Superadditivity in trade-off capacities of quantum channels

Elton Yechao Zhu, *Member, IEEE*, Quntao Zhuang, Min-Hsiu Hsieh, *Senior Member, IEEE*, and Peter W. Shor

***Abstract*—In this article, we investigate the additivity phenomenon in the dynamic capacity of a quantum channel for trading the resources of classical communication, quantum communication, and entanglement. Understanding such an additivity property is important if we want to optimally use a quantum channel for general communication purposes. However, in a lot of cases, the channel one will be using only has an additive single or double resource capacity, and it is largely unknown if this could lead to a strictly superadditive double or triple resource capacity, respectively. For example, if a channel has additive classical and quantum capacities, can the classical-quantum capacity be strictly superadditive? In this work, we answer such questions affirmatively.**

We give proof-of-principle requirements for these channels to exist. In most cases, we can provide an explicit construction of these quantum channels. The existence of these superadditive phenomena is surprising in contrast to the result that the additivity of both classical-entanglement and classical-quantum capacity regions imply the additivity of the triple resource capacity region for a given channel.

***Index Terms*—Additivity; Quantum Channel Capacity; Trade-off Capacity Regions; Quantum Shannon theory.**

I. Introduction

In studying classical communication, Shannon developed powerful probabilistic tools that connect the theoretic throughput of a channel to an entropic quantity defined on a single use of the channel [1]. Shannon's noiseless channel coding theorem involves a random coding strategy to prove achievability and entropic inequalities that show optimality, *i.e.,* the converse. This methodology has now become standard in proving finite or asymptotic optimal resource conversions in information theory.

Quantum Shannon information starts by mimicking classical information theory: typical sets can be generalized to typical subspaces to prove achievability while various entropic inequalities, such as the quantum data processing inequality, can be used to prove the converse. However, the differences between quantum and classical Shannon information are also significant. On one hand, additional resources available in the quantum domain diversify the allowable capacities, resulting in trade-off regions for the resources that are consumed or generated [2]–[4]. The most common, and useful, quantum resource in communication settings is quantum entanglement. Unlike classical shared randomness, which does not increase a classical channel's capability to send more messages, preshared quantum entanglement will generally increase the throughput of a quantum channel for sending classical messages or quantum messages or both [2], [5]–[9]. It thus makes sense to consider the trade-off capacity regions among these three useful resources: entanglement, classical communication, and quantum communication, and this was done in Ref. [4]. The result in Ref. [4] further shows that a coding strategy that exploits the channel coding of these three resources as a whole performs better than strategies that use each of these three resources individually.

On the other hand, even though single-lettered channel capacity formulas have been found in the classical regime for certain communication tasks, when considering related tasks in the quantum regime, the known formulas are generally no longer tractable and are instead regularized capacity formulas [10]–[14]. In other words, evaluation of these capacity quantities requires optimizing information formulas over an arbitrarily large number of uses of a given channel. This largely blocks our understanding of quantum communication capacities. An extreme example shows the existence of two quantum channels that cannot be used to send a quantum message reliably individually but will have a positive channel capacity when both are used simultaneously [15]. However, there are also several examples showing that when additional resources are used to assist, the corresponding assisted capacity will also become additive. The classical capacity over quantum channels is generally superadditive; however, when assisted by a sufficient amount of entanglement, the entanglement-assisted capacity becomes additive [6], [16]. The quantum capacity also exhibits similar properties. When assisted by either entanglement [2], [3] or an unbounded symmetric side channel [17], its assisted quantum capacity becomes additive.

This superadditive property of quantum channel capacities has accordingly attracted significant attention. Hastings [18] proved that the classical capacity over quantum channels is not additive, a result built upon earlier developments by Hayden-Winter [19] and Shor [20]. Recently, three of us showed a rather perplexing result [21]: when assisted by an insufficient amount of entanglement, a channel's classical capacity could be strictly superadditive regardless of whether the unassisted classical capacity is additive. Further, the additivity property of the entanglement-assisted classical capacity shows a form of phase transition. Even if the channel is additive when assisted

E.Y. Zhu is with the Center of Theoretical Physics and Department of Physics, Massachusetts Institute of Technology, Cambridge, MA, 02139 USA (e-mail: eltonzhu@mit.edu).

Q. Zhuang is with the Research Laboratory of Electronics and Department of Physics, Massachusetts Institute of Technology, Cambridge, MA, 02139 USA (e-mail: quntao@mit.edu).

Min-Hsiu Hsieh is with the Centre of Quantum Software and Information (UTS|QSI⟩), University of Technology Sydney, Australia (e-mail: Min-Hsiu.Hsieh@uts.edu.au).

P.W. Shor is with the Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, 02139 USA (e-mail: shor@math.mit.edu).

by a sufficient amount of entanglement or no entanglement at all, it can still be strictly superadditive when assisted with an insufficient amount of entanglement. This phenomenon indicates that quantum channels behave fundamentally differently from classical channels, and our understanding of their behavior is still quite limited.

Ref. [21] used the idea of a switch channel, with one end of the switch a symmetric classical channel, the other end of the switch a quantum channel constructed in Ref. [18], having a strictly superadditive classical capacity. By design the two channels have the same classical capacity. Hence, without entanglement assistance, only the classical channel is used and the classical capacity is additive. However, as the rate of entanglement assistance increases, the quantum channel dominates the communication protocol. With some delicate concavity argument, one can show that there exists some rate of entanglement assistance such that the classical capacity is indeed strictly superadditive.

This paper is inspired by, and aims to extend Ref. [21]. Will additivity of single or double resource capacities always lead to additivity of a general resource trade-off region? We will study superadditivity in a general framework that considers the three most common resources: entanglement, noiseless classical communication and quantum communication. Our results show that (i) additivity of single resource capacities of a quantum channel does not generally imply additivity of double resource capacities, except for the known result [2] that an additive quantum capacity yields an additive entanglement-assisted quantum capacity region (see Table I); and (ii) additive double resource capacities does not generally imply an additive triple resource capacity, except for the known case [8] that additive classical-entanglement and classical-quantum capacity regions yield an additive triple dynamic capacity (see Table II). These results again demonstrate how complex a quantum channel can be, and further investigation is required.

The paper is structured as follows. Section II introduces the various definitions, notations and previous results on the triple resource quantum Shannon theory. Section III summarizes the various superadditivity results that we establish in the paper. Section IV establishes the switch channel that we use for all our constructions, and expresses the triple resource trade-off formula of the switch channel in terms of those of the sub-channels. Section V gives a detailed construction of all the possible superadditivity phenomena.

II. Preliminaries

In this section, we give definitions of basic entropic quantities used in the paper. We also describe the dynamic capacity theorem. Special cases of this include the various single and double resource capacities. Finally, we define the elementary channels that will be used in our explicit constructions.

A bipartite quantum state ρ_{AB} is a positive semi-definite operator on a Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ with trace one. We define the von Neumann entropy, coherent information and quantum mutual information of ρ_{AB}, respectively, as follows:

$$\begin{aligned} S(AB)_\rho &= -\mathrm{Tr}\,[\rho_{AB}\log\rho_{AB}], \\ I(A\rangle B)_\rho &= S(B)_\rho - S(AB)_\rho, \\ I(A;B)_\rho &= S(A)_\rho + I(A\rangle B)_\rho, \end{aligned}$$

where $S(A)_\rho$ is the von Neumann entropy of the reduced state $\rho_A = \mathrm{Tr}_B[\rho_{AB}]$.

For an ensemble $\{p(x), \sigma^x_{AB}\}_{x\in\mathcal{X}}$, let

$$\sigma_{XAB} = \sum_{x\in\mathcal{X}} p(x)|x\rangle\langle x|_X \otimes \sigma^x_{AB},$$

where $\{|x\rangle\}$ forms a fixed orthonormal (computational) basis in a Hilbert space $\mathcal{H}_X$. We need the following information-theoretic quantities as well:

$$\begin{aligned} I(A\rangle BX)_\sigma &= \sum_x p(x) I(A\rangle B)_{\sigma^x}, &(1)\\ I(A;B|X)_\sigma &= \sum_x p(x) I(A;B)_{\sigma^x}, &(2)\\ I(AX;B)_\sigma &= I(X;B)_\sigma + I(A;B|X)_\sigma, &(3) \end{aligned}$$

where $I(A\rangle BX)_\sigma$ and $I(A;B|X)$ in Eqs. (1) and (2) are the conditional coherent information and the conditional mutual information, respectively. $I(X;B)_\sigma$ in Eq. (3) is the Holevo information of $\sigma_{XB} = \mathrm{Tr}_A[\sigma_{XAB}]$.

A quantum channel $\mathcal{N}$ is a completely positive and trace-preserving map. With it, we can transmit either classical or quantum information or both with possible entanglement assistance between the sender and the receiver [8]. More generally, the authors in Ref. [4] proved the following capacity theorem that involves a noisy quantum channel $\mathcal{N}$ and the three resources mentioned above; namely, classical communication (C), quantum communication (Q) and quantum entanglement (E).

Theorem 1 (CQE trade-off [4]): The dynamic capacity region $\mathcal{C}_{CQE}(\mathcal{N})$ of a quantum channel $\mathcal{N}$ is equal to the following expression:

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

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

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

The state-dependent region $\mathcal{C}^{(1)}_{CQE,\sigma}(\mathcal{N})$ is the set of triples (C,Q,E) of rates such that

$$\begin{aligned} C + 2Q &\le I(AX;B)_\sigma, &(4)\\ Q + E &\le I(A\rangle BX)_\sigma, &(5)\\ C + Q + E &\le I(X;B)_\sigma + I(A\rangle BX)_\sigma. &(6) \end{aligned}$$

The above entropic quantities are with respect to a classical-quantum state (cq state) σ_{XAB}, where

$$\sigma_{XAB} \equiv \sum_x p(x)\,|x\rangle\,\langle x|_X \otimes \mathcal{N}_{A'\to B}\left(\phi^x_{AA'}\right), \tag{7}$$

and the states $\phi^x_{AA'}$ are pure.

The dynamic capacity of a channel $\mathcal{N}$ is always superadditive, *i.e.*,

$$\mathcal{C}_{CQE}(\mathcal{N}) \supseteq \mathcal{C}^{(1)}_{CQE}(\mathcal{N}).$$

We say that the dynamic capacity of a channel $\mathcal{N}$ is *(weakly) additive* if

$$C_{CQE}(\mathcal{N}) = C^{(1)}_{CQE}(\mathcal{N}). \tag{8}$$

Thus if

$$C_{CQE}(\mathcal{N}) \supsetneq C^{(1)}_{CQE}(\mathcal{N}),$$

then the dynamic capacity of $\mathcal{N}$ is strictly superadditive, *i.e.,* non-additive.

We will also be using the following stronger notion of additivity in our proofs. The dynamic capacity of a channel $\mathcal{N}$ is *strongly additive* if

$$C^{(1)}_{CQE}(\mathcal{N} \otimes \Psi) = C^{(1)}_{CQE}(\mathcal{N}) + C^{(1)}_{CQE}(\Psi) \tag{9}$$

for an arbitrary channel Ψ. Here, addition means Minkowski sum[1].

The dynamic capacity region $C_{CQE}(\mathcal{N})$ in Theorem 1 allows us to recover known capacity theorems by choosing certain (C, Q, E) in Eqs. (4)-(6) as follows:

- the classical capacity $C_C(\mathcal{N})$ when choosing $Q = E = 0$ [10], [11];
- the quantum capacity $C_Q(\mathcal{N})$ when choosing $C = E = 0$ [12]–[14];
- the classical and quantum capacity $C_{CQ}(\mathcal{N})$ when choosing $E = 0$ (CQ trade-off) [23];
- the entanglement assisted classical capacity $C_{CE}(\mathcal{N})$ when choosing $Q = 0$ (CE trade-off) [7], [16];
- the entanglement assisted quantum capacity $C_{QE}(\mathcal{N})$ when choosing $C = 0$ (QE trade-off) [2], [3];

The additivity of these special cases is defined similarly as in Eq. (8).

We note that the dynamic capacity region is convex, as a convex combination of any two points in the region can be achieved by a time-sharing strategy, *i.e.,* using the channel for a fraction of uses to achieve one point, and using it for the other fraction to achieve the second point.

Below we will briefly describe a few channels which we will repeatedly use.

A quantum channel $\mathcal{N}_{A\to B}$ can be written as

$$\mathcal{N}_{A\to B}(\rho_A) = \mathrm{Tr}_E\left[U\rho_A U^\dagger\right],$$

where $U_{A\to BE}$ is an isometry from $\mathcal{H}_A$ to $\mathcal{H}_B \otimes \mathcal{H}_E$, commonly called the isometric extension of $\mathcal{N}$. E is usually called the environment. Then the complementary channel $\mathcal{N}^c$ is defined by

$$\mathcal{N}^c_{A\to E}(\rho_A) = \mathrm{Tr}_B\left[U\rho_A U^\dagger\right].$$

Definition 2: A *entanglement-breaking channel* $\mathcal{N}_{A\to B}$ is a quantum channel with the property that $\sigma_{RB} = \mathcal{N}_{A\to B}(\rho_{RA})$ is a separable state for any input state ρ_{RA}, where σ_{RB} is separable if it can be written as

$$\sigma_{RB} = \sum_k p(k)\phi^k_R \otimes \phi^k_B, \tag{10}$$

where $\sum_k p(k) = 1$ and all ϕ^k_R, ϕ^k_B are pure states.

[1]For two sets of position vectors A and B in Euclidean space, their Minkowski sum $A + B$ is obtained by adding each vector in A to each vector in B, *i.e.,* $A + B = \{a + b | a \in A, b \in B\}$ [22].

Definition 3: A *Hadamard channel* is a quantum channel whose complementary channel is entanglement-breaking.

Suppose $\Psi_{A'\to B}$ is a Hadamard channel, with the complementary channel $\Psi^c_{A'\to E}$. Then there is a degrading map $\mathcal{D}_{B\to E}$ [24] such that

$$\Psi^c_{A'\to E} = \mathcal{D}_{B\to E} \circ \Psi_{A'\to B}.$$

Moreover, $\mathcal{D}$ can be decomposed as

$$\mathcal{D}_{B\to E} = \mathcal{D}^2_{Y\to E} \circ \mathcal{D}^1_{B\to Y},$$

where Y is a classical variable.

A Hadamard channel has an additive quantum dynamic capacity region, when tensored with an arbitrary quantum channel [24]. Examples of Hadamard channels include the qubit dephasing channel, $1 \to N$ cloning channels, and the Unruh channel. We'll define the qubit dephasing channel and $1 \to N$ cloning channel below. We refer the interested readers to Ref. [24] for more details and properties of these channels.

Definition 4: The *qubit dephasing channel* Ψ^{dph}_η, with dephasing probability η, is defined as

$$\Psi^{\mathrm{dph}}_\eta(\rho) = (1-\eta)\rho + \eta Z\rho Z,$$

where Z is the Pauli-Z operator.

Definition 5: A $1 \to N$ qubit cloning channel $\Psi^{1\to N}$ is a channel that approximately copies the input qubit state with maximal copy fidelity independent of the input state.

Let $\{|j\rangle_B, j = 0, \ldots, N\}$ be an orthonormal basis of the normalized completely symmetric state for the output system of N qubits, where $|j\rangle_B$ is a uniform superposition of computational basis states with $N - j$ 0's and j 1's. Let $\{|i\rangle_E, i = 0, \ldots, N-1\}$ be an orthonormal basis of of the normalized completely symmetric state for the environment system of $N-1$ qubits, where $|i\rangle_E$ is a uniform superposition of computational basis states with $N-i-1$ 0's and i 1's. An isometric extension of $\Psi^{1\to N}$ has the form

$$\begin{aligned} U_{A\to BE} = &\frac{1}{\sqrt{\Delta_N}}\sum_{i=1}^{N-1}\sqrt{N-1}\,|i\rangle_B\langle 0|_A \otimes |i\rangle_E \\ &+ \frac{1}{\sqrt{\Delta_N}}\sum_{i=1}^{N-1}\sqrt{i+1}\,|i+1\rangle_B\langle 0|_A \otimes |i\rangle_E, \end{aligned}$$

where $\Delta_N \equiv N(N+1)/2$.

Definition 6: The *qubit depolarizing channel* Ψ^{dpo}_p, with depolarizing probability p, is defined as

$$\Psi^{\mathrm{dpo}}_p(\rho) = (1-p)\rho + p\frac{I}{2}.$$

The qubit depolarizing channel is known to have an additive classical capacity [25], but a non-additive quantum capacity when $p = 0.746$ [26].

Definition 7: A *random orthogonal channel* Ψ^{ro} is defined as

$$\Psi^{\mathrm{ro}}(\rho) = \sum_{i=1}^{D} P_i O_i \rho O_i^\intercal,$$

where O_i are chosen from the orthogonal group and the probabilities P_i are roughly equal.

	Imply the additivity of		
Additive capacities	CE	CQ	QE
C	N [21]	N [26]	N [26]
Q	N (SecV-C)	N (SecV-C)	Y [3]
C&Q⇔ C&QE	N (SecV-B)	N (SecV-D)	Y[3]

TABLE I
SUMMARY OF RESULTS FOR DOUBLE RESOURCES. "N" STANDS FOR "DOES NOT IMPLY ADDITIVITY", WHILE "Y" MEANS "IMPLIES ADDITIVITY".

	Imply the additivity of
Additive capacities	CQE
QE	N (SecV-C)
CQ	N (SecV-G, conjecture)
CE	N (SecV-E)
CE&Q⇔CE&QE	N (SecV-F)
CE&CQ	Y[8]

TABLE II
SUMMARY OF RESULTS FOR TRIPLE RESOURCES. "N" STANDS FOR "DOES NOT IMPLY ADDITIVITY", WHILE "Y" MEANS "IMPLIES ADDITIVITY".

For $1 \ll D \ll N$, with N the input dimension, such a channel will have a strictly subadditive minimum output entropy with high probability [18].

Definition 8: Consider an arbitrary channel $\Psi_{C\to B}$. Append a register R to the input, with a set of orthonormal bases $\{|k\rangle\}$. We define a *unitally extended channel* [20], [27] $\Phi_{RC\to B}$ as

$$\Phi_{RC\to B}(\rho_{RC}) = \sum_k U_k \Psi_{C\to B}\left(\langle k|\,\rho_{RC}\,|k\rangle_R\right) U_k^\dagger, \qquad (11)$$

where $\{U_k \in \mathcal{U}(|B|) : k \in \{1,\ldots,|R|\}\}$ form a unitary 1-design.

A set of K unitaries $\{U_k\} \in \mathcal{U}(d)$ form a unitary 1-design if it satisfies the following property

$$\frac{1}{K}\sum_k U_k A U_k^\dagger = \mathrm{Tr}(A)\frac{I}{d} \qquad (12)$$

for any $d \times d$ matrix A.

Note that unital extensions are not unique. For practical purposes, we will mostly use Heisenberg-Weyl operators $\{X(x,z) \in \mathcal{U}(d) : x, z \in \{0,,\ldots,d-1\}$, defined as

$$X(x,z)\,|l\rangle = \exp(i2\pi zl/d)\,|l \oplus x\rangle \qquad (13)$$

on the computational basis states $\{|l\rangle\}_{l\in\{0,\ldots,d\}}$. For notational simplicity, we will denote them as $\{X(j), j \in \{1,\ldots,d^2\}\}$.

If a channel has a strictly subadditive minimum output entropy, its unital extension will have a strictly superadditive classical capacity [20]. Hence a unital extension of a random orthogonal channel will have a strictly superadditive classical capacity with high probability.

III. SUMMARY OF RESULTS

We summarize all of our results here. We will denote a single-resource capacity region by a single letter, e.g., C for $C_C(\mathcal{N})$. We will also use short notation for double and triple-resource trade-off regions, e.g., CE for $C_{CE}(\mathcal{N})$ and CQE for $C_{CQE}(\mathcal{N})$. We will use the arrow notation, with "$\to$" meaning additivity of the capacity on the left-hand side implies additivity of the capacity on the right-hand side , and "$\not\to$" meaning additivity of the capacity on the left-hand side does not imply additivity of the capacity on the right-hand side .

A. Double resources (see table I)

1) CE:
   a) C $\not\to$ CE [21]: There exists a quantum channel $\mathcal{N}$, such that its classical capacity is additive, but its CE trade-off capacity region is non-additive. We will give a simplified construction in Sec V-A.
   b) C&Q $\not\to$ CE: There exists a quantum channel $\mathcal{N}$, such that its classical and quantum capacities are both additive, but its CE trade-off capacity region is non-additive. An explicit construction of $\mathcal{N}$ is given in Sec V-B.
2) QE:
   Q $\to$ QE [3]: For any quantum channel $\mathcal{N}$, if its quantum capacity is additive, then its QE trade-off capacity region is always additive.
3) CQ:
   a) C $\not\to$ CQ [26]: There exists a quantum channel $\mathcal{N}$, such that its classical capacity is additive, but its CQ trade-off capacity region is non-additive.
   The depolarizing channel has a non-additive quantum capacity at some dephasing parameter, and hence a non-additive CQ trade-off capacity, while its classical capacity is additive.
   b) Q $\not\to$ CQ: There exists a quantum channel $\mathcal{N}$, such that its quantum capacity is additive, but its CQ trade-off capacity region is non-additive. A construction of such a quantum channel is given in Sec V-C.
   c) C&Q $\not\to$ CQ: Moreover, there exists a quantum channel $\mathcal{N}$, such that its classical and quantum capacities are additive, but its CQ trade-off capacity region is non-additive. A construction of such a quantum channel is given in Sec V-D.

B. Triple resources (see table II)

1) CE $\not\to$ CQE: There exists a quantum channel $\mathcal{N}$ such that its CE trade-off capacity region is additive, but its dynamic capacity region is non-additive. An example is constructed in Sec V-E.
2) CE&Q $\not\to$ CQE: There exists a quantum channel $\mathcal{N}$ such that its quantum capacity and its CE trade-off capacity region are additive, but its dynamic capacity region is non-additive. An example is constructed in Sec V-F.
3) CQ $\not\to$ CQE: (Conjecture) There exists a quantum channel $\mathcal{N}$ such that its CQ trade-off capacity region is additive, but its dynamic capacity region is non-additive. This is the only case in which we do not have an explicit example. We outline a possible construction in Sec V-G.
4) CE&CQ $\to$ CQE [8]: If a quantum channel $\mathcal{N}$ has additive CE and CQ trade-off capacity regions, then its dynamic capacity region is also additive. This statement is first observed in Ref. [8], and an explicit argument can be found in Ref. [24].

IV. Framework

This section presents technical tools that we require for demonstration of superadditivity in trade-off capacities. We first recall the concept of switch channels.

Definition 9: A *switch channel* $\mathcal{N}_{MC\to B}$ between $\mathcal{N}^0_{C\to B}$ and $\mathcal{N}^1_{C\to B}$ with M being a 1-bit switch register is defined as

$$\begin{aligned}&\mathcal{N}_{MC\to B}(\rho_{MC})\\ =&\mathcal{N}^0_{C\to B}(\langle 0|\rho_{MC}|0\rangle_M)+\mathcal{N}^1_{C\to B}(\langle 1|\rho_{MC}|1\rangle_M).\end{aligned}$$

In quantum information theory, switch channels were first used in Ref. [7] to demonstrate the existence of quantum channels such that the quantum capacity is nonzero, but for which pre-shared entanglement does not improve the classical capacity. Subsequently, they were used in Ref. [29] to show the superadditivity of private information, with an alternative definition. A more complicated version was used in Ref. [28] to demonstrate the uncomputability of quantum capacity. Recently, they were also used in Ref. [21] to show the superadditivity of the classical capacity with limited entanglement assistance.

One immediate difficulty is that, even if $\mathcal{N}^0$ and $\mathcal{N}^1$ are well-studied, the dynamic capacity region of $\mathcal{N}$ may not always have a simple expression in terms of those of $\mathcal{N}^0$ and $\mathcal{N}^1$. This is due to the fact that the switch register M can be in a statistical mixture. However, if $\mathcal{N}^0$ and $\mathcal{N}^1$ are unitally extended channels, then the dynamic capacity region of $\mathcal{N}$ does have a simple expression.

Lemma 10: Consider a switch channel $\mathcal{N}_{A'\to B}$ between $\mathcal{N}^0_{RC\to B}$ and $\mathcal{N}^1_{RC\to B}$, with input partition $A'=MRC$ and M being a switch register. Here $\mathcal{N}^0_{RC\to B}$ and $\mathcal{N}^1_{RC\to B}$ are unital extensions of $\Psi^0_{C\to B}$ and $\Psi^1_{C\to B}$ respectively. Then

$$C^{(1)}_{CQE}(\mathcal{N})=\mathrm{Conv}\left(C^{(1)}_{CQE}\left(\mathcal{N}^0\right)\cup C^{(1)}_{CQE}\left(\mathcal{N}^1\right)\right),$$

where Conv denotes the convex hull of points from the two sets.

If, in addition, the quantum dynamic capacity region for $\mathcal{N}^0$ is strongly additive, then we also have

$$C_{CQE}(\mathcal{N})=\mathrm{Conv}\left(C_{CQE}\left(\mathcal{N}^0\right)\cup C_{CQE}\left(\mathcal{N}^1\right)\right).$$

We note that Lemma 10 also applies to the single and double resource capacity regions. This is because these capacity regions are determined by the same set of entropic quantities [8].

The rest of this section is devoted to the proof of this lemma.

Firstly, we note that switch channels and unitally extended channels fall under a broader class of channels that we call partial classical-quantum channels (partial cq channels).

Definition 11: A channel $\Psi_{RC\to B}$ is a *partial cq channel* if there exists a noiseless classical channel $\Pi_{R\to R}$ with orthonormal basis $\{|j\rangle_R\}$

$$\Pi_{R\to R}(\rho)=\sum_j\langle j|\rho|j\rangle|j\rangle\langle j|_R,$$

such that

$$\Psi_{RC\to B}=\Psi_{RC\to B}\circ\Pi_{R\to R}. \tag{14}$$

If there is no register C, then such channels are classical-quantum channels (cq channels).

For partial cq channels, one can always assume inputs are cq states with respect to the input partition R and C for the purpose of evaluating capacities, as we show in Lemma 12 below.

Lemma 12: If $\Psi_{A'\to B}$ is a partial cq channel with partition $A'=RC$, then the optimal trade-off surface of the 1-shot dynamic capacity region $C^{(1)}_{CQE}(\Psi)$ can be achieved with respect to cq states $\sigma_{XAB}=\Psi_{A'\to B}(\rho_{XAA'})$, where $\rho_{XAA'}$ is of the form

$$\rho_{XAA'}=\sum_{x,j}p(x,j)|x,j\rangle\langle x,j|_X\otimes|j\rangle\langle j|_R\otimes\phi^{xj}_{AC}. \tag{15}$$

For $\rho_{XAA'}$ described above, each input state to the channel $|j\rangle\langle j|_R\otimes\phi^{xj}_{AC}$ has entanglement entropy at most $\log|C|$ between input space RC and ancilla A. We can therefore conclude that at most $\log|C|$ ebits of entanglement is useful for the 1-shot dynamic capacity region. This extends similarly to the n-shot dynamic capacity region, for $n>1$.

Proof. We will show that, for any input state

$$\tilde{\rho}_{XAA'}=\sum_x p(x)|x\rangle\langle x|_X\otimes\phi^x_{AA'}, \tag{16}$$

with its output state

$$\varsigma_{XAB}\equiv\Psi_{A'\to B}(\tilde{\rho}_{XAA'})=\sum_x p(x)|x\rangle\langle x|_X\otimes\varsigma^x_{AB},$$

where $\varsigma^x_{AB}=\Psi_{A'\to B}\left(\phi^x_{AA'}\right)$, there exists a corresponding state $\rho_{XAA'}$, in the form of Eq. (15), which can achieve the same rate, if not better.

In fact, the state $\rho_{XAA'}$ can be obtained by applying $\Pi_{R\to R}$ on $\tilde{\rho}_{XAA'}$ and expanding its classical register X. This can be achieved by the following quantum instrument $\mathcal{T}:R\to RX_R$,

$$\mathcal{T}(\psi_R):=\sum_j\langle j|\psi_R|j\rangle|j\rangle\langle j|_R\otimes|j\rangle\langle j|_{X_R}$$

so that

$$\begin{aligned}\rho_{XAA'}&=\mathcal{T}(\tilde{\rho}_{XAA'})\\&=\sum_{x,j}p(x,j)|x,j\rangle\langle x,j|_X\otimes|j\rangle\langle j|_R\otimes\phi^{xj}_{AC}\end{aligned} \tag{17}$$

where we abuse the notation X to denote XX_R in Eq. (17), $p(x,j)\equiv p(x)p(j|x)$, $p(j|x)\equiv\mathrm{Tr}[(|j\rangle\langle j|_R\otimes I_{AC})\phi^x_{ARC}]$ and $\phi^{xj}_{AC}\equiv\langle j|\phi^x_{ARC}|j\rangle_R/p(j|x)$ is still a pure state.

Let $\sigma_{XAB}=\Psi_{A'\to B}(\rho_{XAA'})$. Then

$$\sigma_{XAB}=\sum_{x,j}p(x,j)|x,j\rangle\langle x,j|_X\otimes\sigma^{xj}_{AB}$$

where

$$\sigma^{xj}_{AB}=\Psi_{A'\to B}\left(|j\rangle\langle j|_R\otimes\phi^{xj}_{AC}\right).$$

It follows that

$$\varsigma^x_{AB}=\sum_j p(j|x)\sigma^{xj}_{AB}. \tag{18}$$

Since the dynamic capacity region is fully determined by the three entropic quantities $I(AX;B)$, $I(A\rangle BX)$ and $I(X;B)$ in

Eqs. (4)-(6), it suffices to show that all three entropic quantities evaluated on σ_{XAB} are greater than those evaluated on ς_{XAB}.

1) First consider $I(A\rangle BX)$.

$$\begin{aligned} I(A\rangle BX)_\sigma &= \sum_{x,j} p(x,j) I(A\rangle B)_{\sigma^{xj}} \\ &= \sum_{x,j} p(x)p(j|x) I(A\rangle B)_{\sigma^{xj}} \\ &\geq \sum_{x} p(x) I(A\rangle B)_{\varsigma^{x}} \\ &= I(A\rangle BX)_\varsigma, \end{aligned} \tag{19}$$

where the inequality is due to Eq. (18) and the convexity of coherent information with respect to inputs.

2) Now consider $I(AX;B)$. Similarly,

$$\begin{aligned} I(AX;B)_\sigma &= S(B)_\sigma + I(B\rangle AX)_\sigma \\ &\geq S(B)_\varsigma + I(B\rangle AX)_\varsigma \\ &= I(AX;B)_\varsigma, \end{aligned}$$

where the inequality is due to $\sigma_B = \varsigma_B$ and

$$I(B\rangle AX)_\sigma \geq I(B\rangle AX)_\varsigma$$

follows similarly as Eq. (19), after swapping A and B.

3) Finally consider $I(X;B)$. Writing $|x,j\rangle$ as $|x\rangle|j\rangle$, it can be shown

$$I(X;B)_\sigma \geq I(X;B)_\varsigma$$

using the data processing inequality when we apply the partial trace map $|x\rangle\langle x| \otimes |j\rangle\langle j| \to |x\rangle\langle x|$ to σ_{XB}.

■

Lemma 13: The optimal trade-off surface of the 1-shot quantum dynamic capacity region of a unitally extended channel can always be achieved with σ_{XAB} such that $S(B)_\sigma = \log(|B|)$. This extends similarly to the n-shot dynamic capacity region for $n > 1$.

Proof. Suppose $\Phi_{RC\to B}$ is unitally extended from $\Psi_{C\to B}$. Since a unitally extended channel $\Phi_{RC\to B}$ is a partial cq channel, by Lemma 12, we can consider states of the form

$$\tilde{\rho}_{XAA'} = \sum_{x,j} p(x,j) |x,j\rangle\langle x,j|_X \otimes |j\rangle\langle j|_R \otimes \phi^{xj}_{AC}.$$

Let $\varsigma_{XAB} = \Phi_{RC\to B}(\tilde{\rho}_{XAA'})$ with $A' \equiv RC$. Then

$$\varsigma_{XAB} = \sum_{x,j} p(x,j) |x,j\rangle\langle x,j|_X \otimes \varsigma^{xjj}_{AB},$$

where $\varsigma^{xjk}_{AB} = X(k)\varsigma^{xj}_{AB}X(k)^\dagger$ and $\varsigma^{xj}_{AB} = \Psi_{C\to B}\left(\phi^{xj}_{AC}\right)$.

We can construct another state of the form in Eq. (15):

$$\rho_{X'AA'} = \sum_{x,j,k} p(x,j,k) |x,j,k\rangle\langle x,j,k|_{X'} \otimes |k\rangle\langle k|_R \otimes \phi^{xj}_{AC}, \tag{20}$$

where $p(x,j,k) = p(x,j)/|R|$, and $\sigma_{X'AB} = \Phi_{RC\to B}(\rho_{X'AA'})$:

$$\sigma_{X'AB} = \sum_{x,j,k} p(x,j,k) |x,j,k\rangle\langle x,j,k|_{X'} \otimes \sigma^{xjk}_{AB},$$

where $\sigma^{xjk}_{AB} = X(k)\sigma^{xj}_{AB}X(k)^\dagger$ and $\sigma^{xj}_{AB} = \Psi_{C\to B}\left(\phi^{xj}_{AC}\right)$. The state $\sigma_{X'AB}$ satisfies

$$\begin{aligned} S(B)_\sigma &= S\left(\sum_{x,j,k} p(x,j,k)\sigma^{xjk}_B\right) \\ &\geq \sum_{x,j} p(x,j) S\left(\frac{1}{|R|}\sum_k \sigma^{xjk}_B\right) \\ &= \log(|B|), \end{aligned}$$

where we've used the 1-design formula (12).

One can verify that the dynamic capacity region with $\sigma_{X'AB}$ is larger than that with ς_{XAB} as follows:

$$\begin{aligned} I(A\rangle BX')_\sigma &= \sum_{x,j,k} p(x,j,k) I(A\rangle B)_{\sigma^{xjk}} \\ &= \sum_{x,j} p(x,j) I(A\rangle B)_{\varsigma^{xjj}} = I(A\rangle BX)_\varsigma \end{aligned} \tag{21}$$

$$\begin{aligned} I(AX';B)_\sigma &= S(B)_\sigma + \sum_{x,j,k} p(x,j,k) I(B\rangle A)_{\sigma^{xjk}} \\ &= \log(|B|) + \sum_{x,j} p(x,j) I(B\rangle A)_{\varsigma^{xjj}} \geq I(AX;B)_\varsigma \end{aligned} \tag{22}$$

$$\begin{aligned} I(X';B)_\sigma &= S(B)_\sigma - \sum_{x,j,k} p(x,j,k) S(B)_{\sigma^{xjk}} \\ &= \log(|B|) - \sum_{x,j} p(x,j) S(B)_{\varsigma^{xjj}} \geq I(X;B)_\varsigma. \end{aligned} \tag{23}$$

The key property used in the above equations is, for any Heisenberg-Weyl operator $X(k)$,

$$S(\sigma_B) = S(X(k)\sigma_B X(k)^\dagger).$$

■

Proof of Lemma 8. Following from Lemma 13 and Eq. (20), we only need to consider states of the form

$$\rho_{XAA'} = \sum_{m=0}^{1} p_m |m\rangle\langle m|_M \otimes \rho^m_{XARC} \tag{24}$$

where $p_m = \sum_{x,k} p(x,m,k)$ and

$$\rho^m_{XARC} = \sum_{x,k} \frac{p(x,m,k)}{p_m} |x,m,k\rangle\langle x,m,k|_X \otimes |k\rangle\langle k|_R \otimes \phi^{xm}_{AC},$$

with $p(x,m,k) = p(x,m,k')$ for all k,k' and $m \in \{0,1\}$.

The corresponding channel output is

$$\sigma_{XAB} = \sum_{m=0}^{1} p_m \sigma^m_{XAB} \tag{25}$$

where

$$\sigma^m_{XAB} = \sum_{x,k} \frac{p(x,m,k)}{p_m} |x,m,k\rangle\langle x,m,k|_X \otimes \sigma^{xmk}_{AB} \tag{26}$$

and

$$\sigma^{xmk}_{AB} = X(k)\Psi^m_{C\to B}\left(\phi^{xm}_{AC}\right)X(k)^\dagger.$$

Then all three of the entropic quantities evaluated on σ_{XAB} in Eq. (25) can be decomposed to the corresponding ones evaluated on σ^m_{XAB} given in Eq. (26):

$$\begin{aligned} I(A\rangle BX)_\sigma &= \sum_{m=0}^{1}\sum_{x,k} p(x,m,k) I(A\rangle B)_{\sigma^{xmk}} \\ &= \sum_{m=0}^{1} p_m I(A\rangle BX)_{\sigma^m}. \end{aligned}$$

Likewise,

$$\begin{aligned} I(AX;B)_\sigma &= \log(|B|) + \sum_{m=0}^{1}\sum_{x,k} p(x,m,k) I(B\rangle A)_{\sigma^{xmk}} \\ &= \sum_{m=0}^{1} p_m I(AX;B)_{\sigma^m} \end{aligned}$$

and

$$I(X;B)_\sigma = \sum_{m=0}^{1} p_m I(X;B)_{\sigma^m}.$$

This means if we consider inputs of the form (24), the triple rate for using $\mathcal{N}$ can always be expressed as a linear combination of the triple rates of $\mathcal{N}^0$ and $\mathcal{N}^1$. It is also clear that any linear combination is achievable by the time-sharing principle. Since using states of the form (24) is optimal, we have

$$\begin{aligned} C^{(1)}_{CQE}(\mathcal{N}) &= \bigcup_{0\le p\le 1} pC^{(1)}_{CQE}\left(\mathcal{N}^0\right) + (1-p)C^{(1)}_{CQE}\left(\mathcal{N}^1\right) \\ &= \text{Conv}\left(C^{(1)}_{CQE}\left(\mathcal{N}^0\right) \cup C^{(1)}_{CQE}\left(\mathcal{N}^1\right)\right). \end{aligned}$$

Here again, addition means Minkowski sum. We have also used the fact that the dynamic capacity region of any channel is convex.

Similarly, we have

$$\begin{aligned} C^{(1)}_{CQE}(\mathcal{N}\otimes\mathcal{N}) =&\text{Conv}\Big(C^{(1)}_{CQE}\left(\mathcal{N}^0\otimes\mathcal{N}^0\right) \cup \\ & C^{(1)}_{CQE}\left(\mathcal{N}^0\otimes\mathcal{N}^1\right) \cup C^{(1)}_{CQE}\left(\mathcal{N}^1\otimes\mathcal{N}^1\right)\Big). \end{aligned}$$

If the quantum dynamic capacity region is strongly additive for $\mathcal{N}^0$, then we have

$$C^{(1)}_{CQE}\left(\mathcal{N}^0\otimes\mathcal{N}^1\right) = C^{(1)}_{CQE}\left(\mathcal{N}^0\right) + C^{(1)}_{CQE}\left(\mathcal{N}^1\right) \tag{27}$$

and

$$C^{(1)}_{CQE}\left(\left(\mathcal{N}^0\right)^{\otimes n}\right) = nC_{CQE}\left(\mathcal{N}^0\right). \tag{28}$$

In this case

$$\begin{aligned} & C^{(1)}_{CQE}\left(\mathcal{N}^0\otimes\mathcal{N}^1\right) \\ =&\frac{1}{2}\left(2C_{CQE}\left(\mathcal{N}^0\right) + C^{(1)}_{CQE}\left(\mathcal{N}^1\otimes\mathcal{N}^1\right)\right) \\ \subseteq&\text{Conv}\left(2C_{CQE}\left(\mathcal{N}^0\right) \cup C^{(1)}_{CQE}\left(\mathcal{N}^1\otimes\mathcal{N}^1\right)\right). \end{aligned}$$

Thus the 1-shot quantum dynamic capacity region for $\mathcal{N}\otimes\mathcal{N}$ can be greatly simplified to

$$C^{(1)}_{CQE}(\mathcal{N}\otimes\mathcal{N}) = \text{Conv}\left(2C_{CQE}\left(\mathcal{N}^0\right) \cup C^{(1)}_{CQE}\left(\mathcal{N}^1\otimes\mathcal{N}^1\right)\right).$$

Similarly,

$$\begin{aligned} & C^{(1)}_{CQE}\left(\mathcal{N}^{\otimes k}\right) \\ =&\text{Conv}\Big(C^{(1)}_{CQE}\left(\left(\mathcal{N}^1\right)^{\otimes k}\right) \cup C^{(1)}_{CQE}\left(\mathcal{N}^0\otimes\left(\mathcal{N}^1\right)^{\otimes k-1}\right) \cup \\ & \cdots \cup C^{(1)}_{CQE}\left(\left(\mathcal{N}^0\right)^{\otimes k-1}\otimes\mathcal{N}^1\right) \cup C^{(1)}_{CQE}\left(\left(\mathcal{N}^0\right)^{\otimes k}\right)\Big). \end{aligned}$$

Each term $C^{(1)}_{CQE}\left(\left(\mathcal{N}^0\right)^{\otimes m}\otimes\left(\mathcal{N}^1\right)^{\otimes k-m}\right)$, $0\le m\le k$, can be upper bounded as

$$\begin{aligned} & C^{(1)}_{CQE}\left(\left(\mathcal{N}^0\right)^{\otimes m}\otimes\left(\mathcal{N}^1\right)^{\otimes k-m}\right) \\ =&mC_{CQE}\left(\mathcal{N}^0\right) + C^{(1)}_{CQE}\left(\left(\mathcal{N}^1\right)^{\otimes k-m}\right) \\ \subseteq&mC_{CQE}\left(\mathcal{N}^0\right) + (k-m)C_{CQE}\left(\mathcal{N}^1\right) \\ \subseteq&k\text{Conv}\left(C_{CQE}\left(\mathcal{N}^0\right) \cup C_{CQE}\left(\mathcal{N}^1\right)\right). \end{aligned}$$

Here the second line follows from the addivity of the dynamic capacity region of $\mathcal{N}^0$. The third line follows from the definition of C_{CQE}. The fourth line follows from the definition of convex hull. Thus $C^{(1)}_{CQE}\left(\mathcal{N}^{\otimes k}\right)$ can also be upper bounded as

$$C^{(1)}_{CQE}\left(\mathcal{N}^{\otimes k}\right) \subseteq k\text{Conv}\left(C_{CQE}\left(\mathcal{N}^0\right) \cup C_{CQE}\left(\mathcal{N}^1\right)\right).$$

and

$$\begin{aligned} C_{CQE}(\mathcal{N}) =&\overline{\bigcup_{k=1}^{\infty}\frac{1}{k}C^{(1)}_{CQE}\left(\mathcal{N}^{\otimes k}\right)} \\ \subseteq&\overline{\text{Conv}\left(C_{CQE}\left(\mathcal{N}^0\right) \cup C_{CQE}\left(\mathcal{N}^1\right)\right)} \\ =&\text{Conv}\left(C_{CQE}\left(\mathcal{N}^0\right) \cup C_{CQE}\left(\mathcal{N}^1\right)\right). \end{aligned}$$

The last equality follows because of the topology of the dynamic capacity region, as we show in Appendix C.

By a time-sharing protocol, it is obvious that

$$C_{CQE}(\mathcal{N}) \supseteq \text{Conv}\left(C_{CQE}\left(\mathcal{N}^0\right) \cup C_{CQE}\left(\mathcal{N}^1\right)\right).$$

Hence

$$C_{CQE}(\mathcal{N}) = \text{Conv}\left(C_{CQE}\left(\mathcal{N}^0\right) \cup C_{CQE}\left(\mathcal{N}^1\right)\right).$$

■

While we've been working with Heisenberg-Weyl operators only, we've only used the unitarity of Heisenberg-Weyl operators and the 1-design property (12) in proving the above lemmas. Hence, lemmas 10 and 13 will hold for any unital extension. [2]

Moreover, unital extensions are preserved under tensor product of channels: if Φ^1 is a unital extension of Ψ^1, and Φ^2 is a unital extension of Ψ^2, then $\Phi^1\otimes\Phi^2$ is also a unital extension of $\Psi^1\otimes\Psi^2$. This follows from the fact that if $\{U_j\}\in\mathcal{U}(d_1)$ and $\{V_k\}\in\mathcal{U}(d_2)$ both satisfy Eq. (12), then $\{U_j\otimes V_k\}\in\mathcal{U}(d_1d_2)$ also satisfies Eq. (12).

[2] Note that we do not even require $\mathcal{N}^0$ and $\mathcal{N}^1$ to have the same unital extension. However, to ensure the input dimensions of $\mathcal{N}^0$ and $\mathcal{N}^1$ are the same, their unital extensions must involve the same number of unitaries. For this reason, we stick with the Heisenberg-Weyl operators most of the time.

V. Explicit Construction of Various Superadditivity Phenomena

With the tools developed in Sec IV, we can now explicitly construct channels that satisfy the superadditivity properties stated in Sec III. All our constructions utilize the switch channel idea. We always assume that $\mathcal{N}$ is a switch channel of two unitally extended channels $\mathcal{N}^0$ and $\mathcal{N}^1$. Further, we assume that

(U) $\mathcal{N}^0$ has a strongly additive dynamic capacity region, when tensored with another arbitrary channel.

In this setting, we can use Lemma 10 and its reduction to various single-resource and two-resource capacities.

In each construction, we first state the properties that $\mathcal{N}^0$ and $\mathcal{N}^1$ need to satisfy, in addition to Property (U). We then show how the desired superadditivity of the switch channel $\mathcal{N}$ follows from these properties. In the end, we explicitly construct channels that satisfy the properties we required.

Before we start, we first propose two families of unital extended channels that satisfy (U). Many of our explicit constructions of $\mathcal{N}^0$ will be chosen from these candidates. The first family comes from unital extensions of Hadamard channels. The following lemma shows that the dynamic capacity of the unitally extended Hadamard channels is also additive.

Lemma 14: The dynamic capacity region is strongly additive for a unital extension of a Hadamard channel.

The proof follows from the proof of strong additivity of the dynamic capacity region of a Hadamard channel [24] and the stucture of optimal input states for unitally extended channels.

The second family is unital extensions of classical channels.

Lemma 15: The dynamic capacity region is strongly additive for a classical channel. The same holds for a unital extension of a classical channel.

The dynamic capacity region of a classical channel and its unital extension only has non-trivial classical information transmission. Hence the strong additivity follows from the strong additivity of classical capacity for classical channels.

The detailed proofs of the above lemmas are left to the Appendices, as they are not essential in understanding the construction.

A. Additive C, Superadditive CE

Here we review the original argument in [21] and recast it in the current framework.

We use $C_P(\mathcal{N})$ when we view $C(\mathcal{N})$ as a function of the amount of entanglement assistance P, where $(C(\mathcal{N}), P)$ are points on the CE trade-off curve of $\mathcal{N}$. When $P = 0$, we return to the classical capacity $C_C(\mathcal{N})$. When P is maximal, we arrive at the classical capacity with unlimited entanglement assistance $C_E(\mathcal{N})$. $C_P^{(1)}(\mathcal{N})$ denotes the 1-shot case.

We require $\mathcal{N}^0$ and $\mathcal{N}^1$ to have the following properties:

(A1) $C_C\left(\mathcal{N}^0\right) = C_C\left(\mathcal{N}^1\right)$.

(A2) $\mathcal{N}^1$ has a non-additive CE trade-off capacity region, *i.e.,*

$$\mathcal{C}_{CE}\left(\mathcal{N}^1\right) \supsetneq \mathcal{C}_{CE}^{(1)}\left(\mathcal{N}^1\right),$$

and $\mathcal{C}_{CE}\left(\mathcal{N}^1\right)$ is strictly concave and non-additive at a boundary point of the trade-off region with entanglement consumption $\bar{P}$.

(A3) $\mathcal{C}_{CE}\left(\mathcal{N}^0\right) \subsetneq \mathcal{C}_{CE}\left(\mathcal{N}^1\right)$ and the CE trade-off capacity region of $\mathcal{N}^0$ is strictly smaller than that of $\mathcal{N}^1$ when entanglement consumption is at $\bar{P}$.

In the C_P notation, property (A2) means at $P = \bar{P}$, $C_P\left(\mathcal{N}^1\right) > C_P^{(1)}\left(\mathcal{N}^1\right)$ and $C_P\left(\mathcal{N}^1\right)$ is strictly concave[3] in P at $P = \bar{P}$. Property (A3) means that $C_P\left(\mathcal{N}^0\right) \leq C_P\left(\mathcal{N}^1\right)$ for all P and $C_P\left(\mathcal{N}^0\right) < C_P\left(\mathcal{N}^1\right)$ at $P = \bar{P}$ [4].

Note that Ref. [21] requires $\mathcal{N}^0$ to be a classical channel. However, that is not necessary here.

These three properties (A1)-(A3), together with (U), will guarantee that (i) the classical capacity of $\mathcal{N}$ is additive; and (ii) the CE trade-off capacity region of $\mathcal{N}$ is non-additive at entanglement consumption rate $\bar{P}$.

Combining property (A1) with (U) yields statement (i):

$$\begin{aligned} C_C(\mathcal{N}) &\overset{\text{Lem.10,(U)}}{=} \max\left\{C_C\left(\mathcal{N}^0\right), C_C\left(\mathcal{N}^1\right)\right\} \\ &\overset{\text{(A1)}}{=} C_C\left(\mathcal{N}^0\right) \overset{\text{(U)}}{=} C_C^{(1)}\left(\mathcal{N}^0\right) \\ &\overset{\text{(A1),(U)}}{=} \max\left\{C_C^{(1)}\left(\mathcal{N}^0\right), C_C^{(1)}\left(\mathcal{N}^1\right)\right\} \overset{\text{Lem.10}}{=} C_C^{(1)}(\mathcal{N}). \end{aligned}$$

Here in the first equality, we've used the reduction of Lemma 10 to the classical capacity, with the fact that $\mathcal{N}^0$ has a strongly additive dynamic capacity region (U). The second equality follows from property (A1). In the third equality, we've again used the fact that $\mathcal{N}^0$ has a strongly additive dynamic capacity region (U). By definition $C_C\left(\mathcal{N}^1\right) \geq C_C^{(1)}\left(\mathcal{N}^1\right)$, thus from (A1) we can get $C_C\left(\mathcal{N}^0\right) = C_C^{(1)}\left(\mathcal{N}^0\right) \geq C_C^{(1)}\left(\mathcal{N}^1\right)$, and the fourth equality follows. The last equality follows from Lemma 10. Similar lines of reasoning are used in subsequent sections. Thus we will only indicate the properties used in each step by the superscript.

Property (A3) and Lemma 10 ensure that

$$\begin{aligned} \mathcal{C}_{CE}(\mathcal{N}) &\overset{\text{Lem.10,(U)}}{=} \text{Conv}\left(\mathcal{C}_{CE}\left(\mathcal{N}^0\right) \cup \mathcal{C}_{CE}\left(\mathcal{N}^1\right)\right) \\ &\overset{\text{(A3)}}{=} \mathcal{C}_{CE}\left(\mathcal{N}^1\right). \end{aligned} \tag{29}$$

Since

$$\mathcal{C}_{CE}^{(1)}(\mathcal{N}) \overset{\text{Lem.10,(U)}}{=} \text{Conv}\left(\mathcal{C}_{CE}\left(\mathcal{N}^0\right) \cup \mathcal{C}_{CE}^{(1)}\left(\mathcal{N}^1\right)\right),$$

there exists $P_0, P_1 \geq 0$ and $p \in [0, 1]$ such that $pP_0 + (1-p)P_1 = \bar{P}$ and

$$C_{\bar{P}}^{(1)}(\mathcal{N}) = pC_{P_0}\left(\mathcal{N}^0\right) + (1-p)C_{P_1}^{(1)}\left(\mathcal{N}^1\right). \tag{30}$$

Statement (ii) follows after considering three different cases.

1) $p = 0$.

$$C_{\bar{P}}^{(1)}(\mathcal{N}) \overset{\text{Eq.(30)}}{=} C_{\bar{P}}^{(1)}\left(\mathcal{N}^1\right) \overset{\text{(A2)}}{<} C_{\bar{P}}\left(\mathcal{N}^1\right) \overset{\text{Eq.(29)}}{=} C_{\bar{P}}(\mathcal{N}),$$

where the inequality follows from the superadditivity part of property (A2).

[3]Here by saying a function f is strictly concave at y, we mean $f(y) > (1-p)f(v) + pf(w)$ for all $v < y < w$ satisfying $(1-p)v + pw = y$, with $p \in (0, 1)$.

[4]We wish to emphasize that as long as there are some points in $\mathcal{C}_{CE}(\mathcal{N}_1)$ that is not included in $\mathcal{C}_{CE}(\mathcal{N}_0)$, the inclusion is strict. Hence the two descriptions are the same.

2) $0 < p < 1$. We have

$$C^{(1)}_{\bar{P}}(\mathcal{N}) \overset{\text{Eq.(30)}}{=} pC^{(1)}_{P_0}\left(\mathcal{N}^0\right) + (1-p)C^{(1)}_{P_1}\left(\mathcal{N}^1\right) \overset{\text{(A3)}}{\leq} pC_{P_0}\left(\mathcal{N}^1\right) + (1-p)C_{P_1}\left(\mathcal{N}^1\right) \overset{\text{(A2)}}{<} C_{\bar{P}}\left(\mathcal{N}^1\right) \overset{\text{Eq.(29)}}{=} C_{\bar{P}}(\mathcal{N}).$$

The second inequality follows from the strict concavity part of property (A2).

3) $p = 1$. Then

$$C^{(1)}_{\bar{P}}(\mathcal{N}) \overset{\text{Eq.(30)}}{=} C_{\bar{P}}\left(\mathcal{N}^0\right) \overset{\text{(A3)}}{<} C_{\bar{P}}\left(\mathcal{N}^1\right) \overset{\text{Eq.(29)}}{=} C_{\bar{P}}(\mathcal{N}).$$

Explicit Construction of $\mathcal{N}$: We quote the following property about concave functions [30]: A concave function $u(y)$ is continuous, differentiable from the left and from the right. The derivative is decreasing, *i.e.,* for $x < y$ we have $u'(x-) \geq u'(x+) \geq u'(y-) \geq u'(y+)$. We use "$\pm$" to denote the right and left derivatives when needed.

We first construct $\mathcal{N}^1$. Choose Ψ^{ro} to be a random orthogonal channel with a strictly subadditive minimum output entropy, and Ψ^{ro} has input dimension N. This is unitally extended to Φ^{ro}. As explained in the remark after Definition 7, Φ^{ro} will have a non-additive classical capacity.

Due to Lemma 12, the useful entanglement assistance is at most $\log(N)$. Thus we restrict to $0 \leq P \leq \log(N)$.

Let

$$\epsilon = C_C\left(\Phi^{\text{ro}}\right) - C^{(1)}_C\left(\Phi^{\text{ro}}\right) > 0. \tag{31}$$

Since[5]

$$C^{(1)}_P\left(\Phi^{\text{ro}}\right) \leq C^{(1)}_C\left(\Phi^{\text{ro}}\right) + P, \tag{32}$$

$$C^{(1)}_E\left(\Phi^{\text{ro}}\right) \leq C^{(1)}_C\left(\Phi^{\text{ro}}\right) + \log(N).$$

Since $C_E\left(\Phi^{\text{ro}}\right) = C^{(1)}_E\left(\Phi^{\text{ro}}\right)$ [7], we can conclude

$$C_E\left(\Phi^{\text{ro}}\right) \leq C_C\left(\Phi^{\text{ro}}\right) + \log(N) - \epsilon.$$

This implies $dC_P\left(\Phi^{\text{ro}}\right)/dP$ cannot always be 1. Thus there exists $\bar{P} \in [0, \log(N))$ such that

$$dC_P\left(\Phi^{\text{ro}}\right)/dP = 1, \quad \forall\ 0 \leq P \leq \bar{P}$$

and

$$dC_P\left(\Phi^{\text{ro}}\right)/dP < 1, \quad \forall P > \bar{P}.$$

Next we discuss different cases of $\bar{P}$.

1) $\bar{P} > 0$. Then $C_P\left(\Phi^{\text{ro}}\right)$ is strictly concave at $\bar{P}$. Furthermore, $C_{\bar{P}}\left(\Phi^{\text{ro}}\right) - C^{(1)}_{\bar{P}}\left(\Phi^{\text{ro}}\right) \geq \epsilon$ since $C_{\bar{P}}\left(\Phi^{\text{ro}}\right) = C_C\left(\Phi^{\text{ro}}\right) + \bar{P}$ but $C^{(1)}_{\bar{P}}\left(\Phi^{\text{ro}}\right) \leq C^{(1)}_C\left(\Phi^{\text{ro}}\right) + \bar{P}$. Thus $\mathcal{N}^1 = \Phi^{\text{ro}}$ satisfies (A2).
2) $\bar{P} = 0$. Let $\mathcal{N}^1 = \Phi^{\text{ro}} \otimes \Phi^{\text{dph}}_\eta$, where Φ^{dph}_η is the unital extension of the qubit dephasing channel.
Since $dC_P\left(\Phi^{\text{ro}}\right)/dP|_{0_+} < 1$, choose $\eta > 0$ small such that $dC_P\left(\Phi^{\text{dph}}_\eta\right)/dP|_{1_-} > dC_P\left(\Phi^{\text{ro}}\right)/dP|_{0_+}$. This is possible, as $C_P\left(\Phi^{\text{dph}}_\eta\right) = C_P\left(\Psi^{\text{dph}}_\eta\right)$ and $dC_P\left(\Psi^{\text{dph}}_\eta\right)/dP|_{1_-} \to 1$ as $\eta \to 0$. This ensures that when $0 < P \leq 1$,

$$C_P\left(\mathcal{N}^1\right) = C_C\left(\Phi^{\text{ro}}\right) + C_P\left(\Phi^{\text{dph}}_\eta\right), \tag{33}$$

where we've also used Lemma 14.
For Φ^{dph}_η, it can be shown that $C_P\left(\Phi^{\text{dph}}_\eta\right)$ is strictly concave in P when $\eta < 1/2$ (see Appendix D). Hence $C_P\left(\mathcal{N}^1\right)$ is also strictly concave with respect to P, for $0 < P \leq 1$. Also, when $P < \epsilon$,

$$C_P\left(\mathcal{N}^1\right) > C_C\left(\Phi^{\text{ro}}\right) + C_C\left(\Phi^{\text{dph}}_\eta\right) > C^{(1)}_C\left(\Phi^{\text{ro}}\right) + C_C\left(\Phi^{\text{dph}}_\eta\right) + P \geq C^{(1)}_P\left(\mathcal{N}^1\right).$$

Here the first inequality comes from Eq. (33) and $C_P\left(\Phi^{\text{dph}}_\eta\right) > C_C\left(\Phi^{\text{dph}}_\eta\right)$ when $P > 0$. The second inequality comes from our assumption $P < \epsilon$ and Eq. (31). The last inequality comes from Eq. (32).
This ensures that $C_P\left(\mathcal{N}^1\right)$ is non-additive. Thus when $0 < P < \min\{1, \epsilon\}$, $C_P\left(\mathcal{N}^1\right)$ is strictly concave and non-additive, satisfying (A2).

For $\mathcal{N}^0$, as long as it is a unital extension of a classical channel with $C_C\left(\mathcal{N}^0\right) = C_C\left(\mathcal{N}^1\right)$, it will automatically satisfy property (A3).

B. Additive C and Q, Superadditive CE

In Section V-A, we constructed a channel $\mathcal{N}$ with an additive classical capacity, but a non-additive CE trade-off capacity region. It's unclear if our construction $\mathcal{N}$ has an additive quantum capacity. To extend the argument, we need to make some modifications to the original construction.

In addition to properties (A1)-(A3), the channels $\mathcal{N}^0$ and $\mathcal{N}^1$ need to satisfy

(B1) $C_Q\left(\mathcal{N}^0\right) \geq C_Q\left(\mathcal{N}^1\right)$.

This ensures that the quantum capacity of $\mathcal{N}$ is also additive:

$$C_Q(\mathcal{N}) \overset{\text{Lem.10,(U)}}{=} \max\left\{C_Q\left(\mathcal{N}^0\right), C_Q\left(\mathcal{N}^1\right)\right\} \overset{\text{(B1)}}{=} C_Q\left(\mathcal{N}^0\right) \overset{\text{(U)}}{=} C^{(1)}_Q\left(\mathcal{N}^0\right) \overset{\text{(B1),(U)}}{=} \max\left\{C^{(1)}_Q\left(\mathcal{N}^0\right), C^{(1)}_Q\left(\mathcal{N}^1\right)\right\} \overset{\text{Lem.10}}{=} C^{(1)}_Q(\mathcal{N}).$$

Explicit Construction of $\mathcal{N}$: We take the channels $\mathcal{N}^0$ and $\mathcal{N}^1$ that were constructed in Sec V-A, and compare their quantum capacities. Since $C_Q\left(\mathcal{N}^0\right) = 0$, we can only have $C_Q\left(\mathcal{N}^0\right) \leq C_Q\left(\mathcal{N}^1\right)$. If

$$C_Q\left(\mathcal{N}^0\right) = C_Q\left(\mathcal{N}^1\right),$$

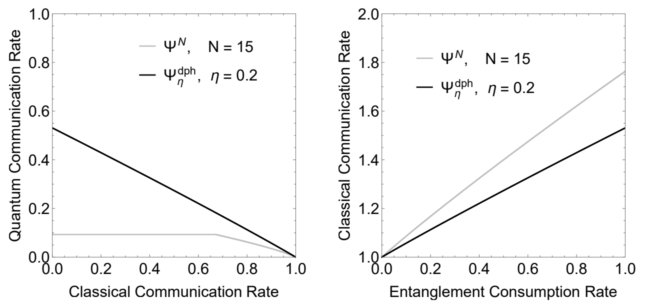
then (B1) is automatically satisfied. Hence we will focus on the case where

$$C_Q\left(\mathcal{N}^0\right) < C_Q\left(\mathcal{N}^1\right).$$

In this case, we call these two channels $\mathit{\Phi}^0$ and $\mathit{\Phi}^1$ respectively. We will construct two new channels $\mathcal{N}^0$ and $\mathcal{N}^1$ that satisfy properties (A1)-(A3) and (B1).

We will use the qubit dephasing channel and $1 \to N$ cloning channel. The expressions of trade-off capacities for these two channels were computed analytically in Ref. [24]. To make the argument work, we will modify them so that the two channels have the same input and output dimension, and the same classical capacity. However, the shape of the trade-off curves are unchanged.

[5]This is implicitly shown in Ref. [8]. Please see Appendix E for a proof.



(a) (b)

Fig. 1. Comparison of trade-off curves between qubit dephasing channel Ψ_η^{dph} and modified $1 \to N$ cloning channel Ψ^N, when $\eta = 0.2$ and $N = 15$. (a) CQ trade-off. (b) CE trade-off.

For the $1 \to N$ cloning channel $\Psi^{1\to N}$, we always tensor an appropriate classical channel, such that the resulting channel has its classical capacity equal to 1, and the output dimension is the same as the input dimension. We denote the resulting channel Ψ^N.

For the dephasing channel, we will tensor a complete depolarizing channel, so that its input and output dimensions match those of Ψ^N. Since tensoring a complete depolarizing channel does not modify the dynamic capacity region of the qubit dephasing channel, we will continue using Ψ_η^{dph} to denote it.

After the above modifications, we observe that for $\eta = 0.2$ and $N = 15$, their trade-off capacities satisfy the following properties (see Fig. 1)

$$C_Q\left(\Psi_\eta^{\text{dph}}\right) > C_Q\left(\Psi^N\right).$$

and

$$C_{CE}\left(\Psi_\eta^{\text{dph}}\right) \subsetneq C_{CE}\left(\Psi^N\right), \tag{34}$$

in the sense that Ψ^N achieves a strictly better classical communication rate than Ψ_η^{dph}, if we have any non-zero amount of entanglement assistance. In the C_P notation, it means $C_P\left(\Psi_\eta^{\text{dph}}\right) < C_P\left(\Psi^N\right)$ for all $P > 0$.

Since unital extensions do not change the CE and CQ trade-off capacity regions of these two channels (see Appendix D), the above properties hold if we replace Ψ_η^{dph} and Ψ^N by their unital extensions Φ_η^{dph} and Φ^N respectively.

Since

$$C_Q\left(\Phi_\eta^{\text{dph}}\right) > C_Q\left(\Phi^N\right),$$

let n be large enough so that

$$nC_Q\left(\Phi_\eta^{\text{dph}}\right) + C_Q\left(\mathit{\Phi}^0\right) \geq nC_Q\left(\Phi^N\right) + C_Q\left(\mathit{\Phi}^1\right).$$

Define

$$\mathcal{N}^0 = \left(\Phi_\eta^{\text{dph}}\right)^{\otimes n} \otimes \mathit{\Phi}^0$$

and

$$\mathcal{N}^1 = \left(\Phi^N\right)^{\otimes n} \otimes \mathit{\Phi}^1.$$

Our choice of n ensures that

$$C_Q\left(\mathcal{N}^0\right) \geq C_Q\left(\mathcal{N}^1\right).$$

We also need to ensure our newly constructed $\mathcal{N}^0$ and $\mathcal{N}^1$ still satisfy properties (A1)-(A3).

As

$$C_C\left(\Phi_\eta^{\text{dph}}\right) = C_C\left(\Phi^N\right) = 1$$

and

$$C_C\left(\mathit{\Phi}^0\right) = C_C\left(\mathit{\Phi}^1\right),$$

we immediately have

$$C_C\left(\mathcal{N}^0\right) = C_C\left(\mathcal{N}^1\right)$$

and property (A1) is satisfied.

The CE trade-off curve of $\Psi^{1\to N}$ is strictly concave for $N \neq 1$ [24], hence property (A2) is also satisfied for $\mathcal{N}^1$.

Property (A3) is satisfied due to Eq. (34).

C. Additive Q, Superadditive CQ

We require $\mathcal{N}^0$ and $\mathcal{N}^1$ to have the following properties:

(C1) $C_Q\left(\mathcal{N}^0\right) \geq C_Q\left(\mathcal{N}^1\right)$.
(C2) $C_C\left(\mathcal{N}^1\right) > C_C^{(1)}\left(\mathcal{N}^1\right)$.
(C3) $C_C\left(\mathcal{N}^1\right) > C_C\left(\mathcal{N}^0\right)$.

These properties (C1)-(C3) will allow us to show that (i) $C_Q(\mathcal{N}) = C_Q^{(1)}(\mathcal{N})$; and (ii) $C_{CQ}(\mathcal{N}) \supsetneq C_{CQ}^{(1)}(\mathcal{N})$.

Statement (i) follows from property (C1) and (U) that $\mathcal{N}^0$ has an additive quantum capacity:

$$\begin{aligned}
C_Q(\mathcal{N}) &\overset{\text{Lem.10,(U)}}{=} \max\left\{C_Q\left(\mathcal{N}^0\right), C_Q\left(\mathcal{N}^1\right)\right\} \\
&\overset{\text{(C1)}}{=} C_Q\left(\mathcal{N}^0\right) \overset{\text{(U)}}{=} C_Q^{(1)}\left(\mathcal{N}^0\right) \\
&\overset{\text{(C1),(U)}}{=} \max\left\{C_Q^{(1)}\left(\mathcal{N}^0\right), C_Q^{(1)}\left(\mathcal{N}^1\right)\right\} \overset{\text{Lem.10}}{=} C_Q^{(1)}(\mathcal{N}).
\end{aligned}$$

Properties (C2) and (C3) together ensure

$$\begin{aligned}
C_C(\mathcal{N}) &\overset{\text{Lem.10,(U)}}{=} \max\left\{C_C\left(\mathcal{N}^0\right), C_C\left(\mathcal{N}^1\right)\right\} \overset{\text{(C3)}}{=} C_C\left(\mathcal{N}^1\right) \\
&\overset{\text{(C2),(C3)}}{>} \max\left\{C_C^{(1)}\left(\mathcal{N}^0\right), C_C^{(1)}\left(\mathcal{N}^1\right)\right\} \overset{\text{Lem.10}}{=} C_C^{(1)}(\mathcal{N}),
\end{aligned}$$

i.e., the classical capacity of $\mathcal{N}$ is non-additive; hence statement (ii) follows.

Explicit Construction of $\mathcal{N}$: Next we construct $\mathcal{N}^0$ and $\mathcal{N}^1$ that satisfy the above properties.

Let Ψ^{ro} be a random orthogonal channel, such that its unital extension has a non-additive classical capacity. For convenience, we also assume Ψ^{ro} has the input dimension $N = 2^n$. Choose η for the qubit dephasing channel Ψ_η^{dph} such that $C_Q\left(\Psi^{\text{ro}}\right) + C_Q\left(\Psi_\eta^{\text{dph}}\right) = m$ for some integer m.

Define

$$\mathcal{N}^1 = \Phi^{\text{ro}} \otimes \Phi_\eta^{\text{dph}},$$

where Φ^{ro} is a unital extension of Ψ^{ro} and Φ_η^{dph} is a unital extension of Ψ_η^{dph}. $\mathcal{N}^1$ has the property that its quantum capacity is $C_Q\left(\mathcal{N}^1\right) = m$, whereas its classical capacity is non-additive, and greater than m.

Define

$$\mathcal{N}^0 = \left(\Phi^{\mathcal{I}}\right)^{\otimes m} \otimes \left(\Phi_1^{\text{dpo}}\right)^{\otimes n+1-m},$$

where $\Phi^{\mathcal{I}}$ is a unital extension of the qubit noiseless channel, and Φ_1^{dpo} is a unital extension of the qubit completely depolarizing channel.

We note that the qubit noiseless channel is a special instance of a qubit dephasing channel. Its classical and quantum capacity are both 1, and this remain unchanged under a unital extension (following Appendix D). For the qubit completely depolarizing channel, it always outputs a maximally mixed state, and this remain unchanged under a unital extension. Thus Φ_1^{dpo} has zero classical and quantum capacity.

As a result, $\mathcal{N}^0$ has its classical and quantum capacity as $C_C\left(\mathcal{N}^0\right) = C_Q\left(\mathcal{N}^0\right) = m$, thus fulfiling the properties (C1) and (C3) above.

D. Additive C and Q, Superadditive CQ

We require $\mathcal{N}^0$ and $\mathcal{N}^1$ to satisfy the following properties:

(D1) $C_C\left(\mathcal{N}^0\right) = C_C\left(\mathcal{N}^1\right)$ and $C_Q\left(\mathcal{N}^0\right) = C_Q\left(\mathcal{N}^1\right)$.

(D2) $\mathcal{N}^1$ has a non-additive CQ trade-off capacity region, meaning
$$\mathcal{C}_{CQ}\left(\mathcal{N}^1\right) \supsetneq \mathcal{C}_{CQ}^{(1)}\left(\mathcal{N}^1\right).$$
$\mathcal{C}_{CQ}\left(\mathcal{N}^1\right)$ is strictly concave and non-additive at a boundary point with classical communication rate $\bar{C}$.

(D3) $\mathcal{C}_{CQ}\left(\mathcal{N}^1\right) \supsetneq \mathcal{C}_{CQ}\left(\mathcal{N}^0\right)$ and the CQ trade-off capacity region of $\mathcal{N}^1$ is strictly larger than that of $\mathcal{N}^0$ when classical communication rate is at $\bar{C}$.

With these properties, we can show that (i) $C_C(\mathcal{N}) = C_C^{(1)}(\mathcal{N})$; (ii) $C_Q(\mathcal{N}) = C_Q^{(1)}(\mathcal{N})$; and (iii) $\mathcal{C}_{CQ}(\mathcal{N}) \supsetneq \mathcal{C}_{CQ}^{(1)}(\mathcal{N})$.

We'll focus on the CQ trade-off curve. Same as in Section V-A, we use a simplified notation $Q_C(\mathcal{N})$ when we view $Q(\mathcal{N})$ as a function of $C(\mathcal{N})$. In the 1-shot scenario, it is denoted by $Q_C^{(1)}(\mathcal{N})$. We'll show there exists $\bar{C} \neq 0$ such that $Q_{\bar{C}}(\mathcal{N}) > Q_{\bar{C}}^{(1)}(\mathcal{N})$.

In the Q_C notation, property (D2) means at $C = \bar{C}$, $Q_C\left(\mathcal{N}^1\right) > Q_C^{(1)}\left(\mathcal{N}^1\right)$ and $Q_C\left(\mathcal{N}^1\right)$ is strictly concave in C at $C = \bar{C}$. Property (D3) implies that $Q_C\left(\mathcal{N}^1\right) > Q_C\left(\mathcal{N}^0\right)$ at $C = \bar{C}$.

Properties (D1) and (U) ensure that
$$\begin{aligned} C_C(\mathcal{N}) &\overset{\text{Lem.10,(U)}}{=} \max\left\{C_C\left(\mathcal{N}^0\right), C_C\left(\mathcal{N}^1\right)\right\} \\ &\overset{\text{(D1)}}{=} C_C\left(\mathcal{N}^1\right) \overset{\text{(U)}}{=} C_C^{(1)}\left(\mathcal{N}^1\right) \\ &\overset{\text{(D1),(U)}}{=} \max\left\{C_C^{(1)}\left(\mathcal{N}^0\right), C_C^{(1)}\left(\mathcal{N}^1\right)\right\} \overset{\text{Lem.10}}{=} C_C^{(1)}(\mathcal{N}) \end{aligned}$$
and similarly
$$\begin{aligned} C_Q(\mathcal{N}) &\overset{\text{Lem.10,(U)}}{=} \max\left\{C_Q\left(\mathcal{N}^0\right), C_Q\left(\mathcal{N}^1\right)\right\} \\ &\overset{\text{(D1)}}{=} C_Q\left(\mathcal{N}^0\right) \overset{\text{(U)}}{=} C_Q^{(1)}\left(\mathcal{N}^0\right) \\ &\overset{\text{(D1),(U)}}{=} \max\left\{C_Q^{(1)}\left(\mathcal{N}^0\right), C_Q^{(1)}\left(\mathcal{N}^1\right)\right\} \overset{\text{Lem.10}}{=} C_Q^{(1)}(\mathcal{N}), \end{aligned}$$
i.e., $\mathcal{N}$ has an additive classical and quantum capacity.

By properties (D3) and (U), we have
$$\begin{aligned} \mathcal{C}_{CQ}(\mathcal{N}) &\overset{\text{Lem.10,(U)}}{=} \text{Conv}\left(\mathcal{C}_{CQ}\left(\mathcal{N}^0\right) \cup \mathcal{C}_{CQ}\left(\mathcal{N}^1\right)\right) \\ &\overset{\text{(D3),(U)}}{=} \mathcal{C}_{CQ}\left(\mathcal{N}^1\right). \end{aligned} \tag{35}$$

Since
$$\mathcal{C}_{CQ}^{(1)}(\mathcal{N}) \overset{\text{Lem.10,(U)}}{=} \text{Conv}\left(\mathcal{C}_{CQ}\left(\mathcal{N}^0\right) \cup \mathcal{C}_{CQ}^{(1)}\left(\mathcal{N}^1\right)\right),$$
there exists C_0, C_1 and $p \in [0, 1]$ such that $pC_0 + (1-p)C_1 = \bar{C}$ and
$$Q_{\bar{C}}^{(1)}(\mathcal{N}) = pQ_{C_0}\left(\mathcal{N}^0\right) + (1-p)Q_{C_1}^{(1)}\left(\mathcal{N}^1\right). \tag{36}$$
Now consider three different cases.

1) $p = 0$. We have
$$Q_{\bar{C}}^{(1)}(\mathcal{N}) \overset{\text{Eq.(36)}}{=} Q_{\bar{C}}^{(1)}\left(\mathcal{N}^1\right) \overset{\text{(D2)}}{<} Q_{\bar{C}}\left(\mathcal{N}^1\right) \overset{\text{Eq.(35)}}{=} Q_{\bar{C}}(\mathcal{N}).$$

2) $0 < p < 1$.
$$\begin{aligned} Q_{\bar{C}}^{(1)}(\mathcal{N}) &\overset{\text{Eq.(36)}}{=} pQ_{C_0}\left(\mathcal{N}^0\right) + (1-p)Q_{C_1}^{(1)}\left(\mathcal{N}^1\right) \\ &\overset{\text{(D3)}}{\leq} pQ_{C_0}\left(\mathcal{N}^1\right) + (1-p)Q_{C_1}\left(\mathcal{N}^1\right) \\ &\overset{\text{(D2)}}{<} Q_{\bar{C}}\left(\mathcal{N}^1\right) \overset{\text{Eq.(35)}}{=} Q_{\bar{C}}(\mathcal{N}). \end{aligned}$$
Here the second inequality follows from the strict concavity part of property (D2).

3) $p = 1$. Then
$$Q_{\bar{C}}^{(1)}(\mathcal{N}) \overset{\text{Eq.(36)}}{=} Q_{\bar{C}}\left(\mathcal{N}^0\right) \overset{\text{(D3)}}{<} Q_{\bar{C}}\left(\mathcal{N}^1\right) \overset{\text{Eq.(35)}}{=} Q_{\bar{C}}(\mathcal{N}).$$

Hence statement (iii) follows.

Explicit Construction: Now we explicitly construct $\mathcal{N}^0$ and $\mathcal{N}^1$.

Choose p such that the qubit depolarizing channel Ψ_p^{dpo} is known to have a non-additive quantum capacity. Consider its unital extension Φ_p^{dpo}. Note that the gradient $dQ_C\left(\Phi_p^{\text{dpo}}\right)/dC$ of the CQ trade-off curve cannot always stay at 0 for the choice of Ψ_p^{dpo} with a positive quantum capacity. It means there exists $0 \leq \bar{C} < C_C\left(\Phi_p^{\text{dpo}}\right)$ such that
$$dQ_C\left(\Phi_p^{\text{dpo}}\right)/dC = 0, \quad \forall 0 \leq C \leq \bar{C}_- \tag{37}$$
and
$$dQ_C\left(\Phi_p^{\text{dpo}}\right)/dC < 0, \quad \forall \bar{C}_+ \leq C \leq C_C\left(\Phi_p^{\text{dpo}}\right).$$

1) $\bar{C} > 0$. In this case, we know $Q_C\left(\Phi_p^{\text{dpo}}\right)$ is strictly concave at $\bar{C}$. Also
$$Q_{\bar{C}}\left(\Phi_p^{\text{dpo}}\right) = Q_0\left(\Phi_p^{\text{dpo}}\right) > Q_0^{(1)}\left(\Phi_p^{\text{dpo}}\right) \geq Q_{\bar{C}}^{(1)}\left(\Phi_p^{\text{dpo}}\right).$$
Here the equality follows from Eq. (37). The first inequality follows because Ψ_p^{dpo} has a non-additive quantum capacity, as both C_Q and $C_Q^{(1)}$ remain unchanged after a unital extension, and Q_C reduces to the quantum capacity at $C = 0$. The second inequality follows as the rate of quantum communication along the CQ trade-off curve must not exceed the quantum capacity.

Choose the noise parameter η for the qubit dephasing channel Ψ_η^{dph} appropriately such that
$$C_Q\left(\Psi_\eta^{\text{dph}}\right) = 1 - C_Q\left(\Psi_p^{\text{dpo}}\right).$$
Define
$$\mathcal{N}^1 = \Phi_p^{\text{dpo}} \otimes \Phi_\eta^{\text{dph}}.$$

It's clear that $\mathcal{N}^1$ is a unitally extended channel of $\Psi_p^{\text{dpo}} \otimes \Psi_\eta^{\text{dph}}$ and has $C_Q\left(\mathcal{N}^1\right) = C_Q\left(\Psi_p^{\text{dpo}} \otimes \Psi_\eta^{\text{dph}}\right) = 1$. The CQ trade-off curve is strictly concave and non-additive at $\bar{C}$. The corresponding Ψ^0 is

$$\Psi^0 = \mathcal{I} \otimes \Psi_1^{\text{dpo}},$$

i.e., a noiseless channel tensor a complete qubit depolarizing channel. $\mathcal{N}^0$ is a unital extension of Ψ^0.

2) $\bar{C} = 0$. Choose η_1 close to 1/2 such that

$$\left.\frac{dQ_C\left(\Phi_{\eta_1}^{\text{dph}}\right)}{dC}\right|_{C_C\left(\Psi_{\eta_1}^{\text{dph}}\right)_-} > \left.\frac{dQ_C\left(\Phi_p^{\text{dpo}}\right)}{dC}\right|_{0_+}.$$

Let

$$\mathcal{N}^1 = \Phi_{\eta_1}^{\text{dph}} \otimes \Phi_p^{\text{dpo}} \otimes \Phi_{\eta_2}^{\text{dph}},$$

where η_2 is chosen such that

$$\begin{aligned} C_Q\left(\mathcal{N}^1\right) =& C_Q\left(\Psi_{\eta_1}^{\text{dph}} \otimes \Psi_p^{\text{dpo}} \otimes \Psi_{\eta_2}^{\text{dph}}\right) \\ =& C_Q\left(\Psi_{\eta_1}^{\text{dph}}\right) + C_Q\left(\Psi_p^{\text{dpo}}\right) + C_Q\left(\Psi_{\eta_2}^{\text{dph}}\right) = 1. \end{aligned}$$

By our choice of η_1, $Q_C\left(\Phi_{\eta_1}^{\text{dph}} \otimes \Phi_p^{\text{dpo}}\right)$ is strictly concave in C for $0 < C < 1$. $Q_C\left(\Phi_{\eta_2}^{\text{dph}}\right)$ is also strictly concave in C. Thus $Q_C\left(\mathcal{N}^1\right)$ is strictly concave in C, for $0 < C < 1$. In this case, the corresponding Ψ^0 is

$$\Psi^0 = \mathcal{I} \otimes \left(\Psi_1^{\text{dpo}}\right)^{\otimes 2},$$

i.e., a noiseless channel tensor two copies of the complete qubit depolarizing channel. $\mathcal{N}^0$ is a unital extension of Ψ^0.

E. Additive CE, Superadditive Q and CQE

Here we construct a channel that has an additive CE trade-off capacity region, but a non-additive quantum capacity, hence a non-additive quantum dynamic capacity region.

Let Ψ^0 be a classical channel and Ψ^1 be the depolarizing channel Ψ_p^{dpo}. p is chosen such that Ψ_p^{dpo} has a non-additive quantum capacity. Also, we require

$$C_C\left(\Psi^0\right) > C_E\left(\Psi^1\right). \tag{38}$$

Now consider the switch channel $\mathcal{N}$, consisting of $\mathcal{N}^0$ and $\mathcal{N}^1$, which are unital extensions of Ψ^0 and Ψ^1. It can be easily shown that unital extension does not change the classical capacity with umlimited entanglement assistance of the qubit depolarizing channel. Thus Eq. (38) implies

$$C_{CE}\left(\mathcal{N}^0\right) \supseteq C_{CE}\left(\mathcal{N}^1\right) \supseteq C_{CE}^{(1)}\left(\mathcal{N}^1\right). \tag{39}$$

Hence

$$\begin{aligned} C_{CE}(\mathcal{N}) &\overset{\text{Lem.10,(U)}}{=} \text{Conv}\left(C_{CE}\left(\mathcal{N}^0\right) \cup C_{CE}\left(\mathcal{N}^1\right)\right) \\ &\overset{\text{Eq.(38)}}{=} C_{CE}\left(\mathcal{N}^0\right) \overset{\text{(U)}}{=} C_{CE}^{(1)}\left(\mathcal{N}^0\right) \\ &\overset{\text{Eq.(38)}}{=} \text{Conv}\left(C_{CE}\left(\mathcal{N}^0\right) \cup C_{CE}^{(1)}\left(\mathcal{N}^1\right)\right) \\ &\overset{\text{Lem.10,(U)}}{=} C_{CE}^{(1)}(\mathcal{N}), \end{aligned} \tag{40}$$

i.e., its CE trade-off capacity region is additive.

Since $C_Q\left(\mathcal{N}^0\right) = 0$, it is clear that the quantum capacity of $\mathcal{N}$ is the same as that of $\mathcal{N}^1$, which is non-additive.

Note that $\mathcal{N}$ is a unitally extended channel. This fact will be implicitly used in Section V-F.

F. Additive CE and Q, Superadditive CQE

Previously in Section V-D, we give an example of a channel with an additive classical and quantum capacity, but whose CQ trade-off capacity region is non-additive. It is unclear if the channel has an additive CE trade-off capacity region, because the CE trade-off capacity region of the depolarizing channel has not been shown to be additive. This is itself an interesting question but we'll not explore it here.

We replace Ψ_p^{dpo} in the original argument of Section V-D by the channel constructed in Section V-E. It's clear that the rest of the argument is not changed and $\mathcal{N}$ still has a non-additive CQ trade-off capacity region.

Now both $\mathcal{N}^0$ and $\mathcal{N}^1$ have an additive CE trade-off capacity region. It's clear that

$$\begin{aligned} C_{CE}(\mathcal{N}) &\overset{\text{Lem.10,(U)}}{=} \text{Conv}\left(C_{CE}\left(\mathcal{N}^0\right) \cup C_{CE}\left(\mathcal{N}^1\right)\right) \\ &\overset{\text{Eq.(40),(U)}}{=} \text{Conv}\left(C_{CE}^{(1)}\left(\mathcal{N}^0\right) \cup C_{CE}^{(1)}\left(\mathcal{N}^1\right)\right) \overset{\text{Lem.10}}{=} C_{CE}^{(1)}(\mathcal{N}), \end{aligned}$$

i.e., the CE trade-off capacity region of $\mathcal{N}$ is additive.

G. Additive CQ, Superadditive CQE (Conjecture)

Our construction in Section V-A has a non-additive CE trade-off capacity region. But most likely its CQ trade-off capacity region is also non-additive. This is because in Section V-A, $\mathcal{N}^0$ is the unital extension of a classical channel, and its CQ trade-off capacity region is trivial. Hence the CQ trade-off capacity region of $\mathcal{N}$ is given by that of $\mathcal{N}^1$, which is most likely non-additive as well.

To achieve an additive CQ trade-off capacity region, we have to substitute $\mathcal{N}^0$ with a channel that has a non-trivial CQ trade-off capacity region.

Recall that our construction in Section V-A requires $\mathcal{N}^0$ and $\mathcal{N}^1$ to have properties (A1)-(A3). These three properties ensure that $\mathcal{N}$ will have a non-additive CE trade-off capacity region, while its classical capacity is still additive.

In extending to a channel with an additive CQ trade-off capacity region, the additional properties we need are

(G1) $C_{CQ}\left(\mathcal{N}^0\right) \supseteq C_{CQ}\left(\mathcal{N}^1\right)$.

Property (G1) and (U) ensure the CQ trade-off capacity region of $\mathcal{N}$ is additive, as

$$\begin{aligned} C_{CQ}(\mathcal{N}) &\overset{\text{Lem.10,(U)}}{=} \text{Conv}\left(C_{CQ}\left(\mathcal{N}^0\right) \cup C_{CQ}\left(\mathcal{N}^1\right)\right) \\ &\overset{\text{(G1)}}{=} C_{CQ}\left(\mathcal{N}^0\right) \overset{\text{(U)}}{=} C_{CQ}^{(1)}\left(\mathcal{N}^0\right) \\ &\overset{\text{(G1)}}{=} \text{Conv}\left(C_{CQ}^{(1)}\left(\mathcal{N}^0\right) \cup C_{CQ}^{(1)}\left(\mathcal{N}^1\right)\right) \\ &\overset{\text{Lem.10}}{=} C_{CQ}^{(1)}(\mathcal{N}). \end{aligned}$$

Unfortunately, we cannot find quantum channels $\mathcal{N}^0$ and $\mathcal{N}^1$ that satisfy all the properties. Hence we do not have an explicit construction in this case. This is because there are very few channels whose dynamic capacity regions we understand. This leaves us with a limited choice of candidates for $\mathcal{N}^0$. However, in principle there is no obstacle and the construction will be readily available once we have a better understanding of quantum channels.

VI. Conclusion

Unlike previous studies on additivity of single resource channel capacity, our work aimed to understand how additivity of single or double resource capacity regions will effect additivity of a general resource trade-off capacity. In contrast to the two known results in the literature; namely, (i) additivity of the quantum capacity implies additivity of the entanglement-assisted quantum capacity region and (ii) additivity of classical-quantum and classical-entanglement capacity regions implies additivity of the three resource capacity region, the additivity of all the remaining situations does not hold. In this work, we identified nearly all possible occurrences where superadditivity could occur in the trade-off quantum dynamic capacity. Furthermore, we provided an explicit construction of quantum channels for most instances. Our main technical tool combines properties of switch channels and unital extension of known quantum channels.

An obvious open question is an explicit construction of a quantum channel whose classical-quantum capacity region is additive, but its triple trade-off capacity is non-additive. Moreover, there are other triple resource trade-off capacity regions [4], [31]. Could similar statements made in this work hold in these scenarios as well?

Acknowledgment

EYZ and PWS are supported by the National Science Foundation under grant Contract Number CCF-1525130. QZ is supported by the Claude E. Shannon Research Assistantship. MH is supported by an ARC Future Fellowship under Grant FT140100574. PWS is supported by the NSF through the STC for Science of Information under grant number CCF0-939370.

Appendix A
Proof of Lemma 14

Proof. Consider $\Phi^0_{RC\to B^0}$ and $\Psi^1_{A^1\to B^1}$, where Φ^0 is a unital extension of a Hadamard channel $\Psi^0_{C\to B^0}$, and Ψ^1 is an arbitrary channel.

The result follows if both the CQ and CE trade-off capacity regions of Φ^0 are additive [4]. To show that the CQ trade-off capacity region is additive for Φ^0, it was shown in Ref. [24] it suffices to prove that

$$f_\lambda\left(\Phi^0\otimes\Psi^1\right) = f_\lambda\left(\Phi^0\right) + f_\lambda\left(\Psi^1\right) \tag{41}$$

for any channel Ψ^1, where

$$f_\lambda(\mathcal{N}) = \max_\rho I(X;B)_\sigma + \lambda I(A\rangle BX)_\sigma. \tag{42}$$

The state σ is the channel output state with ρ being the input state (see, e.g., Theorem 1). The reason why $f_\lambda(\mathcal{N})$ is considered is that, with different values of λ, the function leads to points on the CQ trade-off curve. For a more detailed argument, please see Ref. [24].

In the following, we will only show that $f_\lambda\left(\Phi^0\otimes\Psi^1\right) \le f_\lambda\left(\Phi^0\right) + f_\lambda\left(\Psi^1\right)$ because the other direction is trivial from its definition.

Since $\Phi^0\otimes\Psi^1 : CRA^1 \to B^0B^1$ is a partial cq channel, then by the same argument as that in Lemma 12, $f_\lambda\left(\Phi^0\otimes\Psi^1\right)$ can be achieved with input states of the following form

$$\rho_{XRACA^1} = \sum_{x,j}\frac{p(x)}{|R|}|x,j\rangle\langle x,j|_X\otimes|j\rangle\langle j|_R\otimes\phi^x_{ACA^1},$$

with output states

$$\sigma_{XAB^0B^1} = \sum_{x,j}\frac{p(x)}{|R|}|x,j\rangle\langle x,j|_X\otimes\sigma^{xj}_{AB^0B^1}, \tag{43}$$

where

$$\sigma^{xj}_{AB^0B^1} = \Phi^0\otimes\Psi^1\left(|j\rangle\langle j|_R\otimes\phi^x_{ACA^1}\right).$$

Let $U^0_{C\to B^0E^0}$ and $U^1_{A^1\to B^1E^1}$ be the isometric extensions of Ψ^0 and Ψ^1, and let

$$\begin{aligned}
\varrho_{XACA^1} &= \sum_x p(x)|x\rangle\langle x|_X\otimes\phi^x_{ACA^1}\\
\omega_{XAA^1B^0E^0} &= \left(U^0\otimes I\right)\varrho_{XACA^1}\left(U^0\otimes I\right)^\dagger\\
\varsigma_{XAB^0B^1E^0E^1} &= \left(U^0\otimes U^1\right)\varrho_{XACA^1}\left(U^0\otimes U^1\right)^\dagger.
\end{aligned}$$

Moreover, let

$$\theta_{XYAB^1E^0E^1} = \mathcal{D}^1_{B^0\to Y}\left(\varsigma_{XAB^0B^1E^0E^1}\right),$$

where $\mathcal{D}^2_{Y\to E^0}\circ\mathcal{D}^1_{B^0\to Y} = \mathcal{D}_{B^0\to E^0}$ is a degrading map for the Hadamard channel Ψ^0.

For any state $\sigma_{XAB^0B^1}$ in Eq. (43), we have

$$\begin{aligned}
&f_\lambda\left(\Phi^0\otimes\Psi^1\right)\\
=&I\left(X;B^0B^1\right)_\sigma + \lambda I\left(A\rangle B^0B^1X\right)_\sigma\\
=&S\left(B^0B^1\right)_\sigma + \left[(\lambda-1)S\left(B^0B^1|X\right)_\sigma - \lambda S\left(AB^0B^1|X\right)_\sigma\right]\\
=&S\left(B^0B^1\right)_\varsigma + \left[(\lambda-1)S\left(B^0B^1|X\right)_\varsigma - \lambda S\left(AB^0B^1|X\right)_\varsigma\right],
\end{aligned}$$

where the last equality follows from the same argument used in Eqs. (21) and (23). Then subadditivity of the von Neumann entropy and chain rule yield

$$\begin{aligned}
\le&S\left(B^0\right)_\varsigma + (\lambda-1)S\left(B^0|X\right)_\varsigma - \lambda S\left(E^0|X\right)_\varsigma\\
&+S\left(B^1\right)_\varsigma + (\lambda-1)S\left(B^1|B^0X\right)_\varsigma - \lambda S\left(E^1|E^0X\right)_\varsigma\\
\le&S\left(B^0\right)_\varsigma + (\lambda-1)S\left(B^0|X\right)_\varsigma - \lambda S\left(E^0|X\right)_\varsigma\\
&+S\left(B^1\right)_\theta + (\lambda-1)S\left(B^1|XY\right)_\theta - \lambda S\left(E^1|XY\right)_\theta
\end{aligned}$$

where the last inequality uses the fact that $S\left(B^1|B^0X\right)_\varsigma \leq S\left(B^1|YX\right)_\theta$ due to the existence of $\mathcal{D}^1$ and $S\left(E^1|E^0X\right)_\varsigma \geq S\left(E^1|YX\right)_\theta$ due to the existence of $\mathcal{D}^2$. Finally,

$$\begin{aligned}
&= \left(I\left(X;B^0\right)_\omega + \lambda I\left(AA^1\rangle B^0X\right)_\omega\right) \\
&+ \left(I\left(XY;B^1\right)_\theta + \lambda I\left(AE^0\rangle B^1XY\right)_\theta\right) \\
&\leq f_\lambda\left(\Phi^0\right) + f_\lambda\left(\Psi^1\right)
\end{aligned}$$

because $S(E^0|X)_\varsigma = S(AA^1B^0|X)_\omega$ and $S\left(E^1|XY\right)_\theta = S\left(AB^1E^0|XY\right)_\theta$.

To prove that the CE trade-off capacity region of the channel Φ^0 is additive is equivalent to showing that [24]:

$$g_\lambda\left(\Phi^0\otimes\Psi^1\right) = g_\lambda\left(\Phi^0\right) + g_\lambda\left(\Psi^1\right), \tag{44}$$

where $0 \leq \lambda < 1$,

$$g_\lambda(\mathcal{N}) = \max_\sigma I(AX;B)_\sigma - \lambda S(A|X)_\sigma \tag{45}$$

and σ is of the form given in Eq. (7).

However this proof proceeds similarly; hence, we will omit it. ■

Appendix B
Proof of Lemma 15

Lemma 15: The dynamic capacity region of a classical channel Ψ^0 is strongly additive. It is described by the following relation

$$\begin{aligned}
C + 2Q &\leq C_C\left(\Psi^0\right), \\
Q + E &\leq 0, \\
C + Q + E &\leq C_C\left(\Psi^0\right),
\end{aligned}$$

where $C_C\left(\Psi^0\right)$ is the classical capacity of Ψ^0.

The same holds for a unital extension of a classical channel.

Proof. Consider the 1-shot dynamic capacity region of $\Psi^0_{A^{0\prime}\to B^0}$. By Lemma 12, $C^{(1)}_{CQE}\left(\Psi^0\right)$ can be achieved with respect to cq states $\sigma_{XA^0B^0} = \Psi^0\left(\rho_{XA^0A^{0\prime}}\right)$, where $\rho_{XA^0A^{0\prime}}$ is of the form

$$\rho_{XA^0A^{0\prime}} = \sum_{x,j} p(x,j)\left|x,j\right\rangle\left\langle x,j\right|_X \otimes \left|j\right\rangle\left\langle j\right|_{A^{0\prime}} \otimes \phi^{xj}_{A^0}.$$

Thus

$$\sigma_{XA^0B^0} = \sum_{x,j} p(x,j)\left|x,j\right\rangle\left\langle x,j\right|_X \otimes \sigma^{xj}_{A^0B^0},$$

where

$$\sigma^{xj}_{A^0B^0} = \Psi^0_{A^{0\prime}\to B^0}\left(\left|j\right\rangle\left\langle j\right|_{A^{0\prime}} \otimes \phi^{xj}_{A^0}\right)$$

is now a product state with respect to A^0 and B^0.

The three entropic quantities of interest can be simplied when evaluated with respect to $\sigma_{XA^0B^0}$, as

$$\begin{aligned}
I\left(A^0X;B^0\right)_\sigma &= S\left(B^0\right)_\sigma - \sum_{x,j} p(x,j)S\left(B^0\right)_{\sigma^{xj}} \leq C_C\left(\Psi^0\right), \\
I\left(A^0\rangle B^0X\right)_\sigma &= -\sum_{x,j} p(x,j)S\left(A^0\right)_{\sigma^{xj}} \leq 0, \\
I\left(X;B^0\right)_\sigma &\leq C_C\left(\Psi^0\right).
\end{aligned}$$

It's also clear that those inequalities can be achieved. Thus $C^{(1)}_{CQE}\left(\Psi^0\right)$ is described by

$$\begin{aligned}
C + 2Q &\leq C_C\left(\Psi^0\right), \\
Q + E &\leq 0, \\
C + Q + E &\leq C_C\left(\Psi^0\right).
\end{aligned}$$

Since the classical capacity of a classical channel is additive, the dynamic capacity region of Ψ^0 is additive and is described by the same set of inequalities.

Next we show that the dynamic capacity region is additive for Ψ^0 and Ψ^1, with Ψ^1 arbitrary.

Since $\Psi^0_{A^{0\prime}\to B^0} \otimes \Psi^1_{A^{1\prime}\to B^1}$ is a partial cq channel, its 1-shot dynamic capacity region $C^{(1)}_{CQE}\left(\Psi^0\otimes\Psi^1\right)$ can be achieved with respect to cq states $\sigma_{XAB^0B^1} = \Psi^0_{A^{0\prime}\to B^0} \otimes \Psi^1_{A^{1\prime}\to B^1}\left(\rho_{XAA^{0\prime}A^{1\prime}}\right)$, where $\rho_{XAA^{0\prime}A^{1\prime}}$ is of the form

$$\rho_{XAA^{0\prime}A^{1\prime}} = \sum_{x,j} p(x,j)\left|x,j\right\rangle\left\langle x,j\right|_X \otimes \left|j\right\rangle\left\langle j\right|_{A^{0\prime}} \otimes \phi^{xj}_{AA^{1\prime}}.$$

For $\rho_{XAA^{0\prime}A^{1\prime}}$ of this form, $\sigma_{XAB^0B^1}$ is of the form

$$\sigma_{XAB^0B^1} = \sum_{x,j} p(x,j)\left|x,j\right\rangle\left\langle x,j\right|_X \otimes \sigma^{xj}_{AB^0B^1},$$

with

$$\sigma^{xj}_{AB^0B^1} = \Psi^0_{A^{0\prime}\to B^0} \otimes \Psi^1_{A^{1\prime}\to B^1}\left(\left|j\right\rangle\left\langle j\right|_{A^{0\prime}} \otimes \phi^{xj}_{AA^{1\prime}}\right).$$

For such $\sigma_{XAB^0B^1}$, each of the three entropic quantities have simple upper bounds,

$$\begin{aligned}
I\left(AX;B^0B^1\right)_\sigma &\leq I\left(X;B^0\right)_\sigma + I\left(AX;B^1\right)_\sigma, \\
I\left(A\rangle B^0B^1X\right)_\sigma &= I\left(A\rangle B^1X\right)_\sigma, \\
I\left(X;B^0B^1\right)_\sigma &\leq I\left(X;B^0\right)_\sigma + I\left(X;B^1\right)_\sigma,
\end{aligned}$$

where we've used subadditivity of the von Neumann entropy. Thus the 1-shot dynamic capacity region of $\Psi^0\otimes\Psi^1$ has a simple upper bound

$$C^{(1)}_{CQE}\left(\Psi^0\otimes\Psi^1\right) \subseteq C^{(1)}_{CQE}\left(\Psi^0\right) + C^{(1)}_{CQE}\left(\Psi^1\right).$$

It's trivial to extend it to the dynamic capacity region of $\Psi^0\otimes\Psi^1$

$$C_{CQE}\left(\Psi^0\otimes\Psi^1\right) \subseteq C_{CQE}\left(\Psi^0\right) + C_{CQE}\left(\Psi^1\right).$$

Since the other direction of inclusion is obvious, we have

$$C_{CQE}\left(\Psi^0\otimes\Psi^1\right) = C_{CQE}\left(\Psi^0\right) + C_{CQE}\left(\Psi^1\right).$$

For unital extensions of a classical channel, we observe that, if the Heisenberg-Weyl operators are defined on the standard basis for the output of the channel, then the resulting channel is also a classical channel. Hence the above result applies. ■

Appendix C
Convex hull

Here we show that[6]

$$\begin{aligned}&\overline{\mathrm{Conv}\left(C_{CQE}\left(\mathcal{N}^0\right)\cup C_{CQE}\left(\mathcal{N}^1\right)\right)}\\=&\mathrm{Conv}\left(C_{CQE}\left(\mathcal{N}^0\right)\cup C_{CQE}\left(\mathcal{N}^1\right)\right).\end{aligned}$$

We quote a few properties about convex hull and Minkowski addition that we will use [22]: (i) For two closed sets A and B in $\mathbb{R}^k$, if A is bounded, then $A+B$ is closed. (ii) For two sets A and B in $\mathbb{R}^k$, $\mathrm{Conv}(A+B)=\mathrm{Conv}(A)+\mathrm{Conv}(B)$. (iii) The convex hull of a bounded set in $\mathbb{R}^k$ is also bounded.

First, we note that, by Ref. [4], all points in the 1-shot dynamic capacity region can be achieved by the classically enhanced father protocol, combined with unit protocols, *i.e.,*

$$C^{(1)}_{CQE}\left(\mathcal{N}\right)=\bigcup_{\sigma}C^{(1)}_{CQE,\mathrm{CEF}}\left(\mathcal{N}\right)_\sigma+C_{CQE,\mathrm{unit}},$$

where

$$C^{(1)}_{CQE,\mathrm{CEF}}\left(\mathcal{N}\right)_\sigma=\{I(X;B)_\sigma,\frac{1}{2}I(A;B|X)_\sigma,-\frac{1}{2}I(A;E|X)_\sigma\}$$

is the rate achieved using the classically enhanced father protocol, and σ is of the form in Eq. (7).

The unit protocols are teleportation, superdence-coding and entanglement distribution. $C_{CQE,\mathrm{unit}}$ are all the rates achieved by the unit protocols. They are described by [4]

$$\begin{aligned}C+Q+E&\le 0,\\Q+E&\le 0,\\C+2Q&\le 0.\end{aligned}$$

Clearly $C_{CQE,\mathrm{unit}}$ is convex and closed. Define

$$C^{(1)}_{CQE,\mathrm{CEF}}\left(\mathcal{N}\right)=\bigcup_{\sigma}C^{(1)}_{CQE,\mathrm{CEF}}\left(\mathcal{N}\right)_\sigma .$$

Clearly $C^{(1)}_{CQE,\mathrm{CEF}}\left(\mathcal{N}\right)$ is bounded by the input and output dimensions of $\mathcal{N}$.

Then

$$\begin{aligned}C_{CQE}\left(\mathcal{N}\right)&=\overline{\bigcup_{k=1}\frac{1}{k}\left(C^{(1)}_{CQE,\mathrm{CEF}}\left(\mathcal{N}^{\otimes k}\right)+C_{CQE,\mathrm{unit}}\right)}\\&=\overline{\bigcup_{k=1}\frac{1}{k}C^{(1)}_{CQE,\mathrm{CEF}}\left(\mathcal{N}^{\otimes k}\right)+C_{CQE,\mathrm{unit}}}.\end{aligned}$$

Denote

$$\begin{aligned}A&=\bigcup_{k=1}\frac{1}{k}C^{(1)}_{CQE,\mathrm{CEF}}\left(\mathcal{N}^{\otimes k}\right),\\B&=C_{CQE,\mathrm{unit}}.\end{aligned}$$

Since A is bounded, B is closed, by (i) and (iii), $\overline{A}+B$ is also closed.
Since $A+B\subseteq\overline{A}+B$, and $\overline{A}+B$ is closed, we have

$$\overline{A+B}\subseteq\overline{A}+B.$$

It is also obvious that

$$\overline{A+B}\supseteq\overline{A}+B,$$

hence

$$\overline{A+B}=\overline{A}+B.$$

Denote

$$C_{CQE,\mathrm{CEF}}\left(\mathcal{N}\right)=\overline{\bigcup_{k=1}\frac{1}{k}C^{(1)}_{CQE,\mathrm{CEF}}\left(\mathcal{N}^{\otimes k}\right)}.$$

Then by the above arguments,

$$C_{CQE}\left(\mathcal{N}\right)=C_{CQE,\mathrm{CEF}}\left(\mathcal{N}\right)+C_{CQE,\mathrm{unit}}.$$

Now we apply the above result to $\mathcal{N}^0$ and $\mathcal{N}^1$.

$$\begin{aligned}&\mathrm{Conv}\left(C_{CQE}\left(\mathcal{N}^0\right)\cup C_{CQE}\left(\mathcal{N}^1\right)\right)\\=&\mathrm{Conv}\left(C_{CQE,\mathrm{CEF}}\left(\mathcal{N}^0\right)\cup C_{CQE,\mathrm{CEF}}\left(\mathcal{N}^1\right)+C_{CQE,\mathrm{unit}}\right)\\=&\mathrm{Conv}\left(C_{CQE,\mathrm{CEF}}\left(\mathcal{N}^0\right)\cup C_{CQE,\mathrm{CEF}}\left(\mathcal{N}^1\right)\right)+C_{CQE,\mathrm{unit}}.\end{aligned}$$

In the last line, we used (ii).

Since $C_{CQE,\mathrm{CEF}}\left(\mathcal{N}\right)$ is closed and bounded for any finite dimensional quantum channel $\mathcal{N}$, the same must be true for $C_{CQE,\mathrm{CEF}}\left(\mathcal{N}^0\right)\cup C_{CQE,\mathrm{CEF}}\left(\mathcal{N}^1\right)$. Hence $\mathrm{Conv}\left(C_{CQE,\mathrm{CEF}}\left(\mathcal{N}^0\right)\cup C_{CQE,\mathrm{CEF}}\left(\mathcal{N}^1\right)\right)$ is closed and bounded. Thus $\mathrm{Conv}\left(C_{CQE,\mathrm{CEF}}\left(\mathcal{N}^0\right)\cup C_{CQE,\mathrm{CEF}}\left(\mathcal{N}^1\right)\right)+C_{CQE,\mathrm{unit}}$ is also closed.

Appendix D
Unital extension of the qubit dephasing channel and $1\rightarrow N$ cloning channel

Lemma 16: The CE and CQ trade-off curve of the qubit dephasing channel and $1\rightarrow N$ cloning channels are unchanged after a unital extension.

Proof. Consider the qubit dephasing channel Ψ^{dph}_η and a $1\rightarrow N$ cloning channel $\Psi^{1\rightarrow N}$, and their unital extensions Φ^{dph}_η and $\Phi^{1\rightarrow N}$. The statement of this lemma is equivalent to showing that

$$\begin{aligned}f_\lambda\left(\Psi\right)&=f_\lambda\left(\Phi\right)\quad\forall\lambda\ge 1,\\g_\lambda\left(\Psi\right)&=g_\lambda\left(\Phi\right)\quad\forall 0\le\lambda<1.\end{aligned}$$

for

$$\left(\Psi,\Phi\right)=\left(\Psi^{\mathrm{dph}}_\eta,\Phi^{\mathrm{dph}}_\eta\right),\left(\Psi^{1\rightarrow N},\Phi^{1\rightarrow N}\right).$$

In Lemma 13, we have argued that the 1-shot dynamic capacity region of a unitally extended channel can be achieved with input of the form in Eq. (20). Evaluating $f_\lambda\left(\Phi\right)$ on such states, one obtains

$$\begin{aligned}f_\lambda\left(\Phi\right)&=\log(|B|)+(\lambda-1)S\left(B|X'\right)_\sigma-\lambda S\left(AB|X'\right)_\sigma\\&=\log(|B|)+\sum_{x,j,k}p(x,j,k)\left[(\lambda-1)S(B)_{\sigma^{xjk}}-\lambda S(AB)_{\sigma^{xjk}}\right]\\&=\log(|B|)+\sum_{x,j}p(x,j)\left[(\lambda-1)S(B)_{\sigma^{xj}}-\lambda S(AB)_{\sigma^{xj}}\right]\\&\le\log(|B|)+\max_\sigma\left[(\lambda-1)S\left(B\right)_\sigma-\lambda S\left(AB\right)_\sigma\right],\end{aligned}\tag{46}$$

where

$$\sigma_{AB}=\Psi_{C\rightarrow B}\left(\phi_{AC}\right).\tag{47}$$

[6] $\mathcal{N}^0$ and $\mathcal{N}^1$ are assumed to be finite-dimensional.

For such a $\sigma_{AB} = \Psi(\phi_{AC})$ that achieves Eq. (46), one can construct

$$\rho_{XAA'} = \frac{1}{|R|} |k\rangle \langle k|_X \otimes |k\rangle \langle k|_R \otimes \phi_{AC}. \tag{48}$$

This state will saturate the above inequality.

For Ψ_η^{dph} and $\Psi^{1\to N}$, it can be verified [24] that their f_λ have the same form, *i.e.,*

$$f_\lambda(\Psi) = \log(|B|) + \max_\sigma [(\lambda - 1)S(B)_\sigma - \lambda S(AB)_\sigma], \tag{49}$$

with σ of the form given in Eq. (47).

The same argument also applies to g_λ. ■

The CQ trade-off curve of the qubit dephasing channel was computed in Ref. [23], and the CE trade-off curve was computed in Ref. [8]. The CE and CQ trade-off curves of the $1 \to N$ cloning channel were given in Ref. [24]. Other than the special cases ($\eta = 0, 1/2$ for the dephasing channel, $N = 1$ for the $1 \to N$ cloning channel), it can be verified that their CE and CQ trade-off curves are strictly concave at every point. By Lemma 16, this property is true for their unital extensions.

Appendix E
Proof of Equation (32)

In Ref. [8], it was shown that

$$C_C^{(1)}(\mathcal{N}) = \max_\sigma I(X;B)_\sigma + I(A\rangle BX)_\sigma,$$

where σ is of the form in Eq. (7).

It was also shown that the 1-shot CE trade-off capacity region is described by the set of all $C, E \geq 0$, such that

$$\begin{aligned} C &\leq I(AX;B)_\sigma \\ C &\leq I(X;B)_\sigma + I(A\rangle BX)_\sigma + |E| \end{aligned} \tag{50}$$

In the language of C_P, Eq. (50) means

$$C_P^{(1)}(\mathcal{N}) \leq C_C^{(1)}(\mathcal{N}) + P.$$

References

[1] C. E. Shannon, “A mathematical theory of communication,” *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, July 1948.

[2] I. Devetak, A. W. Harrow, and A. Winter, “A family of quantum protocols,” *Phys. Rev. Lett.*, vol. 93, p. 230504, Dec 2004.

[3] I. Devetak, A. W. Harrow, and A. J. Winter, “A resource framework for quantum Shannon theory,” *IEEE Transactions on Information Theory*, vol. 54, no. 10, pp. 4587–4618, Oct 2008.

[4] M.-H. Hsieh and M. M. Wilde, “Trading classical communication, quantum communication, and entanglement in quantum Shannon theory,” *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4705–4730, Sept 2010.

[5] C. H. Bennett and S. J. Wiesner, “Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states,” *Phys. Rev. Lett.*, vol. 69, pp. 2881–2884, Nov 1992.

[6] 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, pp. 3081–3084, Oct 1999.

[7] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, “Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem,” *IEEE Transactions on Information Theory*, vol. 48, no. 10, pp. 2637–2655, Oct 2002.

[8] M.-H. Hsieh and M. M. Wilde, “Entanglement-assisted communication of classical and quantum information,” *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4682–4704, Sept 2010.

[9] N. Datta and M. H. Hsieh, “One-shot entanglement-assisted quantum and classical communication,” *IEEE Transactions on Information Theory*, vol. 59, no. 3, pp. 1929–1939, March 2013.

[10] A. S. Holevo, “The capacity of the quantum channel with general signal states,” *IEEE Transactions on Information Theory*, vol. 44, no. 1, pp. 269–273, Jan 1998.

[11] B. Schumacher and M. D. Westmoreland, “Sending classical information via noisy quantum channels,” *Phys. Rev. A*, vol. 56, pp. 131–138, Jul 1997.

[12] S. Lloyd, “Capacity of the noisy quantum channel,” *Phys. Rev. A*, vol. 55, pp. 1613–1622, Mar 1997.

[13] P. W. Shor, “The quantum channel capacity and coherent information,” *MSRI Workshop on Quantum Computation*, 2002.

[14] I. Devetak, “The private classical capacity and quantum capacity of a quantum channel,” *IEEE Transactions on Information Theory*, vol. 51, no. 1, pp. 44–55, Jan 2005.

[15] G. Smith and J. Yard, “Quantum communication with zero-capacity channels,” *Science*, vol. 321, no. 5897, pp. 1812–1815, 2008.

[16] P. W. Shor, “The classical capacity achievable by a quantum channel assisted by a limited entanglement.” *Quantum Information and Computation*, vol. 4, no. 6, pp. 537 – 545, 2004.

[17] G. Smith, J. A. Smolin, and A. Winter, “The quantum capacity with symmetric side channels,” *IEEE Transactions on Information Theory*, vol. 54, no. 9, pp. 4208–4217, Sept 2008.

[18] M. B. Hastings, “Superadditivity of communication capacity using entangled inputs,” *Nat Phys*, vol. 5, no. 4, pp. 255–257, 04 2009.

[19] P. Hayden and A. Winter, “Counterexamples to the maximal p-norm multiplicativity conjecture for all p > 1,” *Communications in Mathematical Physics*, vol. 284, no. 1, pp. 263–280, 2008.

[20] P. W. Shor, “Equivalence of additivity questions in quantum information theory,” *Communications in Mathematical Physics*, vol. 246, no. 3, pp. 453–472, 2004.

[21] E. Y. Zhu, Q. Zhuang, and P. W. Shor, “Superadditivity of the classical capacity with limited entanglement assistance,” *Phys. Rev. Lett.*, vol. 119, p. 040503, Jul 2017.

[22] R. Schneider, *Convex Bodies: The Brunn-Minkowski Theory (Encyclopedia of Mathematics and its Applications)*. Cambridge University Press, 1993.

[23] I. Devetak and P. W. Shor, “The capacity of a quantum channel for simultaneous transmission of classical and quantum information,” *Communications in Mathematical Physics*, vol. 256, no. 2, pp. 287–303, 2005.

[24] K. Brádler, P. Hayden, D. Touchette, and M. M. Wilde, “Trade-off capacities of the quantum Hadamard channels,” *Phys. Rev. A*, vol. 81, p. 062312, Jun 2010.

[25] C. King, “Additivity for unital qubit channels,” *Journal of Mathematical Physics*, vol. 43, no. 10, pp. 4641–4653, 2002.

[26] D. P. DiVincenzo, P. W. Shor, and J. A. Smolin, “Quantum-channel capacity of very noisy channels,” *Phys. Rev. A*, vol. 57, pp. 830–839, Feb 1998.

[27] M. Fukuda, “Simplification of additivity conjecture in quantum information theory,” *Quantum Information Processing*, vol. 6, no. 3, pp. 179–186, 2007.

[28] T. Cubitt, D. Elkouss, W. Matthews, M. Ozols, D. Pérez-García, and S. Strelchuk, “Unbounded number of channel uses may be required to detect quantum capacity,” *Nature Communications*, vol. 6, pp. 6739 EP –, Mar 2015, article.

[29] D. Elkouss and S. Strelchuk, “Superadditivity of private information for any number of uses of the channel,” *Phys. Rev. Lett.*, vol. 115, p. 040501, Jul 2015.

[30] W. Rudin, *Principles of Mathematical Analysis*. McGraw-Hill Education, 1976.

[31] M. M. Wilde and M.-H. Hsieh, “The quantum dynamic capacity formula of a quantum channel,” *Quantum Information Processing*, vol. 11, no. 6, pp. 1431–1463, 2012.