

The Platypus of the Quantum Channel Zoo

Felix Leditzky¹, Debbie Leung, Vikesh Siddhu², Graeme Smith, and John A. Smolin

Abstract—Understanding quantum channels and the strange behavior of their capacities is a key objective of quantum information theory. Here we study a remarkably simple, low-dimensional, single-parameter family of quantum channels with exotic quantum information-theoretic features. As the simplest example from this family, we focus on a qutrit-to-qutrit channel that is intuitively obtained by hybridizing together a simple degradable channel and a completely useless qubit channel. Such hybridizing makes this channel’s capacities behave in a variety of interesting ways. For instance, the private and classical capacity of this channel coincide and can be explicitly calculated, even though the channel does not belong to any class for which the underlying information quantities are known to be additive. Moreover, the quantum capacity of the channel can be computed explicitly, given a clear and compelling conjecture is true. This “spin alignment conjecture,” which may be of independent interest, is proved in certain special cases and additional numerical evidence for its validity is provided. Finally, we generalize the qutrit channel in two ways, and the resulting channels and their capacities display similarly rich behavior. In a companion paper, we further show that the qutrit channel demonstrates superadditivity when transmitting quantum information jointly with a variety of assisting channels, in a manner unknown before.

Index Terms—Quantum communication, quantum channels, quantum information science.

Manuscript received 11 May 2022; revised 23 November 2022; accepted 3 February 2023. Date of publication 15 February 2023; date of current version 19 May 2023. This work was supported in part by Army Research Office Multidisciplinary University Research Initiative (ARO MURI) Quantum Network Science under Contract W911NF2120214; in part by NSF under Grant CCF 1652560, Grant PHY 1915407, and Grant 2137953; and in part by Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant. An earlier version of this paper was presented at the 2022 IEEE International Symposium on Information Theory, Aalto University [DOI: 10.1109/ISIT50566.2022.9834678]. (Corresponding author: Felix Leditzky.)

Felix Leditzky was with the Institute for Quantum Computing and the Department of Combinatorics and Optimization, University of Waterloo, Waterloo, ON N2L 3G1, Canada, and also with the Perimeter Institute for Theoretical Physics, Waterloo, ON N2L 2Y5, Canada. He is now with the Department of Mathematics and the Illinois Quantum Information Science and Technology Center (IQUIST), University of Illinois at Urbana–Champaign, Champaign, IL 61820 USA (e-mail: leditzky@illinois.edu).

Debbie Leung is with the Institute for Quantum Computing and the Department of Combinatorics and Optimization, University of Waterloo, Waterloo, ON N2L 3G1, Canada, and also with the Perimeter Institute for Theoretical Physics, Waterloo, ON N2L 2Y5, Canada (e-mail: wcleung@uwaterloo.ca).

Vikesh Siddhu was with the JILA, University of Colorado Boulder, Boulder, CO 80309 USA, and also with the Department of Physics and the Quantum Computing Group, Carnegie Mellon University, Pittsburgh, PA 15213 USA. He is now with the IBM Quantum, IBM T. J. Watson Research Center, Yorktown Heights, NY 10598 USA (e-mail: vsiddhu@protonmail.com).

Graeme Smith is with the JILA, CTQM, and the Department of Physics, University of Colorado Boulder, Boulder, CO 80309 USA (e-mail: graeme.smith@colorado.edu).

John A. Smolin is with the IBM Quantum, IBM T. J. Watson Research Center, Yorktown Heights, NY 10598 USA (e-mail: smolin@us.ibm.com).

Communicated by M. Tomamichel, Associate Editor for Quantum Information Theory.

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TIT.2023.3245985>.

Digital Object Identifier 10.1109/TIT.2023.3245985

I. INTRODUCTION

QUANTUM channels model noisy communication links between quantum parties. The channel noise affecting signals can be mitigated by encoding the messages across many channel uses. The highest rate at which information can be sent reliably is known as a capacity of the channel. Depending on the type of information to be transmitted, we obtain different capacity quantities; for example, a quantum channel may be used to transmit classical, quantum, or private classical information. The capacities in each case are the classical capacity \mathcal{C} , the quantum capacity \mathcal{Q} , and the private classical capacity \mathcal{P} , measured in bits per channel use, qubits per channel use, and private bits per channel use, respectively. The various capacities of a quantum channel quantify its usefulness in the respective communication setting.

There is a variety of synergy effects that may occur during a quantum channel’s transmission of the various types of (quantum, private, classical) information. These include super-additivity of coherent information [2], [3], private information [4], Holevo information [5], superactivation of quantum capacity [6], and private communication at a rate above the quantum capacity [7], [8], [9] (see Section II for details). These nonadditive effects enable exciting and novel communication protocols, but at the same time they obscure a succinct mathematical characterization of the corresponding quantum channel capacities.

Remarkably, such nonadditivities appear to be common; some are exhibited even for simple channels such as the depolarizing channel. On the other hand, for certain classes of quantum channels the information quantities listed above can be additive, thus simplifying their information-theoretical characterization. Unfortunately this is only known to be true in a few special cases. For example, the coherent information of a PPT channel is additive and indeed zero [7]. The only other channel classes with known additive coherent information are the degradable [10] and antidegradable channels, which as a result have a so-called single-letter formula for their quantum capacity. These channels also have the pleasing property of having their private capacity equal to the quantum capacity [11]; as a result, also the private capacity is given by a single-letter formula for these channels. There exists a smattering of channels whose Holevo information is additive, and therefore these channels have a single-letter formula for their classical capacity [12], [13], [14], [15]. However, beyond special examples of proven nonadditivities and additivities, little is known about most capacities of most channels.

The best path towards a deeper understanding of non-additivity effects in quantum information—in fact, a better

understanding of quantum information itself—is to better understand and develop the menagerie of these phenomena. However, clean and clear examples of channels that isolate different aspects of nonadditivity are in short supply. As a result, over the past two decades significant effort has been dedicated to elucidating these phenomena, leading to numerous exciting findings; yet a full understanding still remains elusive. Without such understanding, we lack a theory on how to best communicate with quantum channels, and fail to answer the kinds of questions resolved in classical information theory. This is substantiated by the fact that random codes can be suboptimal; that we cannot evaluate capacities beyond special examples; that the known capacity quantities may not fully capture the communication potential of a noisy channel; and that our understanding of error correction in the quantum setting is incomplete, whether the data is classical, private, or quantum.

This paper studies novel examples of additivity based on combining two well-understood but very different classes of channels. In the simplest case, we combine two qubit-channels in a qutrit channel so that their inputs overlap along one dimension. The superposed identities of the subchannels used in the transmission convey additional information to the receiver and thus becomes an integral part of the communication sent.

Our channel is constructed as follows. We start with a degradable channel, which we know has single-letter quantum and private capacities. We then explicitly break degradability by adding an extra input state that lets the sender transmit quantum information directly to the environment but allows no additional information to be sent to the output. This makes it impossible for the channel output to simulate the environment as required for degradability, and thus takes the channel outside any class known to have additive private or coherent information. Nevertheless, the coherent information appears to remain additive, which we can show up to a very reasonable conjecture. Even more surprisingly, the private information remains additive but takes a *much* larger value than the original channel. This difference in private and quantum capacity is a clear signature of the nondegradability of the channel. The apparent additivity of coherent information must therefore be coming from some new mechanism of additivity, which we seek to understand.

A. Main Results

We study a remarkably simple, single-parameter family of qutrit channels \mathcal{N}_s , along with generalizations \mathcal{M}_d and \mathcal{O} of this channel family to arbitrary dimension d . These channel families exhibit many strange behaviors for quantum communication while having uncomplicated classical and private classical capacities: The classical and private classical capacities can be calculated explicitly because the underlying information quantities (Holevo and private information, respectively) are additive, even though neither channel family belongs to any of the known additivity classes. The same holds true for the quantum capacity: the coherent informations of these families are additive, provided that a certain entropy minimization conjecture is true. We give evidence for the validity of this “spin alignment conjecture” in the main text.

Despite the additivity of the capacities, both channel families have strictly larger private capacity than quantum capacity. The simplicity of our channel families enables us to further generalize the study to a 3-parameter family of channels displaying rich behavior, including separations between all three capacities, which we analyze numerically.

The results in the current paper are even more surprising in the light of additional findings in a companion paper [1]: The coherent information of the channel \mathcal{N}_s tensored with an assisting channel is super-additive, for a large swath of values of s and for some generically chosen assisting channel. The super-additivity can be lifted to quantum capacity for degradable assisting channels if the spin alignment conjecture holds. The assisting channel can have positive or vanishing quantum capacity. The mechanism behind this superadditivity is novel and in particular differs from the known explanation of super-activation [6], [16]. For the d -dimensional generalization \mathcal{M}_d of $\mathcal{N}_{1/2}$, we focus on super-additivity with a $(d-1)$ -dimensional erasure channel and recover results similar to those for \mathcal{N}_s . In addition, even stronger qualitative results can be obtained. First, super-additivity of quantum capacity can be proved unconditionally (without the spin alignment conjecture). Second, the effect holds for all nontrivial values of the erasure probability, for an appropriately large local dimension d .

B. Structure of the Paper

This paper is structured as follows. In Section II we give some general background on quantum channels, their various capacities, and special classes of channels. The main objects of this paper, the channel \mathcal{N}_s and its d -dimensional generalizations \mathcal{M}_d and \mathcal{O} , are defined in Section III. The channel coherent information and the quantum capacity of these channels are discussed in Sections IV and V, respectively. Section VI formulates the spin alignment conjecture relative to which the coherent information of \mathcal{N}_s is additive. We then derive bounds on the quantum capacity of \mathcal{N}_s and \mathcal{M}_d in Section VII, and determine their private and classical capacities exactly in Section VIII. Section IX summarizes and further discusses the results on the capacities of \mathcal{N}_s . Finally, we present the 3-parameter generalization of \mathcal{N}_s and a numerical analysis of its capacities in Section X. MATLAB and Python code used to obtain the numerical results mentioned above will be made available at [17].

II. PRELIMINARIES

A. Quantum Channels

Let \mathcal{H} be a Hilbert space of finite dimension d . Let \mathcal{H}^\dagger be the dual of \mathcal{H} , and $\mathcal{H} \cong \mathcal{H} \otimes \mathcal{H}^\dagger$ be the space of linear operators acting on \mathcal{H} . Let $\mathcal{H}_a, \mathcal{H}_b$, and \mathcal{H}_c be three Hilbert spaces of dimensions d_a, d_b , and d_c respectively. An isometry $E : \mathcal{H}_a \mapsto \mathcal{H}_b \otimes \mathcal{H}_c$, i.e., a map satisfying $E^\dagger E = I_a$ (the identity on \mathcal{H}_a), takes an input Hilbert space \mathcal{H}_a to a subspace of a pair of output spaces $\mathcal{H}_b \otimes \mathcal{H}_c$. This isometry generates a quantum channel pair, $(\mathcal{B}, \mathcal{B}^c)$, i.e., a pair of completely positive trace preserving (CPTP) maps, with superoperators,

$$\mathcal{B}(X) = \text{Tr}_c(EXE^\dagger), \quad \text{and} \quad \mathcal{B}^c(X) = \text{Tr}_b(EXE^\dagger), \quad (1)$$

that take any element $X \in \hat{\mathcal{H}}_a$ to $\hat{\mathcal{H}}_b$ and $\hat{\mathcal{H}}_c$, respectively. Each channel in this pair $(\mathcal{B}, \mathcal{B}^c)$ may be called the complement of the other. The isometry E can be written as $E = \sum_i K_i \otimes |i\rangle_c = \sum_j L_j \otimes |j\rangle_b$ where $K_i : \mathcal{H}_a \mapsto \mathcal{H}_b$ and $L_j : \mathcal{H}_a \mapsto \mathcal{H}_c$ are Kraus operators for the \mathcal{B} and \mathcal{B}^c channels respectively, satisfying $\mathcal{B}(X_a) = \sum_{i=1}^{d_c} K_i X_a K_i^\dagger$ and $\sum_{i=1}^{d_c} K_i^\dagger K_i = I_a$, and similarly for $\{L_j\}_{j=1}^{d_b}$.

If the input of the isometry E is restricted to a subspace $\mathcal{H}_{\bar{a}}$ of \mathcal{H}_a , then such a restricted map is still an isometry on $\mathcal{H}_{\bar{a}}$ and defines a pair of channels $(\bar{\mathcal{B}}, \bar{\mathcal{B}}^c)$, where each channel $\bar{\mathcal{B}}$ and $\bar{\mathcal{B}}^c$ is called a *sub-channel* of \mathcal{B} and \mathcal{B}^c , respectively. When focussing on some quantum channel \mathcal{B} , it is common to denote $\mathcal{H}_a, \mathcal{H}_b$, and \mathcal{H}_c as the channel input, output, and environment respectively. Any CPTP map (together with its complement) may be written as (1) in terms of a suitable isometry E . Another representation of a CPTP map comes from its Choi-Jamiołkowski operator. To define this operator, consider a linear map $\mathcal{B} : \hat{\mathcal{H}}_a \mapsto \hat{\mathcal{H}}_b$ and a maximally entangled state,

$$|\phi\rangle = \frac{1}{\sqrt{d_a}} \sum_i |i\rangle_a \otimes |i\rangle_a, \quad (2)$$

on $\mathcal{H}_a \otimes \mathcal{H}_a$. The *unnormalized* Choi-Jamiołkowski operator of \mathcal{B} is

$$J_{ab}^{\mathcal{B}} = d_a (\mathcal{I}_a \otimes \mathcal{B})(|\phi\rangle\langle\phi|), \quad (3)$$

where \mathcal{I}_a denotes the identity map acting on $\hat{\mathcal{H}}_a$. The linear map \mathcal{B} is CPTP if and only if the above operator is positive semidefinite and its partial trace over \mathcal{H}_b is the identity I_a on \mathcal{H}_a .

B. Quantum Capacity

The quantum capacity $\mathcal{Q}(\mathcal{B})$ of a quantum channel $\mathcal{B} : \hat{\mathcal{H}}_a \rightarrow \hat{\mathcal{H}}_b$ is defined as the largest rate at which quantum information can be sent faithfully through the channel. It can be expressed in terms of an entropic quantity as follows. Let ρ_a denote a density operator (unit trace positive semi-definite operator) on \mathcal{H}_a and for any ρ_a let $\rho_b := \mathcal{B}(\rho_a)$ and $\rho_c := \mathcal{B}^c(\rho_a)$. The coherent information (or entropy bias) of a channel \mathcal{B} at a density operator ρ_a is

$$\Delta(\mathcal{B}, \rho_a) = S(\rho_b) - S(\rho_c), \quad (4)$$

where $S(\rho) = -\text{Tr}(\rho \log \rho)$ (we use \log base 2 by default) is the von-Neumann entropy of ρ . The channel coherent information (sometimes called the *single-letter* coherent information),

$$\mathcal{Q}^{(1)}(\mathcal{B}) = \max_{\rho_a} \Delta(\mathcal{B}, \rho_a), \quad (5)$$

represents an achievable rate for sending quantum information across the channel \mathcal{B} , and hence $\mathcal{Q}(\mathcal{B}) \geq \mathcal{Q}^{(1)}(\mathcal{B})$ [18], [19], [20]. The maximum achievable rate is equal to the quantum capacity of \mathcal{B} , and given by a *multi-letter* formula (sometimes called a regularized expression) [18], [19], [20], [21], [22],

$$\mathcal{Q}(\mathcal{B}) = \sup_{n \in \mathbb{N}} \frac{1}{n} \mathcal{Q}^{(1)}(\mathcal{B}^{\otimes n}), \quad (6)$$

where $\mathcal{B}^{\otimes n}$ represent $n \in \mathbb{N}$ parallel (sometimes called joint) uses of \mathcal{B} . The regularization in (6) is necessary because the

channel coherent information is *super-additive*, i.e., for any two quantum channels \mathcal{B} and \mathcal{B}' used together, the channel coherent information of the joint channel $\mathcal{B} \otimes \mathcal{B}'$ satisfies an inequality,

$$\mathcal{Q}^{(1)}(\mathcal{B} \otimes \mathcal{B}') \geq \mathcal{Q}^{(1)}(\mathcal{B}) + \mathcal{Q}^{(1)}(\mathcal{B}'), \quad (7)$$

which can be strict [3], [6], [23], [24], [25], [26], [27], [28], [29], [30]. The coherent information $\mathcal{Q}^{(1)}(\mathcal{B})$ is said to be *weakly additive* if equality holds in (7) whenever \mathcal{B}' is a tensor power of \mathcal{B} . If this equality holds for arbitrary \mathcal{B}' then $\mathcal{Q}^{(1)}(\mathcal{B})$ is said to be *strongly additive*.

C. Private Capacity

The private capacity $\mathcal{P}(\mathcal{B})$ of a quantum channel $\mathcal{B} : \hat{\mathcal{H}}_a \rightarrow \hat{\mathcal{H}}_b$ is operationally defined as the largest rate at which classical information can be faithfully sent through the channel in such a way that the environment, \mathcal{H}_c , gains no meaningful knowledge about the information being sent. A formula for the private capacity was derived by Cai et al. [20], [31] in terms of a quantity called the channel private information:

$$\mathcal{P}^{(1)}(\mathcal{B}) = \max_{\{p_x, \rho_a^x\}} \left[\Delta(\mathcal{B}, \bar{\rho}_a) - \sum_x p_x \Delta(\mathcal{B}, \rho_a^x) \right]. \quad (8)$$

Here, $\Delta(\mathcal{B}, \rho)$ denotes the coherent information of the channel \mathcal{B} with respect to the state ρ as defined in (4), the maximization is over all quantum state ensembles $\{p_x, \rho_a^x\}_x$, and $\bar{\rho}_a = \sum_x p_x \rho_a^x$ denotes the ensemble average of the density operators $\{\rho_a^x\}$ over the probability distribution $\{p_x\}$ ($p_x > 0$ and $\sum_x p_x = 1$), Restricting the maximization in (8) to an ensemble of pure states ρ_a^x makes $\Delta(\mathcal{B}, \rho_a^x) = 0$, reducing (8) to (5), and resulting in a maximum value which simply equals the channel coherent information $\mathcal{Q}^{(1)}(\mathcal{B})$. This value is at most $\mathcal{P}^{(1)}(\mathcal{B})$, i.e.,

$$\mathcal{Q}^{(1)}(\mathcal{B}) \leq \mathcal{P}^{(1)}(\mathcal{B}). \quad (9)$$

There are channels for which this inequality is strict, thus for such channels $\mathcal{P}^{(1)}(\mathcal{B})$ cannot be obtained using an ensemble of pure states alone.

A channel's private information can also be written as

$$\mathcal{P}^{(1)}(\mathcal{B}) = \max_{\{p_x, \rho_a^x\}} [I(x; b)_\sigma - I(x; c)_\sigma], \quad (10)$$

where the mutual information at a quantum state σ , $I(a; b)_\sigma := S(\sigma_a) + S(\sigma_b) - S(\sigma_{ab})$, is evaluated above at

$$\sigma_{xbe} = (\mathbb{1}_x \otimes J) \left(\sum_x p_x [x]_x \otimes \rho_a^x \right) (\mathbb{1}_x \otimes J)^\dagger, \quad (11)$$

with the label x denoting a classical register, J denoting the \mathcal{B} channel isometry (1), and the notation

$$[\psi] = |\psi\rangle\langle\psi|, \quad (12)$$

represents a dyad onto any ket $|\psi\rangle$. Cai et al. [31] and Devetak [20] proved that the private information $\mathcal{P}^{(1)}(\mathcal{B})$ is an achievable rate for private information transmission, $\mathcal{P}(\mathcal{B}) \geq \mathcal{P}^{(1)}(\mathcal{B})$, and furthermore the private capacity is bounded from above by the regularized private information.

As a result, we have the following coding theorem for the private capacity [20], [31]:

$$\mathcal{P}(\mathcal{B}) = \sup_{n \in \mathbb{N}} \frac{1}{n} \mathcal{P}^{(1)}(\mathcal{B}^{\otimes n}). \quad (13)$$

Except for a few special cases such as degradable channels [10], [11] (for definition see Sec. II-F) the regularization in (13) is required. This requirement arises from super-additivity of $\mathcal{P}^{(1)}$ [4], i.e., $\mathcal{P}^{(1)}$ satisfies an inequality of the form in (7). From (9) it follows that each term, $\mathcal{Q}^{(1)}(\mathcal{B}^{\otimes n})$, in the limit (6) is at most $\mathcal{P}^{(1)}(\mathcal{B}^{\otimes n})$, and thus a channel's quantum capacity $\mathcal{Q}(\mathcal{B})$ is at most its private classical capacity $\mathcal{P}(\mathcal{B})$ (13),

$$\mathcal{Q}(\mathcal{B}) \leq \mathcal{P}(\mathcal{B}). \quad (14)$$

D. Classical Capacity

The classical capacity $\mathcal{C}(\mathcal{B})$ of a quantum channel $\mathcal{B}: \mathcal{H}_a \rightarrow \mathcal{H}_b$ is defined as the largest rate at which classical information can be faithfully sent through the channel. In contrast to private information transmission discussed above, there is no security criterion involving the environment, and hence

$$\mathcal{P}(\mathcal{B}) \leq \mathcal{C}(\mathcal{B}). \quad (15)$$

The classical capacity can be expressed in terms of an information quantity called the Holevo information $\chi(\mathcal{B})$:

$$\begin{aligned} \chi(\mathcal{B}) &= \max_{\{p_x, \rho_a^x\}} \left[S(\bar{\rho}_b) - \sum_x p_x S(\rho_b^x) \right] \\ &= \max_{\{p_x, \rho_a^x\}} I(x; b)_\sigma, \end{aligned} \quad (16)$$

where $\bar{\rho}_a = \sum_x p_x \rho_a^x$ and σ_{xbe} is defined in terms of the quantum state ensemble $\{p_x, \rho_a^x\}_x$ as in (11) above. Quantum states in the maximization (16) of $\chi(\mathcal{B})$ can be chosen to be pure [32]. The classical capacity $\mathcal{C}(\mathcal{B})$ of a quantum channel \mathcal{B} can be expressed as [33] and [34],

$$\mathcal{C}(\mathcal{B}) = \sup_{n \in \mathbb{N}} \frac{1}{n} \chi(\mathcal{B}^{\otimes n}). \quad (17)$$

Again, apart from a few special classes of channels the regularization of the Holevo information in (17) is necessary due to super-additivity of $\chi(\mathcal{B})$ [5], [35]. Comparing (16) with (10) reveals $\mathcal{P}^{(1)}(\mathcal{B}) \leq \chi^{(1)}(\mathcal{P})$, as a result, each term, $\mathcal{P}^{(1)}(\mathcal{B}^{\otimes n})$, in the limit (13) is at most $\chi^{(1)}(\mathcal{P}^{\otimes n})$, and thus a channel's private channel capacity $\mathcal{P}(\mathcal{B})$ is at most its classical capacity $\mathcal{C}(\mathcal{B})$ (15).

E. Entanglement-Assisted Classical Capacity

Finally, we discuss the entanglement-assisted classical capacity $\mathcal{C}_E(\mathcal{B})$ of a quantum channel \mathcal{B} , which is defined as the optimal rate of faithful classical information transmission when the sender and receiver have access to unlimited entanglement assisting the encoding and decoding process. It can be expressed in terms of the channel mutual information $I(\mathcal{B})$, defined as

$$\begin{aligned} I(\mathcal{B}) &= \max_{\rho_a} I(a'; b)_\sigma \\ &= \max_{\rho_a} [S(\sigma_{a'}) + S(\sigma_b) - S(\sigma_{a'b})], \end{aligned} \quad (18)$$

with $\sigma_{a'b} = (\mathcal{I}_{a'} \otimes \mathcal{B})(\psi_{a'a})$ and $|\psi\rangle_{a'a}$ an (arbitrary) purification of the input state ρ_a . Note that $I(a'; b)$ is concave in the input state ρ_a and can therefore be computed efficiently [32], [36]. The entanglement-assisted classical capacity of a quantum channel is equal to its channel mutual information [37]:

$$\mathcal{C}_E(\mathcal{B}) = I(\mathcal{B}), \quad (19)$$

and is an upper bound of $\mathcal{C}(\mathcal{B})$ by definition.

F. Special Channel Classes

If a channel maps the identity element at its input to the identity at its output, then the channel is called a *unital channel*. A channel \mathcal{B} is called *degradable*, and its complement \mathcal{B}^c *anti-degradable*, if there is another channel \mathcal{D} such that $\mathcal{D} \circ \mathcal{B} = \mathcal{B}^c$ [10], [38]. Sometimes this channel \mathcal{D} is called the *degrading map* of the degradable channel \mathcal{B} . For any two channels \mathcal{B}' and \mathcal{B} , each either degradable or anti-degradable, the joint channel $\mathcal{B} \otimes \mathcal{B}'$ has additive coherent information, i.e., equality holds in (7) [10], [39]. For a degradable channel \mathcal{B} , the coherent information $\Delta(\mathcal{B}, \rho_a)$ is concave in ρ_a [40], and thus $\mathcal{Q}^{(1)}(\mathcal{B})$ can be computed with relative ease [36], [41]. As a result the quantum capacity of a degradable channel, which simply equals $\mathcal{Q}^{(1)}(\mathcal{B})$, can also be computed efficiently. An anti-degradable channel has no quantum capacity due to the no-cloning theorem. An instance of an anti-degradable channel is a *measure-and-prepare* or *entanglement-breaking* (EB) channel [42]. An EB channel is one whose Choi-Jamiołkowski operator (3) is separable [42]. The complement of an EB channel is called a *Hadamard channel* [32]. Besides anti-degradable channels, the only other known class of zero-quantum-capacity channels are *entanglement binding* or *positive under partial-transpose* (PPT) channels [7]. A channel is PPT if its Choi-Jamiołkowski operator (3) is positive under partial transpose. If a channel \mathcal{B}^c has zero quantum capacity then its complement \mathcal{B} is called a *more capable* channel. A more capable channel has equal quantum and private capacity. If a more capable channel \mathcal{B} has a complement \mathcal{B}^c with zero private capacity, then \mathcal{B} has additive coherent information in the sense that equality holds in eq. (7) with $\mathcal{B}' = \mathcal{B}$ [43].

III. DEFINITION OF CHANNELS

A. The \mathcal{N}_s Channel

Let $\mathcal{H}_a, \mathcal{H}_b$, and \mathcal{H}_c have dimensions $d_a = d_b = 3$, and $d_c = 2$. Consider an isometry $F_s: \mathcal{H}_a \mapsto \mathcal{H}_b \otimes \mathcal{H}_c$ of the form,

$$\begin{aligned} F_s|0\rangle &= \sqrt{s}|0\rangle \otimes |0\rangle + \sqrt{1-s}|1\rangle \otimes |1\rangle, \\ F_s|1\rangle &= |2\rangle \otimes |0\rangle, \\ F_s|2\rangle &= |2\rangle \otimes |1\rangle, \end{aligned} \quad (20)$$

where $0 \leq s \leq 1/2$. This isometry was introduced previously by one of us in [30] with $|1\rangle$ and $|2\rangle$ in \mathcal{H}_a exchanged. Its Kraus operators are unitarily equivalent to those of a channel \mathcal{L}_a introduced even earlier in [44] and studied further in [45]. The isometry can be written as $F_s = \sum_i K_i \otimes |i\rangle$ where Kraus

operators $K_i : \mathcal{H}_a \mapsto \mathcal{H}_b$ match those in Sec. IV-C of [45] if one permutes the computational basis of \mathcal{H}_a as $|0\rangle_a \mapsto |2\rangle_a \mapsto |1\rangle_a \mapsto |0\rangle_a$, exchanges $|1\rangle_b$ and $|2\rangle_b$, and rewrites s as $\sin^2 \alpha$, $0 \leq \alpha \leq \pi/4$. Through an equation of the form (1) the isometry (20) gives rise to a complementary pair of channels $\mathcal{N}_s : \hat{\mathcal{H}}_a \mapsto \hat{\mathcal{H}}_b$ and $\mathcal{N}_s^c : \hat{\mathcal{H}}_a \mapsto \hat{\mathcal{H}}_c$. This channel pair has two simple properties.

The first property, proved in Sec. III-D, is the existence of degradable sub-channels of \mathcal{N}_s^c , obtained by restricting the channel input to operators on a qubit subspace,

$$\mathcal{H}_{ai} = \text{span}\{|0\rangle, |i\rangle\}, \quad (21)$$

where i is either 1 or 2. This restriction also results in an anti-degradable sub-channel of \mathcal{N}_s^c . Quantum states lying solely in this qubit input sub-space can be used to send an equal amount of quantum and private information to the \mathcal{N}_s output \mathcal{H}_b but such states cannot be used to send any quantum or private information to the \mathcal{N}_s^c output \mathcal{H}_c .

The second channel property, also proved in Sec. III-D, is the presence of a perfect sub-channel of \mathcal{N}_s^c obtained by restricting the channel input to operators on a qubit subspace $\mathcal{H}_{a'}$ spanned by $\{|1\rangle, |2\rangle\}$. Quantum states lying solely in this qubit input subspace $\mathcal{H}_{a'}$ can be used to perfectly send information to the \mathcal{N}_s^c output \mathcal{H}_c while sending no information to the \mathcal{N}_s output \mathcal{H}_b . This perfect transmission of a qubit to \mathcal{H}_c implies that the quantum, private, and classical capacities of \mathcal{N}_s^c are at least 1. Since the dimension of the channel output \mathcal{H}_c is two, all these capacities of \mathcal{N}_s^c are at most 1, thus

$$\mathcal{Q}(\mathcal{N}_s^c) = \mathcal{P}(\mathcal{N}_s^c) = \mathcal{C}(\mathcal{N}_s^c) = 1. \quad (22)$$

Together, the intuitive picture above along with the two simple properties of \mathcal{N}_s and \mathcal{N}_s^c help classify each of these channels. For instance, one can easily infer that each channel in the $(\mathcal{N}_s, \mathcal{N}_s^c)$ pair is neither degradable nor anti-degradable. If \mathcal{N}_s was degradable, then all its sub-channels would also be degradable. However the sub-channel of \mathcal{N}_s obtained by restricting its input to $\mathcal{H}_{a'}$ is anti-degradable. This anti-degradable sub-channel is the complement of a perfect (and hence degradable) sub-channel to \mathcal{H}_c obtained by restricting the channel input to $\mathcal{H}_{a'}$. In similar vein, \mathcal{N}_s^c is not degradable. Consider a sub-channel of \mathcal{N}_s^c obtained by restricting its input to \mathcal{H}_{ai} . This sub-channel is anti-degradable since its complement, a sub-channel of \mathcal{N}_s with input \mathcal{H}_{ai} , is degradable.

Since each channel in the $(\mathcal{N}_s, \mathcal{N}_s^c)$ pair is neither degradable nor anti-degradable, both channels are not EB since EB channels are anti-degradable. A Hadamard channel is the complement of an EB channel, and since each channel in the complementary pair $(\mathcal{N}_s, \mathcal{N}_s^c)$ is not EB, both the channels are not Hadamard channels. Another class of interesting channel are more capable channels. As mentioned in Sec. II-F, a more capable channel has a complement with no quantum capacity. Both \mathcal{N}_s and \mathcal{N}_s^c don't belong to this more capable channel class because these complementary channels both have non-zero quantum capacity. An argument showing that the quantum capacity of \mathcal{N}_s is non-zero was presented previously. An argument showing that the quantum capacity of \mathcal{N}_s is non-zero is given in Sec. IV and also in [30]. Since both \mathcal{N}_s and \mathcal{N}_s^c have non-zero quantum capacity, these channels are not PPT which

have zero quantum capacity. Finally, it is easy to verify that each channel in the $(\mathcal{N}_s, \mathcal{N}_s^c)$ pair is not unital.

The \mathcal{N}_s channel can also be viewed as a hybrid of two simple qubit input channels. The first channel, \mathcal{R}_1 , perfectly maps its input, \mathcal{H}_1 , to the environment \mathcal{H}_c . The second channel, \mathcal{R}_2^s for $s \in [0, 1/2]$, is a degradable channel. To define the first channel, $\mathcal{R}_1 : \hat{\mathcal{H}}_1 \mapsto \hat{\mathcal{H}}_b$, we use $\{|1\rangle, |2\rangle\}$ to label an orthonormal basis of \mathcal{H}_1 . Then an isometry $R_1 : \mathcal{H}_1 \mapsto \mathcal{H}_b \otimes \mathcal{H}_c$ of the form

$$\begin{aligned} R_1|1\rangle &= |2\rangle \otimes |0\rangle \\ R_1|2\rangle &= |2\rangle \otimes |1\rangle \end{aligned} \quad (23)$$

defines $\mathcal{R}_1(X) = \text{Tr}_c(R_1 X R_1^\dagger) = \text{Tr}(X)|2\rangle\langle 2|$. Clearly \mathcal{R}_1 traces out its input to a fixed pure state while sending everything to the environment. Thus \mathcal{R}_1 cannot send any information to its output, $\mathcal{Q}(\mathcal{R}_1) = \mathcal{P}(\mathcal{R}_1) = \mathcal{C}(\mathcal{R}_1) = 0$. To define the second channel \mathcal{R}_2^s , we use $\{|0\rangle, |2\rangle\}$ to label an orthonormal basis of \mathcal{H}_2 . Then $\mathcal{R}_2^s : \hat{\mathcal{H}}_2 \mapsto \hat{\mathcal{H}}_b$ is generated by an isometry, $R_2^s : \mathcal{H}_2 \mapsto \mathcal{H}_b \otimes \mathcal{H}_c$, of the form

$$\begin{aligned} R_2^s|0\rangle &= \sqrt{s} |0\rangle \otimes |0\rangle + \sqrt{1-s} |1\rangle \otimes |1\rangle \\ R_2^s|2\rangle &= |2\rangle \otimes |1\rangle, \end{aligned} \quad (24)$$

where $0 \leq s \leq 1/2$. One can show that \mathcal{R}_2^s is degradable and thus $\mathcal{Q}^{(1)}(\mathcal{R}_2^s) = \mathcal{Q}(\mathcal{R}_2^s) = \mathcal{P}(\mathcal{R}_2^s)$. The channel's classical capacity is one bit. This can be shown by noticing that a bit encoded into inputs $|0\rangle$ and $|2\rangle$ leads to orthogonal outputs achieving a rate of 1, which saturates the maximal entropy determined by the dimensional bound of the input of \mathcal{R}_2^s .

Notice \mathcal{H}_a is the union of \mathcal{H}_1 and \mathcal{H}_2 and the isometry $F_s : \mathcal{H}_a \mapsto \mathcal{H}_b \otimes \mathcal{H}_c$, which gives rise to \mathcal{N}_s , can be written as a hybrid,

$$\begin{aligned} F_s|0\rangle &= R_2^s|0\rangle \\ F_s|1\rangle &= R_1|1\rangle \\ F_s|2\rangle &= R_1|2\rangle = R_2^s|2\rangle, \end{aligned} \quad (25)$$

of isometries R_1 and R_2^s that give rise to \mathcal{R}_1 and \mathcal{R}_2^s , respectively.

B. The \mathcal{M}_d Channel

We now consider a higher dimensional generalization of the isometry in (20) with $s = 1/2$. This generalization, $G : \mathcal{H}_a \mapsto \mathcal{H}_b \otimes \mathcal{H}_c$ with $G^\dagger G = I_a$, operates on Hilbert spaces with dimensions $d_a = d_b = d$, $d_c = d - 1$, and G has the form,

$$\begin{aligned} G|0\rangle &= \frac{1}{\sqrt{d-1}} \sum_{j=0}^{d-2} |j\rangle \otimes |j\rangle \\ G|i\rangle &= |d-1\rangle \otimes |i-1\rangle, \end{aligned} \quad (26)$$

where $1 \leq i \leq d-1$ and $d \geq 3$. When $d = 3$, the isometry above is equivalent to F_s in (20) at $s = 1/2$ and defines a channel pair $(\mathcal{N}_{1/2}, \mathcal{N}_{1/2}^c)$. In general $d \geq 3$ and the isometry (26) defines a pair of channels denoted by $\mathcal{M}_d : \hat{\mathcal{H}}_a \mapsto \hat{\mathcal{H}}_b$ and $\mathcal{M}_d^c : \hat{\mathcal{H}}_a \mapsto \hat{\mathcal{H}}_c$. Let U be any unitary on a $(d-1)$ dimensional Hilbert space. Using U and the standard basis on $\mathcal{H}_a, \mathcal{H}_b$, and \mathcal{H}_c we define unitary operators

$U_a = [0] \oplus U$, $U_b = U \oplus [d-1]$, and $U_c = U^*$ on $\mathcal{H}_a, \mathcal{H}_b$, and \mathcal{H}_c respectively, here $*$ denotes complex conjugation in the standard basis of \mathcal{H}_c . The isometry G (26) is symmetrical in the sense that

$$GU_a = (U_b \otimes U_c)G. \quad (27)$$

Due to this symmetry, both channels, \mathcal{M}_d and \mathcal{M}_d^c , have the property that

$$\begin{aligned} \mathcal{M}_d(U_a \rho U_a^\dagger) &= U_b \mathcal{M}_d(\rho) U_b^\dagger \\ \mathcal{M}_d^c(U_a \rho U_a^\dagger) &= U_c \mathcal{M}_d(\rho) U_c^\dagger, \end{aligned} \quad (28)$$

where ρ is any density operator on \mathcal{H}_a .

The channels \mathcal{M}_d and \mathcal{M}_d^c have properties similar to those of \mathcal{N}_s and \mathcal{N}_s^c , respectively. Like \mathcal{N}_s , the channel \mathcal{M}_d has several degradable sub-channels. Later in Sec. III-D we show that a degradable sub-channel of \mathcal{M}_d is obtained by restricting its input to operators on a qubit subspace \mathcal{H}_{ai} (21) or by restricting the input to a qubit subspace obtained by applying U_a , defined above in (27), to each state in \mathcal{H}_{ai} , where i is some fixed number between 1 and $d-1$. Quantum states lying solely at the input of such a degradable qubit sub-channel can be used to send an equal amount of quantum and private information to the \mathcal{M}_d channel output \mathcal{H}_b but such states cannot be used to send any quantum or private information to the \mathcal{M}_d^c output \mathcal{H}_c .

Like \mathcal{N}_s^c , the channel \mathcal{M}_d^c has a perfect sub-channel to its output \mathcal{H}_c . In Sec. III-D, we show this perfect sub-channel is obtained by restricting the channel input to $\mathcal{H}_{a'}$, a $(d-1)$ -dimensional subspace of \mathcal{H}_a spanned by $\{|i\rangle\}_{i=1}^{d-1}$. Quantum states lying solely in the channel input subspace $\mathcal{H}_{a'}$ can be used to perfectly send information to the \mathcal{M}_d^c output \mathcal{H}_c , while sending no information to the \mathcal{M}_d output \mathcal{H}_b . This perfect transmission, along with arguments similar to those above (22), can be used to show that the capacities of \mathcal{M}_d^c satisfy

$$\mathcal{Q}(\mathcal{M}_d^c) = \mathcal{P}(\mathcal{M}_d^c) = \mathcal{C}(\mathcal{M}_d^c) = \log(d-1). \quad (29)$$

A variety of channel classes were discussed in Sec. II-F. These classes include degradable channels, anti-degradable channels, EB channels, Hadamard channels, and less noisy channels. If a sub-channel of a channel does not belong to any of these classes, then the channel itself does not belong to that same class. Notice $\mathcal{N}_{1/2}$ is a sub-channel of \mathcal{M}_d arising from a restriction of the \mathcal{M}_d channel input to operators on a qutrit subspace spanned by $\{|0\rangle, |1\rangle, |2\rangle\}$. In Sec. III-A, we mentioned that $\mathcal{N}_{1/2}$ is not degradable, anti-degradable, EB, Hadamard, or less noisy. As a result \mathcal{M}_d is not (anti)-degradable, EB, Hadamard, or less noisy. One can easily verify that \mathcal{M}_d , like $\mathcal{N}_{1/2}$, is not a unital channel or a PPT channel. This verification can be done by a direct computation. In the case of ruling out PPT behaviour one can also use the fact that PPT channels have zero quantum capacity, however $\mathcal{N}_{1/2}$, a sub-channel of \mathcal{M}_d has non-zero quantum capacity and thus \mathcal{M}_d also has non-zero quantum capacity. An argument similar to the one above can be used to show that, similar to \mathcal{N}_s^c , the channel \mathcal{M}_d^c is not (anti)-degradable, EB, Hadamard, less noisy, unital or PPT.

C. The General \mathcal{O} Channel

Both isometries in (20) and (26) can be viewed as special cases of an isometry, $H: \mathcal{H}_a \mapsto \mathcal{H}_b \otimes \mathcal{H}_c$, where $d_a = d_b = d$, $d_c = d-1$,

$$H|0\rangle = \sum_{j=0}^{d-2} \mu_j |j\rangle \otimes |j\rangle, \quad H|i\rangle = |d-1\rangle \otimes |i-1\rangle, \quad (30)$$

$1 \leq i \leq d-1$, $\mu_i \in \mathbb{C}$, and $\sum_i |\mu_i|^2 = 1$. Note that complex phases in μ_i can be absorbed in the definition of the standard basis of \mathcal{H}_b , so we may assume that μ_i are real and non-negative. An exchange of μ_i and μ_j ($0 \leq i \neq j \leq d-2$) is equivalent to an exchange of $|i+1\rangle$ and $|j+1\rangle$ in \mathcal{H}_a and an exchange of $|i\rangle$ and $|j\rangle$ in both \mathcal{H}_b and \mathcal{H}_c . All these exchanges can be achieved by performing local unitaries on $\mathcal{H}_a, \mathcal{H}_b$, and \mathcal{H}_c respectively. As a result, we restrict our attention to $\{\mu_i\}$ arranged in ascending order, i.e., $\mu_0 \leq \mu_1 \leq \dots \leq \mu_{d-2}$. In (30), the special case obtained by setting $d=3$, $\mu_0 = \sqrt{s}$ and $\mu_1 = \sqrt{1-s}$ gives (20). On the other hand, when d is more general and all μ_i take equal values $1/\sqrt{d-1}$, we obtain (26).

The isometry H generates a pair of channels $(\mathcal{O}, \mathcal{O}^c)$. Like \mathcal{N}_s in Sec. III-A and \mathcal{M}_d in Sec. III-B, the channel pair \mathcal{O} has several degradable sub-channels. Restricting the input of \mathcal{O} to operators on \mathcal{H}_{ai} (21), i is any fixed number between 1 to $d-1$, results in a degradable sub-channel. The channel complement \mathcal{O}^c , like \mathcal{N}_s^c in Sec. III-A and \mathcal{M}_d^c in Sec. III-B, has a perfect sub-channel. This sub-channel has input dimension $d-1$ and is obtained by restricting the \mathcal{O}^c channel input to operators on $\mathcal{H}_{a'} = \text{span}\{|i\rangle\}_{i=1}^{d-1}$. Proofs for both these properties are in Sec. III-D.

As mentioned earlier in Sec. III-A and Sec. III-B, properties of the type mentioned above imply that quantum states lying in the qubit subspace \mathcal{H}_{ai} at the channel input can be used to send quantum and private information to the \mathcal{O} channel output but not the \mathcal{O}^c output. Additionally, quantum states lying solely in the $(d-1)$ -dimensional $\mathcal{H}_{a'}$ input subspace perfectly send all information to the \mathcal{O}^c output \mathcal{H}_c . Arguments similar to those above (22) can be used to show that

$$\mathcal{Q}(\mathcal{O}^c) = \mathcal{P}(\mathcal{O}^c) = \mathcal{C}(\mathcal{O}^c) = \log(d-1). \quad (31)$$

From (31), it follows that \mathcal{O}^c has non-zero quantum capacity and thus \mathcal{O}^c does not belong to the class of less noisy channels. Using a log-singularity based argument (see Sec. III-D), one can show that $\mathcal{Q}^{(1)}(\mathcal{O}) > 0$. As a result, \mathcal{O} is also outside the class of less noisy channels. Using arguments similar to those used for $(\mathcal{N}_s, \mathcal{N}_s^c)$ in Sec. III-A and for $(\mathcal{M}_d, \mathcal{M}_d^c)$ in Sec. III-B, one can show that both \mathcal{O} and \mathcal{O}^c are not degradable, anti-degradable, EB, Hadamard, unital, or PPT channels.

D. Proofs of Channel Properties

In Sec. III-A, below eq. (20), we claimed that \mathcal{N}_s has degradable sub-channels obtained by restricting the channel input to \mathcal{H}_{ai} in (21). We made a similar claim about \mathcal{M}_d , a d -dimensional generalization of $\mathcal{N}_{1/2}$ defined in Sec. III-B, and about \mathcal{O} , a d -dimensional generalization of both \mathcal{N}_s and \mathcal{M}_d defined in Sec. III-C. We prove all these claims here.

Instead of giving separate proofs for claims about each channel $\mathcal{N}_s, \mathcal{M}_d$, and \mathcal{O} , we give a single proof for the \mathcal{O} channel. From this single proof, correctness of claims for \mathcal{N}_s and \mathcal{M}_d follow as both \mathcal{N}_s and \mathcal{M}_d are special cases of \mathcal{O} (see discussion in Sec. III-C below (30)). Using this single proof, along with the definition of degradability and the symmetry in eq. (28), one can easily show an additional claim in Sec. III-B: restricting the input of \mathcal{M}_d to a qubit sub-space obtained by applying U_a , defined above eq. (27), to each state in \mathcal{H}_{ai} , results in a degradable sub-channel of \mathcal{M}_d .

Let \mathcal{O}_i be a sub-channel of \mathcal{O} obtained by restricting the channel input to a two-dimensional sub-space \mathcal{H}_{ai} (see eq. (21)) of \mathcal{H}_a , where i is a fixed integer between 1 and $d-1$. This sub-channel, \mathcal{O}_i , is degradable. Prior to constructing a degrading map for \mathcal{O}_i , consider an isometry $H_i: \mathcal{H}_{ai} \mapsto \mathcal{H}_b \otimes \mathcal{H}_c$,

$$H_i|0\rangle = H|0\rangle, \quad H_i|i\rangle = H|i\rangle, \quad (32)$$

where H is defined in (30). This isometry H_i generates the \mathcal{O}_i sub-channel and its complement \mathcal{O}_i^c . Let \mathcal{H}_e be a d -dimensional Hilbert space. For any fixed integer i between 1 and $d-1$, consider another isometry $K_i: \mathcal{H}_b \mapsto \mathcal{H}_c \otimes \mathcal{H}_e$ of the form

$$\begin{aligned} K_i|j\rangle &= |j\rangle \otimes |j\rangle \\ K_i|d-1\rangle &= |i-1\rangle \otimes |d-1\rangle, \end{aligned} \quad (33)$$

where $0 \leq j \leq d-2$. The $\mathcal{F}_i: \hat{\mathcal{H}}_b \mapsto \hat{\mathcal{H}}_c$ channel generated by the isometry K_i satisfies

$$\mathcal{F}_i \circ \mathcal{O}_i = \mathcal{O}_i^c. \quad (34)$$

The equality above implies that \mathcal{F}_i is a degrading map for \mathcal{O}_i and thus \mathcal{O}_i is degradable and \mathcal{O}_i^c anti-degradable.

In Sec. III-A we claimed that \mathcal{N}_s^c has perfect sub-channels obtained by restricting the channel input to $\mathcal{H}_{a'}$ (defined below (21)). A similar claim was made about \mathcal{M}_d^c in Sec. III-B, and about \mathcal{O}^c in Sec. III-C. All these claims are proven here by giving a single proof for the \mathcal{O}^c channel. From this proof, correctness of claims for \mathcal{N}_s^c and \mathcal{M}_d^c follow as both \mathcal{N}_s^c and \mathcal{M}_d^c are special cases of \mathcal{O}^c (see discussion in Sec. III-C below (30)). To prove that \mathcal{O}^c has a $(d-1)$ -dimensional sub-channel which perfectly maps its input to \mathcal{H}_c , consider

$$\mathcal{H}_{a'} = \text{span}\{|1\rangle, |2\rangle, \dots, |d-1\rangle\}, \quad (35)$$

a $(d-1)$ -dimensional sub-space of \mathcal{H}_a with projector $P_{a'} = \sum_{i=1}^{d-1} |i\rangle\langle i|$. Restricting the input of \mathcal{O} to $\mathcal{H}_{a'}$ results in a sub-channel $\mathcal{O}': \mathcal{H}_{a'} \mapsto \mathcal{H}_c$ with super-operator

$$\mathcal{O}'(A) = VAV^\dagger, \quad (36)$$

where $V: \mathcal{H}_{a'} \mapsto \mathcal{H}_c$ is a bijection of the form

$$V|i\rangle = |i-1\rangle, \quad 1 \leq i \leq d-1, \quad (37)$$

satisfying

$$VV^\dagger = P_{a'}, \quad \text{and} \quad V^\dagger V = I_c. \quad (38)$$

Since V is a bijection, the sub-channel \mathcal{O}' in (36) perfectly maps its input space $\mathcal{H}_{a'}$ to its $(d-1)$ -dimensional output \mathcal{H}_c .

IV. CHANNEL COHERENT INFORMATION

The coherent information of $\mathcal{N}_s, \mathcal{M}_d$, and \mathcal{O} can be obtained from an optimization of the form (5). In general, this optimization is non-trivial to carry out because the entropy bias (4) is not generally concave in ρ . As a result the coherent information $\mathcal{Q}^{(1)}$ for most channels remains unknown. However, one can show that optimizations for $\mathcal{Q}^{(1)}$ of all three channels $\mathcal{N}_s, \mathcal{M}_d$, and \mathcal{O} can be reduced to a one-parameter concave maximization over a bounded interval. For any $0 \leq s \leq 1/2$,

$$\mathcal{Q}^{(1)}(\mathcal{N}_s) = \max_{\rho_a(u)} \Delta(\mathcal{N}_s, \rho_a(u)), \quad (39)$$

where $\rho_a(u)$ is a one-parameter density operator of the form $\rho_a(u) = (1-u)[0] + u[2]$ and $0 \leq u \leq 1$. For any $d \geq 3$,

$$\mathcal{Q}^{(1)}(\mathcal{M}_d) = \max_{\rho_a(u)} \Delta(\mathcal{M}_d, \rho_a(u)), \quad (40)$$

where $\rho_a(u) = (1-u)[0] + u[i]$, $0 \leq u \leq 1$ and i is any fixed integer between 1 and $d-1$. For any $\mu_0 \leq \mu_1 \leq \dots \leq \mu_{d-2}$,

$$\mathcal{Q}^{(1)}(\mathcal{O}) = \max_{\rho_a(u)} \Delta(\mathcal{O}, \rho_a(u)), \quad (41)$$

where

$$\rho_a(u) = (1-u)[0] + u[d-1], \quad (42)$$

and $0 \leq u \leq 1$. In (39), (40), and (41) the maximization is over a density operator $\rho_a(u)$ which is supported over a subspace of the \mathcal{H}_{ai} form (21). In Secs. III-A, III-B, and III-C, we showed that restricting the channel input to these \mathcal{H}_{ai} subspaces results in a degradable channel. Since the entropy bias of a degradable channel is concave in the channel input, each optimization in (39), (40), and (41) is a concave maximization. In addition, one can also show that all three coherent informations (39), (40), and (41) are strictly positive.

Since \mathcal{N}_s and \mathcal{M}_d are special cases of \mathcal{O} (see discussion below (26)), proving $\mathcal{Q}^{(1)}(\mathcal{O}) > 0$ and the equality (41) also proves $\mathcal{Q}^{(1)}(\mathcal{N}_s) > 0$, $\mathcal{Q}^{(1)}(\mathcal{M}_d) > 0$ and proves the equalities in both (39) and (40). In what follows, Sec. IV-A shows the equality (41), Sec. IV-C shows $\mathcal{Q}^{(1)}(\mathcal{O}) > 0$, Sec. IV-D provides additional insight into certain sub-channels of \mathcal{O} and finally Sec. IV-E extends our results to Rényi entropies.

A. Evaluating $\mathcal{Q}^{(1)}(\mathcal{O})$

Let \mathcal{O}_{d-1} be the sub-channel of \mathcal{O} obtained by restricting the channel input to the sub-space \mathcal{H}_{ai} in (21) where $i = d-1$. In this subsection, we will show that

$$\begin{aligned} \mathcal{Q}^{(1)}(\mathcal{O}) &:= \max_{\rho_a} \Delta(\mathcal{O}, \rho_a) \\ &= \max_{0 \leq u \leq 1} \Delta(\mathcal{O}_{d-1}, \rho_a(u)) \\ &= \mathcal{Q}^{(1)}(\mathcal{O}_{d-1}), \end{aligned} \quad (43)$$

where $\rho_a(u)$ is defined in (42). The above equations can be proved in three steps. The first step exploits the structure of $(\mathcal{O}, \mathcal{O}^c)$ to show that restricting the input ρ_a to a certain block-diagonal form preserves the optimal value in the

second expression in (43). The second step reduces this new optimization further to a one parameter problem (50), using a majorization-based approach. Finally, this one-parameter problem is shown to be equivalent to finding the coherent information of \mathcal{O}_{d-1} . These three steps are detailed as follows.

Recall from (35) that $\mathcal{H}_{a'} = \text{span}\{|1\rangle, |2\rangle, \dots, |d-1\rangle\}$. Let

$$\mathcal{H}_0 := \text{span}\{|0\rangle\}, \quad \text{and} \quad \mathcal{H}_1 := \mathcal{H}_{a'}, \quad (44)$$

so that $\mathcal{H}_a = \mathcal{H}_0 \oplus \mathcal{H}_1$. Let \mathcal{L}_{ij} be the space of linear operators from \mathcal{H}_j to \mathcal{H}_i , i.e.,

$$\mathcal{L}_{ij} = \{E_{ij} \mid E_{ij} : \mathcal{H}_j \mapsto \mathcal{H}_i\}. \quad (45)$$

Any density operator on \mathcal{H}_a can be written as

$$\rho_a = \bigoplus_{i,j} N_{ij}, \quad (46)$$

where i, j are binary and N_{ij} are linear operators from \mathcal{H}_j to \mathcal{H}_i .

Remark 1: Using (30), it is straightforward to verify that, in the standard basis of \mathcal{H}_b , $\mathcal{O}(N_{ii})$ has zero off-diagonal entries, and if $i \neq j$, $\mathcal{O}(N_{ij})$ has zero diagonal entries, while $\mathcal{O}^c(N_{ij})$ is the null operator on \mathcal{H}_c .

For step 1, starting from any ρ_a , we obtain $\tilde{\rho}_a$ by resetting $N_{ij} = 0$ for $i \neq j$. We claim that $\Delta(\mathcal{O}, \rho_a) \leq \Delta(\mathcal{O}, \tilde{\rho}_a)$. To see this, let $\tilde{\rho}_b = \mathcal{O}(\tilde{\rho}_a)$ and $\tilde{\rho}_c = \mathcal{O}^c(\tilde{\rho}_a)$. Using the above remark, $\tilde{\rho}_b$ can be obtained from ρ_b by resetting all the off-diagonal elements to 0, and $\tilde{\rho}_b$ is majorized by ρ_b (see discussion below eq. (54) and Prb.II.5.5 in [46]). Applying Schur-concavity of the von-Neumann entropy (see Sec. IV-B), we have $S(\tilde{\rho}_b) \geq S(\rho_b)$. The above remark also implies that $\tilde{\rho}_c = \rho_c$, so $S(\tilde{\rho}_c) = S(\rho_c)$. Together, this proves the claim $\Delta(\mathcal{O}, \rho_a) \leq \Delta(\mathcal{O}, \tilde{\rho}_a)$. Thus, to maximize the entropy bias (4) we can focus on ρ_a of the form in (46) where $N_{01} = N_{10} = 0$, i.e.,

$$\rho_a = (1-u)[0] + u\sigma, \quad (47)$$

where $0 \leq u \leq 1$ and σ is a density operator on $\mathcal{H}_{a'}$.

For step 2, note that the input ρ_a in (47) gives the outputs

$$\rho_b = (1-u) \left(\sum_{i=0}^{d-2} |\mu_i|^2 [i] \right) + u[d-1], \quad (48)$$

and

$$\rho_c = (1-u) \left(\sum_{i=0}^{d-2} |\mu_i|^2 [i] \right) + uV\sigma V^\dagger, \quad (49)$$

where the channel parameters $\mu_0 \leq \mu_1 \leq \dots \leq \mu_{d-2}$ are fixed. Note that $S(\rho_b)$ only depends on u while $S(\rho_c)$ depends on both u and σ . Thus for any fixed u , the entropy bias $\Delta(\mathcal{O}, \rho_a)$ is maximum when $S(\rho_c)$ is minimum. We will prove in the next subsection that this minimum can always be attained when $\sigma = [d-1]$, for all relevant values of u and the channel parameters $\mu_0 \leq \mu_1 \leq \dots \leq \mu_{d-2}$. As a result,

$$\mathcal{Q}^{(1)}(\mathcal{O}) = \max_{0 \leq u \leq 1} \Delta(\mathcal{O}, \rho_a(u)), \quad (50)$$

where $\rho_a(u)$ is as given in (42), obtained from setting $\sigma = [d-1]$ in (47).

To finish the proof, note that $\rho_a(u)$ is supported on $\mathcal{H}_{a(d-1)}$, so, expanding the maximization to a general density operator Λ supported on $\mathcal{H}_{a(d-1)}$ is nondecreasing:

$$\max_{0 \leq u \leq 1} \Delta(\mathcal{O}, \rho_a(u)) \leq \max_{\Lambda} \Delta(\mathcal{O}, \Lambda). \quad (51)$$

Since Λ is supported on $\mathcal{H}_{a(d-1)}$,

$$\max_{\Lambda} \Delta(\mathcal{O}, \Lambda) = \max_{\Lambda} \Delta(\mathcal{O}_{d-1}, \Lambda), \quad (52)$$

and the RHS of the above is $\mathcal{Q}^{(1)}(\mathcal{O}_{d-1})$. Combining (50), (51), and (52) thus gives

$$\mathcal{Q}^{(1)}(\mathcal{O}) \leq \mathcal{Q}^{(1)}(\mathcal{O}_{d-1}). \quad (53)$$

The opposite inequality holds since \mathcal{O}_{d-1} is a sub-channel of \mathcal{O} , which establishes (43).

B. Majorization and Entropy Minimization

In this subsection, we prove the claim in the previous subsection that ρ_c in (49) has minimum entropy when $\sigma = [d-1]$, for any fixed u and channel parameters $\mu_0 \leq \mu_1 \leq \dots \leq \mu_{d-2}$. We first summarize our main tools, majorization and Schur-concavity of the Shannon entropy. Let us first consider majorization of real vectors. Given a vector \mathbf{x} in \mathbb{R}^t , let \mathbf{x}^\downarrow denote the vector obtained by rearranging the entries of \mathbf{x} in descending order. For any two vectors \mathbf{x}, \mathbf{y} in \mathbb{R}^t , we say that \mathbf{x} is majorized by \mathbf{y} , in symbols $\mathbf{x} \prec \mathbf{y}$, if the inequality,

$$\sum_{j=1}^k \mathbf{x}_j^\downarrow \leq \sum_{j=1}^k \mathbf{y}_j^\downarrow, \quad (54)$$

holds for all $k \leq t$ and becomes an equality at $k = t$. The concept of majorization can be generalized to Hermitian matrices as follows. For any $t \times t$ Hermitian operator N , let $\mathbf{v}(N)$ denote the vector of singular values of N . For two Hermitian operators N, M of the same size, we say that N is majorized by M , $N \prec M$ in symbols, if $\mathbf{v}(N) \prec \mathbf{v}(M)$. The operator $N + M$ satisfies (see Ex.II.1.15 in [46]),

$$\sum_{j=1}^k \mathbf{v}_j^\downarrow(N+M) \leq \sum_{j=1}^k \mathbf{v}_j^\downarrow(N) + \sum_{j=1}^k \mathbf{v}_j^\downarrow(M) \quad (55)$$

for all integers $1 \leq k \leq d$, as a result,

$$\mathbf{v}(N+M) \prec \mathbf{v}^\downarrow(N) + \mathbf{v}^\downarrow(M). \quad (56)$$

The Shannon entropy is Schur-concave, i.e., when a probability vector, \mathbf{p} , is majorized by another, \mathbf{q} , then $h(\mathbf{p}) \geq h(\mathbf{q})$. The von-Neumann entropy of a density operator τ is equal to the Shannon entropy of $\mathbf{v}(\tau)$:

$$S(\tau) = h(\mathbf{v}(\tau)). \quad (57)$$

Thus, like the Shannon entropy, the von-Neumann entropy is Schur-concave, i.e., if $\tau \prec \kappa$ then $S(\tau) \geq S(\kappa)$.

We now return to the problem of minimizing the von-Neumann entropy $S(\rho_c)$. Recalled from (49) that $\rho_c = (1-u) \left(\sum_{i=0}^{d-2} |\mu_i|^2 [i] \right) + uV\sigma V^\dagger$. Let $\sigma' = V\sigma V^\dagger$ and $\Upsilon := \sum_{i=0}^{d-2} |\mu_i|^2 [i]$, so,

$$\rho_c = (1-u)\Upsilon + u\sigma'. \quad (58)$$

Since $S(\rho_c)$ is a concave function, the minimum can be attained when σ' is a pure state. Applying (56) to (58), we obtain

$$\mathbf{v}(\rho_c) \prec (1-u)\mathbf{v}^\perp(\Upsilon) + u\mathbf{v}^\perp(\sigma') \quad (59)$$

where we use $\mathbf{v}(cN) = c\mathbf{v}(N)$ for any real $c > 0$. Since σ' is a pure state, the RHS of (59) is a vector $\mathbf{w} = ((1-u)|\mu_{d-2}|^2 + u, (1-u)|\mu_{d-3}|^2, \dots, (1-u)|\mu_0|^2)$, and from Schur-concavity of the Shannon entropy,

$$S(\rho_c) = h(\mathbf{v}(\rho_c)) \geq h(\mathbf{w}). \quad (60)$$

From the expression of \mathbf{w} , equality is attained when $\sigma' = [d-2]$. So the entropy of ρ_c in (58) is minimized when $\sigma' = [d-2]$. Using the expression of V in (37), $\sigma = [d-1]$.

C. Log-Singularity and Positivity

We shall be using an ϵ log-singularity based method to show $\mathcal{Q}^{(1)}(\mathcal{O}) > 0$ for any $\mu_0 \leq \mu_1 \leq \dots \leq \mu_{d-2}$ and $d \geq 3$. For details of the method see [30]. Let $\rho(\epsilon)$ be a density operator that depends on a real parameter ϵ . The von-Neumann entropy, $S(\epsilon) = -\text{Tr}(\rho(\epsilon) \log \rho(\epsilon))$ is said to have an ϵ log-singularity if one or several eigenvalues of $\rho(\epsilon)$ increase linearly from 0 to leading order in ϵ . As a result of this singularity, $S(\epsilon) \simeq x|\epsilon \log \epsilon|$ where $x > 0$ is called the *rate* of the singularity. The entropy bias, or coherent information of \mathcal{O} at $\rho_a(\epsilon)$, may be concisely denoted by,

$$\Delta(\epsilon) := S_b(\epsilon) - S_c(\epsilon), \quad (61)$$

where $S_b(\epsilon) := S(\rho_b(\epsilon))$ and $S_c(\epsilon) = S(\rho_c(\epsilon))$. If an ϵ log-singularity in $S_b(\epsilon)$ has larger rate than the one in $S_c(\epsilon)$, and $\Delta(0) = 0$, then $\mathcal{Q}^{(1)}(\mathcal{O}) > 0$. In the present case, let $\rho_a(\epsilon)$ be the density operator in (42) where u is replaced with ϵ . At $\epsilon = 0$, $\rho_a(\epsilon)$ is a pure state, hence $\Delta(0) = 0$. In addition, $\rho_c(0)$ has rank $d_c = d-1$, and thus $S_c(\epsilon)$ cannot have an ϵ log-singularity. On the other hand $S_b(\epsilon)$ has an ϵ log-singularity of rate 1 and thus $\mathcal{Q}^{(1)}(\mathcal{O}) > 0$.

D. Sub-Channels

Let \mathcal{O}_i be the sub-channel of \mathcal{O} obtained by restricting the channel input to operators on \mathcal{H}_{ai} (21), where i is some fixed integer between 1 and $d-1$. Using arguments similar to those presented in subsection IV-A, one can show

$$\mathcal{Q}^{(1)}(\mathcal{O}_i) = \max_{\rho_{ai}(u)} \Delta(\mathcal{O}_i, \rho_{ai}(u)), \quad (62)$$

where $\rho_{ai}(u)$ is a one-parameter density operator of the form

$$\rho_{ai}(u) = (1-u)[0] + u[i], \quad (63)$$

and $0 \leq u \leq 1$. Let $S(\rho_{ci})$ be the entropy of $\rho_{ci} = \mathcal{O}_i^c(\rho_{ai})$. Using majorization and Schur concavity arguments (similar to those leading to (60)) and the fact that $\mu_0 \leq \mu_1 \leq \dots \leq \mu_{d-2}$ one can show that

$$S(\rho_{ci}) \geq S(\rho_{cj}), \quad (64)$$

for all $i \leq j$. For any fixed u in (63) the entropy $S(\rho_{bi})$ of $\rho_{bi} = \mathcal{O}_i(\rho_{ai})$, is independent of i . Using (64) and the definition of the entropy bias (4) one obtains,

$$\Delta(\mathcal{O}_i, \rho_{ai}(u)) \leq \Delta(\mathcal{O}_j, \rho_{aj}(u)), \quad (65)$$

for any $i \leq j$. The relation above implies,

$$\mathcal{Q}^{(1)}(\mathcal{O}_i) \leq \mathcal{Q}^{(1)}(\mathcal{O}_j), \quad (66)$$

for all $i \leq j$. Using the above equation along with (43) we get

$$\mathcal{Q}^{(1)}(\mathcal{O}_1) \leq \mathcal{Q}^{(1)}(\mathcal{O}_2) \leq \dots \leq \mathcal{Q}^{(1)}(\mathcal{O}_{d-1}) = \mathcal{Q}^{(1)}(\mathcal{O}) \quad (67)$$

E. Rényi Entropy

Before closing this section, we remark that to obtain (43) we used (a) monotonicity of the von-Neumann entropy $S(\rho_b)$ under block-diagonalization of ρ_a , and (b) concavity of the von-Neumann entropy to argue that pure σ' minimizes $S(\rho_c)$, and finally, we utilized majorization (see Sec. IV-B) to argue that the maximize entropy bias occurs when the input density operator has the form (42). In the definition of (4) if one replaces the von-Neumann entropy with any Rényi entropy,

$$S_\alpha(\rho) = \frac{1}{1-\alpha} \log \text{Tr}(\rho^\alpha), \quad (68)$$

where $0 \leq \alpha \leq \infty$, then the corresponding equation (43) would still hold. We outline the reasoning here. First, monotonicity of the Schur-concave Rényi entropy is unaffected when ρ_a is block-diagonalized. However, unlike the von-Neumann entropy, S_α is not concave for $\alpha > 1$. Thus to prove that the minimum Rényi entropy $S_\alpha(\rho_c)$ (59) also occurs over pure states σ' we employ a different strategy. Write

$$S_\alpha(\rho_c) = S_\alpha(\mathcal{N}(\sigma')) \quad (69)$$

where the quantum channel \mathcal{N} acts as $\mathcal{N}(\sigma') = (1-u)\Upsilon\text{Tr}(\sigma') + u\sigma'$, with $0 \leq u \leq 1$. Next we use the fact that the minimum output Rényi entropy for $\mathcal{N}(\sigma')$ occurs at a pure state input (see Sec. II in [47]). With σ' being restricted to a pure state, one can now use the majorization-based argument to show that the Rényi entropy $S_\alpha(\rho_c)$ is minimum when the input density operator has the form (42). This majorization argument is unaffected when the von-Neumann entropy is replaced by the Rényi entropy which is also Schur-concave.

V. QUANTUM CAPACITY OF \mathcal{O}

Subject to the *spin-alignment conjecture*, introduced in Sec. VI, we show that

$$\begin{aligned} \mathcal{Q}(\mathcal{N}_s) &= \mathcal{Q}^{(1)}(\mathcal{N}_s) \\ \mathcal{Q}(\mathcal{M}_d) &= \mathcal{Q}^{(1)}(\mathcal{M}_d) \\ \mathcal{Q}(\mathcal{O}) &= \mathcal{Q}^{(1)}(\mathcal{O}). \end{aligned} \quad (70)$$

To show these equalities above, we prove the third equality in (70) and infer the other two because \mathcal{N}_s and \mathcal{M}_d are special cases of \mathcal{O} (see Sec. III-C).

Our next step is to compute the coherent information of the channel $\mathcal{O}^{\otimes n}: \mathcal{H}_a^{\otimes n} \mapsto \mathcal{H}_b^{\otimes n}$

$$\begin{aligned} Q^{(1)}(\mathcal{O}^{\otimes n}) &= \max_{\rho} \Delta(\mathcal{O}^{\otimes n}, \rho) \\ &= \max_{\rho} [S(\mathcal{O}^{\otimes n}(\rho)) - S((\mathcal{O}^c)^{\otimes n}(\rho))], \end{aligned} \quad (71)$$

where $n \geq 1$ and ρ is a density operator on $\mathcal{H}_a^{\otimes n} = \mathcal{H}_a^1 \otimes \mathcal{H}_a^2 \otimes \cdots \otimes \mathcal{H}_a^n$, with the superscript i indicating the i^{th} space. Using (44), express

$$\mathcal{H}_a^{\otimes n} = (\mathcal{H}_0 \oplus \mathcal{H}_1)^{\otimes n} = \bigoplus_{\mathbf{b} \in \{0,1\}^n} \mathcal{H}(\mathbf{b}) \quad (72)$$

where $\{0,1\}^n := \{\mathbf{b} \mid \mathbf{b}(i) \in \{0,1\}, 1 \leq i \leq n\}$ is the set of n bit strings and $\mathcal{H}(\mathbf{b}) := \mathcal{H}_{\mathbf{b}(1)} \otimes \mathcal{H}_{\mathbf{b}(2)} \otimes \cdots \otimes \mathcal{H}_{\mathbf{b}(n)}$. Let $\mathcal{L}(\mathbf{s}, \mathbf{t})$ be the space of linear operators from $\mathcal{H}(\mathbf{t})$ to $\mathcal{H}(\mathbf{s})$. Any density operator ρ on $\mathcal{H}_a^{\otimes n}$ can be written as

$$\rho = \bigoplus_{\mathbf{s}, \mathbf{t} \in \{0,1\}^n} N(\mathbf{s}, \mathbf{t}), \quad (73)$$

for some $N(\mathbf{s}, \mathbf{t}) \in \mathcal{L}(\mathbf{s}, \mathbf{t})$. We will obtain the form of $\mathcal{O}^{\otimes n}(\rho)$ and $(\mathcal{O}^c)^{\otimes n}(\rho)$ using the expression of ρ above and the following two remarks (which generalize Remark 1).

Remark 2: If $\mathbf{s} = \mathbf{t}$ then $\mathcal{O}^{\otimes n}(N(\mathbf{s}, \mathbf{t}))$ has zero off-diagonal entries in the basis $\{|0\rangle, \dots, |d-1\rangle\}^{\otimes n}$ of $\mathcal{H}_b^{\otimes n}$. If $\mathbf{s} \neq \mathbf{t}$ then $\mathcal{O}^{\otimes n}(N(\mathbf{s}, \mathbf{t}))$ has zero diagonal entries in the aforementioned basis of $\mathcal{H}_b^{\otimes n}$, while $(\mathcal{O}^c)^{\otimes n}(N(\mathbf{s}, \mathbf{t})) = 0$, the null operator.

To prove remark 2, we first express any operator $N(\mathbf{s}, \mathbf{t})$ as a linear combination of operators of the form,

$$F = \bigotimes_{k=1}^n F_{\mathbf{s}(k)\mathbf{t}(k)}, \quad \text{where } F_{\mathbf{s}(k)\mathbf{t}(k)} \in \mathcal{L}_{\mathbf{s}(k)\mathbf{t}(k)}. \quad (74)$$

To illustrate these notations, for example, take $n = 3, \mathbf{s} = (0, 1, 1), \mathbf{t} = (1, 0, 1)$. Then, each of these operators F has the form $F = F_{01} \otimes F_{10} \otimes F_{11}$ where $F_{01}: \mathcal{H}_1 \mapsto \mathcal{H}_0$, $F_{10}: \mathcal{H}_0 \mapsto \mathcal{H}_1$, and $F_{11}: \mathcal{H}_1 \mapsto \mathcal{H}_1$.

Note that the properties in remark 2 are preserved under linear combinations, so, it suffices to prove the remark for an operator F in (74) instead of $N(\mathbf{s}, \mathbf{t})$. The channels $\mathcal{O}^{\otimes n}$ and $(\mathcal{O}^c)^{\otimes n}$ act on F in (74) as follows,

$$\mathcal{O}^{\otimes n}(F) = \bigotimes_{k=1}^n \mathcal{O}(F_{\mathbf{s}(k)\mathbf{t}(k)}), \quad \text{and} \quad (75)$$

$$(\mathcal{O}^c)^{\otimes n}(F) = \bigotimes_{k=1}^n \mathcal{O}^c(F_{\mathbf{s}(k)\mathbf{t}(k)}). \quad (76)$$

Proof: (of remark 2) If $\mathbf{s} = \mathbf{t}$, for each k , let $\mathbf{s}(k) = \mathbf{t}(k) = i_k \in \{0,1\}$. So, $F_{\mathbf{s}(k)\mathbf{t}(k)} = F_{i_k i_k}$. From Remark 1, $\mathcal{O}(F_{i_k i_k})$ is diagonal in the standard basis of \mathcal{H}_b , so, from (75), $\mathcal{O}^{\otimes n}(F)$ is also diagonal in the standard basis of $\mathcal{H}_b^{\otimes n}$. If $\mathbf{s} \neq \mathbf{t}$, they differ at some position k , so $F_{\mathbf{s}(k)\mathbf{t}(k)} = F_{ij}$ for $i \neq j$. From Remark 1, $\mathcal{O}(F_{\mathbf{s}(k)\mathbf{t}(k)}) = \mathcal{O}(F_{ij})$ has zero diagonal entries, so from (75), $\mathcal{O}^{\otimes n}(F)$ also has zero diagonal entries. Furthermore, from Remark 1 $\mathcal{O}^c(F_{ij}) = 0$, so from (76), $(\mathcal{O}^c)^{\otimes n}(F) = 0$ also. ■

Using (73), we write

$$\mathcal{O}^{\otimes n}(\rho) = \sum_{\mathbf{s} \neq \mathbf{t}} \mathcal{O}^{\otimes n}(N(\mathbf{s}, \mathbf{t})) + \sum_{\mathbf{s}} \mathcal{O}^{\otimes n}(N(\mathbf{s}, \mathbf{s})) \quad (77)$$

$$(\mathcal{O}^c)^{\otimes n}(\rho) = \sum_{\mathbf{s} \neq \mathbf{t}} (\mathcal{O}^c)^{\otimes n}(N(\mathbf{s}, \mathbf{t})) + \sum_{\mathbf{s}} (\mathcal{O}^c)^{\otimes n}(N(\mathbf{s}, \mathbf{s})) \quad (78)$$

From remark 2, the first term on the right side of the equality in (77) has zero diagonal entries, the second term has zero off-diagonal entries, while the first term on the right side of the equality in (78) is zero. Therefore, setting $N(\mathbf{s}, \mathbf{t}) = 0$ for all $\mathbf{s} \neq \mathbf{t}$ has no effect on $(\mathcal{O}^c)^{\otimes n}(\rho)$, nor on the diagonal entries of $\mathcal{O}^{\otimes n}(\rho)$, while all off-diagonal entries of $\mathcal{O}^{\otimes n}(\rho)$ become zero's. This may increase $S(\mathcal{O}^{\otimes n}(\rho))$ but not $S((\mathcal{O}^c)^{\otimes n}(\rho))$ so, when maximizing the entropy difference in the second equality in (71), we may restrict to density operators of the form

$$\rho = \bigoplus_{\mathbf{s} \in \{0,1\}^n} N(\mathbf{s}, \mathbf{s}). \quad (79)$$

To proceed with the analysis, we re-express (79) as a specific convex combination of states. Let M denote a subset of $\{1, \dots, n\}$ and M^c the complement of M in $\{1, \dots, n\}$. Let $|M|$ and $|M^c|$ denote the sizes of M and M^c . We may use the subset M to label some of the channel uses, or some of the input systems, or some of the output systems. For any such subset M , let ω_M denote a density operator acting on the corresponding subset of input spaces $\otimes_{i \in M} \mathcal{H}_a^i$, where \mathcal{H}_a^i denotes the input space of the i -th channel use, and let $|0\rangle\langle 0|^{\otimes M^c}$ denote the pure state $(|0\rangle\langle 0|)^{\otimes |M^c|}$ on the complement set of input spaces $\otimes_{j \in M^c} \mathcal{H}_a^j$. Using this notation, we now show that the density operator in (79) can be written as

$$\rho = \sum_M x_M \omega_M \otimes |0\rangle\langle 0|^{\otimes M^c}. \quad (80)$$

for some density operators ω_M and for some x_M between zero and one such that

$$\sum_M x_M = 1. \quad (81)$$

In the above and throughout, the summation over M is over all subsets of $\{1, \dots, n\}$. For an arbitrary \mathbf{s} , let the k -th entry of \mathbf{s} be i_k . Recall that $N(\mathbf{s}, \mathbf{s})$ is a linear combination of operators of the form $F_{i_1 i_1} \otimes F_{i_2 i_2} \otimes \cdots \otimes F_{i_n i_n}$. Recall also that $\mathcal{H}_0 = \text{span}\{|0\rangle\}$ so $F_{00} \propto |0\rangle\langle 0|$ and for each k where $i_k = 0$, $F_{i_k i_k} \propto |0\rangle\langle 0|$. It means that $N(\mathbf{s}, \mathbf{s}) = \eta_M \otimes |0\rangle\langle 0|^{\otimes M^c}$ where $M = \{i \in \{1, \dots, n\} : s_i = 1\}$ and $M^c = \{i \in \{1, \dots, n\} : s_i = 0\}$ and η_M is an operator acting on $\otimes_{i \in M} \mathcal{H}_a^i$. Substituting this into the expression for ρ in (79), and applying $\rho \geq 0$ and $\text{tr} \rho = 1$ gives $\eta_M \geq 0$ for each M , $\sum_M \text{tr} \eta_M = 1$ so we can write $\eta_M = x_M \omega_M$ for x_M between 0 and 1, $\sum_M x_M = 1$, and each ω_M is a density operator.

Using the form of ρ in (80), one may write

$$\mathcal{O}^{\otimes n}(\rho) = \sum_M x_M (|d-1\rangle\langle d-1|^{\otimes |M|})_M \otimes \Upsilon^{\otimes M^c}, \quad (82)$$

and

$$(\mathcal{O}^c)^{\otimes n}(\rho) = \sum_M x_M \omega'_M \otimes \Upsilon^{\otimes M^c}, \quad (83)$$

where $\omega'_M = V\omega_M V^\dagger$ and $\Upsilon = \sum_{i=0}^{d-2} |\mu_i|^2 |i\rangle\langle i|$ (see (30) and (37)). For fixed channel parameters $\{\mu_i\}$, the entropy $S(\mathcal{O}^{\otimes n}(\rho))$ only depends on $\{x_M\}$ while the entropy $S((\mathcal{O}^c)^{\otimes n}(\rho))$ depends on both ω_M and $\{x_M\}$. Thus for any fixed $\{\mu_i\}$ and $\{x_M\}$, the maximum entropy difference (71) occurs when $S((\mathcal{O}^c)^{\otimes n}(\rho))$ is minimum. This minimum can be obtained by solving the optimization problem

$$\begin{aligned} \min S(\kappa) \\ \kappa &= \sum_M x_M \omega'_M \otimes \Upsilon^{\otimes M^c}, \\ \omega'_M &\geq 0, \\ \text{Tr}(\omega'_M) &= 1, \forall M \subset \{1, \dots, n\}, \end{aligned} \quad (84)$$

where Υ only depends on the fixed $\{\mu_i\}$ and $\{x_M\}$ is also held fixed. In Sec. VI we introduce the *spin alignment conjecture*. Using this conjecture, the minimization problem above has a close form solution $\omega_M = (|d-1\rangle\langle d-1|)^{\otimes |M|}$ for any fixed $\{\mu_i\}$ and $\{x_M\}$. This closed form solution implies that the maximum entropy difference (71) is attained on an input density operator of the form

$$\rho = \sum_M x_M \omega_M \otimes (|0\rangle\langle 0|)^{\otimes M^c}, \quad (85)$$

where $\omega_M = (|d-1\rangle\langle d-1|)^{\otimes |M|}$. The above density operator is supported on a subspace $\mathcal{H}_{a_i}^{\otimes n}$, $i = d-1$ of $\mathcal{H}_a^{\otimes n}$. Using arguments similar to those in below eq. (50) the maximum entropy difference (71),

$$\mathcal{Q}^{(1)}(\mathcal{O}^{\otimes n}) = \max_{\rho} \Delta(\tilde{\mathcal{O}}^{\otimes n}, \rho) = \mathcal{Q}^{(1)}(\mathcal{O}_{d-1}^{\otimes n}). \quad (86)$$

In Sec. III-D, we showed that this sub-channel \mathcal{O}_{d-1} , is degradable. Since \mathcal{O}_{d-1} is degradable, $\mathcal{Q}^{(1)}(\mathcal{O}_{d-1}^{\otimes n}) = n\mathcal{Q}^{(1)}(\mathcal{O}_{d-1})$, and thus the equality (86) simplifies

$$\mathcal{Q}^{(1)}(\mathcal{O}^{\otimes n}) = n\mathcal{Q}^{(1)}(\mathcal{O}_{d-1}). \quad (87)$$

The equality along with (43) gives the desired result

$$\mathcal{Q}^{(1)}(\mathcal{O}^{\otimes n}) = n\mathcal{Q}^{(1)}(\mathcal{O}). \quad (88)$$

VI. SPIN ALIGNMENT CONJECTURE

Since this conjecture may be of independent interest, it is presented in a more self-contained manner, with occasional repetition of information from the previous section. Let $|0\rangle$ and $|1\rangle$ denote the spin up and spin down states of a spin- $\frac{1}{2}$ particle (we just call this a spin). Let

$$Q = s|0\rangle\langle 0| + (1-s)|1\rangle\langle 1|, \quad (89)$$

be a fixed mixed state of a spin, where $0 \leq s \leq 1/2$. For each $M \subset \{1, 2, \dots, n\}$, let M^c be the complement of M , and let $|M|$ and $|M^c|$ be the sizes of M and M^c , respectively. Consider n spins, and view M as a subset of the spins. We use $\omega_M \otimes Q^{\otimes M^c}$ to denote a state on these n spins where each spin

in M^c is in the state Q , and the spins in M are in a joint state given by the density matrix ω_M . Let $\{x_M\}$ be non-negative numbers such that

$$\sum_M x_M = 1, \quad (90)$$

where the sum is over all possible subsets M of $\{1, \dots, n\}$. For an arbitrary set of such numbers $\{x_M\}$, consider n spins in the state $\kappa = \sum_M x_M \omega_M \otimes Q^{\otimes M^c}$ where ω_M are variables. The goal is to minimize the von Neuman entropy of κ , $S(\kappa) = -\text{Tr}(\kappa \log \kappa)$. Formally, the **entropy minimization problem** is given by

$$\begin{aligned} \min S(\kappa) \\ \kappa &= \sum_M x_M \omega_M \otimes Q^{\otimes M^c}, \\ \omega_M &\geq 0, \\ \text{Tr}(\omega_M) &= 1. \end{aligned} \quad (91)$$

We conjecture that the entropy minimization problem has an optimal solution when all spins align with one another as much as possible and aligned with the eigenstate corresponding to the maximum eigenvalue of Q ; in other words, for all $2^n - 1$ possible non-empty subsets M ,

$$\omega_M = |1\rangle\langle 1|^{\otimes |M|}. \quad (92)$$

Note that the minimum of (91) can be attained on pure states ω_M due to concavity of the von Neumann entropy.

We have not come upon a general proof of this conjecture but we have found proofs for various special cases and numerical evidence in other cases. In what follows we briefly mention this evidence for our conjecture.

A. Special Case $n = 1$

For 1 spin, the entropy minimization problem (91) takes the form

$$\begin{aligned} \min S(\kappa) \\ \kappa &= x_1 \omega_1 + x_\phi Q, \\ \omega_1 &\geq 0, \\ \text{Tr}(\omega_1) &= 1, \end{aligned} \quad (93)$$

where ϕ denotes the null set, the subscript 1 is a shorthand for $M = \{1\}$, and x_1 and x_ϕ are arbitrary but fixed non-negative numbers that sum to 1. The above optimization problem (93), generalized to any dimension and any valid density matrix Q , can be solved using Schur-concavity of the von Neumann entropy along with a majorization inequality, see eq. (II.16) on pg. 35 in [46]. The optimal solution ω_1 can always be chosen to be the projector onto any 1-dimensional eigenspace of Q corresponding to its maximum eigenvalue. This proves our conjecture for $n = 1$.

B. Special Case $n = 2$ and $s = 1/2$

When there are two spins and $s = \frac{1}{2}$, $Q = \frac{I}{2}$, the optimization problem in (91) takes the form

$$\begin{aligned} \min S(\kappa) \\ \kappa &= x_{12}\omega_{12} + x_1\omega_1 \otimes \frac{I}{2} + x_2\frac{I}{2} \otimes \omega_2 + x_\phi\frac{I}{2} \otimes \frac{I}{2} \\ \omega_1 &\geq 0, \text{Tr}(\omega_1) = 1, \\ \omega_2 &\geq 0, \text{Tr}(\omega_2) = 1, \\ \omega_{12} &\geq 0, \text{Tr}(\omega_{12}) = 1, \end{aligned} \quad (94)$$

where x_1, x_2, x_{12} and x_ϕ are arbitrary but fixed non-negative numbers that sum to 1. There is enough symmetry to assume without loss of generality that $\omega_1 = |1\rangle\langle 1|, \omega_2 = |1\rangle\langle 1|$ and using the optimal solution for the high-dimensional generalization of the $n = 1$ case (93), $\omega_{12} = |1\rangle\langle 1| \otimes |1\rangle\langle 1|$ can be shown to be an optimal solution.

We mention in passing that when $n = 2$ but s is arbitrary, if we assume $\omega_1, \omega_2, \omega_{12}$ to be diagonal, we can also show the optimality of the conjectured solution, by a detailed case analysis involving majorization and Schur-concavity of the von Neumann entropy.

C. Numerical Evidence

Our conjecture is backed by numerical evidence that we gathered for the optimization problem (91) with $n = 2, \dots, 6$. We randomly sampled probability distributions $\{x_M\}_M$ (see (90)), restricted the optimization to pure states ω_M (see comment after (92)) in order to reduce the number of free parameters, and used both gradient descent-based and global optimization techniques (particle swarm optimization). In all instances, the optimization converged to the conjectured solution in (92), i.e., with all states ω_M being tensor products of (a rotated version of) the eigenvector of Q corresponding to its largest eigenvalue. Of course, it is possible that such convergence is to a local minimum rather than a global minimum, however we have not found minima with values smaller than the conjectured value.

D. Generalization to Higher Dimensions

The entropy minimization problem (91) can be generalized to higher spins (a spin- $\frac{1}{2}$ particle is a qubit and higher-spin particles correspond to qudits):

$$\begin{aligned} \min S(\kappa) \\ \kappa &= \sum_M x_M \omega_M \otimes \Upsilon^{\otimes M^c}, \\ \omega_M &\geq 0, \\ \text{Tr}(\omega_M) &= 1, \end{aligned} \quad (95)$$

where Υ is a fixed density matrix on a qudit, and $\{x_M\}$ a fixed distribution on the subsets of $\{1, \dots, n\}$. As before, the minimum can be attained on pure states ω_M due to concavity of the von Neumann entropy. Let $|\gamma\rangle$ be an eigenvector of Υ corresponding to a maximum eigenvalue. We conjecture that

the entropy is minimized when all the spins are aligned with one another as much as possible; that is,

$$\omega_M = |\gamma\rangle\langle\gamma|^{\otimes |M|}, \quad (96)$$

for all $2^n - 1$ possible non-empty subsets M .

E. Rényi-2 Entropy Variation of the Conjecture Holds

We discuss a side result that illustrates the intuition behind the spin alignment conjecture. Unless otherwise stated, symbols defined here are used only for this subsection. Consider a variation of the minimization problem (95) wherein the von Neumann entropy is replaced by the Rényi-2 entropy, which according to (68), is given by

$$S_2(\rho) = -\log \text{Tr}(\rho^2). \quad (97)$$

In this case, the minimum is attained by states given by (96), and we will outline the proof in this subsection.

Using the expression for Rényi-2 entropy above, the goal to minimize $S_2(\kappa)$ in (95) is equivalent to maximizing

$$\begin{aligned} \text{Tr}(\kappa^2) \\ &= \sum_{MM'} x_M x_{M'} \text{Tr} \left[\left(\omega_M \otimes \Upsilon^{\otimes M^c} \right) \left(\omega_{M'} \otimes \Upsilon^{\otimes M'^c} \right) \right] \end{aligned} \quad (98)$$

where we have used the expression for κ in (95). Since $x_M x_{M'} \geq 0$ for all M, M' , it suffices to show that the states given by (96) maximize each

$$\text{Tr} \left[\left(\omega_M \otimes \Upsilon^{\otimes M^c} \right) \left(\omega_{M'} \otimes \Upsilon^{\otimes M'^c} \right) \right] \quad (99)$$

First, note that (99) can be maximized on pure $\omega_M, \omega_{M'}$ due to convexity. Second, we can group the n spins into 4 systems, where systems 1, 2, 3, 4 include spins in $M \cap M'^c, M \cap M', M^c \cap M',$ and $M^c \cap M'^c$ respectively. Third, the trace over system 4 contributes a multiplicative constant to (99) so we can focus on the maximization on systems 1, 2, 3. The conjecture can be proved if the following lemma holds.

Lemma 1: Consider a tripartite system with parts 1, 2, and 3, and define a maximization problem,

$$\max_{|\mu\rangle, |\nu\rangle} \text{Tr} [(|\mu\rangle\langle\mu|_{12} \otimes \Theta_3) (Z_1 \otimes |\nu\rangle\langle\nu|_{23})], \quad (100)$$

where the subscripts 1, 2, 3 indicate which systems are acted on by the operators, Θ, Z are positive semidefinite operators, and $|\mu\rangle, |\nu\rangle$ are unit vectors. The problem above attains its maximum at

$$|\mu\rangle_{12} = |\theta\rangle_1 \otimes |\xi\rangle_2, \quad |\nu\rangle_{23} = |\xi\rangle_2 \otimes |\zeta\rangle_3, \quad (101)$$

where $|\theta\rangle$ is any eigenvector of Θ corresponding to its maximum eigenvalue, $|\zeta\rangle$ is any eigenvector of Z corresponding to its maximum eigenvalue, and $|\xi\rangle$ is any vector in system 2.

Proof: We can take the computational bases of systems 1 and 3 to be the eigenbases of Θ and Z respectively, so they are diagonal without loss of generality. Let $\Theta = \sum_j \theta_j |j\rangle\langle j|, Z = \sum_i \zeta_i |i\rangle\langle i|$ be their eigen-decompositions. We can always express $|\mu\rangle = \sum_i a_i |i\rangle_1 \otimes |\alpha_i\rangle_2$ and $|\nu\rangle = \sum_j b_j |\beta_j\rangle_2 \otimes |j\rangle_3$ for some nonnegative amplitudes a_i, b_j and unit vectors

$|\alpha_i\rangle$ and $|\beta_j\rangle$ on system 2, such that $\sum_i a_i^2 = \sum_j b_j^2 = 1$. Then,

$$\begin{aligned} & \text{Tr}[(|\mu\rangle\langle\mu|_{12} \otimes \Theta_3)(Z_1 \otimes |\nu\rangle\langle\nu|_{23})] \\ &= \text{Tr} \left[\left(\sum_{i_1, i_2} a_{i_1} a_{i_2} |i_1\rangle\langle i_2|_1 \otimes |\alpha_{i_1}\rangle\langle\alpha_{i_2}|_2 \otimes \sum_j \theta_j |j\rangle\langle j|_3 \right) \right. \\ & \quad \left. \left(\sum_i \zeta_i |i\rangle\langle i|_1 \otimes \sum_{j_1, j_2} b_{j_1} b_{j_2} |\beta_{j_1}\rangle\langle\beta_{j_2}|_2 \otimes |j_1\rangle\langle j_2|_3 \right) \right] \\ &= \sum_i \sum_j a_i^2 \zeta_i b_j^2 \theta_j |\langle\alpha_i|\beta_j\rangle|^2 \\ &\leq \max_{i,j} \zeta_i \theta_j |\langle\alpha_i|\beta_j\rangle|^2 \\ &\leq \max_{i,j} \zeta_i \theta_j. \end{aligned} \quad (102)$$

In the last line above, we use a convexity argument noting that $\{a_i^2 b_j^2\}_{i,j}$ is a probability distribution. The proposed solution $|\mu\rangle_{12} = |\delta\rangle_1 \otimes |\xi\rangle_2$, $|\nu\rangle_{23} = |\xi\rangle_2 \otimes |\gamma\rangle_3$ attains the upper bound $\max_{i,j} \zeta_i \theta_j$ thus must be an optimal solution for the maximization problem, proving the lemma. ■

VII. QUANTUM CAPACITY BOUNDS

A. Upper Bound on $\mathcal{Q}(\mathcal{N}_s)$

We now derive an analytic upper bound and a tighter numerical upper bound on the quantum capacity of \mathcal{N}_s . The analytic bound matches the SDP bound in Prop. 16 of [45] where it appears without proof. We provide a proof showing the SDP bound matches the well-known ‘‘transposition bound’’, which states that $\mathcal{Q}(\mathcal{B}) \leq \log \|\mathcal{T} \circ \mathcal{B}\|_\diamond$ for any channel \mathcal{B} [48]. Here, $\mathcal{T}: X \mapsto X^T$ denotes the transposition map, taken with respect to the same basis used to define the maximally entangled state $|\phi\rangle$ in (2). For a superoperator $\Psi: \hat{\mathcal{H}} \rightarrow \hat{\mathcal{H}}'$ the diamond norm $\|\Psi\|_\diamond$ is defined as

$$\|\Psi\|_\diamond = \sup\{\|(\mathcal{I}_{\mathcal{H}} \otimes \Psi)(X)\|_1 : X \in \hat{\mathcal{H}} \otimes \hat{\mathcal{H}}, \|X\|_1 \leq 1\}. \quad (103)$$

The diamond norm of any linear superoperator can be computed by a semidefinite program [49], which, for $\mathcal{T} \circ \mathcal{B}$, is given by

$$\begin{aligned} \|\mathcal{T} \circ \mathcal{B}\|_\diamond &= \min. \frac{1}{2} (\|Y_a\|_\infty + \|Z_a\|_\infty) \\ & \text{s.t. } Y_{ab}, Z_{ab} \geq 0 \\ & \quad \begin{pmatrix} Y_{ab} & -\mathcal{T}_b(J_{ab}^{\mathcal{B}}) \\ -\mathcal{T}_b(J_{ab}^{\mathcal{B}}) & Z_{ab} \end{pmatrix} \geq 0 \end{aligned} \quad (104)$$

where $J_{ab}^{\mathcal{B}}$ denotes the unnormalized Choi-Jamiołkowski operator of \mathcal{B} (3), and $\mathcal{T}_b = \mathcal{I}_a \otimes \mathcal{T}$ denotes the partial transpose with respect to system b , i.e., the transpose map \mathcal{T} acts on $\hat{\mathcal{H}}_b$.

We compare the transposition bound to another bound on $\mathcal{Q}(\mathcal{B})$ by Wang et al. [50] defined in terms of a quantity $\Gamma(\mathcal{B})$,

which is the solution of the following semidefinite program:

$$\begin{aligned} \Gamma(\mathcal{B}) &= \max. \text{tr } R_{ab} J_{ab}^{\mathcal{B}} \\ & \text{s.t. } R_{ab}, \rho_a \geq 0 \\ & \quad \text{tr } \rho_a = 1 \\ & \quad -\rho_a \otimes \mathbb{1}_b \leq \mathcal{T}_b(R_{ab}) \leq \rho_a \otimes \mathbb{1}_b \end{aligned} \quad (105)$$

The two bounds on the quantum capacity of \mathcal{B} are related as follows:

Proposition 2 (Holevo, Werner [48], Wang et al. [50]): For any quantum channel \mathcal{B} ,

$$\mathcal{Q}(\mathcal{B}) \leq \log \Gamma(\mathcal{B}) \leq \log \|\mathcal{T} \circ \mathcal{B}\|_\diamond, \quad (106)$$

where $\Gamma(\mathcal{B})$ is defined in (105).

The two upper bounds mentioned above yield an analytical upper bound on the quantum capacity of the channel \mathcal{N}_s . Due to unitary equivalence of the \mathcal{N}_s channel to the one in [45] (see discussion below eq. (20)), the SDP bound $\log \Gamma(\mathcal{N}_s)$ matches the one stated without proof in Prop. 16 of [45].

Theorem 3: For $s \in [0, 1/2]$ we have $\log \Gamma(\mathcal{N}_s) = \log \|\mathcal{T} \circ \mathcal{N}_s\|_\diamond = \log(1 + \sqrt{1-s})$, and hence

$$\mathcal{Q}(\mathcal{N}_s) \leq \log(1 + \sqrt{1-s}). \quad (107)$$

Proof: The theorem is proved by asserting that

$$1 + \sqrt{1-s} \leq \Gamma(\mathcal{N}_s) \leq \|\mathcal{T} \circ \mathcal{N}_s\|_\diamond \leq 1 + \sqrt{1-s}, \quad (108)$$

from which the claim follows via Proposition 2.

To prove the first inequality in (108), we pick the following operators (R_{ab}, ρ_a) in the SDP (105) for $\Gamma(\mathcal{N}_s)$:

$$\rho_a = \frac{1}{2} ([0]_a + [2]_a) \quad (109)$$

$$R_{ab} = \frac{1}{2} ([00]_{ab} + [01]_{ab} + |01\rangle\langle 22|_{ab} + |22\rangle\langle 01|_{ab} + [22]_{ab}). \quad (110)$$

It is easy to check that $R_{ab}, \rho_a \geq 0$ and $\text{tr } \rho_a = 1$, and that $\rho_a \otimes \mathbb{1}_b \pm \mathcal{T}_b(R_{ab}) \geq 0$, which ensures that the pair (R_{ab}, ρ_a) is indeed feasible in (105). To compute the objective value $\text{tr } R_{ab} J_{ab}^s$ with $J_{ab}^s \equiv J_{ab}^{\mathcal{N}_s}$, observe that the J_{ab}^s has the form

$$J_{ab}^s = [\psi_1]_{ab} + [\psi_2]_{ab} c \quad (111)$$

$$|\psi_1\rangle_{ab} = \sqrt{s}|00\rangle_{ab} + |12\rangle_{ab} \quad (112)$$

$$|\psi_2\rangle_{ab} = \sqrt{1-s}|01\rangle + |22\rangle. \quad (113)$$

We have $\langle\psi_1|R_{ab}|\psi_1\rangle_{ab} = s/2$ and $\langle\psi_2|R_{ab}|\psi_2\rangle_{ab} = 1 + \sqrt{1-s} - s/2$, and hence

$$\begin{aligned} \Gamma(\mathcal{N}_s) &\geq \text{tr } R_{ab} J_{ab}^s \\ &= \langle\psi_1|R_{ab}|\psi_1\rangle_{ab} + \langle\psi_2|R_{ab}|\psi_2\rangle_{ab} \\ &= 1 + \sqrt{1-s}. \end{aligned} \quad (114)$$

To prove the third inequality in (108), $\|\mathcal{T} \circ \mathcal{N}_s\|_\diamond \leq 1 + \sqrt{1-s}$, we again pick feasible operators in the SDP (104) for $\|\mathcal{T} \circ \mathcal{N}_s\|_\diamond$:

$$\begin{aligned} Y_{ab} = Z_{ab} &= s[00]_{ab} + (1-s)[01]_{ab} + \sqrt{1-s}[02]_{ab} \\ & \quad + [12]_{ab} + [22]_{ab} + |\phi\rangle_{ab} \end{aligned} \quad (115)$$

$$|\phi\rangle_{ab} = \sqrt{\frac{s^2}{1-s}} |10\rangle_{ab} + \sqrt{1-s} |21\rangle_{ab}. \quad (116)$$

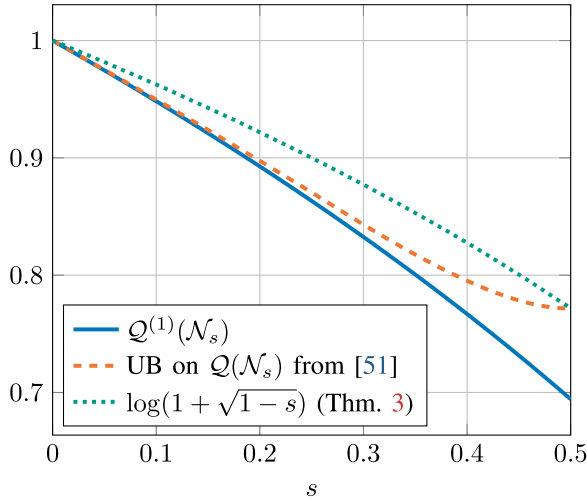


Fig. 1. Lower and upper bounds on the quantum capacity of \mathcal{N}_s . Plotted are the coherent information $\mathcal{Q}^{(1)}(\mathcal{N}_s)$ (solid blue), the upper bound (UB) $\hat{R}_\alpha(\mathcal{N}_s)$ on $\mathcal{Q}(\mathcal{N}_s)$ from [51] with $\alpha = 1 + 2^{-5}$ (dashed orange), and the UB from Theorem 3 (dotted green).

Evidently, $Y_{ab} \geq 0$. Furthermore, the operator $\begin{pmatrix} Y_{ab} & -\mathcal{T}_b(J_{ab}^s) \\ -\mathcal{T}_b(J_{ab}^s) & Y_{ab} \end{pmatrix}$ is unitarily equivalent to the following operator in block-diagonal form:

$$sM \oplus (1-s)M \oplus M \otimes \mathbb{1}_2 \oplus \frac{s}{\sqrt{1-s}}[\phi_1] \oplus \sqrt{1-s}[\phi_2] \quad (117)$$

with the matrix $M = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ and vectors

$$\begin{aligned} |\phi_1\rangle &= -\sqrt{\frac{1-s}{s}}|0\rangle + |1\rangle + \sqrt{\frac{1-s}{s}}|2\rangle \\ |\phi_2\rangle &= \sqrt{\frac{s}{1-s}}|0\rangle + |1\rangle - |2\rangle. \end{aligned} \quad (118)$$

As the operator in (117) is manifestly positive semidefinite, the same holds for

$$\begin{pmatrix} Y_{ab} & -\mathcal{T}_b(J_{ab}^s) \\ -\mathcal{T}_b(J_{ab}^s) & Y_{ab} \end{pmatrix}, \quad (119)$$

showing that Y_{ab} and $Z_{ab} = Y_{ab}$ are feasible in (104). The marginal $Y_a = \text{tr}_b Y_{ab}$ is diagonal with eigenvalues $1 + \sqrt{1-s}$ (of multiplicity 2) and $1 + s/\sqrt{1-s}$. Since $\sqrt{1-s} \geq s/\sqrt{1-s}$ for $s \in [0, 1/2]$, we conclude that $\|Y_a\|_\infty = 1 + \sqrt{1-s}$. Therefore, the SDP (104) for $\|\mathcal{T} \circ \mathcal{N}_s\|_\diamond$ has value at most $\|Y_a\|_\infty = 1 + \sqrt{1-s}$, which concludes the proof of the theorem. ■

While the bound $\log \Gamma(\mathcal{B})$ can be strictly tighter than the transposition bound $\log \|\mathcal{T} \circ \mathcal{B}\|_\diamond$ for certain channels \mathcal{B} [50], Theorem 2 shows that the two bounds in fact coincide for \mathcal{N}_s . The SDP upper bound $\log \Gamma(\mathcal{B})$ was recently improved by Fawzi and Fawzi [51], and evaluating the latter bound (which can again be computed by semidefinite programming) yields an even tighter bound on $\mathcal{Q}(\mathcal{N}_s)$. The two bounds are compared in Fig. 1.

B. Upper Bound on $\mathcal{Q}(\mathcal{M}_d)$

We now derive an analytical upper bound on the quantum capacity of \mathcal{M}_d , which we introduced in Section III-B as a

generalization of $\mathcal{N}_{1/2}$ to arbitrary dimension d of the input and output Hilbert spaces. As a reminder, a channel isometry for \mathcal{M}_d is given by $G: \mathcal{H}_a \rightarrow \mathcal{H}_b \otimes \mathcal{H}_c$ with $\mathcal{H}_a \cong \mathcal{H}_b \cong \mathbb{C}^d$ and $\mathcal{H}_c \cong \mathbb{C}^{d-1}$, defined via the following action on the computational basis $\{|j\rangle_a\}_{j=0}^{d-1}$ of \mathcal{H}_a :

$$\begin{aligned} G: |0\rangle_a &\mapsto \frac{1}{\sqrt{d-1}} \sum_{j=0}^{d-2} |j\rangle_b |j\rangle_c \\ |j\rangle_a &\mapsto |d-1\rangle_b |j-1\rangle_c \quad \text{for } j = 1, \dots, d-1. \end{aligned} \quad (120)$$

The Choi-Jamiołkowski operator $J_{ab}^{\mathcal{M}_d}$ of \mathcal{M}_d is given as follows:

$$\begin{aligned} J_{ab}^{\mathcal{M}_d} &= \sum_{j=0}^{d-2} \left(\frac{1}{d-1} [0j] + [j+1, d-1] + \right. \\ &\quad \left. \frac{1}{\sqrt{d-1}} (|0, j\rangle\langle j+1, d-1| + |j+1, d-1\rangle\langle 0, j|) \right), \end{aligned} \quad (121)$$

where we used the notation $[\psi] \equiv |\psi\rangle\langle\psi|$.

Using Proposition 2 and the SDP (104), we now derive an upper bound on the quantum capacity of the channel \mathcal{M}_d .

Theorem 4: For any $d \geq 2$,

$$\mathcal{Q}(\mathcal{M}_d) \leq \log \left(1 + \frac{1}{\sqrt{d-1}} \right). \quad (122)$$

In particular, $\mathcal{Q}(\mathcal{M}_d) \rightarrow 0$ as $d \rightarrow \infty$.

Proof: We will prove this theorem by constructing feasible operators $Y_{ab} = Z_{ab}$ in the SDP (104) for $\|\mathcal{T} \circ \mathcal{M}_d\|_\diamond$ satisfying $\|\text{tr}_b Y_{ab}\|_\infty = 1 + \frac{1}{\sqrt{d-1}}$, from which the claim follows by invoking Proposition 2.

Consider the following ansatz for Y_{ab} :

$$\begin{aligned} Y_{ab} &= \frac{1}{d-1} \sum_{j=0}^{d-2} [0] \otimes [j] + \frac{1}{\sqrt{d-1}} [0] \otimes [d-1] \\ &\quad + \sum_{j=1}^{d-1} [j] \otimes [d-1] + [\Xi], \end{aligned} \quad (123)$$

$$\text{with } |\Xi\rangle = \frac{1}{\sqrt{d-1}} \sum_{j=1}^{d-1} |j\rangle |j-1\rangle. \quad (124)$$

We clearly have $Y_{ab} \geq 0$. Setting $Z_{ab} = Y_{ab}$, consider the second feasibility constraint in (104),

$$\begin{pmatrix} Y_{ab} & -\mathcal{T}_b(J_{ab}^{\mathcal{M}_d}) \\ -\mathcal{T}_b(J_{ab}^{\mathcal{M}_d}) & Y_{ab} \end{pmatrix} \geq 0. \quad (125)$$

Taking the Schur complement, this is equivalent to the constraint

$$Y_{ab} \geq \mathcal{T}_b(J_{ab}^{\mathcal{M}_d}) Y_{ab}^{-1} \mathcal{T}_b(J_{ab}^{\mathcal{M}_d}) \quad (126)$$

with the inverse taken on the support of Y_{ab} . Using the form of $J_{ab}^{\mathcal{M}_d}$ in (121), it is straightforward to check that $Y_{ab} = \mathcal{T}_b(J_{ab}^{\mathcal{M}_d}) Y_{ab}^{-1} \mathcal{T}_b(J_{ab}^{\mathcal{M}_d})$ so that (126) is satisfied. Finally, it follows from (123) that

$$Y_a = \text{tr}_b Y_{ab} = \left(1 + \frac{1}{\sqrt{d-1}} \right) \mathbb{1}_a, \quad (127)$$

which yields $\|Y_a\|_\infty = 1 + \frac{1}{\sqrt{d-1}}$ and concludes the proof. ■

VIII. PRIVATE AND CLASSICAL CAPACITIES

The private and classical capacities of the channels \mathcal{N}_s and \mathcal{M}_d can be determined exactly. In the following, we will prove in Theorems 6 and 9 that $\mathcal{P}(\mathcal{N}_s) = \mathcal{C}(\mathcal{N}_s) = 1 = \mathcal{P}(\mathcal{M}_d) = \mathcal{C}(\mathcal{M}_d)$. This is remarkable because neither \mathcal{N}_s nor \mathcal{M}_d belong to any of the special classes of channels for which the private or classical capacity is known to have a single-letter expression (that is, equal to the channel private information or the Holevo information; see Section II for the definitions of these quantities).

The upper bound on the private and classical capacities of the channels employ the SDP technique used in prior work of Wang et al. [45]. That same work applied the technique to a qutrit-to-qutrit channel unitarily equivalent to \mathcal{N}_s (see discussion below eq. (20)). Hence, these prior upper and lower bounds on \mathcal{C} (see Prop. 15 in [45]) and \mathcal{C}_E (see Prop. 1 in [44]) imply our bounds on $\mathcal{C}(\mathcal{N}_s)$ in Th. 6 and $\mathcal{C}_E(\mathcal{N}_s)$ in Th. 8. On the other hand, our lower bound on \mathcal{P} , first inequality in (129) below, correctly proves the bound previously stated in [45, Prop. 16], whose proof contained a typo. The corresponding results for the \mathcal{M}_d channel do not follow from these prior works.

A. Capacities of \mathcal{N}_s

We will obtain a lower bound for the channel private information and the Holevo information using the following equiprobable ensemble of two quantum states:

$$\rho_a^1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \rho_a^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 1-s \end{pmatrix}. \quad (128)$$

For this ensemble, the quantity $\Delta(\mathcal{B}, \bar{\rho}_a) - \sum_x p_x \Delta(\mathcal{B}, \rho_a^x)$ evaluates to 1 which gives a lower bound to the channel private information $\mathcal{P}^{(1)}(\mathcal{N}_s) \geq 1$ for any $s \in [0, 1/2]$ (see (8)). Likewise, the Holevo information is 1 for this ensemble, giving a lower bound $1 \leq \chi(\mathcal{N}_s)$ by (16). Using [20], [31] and (15), we obtain chains of inequalities

$$\begin{aligned} 1 &\leq \mathcal{P}^{(1)}(\mathcal{N}_s) \leq \mathcal{P}(\mathcal{N}_s) \leq \mathcal{C}(\mathcal{N}_s), \\ 1 &\leq \mathcal{P}^{(1)}(\mathcal{N}_s) \leq \chi(\mathcal{N}_s) \leq \mathcal{C}(\mathcal{N}_s). \end{aligned} \quad (129)$$

We will now show that $\mathcal{C}(\mathcal{N}_s) \leq 1$ so we have equalities throughout the above. To this end, we employ a semidefinite programming upper bound on the classical capacity derived by Wang et al. [45]:

Proposition 5 ([45]): For any quantum channel \mathcal{B} ,

$$\mathcal{C}(\mathcal{B}) \leq \log \beta(\mathcal{B}), \quad (130)$$

where $\beta(\mathcal{B})$ is the solution of the following SDP:

$$\begin{aligned} \beta(\mathcal{B}) &= \min. \quad \text{tr } S_b \\ &\text{s.t. } R_{ab}, S_b \text{ Hermitian} \\ &\quad -R_{ab} \leq \mathcal{T}_b(J_{ab}^{\mathcal{B}}) \leq R_{ab} \\ &\quad -\mathbb{1}_a \otimes S_b \leq \mathcal{T}_b(R_{ab}) \leq \mathbb{1}_a \otimes S_b \end{aligned} \quad (131)$$

Moreover, $\log \beta(\mathcal{B})$ is a strong converse bound: if the classical information transmission rate exceeds $\log \beta(\mathcal{B})$, the transmission error converges to 1 exponentially fast.

We now state and prove the main result of this section:

Theorem 6: For all $s \in [0, 1/2]$,

$$\mathcal{P}(\mathcal{N}_s) = \mathcal{C}(\mathcal{N}_s) = 1. \quad (132)$$

Moreover, both the classical and the private capacity of \mathcal{N}_s satisfy the strong converse property.

Proof: Due to (129) and Proposition 5, it suffices to show that $\beta(\mathcal{N}_s) \leq 2$ for all $s \in [0, 1/2]$. To this end, let

$$\begin{aligned} R_{ab} &= s[00]_{ab} + (1-s)[01]_{ab} + [02]_{ab} \\ &\quad + [\psi]_{ab} + [12]_{ab} + [22]_{ab} \\ S_b &= s[0]_b + (1-s)[1]_b + [2]_b, \end{aligned} \quad (133)$$

where $|\psi\rangle_{ab} = \sqrt{s}|10\rangle + \sqrt{1-s}|21\rangle$. We now check that (R_{ab}, S_b) is feasible for the SDP (131). Both R_{ab} and S_b are Hermitian by construction, and one readily checks that $\mathbb{1}_a \otimes S_b \pm \mathcal{T}_b(R_{ab}) \geq 0$. Recalling the form of $J_{ab}^s \equiv J_{ab}^{\mathcal{N}_s}$ from (111), it also follows that $R_{ab} \pm \mathcal{T}_b(J_{ab}^s) \geq 0$. These observations establish feasibility of (R_{ab}, S_b) in (131). Furthermore $\text{tr } S_b = 2$, hence $\beta(\mathcal{N}_s) \leq 2$. (As a side remark, we note that (R_{ab}, S_b) is in fact optimal for (131) since $1 \leq \mathcal{C}(\mathcal{N}_s) \leq \log \beta(\mathcal{N}_s) \leq 1$ by (129) and Proposition 5.)

Since $\log \beta(\mathcal{N}_s)$ is a strong converse bound and $\mathcal{C}(\mathcal{N}_s) = \log \beta(\mathcal{N}_s) = 1$, the classical capacity of \mathcal{N}_s satisfies the strong converse property. This also holds for the private capacity $\mathcal{P}(\mathcal{N}_s) = \mathcal{C}(\mathcal{N}_s) = 1$, which can be seen as follows (we refer to [32] for precise definitions). Let $\varepsilon_{\mathcal{C}}$ denote the error for a classical information transmission code. The error $\varepsilon_{\mathcal{P}}$ for a private information transmission code is defined as $\varepsilon_{\mathcal{P}} = \max(\varepsilon_{\mathcal{C}}, \varepsilon_{\text{env}})$, where ε_{env} is an additional error parameter controlling how much information the environment gains about Alice's input. Assume now that we have a private information transmission code with rate $r_{\mathcal{P}} > 1$ and error $\varepsilon_{\mathcal{P}}$. Then this code can also be regarded as a classical information transmission code with rate $r_{\mathcal{C}} = r_{\mathcal{P}} > 1$ and error $\varepsilon_{\mathcal{C}} \leq \varepsilon_{\mathcal{P}}$. The strong converse property of $\mathcal{C}(\mathcal{N}_s)$ implies that $\varepsilon_{\mathcal{C}} \rightarrow 1$ as the code blocklength n increases, from which $\varepsilon_{\mathcal{P}} \rightarrow 1$ follows. Hence, $\mathcal{P}(\mathcal{N}_s)$ also satisfies the strong converse property, which concludes the proof. ■

Theorems 3 and 6 imply that the quantum and private capacities of \mathcal{N}_s are strictly separated:

Corollary 7: $\mathcal{Q}(\mathcal{N}_s) < \mathcal{P}(\mathcal{N}_s)$ for all $s \in (0, 1/2]$.

It was recently shown in [52] that the quantity $\log \beta(\cdot)$ from Proposition 5 also serves as an upper bound on the classical capacity of a quantum channel assisted by a classical feedback channel, denoted $\mathcal{C}_{\leftarrow}(\cdot)$. Hence, Theorem 6 immediately implies that the feedback-assisted classical capacity of \mathcal{N}_s is equal to its classical capacity, $\mathcal{C}(\mathcal{N}_s) = 1 = \mathcal{C}_{\leftarrow}(\mathcal{N}_s)$ for all $s \in [0, 1/2]$. Moreover, the same argument applies to the private capacity of a quantum channel assisted by a public feedback channel.

Finally, we discuss the entanglement-assisted capacity of the channel \mathcal{N}_s . Similar to the classical capacity, it is independent of the parameter s :

Theorem 8: For all $s \in [0, 1/2]$,

$$\mathcal{C}_E(\mathcal{N}_s) = 2. \quad (134)$$

Proof: It is easy to check that the state $\sigma_{ab} = (\mathcal{I}_a \otimes \mathcal{N}_s)(|\psi_{aa'}\rangle)$ with

$$|\psi\rangle_{aa'} = \frac{1}{\sqrt{2}}|00\rangle_{aa'} + \frac{1}{\sqrt{2}}(\sqrt{s}|11\rangle_{aa'} + \sqrt{1-s}|22\rangle_{aa'}) \quad (135)$$

achieves $I(A; B)_\sigma = 2$. The optimality of $|\psi_{aa'}\rangle$ can be verified using the methods of [36]. ■

B. Capacities of \mathcal{M}_d

We proved in Theorem 6 in Section VIII-A that one use of the channel \mathcal{N}_s can faithfully transmit 1 private bit (and thus also 1 classical bit) regardless of the value of s . We now prove that the d -dimensional generalization \mathcal{M}_d of $\mathcal{N}_{1/2}$ defined in (120) retains unit private and classical capacity for any local dimension d :

Theorem 9: For all $d \geq 2$,

$$\mathcal{P}(\mathcal{M}_d) = \mathcal{C}(\mathcal{M}_d) = 1. \quad (136)$$

Moreover, both the classical and private capacity of \mathcal{M}_d satisfy the strong converse property.

Proof: The proof strategy is similar to the one used in Theorem 6. First, consider an equiprobable ensemble with the following two quantum states,

$$\rho_a^0 = [0]_a, \quad \rho_a^1 = \frac{1}{d-1} \sum_{j=1}^{d-1} [j]_a, \quad (137)$$

and form the cqg state

$$\rho_{xbc} = (\mathcal{I}_x \otimes G)\rho_{xa}(\mathcal{I}_x \otimes G)^\dagger, \quad (138)$$

where $G: \mathcal{H}_a \rightarrow \mathcal{H}_b \otimes \mathcal{H}_c$ is the channel isometry for \mathcal{M}_d defined in (120), and $\rho_{xa} = \frac{1}{2}([0]_x \otimes \rho_a^0 + [1]_x \otimes \rho_a^1)$. It is straightforward to check that $I(X; B)_\sigma = 1$ and $I(X; E)_\sigma = 0$, from which we obtain

$$1 \leq \mathcal{P}^{(1)}(\mathcal{M}_d) \leq \mathcal{P}(\mathcal{M}_d) \leq \mathcal{C}(\mathcal{M}_d). \quad (139)$$

The claim of the theorem now follows by showing that $\mathcal{C}(\mathcal{M}_d) \leq 1$ for all d . To this end, we once again employ the upper bound $\log \beta(\mathcal{M}_d)$ from Proposition 5. Consider the following Hermitian operators R_{ab} and S_b (analogous to the operators in (133)):

$$R_{ab} = \frac{1}{d-1} \sum_{j=0}^{d-2} [0]_a \otimes [j]_b + \sum_{j=0}^{d-1} [j]_a \otimes [d-1] + [\psi]_{ab},$$

$$S_b = \frac{1}{d-1} \sum_{j=0}^{d-2} [j]_a + [d-1]_a, \quad (140)$$

where $|\psi\rangle_{ab} = \frac{1}{\sqrt{d-1}} \sum_{j=1}^{d-1} |j\rangle_a \otimes |j-1\rangle_b$. One readily checks that R_{ab} and S_b are feasible in the SDP (131), that is, $R_{ab} \pm \mathcal{T}_b(J_{ab}^{\mathcal{M}_d}) \geq 0$ and $\mathbb{1}_a \otimes S_b \pm \mathcal{T}_b(R_{ab}) \geq 0$. Furthermore, $\text{tr} S_b = 2$ for any d , and hence $\beta(\mathcal{M}_d) \leq 2$. Using Proposition 5, we conclude $\mathcal{C}(\mathcal{M}_d) \leq \log \beta(\mathcal{M}_d) \leq 1$, which together with (139) gives

$$1 \leq \mathcal{P}(\mathcal{M}_d) \leq \mathcal{C}(\mathcal{M}_d) \leq 1. \quad (141)$$

The strong converse property for $\mathcal{C}(\mathcal{M}_d)$ and $\mathcal{P}(\mathcal{M}_d)$ follows in the same way as in the proof of Theorem 6. ■

By the same argument as in the remark after Corollary 7, Theorem 9 implies that the feedback-assisted private and classical capacities of \mathcal{M}_d are equal to their unassisted counterparts.

IX. DISCUSSION OF CAPACITIES OF THE PLATYPUS CHANNELS

We now summarize the findings of Sections IV, VII and VIII on the quantum capacity \mathcal{Q} , private capacity \mathcal{P} and classical capacity \mathcal{C} of the platypus channels \mathcal{N}_s (with $s \in (0, 1/2)$) and \mathcal{M}_d (for $d \geq 3$):

$$0 < \mathcal{Q}^{(1)}(\mathcal{N}_s) \stackrel{?}{=} \mathcal{Q}(\mathcal{N}_s) \leq \log(1 + \sqrt{1-s})$$

$$< 1 = \mathcal{P}^{(1)}(\mathcal{N}_s) = \mathcal{P}(\mathcal{N}_s) = \chi(\mathcal{N}_s) = \mathcal{C}(\mathcal{N}_s), \quad (142)$$

$$0 < \mathcal{Q}^{(1)}(\mathcal{M}_d) \stackrel{?}{=} \mathcal{Q}(\mathcal{M}_d) \leq \log\left(1 + \frac{1}{\sqrt{d-1}}\right)$$

$$< 1 = \mathcal{P}^{(1)}(\mathcal{M}_d) = \mathcal{P}(\mathcal{M}_d) = \chi(\mathcal{M}_d) = \mathcal{C}(\mathcal{M}_d). \quad (143)$$

In the above eq. 142, the left-most equality labeled by “?” is the conjectured weak additivity of the single-letter coherent information, $\mathcal{Q}^{(1)}(\mathcal{N}_s)$, which would be implied by the validity of the “spin alignment conjecture” described in Section VI. The next inequality is Theorem 3. Finally, the four equalities on the RHS of (142) come from Theorem 6 and (129). Likewise, in eq. 143 the conjectured equality labeled by “?” would be implied by the validity of the (higher-dimensional version of the) spin-alignment conjecture in Section VI-D, and the following inequality and equalities are obtained via Theorems 4 and 9, respectively. For both \mathcal{N}_s and \mathcal{M}_d , the private and classical capacity have the strong converse property, as proved in Theorems 6 and 9, respectively. These findings are remarkable for various reasons:

- The private information $\mathcal{P}^{(1)}(\cdot)$ is additive for both \mathcal{N}_s and \mathcal{M}_d . The only known classes of quantum channels with additive private information are (a) “less noisy channels” \mathcal{B} whose complementary channels have vanishing private capacity, $\mathcal{P}(\mathcal{B}^c) = 0$ [43] and of which degradable channels [10], [11] are special cases; (b) anti-degradable channels; and (c) direct sums of partial traces (DSPT), a special case of the ternary ring of operators (TRO) channels [53]. We know from Section III that $\mathcal{P}(\mathcal{N}_s^c) = 1$ for all $s \in [0, 1/2]$ and $\mathcal{P}(\mathcal{M}_d^c) = \log(d-1)$ for $d \geq 3$. Hence, neither \mathcal{N}_s nor \mathcal{M}_d are less noisy. Moreover, clearly neither of these channels is a direct sum of partial traces, so that both \mathcal{N}_s and \mathcal{M}_d fall outside all known classes of channels with additive private information.
- The Holevo information $\chi(\cdot)$ is additive for both \mathcal{N}_s and \mathcal{M}_d . Again, both channels fall outside of all the known classes of channels with additive Holevo information: (a) entanglement-breaking channels [14]; (b) unital qubit channels [12]; (c) depolarizing channels [13]; (d) Hadamard channels [15], [54]; (e) DSPT channels [53]; and (f) erasure channels [55]. Since entanglement-breaking channels have vanishing quantum

capacity and both \mathcal{N}_s and \mathcal{M}_d have positive quantum capacity, the platypus channels are not entanglement-breaking. They clearly do not belong to classes (b) and (c) either. A quantum channel is Hadamard if its complementary channel is entanglement-breaking [32]. Since $Q(\mathcal{N}_s^c) = 1$ and $Q(\mathcal{M}_d^c) = \log(d - 1)$, the complements of \mathcal{N}_s and \mathcal{M}_d cannot be entanglement-breaking, so that neither channel is Hadamard. Finally, neither \mathcal{N}_s nor \mathcal{M}_d are a DSPT nor an erasure channel.

- The quantum capacity of both \mathcal{N}_s and \mathcal{M}_d is strictly smaller than their respective private capacities, for all $s \in (0, 1/2]$ and $d \geq 3$. There are not too many examples of this phenomenon. The first known class is the Horodecki channels, for which the quantum capacity vanishes and the private capacity is strictly positive [8], [56], [57]. The smallest such example has input and output dimensions $d_a = d_b = 3, d_c = 4$, and the separation is typically small. The second class is the so-called “half-rocket channels,” with quantum capacity between 0.6 and 1 but private capacity $\log d$ where the input and output dimensions are $d_a = d^2, d_b = d_c = d^6 - d^4$. This class exhibits an extensive separation of the two capacities [9]. In comparison, \mathcal{N}_s is the smallest known channel with $d_a = d_b = 3, d_c = 2$ exhibiting the separation, and the separation is quite large (at least $\approx s/2$ for most s of interest). A separation of the information quantities $\mathcal{Q}^{(1)}(\cdot)$ and $\mathcal{P}^{(1)}(\cdot)$ was observed for certain channels for which these quantities are also superadditive, such as the depolarizing channel [3], [4] or the dephrasing channel [25]. However, due to superadditivity in these channels we do not know their exact capacities and the true separations between them.
- Reference [8] shows that a quantum state can yield a bit of classical information that is private from the environment E if and only if it is of the form

$$\gamma_{K_A K_B S_A S_B} = U([\phi]_{K_A K_B} \otimes \sigma_{S_A S_B}) U^\dagger \quad (144)$$

where $K_A K_B$ are called the key systems, $S_A S_B$ are called the shield systems, $|\phi\rangle$ is the maximally entangled state on $K_A K_B$, σ is an arbitrary state on $S_A S_B$ and U is a controlled unitary of the form $\sum_{i,j} [ij]_{K_A K_B} \otimes (U_{ij})_{S_A S_B}$, with each U_{ij} a unitary that depends on i and j . The key is shared between two users Alice and Bob. Alice is in possession of $K_A S_A$, Bob is in possession of $K_B S_B$, and they can generate a key by measuring along the computational basis of $K_A K_B$ independently. Furthermore, if the one-way distillable key of the state is strictly greater than the one-way distillable entanglement, each of the shield systems must be nontrivial.

Since \mathcal{N}_s has 3 dimensional input and output and can send one bit privately with a single use, it can be used to make a 3×3 dimensional state shared by Alice and Bob that encodes one private bit. Furthermore, this state is not of the form Eq. (144). To see this, suppose the contrary. By the quantum capacity bound of \mathcal{N}_s , this state has one-way distillable entanglement strictly less than 1, so each of the shield systems must be nontrivial with at least 2 dimensions. Meanwhile, the key systems

have 4 dimensions jointly, so, the total dimension exceeds 9 which is a contradiction.

The resolution is that this state is locally equivalent to a standard p-bit with each of K_A, K_B, S_A, S_B being a qubit, but for which the local ranks of both $K_A S_A$ and $K_B S_B$ are 3. So, our state is a p-bit, but one which has been embedded into smaller dimensional spaces than would be possible generically.

In more detail, here is the protocol to create the 3×3 state which distributes a private key between Alice and Bob. Alice prepares the state

$$\frac{1}{\sqrt{2}} |0\rangle_A |0\rangle_{\tilde{A}} + \frac{\sqrt{s}}{\sqrt{2}} |1\rangle_A |1\rangle_{\tilde{A}} + \frac{\sqrt{1-s}}{\sqrt{2}} |2\rangle_A |2\rangle_{\tilde{A}} \quad (145)$$

and applies F_S (the isometry giving rise to \mathcal{N}_s) to \tilde{A} resulting in the state

$$|\nu\rangle_{ABE} = \frac{1}{\sqrt{2}} |0\rangle_A (\sqrt{s}|0\rangle_B |0\rangle_E + \sqrt{1-s}|1\rangle_B |1\rangle_E) + \frac{\sqrt{s}}{\sqrt{2}} |1\rangle_A |2\rangle_B |0\rangle_E + \frac{\sqrt{1-s}}{\sqrt{2}} |2\rangle_A |2\rangle_B |1\rangle_E. \quad (146)$$

The systems A, B are neither key systems nor shield systems. Consider the local isometries:

$$\begin{aligned} |0\rangle_A &\rightarrow |0\rangle_{K_A} |0\rangle_{S_A}, & |0\rangle_B &\rightarrow |0\rangle_{K_B} |0\rangle_{S_B}, \\ |1\rangle_A &\rightarrow |1\rangle_{K_A} |0\rangle_{S_A}, & |1\rangle_B &\rightarrow |0\rangle_{K_B} |1\rangle_{S_B}, \\ |2\rangle_A &\rightarrow |1\rangle_{K_A} |1\rangle_{S_A}, & |2\rangle_B &\rightarrow |1\rangle_{K_B} |0\rangle_{S_B}. \end{aligned} \quad (147)$$

Applying the above local isometries to $|\nu\rangle_{ABE}$ results in

$$\begin{aligned} &\frac{1}{\sqrt{2}} |0\rangle_{K_A} |0\rangle_{S_A} (\sqrt{s}|0\rangle_{K_B} |0\rangle_{S_B} |0\rangle_E \\ &\quad + \sqrt{1-s}|0\rangle_{K_B} |1\rangle_{S_B} |1\rangle_E) \\ &\quad + \frac{\sqrt{s}}{\sqrt{2}} |1\rangle_{K_A} |0\rangle_{S_A} |1\rangle_{K_B} |0\rangle_{S_B} |0\rangle_E \\ &\quad + \frac{\sqrt{1-s}}{\sqrt{2}} |1\rangle_{K_A} |1\rangle_{S_A} |1\rangle_{K_B} |0\rangle_{S_B} |1\rangle_E \\ &= \frac{1}{\sqrt{2}} |0\rangle_{K_A} |0\rangle_{K_B} (\sqrt{s}|0\rangle_{S_A} |0\rangle_{S_B} |0\rangle_E \\ &\quad + \sqrt{1-s}|0\rangle_{S_A} |1\rangle_{S_B} |1\rangle_E) \\ &\quad + \frac{1}{\sqrt{2}} |1\rangle_{K_A} |1\rangle_{K_B} (\sqrt{s}|0\rangle_{S_A} |0\rangle_{S_B} |0\rangle_E \\ &\quad + \sqrt{1-s}|1\rangle_{S_A} |0\rangle_{S_B} |1\rangle_E) \end{aligned} \quad (148)$$

If we trace out E from the above, we get a state of the form Eq. (144) where $U_{00} = U_{01} = U_{10} = I$, U_{11} is the swap operator, and $\sigma = |00\rangle\langle 00| + |01\rangle\langle 01|$.

- Both the private and the classical capacity of \mathcal{N}_s and \mathcal{M}_d satisfy the strong converse property. For the private capacity, the strong converse property is only known for (a) a subclass of degradable channels called generalized dephasing channels [58] (and for these channels, the quantum and private capacities coincide, $\mathcal{Q} = \mathcal{P}$); (b) DSPT channels [53]. Both \mathcal{N}_s and \mathcal{M}_d are provably non-degradable and not of DSPT form, and hence fall outside both classes. For the classical capacity, the strong

converse is known for a number of channel classes: (a) erasure channels [59]; (b) depolarizing channels, unital qubit channels, and the Holevo-Werner channel [60]; (c) entanglement-breaking and Hadamard channels [61]; (d) DSPT channels [53]. By the arguments made above, neither \mathcal{N}_s nor \mathcal{M}_d belong to any of these classes.

- Finally, the coherent information $\mathcal{Q}^{(1)}(\cdot)$ is additive for both \mathcal{N}_s and \mathcal{M}_d relative to the corresponding version of the spin alignment conjecture. The known classes of channels with additive coherent information are (a) less noisy channels [43] (which includes degradable channels [10]); (b) anti-degradable channels; (c) PPT channels [7]; (d) DSPT channels [53]. The channel \mathcal{N}_s is neither degradable nor anti-degradable. Since PPT channels have vanishing quantum capacity, \mathcal{N}_s and \mathcal{M}_d cannot be PPT either.
- The \mathcal{N}_s channel is unitarily equivalent to a qutrit-qutrit channel \mathcal{L}_α that was introduced in [44] to study zero-error capacities. In the follow-up work [45], the authors showed that the private and classical capacity of \mathcal{L}_α coincide (see Sec. VIII for a more detailed discussion) and satisfy the strong converse property, also noting that \mathcal{L}_α (and hence also \mathcal{N}_s) does not belong to any of the known classes of channels with that property for the private or classical capacity listed above. They further announced (without proof) an analytical upper bound on the quantum capacity of \mathcal{L}_α separating it from the private capacity. This bound coincides with our upper bound, for which we give a full proof in Thm. 3. In our independent study we construct the related channel \mathcal{N}_s as a hybrid of two simple channels (see Sec. III-A) and analyze in detail the additivity properties of the various information quantities of \mathcal{N}_s . Furthermore, we extend the channel construction to a larger family of channels of arbitrary dimension with similar information-theoretic properties (see Sec. III-B and III-C). In the process, we also give full proofs of some of the statements announced in [45] about the capacities of the unitarily equivalent channel \mathcal{L}_α .

X. A 3-PARAMETER GENERALIZATION OF F_s

We now discuss generalizations of the channel \mathcal{N}_s and their capacities, to further our understanding of the phenomena exhibited by \mathcal{N}_s . The one-parameter isometry $F_s : \mathcal{H}_a \mapsto \mathcal{H}_b \otimes \mathcal{H}_c$ in (20) for $0 \leq s \leq 1$ has input dimension $d_a = 3$, output dimension $d_b = 3$ and environment dimension $d_c = 2$. The isometry F_s can be generalized by adding two additional parameters, $\mu, \nu \in [0, 1]$ without changing the input, output, and environment spaces, leading to an isometry $V_{s,\mu,\nu} : \mathcal{H}_a \mapsto \mathcal{H}_b \otimes \mathcal{H}_c$ that acts as

$$\begin{aligned} V_{s,\mu,\nu} |0\rangle &= \sqrt{s} |0\rangle_b \otimes |0\rangle_c + \sqrt{1-s} |1\rangle_b \otimes |1\rangle_c, \\ V_{s,\mu,\nu} |1\rangle &= \sqrt{\nu} |1\rangle_b \otimes |0\rangle_c + \sqrt{1-\nu} |2\rangle_b \otimes |1\rangle_c, \\ V_{s,\mu,\nu} |2\rangle &= \sqrt{\mu} |2\rangle_b \otimes |0\rangle_c + \sqrt{1-\mu} |0\rangle_b \otimes |1\rangle_c. \end{aligned} \quad (149)$$

We use $\mathcal{W}_{s,\mu,\nu}$ to denote the resulting channel from $\hat{\mathcal{H}}_a$ to $\hat{\mathcal{H}}_b$. The isometry $V_{s,\mu,\nu}$ becomes F_s (20) when $\nu = \mu = 1$, so $V_{s,\mu,\nu}$ and $\mathcal{W}_{s,\mu,\nu}$ indeed generalize F_s and \mathcal{N}_s respectively.

We study the degradability of the channel $\mathcal{W}_{s,\mu,\nu}$ using the framework of [62]. We call a channel *pcubed* if it is generated by a pcubed isometry. We call an isometry pcubed if there exists a basis of the input space that is mapped by the isometry to product states of the output space and the environment space. This special basis of the input space is not required to be orthogonal. It is straightforward to verify that, when $0 < s, \mu, \nu < 1$, the isometry $V_{s,\mu,\nu}$ is a *pcubed* isometry:

$$V_{s,\mu,\nu} |\alpha_i\rangle = |\beta_i\rangle \otimes |\gamma_i\rangle \quad \text{for } 0 \leq i \leq 2, \quad (150)$$

where

$$\begin{aligned} |\alpha_i\rangle &= a_0(|0\rangle + \omega^i k_1 |1\rangle + \omega^{2i} k_2 |2\rangle), \\ |\beta_i\rangle &= b_0(|0\rangle + \omega^i l_1 |1\rangle + \omega^{2i} l_2 |2\rangle), \\ |\gamma_i\rangle &= c_0(|0\rangle + \omega^{2i} r |2\rangle), \end{aligned} \quad (151)$$

a_0, b_0 , and c_0 are constants that respectively normalize $|\alpha_i\rangle, |\beta_i\rangle, |\gamma_i\rangle$, ω is a cube root of unity, and k_1, k_2, l_1, l_2 and r are non-negative numbers related to s, μ, ν as follows:

$$\begin{aligned} k_1 &= \frac{1}{r} \sqrt{\frac{1-s}{\nu}}, \quad l_1 = \frac{1}{r} \sqrt{\frac{1-s}{s}} \\ k_2 &= r \sqrt{\frac{s}{1-\mu}}, \quad l_2 = r \sqrt{\frac{\mu}{1-\mu}} \\ r &= \left(\frac{(1-s)(1-\mu)(1-\nu)}{s\mu\nu} \right)^{1/6}. \end{aligned} \quad (152)$$

The inner products between the states witnessing the pcubed channel also characterize its degradability. To this end, let A, B, C be the Gram matrices for the sets $\{|\alpha_i\rangle\}, \{|\beta_i\rangle\}$, and $\{|\gamma_i\rangle\}$, respectively:

$$A_{jk} = \langle \alpha_j | \alpha_k \rangle, \quad B_{jk} = \langle \beta_j | \beta_k \rangle, \quad C_{jk} = \langle \gamma_j | \gamma_k \rangle. \quad (153)$$

Each of A, B, C has the form

$$M = \begin{pmatrix} 1 & m & m^* \\ m^* & 1 & m \\ m & m^* & 1 \end{pmatrix}, \quad (154)$$

where m^* is the complex conjugate of m , and $M = A, B, C$ respectively when m is set to be

$$\begin{aligned} a &= |a_0|^2 (1 + \omega k_1^2 + \omega^2 k_2^2) \\ b &= |b_0|^2 (1 + \omega l_1^2 + \omega^2 l_2^2) \\ c &= |c_0|^2 (1 + \omega^2 r^2), \end{aligned} \quad (155)$$

respectively. As a side remark, since $V_{s,\mu,\nu}$ is an isometry, it follows that $A = B * C$ where $*$ denotes the elementwise or Hadamard Product of two matrices.

The channel $\mathcal{W}_{s,\mu,\nu}$ is degradable if and only if there is a Gram matrix D satisfying

$$B = C * D. \quad (156)$$

To see this, when such a Gram matrix D exists, there are normalized kets $\{|\delta_i\rangle\}$ in some auxiliary Hilbert space \mathcal{H}_d such that $D_{jk} = \langle \delta_j | \delta_k \rangle$. A possible degrading map can be generated by the pcubed isometry from $\mathcal{H}_b \mapsto \mathcal{H}_c \otimes \mathcal{H}_d$ taking $|\beta_i\rangle$ to $|\gamma_i\rangle \otimes |\delta_i\rangle$. The converse follows from (150) since the degradable map must take $|\beta_i\rangle$ to $|\gamma_i\rangle$, and must be generated by an isometry. (See Sec. III-C in [62] for a detailed discussion.)

A. The Isometry $V_{s,\mu,1-\mu}$

We now consider a two parameter subclass of isometries, $W_{s,\mu}$, obtained from setting $\nu = 1 - \mu$ in $V_{s,\mu,\nu}$, with $s \in [0, 1/2]$ and $\mu \in [0, 1]$. Following (149),

$$\begin{aligned} W_{s,\mu} |0\rangle_a &= \sqrt{s} |0\rangle_b \otimes |0\rangle_c + \sqrt{1-s} |1\rangle_b \otimes |1\rangle_c, \\ W_{s,\mu} |1\rangle_a &= \sqrt{1-\mu} |1\rangle_b \otimes |0\rangle_c + \sqrt{\mu} |2\rangle_b \otimes |1\rangle_c, \\ W_{s,\mu} |2\rangle_a &= \sqrt{\mu} |2\rangle_b \otimes |0\rangle_c + \sqrt{1-\mu} |0\rangle_b \otimes |1\rangle_c. \end{aligned} \quad (157)$$

The resulting channel $\mathcal{W}_{s,\mu} = \text{tr}_c(W_{s,\mu} \cdot W_{s,\mu}^\dagger)$ has two Kraus operators

$$\begin{aligned} K_0 &= \begin{pmatrix} \sqrt{s} & 0 & 0 \\ 0 & \sqrt{1-\mu} & 0 \\ 0 & 0 & \sqrt{\mu} \end{pmatrix} \\ K_1 &= \begin{pmatrix} 0 & 0 & \sqrt{1-\mu} \\ \sqrt{1-s} & 0 & 0 \\ 0 & \sqrt{\mu} & 0 \end{pmatrix}. \end{aligned} \quad (158)$$

As a side remark, when $\mu = 1$, the channel $\mathcal{W}_{s,1}$ is unitarily equivalent to \mathcal{N}_s : we have $\mathcal{N}_s = \mathcal{W}_{s,1}(U \cdot U^\dagger)$, where the unitary U swaps $|1\rangle_a$ and $|2\rangle_a$ at the input.

For the rest of the discussion we focus on $s = 1/2$. We will first evaluate the capacities of $\mathcal{W}_{1/2,1/2}$. Then, we study the degradability of $\mathcal{W}_{1/2,\mu}$, followed by a detailed numerical analysis of its capacities.

When $\mu = 1/2$, $K_0 = \frac{1}{\sqrt{2}} \mathbb{1}$, and K_1 is proportional to the qutrit- X Heisenberg-Weyl operator, which acts as $X|i\rangle = |i+1 \pmod 3\rangle$. Hence, $\mathcal{W}_{1/2,1/2}$ is the qutrit X -dephasing channel with dephasing probability $1/2$, which fixes the eigenbasis of X :

$$\begin{aligned} |\psi_0\rangle_a &= \frac{1}{\sqrt{3}}(|0\rangle_a + |1\rangle_a + |2\rangle_a) \\ |\psi_1\rangle_a &= \frac{1}{\sqrt{3}}(|0\rangle_a + \omega|1\rangle_a + \omega^2|2\rangle_a) \\ |\psi_2\rangle_a &= \frac{1}{\sqrt{3}}(|0\rangle_a + \omega^2|1\rangle_a + \omega|2\rangle_a). \end{aligned} \quad (159)$$

We can now evaluate the capacities. The above invariant basis can be used to transmit $\log d$ bits, and this code saturates the dimension bound for the classical capacity. Thus, the Holevo information is additive and $\chi(\mathcal{W}_{1/2,1/2}) = \mathcal{C}(\mathcal{W}_{1/2,1/2}) = \log d$. For the quantum capacity, one may evaluate the SDP upper bound given in (105) for $\mathcal{W}_{1/2,1/2}$ giving a value of $\log(3/2)$. This upper bound is in turn achieved by the input state $\rho_a = \frac{1}{3} \mathbb{1}$ giving coherent information $\Delta(\mathcal{W}_{1/2,1/2}, \rho_a) = \log(3/2)$. Furthermore, $\mathcal{W}_{1/2,1/2}$ is degradable¹ so the private and quantum capacities coincide. Finally, the channel mutual information $I(\mathcal{W}_{1/2,1/2})$ is equal to $\log(9/2)$, since the input $\rho_a = \frac{1}{3} \mathbb{1}$ simultaneously maximizes the first term and separately the second and third terms combined in (18).

¹The channel $\mathcal{W}_{1/2,1/2}$ can be generated by the isometry which attaches $|+\rangle_c$ to system a , and conditioned on system c being in the state $|1\rangle$, the unitary K_1 is applied to system a which is then relabeled as system b . This isometry applied to system b generates a valid degrading map.

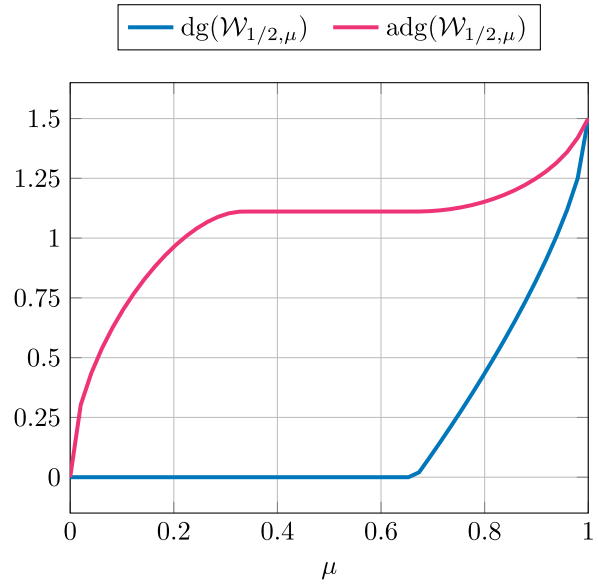


Fig. 2. Degradability parameter $\text{dg}(\mathcal{W}_{1/2,\mu})$ (blue) and antidegradability parameter $\text{adg}(\mathcal{W}_{1/2,\mu})$ (magenta), which are computed using the SDPs from [65]. A quantum channel \mathcal{B} is degradable iff $\text{dg}(\mathcal{B}) = 0$, and anti-degradable iff $\text{adg}(\mathcal{B}) = 0$.

Altogether, for the quantum and private capacities we have

$$\begin{aligned} \mathcal{Q}(\mathcal{W}_{1/2,1/2}) &= \mathcal{Q}^{(1)}(\mathcal{W}_{1/2,1/2}) \\ &= \mathcal{P}^{(1)}(\mathcal{W}_{1/2,1/2}) \\ &= \mathcal{P}(\mathcal{W}_{1/2,1/2}) = \log \frac{3}{2}, \end{aligned} \quad (160a)$$

while for the classical capacities we have

$$\mathcal{C}(\mathcal{W}_{1/2,1/2}) = \chi(\mathcal{W}_{1/2,1/2}) = \log 3, \quad (160b)$$

$$\mathcal{C}_E(\mathcal{W}_{1/2,1/2}) = \log \frac{9}{2}. \quad (160c)$$

Keeping $s = 1/2$, we now return to general $\mu \in [0, 1]$ in studying $\mathcal{W}_{1/2,\mu}$. We first use the pced framework to study degradability of this subclass of channels. The values of a, b , and c in (155) are given by

$$a = \frac{1-2\mu}{2(2-\mu)}, \quad b = \omega^2 \frac{2\mu-1}{2-\mu}, \quad c = -\omega/2. \quad (161)$$

Setting m in (154) to a, b , and c gives the Gram matrices A, B , and C respectively. The channel $\mathcal{W}_{1/2,\mu}$ is degradable iff $B = C * D$ for some Gram matrix D (see (156)). If such a matrix exists, it has the form given by M in (154) with $m = d = 2\omega(1-2\mu)/(2-\mu)$, and thus is a valid Gram matrix (being positive semi-definite) when $1+2d^3-3|d|^2 \geq 0$ (see eq. (57) in [62]). Setting $d = 2\omega(1-2\mu)/(2-\mu)$ gives $\mu \leq 2/3$. Thus, $\mathcal{W}_{1/2,\mu}$ is degradable iff $\mu \leq 2/3$.

We now evaluate lower and upper bounds on the quantum, private, and classical capacities of the channel $\mathcal{W}_{1/2,\mu}$ for $\mu \in [0, 1]$, collected in Fig. 3. The bounds on the quantum capacity $\mathcal{Q}(\mathcal{W}_{1/2,\mu})$ are obtained by numerically optimizing the single-letter coherent information $\mathcal{Q}^{(1)}(\mathcal{W}_{1/2,\mu})$ (solid blue line in Fig. 3) and the SDP upper bound (105) (dashed blue line in Fig. 3). Interestingly, the (possibly tighter) SDP

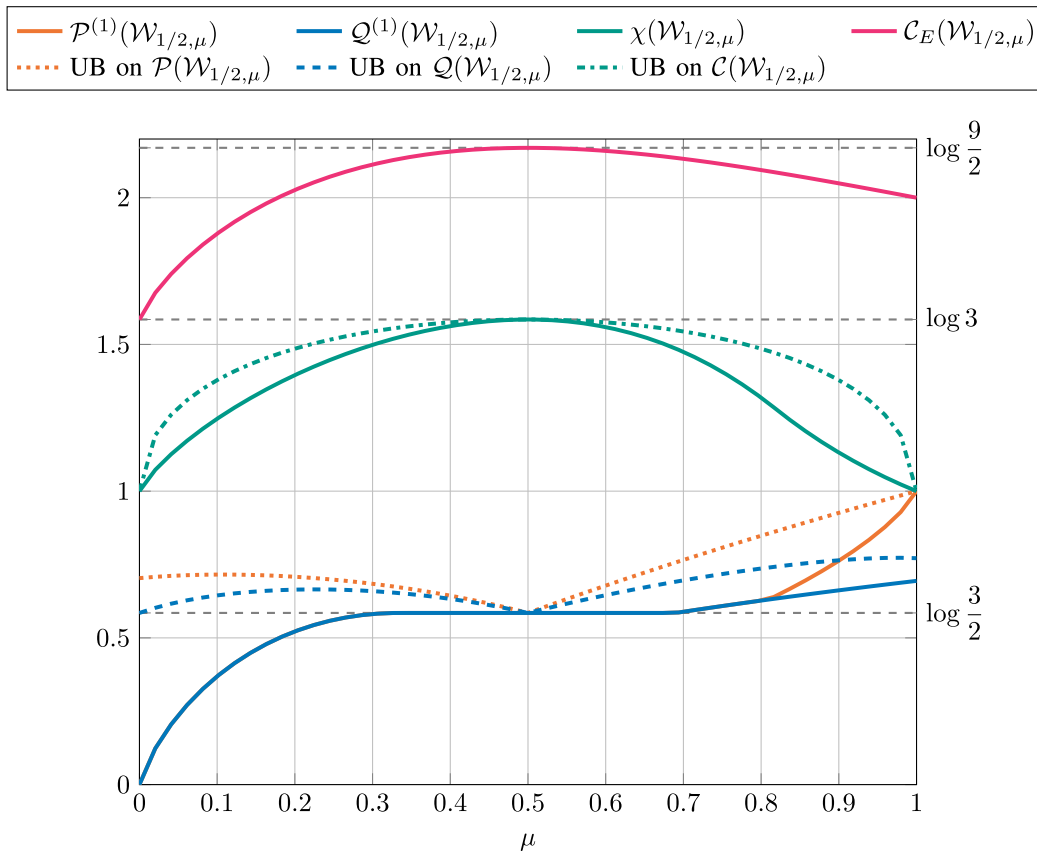


Fig. 3. Lower and upper bounds (UB) on the capacities of the quantum channel $\mathcal{W}_{1/2,\mu}$ defined via the isometry (157). The quantum capacity $\mathcal{Q}(\mathcal{W}_{1/2,\mu})$ is bounded from below by the single-letter coherent information $\mathcal{Q}^{(1)}(\mathcal{W}_{1/2,\mu})$ (solid blue) and from above by the SDP bound (105) derived in [50] (dashed blue). The private capacity $\mathcal{P}(\mathcal{W}_{1/2,\mu})$ is bounded from below by the private information $\mathcal{P}^{(1)}(\mathcal{W}_{1/2,\mu})$ (solid orange) and from above by the bound given by Lemma 11 (dotted orange). The classical capacity $\mathcal{C}(\mathcal{W}_{1/2,\mu})$ is bounded from below by the Holevo information $\chi(\mathcal{W}_{1/2,\mu})$ (solid green) and from above by the SDP bound (131) derived in [45] (dash-dotted green). We also plot the entanglement-assisted classical capacity $\mathcal{C}_E(\mathcal{W}_{1/2,\mu})$ (solid magenta), computed using the technique in [36]. For $\mu = \frac{1}{2}$ the channel $\mathcal{W}_{1/2,1/2}$ is a dephasing channel, for which the special values of the capacities from (160) are marked on the right-hand side.

upper bound from [51] coincides with the SDP bound (105) from [50] for all $\mu \in [0, 1]$. To bound the classical capacity $\mathcal{C}(\mathcal{W}_{1/2,\mu})$, we numerically optimize the single-letter Holevo information $\chi(\mathcal{W}_{1/2,\mu})$ (solid green line in Fig. 3) and evaluate the SDP upper bound (131) (dashed green line in Fig. 3). The entanglement-assisted capacity $\mathcal{C}_E(\mathcal{W}_{1/2,\mu})$ is computed using the technique developed in [36] (solid magenta line in Fig. 3). Finally, for the private capacity $\mathcal{P}(\mathcal{W}_{1/2,\mu})$ we numerically optimize the single-letter private information $\mathcal{P}^{(1)}(\mathcal{W}_{1/2,\mu})$ (solid orange line in Fig. 3). To obtain an upper bound on the private capacity we employ the following recent result by Fawzi and Fawzi [51] providing a bound on the private capacity of a quantum channel in terms of a conic program:

Proposition 10 ([51]): Let $\mathcal{B}: \mathcal{H}_a \rightarrow \mathcal{H}_b$ be a quantum channel with (unnormalized) Choi-Jamiołkowski operator $J_{ab}^{\mathcal{B}}$. Let furthermore $l \in \mathbb{N}$ and set $\alpha = 1 + 2^{-l}$. Then we have

$$\mathcal{P}(\mathcal{B}) \leq \hat{E}_\alpha(\mathcal{B}), \tag{162}$$

where $\hat{E}_\alpha(\mathcal{B}) = l2^l - (2^l + 1) \log(2^l + 1) + (2^l + 1) \log T_\alpha(\mathcal{B})$, and $T_\alpha(\mathcal{B})$ is the solution of the following conic program:

$$T_\alpha(\mathcal{B}) = \max. \text{tr} \left[J_{ab}^{\mathcal{B}} \left(K^{\text{H}} - \sum_{i=1}^l W_i \right) \right]$$

$$\begin{aligned} \text{s.t. } & \{W_i\}_{i=1}^l \in \text{Herm}(\mathcal{H}_a \otimes \mathcal{H}_b) \\ & K, \{Z_i\}_{i=0}^l \in \mathcal{L}(\mathcal{H}_a \otimes \mathcal{H}_b) \\ & \rho_a \in \text{Herm}(\mathcal{H}_a), \text{tr } \rho_a = 1 \\ & \begin{pmatrix} \rho_a \otimes \mathbb{1}_b & K \\ K^\dagger & Z_l^{\text{H}} \end{pmatrix} \geq 0 \\ & \begin{pmatrix} W_i & Z_i \\ Z_i^\dagger & Z_{i-1}^{\text{H}} \end{pmatrix} \geq 0 \quad \text{for all } i = 1, \dots, l \\ & \rho_a \otimes \mathbb{1}_b - Z_0^{\text{H}} \text{ block-positive.} \end{aligned} \tag{163}$$

In the above, $X^{\text{H}} = X + X^\dagger$, and an operator X_{ab} is block-positive (with respect to the bipartition $a : b$) if $(\langle \psi_a | \otimes \langle \phi_b |) X_{ab} (| \psi_a \rangle \otimes | \phi_b \rangle) \geq 0$ for all $|\psi_a\rangle \in \mathcal{H}_a$ and $|\phi_b\rangle \in \mathcal{H}_b$. Hence, block-positive bipartite states are the states Choi-Jamiołkowski operators of positive maps (see, e.g., [63]).

The conic program in Proposition 10 only reduces to an SDP if $d_a d_b \leq 6$, whereas our channel $\mathcal{W}_{1/2,\mu}$ has qutrit input and output, $d_a = d_b = 3$. However, the following strategy suggested to us by Hamza Fawzi [64] may be employed to obtain an (SDP-computable) upper bound on the quantity $T_\alpha(\mathcal{B})$ (and hence $\hat{E}_\alpha(\mathcal{B})$), in turn giving an upper bound on the private capacity of a channel \mathcal{B} via Proposition 10:

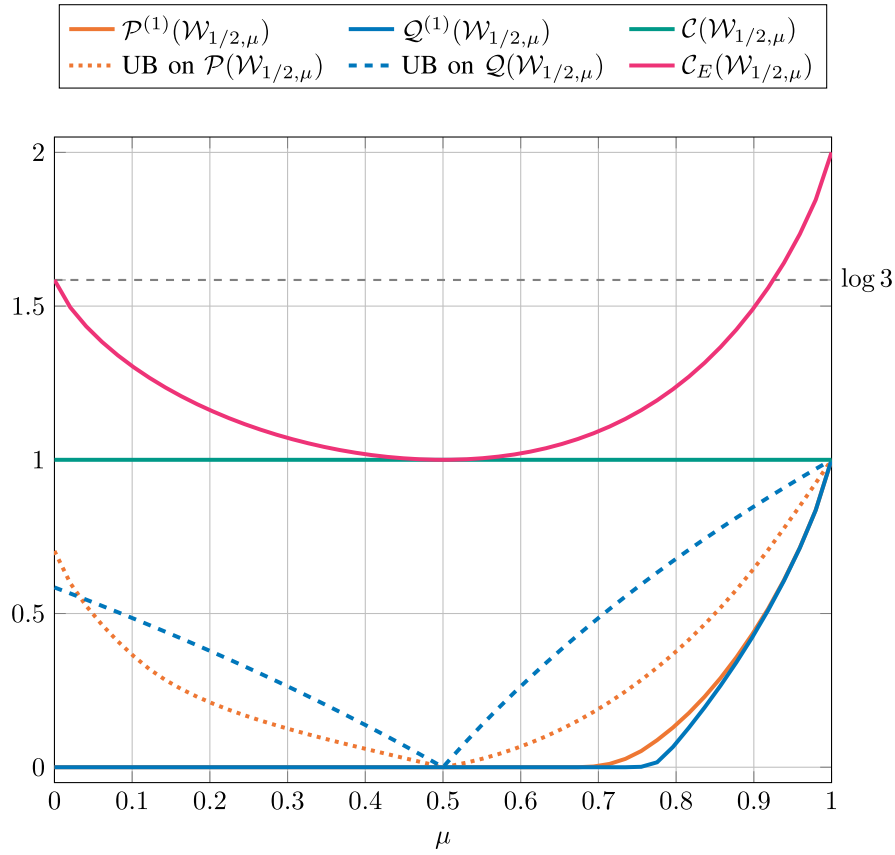


Fig. 4. Lower and upper bounds (UB) on the capacities of the complementary channel $\mathcal{W}_{1/2, \mu}^c$ defined via the isometry (157). The quantum capacity $\mathcal{Q}(\mathcal{W}_{1/2, \mu}^c)$ is bounded from below by the single-letter coherent information $\mathcal{Q}^{(1)}(\mathcal{W}_{1/2, \mu}^c)$ (solid blue) and from above by the SDP bound (105) derived in [50] (dashed blue). The private capacity $\mathcal{P}(\mathcal{W}_{1/2, \mu}^c)$ is bounded from below by the private information $\mathcal{P}^{(1)}(\mathcal{W}_{1/2, \mu}^c)$ (solid orange) and from above by the bound given by Lemma 11 (dotted orange). The Holevo information $\chi(\mathcal{W}_{1/2, \mu}^c)$ coincides with the SDP bound (131) derived in [45], and is hence (numerically) equal to the classical capacity $\mathcal{C}(\mathcal{W}_{1/2, \mu}^c)$ (solid green). We also plot the entanglement-assisted classical capacity $\mathcal{C}_E(\mathcal{W}_{1/2, \mu}^c)$ (solid magenta), computed using the technique in [36].

Lemma 11 ([51], [64]): With the same notation as in Proposition 10, we have the following bound on the private capacity of a quantum channel $\mathcal{B}: \hat{\mathcal{H}}_a \rightarrow \hat{\mathcal{H}}_b$:

$$\mathcal{P}(\mathcal{B}) \leq \min_{\{\phi_a^i\}_{i=1}^N} F_\alpha(\mathcal{B}, \{\phi_a^i\}_{i=1}^N), \quad (164)$$

where $N \in \mathbb{N}$ is some fixed natural number, the minimization is over sets of pure states $\phi_a^i \in \mathcal{H}_a$, $i = 1, \dots, N$, and the quantity $F_\alpha(\mathcal{B}, \{\phi_a^i\}_{i=1}^N)$ is defined as

$$F_\alpha(\mathcal{B}) = l2^l - (2^l + 1) \log(2^l + 1) + (2^l + 1) \log U_\alpha(\mathcal{B}, \{\phi_a^i\}_{i=1}^N). \quad (165)$$

In the above, $U_\alpha(\mathcal{B}, \{\phi_a^i\}_{i=1}^N)$ is the solution of the following semidefinite program:

$$\begin{aligned} \max. \quad & \text{tr} \left[J_{ab}^{\mathcal{B}} \left(K^{\text{H}} - \sum_{i=1}^l W_i \right) \right] \\ \text{s.t.} \quad & \{W_i\}_{i=1}^l \in \text{Herm}(\mathcal{H}_a \otimes \mathcal{H}_b) \\ & K, \{Z_i\}_{i=0}^l \in \mathcal{L}(\mathcal{H}_a \otimes \mathcal{H}_b) \\ & \rho_a \in \text{Herm}(\mathcal{H}_a), \text{tr} \rho_a = 1 \\ & \begin{pmatrix} \rho_a \otimes \mathbb{1}_b & K \\ K^\dagger & Z_l^{\text{H}} \end{pmatrix} \geq 0 \end{aligned}$$

$$\begin{aligned} \begin{pmatrix} W_i & Z_i \\ Z_i^\dagger & Z_{i-1}^{\text{H}} \end{pmatrix} &\geq 0 \quad \text{for all } i = 1, \dots, l \\ \text{tr}_a (\sigma_{ab} ((\phi_a^i)^T \otimes \mathbb{1}_b)) &\geq 0 \quad \text{for } i = 1, \dots, N, \end{aligned} \quad (166)$$

where $\sigma_{ab} = \rho_a \otimes \mathbb{1}_b - Z_0^{\text{H}}$.

Proof: The block positivity constraint on $\sigma_{ab} = \rho_a \otimes \mathbb{1}_b - Z_0^{\text{H}}$ translates via the Choi isomorphism to positivity of the map $\Psi: \hat{\mathcal{H}}_a \rightarrow \hat{\mathcal{H}}_b$ whose Choi-Jamiołkowski operator is σ_{ab} , i.e., $\Psi(\chi_a) \geq 0$ for all pure state $|\chi\rangle_a \in \mathcal{H}_a$. Relaxing this positivity constraint to only requiring $\Psi(\phi_a^i) \geq 0$ for some fixed pure states ϕ_a^i , $i = 1, \dots, N$ now yields a maximization over a larger set of Choi-Jamiołkowski operators resp. maps Ψ , and hence we obtain $U_\alpha(\mathcal{B}, \{\phi_a^i\}_{i=1}^N) \geq T_\alpha(\mathcal{B})$, where $T_\alpha(\mathcal{B})$ is defined in Proposition 10. Strengthening this bound by minimizing over sets of pure states $\{\phi_a^i\}_{i=1}^N$ for fixed $N \in \mathbb{N}$ finishes the proof. ■

Any choice of $N \in \mathbb{N}$ and pure states $\{\phi_a^i\}_{i=1}^N$ yields a feasible point in the minimization in Lemma 11, and hence an upper bound on the private capacity of \mathcal{B} . The states $\{\phi_a^i\}_{i=1}^N$ can for example be sampled from the Haar measure. For our purposes, choosing (a union of) mutually unbiased bases yields a tighter upper bound. More precisely, we choose the computational basis $\{|i\rangle_a\}_{i=0}^2$ and the X -eigenbasis $\{|\psi_i\rangle_a\}_{i=0}^2$, defined in (159). The resulting bound on $\mathcal{P}(\mathcal{W}_{1/2, \mu}^c)$ is plotted

in Fig. 3 (dotted orange line). We also performed a similar numerical analysis for the capacities of the complementary channel $\mathcal{W}_{s,\mu}^c$ in Figure 4.

Figure 3 reveals a number of interesting properties of the 1-parameter channel family $\mathcal{W}_{1/2,\mu}$:

- For $\mu \lesssim 0.8$, the coherent information (solid blue line in Fig. 3) and private information (solid orange line in Fig. 3) coincide. The channel is degradable for $\mu \lesssim 0.65$ and anti-degradable for $\mu = 0$, see Fig. 2.
- For $\mu \gtrsim 0.8$, the private information (solid orange line in Fig. 3) is strictly larger than the coherent information (solid blue line in Fig. 3). For $\mu \gtrsim 0.9$ the private information exceeds the SDP upper bound on the quantum capacity (dashed blue line in Fig. 3), hence giving a provable separation between quantum and private capacity.
- The upper bound on $\mathcal{P}(\mathcal{W}_{1/2,\mu})$ derived via Lemma 11 (dotted orange line in Fig. 3) clearly separates the private capacity from the classical capacity for all $\mu < 1$.
- At $\mu \approx 0.8$ the Holevo information (solid green line in Fig. 3) has an inflection point, changing from concave to convex. For $\mu \lesssim 0.8$ the optimal Holevo information is achieved by an ensemble of three pure states, whereas for $\mu \gtrsim 0.8$ four pure states are needed. This may be a signature of super-additivity of Holevo information, and will be further investigated in future work.

XI. CONCLUDING REMARKS

We have studied families of channels that are very simple, yet still nontrivial in terms of their capacities. Our goal is to better understand the boundary between trivially solvable and incomprehensibly complex behavior. Our main results demonstrate how intricately narrow this boundary can be—with complex quantum effects arising in seemingly innocent and generic settings. We conclude this paper by highlighting some of these results.

The primary example of quantum channel in this study is obtained by combining two very simple channels. As mentioned at the end of Sec. III-A, we construct the channel of interest \mathcal{N}_s by “hybridizing” a degradable channel \mathcal{N}_2 and a completely useless channel \mathcal{N}_1 . The quantum, private, and classical capacities of \mathcal{N}_1 are all zero. Meanwhile, the coherent information and various capacities of the degradable channel \mathcal{N}_2 can be evaluated: $\mathcal{Q}^{(1)}(\mathcal{N}_2) = \mathcal{Q}(\mathcal{N}_2) = \mathcal{P}(\mathcal{N}_2) < 1 = \mathcal{C}(\mathcal{N}_2)$. Both channels have 2-dimensional inputs, and nothing extraordinary on their own. We “stitch” these channels together, via a common joint output state to the receiver and the environment. The resulting hybrid channel \mathcal{N}_s has capacities bearing interesting relationships with those of \mathcal{N}_2 : their coherent informations are identical, $\mathcal{Q}^{(1)}(\mathcal{N}_s) = \mathcal{Q}^{(1)}(\mathcal{N}_2)$, and conditioned on the spin alignment conjecture, $\mathcal{Q}(\mathcal{N}_s) = \mathcal{Q}(\mathcal{N}_2)$ (see Fig. 1). Meanwhile, both the private and classical capacity of \mathcal{N}_s are equal to the classical capacity of \mathcal{N}_2 . In other words, starting from \mathcal{N}_2 and “stitching” onto it a completely useless channel \mathcal{N}_1 boosts the private capacity of \mathcal{N}_2 from its quantum capacity to its strictly larger classical capacity, while all other quantities remain the same.

Both the classical and private capacities of \mathcal{N}_s are, quite intuitively, equal to 1. This value as an upper bound comes from the uselessness of the input state $|2\rangle$ for sending classical information in addition to what can already be sent using the states $|0\rangle$ and $|1\rangle$, a property inherited from the uselessness of \mathcal{N}_1 . The remaining input space is 2-dimensional so the classical capacity cannot exceed 1. This value as a lower bound comes from a simple, single-letter, perfect, private code using two signalling states, $|0\rangle$, along with a mixture of $|1\rangle$ and $|2\rangle$. These give rise to respective orthogonal states to the output but identical states to the environment. Note that this private code is made possible by the stitching of \mathcal{N}_1 to \mathcal{N}_2 ; in fact, this quantum “stitch” contributes to a minimal shield in the p-bit framework discussed in Sec. IX. Since the private classical rate cannot exceed the classical capacity, both private and classical capacities must be 1. It is highly non-trivial to evaluate these capacities rigorously; \mathcal{N}_s does not belong to any known class of channels with additive Holevo and private informations. These capacity calculations also prove that that channel’s Holevo and private information are additive in the sense $\chi(\mathcal{N}_s) = \mathcal{C}(\mathcal{N}_s)$ and $\mathcal{P}^{(1)}(\mathcal{N}_s) = \mathcal{P}(\mathcal{N}_s)$. This additivity however is shown using arguments that are very different from those used in prior works.

Furthermore, in Sec. VIII, we not only find the classical and private capacity of \mathcal{N}_s and its higher dimensional analogue \mathcal{M}_d , we also provide a strong converse bound for each of these capacities. The bound shows that classical or private transmission rates exceeding the capacity attracts an error that converges to 1 exponentially. Such strong converse bounds are unavailable for most quantum channels even when one can compute their capacities.

The quantum capacity of \mathcal{N}_s is somewhat more complicated, but apparently can also be understood. In particular, by restricting to the input space of \mathcal{N}_2 , the quantum capacities are related as $\mathcal{Q}(\mathcal{N}_2) \leq \mathcal{Q}(\mathcal{N}_s)$. Relative to the spin alignment conjecture, we find that $\mathcal{Q}(\mathcal{N}_s) = \mathcal{Q}(\mathcal{N}_2)$ where $\mathcal{Q}(\mathcal{N}_2) = \mathcal{Q}^{(1)}(\mathcal{N}_2)$ because of its degradability. The capacity of \mathcal{N}_s can be understood this way even though it is neither degradable nor antidegradable. A rigorous proof that the capacity of \mathcal{N}_s is additive would follow from the spin-alignment conjecture. Such a proof would be qualitatively different from prior additivity proofs. Our hope is that this spin-alignment conjecture, which is at heart about the geometric structure of states minimizing entropy, will lead to further progress on additivity questions in information theory.

We finally have a family of channels that are not of the usual tractable types and yet for which we know the classical, private, and quantum capacities. Underlying this superficial simplicity, the private capacity is much higher than the quantum capacity, a signature for novel quantum effects at play. These channels, and their higher dimensional generalizations, continue to surprise. One may have thought that the weak additivities observed here would portend strong additivity. But nothing could be further from the truth. In a companion paper [1], we find that the coherent information of the channel \mathcal{N}_s tensored with an assisting channel is super-additive, for a large swath of values of s and for some generically chosen assisting channel. The super-additivity can

be lifted to quantum capacity for degradable assisting channels and if the spin alignment conjecture holds. The assisting channel can have positive or vanishing quantum capacity. The mechanism behind this superadditivity is novel and in particular differs from the known explanation of super-activation [6], [16]. Additional super-additivity of quantum capacity that is unconditional on the spin alignment conjecture can be proved for the d -dimensional generalization \mathcal{M}_d of $\mathcal{N}_{1/2}$, when it is used with a $(d-1)$ -dimensional erasure channel for all nontrivial values of the erasure probability! In contrast, the pan-additivity of our channels is a fascinating and amazing progress.

APPENDIX

REPRESENTATIONS OF THE PLATYPUS CHANNEL FAMILY

A. \mathcal{N}_s Channel

Let $\mathcal{H}_a \cong \mathcal{H}_b \cong \mathbb{C}^3$ and $\mathcal{H}_c \cong \mathbb{C}^2$, and $s \in [0, 1/2]$. The platypus channel $\mathcal{N}_s: \hat{\mathcal{H}}_a \rightarrow \hat{\mathcal{H}}_b$ can be defined as follows:

1) *Isometry*: $\mathcal{N}_s(X_a) = \text{tr}_c F_s X_a F_s^\dagger$ with $F_s: \mathcal{H}_a \mapsto \mathcal{H}_b \otimes \mathcal{H}_c$ defined as

$$\begin{aligned} F_s: |0\rangle_a &\mapsto \sqrt{s} |0\rangle_b |0\rangle_c + \sqrt{1-s} |1\rangle_b |1\rangle_c \\ |1\rangle_a &\mapsto |2\rangle_b |0\rangle_c \\ |2\rangle_a &\mapsto |2\rangle_b |1\rangle_c. \end{aligned} \quad (167)$$

2) *Choi Operator*: $J_{ab}^{\mathcal{N}_s} = d_a(\mathcal{I}_a \otimes \mathcal{N}_s)([\phi]): \mathcal{H}_a \otimes \mathcal{H}_b \rightarrow \mathcal{H}_a \otimes \mathcal{H}_b$ (0's are replaced by .'s for readability),

$$J_{ab}^s = \begin{pmatrix} s & \cdot & \cdot & \cdot & \cdot & \sqrt{s} & \cdot & \cdot \\ \cdot & 1-s & \cdot & \cdot & \cdot & \cdot & \cdot & \sqrt{1-s} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \sqrt{s} & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \sqrt{1-s} & \cdot & \cdot & \cdot & \cdot & \cdot & 1 \end{pmatrix} \quad (168)$$

3) *Kraus Operators*: $\mathcal{N}_s(X_a) = N_0 X_a N_0^\dagger + N_1 X_a N_1^\dagger$ with $N_i: \mathcal{H}_a \rightarrow \mathcal{H}_b$ defined as

$$N_0 = \begin{pmatrix} \sqrt{s} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad N_1 = \begin{pmatrix} 0 & 0 & 0 \\ \sqrt{1-s} & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (169)$$

B. \mathcal{M}_d Channel

For $d \in \mathbb{N}$, $d \geq 3$ let $\mathcal{H}_a \cong \mathcal{H}_b \cong \mathbb{C}^d$ and $\mathcal{H}_c \cong \mathbb{C}^{d-1}$. The platypus channel $\mathcal{M}_d: \hat{\mathcal{H}}_a \rightarrow \hat{\mathcal{H}}_b$ can be defined as follows:

1) *Isometry*: $\mathcal{M}_d(X_a) = \text{tr}_c G X_a G^\dagger$ with $G: \mathcal{H}_a \rightarrow \mathcal{H}_b \otimes \mathcal{H}_c$ defined as

$$\begin{aligned} G: |0\rangle_a &\mapsto \frac{1}{\sqrt{d-1}} \sum_{j=0}^{d-2} |j\rangle_b |j\rangle_c \\ |j\rangle_a &\mapsto |d-1\rangle_b |j-1\rangle_c \quad \text{for } j = 1, \dots, d-1. \end{aligned} \quad (170)$$

2) *Choi Operator*: $J_{ab}^{\mathcal{M}_d} = d_a(\mathcal{I}_a \otimes \mathcal{M}_d)([\phi]): \mathcal{H}_a \otimes \mathcal{H}_b \rightarrow \mathcal{H}_a \otimes \mathcal{H}_b$,

$$\begin{aligned} J_{ab}^{\mathcal{M}_d} &= \sum_{j=0}^{d-2} \left(\frac{1}{d-1} [0j] + [j+1, d-1] \right. \\ &\quad \left. + \frac{1}{\sqrt{d-1}} (|0, j\rangle\langle j+1, d-1| + |j+1, d-1\rangle\langle 0, j|) \right), \end{aligned} \quad (171)$$

where $[\psi] \equiv |\psi\rangle\langle\psi|$.

3) *Kraus Operators*: $\mathcal{M}_d(X_a) = \sum_{k=0}^{d-2} M_k X_a M_k^\dagger$ with $M_k: \mathcal{H}_a \rightarrow \mathcal{H}_b$ for $k = 0, \dots, d-2$ defined as

$$M_k = \frac{1}{\sqrt{d-1}} |k\rangle_b \langle 0|_a + |d-1\rangle_b \langle k+1|_a. \quad (172)$$

C. \mathcal{O} Channel

For $d \in \mathbb{N}$, $d \geq 2$ let $\mathcal{H}_a \cong \mathcal{H}_b \cong \mathbb{C}^d$ and $\mathcal{H}_c \cong \mathbb{C}^{d-1}$. For $0 \leq j \leq d-2$ let $\mu_j \in \mathbb{C}$ with $\sum_{j=0}^{d-2} |\mu_j|^2 = 1$. The platypus channel $\mathcal{O}: \hat{\mathcal{H}}_a \rightarrow \hat{\mathcal{H}}_b$ can be defined as follows:

1) *Isometry*: $\mathcal{O}(X_a) = \text{tr}_c H X_a H^\dagger$ with $H: \mathcal{H}_a \rightarrow \mathcal{H}_b \otimes \mathcal{H}_c$ defined as

$$\begin{aligned} H: |0\rangle_a &\mapsto \sum_{j=0}^{d-2} \mu_j |j\rangle_b |j\rangle_c \\ |j\rangle_a &\mapsto |d-1\rangle_b |j-1\rangle_c \quad \text{for } j = 1, \dots, d-1. \end{aligned} \quad (173)$$

2) *Kraus Operators*: $\mathcal{O}(X_a) = \sum_{k=0}^{d-2} O_k X_a O_k^\dagger$ with $O_k: \mathcal{H}_a \rightarrow \mathcal{H}_b$ for $k = 0, \dots, d-2$ defined as

$$O_k = \mu_k |k\rangle_b \langle 0|_a + |d-1\rangle_b \langle k+1|_a. \quad (174)$$

ACKNOWLEDGMENT

The authors thank Mark M. Wilde for helpful comments regarding feedback-assisted capacities.

REFERENCES

- [1] F. Leditzky, D. Leung, V. Siddhu, G. Smith, and J. A. Smolin, "Generic nonadditivity of quantum capacity in simple channels," *Phys. Rev. Lett.*, 2022.
- [2] P. W. Shor and J. A. Smolin, "Quantum error-correcting codes need not completely reveal the error syndrome," 1996, *arXiv:quant-ph/9604006*.
- [3] D. P. DiVincenzo, P. W. Shor, and J. A. Smolin, "Quantum-channel capacity of very noisy channels," *Phys. Rev. A, Gen. Phys.*, vol. 57, no. 2, p. 830, 1998.
- [4] G. Smith, J. M. Renes, and J. A. Smolin, "Structured codes improve the Bennett-Brassard-84 quantum key rate," *Phys. Rev. Lett.*, vol. 100, no. 17, Apr. 2008, Art. no. 170502.
- [5] M. B. Hastings, "Superadditivity of communication capacity using entangled inputs," *Nature Phys.*, vol. 5, no. 4, pp. 255–257, 2009.
- [6] G. Smith and J. T. Yard, "Quantum communication with zero-capacity channels," *Science*, vol. 321, no. 5897, pp. 1812–1815, Sep. 2008.
- [7] P. Horodecki, M. Horodecki, and R. Horodecki, "Binding entanglement channels," *J. Mod. Opt.*, vol. 47, nos. 2–3, pp. 347–354, Feb. 2000.
- [8] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, "Secure key from bound entanglement," *Phys. Rev. Lett.*, vol. 94, no. 16, Apr. 2005, Art. no. 160502.
- [9] D. Leung, K. Li, G. Smith, and J. A. Smolin, "Maximal privacy without coherence," *Phys. Rev. Lett.*, vol. 113, no. 3, Jul. 2014, Art. no. 030502.
- [10] I. Devetak and P. W. Shor, "The capacity of a quantum channel for simultaneous transmission of classical and quantum information," *Commun. Math. Phys.*, vol. 256, no. 2, pp. 287–303, 2005.
- [11] G. Smith, "Private classical capacity with a symmetric side channel and its application to quantum cryptography," *Phys. Rev. A, Gen. Phys.*, vol. 78v, Art. no. 022306.

- [12] C. King, "Additivity for unital qubit channels," *J. Math. Phys.*, vol. 43, no. 10, pp. 4641–4653, 2002. [Online]. Available: <http://scitation.aip.org/content/aip/journal/jmp/43/10/10.1063/1.1500791>
- [13] C. King, "The capacity of the quantum depolarizing channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 1, pp. 221–229, Jan. 2003.
- [14] P. W. Shor, "Additivity of the classical capacity of entanglement-breaking quantum channels," *J. Math. Phys.*, vol. 43, no. 9, pp. 4334–4340, 2002.
- [15] C. King, K. Matsumoto, M. Nathanson, and M. B. Ruskai, "Properties of conjugate channels with applications to additivity and multiplicativity," *Markov Process Rel. Fields*, vol. 13, nos. 1–2, pp. 391–423, 2007. [Online]. Available: <http://math-mpfrf.org/journal/articles/id1123/>
- [16] J. Oppenheim, "For quantum information, two wrongs can make a right," *Science*, vol. 321, no. 5897, pp. 1783–1784, Sep. 2008.
- [17] (2022). *GitHub Repository With Supplementary Code*. [Online]. Available: <https://github.com/felixled/platypus>
- [18] S. Lloyd, "Capacity of the noisy quantum channel," *Phys. Rev. A, Gen. Phys.*, vol. 55, no. 3, p. 1613, Mar. 1997.
- [19] P. W. Shor, "Quantum error correction," in *Proc. MSRI Workshop Quantum Comput.*, 2002. [Online]. Available: <https://www.msri.org/workshops/203/schedules/1181>
- [20] I. Devetak, "The private classical capacity and quantum capacity of a quantum channel," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 44–55, Jan. 2005.
- [21] H. Barnum, M. A. Nielsen, and B. Schumacher, "Information transmission through a noisy quantum channel," *Phys. Rev. A, Gen. Phys.*, vol. 57, no. 6, pp. 4153–4175, 1998, doi: [10.1103/PhysRevA.57.4153](https://doi.org/10.1103/PhysRevA.57.4153).
- [22] H. Barnum, E. Knill, and M. A. Nielsen, "On quantum fidelities and channel capacities," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1317–1329, Jul. 2000.
- [23] J. Fern and K. B. Whaley, "Lower bounds on the nonzero capacity of Pauli channels," *Phys. Rev. A, Gen. Phys.*, vol. 78, no. 6, Dec. 2008, Art. no. 062335.
- [24] G. Smith and J. A. Smolin, "Degenerate quantum codes for Pauli channels," *Phys. Rev. Lett.*, vol. 98, no. 3, Jan. 2007, Art. no. 030501.
- [25] F. Leditzky, D. Leung, and G. Smith, "Dephasing channel and super-additivity of coherent information," *Phys. Rev. Lett.*, vol. 121, no. 16, Oct. 2018, Art. no. 160501.
- [26] J. Bausch and F. Leditzky, "Error thresholds for arbitrary Pauli noise," *SIAM J. Comput.*, vol. 50, no. 4, pp. 1410–1460, Jan. 2021.
- [27] J. Bausch and F. Leditzky, "Quantum codes from neural networks," *New J. Phys.*, vol. 22, no. 2, Feb. 2020, Art. no. 023005.
- [28] V. Siddhu and R. B. Griffiths, "Positivity and nonadditivity of quantum capacities using generalized erasure channels," *IEEE Trans. Inf. Theory*, vol. 67, no. 7, pp. 4533–4545, Jul. 2021.
- [29] V. Siddhu, "Leaking information to gain entanglement," 2020, *arXiv:2011.15116*.
- [30] V. Siddhu, "Entropic singularities give rise to quantum transmission," *Nature Commun.*, vol. 12, no. 1, p. 5750, Oct. 2021.
- [31] N. Cai, A. Winter, and R. W. Yeung, "Quantum privacy and quantum wiretap channels," *Problems Inf. Transmiss.*, vol. 40, no. 4, pp. 318–336, 2004.
- [32] M. M. Wilde, *Quantum Information Theory*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2016.
- [33] A. S. Holevo, "The capacity of the quantum channel with general signal states," *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 269–273, Jan. 1998.
- [34] B. Schumacher and M. D. Westmoreland, "Sending classical information via noisy quantum channels," *Phys. Rev. A, Gen. Phys.*, vol. 56, no. 1, pp. 131–138, Jul. 1997.
- [35] P. W. Shor, "Equivalence of additivity questions in quantum information theory," *Commun. Math. Phys.*, vol. 246, no. 3, p. 473, Apr. 2004.
- [36] H. Fawzi and O. Fawzi, "Efficient optimization of the quantum relative entropy," *J. Phys. A, Math. Theor.*, vol. 51, no. 15, Apr. 2018, Art. no. 154003.
- [37] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, "Entanglement-assisted classical capacity of noisy quantum channels," *Phys. Rev. Lett.*, vol. 83, no. 15, pp. 3081–3084, Oct. 1999.
- [38] T. S. Cubitt, M. B. Ruskai, and G. Smith, "The structure of degradable quantum channels," *J. Math. Phys.*, vol. 49, no. 10, Oct. 2008, Art. no. 102104.
- [39] F. Leditzky, N. Datta, and G. Smith, "Useful states and entanglement distillation," *IEEE Trans. Inf. Theory*, vol. 64, no. 7, pp. 4689–4708, Jul. 2018.
- [40] J. Yard, P. Hayden, and I. Devetak, "Capacity theorems for quantum multiple-access channels: Classical-quantum and quantum-quantum capacity regions," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 3091–3113, Jul. 2008.
- [41] N. Ramakrishnan, R. Iten, V. B. Scholz, and M. Berta, "Computing quantum channel capacities," *IEEE Trans. Inf. Theory*, vol. 67, no. 2, pp. 946–960, Feb. 2021.
- [42] M. Horodecki, P. W. Shor, and M. B. Ruskai, "Entanglement breaking channels," *Rev. Math. Phys.*, vol. 15, no. 6, pp. 629–641, 2003.
- [43] S. Watanabe, "Private and quantum capacities of more capable and less noisy quantum channels," *Phys. Rev. A, Gen. Phys.*, vol. 85, no. 1, Jan. 2012, Art. no. 012326.
- [44] X. Wang and R. Duan, "Separation between quantum Lovász number and entanglement-assisted zero-error classical capacity," *IEEE Trans. Inf. Theory*, vol. 64, no. 3, pp. 1454–1460, Mar. 2018.
- [45] X. Wang, W. Xie, and R. Duan, "Semidefinite programming strong converse bounds for classical capacity," *IEEE Trans. Inf. Theory*, vol. 64, no. 1, pp. 640–653, Jan. 2018.
- [46] R. Bhatia, *Matrix Analysis*. New York, NY, USA: Springer-Verlag, 1997.
- [47] G. Gour and T. Kemp, "The minimum Rényi entropy output of a quantum channel is locally additive," *Lett. Math. Phys.*, vol. 107, no. 6, pp. 1131–1155, Jun. 2017.
- [48] A. S. Holevo and R. F. Werner, "Evaluating capacities of bosonic Gaussian channels," *Phys. Rev. A, Gen. Phys.*, vol. 63, Feb. 2001, Art. no. 032312.
- [49] J. Watrous, "Simpler semidefinite programs for completely bounded norms," *Chicago J. Theor. Comput. Sci.*, vol. 8, pp. 1–19, Jan. 2013.
- [50] X. Wang, K. Fang, and R. Duan, "Semidefinite programming converse bounds for quantum communication," *IEEE Trans. Inf. Theory*, vol. 65, no. 4, pp. 2583–2592, Apr. 2018.
- [51] K. Fang and H. Fawzi, "Geometric Rényi divergence and its applications in quantum channel capacities," *Commun. Math. Phys.*, vol. 384, pp. 1615–1677, May 2021.
- [52] D. Ding, S. Khatri, Y. Quek, P. W. Shor, X. Wang, and M. M. Wilde, "Bounding the forward classical capacity of bipartite quantum channels," *IEEE Trans. Inf. Theory*, early access, Jan. 3, 2023, doi: [10.1109/TIT.2022.3233924](https://doi.org/10.1109/TIT.2022.3233924).
- [53] L. Gao, M. Junge, and N. LaRacune, "Capacity estimates via comparison with TRO channels," *Commun. Math. Phys.*, vol. 364, no. 1, pp. 83–121, Nov. 2018.
- [54] C. King, "An application of the Lieb-Thirring inequality in quantum information theory," in *Proc. XIVth Int. Congr. Math. Phys.*, Mar. 2006, pp. 486–490.
- [55] C. H. Bennett, D. P. DiVincenzo, and J. A. Smolin, "Capacities of quantum erasure channels," *Phys. Rev. Lett.*, vol. 78, pp. 3217–3220, Apr. 1997.
- [56] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, "General paradigm for distilling classical key from quantum states," *IEEE Trans. Inf. Theory*, vol. 55, no. 4, pp. 1898–1929, Apr. 2009.
- [57] M. Ozols, G. Smith, and J. A. Smolin, "Bound entangled states with a private key and their classical counterpart," *Phys. Rev. Lett.*, vol. 112, Mar. 2014, Art. no. 110502.
- [58] M. Tomamichel, M. M. Wilde, and A. Winter, "Strong converse rates for quantum communication," *IEEE Trans. Inf. Theory*, vol. 63, no. 1, pp. 715–727, Jan. 2017.
- [59] M. M. Wilde and A. Winter, "Strong converse for the quantum capacity of the erasure channel for almost all codes," in *Proc. 9th Conf. Theory Quantum Comput., Commun. Cryptogr. (TQC)*, 2014, pp. 52–66.
- [60] R. König and S. Wehner, "A strong converse for classical channel coding using entangled inputs," *Phys. Rev. Lett.*, vol. 103, no. 7, 2009, Art. no. 070504.
- [61] M. M. Wilde, A. Winter, and D. Yang, "Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy," *Commun. Math. Phys.*, vol. 331, no. 2, pp. 593–622, Jul. 2014.
- [62] V. Siddhu and R. B. Griffiths, "Degradable quantum channels using pure-state to product-of-pure-state isometries," *Phys. Rev. A, Gen. Phys.*, vol. 94, no. 5, Nov. 2016, Art. no. 052331.
- [63] M. M. Wolf, *Quantum Channels and Operations—Guided Tour* (Lecture Notes). 2012. [Online]. Available: <https://www-m5.ma.tum.de/foswiki/pub/M5/Allgemeines/MichaelWolf/QChannelLecture.pdf>
- [64] H. Fawzi, private communication, 2020.
- [65] D. Sutter, V. B. Scholz, A. Winter, and R. Renner, "Approximate degradable quantum channels," *IEEE Trans. Inf. Theory*, vol. 63, no. 12, pp. 7832–7844, Dec. 2017.

Felix Leditzky received the Ph.D. degree from the University of Cambridge in 2017. From 2016 to 2019, he was a Post-Doctoral Research Associate with the JILA, University of Colorado Boulder. From 2019 to 2020, he was a Joint Post-Doctoral Fellow with the Institute for Quantum Computing, University of Waterloo, and the Perimeter Institute for Theoretical Physics, Waterloo, ON, Canada. He is currently an Assistant Professor with the Department of Mathematics, University of Illinois at Urbana-Champaign, and a member of the Illinois Quantum Information Science and Technology Center. His research interests include mathematical and information-theoretic aspects of quantum information theory.

Debbie Leung received the Ph.D. degree in physics from Stanford University in 2000. She joined the Institute for Quantum Computing (IQC) and the Department of Combinatorics and Optimization, University of Waterloo, in 2005. Before that, she was a Tolman Post-Doctoral Fellow with the Institute for Quantum Information, California Institute of Technology. She held a post-doctoral position with the Physics of Information Group, IBM Thomas J. Watson Research Center. She is currently a Professor with the University of Waterloo.

Vikesh Siddhu received the dual B.S. and M.S. degrees in physics from the Indian Institute of Science Education and Research Mohali, India, in 2013, and the M.S. and Ph.D. degrees in physics from Carnegie Mellon University (CMU), Pittsburgh, PA, USA, in 2015 and 2020, respectively. He is currently a Post-Doctoral Researcher at the IBM Quantum, IBM T. J. Watson Research Center. Previously, he held a post-doctoral positions at the JILA, University of Colorado/NIST, and the Tepper School of Business, CMU. His research interests include quantum information theory, (non) convex optimization, and related subjects.

Graeme Smith received the Ph.D. degree in theoretical physics from Caltech in 2006. He is currently an Associate Professor of physics with the University of Colorado Boulder and a JILA Fellow. He was previously a Research Staff Member (2010–2016) and a Postdoctoral Researcher (2007–2010) at the IBM's T. J. Watson Research Center. He was a Research Associate in computer science at the University of Bristol from 2006 to 2007. His research interests include computation and communication in noisy settings, quantum Shannon theory, coding theory, cryptography, quantum estimation and detection, optical communications, and the physics of information. He was awarded an NSF CAREER Award in 2016.

John A. Smolin received the S.B. degree in physics from the Massachusetts Institute of Technology in 1989 and the Ph.D. degree in physics from the University of California at Los Angeles in 1996. After receiving his doctorate, he has been at the IBM Research as a Postdoctoral Researcher and later a member of the Research Staff ever since. He is currently a Research Staff Member at the IBM and an APS Fellow. He and Charles H. Bennett built the first quantum cryptography apparatus at IBM in 1989. His current research interests include quantum information theory and quantum computation, with the occasional misguided foray into the foundations of quantum mechanics.