\frac{3n}{2}}}\sum_{q,s\in(0,1)^n}\left[\otimes_{i=1}^n\left(i\mathsf{Y}_i^{(k)}\mathsf{X}_i^{(k)}\right)^{s_i+q_i+\delta_{ij}}\left(\mathbb{1}+\mathsf{Z}_i^{(k)}\right)|\psi\rangle^{\mathrm{CA}}\right]|q\rangle^{\mathrm{C}''}|s\rangle^{\mathrm{A}''}, \tag{C44}$$

and at the end of the circuit following the action of the two last Hadamards this state transforms to

$$\frac{1}{2^{\frac{5n}{2}}}\sum_{\bar{q},\bar{s},q,s\in(0,1)^n}(-1)^{\bar{q}_iq_i+\bar{s}_is_i}\left[\otimes_{i=1}^n\left(i\mathsf{Y}_i^{(k)}\mathsf{X}_i^{(k)}\right)^{q_i+s_i+\delta_{ij}}\left(\mathbb{1}+\mathsf{Z}_i^{(k)}\right)|\psi\rangle^{\mathrm{CA}}\right]|\bar{q}\rangle^{\mathrm{C}''}|s\rangle^{\mathrm{A}''}. \tag{C45}$$

Here the same reasoning like the one preceding Eq. (C36) can be applied, the only difference being factor $(i\mathsf{Y}_i^{(k)}\mathsf{X}_i^{(k)})^{\delta_{ij}}$. This factor changes the sign of terms in (C36) which correspond to any string $\bar{q}$ for which $\bar{q}_j=1$. The final form of the output of the circuit when input is $\mathsf{Y}_j^{(k)}|\psi\rangle$ can be written as

$$\frac{1}{2^{\frac{3n}{2}}}\sum_{\bar{q}\in(0,1)^n}\left[\otimes_{i=1}^n(-1)^{\bar{q}_j}\left(\mathbb{1}+(-1)^{\bar{q}_i}i\mathsf{Y}_i^{(k)}\mathsf{X}_i^{(k)}\right)\left(\mathbb{1}+\mathsf{Z}_i^{(k)}\right)|\psi\rangle^{\mathrm{CA}}\right]\sigma_y^{\mathrm{C}'_j}\left[\otimes_{i=1}^n|\Phi^+\rangle^{\mathrm{C}'_i\mathrm{A}'_i}\right]|\bar{q}\bar{q}\rangle^{\mathrm{C}''\mathrm{A}''}, \tag{C46}$$

which is equivalent to the formulation from Lemma 2:

$$U\left[\mathsf{Y}_j^{\mathrm{C}}|\psi\rangle^{\mathrm{CA}}\otimes|00\rangle\right]=\sigma_z^{\mathrm{C}''_j}|\xi\rangle\left[\sigma_y{}^{\mathrm{C}'_j}\otimes_{i=1}^n|\Phi^+\rangle^{\mathrm{C}'_i\mathrm{A}'_i}\right], \tag{C47}$$

which completes the proof.

APPENDIX D: PROOF OF LEMMA 3

Correlations $\langle\psi|\mathsf{S}_{l,a}|\psi\rangle=\langle\psi|\mathsf{S}_{l,a}|\psi\rangle=\frac{1}{4}$ for every $l\in\{1,\ldots,m\}$ and $a\in\{0,1,2,3\}$, given in Table I, imply that the norm of states $\mathsf{S}_{l,a}|\psi\rangle$ and $\mathsf{T}_{l,a}|\psi\rangle$ is equal to $\frac{1}{2}$. These correlations allow us to write

$$\mathsf{S}_{l,0}|\psi\rangle\sim\tfrac{1}{4}\left(|\psi\rangle+\mathsf{Z}_{2l-1}^{(k)}\mathsf{Z}_{2l}^{(k)}|\psi\rangle+\mathsf{X}_{2l-1}^{(k)}\mathsf{X}_{2l}^{(k)}|\psi\rangle-\mathsf{Y}_{2l-1}^{(k)}\mathsf{Y}_{2l}^{(k)}|\psi\rangle\right). \tag{D1}$$

Since states $|\psi\rangle$, $\mathsf{Z}_{2l-1}^{(k)}\mathsf{Z}_{2l}^{(k)}|\psi\rangle$, $\mathsf{X}_{2l-1}^{(k)}\mathsf{X}_{2l}^{(k)}|\psi\rangle$, and $\mathsf{Y}_{2l-1}^{(k)}\mathsf{Y}_{2l}^{(k)}|\psi\rangle$ all have unit norm and are mutually orthogonal they can be seen as a part of the basis of all states from $\mathcal{H}^C\otimes\mathcal{H}^A$. Moreover, $\mathsf{S}_{l,0}|\psi\rangle$ has the same norm as the expression from the right-hand side of $\sim$ in Eq. (D1) which implies that

$$\mathsf{S}_{l,0}|\psi\rangle=\tfrac{1}{4}\left(|\psi\rangle+\mathsf{Z}_{2l-1}^{(k)}\mathsf{Z}_{2l}^{(k)}|\psi\rangle+\mathsf{X}_{2l-1}^{(k)}\mathsf{X}_{2l}^{(k)}|\psi\rangle-\mathsf{Y}_{2l-1}^{(k)}\mathsf{Y}_{2l}^{(k)}|\psi\rangle\right). \tag{D2}$$

The same reasoning leads to the following set of equations:

$$\mathsf{S}_{l,1}|\psi\rangle=\tfrac{1}{4}\left(|\psi\rangle+\mathsf{Z}_{2l-1}^{(k)}\mathsf{Z}_{2l}^{(k)}|\psi\rangle-\mathsf{X}_{2l-1}^{(k)}\mathsf{X}_{2l}^{(k)}|\psi\rangle+\mathsf{Y}_{2l-1}^{(k)}\mathsf{Y}_{2l}^{(k)}|\psi\rangle\right), \tag{D3}$$

$$\mathsf{S}_{l,2}|\psi\rangle=\tfrac{1}{4}\left(|\psi\rangle-\mathsf{Z}_{2l-1}^{(k)}\mathsf{Z}_{2l}^{(k)}|\psi\rangle+\mathsf{X}_{2l-1}^{(k)}\mathsf{X}_{2l}^{(k)}|\psi\rangle+\mathsf{Y}_{2l-1}^{(k)}\mathsf{Y}_{2l}^{(k)}|\psi\rangle\right), \tag{D4}$$

$$\mathsf{S}_{l,3}|\psi\rangle=\tfrac{1}{4}\left(|\psi\rangle-\mathsf{Z}_{2l-1}^{(k)}\mathsf{Z}_{2l}^{(k)}|\psi\rangle-\mathsf{X}_{2l-1}^{(k)}\mathsf{X}_{2l}^{(k)}|\psi\rangle-\mathsf{Y}_{2l-1}^{(k)}\mathsf{Y}_{2l}^{(k)}|\psi\rangle\right), \tag{D5}$$

$$\mathsf{T}_{l,0}|\psi\rangle=\tfrac{1}{4}\left(|\psi\rangle+\mathsf{Z}_{2l}^{(k)}\mathsf{Z}_{2l+1}^{(k)}|\psi\rangle+\mathsf{X}_{2l}^{(k)}\mathsf{X}_{2l+1}^{(k)}|\psi\rangle-\mathsf{Y}_{2l}^{(k)}\mathsf{Y}_{2l+1}^{(k)}|\psi\rangle\right), \tag{D6}$$

$$\mathsf{T}_{l,1}|\psi\rangle=\tfrac{1}{4}\left(|\psi\rangle+\mathsf{Z}_{2l}^{(k)}\mathsf{Z}_{2l+1}^{(k)}|\psi\rangle-\mathsf{X}_{2l}^{(k)}\mathsf{X}_{2l+1}^{(k)}|\psi\rangle+\mathsf{Y}_{2l}^{(k)}\mathsf{Y}_{2l+1}^{(k)}|\psi\rangle\right), \tag{D7}$$

$$\mathsf{T}_{l,2}|\psi\rangle=\tfrac{1}{4}\left(|\psi\rangle-\mathsf{Z}_{2l}^{(k)}\mathsf{Z}_{2l+1}^{(k)}|\psi\rangle+\mathsf{X}_{2l}^{(k)}\mathsf{X}_{2l+1}^{(k)}|\psi\rangle+\mathsf{Y}_{2l}^{(k)}\mathsf{Y}_{2l+1}^{(k)}|\psi\rangle\right), \tag{D8}$$

$$\mathsf{T}_{l,3}|\psi\rangle=\tfrac{1}{4}\left(|\psi\rangle-\mathsf{Z}_{2l}^{(k)}\mathsf{Z}_{2l+1}^{(k)}|\psi\rangle-\mathsf{X}_{2l}^{(k)}\mathsf{X}_{2l+1}^{(k)}|\psi\rangle-\mathsf{Y}_{2l}^{(k)}\mathsf{Y}_{2l+1}^{(k)}|\psi\rangle\right). \tag{D9}$$

Equations (D2)–(D5) are equivalent to the following set of equations:

$$\mathsf{Z}_{2l-1}^{(k)}\mathsf{Z}_{2l}^{(k)}|\psi\rangle=(\mathsf{S}_{l,0}+\mathsf{S}_{l,1}-\mathsf{S}_{l,2}-\mathsf{S}_{l,3})|\psi\rangle, \tag{D10a}$$

$$\mathsf{X}_{2l-1}^{(k)}\mathsf{X}_{2l}^{(k)}|\psi\rangle=(\mathsf{S}_{l,0}-\mathsf{S}_{l,1}+\mathsf{S}_{l,2}-\mathsf{S}_{l,3})|\psi\rangle, \tag{D10b}$$

$$\mathsf{Y}_{2l-1}^{(k)}\mathsf{Y}_{2l}^{(k)}|\psi\rangle=(-\mathsf{S}_{l,0}+\mathsf{S}_{l,1}+\mathsf{S}_{l,2}-\mathsf{S}_{l,3})|\psi\rangle. \tag{D10c}$$

Based on the last set of equations and the fact that $\{\mathsf{S}_{l,a}\}_{l,a}$ is orthogonal set of projectors which all commute with all the operators from $\{\mathsf{Z}_j^{(k)}, \mathsf{X}_j^{(k)}\}_{j,k}$, one can show that

$$\begin{aligned}\mathsf{X}_{2l-1}^{(k)}\mathsf{X}_{2l}^{(k)}\mathsf{Z}_{2l-1}^{(k)}\mathsf{Z}_{2l}^{(k)}|\psi\rangle &= \mathsf{X}_{2l-1}^{(k)}\mathsf{X}_{2l}^{(k)}(\mathsf{S}_{l,0}+\mathsf{S}_{l,1}-\mathsf{S}_{l,2}-\mathsf{S}_{l,3})|\psi\rangle\\ &= (\mathsf{S}_{l,0}+\mathsf{S}_{l,1}-\mathsf{S}_{l,2}-\mathsf{S}_{l,3})(\mathsf{S}_{l,0}-\mathsf{S}_{l,1}+\mathsf{S}_{l,2}-\mathsf{S}_{l,3})|\psi\rangle\\ &= (\mathsf{S}_{l,0}-\mathsf{S}_{l,1}-\mathsf{S}_{l,2}+\mathsf{S}_{l,3})|\psi\rangle\\ &= -\mathsf{Y}_{2l-1}^{(k)}\mathsf{Y}_{2l}^{(k)}|\psi\rangle. \end{aligned} \tag{D11}$$

Starting from Eqs. (D6)–(D9) one can obtain

$$\mathsf{X}_{2l}^{(k)}\mathsf{X}_{2l+1}^{(k)}\mathsf{Z}_{2l}^{(k)}\mathsf{Z}_{2l+1}^{(k)}|\psi\rangle = -\mathsf{Y}_{2l}^{(k)}\mathsf{Y}_{2l+1}^{(k)}|\psi\rangle. \tag{D12}$$

Equations (D11) and (D12) hold for every k and every l. Let us take $l=1$ and check how Eq. (D11) affects vector $|\xi_{\bar{q}}\rangle = \otimes_{i=1}^{n}(\mathbb{1}+(-1)^{\bar{q}_i}i\mathsf{Y}_i^{(k)}\mathsf{X}_i^{(k)})(\mathbb{1}+\mathsf{Z}_i^{(k)})|\psi\rangle$. Let us write it in the following form:

$$|\xi_{\bar{q}}\rangle = L_{\text{rest}}\otimes\big(\mathbb{1}+(-1)^{\bar{q}_1}i\mathsf{Y}_1^{(k)}\mathsf{X}_1^{(k)}\big)\big(\mathbb{1}+\mathsf{Z}_1^{(k)}\big)\otimes\big(\mathbb{1}+(-1)^{\bar{q}_2}i\mathsf{Y}_2^{(k)}\mathsf{X}_2^{(k)}\big)\big(\mathbb{1}+\mathsf{Z}_2^{(k)}\big)|\psi\rangle,$$

where $L_{\text{rest}} = \otimes_{i=3}^{n}(\mathbb{1}+(-1)^{\bar{q}_i}i\mathsf{Y}_i^{(k)}\mathsf{X}_i^{(k)})(\mathbb{1}+\mathsf{Z}_i^{(k)})$. Let us assume $\bar{q}_1\neq\bar{q}_2$ and omit L_{rest} for the sake of shorter exposition. Then $|\xi_{\bar{q}}\rangle$ reads

$$\begin{aligned}&|\psi\rangle \pm i\mathsf{Y}_2^{(k)}\mathsf{X}_2^{(k)}|\psi\rangle + \mathsf{Z}_2^{(k)}|\psi\rangle \pm i\mathsf{Y}_2^{(k)}\mathsf{X}_2^{(k)}\mathsf{Z}_2^{(k)}|\psi\rangle \mp i\mathsf{Y}_1^{(k)}\mathsf{X}_1^{(k)}|\psi\rangle + \mathsf{Y}_1^{(k)}\mathsf{X}_1^{(k)}\mathsf{Y}_2^{(k)}\mathsf{X}_2^{(k)}|\psi\rangle \mp i\mathsf{Y}_1^{(k)}\mathsf{X}_1^{(k)}\mathsf{Z}_2^{(k)}|\psi\rangle\\ &+\mathsf{Y}_1^{(k)}\mathsf{X}_1^{(k)}\mathsf{Y}_2^{(k)}\mathsf{X}_2^{(k)}\mathsf{Z}_2^{(k)}|\psi\rangle + \mathsf{Z}_1^{(k)}|\psi\rangle \pm i\mathsf{Z}_1^{(k)}\mathsf{Y}_2^{(k)}\mathsf{X}_2^{(k)}|\psi\rangle + \mathsf{Z}_1^{(k)}\mathsf{Z}_2^{(k)}|\psi\rangle \pm i\mathsf{Z}_1^{(k)}\mathsf{Y}_2^{(k)}\mathsf{X}_2^{(k)}\mathsf{Z}_2^{(k)}|\psi\rangle + \mp i\mathsf{Y}_1^{(k)}\mathsf{X}_1^{(k)}\mathsf{Z}_1^{(k)}|\psi\rangle\\ &+\mathsf{Y}_1^{(k)}\mathsf{X}_1^{(k)}\mathsf{Z}_1^{(k)}\mathsf{Y}_2^{(k)}\mathsf{X}_2^{(k)}|\psi\rangle + \mp i\mathsf{Y}_1^{(k)}\mathsf{X}_1^{(k)}\mathsf{Z}_1^{(k)}\mathsf{Z}_2^{(k)}|\psi\rangle + \mathsf{Y}_1^{(k)}\mathsf{X}_1^{(k)}\mathsf{Z}_1^{(k)}\mathsf{Y}_2^{(k)}\mathsf{X}_2^{(k)}\mathsf{Z}_2^{(k)}|\psi\rangle.\end{aligned}$$

This expression can be written as a sum of expressions, each equal to 0. To show this let us rearrange Eq. (D11) for the case $l=1$. It can be written in eight different ways, which are given below:

$$\begin{aligned}&|\psi\rangle + \mathsf{Y}_1^{(k)}\mathsf{X}_1^{(k)}\mathsf{Z}_1^{(k)}\mathsf{Y}_2^{(k)}\mathsf{X}_2^{(k)}\mathsf{Z}_2^{(k)}|\psi\rangle = 0, \quad \mathsf{Y}_2^{(k)}\mathsf{X}_2^{(k)}|\psi\rangle + \mathsf{Y}_1^{(k)}\mathsf{X}_1^{(k)}\mathsf{Z}_1^{(k)}\mathsf{Z}_2^{(k)}|\psi\rangle = 0,\\ &\mathsf{Z}_2^{(k)}|\psi\rangle + \mathsf{Y}_1^{(k)}\mathsf{X}_1^{(k)}\mathsf{Z}_1^{(k)}\mathsf{Y}_2^{(k)}\mathsf{X}_2^{(k)}|\psi\rangle = 0, \quad \mathsf{Y}_2^{(k)}\mathsf{X}_2^{(k)}\mathsf{Z}_2^{(k)}|\psi\rangle + \mathsf{Y}_1^{(k)}\mathsf{X}_1^{(k)}\mathsf{Z}_1^{(k)}|\psi\rangle = 0,\\ &\mathsf{Y}_1^{(k)}\mathsf{X}_1^{(k)}|\psi\rangle + \mathsf{Z}_1^{(k)}\mathsf{Y}_2^{(k)}\mathsf{X}_2^{(k)}\mathsf{Z}_2^{(k)}|\psi\rangle = 0, \quad \mathsf{Y}_1^{(k)}\mathsf{X}_1^{(k)}\mathsf{Y}_2^{(k)}\mathsf{X}_2^{(k)}|\psi\rangle + \mathsf{Z}_1^{(k)}\mathsf{Z}_2^{(k)}|\psi\rangle = 0,\\ &\mathsf{Y}_1^{(k)}\mathsf{X}_1^{(k)}\mathsf{Z}_2^{(k)}|\psi\rangle + \mathsf{Z}_1^{(k)}\mathsf{Y}_2^{(k)}\mathsf{X}_2^{(k)}|\psi\rangle = 0, \quad \mathsf{Y}_1^{(k)}\mathsf{X}_1^{(k)}\mathsf{Y}_2^{(k)}\mathsf{X}_2^{(k)}\mathsf{Z}_2^{(k)}|\psi\rangle + \mathsf{Z}_1^{(k)}|\psi\rangle = 0.\end{aligned} \tag{D13}$$

All these equations are obtained from Eq. (D11) by using commutation relations (C19), expressions (C14), anticommutation relations (C11), and the fact that operators $P_i^{(k)}$ for $P\in\{X,Y,Z\}$ are reflections, defined by property $P_i^{(k)^2}=\mathbb{1}$ on the support of $|\psi\rangle$.

Premise $\bar{q}_1\neq\bar{q}_2$ leads to conclusion $|\xi_{\bar{q}}\rangle=0$. In a completely analogous way, starting from Eq. (D11) one can show that $|\xi_{\bar{q}}\rangle=0$ if there exists l such that $\bar{q}_{2l-1}\neq\bar{q}_{2l}$. Similarly, Eq. (D12) can be used to prove that $|\xi_{\bar{q}}\rangle=0$ if there exists l such that $\bar{q}_{2l}\neq\bar{q}_{2l+1}$. The only two states $|\bar{q}\rangle$ which satisfy $\bar{q}_{2l-1}=\bar{q}_{2l}=\bar{q}_{2l+1}$ are $|\bar{q}\rangle=|0\cdots0\rangle$ and $|\bar{q}\rangle=|1\cdots1\rangle$. This means that

$$|\xi\rangle = |\xi_0\rangle\otimes|0\cdots0\rangle + |\xi_1\rangle\otimes|1\cdots1\rangle, \tag{D14}$$

which is exactly what had to be proven.

APPENDIX E: ROBUST VERSION OF LEMMA 3

In this Appendix we show how one can derive a noise robust version of Lemma 3 given a noise robust version of Lemma 2. Specifically, we show that if each of the probabilities differ by at most η from the values in Table I one has

$$\big\|U[|\psi\rangle\otimes|00\rangle] - |\tilde{\xi}\rangle\otimes\big[\otimes_{i=1}^{n}|\Phi^+\rangle\big]\big\| \leqslant O(\epsilon^m)+O(\sqrt{\eta}), \tag{E1}$$

$$\big\|U\big[\Pi_j\mathsf{Z}_j^{\mathrm{C}}|\psi\rangle\otimes|00\rangle\big] - |\tilde{\xi}\rangle\otimes\big[\otimes_j{\sigma_{\mathrm{z}}}^{\mathrm{C}'_j}\otimes_{i=1}^{n}|\Phi^+\rangle\big]\big\| \leqslant O(\epsilon^m)+O(\sqrt{\eta}), \tag{E2}$$

$$\big\|U\big[\Pi_j\mathsf{X}_j^{\mathrm{C}}|\psi\rangle\otimes|00\rangle\big] - |\tilde{\xi}\rangle\otimes\big[\otimes_j{\sigma_{\mathrm{x}}}^{\mathrm{C}'_j}\otimes_{i=1}^{n}|\Phi^+\rangle\big]\big\| \leqslant O(\epsilon^m)+O(\sqrt{\eta}), \tag{E3}$$

$$\big\|U\big[\Pi_j\mathsf{Y}_j^{\mathrm{C}}|\psi\rangle\otimes|00\rangle\big] - {\sigma_{\mathrm{z}}}^{\mathrm{C}''_j}|\tilde{\xi}\rangle\otimes\big[\otimes_j{\sigma_{\mathrm{y}}}^{\mathrm{C}'_j}\otimes_{i=1}^{n}|\Phi^+\rangle\big]\big\| \leqslant O(\epsilon^m)+O(\sqrt{\eta}), \tag{E4}$$

where $|\tilde{\xi}\rangle$ is the state given in Eq. (32),

$$|\tilde{\xi}\rangle = |\xi_0\rangle\otimes|0\cdots0\rangle + |\xi_1\rangle\otimes|1\cdots1\rangle \tag{E5}$$

and the scaling ϵ^m (for some m) follows from a robust self-test of Lemma 2 (see following) given nonmaximal violation of the triple CHSH Bell inequalities. Here we focus on proving (E1); a similar technique can be applied to the remaining three equations. First, note that by writing $|\Phi_n^+\rangle = \otimes_{i=1}^n |\Phi^+\rangle$ and using the triangle inequality we have

$$\|U[|\psi\rangle \otimes |00\rangle] - |\tilde{\xi}\rangle \otimes |\Phi_n^+\rangle\| = \|U[|\psi\rangle \otimes |00\rangle] - |\xi\rangle \otimes |\Phi_n^+\rangle + |\xi\rangle \otimes |\Phi_n^+\rangle - |\xi\rangle \otimes |\Phi_n^+\rangle\| \tag{E6}$$

$$\leqslant \|U[|\psi\rangle \otimes |00\rangle] - |\tilde{\xi}\rangle \otimes |\Phi_n^+\rangle\| + \||\tilde{\xi}\rangle \otimes |\Phi_n^+\rangle - |\xi\rangle \otimes |\Phi_n^+\rangle\|, \tag{E7}$$

where $|\xi\rangle$ is taken to be the state appearing in Lemma 2. The first term now gives the bound of order ϵ^m that follows from the robust self-test of Lemma 2. We now focus on the second term, that is, we need to bound

$$\||\xi\rangle - (|\xi_0\rangle \otimes |0\cdots 0\rangle + |\xi_1\rangle \otimes |1\cdots 1\rangle)\| \sim O(\sqrt{\eta}).$$

Given that there is a positive η such that observed probabilities are at most η far from the values given in Table I, let us upper bound the following expression:

$$\left\| \mathsf{S}_{l,0}|\psi\rangle - \frac{\mathbb{I} + \mathsf{Z}_{2l-1}\mathsf{Z}_{2l} + \mathsf{X}_{2l-1}\mathsf{X}_{2l} - \mathsf{Y}_{2l-1}\mathsf{Y}_{2l}}{4}|\psi\rangle \right\|. \tag{E8}$$

By definition it is equal to

$$\Bigg(\langle \mathsf{S}_{l,0}\rangle - \frac{\langle \mathsf{S}_{l,0}\rangle + \langle \mathsf{S}_{l,0}\mathsf{Z}_{2l-1}\mathsf{Z}_{2l}\rangle + \langle \mathsf{S}_{l,0}\mathsf{X}_{2l-1}\mathsf{X}_{2l}\rangle - \langle \mathsf{S}_{l,0}\mathsf{Y}_{2l-1}\mathsf{Y}_{2l}\rangle}{2} + \frac{\langle \mathbb{I}\rangle}{4} + \frac{\langle \mathsf{Z}_{2l-1}\mathsf{Z}_{2l}\rangle + \langle \mathsf{X}_{2l-1}\mathsf{X}_{2l}\rangle - \langle \mathsf{Y}_{2l-1}\mathsf{Y}_{2l}\rangle}{8}$$
$$+ \frac{\langle \mathsf{Z}_{2l-1}\mathsf{Z}_{2l}\mathsf{X}_{2l-1}\mathsf{X}_{2l}\rangle + \langle \mathsf{X}_{2l-1}\mathsf{X}_{2l}\mathsf{Z}_{2l-1}\mathsf{Z}_{2l}\rangle - \langle \mathsf{Z}_{2l-1}\mathsf{Z}_{2l}\mathsf{Y}_{2l-1}\mathsf{Y}_{2l}\rangle}{16}$$
$$+ \frac{-\langle \mathsf{Y}_{2l-1}\mathsf{Y}_{2l}\mathsf{Z}_{2l-1}\mathsf{Z}_{2l}\rangle - \langle \mathsf{X}_{2l-1}\mathsf{X}_{2l}\mathsf{Y}_{2l-1}\mathsf{Y}_{2l}\rangle - \langle \mathsf{Y}_{2l-1}\mathsf{Y}_{2l}\mathsf{X}_{2l-1}\mathsf{X}_{2l}\rangle}{16} \Bigg)^{1/2}. \tag{E9}$$

Observe now that

$$\begin{aligned}
&\left|(\langle\psi| \otimes \langle 00|)\mathsf{Z}_{2l-1}\mathsf{Z}_{2l}^{\mathrm{C}}|\psi\rangle \otimes |00\rangle\right| \\
&\quad = \left|(\langle\psi| \otimes \langle 00|)U^\dagger U\left(\mathsf{Z}_{2l-1}\mathsf{Z}_{2l}^{\mathrm{C}}|\psi\rangle \otimes |00\rangle\right)\right| \\
&\quad = \left|(\langle\psi| \otimes \langle 00|)U^\dagger\left[U\left(\mathsf{Z}_{2l-1}\mathsf{Z}_{2l}^{\mathrm{C}}|\psi\rangle \otimes |00\rangle\right) - |\tilde{\xi}\rangle \otimes \left[\sigma_{\mathrm{z}}^{\mathrm{C}'_{2l-1}} \otimes \sigma_{\mathrm{z}}^{\mathrm{C}'_{2l}} \otimes_{i=1}^n |\Phi^+\rangle\right] + |\tilde{\xi}\rangle \otimes \left[\sigma_{\mathrm{z}}^{\mathrm{C}'_{2l-1}} \otimes \sigma_{\mathrm{z}}^{\mathrm{C}'_{2l}} \otimes_{i=1}^n |\Phi^+\rangle\right]\right]\right| \\
&\quad \leqslant \|U[|\psi\rangle \otimes |00\rangle]\| \left\|U\left(\mathsf{Z}_{2l-1}\mathsf{Z}_{2l}^{\mathrm{C}}|\psi\rangle \otimes |00\rangle\right) - |\tilde{\xi}\rangle \otimes \left[\sigma_{\mathrm{z}}^{\mathrm{C}'_{2l-1}} \otimes \sigma_{\mathrm{z}}^{\mathrm{C}'_{2l}} \otimes_{i=1}^n |\Phi^+\rangle\right]\right\| \\
&\qquad + \left|\left((\langle\psi| \otimes \langle 00|)U^\dagger - \langle\tilde{\xi}| \otimes \left[\otimes_{i=1}^n \langle\Phi^+|\right] + \langle\tilde{\xi}| \otimes \left[\otimes_{i=1}^n \langle\Phi^+|\right]\right)|\tilde{\xi}\rangle \otimes \left[\sigma_{\mathrm{z}}^{\mathrm{C}'_{2l-1}} \otimes \sigma_{\mathrm{z}}^{\mathrm{C}'_{2l}} \otimes_{i=1}^n |\Phi^+\rangle\right]\right| \\
&\quad \leqslant O(\epsilon^m) + \left\|U[|\psi\rangle \otimes |00\rangle] - |\tilde{\xi}\rangle \otimes \left[\otimes_{i=1}^n |\Phi^+\rangle\right]\right\| \left\||\tilde{\xi}\rangle \otimes \left[\otimes_j \sigma_{\mathrm{z}}^{\mathrm{C}'_j} \otimes_{i=1}^n |\Phi^+\rangle\right]\right\| \leqslant O(\epsilon^m).
\end{aligned} \tag{E10}$$

In the first line we just added a unitary which does not change the inner product, while in the second line we just added a zero term. In the third line we used triangle and Cauchy-Schwartz inequalities. In the fourth line we again added a zero term and used again triangle and Cauchy-Schwartz inequalities to obtain the fifth line. Using the same sequence of steps the equivalent bound can be obtained for inner products of $\langle \mathsf{X}_{2l-1}\mathsf{X}_{2l}\rangle$ and $\langle \mathsf{Y}_{2l-1}\mathsf{Y}_{2l}\rangle$ and also for all inner products from the third and fourth line of (E9). All these inner products have absolute value as the one derived in Eq. (E10). Finally, to bound the first line from (E9) let us assume the worst case correction of Table I, i.e.,

$$\langle \mathsf{S}_{0,l}\rangle = \tfrac{1}{4} + \eta, \quad \langle \mathsf{S}_{0,l}\mathsf{Z}_{2l-1}\mathsf{Z}_{2l}\rangle = \tfrac{1}{4} - \eta, \quad \langle \mathsf{S}_{0,l}\mathsf{X}_{2l-1}\mathsf{X}_{2l}\rangle = \tfrac{1}{4} - \eta, \quad \langle \mathsf{S}_{0,l}\mathsf{Y}_{2l-1}\mathsf{Y}_{2l}\rangle = -\tfrac{1}{4} + \eta.$$

In this case the value of the first line from (E9) is equal to 2η. By summing all the terms we obtain for (E8)

$$\left\| \mathsf{S}_{l,0}|\psi\rangle - \frac{\mathbb{I} + \mathsf{Z}_{2l-1}\mathsf{Z}_{2l} + \mathsf{X}_{2l-1}\mathsf{X}_{2l} - \mathsf{Y}_{2l-1}\mathsf{Y}_{2l}}{4}|\psi\rangle \right\| \leqslant O(\eta^{1/2} + \epsilon^m). \tag{E11}$$

Similar robust versions of Eqs. (D3)–(D9) can be obtained, each having the same robustness bound. Furthermore, using triangle inequality and relations analogous to (E11) the following bounds can be obtained:

$$\begin{aligned}
&\left\|\mathsf{Z}_{2l-1}^{(k)}\mathsf{Z}_{2l}^{(k)}|\psi\rangle - (\mathsf{S}_{l,0} + \mathsf{S}_{l,1} - \mathsf{S}_{l,2} - \mathsf{S}_{l,3})|\psi\rangle\right\| \\
&\quad = \left\|\mathsf{Z}_{2l-1}^{(k)}\mathsf{Z}_{2l}^{(k)}|\psi\rangle - \mathsf{Z}_{2l-1}\mathsf{Z}_{2l}|\psi\rangle + \mathsf{Z}_{2l-1}\mathsf{Z}_{2l}|\psi\rangle - (\mathsf{S}_{l,0} + \mathsf{S}_{l,1} - \mathsf{S}_{l,2} - \mathsf{S}_{l,3})|\psi\rangle\right\| \\
&\quad \leqslant \left\|\mathsf{Z}_{2l-1}^{(k)}\mathsf{Z}_{2l}^{(k)}|\psi\rangle - \mathsf{Z}_{2l-1}\mathsf{Z}_{2l}|\psi\rangle\right\| + \|\mathsf{Z}_{2l-1}\mathsf{Z}_{2l}|\psi\rangle - (\mathsf{S}_{l,0} + \mathsf{S}_{l,1} - \mathsf{S}_{l,2} - \mathsf{S}_{l,3})|\psi\rangle\| \\
&\quad \leqslant O(\epsilon^m) + \left\|\frac{\mathsf{Z}_{2l-1}\mathsf{Z}_{2l} + \mathbb{I} + \mathsf{X}_{2l-1}\mathsf{X}_{2l} - \mathsf{Y}_{2l-1}\mathsf{Y}_{2l}}{4}|\psi\rangle + \frac{\mathsf{Z}_{2l-1}\mathsf{Z}_{2l} + \mathbb{I} - \mathsf{X}_{2l-1}\mathsf{X}_{2l} + \mathsf{Y}_{2l-1}\mathsf{Y}_{2l}}{4}|\psi\rangle\right.
\end{aligned}$$

$$
\begin{aligned}
&+\frac{\mathsf{Z}_{2l-1}\mathsf{Z}_{2l}-\mathbb{I}+\mathsf{X}_{2l-1}\mathsf{X}_{2l}+\mathsf{Y}_{2l-1}\mathsf{Y}_{2l}}{4}|\psi\rangle+\frac{\mathsf{Z}_{2l-1}\mathsf{Z}_{2l}-\mathbb{I}-\mathsf{X}_{2l-1}\mathsf{X}_{2l}-\mathsf{Y}_{2l-1}\mathsf{Y}_{2l}}{4}|\psi\rangle-(\mathsf{S}_{l,0}+\mathsf{S}_{l,1}-\mathsf{S}_{l,2}-\mathsf{S}_{l,3})|\psi\rangle\Big\|\\
&\leqslant O(\epsilon^m)+\left\|\left(\mathsf{S}_{l,0}-\frac{\mathsf{Z}_{2l-1}\mathsf{Z}_{2l}+\mathbb{I}+\mathsf{X}_{2l-1}\mathsf{X}_{2l}-\mathsf{Y}_{2l-1}\mathsf{Y}_{2l}}{4}\right)|\psi\rangle\right\|+\left\|\left(\mathsf{S}_{l,1}-\frac{\mathsf{Z}_{2l-1}\mathsf{Z}_{2l}+\mathbb{I}-\mathsf{X}_{2l-1}\mathsf{X}_{2l}+\mathsf{Y}_{2l-1}\mathsf{Y}_{2l}}{4}\right)|\psi\rangle\right\|\\
&+\left\|\left(\mathsf{S}_{l,2}+\frac{\mathsf{Z}_{2l-1}\mathsf{Z}_{2l}-\mathbb{I}-\mathsf{X}_{2l-1}\mathsf{X}_{2l}-\mathsf{Y}_{2l-1}\mathsf{Y}_{2l}}{4}\right)|\psi\rangle\right\|+\left\|\left(\mathsf{S}_{l,3}+\frac{\mathsf{Z}_{2l-1}\mathsf{Z}_{2l}-\mathbb{I}+\mathsf{X}_{2l-1}\mathsf{X}_{2l}+\mathsf{Y}_{2l-1}\mathsf{Y}_{2l}}{4}\right)|\psi\rangle\right\|\\
&\leqslant O(\eta^{1/2}+\epsilon^m), \qquad \text{(E12)}
\end{aligned}
$$

and similarly

$$\left\|\mathsf{X}^{(k)}_{2l-1}\mathsf{X}^{(k)}_{2l}|\psi\rangle-(\mathsf{S}_{l,0}-\mathsf{S}_{l,1}+\mathsf{S}_{l,2}-\mathsf{S}_{l,3})|\psi\rangle\right\|\leqslant O(\eta^{1/2}+\epsilon^m), \qquad \text{(E13)}$$

$$\left\|\mathsf{Y}^{(k)}_{2l-1}\mathsf{Y}^{(k)}_{2l}|\psi\rangle-(-\mathsf{S}_{l,0}+\mathsf{S}_{l,1}+\mathsf{S}_{l,2}-\mathsf{S}_{l,3})|\psi\rangle\right\|\leqslant O(\eta^{1/2}+\epsilon^m). \qquad \text{(E14)}$$

The robust analog of (D11) is obtained through the following chain of inequalities:

$$
\begin{aligned}
&\left\|\mathsf{X}^{(k)}_{2l-1}\mathsf{X}^{(k)}_{2l}\mathsf{Z}^{(k)}_{2l-1}\mathsf{Z}^{(k)}_{2l}|\psi\rangle+\mathsf{Y}^{(k)}_{2l-1}\mathsf{Y}^{(k)}_{2l}|\psi\rangle\right\|\\
&=\Big\|\mathsf{X}^{(k)}_{2l-1}\mathsf{X}^{(k)}_{2l}\mathsf{Z}^{(k)}_{2l-1}\mathsf{Z}^{(k)}_{2l}|\psi\rangle-\mathsf{X}^{(k)}_{2l-1}\mathsf{X}^{(k)}_{2l}(\mathsf{S}_{l,0}+\mathsf{S}_{l,1}-\mathsf{S}_{l,2}-\mathsf{S}_{l,3})|\psi\rangle\\
&\quad+\mathsf{X}^{(k)}_{2l-1}\mathsf{X}^{(k)}_{2l}(\mathsf{S}_{l,0}+\mathsf{S}_{l,1}-\mathsf{S}_{l,2}-\mathsf{S}_{l,3})|\psi\rangle+\mathsf{Y}^{(k)}_{2l-1}\mathsf{Y}^{(k)}_{2l}|\psi\rangle\Big\|\\
&\leqslant O\left(\eta^{\frac{1}{2}}+\epsilon^m\right)+\left\|\mathsf{X}^{(k)}_{2l-1}\mathsf{X}^{(k)}_{2l}(\mathsf{S}_{l,0}+\mathsf{S}_{l,1}-\mathsf{S}_{l,2}-\mathsf{S}_{l,3})|\psi\rangle+\mathsf{Y}^{(k)}_{2l-1}\mathsf{Y}^{(k)}_{2l}|\psi\rangle\right\|\\
&=O\left(\eta^{\frac{1}{2}}+\epsilon^m\right)+\Big\|\mathsf{X}^{(k)}_{2l-1}\mathsf{X}^{(k)}_{2l}(\mathsf{S}_{l,0}+\mathsf{S}_{l,1}-\mathsf{S}_{l,2}-\mathsf{S}_{l,3})|\psi\rangle-(\mathsf{S}_{l,0}+\mathsf{S}_{l,1}-\mathsf{S}_{l,2}-\mathsf{S}_{l,3})(\mathsf{S}_{l,0}-\mathsf{S}_{l,1}+\mathsf{S}_{l,2}-\mathsf{S}_{l,3})|\psi\rangle\\
&\quad+(\mathsf{S}_{l,0}+\mathsf{S}_{l,1}-\mathsf{S}_{l,2}-\mathsf{S}_{l,3})(\mathsf{S}_{l,0}-\mathsf{S}_{l,1}+\mathsf{S}_{l,2}-\mathsf{S}_{l,3})|\psi\rangle+\mathsf{Y}^{(k)}_{2l-1}\mathsf{Y}^{(k)}_{2l}|\psi\rangle\Big\|\leqslant O\left(\eta^{\frac{1}{2}}+\epsilon^m\right). \qquad \text{(E15)}
\end{aligned}
$$

To obtain the first inequality we used (E12) and the fact that multiplication by a unitary ($\mathsf{X}^{(k)}_{2l-1}\mathsf{X}^{(k)}_{2l}$) does not change the norm. The last inequality is the consequence of (E13) and (E14) and the fact that $\mathsf{S}_{l,0}+\mathsf{S}_{l,1}-\mathsf{S}_{l,2}-\mathsf{S}_{l,3}$ is a unitary operator. In a similar manner one can obtain

$$\left\|\mathsf{X}^{(k)}_{2l}\mathsf{X}^{(k)}_{2l+1}\mathsf{Z}^{(k)}_{2l}\mathsf{Z}^{(k)}_{2l+1}|\psi\rangle+\mathsf{Y}^{(k)}_{2l}\mathsf{Y}^{(k)}_{2l+1}|\psi\rangle\right\|\leqslant O(\eta^{1/2}+\epsilon^m). \qquad \text{(E16)}$$

Finally, to obtain (D14) for 2^n-2 different values of l one of two inequalities (E15) and (E16) is used eight times [see (D13)], thus leading to the final bound.

$$\||\xi\rangle-|\xi_0\rangle\otimes|0\cdots0\rangle-|\xi_1\rangle\otimes|1\cdots1\rangle\|\leqslant O[n^{1/2}(\eta^{1/2}+\epsilon^m)].$$

APPENDIX F: ENTANGLEMENT CERTIFICATION PROOFS: QUBITS

1. Positivity of $\mathcal{I}$ for separable states: qubits

Our aim is to prove that under maximal violation in step (ii) of the protocol

$$\mathcal{I}=\sum_{cduw}\omega^{zw}_{cd}\;p(c,+,+,d|z,x=\star,y=\star,w)\geqslant 0 \qquad \text{(F1)}$$

holds for all separable ϱ^{AB}. First, note that the projectors for Charlie's measurement can be compactly written

$$\Pi^{\mathrm{C}'\mathrm{C}''}_{c|z}=U^\dagger_{\mathrm{C}}\sum_j\left(\pi^{\mathrm{C}'}_{c|z}\right)^{T_j}\otimes|j\rangle\langle j|^{\mathrm{C}''}U_{\mathrm{C}}, \qquad \text{(F2)}$$

where U_{C} is the local unitary from lemma 1 and $\pi_{c|z}$ are projectors onto the Pauli eigenvectors, i.e., $\pi_{c|z}=\frac{1}{2}[\mathbb{1}+c\sigma_z]$ for $\sigma_z=\sigma_{\mathrm{z}},\sigma_{\mathrm{x}},\sigma_{\mathrm{y}}$. Thus, at maximum violation, the (subnormalized) states that Alice receives in the A_0 spaces conditional on a certain c,z are given by

$$\tau_{c|z}=\frac{1}{2}U^\dagger_{\mathrm{A}}\left[\sum_j\varrho^j_\xi\otimes\left(\pi^{\mathrm{A}'_0}_{c|z}\right)^{T_j}\right]U_{\mathrm{A}}, \qquad \text{(F3)}$$

where

$$\varrho^j_\xi=\mathrm{tr}_{\mathrm{C}''\mathrm{C}\mathrm{C}'}[|j\rangle\langle j|^{\mathrm{C}''}\otimes|\xi\rangle\langle\xi|^{\mathrm{C}''\mathrm{C}\mathrm{A}''_0\mathrm{A}_0}]. \qquad \text{(F4)}$$

Here we have used the property $\mathrm{tr}_{\mathrm{C}}[|\Phi^+\rangle\langle\Phi^+|C\otimes\mathbb{1}]=C^T$. We thus have

$$
\begin{aligned}
&p(c,+,+,d|z,x=\star,y=\star,w)\\
&=\mathrm{tr}[\mathsf{M}^{\mathrm{A}_0\mathrm{A}}_{+|\star}\otimes\mathsf{M}^{\mathrm{B}_0\mathrm{B}}_{+|\star}\,\tau_{c|z}\otimes\varrho^{\mathrm{AB}}\otimes\tau_{d|w}] \qquad \text{(F5)}\\
&=\sum_{j,k}\mathrm{tr}\left[\mathsf{A}\otimes\mathsf{B}\,\varrho^j_\xi\otimes\left(\pi^{\mathrm{A}'_0}_{c|z}\right)^{T_j}\otimes\varrho^{\mathrm{AB}}\otimes\left(\pi^{\mathrm{B}'_0}_{d|w}\right)^{T_k}\otimes\varrho^k_\xi\right], \qquad \text{(F6)}
\end{aligned}
$$

where $\mathsf{A}=\frac{1}{2}U_{\mathrm{A}}\mathsf{M}^{\mathrm{A}_0\mathrm{A}}_{+|\star}U^\dagger_{\mathrm{A}}$, $\mathsf{B}=\frac{1}{2}U_{\mathrm{B}}\mathsf{M}^{\mathrm{B}\mathrm{B}_0}_{+|\star}U^\dagger_{\mathrm{B}}$. Now, assume that ϱ^{AB} is product so that $\varrho^{\mathrm{AB}}=\sigma^{\mathrm{A}}\otimes\sigma^{\mathrm{B}}$ (mixtures of such

states will be considered later). Then the above takes the form

$$\sum_{j,k} \mathrm{tr}\left[\pi_{c|z}^{T_j} \otimes \pi_{d|w}^{T_k} \mathsf{A}_j \otimes \mathsf{B}_k\right], \tag{F7}$$

where

$$\mathsf{A}_j = \mathrm{tr}_{\mathrm{AA_0A_0''}}\left[\mathsf{A}\, \varrho_\xi^j \otimes \mathbb{1}_{\mathrm{A_0'}} \otimes \sigma^{\mathrm{A}}\right],$$
$$\mathsf{B}_k = \mathrm{tr}_{\mathrm{BB_0B_0''}}\left[\mathsf{B}\, \sigma^{\mathrm{B}} \otimes \mathbb{1}_{\mathrm{B_0'}} \otimes \varrho_\xi^k\right]. \tag{F8}$$

Note that A_j and B_k are positive operators since A_j can be seen as a positive map applied to σ^{A}. Using this we may now write $\mathcal{I}$ as

$$\mathcal{I} = \sum_{jk} \sum_{cdzw} \omega_{cd}^{zw}\, \mathrm{tr}\left[\pi_{c|z}^{T_j} \otimes \pi_{d|w}^{T_k} \mathsf{A}_j \otimes \mathsf{B}_k\right] \tag{F9}$$

$$= \sum_{jk} \sum_{cdzw} \omega_{cd}^{zw}\, \mathrm{tr}\left[\pi_{c|z} \otimes \pi_{d|w}\, \mathsf{A}_j^{T_j} \otimes \mathsf{B}_k^{T_k}\right] \tag{F10}$$

$$= \sum_{jk} \mathrm{tr}\left[\mathcal{W}\, \mathsf{A}_j^{T_j} \otimes \mathsf{B}_k^{T_k}\right] \geqslant 0, \tag{F11}$$

where the second equality follows from $\mathrm{tr}[X] = \mathrm{tr}[X^T]$, and the final inequality follows from the fact that $\mathsf{A}_j^{T_j}$ and $\mathsf{B}_k^{T_k}$ are positive operators and thus $\mathsf{A}_j^{T_j} \otimes \mathsf{B}_k^{T_k}$ is a unnormalized product state. Since $\mathcal{I}$ is linear in ϱ^{AB} one also has $\mathcal{I} \geqslant 0$ for mixtures of product states and thus all separable states.

2. Positivity of $\mathcal{I}$ for separable states: arbitrary dimension

The proof follows the same structure as for the qubit case. As a consequence of Lemma 3, we have that Alice receives the subnormalized steered states conditioned on $\mathbf{z}$, $\mathbf{c}$:

$$\tau_{\mathbf{c},\mathbf{z}} = \frac{1}{d}\, U_{\mathrm{A}}^\dagger \left[\sum_{j=0}^{1} \varrho_\xi^j \otimes \left(\pi_{\mathbf{c}|\mathbf{z}}^{\mathrm{A_0'}}\right)^{T^j}\right] U_{\mathrm{A}}, \tag{F12}$$

where we define

$$\pi_{\mathbf{c}|\mathbf{z}}^{\mathrm{A_0'}} = \otimes_i\, \pi_{c_i|z_i}^{\mathrm{A_{0i}'}} \quad \text{and}$$
$$\varrho_\xi^j = \mathrm{tr}_{\mathrm{C''CC'}}\left[(\otimes_i |j\rangle\langle j|^{\mathrm{C_i''}})\, |\xi\rangle\langle\xi|^{\mathrm{C''CA_0''A_0}}\right], \tag{F13}$$

and Bob has analogous states conditioned on Daisy's input and output. Now, the probabilities are given by

$$p(\mathbf{c}, +, +, \mathbf{d}|\mathbf{z}, x = \star, y = \star, \mathbf{w})$$
$$= \mathrm{tr}\left[\mathsf{M}_{+|\star}^{\mathrm{A_0A}} \otimes \mathsf{M}_{+|\star}^{\mathrm{B_0B}}\, \tau_{\mathbf{c}|\mathbf{z}} \otimes \varrho^{\mathrm{AB}} \otimes \tau_{\mathbf{d}|\mathbf{w}}\right] \tag{F14}$$

$$= \sum_{j,k} \mathrm{tr}\left[\mathsf{A} \otimes \mathsf{B}\, \varrho_\xi^j \otimes \left(\pi_{\mathbf{c}|\mathbf{z}}^{\mathrm{A_0'}}\right)^{T_j} \otimes \varrho^{\mathrm{AB}} \otimes \left(\pi_{\mathbf{d}|\mathbf{w}}^{\mathrm{B_0'}}\right)^{T_k} \otimes \varrho_\xi^k\right], \tag{F15}$$

and $A = \frac{1}{d} U_{\mathrm{A}} \mathsf{M}_{+|\star}^{\mathrm{A_0A}} U_{\mathrm{A}}^\dagger$, $\mathsf{B} = \frac{1}{d} U_{\mathrm{B}} \mathsf{M}_{+|\star}^{\mathrm{BB_0}} U_{\mathrm{B}}^\dagger$. For separable $\varrho^{\mathrm{AB}} = \sigma^{\mathrm{A}} \otimes \sigma^{\mathrm{B}}$ this takes the form

$$p(\mathbf{c}, +, +, \mathbf{d}|\mathbf{z}, x = \star, y = \star, \mathbf{w})$$
$$= \sum_{j,k} \mathrm{tr}\left[\pi_{\mathbf{c}|\mathbf{z}}^{T_j} \otimes \pi_{\mathbf{d}|\mathbf{w}}^{T_k} \mathsf{A}_j \otimes \mathsf{B}_k\right], \tag{F16}$$

where again we have the positive operators

$$\mathsf{A}_j = \mathrm{tr}_{\mathrm{AA_0A_0''}}\left[\mathsf{A}\, \varrho_\xi^j \otimes \mathbb{1}_{\mathrm{A_0'}} \otimes \sigma^{\mathrm{A}}\right],$$
$$\mathsf{B}_k = \mathrm{tr}_{\mathrm{BB_0B_0''}}\left[\mathsf{B}\, \sigma^{\mathrm{B}} \otimes \mathbb{1}_{\mathrm{B_0'}} \otimes \varrho_\xi^k\right]. \tag{F17}$$

Hence we find

$$\mathcal{I} = \sum_{jk} \sum_{\mathbf{cdzw}} \omega_{\mathbf{cd}}^{\mathbf{zw}}\, \mathrm{tr}\left[\pi_{\mathbf{c}|\mathbf{z}}^{T_j} \otimes \pi_{\mathbf{d}|\mathbf{w}}^{T_k} \mathsf{A}_j \otimes \mathsf{B}_k\right] \tag{F18}$$

$$= \sum_{jk} \sum_{\mathbf{cdzw}} \omega_{\mathbf{cd}}^{\mathbf{zw}}\, \mathrm{tr}\left[\pi_{\mathbf{c}|\mathbf{z}} \otimes \pi_{\mathbf{d}|\mathbf{w}}\, \mathsf{A}_j^{T_j} \otimes \mathsf{B}_k^{T_k}\right] \tag{F19}$$

$$= \sum_{jk} \mathrm{tr}\left[\mathcal{W}\, \mathsf{A}_j^{T_j} \otimes \mathsf{B}_k^{T_k}\right] \geqslant 0. \tag{F20}$$

Again, due to the linearity of $\mathcal{I}$ in ϱ^{AB}, one has $\mathcal{I} \geqslant 0$ for all separable states, completing the proof.

APPENDIX G: ROBUST ENTANGLEMENT CERTIFICATION

In this Appendix we prove a relation (61) from the main text. We start from robust self-testing statements for Lemma 3:

$$\begin{aligned}
\left\| U[|\psi\rangle \otimes |00\rangle] - |\xi\rangle \otimes \left[\otimes_{i=1}^n |\Phi^+\rangle^{\mathrm{C_i'A_i'}}\right] \right\| &\leqslant \theta,\\
\left\| U\left[\mathsf{Z}_j^{\mathrm{C}}|\psi\rangle \otimes |00\rangle\right] - |\xi\rangle \otimes \left[\sigma_{\mathrm{z}}^{\mathrm{C_j'}} \otimes_{i=1}^n |\Phi^+\rangle^{\mathrm{C_i'A_i'}}\right] \right\| &\leqslant \theta,\\
\left\| U\left[\mathsf{X}_j^{\mathrm{C}}|\psi\rangle \otimes |00\rangle\right] - |\xi\rangle \otimes \left[\sigma_{\mathrm{x}}^{\mathrm{C_j'}} \otimes_{i=1}^n |\Phi^+\rangle^{\mathrm{C_i'A_i'}}\right] \right\| &\leqslant \theta,\\
\left\| U\left[\mathsf{Y}_j^{\mathrm{C}}|\psi\rangle \otimes |00\rangle\right] - \sigma_{\mathrm{z}}^{\mathrm{C_j''}}|\xi\rangle \otimes \left[\sigma_{\mathrm{y}}^{\mathrm{C_j'}} \otimes_{i=1}^n |\Phi^+\rangle^{\mathrm{C_i'A_i'}}\right] \right\| &\leqslant \theta,
\end{aligned} \tag{G1}$$

and similarly for Daisy's measurements. These inequalities imply

$$\begin{aligned}
U[|\psi\rangle \otimes |00\rangle] &= |\xi\rangle \otimes \left[\otimes_{i=1}^n |\Phi^+\rangle^{\mathrm{C_i'A_i'}}\right] + |\hat{\Omega}\rangle,\\
U[\mathsf{Z}_j|\psi\rangle \otimes |00\rangle] &= |\xi\rangle \otimes \left[\sigma_{\mathsf{Z}}^{\mathrm{C_j'}} \otimes_{i=1}^n |\Phi^+\rangle^{\mathrm{C_i'A_i'}}\right] + |\hat{\Omega}_{\mathsf{Z}_j}\rangle,\\
U[\mathsf{X}_j|\psi\rangle \otimes |00\rangle] &= |\xi\rangle \otimes \left[\sigma_{\mathsf{X}}^{\mathrm{C_j'}} \otimes_{i=1}^n |\Phi^+\rangle^{\mathrm{C_i'A_i'}}\right] + |\hat{\Omega}_{\mathsf{X}_j}\rangle,\\
U[\mathsf{Y}_j|\psi\rangle \otimes |00\rangle] &= |\xi\rangle \otimes \left[\sigma_{\mathsf{Y}}^{\mathrm{C_j'}} \otimes_{i=1}^n |\Phi^+\rangle^{\mathrm{C_i'A_i'}}\right] + |\hat{\Omega}_{\mathsf{Y}_j}\rangle,
\end{aligned} \tag{G2}$$

where $|\hat{\Omega}\rangle$, $|\hat{\Omega}_{\mathsf{Z}_j}\rangle$, $|\hat{\Omega}_{\mathsf{X}_j}\rangle$ all have vector norm smaller than or equal to θ. Let us concentrate on the first two equations from (G2) to get

$$U\left[\frac{\mathbb{I} \pm \mathsf{Z}_j}{2}|\psi\rangle \otimes |00\rangle\right]$$
$$= |\xi\rangle \otimes \left[\frac{\mathbb{I} \pm \sigma_{\mathsf{Z}}^{\mathrm{C_j'}}}{2} \otimes_{i=1}^n |\Phi^+\rangle^{\mathrm{C_i'A_i'}}\right] + |\Omega_{\mathsf{Z}_j}^{\pm}\rangle, \tag{G3}$$

where

$$|\Omega_{\mathsf{Z}_j}^{\pm}\rangle = \tfrac{1}{2}\left(|\hat{\Omega}\rangle \pm |\hat{\Omega}_{\mathsf{Z}_j}\rangle\right)$$

is such that

$$\left\| |\Omega_{\mathsf{Z}_j}^{\pm}\rangle \right\| \leqslant \tfrac{1}{2}\left(\left\| |\hat{\Omega}_{\mathsf{Z}_j}\rangle \right\| + \| |\hat{\Omega}\rangle \|\right) = \theta, \tag{G4}$$

due to the triangle inequality. Let us also recall that

$$\left\| |\xi\rangle \otimes \left[\frac{\mathbb{I} \pm \sigma_{\mathsf{Z}}^{\mathrm{C}'_j}}{2} \otimes_{i=1}^{n} |\Phi^+\rangle^{\mathrm{C}'_i \mathrm{A}'_i} \right] \right\| = \| |\psi_{Z,\pm}\rangle \| = \frac{1}{\sqrt{2}}. \tag{G5}$$

The subnormalized state Alice receives after Charlie measures Z_j and obtains ± 1 is

$$\begin{aligned} \hat{\tau}_{\mathsf{Z}_j,\pm} &= \mathrm{tr}_{C''CC'} \left[U_A^\dagger \left(|\psi_{Z,\pm}\rangle\langle\psi_{Z,\pm}| + |\Omega^\pm_{\mathsf{Z}_j}\rangle\langle\Omega^\pm_{\mathsf{Z}_j}| \right. \right. \\ &\quad \left. \left. + |\psi_{Z,\pm}\rangle\langle\Omega^\pm_{\mathsf{Z}_j}| + |\Omega^\pm_{\mathsf{Z}_j}\rangle\langle\psi_{Z,\pm}| \right) U_A \right] \\ &= \mathrm{tr}_{C''CC'} \left[U_A^\dagger \left(|\psi_{Z,\pm}\rangle\langle\psi_{Z,\pm}| + \Delta^\pm_{\mathsf{Z}_j} \right) U_A \right]. \end{aligned} \tag{G6}$$

It is useful to estimate trace norm $\|M\|_1 = \mathrm{tr}\,|M|$ of operator $\Delta^\pm_{\mathsf{Z}_j}$. For that purpose we use triangle inequality

$$\begin{aligned} \left\| \Delta^\pm_{\mathsf{Z}_j} \right\|_1 &= \left\| |\Omega^\pm_{\mathsf{Z}_j}\rangle\langle\Omega^\pm_{\mathsf{Z}_j}| + |\psi_{Z,\pm}\rangle\langle\Omega^\pm_{\mathsf{Z}_j}| + |\Omega^\pm_{\mathsf{Z}_j}\rangle\langle\psi_{Z,\pm}| \right\|_1 \\ &\leqslant \left\| |\Omega^\pm_{\mathsf{Z}_j}\rangle\langle\Omega^\pm_{\mathsf{Z}_j}| \right\|_1 + \left\| |\psi_{Z,\pm}\rangle\langle\Omega^\pm_{\mathsf{Z}_j}| \right\|_1 \\ &\quad + \left\| |\Omega^\pm_{\mathsf{Z}_j}\rangle\langle\psi_{Z,\pm}| \right\|_1. \end{aligned} \tag{G7}$$

Let us now estimate the trace norm of each term separately, starting from the first term

$$\begin{aligned} \left\| |\Omega^\pm_{\mathsf{Z}_j}\rangle\langle\Omega^\pm_{\mathsf{Z}_j}| \right\|_1 &= \mathrm{tr}\left(\left| |\Omega^\pm_{\mathsf{Z}_j}\rangle\langle\Omega^\pm_{\mathsf{Z}_j}| \right| \right) = \mathrm{tr}\left(|\Omega^\pm_{\mathsf{Z}_j}\rangle\langle\Omega^\pm_{\mathsf{Z}_j}| \right) \\ &= \mathrm{tr}\left(\langle\Omega^\pm_{\mathsf{Z}_j}|\Omega^\pm_{\mathsf{Z}_j}\rangle \right) \leqslant \theta^2. \end{aligned}$$

The first equality is just the definition of the trace norm, the second uses positivity of $|\Omega^\pm_{\mathsf{Z}_j}\rangle\langle\Omega^\pm_{\mathsf{Z}_j}|$, and the inequality follows from (G4). Trace norm of the second term from (G7) can be bounded in the following way:

$$\left\| |\psi_{Z,\pm}\rangle\langle\Omega^\pm_{\mathsf{Z}_j}| \right\|_1 = \mathrm{tr}\left(\sqrt{|\psi_{Z,\pm}\rangle\langle\Omega^\pm_{\mathsf{Z}_j}|\,|\Omega^\pm_{\mathsf{Z}_j}\rangle\langle\psi_{Z,\pm}|} \right) \leqslant \frac{\theta}{\sqrt{2}},$$

where the inequality follows from (G4) and norm of $|\psi_{Z,\pm}\rangle$. Finally, the trace norm of third term from (G7) is

$$\begin{aligned} \left\| |\Omega^\pm_{\mathsf{Z}_j}\rangle\langle\psi_{Z,\pm}| \right\|_1 &= \mathrm{tr}\left(\sqrt{|\Omega^\pm_{\mathsf{Z}_j}\rangle\langle\psi_{Z,\pm}|\psi_{Z,\pm}\rangle\langle\Omega^\pm_{\mathsf{Z}_j}|} \right) \\ &= \frac{1}{\sqrt{2}} \mathrm{tr}\left(\sqrt{|\Omega^\pm_{\mathsf{Z}_j}\rangle\langle\Omega^\pm_{\mathsf{Z}_j}|} \right) \leqslant \frac{\theta}{\sqrt{2}}. \end{aligned}$$

To get the last inequality we used the relation

$$\mathrm{tr}\left(\sqrt{|\Omega^\pm_{\mathsf{Z}_j}\rangle\langle\Omega^\pm_{\mathsf{Z}_j}|} \right) = \mathrm{tr}\left(\sqrt{\langle\Omega^\pm_{\mathsf{Z}_j}|\Omega^\pm_{\mathsf{Z}_j}\rangle \frac{|\Omega^\pm_{\mathsf{Z}_j}\rangle\langle\Omega^\pm_{\mathsf{Z}_j}|}{\langle\Omega^\pm_{\mathsf{Z}_j}|\Omega^\pm_{\mathsf{Z}_j}\rangle}} \right) \leqslant \theta,$$

where the last inequality comes from the fact that $|\Omega^\pm_{\mathsf{Z}_j}\rangle\langle\Omega^\pm_{\mathsf{Z}_j}|/\langle\Omega^\pm_{\mathsf{Z}_j}|\Omega^\pm_{\mathsf{Z}_j}\rangle$ is a projector. Finally, (G7) reduces to

$$\left\| \Delta^\pm_{\mathsf{Z}_j} \right\|_1 \leqslant \sqrt{2}\theta + \theta^2. \tag{G8}$$

An equivalent bound can be obtained when Charlie measures X_j or Y_j. By rewriting (G7), we can see that Alice's steered states have the following form:

$$\hat{\tau}_{\mathbf{c}|\mathbf{z}} = \tau_{\mathbf{c}|\mathbf{z}} + \Delta_{\mathbf{c}|\mathbf{z}} \quad \forall \mathbf{c}, \mathbf{z},$$

where $\tau_{\mathbf{c}|\mathbf{z}}$ are the ideal steered states given in Eq. (F12). Depending on $\mathbf{c}$ and $\mathbf{z}$ the operators $\Delta_{\mathbf{c}|\mathbf{z}}$ are obtained by tracing out Charlie's system from the corresponding $\Delta^\pm_{\mathsf{P}_j}$, with $\mathsf{P} \in \{\mathsf{Z}, \mathsf{X}, \mathsf{Y}\}$. For every $\mathbf{c}$ and $\mathbf{z}$ the correction states $\Delta_{\mathbf{c}|\mathbf{z}}$ have bounded trace norm

$$\|\Delta_{\mathbf{c}|\mathbf{z}}\|_1 = \left\| \mathrm{tr}_{CC'C''} \left(\Delta^\pm_{\mathsf{P}_j} \right) \right\|_1 \leqslant \left\| \Delta^\pm_{\mathsf{P}_j} \right\|_1 \leqslant \sqrt{2}\theta + \theta^2. \tag{G9}$$

The first inequality comes from the fact that trace norm cannot increase by performing partial trace [44]. Similarly, Bob's steered states have form

$$\begin{aligned} \hat{\tau}_{\mathbf{d}|\mathbf{w}} &= \tau_{\mathbf{d}|\mathbf{w}} + \Delta_{\mathbf{d}|\mathbf{w}}, \\ \|\Delta_{\mathbf{d}|\mathbf{w}}\|_1 &\leqslant \sqrt{2}\theta + \theta^2. \end{aligned} \tag{G10}$$

Equipped with characterization of Alice's and Bob's steered states let us estimate the lowest value of $\mathcal{I}$ from (57) when evaluated on a separable state $\varrho^{AB} = \sum_\lambda p_\lambda \varrho^A_\lambda \otimes \varrho^B_\lambda$:

$$\begin{aligned} \mathcal{I} &= \sum_\lambda p_\lambda \sum_{\mathbf{c},\mathbf{d},\mathbf{z},\mathbf{w}} \omega^{\mathbf{z},\mathbf{w}}_{\mathbf{c},\mathbf{d}} \,\mathrm{tr}\left[\mathsf{M}^{\mathrm{A}_0\mathrm{A}}_{+|\star} \otimes \mathsf{M}^{\mathrm{B}_0\mathrm{B}}_{+|\star} \,\hat{\tau}_{\mathbf{c}|\mathbf{z}} \otimes \varrho^{\mathrm{A}}_\lambda \otimes \varrho^{\mathrm{B}}_\lambda \otimes \hat{\tau}_{\mathbf{d}|\mathbf{w}} \right] \\ &= \sum_\lambda p_\lambda \sum_{\mathbf{c},\mathbf{d},\mathbf{z},\mathbf{w}} \omega^{\mathbf{z},\mathbf{w}}_{\mathbf{c},\mathbf{d}} \,\mathrm{tr}\left[\mathsf{M}^{\mathrm{A}_0\mathrm{A}}_{+|\star} \otimes \mathsf{M}^{\mathrm{B}_0\mathrm{B}}_{+|\star} \,(\tau_{\mathbf{c}|\mathbf{z}} + \Delta_{\mathbf{c}|\mathbf{z}}) \otimes \varrho^{\mathrm{A}}_\lambda \otimes \varrho^{\mathrm{B}}_\lambda \otimes (\tau_{\mathbf{d}|\mathbf{w}} + \Delta_{\mathbf{d}|\mathbf{w}}) \right] \\ &= \mathcal{I}_{\text{noiseless}} + \sum_\lambda p_\lambda \sum_{\mathbf{c},\mathbf{d},\mathbf{z},\mathbf{w}} \omega^{\mathbf{z},\mathbf{w}}_{\mathbf{c},\mathbf{d}} \Big\{ \mathrm{tr}\left[\mathsf{M}^{\mathrm{A}_0\mathrm{A}}_{+|\star} \Delta_{\mathbf{c}|\mathbf{z}} \otimes \varrho^{\mathrm{A}}_\lambda \right] \mathrm{tr}\left[\mathsf{M}^{\mathrm{B}_0\mathrm{B}}_{+|\star} \tau_{\mathbf{d}|\mathbf{w}} \otimes \varrho^{\mathrm{B}}_\lambda \right] \\ &\quad + \mathrm{tr}\left[\mathsf{M}^{\mathrm{A}_0\mathrm{A}}_{+|\star} \tau_{\mathbf{c}|\mathbf{z}} \otimes \varrho^{\mathrm{A}}_\lambda \right] \mathrm{tr}\left[\mathsf{M}^{\mathrm{B}_0\mathrm{B}}_{+|\star} \Delta_{\mathbf{d}|\mathbf{w}} \otimes \varrho^{\mathrm{B}}_\lambda \right] + \mathrm{tr}\left[\mathsf{M}^{\mathrm{A}_0\mathrm{A}}_{+|\star} \Delta_{\mathbf{c}|\mathbf{z}} \otimes \varrho^{\mathrm{A}}_\lambda \right] \mathrm{tr}\left[\mathsf{M}^{\mathrm{B}_0\mathrm{B}}_{+|\star} \Delta_{\mathbf{d}|\mathbf{w}} \otimes \varrho^{\mathrm{B}}_\lambda \right] \Big\}. \end{aligned} \tag{G11}$$

$\mathcal{I}_{\text{noiseless}}$ is the value $\mathcal{I}$ would have in the ideal case $\theta = 0$. To estimate how negative the total value of $\mathcal{I}$ given in Eq. (G11) can be, we assume the worst case, i.e., $\mathcal{I}_{\text{noiseless}} = 0$ and all other contributions give negative contribution. To bound the absolute value of those contributions note that

$$\begin{aligned} \left| \mathrm{tr}\left[\mathsf{M}^{\mathrm{A}_0\mathrm{A}}_{+|\star} \Delta_{\mathbf{c}|\mathbf{z}} \otimes \varrho^{\mathrm{A}}_\lambda \right] \right| &\leqslant \mathrm{tr}\left| \mathsf{M}^{\mathrm{A}_0\mathrm{A}}_{+|\star} \Delta_{\mathbf{c}|\mathbf{z}} \otimes \varrho^{\mathrm{A}}_\lambda \right| \\ &= \left\| \mathsf{M}^{\mathrm{A}_0\mathrm{A}}_{+|\star} \Delta_{\mathbf{c}|\mathbf{z}} \otimes \varrho^{\mathrm{A}}_\lambda \right\|_1 \\ &\leqslant \left\| \mathsf{M}^{\mathrm{A}_0\mathrm{A}}_{+|\star} \right\|_\infty \left\| \Delta_{\mathbf{c}|\mathbf{z}} \otimes \varrho^{\mathrm{A}}_\lambda \right\|_1 \\ &\leqslant \mathrm{tr}(|\Delta_{\mathbf{c}|\mathbf{z}}|)\,\mathrm{tr}\left(\varrho^{\mathrm{A}}_\lambda \right) \\ &= \|\Delta_{\mathbf{c}|\mathbf{z}}\|_1 \leqslant \sqrt{2}\theta + \theta^2. \end{aligned} \tag{G12}$$

The first line follows from the inequality $|\mathrm{tr}(A)| \leqslant \mathrm{tr}\,|A|$. To obtain the third line we used Hölder's inequality $\mathrm{tr}(AB) \leqslant \|A\|_\infty \|B\|_1$ [45,46]. The fourth lines uses the fact that infinite

Schatten norm of $\mathsf{M}_{+|\star}^{\mathrm{A}_0\mathrm{A}}$ is its maximal eigenvalue which cannot be larger than one. Finally in the fifth line we used the fact that $\varrho_\lambda^{\mathrm{A}}$ is a normalized state and (G9). By using the same argumentation one can show that

$$\left|\operatorname{tr}\left[\mathsf{M}_{+|\star}^{\mathrm{A}_0\mathrm{A}}\tau_{\mathbf{d}|\mathbf{w}}\otimes\varrho_\lambda^{\mathrm{A}}\right]\right|\leqslant\tfrac{1}{2}. \tag{G13}$$

If we plug (G12) and (G13) and their analogs obtained by transforming $(\mathbf{z},\mathbf{c})\leftrightarrow(\mathbf{w},\mathbf{d})$ into (G11) we have that in the worst case

$$\mathcal{I}\sim O(\theta).$$

APPENDIX H: ENTANGLEMENT CERTIFICATION OF TWO-QUBIT WERNER STATES

Here we analyze the effect of noise in the auxiliary states when certifying the entanglement of the two-qubit Werner states

$$\varrho_W(p)=p|\Phi^+\rangle\langle\Phi^+|+(1-p)\tfrac{\mathbb{1}}{4}, \tag{H1}$$

where $|\Phi^+\rangle=\frac{1}{\sqrt{2}}(|00\rangle+|11\rangle)$. Note that the optimal entanglement witness for these state under white noise is

$$W=\sigma_z\otimes\sigma_z+\sigma_y\otimes\sigma_y+\sigma_x\otimes\sigma_x+\mathbb{I}. \tag{H2}$$

One has $\mathrm{Tr}[W\rho_{\mathrm{SEP}}]\geqslant 0$ and $\mathrm{Tr}[W\varrho_{\mathrm{W}}(p)]=1-3p$. From Eqs. (G11)–(G13) and taking the worst case, one can certify entanglement using the above witness if

$$\mathcal{I}<-12[(\sqrt{2}\theta+\theta^2)^2+\sqrt{2}\theta+\theta^2], \tag{H3}$$

where θ quantifies the robustness of self-testing [see (G1)]. Here the number 12 comes from the number of terms in the decomposition of the witness (H2) into products of Pauli projectors. Let us assume that the auxiliary states are also Werner states with visibility η. Assuming noiseless measurements, one would expect to observe a value

$$\mathcal{I}=\tfrac{1}{16}\big((1-3p)\eta^2+2\eta(1-\eta)+(1-\eta)^2\tfrac{1}{4}\big), \tag{H4}$$

since there is probability η^2 that the auxiliary states both produce a maximally entangled state and if one or no maximally entangled states are produced in the auxiliary states the value of the inequality will be $1/16$ or $1/64$, respectively. Thus, one is able to certify entanglement if

$$\mathcal{I}=\tfrac{1}{16}\big((1-3p)\eta^2+2\eta(1-\eta)+(1-\eta)^2\tfrac{1}{4}\big)<-12[(\sqrt{2}\theta+\theta^2)^2+\sqrt{2}\theta+\theta^2]. \tag{H5}$$

This inequality gives the condition that needs to be satisfied in order to be able to certify the entanglement of the state (H1). Note that θ will implicitly depend on η through some robust self-testing statement. Given a particular η, one therefore needs to ensure that there is a robust self-testing statement with corresponding θ smaller than some critical θ_{crit} given by (H5). In Fig. 5 we plot the values of θ_{crit} for different values of η and taking $p=0.6$ (note that the state has a local hidden variable model in the standard Bell scenario for this visibility [47,48]). For $\eta=1$ we have $\theta=0$ which is below θ_{crit}. As one decreases η, at some point the θ given by the robust self-testing statement will be above the critical value and the method will not work. The question is then for which value of η does this happen? Given the small values of θ_{crit} this will likely happen for a value of η very close to 1. We do not go further into the analysis here; to get precise numbers one could use the methods we present in Appendix B or for better results try to extend the method in Ref. [33] to the self-testing of measurements. We note however that very high visibilities can be achieved experimentally, e.g., using photonic setup visibilities of above 0.999 [49] and 0.997 [50] have been reported.

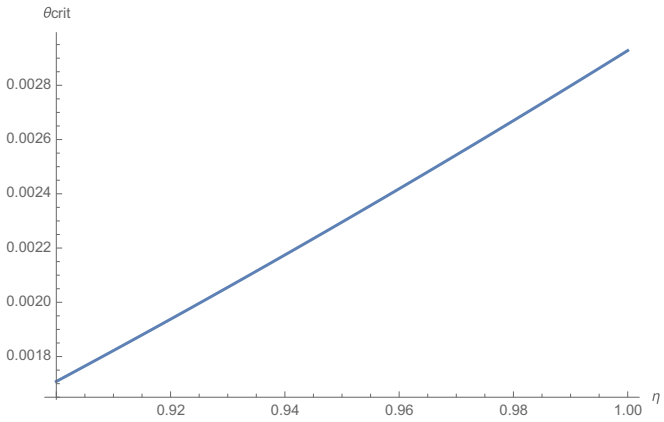

FIG. 5. Critical robustness of self-testing needed to certify the entanglement of the Werner state with noise parameter 0.6 as a function of the visibility of the auxiliary states.

[1] D. Rosset, R. Ferretti-Schobitz, J. D. Bancal, N. Gisin, and Y.-C. Liang, Imperfect measurement settings: Implications for quantum state tomography and entanglement witnesses, Phys. Rev. A **86**, 062325 (2012).

[2] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov, Hacking commercial quantum cryptography systems by tailored bright illumination, Nat. Photon. **4**, 686 (2010).

[3] J. S. Bell, On the Einstein Podolsky Rosen paradox, Physics **1**, 195 (1964).

[4] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).

[5] R. Colbeck, Quantum and relativistic protocols for secure multi-party computation, Ph.D. Thesis, University of Cambridge, 2006, arXiv:0911.3814.

[6] A. Acín and L. Masanes, Certified randomness in quantum physics, Nature (London) **540**, 213 (2016).

[7] S. Pironio *et al.*, Random numbers certified by Bell's theorem, Nature (London) **464**, 1021 (2010).

[8] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Device-Independent Security of Quantum Cryptography Against Collective Attacks, Phys. Rev. Lett. **98**, 230501 (2007).

[9] J. Barrett, L. Hardy, and A. Kent, No Signaling and Quantum Key Distribution, Phys. Rev. Lett. **95**, 010503 (2005).

[10] U. Vazirani and T. Vidick, Fully Device-Independent Quantum Key Distribution, Phys. Rev. Lett. **113**, 140501 (2014).

[11] R. Arnon-Friedman, R. Renner, and T. Vidick, Simple and tight device-independent security proofs, arXiv:1607.01797.

[12] N. Brunner, S. Pironio, A. Acín, N. Gisin, A. A. Methot, and V. Scarani, Testing the Hilbert Space Dimension, Phys. Rev. Lett. **100**, 210503 (2008).

[13] R. F. Werner, Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model, Phys. Rev. A **40**, 4277 (1989).

[14] J. Barrett, Nonsequential positive-operator-valued measurements on entangled mixed states do not always violate a Bell inequality, Phys. Rev. A **65**, 042302 (2002).

[15] R. Augusiak, M. Demianowicz, and A. Acín, Local hidden-variable models for entangled quantum states, J. Phys. A: Math. Theor. **47**, 424002 (2014).

[16] F. Buscemi, All Entangled Quantum States are Nonlocal, Phys. Rev. Lett. **108**, 200401 (2012).

[17] C. Branciard, D. Rosset, Y.-C. Liang, and N. Gisin, Measurement-Device-Independent Entanglement Witnesses for All Entangled Quantum States, Phys. Rev. Lett. **110**, 060405 (2013).

[18] E. Verbanis, A. Martin, D. Rosset, C. C. W. Lim, R. T. Thew, and H. Zbinden, Resource-Efficient Measurement Device Independent Entanglement Witness, Phys. Rev. Lett. **116**, 190501 (2016).

[19] D. Mayers and A. Yao, Self-testing quantum apparatus, Quant. Inf. Comput. **4**, 273 (2004).

[20] T. H. Yang and M. Navascués, Robust self-testing of unknown quantum systems into any entangled two-qubit states, Phys. Rev. A **87**, 050102 (2013).

[21] C. Bamps and S. Pironio, Sum-of-squares decompositions for a family of Clauser-Horne-Shimony-Holt-like inequalities and their application to self-testing, Phys. Rev. A **91**, 052111 (2015).

[22] I. Šupić, A. Coladangelo, R. Augusiak, and A. Acín, Self-testing multipartite entangled states through projections onto two systems, New J. Phys. **20**, 083041 (2018).

[23] M. McKague, T. H. Yang, and V. Scarani, Robust self-testing of the singlet, J. Phys. A: Math. Theor. **45**, 455304 (2014).

[24] M. McKague, Self-testing in parallel with CHSH, Quantum, **1**, 1 (2017).

[25] X. Wu, J. D. Bancal, M. McKague, and V. Scarani, Device-independent parallel self-testing of two singlets, Phys. Rev. A **93**, 062121 (2016).

[26] M. McKague, Self-testing in parallel, New J. Phys. **18**, 045013 (2016).

[27] A. Coladangelo, K. T. Goh, and V. Scarani, All pure bipartite entangled states can be self-tested, Nat. Commun. **8**, 15485 (2017).

[28] M. McKague and M. Mosca, *Generalized Self-testing and the Security of the 6-State Protocol Theory of Quantum Computation, Communication, and Cryptography: 5th Conference* (Lecture Notes in Computer Sciences Vol. 6519), pp. 113–130 (Springer, 2011).

[29] J. Kaniewski, Self-testing of binary observables based on commutation, Phys. Rev. A **95**, 062323 (2017).

[30] J. Bowles, I. Šupić, D. Cavalcanti, and A. Acín, Device-independent entanglement certification of all entangled states, Phys. Rev. Lett. **121**, 180503 (2018).

[31] A. Coladangelo, A. Grilo, S. Jeffery, and T. Vidick, Verifier-on-a-leash: New schemes for verifiable delegated quantum computation, with quasilinear resources, arXiv:1708.07359.

[32] M. McKague, Quantum information processing with adversarial devices, Ph.D. Thesis, University of Waterloo, Ontario, Canada, arXiv:1006.2352.

[33] J. Kaniewski, Analytic and (nearly) Optimal Self-Testing Bounds for the Clauser-Holt-Shimony-Horne and Mermin Inequalities, Phys. Rev. Lett. **117**, 070402 (2016).

[34] A. Acín, S. Pironio, T. Vértesi, and P. Wittek, Optimal randomness certification from one entangled bit, Phys. Rev. A **93**, 040102 (2016).

[35] O. Andersson, P. Badziag, I. Bengtsson, I. Dumitru, and A. Cabello, Self-testing properties of Gisin's elegant Bell inequality, Phys. Rev. A **96**, 032119 (2017).

[36] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, Phys. Rev. Lett. **23**, 880 (1969).

[37] A. Coladangelo, Parallel self-testing of (tilted) EPR pairs via copies of (tilted) CHSH, Quantum Inf. Comput. **17**, 831 (2017).

[38] M. Coudron and A. Natarajan, The parallel-repeated magic square game is rigid, arXiv:1609.06306.

[39] F. Dupuis, O. Fawzi, and R. Renner, Entropy accumulation, arXiv:1607.01796.

[40] C. Ci W. Lim, C. Portmann, M. Tomamichel, R. Renner, and N. Gisin, Device-Independent Quantum Key Distribution with Local Bell Test, Phys. Rev. X **3**, 031006 (2013).

[41] H.-K. Lo, M. Curty, and B. Qi, Measurement-Device-Independent Quantum Key Distribution, Phys. Rev. Lett. **108**, 130503 (2012).

[42] S. L. Braunstein and S. Pirandola, Side-Channel-Free Quantum Key Distribution, Phys. Rev. Lett. **108**, 130502 (2012).

[43] I. Šupić, R. Augusiak, A. Salavrakos, and A. Acín, Self-testing protocols based on the chained Bell inequalities, New J. Phys. **18**, 035013 (2016).

[44] A. Rastegin, Relations for certain symmetric norms and anti-norms before and after partial trace, J. Stat. Phys. **148**, 1040 (2012).

[45] L. J. Rogers, *An Extension of a Certain Theorem in Inequalities* (Messenger of Mathematics, New Series, 1888), XVII (10), pp. 145–150.

[46] O. Holder, Ueber einen Mittelwertsatz, Nachrichten von der Königl, Gesellschaft der Wissenschaften und der Georg-Augusts-Universität zu Göttingen, Band (in German) **2**, 38 (1889).

[47] A. Acín, N. Gisin, and B. Toner, Grothendieck's constant and local models for noisy entangled quantum states, Phys. Rev. A **73**, 062105 (2006).

[48] F. Hirsch, M. T. Quintino, T. Vértesi, M. Navascués, and N. Brunner, Better local hidden variable models for two-qubit

Werner states and an upper bound on the Grothendieck constant, Quantum **1**, 3 (2017).

[49] H. S. Poh, S. K. Joshi, and A. Cerè, Adán Cabello and Christian Kurtsiefer, Approaching Tsirelson's Bound in a Photon Pair Experiment, Phys. Rev. Lett. **115**, 180408 (2015).

[50] S. Gómez, A. Mattar, E. S. Gómez, D. Cavalcanti, O. Jiménez Farías, A. Acín, and G. Lima, Experimental nonlocality-based randomness generation with nonprojective measurements, Phys. Rev. A **97**, 040102 (2018).