Correlations between outcomes of random measurements

Minh Cong Tran,[1,*] Borivoje Dakić,[2,3] Wiesław Laskowski,[4] and Tomasz Paterek[1,5,6]
[1]*School of Physical and Mathematical Sciences, Nanyang Technological University, 637371 Singapore*
[2]*Faculty of Physics, University of Vienna, Boltzmanngasse 5, A-1090 Vienna, Austria*
[3]*Institute for Quantum Optics and Quantum Information, Austrian Academy of Sciences, Boltzmanngasse 3, A-1090 Vienna, Austria*
[4]*Institute of Theoretical Physics and Astrophysics, Faculty of Mathematics, Physics and Informatics, University of Gdańsk, 80-308 Gdańsk, Poland*
[5]*Centre for Quantum Technologies, National University of Singapore, 117543 Singapore*
[6]*MajuLab, CNRS-UNS-NUS-NTU International Joint Research Unit, UMI 3654, Singapore*
(Received 27 May 2016; published 3 October 2016)

We recently showed that multipartite correlations between outcomes of random observables detect quantum entanglement in all pure and some mixed states. In this followup article we further develop this approach, derive a maximal amount of such correlations, and show that they are not monotonous under local operations and classical communication. Nevertheless, we demonstrate their usefulness in entanglement detection with a single random observable per party. Finally we study convex-roof extension of the correlations and provide a closed-form necessary and sufficient condition for entanglement in rank-2 mixed states and a witness in general.

DOI: 10.1103/PhysRevA.94.042302

I. INTRODUCTION

The Bell singlet state is a paradigmatic example of an entangled state. This is usually demonstrated by noting that the entropy of the pair of particles is smaller than the entropy of each particle, a possibility forbidden in classical objects. At the same time the singlet state is famous for its correlations. Indeed, two observers measuring the same spin direction will always find their outcomes opposite. This holds independently of a particular measurement direction, in agreement with the total spin being zero. Furthermore, even for spin directions that differ quantum mechanics predicts high probability of opposite outcomes. One might therefore ask if correlations between randomly chosen observables reveal entanglement. We have recently shown that such "random correlations" are indeed a feature of entangled pure states [1]: A pure N-particle state is entangled if and only if the squared N-partite correlation functions averaged over uniform choices of local observables exceed a certain bound.

In this followup paper we extend our approach in several ways. In Sec. II we focus on pure states in arbitrary dimensions and derive explicitly equivalence between entanglement and the random correlations in general. We then study maximal amount of random correlations in a pure state and find that it is achieved (nonuniquely) by the Greenberger-Horne-Zeilinger (GHZ) states of an odd number of qubits. (We conjecture that GHZ states give rise to maximal random correlations in general.) It turns out that random correlations of the two-dimensional (2D) cluster states scale intermediately as expected from entanglement of resources for universal quantum computing [2,3]. All this suggests that random correlations might be a proper entanglement monotone. We show that this is true for bipartite systems and provide explicit counterexamples for a five-qubit system. Nevertheless, the random correlations are helpful as entanglement witnesses which we demonstrate on a vivid example where entanglement is detected with *one* random observable per party.

In Sec. III we move to mixed states and consider the convex roof extension of random correlations. We prove a necessary and sufficient condition for entanglement in rank-2 states and present an entanglement witness for general states. The witness is illustrated on an explicit example where it detects all entangled states of a certain family.

II. PURE STATES

Let us briefly summarize previous results regarding random correlations and a related notion of the length of correlations (not to be confused with the length in physical space). We start with two-level systems, qubits. Any N-qubit density matrix can be represented in terms of Pauli matrices as

$$\rho = \frac{1}{2^N} \sum_{\mu_1,\ldots,\mu_N=0}^{3} T_{\mu_1\ldots\mu_N} \sigma_{\mu_1} \otimes \cdots \otimes \sigma_{\mu_N}, \tag{1}$$

where $\sigma_1,\sigma_2,\sigma_3$ are the Pauli matrices and σ_0 is the identity. The real coefficients $T_{\mu_1\ldots\mu_N} \in [-1,1]$ form an extended correlation tensor which is just an alternative to the density matrix representation of a quantum state. If each party chooses to measure their qubit along a local direction $\vec{u}_n$ then expectation of the product of their measurement outcomes, the so-called correlation function, is given by

$$E(\vec{u}_1,\ldots,\vec{u}_N) = \sum_{j_1,\ldots,j_N=1}^{3} T_{j_1\ldots j_N} (\vec{u}_1)_{j_1} \ldots (\vec{u}_N)_{j_N}, \tag{2}$$

where $(\vec{u}_n)_{j_n}$ is the j_nth component of vector $\vec{u}_n$. We define *random correlations* as the expectation value of squared correlation functions averaged over uniform choices of settings for each individual observer,

$$\mathcal{R} \equiv \frac{1}{(4\pi)^N} \int d\vec{u}_1 \ldots \int d\vec{u}_N \; E^2(\vec{u}_1,\ldots,\vec{u}_N), \tag{3}$$

where $d\vec{u}_n = \sin\theta_n d\theta_n d\phi_n$ is the usual measure on the unit sphere. To estimate $\mathcal{R}$, it would seem that we have to take into

*Present address: Joint Center for Quantum Information and Computer Science, University of Maryland, College Park, Maryland 20742, USA.

account all local directions but in fact it is sufficient to consider only a set of orthogonal axes for each party [1]:

$$\mathcal{R} = \frac{1}{3^N} \sum_{\vec{u}_1,\ldots,\vec{u}_N=\vec{x},\vec{y},\vec{z}} E^2(\vec{u}_1,\ldots,\vec{u}_N) \equiv \frac{\mathcal{C}}{3^N}, \tag{4}$$

where *length of correlations* $\mathcal{C}$ is defined as the sum of squared correlations measured along a complete set of orthogonal local axes. Note that no common reference frame is required; each observer is allowed a different Cartesian coordinate system. The name "length of correlations" refers to the fact that $\mathcal{C}$ is a squared norm of the correlation tensor (with components having solely correlation functions between all N qubits). It has been shown in Refs. [1,4–6] that $\mathcal{C} > 1$ if and only if the system is in a pure entangled state. In Appendix A we present a simple alternative proof for pure systems of two and three qubits that follows directly from the Schmidt decomposition. This line of reasoning does not extend to a higher number of qubits because the restrictions brought forward by the Schmidt decomposition do not engage a sufficient number of subsystems.

A. Higher dimensions

We now extend our criteria for entanglement to higher dimensions, i.e., qudits of dimension d. The final result is already presented in Ref. [1], but we would like to clarify which quantities exactly we consider and discuss explicitly some subtleties. The step-by-step derivation is presented below.

We replace the Pauli matrices with a complete orthogonal basis consisting of identity and d^2-1 traceless operators σ_j such that for all $j,k = 1,\ldots,d^2-1$

$$\mathrm{Tr}\,(\sigma_j) = 0, \tag{5}$$

$$\mathrm{Tr}\,(\sigma_j \sigma_k^\dagger) = d\,\delta_{jk}. \tag{6}$$

Various concrete realizations can be taken here, e.g., generalized Gell-Mann operators (Hermitian basis) or the Weyl-Heisenberg operators (unitary basis). For completeness we write them in Appendix B.

An arbitrary state ρ of N qudits can be decomposed using these operators in a way similar to the Bloch representation:

$$\rho = \frac{1}{d^N} \sum_{\mu_1,\ldots,\mu_N=0}^{d^2-1} T_{\mu_1\ldots\mu_N} \sigma_{\mu_1} \otimes \cdots \otimes \sigma_{\mu_N}, \tag{7}$$

with $\sigma_0 = \mathbb{I}$ being the identity and σ_j the operators defined above. As before, the coefficients are

$$T_{\mu_1\ldots\mu_N} = \mathrm{Tr}\big(\rho\, \sigma_{\mu_1}^\dagger \otimes \cdots \otimes \sigma_{\mu_N}^\dagger\big). \tag{8}$$

However, these coefficients are in general complex valued as σ's are not required to be Hermitian. We therefore define the *length of correlations* with the help of absolute value:

$$\mathcal{C} \equiv \sum_{j_1,\ldots,j_N=1}^{d^2-1} \big|T_{j_1\ldots j_N}\big|^2. \tag{9}$$

As proven in Appendix C, the length of correlations $\mathcal{C}$ is invariant under the choice of local basis as long as (5) and (6) are satisfied. Note that this is more general than invariance under local unitary transformations. For example, the Gell-Mann basis cannot be obtained from the Weyl-Heisenberg basis by local unitaries because the corresponding operators have different eigenvalues. Nevertheless, the length of correlations is the same in both bases provided they are suitably normalized.

Similar to the case of qubits, the length of correlations can also be used to identify entanglement in pure states.

Theorem 1. Pure state $|\Psi\rangle$ is entangled if and only if

$$\mathcal{C} > (d-1)^N. \tag{10}$$

Proof. The proof follows the same lines as for qubits and is presented in Appendix D. ∎

It should also be clear how to extend this theorem to cover subsystems of different dimensions, e.g., 2×3.

We will now define random correlations for qudits and show that they are proportional to $\mathcal{C}$. The generalization is obtained by requesting that random correlations are the average of squared correlations over uniform choices of unitary matrices U_n for each observer:

$$\mathcal{R} \equiv \frac{1}{W^N} \int dU_1 \ldots \int dU_N\ |E(U_1,\ldots,U_N)|^2, \tag{11}$$

where W is a normalization constant, i.e., $\int dU_n = W$, and the unitary-dependent correlation function reads

$$E(U_1,\ldots,U_N) = \mathrm{Tr}\left[\rho \bigotimes_{n=1}^{N} U_n^\dagger \sigma_1^\dagger U_n\right]. \tag{12}$$

Here we have chosen exemplary operator σ_1 as an initial observable of the averaging. In Appendix E we prove that $\mathcal{R}$ is independent of the choice of this initial operator as long as it is traceless and normalized. With these definitions we link random correlations and entanglement as follows.

Theorem 2. For any pure state of N qudits,

$$\mathcal{R} = \frac{\mathcal{C}}{(d^2-1)^N}. \tag{13}$$

Proof. Note that the length of correlations $\mathcal{C}$ is a sum of $(d^2-1)^N$ squared correlation functions, each of which is equal to $\mathcal{R}$ after averaging over local unitary transformations (see Appendix E). Since $\mathcal{C}$ is invariant under such transformations, overall these averages still sum up to $\mathcal{C}$. Therefore $\mathcal{C}$ is just a multiple of $\mathcal{R}$. ∎

B. Maximal random correlations

Here we examine states that maximize $\mathcal{C}$ (and therefore $\mathcal{R}$). We focus on multiple qubits. Let us begin by proving the upper bound on the length of correlations.

Theorem 3. Every pure state of odd number of qubits N satisfies

$$\mathcal{C} \leqslant 2^{N-1} \quad (N \text{ is odd}). \tag{14}$$

Proof. Let us denote by C_k the sum of squared correlations between any k observers. In particular, for a system of N qubits in a state $|\psi\rangle$ we have $C_N = \mathcal{C}(\psi)$. The purity condition of $|\psi\rangle$ implies

$$1 + C_1 + C_2 + C_3 + \cdots + C_{N-1} + C_N = 2^N, \tag{15}$$

whereas Ref. [7] demonstrates for N odd

$$1 - C_1 + C_2 - C_3 + \cdots + C_{N-1} - C_N = 0. \tag{16}$$

Summing up (15) and (16) we obtain

$$C_1 + C_3 + \cdots + C_N = 2^{N-1}. \tag{17}$$

The theorem follows by noting that each C_k is non-negative. ■

The bound of (14) is tight and it is achieved, e.g., by the GHZ state

$$|\mathrm{GHZ}\rangle_N = \frac{1}{\sqrt{2}}(|0\rangle^{\otimes N} + |1\rangle^{\otimes N}). \tag{18}$$

Although the theorem works only for states of odd number of qubits N, a similar bound, of value $2^{N-1}+1$, is observed for GHZ states of even N. We conjecture that this is indeed the maximum possible value of $\mathcal{C}$ for any even N. We have been uniformly sampling pure states randomly over respective spaces and so far no counterexample to the conjecture has been found. We also note that GHZ states are not the only states with maximal value of $\mathcal{C}$. For example, for $N = 4$ qubits the same value is attained by the double singlet state

$$|\Psi^-\rangle|\Psi^-\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}} \otimes \frac{|01\rangle - |10\rangle}{\sqrt{2}}. \tag{19}$$

C. Random correlations of cluster states

As another concrete example we calculate the length of correlations for 2D cluster states [8,9]. Since they are universal for quantum computing their geometric measure of entanglement displays intermediate values in agreement with findings that too highly entangled states are useless for universal quantum computing [2,3]. We find the same behavior of the length of correlations.

Consider a square lattice of size n with each node connected to its nearest neighbors. At each node a, let there be a qubit associated with it and an operator

$$K_a = \sigma_z^{(a)} \bigotimes_{b\in\mathcal{N}(a)} \sigma_x^{(b)}, \tag{20}$$

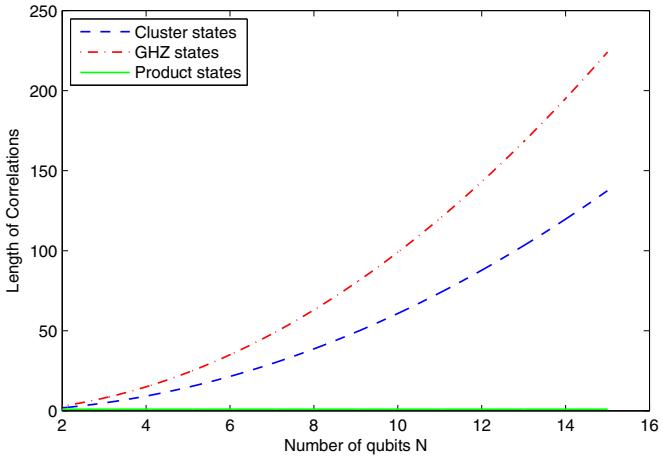

FIG. 1. Length of correlations for 2D cluster states (blue dashed line) and GHZ states (red dash-dot line). Here the length of correlations has the same features as the geometric measure of entanglement. However, see Sec. II D. The curve is obtained by fitting an exponential function to numerical data.

where the superscripts a and b show on which qubits the Pauli matrices act. The tensor product is taken over the set $\mathcal{N}(a)$ of nodes neighboring with a. The cluster state $|C\rangle$ is defined as a common eigenstate of operators K_a [8]:

$$K_a|C\rangle = |C\rangle, \quad \forall a. \tag{21}$$

We compute the length of correlations for the $n \times n$ cluster states for small n and extrapolate to larger n. As seen in Fig. 1, random correlations of cluster states are halfway between product states and GHZ states, so that they mimic behavior of the geometric measure of entanglement.

D. Does $\mathcal{C}$ measure entanglement?

Since $\mathcal{C}$ and $\mathcal{R}$ perfectly distinguish pure entangled states from disentangled ones, and in calculations of concrete examples they display proportionality to the geometric measure of entanglement, we ask if they are entanglement monotones in general. We prove that they are the monotones for bipartite systems. However, this does not generalize to multipartite systems as we will show on counterexamples below.

1. Random correlations may increase on average under local operations

Consider the following state of five qubits:

$$|\Psi\rangle = \frac{|0\rangle|\mathrm{GHZ}\rangle_4 + |1\rangle|D_4^2\rangle}{\sqrt{2}}, \tag{22}$$

where $|\mathrm{GHZ}\rangle_4$ is the GHZ state of four qubits, defined in (18), and $|D_4^2\rangle$ is the four-qubit Dicke state,

$$|D_4^2\rangle = \frac{1}{\sqrt{6}}(|1100\rangle + |1010\rangle + |1001\rangle, + |0110\rangle + |0101\rangle + |0011\rangle). \tag{23}$$

It is straightforward to verify that the length of correlations of $|\Psi\rangle$ is $\mathcal{C}(\Psi) = 8$. If a projective measurement in the computational basis is performed on the first qubit, the state will collapse to either $|0\rangle|\mathrm{GHZ}\rangle$ or $|1\rangle|D_4^2\rangle$, both of which have length of correlations equal to 9. Thus $\mathcal{C}$ increases after such local measurement independently of the actual measurement outcome. Thus, it is not a legitimate entanglement measure [10]. This five-qubit example is the simplest that we were able to find. Therefore, in principle, the random correlations could still be entanglement monotone for systems of three and four qubits.

2. Random correlations and LOCC conversion between pure states

The previous section disproves a strong version of monotonicity of random correlations under local operations and classical communication (LOCC). There is still a possibility that $\mathcal{R}$ is a monotone not on average, i.e., $\mathcal{R}$ could be a monotone under LOCC operations that map pure states to pure states. We show here that this weaker form of monotonicity also does not hold for $\mathcal{R}$ when applied to multipartite systems.

For bipartite systems the following statement holds.

Theorem 4. For pure bipartite states, $|\psi\rangle$ can be converted by LOCC to $|\phi\rangle$ if and only if $\mathcal{C}(\psi) \geqslant \mathcal{C}(\phi)$.

Proof. From Nielsen's theorem [11], $|\psi\rangle$ can be converted by LOCC to $|\phi\rangle$ if and only if the Schmidt probabilities $p_j(\psi)$ of $|\psi\rangle$ are majorized by $p_j(\phi)$ of $|\phi\rangle$:

$$\sum_{j=1}^{k} p_j(\psi) \leqslant \sum_{j=1}^{k} p_j(\phi), \tag{24}$$

for any $k = 1, \ldots, d$ and the Schimdt probabilities are arranged in decreasing order. From purity condition and the Schmidt decomposition, we find $\mathcal{C}(\psi) = d^2 + 1 - 2d \sum_j p_j^2(\psi)$. Since the square function is strictly convex in $\mathbb{R}^+$, by Karamata's inequality $p_j(\psi)$ is majorized by $p_j(\phi)$ if and only if $\sum_j p_j^2(\psi) \leqslant \sum_j p_j^2(\phi)$, and hence if and only if $\mathcal{C}(\psi) \geqslant \mathcal{C}(\phi)$. ■

We give the following counterexample for multipartite systems. Consider the pair of states:

$$|\psi\rangle = \frac{|0\rangle|\psi^-\rangle|\psi^-\rangle + |1\rangle|\psi^+\rangle|\psi^+\rangle}{\sqrt{2}}, \tag{25}$$

$$|\phi\rangle = |0\rangle|\psi^-\rangle|\psi^-\rangle, \tag{26}$$

where $|\psi^\pm\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle)$ are the two Bell states. Starting with $|\psi\rangle$ we measure the first qubit in the standard basis and, depending on the outcome, apply suitable local unitaries on say the second and fourth qubit to obtain $|\phi\rangle$. However, $\mathcal{C}(\psi) = 8$ whereas $\mathcal{C}(\phi) = 9$.

E. Witnessing entanglement with single random setting per party

Although random correlations do not measure entanglement we argue here that they are useful for entanglement detection. In particular, they allow us to detect quantum entanglement with a high level of confidence even with a single random measurement setting per party [1]. Such advantage is relevant to experiments. For example, the small count rates in a multiphoton experiment (e.g., Ref. [12]) make the measurement of every next setting expensive. The strategy presented here reduces the number of settings required to detect entanglement to its ultimate minimum.

In principle, to determine $\mathcal{R}$, an infinite number of measurements has to be performed both in terms of K, the resources needed to estimate correlation functions, and in terms of M, the resources needed for averaging over random settings. In Ref. [1] we introduced entanglement witness [13,14] that takes the finiteness of K and M into account, and here we demonstrate explicitly the effect of finite K on this witness. For a single random setting, $M = 1$, the witness reads

$$\mathcal{R}_K > 1/3^N + \delta \;\Rightarrow\; \text{likely } \psi \text{ is ent}, \tag{27}$$

where $1/3^N$ is the random correlation of the product state of N qubits and δ is used to set the confidence level of entanglement detection. Namely, if estimated correlation $\mathcal{R}_K$ is far from what is expected for a product state, most likely we are measuring an entangled state. In our calculations, the confidence level is set by requiring that random correlation of a product state is smaller than $1/3^N + \delta$ with probability 95.4%. Table I shows the probability to detect entanglement in GHZ states (and states that can be reached from GHZ by local unitaries) of N qubits with both finite and infinite K. For sufficiently big K the chance of detection grows with N. For finite K the detection probability grows for small N and then starts decaying. This is because the random correlation of any state is exponentially small in N, and therefore there exists critical N for which the error $1/\sqrt{K}$ in estimation of the average due to finite K is comparable with the bound of Eq. (27). As illustration, Table I shows that $K = 1000$ trials is essentially infinity for up to $N = 5$ qubits. For $N = 6$, the bound of the entanglement witness is $\approx$0.01 and indeed matches random correlation of the six-qubit GHZ state, $\approx$0.04, reduced by the error $1/\sqrt{K} \approx 0.03$.

TABLE I. Probability of detecting N-qubit GHZ entanglement with a single random measurement per party at confidence level of 95.4%. K gives the number of trials after which the correlation function is estimated.

N	3	4	5	6	7	8	9	10
$K = 1000$	26%	44%	47%	57%	52%	48%	41%	34%
$K \to \infty$	26%	44%	48%	63%	67%	77%	80%	86%

III. MIXED STATES

So far we have only considered the length of correlations in pure states. Although the previous definition of $\mathcal{C}$, as given in Eq. (4), suits a mixed state ρ it no longer identifies entanglement with certainty as it does for pure states. Clearly, $\mathcal{C}$ can be even less than unity for mixed states. Nevertheless, a necessary and sufficient condition can still be established for entanglement in rank-2 states.

Our approach is to define a new quantity via convex roof extension of the length of correlations:

$$\mathcal{E}(\rho) = \min_{\{\mu_k, \Psi_k\}} \sum_k \mu_k \mathcal{C}(\Psi_k), \tag{28}$$

where the sum is minimized over all possible pure-state decompositions $\{\mu_k, \Psi_k\}$ of ρ, i.e., $\rho = \sum_k \mu_k |\Psi_k\rangle\langle\Psi_k|$. A state ρ is entangled if and only if $\mathcal{E}(\rho) > 1$. As in other convex roof constructions, calculation of $\mathcal{E}(\rho)$ is generally a challenge. However, for certain families of mixed states, explicit formulas for $\mathcal{E}(\rho)$ can be found.

A. Necessary and sufficient condition for entanglement in rank-2 states

Rank-2 states are mixed states which belong to a subspace spanned by only two distinct pure states. They possess properties similar to those of a single qubit that notably simplify the minimization problem of (28). In what follows, we shall acquire a similar technique to that presented by Osborne [15] to evaluate the entanglement of an arbitrary mixed state of rank 2.

Theorem 5. For a multipartite mixed state of rank 2

$$\mathcal{E}(\rho) = \mathcal{C}(\rho) + \tfrac{1}{2}[1 - \mathrm{Tr}(\rho^2)]w_{\min}, \tag{29}$$

where $\mathcal{C}(\rho)$ is given in Eq. (4) and $w_{\min}$ is the lowest eigenvalue of 3×3 matrix defined in the proof.

Proof. See Appendix F. ■

The advantage of Theorem 5 lies in its computability. As a demonstration, we prove that a nontrivial mixture of a product

state and an entangled pure state is always entangled [16,17]. It turns out that $\mathcal{E}(\rho)$ behaves similarly to entanglement quantifiers. Since every product state can be brought to $|00\ldots0\rangle$ using a local unitary transformation, let us write such mixture in the most general form as

$$\rho = p|00\ldots0\rangle\langle 00\ldots0| + (1-p)|\Phi\rangle\langle\Phi|, \tag{30}$$

where p is a probability and $|\Phi\rangle$ is a general pure state. A compact formula for $\mathcal{E}(\rho)$ could be found if we further restrict the state $|\Phi\rangle$ to a superposition of $|00\ldots0\rangle$ and another product state $|\alpha\rangle$ orthogonal to $|00\ldots0\rangle$. Direct application of Theorem 5 shows

$$\mathcal{E}(\rho) = 1 + (1-p)^2[\mathcal{C}(\Phi) - 1]. \tag{31}$$

If $|\Phi\rangle$ is entangled, its length of correlations $\mathcal{C}(\Phi) > 1$. In this case also $\mathcal{E}(\rho) > 1$ and the mixture is entangled for all nontrivial values of p.

B. Witness for general states

We extend the idea used in Theorem 5 to mixed states of arbitrary rank. By following the same steps as in the preceding proof, with the Pauli matrices replaced by generalized Gell-Mann matrices, we obtain a lower bound of the following theorem.

Theorem 6. For a multipartite mixed state of rank m,

$$\mathcal{E}(\rho) \geqslant \mathcal{W}(\rho) \equiv \mathcal{C}(\rho) + \frac{w_{\min}}{m^2}[1 - \mathrm{Tr}(\rho^2)], \tag{32}$$

where all the quantities are defined in analogy to Theorem 5.

This is no longer a necessary and sufficient condition for entanglement because there might not be a physical pure state decomposition that achieves the minimum similar to Eq. (F8).

Nevertheless, this witness is of some interest as demonstrated by the following example where it detects all the entangled states of a certain family. Consider three qubits in the mixed state,

$$\rho = (1-p)|W\rangle\langle W| + p\,\rho_n, \tag{33}$$

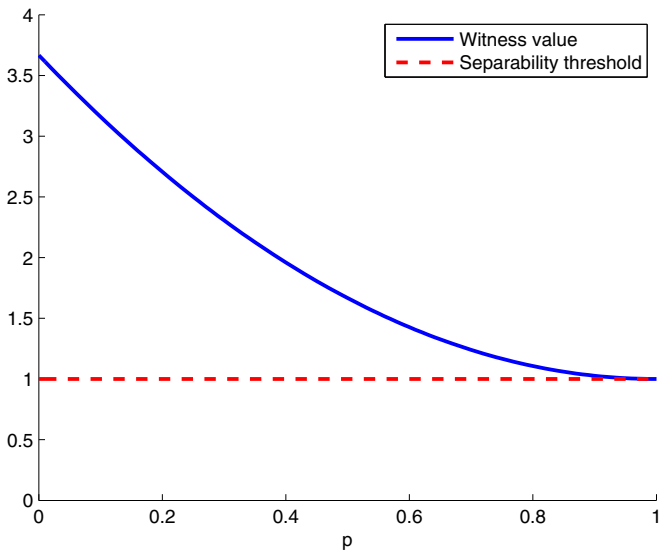

FIG. 2. Witness (32) detects entanglement of all the entangled states of the family given in Eq. (33).

with

$$|W\rangle = \tfrac{1}{\sqrt{3}}(|100\rangle + |010\rangle + |001\rangle),$$
$$\rho_n = \tfrac{1}{3}(|100\rangle\langle 100| + |010\rangle\langle 010| + |001\rangle\langle 001|). \tag{34}$$

It is straightforward to verify that ρ is always of rank 3 except for trivial values of $p = 0$ or 1. Our witness reveals that the mixed state is entangled for all $p < 1$; see Fig. 2.

IV. CONCLUSION

In conclusion, we showed that a multipartite pure quantum state of any dimensions is entangled if and only if it gives rise to higher squared correlations in random measurements. Only correlations between all the parties are relevant. Alternatively, the sum of all squared components of the correlation tensor is higher in all entangled pure quantum states than in product states. Additionally to various features discussed in the main text, this provides understanding why certain pure entangled states do not violate any two-setting Bell inequalities for correlation functions [18–20]. Conditions for violation of such inequalities involve a plane of correlation tensor defined by the two settings. There exist states which have bounded correlations in every plane of the correlation tensor, but when all squared correlations are summed up the state is revealed as entangled.

ACKNOWLEDGMENTS

We warmly thank M. Horodecki for stimulating discussions. This work was supported by the National Research Foundation, Ministry of Education of Singapore Grant No. RG98/13, start-up grant of the Nanyang Technological University, NCN Grant No. 2012/05/E/ST2/02352, European Commission Project RAQUEL, and Austrian Science Fund (FWF) through the Individual Project (No. 2462).

APPENDIX A: RANDOM CORRELATIONS AND SCHMIDT DECOMPOSITION

Theorem 7. A pure state of two and three qubits is entangled iff its length of correlations is greater than 1.

Proof. We will only use Schmidt decompositions and purity conditions. For a pure state Ψ of two qubits, the purity condition, $\mathrm{Tr}(\rho^2) = 1$, requires that

$$\mathcal{C}(\Psi) + |\vec{a}|^2 + |\vec{b}|^2 = 3, \tag{A1}$$

where $\vec{a},\vec{b}$ are the local Bloch vectors and $\mathcal{C}(\Psi)$ is the length of correlations [see Eq. (4)]. From Schmidt decomposition one has $|\vec{a}|^2 = |\vec{b}|^2 \leqslant 1$. Thus the length of correlations must satisfy $\mathcal{C}(\Psi) \geqslant 1$. The only case that $\mathcal{C}(\Psi) = 1$ is when both $|\vec{a}|^2 = |\vec{b}|^2 = 1$, which means Ψ is a product state.

For three qubits in a pure state Ψ, let us denote by ρ_i the reduced state of the ith subsystem and by ρ_{ij} the reduced state of the ith and jth subsystems together. Schmidt decomposition requires that $\mathrm{Tr}(\rho_i^2) = \mathrm{Tr}(\rho_{jk}^2)$ for every $i \neq j \neq k$. In terms of correlations this gives

$$\tfrac{1}{2}(1 + |\vec{v}_i|^2) = \tfrac{1}{4}[1 + |\vec{v}_j|^2 + |\vec{v}_k|^2 + \mathcal{C}(\rho_{jk})], \tag{A2}$$

where, e.g., $\vec{v}_i$ are the Bloch vectors of ρ_i and $\mathcal{C}(\rho_{jk})$ is the length of correlations of the state ρ_{jk}. Note that there are three equations of the form (A2). After summing them up all the Bloch vectors cancel out and we find

$$\mathcal{C}(\rho_{12}) + \mathcal{C}(\rho_{13}) + \mathcal{C}(\rho_{23}) = 3. \tag{A3}$$

Let us recall the purity condition for Ψ:

$$|\vec{v}_1|^2 + |\vec{v}_2|^2 + |\vec{v}_3|^2 + \mathcal{C}(\Psi) + \mathcal{C}(\rho_{12}) + \mathcal{C}(\rho_{13}) + \mathcal{C}(\rho_{23}) = 7. \tag{A4}$$

From (A3) and (A4) together with the fact that the length of Bloch vectors is upper bounded by unity, we have $\mathcal{C}(\Psi) \geqslant 1$. In addition, $\mathcal{C}(\Psi)$ is equal to 1 if and only if all three local Bloch vectors are normalized, i.e., Ψ is a product state. ■

APPENDIX B: OPERATOR BASES

Two most often used operator bases that satisfy Eqs. (5) and (6) are as follows.

The generalized Gell-Mann matrices (Hermitian basis) can be divided into three classes:

$$G^+_{mn} = \sqrt{\tfrac{d}{2}}(|m\rangle\langle n| + |n\rangle\langle m|), \tag{B1}$$

$$G^-_{mn} = \sqrt{\tfrac{d}{2}}(-i|m\rangle\langle n| + i|n\rangle\langle m|),$$

$$\lambda_l = \sqrt{\tfrac{d}{(l+1)(l+2)}}\left(\sum_{j=0}^{l}|j\rangle\langle j| - (l+1)|l+1\rangle\langle l+1|\right), \tag{B2}$$

where $0 \leqslant m < n \leqslant d-1$, and $0 \leqslant l \leqslant d-2$ and d is the dimension of the Hilbert space of pure states.

The Weyl-Heisenberg matrices (unitary basis) read

$$W_{mn} = X^m Z^n, \quad m,n = 0,\ldots,d-1, \tag{B3}$$

$$Z = \sum_k \exp(i2\pi k/d)|k\rangle\langle k|, \tag{B4}$$

$$X = \sum_k |(k+1) \mod d\rangle\langle k|. \tag{B5}$$

APPENDIX C: INVARIANCE OF THE LENGTH OF CORRELATIONS

Theorem 8. The length of correlations, Eq. (9), is invariant under the choice of local basis satisfying Eqs. (5) and (6).

Proof. Denote by $\mathcal{C}$ the length of correlations calculated in a basis $\{\sigma_j\}$ for each of the N qudits. Without loss of generality, we shall prove that the same value is obtained if the local basis of the first qudit is changed to $\{\sigma'_j\}$. Since only the traceless operators enter the definition of the length of correlations, i.e., $j = 1,\ldots,d^2-1$, and both bases are complete we have

$$\sigma'_j = \sum_{k=1}^{d^2-1}\alpha_{jk}\sigma_k, \tag{C1}$$

where α_{jk} are complex coefficients forming unitary matrix α because

$$\sum_k \alpha_{jk}\alpha^*_{lk} = \frac{1}{d}\mathrm{Tr}[\sigma'_j(\sigma'_l)^\dagger] = \delta_{jl}. \tag{C2}$$

In matrix form, $\alpha\alpha^\dagger = \mathbb{I}$. Denote by $\mathcal{C}'$ the length of correlations evaluated in the $\{\sigma'_j\}$ basis. Let σ_M denote the tensor product of $\sigma_{j_2}\otimes\cdots\otimes\sigma_{j_N}$ for the last $N-1$ qudits. We have

$$\begin{aligned}
\mathcal{C}' &= \sum_M\sum_j |\mathrm{Tr}(\rho\,\sigma'_j\otimes\sigma_M)|^2 \\
&= \sum_M\sum_j \left|\sum_k \alpha_{jk}\mathrm{Tr}(\rho\,\sigma_k\otimes\sigma_M)\right|^2 \\
&= \sum_M\sum_j \left|\sum_k \alpha_{jk}T_{kM}\right|^2 \\
&= \sum_M\sum_j \left(\sum_k |\alpha_{jk}|^2|T_{kM}|^2 + \sum_{k\neq l}\alpha_{jk}\alpha^*_{jl}T_{kM}T^*_{lM}\right) \\
&= \sum_M\sum_k \left(\sum_j |\alpha_{jk}|^2\right)|T_{kM}|^2 \\
&\quad + \sum_M\sum_{k\neq l}\left(\sum_j \alpha_{jk}\alpha^*_{jl}\right)T_{kM}T^*_{lM} \\
&= \sum_M\sum_k |T_{kM}|^2 = \mathcal{C},
\end{aligned} \tag{C3}$$

where in the second last equality we used (C2) and $\alpha^\dagger\alpha = (\alpha^{-1}\alpha)(\alpha^\dagger\alpha) = \alpha^{-1}(\alpha\alpha^\dagger)\alpha = \alpha^{-1}\alpha = \mathbb{I}$. ■

APPENDIX D: LENGTH OF CORRELATIONS AND ENTANGLEMENT

Here we prove Theorem 1 of the main text: *Pure state $\rho = |\Psi\rangle\langle\Psi|$ is entangled if and only if*

$$\mathcal{C} > (d-1)^N. \tag{D1}$$

Proof. Clearly, if $|\Psi\rangle$ is a product state, i.e., $|\Psi\rangle = |\Psi_1\rangle\otimes\cdots\otimes|\Psi_N\rangle$, then the length of correlations factors and we obtain

$$\mathcal{C}(\Psi) = \prod_{i=1}^{N}\mathcal{C}(\Psi_i) = (d-1)^N. \tag{D2}$$

For the proof in the other direction consider two copies of the state. In general, we can write the length of correlations as $\mathcal{C} = \mathrm{Tr}(\rho\otimes\rho\,\mathcal{S})$, where $\mathcal{S}$ is defined as

$$\mathcal{S} = S^{11'}\otimes\cdots\otimes S^{NN'}, \tag{D3}$$

with $S^{nn'} = \sum_{j_n=1}^{d^2-1}\sigma_{j_n}\otimes\sigma_{j_n}$, and the superscript denotes pairs of qubits S acts upon. It can be directly verified that choosing σ's as the Gell-Mann operators (without loss of generality as $\mathcal{C}$ is basis independent; see Appendix C) $S^{nn'}$ has

(1) d eigenstates $|\alpha_j\rangle = |jj\rangle$ $(0 \leqslant j \leqslant d-1)$ with eigenvalue $d-1$,

(2) $\frac{d(d-1)}{2}$ eigenstates $|\beta_{ij}\rangle = |ij\rangle + |ji\rangle$ $(0 \leqslant i < j \leqslant d-1)$ with eigenvalue $d-1$,

(3) $\frac{d(d-1)}{2}$ eigenstates $|\gamma_{ij}\rangle = |ij\rangle - |ji\rangle$ $(0 \leqslant i < j \leqslant d-1)$ with eigenvalue $-d-1$.

Since all the eigenvalues of S are $\pm(d-1)$, all the eigenvalues of $\mathcal{S}$ are of the form $(-1)^k(d+1)^k(d-1)^{N-k}$ for $k=0,1,\ldots,N$. However, since $-(d+1)$ corresponds to antisymmetric eigenstates, it must occur in pairs, i.e., k must be even. This follows from the fact that $\mathcal{S}$ is calculated for two identical copies of the state and therefore has no antisymmetric component. Thus the smallest eigenvalue of $\mathcal{S}$ is $(d-1)^N$. It therefore follows that $\mathcal{C} \geqslant (d-1)^N$. The equality is when $|\Psi\rangle \otimes |\Psi\rangle$ lies in the space spanned by the eigenstates $|\alpha_j\rangle$ and $|\beta_{ij}\rangle$. Such state is symmetric with respect to the exchange of *any* qudit j and its copy j'. Let a general expansion of $|\Psi\rangle$ in the standard basis be

$$|\Psi\rangle = \sum_{j_1\ldots j_N} \eta_{j_1\ldots j_N} |j_1 \ldots j_N\rangle, \tag{D4}$$

where $\eta_{j_1\ldots j_N}$ are the complex coefficients. We focus on the first two qudits and write $|\Psi\rangle = \sum \eta_{j_1 j_2|J} |j_1 j_2 J\rangle$ where J stands for a sequence $j_3 j_4 \ldots j_N$ for the last $N-2$ qudits. The state of the two copies of $|\Psi\rangle$ is

$$|\Psi\rangle \otimes |\Psi\rangle = \sum_{j_1,j_2,J} \sum_{j_1',j_2',J'} \eta_{j_1 j_2|J} \eta_{j_1' j_2'|J'} |j_1 j_2 J\rangle \left| j_1' j_2' J' \right\rangle. \tag{D5}$$

We now exchange the first qudit:

$$|\Psi\rangle \otimes |\Psi\rangle = \sum_{j_1,j_2,J} \sum_{j_1',j_2',J'} \eta_{j_1' j_2|J} \eta_{j_1 j_2'|J'} |j_1 j_2 J\rangle \left| j_1' j_2' J' \right\rangle. \tag{D6}$$

Comparing Eqs. (D5) and (D6), we obtain relations between the coefficients η:

$$\eta_{j_1 j_2|J} \eta_{j_1' j_2'|J'} = \eta_{j_1' j_2|J} \eta_{j_1 j_2'|J'}, \tag{D7}$$

which holds for any $j_1, j_1', j_2, j_2', J, J'$. In particular, for $J = J', j_1' = 1$ we find

$$\frac{\eta_{j_1 j_2|J}}{\eta_{1 j_2|J}} = \frac{\eta_{j_1 j_2'|J}}{\eta_{1 j_2'|J}} \equiv k_{j_1} \text{ independent of } j_2, j_2'. \tag{D8}$$

Using these relations we can rewrite the state $|\Psi\rangle$ as

$$|\Psi\rangle = \sum_{j_1,j_2,J} \eta_{j_1 j_2|J} |j_1 j_2 J\rangle \tag{D9}$$

$$= \sum_{j_2,J} \left(\sum_{j_1} \eta_{j_1 j_2|J} |j_1\rangle \right) |j_2 J\rangle \tag{D10}$$

$$= \sum_{j_2,J} \left(\sum_{j_1} k_{j_1} \eta_{1 j_2|J} |j_1\rangle \right) |j_2 J\rangle \tag{D11}$$

$$= \left(\sum_{j_1} k_{j_1} |j_1\rangle \right) \otimes \left(\sum_{j_2,J} \eta_{1,j_2|J} |j_2 J\rangle \right). \tag{D12}$$

Thus $|\Psi\rangle$ is a tensor product of a pure state for the first qudit and another pure state for the last $N-1$ qudits. By applying this argument iteratively we find that $|\Psi\rangle$ is fully separable. ■

APPENDIX E: RANDOM CORRELATIONS DO NOT DEPEND ON THE INITIAL OPERATOR IN AVERAGING

We present two lemmas before moving to the main theorem.

Lemma 1. For traceless and trace-orthogonal operators σ_1 and σ_2, and arbitrary state ρ of two qudits,

$$A \equiv \int dU \, \mathrm{Tr}(U \otimes U\rho\, U^\dagger \otimes U^\dagger \sigma_1^\dagger \otimes \sigma_2) \tag{E1}$$

$$= 0. \tag{E2}$$

Proof. By bringing the integral inside the trace one recognizes the Werner state

$$\rho_W = \int dU \, U \otimes U\rho\, U^\dagger \otimes U^\dagger. \tag{E3}$$

All such states can be written in the form [21]

$$\rho_W = \frac{1}{d^2 - d\alpha}(\mathbb{I} - \alpha P), \tag{E4}$$

where $\alpha \in [-1,1]$ and P is the swap operator

$$P = \sum_{i,j} |ij\rangle\langle ji|. \tag{E5}$$

Since $\sigma_1^\dagger \otimes \sigma_2$ is traceless, we have

$$A = \frac{\alpha}{d^2 - \alpha d} \mathrm{Tr}(P \sigma_1^\dagger \otimes \sigma_2). \tag{E6}$$

It can be directly verified that

$$\mathrm{Tr}(P\sigma_1^\dagger \otimes \sigma_2) = \mathrm{Tr}(\sigma_1^\dagger \sigma_2). \tag{E7}$$

The lemma follows from orthogonality of σ's. ■

Lemma 2. For traceless and trace-orthogonal operators σ_1 and σ_2,

$$B \equiv \int dU \, \mathrm{Tr}(U^\dagger \otimes U^\dagger \sigma_1^\dagger \otimes \sigma_2 U \otimes U) \tag{E8}$$

$$= 0. \tag{E9}$$

Proof. We shall prove that B has all matrix elements $\langle mn|B|kl\rangle = 0$. First write

$$\langle mn|B|kl\rangle = \int dU \, \mathrm{Tr}(|kl\rangle\langle mn| U^\dagger \otimes U^\dagger \sigma_1^\dagger \otimes \sigma_2 U \otimes U). \tag{E10}$$

The diagonal elements, i.e., $k=m$ and $l=n$, vanish due to Lemma 1, because $|kl\rangle\langle kl|$ is a valid density matrix. For off-diagonal elements, i.e., $k \neq m$ or $l \neq n$, apply Lemma 1 to states $\rho_1 = \frac{1}{2}(|kl\rangle + |mn\rangle)(\langle kl| + \langle mn|)$ and $\rho_2 = \frac{1}{2}(|kl\rangle + i|mn\rangle)(\langle kl| - i\langle mn|)$ to obtain respectively

$$0 = \langle mn|B|kl\rangle + \langle kl|B|mn\rangle, \tag{E11}$$

$$0 = \langle mn|B|kl\rangle - \langle kl|B|mn\rangle. \tag{E12}$$

The sum and difference reveal that all off-diagonal elements vanish. ■

Theorem 9. Random correlations

$$\mathcal{R}(\vec{\lambda}) = \frac{1}{W^N} \int dU_1 \ldots \int dU_N \left| \mathrm{Tr}\left(\rho \bigotimes_{n=1}^{N} U_n^\dagger \lambda_n^\dagger U_n \right) \right|^2 \tag{E13}$$

do not depend on the choice of operators λ_n such that $\mathrm{Tr}(\lambda_n) = 0$ and $\mathrm{Tr}(\lambda_n^2) = d$.

Proof. Without loss of generality let us focus on the first subsystem and denote by J the sequence of the last $N-1$ qudits. We therefore write

$$\mathcal{R}(\sigma_1) \equiv \frac{1}{W^N}\int dU_J \int dU_1 \times |\mathrm{Tr}((U_1^\dagger \otimes U_J^\dagger)(\sigma_1^\dagger \otimes \sigma_J^\dagger)(U_1 \otimes U_J))|^2. \quad \text{(E14)}$$

The expression involving the trace can be linearized using the second copy of the state ρ as follows:

$$\begin{aligned}&|\mathrm{Tr}(\rho(U_1^\dagger \otimes U_J^\dagger)(\sigma_1^\dagger \otimes \sigma_J^\dagger)(U_1 \otimes U_J))|^2\\ &\quad = \mathrm{Tr}((\rho \otimes \rho)(U_1^\dagger \otimes U_J^\dagger \otimes U_1^\dagger \otimes U_J^\dagger)\\ &\qquad (\sigma_1^\dagger \otimes \sigma_J^\dagger \otimes \sigma_1 \otimes \sigma_J)(U_1 \otimes U_J \otimes U_1 \otimes U_J)). \quad \text{(E15)}\end{aligned}$$

Without loss of generality (see Appendix C) we will now consider the Weyl-Heisenberg basis. Since all the operators within this basis are related by a unitary we have

$$\mathcal{R}(\sigma_1) = \mathcal{R}(\sigma_2) = \cdots = \mathcal{R}(\sigma_{d^2-1}). \quad \text{(E16)}$$

Take now operator λ from the thesis and decompose it in this basis:

$$\lambda = \sum_{j=1}^{d^2-1} \gamma_j \sigma_j, \quad \text{with} \quad \sum_{j=1}^{d^2-1} |\gamma_j|^2 = 1, \quad \text{(E17)}$$

where the first equation follows from $\mathrm{Tr}(\lambda) = 0$, and the second from $\mathrm{Tr}(\lambda^2) = d$. The random correlations calculated with λ as the initial operator satisfy

$$\mathcal{R}(\lambda) = \sum_{j,k=1}^{d^2-1} \gamma_j \gamma_k^* \int dU_J \int dU_1 \mathrm{Tr}\Big(\rho \otimes \rho U_1^\dagger \otimes U_J^\dagger \otimes U_1^\dagger \otimes U_J^\dagger \sigma_k^\dagger \otimes \sigma_J^\dagger \otimes \sigma_j \otimes \sigma_J U_1 \otimes U_J \otimes U_1 \otimes U_J\Big) \quad \text{(E18)}$$

$$= \sum_{j,k} \gamma_j \gamma_k^* \int dU_J \mathrm{Tr}\left[\rho \otimes \rho\left(\int dU_1 U_1^\dagger \otimes U_1^\dagger \sigma_k^\dagger \otimes \sigma_j U_1 \otimes U_1\right) \otimes U_J^\dagger \otimes U_J^\dagger \sigma_J^\dagger \otimes \sigma_J U_J \otimes U_J\right] \quad \text{(E19)}$$

$$= \sum_{j} |\gamma_j|^2 \int dU_J \mathrm{Tr}\left[\rho \otimes \rho\left(\int dU_1 U_1^\dagger \otimes U_1^\dagger \sigma_j^\dagger \otimes \sigma_j U_1 \otimes U_1\right) \otimes U_J^\dagger \otimes U_J^\dagger \sigma_J^\dagger \otimes \sigma_J U_J \otimes U_J\right] \quad \text{(E20)}$$

$$= \sum_{j} |\gamma_j|^2 \mathcal{R}(\sigma_j) = \mathcal{R}(\sigma_1), \quad \text{(E21)}$$

where in the second line we isolated the first particle from the principal system and the first particle from the copy, in the third line we use Lemma 2, and in the last line Eq. (E16). ■

APPENDIX F: THEOREM ABOUT THE CONVEX ROOF EXTENSION

Here we prove Theorem 5 of the main text: *For a multipartite mixed state of rank 2,*

$$\mathcal{E}(\rho) = \mathcal{C}(\rho) + \tfrac{1}{2}[1 - \mathrm{Tr}(\rho^2)]w_{\min}, \quad \text{(F1)}$$

where $\mathcal{C}(\rho)$ is given in Eq. (4) and $w_{\min}$ is the lowest eigenvalue of the 3×3 matrix defined in the proof below.

Proof. By the definition of rank-2 states ρ can be written as a mixture of $|\tilde{0}\rangle$ and $|\tilde{1}\rangle$, which are two N-qubit pure states. Without loss of generality assume that they are mutually orthogonal. The length of correlations of a pure state Ψ can be written as expectation value of an operator $\mathcal{S}$ in the two-copy state $|\Psi\rangle|\Psi\rangle$ (see Appendix D). Therefore, we can write $\tilde{\mathcal{E}}$ for a particular decomposition $\{\mu_k, \Psi_k\}$ of ρ as

$$\tilde{\mathcal{E}} = \sum_k \mu_k \mathrm{Tr}(\mathcal{S}\,\Pi_k), \quad \text{(F2)}$$

where we denote $\Pi_k = |\Psi_k\rangle\langle\Psi_k| \otimes |\Psi_k\rangle\langle\Psi_k|$. Since all the pure states Ψ_k are within the subspace spanned by $|\tilde{0}\rangle, |\tilde{1}\rangle$, only the projection of $\mathcal{S}$ onto this subspace will contribute to the trace in (F2). Let us therefore introduce 4×4 matrix $\tilde{\mathcal{S}}$ with matrix elements $\langle \tilde{i}\,\tilde{j}|\mathcal{S}|\tilde{m}\tilde{n}\rangle$, where $i, j, m, n = 0, 1$. Similarly, by introducing 4×4 matrix $\tilde{\Pi}_k$ with elements $\langle \tilde{i}\,\tilde{j}|\Pi_k|\tilde{m}\tilde{n}\rangle$, we can rewrite (F2) as

$$\tilde{\mathcal{E}} = \sum_k \mu_k \mathrm{Tr}(\tilde{\mathcal{S}}\,\tilde{\Pi}_k). \quad \text{(F3)}$$

We now represent operators with the tilde in terms of Pauli matrices operating in the support of ρ:

$$\tilde{\Pi}_k = \tfrac{1}{2}(\mathbb{I} + \vec{r}_k \cdot \vec{\sigma}) \otimes \tfrac{1}{2}(\mathbb{I} + \vec{r}_k \cdot \vec{\sigma}), \quad \text{(F4)}$$

$$\tilde{\mathcal{S}} = \frac{1}{4}\left(s_0 \mathbb{I} \otimes \mathbb{I} + \vec{s} \cdot \vec{\sigma} \otimes \mathbb{I} + \mathbb{I} \otimes \vec{s} \cdot \vec{\sigma} + \sum_{i=1}^{3} w_i \sigma_i \otimes \sigma_i\right), \quad \text{(F5)}$$

where $s_0 = \mathrm{Tr}(\tilde{\mathcal{S}})$ and $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ is a vector of standard Pauli matrices chosen in such a way that $\tilde{\mathcal{S}}$ is diagonal with real entries ordered as $w_1 \geqslant w_2 \geqslant w_3$. Note that the "local" parts described by vector $\vec{s}$ are the same since $\tilde{\mathcal{S}}$ is symmetric with respect to the exchange of the two copies. In this representation Eq. (F3) takes the form

$$\tilde{\mathcal{E}} = \frac{1}{4}\left[s_0 + 2\vec{s} \cdot \vec{\rho} + \sum_k \mu_k\left(w_1 x_k^2 + w_2 y_k^2 + w_3 z_k^2\right)\right], \quad \text{(F6)}$$

where $\vec{\rho} = (\rho_x, \rho_y, \rho_z)$ is the Bloch vector representing the state ρ in the subspace spanned by $|\tilde{0}\rangle$ and $|\tilde{1}\rangle$, and we have introduced components of vectors $\vec{r}_k = (x_k, y_k, z_k)$. Since $|\Psi_k\rangle$ are pure states, the purity condition implies $z_k^2 = 1 - x_k^2 - y_k^2$ for all k in the decomposition. Equation (F6) becomes

$$\tilde{\mathcal{E}} = \frac{1}{4}\Big[s_0 + 2\vec{s}\cdot\vec{\rho} + w_3 + (w_1 - w_3)\Big(\sum_k \mu_k x_k^2\Big) + (w_2 - w_3)\Big(\sum_k \mu_k y_k^2\Big)\Big]. \tag{F7}$$

Note that the sums on the right-hand side are quadratic and thus convex. Therefore $\sum_k \mu_k x_k^2 \geqslant (\sum_k \mu_k x_k)^2 = \rho_x^2$ and similarly $\sum_k \mu_k y_k^2 \geqslant \rho_y^2$. Both inequalities can be saturated if $x_k = \rho_x$ and $y_k = \rho_y$ for all k. We therefore have solved the minimization problem of the earlier convex-roof construction

$$\mathcal{E}(\rho) = \min_{\{\mu_k, \Psi_k\}} \tilde{\mathcal{E}} = \tfrac{1}{4}\big[s_0 + 2\vec{s}\cdot\vec{\rho} + w_3 + (w_1 - w_3)\rho_x^2 + (w_2 - w_3)\rho_y^2\big]. \tag{F8}$$

Taking into account that

$$\mathcal{C}(\rho) = \tfrac{1}{4}\big[s_0 + 2\vec{s}\cdot\vec{\rho} + w_1\rho_x^2 + w_2\rho_y^2 + w_3\rho_z^2\big], \tag{F9}$$

$$\mathrm{Tr}(\rho^2) = \tfrac{1}{2}\big(1 + \rho_x^2 + \rho_y^2 + \rho_z^2\big), \tag{F10}$$

we may simplify (F8) to

$$\mathcal{E}(\rho) = \mathcal{C}(\rho) + \tfrac{1}{2}[1 - \mathrm{Tr}(\rho^2)]w_{\min}, \tag{F11}$$

and the theorem follows. ■

[1] M. C. Tran, B. Dakić, F. Arnault, W. Laskowski, and T. Paterek, Phys. Rev. A **92**, 050301(R) (2015).
[2] D. Gross, S. T. Flammia, and J. Eisert, Phys. Rev. Lett. **102**, 190501 (2009).
[3] M. J. Bremner, C. Mora, and A. Winter, Phys. Rev. Lett. **102**, 190502 (2009).
[4] Ali Saif M. Hassan and P. S. Joag, Quantum Inf. Comput. **8**, 773 (2008).
[5] Ali Saif M. Hassan and P. S. Joag, Phys. Rev. A **77**, 062334 (2008).
[6] Ali Saif M. Hassan and P. S. Joag, Phys. Rev. A **80**, 042302 (2009).
[7] C. Eltschka and J. Siewert, Phys. Rev. Lett. **114**, 140402 (2015).
[8] H. J. Briegel and R. Raussendorf, Phys. Rev. Lett. **86**, 910 (2001).
[9] R. Raussendorf and H. J. Briegel, Phys. Rev. Lett. **86**, 5188 (2001).
[10] This section gives a counterexample to Prop. 6 of Ref. [6]. We have informed the authors and they are preparing a corrigendum.
[11] M. A. Nielsen, Phys. Rev. Lett. **83**, 436 (1999).
[12] X.-C. Yao, T.-X. Wang, P. Xu, H. Lu, G.-S. Pan, X.-H. Bao, C.-Z. Peng, C.-Y. Lu, Y.-A. Chen, and J.-W. Pan, Nat. Photon. **6**, 225 (2012).
[13] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1 (1996).
[14] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009).
[15] T. J. Osborne, Phys. Rev. A **72**, 022309 (2005).
[16] P. Horodecki, J. A. Smolin, B. M. Terhal, and A. V. Thapliyal, Theor. Comput. Sci. **292**, 589 (2003).
[17] M. C. Tran, W. Laskowski, and T. Paterek, J. Phys. A **47**, 424025 (2014).
[18] R. F. Werner and M. M. Wolf, Phys. Rev. A **64**, 032112 (2001).
[19] M. Żukowski and Č. Brukner, Phys. Rev. Lett. **88**, 210401 (2002).
[20] M. Żukowski, Č. Brukner, W. Laskowski, and M. Wieśniak, Phys. Rev. Lett. **88**, 210402 (2002).
[21] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).