1-1-2012

The quantum dynamic capacity formula of a quantum channel

Mark M. Wilde
Université McGill

Min Hsiu Hsieh
Japan Science and Technology Agency

Recommended Citation

The quantum dynamic capacity formula of a quantum channel

Mark M. Wilde · Min-Hsiu Hsieh

October 31, 2018

Abstract The dynamic capacity theorem characterizes the reliable communication rates of a quantum channel when combined with the noiseless resources of classical communication, quantum communication, and entanglement. In prior work, we proved the converse part of this theorem by making contact with many previous results in the quantum Shannon theory literature. In this work, we prove the theorem with an "ab initio" approach, using only the most basic tools in the quantum information theorist's toolkit: the Alicki-Fannes' inequality, the chain rule for quantum mutual information, elementary properties of quantum entropy, and the quantum data processing inequality. The result is a simplified proof of the theorem that should be more accessible to those unfamiliar with the quantum Shannon theory literature. We also demonstrate that the "quantum dynamic capacity formula" characterizes the Pareto optimal trade-off surface for the full dynamic capacity region. Additivity of this formula simplifies the computation of the trade-off surface, and we prove that its additivity holds for the quantum Hadamard channels and the quantum erasure channel. We then determine exact expressions for and plot the dynamic capacity region of the quantum dephasing channel, an example from the Hadamard class, and the quantum erasure channel.

PACS 03.67.Hk · 03.67.Pp

1 Introduction

Quantum Shannon theory is the study of the transmission capabilities of a noisy resource when a large number of independent and identically distributed (IID) copies of the resource are available.[1] An important task in this area of study is to determine how noiseless resources interact with a noisy quantum channel. That is, we would like to know the reliable communication rates if a sender can use noiseless resources in addition to a quantum channel to generate other noiseless resources. In prior work, we have studied one such setting, where a sender and receiver generate or consume classical communication, quantum communication, and entanglement along with the consumption of a noisy quantum channel [16, 15, 17]. The result of these efforts was a characterization of the "dynamic capacity region" of a noisy quantum channel.[2]

One of the shortcomings of the characterization of the dynamic capacity region in Refs. [16, 15, 17] is that its computation for a general quantum channel requires regularized formulas (which are over an infinite number of channel uses). Though, later, Brádler *et al.* demonstrated that the quantum

[1] The first few chapters of Yard's thesis provide an introductory and accessible overview of the subject [27].

[2] "Dynamic" in this context and throughout this paper refers to the fact that a noisy channel is a dynamic resource, as opposed to a "static" resource such as a shared bipartite state.

Hadamard channels [18] are a natural class of channels for which the computation of one octant of the region simplifies [9] because the structure of these channels appears to be "just right" for this to hold. The proof considers special quadrants of the dynamic capacity region and employs several *reductio ad absurdum* arguments to characterize one octant of the full region. The work of Brádler *et al.* showed that we can claim a complete understanding of the abilities of an entanglement-assisted quantum Hadamard channel for the transmission of classical and quantum information.

The aim of the present work is two-fold: 1) to simplify the proof of the converse part of the dynamic capacity theorem in Ref. [17] and 2) to show that there is one important formula to consider for any task involving noiseless classical communication, noiseless quantum communication, noiseless entanglement, and many uses of a noisy quantum channel. Our previous proof of the converse in Ref. [17] relies extensively on prior literature in quantum Shannon theory, perhaps making our ideas inaccessible to an audience unfamiliar with this increasingly "tangled web." Here, we apply an "ab initio" approach to the proof, using only four tools from quantum information theory: the Alicki-Fannes' inequality [1], the chain rule for quantum mutual information, elementary properties of quantum entropy, and the quantum data processing inequality [23]. As such, the proof here should be more accessible to a broader audience and more streamlined because it is "disentangled" from the complex web that the quantum Shannon theory literature has become.

We also propose a new formula that characterizes any task involving classical communication, quantum communication, and entanglement in dynamic quantum Shannon theory.[3] For this reason, we call this formula the "quantum dynamic capacity formula." In particular, additivity of this formula implies a complete understanding of any task in dynamic quantum Shannon theory involving the three fundamental noiseless resources. We find a simplified, direct proof that additivity holds for the Hadamard class of channels and for a quantum erasure channel. The additivity proof for the quantum erasure channel is different from that of the Hadamard channel—it exploits the particular structure of the quantum erasure channel.

We structure this work as follows. The next section reviews the minimal tools from quantum information theory necessary to understand the rest of the paper. Section 3 outlines the information processing task considered in this paper, defines what it means for a rate triple to be achievable, and provides a definition of the dynamic capacity region of a quantum channel. Section 4 states the dynamic capacity theorem, and Section 5 contains a brief review of the proof of the achievability part of the theorem. The main protocol for proving this part is the "classically-enhanced father protocol," whose detailed proof we gave in Ref. [16]. Section 6 contains the converse proof, where we proceed with the minimal tools stated above. We then show in Section 7 how the quantum dynamic capacity formula characterizes the optimization task for computing the Pareto optimal trade-off surface for the dynamic capacity region. Section 8 proves that the quantum dynamic capacity formula is additive for the Hadamard class of channels, and in Section 9, we directly compute and plot the region for a qubit dephasing channel, which is a channel that falls within the Hadamard class. Section 10 then shows that the quantum dynamic capacity is additive for the quantum erasure channel, and we compute and plot the region for this channel also. Finally, we conclude with a brief discussion.

2 Definitions and notation

We first establish some definitions and notation that we employ throughout the paper and review a few important properties of quantum entropy. Let Φ^{AB} denote the maximally entangled state shared between two parties:

$$|\Phi\rangle^{AB} \equiv \frac{1}{\sqrt{D}} \sum_{i=1}^{D} |i\rangle^A |i\rangle^B .$$

An ebit corresponds to the special case where $D = 2$. Let $\overline{\Phi}^{M_A M_B}$ denote the maximally correlated state shared between two parties:

$$\overline{\Phi}^{M_A M_B} \equiv \frac{1}{D} \sum_{i=1}^{D} |i\rangle\langle i|^{M_A} \otimes |i\rangle\langle i|^{M_B} .$$

---

[3] "Dynamic quantum Shannon theory" refers to the setting in which a sender and a receiver have access to many uses of a quantum channel connecting them.

A common randomness bit corresponds to the special case where $D = 2$.

A completely-positive trace-preserving (CPTP) map $\mathcal{N}^{A' \to B}$ is the most general map we consider that maps from a quantum system A' to another quantum system B [22]. It acts as follows on any density operator ρ:

$$\mathcal{N}^{A' \to B}(\rho) = \sum_k A_k \rho A_k^\dagger,$$

where the operators A_k satisfy the condition $\sum_k A_k^\dagger A_k = I$. A quantum channel admits an isometric extension $U_{\mathcal{N}}^{A' \to BE}$, which is a unitary embedding into a larger Hilbert space. One recovers the original channel by taking a partial trace over the "environment" system E.

We consider a three-dimensional capacity region throughout this work (as in Ref. [17]), whose points (C, Q, E) correspond to rates of classical communication, quantum communication, and entanglement generation/consumption, respectively. For example, the teleportation protocol consumes two classical bits and an ebit to generate a noiseless qubit [3]. Thus, we write it as the following rate triple:

$$(-2, 1, -1),$$

where we indicate consumption of a resource with a negative sign and generation of a resource with a positive sign. Also, the super-dense coding protocol consumes a noiseless qubit channel and an ebit to generate two classical bits [6]. It corresponds to the rate triple:

$$(2, -1, -1).$$

Another protocol that we exploit is entanglement distribution. It uses a noiseless qubit channel to establish a noiseless ebit and corresponds to

$$(0, -1, 1).$$

The entropy $H(A)_\rho$ of a density operator ρ^A on some quantum system A is as follows [22]:

$$H(A)_\rho \equiv -\text{Tr}\left\{\rho^A \log \rho^A\right\},$$

where the logarithm is base two. The entropy can never exceed the logarithm of the dimension of A. The quantum mutual information $I(A; B)_\sigma$ of a bipartite state σ^{AB} is as follows:

$$I(A; B)_\sigma \equiv H(A)_\sigma + H(B)_\sigma - H(AB)_\sigma.$$

Observe that the quantum mutual information $I(M_A; M_B)$ of the state $\overline{\Phi}^{M_A M_B}$ is equal to $\log D$ bits, and the quantum mutual information $I(A; B)$ of the state Φ^{AB} is equal to $2 \log D$ qubits. If one system is classical, then the quantum mutual information can never be greater than the logarithm of the dimension of the classical system. If both systems are quantum, then the quantum mutual information can never be greater than twice the minimum of the logarithms of the dimensions of the two quantum systems. The quantum mutual information vanishes if the bipartite state σ^{AB} is a product state. The conditional quantum mutual information for three quantum systems A, B, and C is as follows:

$$I(A; B|C) \equiv H(AC) + H(BC) - H(C) - H(ABC),$$

and is always non-negative due to strong subadditivity [20]. The coherent information $I(A \rangle B)_\sigma$ of a state σ^{AB} is as follows [23]:

$$I(A \rangle B)_\sigma \equiv H(B)_\sigma - H(AB)_\sigma.$$

Observe that the coherent information $I(A \rangle B)$ of the state Φ^{AB} is equal to $\log D$ qubits. The chain rule for quantum mutual information gives the following relation for any three quantum systems A, B, and C:

$$\begin{align} I(AB; C) &= I(B; C) + I(A; C|B) \tag{1}\\ &= I(A; C) + I(B; C|A). \end{align}$$

A classical-quantum state σ^{XABE} of the following form plays an important role throughout this paper:

$$\sigma^{XABE} \equiv \sum_x p_X(x) |x\rangle \langle x|^X \otimes U_{\mathcal{N}}^{A' \to BE}(\phi_x^{AA'}),$$

where the states $\phi_x^{AA'}$ are pure bipartite states and $U_{\mathcal{N}}^{A' \to BE}$ is the isometric extension of some noisy channel $\mathcal{N}^{A' \to B}$. Applying the above chain rule gives the following relation:

$$I(AX;B)_\sigma = I(X;B)_\sigma + I(A;B|X)_\sigma. \qquad (2)$$

Additionally, one can readily check that the following relation holds for a state of the above form:

$$I(A\rangle BX)_\sigma = \frac{1}{2}I(A;B|X)_\sigma - \frac{1}{2}I(A;E|X)_\sigma. \qquad (3)$$

The Alicki-Fannes' inequality is a statement of the continuity of coherent information [1], and a simple variant of it gives continuity of quantum mutual information. First, suppose that two bipartite states ρ^{AB} and σ^{AB} are ϵ-close in trace norm:

$$\left\| \rho^{AB} - \sigma^{AB} \right\|_1 \le \epsilon.$$

Then the Alicki-Fannes' inequality states that their respective coherent informations are close:

$$|I(A\rangle B)_\rho - I(A\rangle B)_\sigma| \le 4\epsilon \log|A| + 2H_2(\epsilon),$$

where $|A|$ is the dimension of the system A and H_2 is the binary entropy function. A simple "tweak" of the above inequality shows that the quantum mutual informations are close [16]:

$$|I(A;B)_\rho - I(A;B)_\sigma| \le 5\epsilon \log|A| + 3H_2(\epsilon).$$

The quantum data processing inequality states that quantum correlations can never increase under the application of a noisy map [23, 22]. There are two important manifestations of it. Suppose that Alice possesses two quantum systems A and A' in her lab. She sends A' through a noisy quantum channel to Bob and he receives it in some system B. Then the following inequality applies to the coherent information:

$$I(A\rangle A') \ge I(A\rangle B),$$

demonstrating that quantum correlations, as measured by the coherent information, can only decrease under noisy processing. The following inequality also applies to the quantum mutual information:

$$I(A;A') \ge I(A;B),$$

demonstrating a similar notion for the quantum mutual information. Now suppose that Alice sends A to Christabelle and she receives it in some system C. Then the following inequality applies as well:

$$I(A;A') \ge I(C;A') \ge I(C;B),$$

but a similar inequality does not hold for the coherent information, due to its asymmetry. We make extensive use of quantum data processing in our proofs.

3 The information processing task

We are interested in the most general protocol that generates classical communication, quantum communication, and entanglement by consuming many uses of a noisy quantum channel $\mathcal{N}^{A' \to B}$ and the same respective resources (see Figure 1). We say that such a protocol is "catalytic" because we are allowing it to consume the same resources that it generates, though we are keeping track of the net rates of consumption or generation.

The protocol begins with Alice possessing two classical registers (each labeled by M and of dimension $2^{n\bar{C}}$), a quantum register A_1 of dimension $2^{n\bar{Q}}$ entangled with a reference system R, and another quantum register T_A of dimension $2^{n\bar{E}}$ that contains her half of the shared entanglement with Bob:

$$\omega^{MMRA_1T_AT_B} \equiv \overline{\Phi}^{MM} \otimes \Phi^{RA_1} \otimes \Phi^{T_AT_B}.$$

She passes one of the classical registers and the registers A_1 and T_A into a CPTP encoding map $\mathcal{E}^{MA_1T_A \to A'^n S_A L A_2}$ that outputs a quantum register S_A of dimension $2^{n\bar{E}}$ and a quantum register A_2 of dimension $2^{n\bar{Q}}$, a classical register L of dimension $2^{n\tilde{C}}$, and many quantum systems A'^n for input to the
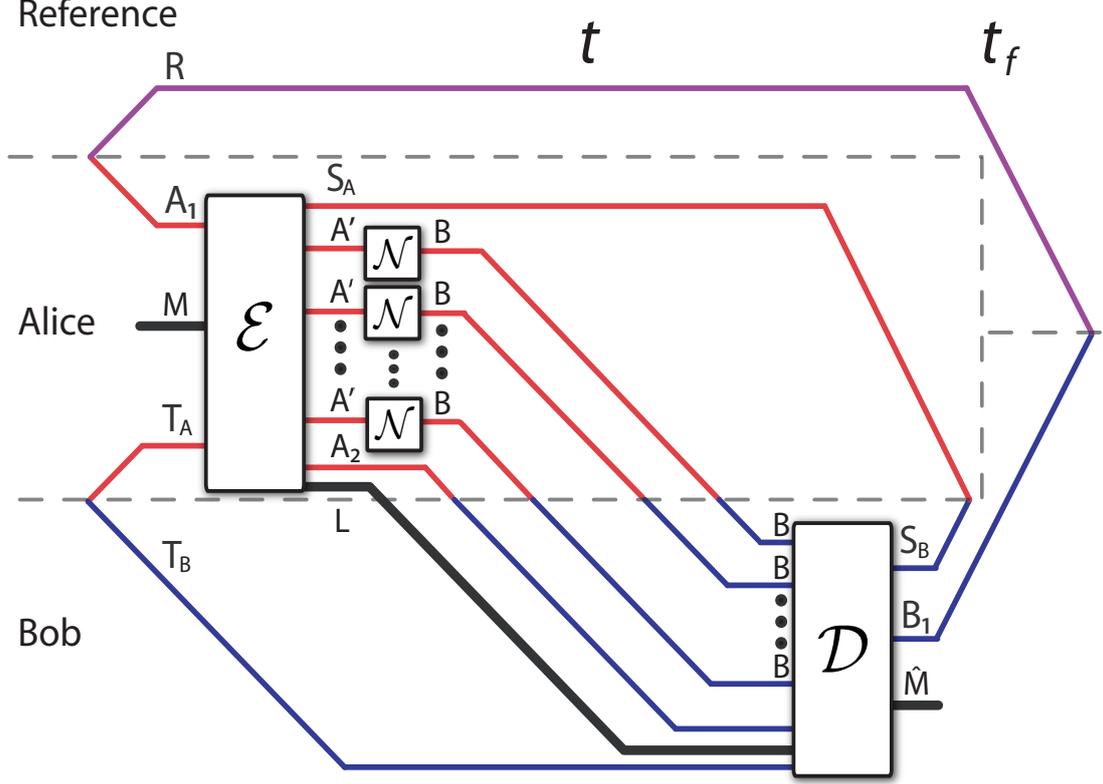
Fig. 1 (Color online) The most general protocol for generating classical communication, quantum communication, and entanglement with the help of the same respective resources and many uses of a noisy quantum channel. Alice begins with her classical register M, her quantum register A_1, and her half of the shared entanglement in register T_A. She encodes according to some CPTP map $\mathcal{E}$ that outputs a quantum register S_A, many registers A'^n, a quantum register A_2, and a classical register L. She inputs A'^n to many uses of the noisy channel $\mathcal{N}$ and transmits A_2 over a noiseless quantum channel and L over a noiseless classical channel. Bob receives the channel outputs B^n, the quantum register A_2, and the classical register L and performs a decoding $\mathcal{D}$ that recovers the quantum information and classical message. The decoding also generates entanglement with system S_A. Many protocols are a special case of the above one. For example, the protocol is entanglement-assisted communication of classical and quantum information [16] if the registers L, S_A, S_B, and A_2 are null.

channel. The register S_A is for creating entanglement with Bob. The state after the encoding map $\mathcal{E}$ is as follows:

$$\omega^{MA'^n S_A L A_2 R T_B} \equiv \mathcal{E}^{MA_1 T_A \to A'^n S_A L A_2}(\omega^{MMRA_1 T_A T_B}).$$

She sends the systems A'^n through many uses $\mathcal{N}^{A'^n \to B^n}$ of the noisy channel $\mathcal{N}^{A' \to B}$, transmits L over a noiseless classical channel, and transmits A_2 over a noiseless quantum channel, producing the following state:

$$\omega^{MB^n S_A L A_2 R T_B} \equiv \mathcal{N}^{A'^n \to B^n}(\omega^{MA'^n S_A L A_2 R T_B}).$$

The above state is a state of the form in (15) with $A \equiv RT_B A_2 S_A$ and $X \equiv ML$. Bob then applies a map $\mathcal{D}^{B^n A_2 T_B L \to B_1 S_B \hat{M}}$ that outputs a quantum system B_1, a quantum system S_B, and a classical register $\hat{M}$. Let ω' denote the final state. The following condition holds for a good protocol:

$$\left\| \overline{\Phi}^{M\hat{M}} \otimes \Phi^{RB_1} \otimes \Phi^{S_A S_B} - (\omega')^{MB_1 S_B \hat{M} S_A R} \right\|_1 \leq \epsilon, \tag{4}$$

implying that Alice and Bob establish maximal classical correlations in M and $\hat{M}$ and maximal entanglement between S_A and S_B. The above condition also implies that the coding scheme preserves the entanglement with the reference system R. The net rate triple for the protocol is as follows: $\left(\bar{C} - \tilde{C}, \bar{Q} - \tilde{Q}, \bar{E} - \tilde{E} \right)$.

The protocol generates a resource if its corresponding rate is positive, and it consumes a resource if its corresponding rate is negative. Such a protocol defines an (n, C, Q, E, ϵ) code with

$$C = \bar{C} - \tilde{C}, \tag{5}$$

$$Q = \bar{Q} - \tilde{Q}, \tag{6}$$

$$E = \bar{E} - \tilde{E}. \tag{7}$$

Definition 1 (Achievability) A rate triple (C, Q, E) is achievable if there exists an (n, C, Q, E, ϵ) code with error (as defined in (4)) smaller than ϵ for all $\epsilon > 0$ and sufficiently large n.

Definition 2 (Dynamic Capacity Region) The dynamic capacity region $\mathcal{C}_{\mathrm{CQE}}(\mathcal{N})$ of a noisy quantum channel $\mathcal{N}^{A' \to B}$ is a three-dimensional region in the (C, Q, E) space defined by the closure of the set of all achievable rate triples (C, Q, E).

4 The dynamic capacity theorem

The dynamic capacity theorem gives bounds on the reliable communication rates of a noisy quantum channel when combined with the noiseless resources of classical communication, quantum communication, and shared entanglement [17]. The theorem applies regardless of whether a protocol consumes the noiseless resources or generates them.

Theorem 1 (Dynamic Capacity) *The dynamic capacity region $\mathcal{C}_{\mathrm{CQE}}(\mathcal{N})$ of a quantum channel $\mathcal{N}$ is equal to the following expression:*

$$\mathcal{C}_{\mathrm{CQE}}(\mathcal{N}) = \overline{\bigcup_{k=1}^{\infty} \frac{1}{k} \mathcal{C}_{\mathrm{CQE}}^{(1)}(\mathcal{N}^{\otimes k})}, \tag{8}$$

where the overbar indicates the closure of a set. The "one-shot" region $\mathcal{C}_{\mathrm{CQE}}^{(1)}(\mathcal{N})$ is the union of the "one-shot, one-state" regions $\mathcal{C}_{\mathrm{CQE},\sigma}^{(1)}(\mathcal{N})$:

$$\mathcal{C}_{\mathrm{CQE}}^{(1)}(\mathcal{N}) \equiv \bigcup_{\sigma} \mathcal{C}_{\mathrm{CQE},\sigma}^{(1)}(\mathcal{N}).$$

The "one-shot, one-state" region $\mathcal{C}_{\mathrm{CQE},\sigma}^{(1)}(\mathcal{N})$ is the set of all rates C, Q, and E, such that

$$C + 2Q \leq I(AX; B)_\sigma, \tag{9}$$

$$Q + E \leq I(A \rangle BX)_\sigma, \tag{10}$$

$$C + Q + E \leq I(X; B)_\sigma + I(A \rangle BX)_\sigma. \tag{11}$$

The above entropic quantities are with respect to a classical-quantum state σ^{XAB} where

$$\sigma^{XAB} \equiv \sum_x p(x) |x\rangle \langle x|^X \otimes \mathcal{N}^{A' \to B}(\phi_x^{AA'}), \tag{12}$$

and the states $\phi_x^{AA'}$ are pure. It is implicit that one should consider states on A'^k instead of A' when taking the regularization in (8).

The above theorem is a "multi-letter" capacity theorem because of the regularization in (8). Though, we show in Sections 8 and 10 that the regularization is not necessary for the Hadamard class of channels or the quantum erasure channels, respectively. We prove the above theorem in two parts:

1. The direct coding theorem below shows that combining the "classically-enhanced father protocol" with teleportation, super-dense coding, and entanglement distribution achieves the above region.
2. The converse theorem demonstrates that any coding scheme cannot do better than the regularization in (8), in the sense that a scheme with vanishing error should have its rates below the above amounts. We prove the converse theorem directly in "one fell swoop," by employing a catalytic, information-theoretic approach. The converse proof is different from our earlier one [17] because we employ straightforward information-theoretic arguments instead of making contact with prior quantum Shannon theoretic literature.

5 Dynamic achievable rate region

The unit resource achievable region is what Alice and Bob can achieve with the protocols entanglement distribution, teleportation, and super-dense coding [17]. It is the cone of the rate triples corresponding to these protocols:

$$\{\alpha\,(0, -1, 1) + \beta\,(2, -1, -1) + \gamma\,(-2, 1, -1) : \alpha, \beta, \gamma \geq 0\}.$$

We can also write any rate triple (C, Q, E) in the unit resource capacity region with a matrix equation:

$$\begin{bmatrix} C \\ Q \\ E \end{bmatrix} = \begin{bmatrix} 0 & 2 & -2 \\ -1 & -1 & 1 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix}. \tag{13}$$

The inverse of the above matrix is as follows:

$$\begin{bmatrix} -\frac{1}{2} & -1 & 0 \\ 0 & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix},$$

and gives the following set of inequalities for the unit resource achievable region:

$$C + 2Q \leq 0,$$
$$Q + E \leq 0,$$
$$C + Q + E \leq 0,$$

by inverting the matrix equation in (13) and applying the constraints $\alpha, \beta, \gamma \geq 0$.

Now, let us include the classically-enhanced father protocol [16]. Ref. [16] proved that we can achieve the following rate triple by channel coding over a noisy quantum channel $\mathcal{N}^{A' \to B}$:

$$\left(I\,(X; B)_\sigma\,, \frac{1}{2} I\,(A; B|X)_\sigma\,, -\frac{1}{2} I\,(A; E|X)_\sigma\right),$$

for any state σ^{XABE} of the form:

$$\sigma^{XABE} \equiv \sum_x p_X\,(x)\,|x\rangle\,\langle x|^X \otimes U_{\mathcal{N}}^{A' \to BE}(\phi_x^{AA'}), \tag{14}$$

where $U_{\mathcal{N}}^{A' \to BE}$ is an isometric extension of the quantum channel $\mathcal{N}^{A' \to B}$. Specifically, we showed in Ref. [16] that one can achieve the above rates with vanishing error in the limit of large blocklength. Thus the achievable rate region is the following translation of the unit resource achievable region in (13):

$$\begin{bmatrix} C \\ Q \\ E \end{bmatrix} = \begin{bmatrix} 0 & 2 & -2 \\ -1 & -1 & 1 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} + \begin{bmatrix} I\,(X; B)_\sigma \\ \frac{1}{2} I\,(A; B|X)_\sigma \\ -\frac{1}{2} I\,(A; E|X)_\sigma \end{bmatrix}.$$

We can now determine bounds on an achievable rate region that employs the above coding strategy. We apply the inverse of the matrix in (13) to the LHS and RHS. Then using (2), (3), and the constraints $\alpha, \beta, \gamma \geq 0$, we obtain the inequalities in (9-11), corresponding exactly to the one-shot, one-state region in Theorem 1. Taking the union over all possible states σ in (14) and taking the regularization gives the full dynamic achievable rate region.

6 Catalytic and information theoretic converse proof

This section begins one of the main contributions of this work. We provide a catalytic, information theoretic converse proof of the dynamic capacity region, showing that (8) gives a multi-letter characterization of it. The catalytic approach means that we are considering the most general protocol that *consumes and generates* classical communication, quantum communication, and entanglement in addition to the uses of the noisy quantum channel. This approach has the advantage that we can prove the converse theorem in "one fell swoop" rather than considering one octant of the (C, Q, E) space at a time as we did in Ref. [17]. Additionally, we do not need to make contact with prior work in quantum Shannon theory. We employ the Alicki-Fannes' inequality, the chain rule for quantum mutual information, elementary properties of quantum entropy, and the quantum data processing inequality to prove the converse.

There are some subtleties in our proof for the converse theorem. We prove that the bounds in (9-11) hold for common randomness generation instead of classical communication because a capacity for generating common randomness can only be better than that for generating classical communication (classical communication can generate common randomness). We also consider a protocol that preserves entanglement with a reference system instead of one that generates quantum communication. Barnum *et al.* showed that this task is equivalent to the transmission of quantum information [2].

We prove that the converse theorem holds for a state of the following form

$$\sigma^{XAB} \equiv \sum_x p(x) |x\rangle \langle x|^X \otimes \mathcal{N}^{A' \to B}(\rho_x^{AA'}), \tag{15}$$

where the states $\rho_x^{AA'}$ are mixed, rather than proving it for a state of the form in (14). Then we show in Section 6.1 that it is not necessary to consider an ensemble of mixed states—i.e., we can do just as well with an ensemble of pure states, giving the statement of Theorem 1.

We begin by proving the first bound in (9). All system labels are as given in Section 3, and we consider the most general protocol as outlined in that section. Consider the following chain of inequalities:

$$
\begin{aligned}
n\left(\bar{C} + 2\bar{Q}\right) &= I(M; \hat{M})_{\overline{\Phi}} + I\left(R; B_1\right)_{\Phi} \\
&\leq I(M; \hat{M})_{\omega'} + I\left(R; B_1\right)_{\omega'} + n\delta' \\
&\leq I\left(M; B^n A_2 L T_B\right)_\omega + I\left(R; B^n A_2 L T_B\right)_\omega \\
&\leq I\left(M; B^n A_2 L T_B\right)_\omega + I\left(R; B^n A_2 L T_B M\right)_\omega \\
&= I\left(M; B^n A_2 L T_B\right)_\omega + I\left(R; B^n A_2 L T_B | M\right)_\omega + I\left(R; M\right)_\omega.
\end{aligned}
$$

The first equality holds by evaluating the quantum mutual informations on the respective states $\overline{\Phi}^{M\hat{M}}$ and Φ^{RB_1}. The first inequality follows from the condition in (4) and an application of the Alicki-Fannes' inequality where δ' vanishes as $\epsilon \to 0$. We suppress this term in the rest of the inequalities for convenience. The second inequality follows from quantum data processing, and the third follows from another application of quantum data processing. The second equality follows by applying the mutual information chain rule in (1). We continue below:

$$
\begin{aligned}
&= I\left(M; B^n A_2 L T_B\right)_\omega + I\left(R; B^n A_2 L T_B | M\right)_\omega \\
&= I\left(M; B^n A_2 L T_B\right)_\omega + I\left(R A_2 L T_B; B^n | M\right)_\omega + I\left(R; A_2 L T_B | M\right)_\omega - I\left(B^n; A_2 L T_B | M\right)_\omega \\
&= I\left(M; B^n\right)_\omega + I\left(M; A_2 L T_B | B^n\right)_\omega + I\left(R A_2 L T_B; B^n | M\right)_\omega + I\left(R; A_2 L T_B | M\right)_\omega - I\left(B^n; A_2 L T_B | M\right)_\omega \\
&= I\left(R A_2 L M T_B; B^n\right)_\omega + I\left(M; A_2 L T_B | B^n\right)_\omega + I\left(R; A_2 L T_B | M\right)_\omega - I\left(B^n; A_2 L T_B | M\right)_\omega \\
&\leq I\left(R A_2 T_B S_A L M; B^n\right)_\omega + I\left(M; A_2 L T_B | B^n\right)_\omega + I\left(R; A_2 L T_B | M\right)_\omega - I\left(B^n; A_2 L T_B | M\right)_\omega \\
&= I\left(AX; B^n\right)_\omega + I\left(M; A_2 L T_B | B^n\right)_\omega + I\left(R; A_2 L T_B | M\right)_\omega - I\left(B^n; A_2 L T_B | M\right)_\omega.
\end{aligned}
$$

The first equality follows because $I\left(R; M\right)_\omega = 0$ for this protocol. The second equality follows from applying the chain rule for quantum mutual information to the term $I\left(R; B^n A_2 L T_B | M\right)_\omega$, and the third is another application of the chain rule to the term $I\left(M; B^n A_2 L T_B\right)_\omega$. The fourth equality follows by combining $I\left(M; B^n\right)_\omega$ and $I\left(R A_2 L T_B; B^n | M\right)_\omega$ with the chain rule. The inequality follows from an application of quantum data processing. The final equality follows from the definitions $A \equiv R T_B A_2 S_A$

and $X \equiv ML$. We now focus on the term $I\left(M; A_2 L T_B | B^n\right)_\omega + I\left(R; A_2 L T_B | M\right)_\omega - I\left(B^n; A_2 L T_B | M\right)_\omega$ and show that it is less than $n\left(\tilde{C} + 2\tilde{Q}\right)$:

$$
\begin{aligned}
& I\left(M; A_2 L T_B | B^n\right)_\omega + I\left(R; A_2 L T_B | M\right)_\omega - I\left(B^n; A_2 L T_B | M\right)_\omega \\
&= I\left(M; A_2 L T_B B^n\right)_\omega + I\left(R; A_2 L T_B M\right)_\omega - I\left(B^n; A_2 L T_B M\right)_\omega - I\left(R; M\right)_\omega \\
&= I\left(M; A_2 L T_B B^n\right)_\omega + I\left(R; A_2 L T_B M\right)_\omega - I\left(B^n; A_2 L T_B M\right)_\omega \\
&= H\left(A_2 L T_B B^n\right)_\omega + H\left(R\right)_\omega - H\left(R A_2 L T_B | M\right)_\omega - H\left(B^n\right)_\omega \\
&= H\left(A_2 L T_B B^n\right)_\omega - H\left(A_2 L T_B | M R\right)_\omega - H\left(B^n\right)_\omega \\
&= H\left(A_2 L T_B | B^n\right)_\omega - H\left(A_2 L T_B | M R\right)_\omega .
\end{aligned}
$$

The first equality follows by applying the chain rule for quantum mutual information. The second equality follows because $I\left(R; M\right)_\omega = 0$ for this protocol. The third equality follows by expanding the quantum mutual informations. The next two inequalities follow from straightforward entropic manipulations and that $H\left(R\right)_\omega = H\left(R | M\right)_\omega$ for this protocol. We continue below:

$$
\begin{aligned}
&= H\left(A_2 L | B^n\right)_\omega + H\left(T_B | B^n A_2 L\right)_\omega - H\left(T_B | M R\right)_\omega - H\left(A_2 L | T_B M R\right)_\omega \\
&= H\left(A_2 L | B^n\right)_\omega + H\left(T_B | B^n A_2 L\right)_\omega - H\left(T_B\right)_\omega - H\left(A_2 L | T_B M R\right)_\omega \\
&= H\left(A_2 L | B^n\right)_\omega - I\left(T_B; B^n A_2 L\right)_\omega - H\left(A_2 L | T_B M R\right) \\
&\leq H\left(A_2 L\right)_\omega - H\left(A_2 L | T_B M R\right)_\omega \\
&= I\left(A_2 L; T_B M R\right)_\omega \\
&= I\left(L; T_B M R\right)_\omega + I\left(A_2; T_B M R | L\right)_\omega \\
&\leq n\left(\tilde{C} + 2\tilde{Q}\right).
\end{aligned}
$$

The first two equalities follow from the chain rule for entropy and the second exploits that $H\left(T_B | M R\right) = H\left(T_B\right)$ for this protocol. The third equality follows from the definition of quantum mutual information. The inequality follows from subadditivity of entropy and that $I\left(T_B; B^n A_2 L\right)_\omega \geq 0$. The fourth equality follows from the definition of quantum mutual information and the next equality follows from the chain rule. The final inequality follows because the quantum mutual information $I\left(L; T_B M R\right)_\omega$ can never be larger than the logarithm of the dimension of the classical register L and because the quantum mutual information $I\left(A_2; T_B M R | L\right)_\omega$ can never be larger than twice the logarithm of the dimension of the quantum register A_2. Thus the following inequality applies

$$
n\left(\bar{C} + 2\bar{Q}\right) \leq I\left(AX; B^n\right)_\omega + n\left(\tilde{C} + 2\tilde{Q}\right) + n\delta',
$$

demonstrating that (9) holds for the net rates.

We now prove the second bound in (10). Consider the following chain of inequalities:

$$
\begin{aligned}
n\left(\bar{Q} + \bar{E}\right) &= I\left(R \rangle B_1\right)_\Phi + I\left(S_A \rangle S_B\right)_\Phi \\
&= I\left(R S_A \rangle B_1 S_B\right)_{\Phi \otimes \Phi} \\
&\leq I\left(R S_A \rangle B_1 S_B\right)_{\omega'} + n\delta' \\
&\leq I\left(R S_A \rangle B_1 S_B M\right)_{\omega'} \\
&\leq I\left(R S_A \rangle B^n A_2 T_B L M\right)_\omega \\
&= H\left(B^n A_2 T_B | L M\right)_\omega - H\left(R S_A B^n A_2 T_B | L M\right)_\omega \\
&\leq H\left(B^n | L M\right)_\omega + H\left(A_2 | L M\right)_\omega + H\left(T_B | L M\right)_\omega - H\left(R S_A B^n A_2 T_B | L M\right)_\omega \\
&\leq I\left(R S_A A_2 T_B \rangle B^n L M\right)_\omega + n\left(\tilde{Q} + \tilde{E}\right) \\
&= I\left(A \rangle B^n X\right)_\omega + n\left(\tilde{Q} + \tilde{E}\right).
\end{aligned}
$$

The first equality follows by evaluating the coherent informations of the respective states $\Phi^{R B_1}$ and $\Phi^{S_A S_B}$. The second equality follows because $\Phi^{R B_1} \otimes \Phi^{T_A T_B}$ is a product state. The first inequality follows from the condition in (4) and an application of the Alicki-Fannes' inequality with δ' vanishing when $\epsilon \to 0$. We suppress the term $n\delta'$ in the following lines. The next two inequalities follow from

quantum data processing. The third equality follows from the definition of coherent information. The fourth inequality follows from subadditivity of entropy. The fifth inequality follows from the definition of coherent information and the fact that the entropy can never be larger than the logarithm of the dimension of the corresponding system. The final equality follows from the definitions $A \equiv RT_B A_2 S_A$ and $X \equiv ML$. Thus the following inequality applies

$$n \left(\bar{Q} + \bar{E} \right) \leq I \left(A \rangle B^n X \right) + n \left(\tilde{Q} + \tilde{E} \right),$$

demonstrating that (10) holds for the net rates.

We prove the last bound in (11). Consider the following chain of inequalities:

$$
\begin{aligned}
n \left(\bar{C} + \bar{Q} + \bar{E} \right) &= I(M; \hat{M})_{\overline{\Phi}} + I \left(RS_A \rangle B_1 S_B \right)_{\Phi \otimes \Phi} \\
&\leq I(M; \hat{M})_{\omega'} + I \left(RS_A \rangle B_1 S_B \right)_{\omega'} + n\delta' \\
&\leq I(M; B^n A_2 T_B L)_{\omega} + I \left(RS_A \rangle B^n A_2 T_B LM \right)_{\omega} \\
&= I \left(ML; B^n A_2 T_B \right)_{\omega} + I \left(M; L \right)_{\omega} - I \left(A_2 B^n T_B; L \right)_{\omega} \\
&\quad + H \left(B^n | LM \right) + H \left(A_2 T_B | B^n LM \right)_{\omega} - H \left(RS_A A_2 T_B B^n | LM \right)_{\omega} \\
&= I \left(ML; B^n \right)_{\omega} + I \left(ML; A_2 T_B | B^n \right)_{\omega} + I \left(M; L \right)_{\omega} - I \left(A_2 B^n T_B; L \right)_{\omega} \\
&\quad + H \left(A_2 T_B | B^n LM \right)_{\omega} + I \left(RS_A A_2 T_B \rangle B^n LM \right)_{\omega}.
\end{aligned}
$$

The first equality follows from evaluating the mutual information of the state $\overline{\Phi}^{M\hat{M}}$ and the coherent information of the product state $\Phi^{RB_1} \otimes \Phi^{S_A S_B}$. The first inequality follows from the condition in (4) and an application of the Alicki-Fannes' inequality with δ' vanishing when $\epsilon \to 0$. We suppress the term $n\delta'$ in the following lines. The second inequality follows from quantum data processing. The second equality follows from applying the chain rule for quantum mutual information to $I(M; B^n A_2 T_B L)_{\omega}$ and by expanding the coherent information $I \left(RS_A \rangle B^n A_2 T_B LM \right)_{\omega}$. The third equality follows from applying the chain rule for quantum mutual information to $I \left(ML; B^n A_2 T_B \right)_{\omega}$ and from the definition of coherent information. We continue below:

$$
\begin{aligned}
&= I \left(ML; B^n \right)_{\omega} + I \left(RS_A A_2 T_B \rangle B^n LM \right)_{\omega} \\
&\quad + I \left(ML; A_2 T_B | B^n \right)_{\omega} + I \left(M; L \right)_{\omega} - I \left(A_2 B^n T_B; L \right)_{\omega} + H \left(A_2 T_B | B^n LM \right)_{\omega} \\
&= I \left(ML; B^n \right)_{\omega} + I \left(RS_A A_2 T_B \rangle B^n LM \right)_{\omega} \\
&\quad + H \left(A_2 T_B | B^n \right)_{\omega} + I \left(M; L \right)_{\omega} - I \left(A_2 B^n T_B; L \right)_{\omega} \\
&\leq I \left(ML; B^n \right)_{\omega} + I \left(RS_A A_2 T_B \rangle B^n LM \right)_{\omega} + n \left(\tilde{C} + \tilde{Q} + \tilde{E} \right) \\
&= I \left(X; B^n \right)_{\omega} + I \left(A \rangle B^n X \right)_{\omega} + n \left(\tilde{C} + \tilde{Q} + \tilde{E} \right).
\end{aligned}
$$

The first equality follows by rearranging terms. The second equality follows by canceling terms. The inequality follows from subadditivity of the entropy $H \left(A_2 T_B | B^n \right)_{\omega}$, the fact that the entropy $H \left(A_2 T_B | B^n \right)_{\omega}$ can never be larger than the logarithm of the dimension of the systems $A_2 T_B$, that the mutual information $I \left(M; L \right)_{\omega}$ can never be larger than the logarithm of the dimension of the classical register L, and because $I \left(A_2 B^n T_B; L \right)_{\omega} \geq 0$. The last equality follows from the definitions $A \equiv RT_B A_2 S_A$ and $X \equiv ML$. Thus the following inequality holds

$$n \left(\bar{C} + \bar{Q} + \bar{E} \right) \leq I \left(X; B^n \right)_{\omega} + I \left(A \rangle B^n X \right)_{\omega} + n \left(\tilde{C} + \tilde{Q} + \tilde{E} \right) + n\delta',$$

demonstrating that the inequality in (11) applies to the net rates. This concludes the catalytic proof of the converse theorem.

6.1 Pure state ensembles are sufficient

We prove that it is sufficient to consider an ensemble of pure states as in the statement of Theorem 1 rather than an ensemble of mixed states as in (15) in the proof of our converse theorem. Our argument relies on a classic trick exploited in quantum Shannon theory [11]. We first determine a spectral decomposition of

the mixed state ensemble, model the index of the pure states in the decomposition as a classical variable Y, and then place this classical variable Y in a classical register. It follows that the communication rates can only improve, and it is sufficient to consider an ensemble of pure states.

Consider that each mixed state in the ensemble in (15) admits a spectral decomposition of the following form:

$$\rho_x^{AA'} = \sum_y p\left(y|x\right) \psi_{x,y}^{AA'}.$$

We can thus represent the ensemble as follows:

$$\rho^{XAB} \equiv \sum_{x,y} p(x)p\left(y|x\right) |x\rangle \langle x|^X \otimes \mathcal{N}^{A'\rightarrow B}(\psi_{x,y}^{AA'}). \tag{16}$$

The inequalities in (9-11) for the dynamic capacity region involve the mutual information $I(AX;B)_\rho$, the Holevo information $I(X;B)_\rho$, and the coherent information $I(A\rangle BX)_\rho$. As we show below, each of these entropic quantities can only improve in each case if make the variable y be part of the classical variable. This improvement then implies that it is only necessary to consider pure states in the dynamic capacity theorem.

Let θ^{XYAB} denote an augmented state of the following form:

$$\theta^{XYAB} \equiv \sum_x p(x)p\left(y|x\right) |x\rangle \langle x|^X \otimes |y\rangle \langle y|^Y \otimes \mathcal{N}^{A'\rightarrow B}(\psi_{x,y}^{AA'}). \tag{17}$$

This state is actually a state of the form in (12) if we subsume the classical variables X and Y into one classical variable. The following three inequalities each follow from an application of the quantum data processing inequality:

$$I(X;B)_\rho = I(X;B)_\theta \le I(XY;B)_\theta, \tag{18}$$

$$I(AX;B)_\rho = I(AX;B)_\theta \le I(AXY;B)_\theta \tag{19}$$

$$I(A\rangle BX)_\rho = I(A\rangle BX)_\theta \le I(A\rangle BXY)_\theta. \tag{20}$$

Each of these inequalities proves the desired result for the respective Holevo information, mutual information, and coherent information, and it suffices to consider an ensemble of pure states in Theorem 1.

7 The quantum dynamic capacity formula

We introduce the quantum dynamic capacity formula and show how additivity of it implies that the computation of the Pareto optimal trade-off surface of the capacity region requires just a single channel use, rather than an infinite number of them (as in regularized formulas). The Pareto optimal trade-off surface consists of all points in the capacity region that are Pareto optimal, in the sense that it is not possible to make improvements in one resource without offsetting another resource (these are essentially the boundary points of the region in our case). We then show how several important capacity formulas in the quantum Shannon theory literature are special cases of the quantum dynamic capacity formula.

Definition 3 (Quantum Dynamic Capacity Formula) The quantum dynamic capacity formula of a quantum channel $\mathcal{N}$ is as follows:

$$D_{\lambda,\mu}\left(\mathcal{N}\right) \equiv \max_\sigma I\left(AX;B\right)_\sigma + \lambda I\left(A\rangle BX\right)_\sigma + \mu\left(I\left(X;B\right)_\sigma + I\left(A\rangle BX\right)_\sigma\right), \tag{21}$$

where $\lambda, \mu \ge 0$.

Definition 4 The regularized quantum dynamic capacity formula is as follows:

$$D_{\lambda,\mu}^{\text{reg}}\left(\mathcal{N}\right) \equiv \lim_{n\rightarrow\infty} \frac{1}{n} D_{\lambda,\mu}\left(\mathcal{N}^{\otimes n}\right).$$

11

Lemma 1 *Suppose the quantum dynamic capacity formula is additive for any two channels $\mathcal{N}$ and $\mathcal{M}$:*

$$D_{\lambda,\mu}\left(\mathcal{N}\otimes\mathcal{M}\right)=D_{\lambda,\mu}\left(\mathcal{N}\right)+D_{\lambda,\mu}\left(\mathcal{M}\right).$$

Then the regularized quantum dynamic capacity formula for $\mathcal{N}$ is equal to the quantum dynamic capacity formula for $\mathcal{N}$:

$$D_{\lambda,\mu}^{reg}\left(\mathcal{N}\right)=D_{\lambda,\mu}\left(\mathcal{N}\right).$$

In this sense, the regularized formula "single-letterizes" and it is not necessary to take the limit.

Proof We prove the result using induction on n. The base case for $n=1$ is trivial. Suppose the result holds for n: $D_{\lambda,\mu}(\mathcal{N}^{\otimes n})=nD_{\lambda,\mu}(\mathcal{N})$. Then the following chain of equalities proves the inductive step:

$$\begin{aligned}
D_{\lambda,\mu}(\mathcal{N}^{\otimes n+1}) &= D_{\lambda,\mu}(\mathcal{N}\otimes\mathcal{N}^{\otimes n})\\
&= D_{\lambda,\mu}(\mathcal{N})+D_{\lambda,\mu}(\mathcal{N}^{\otimes n})\\
&= D_{\lambda,\mu}(\mathcal{N})+nD_{\lambda,\mu}(\mathcal{N}).
\end{aligned}$$

The first equality follows by expanding the tensor product. The second critical equality follows from the assumption that the formula is additive. The final equality follows from the induction hypothesis.

Theorem 2 *Single-letterization of the quantum dynamic capacity formula implies that the computation of the Pareto optimal trade-off surface of the quantum dynamic capacity region requires an optimization over a single channel use.*

Proof We employ ideas from Ref. [7] for the proof. We would like to characterize all the points in the capacity region that are Pareto optimal. Such a task is standard vector optimization in the theory of Pareto trade-off analysis (see Section 4.7 of Ref. [7]). We can phrase the optimization task as the following scalarization of the vector optimization task:

$$\max_{C,Q,E,p(x),\phi_x} w_C C + w_Q Q + w_E E \tag{22}$$

subject to

$$C + 2Q \le I(AX;B^n)_\sigma, \tag{23}$$
$$Q + E \le I(A\rangle B^n X)_\sigma, \tag{24}$$
$$C + Q + E \le I(X;B^n)_\sigma + I(A\rangle B^n X)_\sigma, \tag{25}$$

where the maximization is over all C, Q, and E and over probability distributions $p_X(x)$ and bipartite states $\phi_x^{AA'^n}$. The geometric interpretation of the scalarization task is that we are trying to find a supporting plane of the dynamic capacity region where the weight vector (w_C, w_Q, w_E) is the normal vector of the plane and the value of its inner product with (C, Q, E) characterizes the offset of the plane.

The Lagrangian of the above optimization problem is

$$\begin{aligned}
\mathcal{L}\left(C,Q,E,p_X(x),\phi_x^{AA'^n},\lambda_1,\lambda_2,\lambda_3\right) &\equiv w_C C + w_Q Q + w_E E + \lambda_1\left(I\left(AX;B^n\right)_\sigma-(C+2Q)\right)\\
&\quad + \lambda_2\left(I\left(A\rangle B^n X\right)_\sigma-(Q+E)\right)\\
&\quad + \lambda_3\left(I\left(X;B^n\right)_\sigma+I\left(A\rangle B^n X\right)_\sigma-(C+Q+E)\right),
\end{aligned}$$

and the Lagrange dual function g [7] is

$$g\left(\lambda_1,\lambda_2,\lambda_3\right)\equiv\sup_{C,Q,E,p(x),\phi_x^{AA'^n}}\mathcal{L}\left(C,Q,E,p_X(x),\phi_x^{AA'^n},\lambda_1,\lambda_2,\lambda_3\right),$$

12

where $\lambda_1, \lambda_2, \lambda_3 \geq 0$. The optimization task simplifies if the Lagrange dual function does. Thus, we rewrite the Lagrange dual function as follows:

$$
\begin{aligned}
g\left(\lambda_1, \lambda_2, \lambda_3\right) = & \sup_{C,Q,E,p(x),\phi_x^{AA'^n}} w_C C + w_Q Q + w_E E + \lambda_1 \left(I\left(AX; B^n\right)_\sigma - (C + 2Q)\right) \\
& + \lambda_2 \left(I\left(A \rangle B^n X\right)_\sigma - (Q + E)\right) \\
& + \lambda_3 \left(I\left(X; B^n\right)_\sigma + I\left(A \rangle B^n X\right)_\sigma - (C + Q + E)\right) \\
= & \sup_{C,Q,E,p(x),\phi_x^{AA'^n}} \left(w_C - \lambda_1 - \lambda_3\right) C + \left(w_Q - 2\lambda_1 - \lambda_2 - \lambda_3\right) Q + \left(w_E - \lambda_2 - \lambda_3\right) E \\
& + \lambda_1 \left(I\left(AX; B^n\right)_\sigma + \frac{\lambda_2}{\lambda_1} I\left(A \rangle B^n X\right)_\sigma + \frac{\lambda_3}{\lambda_1}\left(I\left(X; B^n\right)_\sigma + I\left(A \rangle B^n X\right)_\sigma\right)\right) \\
= & \sup_{C,Q,E} \left(w_C - \lambda_1 - \lambda_3\right) C + \left(w_Q - 2\lambda_1 - \lambda_2 - \lambda_3\right) Q + \left(w_E - \lambda_2 - \lambda_3\right) E \\
& + \lambda_1 \left(\max_{p(x),\phi_x^{AA'^n}} I\left(AX; B^n\right)_\sigma + \frac{\lambda_2}{\lambda_1} I\left(A \rangle B^n X\right)_\sigma + \frac{\lambda_3}{\lambda_1}\left(I\left(X; B^n\right)_\sigma + I\left(A \rangle B^n X\right)_\sigma\right)\right).
\end{aligned}
$$

The first equality follows by definition. The second equality follows from some algebra, and the last follows because the Lagrange dual function factors into two separate optimization tasks: one over C, Q, and E and another that is equivalent to the quantum dynamic capacity formula with $\lambda = \lambda_2/\lambda_1$ and $\mu = \lambda_3/\lambda_1$. Thus, the computation of the Pareto optimal trade-off surface requires just a single use of the channel if the quantum dynamic capacity formula in (21) single-letterizes.

7.1 Special cases of the quantum dynamic capacity formula

We now show how several capacity formulas of a quantum channel, including the entanglement-assisted classical capacity [5], the Lloyd-Shor-Devetak (LSD) formula for the quantum capacity [21,25,10], and the Holevo-Schumacher-Westmoreland (HSW) formula for the classical capacity [14,24] are special cases of the quantum dynamic capacity formula.

We first give a geometric interpretation of these special cases before proceeding to the proofs. Recall that the dynamic capacity region has the simple interpretation as a translation of the three-faced unit resource capacity region along the classically-enhanced father trade-off curve (see Figure 2 for the example of the region of the dephasing channel). Any particular weight vector (w_C, w_Q, w_E) in (22) gives a set of parallel planes that slice through the (C, Q, E) space, and the goal of the scalar optimization task is to find one of these planes that is a supporting plane, intersecting a point (or a set of points) on the trade-off surface of the dynamic capacity region. We consider three special planes:

1. The first corresponds to the plane containing the vectors of super-dense coding and teleportation. The normal vector of this plane is $(1, 2, 0)$, and suppose that we set the weight vector in (22) to be this vector. Then the optimization program finds the set of points on the trade-off surface such that a plane with this normal vector is a supporting plane for the region. The optimization program singles out (23), and we can think of this as being equivalent to setting $\lambda_2, \lambda_3 = 0$ in the Lagrange dual function. We show below that the optimization program becomes equivalent to finding the entanglement-assisted capacity [5], in the sense that the quantum dynamic capacity formula becomes the entanglement-assisted capacity formula.
2. The next plane contains the vectors of teleportation and entanglement distribution. The normal vector of this plane is $(0, 1, 1)$. Setting the weight vector in (22) to be this vector makes the optimization program single out (24), and we can think of this as being equivalent to setting $\lambda_1, \lambda_3 = 0$ in the Lagrange dual function. We show below that the optimization program becomes equivalent to finding the quantum capacity [21,25,10], in the sense that the quantum dynamic capacity formula becomes the LSD formula for the quantum capacity.
3. A third plane contains the vectors of super-dense coding and entanglement distribution. The normal vector of this plane is $(1, 1, 1)$. Setting the weight vector in (22) to be this vector makes the optimization program single out (25), and we can think of this as being equivalent to setting $\lambda_1, \lambda_2 = 0$ in the Lagrange dual function. We show below that the optimization becomes equivalent to finding the

classical capacity [21, 25, 10], in the sense that the quantum dynamic capacity formula becomes the HSW formula for the classical capacity.

Corollary 1 *The quantum dynamic capacity formula is equivalent to the entanglement-assisted classical capacity formula when $\lambda, \mu = 0$, in the sense that*

$$\max_{\sigma} I\left(AX; B\right) = \max_{\phi^{AA'}} I\left(A; B\right).$$

Proof The inequality $\max_{\sigma} I\left(AX; B\right) \geq \max_{\phi^{AA'}} I\left(A; B\right)$ follows because the state σ is of the form in (14) and we can always choose $p_X\left(x\right) = \delta_{x, x_0}$ and $\phi_{x_0}^{AA'}$ to be the state that maximizes $I\left(A; B\right)$.

We now show the other inequality $\max_{\sigma} I\left(AX; B\right) \leq \max_{\phi^{AA'}} I\left(A; B\right)$. First, consider that the following chain of equalities holds for any state ϕ^{ABE} resulting from the isometric extension of the channel:

$$\begin{aligned} I\left(A; B\right) &= H\left(B\right) + H\left(A\right) - H\left(AB\right) \\ &= H\left(B\right) + H\left(BE\right) - H\left(E\right) \\ &= H\left(B\right) + H\left(B|E\right). \end{aligned}$$

In this way, we see that the mutual information is purely a function of the channel input density operator $\text{Tr}_A\left\{\phi^{AA'}\right\}$. Then consider any state σ of the form in (14). The following chain of inequalities holds

$$\begin{aligned} I\left(AX; B\right)_{\sigma} &= H\left(A|X\right)_{\sigma} + H\left(B\right)_{\sigma} - H\left(E|X\right)_{\sigma} \\ &= H\left(BE|X\right)_{\sigma} + H\left(B\right)_{\sigma} - H\left(E|X\right)_{\sigma} \\ &= H\left(B|EX\right)_{\sigma} + H\left(B\right)_{\sigma} \\ &\leq H\left(B|E\right)_{\sigma} + H\left(B\right)_{\sigma} \\ &\leq \max_{\phi^{AA'}} I\left(A; B\right). \end{aligned}$$

The first equality follows by expanding the mutual information. The second equality follows because the state on ABE is pure when conditioned on X. The third equality follows from the entropy chain rule. The first inequality follows from strong subadditivity, and the last follows because the state after tracing out systems X and A is a particular state that arises from the channel and cannot be larger than the maximum.

Corollary 2 *The quantum dynamic capacity formula is equivalent to the LSD quantum capacity formula in the limit where $\lambda \to \infty$ and μ is fixed, in the sense that*

$$\max_{\sigma} I\left(A\rangle BX\right) = \max_{\phi^{AA'}} I\left(A\rangle B\right).$$

Proof The inequality $\max_{\sigma} I\left(A\rangle BX\right) \geq \max_{\phi^{AA'}} I\left(A\rangle B\right)$ follows because the state σ is of the form in (14) and we can always choose $p_X\left(x\right) = \delta_{x, x_0}$ and $\phi_{x_0}^{AA'}$ to be the state that maximizes $I\left(A\rangle B\right)$.

The inequality $\max_{\sigma} I\left(A\rangle BX\right) \leq \max_{\phi^{AA'}} I\left(A\rangle B\right)$ follows because $I\left(A\rangle BX\right) = \sum_x p_X\left(x\right) I\left(A\rangle B\right)_{\phi_x}$ and the maximum is always greater than the average.

Corollary 3 *The quantum dynamic capacity formula is equivalent to the HSW classical capacity formula in the limit where $\mu \to \infty$ and λ is fixed, in the sense that*

$$\max_{\sigma} I\left(A\rangle BX\right)_{\sigma} + I\left(X; B\right)_{\sigma} = \max_{\{p_X(x), \psi_x\}} I\left(X; B\right).$$

The inequality $\max_{\sigma} I\left(A\rangle BX\right)_{\sigma} + I\left(X; B\right)_{\sigma} \geq \max_{\{p_X(x), \psi_x\}} I\left(X; B\right)$ follows by choosing σ to be the pure ensemble that maximizes $I\left(X; B\right)$ and noting that $I\left(A\rangle BX\right)_{\sigma}$ vanishes for a pure ensemble.

We now prove the inequality $\max_\sigma I(A\rangle BX)_\sigma + I(X;B)_\sigma \leq \max_{\{p_X(x),\psi_x\}} I(X;B)$. Consider a state ω^{XYBE} obtained by performing a von Neumann measurement on the A system of the state σ^{XABE}. Then

$$
\begin{aligned}
I(A\rangle BX)_\sigma + I(X;B)_\sigma &= H(B)_\sigma - H(E|X)_\sigma \\
&= H(B)_\omega - H(E|X)_\omega \\
&\leq H(B)_\omega - H(E|XY)_\omega \\
&= H(B)_\omega - H(B|XY)_\omega \\
&= I(XY;B)_\omega \\
&\leq \max_{\{p_X(x),\psi_x\}} I(X;B).
\end{aligned}
$$

The first equality follows by expanding the conditional coherent information and the Holevo information. The second equality follows because the measured A system is not involved in the entropies. The first inequality follows because conditioning does not increase entropy. The third equality follows because the state ω is pure when conditioned on X and Y. The fourth equality follows by definition, and the last inequality follows for clear reasons.

8 Single-letter dynamic capacity region for the quantum Hadamard channels

Below we show that the regularization in (8) is not necessary if the quantum channel is a Hadamard channel. This result holds because a Hadamard channel has a special structure. The development of the proof is similar to that in Ref. [9], but simplified because we obtain the single-letter result more directly.

Theorem 3 *The dynamic capacity region $\mathcal{C}_{\mathrm{CQE}}(\mathcal{N}_H)$ of a quantum Hadamard channel $\mathcal{N}_H$ is equal to its one-shot region $\mathcal{C}_{\mathrm{CQE}}^{(1)}(\mathcal{N}_H)$.*

The proof of the above theorem follows in two parts: 1) the below lemma shows the quantum dynamic capacity formula is additive when one of the channels is Hadamard and 2) the induction argument in Lemma 1 that proves single-letterization.

Lemma 2 *The following additivity relation holds for a Hadamard channel $\mathcal{N}_H$ and any other channel $\mathcal{N}$:*

$$
D_{\lambda,\mu}(\mathcal{N}_H \otimes \mathcal{N}) = D_{\lambda,\mu}(\mathcal{N}_H) + D_{\lambda,\mu}(\mathcal{N}).
$$

Proof We first note that the inequality $D_{\lambda,\mu}(\mathcal{N}_H \otimes \mathcal{N}) \geq D_{\lambda,\mu}(\mathcal{N}_H) + D_{\lambda,\mu}(\mathcal{N})$ holds for any two channels simply by selecting the state σ in the maximization to be a tensor product of the ones that individually maximize $D_{\lambda,\mu}(\mathcal{N}_H)$ and $D_{\lambda,\mu}(\mathcal{N})$.

So we prove that the non-trivial inequality $D_{\lambda,\mu}(\mathcal{N}_H \otimes \mathcal{N}) \leq D_{\lambda,\mu}(\mathcal{N}_H) + D_{\lambda,\mu}(\mathcal{N})$ holds when the first channel is a Hadamard channel. Since the first channel is Hadamard, it is degradable and its degrading map has a particular structure: there are maps $\mathcal{D}_1^{B_1 \to Y}$ and $\mathcal{D}_2^{Y \to E_1}$ where Y is a classical register and such that the degrading map is $\mathcal{D}_2^{Y \to E_1} \circ \mathcal{D}_1^{B_1 \to Y}$ [9,18]. Suppose the state we are considering to input to the tensor product channel is

$$
\rho^{XAA_1'A_2'} \equiv \sum_x p_X(x) |x\rangle \langle x|^X \otimes \phi_x^{AA_1'A_2'},
$$

and this state is the one that maximizes $D_{\lambda,\mu}(\mathcal{N}_H \otimes \mathcal{N})$. Suppose that the output of the first channel is

$$
\theta^{XAB_1E_1A_2'} \equiv U_{\mathcal{N}_H}^{A_1' \to B_1 E_1}(\rho^{XAA_1'A_2'}),
$$

and the output of the second channel is

$$
\omega^{XAB_1E_1B_2E_2} \equiv U_{\mathcal{N}}^{A_2' \to B_2 E_2}(\theta^{XAB_1E_1A_2'}).
$$

Finally, we define the following state as the result of applying the first part of the Hadamard degrading map (a von Neumann measurement) to ω:

$$
\sigma^{XYAE_1B_2E_2} \equiv \mathcal{D}_1^{B_1 \to Y}(\omega^{XAB_1E_1B_2E_2}).
$$

In particular, the state σ on systems $AE_1B_2E_2$ is pure when conditioned on X and Y. Then the following chain of inequalities holds:

$$
\begin{aligned}
D_{\lambda,\mu} \left(\mathcal{N}_H \otimes \mathcal{N} \right) &= I \left(AX; B_1 B_2 \right)_\omega + \lambda I \left(A \rangle B_1 B_2 X \right)_\omega + \mu \left(I \left(X; B_1 B_2 \right)_\omega + I \left(A \rangle B_1 B_2 X \right)_\omega \right) \\
&= H \left(B_1 B_2 E_1 E_2 | X \right)_\omega + \lambda H \left(B_1 B_2 | X \right)_\omega + \left(\mu + 1 \right) H \left(B_1 B_2 \right)_\omega - \left(\lambda + \mu + 1 \right) H \left(E_1 E_2 | X \right)_\omega \\
&= H \left(B_1 E_1 | X \right)_\omega + \lambda H \left(B_1 | X \right)_\omega + \left(\mu + 1 \right) H \left(B_1 \right)_\omega - \left(\lambda + \mu + 1 \right) H \left(E_1 | X \right)_\omega + \\
&\quad H \left(B_2 E_2 | B_1 E_1 X \right)_\omega + \lambda H \left(B_2 | B_1 X \right)_\omega + \left(\mu + 1 \right) H \left(B_2 | B_1 \right)_\omega - \left(\lambda + \mu + 1 \right) H \left(E_2 | E_1 X \right)_\omega \\
&\leq H \left(B_1 E_1 | X \right)_\theta + \lambda H \left(B_1 | X \right)_\theta + \left(\mu + 1 \right) H \left(B_1 \right)_\theta - \left(\lambda + \mu + 1 \right) H \left(E_1 | X \right)_\theta + \\
&\quad H \left(B_2 E_2 | Y X \right)_\sigma + \lambda H \left(B_2 | Y X \right)_\sigma + \left(\mu + 1 \right) H \left(B_2 \right)_\sigma - \left(\lambda + \mu + 1 \right) H \left(E_2 | Y X \right)_\sigma \\
&= I \left(AA'_2 X; B_1 \right)_\theta + \lambda I \left(AA'_2 \rangle B_1 X \right)_\theta + \mu \left(I \left(X; B_1 \right)_\theta + I \left(AA'_2 \rangle B_1 X \right)_\theta \right) + \\
&\quad I \left(AE_1 Y X; B_2 \right)_\sigma + \lambda I \left(AE_1 \rangle B_2 Y X \right)_\sigma + \mu \left(I \left(Y X; B_2 \right)_\sigma + I \left(AE_1 \rangle B_2 Y X \right)_\sigma \right) \\
&\leq D_{\lambda,\mu} \left(\mathcal{N}_H \right) + D_{\lambda,\mu} \left(\mathcal{N} \right).
\end{aligned}
$$

The first equality follows by evaluating the quantum dynamic capacity formula $D_{\lambda,\mu} \left(\mathcal{N}_H \otimes \mathcal{N} \right)$ on the state ρ. The next two equalities follow by rearranging entropies and because the state ω on systems $AB_1E_1B_2E_2$ is pure when conditioned on X. The inequality in the middle is the crucial one and follows from the Hadamard structure of the channel: we exploit monotonicity of conditional entropy under quantum operations so that $H \left(B_2 | B_1 X \right)_\omega \leq H \left(B_2 | Y X \right)_\sigma$ and $H \left(E_2 | Y X \right)_\sigma \leq H \left(E_2 | E_1 X \right)_\omega$. The next equality follows by rearranging entropies and the final one follows because θ is a state of the form (14) for the first channel while σ is a state of the form (14) for the second channel.

9 The dynamic capacity region of a dephasing channel

The below theorem shows that the full dynamic capacity region admits a particularly simple form when the noisy quantum channel is a qubit dephasing channel Δ_p where

$$
\begin{aligned}
\Delta_p \left(\rho \right) &\equiv \left(1 - p \right) \rho + p \Delta \left(\rho \right), \\
\Delta \left(\rho \right) &\equiv \langle 0 | \rho | 0 \rangle | 0 \rangle \langle 0 | + \langle 1 | \rho | 1 \rangle | 1 \rangle \langle 1 |.
\end{aligned}
$$

A dephasing channel is an example of a quantum Hadamard channel [9].[4] Figure 2 plots this region for the case of a dephasing channel with dephasing parameter $p = 0.2$. The proof of the following theorem exploits the same techniques as in Ref. [9].

Theorem 4 *The dynamic capacity region $\mathcal{C}_{CQE}(\Delta_p)$ of a dephasing channel with dephasing parameter p is the set of all C, Q, and E such that*

$$
C + 2Q \leq 1 + H_2 \left(\nu \right) - H_2(\gamma \left(\nu, p \right)), \tag{26}
$$

$$
Q + E \leq H_2 \left(\nu \right) - H_2(\gamma \left(\nu, p \right)), \tag{27}
$$

$$
C + Q + E \leq 1 - H_2(\gamma \left(\nu, p \right)), \tag{28}
$$

where $\nu \in [0, 1/2]$, H_2 is the binary entropy function, and

$$
\gamma \left(\nu, p \right) \equiv \frac{1}{2} + \frac{1}{2} \sqrt{1 - 16 \cdot \frac{p}{2} \left(1 - \frac{p}{2} \right) \nu (1 - \nu)}.
$$

Proof We first notice that it suffices to consider an ensemble of pure states whose reductions to A' are diagonal in the dephasing basis (as in Lemma 11 of Ref. [9]). Next we prove below that it is sufficient to consider an ensemble of the following form to characterize the boundary points of the region:

$$
\frac{1}{2} | 0 \rangle \langle 0 |^X \otimes \psi_0^{AA'} + \frac{1}{2} | 1 \rangle \langle 1 |^X \otimes \psi_1^{AA'}, \tag{29}
$$

---

[4] Brádler showed that cloning channels and an Unruh channel are also in the Hadamard class [8].
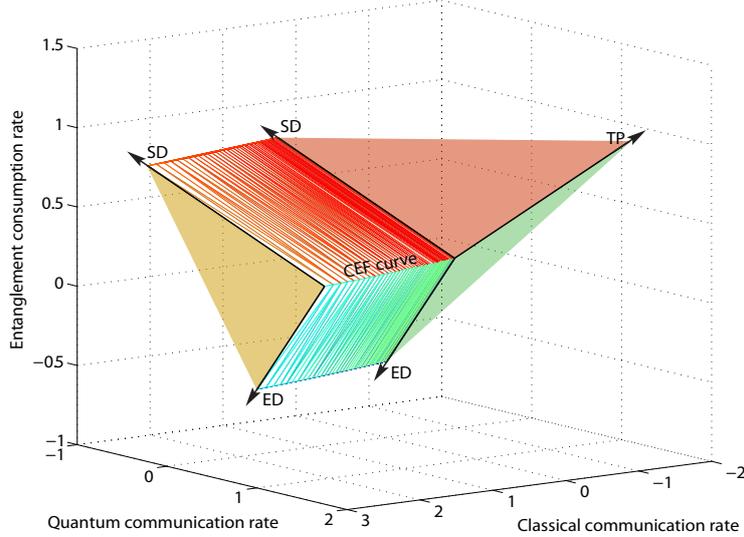
Fig. 2 (Color online) A plot of the dynamic capacity region for a qubit dephasing channel with dephasing parameter $p = 0.2$. The plot shows that the classically-enhanced father (CEF) trade-off curve lies along the boundary of the dynamic capacity region. The rest of the region is simply the combination of the CEF points with the unit protocols teleportation (TP), super-dense coding (SD), and entanglement distribution (ED).

where $\psi_0^{AA'}$ and $\psi_1^{AA'}$ are pure states, defined as follows for $\nu \in [0, 1/2]$:

$$\text{Tr}_A \left\{ \psi_0^{AA'} \right\} = \nu \, |0\rangle \langle 0|^{A'} + (1 - \nu) \, |1\rangle \langle 1|^{A'}, \tag{30}$$

$$\text{Tr}_A \left\{ \psi_1^{AA'} \right\} = (1 - \nu) \, |0\rangle \langle 0|^{A'} + \nu \, |1\rangle \langle 1|^{A'}. \tag{31}$$

We now prove the above claim. We assume without loss of generality that the dephasing basis is the computational basis. Consider a classical-quantum state with a finite number N of conditional density operators $\phi_x^{AA'}$ whose reduction to A' is diagonal:

$$\rho^{XAA'} \equiv \sum_{x=0}^{N-1} p_X(x) \, |x\rangle\langle x|^X \otimes \phi_x^{AA'}.$$

We can form a new classical-quantum state with double the number of conditional density operators by "bit-flipping" the original conditional density operators:

$$\sigma^{XAA'} \equiv \frac{1}{2} \sum_{x=0}^{N-1} p_X(x) \left(|x\rangle\langle x|^X \otimes \phi_x^{AA'} + |x+N\rangle\langle x+N|^X \otimes X^{A'} \phi_x^{AA'} X^{A'} \right),$$

where X is the σ_X "bit-flip" Pauli operator. Consider the following chain of inequalities that holds for all $\lambda, \mu \geq 0$:

$$
\begin{aligned}
& I(AX; B)_\rho + \lambda I(A\rangle BX)_\rho + \mu \left(I(X; B)_\rho + I(A\rangle BX)_\rho \right) \\
&= H(A|X)_\rho + (\mu + 1) H(B)_\rho + \lambda H(B|X)_\rho - (\lambda + \mu + 1) H(E|X)_\rho \\
&\leq (\mu + 1) H(B)_\sigma + H(A|X)_\sigma + \lambda H(B|X)_\sigma - (\lambda + \mu + 1) H(E|X)_\sigma \\
&= (\mu + 1) + H(A|X)_\sigma + \lambda H(B|X)_\sigma - (\lambda + \mu + 1) H(E|X)_\sigma \\
&= (\mu + 1) + \sum_x p_X(x) \left[H(A)_{\phi_x} + \lambda H(B)_{\phi_x} - (\lambda + \mu + 1) H(E)_{\phi_x} \right] \\
&\leq (\mu + 1) + \max_x \left[H(A)_{\phi_x} + \lambda H(B)_{\phi_x} - (\lambda + \mu + 1) H(E)_{\phi_x} \right] \\
&= (\mu + 1) + H(A)_{\phi_x^*} + \lambda H(B)_{\phi_x^*} - (\lambda + \mu + 1) H(E)_{\phi_x^*}.
\end{aligned}
$$

17

The first equality follows by standard entropic manipulations. The second equality follows because the conditional entropy $H(B|X)$ is invariant under a bit-flipping unitary on the input state that commutes with the channel: $H(B)_{X\rho_x^B X} = H(B)_{\rho_x^B}$. Furthermore, a bit flip on the input state does not change the eigenvalues for the output of the dephasing channel's complementary channel: $H(E)_{\mathcal{N}^c(X\rho_x^{A'} X)} = H(E)_{\mathcal{N}^c(\rho_x^{A'})}$. The first inequality follows because entropy is concave, i.e., the local state σ^B is a mixed version of ρ^B. The third equality follows because $H(B)_{\sigma^B} = H\left(\sum_x \frac{1}{2} p_X(x)\left(\rho_x^B + X\rho_x^B X\right)\right) = H\left(\frac{1}{2}\sum_x p_X(x) I\right) = 1$. The fourth equality follows because the system X is classical. The second inequality follows because the maximum value of a realization of a random variable is not less than its expectation. The final equality simply follows by defining ϕ_x^* to be the conditional density operator on systems A, B, and E that arises from sending through the channel a state whose reduction to A' is of the form $\nu |0\rangle\langle 0|^{A'} + (1-\nu)|1\rangle\langle 1|^{A'}$. Thus, an ensemble of the kind in (29) is sufficient to attain a point on the boundary of the region.

Evaluating the entropic quantities in Theorem 1 on a state of the above form then gives the expression for the region in Theorem 4.

10 Single-letter dynamic capacity region for the quantum erasure channels

Below we show that the regularization in (8) is not necessary if the quantum channel is a quantum erasure channel. The quantum erasure channel also has a special structure, but the proof proceeds differently from that for a quantum Hadamard channel.

A quantum erasure channel with erasure parameter ϵ is the following map [13]:

$$\mathcal{N}_\epsilon(\rho) \equiv (1-\epsilon)\rho + \epsilon |e\rangle\langle e|.$$

Notice that the receiver can perform a measurement $\{|0\rangle\langle 0| + |1\rangle\langle 1|, |e\rangle\langle e|\}$ and can learn whether the channel erased the state. The receiver can do this without disturbing the state in any way. An isometric extension $U_{\mathcal{N}_\epsilon}^{A'\to BE}$ of it acts as follows on a purification $|\psi\rangle^{AA'}$ of the state $\rho^{A'}$:

$$U_{\mathcal{N}_\epsilon}^{A'\to BE} |\psi\rangle^{AA'} = \sqrt{1-\epsilon}\,|\psi\rangle^{AB} |e\rangle^E + \sqrt{\epsilon}\,|\psi\rangle^{AE} |e\rangle^B.$$

In the above representation, we see that the erasure channel has the interpretation that it hands the input to Bob with probability $1-\epsilon$ while giving an erasure flag $|e\rangle$ to Eve, and it hands the input to Eve with probability ϵ while giving the erasure flag to Bob.

Theorem 5 *The dynamic capacity region $\mathcal{C}_{\mathrm{CQE}}(\mathcal{N}_\epsilon)$ of a quantum erasure channel $\mathcal{N}_\epsilon$ is the set of all C, Q, and E such that*

$$C + 2Q \leq (1-\epsilon)(1 + H_2(p)),$$
$$Q + E \leq (1-2\epsilon)H_2(p),$$
$$C + Q + E \leq 1 - \epsilon - \epsilon H_2(p),$$

where $p \in [0, 1/2]$.

Figure 3 plots the dynamic capacity region of a quantum erasure channel with erasure parameter $\epsilon = 1/4$. It turns out that time-sharing is the optimal strategy here, and there is not an interesting trade-off curve for the quantum erasure channel.

We in fact proved Theorem 5 in Ref. [17] by employing a *reductio ad absurdum* argument reminiscent of the earliest arguments for proving capacities of quantum erasure channels [4]. This approach gives the correct answer, but suffers from two shortcomings:

1. We do not learn much about how to exploit the structure of the quantum erasure channel with the *reductio ad absurdum* approach. As an example, Smith and Yard exploited the simple structure of the quantum erasure channel and discovered far reaching consequences [26]. In particular, they discovered that the quantum capacity (and for that matter, any future proposed quantum capacity formula) can never be generally additive, by combining the erasure channel with another one.
2. The *reductio ad absurdum* argument rests on the assumption that several known capacity formulas are continuous as a function of channels. Leung and Smith later showed that the known formulas are indeed continuous [19], redeeming the original argument in Ref. [4].
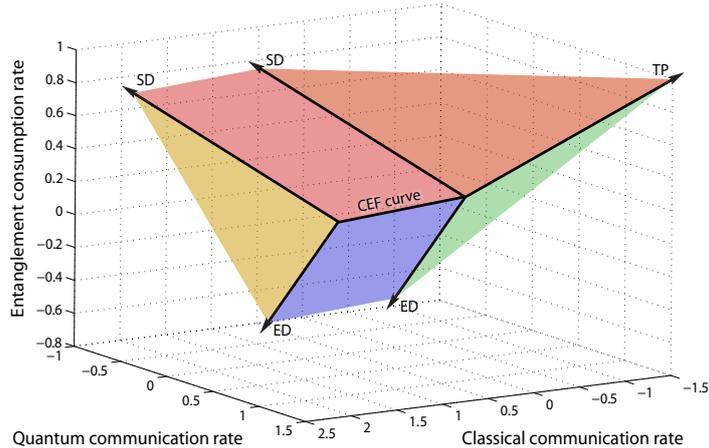
Fig. 3 (Color online) A plot of the dynamic capacity region for a qubit erasure channel with erasure parameter $\epsilon = 1/4$. The plot shows that the classically-enhanced father (CEF) trade-off curve lies along the boundary of the dynamic capacity region and it is not actually a curve but rather a line because time-sharing is optimal. The rest of the region is simply the combination of the CEF points with the unit protocols teleportation (TP), super-dense coding (SD), and entanglement distribution (ED).

Here, we prove Theorem 5 above by carefully studying the structure of the quantum erasure channel and its additivity properties for the full dynamic capacity region. We prove the theorem in a few steps. First, we prove that the classical capacity of the quantum erasure channel admits a single-letter formula.[5] We then simplify the quantum dynamic capacity formula in (21) for the case of a quantum erasure channel and find that it is only necessary to consider certain values of the parameters μ and λ when we optimize. The proof of Lemma 7 exploits these conditions to show that the quantum dynamic capacity formula is additive for the case of two quantum erasure channels. It then follows by a trivial induction step (the same as in Lemma 1) that the full dynamic region single-letterizes and is of the form in Theorem 5.

Lemma 3 (Bennett *et al.* [4]) *The Holevo information of a quantum erasure channel is equal to* $1 - \epsilon$:

$$\chi(\mathcal{N}_\epsilon) \equiv \max_{\sigma^{XB}} I(X;B) = 1 - \epsilon$$

Proof We begin with an input ensemble of the following form:

$$\rho^{XA'} \equiv \sum_x p_X(x) |x\rangle \langle x|^X \otimes \phi_x^{A'},$$

where the states $\phi_x^{A'}$ are pure (it suffices to consider pure state ensembles for classical capacity). Feeding the A' system into the quantum erasure channel leads to a classical-quantum state $\sigma^{XB} = \sum_x p_X(x) |x\rangle \langle x|^X \otimes \mathcal{N}_\epsilon^{A' \to B}(\phi_x^{A'})$. Bob can then measure his system B to learn whether the channel erases the qubit. Let X_E denote a classical register where Bob places the result of the measurement so that we then have the state σ^{XBX_E}. It holds that any entropy evaluated on system B is equal to the joint entropy of B and X_E because this measurement does not disturb the state in any way. Consider the following chain of inequalities:

$$
\begin{aligned}
I(X;B)_\sigma &= H(B)_\sigma - H(B|X)_\sigma \\
&= H(BX_E)_\sigma - H(BX_E|X)_\sigma \\
&= H(B|X_E)_\sigma - H(B|X_EX)_\sigma \\
&= (1-\epsilon)H(A')_\rho - (1-\epsilon)H(A'|X)_\rho \\
&= (1-\epsilon)I(X;A')_\rho \\
&\leq (1-\epsilon).
\end{aligned}
$$

---

[5] The proof already appears in Ref. [4], but it again suffers from the aforementioned shortcomings.

19

The first equality follows by expanding the mutual information and the second equality follows from the above fact regarding the joint entropy of B and X_E. The third equality follows by expanding and canceling terms. The fourth equality follows by conditioning on the classical erasure flag register X_E and realizing that the entropy of Bob's system B is the entropy of the input state with probability $1 - \epsilon$ and otherwise is the entropy of the erasure state $|e\rangle$ (this latter entropy vanishes because this state is pure). Thus, the above sequence of steps reduces an optimization problem on the output of the channel to a simple optimization over the input ensemble. The Holevo information is then $1 - \epsilon$ because the quantity $I(X; A')_\rho$ can never be larger than unity for the case of a qubit erasure channel (it reaches unity for an ensemble of two orthogonal pure states chosen with equal probability).

Lemma 4 *The following additivity lemma holds for a quantum erasure channel $\mathcal{N}_\epsilon$:*

$$\chi(\mathcal{N}_\epsilon \otimes \mathcal{N}_\epsilon) = \chi(\mathcal{N}_\epsilon) + \chi(\mathcal{N}_\epsilon) = 2(1 - \epsilon).$$

Proof It suffices to prove the inequality $\chi(\mathcal{N}_\epsilon \otimes \mathcal{N}_\epsilon) \leq \chi(\mathcal{N}_\epsilon) + \chi(\mathcal{N}_\epsilon)$ because the other inequality holds trivially. We define the following ensemble of states for the tensor product channel $\mathcal{N}_\epsilon \otimes \mathcal{N}_\epsilon$:

$$\rho^{XA_1'A_2'} \equiv \sum_x p_X(x) |x\rangle\langle x|^X \otimes \phi_x^{A_1'A_2'}. \tag{32}$$

We can also define the following augmented ensemble based on the above one:

$$\sigma^{XIJA_1'A_2'} \equiv \frac{1}{16} \sum_{x,i,j} p_X(x) |x\rangle\langle x|^X \otimes |i\rangle\langle i|^I \otimes |j\rangle\langle j|^J \otimes (\sigma_i^{A_1'} \otimes \sigma_j^{A_2'}) \phi_x^{A_1'A_2'} (\sigma_i^{A_1'} \otimes \sigma_j^{A_2'}), \tag{33}$$

where $\sigma_0 \equiv I$, $\sigma_1 \equiv \sigma_X$, $\sigma_2 \equiv \sigma_Y$, and $\sigma_3 \equiv \sigma_Z$. In particular, note that we obtain the maximally mixed state when tracing over classical registers I and J. Let $\omega^{XB_1B_2}$ and $\theta^{XIJB_1B_2}$ denote the states obtained by sending systems A_1' and A_2' of the respective states $\rho^{XA_1'A_2'}$ and $\sigma^{XIJA_1'A_2'}$ through two uses of the quantum erasure channel. Let $X_{E,1}$ and $X_{E,2}$ denote the classical variables Bob obtains by determining whether the channel erased his states (they also denote the registers where he places the results). Consider the following chain of inequalities that holds for any state $\omega^{XB_1B_2}$:

$$\begin{aligned}
I(X; B_1B_2)_\omega &= H(B_1B_2)_\omega - H(B_1B_2|X)_\omega \\
&= H(B_1B_2|X_{E,1}X_{E,2})_\omega - H(B_1B_2|X_{E,1}X_{E,2}X)_\omega \\
&= (1-\epsilon)^2 H(A_1'A_2')_\rho + (1-\epsilon)\epsilon \left(H(A_1')_\rho + H(A_2')_\rho \right) \\
&\quad - (1-\epsilon)^2 H(A_1'A_2'|X)_\rho - (1-\epsilon)\epsilon \left(H(A_1'|X)_\rho + H(A_2'|X)_\rho \right) \\
&\leq 2(1-\epsilon) - (1-\epsilon)^2 H(A_1'A_2'|XIJ)_\sigma - (1-\epsilon)\epsilon \left(H(A_1'|XIJ)_\sigma + H(A_2'|XIJ)_\sigma \right) \\
&\leq 2(1-\epsilon).
\end{aligned}$$

The first equality holds by expanding the mutual information. The next equality holds because an "erasure measurement" does not change entropy. The third equality follows by exploiting the properties of the erasure channels. The first inequality holds because the unconditional entropies of the A' systems on the state ρ are always less than those for the state σ. The final inequality follows because the entropies in the previous line are non-negative. The statement of the theorem then follows.

Lemma 5 *The quantum dynamic capacity formula in (21) simplifies as follows for a quantum erasure channel $\mathcal{N}_\epsilon$:*

$$D_{\lambda,\mu}(\mathcal{N}_\epsilon) \equiv \max_{p \in [0,1/2]} (1-\epsilon)(1 + H_2(p)) + \lambda(1 - 2\epsilon) H_2(p) + \mu((1-\epsilon) - \epsilon H_2(p)). \tag{34}$$

Thus, the "one-letter" dynamic capacity region of a quantum erasure channel is as Theorem 5 states.

Proof We exploit the following classical-quantum states:

$$\rho^{XAA'} \equiv \sum_x p_X(x) |x\rangle\langle x|^X \otimes \phi_x^{AA'},$$

$$\sigma^{XIAA'} \equiv \sum_{x,i} \frac{1}{4} p_X(x) |x\rangle\langle x|^X \otimes |i\rangle\langle i|^I \otimes (\sigma_i^{A'}) \phi_x^{AA'} (\sigma_i^{A'}),$$

and let ρ^{XABE} and σ^{XIABE} be the states obtained by transmitting the A' system through the isometric extension of the erasure channel. Let $\sigma_x^{A'} \equiv \mathrm{Tr}_A\left\{\phi_x^{AA'}\right\}$. Furthermore, let the eigenvalues of the state $\sigma_x^{A'}$ with highest entropy on system A' be p and $1-p$. Consider that the following chain of inequalities holds for any state ρ^{XABE}:

$$
\begin{aligned}
& I\left(AX;B\right)_\rho + \lambda I\left(A\rangle BX\right)_\rho + \mu\left(I\left(X;B\right)_\rho + I\left(A\rangle BX\right)_\rho\right) \\
&= H\left(A|X\right)_\rho + (\mu+1)H\left(B\right)_\rho + \lambda H\left(B|X\right)_\rho - (\lambda+\mu+1)H\left(E|X\right)_\rho \\
&= H\left(A'|X\right)_\rho + (\mu+1)H\left(B|X_E\right)_\rho + \lambda H\left(B|X_E X\right)_\rho - (\lambda+\mu+1)H\left(E|X_E X\right)_\rho \\
&= H\left(A'|X\right)_\rho + (\mu+1)(1-\epsilon)H\left(A'\right)_\rho + \lambda(1-\epsilon)H\left(A'|X\right)_\rho - (\lambda+\mu+1)\epsilon H\left(A'|X\right)_\rho \\
&\leq (\mu+1)(1-\epsilon) + (1+\lambda(1-\epsilon)-(\lambda+\mu+1)\epsilon)H\left(A'|XI\right)_\sigma \\
&= (\mu+1)(1-\epsilon) + (1-\epsilon+\lambda(1-2\epsilon)-\mu\epsilon)\sum_x p_X(x)H\left(A'\right)_{\sigma_x} \\
&\leq (\mu+1)(1-\epsilon) + (1-\epsilon+\lambda(1-2\epsilon)-\mu\epsilon)H\left(A'\right)_{\sigma_x^*} \\
&= (1-\epsilon)(1+H_2(p)) + \mu(1-\epsilon-\epsilon H_2(p)) + \lambda(1-2\epsilon)H_2(p).
\end{aligned}
$$

The first equality follows by standard entropic manipulations. The second equality follows by incorporating the classical erasure flag variable. The third equality follows by exploiting the properties of the quantum erasure channel. The first inequality follows because the unconditional entropy of the state ρ is always less than that of the state σ. The next equality follows by expanding the conditional entropy. The second inequality follows because an average is always less than a maximum. The final equality follows by plugging in the eigenvalues of σ_x^*. The form of the quantum dynamic capacity formula then follows because this chain of inequalities holds for any input ensemble.

Lemma 6 *It suffices to consider the set of $\lambda, \mu \geq 0$ for which*

$$(1-\epsilon) + \lambda(1-2\epsilon) \geq \mu\epsilon.$$

Otherwise, we are just maximizing the classical capacity, which we know from Lemma 3 is equal to $1-\epsilon$.

Proof Consider rewriting the expression in (34) as follows:

$$\max_{p\in[0,1/2]} (1-\epsilon) + \mu(1-\epsilon) + [(1-\epsilon)+\lambda(1-2\epsilon)-\mu\epsilon]H_2(p).$$

Suppose that the expression in square brackets is negative, i.e.,

$$(1-\epsilon) + \lambda(1-2\epsilon) < \mu\epsilon.$$

Then the maximization over p simply chooses $p=0$ so that $H_2(p)$ vanishes and the negative term disappears. The resulting expression for the quantum dynamic capacity formula is

$$(1-\epsilon) + \mu(1-\epsilon),$$

which corresponds to the following region

$$
\begin{aligned}
C + 2Q &\leq 1-\epsilon, \\
Q + E &\leq 0, \\
C + Q + E &\leq 1-\epsilon.
\end{aligned}
$$

The above region is equivalent to a translation of the unit resource capacity region to the classical capacity rate triple $(1-\epsilon, 0, 0)$. Thus, it suffices to restrict the parameters λ and μ as above for the quantum erasure channel.

Lemma 7 *The following additivity relation holds for two quantum erasure channels $\mathcal{N}_\epsilon$ with the same erasure parameter ϵ:*

$$D_{\lambda,\mu}(\mathcal{N}_\epsilon \otimes \mathcal{N}_\epsilon) = D_{\lambda,\mu}(\mathcal{N}_\epsilon) + D_{\lambda,\mu}(\mathcal{N}_\epsilon).$$

Proof We prove the non-trivial inequality $D_{\lambda,\mu}(\mathcal{N}_\epsilon \otimes \mathcal{N}_\epsilon) \leq D_{\lambda,\mu}(\mathcal{N}_\epsilon) + D_{\lambda,\mu}(\mathcal{N}_\epsilon)$. We define the following states:

$$\rho^{XAA_1'A_2'} \equiv \sum_x p_X(x) |x\rangle\langle x|^X \otimes \phi_x^{AA_1'A_2'},$$

$$\omega^{XAB_1E_1B_2E_2} \equiv U_{\mathcal{N}_\epsilon}^{A_1' \to B_1 E_1} \otimes U_{\mathcal{N}_\epsilon}^{A_2' \to B_2 E_2} (\rho^{XAA_1'A_2'}),$$

and we suppose that $\rho^{XAA_1A_2}$ is the state that maximizes $D_{\lambda,\mu}(\mathcal{N}_\epsilon \otimes \mathcal{N}_\epsilon)$. Consider the following equality:

$$I(AX;B_1B_2)_\omega + \lambda I(A\rangle B_1 B_2 X)_\omega + \mu(I(X;B_1B_2)_\omega + I(A\rangle B_1 B_2 X)_\omega)$$
$$= H(A|X)_\omega + H(B_1 B_2)_\omega - H(E_1 E_2 | X)_\omega + \lambda(H(B_1 B_2 | X)_\omega - H(E_1 E_2 | X)_\omega)$$
$$+ \mu(H(B_1 B_2)_\omega - H(E_1 E_2 | X)_\omega).$$

It follows simply by rewriting entropies. We continue below:

$$= H(A_1' A_2' | X)_\rho + (1-\epsilon)^2 H(A_1' A_2')_\rho + \epsilon(1-\epsilon)\left(H(A_1')_\rho + H(A_2')_\rho\right)$$
$$- \left[\epsilon^2 H(A_1' A_2' | X)_\rho + \epsilon(1-\epsilon)\left(H(A_1'|X)_\rho + H(A_2'|X)_\rho\right)\right]$$
$$+ \lambda\left[(1-\epsilon)^2 H(A_1' A_2' | X)_\rho + \epsilon(1-\epsilon)\left(H(A_1'|X)_\rho + H(A_2'|X)_\rho\right)\right]$$
$$- \lambda\left[\epsilon^2 H(A_1' A_2' | X)_\rho + \epsilon(1-\epsilon)\left(H(A_1'|X)_\rho + H(A_2'|X)_\rho\right)\right]$$
$$+ \mu\left[(1-\epsilon)^2 H(A_1' A_2')_\rho + \epsilon(1-\epsilon)\left(H(A_1')_\rho + H(A_2')_\rho\right)\right]$$
$$- \mu\left[\epsilon^2 H(A_1' A_2' | X)_\rho + \epsilon(1-\epsilon)\left(H(A_1'|X)_\rho + H(A_2'|X)_\rho\right)\right].$$

The above equality follows by exploiting the properties of the quantum erasure channel. Continuing, the above quantity is less than the following one:

$$\leq 2(1-\epsilon) + (1-\epsilon^2) H(A_1' A_2' | XIJ)_\sigma + \epsilon(1-\epsilon)(H(A_1'|XIJ)_\sigma + H(A_2'|XIJ)_\sigma)$$
$$+ \lambda(1-2\epsilon) H(A_1' A_2' | XIJ)_\sigma$$
$$+ \mu\left[2(1-\epsilon) - \epsilon^2 H(A_1' A_2' | XIJ)_\sigma - \epsilon(1-\epsilon)(H(A_1'|XIJ)_\sigma - H(A_2'|XIJ)_\sigma)\right]$$
$$= 2(1-\epsilon) + (1-\epsilon)(H(A_1'|XIJ)_\sigma + H(A_2'|XIJ)_\sigma) + \lambda(1-2\epsilon)(H(A_1'|XIJ)_\sigma + H(A_2'|XIJ)_\sigma)$$
$$+ \mu[2(1-\epsilon) - \epsilon(H(A_1'|XIJ)_\sigma - H(A_2'|XIJ)_\sigma)]$$
$$- \left[(1-\epsilon^2 + \lambda(1-2\epsilon) - \mu\epsilon^2) I(A_1'; A_2' | XIJ)_\sigma\right]$$
$$\leq D_{\lambda,\mu}(\mathcal{N}_\epsilon) + D_{\lambda,\mu}(\mathcal{N}_\epsilon) - \left[(1-\epsilon^2 + \lambda(1-2\epsilon) - \mu\epsilon^2) I(A_1'; A_2' | XIJ)_\rho\right]$$
$$\leq D_{\lambda,\mu}(\mathcal{N}_\epsilon) + D_{\lambda,\mu}(\mathcal{N}_\epsilon).$$

The first inequality follows from similar proofs we have seen for a state σ of the form in (33). The first equality follows by rearranging terms. The second inequality follows from the form of $D_{\lambda,\mu}$ in (34). The final inequality follows because Lemma 6 states that it is sufficient to consider $(1-\epsilon) + \lambda(1-2\epsilon) \geq \mu\epsilon$. Note that this condition implies that

$$1 - \epsilon^2 + \lambda(1-2\epsilon) \geq \mu\epsilon^2,$$

and hence that the quantity in square brackets in the line above the last one is positive.

11 Conclusion

We found a purely information theoretic approach to proving the converse part of the dynamic capacity region. This technique should be simpler to understand for those unfamiliar with the quantum Shannon theory literature. We also phrased the optimization task for the full dynamic capacity region in terms of the quantum dynamic capacity formula in (21) and proved its additivity (and hence single-letterization

of the dynamic capacity region) for the quantum Hadamard channels and quantum erasure channels. We note some open problems below.

There might be room for improvement in our formulas that characterize the dynamic capacity region when the channel is not of the Hadamard class or a quantum erasure channel. Though, our characterization has the simple interpretation as the regularization of what one can achieve with the classically-enhanced father protocol [16] combined with teleportation, super-dense coding, and entanglement distribution. It is difficult to imagine a simpler characterization than this one, despite its multi-letter nature for the general case.

We would like to find other channels besides the Hadamard or erasure channels for which the region single-letterizes. We conjecture that additivity of the quantum dynamic capacity formula in (21) holds for channels that have hybrid Hadamard-erasure behavior such as the phase erasure channel in Ref. [4]. It would also be interesting to find channels that are not hybrid Hadamard-erasure for which additivity of (21) holds.

There is also one interesting speculation to muse over that Professor David Avis suggested to us. Do each of the inequalities in Theorem 1 correspond to some fundamental physical law? This might shed further connections between information theory and physics that have not been elucidated yet.

References

1. Robert Alicki and Mark Fannes. Continuity of quantum conditional information. *Journal of Physics A: Mathematical and General*, 37(5):L55–L57, 2004.
2. Howard Barnum, Emanuel Knill, and Michael A. Nielsen. On quantum fidelities and channel capacities. *IEEE Transactions on Information Theory*, 46:1317–1329, 2000.
3. Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Physical Review Letters*, 70:1895–1899, 1993.
4. Charles H. Bennett, David P. DiVincenzo, and John A. Smolin. Capacities of quantum erasure channels. *Physical Review Letters*, 78(16):3217–3220, April 1997.
5. Charles H. Bennett, Peter W. Shor, John A. Smolin, and Ashish V. Thapliyal. Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem. *IEEE Transactions on Information Theory*, 48:2637, 2002.
6. Charles H. Bennett and Stephen J. Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Physical Review Letters*, 69:2881–2884, 1992.
7. Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, The Edinburgh Building, Cambridge, CB2 8RU, UK, 2004.
8. Kamil Brádler. An infinite sequence of additive channels: the classical capacity of cloning channels. *IEEE Transactions on Information Theory*, 57(8):5497–5503, August 2011. arXiv:0903.1638.
9. Kamil Brádler, Patrick Hayden, Dave Touchette, and Mark M. Wilde. Trade-off capacities of the quantum Hadamard channels. *Physical Review A*, 81:062312, June 2010. arXiv:1001.1732.
10. Igor Devetak. The private classical capacity and quantum capacity of a quantum channel. *IEEE Transactions on Information Theory*, 51(1):44–55, 2005.
11. Igor Devetak, Aram W. Harrow, and Andreas Winter. A resource framework for quantum Shannon theory. *IEEE Transactions on Information Theory*, 54(10):4587–4618, October 2008.
12. Igor Devetak and Peter W. Shor. The capacity of a quantum channel for simultaneous transmission of classical and quantum information. *Communications in Mathematical Physics*, 256(2):287–303, 2005.
13. M. Grassl, Th. Beth, and T. Pellizzari. Codes for the quantum erasure channel. *Physical Review A*, 56(1):33–38, July 1997.
14. Alexander S. Holevo. The capacity of the quantum channel with general signal states. *IEEE Transactions on Information Theory*, 44:269–273, 1998.
15. Min-Hsiu Hsieh and Mark M. Wilde. *Theory of Quantum Computation, Communication, and Cryptography*, volume 5906 of *Lecture Notes in Computer Science*, chapter Optimal Trading of Classical Communication, Quantum Communication, and Entanglement, pages 85–93. Springer-Verlag, May 2009.

16. Min-Hsiu Hsieh and Mark M. Wilde. Entanglement-assisted communication of classical and quantum information. *IEEE Transactions on Information Theory*, 56(9):4682–4704, September 2010. arXiv:0811.4227.

17. Min-Hsiu Hsieh and Mark M. Wilde. Trading classical communication, quantum communication, and entanglement in quantum Shannon theory. *IEEE Transactions on Information Theory*, 56(9):4705–4730, September 2010. arXiv:0901.3038.

18. Christopher King, Keiji Matsumoto, Michael Nathanson, and Mary Beth Ruskai. Properties of conjugate channels with applications to additivity and multiplicativity. *Markov Processes and Related Fields*, 13(2):391–423, 2007. J. T. Lewis memorial issue.

19. Debbie Leung and Graeme Smith. Continuity of quantum channel capacities. *Communications in Mathematical Physics*, 292(1):201–215, November 2009.

20. Elliot H. Lieb and Mary Beth Ruskai. Proof of strong subadditivity of quantum-mechanical entropy. *Journal of Mathematical Physics*, 14:1938, 1973.

21. Seth Lloyd. The capacity of a noisy quantum channel. *Physical Review A*, 55:1613–1622, 1997.

22. Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, New York, 2000.

23. Benjamin Schumacher and Michael A. Nielsen. Quantum data processing and error correction. *Physical Review A*, 54:2629–2635, 1996.

24. Benjamin Schumacher and Michael D. Westmoreland. Sending classical information via noisy quantum channels. *Physical Review A*, 56:131–138, 1997.

25. Peter W. Shor. The quantum channel capacity and coherent information. MSRI workshop on quantum computation, 2002.

26. Graeme Smith and Jon Yard. Quantum communication with zero-capacity channels. *Science*, 321:1812–1815, September 2008.

27. Jon Yard. *Simultaneous classical-quantum capacities of quantum multiple access channels*. PhD thesis, Stanford University, Stanford, CA, 2005. quant-ph/0506050.