Certifying the purity of quantum states with temporal correlations

Cornelia Spee

Institute for Quantum Optics and Quantum Information (IQOQI), Austrian Academy of Sciences, Boltzmanngasse 3, 1090 Vienna, Austria and Naturwissenschaftlich-Technische Fakultät, Universität Siegen, Walter-Flex-Straße 3, 57068 Siegen, Germany

(Received 9 October 2019; revised 7 May 2020; accepted 11 June 2020; published 17 July 2020)

Correlations obtained from sequences of measurements have been employed to distinguish among different physical theories or to witness the dimension of a system. In this work, we show that they can also be used to establish semi-device-independent lower bounds on the purity of the initial quantum state or even on one of the postmeasurement states. For single systems, this provides information on the quality of the preparation procedures of pure states or the implementation of measurements with anticipated pure postmeasurement states. For joint systems, one can combine our bound with results from entanglement theory to infer an upper bound on the concurrence based on the temporal correlations observed on a subsystem.

DOI: 10.1103/PhysRevA.102.012420

I. INTRODUCTION

Many applications in quantum information theory such as teleportation [1] and measurement-based quantum computation [2] use as a resource pure entangled states. That is, ideally, the corresponding protocols are applied to the respective pure resource state and deviations from this resource may result in errors [1,3] and lead to the need of entanglement purification [4–6] or fault-tolerant implementations (see, e.g., [3,7,8]) if one also takes into account imperfections after the preparation procedure. Due to interactions with the environment, in experiments often mixed states are prepared instead of the desired pure state. By knowing how much the prepared state differs from a pure state, one obtains some intuition on the quality of the preparation process without using full tomography. However, it should be noted that the purity only provides information about how much the prepared state deviates from a pure one (which might not necessarily be the desired one).

The purity of a quantum state can be quantified via

$$\mathcal{P}(\varrho) = \mathrm{tr}[(\varrho)^2]. \tag{1}$$

The purity attains its maximal value of 1 for pure states and its minimal value of $1/d$ for the maximally mixed state for d-dimensional systems. It is related to the linear entropy $S_L(\varrho) = 1 - \mathcal{P}(\varrho)$ and the Renyi-2 entropy [9] $\mathcal{H}_2(\varrho) = -\log_2[\mathcal{P}(\varrho)]$. The purity (or nonuniformity) of quantum states has been also studied from a resource-theoretic point of view [10–12]. Moreover, the task of distilling local pure states via a subclass of local operations and classical communication has been considered [13,14].

It is well known that the purity (of subsystems) of bipartite systems and their entanglement are connected. States for which the purity of the whole system is sufficiently small have to be separable, as there exists a set containing only separable states around the maximally mixed state which has a finite volume [15,16]. For two-qubit pure states, any entanglement measure can be written as a function of the purity of one of its subsystems, as in this case the purity uniquely determines the set of Schmidt coefficients and any entanglement measure for bipartite pure states is a function of the Schmidt coefficients [17]. The optimal strategy to estimate the entanglement of an unknown two-qubit pure state from n copies of this state has been shown to correspond to the estimation of the purity of the single-qubit reduced state and an explicit optimal protocol to do so has been proposed [18]. Therefore, this scheme only requires local measurements of one of the parties (but which act nonlocally on the different copies). Moreover, in the asymptotic regime, separable measurements of one of the parties assisted by classical communication among the copies can be shown to perform optimally [19].

For mixed (or higher-dimensional) states, the relation among entanglement and subsystem purity is no longer a one-to-one correspondence; however, for example, lower [20] and upper [21] bounds based on the purity of a subsystem and total system have been shown for the concurrence $C(\varrho)$ [22,23], which is an entanglement measure. In particular, it has been shown that [20]

$$\max_{X\in\{A,B\}} 2\{\mathrm{tr}[(\varrho)^2] - \mathrm{tr}[(\varrho_X)^2]\} \leqslant [C(\varrho)]^2 \tag{2}$$

and [21]

$$[C(\varrho)]^2 \leqslant \min_{X\in\{A,B\}} 2\{1 - \mathrm{tr}[(\varrho_X)^2]\}, \tag{3}$$

where ϱ_X is the reduced state of subsystem X. The first bound captures quantitatively the observation that only for entangled states, the reduced states can be more mixed than the state of the whole system [24]. The upper and lower bound on the concurrence given above can be determined in an experiment by measuring local observables using two identical copies of the state ϱ [20,21]. The purity (or Renyi-n entropies) of a system can also be experimentally measured by employing two copies of the state (see, e.g., [25–29] and references therein) or by performing randomized measurements [30–33]. It has been shown that if one uses two copies, a nonlocal unitary among them, and a local two-outcome measurement on only one of the copies but no ancilla or randomized measurements, it is only possible to extract the purity in case the dimension is odd [34]. The task of discriminating pure and mixed states has been considered [35,36], which also leads to schemes to estimate the purity. These are either based on maximum confidence discrimination [35] or an uncertainty relation [36] and require nonlocal measurements among the copies or control over the measured observable. Moreover, measurement schemes that allow one to determine the purity of single-mode Gaussian states have been proposed [37] and the relation among the (global and local) purities and entanglement of Gaussian states has been studied [38].

By performing tomography on the system, one could reconstruct the state and calculate the purity of the system. In particular, there exist adaptive schemes which do not rely on any assumption on the states [39,40] or which are designed for pure states and in which the assumption of purity can be certified from the observed data [41,42]. However, it should be noted that as in any tomographic approach, the measurements are required to be characterized (at least to some extent). The relation of the scaling of the accuracy in device-dependent adaptive process tomography and the purity of the measured state has been studied [43].

Device-independent bounds on the linear entropy (of the total system) or the concurrence can also be obtained from the value of violation of a Bell inequality [44,45]. Moreover, device-independent entropy witnesses based on dimension witnesses have been proposed in the context of prepare-and-measure scenarios [46] and sector lengths which are related to the average purity of reduced states have been studied (see, e.g., [47–49] and references therein).

Here we propose to use the temporal correlations obtained from sequences of measurements on a single copy to deduce a semi-device-independent lower bound on the purity. This approach relies only on the assumption of the dimension of the measured (sub)system and that measurements can be repeated (see below for more details). Note that even though, for a single-qubit system, less measurements are required in a tomographic approach than in our approach, such schemes require knowledge about the measurements that are implemented. Moreover, our approach does not require one to prepare two identical copies of the state at the same time and to act nonlocally on the subsystems of different copies [50]. It is straightforward to see from the equations above that a lower bound on the purity of a (sub)system provides an upper bound on the linear entropy or the concurrence. Moreover, it has been shown that a lower bound on the purity implies a lower bound on the accessible information [51].

Our approach uses sequential measurements and is conceptually different from the ones previously studied. In particular, we can also give a lower bound on the maximal purity of the postmeasurement state at the second time step for one of the outcomes provided the purity of the initial state is known.

II. USING TEMPORAL CORRELATIONS TO OBTAIN A LOWER BOUND ON THE PURITY

We will consider, in the following, sequences of general measurements acting on a single (sub)system whose (reduced) state is ϱ_{in} (see Fig. 1). To be more precise, we will examine the correlations $p(ab|xy)$ which correspond to the probability

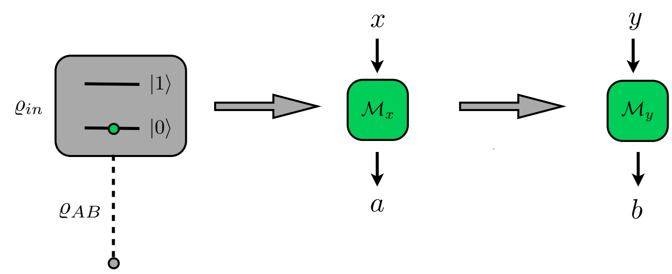

FIG. 1. This figure shows schematically the scenario considered here. Sequences of measurements are performed on the qubit state ϱ_{in}, which may correspond to the reduced density matrix of ϱ_{AB}, which in turn describes a composite system. One observes temporal correlations $p(ab|xy)$ with measurement outcomes denoted by a, b and measurement settings by x, y.

for obtaining outcome a in a first time step if one performs measurement x, and then observing outcome b in a second time step if measurement y is performed.

We will assume that one can use the same measurement apparatus at different time steps and the labeling of measurement settings does not change, e.g., in case $x = y$, one performs the same measurement twice, however, the outcomes do not need to be the same. The only further assumption will be that in the following, the (sub)system that is measured is a two-dimensional system. In particular, we will not restrict the type of measurements, i.e., arbitrary instruments [52] are allowed. This scenario has also been considered in [53–55].

We consider the following quantity:

$$\mathcal{B}_1 = p(++|00) + p(++|11) + p(+-|01) + p(+-|10).$$

It has been shown in [53] that one can provide a (nontrivial) upper bound on $\mathcal{B}_1$ for general measurements on a qubit, which allows one to employ $\mathcal{B}_1$ as a dimension witness. As we show here, it can also be used to witness the purity.

It can be proven that for any choice of measurements, the maximum will be attained for pure initial and postmeasurement states and that the maximum attainable value for fixed purity of the initial state will be monotonically increasing with increasing purity (see Appendix B). These relations are the key to use temporal correlations for obtaining lower bounds on the purity. In particular, it implies that in order to observe a certain value, the system has to have at least a certain amount of purity. In Appendix A, we will show that this key idea can, in principle, also be used to obtain lower bounds on the purity of the initial state for higher-dimensional systems and that it is essentially possible to employ $\mathcal{B}_1$ for this purpose. More precisely, we show for arbitrary finite dimension that the maximal attainable value of any linear function of correlations of two time steps is a monotonically increasing function of the initial purity and that the maximal value of $\mathcal{B}_1$ for arbitrary measurements is not constant as a function of the purity.

For a qubit, here we provide the explicit (analytic) relation between the maximal attainable value of $\mathcal{B}_1$ and the purity. In order to ease the notation (and as it appears naturally in the derivations), we will from here on mainly refer to the length of the Bloch vector instead of the purity. That is, we will use the Bloch decomposition for

$$\varrho_{\text{in}} = \tfrac{1}{2}(\mathbb{1} + p\,\vec{\alpha}_{\text{in}} \cdot \vec{\sigma}), \tag{4}$$

with $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$, σ_i being the Pauli matrices, $0 \leqslant p \leqslant 1$, $\vec{\alpha}_{\text{in}} \in \boldsymbol{R}^3$, and $|\vec{\alpha}_{\text{in}}| = 1$. With this, p is the length of the Bloch vector and the purity of the initial state is given by $\mathcal{P}(\varrho_{\text{in}}) = 1/2(1 + p^2)$. Note that the purity $\mathcal{P}$ is monotonically increasing as a function of p (and vice versa). Let us then denote by $B_1(p)$ the maximal attainable value of $\mathcal{B}_1$ for a given length of the Bloch vector, $p = \sqrt{2\mathcal{P} - 1}$, of the initial state ϱ_{in} and arbitrary choice of measurements. Then it holds that

$$B_1(p) = 1/2(5 + p). \tag{5}$$

This relation follows from Theorem 1 and we will discuss below how to derive it from this theorem.

The measurements that attain the maximum of $\mathcal{B}_1$ are the same, independent of the purity. In particular, the following protocol allows one to attain $B_1(p)$. One of the measurements announces deterministically the outcome "+" and then prepares the state $1/2(\mathbb{1} - \vec{\alpha}_{\text{in}} \cdot \vec{\sigma})$. The other measurement measures the observable $\vec{\alpha}_{\text{in}} \cdot \vec{\sigma}$.

If one obtains in an experiment a value for $\mathcal{B}_1$, denoted here and in the following by $\mathcal{B}_1^{\text{exp}}$, one can straightforwardly deduce a lower bound on the purity of the measured initial state. This is due to the fact that $B_1(p)$ is a monotonically increasing function of $\mathcal{P}$ [see Eq. (5)] and $B_1(p) \geqslant \mathcal{B}_1^{\text{exp}}$ if the purity of the initial state that is measured in the experiment is given by $\mathcal{P} = 1/2(1 + p^2)$. The last relation captures that in an experiment, the measurements that are implemented do not need to be the optimal ones that allow one to attain $\mathcal{B}_1(p)$. With this, one obtains that in order to observe $\mathcal{B}_1^{\text{exp}}$, a certain amount of purity is required. In particular, we obtain the following observation:

Observation 1. Let $\mathcal{B}_1^{\text{exp}}$ be the value for $\mathcal{B}_1$ obtained in an experiment by performing sequences of measurements on the state ϱ_{in}. Then it holds, for the purity $\mathcal{P}$ of ϱ_{in}, that

$$\mathcal{P} \geqslant \frac{(2\,\mathcal{B}_1^{\text{exp}} - 5)^2 + 1}{2}. \tag{6}$$

Hence, temporal correlations allow one to witness the initial purity.

Knowing the purity of the initial state, it is also possible to deduce a lower bound on the maximal purity of the postmeasurement state occurring at the second time step for outcome "+". To be more precise, one can provide a lower bound on the state measured in the second time step, which here and in the following we will refer to as the postmeasurement state. Let p be the length of the Bloch vector of the initial state ϱ_{in} and $w_{+|i}$ the one of the postmeasurement state that is obtained after performing measurement $i \in \{0, 1\}$ on ϱ_{in} and observing outcome "+." Then, one can determine the maximum $B_1(p, w_{+|0}, w_{+|1})$ that is attainable with all measurements and states that respect the imposed purities. One can show that $B_1(p, w_{+|0}, w_{+|1})$ is monotonically increasing as a function of $\mathcal{W}_{+|i} = 1/2(1 + w_{+|i}^2)$ (assuming the other purities fixed but arbitrary). Moreover, in an experiment leading to $\mathcal{B}_1^{\text{exp}}$ in which the states occur with the respective purities, it might be that one deviates from the optimal protocol. Hence, it holds

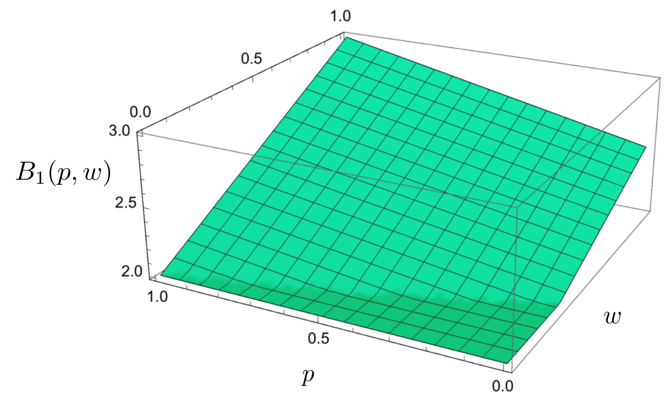

FIG. 2. This figure shows the maximal attainable value of $\mathcal{B}_1$ as function of a given Bloch vector length of the initial state [i.e., purity $\mathcal{P} = 1/2(1 + p^2)$] and given Bloch vector length of the postmeasurement states for both measurements corresponding to outcome "+" [i.e., purity $\mathcal{W} = 1/2(1 + w^2)$]; see Theorem 1.

for $w_{\max} = \max_{i\in\{0,1\}} w_{+|i}$ that

$$B_1(p, w_{\max}, w_{\max}) \geqslant B_1(p, w_{+|0}, w_{+|1}) \geqslant \mathcal{B}_1^{\text{exp}}.$$

It only remains to determine $B_1(p, w) \equiv B_1(p, w, w)$ to provide an explicit lower bound on the maximal purity of the postmeasurement states of outcome "+" depending on the purity of the input state. In the following theorem, we provide a closed formula for $B_1(p, w)$ (see, also, Fig. 2).

Theorem 1. Let $\mathcal{P}$ be the purity of the initial state and $\mathcal{W}$ the purity of the postmeasurement states that occur for measurement $i \in \{0, 1\}$ observing outcome "+." Then, for a two-dimensional system, the maximal value of $\mathcal{B}_1$, $B_1(p, w)$, that can be obtained for arbitrary initial states and measurements that respect these constraints on the purities, is given by

$$B_1(p, w) = \begin{cases} 2, & 0 \leqslant w \leqslant \frac{1-p}{3+p} \\ 1 + \frac{1+w}{2} + \frac{(1+p)(1+w)}{4}, & \frac{1-p}{3+p} < w \leqslant 1, \end{cases}$$

where $w = \sqrt{2\mathcal{W} - 1}$ and $p = \sqrt{2\mathcal{P} - 1}$ are the length of the Bloch vector for the respective purity.

The proof of this theorem can be found in Appendix B.

This theorem allows one to deduce a lower bound on the maximal purity of the postmeasurement states provided that the purity of the initial state is known. In particular, we have that for $\mathcal{B}_1^{\text{exp}} \leqslant 2$, we cannot deduce a lower bound; however, if $\mathcal{B}_1^{\text{exp}} > 2$, it follows from the theorem above and $B_1(p, w_{\max}) \geqslant \mathcal{B}_1^{\text{exp}}$ that

$$\mathcal{W}_{\max} \geqslant \frac{14 + 4(\mathcal{B}_1^{\text{exp}})^2 + \mathcal{P} + 5\sqrt{2\,\mathcal{P} - 1}}{4 + \mathcal{P} + 3\sqrt{2\,\mathcal{P} - 1}} - \frac{2\,\mathcal{B}_1^{\text{exp}}(7 + \sqrt{2\,\mathcal{P} - 1})}{4 + \mathcal{P} + 3\sqrt{2\,\mathcal{P} - 1}}. \tag{7}$$

Moreover, note that as $B_1(p, w)$ is monotonically increasing as a function of $\mathcal{W}$, we also have that

$$B_1(p) = \max_{0\leqslant w\leqslant 1} B_1(p, w) = B_1(p, 1) = B_1(p) = \frac{5 + p}{2}, \tag{8}$$

which allows one to bound the purity of the initial state as argued above [see Eq. (5) and Observation 1].

In an experiment, one may not be able to perfectly implement the measurements that realize the tight bound in Eq. (5), but one may have postmeasurement states which are not perfectly pure. Theorem 1 allows one to deduce how robust the estimation of the initial purity is with respect to not perfectly pure postmeasurement states as it also specifies for this case the maximal attainable value. Consider the case that the maximal length of the Bloch vectors of the postmeasurement state is given by $1-\epsilon$. Then, it holds that $B_1(p, 1-\epsilon) = B_1(p) - \frac{3+p}{4}\epsilon$, i.e., the deviation is, at most, ϵ and also linear in this parameter. Note further that in case one can estimate a lower bound on the purity of the postmeasurement state, one can also use Theorem 1 to improve the lower bound on the purity of the initial state (by using the lower bound for the respective w in Theorem 1). As this requires information about the measurement, the so-obtained bound is no longer semi-device independent.

III. UPPER BOUND ON THE CONCURRENCE BASED ON THE PURITY

As mentioned before, it is well known that for bipartite pure states, there is a close connection between entanglement and the purity of the reduced state of a single party. In particular, the reduced state is pure only for product states, whereas for maximally entangled states, it is maximally mixed. For mixed states and on a more quantitative level, entanglement measures such as the concurrence are defined as the convex roof extension of a function of the local purity. More precisely, the concurrence [22,23] is given by

$$C(\rho) = \inf \sum_i q_i C(|\psi_i\rangle), \tag{9}$$

where the infimum is taken over all pure state decompositions, $\rho = \sum_i q_i |\psi_i\rangle\langle\psi_i|$, and $C(|\psi_i\rangle) = \sqrt{2\{1-\mathrm{tr}[(\rho_A^i)^2]\}}$ with $\rho_A^i = \mathrm{tr}_B(|\psi_i\rangle\langle\psi_i|)$. It seems therefore natural to consider the relation among the concurrence and the purity of the reduced state more closely in order to obtain a bound on the concurrence. The following result will allow us to provide an upper bound on the concurrence based on the observed temporal correlations. For two-qubit states ϱ_{AB} with $\rho_A = \mathrm{tr}_B(\varrho_{AB})$ and $\rho_B = \mathrm{tr}_A(\varrho_{AB})$, it holds that [21]

$$C(\varrho) \leqslant \min_{X\in\{A,B\}} \sqrt{2\{1-\mathrm{tr}[(\varrho_X)^2]\}}. \tag{10}$$

This bound has already been observed for arbitrary bipartite d-dimensional states in [21]. For completeness, we will nevertheless present in Appendix C an (alternative, but similar) proof for two-qubit states. Combining this with the lower bound on the purity based on temporal correlations (see Observation 1), we can state the following observation:

Observation 2. Let ϱ_{AB} be a two-qubit state and $\mathcal{B}_1^{\mathrm{exp}}$ the experimental value for $\mathcal{B}_1$ obtained for sequences of measurements on one of the subsystems. Then it holds for the concurrence $C(\varrho_{AB})$ that

$$C(\varrho_{AB}) \leqslant \sqrt{1-\left(2\,\mathcal{B}_1^{\mathrm{exp}}-5\right)^2}. \tag{11}$$

Moreover, it has also been shown in [21] that for multipartite states, $C(\varrho) \leqslant 2^{1-n/2}\sqrt{2^n-2-\sum_i \mathrm{tr}[(\varrho_i)^2]}$. Here, $C(\varrho)$ is a generalization of the concurrence to the multipartite case defined by $C(\psi) = 2^{1-n/2}\sqrt{2^n-2-\sum_i \mathrm{tr}[(\varrho_i)^2]}$ [56,57], where n is the number of parties, ϱ_i are the single-party density matrices, and $C(\varrho)$ is obtained via the convex roof extension from $C(\psi)$ [see Eq. (9)].

Hence, for a multipartite system also, one can first obtain, from the correlations that arise from sequences of local measurements on subsystems, a semi-device-independent lower bound on the purity of the subsystems and, with this, then an upper bound on the concurrence of the joint system.

IV. SUMMARY AND OUTLOOK

In this work, we considered sequential measurements on a qubit. We showed that one can deduce from the observed correlations a lower bound on the purity of the initial state of the qubit. In case the qubit is part of a two-qubit system, this provides an upper bound on the concurrence. Moreover, provided that the purity of the initial state is known, our approach allows one to obtain a lower bound on the maximal purity of the postmeasurement states occurring at the second time step for one of the outcomes. Our result shows that it is possible to use temporal correlations for bounds on the purity and the concurrence by explicitly considering the example of a qubit. Moreover, we proved that also for higher-dimensional systems, it is essentially possible to employ temporal correlations in order to establish bounds on the initial purity. It would be relevant to pursue our investigation of higher-dimensional systems and provide explicit purity witnesses. Moreover, it would be interesting to see whether longer sequences allow, in principle, for a better performance as has been observed for the case of dimension witnesses [55].

ACKNOWLEDGMENTS

I thank Otfried Gühne, Costantino Budroni, and Nikolai Wyderka for useful discussions. Moreover, I would like to thank Otfried Gühne for careful reading of the manuscript. This work has been supported by the Austrian Science Fund (FWF): J 4258-N27, the ERC (Consolidator Grant No. 683107/TempoQ), and the DAAD (Project No. 57445566).

APPENDIX A: TEMPORAL CORRELATIONS ALLOW TO WITNESS THE PURITY FOR *d*-DIMENSIONAL SYSTEMS

In this Appendix, we show that it is essentially possible to use temporal correlations for providing lower bounds on the purity for d-dimensional systems. In particular, we will prove that one can construct functions of the correlations whose maximum for arbitrary measurements is monotonically increasing as a function of the purity of the initial state. Hence, in order to observe a certain value of this function, one has to have a certain amount of purity and, therefore, these can be used to provide lower bounds on the purity. Moreover, we will show that in principle, the quantity $\mathcal{B}_1$ could also be used to gain information about the purity for d-dimensional systems.

Proof. In order to do so, we consider a quantity $\mathcal{R} = \sum \alpha_{abxy} p(ab|xy)$, which is linear in the correlations and therefore also in the initial state. More precisely, it holds that $p(ab|xy) = p(a|x)p(b|axy) = \mathrm{tr}(\mathcal{E}_{a|x}\varrho_{\mathrm{in}})p(b|axy)$, with $\mathcal{E}_{a|x}$ being the effect for the measurement in the first time step. As we will here and in the following restrict to two time steps and furthermore be interested in the optimal measurements, it is possible to use as a description of the measurements their effects and postmeasurement states (instead of describing them via instruments). The effects $\mathcal{E}_{a|x}$ are positive semidefinite matrices with the property that $\sum_a \mathcal{E}_{a|x} = \mathbb{1}$. They allow one to calculate the probability for obtaining outcome "a" by implementing measurement x on a state ρ via $\mathrm{tr}(\rho\mathcal{E}_{a|x})$. Note that for any combination of effects and postmeasurement states, there exists a valid instrument that realizes them when being applied to an arbitrary ϱ_{in}. This can be achieved by first applying an instrument with the desired effects and then preparing the system in the desired postmeasurement state. Note that in case more time steps are considered, not any sequence of postmeasurement states is possible. However, note that by considering the description in terms of instruments and the Heisenberg picture, the following argumentation can be straightforwardly generalized to longer sequences.

Now let us assume that for a given purity of the initial state $\mathcal{P}$, one knows the optimal protocol that maximizes $\mathcal{R}$. We then have $\mathcal{P} = \sum_i q_i^2$, with q_i being the eigenvalues of the optimal ϱ_{in}. Let us denote the eigenbasis for ϱ_{in} by $\{|i\rangle\}$, the corresponding optimal effects by $\tilde{\mathcal{E}}_{a|x}$, and the maximal attainable value by $R(\mathcal{P})$. Then, one obtains

$$R(\mathcal{P}) = \sum_{i,a,x,b,y} q_i \alpha_{abxy} \mathrm{tr}(\tilde{\mathcal{E}}_{a|x}|i\rangle\langle i|)p(b|axy). \qquad \text{(A1)}$$

Note that there always exists a state $|k\rangle$ in the eigenbasis for which

$$\sum_{a,x,b,y} \alpha_{abxy} \mathrm{tr}(\tilde{\mathcal{E}}_{a|x}|k\rangle\langle k|)p(b|axy) \geqslant \sum_{a,x,b,y} \alpha_{abxy} \mathrm{tr}(\tilde{\mathcal{E}}_{a|x}|j\rangle\langle j|)p(b|axy) \qquad \text{(A2)}$$

for all $|j\rangle$. Note further that as we assume the optimal strategy, one can choose, without loss of generality, that $q_k \geqslant q_j$. This is due to the fact that if $q_k < q_j$ for some $|j\rangle$ for which the inequality in Eq. (A2) is strict, one could apply a unitary exchanging $|k\rangle$ and $|j\rangle$ before and after the supposedly optimal measurements, and obtain a higher value for $\mathcal{R}$, which contradicts our assumption that we are implementing the optimal protocol. In case the inequality is an equality, we can simply relabel $|k\rangle$ to obtain $q_k \geqslant q_j$.

It then remains to show that by increasing the purity, it is possible to increase the maximal value of $\mathcal{R}$. Let us first consider the case that the inequality in Eq. (A2) is strict for at least one $|j\rangle$ with $q_j \neq 0$, which we will denote by $|l\rangle$. Then, for any purity $\mathcal{Q} > \mathcal{P}$, one can find a value $\epsilon > 0$ such that $\mathcal{Q} = (q_k + \epsilon)^2 + (q_l - \epsilon)^2 + \sum_{i \neq k,l} q_i^2$, i.e., $\epsilon = \frac{q_l - q_k + \sqrt{2(\mathcal{Q}-\mathcal{P}) + (q_k - q_l)^2}}{2}$. We will then use the notation $\tilde{q}_i = q_i$ for $i \notin \{k, l\}$, $\tilde{q}_k = q_k + \epsilon$, and $\tilde{q}_l = q_l - \epsilon$. It is then straightforward to see that choosing the same effects and imposing the same postmeasurement states as before (e.g., by considering some measure-and-prepare channel), one obtains

$$\sum_{a,x,b,y,i} \tilde{q}_i \alpha_{abxy} \mathrm{tr}(\tilde{\mathcal{E}}_{a|x}|i\rangle\langle i|)p(b|axy) > R(\mathcal{P}). \qquad \text{(A3)}$$

Hence, we have shown that for any purity $\mathcal{Q} > \mathcal{P}$, there exists a strategy (measurements) that allows one to exceed the maximal attainable value $R(\mathcal{P})$ in case inequality (A2) is strict for at least one $|j\rangle$ for which $q_j \neq 0$.

Note that in case the inequality (A2) is an equality, for some set $j \in \mathcal{J} = \{j_1, \ldots, j_n\}$ and $q_i = 0$ for $i \notin \mathcal{J}$, it can be easily seen that this implies that for any purity which can be realized with a density matrix of rank n or smaller, the value $R(\mathcal{P})$ can be attained. Note further that this implies that for any purity $\mathcal{Q}$ which corresponds to a density matrix ρ_n of rank n and $\mathcal{Q} < \mathcal{P}$, the optimal strategy has the property that for the whole eigenbasis (with nonzero eigenvalue) of ρ_n, one obtains an equality as otherwise the maximal attainable value of $\mathcal{R}$ has to strictly increase with increasing purity, as we just have shown. However, this implies that also with higher purity, this value is attainable and therefore it has to hold that in this case, we have that $R(\mathcal{Q}) = R(\mathcal{P})$. Increasing now $\mathcal{P}$, one obtains that the maximal attainable value is at least $R(\mathcal{P})$ and either remains constant or increases by the argumentation given before.

In summary, we have that for both cases, the maximal attainable value of $\mathcal{R}$ is monotonically increasing as a function of the purity. Hence, in case it is not constant, for all purities, $\mathcal{R}$ can be used to obtain some information on the purity. It is obvious that for any dimension d, there exists some quantity $\mathcal{R}$ whose maximum for given purity does not remain constant for all purities.

As an example, consider $\mathcal{B}_1$ for which one can show that for the maximally mixed state, the maximum is upper bounded by $\max[3, 4(1 - 1/d)]$ but the maximal value for a pure state corresponds to 4 for $d \geqslant 3$. This implies that also for higher dimensions, the maximal attainable value of $\mathcal{B}_1$ is not constant as a function of the purity. In order to see the upper bound on $\mathcal{B}_1$ for the maximally mixed state, note first that one can use an analogous argumentation as before to show that for fixed purity of the initial state, the maximal attainable value of $\mathcal{B}_1$ is monotonically increasing as a function of the purity of the postmeasurement states. Hence, the optimal postmeasurement states are either pure or can be chosen to be pure. As then, there are only two pure postmeasurement states appearing in the quantity; this implies that only a two-dimensional subspace is relevant for the measurements in the second time step. Moreover, considering the first time step, it is then straightforward to see that in order to obtain the maximum, the diagonal terms in the effects for outcome "+" should be one in the orthogonal complement to this subspace and terms mixing the qubit subspace and its complement are chosen to be zero (in order to not introduce further constraints on the two-dimensional subspace due to positivity). We then use that one can parametrize the restriction of the effects to the two-dimensional subspace and the states as in [53]. That is, one can use, for such effects, the parametrization $\mathcal{E}_{+|i} = a_i(\mathbb{1}_2 + b_i\vec{\sigma} \cdot \vec{c}_i) + \mathbb{1}_{d-2}$, where $\vec{c}_i \in \boldsymbol{R}^3$, $|\vec{c}_i| = 1$, $0 \leqslant a_i \leqslant 1/(1 + b_i)$, $0 \leqslant b_i \leqslant 1$, $\mathbb{1}_x$ denotes the x-dimensional identity and $\vec{\sigma}$ ($\mathbb{1}_2$) the vector of Pauli matrices (the identity) in the qubit subspace, respectively. Using then that the initial state is maximally mixed, one can show, analogously to [53], that the maximum of $\mathcal{B}_1$ is smaller or equal to $\max[3, 4(1 - 1/d)]$ or it is attained when the effects are projective. More precisely, consider first the points where the derivative with respect to a_0 (assuming all other parameters to be fixed but arbitrary) vanishes, i.e.,

$$\begin{aligned} a_0 \frac{d\,\mathcal{B}_1}{da_0} = &\left[p(+|0) - \frac{d-2}{d}\right][p(+|+00) + p(-|+01)] \\ &+ p(+|1)[p(-|+10) - 1] \\ &+ p(+|0)p(+|+00) = 0, \end{aligned}$$

where we used $p(+b|xy) = p(+|x)p(b|+xy)$. Hence, at the points where the derivative vanishes, it holds that

$$\begin{aligned} p(+|0)&[p(+|+00) + p(-|+01)] + p(+|1)[p(-|+10) \\ &= p(+|1) + \frac{d-2}{d}[p(+|+00) + p(-|+01)] \\ &\quad - p(+|1)p(+|+00) \leqslant 3 - \frac{4}{d}, \end{aligned} \tag{A4}$$

and, therefore, $\mathcal{B}_1 \leqslant 4(1 - 1/d)$ at these points. It remains to consider the boundary $a_0 = 0$ and $a_0 = 1/(1 + b_0)$. It is easy to see that for $a_0 = 0$, it holds that $\mathcal{B}_1 \leqslant 2$. Before considering the case $a_0 = 1/(1 + b_0)$, note that the quantity $\mathcal{B}_1$ is symmetric with respect to the exchange of measurement setting 0 and 1. Hence, in order to possibly achieve a value for $\mathcal{B}_1$ that is greater than $4(1 - 1/d)$, measurements with $a_1 = 1/(1 + b_1)$ have to be used. As can be easily seen, the choice $a_i = 1/(1 + b_i)$ corresponds to measurements for which the effect corresponding to the outcome "−" is proportional to a projector, i.e.,

$$\mathcal{E}_{-|i} = \frac{u_i}{2}(\mathbb{1}_2 - \vec{\sigma} \cdot \vec{c}_i), \tag{A5}$$

with $0 \leqslant u_i \leqslant 1$. One can then use exactly the same argumentation as in Appendix C1 of [53] to show that either $\mathcal{B}_1 \leqslant 3$ or the effects have to be projective, i.e., one shows that at the points where the derivative with respect to u_i vanishes, it holds that $\mathcal{B}_1 \leqslant 3$ and that the same holds true for the boundary point $u_i = 0$. Considering then projective effects and the optimal choice of postmeasurement states for such effects as in Appendix C1 of [53], the quantity depends on one remaining parameter, i.e., the angle between the Bloch vectors in the restriction to the two-dimensional subspace of the two measurements. It is then straightforward to see that the maximum attainable value with projective effects is given by $4(1 - 1/d)$, which is obtained for $\vec{c}_0 = -\vec{c}_1$. In summary, we have shown that for the maximally mixed state, it is not possible to exceed $\max[3, 4(1 - 1/d)]$. In particular, for $d \geqslant 4$, we have that $4(1 - 1/d) \geqslant 3$ and, in this case, the bound can be reached. For pure initial states, one can attain a value of 4 in case $d \geqslant 3$ (see [53]). Hence, we have that the maximal attainable value $\mathcal{B}_1$ is not constant but, due to the argumentation above, at least on some interval(s), strictly increasing with increasing purity. This concludes the proof that temporal correlations can be used to build witnesses for the purity of d-dimensional states. ■

APPENDIX B: PROOF OF THEOREM 1

In this Appendix, we will first show that $B_1(p, w_{+|0}, w_{+|1})$ (as defined in the main text and below) is monotonically increasing as a function of $w_{+|1}$ (and, therefore, also $\mathcal{W}_{+|i}$). Moreover, we will prove Theorem 1.

Recall first that $B_1(p, w_{+|0}, w_{+|1})$ is the maximal value for $\mathcal{B}_1$ that is attainable with arbitrary (time-independent) measurements for a given purity $\mathcal{P} = 1/2(1 + p^2)$ of the initial state and fixed purity of the states that are measured at the second time step $\mathcal{W}_{+|i} = 1/2(1 + w_{+|i}^2)$ if, in the first time step measurement, i is performed and outcome "+" is obtained. Analogously to the proof in Appendix A, one can show in general that for two time steps, the maximum (with respect to all other parameters but the purities of the states which are assumed to be fixed apart from the purity of one postmeasurement state) of any linear function of temporal correlations is also a monotonically increasing function of the purity for the postmeasurement states. This implies that, in particular, $B_1(p, w_{+|0}, w_{+|1})$ is monotonically increasing as a function of $\mathcal{W}_{+|i}$ (assuming that all other parameters are fixed but arbitrary).

Alternatively, one can also show that $B_1(p, w_{+|0}, w_{+|1})$ is monotonically increasing as a function of $\mathcal{W}_{+|i}$ by considering the derivative with respect to $w_{+|i}$.

In order to do so, we parametrize the effects via $\mathcal{E}_{+|x} = r_x\mathbb{1} + q_x\vec{v}_x \cdot \vec{\sigma}$ and $\mathcal{E}_{-|x} = \mathbb{1} - \mathcal{E}_{+|x}$ for $x \in \{0, 1\}$ with $0 \leqslant q_x \leqslant r_x \leqslant 1 - q_x$, $\vec{v}_x \in \boldsymbol{R}^3$, $|\vec{v}_x| = 1$, and $\vec{\sigma}$ is a vector containing the Pauli matrices. As mentioned in the main text, one can use the Bloch decomposition to parametrize states with fixed purity, i.e.,

$$\varrho = 1/2(\mathbb{1} + w\vec{\alpha} \cdot \vec{\sigma}), \tag{B1}$$

with $|\vec{\alpha}| = 1$ and $0 \leqslant w \leqslant 1$. The purity $\mathcal{W}$ is then related to the length of the Bloch vector w via

$$\mathcal{W} = \tfrac{1}{2}(1 + w^2). \tag{B2}$$

Using this parametrization for the states, one can analogously, as in Appendix C of [53], determine the initial and postmeasurement states that maximize $\mathcal{B}_1$ for given effects and purities. More precisely, the Bloch vectors for the postmeasurement states of measurement $i \in \{0, 1\}$ are proportional to $(-1)^i(q_0\vec{v}_0 - q_1\vec{v}_1)$ and the Bloch vector for the initial state has to be chosen proportional to $q_0X_0\vec{v}_0 + q_1X_1\vec{v}_1$, with $X_0 = 1 + r_0 - r_1 + w_{+|0}\sqrt{q_0^2 + q_1^2 - 2q_0q_1\vec{v}_0 \cdot \vec{v}_1}$ and $X_1 = 1 + r_1 - r_0 + w_{+|1}\sqrt{q_0^2 + q_1^2 - 2q_0q_1\vec{v}_0 \cdot \vec{v}_1}$. This is due to the fact that the choice of $\vec{\alpha}$ which maximizes $\vec{\alpha} \cdot \vec{\beta}$ under the constraint that the length of $\vec{\alpha}$ is fixed is given by $\vec{\alpha} = c\vec{\beta}$, with $c \geqslant 0$ and c ensuring the correct length of the vector. For the optimal choice of states (and arbitrary effects), one observes that $\mathcal{B}_1$ is a linear function of p and one can show that $d\,\mathcal{B}_1\,/dw_{+|i} \geqslant 0$. Hence, $\mathcal{B}_1$ is monotonically increasing as a function of $w_{+|i}$, which implies that it is also a monotonically increasing function of $\mathcal{W}_{+|i}$. In particular, we have that

$$B_1(p, w_{\max}, w_{\max}) \geqslant B_1(p, w_{+|0}, w_{+|1}),$$

where $w_{\max} = \max_{i \in \{0,1\}} w_{+|i}$.

In the following, we will use the notation $B_1(p,w) \equiv B_1(p,w,w)$. We will next show Theorem 1. In order to improve readability, we repeat the theorem here.

Theorem 1. Let $\mathcal{P}$ be the purity of the initial state and $\mathcal{W}$ the purity of the postmeasurement states that occur for measurement $i \in \{0,1\}$ observing outcome "+." Then, for a two-dimensional system, the maximal value of $\mathcal{B}_1$, $B_1(p,w)$, that can be obtained for arbitrary initial states and measurements that respect these constraints on the purities, is given by

$$B_1(p,w) = \begin{cases} 2, & 0 \leqslant w \leqslant \frac{1-p}{3+p} \\ 1 + \frac{1+w}{2} + \frac{(1+p)(1+w)}{4}, & \frac{1-p}{3+p} < w \leqslant 1, \end{cases}$$

where $w = \sqrt{2\mathcal{W}-1}$ and $p = \sqrt{2\mathcal{P}-1}$ are the length of the Bloch vector for the respective purity.

Proof. Note first that by deterministically assigning outcome "+" for both measurements independent of the state that is measured, one obtains that $\mathcal{B}_1 = 2$. Moreover, by using the following protocol, one can obtain $\mathcal{B}_1 = 1 + \frac{1+w}{2} + \frac{(1+p)(1+w)}{4}$. Let the initial state have a Bloch vector of length p pointing in the z direction, i.e., in Eq. (B1), we have that $\vec{\alpha} = (0,0,1)$. One of the measurements is chosen to be of the form that one deterministically announces "+" and prepares the state with Bloch vector pointing in the $-z$ direction, i.e., $\vec{\alpha} = (0,0,-1)$, and of length w. The other measurement performs a projective measurement in the computational basis $1/2(\mathbb{1} \pm \sigma_z)$ with associated outcome "±" and then prepares the state with $\vec{\alpha} = (0,0,1)$ and length w. Hence, the values for $\mathcal{B}_1$ given in the theorem above are attainable. Moreover, note that $2 \geqslant 1 + \frac{1+w}{2} + \frac{(1+p)(1+w)}{4}$ if and only if $w \leqslant \frac{1-p}{3+p}$. It remains to show that $\mathcal{B}_1$ for given p and w cannot exceed $\max[2, 1 + \frac{1+w}{2} + \frac{(1+p)(1+w)}{4}]$.

In order to do so, we note first that it can be easily seen using the same argumentation [58] as in [53] that either $B_1(p,w) \leqslant 2$ or for both measurements the effects for outcome "−" are proportional to projectors. That is, one uses the parametrization

$$\mathcal{E}_{+|i} = a_i(\mathbb{1}_2 + b_i\vec{\sigma}\cdot\vec{c}_i) \tag{B3}$$

with $0 \leqslant a_i \leqslant 1/(1+b_i)$, $0 \leqslant a_i \leqslant 1$, and $\vec{c}_i$ being a real vector of unit length, and considers the critical points with respect to a_i (all other parameter are assumed to be fixed but arbitrary). One observes that at the points where the derivative vanishes, Eq. (A4) with $d = 2$ holds. Hence, at these points, $\mathcal{B}_1 \leqslant 2$. It remains to consider the boundary of the interval $0 \leqslant a_i \leqslant 1/(1+b_i)$. For $a_i = 0$, one also obtains $\mathcal{B}_1 \leqslant 2$. For $a_i = 1/(1+b_i)$, the effect corresponding to outcome "−" is proportional to a projector.

Using then the parametrization of effects as given in Eq. (A5) and considering the points where the gradient with respect to u_0 and u_1 (assuming again all other parameters to be fixed but arbitrary) vanishes, we obtain

$$\sum_i u_i \frac{\partial \mathcal{B}_1}{\partial u_i} = 0. \tag{B4}$$

This is equivalent to

$$\mathcal{B}_1 = \tfrac{1}{2}[p(+|0) + p(+|1) + p(+|+00) + p(-|+01) + p(+|+11) + p(-|+10)], \tag{B5}$$

where we used that one can write $p(+b|xy) = p(+|x)p(b|+xy)$. By maximizing the right-hand side of this equation, one can obtain an upper bound on $\mathcal{B}_1$ at the points where the gradient vanishes. Note first that the expression is a linear function in the parameters u_i and therefore is maximal at one of the boundary points $u_i = 0$ or $u_i = 1$. If $u_0 = u_1 = 0$, then, independent of the measured states, outcome "−" never occurs and therefore the right-hand side is upper bounded by two. In case $u_0 = u_1 = 1$, the effects are projective and by choosing the optimal initial and postmeasurement states analogous to [53], we get, for the right-hand side,

$$\tfrac{1}{4}[6 + 2w\sqrt{2-2x} + p\sqrt{2+2x}], \tag{B6}$$

where, here and in the following, x corresponds to the angle between the Bloch vectors of the effects for outcome "+" of the two measurements, i.e., $\vec{c}_0 \cdot \vec{c}_1 = x$. More precisely, one chooses the vector $\vec{\alpha}$ in Eq. (B1) for the initial state proportional to $\vec{c}_0 + \vec{c}_1$ and for the postmeasurement states proportional to $\pm(\vec{c}_0 - \vec{c}_1)$. This choice is optimal, as the maximal value of an expression of the form $\vec{\alpha}\cdot\vec{v}$ is attained if $\vec{\alpha}$ is chosen parallel to $\vec{v}$. One can then easily show that Eq. (B6) is maximized for the point where the derivative with respect to x vanishes (and not at the boundary given by $x \in \{\pm 1\}$), i.e., $x = (p^2 - 4w^2)/(p^2 + 4w^2)$. This results in a maximal value for the right-hand side that is strictly smaller than $1 + \frac{1+w}{2} + \frac{(1+p)(1+w)}{4}$ for all possible values of p and w. It remains then to consider $u_0 = 0$ and $u_1 = 1$ as Eq. (B5) is symmetric with respect to the exchange of measurements 0 and 1. It can be easily seen that in this case, the right-hand side of Eq. (B5) is, at most,

$$\frac{1}{2}\left[3 + \frac{1+p}{2} + w\right] \leqslant 1 + \frac{1+w}{2} + \frac{(1+p)(1+w)}{4}. \tag{B7}$$

In summary, we have seen that for the points where the gradient with respect to u_i vanishes, $\mathcal{B}_1 \leqslant 2$ or $\mathcal{B}_1 \leqslant 1 + \frac{1+w}{2} + \frac{(1+p)(1+w)}{4}$ for the given purities, which implies in particular that $\mathcal{B}_1 \leqslant \max(2, 1 + \frac{1+w}{2} + \frac{(1+p)(1+w)}{4})$. In order to prove the theorem, it therefore remains to show that this upper bound also holds true at the boundary of the domain $0 \leqslant u_i \leqslant 1$, i.e., the effect for one of the measurements is either projective (case A) or the identity (case B). Note that $\mathcal{B}_1$ is symmetric regarding the exchange of the measurements. Let us first discuss case A and choose, without loss of generality, $u_1 = 1$, i.e., measurement 1 is projective. At the points where the derivative with respect to u_0 vanishes, one obtains

$$\begin{aligned}\mathcal{B}_1 &= \{[p(+|0) - 1][1 - p(+|+00)] + 1 \\ &\quad + p(-|+01) + p(+|1)p(+|+11)\} \\ &\leqslant 1 + \frac{1+w}{2} + \frac{(1+p)(1+w)}{4},\end{aligned} \tag{B8}$$

for all p and w if $u_1 = 1$. It remains to consider, for case A, the boundary points $u_0 = 0$ and $u_0 = 1$. The case $u_0 = 0$ corresponds to a deterministic assignment of outcome "+" and is included in case B. By choosing analogously to before (see also [53]) the optimal states for the measurements with $u_0 = u_1 = 1$, one obtains that in this case,

$$\tfrac{1}{4}(2 + w\sqrt{2-2x})(2 + p\sqrt{2+2x}). \tag{B9}$$

It can be shown that at the critical points, this function is smaller or equal to $1 + \frac{1+w}{2} + \frac{(1+p)(1+w)}{4}$, and therefore with projective effects one cannot exceed this value. We will proceed with case B and choose, without loss of generality, $u_0 = 0$. It is then immediate to see that the Bloch vectors of the optimal states have to be chosen parallel or antiparallel to the Bloch vector of measurement 1. For this choice of states and measurements, one obtains

$$\mathcal{B}_1 = \tfrac{1}{4}[8 + (1-p)(1-w)u_1^2 + 2u_1(-1 + p + 2w)].$$

It can be checked that for the boundary points $u_1 = 0$ and $u_1 = 1$, this implies that $\mathcal{B}_1 = 2$ and $\mathcal{B}_1 = 1 + \frac{1+w}{2} + \frac{(1+p)(1+w)}{4}$. Moreover, it can be easily seen that the point where the derivative with respect to u_1 vanishes corresponds to a minimum. In summary, this concludes the proof that $\mathcal{B}_1$, for given length of the Bloch vectors of the states w and p, is upper bounded by $\max(2, 1 + \frac{1+w}{2} + \frac{(1+p)(1+w)}{4})$. Recall that this bound is tight and that $2 \geqslant 1 + \frac{1+w}{2} + \frac{(1+p)(1+w)}{4}$ if and only if $w \leqslant \frac{1-p}{3+p}$, which proves the theorem. ∎

APPENDIX C: PROOF OF THE UPPER BOUND ON THE CONCURRENCE BASED ON THE PURITY OF A SUBSYSTEM

It should be noted that the upper bound on the concurrence given by $C(\varrho_{AB}) \leqslant \min_{X\in\{A,B\}} \sqrt{2\{1 - \mathrm{tr}[(\varrho_X)^2]\}}$ has already been proven for arbitrary bipartite d-dimension systems in [21]. For the sake of completeness, here we provide an (alternative, but similar) proof for two-qubit states.

Proof. We will use, in the following, that in the two-qubit case, it has been proven in [59] that for any ϱ, there exists some decomposition into pure states, $\varrho = \sum_{i=1}^{r} p_i|\phi_i\rangle\langle\phi_i|$, such that $C(\varrho) = C(|\phi_i\rangle)$ for all $i \in \{1,\dots,r\}$. Moreover, recall that it holds, for the pure states $|\phi_i\rangle$, that $C(|\phi_i\rangle) = \sqrt{2\{1 - \mathrm{tr}[(\varrho_A^i)^2]\}}$ with $\varrho_A^i = \mathrm{tr}_B(|\phi_i\rangle\langle\phi_i|)$. Note that due to $C(|\phi_i\rangle) = C(|\phi_j\rangle)$, we have, therefore,

$$\mathrm{tr}\big[\big(\varrho_A^i\big)^2\big] = \mathrm{tr}\big[\big(\varrho_A^j\big)^2\big] \equiv Q(\varrho). \tag{C1}$$

From this equation, it follows that

$$\begin{aligned}\mathrm{tr}[(\varrho_A)^2] &= \sum_{i,j} p_i p_j \mathrm{tr}\big(\varrho_A^i \varrho_A^j\big)\\ &\leqslant \sum_{i,j} p_i p_j \sqrt{\mathrm{tr}\big[\big(\varrho_A^i\big)^2\big]}\sqrt{\mathrm{tr}\big[\big(\varrho_A^j\big)^2\big]}\\ &= \sum_{i,j} p_i p_j \mathrm{tr}\big[\big(\varrho_A^i\big)^2\big] = Q(\varrho).\end{aligned} \tag{C2}$$

The inequality arises from the Cauchy-Schwarz inequality (using the Hilbert-Schmidt inner product for each summand) and then we use Eq. (C1) and $\sum p_i = 1$. Hence, we have that

$$C(\varrho) = C(|\phi_i\rangle) = \sqrt{2[1 - Q(\varrho)]} \leqslant \sqrt{2\{1 - \mathrm{tr}[(\varrho_A)^2]\}}. \tag{C3}$$

One can show analogously that the bound also holds true for ϱ_B, which proves the statement. ∎

[1] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Phys. Rev. Lett. **70**, 1895 (1993).

[2] R. Raussendorf and H. J. Briegel, Phys. Rev. Lett. **86**, 5188 (2001).

[3] M. A. Nielsen and C. M. Dawson, Phys. Rev. A **71**, 042323 (2005).

[4] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Phys. Rev. Lett. **76**, 722 (1996); **78**, 2031(E) (1997).

[5] W. Dür, H. Aschauer, and H. J. Briegel, Phys. Rev. Lett. **91**, 107903 (2003); H. Aschauer, W. Dür, and H. J. Briegel, Phys. Rev. A **71**, 012319 (2005).

[6] C. Kruszynska, A. Miyake, H. J. Briegel, and W. Dür, Phys. Rev. A **74**, 052316 (2006).

[7] R. Raussendorf, J. Harrington, and K. Goyal, Ann. Phys. **321**, 2242 (2006).

[8] C. M. Dawson, H. L. Haselgrove, and M. A. Nielsen, Phys. Rev. A **73**, 052306 (2006).

[9] A. Reńyi, in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Contributions to the Theory of Statistics (University of California Press, Berkeley, Los Angeles, 1961), Vol. 1, p. 547.

[10] M. Horodecki, P. Horodecki, and J. Oppenheim, Phys. Rev. A **67**, 062104 (2003).

[11] G. Gour, M. P. Müller, V. Narasimhachar, R. W. Spekkens, and N. Y. Halpern, Phys. Rep. **583**, 1 (2015).

[12] A. Streltsov, H. Kampermann, S. Wölk, M. Gessner, and D. Bruß, New J. Phys. **20**, 053058 (2018).

[13] M. Horodecki, K. Horodecki, P. Horodecki, R. Horodecki, J. Oppenheim, A. Sen (De), and U. Sen, Phys. Rev. Lett. **90**, 100402 (2003).

[14] I. Devetak, Phys. Rev. A **71**, 062303 (2005).

[15] K. Życzkowski, P. Horodecki, A. Sanpera, and M. Lewenstein, Phys. Rev. A **58**, 883 (1998).

[16] L. Gurvits and H. Barnum, Phys. Rev. A **66**, 062311 (2002).

[17] M. A. Nielsen, Phys. Rev. Lett. **83**, 436 (1999).

[18] A. Acín, R. Tarrach, and G. Vidal, Phys. Rev. A **61**, 062307 (2000).

[19] E. Bagan, M. A. Ballester, R. Muñoz-Tapia, and O. Romero-Isart, Phys. Rev. Lett. **95**, 110504 (2005).

[20] F. Mintert and A. Buchleitner, Phys. Rev. Lett. **98**, 140505 (2007).

[21] C.-J. Zhang, Y.-X. Gong, Y.-S. Zhang, and G.-C. Guo, Phys. Rev. A **78**, 042308 (2008).

[22] S. Hill and W. K. Wootters, Phys. Rev. Lett. **78**, 5022 (1997).

[23] P. Rungta, V. Buzek, C. M. Caves, M. Hillery, and G. J. Milburn, Phys. Rev. A **64**, 042315 (2001).

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

[25] A. K. Ekert, C. M. Alves, D. K. L. Oi, M. Horodecki, P. Horodecki, and L. C. Kwek, Phys. Rev. Lett. **88**, 217901 (2002).

[26] C. M. Alves and D. Jaksch, Phys. Rev. Lett. **93**, 110501 (2004).

[27] A. J. Daley, H. Pichler, J. Schachenmayer, and P. Zoller, Phys. Rev. Lett. **109**, 020505 (2012).
[28] H. Pichler, L. Bonnes, A. J. Daley, A. M. Läuchli, and P. Zoller, New J. Phys. **15**, 063003 (2013).
[29] R. Islam, R. Ma, P. M. Preiss, M. E. Tai, A. Lukin, M. Rispoli, and M. Greiner, Nature (London) **528**, 77 (2015).
[30] S. J. van Enk and C. W. J. Beenakker, Phys. Rev. Lett. **108**, 110503 (2012).
[31] A. Elben, B. Vermersch, M. Dalmonte, J. I. Cirac, and P. Zoller, Phys. Rev. Lett. **120**, 050406 (2018).
[32] A. Elben, B. Vermersch, C. F. Roos, and P. Zoller, Phys. Rev. A **99**, 052323 (2019).
[33] T. Brydges, A. Elben, P. Jurcevic, B. Vermersch, C. Maier, B. P. Lanyon, P. Zoller, R. Blatt, and C. F. Roos, Science **364**, 260 (2019).
[34] T. Tanaka, G. Kimura, and H. Nakazato, Phys. Rev. A **87**, 012303 (2013).
[35] C. Zhang, G. Wang, and M. Ying, Phys. Rev. A **75**, 062306 (2007).
[36] S. Mal, T. Pramanik, and A. S. Majumdar, Phys. Rev. A **87**, 012105 (2013).
[37] M. G. A. Paris, F. Illuminati, A. Serafini, and S. De Siena, Phys. Rev. A **68**, 012314 (2003).
[38] G. Adesso, A. Serafini, and F. Illuminati, Open Syst. Inf. Dyn. **12**, 189 (2005).
[39] D. Ahn, Y. S. Teo, H. Jeong, F. Bouchard, F. Hufnagel, E. Karimi, D. Koutný, J. Řeháček, Z. Hradil, G. Leuchs, and L. L. Sánchez-Soto, Phys. Rev. Lett. **122**, 100404 (2019).
[40] X.-D. Yu and O. Gühne, Phys. Rev. A **99**, 062310 (2019).
[41] D. Goyeneche, G. Cañas, S. Etcheverry, E. S. Gómez, G. B. Xavier, G. Lima, and A. Delgado, Phys. Rev. Lett. **115**, 090401 (2015).
[42] Q. P. Stefano, L. Rebón, S. Ledesma, and C. Iemmi, arXiv:1903.05709.
[43] A. Zhang, Y. Zhang, F. Xu, L. Li, and L. Zhang, J. Phys. A: Math. Theor. **51**, 395304 (2018).
[44] F. Verstraete and M. M. Wolf, Phys. Rev. Lett. **89**, 170401 (2002).
[45] E. Santos, Phys. Rev. A **69**, 022305 (2004); **70**, 059901(E) (2004).
[46] R. Chaves, J. B. Brask, and N. Brunner, Phys. Rev. Lett. **115**, 110501 (2015).
[47] H. Aschauer, J. Calsamiglia, M. Hein, and H. J. Briegel, Quantum Inf. Comput. **4**, 383 (2004).
[48] N. Wyderka and O. Gühne, arXiv:1905.06928.
[49] C. Eltschka and J. Siewert, Quantum **4**, 229 (2020).
[50] The observables that one needs to measure in order to obtain the bounds in Eqs. (2) and (3) are local with respect to the splitting $A_1A_2|B_1B_2$, where A_i (B_i) refers to party A (B) of copy i, respectively, but are nonlocal with respect to the splitting $A_1|A_2$ and/or $B_1|B_2$ [20,21].
[51] M. Dall'Arno and F. Buscemi, IEEE Trans. Inf. Theory **65**, 2614 (2019).
[52] Instruments are collections of completely positive maps (each associated to one outcome) for which the sum over all of these maps results in a trace-preserving map (see, e.g., [60]).
[53] J. Hoffmann, C. Spee, O. Gühne, and C. Budroni, New J. Phys. **20**, 102001 (2018).
[54] C. Budroni, G. Fagundes, and M. Kleinmann, New J. Phys. **21**, 093018 (2019).
[55] C. Spee, H. Siebeneich, T. F. Gloger, P. Kaufmann, M. Johanning, M. Kleinmann, C. Wunderlich, and O. Gühne, New J. Phys. **22**, 023028 (2020).
[56] A. R. R. Carvalho, F. Mintert, and A. Buchleitner, Phys. Rev. Lett. **93**, 230501 (2004).
[57] L. Aolita and F. Mintert, Phys. Rev. Lett. **97**, 050501 (2006).
[58] In [53], we argued that either the maximum of $\mathcal{B}_1$ is smaller than or equal to 3 or the effects for each measurement corresponding to outcome "−" have rank 1. Examining the points where the derivative with respect to a_i vanishes more closely, it can be easily seen that $\mathcal{B}_1 \leqslant 2$ at that points for any given states (as is explained in the following).
[59] W. K. Wootters, Phys. Rev. Lett. **80**, 2245 (1998).
[60] T. Heinosaari and M. Ziman, *The Mathematical Language of Quantum Mechanics* (Cambridge University Press, Cambridge, 2011).