

Overcoming a limitation of deterministic dense coding with a nonmaximally entangled initial state

P. S. Bourdon*

Department of Mathematics, Washington and Lee University, Lexington, Virginia 24450, USA

E. Gerjuoy†

Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA

(Received 29 June 2009; published 16 February 2010)

Under two-party deterministic dense coding, Alice communicates (perfectly distinguishable) messages to Bob via a qudit from a pair of entangled qudits in pure state $|\Psi\rangle$. If $|\Psi\rangle$ represents a maximally entangled state (i.e., each of its Schmidt coefficients is $\sqrt{1/d}$), then Alice can convey to Bob one of d^2 distinct messages. If $|\Psi\rangle$ is not maximally entangled, then Ji *et al.* [Phys. Rev. A **73**, 034307 (2006)] have shown that under the original deterministic dense-coding protocol, in which messages are encoded by unitary operations performed on Alice's qudit, it is impossible to encode $d^2 - 1$ messages. Encoding $d^2 - 2$ messages is possible; see, for example, the numerical studies by Mozes *et al.* [Phys. Rev. A **71**, 012311 (2005)]. Answering a question raised by Wu *et al.* [Phys. Rev. A **73**, 042311 (2006)], we show that when $|\Psi\rangle$ is not maximally entangled, the communications limit of $d^2 - 2$ messages persists even when the requirement that Alice encode by unitary operations on her qudit is weakened to allow encoding by more general quantum operators. We then describe a dense-coding protocol that can overcome this limitation with high probability, assuming the largest Schmidt coefficient of $|\Psi\rangle$ is sufficiently close to $\sqrt{1/d}$. In this protocol, $d^2 - 2$ of the messages are encoded via unitary operations on Alice's qudit, and the final $(d^2 - 1)$ -th message is encoded via a non-trace-preserving quantum operation.

DOI: [10.1103/PhysRevA.81.022314](https://doi.org/10.1103/PhysRevA.81.022314)

PACS number(s): 03.67.Hk, 03.65.Ud, 03.67.Mn

I. INTRODUCTION

We assume that Alice and Bob, located some distance apart, each initially control one qudit from an entangled pair in pure state $|\Psi\rangle$. Deterministic dense coding, originated by Bennett and Wiesner [1], allows Alice to send to Bob, via her qudit, one of up to d^2 perfectly distinguishable messages. Under the original protocol for deterministic dense coding, to send a message to Bob, Alice first applies to her qudit a local unitary operation selected from a group of possible "encoding unitaries." She then sends her qudit via a d -dimensional noiseless quantum channel to Bob, who—with both qudits now in his possession—performs a measurement that reveals the particular encoding unitary operation that Alice had performed; that is, it reveals the particular message Alice had chosen to send. So that these messages are perfectly distinguishable by Bob (the hallmark of *deterministic* dense coding), the set of encoding unitaries that Alice may apply to her qudit must produce a set of orthogonal states in the state-space of the two-qudit system that Alice and Bob share. Since that state-space is of dimension d^2 , she can send at most d^2 messages to Bob by this process. She can send as many as d^2 messages only when $|\Psi\rangle$ is a maximally entangled state, meaning that its Schmidt coefficients are identical (each equaling $\sqrt{1/d}$).

When $|\Psi\rangle$ is not maximally entangled, Alice cannot send d^2 messages via the original dense-coding protocol; in fact, Ji *et al.* [2] establish that she can send at most $d^2 - 2$ messages (a result suggested by numerical data in [3]). Wu *et al.* ask ([4], p. 10) whether Alice's inability to create $d^2 - 1$ messages (when $|\Psi\rangle$ is not maximally entangled) reflects a limitation

of the type of unitary encoding employed in the original protocol. We answer this question here, showing that the limitation of $d^2 - 2$ messages persists even when Alice uses as encoding operations the most general quantum operations possible. Alice performs a quantum operation on her qudit via the following three-step process: (i) she pairs her qudit with an ancillary quantum particle (of dimension ≥ 2 but otherwise arbitrary), (ii) she then applies a unitary operation to the pair, and (iii) she then either measures the ancilla or chooses not to measure the ancilla. Mathematically speaking, if in step (iii), she chooses not to measure the ancilla, then the quantum operation Alice has performed is trace preserving. If in step (iii), she does measure the ancilla, then her quantum operation may not preserve trace. For further information about quantum operators, the reader may consult Ref. ([5], Chapter 8).

This article is organized as follows. In the next section, we describe the mathematical framework for our results, providing a discussion of deterministic dense coding with quantum encoding operators. In Sec. III, we show that when the initial state $|\Psi\rangle$ of their two-qudit system is not maximally entangled, Alice cannot send $d^2 - 1$ messages to Bob under deterministic dense coding even when she may use quantum operators to encode messages. However, in Sec. IV, we introduce a new dense-coding procedure that allows some nonmaximally entangled states $|\Psi\rangle$ to support, with high probability, communication of $d^2 - 1$ perfectly distinguishable messages, with $d^2 - 2$ of the messages encoded by unitary operations on Alice's qudit and the final $(d^2 - 1)$ -th message encoded via a (non-trace-preserving) quantum operation. In fact, for any probability p as close as desired to 1, we show there is a nonmaximally entangled state $|\Psi\rangle$ that will support, with probability exceeding p , communication of $d^2 - 1$ messages under our protocol. Our protocol is designed so that Bob will never misinterpret a message; rather, there is a small chance he will receive no message. He will receive

*pbourdon@wlu.edu

†gerjuoy@pitt.edu

no message only when Alice wishes to send the $(d^2 - 1)$ -th message and her encoding procedure fails. We conclude the article with a detailed example illustrating how Alice and Bob can, with probability exceeding 97%, use a two-qubit system in state $|\Psi\rangle = \frac{9}{4\sqrt{10}}|00\rangle + \frac{\sqrt{79}}{4\sqrt{10}}|11\rangle$ as a resource for the communication of three perfectly distinguishable messages (via a two-dimensional noiseless quantum channel). In this example, the probability of success rises to more than 98% if Alice and Bob are willing to tolerate a small chance of Bob's incorrectly interpreting a message.

We note that the $d = 2$ case of the result of Sec. III of this article appears in Appendix A of [6].

II. BACKGROUND

A. Initial state

Let $H = H_A \otimes H_B$ be the state space of the two-qudit system that Alice and Bob share, where $(|0\rangle_A, |1\rangle_A, \dots, |d-1\rangle_A)$ is an orthonormal basis for H_A and $(|0\rangle_B, |1\rangle_B, \dots, |d-1\rangle_B)$ is an orthonormal basis for H_B . Note H has orthonormal basis $\mathcal{B} = \{|ij\rangle : 0 \leq i, j \leq d-1\}$, where we have used $|ij\rangle$ as a convenient substitute for $|i\rangle_A |j\rangle_B$. We assume that the initial state $|\Psi\rangle$ of Alice and Bob's two-qudit system has the Schmidt representation

$$|\Psi\rangle = \sum_{j=0}^{d-1} \sqrt{\lambda_j} |jj\rangle, \quad (1)$$

where the Schmidt coefficients $\sqrt{\lambda_0}, \sqrt{\lambda_1}, \dots, \sqrt{\lambda_{d-1}}$ have squares summing to 1 (which assures normalization) and where we assume, without loss of generality, that

$$\lambda_0 \geq \lambda_1 \geq \dots \geq \lambda_{d-1} \geq 0. \quad (2)$$

We frequently describe the initial state in terms of its density operator $|\Psi\rangle\langle\Psi|$.

Wu *et al.* ([4], Sec. IV.B) establish that when $|\Psi\rangle$ of Eq. (1) allows Alice to send to Bob L perfectly distinguishable messages (using any encoding scheme), then

$$\lambda_0 \leq \frac{d}{L}. \quad (3)$$

Suppose, for example, that $L > d(d-1)$; then the preceding inequality yields $\lambda_0 < 1/(d-1)$. Because $\sqrt{\lambda_0}$ is the largest Schmidt coefficient of $|\Psi\rangle$ and $\sum_{j=0}^{d-1} \lambda_j = 1$, it follows that if $\lambda_0 < 1/(d-1)$, then every Schmidt coefficient in Eq. (1) is nonzero. In particular, if we assume Alice is able to deterministically send $L = d^2 - 1$ messages, then, because $d^2 - 1 > d(d-1)$, all Schmidt coefficients of $|\Psi\rangle$ must be nonzero. *For the remainder of this article, we assume all Schmidt coefficients of $|\Psi\rangle$ are nonzero.*

B. Encoding operations

For any vector space W , we let $\mathcal{L}(W)$ denote the vector space of all linear operators on W .

Recall that Alice encodes messages for Bob, physically speaking, by applying a unitary operation to either (i) her qudit or, more generally, (ii) to her qudit paired with an ancillary particle, perhaps measuring the ancilla afterward. In either case, Alice's encoding action may be represented

mathematically by a quantum operator applied to $|\Psi\rangle\langle\Psi| \in \mathcal{L}(H)$:

$$|\Psi\rangle\langle\Psi| \mapsto \sum_{j=0}^{N-1} (K^{(j)} \otimes I_B) |\Psi\rangle\langle\Psi| (K^{(j)} \otimes I_B)^\dagger, \quad (4)$$

where N is a positive integer and the $K^{(j)}$'s are *Kraus operators* in $\mathcal{L}(H_A)$ satisfying

$$\sum_{j=0}^{N-1} (K^{(j)})^\dagger K^{(j)} \leq I_A. \quad (5)$$

In case (i), the sum on the right of relation (4) has only one summand and $K^{(0)} = U$, where U is a unitary operator on H_A . Thus, in case (i) the inequality (5) is an equality. It is also an equality in case (ii) (see, e.g., Appendix B) provided Alice does not measure the ancilla.

As we explain in Sec. II E, it is easy to see that Alice's ability to measure the ancilla can never be used to increase the number of messages she can send to Bob through deterministic dense coding. Thus, for now, we assume that Alice does not measure the ancillary particle, which means that in either case (i) or (ii), Alice's encoding action is described by a quantum operator \mathcal{E} having the *operator sum representation*

$$\mathcal{E}(\rho) = \sum_{j=0}^{N-1} (K^{(j)} \otimes I_B) \rho (K^{(j)} \otimes I_B)^\dagger, \quad \rho \in \mathcal{L}(H), \quad (6)$$

where the $K^{(j)}$'s satisfy

$$\sum_{j=0}^{N-1} (K^{(j)})^\dagger K^{(j)} = I_A, \quad (7)$$

making \mathcal{E} *trace preserving*.

The quantum operator \mathcal{E} of Eq. (6) has many different operator sum representations (see, e.g., [5], Theorem 8.2). Among these representations, there is one for which the number N of Kraus-operator elements assumes its minimum possible value m . This number m is the *Kraus rank* of \mathcal{E} , and it is easy to see that any group of exactly m Kraus-operator elements representing \mathcal{E} must be linearly independent in $\mathcal{L}(H_A)$ [which is equivalent to the linear independence of $K^{(0)} \otimes I_B, K^{(1)} \otimes I_B, \dots, K^{(m-1)} \otimes I_B$ in $\mathcal{L}(H)$].

We now state an important lemma for our work; its (short) proof occupies Appendix A.

Lemma II.1. Suppose that all the Schmidt coefficients of $|\Psi\rangle$ in the representation (1) are nonzero and that $K^{(0)}, K^{(1)}, \dots, K^{(m-1)}$ are linearly independent in $\mathcal{L}(H_A)$; then $(K^{(0)} \otimes I_B)|\Psi\rangle, (K^{(1)} \otimes I_B)|\Psi\rangle, \dots, (K^{(m-1)} \otimes I_B)|\Psi\rangle$ are linearly independent vectors in H .

C. Perfect distinguishability

To send L perfectly distinguishable messages to Bob, Alice must be able to perform L encoding operations on her qudit (perhaps paired with an ancilla) with each such operation producing a message that Bob will recognize and never mistake for a message corresponding to another encoding operation. Each of the L encoding operations results in a density-operator description of the two-qudit system that Alice and Bob

share: $\rho_0, \rho_1, \dots, \rho_{L-1}$. The perfect distinguishability of the corresponding messages means that

$$\rho_i \rho_j = 0 \quad \text{whenever} \quad i \neq j; \quad (8)$$

that is, the supports of the density operators representing messages must be orthogonal. (For a formal proof that orthogonality of the supports provides perfect distinguishability, see Ref. ([7], Theorem 1).)

Observe that the support of the density operator $\mathcal{E}(|\Psi\rangle\langle\Psi|)$ representing the message encoded by the trace-preserving quantum operator (6) is precisely the linear span of

$$\{(K^{(0)} \otimes I_B)|\Psi\rangle, (K^{(1)} \otimes I_B)|\Psi\rangle, \dots, (K^{(N-1)} \otimes I_B)|\Psi\rangle\}. \quad (9)$$

Thus, the perfect distinguishability of the messages Alice produces via her quantum encoding operations amounts to the following at the Kraus-operator level: Messages produced by distinct quantum operations \mathcal{E}_1 and \mathcal{E}_2 are perfectly distinguishable if and only if whenever K_1 is a Kraus operator for \mathcal{E}_1 and K_2 is a Kraus operator for \mathcal{E}_2 , then $\langle\Psi|(K_1 \otimes I_B)^\dagger(K_2 \otimes I_B)|\Psi\rangle = 0$.

D. Decoding messages

Suppose that Alice is able to encode L perfectly distinguishable messages, represented by density operators $\rho_0, \rho_1, \dots, \rho_{L-1}$. For each $j = 0, 1, \dots, L - 1$, let S_j be the support of ρ_j , so that by Eq. (8), $\{S_j\}_{j=0}^{L-1}$ is a collection of pairwise orthogonal subspaces of H . To send message j to Bob, Alice performs the quantum operation \mathcal{E} that creates the state ρ_j and sends her qudit to Bob through a noiseless channel, keeping any ancillary particle she may have used in executing \mathcal{E} . To decode Alice’s message, Bob simply performs a projective measurement described by the observable

$$\sum_{j=0}^{L-1} j P_{S_j},$$

which is equivalent to the observable $\sum_{j=0}^{L-1} j(P_{S_j} \otimes I_a)$, where I_a is the identity on the Hilbert space H_a of the ancilla. The pairwise orthogonality of the subspaces S_j ensures Bob will measure j precisely when message j has been sent, that is, precisely when Alice has created the “message state” ρ_j .

E. Ancilla measurement

Suppose that Alice pairs her qudit A with an ancillary N -level particle a . The Hilbert space for the Aa pair is $H_A \otimes H_a$, with orthonormal basis $\{|i\rangle_A |j\rangle_a : i = 0, 1, 2, \dots, d - 1; j = 0, 1, 2, \dots, N - 1\}$. Assume, as before, that Alice’s particle A is entangled with Bob’s particle B and that their two-qudit system is in state $|\Psi\rangle$ given by Eq. (1). Assume that a is in state $|0\rangle_a$. Suppose Alice performs a unitary operation U on the pair Aa . The effect of Alice’s unitary operator U on the state $|\Psi\rangle|0\rangle_a$ may be described as follows (see Appendix B):

$$|\Psi\rangle|0\rangle_a \mapsto \sum_{j=0}^{N-1} (K^{(j)} \otimes I_B)|\Psi\rangle \otimes |j\rangle_a, \quad (10)$$

where the $K^{(j)}$ ’s are Kraus operators that satisfy Eq. (7).

Forming the density operator corresponding to relation (10) and taking the partial trace over the ancillary system produces the density operator on the right of relation (4), which describes the message state that Alice creates using U . Denote by \mathcal{E} the corresponding trace-preserving quantum operation on $\mathcal{L}(H)$ —it has the form (6), and the density operator $\mathcal{E}(|\Psi\rangle\langle\Psi|)$ for the encoded message has support equal to the linear span of the set (9). Suppose that Alice applies U and then performs a measurement of the ancilla a described by the collection $\{M_x\}$ of measurement operators. Assuming y is the outcome of the measurement and recalling that the application of U to the state $|\Psi\rangle|0\rangle_a$ yields the state (10), we see the state of the ABa system after measurement is

$$c \sum_{j=0}^{N-1} (K^{(j)} \otimes I_B)|\Psi\rangle \otimes M_y |j\rangle_a, \quad (11)$$

where c is a normalizing constant. Because $M_y |j\rangle_a = \sum_{i=0}^{N-1} \alpha_{ij}^{(y)} |i\rangle_a$ for some collection of scalars $\alpha_{ij}^{(y)}$, expression (11) may be written

$$c \sum_{i=0}^{N-1} \left[\sum_{j=0}^{N-1} \alpha_{ij}^{(y)} (K^{(j)} \otimes I_B)|\Psi\rangle \right] \otimes |i\rangle_a,$$

which corresponds to a density operator ρ on $\mathcal{L}(H)$ whose support will be contained in that of $\mathcal{E}(|\Psi\rangle\langle\Psi|)$, because the support of ρ consists of linear combinations of vectors of the form $\sum_{j=0}^{N-1} \alpha_{ij}^{(y)} (K^{(j)} \otimes I_B)|\Psi\rangle$, each one of which is in the support of $\mathcal{E}(|\Psi\rangle\langle\Psi|)$. Hence if Alice measures the ancilla before she sends her message to Bob, then he will still receive the intended message. Equally important is that because the measurement operators $\{M_x\}$ satisfy the completeness relation $\sum_x M_x^\dagger M_x = I_a$, Alice cannot predetermine some proper subspace S of the support of $\mathcal{E}(|\Psi\rangle\langle\Psi|)$ and use measurement of the ancilla to produce *with certainty* a state of the AB system whose density-operator description ρ has support contained in S . It follows that Alice cannot use measurement of ancillary particles during the encoding process to increase the number of messages she may send to Bob via deterministic dense coding.

We now turn to our main results.

III. ALICE CANNOT ENCODE $d^2 - 1$ MESSAGES WHEN $|\Psi\rangle$ IS NOT MAXIMALLY ENTANGLED

We suppose that Alice can use quantum operators $\mathcal{E}_0, \mathcal{E}_1, \dots, \mathcal{E}_{d^2-2}$ to encode $d^2 - 1$ perfectly distinguishable messages for Bob and prove that their initial two-qudit state $|\Psi\rangle$ must be maximally entangled. We have established that we may assume, without loss of generality, that Alice’s encoding quantum operators are trace-preserving. For $j = 0, 1, \dots, d^2 - 2$, let m_j be the Kraus rank of \mathcal{E}_j . Thanks to Lemma II.1 and the discussion following Eq. (3), the support of the density operator $\mathcal{E}_j(|\Psi\rangle\langle\Psi|)$ must have dimension m_j . We know that perfect distinguishability of messages means that the supports of the density operators $\mathcal{E}_0(|\Psi\rangle\langle\Psi|), \mathcal{E}_1(|\Psi\rangle\langle\Psi|), \dots, \mathcal{E}_{d^2-2}(|\Psi\rangle\langle\Psi|)$ must be pairwise orthogonal. These supports are subspaces of the d^2 dimensional space

$H = H_A \otimes H_B$. Thus,

$$\sum_{j=0}^{d^2-2} m_j \leq d^2. \quad (12)$$

Because $m_j \geq 1$ for each j , the preceding inequality shows that $m_j = 1$ for all but 1 of the j 's, and for the remaining j value, either $m_j = 1$ or $m_j = 2$. Note that if $m_j = 1$, then because \mathcal{E}_j is trace preserving [i.e., its Kraus-operator elements satisfy (7)], we see that \mathcal{E}_j is an original-protocol unitary encoding operation. Thus, if $m_j = 1$ for every j , then Alice can send to Bob $d^2 - 1$ messages via original-protocol unitary encoding, and, as we indicated earlier, Ji *et al.* have shown that in this case $|\Psi\rangle$ must be maximally entangled. Thus, to complete the argument, we must show that $|\Psi\rangle$ must also be maximally entangled in case $m_j = 2$ for some j and the rest of the m_j 's equal 1.

Without loss of generality, we assume that $m_{d^2-2} = 2$ so that \mathcal{E}_{d^2-2} may be expressed in the form

$$\mathcal{E}_{d^2-2}(\rho) = (K^{(0)} \otimes I_B)\rho(K^{(0)} \otimes I_B)^\dagger + (K^{(1)} \otimes I_B)\rho(K^{(1)} \otimes I_B)^\dagger,$$

where $K^{(0)}$ and $K^{(1)}$ are linearly independent Kraus operators satisfying

$$(K^{(0)})^\dagger K^{(0)} + (K^{(1)})^\dagger K^{(1)} = I_A. \quad (13)$$

Each of the remaining encoding operators $\mathcal{E}_0, \mathcal{E}_1, \dots, \mathcal{E}_{d^2-3}$ is an original protocol unitary operation:

$$\mathcal{E}_j(\rho) = (U^{(j)} \otimes I_B)\rho(U^{(j)} \otimes I_B)^\dagger,$$

where for $j = 0, 1, \dots, d^2 - 3$, $U^{(j)}$ is a unitary operator on H_A . Perfect distinguishability of the messages produced by $\mathcal{E}_0, \dots, \mathcal{E}_{d^2-3}$, that is, pairwise orthogonality of support sets of the corresponding density operators, means that

$$\{(U^{(0)} \otimes I_B)|\Psi\rangle, (U^{(1)} \otimes I_B)|\Psi\rangle, \dots, (U^{(d^2-3)} \otimes I_B)|\Psi\rangle\} \quad (14)$$

is an orthogonal set in H . Moreover, the support of $\mathcal{E}_{d^2-2}(|\Psi\rangle\langle\Psi|)$, representing the ‘‘final message,’’ must be orthogonal to the subspace of H spanned by the vectors in the set (14). Equivalently, each of

$$|\phi_0\rangle \equiv (K^{(0)} \otimes I_B)|\Psi\rangle \quad \text{and} \quad |\phi_1\rangle \equiv (K^{(1)} \otimes I_B)|\Psi\rangle \quad (15)$$

is orthogonal to every vector in the set (14).

In fact, we can even assume the two vectors $|\phi_0\rangle$ and $|\phi_1\rangle$ representing the $(d^2 - 1)$ -th message are orthogonal to each other. We justify this claim in Appendix C. Thus, henceforth we assume that the two Kraus states $|\phi_0\rangle$ and $|\phi_1\rangle$ representing the $(d^2 - 1)$ -th message are orthogonal to each other as well as to each element of the set (14). Note well that $|\phi_0\rangle$ and $|\phi_1\rangle$ are linearly independent vectors (by Lemma II.1); thus, in particular, neither is the zero vector.

Though they are orthogonal, $|\phi_0\rangle$ and $|\phi_1\rangle$ are not normalized. Because $\langle\phi_0|\phi_0\rangle + \langle\phi_1|\phi_1\rangle = 1$ [via Eq. (13)], if we set

$$x = \langle\phi_0|\phi_0\rangle, \quad (16)$$

then $0 < x < 1$ (both $|\phi_0\rangle$ and $|\phi_1\rangle$ are nonzero) and the pair of ‘‘Kraus states’’ spanning the support of $\mathcal{E}_{d^2-2}(|\Psi\rangle\langle\Psi|)$ are

$$|\phi_0\rangle/\sqrt{x} \quad \text{and} \quad |\phi_1\rangle/\sqrt{1-x}.$$

We assume, without loss of generality, that $0 < x \leq 1/2$ (otherwise we can just switch labels on $|\phi_0\rangle$ and $|\phi_1\rangle$).

We complete the proof by establishing that the existence of the following orthonormal subset of H forces all Schmidt coefficients of $|\Psi\rangle$ to have the same value, so that $|\Psi\rangle$ is maximally entangled:

$$\{(U^{(0)} \otimes I_B)|\Psi\rangle, (U^{(1)} \otimes I_B)|\Psi\rangle, \dots, (U^{(d^2-3)} \otimes I_B)|\Psi\rangle, |\phi_0\rangle/\sqrt{x}, |\phi_1\rangle/\sqrt{1-x}\}. \quad (17)$$

We view the operators $K^{(q)}$ of relation (15) as well as the operators $U^{(n)}$ of (14)—all of which are operators on the d -dimensional Hilbert space H_A —as $d \times d$ matrices with respect to the basis $(|0\rangle_A, |1\rangle_A, \dots, |d-1\rangle_A)$, and we let $u_{ij}^{(n)}$ and $k_{ij}^{(q)}$ denote the entries of these matrices. For $j = 0, 1, 2, \dots, d-1$ and $q = 0, 1$, let

$${}_j K^{(q)}$$

denote the j -th column of the $d \times d$ matrix $K^{(q)}$. Note well that the initial column of $K^{(q)}$ is ${}_0 K^{(q)}$. From Eq. (13), we conclude

$$\langle{}_i K^{(0)}|{}_j K^{(0)}\rangle + \langle{}_i K^{(1)}|{}_j K^{(1)}\rangle = \delta_{ij}. \quad (18)$$

Define $d \times d$ matrices as

$$E = \frac{1}{\sqrt{x}} K^{(0)} \quad \text{and} \quad W = \frac{1}{\sqrt{1-x}} K^{(1)}. \quad (19)$$

We now order the basis for the Hilbert space $H_A \otimes H_B$ of the Alice-Bob system:

$$\mathcal{B} = (|00\rangle, |10\rangle, |2, 0\rangle, \dots, |d-1, 0\rangle, |01\rangle, |11\rangle, \dots, |d-1, 1\rangle, \dots, |d-1, d-1\rangle). \quad (20)$$

Thus the basis elements are listed in d groups of d elements with the ordering of the groups determined by the second of the pair $|ij\rangle$ and the ordering within the groups determined by the first of the pair. This is the ordering used by Gerjuoy *et al.* in [8] to form an *augmented message matrix* for an original-protocol unitary encoding of messages. Gerjuoy *et al.* use the augmented message matrix (for example) to present an alternate proof of the result of Ji *et al.* establishing that $d^2 - 1$ messages cannot be produced by original-protocol unitary encoding. We use a similar matrix M later.

View $|\phi_0\rangle/\sqrt{x}$ and $|\phi_1\rangle/\sqrt{1-x}$ as column vectors consisting, respectively, of the coordinates of $|\phi_0\rangle/\sqrt{x}$ and $|\phi_1\rangle/\sqrt{1-x}$ with respect to the basis \mathcal{B} :

$$\left[\frac{|\phi_0\rangle}{\sqrt{x}} \right]_{ij} = \frac{\sqrt{\lambda_j}}{\sqrt{x}} k_{ij}^{(0)} \quad \text{and} \quad \left[\frac{|\phi_1\rangle}{\sqrt{1-x}} \right]_{ij} = \frac{\sqrt{\lambda_j}}{\sqrt{1-x}} k_{ij}^{(1)}.$$

The natural ordering for the entries of these vectors is provided by the ordering of the basis pairings in \mathcal{B} . Thus, 00 is the initial entry (followed by 10; 20; \dots ; $d-1, 0$), the $d+1$ entry is 01, and the final entry is $d-1, d-1$. Thus the first d entries of $|\phi_0\rangle/\sqrt{x}$ constitute the column $(\sqrt{\lambda_0}/\sqrt{x})({}_0 K^{(0)})$, the next d constitute $(\sqrt{\lambda_1}/\sqrt{x})({}_1 K^{(0)})$, and so on. Similarly, the first d entries of $|\phi_1\rangle/\sqrt{1-x}$ constitute $(\sqrt{\lambda_0}/\sqrt{1-x})({}_0 K^{(1)})$, the next d constitute $(\sqrt{\lambda_1}/\sqrt{1-x})({}_1 K^{(1)})$, and so on.

Form the $d^2 \times d^2$ matrix M whose first $d^2 - 2$ columns are, in order, the coordinates with respect to \mathcal{B} of $(U^{(n)} \otimes I_B)|\Psi\rangle$, $n = 0, 1, \dots, d^2 - 3$, and whose final two columns are $|\phi_0\rangle/\sqrt{x}$ (penultimate) and $|\phi_1\rangle/\sqrt{1-x}$. Thus M is the natural $d^2 \times d^2$ matrix corresponding to the orthonormal set (17). Because the columns of M constitute an orthonormal basis of \mathbb{C}^{d^2} , M is unitary.

Using the structure of M , we now prove that if $0 \leq i, j \leq d - 1$ and $i \neq j$, then

$$\frac{\sqrt{\lambda_i \lambda_j}}{x} \langle_i K^{(0)} |_j K^{(0)} \rangle + \frac{\sqrt{\lambda_i \lambda_j}}{1-x} \langle_i K^{(1)} |_j K^{(1)} \rangle = 0. \quad (21)$$

Because the matrix M is unitary, the inner product of each pair of distinct rows of M is 0. Thus if i, j are distinct elements of $\{0, 1, \dots, d - 1\}$, then upon taking the inner products of rows labeled by si and s_j and next summing over s , we have

$$\begin{aligned} 0 &= \sum_{s=0}^{d-1} \left[\sqrt{\lambda_i \lambda_j} \sum_{n=0}^{d^2-3} (u_{si}^{(n)})^* u_{sj}^{(n)} \right] + \sum_{s=0}^{d-1} \frac{\sqrt{\lambda_i \lambda_j}}{x} (k_{si}^{(0)})^* k_{sj}^{(0)} \\ &\quad + \sum_{s=0}^{d-1} \frac{\sqrt{\lambda_i \lambda_j}}{1-x} (k_{si}^{(1)})^* k_{sj}^{(1)} \\ &= \sum_{n=0}^{d^2-3} \left[\sqrt{\lambda_i \lambda_j} \sum_{s=0}^{d-1} (u_{si}^{(n)})^* u_{sj}^{(n)} \right] + \frac{\sqrt{\lambda_i \lambda_j}}{x} \langle_i K^{(0)} |_j K^{(0)} \rangle \\ &\quad + \frac{\sqrt{\lambda_i \lambda_j}}{1-x} \langle_i K^{(1)} |_j K^{(1)} \rangle \\ &= \frac{\sqrt{\lambda_i \lambda_j}}{x} \langle_i K^{(0)} |_j K^{(0)} \rangle + \frac{\sqrt{\lambda_i \lambda_j}}{1-x} \langle_i K^{(1)} |_j K^{(1)} \rangle, \end{aligned}$$

where the final equality holds because for each $n \in \{0, 1, 2, \dots, d^2 - 3\}$, the matrix $U^{(n)}$ is unitary (in particular, its columns are orthogonal). We have proved Eq. (21). Continuing to assume $i \neq j$, we now combine Eqs. (21) and (18) and record the result in matrix-equation form:

$$\begin{pmatrix} 1 & 1 \\ \frac{\sqrt{\lambda_i \lambda_j}}{x} & \frac{\sqrt{\lambda_i \lambda_j}}{1-x} \end{pmatrix} \begin{pmatrix} \langle_i K^{(0)} |_j K^{(0)} \rangle \\ \langle_i K^{(1)} |_j K^{(1)} \rangle \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (22)$$

The determinant of the matrix of coefficients on the left of the preceding equation is $\sqrt{\lambda_i \lambda_j} [(1/(1-x)) - (1/x)]$. There are two possibilities: (i) either this determinant is 0, in which case $x = 1/2$ (because $\lambda_i \lambda_j$ is nonzero), or (ii) this determinant is nonzero, in which case Eq. (22) shows that we must have

$$\langle_i K^{(0)} |_j K^{(0)} \rangle = 0 \quad \text{and} \quad \langle_i K^{(1)} |_j K^{(1)} \rangle = 0 \quad (23)$$

for all pairs of distinct i and j in $\{0, 1, 2, \dots, d - 1\}$. We show that in both cases (i) and (ii), $|\Psi\rangle$ must be a maximally entangled state, completing the argument.

A. Case (i): $x = 1/2$

Unitarity of M implies each of its rows has length one; thus, for every i and j in $\{0, 1, 2, \dots, d - 1\}$, we have $\sum_{n=0}^{d^2-1} |M_{ij,n}|^2 = 1$; equivalently,

$$\lambda_j \left(\sum_{n=0}^{d^2-3} |u_{ij}^{(n)}|^2 + \frac{|k_{ij}^{(0)}|^2}{x} + \frac{|k_{ij}^{(1)}|^2}{1-x} \right) = 1. \quad (24)$$

Restricting attention to diagonal entries, we see that Eq. (13) reveals

$$\sum_{i=0}^{d-1} (|k_{ij}^{(0)}|^2 + |k_{ij}^{(1)}|^2) = 1$$

for each $j \in \{0, 1, \dots, d - 1\}$. Thus, if we set

$$b_j = \sum_{i=0}^{d-1} |k_{ij}^{(0)}|^2, \quad \text{then} \quad \sum_{i=0}^{d-1} |k_{ij}^{(1)}|^2 = 1 - b_j. \quad (25)$$

Now sum both sides of Eq. (24) from $i = 0$ to $d - 1$ and use the unitarity of $U^{(n)}$ for each n to see that

$$\lambda_j \left(d^2 - 2 + \frac{b_j}{x} + \frac{1 - b_j}{1 - x} \right) = d \quad (26)$$

for each j . Because we are assuming $x = 1/2$ for case (i), the expression in parentheses on the left of the preceding equation simplifies to d^2 , and the equation yields $\lambda_j d^2 = d$ for every j , that is, $\lambda_j = 1/d$ for every j . We have shown in this case that $|\Psi\rangle$ is maximally entangled.

B. Case (ii): $0 < x < 1/2$

Equation (23) holds in this case so that distinct columns of the Kraus matrix $K^{(0)}$ are orthogonal, and the same is true of $K^{(1)}$. Thus, recalling the definitions of the matrices E and W from Eqs. (19) as well as the notation introduced in relation (25), we have

$$E^\dagger E = \begin{pmatrix} \frac{b_0}{x} & 0 & 0 & \dots & 0 \\ 0 & \frac{b_1}{x} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & \frac{b_{d-1}}{x} \end{pmatrix} \quad (27)$$

and

$$W^\dagger W = \begin{pmatrix} \frac{1-b_0}{1-x} & 0 & 0 & \dots & 0 \\ 0 & \frac{1-b_1}{1-x} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & \frac{1-b_{d-1}}{1-x} \end{pmatrix}. \quad (28)$$

We need to exploit further the structure of the unitary matrix M . Note that Eq. (24) may be rewritten as

$$\sum_{n=0}^{d^2-3} |u_{ij}^{(n)}|^2 + |k_{ij}^{(0)}|^2/x + |k_{ij}^{(1)}|^2/(1-x) = \frac{1}{\lambda_j}. \quad (29)$$

Fix i in Eq. (29) and sum both sides from $j = 0$ to $j = d - 1$; use the fact that the rows of each $U^{(n)}$ matrix all have length one to obtain

$$\frac{1}{x} \sum_{j=0}^{d-1} |k_{ij}^{(0)}|^2 + \frac{1}{1-x} \sum_{j=0}^{d-1} |k_{ij}^{(1)}|^2 = \sum_{j=0}^{d-1} \frac{1}{\lambda_j} - (d^2 - 2). \quad (30)$$

Note well that it follows from the preceding equation that the diagonal entries of the matrix

$$EE^\dagger + WW^\dagger \quad (31)$$

all have the common value $\sum_{j=0}^{d-1} \frac{1}{\lambda_j} - (d^2 - 2)$. We claim that every off-diagonal entry of $EE^\dagger + WW^\dagger$ is 0.

Let p and q be distinct integers in $\{0, 1, 2, \dots, d-1\}$. Fix $j \in \{0, 1, 2, \dots, d-1\}$ and take the inner product of rows q, j and p, j of M to obtain

$$0 = \lambda_j \left[\sum_{n=0}^{d^2-3} u_{pj}^{(n)} (u_{qj}^{(n)})^* + \frac{1}{x} k_{pj}^{(0)} (k_{qj}^{(0)})^* + \frac{1}{1-x} k_{pj}^{(1)} (k_{qj}^{(1)})^* \right]$$

so that

$$0 = \sum_{n=0}^{d^2-3} u_{pj}^{(n)} (u_{qj}^{(n)})^* + \frac{1}{x} k_{pj}^{(0)} (k_{qj}^{(0)})^* + \frac{1}{1-x} k_{pj}^{(1)} (k_{qj}^{(1)})^*.$$

Now sum both sides of the preceding equation from $j = 0$ to $j = d-1$ and use the orthogonality of row p of $U^{(n)}$ and row q of $U^{(n)}$ for each n to obtain

$$\begin{aligned} 0 &= \sum_{j=0}^{d-1} \left[\frac{1}{x} k_{pj}^{(0)} (k_{qj}^{(0)})^* + \frac{1}{1-x} k_{pj}^{(1)} (k_{qj}^{(1)})^* \right] \\ &= (EE^\dagger + WW^\dagger)_{pq}, \end{aligned}$$

and it follows that $EE^\dagger + WW^\dagger$ is a diagonal matrix. Using our earlier observation that all the diagonal entries of $EE^\dagger + WW^\dagger$ have common value $\gamma \equiv \sum_{j=0}^{d-1} \frac{1}{\lambda_j} - (d^2 - 2)$, we see that

$$EE^\dagger + WW^\dagger = \gamma I. \tag{32}$$

The conclusion of our argument relies upon the following observation arising from the *polar decomposition* (see, e.g., Theorem 2.3 on page 78 of [5]). For any $n \times n$ matrix Y , we know there is a unitary matrix U such that

$$Y = U\sqrt{Y^\dagger Y} = \sqrt{Y Y^\dagger}U.$$

Thus

$$\sqrt{Y Y^\dagger} = U\sqrt{Y^\dagger Y}U^\dagger,$$

and squaring both sides of the preceding equation yields

$$Y Y^\dagger = U Y^\dagger Y U^\dagger.$$

Thus $Y Y^\dagger$ and $Y^\dagger Y$ are unitarily equivalent, and they have the same eigenvalues counting multiplicities. In fact, suppose that λ is an eigenvalue for $Y Y^\dagger$ with corresponding eigenvector \mathbf{v} . Then

$$\begin{aligned} 0 &= (Y Y^\dagger - \lambda I)\mathbf{v} = (U Y^\dagger Y U^\dagger - \lambda I)\mathbf{v} \\ &= U(Y^\dagger Y - \lambda I)U^\dagger \mathbf{v}. \end{aligned}$$

We see that $0 = U(Y^\dagger Y - \lambda I)U^\dagger \mathbf{v}$, and multiplying both sides on the left by U^\dagger yields

$$0 = (Y^\dagger Y - \lambda I)U^\dagger \mathbf{v}.$$

Thus $U^\dagger \mathbf{v}$ is an eigenvector for $Y^\dagger Y$ with eigenvalue λ . The bottom line is that $Y^\dagger Y$ and $Y Y^\dagger$ always have the same eigenvalues counting multiplicities. We apply this fact to the pairs $E^\dagger E$ and EE^\dagger and $W^\dagger W$ and WW^\dagger .

Equation (27) shows us that the set of eigenvalues of $E^\dagger E$ is

$$\{b_0/x, b_1/x, \dots, b_{d-1}/x\}$$

and Eq. (28) shows us that the set of eigenvalues of $W^\dagger W$ is

$$\left\{ \frac{1-b_0}{1-x}, \frac{1-b_1}{1-x}, \dots, \frac{1-b_{d-1}}{1-x} \right\}. \tag{33}$$

Because for $j = 0, 1, 2, \dots, d-1$, b_j/x is an eigenvalue of $E^\dagger E$, it is also an eigenvalue of EE^\dagger . Let \mathbf{v}_j be an eigenvector for EE^\dagger with corresponding eigenvalue b_j/x . By applying both sides of Eq. (32) to \mathbf{v}_j and doing a bit of rearranging, we obtain

$$WW^\dagger \mathbf{v}_j = \left(\gamma - \frac{b_j}{x} \right) \mathbf{v}_j.$$

Thus $\gamma - (b_j/x)$ is an eigenvalue of WW^\dagger . In fact, it is easy to see the set of eigenvalues of WW^\dagger is precisely

$$\left\{ \gamma - \frac{b_0}{x}, \gamma - \frac{b_1}{x}, \dots, \gamma - \frac{b_{d-1}}{x} \right\}.$$

From Eq. (26), we have

$$d^2 - 2 + \frac{b_j}{x} + \frac{1-b_j}{1-x} = \frac{d}{\lambda_j},$$

so that

$$b_j = -\frac{x}{1-2x} + \frac{x(1-x)}{1-2x} \left(\frac{d}{\lambda_j} + 2 - d^2 \right). \tag{34}$$

Because the sequence λ_j decreases with j , the preceding equation shows that b_j increases with j :

$$b_0 \leq b_1 \leq \dots \leq b_{d-1}. \tag{35}$$

Thus, the work of the preceding paragraph shows that

$$\gamma - \frac{b_0}{x}$$

must be the largest eigenvalue of WW^\dagger ; equivalently, the largest eigenvalue of $W^\dagger W$ (because WW^\dagger and $W^\dagger W$ share the same eigenvalues). Hence, recalling our earlier listing (33) of the eigenvalues and (35), we must have

$$\frac{1-b_0}{1-x} = \gamma - \frac{b_0}{x}.$$

By rearranging and using $\gamma = \sum_{j=0}^{d-1} \frac{1}{\lambda_j} - (d^2 - 2)$, we see that the preceding equation yields

$$\sum_{j=0}^{d-1} \frac{1}{\lambda_j} - (d^2 - 2) = \frac{1-b_0}{1-x} + \frac{b_0}{x}.$$

Now substitute the left-hand side of the preceding equation for the right-hand side in Eq. (26) in the $j = 0$ case:

$$\lambda_0 \left\{ d^2 - 2 + \left[\sum_{j=0}^{d-1} \frac{1}{\lambda_j} - (d^2 - 2) \right] \right\} = d,$$

so that

$$\sum_{j=0}^{d-1} \frac{\lambda_0}{\lambda_j} = d.$$

We know that $\lambda_0 \geq \lambda_j$ for $j = 0, 1, 2, \dots, d-1$, so that $\lambda_0/\lambda_j \geq 1$ for each j . From the preceding equation, we conclude that $\lambda_0/\lambda_j = 1$ for all j ; that is, all Schmidt coefficients of $|\Psi\rangle$ are equal, which means $|\Psi\rangle$ must be maximally entangled, as was to be proved.

IV. ALICE CAN ENCODE $d^2 - 1$ MESSAGES WITH HIGH PROBABILITY FOR CERTAIN NONMAXIMALLY ENTANGLED STATES

We continue to assume that Alice and Bob share a two-qudit system in state $|\Psi\rangle$ of (1) with largest Schmidt coefficient $\sqrt{\lambda_0}$. In this section, we show that for certain nonmaximally entangled states (those for which λ_0 is small but still exceeds $1/d$), Alice can encode $d^2 - 2$ perfectly distinguishable messages via original-protocol unitary encoding and, with high probability, can use a non-trace-preserving quantum operation to encode a $(d^2 - 1)$ -th message perfectly distinguishable from those encoded according to the original protocol.

We assume throughout this section that $|\Psi\rangle$ of (1) is not maximally entangled ($\lambda_0 > 1/d$) yet it is entangled enough to permit Alice to send to Bob $d^2 - 2$ perfectly distinguishable messages via unitary encoding operators $U^{(n)}$ on H_A , $n = 0, 1, \dots, d^2 - 3$. The result (3) of Wu *et al.* shows that $\lambda_0 \leq [d/(d^2 - 2)]$. We assume that λ_0 is strictly less than $d/(d^2 - 2)$ and hence $\lambda_j < d/(d^2 - 2)$ for $j = 0, 1, 2, \dots, d - 1$. It follows that all Schmidt coefficients of $|\Psi\rangle$ must be nonzero (consistent with our standing assumption); in addition, we have for all j that

$$R_j \equiv d - (d^2 - 2)\lambda_j \tag{36}$$

is positive. Note well that R_j increases with j : $R_0 \leq R_1 \leq \dots \leq R_{d-1}$.

We continue to assume that the natural basis \mathcal{B} for the Alice-Bob system is ordered as in Eq. (20).

We construct a $d^2 \times d^2$ matrix M as follows. Just as in the preceding section, the coordinate vectors of $(U^{(n)} \otimes I_B)|\Psi\rangle$, $n = 0, 1, \dots, d^2 - 3$, relative to the basis \mathcal{B} , give the first $d^2 - 2$ columns of M . By adding two appropriately chosen columns, we can extend these $d^2 - 2$ orthonormal columns to form a $d^2 \times d^2$ unitary matrix M . Let \mathbf{v} be the penultimate column and \mathbf{w} be the final column of this unitary matrix M . Label the entries of \mathbf{v} and \mathbf{w} according to the ordering of \mathcal{B} (just as we labeled the entries of the column vectors for $|\phi_0\rangle/\sqrt{x}$ and $|\phi_1\rangle/\sqrt{1-x}$ in the preceding section) so that, for example, the entries of \mathbf{v} are $v_{00}, v_{10}, v_{20}, \dots, v_{d-1,0}, v_{01}, v_{11}, \dots, v_{d-1,1}, \dots, v_{d-1,d-1}$.

Form a $d \times d$ matrix T such that

$$t_{ij} = \frac{\sqrt{\lambda_{d-1}}}{\sqrt{\lambda_j} \sqrt{R_{d-1}}} v_{ij},$$

where R_j is defined by relation (36). For example, if $d = 3$, we have

$$T = \begin{pmatrix} \frac{\sqrt{\lambda_2}}{\sqrt{R_2\lambda_0}} v_{00} & \frac{\sqrt{\lambda_2}}{\sqrt{R_2\lambda_1}} v_{01} & \frac{1}{\sqrt{R_2}} v_{02} \\ \frac{\sqrt{\lambda_2}}{\sqrt{R_2\lambda_0}} v_{10} & \frac{\sqrt{\lambda_2}}{\sqrt{R_2\lambda_1}} v_{11} & \frac{1}{\sqrt{R_2}} v_{12} \\ \frac{\sqrt{\lambda_2}}{\sqrt{R_2\lambda_0}} v_{20} & \frac{\sqrt{\lambda_2}}{\sqrt{R_2\lambda_1}} v_{21} & \frac{1}{\sqrt{R_2}} v_{22} \end{pmatrix}.$$

Similarly, define the $d \times d$ matrix Y by

$$y_{ij} = \frac{\sqrt{\lambda_{d-1}}}{\sqrt{\lambda_j} \sqrt{R_{d-1}}} w_{ij}.$$

Claim. $T^\dagger T + Y^\dagger Y \leq I$, that is, $I - T^\dagger T - Y^\dagger Y$ is a positive matrix.

(Here I is the $d \times d$ identity matrix.) We now justify this claim. Using the structure of the matrix M , we sum the squared magnitudes of rows $0, j$ through $d - 1, j$ to get

$$(d^2 - 2)\lambda_j + \sum_{i=0}^{d-1} |v_{ij}|^2 + \sum_{i=0}^{d-1} |w_{ij}|^2 = d,$$

where we have used the facts that the rows of M have length one and that the entries $0, j$ through $d - 1, j$ in any one of the first $d^2 - 2$ columns of M constitute the column of a unitary matrix. Thus,

$$\sum_{i=0}^{d-1} |v_{ij}|^2 + \sum_{i=0}^{d-1} |w_{ij}|^2 = d - (d^2 - 2)\lambda_j = R_j. \tag{37}$$

Note that for $j = 0, 1, \dots, d - 1$,

$$\begin{aligned} (T^\dagger T + Y^\dagger Y)_{jj} &= \frac{\lambda_{d-1}}{\lambda_j} \frac{1}{R_{d-1}} \left(\sum_{i=0}^{d-1} |v_{ij}|^2 + \sum_{i=0}^{d-1} |w_{ij}|^2 \right) \\ &= \frac{\lambda_{d-1}}{\lambda_j} \frac{R_j}{R_{d-1}}. \end{aligned}$$

Because λ_j decreases with j while R_j increases with j and both are positive, we see that

$$1 \geq (T^\dagger T + Y^\dagger Y)_{jj} > 0$$

for each j ; moreover,

$$1 = (T^\dagger T + Y^\dagger Y)_{d-1,d-1}.$$

Thus, the diagonal entries of $I - T^\dagger T - Y^\dagger Y$ are nonnegative with the final entry being 0. We assert that the off-diagonal entries of $I - T^\dagger T - Y^\dagger Y$ are all zeros, equivalently, that the off-diagonal entries of $T^\dagger T + Y^\dagger Y$ are all zeros. Let r and s be distinct elements of $\{0, 1, \dots, d - 1\}$. Because the matrix M is unitary, the inner product of each pair of distinct rows of M is 0. Thus, the inner product of the ir and is rows of the matrix M vanishes, as therefore does the sum over i of these inner products (where we are assuming of course that $r \neq s$). Accordingly, recalling our explanation earlier in this section of how the matrix M is constructed, we obtain

$$\begin{aligned} 0 &= \sum_{i=0}^{d-1} \left[\sqrt{\lambda_r \lambda_s} \sum_{n=0}^{d^2-3} (u_{ir}^{(n)})^* u_{is}^{(n)} \right] + \sum_{i=0}^{d-1} v_{ir}^* v_{is} + \sum_{i=0}^{d-1} w_{ir}^* w_{is} \\ &= \sqrt{\lambda_r \lambda_s} \sum_{n=0}^{d^2-3} \sum_{i=0}^{d-1} (u_{ir}^{(n)})^* u_{is}^{(n)} + \sum_{i=0}^{d-1} v_{ir}^* v_{is} + \sum_{i=0}^{d-1} w_{ir}^* w_{is} \\ &= \sum_{i=0}^{d-1} v_{ir}^* v_{is} + \sum_{i=0}^{d-1} w_{ir}^* w_{is}, \end{aligned}$$

where the final equality holds because for each $n \in \{0, 1, 2, \dots, d^2 - 3\}$, the matrix $U^{(n)}$ is unitary (in particular, its columns are orthogonal). We have $0 = \sum_{i=0}^{d-1} v_{ir}^* v_{is} + \sum_{i=0}^{d-1} w_{ir}^* w_{is}$ so that

$$\begin{aligned} 0 &= \frac{\lambda_{d-1}}{R_{d-1} \sqrt{\lambda_r \lambda_s}} \left(\sum_{i=0}^{d-1} v_{ir}^* v_{is} + \sum_{i=0}^{d-1} w_{ir}^* w_{is} \right) \\ &= (T^\dagger T + Y^\dagger Y)_{rs}, \end{aligned}$$

as desired.

We have shown $I - T^\dagger T - Y^\dagger Y$ is positive: it is a diagonal matrix with nonnegative entries along the diagonal.

We view T and Y as operators on H_A . It is easy to see that $(T \otimes I_B)|\Psi\rangle$ and $(Y \otimes I_B)|\Psi\rangle$ are each orthogonal to the messages $(U^{(n)} \otimes I_B)|\Psi\rangle$, $n = 0, 1, \dots, d^2 - 3$. The coordinate vector of $(T \otimes I_B)|\Psi\rangle$ relative to \mathcal{B} is just $\frac{\sqrt{\lambda_{d-1}}}{\sqrt{R_{d-1}}}\mathbf{v}$, and that for $(Y \otimes I_B)|\Psi\rangle$ is $\frac{\sqrt{\lambda_{d-1}}}{\sqrt{R_{d-1}}}\mathbf{w}$. Because \mathbf{v} and \mathbf{w} are each orthogonal to the first $d^2 - 2$ columns of M , so are these scalar multiples of \mathbf{v} and \mathbf{w} . Note that neither of the vectors $\frac{\sqrt{\lambda_{d-1}}}{\sqrt{R_{d-1}}}\mathbf{v}$ or $\frac{\sqrt{\lambda_{d-1}}}{\sqrt{R_{d-1}}}\mathbf{w}$ is the zero vector (so that neither T nor Y is the zero matrix; in fact, it is easy to see T and Y are linearly independent).

Define C to be the square root of $I - T^\dagger T - Y^\dagger Y$. Note C is a diagonal matrix whose diagonal entries are square roots of the diagonal entries of $I - T^\dagger T - Y^\dagger Y$. We have

$$T^\dagger T + Y^\dagger Y + C^\dagger C = I. \quad (38)$$

We have already noted that $(C^\dagger C)_{d-1, d-1} = 0$. Now, for $j = 0, 1, \dots, d - 2$, we have

$$(C^\dagger C)_{jj} = 1 - \frac{\lambda_{d-1}}{\lambda_j} \frac{R_j}{R_{d-1}} = \frac{d(\lambda_j - \lambda_{d-1})}{\lambda_j [d - (d^2 - 2)\lambda_{d-1}]} \quad (39)$$

$$\leq \frac{d^2(\lambda_j - \lambda_{d-1})}{2\lambda_j}, \quad (40)$$

where to obtain the second equality we have used Eq. (36) and to obtain the final inequality we have used $\lambda_{d-1} \leq 1/d$. Because $\lambda_0 \geq \lambda_j \geq \lambda_{d-1} \geq 1 - (d - 1)\lambda_0$, both λ_j and λ_{d-1} approach $1/d$ as λ_0 approaches $1/d$. Thus, the overestimate (40) for $(C^\dagger C)_{jj}$ shows that all diagonal entries of $C^\dagger C$ approach 0 as λ_0 approaches $1/d$. For $j = 0, 1, \dots, d - 1$, let $\gamma_j = (C^\dagger C)_{jj}$, so that C is a $d \times d$ diagonal matrix with diagonal entries $\sqrt{\gamma_0}, \dots, \sqrt{\gamma_{d-1}}$, and, as we just discussed, for each j , $\gamma_j \rightarrow 0$ as $\lambda_0 \rightarrow 1/d$.

Because T , Y , and C satisfy the Kraus-operator condition (38), a process, described on page 365 of Ref. [5], for example, establishes that Alice can pair her qudit with an ancillary qutrit a and perform a unitary operator \tilde{U} on the pair to cast the ABa triple, in initial state $|\Psi\rangle|0\rangle_a$, into the state (10) with $N = 3$, $K^{(0)} = T$, $K^{(1)} = Y$, and $K^{(2)} = C$:

$$(T \otimes I_B)|\Psi\rangle \otimes |0\rangle_a + (Y \otimes I_B)|\Psi\rangle \otimes |1\rangle_a + (C \otimes I_B)|\Psi\rangle \otimes |2\rangle_a. \quad (41)$$

The corresponding reduced-density operator description of the resulting state of the AB system is $\mathcal{E}(|\Psi\rangle\langle\Psi|)$, where

$$\mathcal{E}(\rho) = T\rho T^\dagger + Y\rho Y^\dagger + C\rho C^\dagger. \quad (42)$$

Because we are assuming that $|\Psi\rangle$ is not maximally entangled, C cannot be the zero matrix. If it were, this would contradict the work of Sec. III, because Alice could then use \mathcal{E} of Eq. (42) to encode a $(d^2 - 1)$ -th message for Bob perfectly distinguishable from the unitary messages represented by $(U^{(n)} \otimes I_B)|\Psi\rangle$, $n = 0, 1, \dots, d^2 - 3$ (because $(T \otimes I_B)|\Psi\rangle$ and $(Y \otimes I_B)|\Psi\rangle$ are orthogonal to the ‘‘unitary messages’’).

We now show that under appropriate conditions, Alice’s measurement of the ancilla a after applying \tilde{U} to the Aa pair, can, with high probability, create a $(d^2 - 1)$ -th message for Bob perfectly distinguishable from the initial $d^2 - 2$ messages. Alice and Bob can agree (say, before they part company) that Bob will decode messages from Alice via the observable

$$\sum_{j=0}^{d^2-2} j P_{S_j} \quad (43)$$

where, for $j = 0, 1, \dots, d^2 - 3$, S_j is the one-dimensional subspace of H spanned by $(U^{(j)} \otimes I_B)|\Psi\rangle$, and, for $j = d^2 - 2$, S_j is the two-dimensional subspace on H spanned by $(T \otimes I_B)|\Psi\rangle$ and $(Y \otimes I_B)|\Psi\rangle$ (and P stands for projection). Because the subspaces S_j are pairwise orthogonal, Bob will receive perfectly distinguishable messages as long as Alice either encodes via some selected one of the original-protocol unitaries or else encodes via a quantum operation (in this case non-trace-preserving) that yields a state of the AB system described by a density operator whose support is S_{d^2-2} . Alice has no trouble producing a unitary message by applying a unitary operation to her qudit alone. She can thereby produce $d^2 - 2$ perfectly distinguishable messages. To (attempt to) produce the final $(d^2 - 1)$ -th message, Alice applies the unitary operation \tilde{U} to the qudit-qutrit pair Aa , as described in the preceding paragraph, casting the ABa system into the state (41). Then she performs the projective measurement on ABa corresponding to $P_{|2\rangle_a} \equiv (I_A \otimes I_B \otimes |2\rangle_a\langle 2|_a)$; she will measure 1 (i.e., the state $|2\rangle_a$) with probability

$$p_1 = \|(C \otimes I_B)|\Psi\rangle\|^2 = \left\| \sum_{j=0}^{d-1} \sqrt{\lambda_j} \sqrt{\gamma_j} |jj\rangle \right\|^2 = \sum_{j=0}^{d-1} \lambda_j \gamma_j, \quad (44)$$

where the second equality follows from the fact that C is a diagonal matrix (with diagonal entries $\sqrt{\gamma_j}$ and where we use $\|\cdot\|$ to denote vector length. Upon substituting our overestimate (40) for γ_j into Eq. (44) and using $\sum_{j=0}^{d-1} \lambda_j = 1$, we obtain $p_1 \leq (d^2/2) - (d^3/2)\lambda_{d-1}$, which, by applying $\lambda_{d-1} \geq 1 - (d - 1)\lambda_0$, yields

$$p_1 \leq \frac{d^3(d - 1)}{2} \left(\lambda_0 - \frac{1}{d} \right). \quad (45)$$

Thus, for λ_0 sufficiently close to $1/d$, the probability p_1 that Alice will measure 1 via the projective measurement $P_{|2\rangle_a}$ approaches 0. Thus, with probability $1 - p_1$ (approaching 1 as $\lambda_0 \rightarrow 1/d$), the measurement $P_{|2\rangle_a}$ will yield 0, casting the ABa system into the state

$$\frac{1}{\sqrt{1 - p_1}} [(T \otimes I_B)|\Psi\rangle \otimes |0\rangle_a + (Y \otimes I_B)|\Psi\rangle \otimes |1\rangle_a]. \quad (46)$$

At this point, Alice can send her qudit to Bob (through a noiseless quantum channel) and using his observable modeled by (43), Bob will receive message $d^2 - 1$ with certainty.

Observe that the reduced-density operator corresponding to relation (46) for the state of the AB system is

$$\frac{1}{1 - p_1} [(T \otimes I_B)|\Psi\rangle\langle\Psi|(T \otimes I_B)^\dagger + (Y \otimes I_B)|\Psi\rangle\langle\Psi|(Y \otimes I_B)^\dagger],$$

which is associated with the non-trace-preserving (because $C \neq 0$) quantum operator

$$\rho \mapsto (T \otimes I_B)\rho(T \otimes I_B)^\dagger + (Y \otimes I_B)\rho(Y \otimes I_B)^\dagger. \quad (47)$$

Thus, with probability $1 - p_1$, Alice can use the non-trace-preserving quantum operator defined by relation (47) to encode a $(d^2 - 1)$ -th message for Bob, and he will never mistake this message for any of the original-protocol unitary messages she may encode. A nice feature of this dense-coding scheme, which allows Alice to send $d^2 - 1$ perfectly distinguishable messages to Bob with high probability (and the first $d^2 - 2$ of those messages with certainty), is that if Alice does measure the ancilla to be in the undesirable state $|2\rangle_a$, then she can choose not to send her qudit to Bob because he would not be assured of receiving her intended message. If Alice does wish to send the $(d^2 - 1)$ -th message to Bob, we have shown that the probability of failure p_1 of the protocol has the upper bound (45), which gives an indication of how p_1 decreases to 0 as λ_0 approaches $1/d$. We note that the overestimate (45) of p_1 does not typically provide a sharp bound on the probability of failure. Suppose, for example, all Schmidt coefficients of $|\Psi\rangle$, except the largest, are equal, implying $\lambda_j = \lambda_{d-1} = (1 - \lambda_0)/(d - 1)$ for $j = 1, 2, \dots, d - 2$. Then, recalling $(C^\dagger C)_{jj} = \gamma_j$, we see that Eq. (39) shows that $\gamma_j = 0$ for $j = 1, 2, \dots, d - 1$; thus Eq. (39) combined with Eq. (44) yields

$$p_1 = \frac{d^2(\lambda_0 - 1/d)}{d^2(\lambda_0 - 1/d) + 2(1 - \lambda_0)} \leq \frac{d^3}{2(d - 1)} \left(\lambda_0 - \frac{1}{d} \right), \quad (48)$$

where the inequality holds because $d^2(\lambda_0 - 1/d) + 2(1 - \lambda_0)$ increases with λ_0 , taking the value $2(d - 1)/d$ when $\lambda_0 = 1/d$. Because λ_0 cannot exceed $d/(d^2 - 2)$ (as explained at the opening of this section), the bound on the probability of failure given by Eq. (48) cannot exceed $d^2/[(d - 1)(d^2 - 2)]$, which decreases asymptotically with increasing d as $1/d$. If we assume that $\lambda_0 = d/(d^2 - 1)$ [the value of the bound (3) for $L = d^2 - 1$], so that $\lambda_0 - 1/d$ is roughly 50% of its maximum allowed value, then the bound (48) yields small error probabilities even for small d (e.g., 28% for $d = 3$, dropping to 8.5% for $d = 7$).

We conclude with a concrete example presenting a situation in which Alice has more than a 97% chance of communicating three perfectly distinguishable messages to Bob using a system of two less than maximally entangled qubits. Suppose Alice and Bob share a two-qubit system in state

$$|\Psi\rangle = \frac{9}{4\sqrt{10}}|00\rangle + \frac{\sqrt{79}}{4\sqrt{10}}|11\rangle, \quad (49)$$

for which $\lambda_0 = 81/160$ and $\lambda_1 = 79/160$. For the preceding state, Alice may choose unitary encoding operations on her qubit A corresponding to the identity $I_A = |0\rangle\langle 0| + |1\rangle\langle 1|$ and the shift operator $X = |1\rangle\langle 0| + |0\rangle\langle 1|$. Here's a matrix M for this situation, whose first two columns, respectively, are the coordinates of $(I_A \otimes I_B)|\Psi\rangle$ and $(X \otimes I_B)|\Psi\rangle$ relative to $\mathcal{B} = (|00\rangle, |10\rangle, |01\rangle, |11\rangle)$ and whose remaining columns extend the first two to an orthonormal basis of \mathbb{C}^4 :

$$M = \begin{bmatrix} \frac{9}{40}\sqrt{10} & 0 & \frac{1}{40}\sqrt{395} & \frac{1}{40}\sqrt{395} \\ 0 & \frac{9}{40}\sqrt{10} & \frac{1}{40}\sqrt{395} & -\frac{1}{40}\sqrt{395} \\ 0 & \frac{1}{40}\sqrt{79}\sqrt{10} & -\frac{9}{40}\sqrt{5} & \frac{9}{40}\sqrt{5} \\ \frac{1}{40}\sqrt{79}\sqrt{10} & 0 & -\frac{9}{40}\sqrt{5} & -\frac{9}{40}\sqrt{5} \end{bmatrix}.$$

The matrices T , Y , and C are

$$T = \begin{bmatrix} \frac{79}{162} & -1/2 \\ \frac{79}{162} & -1/2 \end{bmatrix}, \quad Y = \begin{bmatrix} \frac{79}{162} & 1/2 \\ -\frac{79}{162} & -1/2 \end{bmatrix},$$

$$C = \begin{bmatrix} \sqrt{\frac{320}{6561}} & 0 \\ 0 & 0 \end{bmatrix}.$$

Note γ_0 , the upper left entry of $C^\dagger C$, is $320/6561$ whereas γ_1 , the lower right entry, is 0 (no surprise). The probability that Alice will measure the ancillary qubit a to be in state $|2\rangle_a$ after applying the unitary \tilde{U} to the Aa system (as described previously) is $\lambda_0\gamma_0 = 81/160 \cdot 320/6561 = 2/81$. Thus, the probability that this measurement will cast the system into the desirable state (46) is $79/81 \approx 97.5\%$. Thus there is greater than a 97% chance that Alice and Bob can use the less-than-maximally-entangled state (49) as a resource for the communication of three perfectly distinguishable messages. Specifically, Alice can communicate message 0, generated by applying I_A to A , with certainty; message 1, generated by applying X to A , with certainty; and message 2, generated by applying \tilde{U} to Aa and then measuring a , with probability $79/81$. If in attempting to encode message 2, Alice applies \tilde{U} to Aa and then observes a to be in the undesirable state $|2\rangle_a$, then the state of the system she shares with Bob is $(C \otimes I_B)|\Psi\rangle/\|C \otimes I_B|\Psi\rangle\| = |00\rangle$, and she should not send her qubit to Bob because he would misinterpret her intended message (message 2) as message 0.

If Alice and Bob are willing to tolerate a small chance of Bob's misinterpreting a message, they will be better served if Alice encodes message 2 by simply applying \tilde{U} to Aa and then sending her qubit to Bob *without measuring the ancilla* a . In this example, the projectors that Bob would use for decoding the message that Alice sends are

$$P_{S_0} = |\Psi\rangle\langle\Psi|,$$

$$P_{S_1} = (X \otimes I_B)|\Psi\rangle\langle\Psi|(X \otimes I_B)^\dagger,$$

and

$$P_{S_2} = \frac{1}{p_T}(T \otimes I_B)|\Psi\rangle\langle\Psi|(T \otimes I_B)^\dagger + \frac{1}{p_Y}(Y \otimes I_B)|\Psi\rangle\langle\Psi|(Y \otimes I_B)^\dagger,$$

where $p_T = \|(T \otimes I_B)|\Psi\rangle\|^2 = 79/162$ and $p_Y = \|(Y \otimes I_B)|\Psi\rangle\|^2 = 79/162$. Suppose that Alice performs only the

unitary \tilde{U} on the the Aa qubit-qutrit system *and does not measure the ancilla*. If Alice then sends her qubit to Bob (representing transmission of message 2), then Bob would measure message 2 with probability

$$\begin{aligned} & \text{tr}[P_{S_2}(T \otimes I_B)|\Psi\rangle\langle\Psi|(T \otimes I_B)^\dagger + P_{S_2}(Y \otimes I_B)|\Psi\rangle\langle\Psi| \\ & \quad \times (Y \otimes I_B)^\dagger + P_{S_2}(C \otimes I_B)|\Psi\rangle\langle\Psi|(C \otimes I_B)^\dagger] \\ & = \text{tr}[(T \otimes I_B)|\Psi\rangle\langle\Psi|(T \otimes I_B)^\dagger + (Y \otimes I_B)|\Psi\rangle\langle\Psi| \\ & \quad \times (Y \otimes I_B)^\dagger + P_{S_2}(C \otimes I_B)|\Psi\rangle\langle\Psi|(C \otimes I_B)^\dagger] \\ & = 79/80 = 98.75\%. \end{aligned}$$

Bob will always decode messages 0 and 1 correctly.

ACKNOWLEDGMENTS

The authors are indebted to Scott Cohen, Vlad Gheorghiu, Robert Griffiths, and Thomas Williams for useful discussions.

APPENDIX A: PROOF OF LEMMA II.1

Lemma II.1. Suppose that all the Schmidt coefficients of $|\Psi\rangle$ in the representation (1) are nonzero and that $K^{(0)}, K^{(1)}, \dots, K^{(m-1)}$ are linearly independent in $\mathcal{L}(H_A)$; then $(K^{(0)} \otimes I)|\Psi\rangle, (K^{(1)} \otimes I)|\Psi\rangle, \dots, (K^{(m-1)} \otimes I)|\Psi\rangle$ are linearly independent vectors in H .

Proof. Suppose that $\alpha_0, \alpha_1, \dots, \alpha_{m-1}$ are scalars such that

$$\sum_{p=0}^{m-1} \alpha_p (K^{(p)} \otimes I)|\Psi\rangle = 0.$$

Then, letting $k_{ij}^{(p)} = \langle i|K^{(p)}|j\rangle$, we obtain

$$\sum_{i,j=0}^{d-1} \left(\sqrt{\lambda_j} \sum_{p=0}^{m-1} \alpha_p k_{ij}^{(p)} \right) |ij\rangle = 0,$$

which implies, upon taking the inner product of both sides of the preceding equation with $|ij\rangle$, that for each i and j

$$\sum_{p=0}^{m-1} \alpha_p k_{ij}^{(p)} = 0 \quad (\text{A1})$$

(because $\sqrt{\lambda_j}$ is nonzero for each j). Note that Eq. (A1) says

$$\sum_{p=0}^{m-1} \alpha_p K^{(p)} = 0,$$

which implies $\alpha_p = 0$ for all p because $K^{(0)}, K^{(1)}, \dots, K^{(m-1)}$ are linearly independent. Thus, $(K^{(0)} \otimes I)|\Psi\rangle, (K^{(1)} \otimes I)|\Psi\rangle$, and $(K^{(m-1)} \otimes I)|\Psi\rangle$ are linearly independent.

APPENDIX B: KRAUS OPERATOR REPRESENTATIONS

Suppose that Alice pairs her qudit A with an ancillary N -level particle a . The Hilbert space for the Aa pair is $H_A \otimes H_a$ with orthonormal basis $\mathcal{B}_{Aa} \equiv \{|i\rangle_A |r\rangle_a : i = 0, 1, 2, \dots, d-1; r = 0, 1, 2, \dots, N-1\}$. Assume, that Alice's particle A is entangled with Bob's particle B and their two-qudit system is in state $|\Psi\rangle$ given by Eq. (1). Assume that a is in state $|0\rangle_a$. Suppose the physical equivalent of a unitary operator U on $H_A \otimes H_a$ is applied to the pair Aa .

We can express U in terms of its action on the the basis elements in \mathcal{B}_{Aa} as follows:

$$\begin{aligned} U & = \sum_{i,j=0}^{d-1} \sum_{r,s=0}^{N-1} u_{ir,js} |i\rangle_A |r\rangle_a \langle j|_A \langle s|_a \\ & = \sum_{i,j=0}^{d-1} \sum_{r,s=0}^{N-1} u_{ir,js} (|i\rangle_A \langle j|_A \otimes |r\rangle_a \langle s|_a), \end{aligned} \quad (\text{B1})$$

where the scalars $u_{ir,js}$ must satisfy the following condition because of the unitarity of U : for $0 \leq j, j' \leq d-1$ and $0 \leq s, s' \leq N-1$,

$$\sum_{i=0}^{d-1} \sum_{r=0}^{N-1} u_{ir,js}^* u_{ir,j's'} = \delta_{j,j'} \delta_{s,s'}. \quad (\text{B2})$$

Letting I_B be the identity on Bob's Hilbert space H_B , we see that the effect of Alice's unitary operator U on the state $|\Psi\rangle|0\rangle_a$ may be described as follows:

$$\begin{aligned} & \left[\sum_{i,j=0}^{d-1} \sum_{r,s=0}^{N-1} u_{ir,js} (|i\rangle_A \langle j|_A \otimes I_B) \otimes (|r\rangle_a \langle s|_a) \right] |\Psi\rangle \otimes |0\rangle_a \\ & = \sum_{i,j=0}^{d-1} \sum_{r,s=0}^{N-1} u_{ir,js} (|i\rangle_A \langle j|_A \otimes I_B) |\Psi\rangle \otimes (|r\rangle_a \langle s|_a) |0\rangle_a \\ & = \sum_{r=0}^{N-1} \sum_{i,j=0}^{d-1} u_{ir,j0} (|i\rangle_A \langle j|_A \otimes I_B) |\Psi\rangle \otimes |r\rangle_a \\ & = \sum_{r=0}^{N-1} \left[\left(\sum_{i,j=0}^{d-1} u_{ir,j0} |i\rangle_A \langle j|_A \right) \otimes I_B \right] |\Psi\rangle \otimes |r\rangle_a \\ & = \sum_{r=0}^{N-1} (K^{(r)} \otimes I_B) |\Psi\rangle \otimes |r\rangle_a, \end{aligned}$$

where $K^{(r)} = \sum_{i,j=0}^{d-1} u_{ir,j0} |i\rangle_A \langle j|_A$ are Kraus operators that satisfy $\sum_{r=0}^{N-1} (K^{(r)})^\dagger K^{(r)} = I_A$:

$$\begin{aligned} \sum_{r=0}^{N-1} (K^{(r)})^\dagger K^{(r)} & = \sum_{r=0}^{N-1} \left(\sum_{i,j=0}^{d-1} u_{ir,j0}^* |j\rangle_A \langle i|_A \right) \\ & \quad \times \left(\sum_{i',j'=0}^{d-1} u_{i'r,j'0} |i'\rangle_A \langle j'|_A \right) \\ & = \sum_{r=0}^{N-1} \sum_{i,j,j'=0}^{d-1} u_{ir,j0}^* u_{i'r,j'0} |j\rangle_A \langle j'|_A \\ & = \sum_{j,j'=0}^{d-1} \sum_{i=0}^{d-1} \sum_{r=0}^{N-1} u_{ir,j0}^* u_{i'r,j'0} |j\rangle_A \langle j'|_A \\ & = \sum_{j,j'=0}^{d-1} \delta_{j,j'} |j\rangle_A \langle j'|_A \\ & = \sum_{j=0}^{d-1} |j\rangle_A \langle j|_A \\ & = I_A, \end{aligned}$$

where we have used Eq. (B2) to obtain the fourth equality. Thus, the process of pairing A with the N -level particle a and applying U to the pair always casts the three-particle system ABa into a state of the form (10), where the Kraus operators $K^{(0)}, \dots, K^{(N-1)}$ satisfy Eq. (7). The quantum operator \mathcal{E} associated with this process will have the form (6), and $N \geq m$, where m is the Kraus rank of \mathcal{E} .

APPENDIX C: ORTHOGONALITY OF KRAUS STATES

Let

$$\mathcal{E}(\rho) = (K^{(0)} \otimes I_B)\rho(K^{(0)} \otimes I_B)^\dagger + (K^{(1)} \otimes I_B)\rho(K^{(1)} \otimes I_B)^\dagger, \quad \rho \in \mathcal{L}(H),$$

where $K^{(0)}$ and $K^{(1)}$ are linearly independent Kraus operators on $\mathcal{L}(H_A)$ satisfying Eq. (13). We prove that there are Kraus operators $R^{(0)}$ and $R^{(1)}$ on $\mathcal{L}(H_A)$ such that (i) the quantum operator \mathcal{E} is also given by $\mathcal{E}(\rho) = (R^{(0)} \otimes I_B)\rho(R^{(0)} \otimes I_B)^\dagger + (R^{(1)} \otimes I_B)\rho(R^{(1)} \otimes I_B)^\dagger$, and (ii) the vectors

$$(R^{(0)} \otimes I_B)|\Psi\rangle \quad \text{and} \quad (R^{(1)} \otimes I_B)|\Psi\rangle \quad (\text{C1})$$

are orthogonal vectors in H [where $|\Psi\rangle$ is given by Eq. (1) and all of its Schmidt coefficients are nonzero].

Set $|\phi_0\rangle = (K^{(0)} \otimes I_B)|\Psi\rangle$ and $|\phi_1\rangle = (K^{(1)} \otimes I_B)|\Psi\rangle$. Note neither $|\phi_0\rangle$ nor $|\phi_1\rangle$ is the zero vector because they are linearly independent (by Lemma II.1). Assume that $|\phi_0\rangle$ and $|\phi_1\rangle$ are not orthogonal states. (If they are, then we are done—set $R^{(0)} = K^{(0)}$ and $R^{(1)} = K^{(1)}$.) For $\alpha = 0, 1$, define

$$R^{(\alpha)} = \sum_{\beta=0}^1 v_{\alpha\beta} K^{(\beta)}, \quad (\text{C2})$$

where the $v_{\alpha\beta}$'s are complex numbers that constitute a 2×2 matrix V . It is easy to check that if V is unitary, then for any operator $\rho \in \mathcal{L}(H)$,

$$\sum_{j=0}^1 (R^{(j)} \otimes I_B)\rho(R^{(j)} \otimes I_B)^\dagger = \sum_{j=0}^1 (K^{(j)} \otimes I_B)\rho(K^{(j)} \otimes I_B)^\dagger.$$

Thus, to complete our proof, we need show only that there is a unitary matrix V such that the operators $R^{(0)}$ and $R^{(1)}$ defined by Eq. (C2) are also such that the vectors (C1) are orthogonal in H .

Let μ, ν , and θ denote real numbers, and observe that

$$V = \begin{bmatrix} e^{i\mu} \cos \theta & -e^{i\nu} \sin \theta \\ e^{-i\nu} \sin \theta & e^{-i\mu} \cos \theta \end{bmatrix}$$

is unitary. Thus, using definition (C2), with the scalars $v_{\alpha\beta}$ determined by the preceding matrix, we have

$$\begin{aligned} (R^{(0)} \otimes I)|\Psi\rangle &= (e^{i\mu} \cos \theta)|\phi_0\rangle - (e^{i\nu} \sin \theta)|\phi_1\rangle \\ &= (e^{i\mu} \cos \theta)[|\phi_0\rangle - (e^{i(\nu-\mu)} \tan \theta)|\phi_1\rangle] \end{aligned} \quad (\text{C3})$$

$$\begin{aligned} (R^{(1)} \otimes I)|\Psi\rangle &= (e^{-i\nu} \sin \theta)|\phi_0\rangle + (e^{-i\mu} \cos \theta)|\phi_1\rangle \\ &= (e^{-i\mu} \cos \theta)[(e^{i(\mu-\nu)} \tan \theta)|\phi_0\rangle + |\phi_1\rangle]. \end{aligned} \quad (\text{C4})$$

If we can choose the scalars μ, ν , and θ such that $(R^{(0)} \otimes I)|\Psi\rangle$ and $(R^{(1)} \otimes I)|\Psi\rangle$ are orthogonal in H , our proof is complete.

Observe from Eqs. (C3) and (C4) that if we desire that $(R^{(0)} \otimes I)|\Psi\rangle$ and $(R^{(1)} \otimes I)|\Psi\rangle$ be orthogonal, then, because $|\phi_0\rangle$ and $|\phi_1\rangle$ are not orthogonal, we must have $\cos(\theta) \neq 0$, which justifies our factorization on the right of Eqs. (C3) and (C4). Using these factored forms from (C3) and (C4) and writing $\mu - \nu = \xi$, we compute the inner product of $(R^{(0)} \otimes I)|\Psi\rangle$ and $(R^{(1)} \otimes I)|\Psi\rangle$, obtaining the following necessary and sufficient condition for the inner product to be 0:

$$\begin{aligned} &-(e^{2i\xi} \tan^2 \theta)\langle\phi_1|\phi_0\rangle + (e^{i\xi} \tan \theta)(\langle\phi_0|\phi_0\rangle \\ &\quad - \langle\phi_1|\phi_1\rangle) + \langle\phi_0|\phi_1\rangle = 0. \end{aligned} \quad (\text{C5})$$

In the preceding equation, the scalar products $\langle\phi_\alpha|\phi_\beta\rangle$ ($\alpha, \beta = 0, 1$), like the $|\phi_\alpha\rangle$'s themselves, can be considered known. Thus, because $\langle\phi_1|\phi_0\rangle$ is nonzero, Eq. (C5) is a quadratic equation in the complex variable $z = e^{i\xi} \tan \theta$, so that to produce the desired orthogonal states we simply choose values of θ and ξ making $e^{i\xi} \tan \theta$ a root of this quadratic.

- [1] C. H. Bennett and S. J. Wiesner, Phys. Rev. Lett. **69**, 2881 (1992).
 [2] Z. Ji, Y. Feng, R. Duan, and M. Ying, Phys. Rev. A **73**, 034307 (2006).
 [3] S. Mozes, J. Oppenheim, and B. Reznik, Phys. Rev. A **71**, 012311 (2005).
 [4] S. Wu, S. M. Cohen, Y. Sun, and R. B. Griffiths, Phys. Rev. A **73**, 042311 (2006).

- [5] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
 [6] M. R. Beran and S. M. Cohen, Phys. Rev. A **78**, 062337 (2008).
 [7] Y. Feng, R. Duan, and M. Ying, Phys. Rev. A **70**, 012308 (2004).
 [8] E. Gerjuoy, H. T. Williams, and P. S. Bourdon, Phys. Rev. A **79**, 042315 (2009).