Optimal one-shot entanglement sharing

Vikesh Siddhu and John Smolin
IBM Quantum, IBM T.J. Watson Research Center, Yorktown Heights, New York 10598, USA

(Received 20 January 2023; revised 22 August 2023; accepted 23 August 2023; published 28 September 2023)

Sharing entanglement across quantum interconnects is fundamental for quantum information processing. We discuss a practical setting where this interconnect, modeled by a quantum channel, is used once with the aim of sharing high-fidelity entanglement. For any channel, we provide methods to easily find both this maximum fidelity and optimal inputs that achieve it. Unlike most metrics for sharing entanglement, this maximum fidelity can be shown to be multiplicative. This ensures a complete understanding in the sense that the maximum fidelity and optimal inputs found in our one-shot setting extend even when the channel is used multiple times, possibly with other channels. Optimal inputs need not be fully entangled. We find that the minimum entanglement in these optimal inputs can even vary discontinuously with channel noise. Generally, noise parameters are hard to identify and remain unknown for most channels. However, for all qubit channels with qubit environments, we provide a rigorous noise parametrization, which we explain in terms of no cloning. This noise parametrization and a channel representation that we call the standard Kraus decomposition have pleasing properties that make them useful more generally.

DOI: 10.1103/PhysRevA.108.032617

I. INTRODUCTION

Quantum computation and communication requires faithful transmission of quantum information between various separated parties. These parties may be closely separated quantum computing nodes or widely separated receivers and transmitters of quantum states. The former appear in models of a quantum intranet [1] while the latter appear in discussions of a quantum internet [2,3]. Noise in these and other such setups hinders their use. A dominant source of noise is the quantum interconnect carrying quantum information between parties. This interconnect is modeled mathematically by a quantum channel, a completely positive trace-preserving map. Quantum information sent and processed across this channel is equivalent to entanglement shared and processed using the channel [4]. Without investigating methods, metrics, protocols, and characteristics of sharing entanglement across quantum channels, our understanding and ability to control and scale quantum computation and communication remains partial.

The most well-studied setting for sharing entanglement allows asymptotically many channel uses [5,6]. Across all channels used together, local pre- and postprocessing of entanglement is allowed along with classical communication from channel input to output. Using these allowed operations, the largest number of fully entangled states, per channel use, shared with asymptotically vanishing error defines the quantum capacity of the channel. Studies of this metric reveal that while theoretically beautiful [7–11], a channel's quantum capacity is hard to compute and nontrivial to understand in general [12,13]. Both of these features come from superadditivity. Superadditivity of quantum capacity implies that the quantum capacity of several channels used jointly is not completely specified by the quantum capacity of each channel [14,15].

Asymptotic channel capacities provide rich conceptual and practical difficulties. For these reasons, it is desirable to study entanglement transmission with as little encoding and decoding as possible. The simplest setting here is a single use of a channel (which can itself be joint uses of many channels) with no postprocessing. This setting need not allow sharing of noiseless entanglement. Thus, one may define a metric for sharing entanglement with some acceptable level of noise. One such metric, called the one-shot quantum capacity, is roughly the largest fully entangled state that can be shared across a channel with at most a fixed but arbitrary error [16]. This one-shot capacity, its connection to asymptotic capacities, and a method for understanding and achieving these have been recently explored [17–27]. However, we do not fully understand notions of additivity for this capacity; ways of computing and explicit protocols for achieving the one-shot capacity are not completely known.

A key metric in the one-shot setting is the highest fidelity between the state shared across the channel and a maximally entangled state [5,28]. This fidelity characterizes the optimal performance of various teleportation-based tasks [29]. The optimal fidelity between a pure entangled state shared across the channel and a maximally entangled state is known [30,31]. Surprisingly, the optimal pure state input need not be maximally entangled, which is consistent with fidelity not being an entanglement monotone.

The one-shot setting is augmented by postprocessing using one round of local operations and two-way classical communication (2-LOCC) [31–34]. However, in this setting it is unknown if the optimal fidelity is multiplicative (analog of additivity in this setting). There is no known method for computing or explicit protocol for achieving this optimal fidelity in general. The only exception is qubit channels, where optimal protocols use pure state inputs and do not require 2-LOCC

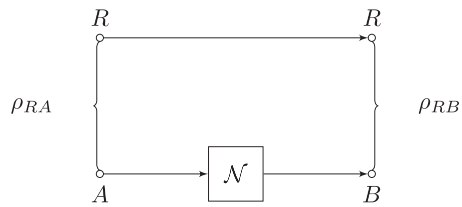

FIG. 1. Diagram representing one-shot entanglement passing.

[32,34]. Surprisingly, the behavior of such optimal protocols for the simplest of qubit channels is not fully known.

One way to understand a metric for sharing entanglement across a specific channel is to study variation in the metric with the amount of noise in the channel. Surprisingly, even for the simplest qubit channels, noise parameters are only partially understood.

Results

In this work, we introduce and solve the problem for sharing entanglement in a one-shot setting where an arbitrary mixed state ρ_{RA} may be prepared across a reference system R and channel input A. This input is sent via a fixed channel $\mathcal{N} : A \mapsto B$ (see Fig. 1) to achieve the maximum fidelity $\mathcal{O}(\mathcal{N})$ between the channel output ρ_{RB} and a maximally entangled state across R and B. We reformulate $\mathcal{O}(\mathcal{N})$ via a semidefinite program [35]. Our first main result is to express $\mathcal{O}(\mathcal{N})$ in two useful ways (see Sec. IV A with Theorem IV A): first, using what we define (in Sec. III A) as a channel's standard Kraus decomposition, and second, in terms of the operator norm of a channel's Choi-Jamiołkowski operator. Next, we show that the maximum fidelity $\mathcal{O}$ is multiplicative (see Theorem IV B). Not only can $\mathcal{O}(\mathcal{N})$ be achieved using pure states, but, in certain cases, it can also be achieved using a variety of mixed states. We give a recipe to construct these pure and mixed states. For all extremal [see the definition below Eq. (40)] qubit channels, we compute optimal inputs and the minimum amount of entanglement $\mathcal{E}$ required to create these inputs. Somewhat surprisingly, the minimum entanglement $\mathcal{E}$ is found to be discontinuous in these noise parameters. Typically, $\mathcal{E}$ is less than its maximal value of 1, but $\mathcal{O}$ is high enough for the channel to be useful for teleportation, even if the channel has no quantum capacity (see Sec. V A). For very noisy qubit Pauli channels, we find separable inputs that achieve the same fidelity as maximally entangled ones found previously (see Sec. V B). We also find that optimal inputs for qutrit channels have a much richer structure than qubit channels (see Sec. V C). Noise parameters for general channels remain unknown. We find rigorous noise parameters for all extremal qubit channels (see Sec. III C), a result that may be of independent interest.

Unlike other metrics in settings for entanglement sharing, $\mathcal{O}$ is multiplicative. Thus, even when a channel $\mathcal{N}$ is used multiple times, possibly with other channels, its maximum fidelity $\mathcal{O}(\mathcal{N})$ fully characterizes its ability for sharing high fidelity entanglement without postprocessing. Our results also give rigorous lower bounds on entanglement fidelities that can be achieved when allowing for multiple rounds of 2-LOCC. These bounds are tight for one round of 2-LOCC using qubit channels. Characterization of the noise parameters for all extremal qubit channels presented here paves the way for a stronger understanding of quantum channels and quantum protocols across channels.

II. PRELIMINARIES

Let $\mathbf{x}$ denote a vector in n-dimensional real space, $\mathbb{R}^n$, let $\mathbf{x}_i$ denote the $(i+1)$th coordinate of $\mathbf{x}$, and let $|\mathbf{x}_i|$ denote its absolute value. Coordinates of $\mathbf{x}$ rearranged in decreasing order give $\mathbf{x}^\downarrow$, a vector satisfying $\mathbf{x}^\downarrow_0 \geqslant \mathbf{x}^\downarrow_1 \geqslant \cdots \geqslant \mathbf{x}^\downarrow_{n-1}$. The Euclidean norm of $\mathbf{x}$ is $|\mathbf{x}| := \sqrt{\sum_i \mathbf{x}_i^2}$. Let $|\psi\rangle$ denote a ket in a Hilbert space $\mathcal{H}$ of finite dimension d, and let $||\psi\rangle| := \sqrt{\langle\psi|\psi\rangle}$ denote its norm. A pure quantum state is represented by a ket with unit norm. Let $\mathcal{L}(\mathcal{H})$ denote the space of linear operators on $\mathcal{H}$. For any two quantum states $|\psi\rangle$ and $|\phi\rangle$, the dyad $|\psi\rangle\langle\phi| \in \mathcal{L}(\mathcal{H})$ and the projector onto $|\psi\rangle$, $|\psi\rangle\langle\psi| \in \mathcal{L}(\mathcal{H})$. The Frobenius inner product between two operators N and O in $\mathcal{L}(\mathcal{H})$ is

$$\langle N, O\rangle := \mathrm{Tr}(N^\dagger O), \tag{1}$$

where $N^\dagger$ represents the adjoint (conjugate transpose) of N. A Hermitian operator $H \in \mathcal{L}(\mathcal{H})$, satisfying $H = H^\dagger$, represents an observable. This operator has an eigendecomposition,

$$H = \sum_i \mathbf{x}_i |\psi_i\rangle\langle\psi_i|, \tag{2}$$

where $\mathbf{x}_i \in \mathbb{R}$ is an eigenvalue of H corresponding to eigenvector $|\psi_i\rangle$, and the collection of eigenvectors $\{|\psi_i\rangle\}$ form an orthonormal basis of $\mathcal{H}$, $\langle\psi_i|\psi_j\rangle = \delta_{ij}$, where δ_{ij} is the Kronecker delta function. Support of H is the subspace spanned by its eigenvectors with nonzero eigenvalues. In (2), if $\mathbf{x}_i \geqslant 0$ for all i, then we say H is positive-semidefinite (PSD), $H \succeq 0$. An optimization, over PSD matrices, of the form

$$\begin{aligned} &\text{maximize } \mathrm{Tr}(A_0 H) \\ &\text{subject to } \mathrm{Tr}(A_i H) = c_i, \quad \forall\, 1 \leqslant i \leqslant n, \\ &\text{and } H \succeq 0, \end{aligned} \tag{3}$$

where A_i are Hermitian, is called a semidefinite program (see Sec. 1.2.3 in [49] and citations to and within [35]). The square root of a PSD operator H, $\sqrt{H}$, is obtained by replacing $\mathbf{x}_i$ in (2) with $\sqrt{\mathbf{x}_i}$. For any operator $O \in \mathcal{L}(\mathcal{H})$,

$$\begin{aligned} &||O|| := \max_{||\psi\rangle| \leqslant 1} |O|\psi\rangle|, \\ &||O||_1 := \mathrm{Tr}(\sqrt{OO^\dagger}), \quad \text{and} \quad ||O||_2 := \sqrt{\mathrm{Tr}(OO^\dagger)}, \end{aligned} \tag{4}$$

denote the spectral norm, the trace norm, and the Frobenius norm, respectively. For H in (2),

$$||H|| = |\mathbf{x}^\downarrow_0|, \quad ||H||_1 = \sum_i |\mathbf{x}_i|, \quad \text{and} \quad ||H||_2 = |\mathbf{x}|. \tag{5}$$

A density operator $\rho \in \mathcal{L}(\mathcal{H})$ is a positive-semidefinite operator with unit trace, $\mathrm{Tr}(\rho) = 1$, which represents a mixed quantum state. Its von-Neumann entropy,

$$S(\rho) = -\mathrm{Tr}(\rho \log_2 \rho), \tag{6}$$

where log is base 2. The fidelity between two density operators ρ and σ,

$$F(\rho,\sigma) := ||\sqrt{\rho}\sqrt{\sigma}||_1. \tag{7}$$

Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be two Hilbert spaces of dimensions d_A and d_B, respectively, and let $\mathcal{H}_{AB}$ denote the tensor product space $\mathcal{H}_A \otimes \mathcal{H}_B$. Given a pure state $|\psi\rangle_{AB} \in \mathcal{H}_{AB}$, density operators

$$\psi_A = \mathrm{Tr}_B(|\psi\rangle\langle\psi|) \quad \text{and} \quad \psi_B = \mathrm{Tr}_A(|\psi\rangle\langle\psi|) \tag{8}$$

denote the partial trace of $|\psi\rangle\langle\psi|$ over $\mathcal{H}_B$ and $\mathcal{H}_A$, respectively. The entanglement of formation of a pure state $|\psi\rangle_{AB}$,

$$E_f(|\psi\rangle_{AB}) = S(\psi_A), \tag{9}$$

and for a mixed state ρ_{AB},

$$E_f(\rho_{AB}) = \min \sum_i p_i E_f(|\psi_i\rangle_{AB}), \tag{10}$$

is the minimum average entanglement E_f over all pure state decompositions, $\rho_{AB} = \sum_i p_i|\psi_i\rangle\langle\psi_i|$, $p_i \geqslant 0$ and $\sum_i p_i = 1$.

Let $\mathcal{A} = \{|a_i\rangle\}$ and $\mathcal{B} = \{|b_j\rangle\}$ be orthonormal bases, of $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively, i.e.,

$$\langle a_i|a_j\rangle = \langle b_i|b_j\rangle = \delta_{ij}. \tag{11}$$

Using these bases $\mathcal{A}$ and $\mathcal{B}$ we can represent any linear operator $L : \mathcal{H}_A \mapsto \mathcal{H}_B$ as a matrix,

$$L = \sum_{ij}[L]_{ij}|b_i\rangle\langle a_j|, \tag{12}$$

with elements $[L]_{ij}$. We can define two basis-dependent linear maps,

$$L^* = \sum_{ij}[L]^*_{ij}|b_i\rangle\langle a_j| \quad \text{and} \quad L^T = \sum_{ij}[L]_{ij}|a_j\rangle\langle b_i|, \tag{13}$$

representing complex conjugate and transpose, respectively. In contrast to L^* and L^T, the adjoint $L^\dagger = (L^*)^T = (L^T)^*$ is basis-independent. If $\mathcal{H}_A$ and $\mathcal{H}_B$ have the same dimension d, then one can choose $\mathcal{A}$ and $\mathcal{B}$ to be the same, say the standard basis $\{|i\rangle\}$, and construct an identity map $I_{A\leftarrow B} : \mathcal{H}_B \mapsto \mathcal{H}_A$,

$$I_{A\leftarrow B}|i\rangle_B = |i\rangle_A. \tag{14}$$

This subscript notation $A \leftarrow B$ is dropped shortly after defining how the identity map above is used to map a ket $|\phi\rangle_B \in \mathcal{H}_B$, an operator $O_B \in \mathcal{L}(\mathcal{H}_B)$, and part of an operator $L_{AB} \in \mathcal{L}(\mathcal{H}_{AB})$ to

$$|\phi\rangle_A = I_{A\leftarrow B}|\psi\rangle_B, \quad O_A = I_{A\leftarrow B}O_B I_{B\leftarrow A}, \quad \text{and}$$
$$L_{AA} = (I_{A\leftarrow B}\otimes I_A)L_{BA}(I_{B\leftarrow A}\otimes I_A), \tag{15}$$

respectively; here I_A is the identity on the $\mathcal{H}_A$ space. Later, these mappings are done implicitly by simply replacing the subscripts in an obvious way.

Operator-ket duality

Operator-ket duality is the concept of fixing an orthonormal basis $\mathcal{A} = \{|a_i\rangle\}$ of $\mathcal{H}_A$ and using an unnormalized maximally entangled state on $\mathcal{H}_A \otimes \mathcal{H}_A$,

$$|\gamma\rangle_{AA} = \sum_i |a_i\rangle \otimes |a_i\rangle, \tag{16}$$

to associate with any linear operator $K : \mathcal{H}_A \mapsto \mathcal{H}_B$ a ket, $|\psi\rangle_{AB} = (I_A \otimes K)|\gamma\rangle$, obtained by acting K on one-half of $|\gamma\rangle$. Conversely, for fixed orthonormal basis $\mathcal{A}$, one associates with any ket $|\psi\rangle_{AB}$ a linear operator

$$K = \sum_i |\chi_i\rangle\langle a_i|, \quad \text{where} \quad |\chi_i\rangle_B = (\langle a_i|_A \otimes I_B)|\psi\rangle_{AB}. \tag{17}$$

In analogy to the discussion above, fixing an orthonormal basis $\mathcal{B} = \{|b_j\rangle\}$ of $\mathcal{H}_B$ one associates with the ket $|\psi\rangle_{AB}$ an operator $L : \mathcal{H}_B \mapsto \mathcal{H}_A$. This operator $L = K^T$, where the transpose operation is taken using basis $\mathcal{A}$ and $\mathcal{B}$ as described in (13).

In what follows, we use the notation $|K\rangle \in \mathcal{H}_{AB}$ for a ket associated with the operator $K : \mathcal{H}_A \mapsto \mathcal{H}_B$ through the operator-ket duality above where basis $\mathcal{A}$ is fixed. This ket and operator pair satisfy

$$|K\rangle_{AB} = (I \otimes K)|\gamma\rangle_{AA}. \tag{18}$$

For any two maps K and K' from $\mathcal{H}_A$ to $\mathcal{H}_B$ and associated kets $|K\rangle_{AB}$ and $|K'\rangle_{AB}$, respectively, one can show that

$$\langle K, K'\rangle = \langle K|K'\rangle. \tag{19}$$

Using the orthonormal basis $\mathcal{B}$ of $\mathcal{H}_B$, one can associate with $K^\dagger : \mathcal{H}_B \mapsto \mathcal{H}_A$ the ket $|K^\dagger\rangle_{BA}$. In this ket, swapping the spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ [see the discussion below (14)] gives $|K^\dagger\rangle_{AB}$, which then satisfies

$$|K^\dagger\rangle_{AB} = |K\rangle^*_{AB}, \tag{20}$$

where complex conjugation of any ket $|\chi\rangle_{AB} = \sum_{ij} c_{ij}|a_i\rangle \otimes |b_j\rangle$ is defined using basis $\mathcal{A}$ and $\mathcal{B}$ as $|\chi\rangle^*_{AB} = \sum_{ij} c^*_{ij}|a_i\rangle \otimes |b_j\rangle$.

III. QUANTUM CHANNELS

Let $\mathcal{H}_A$, $\mathcal{H}_B$, and $\mathcal{H}_C$ be three Hilbert spaces, and let $V : \mathcal{H}_A \mapsto \mathcal{H}_B \otimes \mathcal{H}_C$ be an isometry, i.e., $V^\dagger V = I_A$. This isometry defines a pair of quantum channels $\mathcal{N}$ and $\mathcal{N}^c$, i.e., a pair of completely positive trace preserving (CPTP) maps with superoperators

$$\mathcal{N}(O) = \mathrm{Tr}_C(VOV^\dagger) \quad \text{and} \quad \mathcal{N}^c(O) = \mathrm{Tr}_B(VOV^\dagger), \tag{21}$$

taking $O \in \mathcal{L}(\mathcal{H}_A)$ to be $\mathcal{L}(\mathcal{H}_B)$ and $\mathcal{L}(\mathcal{H}_C)$, respectively. The quantum channel $\mathcal{N}$ is called degradable and $\mathcal{N}^c$ is antidegradable if there exists a quantum channel $\mathcal{D}$ such that $\mathcal{D} \circ \mathcal{N} = \mathcal{N}^c$ [11].

Let $\mathcal{I}_A$ be the identity map from $\mathcal{L}(\mathcal{H}_A)$ to itself. Using an unnormalized maximally entangled state $|\gamma\rangle_{AA}$ (16) we define the Choi-Jamiołkowski [36,37] operator of the linear map $\mathcal{N}$ as

$$J^{\mathcal{N}}_{AB} = \mathcal{I}_A \otimes \mathcal{N}(|\gamma\rangle\langle\gamma|) = \sum_{ij} |a_i\rangle\langle a_j| \otimes \mathcal{N}(|a_i\rangle\langle a_j|). \tag{22}$$

This operator contains all information about $\mathcal{N}$. For instance,

$$\mathcal{N}(|a_i\rangle\langle a_j|) = (\langle a_i| \otimes I_B)J^{\mathcal{N}}_{AB}(|a_j\rangle \otimes I_B), \tag{23}$$

$\mathcal{N}$ is completely positive (CP) if and only if $J^{\mathcal{N}}_{AB}$ is positive-semidefinite, and

$$\mathrm{Tr}_B\big(J^{\mathcal{N}}_{AB}\big) = I_A \tag{24}$$

if and only if $\mathcal{N}$ is trace-preserving; $\mathrm{Tr}(\mathcal{N}(O)) = \mathrm{Tr}(O)$ for all O. Equivalently, a linear map $\mathcal{N} : \mathcal{L}(\mathcal{H}_A) \mapsto \mathcal{L}(\mathcal{H}_B)$ is CP if and only if it can be written in the form

$$\mathcal{N}(O) = \sum_i K_i O K_i^\dagger, \tag{25}$$

where $K_i : \mathcal{H}_A \mapsto \mathcal{H}_B$ is a linear operator, and the collection $\{K_i\}$ are called Kraus operators. The map in (25) is trace-preserving when these Kraus operators satisfy $\sum_i K_i^\dagger K_i = I_A$. When $\mathcal{N}$ is unital, i.e., $\mathcal{N}(I_A) = I_B$, the Kraus operators satisfy $\sum_i K_i K_i^\dagger = I_B$. If $\mathcal{H}_A$ and $\mathcal{H}_B$ have the same dimension, then they are isomorphic to one another and can be denoted by $\mathcal{H}$. If these Kraus operators on $\mathcal{H}$ are Hermitian operators (or normal operators), then the channel is automatically unital.

A. A standard Kraus decomposition

For a given channel $\mathcal{N} : \mathcal{L}(\mathcal{H}_A) \mapsto \mathcal{L}(\mathcal{H}_B)$, the set of Kraus operators is not unique. However, one can construct what can be called a *standard Kraus decomposition* with some pleasing properties.

Consider the eigendecomposition of the Choi-Jamiołkowski operator in (22),

$$J_{AB}^{\mathcal{N}} = \sum_i \mathbf{e}_i^\downarrow |L_i\rangle\langle L_i|, \tag{26}$$

where eigenvalues $\mathbf{e}_0^\downarrow \geqslant \mathbf{e}_1^\downarrow \geqslant \cdots \geqslant \mathbf{e}_{d_A d_B - 1}^\downarrow \geqslant 0$ and eigenvectors $\{|L_i\rangle\}$ form an orthonormal basis of $\mathcal{H}_{AB}$,

$$\langle L_i | L_j \rangle = \delta_{ij}. \tag{27}$$

Applying operator-ket duality using orthonormal basis $\mathcal{A} = \{|a_i\rangle\}$ to kets $\{|L_i\rangle\}$ results in a collection of orthonormal operators $\{L_i\}$ that map $\mathcal{H}_A$ to $\mathcal{H}_B$ (see Sec. II A). Using these operators, define $K_i : \mathcal{H}_A \mapsto \mathcal{H}_B$,

$$K_i := \sqrt{\mathbf{e}_i^\downarrow} L_i, \tag{28}$$

and notice from operator-ket duality we get

$$|K_i\rangle := \sqrt{\mathbf{e}_i^\downarrow} |L_i\rangle. \tag{29}$$

Lemma 1. Operators $\{K_i\}$ form a Kraus decomposition of $\mathcal{N}$,

$$\mathcal{N}(O) = \sum_i K_i O K_i^\dagger. \tag{30}$$

Proof. In (26), use (18) to obtain

$$J_{AB}^{\mathcal{N}} = \sum_i \mathbf{e}_i^\downarrow (I_A \otimes L_i)|\gamma\rangle\langle\gamma|(I_A \otimes L_i)^\dagger \tag{31}$$

$$= \sum_i (I_A \otimes K_i)|\gamma\rangle\langle\gamma|(I_A \otimes K_i)^\dagger, \tag{32}$$

where the second inequality uses (28). This second inequality, together with (23), gives

$$\mathcal{N}(|a_k\rangle\langle a_l|) = \sum_i K_i(|a_k\rangle\langle a_l|)K_i^\dagger. \tag{33}$$

This equality, together with linearity of $\mathcal{N}$, proves this lemma. ■

Using (26), (27), (28), and (29), one can show that the Kraus operators $\{K_i\}$ satisfy

$$\langle K_i, K_j \rangle = \langle K_i, K_i \rangle \delta_{ij} \quad \text{and} \quad \langle K_i, K_i \rangle \geqslant \langle K_j, K_j \rangle, \tag{34}$$

where $i \leqslant j$ and we use $\langle K_i, K_i \rangle = \mathbf{e}_i^\downarrow$. In addition to being orthogonal and ordered in the way captured by the above equation, the Kraus operators $\{K_i\}$ have several other useful properties. The total number of nonzero operators $\{K_i\}$ is the rank of the Choi-Jamiołkowsi operator $\mathcal{J}_{AB}^{\mathcal{N}}$. This rank is the minimum number of Kraus operators required to represent the channel $\mathcal{N}$. When the eigenvalues of $J_{AB}^{\mathcal{N}}$ are distinct, the norm $\langle K_i, K_i \rangle$ of each Kraus operator is simply the $(i+1)$th largest eigenvalue of $\mathcal{J}_{AB}^{\mathcal{N}}$. From these Kraus operators, one can obtain the Choi-Jamiołkowsi operator (22),

$$\mathcal{J}_{AB}^{\mathcal{N}} = \sum_i |K_i\rangle\langle K_i|, \tag{35}$$

where we have applied operator-ket duality (see Sec. II A) to convert operators $K_i : \mathcal{H}_A \mapsto \mathcal{H}_B$ to kets $|K_i\rangle \in \mathcal{H}_{AB}$ using basis $\mathcal{A} = \{|a_i\rangle\}$. Notice that $|K_i\rangle$ is an unnormalized eigenvector of $\mathcal{J}_{AB}^{\mathcal{N}}$ with eigenvalue $\langle K_i, K_i \rangle$. We call $\{K_i\}$ in Lemma 1 a standard Kraus decomposition.

B. Dual channel

Given a map $\mathcal{N} : \mathcal{L}(\mathcal{H}_A) \mapsto \mathcal{L}(\mathcal{H}_B)$, its dual $\mathcal{N}^\dagger : \mathcal{L}(\mathcal{H}_B) \mapsto \mathcal{L}(\mathcal{H}_A)$ is defined via

$$\mathrm{Tr}(\mathcal{N}^\dagger(O)\rho) = \mathrm{Tr}(O\mathcal{N}(\rho)), \tag{36}$$

where $\rho \in \mathcal{L}(\mathcal{H}_A)$ and $O \in \mathcal{L}(\mathcal{H}_B)$. (This definition of dual map (36), common in quantum information [see Def. (6.2) in [38] or below Eq. (1.44) in [39]], differs from another, $\langle \mathcal{N}^\dagger(O), \rho \rangle = \langle O, \mathcal{N}(\rho) \rangle$, found in mathematics literature. The two definitions coincide for maps satisfying $\mathcal{N}(\rho^\dagger) = (\mathcal{N}(\rho))^\dagger$, but they can differ when this property is not satisfied. For example, if $\mathcal{N}(\rho) = c\rho$, and c is complex, then the two definitions give different dual maps.) A quantum channel $\mathcal{N}$ evolves a quantum state ρ, and its dual channel $\mathcal{N}^\dagger$ evolves an observable O. The right side of the above equality represents the expectation value of the evolved quantum state $\mathcal{N}(\rho)$ with respect to a fixed observable O, while the left side of the equality gives the expectation value of a fixed state ρ with respect to the evolved observable $\mathcal{N}^\dagger(O)$. If $\mathcal{N}$ is CP and has Kraus decomposition (25), then $\mathcal{N}^\dagger$ is also CP with Kraus operators $\{K_i^\dagger\}$, and if $\mathcal{N}$ is trace-preserving, then $\mathcal{N}^\dagger$ is unital (see Chap. 6 in [38]). A CP map $\mathcal{N}$ with standard Kraus operators $\{K_i\}$ has dual map $\mathcal{N}^\dagger$ with standard Kraus operators $\{K_i^\dagger\}$ since

$$\langle K_i^\dagger, K_j^\dagger \rangle = (\langle K_i, K_j \rangle)^*. \tag{37}$$

The Choi-Jamiołkowsi operator (22) of the dual channel,

$$J_{BA}^{\mathcal{N}^\dagger} = \sum_i |K_i^\dagger\rangle\langle K_i^\dagger|, \tag{38}$$

where $\{|K_i^\dagger\rangle\}$ in $\mathcal{H}_{BA}$ are defined via operator-ket duality using basis $\mathcal{B} = \{|b_j\rangle\}$. We aim to compare $J_{BA}^{\mathcal{N}^\dagger}$, an operator on $\mathcal{H}_{BA}$, with $J_{AB}^{\mathcal{N}}$, an operator on a different space $\mathcal{H}_{AB}$. To carry out the comparison, interchange B and A in (38) and use (20) and (35) to get

$$J_{AB}^{\mathcal{N}} = \left(J_{AB}^{\mathcal{N}^\dagger}\right)^*. \tag{39}$$

The Choi-Jamiołkowsi operator of a channel and its dual can be taken to be complex conjugates of one another.

C. Extremal qubit channels

The set of quantum channels from $\mathcal{L}(\mathcal{H}_A)$ to $\mathcal{L}(\mathcal{H}_B)$ is convex, i.e., if $\mathcal{N}$ and $\mathcal{M}$ are quantum channels, then

$$\mathcal{K} = \lambda\mathcal{N} + (1-\lambda)\mathcal{M} \tag{40}$$

is a quantum channel for any $0 \leqslant \lambda \leqslant 1$. Any quantum channel $\mathcal{K}$ is extremal, i.e., it is an extreme point of the set of quantum channels, if equality of the type (40) holds only when $\lambda = 0$ or 1, or the only channels $\mathcal{N}$ and $\mathcal{M}$ satisfying the equality both equal $\mathcal{K}$.

A quantum channel $\mathcal{N} : \mathcal{L}(\mathcal{H}_A) \mapsto \mathcal{L}(\mathcal{H}_B)$ is called a qubit channel when $\mathcal{H}_A$ and $\mathcal{H}_B$ are two-dimensional. For these two-dimensional spaces, we can use the standard basis $\{|i\rangle\}$, where $i \in \{0, 1\}$, to define Pauli operators,

$$X = |0\rangle\langle 1| + |1\rangle\langle 0|, \quad Y = -i|0\rangle\langle 1| + i|1\rangle\langle 0|, \quad \text{and}$$
$$Z = |0\rangle\langle 0| - |1\rangle\langle 1|. \tag{41}$$

Extreme points of qubit channels are studied in various works [40–44]. A qubit channel is extremal if it has a single Kraus operator, given by a unitary operator, or it has two Kraus operators, each not proportional to a unitary operator (see Cor. 15 in [44]). Up to local unitaries at the channel input and output, a qubit channel $\mathcal{N}$ with two Kraus operators can be written as [42]

$$\mathcal{N}(O) = K_0 O K_0^\dagger + K_1 O K_1^\dagger, \tag{42}$$

where

$$K_0 = \begin{pmatrix} \cos(\frac{v-u}{2}) & 0 \\ 0 & \cos(\frac{v+u}{2}) \end{pmatrix},$$
$$K_1 = \begin{pmatrix} 0 & \sin(\frac{v+u}{2}) \\ \sin(\frac{v-u}{2}) & 0 \end{pmatrix}, \tag{43}$$

are expressed in the standard basis $\{|i\rangle\}$ at $\mathcal{H}_A$ and $\mathcal{H}_B$, $u \in [0, 2\pi]$ and $v \in [0, \pi)$.

While u and v parametrize the channel (42), they do not necessarily represent noise parameters that have a monotonic relationship with the amount of noise introduced by the channel. In certain special cases, noise parameters can be arrived at intuitively. For instance, when $u = 0$,

$$\mathcal{N}(O) = \cos^2\left(\frac{v}{2}\right)O + \sin^2\left(\frac{v}{2}\right)XOX \tag{44}$$

is a qubit dephasing channel with dephasing probability $\sin^2(v/2)$. (Notice that the dephasing channel is not extremal since each of its Kraus operators is proportional to a unitary operator [see the discussion above (42)].) By performing a unitary, X, at the input channel input $\mathcal{H}_A$, this dephasing channel (44) can be converted to another dephasing channel with dephasing probability $1 - \sin^2(v/2)$. Thus a dephasing probability of half gives maximum dephasing. This dephasing probability is an intuitive noise parameter in the sense that as this probability is increased from zero to a half, the channel becomes noisier.

Another special case is when $u + v = 2\pi$. Here, if kets $|0\rangle$ and $|1\rangle$ are interchanged at the channel input and output, $\mathcal{N}$ becomes a qubit amplitude damping channel. The qubit amplitude damping channel fixes $|0\rangle\langle 0|$ but $|1\rangle\langle 1|$ decays to $|0\rangle\langle 0|$ with probability $\sin^2 v$. Intuitively, this damping probability is a noise parameter in the sense that as the damping probability is increased from 0 to 1, the channel becomes noisier. Except for these special cases of dephasing and amplitude damping, suitable noise parameters are not necessarily easy to guess.

As discussed above, when $\mathcal{N}$ represents amplitude damping noise, the noise parameter is the damping probability. In all other cases, this qubit channel $\mathcal{N}$ can be generated from an isometry (see the discussion in Sec. III) of a special form. An isometry of this *pcubed* form [45],

$$V|\alpha_i\rangle = |\beta_i\rangle \otimes |\gamma_i\rangle, \tag{45}$$

where $i \in \{0, 1\}$, takes some special input *pure* states $\{|\alpha_i\rangle\}$ that are not necessarily orthogonal but form a basis of $\mathcal{H}_A$, to a *product of pure* states $\{|\beta_i\rangle\}$ at the $\mathcal{H}_B$ output and $\{|\gamma_i\rangle\}$ at the $\mathcal{H}_C$ output. The Gram matrices G_A, G_B, and G_C of $\{|\alpha_i\rangle\}$, $\{|\beta_j\rangle\}$, and $\{|\gamma_k\rangle\}$, respectively, satisfy

$$[G_A]_{ij} = \langle\alpha_i|\alpha_j\rangle = \langle\beta_i|\beta_j\rangle\langle\gamma_i|\gamma_j\rangle = [G_B]_{ij}[G_C]_{ij} \tag{46}$$

if and only if V is an isometry, i.e., $V^\dagger V = I_A$ [45]. These matrices take the form

$$G_A = \begin{pmatrix} 1 & a \\ a & 1 \end{pmatrix}, \quad G_B = \begin{pmatrix} 1 & b \\ b & 1 \end{pmatrix}, \quad \text{and} \quad G_C = \begin{pmatrix} 1 & c \\ c & 1 \end{pmatrix}, \tag{47}$$

where $-1 < a < 1$, $-1 \leqslant b \leqslant 1$, $-1 \leqslant c \leqslant 1$, and $a = bc$. The parameters b and c completely specify the isometry V in (45) and thus the channel $\mathcal{N}$. One may parametrize $|\alpha_i\rangle$ using the standard basis as

$$|\alpha_i\rangle = \sqrt{\frac{1+a}{2}}|0\rangle + (-1)^i\sqrt{\frac{1-a}{2}}|1\rangle. \tag{48}$$

In this parametrization, replacing a with b gives $|\beta_i\rangle$ and replacing a with c gives $|\gamma_i\rangle$. The parameters b and c are related to u and v in (43) as follows:

$$\sin^2 v = \frac{1-c^2}{1-(bc)^2} \quad \text{and} \quad \cos^2 u = \frac{1-b^2}{1-(bc)^2}, \tag{49}$$

where $|bc| \neq 1$. The Kraus operators in (43) can be written as

$$K_0 = \begin{pmatrix} \sqrt{\frac{(1+b)(1+c)}{2(1+bc)}} & 0 \\ 0 & \sqrt{\frac{(1-b)(1+c)}{2(1-bc)}} \end{pmatrix} \quad \text{and}$$
$$K_1 = \begin{pmatrix} 0 & \sqrt{\frac{(1+b)(1-c)}{2(1-bc)}} \\ \sqrt{\frac{(1-b)(1-c)}{2(1+bc)}} & 0 \end{pmatrix}. \tag{50}$$

While these Kraus operators look more complicated than those in (43), several other channel properties simplify when using the parameters b and c. For instance, the channel $\mathcal{N}$ with parameters b and c is degradable if $|b/c| < 1$, otherwise $|b/c| \geqslant 1$ and the channel is antidegradable [45].

In general, $-1 \leqslant b \leqslant 1$ and $-1 \leqslant c \leqslant 1$, however one can simplify the parameter space. In the discussion above, replacing b with $-b$ while keeping c fixed results in a new channel $\tilde{\mathcal{N}}$ which is equivalent to $\mathcal{N}$ up to local unitaries at the channel input and output. To see this, notice this replacement defines a new isometry $\tilde{V}$ of the *pcubed* form,

$$\tilde{V}|\tilde{\alpha}_i\rangle = |\tilde{\beta}_i\rangle \otimes |\gamma_i\rangle, \tag{51}$$

where $|\tilde{\alpha}_i\rangle$ and $|\tilde{\beta}_i\rangle$ are kets obtained from $|\alpha_i\rangle$ and $|\beta_i\rangle$ [see the definition below Eq. (48)] by replacing a and b with $-a$ and $-b$, respectively. This new isometry $\tilde{V}$ is related to V in (45) via local unitaries as follows:

$$(I_C \otimes X_B)\tilde{V} = VX_A, \tag{52}$$

where X is defined in (41). In a similar vein, a channel with parameters b and c is equivalent up to local unitaries to a channel with parameters b and $-c$. These equivalences allow us to restrict the parameter space $-1 \leqslant b \leqslant 1$ and $-1 \leqslant c \leqslant 1$ to the positive quadrant $0 \leqslant b \leqslant 1$ and $0 \leqslant c \leqslant 1$.

We show that any channel $\mathcal{N}$ with parameters b and c can simulate another channel $\mathcal{N}'$ with parameters b and $c' \leqslant c$, in the sense

$$\mathcal{N}' = \mathcal{N} \circ \mathcal{M}, \tag{53}$$

where $\mathcal{M}$ is a quantum channel. Proof of the above equation is easy to see from a *pcubed* point of view. Let $\mathcal{N} : \mathcal{L}(\mathcal{H}_A) \mapsto \mathcal{L}(\mathcal{H}_B)$ be generated by the isometry in (45), let $\mathcal{N}' : \mathcal{L}(\mathcal{H}_A) \mapsto \mathcal{L}(\mathcal{H}_B)$ be generated by an isometry $V' : \mathcal{H}_A \mapsto \mathcal{H}_B \otimes \mathcal{H}_{C'}$ of the same form as V in (45), but

$$V'|\alpha'_i\rangle = |\beta_i\rangle \otimes |\gamma'_i\rangle, \tag{54}$$

where $c' = \langle\gamma'_0|\gamma'_1\rangle$ and $a' = \langle\alpha'_0|\alpha'_1\rangle = bc'$. The $\mathcal{M} : \mathcal{L}(\mathcal{H}_A) \mapsto \mathcal{L}(\mathcal{H}_A)$ channel in (53) is generated by an isometry $W : \mathcal{H}_A \mapsto \mathcal{H}_A \otimes \mathcal{H}_D$ of the form (45) with

$$W|\alpha'_i\rangle = |\alpha_i\rangle \otimes |\delta_i\rangle, \tag{55}$$

where $d := \langle\delta_0|\delta_1\rangle = c'/c$ takes values between 0 and 1 since $0 \leqslant c' \leqslant c$. The relationship in (53) ensures that $\mathcal{N}'$ is noisier than $\mathcal{N}$. As a result, for fixed b, if one decreases c then the channel $\mathcal{N}$ becomes noisier.

This parameter c captures the lack of distinguishability between pure states arriving at the environment. If c is decreased, more information flows to the environment. The no-cloning theorem [46–48] indicates that such a flow to the environment must come at the cost of information flow to the output. Thus $\mathcal{N}$ becomes noisier with decreasing c. We shall be interested in using c as the noise parameter with b fixed. In the limiting $b = 0$ case, $\mathcal{N}$ becomes the qubit dephasing channel (44) with dephasing probability $(1 - c)/2$. Here, decreasing c from 1 to 0 increases the dephasing probability from 0 to half.

IV. OPTIMAL ENTANGLEMENT SHARING

A. High fidelity entanglement

Consider two parties, Alice and Bob, connected by some quantum channel $\mathcal{N} : \mathcal{L}(\mathcal{H}_A) \mapsto \mathcal{L}(\mathcal{H}_B)$, where $\mathcal{H}_A$ and $\mathcal{H}_B$ have the same dimension d. Suppose Alice has access to a second d-dimensional system with Hilbert space $\mathcal{H}_R$. What bipartite state ρ_{RA} should Alice prepare such that sharing with Bob one half of this state across the channel $\mathcal{N}$ results in a state ρ_{RB} with the highest fidelity $F(\rho_{RB}, \phi_{RB})$ to a maximally entangled state,

$$|\phi\rangle_{RB} = \frac{1}{\sqrt{d}}|\gamma\rangle_{RB}, \tag{56}$$

between reference $\mathcal{H}_R$ and output $\mathcal{H}_B$? The optimal state prepared by Alice, which we denote by Λ_{RA}, and the maximum fidelity,

$$\mathcal{O}(\mathcal{N}) := F(\Lambda_{RB}, \phi_{RB}), \tag{57}$$

have been characterized previously in terms of the channel's Choi-Jamiołkowski operator [30,31,34] when ρ_{RA} is pure. For possibly mixed ρ_{RA}, our reformulation of these results in terms of the standard Kraus decomposition of a channel and the operator norm of the channel's Choi-Jamiołkowski operator agree with these previous characterizations. We extend these results by finding families of mixed input states Λ_{RB} that achieve $\mathcal{O}(\mathcal{N})$. This reformulation and extension is used later in our study. We begin our reformulation using a semidefinite program

$$\begin{aligned} &\text{maximize } F(\rho_{RB}, \phi_{RB}) \\ &\text{subject to } \rho_{RB} = \mathcal{I}_R \otimes \mathcal{N}(\rho_{RA}), \\ &\rho_{RA} \succeq 0, \quad \text{Tr}(\rho_{RA}) = 1. \end{aligned} \tag{58}$$

The optimum value of the above program gives $\mathcal{O}(\mathcal{N})$, and the density operator which achieves this optimum gives Λ_{RA}. The following Theorem captures the solution to the above problem:

Theorem 1. Given a channel $\mathcal{N}$ with standard Kraus operators $\{K_i\}$,

$$\mathcal{O}(\mathcal{N}) = \frac{1}{d}\langle K_0, K_0\rangle = \frac{1}{d}\left|\left|J^{\mathcal{N}}_{RB}\right|\right| = F(\Lambda_{RB}, \phi_{RB}), \tag{59}$$

where the input Λ_{RA} has support in the span of $\{|K_i^\dagger\rangle_{RA}\}$ satisfying $\langle K_i, K_i\rangle = \langle K_0, K_0\rangle$.

Proof. Using Eq. (7) along with the fact that ϕ_{RB} is a pure state, one writes $F(\rho_{RB}, \phi_{RB})$ as an inner product $\langle\rho_{RB}, \phi_{RB}\rangle$. This inner product is rewritten as $\langle\mathcal{I}_R \otimes \mathcal{N}(\rho_{RA}), \phi_{RB}\rangle$ using the first equality constraint in (58). This rewriting can be reduced to $\langle\rho_{RA}, (\mathcal{I}_R \otimes \mathcal{N})^\dagger(\phi_{RB})\rangle$ using definition (36) of the dual channel. Using the discussion below (36), or otherwise, one can show that the dual of the tensor product of two channels is the tensor product of the dual of individual channels. Thus $\langle\rho_{RA}, (\mathcal{I}_R \otimes \mathcal{N})^\dagger(\phi_{RB})\rangle = \langle\rho_{RA}, \mathcal{I}_R \otimes \mathcal{N}^\dagger(\phi_{RB})\rangle$, where we used the fact that $\mathcal{I}_R^\dagger$ is $\mathcal{I}_R$. Next, notice $(\mathcal{I}_R \otimes \mathcal{N}^\dagger)\phi_{RB}$ is just $\mathcal{J}^{\mathcal{N}^\dagger}_{RA}/d$ (22). Using these observations, rewrite (58) as

$$\begin{aligned} &\text{maximize } \frac{1}{d}\langle\rho_{RA}, \mathcal{J}^{\mathcal{N}^\dagger}_{RA}\rangle \\ &\text{subject to } \rho_{RA} \succeq 0, \quad \text{Tr}(\rho_{RA}) = 1. \end{aligned} \tag{60}$$

Solution to this semidefinite program is $(1/d)$ times the maximum eigenvalue of $J^{\mathcal{N}^\dagger}_{RA}$ obtained by setting $\rho_{RA} = \Lambda_{RA}$, where Λ_{RA} is any density operator with support on the eigenspace of this maximum eigenvalue. This largest eigenvalue can be written as $\langle K_0^\dagger, K_0^\dagger\rangle = \langle K_0, K_0\rangle$ using (34) and

(37). The largest eigenvalue can also be written as the spectral norm, $||J^{\mathcal{N}}_{RB}||$, by applying definition (5). The support of the largest eigenvalue, $\langle K_0, K_0\rangle$, of $J^{\mathcal{N}^\dagger}_{RA}$ is the span of the collection of eigenvectors corresponding to this value. This collection contains eigenvectors $|K_i^\dagger\rangle$ of $J^{\mathcal{N}^\dagger}_{RA}$ [see (38)] with eigenvalue $\langle K_i^\dagger, K_i^\dagger\rangle$ equaling the largest eigenvalue $\langle K_0^\dagger, K_0^\dagger\rangle$. The eigenvalues of $J^{\mathcal{N}^\dagger}_{RA}$ can be shown to equal corresponding eigenvalues of $J^{\mathcal{N}}_{RA}$ using (39), i.e., one can show that $\langle K_i^\dagger, K_i^\dagger\rangle = \langle K_i, K_i\rangle$. ■

The fidelity between a fixed state ρ_{AB} and a fully entangled state, maximized over all possible fully entangled states, is called the *fully entangled fraction* [5,29]

$$F_e(\rho_{AB}) = \max_{U_A} F(\rho_{AB}, (U_A\otimes I_B)\phi_{AB}(U_A\otimes I_B)^\dagger), \tag{61}$$

where U_A is a unitary operator on $\mathcal{H}_A$.

Lemma 2. The *largest* fully entangled fraction obtained by sending one-half of a mixed state ρ_{RA} across the channel $\mathcal{N}$ maximized over all ρ_{RA} equals $\mathcal{O}(\mathcal{N})$ (57).

Proof. Notice that the largest fully entangled fraction can be found by modifying the optimization problem (58) as follows: replace ϕ_{RB} with $\chi_{RB} = (U_R\otimes I_B)\phi_{RB}(U_R\otimes I_B)^\dagger$ and optimize over both unitary matrices U_R and density operators ρ_{RA}. Notice, in this larger optimization problem, one can simplify the objective function $F(\rho_{RB}, \chi_{RB}) = F(\rho'_{RB}, \phi_{RB})$, where $\rho'_{RB} = (U_R\otimes I_B)^\dagger\rho_{RB}(U_R\otimes I_B)$. Since $\rho'_{RB} = \mathcal{I}\otimes\mathcal{N}(\rho'_{RA})$, where $\rho'_{RA} = (U_R\otimes I_A)^\dagger\rho_{RA}(U_R\otimes I_A)$, one can rephrase this optimization at hand purely in terms of a single new variable ρ'_{RA}, satisfying $\rho'_{RA}\succeq 0$ and $\mathrm{Tr}(\rho'_{RA}) = 1$. In this rephrasing variable, U_R no longer participates. However, the new problem in terms of ρ'_{RA} is identical to (57). ■

The above result generalizes to a mixed state that was implicitly found for pure states in the proof of Lemma 2 in [31].

Let Λ_{RA} be the state in Theorem 1. We are interested in the minimum amount of entanglement over all states of this type. To capture this minimum, we use entanglement of formation (10). When Λ_{RA} is a unique pure state, we write the *input entanglement*

$$\mathcal{E}(\mathcal{N}) = S(\sigma_A), \tag{62}$$

and when Λ_{RA} can be chosen to be mixed, we write

$$\mathcal{E}(\mathcal{N}) = \min_{\Lambda_{RA}} E_f(\Lambda_{RA}), \tag{63}$$

where Λ_{RA} are states in Theorem 1. When Λ_{RA} can be chosen to be a separable state, $\mathcal{E}(\mathcal{N}) = 0$.

B. Multiplicativity

Suppose Alice and Bob are connected by two independent channels, which may be the same or different. What state should Alice prepare such that sending one-half of it across the joint channel results in Alice and Bob sharing a joint state with maximum fidelity to a fully entangled state? What is this maximum fidelity? Can one hope to use correlations across the two channels connecting Alice and Bob to get more fidelity than what can be achieved without using any correlation across the channels? Variants of these natural questions have been asked about the transmission of information across asymptotically many uses of quantum channels. Those questions have been hard to answer. Here we mathematically formulate and answer the questions we posed above.

Let the two channels connecting Alice and Bob be $\mathcal{N}_1 : \mathcal{L}(\mathcal{H}_{A1})\mapsto\mathcal{L}(\mathcal{H}_{B1})$ and $\mathcal{N}_2 : \mathcal{L}(\mathcal{H}_{A2})\mapsto\mathcal{L}(\mathcal{H}_{B2})$; here $d_{A1} = d_{B1}$ and $d_{A2} = d_{B2}$. For each channel input $\mathcal{H}_{A1}$ and $\mathcal{H}_{A2}$, define auxiliary spaces $\mathcal{H}_{R1}$ and $\mathcal{H}_{R2}$. Let $\mathcal{I}_{R1}$ and $\mathcal{I}_{R2}$ be identity maps on these auxiliary spaces, $\mathcal{L}(\mathcal{H}_{R1})$ and $\mathcal{L}(\mathcal{H}_{R2})$, respectively, $\mathcal{H}_A := \mathcal{H}_{A1}\otimes\mathcal{H}_{A2}$, $\mathcal{H}_B := \mathcal{H}_{B1}\otimes\mathcal{H}_{B2}$, $\mathcal{H}_R = \mathcal{H}_{R1}\otimes\mathcal{H}_{R2}$, $\mathcal{N} = \mathcal{N}_1\otimes\mathcal{N}_2$, and $\mathcal{I}_R = \mathcal{I}_{R1}\otimes\mathcal{I}_{R2}$. If Alice prepares a state that does not correlate inputs to the two channels $\mathcal{I}_{R1}\otimes\mathcal{N}_1$ and $\mathcal{I}_{R2}\otimes\mathcal{N}_2$, then the maximum fidelity with a fully entangled state across auxiliary space $\mathcal{H}_R$ and the channel output $\mathcal{H}_B$ can be found as follows:

$$\begin{aligned}
&\text{maximize } F(\rho_{RB}, \phi_{RB})\\
&\text{subject to } \rho_{RB} = (\mathcal{I}_R\otimes\mathcal{N})\rho_{RA},\\
&\rho_{RA} = \rho_{R1A1}\otimes\rho_{R2A2},\\
&\rho_{RA}\succeq 0,\quad \mathrm{Tr}(\rho_{RA}) = 1.
\end{aligned}\tag{64}$$

The optimum of the above problem is simply $O(\mathcal{N}_1)O(\mathcal{N}_2)$. It is obtained at $\Lambda_{RA} = \Lambda_{R1A1}\otimes\Lambda_{R2A2}$, where Λ_{R1A1} and Λ_{R2A2} are optima to optimizations of the form (58) for $\mathcal{N}_1$ and $\mathcal{N}_2$, respectively. On the other hand, if Alice prepares a state that may correlate the inputs to $\mathcal{I}_{R1}\otimes\mathcal{N}_1$ and $\mathcal{I}_{R2}\otimes\mathcal{N}_2$, then the maximum fidelity $\mathcal{O}(\mathcal{N}_1\otimes\mathcal{N}_2)$ is found by solving (64) without the product constraint, $\rho_{RA} = \rho_{R1A1}\otimes\rho_{R2A2}$. This fidelity maximum $O(\mathcal{N}_1\otimes\mathcal{N}_2)$ can be higher,

$$O(\mathcal{N}_1\otimes\mathcal{N}_2)\geqslant O(\mathcal{N}_1)O(\mathcal{N}_2), \tag{65}$$

since the optimum $O(\mathcal{N}_1)O(\mathcal{N}_2)$ of (64) bounds from below the optimum of (64) without the product constraint, $\rho_{RA} = \rho_{R1A1}\otimes\rho_{R2A2}$.

Theorem 2. The maximum fidelity $O(\mathcal{N}_1\otimes\mathcal{N}_2)$ is multiplicative, i.e., equality holds in (65),

$$O(\mathcal{N}_1\otimes\mathcal{N}_2) = O(\mathcal{N}_1)O(\mathcal{N}_2). \tag{66}$$

Proof. Let $\mathcal{N}_1$ and $\mathcal{N}_2$ have standard Kraus decomposition $\{J_q\}$ and $\{K_r\}$, respectively. Using Theorem 1, we write

$$\mathcal{O}(\mathcal{N}_1) = \frac{1}{d_{A1}}\langle J_0, J_0\rangle \quad\text{and}\quad \mathcal{O}(\mathcal{N}_2) = \frac{1}{d_{A1}}\langle K_0, K_0\rangle. \tag{67}$$

A standard Kraus decomposition $\{L_p\}$ for $\mathcal{N}_1\otimes\mathcal{N}_2$ can be chosen such that each L_p is of the form $J_q\otimes K_r$ for some q and r. When $q = r = 0$, then p can be chosen to be 0,

$$L_0 = J_0\otimes K_0, \tag{68}$$

since

$$\langle L_0, L_0\rangle = \langle J_0, J_0\rangle\langle K_0, K_0\rangle \geqslant \langle J_q, J_q\rangle\langle K_r, K_r\rangle = \langle L_p, L_p\rangle \tag{69}$$

for all q, r and corresponding p. Using (68) and Theorem 1 on $\mathcal{N}_1\otimes\mathcal{N}_2$ gives

$$\mathcal{O}(\mathcal{N}_1\otimes\mathcal{N}_2) = \frac{1}{d_{A1}d_{A2}}\langle L_0, L_0\rangle. \tag{70}$$

The above equality, together with (67) and (69), proves the result.

Alternatively, notice

$$J^{\mathcal{N}}_{RB} = \sum_p |L_p\rangle\langle L_p| = \sum_{qr} |J_q\rangle\langle J_q| \otimes |K_r\rangle\langle K_r|$$
$$= J^{\mathcal{N}_1}_{R1B1} \otimes J^{\mathcal{N}_2}_{R2B2}, \tag{71}$$

where the first equality follows from (35). Using Theorem 1, write

$$\mathcal{O}(\mathcal{N}_1) = \frac{1}{d_{A1}} ||J^{\mathcal{N}_1}_{R1B1}||,$$
$$\mathcal{O}(\mathcal{N}_2) = \frac{1}{d_{A2}} ||J^{\mathcal{N}_2}_{R2B2}||, \quad \text{and}$$
$$\mathcal{O}(\mathcal{N}_1 \otimes \mathcal{N}_2) = \frac{1}{d_A} ||J^{\mathcal{N}}_{RB}||, \tag{72}$$

where $d_A = d_{A1} d_{A2}$. The operator norm is submultiplicative (see Sec. 1.1.3 in [49]),

$$||AB|| \leqslant ||A|| \times ||B||, \tag{73}$$

which implies

$$||A \otimes B|| \leqslant ||A|| \times ||B||. \tag{74}$$

Using the above equation along with (65) and (72) also proves the result. ■

V. APPLICATIONS

A. Extremal qubit channels

A qubit channel $\mathcal{N}$ has $d_A = d_B = 2$. If the channel has one Kraus operator, then the channel is simply a conjugation with a unitary matrix and $\mathcal{O}(\mathcal{N}) = 1$. The next simplest qubit channel has two Kraus operators, given in (42). One special case of this channel is the qubit amplitude damping channel. Kraus operators for this amplitude channel can be written as

$$K_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix} \quad \text{and} \quad K_1 = \begin{pmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{pmatrix}, \tag{75}$$

where $0 \leqslant p \leqslant 1$ is the probability that the state $|1\rangle\langle 1|$ damps to $|0\rangle\langle 0|$. A simple calculation shows that these Kraus operators constitute a standard Kraus decomposition of $\mathcal{N}$. Using this decomposition in Theorem 1, we find

$$\mathcal{O}(\mathcal{N}) = 1 - p/2 \quad \text{and} \quad \Lambda_{RA} = \frac{|K_0\rangle\langle K_0|}{\langle K_0, K_0\rangle}, \tag{76}$$

a result that agrees with [33]. In general, the amount of entanglement generated at the input [see the definition in (62)],

$$E(\mathcal{N}) = h\left(\frac{1}{2-p}\right), \tag{77}$$

where $h(x) := -x \log x - (1-x)\log(1-x)$ is the binary entropy function, with log base 2. This value is nonzero unless $p = 1$, where $E(\mathcal{N}) = 0$ and Λ_{RA} in (76) is a product state.

When the qubit channel $\mathcal{N}$ with two Kraus operators is not an amplitude damping channel, the channel Kraus operators take the form (50). These Kraus operators $\{K_0, K_1\}$ have two parameters $0 \leqslant b \leqslant 1$ and $0 \leqslant c \leqslant 1$. If b is fixed and c is decreased from 1, the channel becomes more noisy [see the discussion containing (53)]. Operators $\{K_0, K_1\}$ form a standard Kraus decomposition. Using them in Theorem IV A gives

$$\mathcal{O}(\mathcal{N}) = \frac{(1+c)(1-b^2c)}{2(1-b^2c^2)} \quad \text{and}$$
$$\Lambda_{RA} = \begin{cases} \frac{|K_0^\dagger\rangle\langle K_0^\dagger|}{\langle K_0^\dagger, K_0^\dagger\rangle} & \text{if} \quad b \neq 1 \text{ and } c \neq 0, \\ \sum_{ij} f_{ij} |K_i^\dagger\rangle\langle K_j^\dagger| & \text{if} \quad b = 1 \text{ or } c = 0, \end{cases} \tag{78}$$

where complex numbers f_{ij} are free except that they result in a valid density operator Λ_{RA} (see Fig. 3). At $b = 1$ or $c = 0$, Λ_{RA} is supported on a two-dimensional space spanned by $\{|K_0^\dagger\rangle_{RA}, |K_1^\dagger\rangle_{RA}\}$. This two-dimensional space is a subspace of a two-qubit space $\mathcal{H}_{RA}$. Quite generally, such a subspace has at least one product state (see the Lemma in [50]), but typically there are two [42,45]. In the $c = 0$ case, these product states take the simple form

$$|+\rangle_R \otimes |\psi_+\rangle_A \quad \text{and} \quad |-\rangle_R \otimes |\psi_-\rangle_A, \tag{79}$$

where $|\psi_+\rangle_A = \frac{1}{\sqrt{2}}(\sqrt{1+b}|0\rangle + \sqrt{1-b}|1\rangle)$, $|\psi_-\rangle_A = \frac{1}{\sqrt{2}}(\sqrt{1+b}|0\rangle - \sqrt{1-b}|1\rangle)$, $|+\rangle_A = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, and $|-\rangle_A = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$.

At $b = 1$ or $c = 0$ one can choose Λ_{RA} to be a projector onto a product state. As a result, at $b = 1$ or $c = 0$, the input

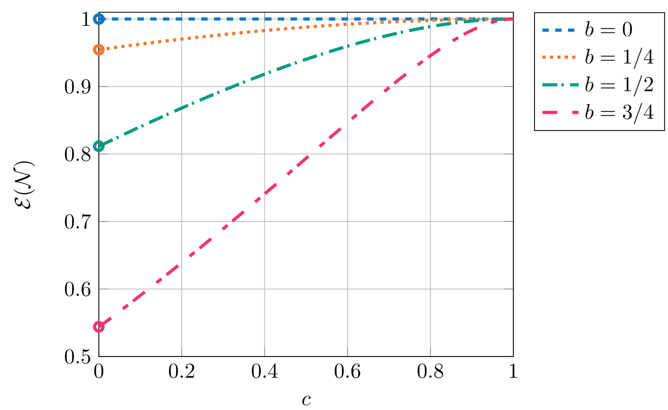

FIG. 2. Plot of $\mathcal{E}(\mathcal{M})$ as a function of c for various b values. The open circle indicates that the value is zero.

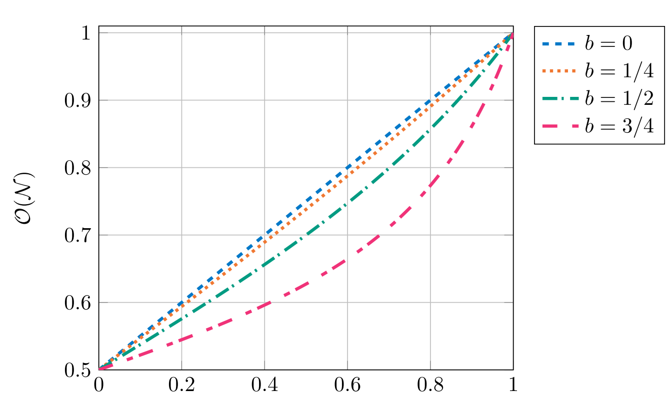

FIG. 3. Plot of $\mathcal{O}(\mathcal{M})$ as a function of c for various b values.

entanglement, defined in (9), is zero. In general,

$$E(\mathcal{N}) = \begin{cases} 0 & \text{if} \quad b = 1 \text{ or } c = 0, \\ h\left(\frac{(1+b)(1-bc)}{2(1-b^2c)}\right) & \text{otherwise,} \end{cases} \tag{80}$$

where the expressions for $\mathcal{E}(\mathcal{N})$ at $b \neq 1$ and $c \neq 0$ come from using the form of Λ_{RA} in (78). In Fig. 2 we fix b and plot $E(\mathcal{N})$ as a function of c; increasing c makes $\mathcal{N}$ less noisy [see the discussion containing Eq. (53)]. In these plots, as c is increased from zero, the value of $\mathcal{E}(\mathcal{N})$ increases discontinuously from 0, at $c = 0$, and continues to increase monotonically until $c = 1$, where $\mathcal{N}$ becomes a perfect channel. Across various plots with fixed b, we notice that increasing b decreases $\mathcal{E}(\mathcal{N})$, which ultimately goes to zero as $b \mapsto 1$ for all $bc \neq 1$.

All these features mentioned above are intriguing. In the parameter range $0 < c < 1$, one finds an expected result [34] that the minimum amount of entanglement needed at the input to have maximum fidelity with a fully entangled output is strictly less than 1. In particular, if one generates more than $\mathcal{E}(\mathcal{N}) < 1$ entanglement at the input, the fidelity with a maximally entangled output is strictly less. The key addition here is the quantification of the amount of entanglement and a parametrization of the channel in such a way that the amount of entanglement is monotone in the noise parameters of the channel.

Next, at $c = 0$, there is a discontinuous change in $\mathcal{E}(\mathcal{N})$ which starts at zero and then takes a large finite value $\simeq h((1+b)/2)$. From a mathematical standpoint, the discontinuity arises because the solution to the optimization (58) becomes degenerate and this degeneracy allows more freedom in choosing optimum inputs. Due to the structure of qubit channels, this input can be chosen to be separable, as mentioned in the discussion containing (79).

B. Qubit Pauli channels

A qubit Pauli channel $\mathcal{N} : \mathcal{H}_A \mapsto \mathcal{H}_B$ can be written as

$$\mathcal{N}(\rho) = \sum_i p_i \sigma_i \rho \sigma_i^\dagger, \tag{81}$$

where $p_i \geqslant 0$, $\sum_i p_i = 1$, and the Kraus operators $\{\sqrt{p_i}\sigma_i\}$, $\sigma_i : \mathcal{H}_A \mapsto \mathcal{H}_B$, are proportional to Pauli matrices. These matrices can be written in the standard $\{|0\rangle, |1\rangle\}$ basis of $\mathcal{H}_A$ and $\mathcal{H}_B$ as

$$\sigma_0 = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

$$\sigma_2 = Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \text{and} \quad \sigma_3 = Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{82}$$

Without loss of generality, we can assume $p_0 \geqslant p_i$ for all $i \in \{1, 2, 3\}$. This assumption comes from the following argument. Assume $p_i \geqslant p_j$ for some $i \neq 0$ and all $j \in \{0, 1, 2, 3\}$. Then conjugating the input ρ with σ_i will still result in a Pauli channel (81). However, this resulting channel will have $p_0 \geqslant p_i$ for all $i \in \{1, 2, 3\}$.

For qubit Pauli channels, the value of $\mathcal{O}(\mathcal{N})$ and the fact that it can be achieved using a maximally entangled input state Λ_{RA} was found in [34]; however, we note later that one can also achieve $\mathcal{O}(\mathcal{N})$ using a separable pure state when $\mathcal{N}$ is very noisy. Since the Pauli matrices are orthogonal to each other, in a standard Kraus decomposition, $\{K_i\}$, of $\mathcal{N}$ we can always choose each K_i to be $\sqrt{p_j}\sigma_j$ for some j. As $p_0 \geqslant p_i$, $K_0 = \sqrt{p_0}\sigma_0$, and from Theorem 1 we get

$$\mathcal{O}(\mathcal{N}) = p_0. \tag{83}$$

When $p_0 > p_j$ for all j, $\Lambda_{RA} = |\sigma_0^\dagger\rangle\langle\sigma_0^\dagger|/2$, i.e., Λ_{RA} is a projector onto a maximally entangled state and thus $\mathcal{E}(\mathcal{N}) = 1$. However, if for some i, $p_0 = p_i$ then Λ_{RA} is any density operator with support in a space spanned by $\{|\sigma_0^\dagger\rangle_{RA}, |\sigma_i^\dagger\rangle_{RA}\}$. This space is a two-dimensional subspace of a two-qubit space $\mathcal{H}_{RA}$. Following the discussion containing (79), this subspace contains at least one product state. As a result, for any i if $p_0 = p_i$ we can choose Λ_{RA} to be a product state and thus $\mathcal{E}(\mathcal{N}) = 0$. Consequently,

$$\mathcal{E}(\mathcal{N}) = \begin{cases} 1 & \text{if} \quad p_0 > p_i \ \forall i, \\ 0 & \text{if} \quad p_0 = p_i \text{ for some } i. \end{cases} \tag{84}$$

When $p_0 = p_i$, the best fidelity with a maximally entangled state at the output is achieved by sending a separable input Λ_{RA}. Consequently, the output Λ_{RB} is also separable. This separable output is expected to have a small fidelity with a fully entangled state. This expectation is met, the condition $p_0 = p_i$ together with $\sum_i p_i = 1$ forces $p_0 \leqslant 1/2$, and thus $\mathcal{O}(\mathcal{N}) \leqslant 1/2$. Such a value of half for fidelity with a maximally entangled state $|\phi\rangle_{AB}$ is considered small since this value of half can be achieved by a simple separable state $\rho_{RB} = \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|)$.

One may wonder which qubit Pauli channels satisfy $p_0 = p_i \geqslant p_j$. Any qubit Pauli channel of this type is antidegradable. In general, $\mathcal{N}$ in (81) with $p_0 \geqslant p_i$ is antidegradable [50–52] if and only if

$$p_1 + p_2 + p_3 + \sqrt{p_1 p_2} + \sqrt{p_1 p_3} + \sqrt{p_2 p_3} \geqslant 1/2. \tag{85}$$

We are interested in the case where $p_0 = p_i$ for some i. The above condition remains unaffected when permuting p_i and p_j, thus we let $p_0 = p_1 = p$, denote p_2 by q, and then $p_3 = 1 - 2p - q$. Using these substitutions on the left side of (85), together with $1 \geqslant p \geqslant q \geqslant 0$ and $p \geqslant 1 - 2p - q$, we find that the above inequality (85) is always satisfied. Thus $p_0 = p_i \geqslant p_j$ implies that the qubit Pauli channel $\mathcal{N}$ is antidegradable.

Pauli channels (81) have a key property: up to local unitaries at the channel input and output, any unital qubit channel can always be written as a Pauli channel [42]. An interesting observation about qubit channels is that Λ_{RA} in Theorem 1 can be chosen to be a maximally entangled state if and only if $\mathcal{N}$ is unital [34]. It is interesting for that reason to ask if such a result holds in higher dimension. In this next section, we find that it does not.

In the case of qubit Pauli channels, but also for extremal qubit channels, we found that it is possible to find separable input states Λ_{RA} that achieve the most fidelity with a fully entangled state at the channel output. This separable state appeared when a qubit channel $\mathcal{N}$'s standard Kraus decomposition $\{K_i\}$ satisfied the condition $\langle K_0, K_0\rangle = \langle K_j, K_j\rangle$ for at least one $j \neq 0$. Using Eq. (35), this condition reduces to the channel's Choi-Jamiołkowsi operator $J_{RB}^{\mathcal{N}}$ having its largest eigenvalue be degenerate. In general, we have the following lemma.

Lemma 3. If $\mathcal{N}$ is a qubit channel and the largest eigenvalue of $J_{RB}^{\mathcal{N}}$ is degenerate, then Λ_{RA} in Theorem 1 can be chosen to be separable.

Proof. Let $\{K_i\}$ be a standard Kraus decomposition of $\mathcal{N}$. Since $J_{RB}^{\mathcal{N}}$ is degenerate, $\langle K_0, K_0\rangle = \langle K_1, K_1\rangle$ and Λ_{RA} has support in the span of $\{|K_0^\dagger\rangle, |K_1^\dagger\rangle\}$. This support is a two-dimensional subspace of a two-qubit space, and thus contains a product state. Hence Λ_{RA} can be chosen to be a projector onto this product state. ■

While it may be tempting to conjecture that the above result holds in higher-dimensional channels, we show in the next section that it does not.

C. Some qutrit channels

We construct two qutrit channels. The first channel, $\mathcal{M}$, is not unital but its optimal input state Λ_{RA}, defined in Theorem 1, is unique and maximally entangled. The second channel, $\mathcal{P}$, is unital, however its optimal input state Λ_{RA} is neither maximally entangled nor separable. Using the second channel, we demonstrate that when the largest eigenvalue of $J_{RB}^{\mathcal{N}}$ is degenerate, Λ_{RA} can still be entangled. The demonstration contrasts with Lemma 3.

Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be three-dimensional Hilbert spaces. Let $\mathcal{M} : \mathcal{L}(\mathcal{H}_A) \mapsto \mathcal{L}(\mathcal{H}_B)$ be a channel with Kraus operators

$$K_0 = \sqrt{\lambda} I, \quad K_1 = \sqrt{1-\lambda}(|0\rangle\langle 1| + |1\rangle\langle 0|), \quad \text{and}$$
$$K_2 = \sqrt{1-\lambda}|1\rangle\langle 2|, \tag{86}$$

where $0 \leqslant \lambda \leqslant 1$. This channel $\mathcal{M}$ is not unital, except when $\lambda = 1$. When $2/5 < \lambda < 1$, $\{K_i\}$ is a standard Kraus decomposition of $\mathcal{M}$ with $\langle K_0, K_0\rangle > \langle K_i, K_i\rangle$ for all $i \neq 0$. From Theorem 1 we find

$$\mathcal{O}(\mathcal{M}) = \lambda, \quad \Lambda_{RA} = \tfrac{1}{3}|I\rangle\langle I|, \quad \text{and} \quad \mathcal{E}(\mathcal{M}) = \log_2 3. \tag{87}$$

Thus when $2/5 < \lambda < 1$, the input Λ_{RA} is unique, and it is maximally entangled, however the channel $\mathcal{M}$ is not unital.

Let $\mathcal{P} : \mathcal{L}(\mathcal{H}_A) \mapsto \mathcal{L}(\mathcal{H}_B)$ be a qutrit channel with Kraus operators

$$L_0 = \sqrt{\frac{z+2}{4}}(|0\rangle\langle 1| + |1\rangle\langle 0|),$$
$$L_1 = \sqrt{\frac{1-z}{2}}(|1\rangle\langle 2| + |2\rangle\langle 1|),$$
$$L_2 = \sqrt{\frac{1-z}{2}}(|0\rangle\langle 2| + |2\rangle\langle 0|), \quad \text{and}$$
$$L_3 = \sqrt{\frac{z}{4}}(|0\rangle\langle 0| + |1\rangle\langle 1| - 2|2\rangle\langle 2|), \tag{88}$$

where $0 \leqslant z \leqslant 1$. Since each Kraus operator L_i is Hermitian, $\mathcal{P}$ is unital [see the discussion below (25)]. Kraus operators $\{L_i\}$ are standard and thus Theorem 1 immediately gives $\mathcal{O}(\mathcal{M}) = (z+2)/6$. When $z \neq 0$,

$$\Lambda_{AR} = |L_0^\dagger\rangle\langle L_0^\dagger|, \tag{89}$$

where $|L_0^\dagger\rangle_{RA} = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$ is not a maximally entangled state of two qutrits. When $z = 0$, $L_3 = 0$, $\langle L_0|L_0\rangle = \langle L_1|L_1\rangle = \langle L_2|L_2\rangle$ and thus largest eigenvalue of $J_{RB}^{\mathcal{M}}$ has a degenerate spectrum. In this case, Λ_{RA} has support in a subspace $\mathcal{S}$ spanned by $\{|L_0^\dagger\rangle_{RA}, |L_1^\dagger\rangle_{RA}, |L_2^\dagger\rangle_{RA}\}$. This subspace only contains nonproduct vectors, i.e., it is *completely entangled* in the sense of Parthasarathy (see Definition 1.2 in [53]). Consequently, any density operator Λ_{RA} supported on this subspace is entangled.

VI. DISCUSSION

In this work, we considered a one-shot setting where one-half of any bipartite mixed state may be sent across a single use of a fixed channel $\mathcal{N}$. The goal in this setting is to share a state with maximum fidelity $\mathcal{O}(\mathcal{N})$ to a fully entangled state. Interestingly, maximum fidelity $\mathcal{O}$ defined in the one-shot setting fully characterizes the ability of any channel to share high fidelity entanglement over multiple channel uses, possibly used in parallel with other channels. This extension follows from the multiplicative nature of $\mathcal{O}$, proved in Sec. IV B.

Using a semidefinite program, we reformulate the maximum fidelity, found previously for pure state inputs [30,31,34]. The first reformulation, see Theorem 1 and its proof, makes greater use of a channel's Kraus operators rather than its Choi-Jamiołkowski operator, as done previously. In particular, optimal input(s) achieving $\mathcal{O}(\mathcal{N})$ are simply linear combinations of flattened versions of a channel's standard Kraus operators with the largest norm, and the optimal value $\mathcal{O}(\mathcal{N})$ is this largest norm itself. These two channel representations are formally equivalent (see Sec. III for brief discussion), however the Kraus decomposition can sometimes be easier to work with and can provide different insights when discussing maximum fidelity $\mathcal{O}(\mathcal{N})$, but perhaps in other cases as well. In the present case, the standard Kraus operators (see Sec. III A for definition) simplifies the search for and broadens the types of channel inputs Λ_{RA} which achieve $\mathcal{O}$.

One way in which we have broadened the search for optimal inputs Λ_{RA} is to identify channels $\mathcal{N}$ for which Λ_{RA} can be chosen to be separable. This choice appears in two notable cases. The first case is when $\mathcal{N}$ is an extremal qubit channel. Here, separability of Λ_{RA} leads to a discontinuous jump in the minimal amount of entanglement $\mathcal{E}(\mathcal{N})$ generated to achieve maximum fidelity with a fully entangled state (see the discussion with Fig. 2). A second notable case where Λ_{RA} can be chosen to be separable is for noisy unital qubit channels where the input may be ordinarily chosen to be fully entangled [see the discussion containing Eq. (84)]. These findings motivate a characterization of channels $\mathcal{N}$ for which Λ_{RA} is possibly separable, i.e., $\mathcal{E}(\mathcal{N}) = 0$. One typically expects such channels not to be useful for sharing entanglement in the type of one-shot setting discussed in Sec. IV A. One example of such channels is in Lemma 3. The lemma extends to channels with Choi-Jamiołkowsi operator $J_{AB}^{\mathcal{N}}$ having a greater than $(d-1)^2$-fold degeneracy in their largest eigenvalue. The support of this largest eigenvalue subspace always has a product state (proof for this can be constructed using Prop 1.4 in [53]), and thus Λ_{RA} can be chosen to be a product state and $\mathcal{E}(\mathcal{N}) = 0$. On the other hand, we also find a channel whose Choi-Jamiołkowsi operator has a degeneracy in its largest eigenvalue, but the optimal input for the channel must be entangled.

Another way in which we have broadened the search for optimal inputs Λ_{RA} is to consider an extension of results found previously. For qubit channels, a fully entangled input was known to achieve $\mathcal{O}$ if and only if the channel was unital. In higher dimensions, we find this result no longer holds. We construct a unital qutrit channel for which the optimal input must be less than fully entangled. We also construct a qutrit channel that is not unital, but for which a fully entangled input is necessary to obtain the largest overlap.

Our second reformulation of $\mathcal{O}(\mathcal{N})$ in Theorem 1 notes that it equals the operator norm of the channel's Choi-Jamiołkowski operator, up to normalization. This observation can not only simplify discussions about $\mathcal{O}(\mathcal{N})$ (for instance, see the proof of Theorem 2), it also gives the operator norm of the Choi-Jamiołkowski operator a simple interpretation.

The single channel use setting discussed here can be extended by allowing the reference system and the channel output system to be processed using local operations and one-way or two-way classical communication, labeled 1-LOCC and 2-LOCC, respectively. Building on ideas in [32,54], it has been shown for qubit channels that a maximum fully entangled fraction allowing a single round of 2-LOCC, $\mathcal{O}'$, equals $\mathcal{O}$ [34]. Understanding $\mathcal{O}'$ in higher-dimensional channels while exploring optimal protocols and multiplicativity of $\mathcal{O}'$ may form an interesting direction of future work. Extending our work to a setting where the reference system also becomes noisy may be interesting. Prior discussions [56,57] on this setting connect with entanglement annihilating channels [55]. Another direction can come from extending results in Sec. V B where we show that that a set of qubit Pauli channels with $\mathcal{E}(\mathcal{N}) = 0$ also have no quantum capacity $\mathcal{Q}$. It could be interesting to study the relation of $\mathcal{O}$ and $\mathcal{E}$ to $\mathcal{Q}$.

Along the way to analyzing the maximum fidelity, we found it useful to study extremal qubit channels. These simple channels can be considered the most basic qubit channels. However, to our knowledge, noise parameters for these channels have not been adequately discussed. In Sec. III C, we show that the *pcubed* point of view allows one to identify noise parameters for this channel in such a way that the channel becomes demonstrably noisier as a parameter is varied monotonically. The hope is that such an identification makes this channel class a better understood and nontrivial testbed for ideas in quantum information science. We also flesh out two useful properties of general channels. First, in Sec. III A, we show the existence of a standard Kraus decomposition where the Kraus operators are orthogonal and their norm is ordered. Second, in Sec. III B, we show how the Choi-Jamiołkowski operator of a channel and its dual can always be taken to be complex conjugates of each other.

ACKNOWLEDGMENTS

V.S. thanks Felix Leditzky for helpful discussions, Sergey Filippov for bringing Refs. [56,57] to his attention, and Chloe Kim and Dina Abdelhadi for useful comments.

[1] J. M. Chow, Quantum intranet, IET Quantum Commun. **2**, 26 (2021).

[2] H. J. Kimble, The quantum internet, Nature (London) **453**, 1023 (2008).

[3] S. Wehner, D. Elkouss, and R. Hanson, Quantum internet: A vision for the road ahead, Science **362**, eaam9288 (2018).

[4] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Teleporting an Unknown Quantum State via dual Classical and Einstein-Podolsky-Rosen Channels, Phys. Rev. Lett. **70**, 1895 (1993).

[5] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Mixed-state entanglement and quantum error correction, Phys. Rev. A **54**, 3824 (1996).

[6] M. M. Wilde, *Quantum Information Theory*, 2nd ed. (Cambridge University Press, Cambridge, 2017).

[7] S. Lloyd, Capacity of the noisy quantum channel, Phys. Rev. A **55**, 1613 (1997).

[8] P. W. Shor, Quantum error correction, Nov. (2002), http://www.msri.org/workshops/203/schedules/1181.

[9] N. Cai, A. Winter, and R. W. Yeung, Quantum privacy and quantum wiretap channels, Probl. Inf. Transmiss. **40**, 318 (2004).

[10] I. Devetak, The private classical capacity and quantum capacity of a quantum channel, IEEE Trans. Inf. Theor. **51**, 44 (2005).

[11] I. Devetak and P. W. Shor, The capacity of a quantum channel for simultaneous transmission of classical and quantum information, Commun. Math. Phys. **256**, 287 (2005).

[12] G. Smith and J. Yard, Quantum communication with zero-capacity channels, Science **321**, 1812 (2008).

[13] T. Cubitt, D. Elkouss, W. Matthews, M. Ozols, D. Pérez-García, and S. Strelchuk, Unbounded number of channel uses may be required to detect quantum capacity, Nat. Commun. **6**, 6739 (2015).

[14] V. Siddhu, Entropic singularities give rise to quantum transmission, Nat. Commun. **12**, 5750 (2021).

[15] F. Leditzky, D. Leung, V. Siddhu, G. Smith, and J. A. Smolin, Generic Nonadditivity of Quantum Capacity in Simple Channels, Phys. Rev. Lett. **130**, 200801 (2023).

[16] F. Buscemi and N. Datta, The quantum capacity of channels with arbitrarily correlated noise, IEEE Trans. Inf. Theor. **56**, 1447 (2010).

[17] N. Datta and M.-H. Hsieh, The apex of the family tree of protocols: optimal rates and resource inequalities, New J. Phys. **13**, 093042 (2011).

[18] N. Datta and M.-H. Hsieh, One-shot entanglement-assisted quantum and classical communication, IEEE Trans. Inf. Theor. **59**, 1929 (2013).

[19] W. Matthews and S. Wehner, Finite blocklength converse bounds for quantum channels, IEEE Trans. Inf. Theor. **60**, 7317 (2014).

[20] S. Beigi, N. Datta, and F. Leditzky, Decoding quantum information via the Petz recovery map, J. Math. Phys. **57**, 082203 (2016).

[21] M. Tomamichel, M. Berta, and J. M. Renes, Quantum coding with finite resources, Nat. Commun. **7**, 11419 (2016).

[22] C. Pfister, M. Adriaan Rol, A. Mantri, M. Tomamichel, and S. Wehner, Capacity estimation and verification of quantum

channels with arbitrarily correlated errors, Nat. Commun. **9**, 27 (2018).

[23] A. Anshu, R. Jain, and N. A. Warsi, Building blocks for communication over noisy quantum networks, IEEE Trans. Inf. Theor. **65**, 1287 (2019).

[24] X. Wang, K. Fang, and R. Duan, Semidefinite programming converse bounds for quantum communication, IEEE Trans. Inf. Theor. **65**, 2583 (2018).

[25] F. Salek, A. Anshu, M.-H. Hsieh, R. Jain, and J. R. Fonollosa, One-shot capacity bounds on the simultaneous transmission of classical and quantum information, IEEE Trans. Inf. Theor. **66**, 2141 (2020).

[26] Y. Nakata, E. Wakakuwa, and H. Yamasaki, One-shot quantum error correction of classical and quantum information, Phys. Rev. A **104**, 012408 (2021).

[27] S. Khatri and M. M. Wilde, Principles of quantum communication theory: A modern approach (2020), arXiv:2011.04672.

[28] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels, Phys. Rev. Lett. **76**, 722 (1996).

[29] M. Horodecki, P. Horodecki, and R. Horodecki, General teleportation channel, singlet fraction, and quasidistillation, Phys. Rev. A **60**, 1888 (1999).

[30] F. Verstraete and H. Verschelde, On quantum channels, arXiv:quant-ph/0202124.

[31] R. Pal and S. Bandyopadhyay, Entanglement sharing via qudit channels Nonmaximally entangled states may be necessary for one-shot optimal singlet fraction and negativity, Phys. Rev. A **97**, 032322 (2018).

[32] F. Verstraete and H. Verschelde, Optimal Teleportation with a Mixed State of Two Qubits, Phys. Rev. Lett. **90**, 097901 (2003).

[33] S. Bandyopadhyay and A. Ghosh, Optimal fidelity for a quantum channel may be attained by nonmaximally entangled states, Phys. Rev. A **86**, 020304(R) (2012).

[34] R. Pal, S. Bandyopadhyay, and S. Ghosh, Entanglement sharing through noisy qubit channels: One-shot optimal singlet fraction, Phys. Rev. A **90**, 052304 (2014).

[35] L. Vandenberghe and S. Boyd, Semidefinite programming, SIAM Rev. **38**, 49 (1996).

[36] A. Jamiołkowski, Linear transformations which preserve trace and positive semidefiniteness of operators, Rep. Math. Phys. **3**, 275 (1972).

[37] M.-D. Choi, Completely positive linear maps on complex matrices, Lin. Alg. Appl. **10**, 285 (1975).

[38] A. S. Holevo, *Quantum Systems, Channels, Information: A Mathematical Introduction* (De Gruyter, Berlin, Boston, 2013), Vol. 16.

[39] M. M. Wolf, *Quantum Channels and Operations–Guided Tour*, Lecture Notes, 2012, https://www.math.cit.tum.de/math/forschung/gebiete/mathematical-physics/.

[40] A. Fujiwara and P. Algoet, Affine parameterization of quantum channels, in *Proceedings of the 1998 IEEE International Symposium on Information Theory (Cat. No. 98CH36252), Cambridge, 1998* (IEEE, Piscataway, NJ, 1998), p. 87.

[41] C.-S. Niu and R. B. Griffiths, Two-qubit copying machine for economical quantum eavesdropping, Phys. Rev. A **60**, 2764 (1999).

[42] M. Beth Ruskai, S. Szarek, and E. Werner, An analysis of completely-positive trace-preserving maps on m2, Lin. Alg. Appl. **347**, 159 (2002).

[43] M. M. Wolf and D. Pérez-García, Quantum capacities of channels with small environment, Phys. Rev. A **75**, 012303 (2007).

[44] S. Friedland and R. Loewy, On the extreme points of quantum channels, arXiv:1309.5898.

[45] V. Siddhu and R. B. Griffiths, Degradable quantum channels using pure-state to product-of-pure-state isometries, Phys. Rev. A **94**, 052331 (2016).

[46] J. L. Park, The concept of transition in quantum mechanics, Found. Phys. **1**, 23 (1970).

[47] D. Dieks, Communication by epr devices, Phys. Lett. A **92**, 271 (1982).

[48] W. K. Wootters and W. H. Zurek, A single quantum cannot be cloned, Nature (London) **299**, 802 (1982).

[49] J. Watrous, *The Theory of Quantum Information*, 1st ed. (Cambridge University Press, Cambridge, 2018).

[50] C.-S. Niu and R. B. Griffiths, Optimal copying of one quantum bit, Phys. Rev. A **58**, 4377 (1998).

[51] N. J. Cerf, Pauli Cloning of a Quantum Bit, Phys. Rev. Lett. **84**, 4497 (2000).

[52] T. S. Cubitt, M. B. Ruskai, and G. Smith, The structure of degradable quantum channels, J. Math. Phys. **49**, 102104 (2008).

[53] K. R. Parthasarathy, On the maximal dimension of a completely entangled subspace for finite level quantum systems, Proc. Math. Sci. **114**, 365 (2004).

[54] F. Verstraete and H. Verschelde, Fidelity of mixed states of two qubits, Phys. Rev. A **66**, 022307 (2002).

[55] L. Moravčíková and M. Ziman, Entanglement-annihilating and entanglement-breaking channels, J. Phys. A **43**, 275306 (2010).

[56] S. N. Filippov, T. Rybár, and M. Ziman, Local two-qubit entanglement-annihilating channels, Phys. Rev. A **85**, 012303 (2012).

[57] S. N. Filippov, V. V. Frizen, and D. V. Kolobova, Ultimate entanglement robustness of two-qubit states against general local noises, Phys. Rev. A **97**, 012322 (2018).