# Trade-off capacities of the quantum Hadamard channels

Kamil Brádler, Patrick Hayden, Dave Touchette, and Mark M. Wilde
*School of Computer Science, McGill University, Montreal, Québec, H3A 2A7, Canada*
(Received 12 April 2010; published 14 June 2010)

Coding theorems in quantum Shannon theory express the ultimate rates at which a sender can transmit information over a noisy quantum channel. More often than not, the known formulas expressing these transmission rates are intractable, requiring an optimization over an infinite number of uses of the channel. Researchers have rarely found quantum channels with a tractable classical or quantum capacity, but when such a finding occurs, it demonstrates a complete understanding of that channel's capabilities for transmitting classical or quantum information. Here we show that the three-dimensional capacity region for entanglement-assisted transmission of classical and quantum information is tractable for the Hadamard class of channels. Examples of Hadamard channels include generalized dephasing channels, cloning channels, and the Unruh channel. The generalized dephasing channels and the cloning channels are natural processes that occur in quantum systems through the loss of quantum coherence or stimulated emission, respectively. The Unruh channel is a noisy process that occurs in relativistic quantum information theory as a result of the Unruh effect and bears a strong relationship to the cloning channels. We give exact formulas for the entanglement-assisted classical and quantum communication capacity regions of these channels. The coding strategy for each of these examples is superior to a naïve time-sharing strategy, and we introduce a measure to determine this improvement.

DOI: 10.1103/PhysRevA.81.062312 PACS number(s): 03.67.Hk, 03.67.Pp, 04.62.+v

## I. INTRODUCTION

One of the aims of quantum Shannon theory is to characterize the ultimate limits on the transmission of information over a noisy quantum channel. Holevo, Schumacher, and Westmoreland contributed the first seminal result in this direction by providing a lower bound for the ultimate limit of a noisy quantum channel to transmit classical information [1,2], a result now known as the HSW coding theorem. Lloyd, Shor, and Devetak then contributed increasingly rigorous proofs of the quantum channel coding theorem [3–5], now known as the LSD coding theorem, that provides a lower bound on the ultimate limit for a noisy quantum channel to transmit quantum information. Other expository proofs appeared later, providing insight into the nature of quantum coding [6–10]. Bennett *et al.* [11] and Barnum *et al.* [12] also showed that the capacity of a quantum channel for transmitting quantum information is the same as that channel's capacity for generating shared entanglement between sender and receiver. These three results form the core of the dynamic, single-resource quantum Shannon theory, where a sender exploits a noisy quantum channel to establish a single noiseless resource, namely, classical communication, quantum communication, or shared entanglement, with a receiver.

A formula for the capacity of a channel gives a "single-letter" characterization if the computation of the capacity requires an optimization over only a single use of the channel, and the formula gives a "multiletter" characterization otherwise. A single-letter characterization implies that the computation of the capacity is tractable for a fixed input dimension of the channel, whereas a multiletter characterization typically requires an optimization over an infinite number of uses of the channel and is therefore intractable in this case. This "single-letterization" issue does not play a central role in classical information theory for the most basic task of communication over a noisy classical channel because single-letterization occurs naturally in Shannon's original analysis for classical, memoryless channels [13]. But this issue plays a prominent role in the domain of quantum Shannon theory even for the most basic communication tasks. Our knowledge so far indicates that the computation of the classical capacity is intractable in the general case [14–17], with the same seeming to hold generally for the quantum capacity [18,19]. These results underscore our incomplete understanding of the nature of quantum information, but they also have the surprising and "uniquely quantum" respective consequences that the strong correlations present in entangled uses of a quantum channel boost the classical capacity and that the degeneracy property of quantum codes can enhance the quantum capacity.

Thus, in hindsight, we might now say that any channel with a single-letter capacity formula is a "rare gem" in quantum Shannon theory given that the known formulas for capacities generally give multiletter characterizations. In fact, researchers once conjectured that the HSW formula for the classical capacity would generally give a single-letter characterization [20,21], until the recent result of Hastings [14]. Researchers have found several examples of these gems for the classical capacity: the identity channel [22], unital qubit channels [23], erasure channels [24], Hadamard channels [25], entanglement-breaking channels [26], depolarizing channels [27], transpose-depolarizing channels [28], shifted depolarizing channels [29], cloning channels [30], and the so-called "Unruh" channel [30]. Researchers have found fewer such exemplary single-letter gems for the quantum capacity: erasure channels [24], degradable channels [31,32], conjugate degradable channels [33], and amplitude damping channels [34]. We do not yet have a general method for determining whether a given channel's capacity admits a single-letter characterization using known formulas—the techniques for proving single-letterization of the above examples are all *ad hoc*, varying from case to case. Additionally, we can observe that it is an even rarer gem for a channel to admit a single-letter characterization for both the classical and

quantum capacity. This "single-letter overlap" occurs for erasure channels [24], generalized dephasing channels [31], cloning channels [30,33], and the Unruh channel [30,33,35], but the reasons for single-letterization of the classical and quantum capacities have no obvious connection.

After the early work in quantum Shannon theory, several researchers considered how different noiseless resources such as entanglement, classical communication, or quantum communication might trade off against one another together with a noisy channel. The first findings in this direction were those of Bennett *et al.* [36,37], who showed that unlimited, shared, noiseless entanglement can boost the classical capacity of a noisy quantum channel, generalizing the superdense coding effect [38]. Perhaps even more surprising was that the formula for classical capacity with unlimited entanglement assistance admits a single-letter characterization, marking the first time that we could say there is a problem in quantum Shannon theory that we truly understand. Shor then refined this result by considering the classical capacity of a channel assisted by a finite amount of shared entanglement [39]. He calculated a trade-off curve that determines how a sender can optimally trade the consumption of noiseless entanglement with the generation of noiseless classical communication. This trade-off curve also bounds a rate region consisting of rates of entanglement consumption and generated classical communication. Unfortunately, the formulas for the rate region do not give a single-letter characterization in the general case.

Shor's result then inspired Devetak and Shor to consider a scenario where a sender exploits a noisy quantum channel to simultaneously transmit both noiseless classical and quantum information [31], a scenario later dubbed "classically enhanced quantum coding" [40,41] after schemes formulated in the theory of quantum error correction [42,43]. Devetak and Shor provided a multiletter characterization of the classically enhanced quantum capacity region for general channels, but were able to show that both generalized dephasing channels and erasure channels admit single-letter capacity regions. We must emphasize that single-letterization of both the classical and quantum capacities of a noisy quantum channel does not immediately imply the single letterization of the classically enhanced quantum capacity region (though the latter does imply the former). That is, the proof that the region single-letterizes requires a different technique that has no obvious connection to the techniques used to prove the single letterization of the individual capacities. The additional benefit of the Devetak-Shor classically enhanced quantum coding scheme is that it beats a time-sharing strategy[1] for some channels.

We might say that the previous scenarios are a part of the dynamic, double-resource quantum Shannon theory, where a sender can exploit a noisy quantum channel to generate two noiseless resources, or a sender can exploit a noisy quantum channel in addition to a noiseless resource to generate another noiseless resource. This theory culminated with the landmark work of Devetak *et al.* that provided a multiletter characterization for virtually every permutation of two resources and a noisy quantum channel which one can consider [44,45], but they neglected to search for channels with single-letter characterizations of the double-resource capacity regions. Other researchers concurrently considered how noiseless resources might trade off against each other in tasks outside of the dynamic, double-resource quantum Shannon theory, such as quantum compression [46–48], remote state preparation [49,50], and hybrid quantum memories [51].

The next natural step in this line of inquiry was to consider the dynamic, triple-resource quantum Shannon theory. Hsieh and Wilde did so by providing a multiletter characterization of an entanglement-assisted quantum channel's ability to transmit both classical and quantum information [40,41,52]. In addition, they found that the formulas for a generalized dephasing channel, an erasure channel, and the trivial completely depolarizing channel all give a single-letter triple trade-off capacity region (though they omitted the full proof for the generalized dephasing channels). Thus, these findings provided a complete understanding of the ultimate performance limits of any scheme for entanglement-assisted classical and quantum error correction [42,43,53–56], at least for the aforementioned channels. The authors of Ref. [40] also constructed a new protocol, dubbed the "classically enhanced father protocol," that outperforms a time-sharing strategy for transmitting both classical and quantum information over an entanglement-assisted quantum channel.

In this paper we contribute a class of "single-letter gems" to the dynamic, triple-resource quantum Shannon theory. We provide single-letter formulas for the triple trade-off capacity region of the Hadamard class of channels. We then compute and plot examples of the triple trade-off regions for the generalized dephasing channels, the cloning channels, and the Unruh channel, all of which are members of the class of Hadamard channels. The generalized dephasing channel represents a natural mechanism for decoherence in physical systems such as superconducting qubits [57], the cloning channel represents a natural process that occurs during stimulated emission [58–60], and the Unruh channel arises in relativistic quantum information theory [30,33,35], bearing connections to the process of black-hole stimulated emission [61]. The proof technique to determine the formulas for the cloning channel extends naturally to the formulas for the Unruh channel by exploiting the insights of Brádler in Ref. [30]. We also find that the coding strategy for each of these channels beats a simple time-sharing strategy, and we introduce a measure to compute the amount by which it beats a time-sharing strategy.

We structure this work as follows. Section II reviews the definitions of noisy quantum channels, classical-quantum states, information-theoretic quantities, and examples of noisy quantum channels that are relevant for our purposes here. Section III then reviews the capacity regions mentioned previously, specifically, the classically enhanced quantum (CQ) capacity region, the entanglement-assisted classical (CE) capacity region, and the entanglement-assisted classical and quantum (CQE) capacity region—we abbreviate a capacity region by the noiseless resources involved: classical communication (C), quantum communication (Q), or entanglement (E). We then present our main result in Sec. IV: the proof that the CQ and CE capacity regions of Hadamard channels

[1]Time sharing in this case is a simple strategy that exploits an HSW code for some fraction of the channel uses and an LSD code for the other fraction of the channel uses.

admit a single-letter characterization. As first observed by Hsieh and Wilde [40], and perhaps surprisingly, the single letterization of these two regions immediately implies the single letterization of the CQE capacity region. Section V computes the CQE capacity region of the qubit dephasing channel, the $1 \to N$ cloning channel, and the Unruh channel, and the next section plots the CQE regions. The final section measures the improvement of the optimal protocols over a time-sharing strategy. Finally, we conclude with some remaining observations and suggestions for future work.

## II. DEFINITIONS AND NOTATION

### A. Quantum channels and classical-quantum states

We first review the notion of a noisy quantum channel, an isometric extension, and a classical-quantum state.

A noisy quantum channel $\mathcal{N}$ is a completely positive trace-preserving (CPTP) convex-linear map. It admits a Kraus representation [62] so that its action on a density operator $\sigma$ is as follows

$$\mathcal{N}(\sigma) = \sum_i N_i \sigma N_i^\dagger, \tag{1}$$

where the Kraus operators form a resolution of the identity $\sum_i N_i^\dagger N_i = I$, ensuring that the map is trace preserving. The notion of an isometric extension of a noisy quantum channel proves to be useful in quantum Shannon theory—the notion is similar to that of a purification of a density operator. Let $U_{\mathcal{N}}^{A' \to BE}$ denote an isometric extension of the noisy map $\mathcal{N}$, defined such that $U_{\mathcal{N}}^\dagger U_{\mathcal{N}} = I$ and

$$\mathcal{N}^{A' \to B}(\sigma) = \mathrm{Tr}_E\{U_{\mathcal{N}} \sigma U_{\mathcal{N}}^\dagger\}.$$

The Kraus operators provide a straightforward method for constructing an isometric extension

$$U_{\mathcal{N}}^{A' \to BE} = \sum_i N_i^{A' \to B} \otimes |i\rangle^E,$$

where the set $\{|i\rangle^E\}$ is an orthonormal set of states. The following relation gives the conjugation of a density operator $\sigma$ by the isometry $U_{\mathcal{N}}$

$$U_{\mathcal{N}} \sigma U_{\mathcal{N}}^\dagger = \sum_{i,j} (N_i \sigma N_j^\dagger)^B \otimes |i\rangle\langle j|^E. \tag{2}$$

Let $\mathcal{N}^c$ denote a complementary channel of $\mathcal{N}$, unique up to isometries on the system $E$, whose action on $\sigma$ is

$$\mathcal{N}^c(\sigma) \equiv \mathrm{Tr}_B\{U_{\mathcal{N}} \sigma U_{\mathcal{N}}^\dagger\}.$$

Observe that the following channel is a valid complementary channel for the channel in (1)

$$\mathcal{N}^c(\sigma) = \sum_{i,j} \mathrm{Tr}\{N_i \sigma N_j^\dagger\} |i\rangle\langle j|^E.$$

A quantum channel is degradable [31,32], a notion imported from classical information theory [63], if there exists a degrading channel $\mathcal{D}^{B \to E}$ that simulates the action of the complementary channel $(\mathcal{N}^c)^{A' \to E}$ so that

$$\forall \sigma \mathcal{D}^{B \to E} \circ \mathcal{N}^{A' \to B}(\sigma) = (\mathcal{N}^c)^{A' \to E}(\sigma).$$

Suppose that Alice possesses an ensemble $\{[p_X(x), \rho_x^{A'}]\}$ of quantum states where $p_X(x)$ is the probability density function for a random variable $X$ and $\rho_x^{A'}$ is a density operator conditional on the realization $x$ of random variable $X$. She can augment this ensemble by correlating a classical variable with each $\rho_x^{A'}$. This procedure produces an augmented ensemble $\{[p_X(x), |x\rangle\langle x|^X \otimes \rho_x^{A'}]\}$, where the states $\{|x\rangle^X\}$ form an orthonormal set. Taking the expectation over the augmented ensemble gives the following classical-quantum state

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

Let $|\phi_x\rangle^{AA'}$ denote purifications of each $\rho_x^{A'}$. The following state is also a classical-quantum state

$$\rho^{XAA'} \equiv \sum_x p_X(x) |x\rangle\langle x|^X \otimes |\phi_x\rangle\langle\phi_x|^{AA'}. \tag{3}$$

Suppose that Alice transmits the $A'$ subsystem through a noisy quantum channel $\mathcal{N}^{A' \to B}$. The state output from the channel is $\rho^{XAB}$ where

$$\rho^{XAB} \equiv \mathcal{N}^{A' \to B}(\rho^{XAA'}) \tag{4}$$

$$= \textstyle\sum_x p_X(x) |x\rangle\langle x|^X \otimes \mathcal{N}^{A' \to B}(|\phi_x\rangle\langle\phi_x|^{AA'}). \tag{5}$$

It is implicit that an identity acts on a system for which there is no label on the CPTP map $\mathcal{N}$. Note that the registers $XAB$ also form a classical-quantum system. For such a multiparty state, we adopt the convention that a state with a superscript unambiguously identifies a density operator. For example, we define the reduced density operator $\rho^A \equiv \mathrm{Tr}_{XB}\{\rho^{XAB}\}$. We define states $|\phi_x\rangle^{ABE}$ as

$$|\phi_x\rangle^{ABE} \equiv U_{\mathcal{N}} |\phi_x\rangle^{AA'}.$$

Observe that the states $|\phi_x\rangle^{ABE}$ purify each $\rho_x^{AB}$. All of the above classical-quantum states are important throughout this paper.

### B. Information-theoretic quantities

We now define some standard information theoretic quantities. The von Neumann entropy of a quantum state $\rho$ is defined as

$$H(\rho) \equiv -\mathrm{Tr}\{\rho \log_2 \rho\}.$$

We write $H(A)_\rho \equiv H(\rho^A)$ for a subsystem $A$ of $\rho$, omitting the subscript if $\rho$ is implicitly clear. Note that the von Neumann entropy is equal to the Shannon entropy for a classical system $X$

$$H(X) \equiv -\sum_x p_X(x) \log_2[p_X(x)].$$

We define the conditional entropy $H(A|B)$, the mutual information $I(A;B)$, and the coherent information $I(A\rangle B)$ as follows for a bipartite state $\rho^{AB}$

$$H(A|B) \equiv H(AB) - H(B),$$
$$I(A;B) \equiv H(A) - H(A|B),$$
$$I(A\rangle B) \equiv -H(A|B).$$

The Holevo quantity $\chi$ for the state $\rho^{XB}$ of a classical-quantum system $XB$ is

$$\chi\left(\left\{p_X(x),\rho_x^B\right\}\right) \equiv I(X;B)_\rho.$$

The following relation holds by the joint entropy theorem [62] when the conditioning system in a conditional entropy is classical

$$H(B|X) = \sum_x p_X(x) H(B)_{\rho_x^B}.$$

### C. Quantum channels

#### 1. Entanglement-breaking channels

A noisy quantum channel $\mathcal{N}$ is entanglement breaking if it outputs a separable state whenever half of any entangled state is the input to the channel [64]. By the methods in Ref. [64], a channel is entanglement breaking if it produces a separable state when its input is half of a maximally entangled state. More precisely, suppose that $|\Phi\rangle^{AA'}$ is the maximally entangled qudit state

$$|\Phi\rangle^{AA'} \equiv \frac{1}{\sqrt{D}} \sum_{i=0}^{D-1} |i\rangle^A |i\rangle^{A'}. \tag{6}$$

A channel $\mathcal{N}^{A'\to B}$ is entanglement breaking if and only if

$$\mathcal{N}^{A'\to B}(\Phi^{AA'}) = \sum_x p_X(x) \rho_x^A \otimes \sigma_x^B,$$

where $p_X(x)$ is an arbitrary probability distribution and $\rho_x^A$ and $\sigma_x^B$ are arbitrary density operators.

We can also think of an entanglement-breaking channel as a quantum-classical-quantum channel [64], in the sense that it first applies a noisy channel $\mathcal{N}_1$, it performs a complete von Neumann measurement of the resulting state, and it finally applies another noisy channel $\mathcal{N}_2$. Consequently, we can write the action of an entanglement-breaking channel $\mathcal{N}_{\text{EB}}$ as follows:

$$\mathcal{N}_{\text{EB}}(\rho) = \sum_x \text{Tr}\{\Lambda_x \rho\} \sigma_x, \tag{7}$$

where $\{\Lambda_x\}$ is a positive-operator-valued measurement (POVM) and the states $\sigma_x$ are arbitrary. Additionally, any entanglement-breaking channel admits a Kraus representation whose Kraus operators are unit rank: $N_i = |\xi_i\rangle^B \langle\varsigma_i|^{A'}$ [64]. Note that the sets $\{|\xi_i\rangle^B\}$ and $\{|\varsigma_i\rangle^{A'}\}$ each do not necessarily form an orthonormal set.

The classical capacity of an entanglement-breaking channel single-letterizes [26], and its quantum capacity vanishes because it destroys entanglement and thus cannot transmit quantum information.

#### 2. Hadamard channels

A Hadamard channel is a quantum channel whose complementary channel is entanglement breaking [25]. We can write its output as the Hadamard product (elementwise multiplication) of a representation of the input density operator with another operator.[2] To briefly review how this comes about, suppose that the complementary channel $(\mathcal{N}^c)^{A'\to E}$ of a channel $\mathcal{N}^{A'\to B}$ is entanglement breaking. Then, using the fact in the previous section that its Kraus operators $|\xi_i\rangle^E \langle\zeta_i|^{A'}$ are unit rank and the construction in (2) for an isometric extension, we can write an isometric extension $U_{\mathcal{N}^c}$ for $\mathcal{N}^c$ as

$$\begin{aligned} U_{\mathcal{N}^c} \sigma U_{\mathcal{N}^c}^\dagger &= \sum_{i,j} |\xi_i\rangle^E \langle\zeta_i|^{A'} \sigma |\zeta_j\rangle^{A'} \langle\xi_j|^E \otimes |i\rangle^B \langle j|^B \\ &= \sum_{i,j} \langle\zeta_i|^{A'} \sigma |\zeta_j\rangle^{A'} |\xi_i\rangle^E \langle\xi_j|^E \otimes |i\rangle^B \langle j|^B. \end{aligned} \tag{8}$$

The sets $\{|\xi_i\rangle^E\}$ and $\{|\zeta_i\rangle^{A'}\}$ each do not necessarily consist of orthonormal states, but the set $\{|i\rangle^B\}$ does. Tracing over the system $E$ gives the original channel from system $A'$ to $B$

$$\mathcal{N}_{\text{H}}^{A'\to B}(\sigma) = \sum_{i,j} \langle\zeta_i|^{A'} \sigma |\zeta_j\rangle^{A'} \langle\xi_j|\xi_i\rangle^E |i\rangle^B \langle j|^B. \tag{9}$$

Let $\Sigma$ denote the matrix with elements $[\Sigma]_{i,j} = \langle\zeta_i|^{A'} \sigma |\zeta_j\rangle^{A'}$, a representation of the input state $\sigma$, and let $\Gamma$ denote the matrix with elements $[\Gamma]_{i,j} = \langle\xi_j|\xi_i\rangle^E$. Then, from (9), it is clear that the output of the channel is the Hadamard product $*$ of $\Sigma$ and $\Gamma^\dagger$ with respect to the basis $\{|i\rangle^B\}$

$$\mathcal{N}_{\text{H}}^{A'\to B}(\sigma) = \Sigma * \Gamma^\dagger.$$

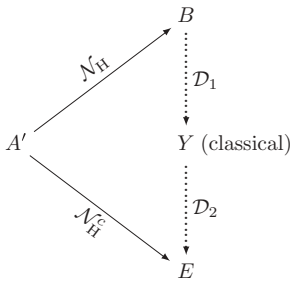
For this reason, such a channel is known as a Hadamard channel.

Hadamard channels are degradable. If Bob performs a von Neumann measurement of his state in the basis $\{|i\rangle^B\}$ and prepares the state $|\xi_i\rangle^E$ conditional on the outcome of the measurement, this procedure simulates the complementary channel $(\mathcal{N}^c)^{A'\to E}$ and also implies that the degrading map $\mathcal{D}^{B\to E}$ is entanglement breaking. To be more precise, the Kraus operators of the degrading map $\mathcal{D}^{B\to E}$ are $\{|\xi_i\rangle^E \langle i|^B\}$ so that

$$\begin{aligned} \mathcal{D}^{B\to E}\left[\mathcal{N}_{\text{H}}^{A'\to B}(\sigma)\right] &= \sum_i |\xi_i\rangle^E \langle i|^B \mathcal{N}^{A'\to B}(\sigma) |i\rangle^B \langle\xi_i|^E \\ &= \sum_i \langle i|^{A'} \sigma |i\rangle^{A'} |\xi_i\rangle\langle\xi_i|^E, \end{aligned}$$

demonstrating that this degrading map effectively simulates the complementary channel $(\mathcal{N}_{\text{H}}^c)^{A'\to E}$. Note that we can view this degrading map as the composition of two maps: A first map $\mathcal{D}_1^{B\to Y}$ performs the von Neumann measurement, leading to a classical variable $Y$, and a second map $\mathcal{D}_2^{Y\to E}$ performs the state preparation, conditional on the value of the classical variable $Y$. We can therefore write $\mathcal{D}^{B\to E} = \mathcal{D}_2^{Y\to E} \circ \mathcal{D}_1^{B\to Y}$. This observation is crucial to our proof of the single letterization of both the CQ and CE capacity regions of Hadamard channels. These structural relationships are summarized in the following commutative diagram:

[2] Strictly speaking, King *et al.* did not refer to such channels as "Hadamard channels," but rather as "channels of the Hadamard form." Here, we loosely refer to them simply as "Hadamard channels."




We show in the next two sections that generalized dephasing channels and cloning channels are both members of the Hadamard class because their complementary channels are entanglement-breaking.

### 3. *Generalized dephasing channels*

Generalized dephasing channels model physical processes where there is no loss of energy but there is a loss of quantum coherence [62], a type of quantum noise that dominates, for example, in superconducting qubits [57]. The respective input and output systems $A'$ and $B$ of such channels are of the same dimension. Let $\{|i\rangle^{A'}\}$ and $\{|i\rangle^{B}\}$ be some respective preferred orthonormal bases for these systems, the first of which we call the dephasing basis. The channel does not affect any state that is diagonal in the dephasing basis, but it mixes coherent superpositions of these basis states.

An isometric extension $U_{\mathcal{N}_{\rm GD}}^{A'\to BE}$ of a generalized dephasing channel $\mathcal{N}_{\rm GD}^{A'\to B}$ has the form

$$U_{\mathcal{N}_{\rm GD}}^{A'\to BE} \equiv \sum_i |i\rangle^B\langle i|^{A'} \otimes |\vartheta_i\rangle^E, \tag{10}$$

where the set $\{|\vartheta_i\rangle^E\}$ is not necessarily an orthonormal set. The output of a generalized dephasing channel is as follows:

$$\mathcal{N}_{\rm GD}(\sigma) = \sum_{i,j} \langle\vartheta_j|\vartheta_i\rangle^E \langle i|\sigma|j\rangle^{A'} |i\rangle\langle j|^B.$$

If the states $|\vartheta_i\rangle^E$ are orthonormal, the channel is a completely dephasing channel $\Delta^{A'\to B}$

$$\Delta^{A'\to B}(\sigma) \equiv \sum_i |i\rangle^B\langle i|^{A'}\sigma|i\rangle^{A'}\langle i|^B.$$

We obtain the complementary channel of a generalized dephasing channel by tracing over Bob's system $B$

$$\mathrm{Tr}_B\left\{U_{\mathcal{N}_{\rm GD}}\sigma U_{\mathcal{N}_{\rm GD}}^\dagger\right\} = \sum_i \langle i|^{A'}\sigma|i\rangle^{A'}|\vartheta_i\rangle\langle\vartheta_i|^E,$$

which we recognize to be an entanglement-breaking channel of the form in (7). Thus, a generalized dephasing channel is a Hadamard channel because its complementary channel is entanglement breaking. In fact, the isometric extension in (10) of the generalized dephasing channel appears remarkably similar to the isometric extension in (8) of the Hadamard channel, with the exception that the states $\{|i\rangle^{A'}\}$ of the generalized dephasing channel form an orthonormal basis.

A completely dephasing channel $\Delta$ commutes with a generalized dephasing channel $\mathcal{N}_{\rm GD}$ because

$$(\mathcal{N}_{\rm GD}\circ\Delta)(\sigma) = (\Delta\circ\mathcal{N}_{\rm GD})(\sigma) = \sum_i \langle i|\sigma|i\rangle^{A'}|i\rangle\langle i|^B.$$

The property $\mathcal{N}_{\rm GD}^c = \mathcal{N}_{\rm GD}^c\circ\Delta$ also holds for the complementary channel

$$\mathcal{N}_{\rm GD}^c(\sigma) = (\mathcal{N}_{\rm GD}^c\circ\Delta)(\sigma) = \sum_i \langle i|\,\rho\,|i\rangle^{A'}|\vartheta_i\rangle\langle\vartheta_i|^E.$$

The simplest example of a generalized dephasing channel is a qubit dephasing channel. The action of the $p$-dephasing qubit channel is

$$\mathcal{N}(\sigma) = (1-p)\sigma + p\Delta(\sigma), \tag{11}$$

where $\Delta$ in this case is

$$\Delta(\sigma) = \tfrac{1}{2}(\sigma + Z\sigma Z),$$

and $Z$ is the Pauli matrix $\sigma_Z$. Hence an isometric extension $U_{\mathcal{N}}^{A'\to BE}$ of the qubit dephasing channel has the form

$$U_{\mathcal{N}}^{A'\to BE} = \sqrt{1-\frac{p}{2}}\, I\otimes|0\rangle^E + \sqrt{\frac{p}{2}}\, Z\otimes|1\rangle^E,$$

where $I$ is the identity operator. Therefore, the following is a complementary channel $\mathcal{N}^c$ of a qubit dephasing channel

$$\mathcal{N}^c(\sigma) = \frac{p}{2}|0\rangle\langle 0|^E + \left(1-\frac{p}{2}\right)|1\rangle\langle 1|^E + \sqrt{\left(1-\frac{p}{2}\right)\frac{p}{2}}\mathrm{Tr}\{\sigma Z\}(|0\rangle\langle 1|^E + |1\rangle\langle 0|^E),$$

and we observe that a bit flip on the input state does not change the eigenvalues of the resulting environment output state.

### 4. *Cloning channels*

The no-cloning theorem forbids the cloning of arbitrary quantum states [65]. However, nothing prevents one from performing approximate cloning provided the fidelity of cloning is not too high. A universal $1\to N$ cloner is a device that approximately copies the input with maximal copy fidelity independent of the input state [66]. We refer to such a device as a cloning channel [30,33]. Such a decoherence process occurs naturally during stimulated emission [58–60].

We focus on $1\to N$ qubit cloning channels where a single qubit serves as the input, and the output is $N$ identical approximate copies on $N$ respective qubit systems. These universal cloners are unitarily covariant [67], in the sense that a unitary $V$ on the input qubit maps to an irreducible representation $R_V$ of $V$ on the output state

$$\mathcal{N}(V\rho V) = R_V\mathcal{N}(\rho)R_V^\dagger.$$

We present the action of an isometric extension of a $1\to N$ cloning channel on a basis $\{|0\rangle,|1\rangle\}$ for a qubit input system $A'$. Let $|j\rangle^B$ be an orthonormal basis of normalized completely symmetric states for the output system $B$ that consists of $N$ qubits

$$\{|j\rangle^B \equiv |N-j,j\rangle\}_{j=0}^N,$$

where $|N-j,j\rangle^B$ denotes a normalized state on an $N$-qubit system that is a uniform superposition of computational basis

states with $N-j$ "zeros" and $j$ "ones." Let $|i\rangle^E$ be an orthonormal basis for the environment $E$

$$\{|i\rangle^E \equiv |N-i-1,i\rangle\}_{i=0}^{N-1},$$

where $|N-i-1,i\rangle^E$ denotes a normalized state on an $(N-1)$-qubit system that is a uniform superposition of computational basis states with $N-i-1$ "zeros" and $i$ "ones." Then an isometric extension $U_{\mathcal{N}_{\mathrm{Cl}}}^{A'\to BE}$ of the $1\to N$ cloning channel $\mathcal{N}$ has the form

$$U_{\mathcal{N}_{\mathrm{Cl}}}^{A'\to BE} \equiv \frac{1}{\sqrt{\Delta_N}}\sum_{i=0}^{N-1}\sqrt{N-i}\,|i\rangle^B\langle 0|^{A'}\otimes|i\rangle^E + \frac{1}{\sqrt{\Delta_N}}\sum_{i=0}^{N-1}\sqrt{i+1}\,|i+1\rangle^B\langle 1|^{A'}\otimes|i\rangle^E, \quad (12)$$

where $\Delta_N \equiv N(N+1)/2$. A set of Kraus operators for the channel $\mathcal{N}_{\mathrm{Cl}}$ is as follows:

$$\left\{\frac{1}{\sqrt{\Delta_N}}(\sqrt{N-i}\,|i\rangle^B\langle 0|^{A'}+\sqrt{i+1}\,|i+1\rangle^B\langle 1|^{A'})\right\}_{i=0}^{N-1},$$

and a set of Kraus operators for the complementary channel $\mathcal{N}_{\mathrm{Cl}}^c$ is as follows:

$$\sqrt{N}|0\rangle^E\langle 0|^{A'},$$
$$\{\sqrt{N-i}|i\rangle^E\langle 0|^{A'}+\sqrt{i}|i-1\rangle^E\langle 1|^{A'}\}_{i=1}^{N-1},$$
$$\sqrt{N}|N-1\rangle^E\langle 1|^{A'}.$$

We can rewrite the Kraus operators for the complementary channel of a $1\to 2$ cloning channel as follows:

$$\left\{\sqrt{\frac{1}{3}}|+\rangle\langle +|,\sqrt{\frac{1}{3}}|-\rangle\langle -|,\sqrt{\frac{1}{3}}|0\rangle\langle 0|,\sqrt{\frac{1}{3}}|1\rangle\langle 1|,\sqrt{\frac{1}{3}}|+_Y\rangle\langle +_Y|\sigma_Z,\sqrt{\frac{1}{3}}|-_Y\rangle\langle -_Y|\sigma_Z\right\},$$

where $|+\rangle$ and $|-\rangle$ are the eigenstates of the Pauli $X$ matrix and $|+_Y\rangle$ and $|-_Y\rangle$ are the eigenstates of the Pauli $Y$ matrix. The representation of the complementary channel with unit rank Kraus operators explicitly demonstrates that it is entanglement breaking (as first observed with a positive partial transpose argument in Ref. [30]), and the other arguments in Ref. [30] demonstrate that every $1\to N$ cloning channel is entanglement breaking. Therefore a $1\to N$ cloning channel is in the class of Hadamard channels.

#### 5. *Unruh channels*

The Unruh channel is a natural channel to consider in the context of quantum field theory [35]. The output of the Unruh channel is the quantum state detected by a uniformly accelerated observer when the input is a dual-rail photonic qubit prepared by a Minkowski observer. The mathematical structure of the output of an Unruh channel $\mathcal{N}_{\mathrm{U}}$ is an infinite-dimensional block-diagonal density matrix where the $N$th block is an instance of a $1\to N$ cloning channel

$$\mathcal{N}_{\mathrm{U}}(\sigma) \equiv \bigoplus_{l=2}^{\infty} p_l(z)S_l(\sigma), \quad (13)$$

where

$$p_l(z) \equiv (1-z)^3 z^{l-2}\Delta_{l-1},$$

the "acceleration parameter" $z\in[0,1)$ is a strictly increasing function of acceleration, $\Delta_{l-1}\equiv(l-1)l/2$, and $S_l$ is the output of a $1\to(l-1)$ cloning channel. The complementary channel $\mathcal{N}_{\mathrm{U}}^c$ is similar, with $S_l$ replaced by $S_l^c$, the complementary channel of a $1\to(l-1)$ cloning channel.

## III. REVIEW OF CAPACITY REGIONS

We review several trade-off capacity regions that have appeared in the quantum Shannon theory literature. We first review the Devetak-Shor result for the CQ capacity region [31]. We then review Shor's results on the CE capacity region [39], followed by a review of the natural generalization to the triple trade-off region for CQE coding [40].

### A. Classically enhanced quantum capacity region

Consider a protocol that exploits a noisy quantum channel $\mathcal{N}$ to transmit both classical and quantum information. The goal of such a protocol is to transmit as much of these resources as possible with vanishing error probability and fidelity approaching unity in the limit of a large number of uses of the channel $\mathcal{N}$. More precisely, the protocol transmits one classical message from a set of $M$ messages and an arbitrary quantum state of dimension $K$ using a large number $n$ uses of the quantum channel. The classical rate of transmission is $C\equiv\frac{\log_2(M)}{n}$ bits per channel use, and the quantum rate of transmission is $Q\equiv\frac{\log_2(K)}{n}$ qubits per channel use. If there is a scheme that transmits classical data at rate $C$ with vanishing error probability and quantum data at rate $Q$ with fidelity approaching unity in the limit of a large number of uses of the channel, we say that the rates $C$ and $Q$ form an achievable rate pair $(C,Q)$.

Devetak and Shor's main result in Ref. [31] is that all achievable rate pairs lie in the following CQ capacity region of $\mathcal{N}$

$$\mathcal{C}_{\mathrm{CQ}}(\mathcal{N}) \equiv \overline{\bigcup_{k=1}^{\infty}\frac{1}{k}\mathcal{C}_{\mathrm{CQ}}^{(1)}(\mathcal{N}^{\otimes k})}, \quad (14)$$

where $\overline{Z}$ is the closure of a set $Z$, and the "one-shot" region $\mathcal{C}^{(1)}(\mathcal{N})$ is as follows:

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

The "one-shot, one-state" region $\mathcal{C}_{\mathrm{CQ},\rho}^{(1)}(\mathcal{N})$ is the set of all $C,Q\geqslant 0$ such that

$$C \leqslant I(X;B)_\rho,$$
$$Q \leqslant I(A\rangle BX)_\rho,$$

where $\rho$ is a classical-quantum state of the following form:

$$\rho^{XABE} \equiv \sum_x p_X(x)|x\rangle\langle x|^X\otimes U_{\mathcal{N}}^{A'\to BE}(\phi_x^{AA'}), \quad (15)$$

the states $\phi_x^{AA'}$ are pure and the dimension of the classical system is finite [31]. For general channels, the multiletter

characterization in (14) is necessary, but for certain channels, such as erasure channels and generalized dephasing channels, the CQ capacity region admits a single-letter characterization [31]. In Sec. IV A, we show that the CQ region for all Hadamard channels admits a single-letter characterization.

### B. Entanglement-assisted classical capacity region

Now consider a protocol that exploits shared noiseless entanglement and a noisy quantum channel $\mathcal{N}$ to transmit classical information. The goal of such a protocol is to transmit as much classical information as possible with vanishing error probability while consuming as little entanglement as possible in the limit of a large number of uses of the channel $\mathcal{N}$. More precisely, the protocol transmits one classical message from a set of $M$ messages using a large number $n$ uses of the quantum channel and a maximally entangled state $\Phi^{T_A T_B}$ of the form in (6) where the sender possesses the system $T_A$, the receiver possesses the system $T_B$, and $D$ is the Schmidt rank of the entanglement. The classical rate of transmission is $C \equiv \frac{\log_2(M)}{n}$ bits per channel use, and the rate of entanglement consumption is $E \equiv \frac{\log_2(D)}{n}$ ebits per channel use. If there is a scheme that transmits classical data at rate $C$ with vanishing error probability and consumes entanglement at rate $E$ in the limit of a large number of uses of the channel, we say that the rates $C$ and $E$ form an achievable rate pair $(C,E)$.

Shor's main result in Ref. [39] is that all achievable rate pairs lie in the following entanglement-assisted classical capacity region of $\mathcal{N}$:

$$\mathcal{C}_{\mathrm{CE}}(\mathcal{N}) = \overline{\bigcup_{k=1}^{\infty} \frac{1}{k} \mathcal{C}^{(1)}_{\mathrm{CE}}(\mathcal{N}^{\otimes k})}, \tag{16}$$

where the "one-shot" region $\mathcal{C}^{(1)}_{\mathrm{CE}}(\mathcal{N})$ is a union of the "one-shot, one-state" regions $\mathcal{C}^{(1)}_{\mathrm{CE},\rho}(\mathcal{N})$. The one-shot, one-state region $\mathcal{C}^{(1)}_{\mathrm{CE},\rho}(\mathcal{N})$ is the set of all $C,E \geqslant 0$ such that

$$C \leqslant I(AX;B)_\rho,$$
$$E \geqslant H(A|X)_\rho,$$

where $\rho$ is again a state of the form in (15).

Hsieh and Wilde later gave a more refined "trapezoidal" characterization of the one-shot, one-state region $\mathcal{C}^{(1)}_{\mathrm{CE},\rho}(\mathcal{N})$ [40], but the one-shot regions $\mathcal{C}^{(1)}_{\mathrm{CE}}(\mathcal{N})$ resulting from both Shor's "rectangular" characterization and Hsieh and Wilde's pentagonal characterization are equivalent and the characterization above suffices for our purposes here.

For general channels, the multiletter characterization in (16) is necessary, but for certain channels, such as the erasure channels and generalized dephasing channels, the CE capacity region admits a single-letter characterization as Ref. [40] stated and as we show in full detail here for the generalized dephasing channels. In fact, in Sec. IV B, we show that the CE region for all Hadamard channels, of which a generalized dephasing channel is an example, admits a single-letter characterization.

### C. Capacity region for entanglement-assisted transmission of classical and quantum information

The natural generalization of the two scenarios we have just considered is CQE communication. The goal of such a protocol is to transmit as much classical information with vanishing error probability and quantum information with fidelity approaching unity while consuming as little entanglement as possible in the limit of a large number of uses of the channel $\mathcal{N}$. More precisely, the protocol transmits one classical message from a set of $M$ messages and an arbitrary quantum state of dimension $K$ using a large number $n$ uses of the quantum channel and a maximally entangled state $\Phi^{T_A T_B}$ of dimension $D$. The classical rate of transmission is $C \equiv \frac{\log_2(M)}{n}$ bits per channel use, the quantum rate of transmission is $Q \equiv \frac{\log_2(K)}{n}$ qubits per channel use, and the rate of entanglement consumption is $E \equiv \frac{\log_2(D)}{n}$ ebits per channel use. If there is a scheme that transmits classical data, transmits quantum data, and consumes entanglement at respective rates $C$, $Q$, and $E$ with vanishing error probability and fidelity approaching unity in the limit of a large number of uses of the channel, we say that the rates $C$, $Q$, and $E$ form an achievable rate triple $(C,Q,E)$.

Hsieh and Wilde's main result in Ref. [40] is that all achievable rate triples lie in the following CQE capacity region for the channel $\mathcal{N}$:

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

where the one-shot region $\mathcal{C}^{(1)}_{\mathrm{CQE}}(\mathcal{N})$ is the union of the "one-shot, one-state" regions $\mathcal{C}^{(1)}_{\mathrm{CQE},\rho}(\mathcal{N})$. The "one-shot, one-state" region $\mathcal{C}^{(1)}_{\mathrm{CQE},\rho}(\mathcal{N})$ is the set of all $C,Q,E \geqslant 0$ such that

$$C + 2Q \leqslant I(AX;B)_\rho, \tag{17}$$
$$Q \leqslant I(A\rangle BX)_\rho + E, \tag{18}$$
$$C + Q \leqslant I(X;B)_\rho + I(A\rangle BX)_\rho + E. \tag{19}$$

One of the observations in Ref. [40] is that there is a way to single-letterize the CQE capacity region provided one can show that both the CQ and CE trade-off curves single-letterize. We review these arguments briefly. Suppose that we have shown that both the CQ and CE trade-off curves single-letterize. Then three surfaces specify the boundary of the CQE capacity region so that we can simplify the characterization in (17)–(19). The first surface to consider is that formed by combining the CE trade-off curve with the "inverse" of the superdense coding protocol. Recall that the superdense coding protocol exploits a noiseless ebit and a noiseless qubit channel to transmit two classical bits [38]. Let $(C_{\mathrm{CE}}(s_1),0,E_{\mathrm{CE}}(s_1))$ denote a parametrization of all points on the CE trade-off curve with respect to some parameter $s_1 \in [0,1/2]$, and recall that each point on the trade-off curve has corresponding entropic quantities of the form $(I(AX;B),0,H(A|X))$. Then the surface formed by combining the CE trade-off curve with the inverse of superdense coding is

$$\left\{(C_{\mathrm{CE}}(s_1) - 2Q, Q, E_{\mathrm{CE}}(s_1) - Q) : s_1 \in [0,1/2], \right.$$
$$\left. 0 \leqslant Q \leqslant \min\left\{\tfrac{1}{2} C_{\mathrm{CE}}(s_1), E_{\mathrm{CE}}(s_1)\right\}\right\}.$$

This surface forms an outer bound for the capacity region. Were it not so, then one could combine points outside it with superdense coding to outperform points on the CE trade-off curve, contradicting the optimality of this trade-off curve.

The next surface to consider is that formed by combining the CQ trade-off curve with the "inverse" of the entanglement distribution protocol. Recall that the entanglement distribution protocol exploits a noiseless qubit channel to establish a shared noiseless ebit. Let $\mathbf{(}C_{\mathrm{CQ}}(s_2), Q_{\mathrm{CQ}}(s_2), 0\mathbf{)}$ denote a parametrization of all points on the CQ trade-off curve with respect to some parameter $s_2 \in [0,1/2]$, and recall that each point on the trade-off curve has corresponding entropic quantities of the form $\mathbf{(}I(X;B), I(A\rangle BX), 0\mathbf{)}$. Then the surface formed by combining the CQ trade-off curve with the inverse of entanglement distribution is

$$\{\mathbf{(}C_{\mathrm{CQ}}(s_2), Q_{\mathrm{CQ}}(s_2) + E, E\mathbf{)} : s_2 \in [0,1/2], \quad E \geqslant 0\}.$$

This surface also forms an outer bound for the capacity region. Were it not so, then one could combine points outside it with entanglement distribution to outperform points on the CQ trade-off curve, contradicting the optimality of this trade-off curve.

The final surface to consider is the following regularization of the plane that (17) specifies

$$C + 2Q \leqslant \frac{1}{n} h(\mathcal{N}^{\otimes n}), \tag{20}$$

for all $n \geqslant 1$, where

$$h(\mathcal{N}) \equiv \max_{\rho} I(AX;B),$$

and $\rho$ is a state of the form in (15). Lemma 1 states that $h(\mathcal{N}^{\otimes n})$ actually single-letterizes for any quantum channel $\mathcal{N}$

$$h(\mathcal{N}^{\otimes n}) = n h(\mathcal{N}),$$

so that the computation of the boundary $h(\mathcal{N}^{\otimes n})/n$ is tractable. Its proof is a consequence of the single letterization of the entanglement-assisted classical capacity [37], but we provide it in Appendix A for completeness.

*Lemma 1*. The plane in (20) admits a single-letter characterization for any noisy quantum channel $\mathcal{N}$

$$h(\mathcal{N}^{\otimes n}) = n h(\mathcal{N}).$$

The above three surfaces all form outer bounds on the CQE capacity region, but is it clear that we can achieve points along the boundaries? To answer this question, we should consider the intersection of the first and second surfaces, found by solving the following equation for $Q$ and $E$:

$$\begin{aligned}&\mathbf{(}C_{\mathrm{CE}}(s_1) - 2Q, Q, E_{\mathrm{CE}}(s_1) - Q\mathbf{)}\\ &= \mathbf{(}C_{\mathrm{CQ}}(s_2), Q_{\mathrm{CQ}}(s_2) + E, E\mathbf{)}.\end{aligned}$$

Using the entropic expressions for the trade-off curves and solving the above equation gives that all points along the intersection have entropic quantities of the following form:

$$\left(I(X;B), \tfrac{1}{2} I(A;B|X), \tfrac{1}{2} I(A;E|X)\right).$$

The authors of Ref. [40] constructed a protocol, dubbed the "classically enhanced father protocol" that can achieve the above rates for CQE communication. Thus, by combining this protocol with superdense coding and entanglement distribution, we can achieve all points inside the first and second surfaces with entanglement consumption below a certain rate. Finally, by combining this protocol with superdense coding, entanglement distribution, and the wasting of entanglement, we can achieve all points that lie inside all three surfaces, and thus we can achieve the full CQE capacity region. We summarize these results as the following theorem.

*Theorem 1*. Suppose the CQ and CE trade-off curves of a quantum channel $\mathcal{N}$ single-letterize. Then the full CQE capacity region of $\mathcal{N}$ single-letterizes

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

We apply the above theorem in the next section. We first show that both the CQ and CE trade-off curves single-letterize for all Hadamard channels, and it then follows that the full CQE capacity region single-letterizes for these channels.

## IV. SINGLE LETTERIZATION OF THE CQ AND CE TRADE-OFF CURVES FOR HADAMARD CHANNELS

### A. CQ trade-off curve

For the CQ region we would like to maximize both the classical and quantum communication rates, but we cannot have both be simultaneously optimal. Thus, we must trade between these resources. If we are willing to reduce the quantum communication rate by a little then we can communicate more classical information and vice versa.

Our main theorem below states and proves that the following function generates points on the CQ trade-off curve for Hadamard channels

$$f_\lambda(\mathcal{N}) \equiv \max_{\rho} I(X;B)_\rho + \lambda I(A\rangle BX)_\rho, \tag{21}$$

where the state $\rho$ is of the form in (15) and $\lambda \geqslant 1$.

*Theorem 2*. For any fixed $\lambda \geqslant 1$, the function in (21) leads to a point

$$\mathbf{(}I(X;B)_\rho, I(A\rangle BX)_\rho\mathbf{)},$$

on the CQ trade-off curve, provided $\rho$ maximizes (21).

*Proof*. This theorem follows from the results of Lemmas 2, 3, and 4 and Corollary 1. ∎

*Lemma 2*. For any fixed $\lambda \geqslant 0$, the function in (21) leads to a point $\mathbf{(}I(X;B)_\rho, I(A\rangle BX)_\rho\mathbf{)}$ on the one-shot CQ trade-off curve in the sense of Theorem 2.

*Proof*. We argue by contradiction. Suppose that a particular state $\rho^{XAB}$ of the form in (15) maximizes (21). Suppose further that it does not lead to a point on the trade-off curve. That is, given the constraint $Q = I(A\rangle BX)_\rho$, there is some other state $\sigma^{XAB}$ of the form in (15) such that $I(A\rangle BX)_\sigma = I(A\rangle BX)_\rho = Q$ where $C = I(X;B)_\sigma > I(X;B)_\rho$. But this result implies the following inequality for all $\lambda \geqslant 0$

$$I(X;B)_\sigma + \lambda I(A\rangle BX)_\sigma > I(X;B)_\rho + \lambda I(A\rangle BX)_\rho,$$

contradicting the fact that the state $\rho^{XAB}$ maximizes (21).

*Lemma 3*. We obtain all points on the one-shot CQ trade-off curve by considering $\lambda \geqslant 1$ in the maximization in (21) because the maximization optimizes only the classical capacity for all $\lambda$ such that $0 \leqslant \lambda < 1$.

*Proof.* Consider a state $\rho^{XABE}$ of the form in (15). Suppose that we perform a von Neumann measurement of the system $A$, resulting in a classical variable $Y$ and let $\sigma^{XYBE}$ denote the resulting state. Then the following chain of inequalities holds for all $\lambda$ such that $0 \leqslant \lambda < 1$

$$\begin{aligned}
&I(X;B)_\rho + \lambda I(A\rangle BX)_\rho \\
&\quad = H(B)_\rho + (\lambda - 1)H(B|X)_\rho - \lambda H(E|X)_\rho \\
&\quad = H(B)_\sigma + (\lambda - 1)H(B|X)_\sigma - \lambda H(E|X)_\sigma \\
&\quad \leqslant H(B)_\sigma + (\lambda - 1)H(B|XY)_\sigma - \lambda H(E|XY)_\sigma. \\
&\quad = H(B)_\sigma + (\lambda - 1)H(B|XY)_\sigma - \lambda H(B|XY)_\sigma \\
&\quad = H(B)_\sigma - H(B|XY)_\sigma \\
&\quad = I(XY;B)_\sigma.
\end{aligned}$$

The first equality follows from definitions and because the state $\rho^{XABE}$ on systems $A$, $B$, and $E$ is pure when conditioned on the classical variable $X$. The second equality follows because the von Neumann measurement does not affect the systems involved in the entropic expressions. The inequality follows because $0 \leqslant \lambda < 1$ and conditioning reduces entropy. The third equality follows because the reduced state of $\sigma^{XYBE}$ on systems $B$ and $E$ is pure when conditioned on both $X$ and $Y$. The last two equalities follow from algebra and the definition of the quantum mutual information.

Thus it becomes clear that the maximization of the original quantity for $0 \leqslant \lambda < 1$ is always less than the classical capacity because $I(XY;B)_\sigma \leqslant \max_\rho I(X;B)$. It then follows that the trade-off curve really starts when $\lambda \geqslant 1$. ∎

We remark that the above proof gives an alternate mathematical justification for considering only $\lambda \geqslant 1$ in the maximization than does the original operational reason given in Ref. [31].

The following lemma is the crucial one that leads to our main result in this section: the single letterization of the CQ trade-off curve for Hadamard channels.

*Lemma 4.* The following additivity relation holds for a Hadamard channel $\mathcal{N}_1$ and any other channel $\mathcal{N}_2$

$$f_\lambda(\mathcal{N}_1 \otimes \mathcal{N}_2) = f_\lambda(\mathcal{N}_1) + f_\lambda(\mathcal{N}_2).$$

*Proof.* The inequality $f_\lambda(\mathcal{N}_1 \otimes \mathcal{N}_2) \geqslant f_\lambda(\mathcal{N}_1) + f_\lambda(\mathcal{N}_2)$ trivially holds for all quantum channels because the maximization on the right-hand side (rhs) is a restriction of the maximization in the left-hand side (lhs) to a tensor product of states of the form in (15). Therefore we prove the nontrivial inequality $f_\lambda(\mathcal{N}_1 \otimes \mathcal{N}_2) \leqslant f_\lambda(\mathcal{N}_1) + f_\lambda(\mathcal{N}_2)$ when $\mathcal{N}_1$ is a Hadamard channel.

Suppose that $\mathcal{N}_1^{A_1\to B_1}$ is a Hadamard channel and $\mathcal{N}_2^{A_2\to B_2}$ is any quantum channel with respective complementary channels $(\mathcal{N}_1^c)^{A_1\to E_1}$ and $(\mathcal{N}_2^c)^{A_2\to E_2}$. Then $\mathcal{N}_1^c$ is entanglement breaking because $\mathcal{N}_1$ is a Hadamard channel. Therefore there exists a degrading map $\mathcal{D}^{B_1\to E_1}$ because $\mathcal{N}_1^c$ is entanglement breaking. The degrading map consists of a measurement with classical output on system $Y$ followed by the preparation of a state on $E_1$ depending on the measurement outcome (as described in Sec. II C 2). We can therefore write $\mathcal{D}^{B_1\to E_1} = \mathcal{D}_2^{Y\to E_1} \circ \mathcal{D}_1^{B_1\to Y}$.

Let

$$\begin{aligned}
\psi^{XAA_1A_2} &\equiv \sum_x p_X(x)|x\rangle\langle x|^X \otimes |\phi_x\rangle\langle\phi_x|^{AA_1A_2}, \\
\theta^{XAB_1E_1A_2} &\equiv U_1\psi U_1^\dagger, \\
\rho^{XAB_1E_1B_2E_2} &\equiv (U_1 \otimes U_2)\psi(U_1 \otimes U_2)^\dagger,
\end{aligned} \tag{22}$$

where $U_j^{A_j\to B_jE_j}$ is the isometric extension of $\mathcal{N}_j$. Suppose further that $\rho$ is the state that maximizes $f_\lambda(\mathcal{N}_1 \otimes \mathcal{N}_2)$. Also, let $\sigma \equiv \mathcal{D}_1(\rho)$ and note that this state is of the following form:

$$\begin{aligned}
\sigma^{XYAB_2E_1E_2} \equiv \sum_x p_X(x)p_{Y|X}(y|x)|x\rangle\langle x|^X \otimes |y\rangle\langle y|^Y \\
\otimes |\phi_{x,y}\rangle\langle\phi_{x,y}|^{AB_2E_1E_2}.
\end{aligned}$$

Then the following inequalities hold for all $\lambda \geqslant 1$:

$$\begin{aligned}
f_\lambda(\mathcal{N}_1 \otimes \mathcal{N}_2) &= I(X;B_1B_2)_\rho + \lambda I(A\rangle B_1B_2X)_\rho \\
&= H(B_1B_2)_\rho + (\lambda - 1)H(B_1B_2|X)_\rho - \lambda H(E_1E_2|X)_\rho \\
&\leqslant H(B_1)_\rho + (\lambda - 1)H(B_1|X)_\rho - \lambda H(E_1|X)_\rho + H(B_2)_\rho + (\lambda - 1)H(B_2|B_1X)_\rho - \lambda H(E_2|E_1X)_\rho \\
&\leqslant H(B_1)_\rho + (\lambda - 1)H(B_1|X)_\rho - \lambda H(E_1|X)_\rho + H(B_2)_\sigma + (\lambda - 1)H(B_2|YX)_\sigma - \lambda H(E_2|YX)_\sigma \\
&= \mathbf{(}I(X;B_1)_\theta + \lambda I(AA_2\rangle B_1X)_\theta\mathbf{)} + \mathbf{(}I(XY;B_2)_\sigma + \lambda I(AE_1\rangle B_2XY)_\sigma\mathbf{)} \\
&\leqslant f_\lambda(\mathcal{N}_1) + f_\lambda(\mathcal{N}_2).
\end{aligned}$$

The first equality follows by definition and the assumption that $\rho$ maximizes $f_\lambda(\mathcal{N}_1 \otimes \mathcal{N}_2)$. The second equality follows from entropic manipulations and the fact that $H(AB_1B_2|X) = H(E_1E_2|X)$ for the state $\rho$. The first inequality follows from subadditivity of entropy and the chain rule [62]. The second inequality uses two applications of the monotonicity of conditional entropy with respect to quantum channels acting on the conditioned system. Specifically, $H(B_2|B_1X)_\rho \leqslant H(B_2|YX)_\sigma$ because of the existence of the map $\mathcal{D}_1$ while $H(E_2|YX)_\sigma \leqslant H(E_2|E_1X)_\rho$ because of the existence of the map $\mathcal{D}_2$. The third equality follows because $H(E_1|X)_\rho = H(AA_2B_1|X)_\theta$ and $H(E_2|YX)_\rho = H(AE_1B_2|YX)_\sigma$. The final inequality follows because $\theta$ and $\sigma$ are both states of the form in (15). ∎

*Corollary 1.* The one-shot CQ trade-off curve is equal to the regularized CQ trade-off curve when the noisy quantum channel $\mathcal{N}$ is a Hadamard channel

$$f_\lambda(\mathcal{N}^{\otimes n}) = nf_\lambda(\mathcal{N}).$$

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

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

The first equality follows by expanding the tensor product. The second critical equality follows from the application of Lemma 4 because $\mathcal{N}$ is a Hadamard channel. The final equality follows from the induction hypothesis. ∎

There is one last point that we should address concerning the CQ trade-off curve. There is the possibility in this trade-off problem that the parameter $\lambda$ does not parametrize all points on the trade-off curve, potentially leading to a gap in the trade-off curve. We address this concern in Appendix A by proving that one gets the entire trade-off curve by varying $\lambda$ and taking the convex hull of the resulting points. A similar proof holds for the CE trade-off curve.

### B. CE trade-off curve

For the CE region we would like to maximize both the classical communication rate while minimizing the entanglement consumption rate, but we cannot have both be simultaneously optimal. Thus, we must trade between these resources. If we are willing to reduce the classical communication rate by little, then the protocol does not require as much entanglement consumption.

Our main theorem below states that the following function generates points on the CE trade-off curve

$$g_\lambda(\mathcal{N}) \equiv \max_\rho I(AX;B)_\rho - \lambda H(A|X)_\rho, \tag{23}$$

where the state $\rho$ is of the form in (15) and $0 \leqslant \lambda \leqslant 1$.

*Theorem 3*. For $0 \leqslant \lambda \leqslant 1$, the function in (23) leads to a point

$$\mathbf{(}I(AX;B)_\rho, H(A|X)_\rho\mathbf{)},$$

on the CE trade-off curve, provided $\rho$ maximizes (23).

*Proof*. The proof of this theorem proceeds similarly to that of Theorem 2 in the previous section. It follows from the results of Lemmas 8, 9 and 10 and Corollary 2 in Appendix B. ∎

## V. PARAMETRIZATION OF THE TRADE-OFF CURVES

The results in the previous section demonstrate that the CQE capacity region for all Hadamard channels single-letterizes. These results imply that we can actually compute the CQE capacity region for these channels, by the arguments in Sec. III C. In this section we consider several instances of Hadamard channels and show how we can exactly characterize their corresponding CQE capacity region. We first consider the qubit dephasing channel and the $1 \to N$ cloning channel, and then show how to apply the results for the $1 \to N$ cloning channel to the Unruh channel.

### A. Qubit dephasing channel

This section briefly recalls the parametrizations for the CQ and CE trade-off curves for a qubit dephasing channel. Devetak and Shor gave a particular parametrization for the CQ trade-off curve for the case of a qubit dephasing channel in Appendix B of Ref. [31], and Hsieh and Wilde followed by giving a parametrization for the CE trade-off curve for qubit dephasing channels [40]. For the purposes of completion and comparison, we recall these parametrizations below, and Appendices D and E of this paper provide full proofs of these assertions.

*Theorem 4*. All points on the CQ trade-off curve for a $p$-dephasing qubit channel have the following form:

$$\mathbf{(}1 - H_2(\mu), H_2(\mu) - H_2[\gamma(\mu,p)]\mathbf{)},$$

and all points on the CE trade-off curve for a $p$-dephasing qubit channel have the following form:

$$[1 + H_2(\mu) - H_2\mathbf{(}\gamma(\mu,p)\mathbf{)}, H_2(\mu)],$$

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

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

Figure 1 plots both the CQ and CE trade-off curves for several $p$-dephasing qubit channels, where $p$ varies from zero to one in increments of $1/10$. The figure demonstrates that both classically enhanced quantum coding and entanglement-assisted classical coding beat a time-sharing strategy for any $p$ such that $0 < p < 1$.

We do not plot the full CQE capacity region for the dephasing qubit channel but instead point the interested reader to Fig. 4 of Ref. [40].

### B. $1 \to N$ cloning channels

We now compute exact expressions for the CQ and CE trade-off curves, plot them for several $1 \to N$ cloning channels, and plot the CQE capacity region for a particular $1 \to N$ cloning channel. The following theorem states these expressions, and the lemmas and subsequent calculation following it provide a proof.

*Theorem 5*. All points on the CQ trade-off curve for a $1 \to N$ cloning channel have the following form:

$$\left[\log_2(N+1) - H\left(\frac{\lambda_i(\mu)}{\Delta_N}\right), H\left(\frac{\lambda_i(\mu)}{\Delta_N}\right) - H\left(\frac{\eta_i(\mu)}{\Delta_N}\right)\right],$$

and all points on the CE trade-off curve for a $1 \to N$ cloning channel have the following form:

$$\mathbf{[}\log_2(N+1) + H_2(\mu) - H[\eta_i(\mu)/\Delta_N],\ H_2(\mu)\mathbf{]},$$

where $H$ is the entropy function $H(\cdot) \equiv -\sum_i (\cdot)\log_2(\cdot)$,

$$\begin{aligned} \Delta_N &\equiv N(N+1)/2, \\ \lambda_i(\mu) &\equiv (N-2i)\mu + i \quad \text{for} \quad 0 \leqslant i \leqslant N, \\ \eta_i(\mu) &\equiv (N-1-2i)\mu + i + 1 \quad \text{for} \quad 0 \leqslant i \leqslant N-1, \\ \mu &\in [0,1/2]. \end{aligned}$$

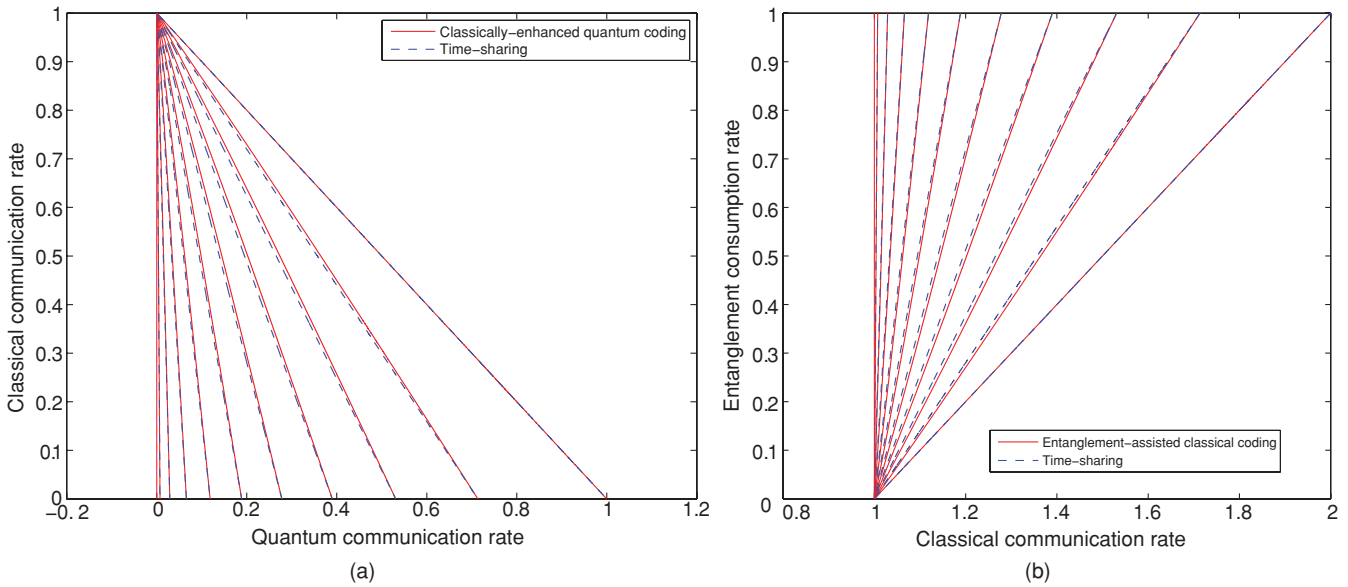



FIG. 1. (Color online) Plot of (a) the CQ trade-off curve and (b) the CE trade-off curve for a $p$-dephasing qubit channel for $p = 0, 0.1, 0.2, \ldots, 0.9, 1$. The trade-off curves for $p = 0$ correspond to those of a noiseless qubit channel and are the rightmost trade-off curve in each plot. The trade-off curves for $p = 1$ correspond to those for a classical channel, and are the leftmost trade-off curves in each plot. Each trade-off curve between these two extremes beats a time-sharing strategy, but these two extremes do not beat time sharing.

*Lemma 5*. An ensemble of the following form parametrizes all points on the CQ trade-off curve for a $1 \to N$ cloning channel:

$$\tfrac{1}{2}|0\rangle\langle 0|^X \otimes \psi_0^{AA'} + \tfrac{1}{2}|1\rangle\langle 1|^X \otimes \psi_1^{AA'}, \tag{24}$$

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

$$\mathrm{Tr}_A\{\psi_0^{AA'}\} = \mu|0\rangle\langle 0|^{A'} + (1-\mu)|1\rangle\langle 1|^{A'}, \tag{25}$$

$$\mathrm{Tr}_A\{\psi_1^{AA'}\} = (1-\mu)|0\rangle\langle 0|^{A'} + \mu|1\rangle\langle 1|^{A'}. \tag{26}$$

*Proof*. Consider a classical-quantum state with a finite number of conditional density operators $\rho_x^{A'}$

$$\rho^{XA'} = \sum_x p_X(x)|x\rangle\langle x|^X \otimes \rho_x^{A'}.$$

Let $\mathcal{N}_{\mathrm{Cl}}^{A'\to B}$ denote the cloning channel, $U_{\mathcal{N}_{\mathrm{Cl}}}^{A'\to BE}$ an isometric extension, $(\mathcal{N}_{\mathrm{Cl}}^c)^{A'\to E}$ a complementary channel, and let $\rho^{XBE}$ denote the state resulting from applying the isometric extension $U_{\mathcal{N}_{\mathrm{Cl}}}^{A'\to BE}$ of the cloning channel to system $A'$ of $\rho^{XA'}$.

We can form a new classical-quantum state with quadruple the number of conditional density operators by applying all Pauli matrices to the original conditional density operators:

$$\sigma^{XJA'} \equiv \sum_x \sum_{j=0}^{3} \frac{1}{4} p_X(x)|x\rangle\langle x|^X \otimes |j\rangle\langle j|^J \otimes \sigma_j \rho_x^{A'} \sigma_j,$$

where $\sigma_j$ labels the four Pauli matrices $\sigma_0 \equiv I$, $\sigma_1 \equiv X$, $\sigma_2 \equiv Y$, $\sigma_3 \equiv Z$. Let $\sigma^{XJBE}$ denote the state resulting from sending $A'$ through the isometric extension $U_{\mathcal{N}_{\mathrm{Cl}}}^{A'\to BE}$ of the cloning channel.

Recall from Sec. II C4 that the cloning channel is covariant. In fact, the following relationships hold for any input density operator $\sigma$ and any unitary $V$ acting on the input system $A'$

$$\mathcal{N}_{\mathrm{Cl}}(V\sigma V^\dagger) = R_V \mathcal{N}_{\mathrm{Cl}}(\sigma) R_V^\dagger,$$
$$\mathcal{N}_{\mathrm{Cl}}^c(V\sigma V^\dagger) = S_V \mathcal{N}_{\mathrm{Cl}}^c(\sigma) S_V^\dagger,$$

where $R_V$ and $S_V$ are higher-dimensional irreducible representations of the unitary $V$ on the respective systems $B$ and $E$. The state $\sigma^B$ is equal to the maximally mixed state on the symmetric subspace for the following reasons:

$$\sigma^B = \mathcal{N}_{\mathrm{Cl}}(\sigma^{A'}) = \mathcal{N}_{\mathrm{Cl}}\left(\frac{I^{A'}}{2}\right) = \mathcal{N}_{\mathrm{Cl}}\left(\int V\omega V^\dagger \, \mathrm{d}V\right)$$
$$= \int R_V \mathcal{N}(\omega) R_{V^\dagger} \, \mathrm{d}V = \frac{1}{N+1}\sum_{i=0}^{N} |i\rangle\langle i|^B, \tag{27}$$

where the third equality exploits the linearity and covariance of the cloning channel $\mathcal{N}_{\mathrm{Cl}}$.

Then the following chain of inequalities holds for all $\lambda \geqslant 1$:

$$\begin{aligned}
&I(X;B)_\rho + \lambda I(A\rangle BX)_\rho \\
&\quad = H(B)_\rho + (\lambda-1)H(B|X)_\rho - \lambda H(E|X)_\rho \\
&\quad = H(B)_\rho + (\lambda-1)H(B|XJ)_\sigma - \lambda H(E|XJ)_\sigma \\
&\quad \leqslant H(B)_\sigma + (\lambda-1)H(B|XJ)_\sigma - \lambda H(E|XJ)_\sigma \\
&\quad = \log_2(N+1) + (\lambda-1)H(B|XJ)_\sigma - \lambda H(E|XJ)_\sigma \\
&\quad = \log_2(N+1) + \sum_x p_X(x)[(\lambda-1)H(B)_{\rho_x} - \lambda H(E)_{\rho_x}] \\
&\quad \leqslant \log_2(N+1) + \max_x[(\lambda-1)H(B)_{\rho_x} - \lambda H(E)_{\rho_x}] \\
&\quad = \log_2(N+1) + (\lambda-1)H(B)_{\rho_x^*} - \lambda H(E)_{\rho_x^*}.
\end{aligned} \tag{28}$$

The first equality follows by standard entropic manipulations. The second equality follows because the conditional entropies are invariant under unitary transformations

$$H(B)_{R_{\sigma_j}\rho_x^B R_{\sigma_j}^\dagger} = H(B)_{\rho_x^B},$$
$$H(E)_{S_{\sigma_j}\rho_x^E S_{\sigma_j}^\dagger} = H(E)_{\rho_x^E},$$

where $R_{\sigma_j}$ and $S_{\sigma_j}$ are higher-dimensional representations of $\sigma_j$ on systems $B$ and $E$, respectively. The first inequality follows because entropy is concave, i.e., the local state $\sigma^B$ is a mixed version of $\rho^B$. The third equality follows because (27) implies that $H(B)_{\sigma^B} = \log_2(N+1)$. The fourth equality follows because the systems $X$ and $J$ are classical. The second inequality follows because the maximum value of a realization of a random variable is not less than its expectation. The final equality simply follows by defining $\rho_x^*$ to be the conditional density operator on systems $B$ and $E$ that arises from sending an arbitrary state through the channel.

The entropies $H(B)_{\rho_x^*}$ and $H(E)_{\rho_x^*}$ depend only on the eigenvalues of the input state $\rho_x^*$ by the covariance of both the cloning channel and its complement. We can therefore choose $\rho_x^*$ to be diagonal in the $\{|0\rangle,|1\rangle\}$ basis of $A'$. The ensemble defined to consist of the purifications of $\rho_x^*$ and $\sigma_x\rho_x^*\sigma_x^*$ assigned equal probabilities then saturates the upper bound on $I(X;B)_\rho + \lambda I(A\rangle BX)_\rho$ in (28), which concludes the proof. ∎

*Proof of CQ trade-off in Theorem 5*. We simply need to compute the Holevo information $I(X;B)$ and the coherent information $I(A\rangle BX)$ for an ensemble of the form in the statement of Lemma 5.

The purifications of the states in (25) and (26) are the same as in (D4) and (D5). These states lead to classical-quantum states of the form in (24). An isometric extension $U_{\mathcal{N}_{\text{Cl}}}$ of the $1 \to N$ cloning channel acts as follows on these states:

$$|\psi_0\rangle^{ABE} = \frac{1}{\sqrt{\Delta_N}}\left[\sum_{i=0}^{N-1}(\sqrt{\mu}\sqrt{N-i}\,|0\rangle^A|i\rangle^B + \sqrt{1-\mu}\sqrt{i+1}\,|1\rangle^A|i+1\rangle^B)|i\rangle^E\right],$$
$$|\psi_1\rangle^{ABE} = \frac{1}{\sqrt{\Delta_N}}\left[\sum_{i=0}^{N-1}(\sqrt{1-\mu}\sqrt{N-i}\,|0\rangle^A|i\rangle^B + \sqrt{\mu}\sqrt{i+1}\,|1\rangle^A|i+1\rangle^B)|i\rangle^E\right].$$

The state at the output of the isometric extension is as follows:

$$\rho^{XABE} \equiv \tfrac{1}{2}\left[|0\rangle\langle 0|^X \otimes \psi_0^{ABE} + |1\rangle\langle 1|^X \otimes \psi_1{}^{ABE}\right].$$

Defining $\lambda_i(\mu) \equiv (N-2i)\mu + i$ and $\Delta_N = N(N+1)/2$ gives

$$\rho^{XB} = \frac{1}{2}\left[|0\rangle\langle 0|^X \otimes \psi_0^B + |1\rangle\langle 1|^X \otimes \psi_1^B\right],$$
$$\psi_0^B = \sum_{i=0}^{N}\frac{\lambda_i(\mu)}{\Delta_N}|i\rangle\langle i|^B, \tag{29}$$
$$\psi_1^B = \sum_{i=0}^{N}\frac{\lambda_i(\mu)}{\Delta_N}|N-i\rangle\langle N-i|^B,$$

$$\rho^B = \frac{1}{N+1}\sum_{i=0}^{N}|i\rangle\langle i|^B. \tag{30}$$

We can then compute the entropies $H(B)$ and $H(B|X)$

$$H(B) = \log_2(N+1),$$
$$H(B|X) = H[\lambda_i(\mu)/\Delta_N],$$

giving the following expression for the Holevo information:

$$I(X;B) = H(B) - H(B|X) = \log_2(N+1) - H[\lambda_i(\mu)/\Delta_N], \tag{31}$$

The above expression coincides with the expression for the classical capacity of the $1 \to N$ cloning channel when $\mu = 0$ (Corollary 2 in Ref. [30])

$$I(X;B)_{\mu=0} = 1 - \log_2 N + \frac{1}{\Delta_N}\sum_{i=0}^{N} i\log_2 i.$$

The expression in (31) vanishes when $\mu = \frac{1}{2}$.

We now compute the coherent information $I(A\rangle BX) = H(B|X) - H(E|X)$. The following states are important in this computation:

$$\rho^{XE} = \frac{1}{2}\left[|0\rangle\langle 0|^X \otimes \psi_0^E + |1\rangle\langle 1|^X \otimes \psi_1^E\right],$$
$$\psi_0^E = \sum_{i=0}^{N-1}\frac{\eta_i(\mu)}{\Delta_N}|i\rangle\langle i|^E, \tag{32}$$

$$\psi_1^E = \sum_{i=0}^{N-1}\frac{\eta_i(\mu)}{\Delta_N}|N-1-i\rangle\langle N-1-i|^E, \tag{33}$$

where $\eta_i(\mu) \equiv (N-1-2i)\mu + i + 1$. We can then compute the entropy $H(E|X)$:

$$H(E|X) = H[\eta_i(\mu)/\Delta_N],$$

and the coherent information is as follows:

$$I(A\rangle BX) = H[\lambda_i(\mu)/\Delta_N] - H[\eta_i(\mu)/\Delta_N].$$

The above expression vanishes when $\mu = 0$. It coincides with the quantum capacity of a $1 \to N$ cloning channel when $\mu = \frac{1}{2}$ (Eq. (30) in Ref. [33])

$$I(A\rangle BX)_{\mu=1/2} = \log_2\left(\frac{N+1}{N}\right).$$ ∎

We now turn to the proof of the CE trade-off curve for the $1 \to N$ cloning channel.

*Lemma 6*. An ensemble of the same form as in Lemma 5 parametrizes all points on the CE trade-off curve for a $1 \to N$ cloning channel.

*Proof*. The proof of this lemma proceeds along similar lines as the proof of Lemma 5, using the same states and ideas concerning the cloning channel. We highlight the major differences by giving the following chain of inequalities that holds for all $\lambda$ such that $0 \leqslant \lambda < 1$.

$$I(AX;B)_\rho - \lambda H(A|X)_\rho$$
$$= (1-\lambda)H(A|X)_\rho + H(B)_\rho - H(E|X)_\rho$$
$$= H(B)_\rho + (1-\lambda)H(A|XJ)_\sigma - H(E|XJ)_\sigma$$

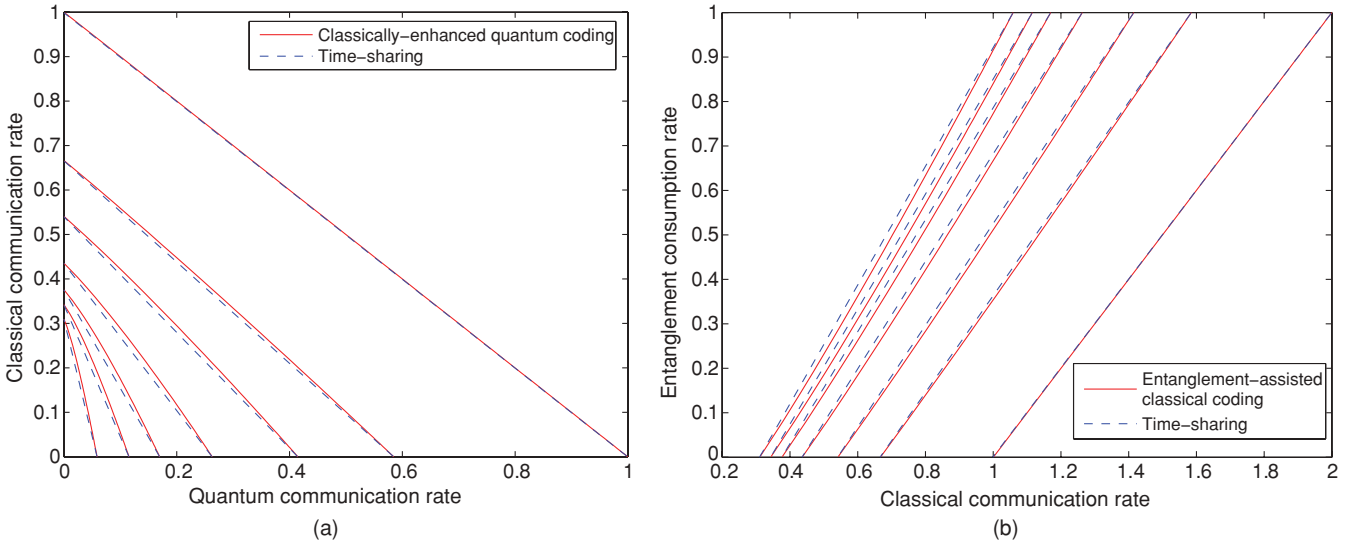



FIG. 2. (Color online) Plot of (a) the CQ trade-off curve and (b) the CE trade-off curve of a $1 \to N$ cloning channel for $N = 1, 2, 3, 5, 8, 12$, and 24. The trade-off curves for $N = 1$ correspond to those for the noiseless qubit channel and are the rightmost trade-off curves in each plot. In both plots, proceeding left from the $N = 1$ curve, we obtain trade-off curves for $N = 2, 3, 5, 8, 12$, and 24 and notice that they all beat a time-sharing strategy by a larger relative proportion when $N$ increases.

$$
\begin{aligned}
&\leqslant H(B)_\sigma + (1-\lambda)H(A|XJ)_\sigma - H(E|XJ)_\sigma \\
&= \log_2(N+1) + (1-\lambda)H(A|XJ)_\sigma - H(E|XJ)_\sigma \\
&= \log_2(N+1) + \sum_x p_X(x)[(1-\lambda)H(A)_{\psi_x} - H(E)_{\psi_x}] \\
&\leqslant \log_2(N+1) + \max_x[(1-\lambda)H(A)_{\psi_x} - H(E)_{\psi_x}] \\
&= \log_2(N+1) + (1-\lambda)H(A)_{\psi_x^*} - H(E)_{\psi_x^*}.
\end{aligned}
$$

We do not provide justifications for the above chain of inequalities because it follows for similar reasons to the chain of inequalities in Lemma 5. ■

*Proof of CE trade off in Theorem 5.* The proof follows by noting that $I(AX;B) = H(A|X) + H(B) - H(E|X)$, $H(A|X) = H_2(\mu)$, and that we have already computed $H(B)$ and $H(E|X)$ in the proof of the CQ trade off in Theorem 5. ■

Figure 2 plots the CQ and CE trade-off curves for a $1 \to N$ cloning channel, for $N = 1, 2, 3, 5, 8, 12$, and 24. The figure demonstrates that both classically enhanced quantum coding and entanglement-assisted classical coding beat a time-sharing strategy for a cloning channel when $N > 1$.

Figure 3 plots the full CQE capacity region for a $1 \to 10$ cloning channel. The figure demonstrates that the classically enhanced father protocol, combined with entanglement distribution and superdense coding beats a time-sharing strategy because the first two surfaces described in Sec. III C are strictly concave for the cloning channels when $N > 1$.

### C. Unruh channel

We now compute the trade-off curves for the Unruh channel, defined in (13). Capacity questions are directly linked to those of the cloning channel because the mathematical structure of the output of the Unruh channel is an infinite-dimensional block-diagonal density matrix, with each block containing an occurrence of a $1 \to N$ cloning channel. In fact, maximizing the rate of transmission is equivalent to maximizing it in each block. This observation gives us the following theorem, and its proof appears in Appendix F.

*Theorem 6.* All points on the CQ trade-off curve for an Unruh channel have the following form:

$$(C_{\mathrm{CQ}}(z,\mu),\ Q_{\mathrm{CQ}}(z,\mu)),$$

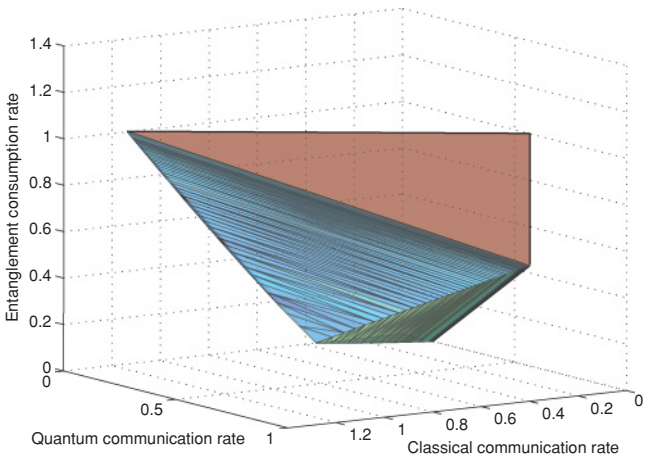



FIG. 3. (Color online) The figure plots the CQE capacity region for a $1 \to 10$ cloning channel. It features three distinct surfaces. The first is the flat vertical plane that arises from the bound $R + 2Q \leqslant \log_2(\frac{N+1}{N}) + 1$, which is the entanglement-assisted classical capacity of a $1 \to N$ cloning channel. The plane extends infinitely upward because we can always achieve these rate triples simply by wasting entanglement. The second surface is that below and to the left of the plane, formed by combining the CE trade-off curve with the inverse of superdense coding, as described in Sec. III C. The final surface is that below and to the right of the plane, formed by combining the CQ trade-off curve with the inverse of entanglement distribution, as described in Sec. III C.

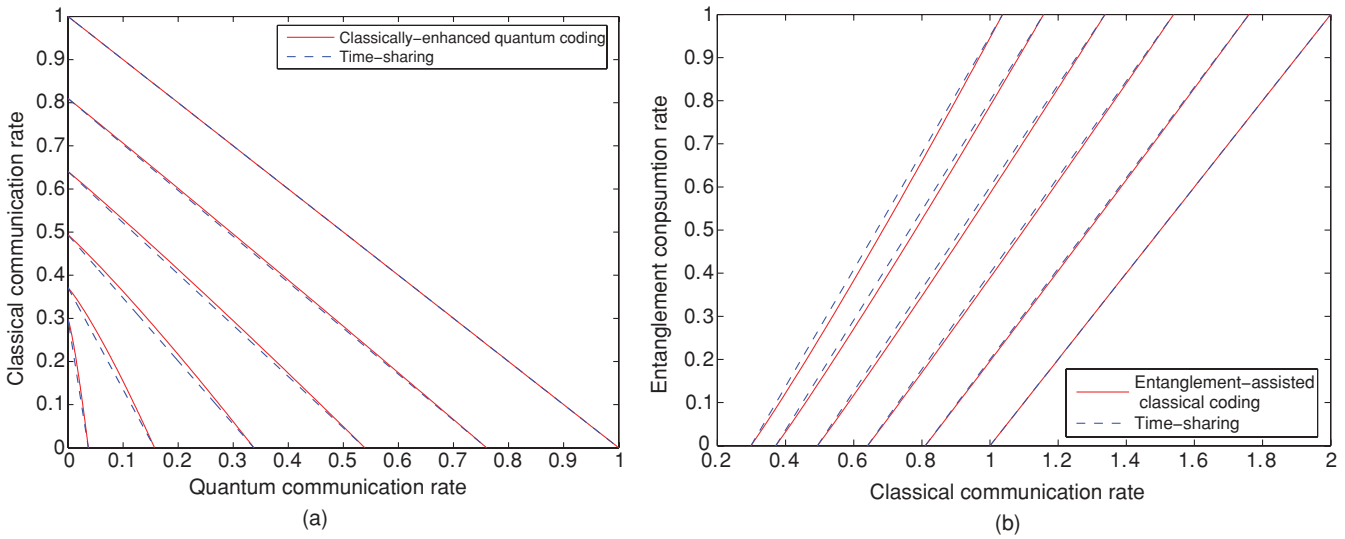



FIG. 4. (Color online) Plot of the (a) CQ trade-off curve and (b) the CE trade-off curve of an Unruh channel for $z = 0$, 0.2, 0.4, 0.6, 0.8, and 0.95. The trade-off curves for $z = 0$ correspond to those for the noiseless qubit channel and are the rightmost trade-off curves in each plot. In both plots, proceeding left from the $z = 0$ curve, we obtain trade-off curves for $z = 0.2$, 0.4, 0.6, 0.8, and 0.95 and notice that they all beat a time-sharing strategy by a larger relative proportion as $z$ increases.

and all points on the CE trade-off curve for an Unruh channel have the following form:

$$(C_{\mathrm{CE}}(z,\mu),\ E_{\mathrm{CE}}(\mu)),$$

where

$$C_{\mathrm{CQ}}(z,\mu) \equiv 1 - \sum_{l=2}^{\infty} p_l(z) \log_2 (l-1) + \sum_{l=2}^{\infty} \frac{p_l(z)}{\Delta_{l-1}} \sum_{i=0}^{l-1} \lambda_i^{(l-1)}(\mu) \log_2 \left[\lambda_i^{(l-1)}(\mu)\right],$$

$$Q_{\mathrm{CQ}}(z,\mu) \equiv -\sum_{l=2}^{\infty} \frac{p_l(z)}{\Delta_{l-1}} \sum_{i=0}^{l-1} \lambda_i^{(l-1)}(\mu) \log_2 \left[\lambda_i^{(l-1)}(\mu)\right] + \sum_{l=2}^{\infty} \frac{p_l(z)}{\Delta_{l-1}} \sum_{i=0}^{l-2} \eta_i^{(l-1)}(\mu) \log_2 \left[\eta_i^{(l-1)}(\mu)\right],$$

$$C_{\mathrm{CE}}(z,\mu) \equiv H_2(\mu) + 1 - \sum_{l=2}^{\infty} p_l(z) \log_2 (l-1) + \sum_{l=2}^{\infty} \frac{p_l(z)}{\Delta_{l-1}} \sum_{i=0}^{l-2} \eta_i^{(l-1)}(\mu) \log_2 \left[\eta_i^{(l-1)}(\mu)\right],$$

$$E_{\mathrm{CE}}(\mu) \equiv H_2(\mu),$$

$$\lambda_i^{(l-1)}(\mu) \equiv (l-1-2i)\mu + i,$$

$$\eta_i^{(l-1)}(\mu) \equiv (l-2-2i)\mu + i + 1,$$

and $H_2$ is the binary entropy function.

Figure 4 plots both the CQ and CE trade-off curves of the Unruh channel for several values of the acceleration parameter $z = 0$, 0.2, 0.4, 0.6, 0.8, 0.95. The figure demonstrates that both classically enhanced quantum coding and entanglement-assisted classical coding beat a time-sharing strategy for an Unruh channel when $z > 0$.

Figure 5 plots the full CQE capacity region for an Unruh channel with acceleration parameter $z = 0.95$. The figure demonstrates that the classically enhanced father protocol,

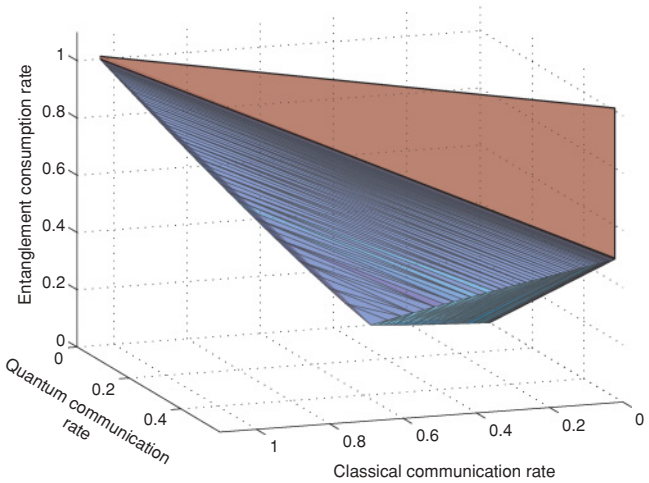



FIG. 5. (Color online) The figure plots the CQE capacity region for an Unruh channel with acceleration parameter $z = 0.95$. It features three distinct surfaces. The first is the flat vertical plane that arises from the bound $R + 2Q \leqslant \mathcal{C}_{\mathrm{EAC}}$, where $\mathcal{C}_{\mathrm{EAC}}$ is the entanglement-assisted classical capacity of an Unruh channel. The plane extends infinitely upward because we can always achieve these rate triples simply by wasting entanglement. The second surface is that below and to the left of the plane, formed by combining the CE trade-off curve with the inverse of superdense coding, as described in Sec. III C. The final surface is that below and to the right of the plane, formed by combining the CQ trade-off curve with the inverse of entanglement distribution, as described in Sec. III C.

combined with entanglement distribution and superdense coding beats a time-sharing strategy because the first two surfaces described in Sec. III C are strictly concave for the Unruh channel when $z > 0$.

## VI. MEASURING THE GAIN OVER A TIME-SHARING STRATEGY

Figures 1, 2, and 4 demonstrate that classically enhanced quantum coding and entanglement-assisted classical coding both beat the time-sharing strategy for the dephasing, cloning, and Unruh channels. The authors of Ref. [31] provided a simple way to compute the benefit of "specialized coding" over the time-sharing strategy, simply by plotting the difference between a trade-off curve and the line of time-sharing as a function of one of the rates in a trade-off scenario.

The previous gain measure illustrates the benefit of specialized coding, but it ignores the relative improvement that specialized coding may give over time-sharing for very noisy channels. As a result, that gain measure tends to zero as one of the capacities tends to zero and thus loses meaning in this asymptotic limit. For example, consider a cloning channel with $N = 1{,}000{,}000$. This channel is particularly noisy for quantum transmission with a low quantum capacity at approximately $1.5 \times 10^{-6}$qubits/channel use, but the classical capacity is approximately 0.27 bits/channel use. Suppose that the sender would like to transmit classical data at a rate of 0.165 bits/channel use to transmit more quantum information. With a time-sharing strategy, the sender can transmit quantum data at approximately the rate $5.9 \times 10^{-7}$qubits/channel use, while classically enhanced quantum coding transmits quantum data at approximately the rate $7.2 \times 10^{-7}$ qubits/channel use. This improvement in transmission appears low in absolute terms, but the relative increase of transmission is substantial, and a measure that captures this relative increase may be more useful for studying the gain.

We suggest an alternate gain measure that highlights the relative improvement and is simple to compute numerically for both the CQ and CE trade-off curves. Let $A_{\mathrm{CQ}}$ denote the area under the CQ trade-off curve and let $A_{\mathrm{TSCQ}}$ denote the area under the line of time sharing. Then the relative gain $G_{\mathrm{CQ}}$ for CQ trading is the ratio of $A_{\mathrm{CQ}}$ to $A_{\mathrm{TSCQ}}$

$$G_{\mathrm{CQ}} \equiv \frac{A_{\mathrm{CQ}}}{A_{\mathrm{TSCQ}}}.$$

The relative gain measure for the CE trade-off curve is similar. Let $A_{\mathrm{CE}}$ denote the area under the CE trade-off curve and let $A_{\mathrm{TSCE}}$ denote the area under the line of time sharing. Then the relative gain $G_{\mathrm{CE}}$ for CE trading is the ratio of $A_{\mathrm{TSCQ}}$ to $A_{\mathrm{CE}}$

$$G_{\mathrm{CE}} \equiv \frac{A_{\mathrm{TSCE}}}{A_{\mathrm{CE}}}.$$

Each of the above relative gains exhibits nontrivial behavior even if one of the capacities vanishes as the noise of a channel increases. These measures are also average gains because the area involves an integration over all points on the trade-off curve. One could generalize these relative gain measures to the CQE capacity region by taking the ratio of the volume

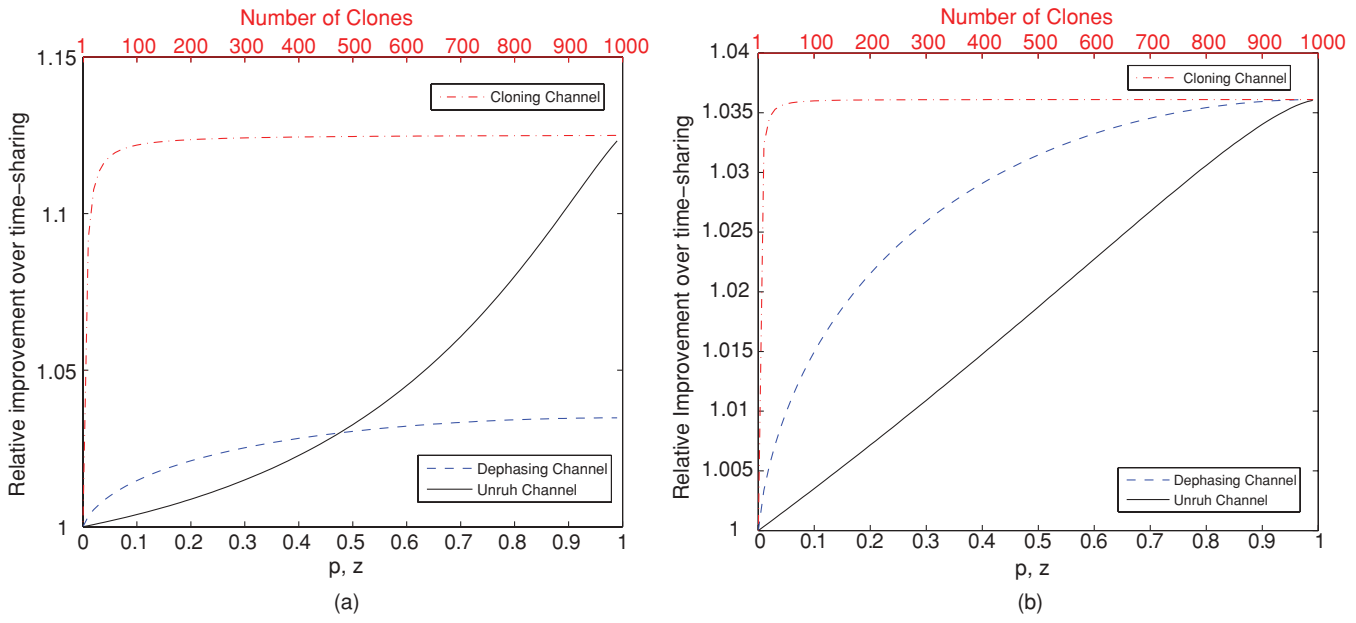



FIG. 6. (Color online) The figures plot (a) the relative gain $G_{\mathrm{CQ}}$ for the CQ trade-off curve and (b) the relative gain $G_{\mathrm{CE}}$ for the CE trade-off curve for the dephasing channel, the cloning channel, and the Unruh channel. The figures plot these relative gains as a function of the dephasing parameter $p \in [0,1]$ for the $p$-dephasing qubit channel (bottom horizontal axis), as a function of the acceleration parameter $z \in [0,1]$ for the Unruh channel (bottom horizontal axis), and as a function of the number of clones $N$ for the $1 \to N$ cloning channel (top horizontal axis). The plot on the left demonstrates that the relative gain $G_{\mathrm{CQ}}$ for the cloning channels is best as $N$ increases, while the Unruh channel features an improved relative gain over a dephasing channel for large accelerations. The plot on the right features different behavior—the relative gain $G_{\mathrm{CE}}$ of the dephasing channel outperforms that for the Unruh channel if we consider the parameters $p$ and $z$ on equal footing, in spite of their drastically different physical interpretations.

enclosed by the bounding surfaces for CQE capacity region to the volume enclosed by surfaces obtained by time sharing.

Figure 6 plots the relative gains $G_{\mathrm{CQ}}$ and $G_{\mathrm{CE}}$ as a function of the dephasing parameter $p$ for the dephasing qubit channel, as a function of the acceleration parameter $z$ for the Unruh channel, and as a function of the number of clones $N$ for the $1 \to N$ cloning channel. The accompanying caption features an interpretation of the results.

One criterion that seems to be necessary to obtain a large relative gain for the CQ trade-off curve is that the quantum capacity of the channel should be much smaller than the classical capacity, so that the area between the inner bounding line of time-sharing and the trade-off curve is relatively large.

## VII. CONCLUSION

We have proven that the CQE capacity region for all Hadamard channels admits a single-letter characterization. Particular examples of the Hadamard channels are generalized dephasing channels and cloning channels, and we have computed exact formulas that specify their CQE capacity regions. Furthermore, we have obtained expressions for the CQE capacity region of an Unruh channel because of its close connection with the cloning channels. The classically enhanced father protocol beats a simple time-sharing strategy for all of these channels, stressing the need for nontrivial coding techniques when trading multiple resources.

It is interesting to ponder the reason why a particular channel obtains an improvement over time sharing. The relative improvements are most significant for the CQ trade off, in which case the cloning and Unruh channels exhibit much more substantial gains than the dephasing channels. In retrospect, it is perhaps surprising that the dephasing channels exhibit any improvement at all. Since these channels can transmit classical data noiselessly, it would be natural to expect that any optimal strategy for sending classical bits would directly exploit this capability. For CQ trade-off coding, that would entail allocating some fraction of channel uses to noiseless classical data transmission and the rest to quantum, which is exactly the time-sharing strategy. The existence of a nontrivial CQ trade off indicates that this strategy is actually not optimal. In contrast, the cloning and Unruh channels are incapable of sending classical data noiselessly when $N > 1$ or $z > 0$ so any communication strategy requires error correction with the attendant opportunity for nontrivial trade-off coding.

Some future directions for this work are in order. It would be desirable to discover other channels for which the full CQE capacity region single-letterizes, but it is likely that the technique for proving single letterization would be completely different. The ideas exploited here are that Hadamard channels are degradable and have entanglement-breaking complementary channels, allowing us to generalize the Devetak-Shor proof technique in Appendix B of Ref. [31]. If other single-letter examples do exist, one should then determine if such a channel obtains an improvement over time sharing and perhaps attempt to uncover a general method that determines if a channel obtains a gain over time sharing.

It may be interesting to explore the static case, where two parties share a bipartite state and exploit this state and some noiseless resources to extract other noiseless resources. One might consider bipartite states that arise from Hadamard channels and attempt to single-letterize the static capacity region. Hsieh and Wilde found formulas for the triple trade-off static capacity region in Ref. [41], but the task of single letterization for the static case is more difficult than that for the dynamic case because one must consider quantum instruments applied to many copies of the bipartite state.

It would also be interesting to consider versions of the Unruh channel other than the original definition in Ref. [35] and determine if the corresponding CQE capacity region can single-letterize (we refer to this channel as "the" Unruh channel, but it may be more appropriate to consider it as "an" Unruh channel). One can consider the CQE capacity region for a qudit Unruh channel by exploiting some of the insights in Ref. [68]. Finally, one might consider a trade-off capacity region for an Unruh channel that includes the resource of private quantum information given that this noiseless resource appears in relativistic quantum information theory, or one could consider trading public classical information, private classical information, and secret key by exploiting the ideas in Refs. [69,70].

## ACKNOWLEDGMENTS

The authors acknowledge useful discussions with Min-Hsiu Hsieh concerning the CQ and CE trade-off curves and the CQE capacity region. K. B. and P. H. acknowledge support from the Office of Naval Research under Grant No. N000140811249. P. H. acknowledges support from the Canada Research Chairs program, CIFAR, FQRNT, MITACS, NSERC, QuantumWorks, and the Sloan Foundation. M. M. W. acknowledges support from the MDEIE (Québec) PSR-SIIRI international collaboration grant.

## APPENDIX A: SINGLE LETTERIZATION OF THE PLANE IN (17)

The following additivity lemma aids in proving the additivity of the plane in (17).

*Lemma 7.* The following additivity relation holds for any two quantum channels $\mathcal{N}_1$ and $\mathcal{N}_2$:

$$h(\mathcal{N}_1 \otimes \mathcal{N}_2) = h(\mathcal{N}_1) + h(\mathcal{N}_2).$$

*Proof.* The inequality $h(\mathcal{N}_1 \otimes \mathcal{N}_2) \geqslant h(\mathcal{N}_1) + h(\mathcal{N}_2)$ trivially holds for all quantum channels because the maximization on the rhs is a restriction of the maximization in the lhs to a tensor product of states of the form in (15). Therefore we prove the nontrivial inequality $h(\mathcal{N}_1 \otimes \mathcal{N}_2) \leqslant h(\mathcal{N}_1) + h(\mathcal{N}_2)$.

Let

$$\psi^{XAA_1A_2} \equiv \sum_x p_X(x)|x\rangle\langle x|^X \otimes |\varphi_x\rangle\langle\varphi_x|^{AA_1A_2},$$

$$\theta^{XAB_1E_1A_2} \equiv U_1\psi U_1^\dagger,$$

$$\omega^{XAA_1B_2E_2} \equiv U_2\psi U_2^\dagger,$$

$$\rho^{XAB_1E_1B_2E_2} \equiv (U_1 \otimes U_2)\psi(U_1 \otimes U_2)^\dagger,$$

where $U_j^{A_j \to B_jE_j}$ is the isometric extension of $\mathcal{N}_j$. Suppose further that $\rho$ is the state that maximizes $h(\mathcal{N}_1 \otimes \mathcal{N}_2)$.

The following chain of inequalities holds for any two channels $\mathcal{N}_1$ and $\mathcal{N}_2$:

$$\begin{aligned}
h(\mathcal{N}_1 \otimes \mathcal{N}_2) &= I(AX; B_1B_2)_\rho \\
&= H(B_1B_2|E_1E_2X)_\rho + H(B_1B_2)_\rho \\
&\leqslant H(B_1|E_1X)_\rho + H(B_1)_\rho + H(B_2|E_2X)_\rho + H(B_2)_\rho, \\
&= H(B_1E_1|X)_\rho - H(E_1|X)_\rho + H(B_1)_\rho + H(B_2E_2|X)_\rho - H(E_2|X)_\rho + H(B_2)_\rho \\
&= H(AA_2|X)_\theta - H(AA_2B_1|X)_\theta + H(B_1)_\theta + H(AA_1|X)_\omega - H(AA_1B_1|X)_\omega + H(B_2)_\omega \\
&= I(AA_2X; B_1)_\theta + I(AA_1X; B_2)_\omega \\
&\leqslant h(\mathcal{N}_1) + h(\mathcal{N}_2).
\end{aligned}$$

The first equality holds because $\rho$ is the state that maximizes $h(\mathcal{N}_1 \otimes \mathcal{N}_2)$. The second equality follows from straightforward entropic manipulations by noting that the state $\rho$ on systems $A$, $B_1$, $B_2$, $E_1$, and $E_2$ is pure when conditioned on $X$. The first inequality follows from an application of strong subadditivity [62] and subadditivity. The next three equalities follow by straightforward entropic manipulations. The last inequality follows from the definition of $h$ in (20) and the fact that both $\theta$ and $\omega$ are states of the form in (15). ■

*Proof of Lemma 1*. The proof follows from Lemma 7 in the same way as Corollary 1 does from Lemma 4. ■

We remark that the above lemma follows from the single letterization of the entanglement-assisted classical capacity [37], but we have included the result for completeness.

## APPENDIX B: PROOF OF THE SINGLE LETTERIZATION OF THE CE TRADE-OFF CURVE FOR HADAMARD CHANNELS

*Lemma 8*. For any fixed $\lambda \geqslant 0$, the function in (23) leads to a point $\mathbf{(}I(AX; B)_\rho, H(A|X)_\rho\mathbf{)}$ on the one-shot CE trade-off curve in the sense of Theorem 3.

*Proof*. We argue by contradiction. Suppose that $\rho^{XAB}$ maximizes (23). Suppose further that it does not lead to a point on the CE trade-off curve. That is, given the constraint $C = I(AX; B)_\rho$, there exists some other state $\sigma^{XAB}$ of the form in (15) such that $I(AX; B)_\sigma = I(AX; B)_\rho = C$, but $E = H(A|X)_\sigma < H(A|X)_\rho$. That is, the state $\sigma$ allows for communication of as much classical information as the state $\rho$ does, but consumes less entanglement. But then the following inequality holds for all $\lambda \geqslant 0$

$$I(AX; B)_\sigma - \lambda H(A|X)_\sigma > I(AX; B)_\rho - \lambda H(A|X)_\rho,$$

contradicting the fact that the state $\rho^{XAB}$ maximizes (23). ■

*Lemma 9*. We get all points on the CE trade-off curve by considering $0 \leqslant \lambda \leqslant 1$ in the maximization in (23) because the maximization optimizes only the classical capacity when $\lambda > 1$.

*Proof*. Consider a state $\rho^{XABE}$ of the form in (15). Suppose that we perform a von Neumann measurement of the system $A$, resulting in a classical variable $Y$, and let $\sigma^{XYBE}$ denote the resulting state. The following chain of inequalities then holds for all $\lambda > 1$:

$$\begin{aligned}
&I(AX; B)_\rho - \lambda H(A|X)_\rho \\
&\quad = H(A|X)_\rho - H(E|X)_\rho + H(B)_\rho - \lambda H(A|X)_\rho \\
&\quad = (1-\lambda)H(BE|X)_\rho - H(E|X)_\rho + H(B)_\rho \\
&\quad = (1-\lambda)H(BE|X)_\sigma - H(E|X)_\sigma + H(B)_\sigma \\
&\quad \leqslant H(B)_\sigma - H(E|X)_\sigma \\
&\quad \leqslant H(B)_\sigma - H(E|XY)_\sigma \\
&\quad = H(B)_\sigma - H(B|XY)_\sigma \\
&\quad = I(XY; B)_\sigma.
\end{aligned}$$

The first equality follows from definitions and from the equality $H(AB|X)_\rho = H(E|X)_\rho$ for the state $\rho^{XABE}$ and the second equality follows from algebra and the fact that $H(BE|X)_\rho = H(A|X)_\rho$. The third equality follows because the von Neumann measurement does not affect the systems in the entropic quantities. The first inequality follows because $\lambda > 1$ and the entropy $H(BE|X)_\rho$ is always positive. The second inequality follows because conditioning reduces entropy. The fourth equality follows because the reduced state of $\sigma^{XYBE}$ on systems $B$ and $E$ is pure when conditioned on both $X$ and $Y$. The last equality follows from the definition of the quantum mutual information.

Thus, it becomes clear that the maximization of the original quantity when $\lambda > 1$ is always less than the classical capacity because $I(XY; B)_\sigma \leqslant \max_\rho I(X; B)$. It then follows that the trade-off curve really occurs for the interval $0 \leqslant \lambda \leqslant 1$. ■

The following lemma is the crucial one that leads to our main result in this section: the single-letterization of the CE trade-off curve for Hadamard channels.

*Lemma 10*. The following additivity relation holds for a Hadamard channel $\mathcal{N}_1$ and any other channel $\mathcal{N}_2$:

$$g_\lambda(\mathcal{N}_1 \otimes \mathcal{N}_2) = g_\lambda(\mathcal{N}_1) + g_\lambda(\mathcal{N}_2).$$

*Proof*. The inequality $g_\lambda(\mathcal{N}_1 \otimes \mathcal{N}_2) \geqslant g_\lambda(\mathcal{N}_1) + g_\lambda(\mathcal{N}_2)$ trivially holds for all quantum channels because the maximization on the rhs is a restriction of the maximization in the lhs to a tensor product of states of the form in (15). Therefore we prove the nontrivial inequality $g_\lambda(\mathcal{N}_1 \otimes \mathcal{N}_2) \leqslant g_\lambda(\mathcal{N}_1) + g_\lambda(\mathcal{N}_2)$ when $\mathcal{N}_1$ is a Hadamard channel.

The channels $\mathcal{N}_1^{A_1\to B_1}$ and $\mathcal{N}_2^{A_2\to B_2}$ and their respective complementary channels $(\mathcal{N}_1^c)^{A_1\to E_1}$ and $(\mathcal{N}_2^c)^{A_2\to E_2}$ have the same properties as they do in the proof of Theorem 1. The state that is the output of the channels is the state $\rho^{XAB_1E_1B_2E_2}$ in (22), but we now define it to be the state that maximizes $g_\lambda(\mathcal{N}_1\otimes\mathcal{N}_2)$. Define $\theta$ as before and $\sigma$ again to be the state after processing system $B_1$ of $\rho$ with $\mathcal{D}_1$.

The following chain of inequalities holds when $\lambda \leqslant 1$:

$$\begin{aligned}
g_\lambda(\mathcal{N}_1\otimes\mathcal{N}_2) &= I(AX;B_1B_2)_\rho - \lambda H(A|X)_\rho \\
&= (1-\lambda)H(B_1B_2E_1E_2|X)_\rho - H(E_1E_2|X)_\rho + H(B_1B_2)_\rho \\
&= (1-\lambda)H(B_1E_1|X)_\rho - H(E_1|X)_\rho + H(B_1)_\rho + (1-\lambda)H(B_2E_2|B_1E_1X)_\rho - H(E_2|E_1X)_\rho + H(B_2|B_1)_\rho \\
&\leqslant (1-\lambda)H(B_1E_1|X)_\rho - H(E_1|X)_\rho + H(B_1)_\rho + (1-\lambda)H(B_2E_2|YX)_\sigma - H(E_2|YX)_\sigma + H(B_2)_\sigma \\
&= (1-\lambda)H(AA_2|X)_\theta - H(AA_2B_1|X)_\theta + H(B_1)_\theta + (1-\lambda)H(AE_1|YX)_\sigma - H(AE_1B_2|YX)_\sigma + H(B_2)_\sigma \\
&= \mathbf{(}I(AA_2X;B_1)_\theta - \lambda H(AA_2|X)_\theta\mathbf{)} + \mathbf{(}I(AE_1XY;B_2)_\sigma - \lambda H(AE_1|XY)_\sigma\mathbf{)} \\
&\leqslant g_\lambda(\mathcal{N}_1) + g_\lambda(\mathcal{N}_2).
\end{aligned}$$

The first equality follows because $\rho$ is the state that maximizes $g_\lambda(\mathcal{N}_1\otimes\mathcal{N}_2)$. The second equality follows from entropic manipulations. The third equality follows from the chain rule. The first and crucial inequality follows from monotonicity of the conditional entropy $H(B_2E_2|B_1E_1X)_\rho$ under the map $\mathcal{D}_1$ and the discarding of system $E_1$, monotonicity of the conditional entropy $H(E_2|E_1X)_\rho$ under the map $\mathcal{D}_2$, and conditioning does not increase entropy. The fourth equality follows because the reduced state of $\theta$ on systems $A$, $A_2$, $B_1$, and $E_1$ is pure when conditioned on $X$, and the reduced state of $\sigma$ on systems $A$, $E_1$, $B_2$, and $E_2$ is pure when conditioned on both $X$ and $Y$. The fifth equality follows from entropic manipulations, and the final inequality follows because $\theta$ and $\sigma$ are both states of the form in (15). ■

*Corollary 2*. The one-shot CE trade-off curve is equal to the regularized CE trade-off curve when the noisy quantum channel $\mathcal{N}$ is a Hadamard channel

$$g_\lambda(\mathcal{N}^{\otimes n}) = n g_\lambda(\mathcal{N}).$$

*Proof*. The proof exploits the same induction technique that Corollary 1 does, but it applies the result of Lemma 10. ■

## APPENDIX C: ON THE PARAMETRIZATION OF THE TRADE-OFF CURVE

*Lemma 11*. $\lambda$ parametrizes all points on the CQ and CE trade-off curves with the possible exception of those lying on segments of constant slope.

*Proof*. We prove the lemma for the case of the CQ trade-off curve. The proof for the CE trade-off curve is similar. Let $[C(t),Q(t)]$ for $0\leqslant t\leqslant 1$ be a parametrization of the trade-off curve with $C(0)$ equal to the classical capacity and $Q(1)$ equal to the quantum capacity. The function $C(t)$ is monotonically decreasing and the function $Q(t)$ is monotonically increasing. The graph of the trade-off curve is convex and therefore has one-sided directional derivatives at all points [71]. It is also monotonically decreasing.

Consider the function $f_\lambda(\mathcal{N})$ where

$$f_\lambda(\mathcal{N}) \equiv \max_{\rho^{XBE}} I(X;B) + \lambda I(A\rangle BX).$$

For any $0\leqslant t\leqslant 1$, we have $f_\lambda(\mathcal{N}) = C(t) + \lambda Q(t)$ if and only if

$$C(t) + \lambda Q(t) \geqslant C(s) + \lambda Q(s). \tag{C1}$$

for all $0\leqslant s\leqslant 1$. Perhaps more instructively, if $s<t$ and $Q(s)\neq Q(t)$, this inequality can be written as

$$\frac{C(s)-C(t)}{Q(s)-Q(t)} \geqslant -\lambda,$$

because of the monotonicity of the functions $C$ and $Q$. Likewise, when $s>t$, it has the form

$$\frac{C(s)-C(t)}{Q(s)-Q(t)} \leqslant -\lambda.$$

If $[C(t),Q(t)]$ is a point on the graph at which the derivative is not constant, then setting $-\lambda$ to be the slope of the graph will lead to Eq. (C1) being satisfied. If the graph is not differentiable at $[C(t),Q(t)]$, then the slope must drop discontinuously at that point. Setting $-\lambda$ to any value in the gap will again lead to the condition being satisfied. ■

At points where the graph is differentiable but the slope is constant, $\lambda$ might not be a good parameter. These points, however, are in the convex hull of the points that $\lambda$ does parametrize.

## APPENDIX D: FORM OF THE CQ TRADE-OFF CURVE FOR QUBIT DEPHASING CHANNELS

We first prove two important lemmas and then prove a theorem that gives the exact form of the CQ trade-off curve.

*Lemma 12*. Let $\mathcal{N}$ be a generalized dephasing channel. In the optimization task for the CQ trade-off curve, it suffices to consider a classical-quantum state with diagonal conditional density operators in the sense that the following inequality holds:

$$I(X;B)_\rho + \lambda I(A\rangle BX)_\rho \leqslant I(X;B)_\theta + \lambda I(A\rangle BX)_\theta,$$

where

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

$U_{\mathcal{N}}^{A'\to BE}$ is an isometric extension of $\mathcal{N}$, $\varphi_x^{A'} = \Delta(\varphi_x^{A'}) = \Delta(\phi_x^{A'})$, and $\Delta$ is the completely dephasing channel.

*Proof.* The proof of this lemma is similar to the proof of Lemma 9 in Ref. [72]. Consider another classical-quantum state $\sigma$ in addition to the two presented in the statement of the theorem

$$\sigma^{XAYE} \equiv \sum_x p_X(x)|x\rangle\langle x|^X \otimes \left(\Delta^{B\to Y} \circ U_{\mathcal{N}}^{A'\to BE}\right)\left(\phi_x^{AA'}\right).$$

Then the following chain of inequalities holds for all $\lambda \geqslant 1$:

$$\begin{aligned} &I(X;B)_\rho + \lambda I(A\rangle BX)_\rho \\ &\quad = H(B)_\rho + (\lambda-1)H(B|X)_\rho - \lambda H(E|X)_\rho \\ &\quad \leqslant H(Y)_\sigma + (\lambda-1)H(Y|X)_\sigma - \lambda H(E|X)_\sigma \\ &\quad = H(B)_\theta + (\lambda-1)H(B|X)_\theta - \lambda H(E|X)_\theta \\ &\quad = I(X;B)_\theta + \lambda I(A\rangle BX)_\theta. \end{aligned}$$

The first equality follows from entropic manipulations. The inequality follows because the entropies $H(B)_\rho$ and $H(B|X)_\rho$ can only increase under a complete dephasing [62]. The second equality follows because $\mathcal{N}\circ\Delta = \Delta\circ\mathcal{N}$ and $\mathcal{N}^c\circ\Delta = \mathcal{N}^c$ for a generalized dephasing channel $\mathcal{N}$, and the final equality follows from entropic manipulations. ∎

*Lemma 13.* An ensemble of the following form parametrizes all points on the CQ trade-off curve for a qubit dephasing channel

$$\tfrac{1}{2}|0\rangle\langle 0|^X \otimes \psi_0^{AA'} + \tfrac{1}{2}|1\rangle\langle 1|^X \otimes \psi_1^{AA'}, \tag{D1}$$

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

$$\mathrm{Tr}_A\left\{\psi_0^{AA'}\right\} = \mu|0\rangle\langle 0|^{A'} + (1-\mu)|1\rangle\langle 1|^{A'}, \tag{D2}$$
$$\mathrm{Tr}_A\left\{\psi_1^{AA'}\right\} = (1-\mu)|0\rangle\langle 0|^{A'} + \mu|1\rangle\langle 1|^{A'}. \tag{D3}$$

*Proof.* We assume without loss of generality that the dephasing basis is the computational basis. Consider a classical-quantum state with a finite number $N$ of diagonal conditional density operators $\rho_x^{A'}$

$$\rho^{XA'} \equiv \sum_{x=0}^{N-1} p_X(x)|x\rangle\langle x|^X \otimes \rho_x^{A'}.$$

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

$$\begin{aligned} \sigma^{XA'} \equiv &\frac{1}{2}\sum_{x=0}^{N-1} p_X(x)|x\rangle\langle x|^X \otimes \rho_x^{A'} \\ &+ \frac{1}{2}\sum_{x=0}^{N-1} p_X(x)|x+N\rangle\langle x+N|^X \otimes X\rho_x^{A'}X, \end{aligned}$$

where $X$ is the $\sigma_X$ "bit-flip" Pauli operator. Consider the following chain of inequalities that holds for all $\lambda \geqslant 1$:

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

The first equality follows by standard entropic manipulations. The second equality follows because the conditional entropy $H(B|X)$ is invariant under a bit-flipping unitary on the input state that commutes with the channel $H(B)_{X\rho_x^B X} = H(B)_{\rho_x^B}$. Furthermore, a bit flip on the input state does not change the eigenvalues for the output of the dephasing channel's complementary channel (as observed in Sec. II C 3) $H(E)_{\mathcal{N}^c(X\rho_x^{A'}X)} = H(E)_{\mathcal{N}^c(\rho_x^{A'})}$. The first inequality follows because entropy is concave (i.e., the local state $\sigma^B$ is a mixed version of $\rho^B$). The third equality follows because $H(B)_{\sigma^B} = H[\sum_x \frac{1}{2}p_X(x)(\rho_x^B + X\rho_x^B X)] = H[\frac{1}{2}\sum_x p_X(x)I] = 1$. The fourth equality follows because the system $X$ is classical. The second inequality follows because the maximum value of a realization of a random variable is not less than its expectation. The final equality simply follows by defining $\rho_x^*$ to be the conditional density operator on systems $B$ and $E$ that arises from sending a diagonal state of the form $\mu|0\rangle\langle 0|^{A'} + (1-\mu)|1\rangle\langle 1|^{A'}$ through the channel. Thus, an ensemble of the kind in (D1) is sufficient to attain a point on the CQ trade-off curve.

In the last step above, we observe that there is a direct correspondence between $\mu$ that parametrizes the ensemble and $\lambda$ that parametrizes a point on the CQ trade-off curve. ∎

*Proof of CQ trade off in Theorem 4.* We simply need to compute the Holevo information $I(X;B)$ and the coherent information $I(A\rangle BX)$ for an ensemble of the form in the statement of Lemma 13, due to the results of Lemmas 12 and 13.

First consider respective purifications of the states in (D2)–(D3)

$$|\psi_0\rangle^{AA'} = \sqrt{\mu}\,|0\rangle^A|0\rangle^{A'} + \sqrt{1-\mu}\,|1\rangle^A|1\rangle^{A'}, \tag{D4}$$
$$|\psi_1\rangle^{AA'} = \sqrt{1-\mu}\,|0\rangle^A|0\rangle^{A'} + \sqrt{\mu}\,|1\rangle^A|1\rangle^{A'}. \tag{D5}$$

The above states lead to a classical-quantum state of the form in (D1). An isometric extension $U_{\mathcal{N}} = \sqrt{1-\frac{p}{2}}I\otimes|0\rangle^E + \sqrt{\frac{p}{2}}Z\otimes|1\rangle^E$ of the qubit dephasing channel acts as follows on the above states:

$$\begin{aligned} |\psi_0\rangle^{ABE} &\equiv U_{\mathcal{N}}|\psi_0\rangle^{AA'} \\ &= \sqrt{\mu}\,|0\rangle^A|0\rangle^B\left(\sqrt{1-\frac{p}{2}}\,|0\rangle^E + \sqrt{\frac{p}{2}}\,|1\rangle^E\right) \\ &\quad + \sqrt{1-\mu}\,|1\rangle^A|1\rangle^B\left(\sqrt{1-\frac{p}{2}}\,|0\rangle^E - \sqrt{\frac{p}{2}}\,|1\rangle^E\right), \end{aligned}$$

$$
\begin{aligned}
|\psi_1\rangle^{ABE} &\equiv U_{\mathcal{N}} |\psi_1\rangle^{AA'} \\
&= \sqrt{1-\mu}\,|0\rangle^A |0\rangle^B \left(\sqrt{1-\frac{p}{2}}\,|0\rangle^E + \sqrt{\frac{p}{2}}\,|1\rangle^E\right) \\
&\quad + \sqrt{\mu}\,|1\rangle^A |1\rangle^B \left(\sqrt{1-\frac{p}{2}}\,|0\rangle^E - \sqrt{\frac{p}{2}}\,|1\rangle^E\right).
\end{aligned}
$$

The classical-quantum state at the output of the channel is as follows:

$$
\begin{aligned}
\rho^{XABE} \equiv \tfrac{1}{2}[&|0\rangle\langle 0|^X \otimes |\psi_0\rangle\langle\psi_0|^{ABE} \\
&+ |1\rangle\langle 1|^X \otimes |\psi_1\rangle\langle\psi_1|^{ABE}]. \qquad \text{(D6)}
\end{aligned}
$$

The following states are useful for computing the entropies $H(B)$ and $H(B|X)$:

$$
\begin{aligned}
\rho^{XB} &= \tfrac{1}{2}\left[|0\rangle\langle 0|^X \otimes \psi_0^B + |1\rangle\langle 1|^X \otimes \psi_1^B\right], \\
\psi_0^B &= \mu|0\rangle\langle 0|^B + (1-\mu)|1\rangle\langle 1|^B, \\
\psi_1^B &= (1-\mu)|0\rangle\langle 0|^B + (\mu)|1\rangle\langle 1|^B, \\
\rho^B &= \tfrac{1}{2}[|0\rangle\langle 0|^B + |1\rangle\langle 1|^B].
\end{aligned}
$$

The Holevo information is then as follows:

$$
I(X;B) = H(B) - H(B|X) = 1 - H_2(\mu).
$$

We now compute the coherent information $I(A\rangle BX) = H(B|X) - H(E|X)$. The following states are important in this computation:

$$
\rho^{XE} = \tfrac{1}{2}\left[|0\rangle\langle 0|^X \otimes \psi_0^E + |1\rangle\langle 1|^X \otimes \psi_1^E\right],
$$

$$
\begin{aligned}
\psi_0^E &= \left(1-\frac{p}{2}\right)|0\rangle\langle 0|^E + \frac{p}{2}|1\rangle\langle 1|^E \\
&\quad + \sqrt{1-\frac{p}{2}}\sqrt{\frac{p}{2}}(2\mu-1)(|0\rangle\langle 1|^E + |1\rangle\langle 0|^E),
\end{aligned}
$$

$$
\begin{aligned}
\psi_1^E &= \left(1-\frac{p}{2}\right)|0\rangle\langle 0|^E + \frac{p}{2}|1\rangle\langle 1|^E \\
&\quad - \sqrt{1-\frac{p}{2}}\sqrt{\frac{p}{2}}(2\mu-1)(|0\rangle\langle 1|^E + |1\rangle\langle 0|^E).
\end{aligned}
$$

We compute the determinants of the density operators $\psi_0^E$ and $\psi_1^E$

$$
\begin{aligned}
\mathrm{Det}\left(\psi_0^E\right) &= \mathrm{Det}\left(\psi_1^E\right) \\
&= \left(1-\frac{p}{2}\right)\frac{p}{2}[1-(2\mu-1)^2] \\
&= 2p\mu\left(1-\frac{p}{2}\right)(1-\mu).
\end{aligned}
$$

These determinants lead to the same eigenvalues for both $\psi_0^E$ and $\psi_1^E$

$$
\begin{aligned}
\lambda_\pm &\equiv \frac{1}{2} \pm \sqrt{\frac{1}{4} - \mathrm{Det}\left(\psi_o^E\right)} \\
&= \frac{1}{2} \pm \frac{1}{2}\sqrt{1 - 16 \times \frac{p}{2}\left(1-\frac{p}{2}\right)\mu(1-\mu)}. \qquad \text{(D7)}
\end{aligned}
$$

Thus, the coherent information is as stated in the theorem $I(A\rangle BX) = H_2(\mu) - H_2(\lambda_+)$. ∎

## APPENDIX E: FORM OF THE CE TRADE-OFF CURVE FOR QUBIT DEPHASING CHANNELS

We first prove two important lemmas and then prove a theorem that gives the exact form of the CE trade-off curve.

*Lemma 14.* Let $\mathcal{N}$ be a generalized dephasing channel. In the optimization task for the CE trade-off curve, it suffices to consider a classical-quantum state with diagonal conditional density operators, in the sense that the following inequality holds when $0 \leqslant \lambda \leqslant 1$:

$$
I(AX;B)_\rho - \lambda H(A|X)_\rho \leqslant I(AX;B)_\theta - \lambda H(A|X)_\theta,
$$

where

$$
\begin{aligned}
\rho^{XABE} &\equiv \sum_x p_X(x)|x\rangle\langle x|^X \otimes U_{\mathcal{N}}^{A'\to BE}\left(\phi_x^{AA'}\right), \\
\theta^{XABE} &\equiv \sum_x p_X(x)|x\rangle\langle x|^X \otimes U_{\mathcal{N}}^{A'\to BE}\left(\varphi_x^{AA'}\right),
\end{aligned}
$$

$U_{\mathcal{N}}^{A'\to BE}$ is an isometric extension of $\mathcal{N}$, $\varphi_x^{A'} = \Delta(\varphi_x^{A'}) = \Delta(\phi_x^{A'})$, and $\Delta$ is the completely dephasing channel.

*Proof.* The proof of this lemma is similar to the proof of Lemma 9 in Ref. [72]. Consider another classical-quantum state $\sigma$ in addition to the two presented in the statement of the theorem

$$
\sigma^{XY_AY_BE} \equiv \sum_x p_X(x)|x\rangle\langle x|^X \otimes \Delta^{B\to Y}\left[U_{\mathcal{N}}^{A'\to BE}\left(\phi_x^{AA'}\right)\right].
$$

Then the following chain of inequalities holds when $0 \leqslant \lambda \leqslant 1$:

$$
\begin{aligned}
&I(AX;B)_\rho - \lambda H(A|X)_\rho \\
&\quad = (1-\lambda)H(A|X)_\rho + H(B)_\rho - H(E|X)_\rho \\
&\quad = (1-\lambda)\sum_x p_X(x)H(A')_{\phi_x} + H(B)_\rho - H(E|X)_\rho \\
&\quad \leqslant (1-\lambda)\sum_x p_X(x)H(A')_{\Delta(\phi_x^{A'})} + H(Y)_\sigma - H(E|X)_\sigma \\
&\quad = (1-\lambda)H(A|X)_\theta + H(B)_\theta - H(E|X)_\theta \\
&\quad = I(AX;B)_\theta - \lambda H(A|X)_\theta.
\end{aligned}
$$

The first equality follows from entropic manipulations. The second equality follows because the system $X$ is classical and the states $\phi_x^{AA'}$ are pure. The inequality follows because the entropies $H(A')$ and $H(B)_\rho$ can only increase under a complete dephasing. The third equality follows because $\mathcal{N} \circ \Delta = \Delta \circ \mathcal{N}$ and $\mathcal{N}^c \circ \Delta = \mathcal{N}^c$ for a generalized dephasing channel $\mathcal{N}$, and the final equality follows from entropic manipulations. ∎

*Lemma* 15. An ensemble of the following form parametrizes all points on the CE trade-off curve for a qubit dephasing channel:

$$
\tfrac{1}{2}\left(|0\rangle\langle 0|^X \otimes \psi_0^{AA'} + |1\rangle\langle 1|^X \otimes \psi_1^{AA'}\right), \qquad \text{(E1)}
$$

where the states $\psi_0$ and $\psi_1$ are the same as they are in the statement of Lemma 13.

*Proof.* The proof proceeds similarly to the proof of Lemma 13, with the same definitions of states $\rho$ and $\sigma$, but with the

following different chain of inequalities for $0 \leqslant \lambda \leqslant 1$:

$$\begin{aligned}
&I(AX;B)_\rho - \lambda H(A|X)_\rho \\
&\quad = (1-\lambda)H(A|X)_\rho + H(B)_\rho - H(E|X)_\rho \\
&\quad = (1-\lambda)H(A|X)_\sigma + H(B)_\rho - H(E|X)_\sigma \\
&\quad \leqslant (1-\lambda)H(A|X)_\sigma + H(B)_\sigma - H(E|X)_\sigma \\
&\quad = (1-\lambda)H(A|X)_\sigma + 1 - H(E|X)_\sigma \\
&\quad = 1 + \sum_x p_X(x)[(1-\lambda)H(A)_{\psi_x} - H(E)_{\psi_x}] \\
&\quad \leqslant 1 + \max_x[(1-\lambda)H(A)_{\psi_x} - H(E)_{\psi_x}] \\
&\quad = 1 + (1-\lambda)H(A)_{\psi_x^*} - H(E)_{\psi_x^*}.
\end{aligned}$$

We do not provide justifications for the above steps because they are identical those in the proof of Lemma 13. ∎

*Proof of CE trade off in Theorem 4.* The proof follows by noting that $I(AX;B) = H(A|X) + H(B) - H(E|X)$, $H(A|X) = H_2(\mu)$, and that we have already computed $H(B)$ and $H(E|X)$ in the first part of the proof of Theorem 4. ∎

## APPENDIX F: CQ AND CE TRADE-OFF CURVES FOR THE UNRUH CHANNEL

*Proof.* The input state in (24) that traces out the CQ curve for the cloning channels also does so for the Unruh channel. We consider the purification of this state, so that the output state is as follows:

$$\rho^{XABE} \equiv \tfrac{1}{2}\left[|0\rangle\langle 0|^X \otimes {\psi_0}^{ABE} + |1\rangle\langle 1|^X \otimes \psi_1^{ABE}\right].$$

Let $\mathcal{N}(|0\rangle\langle 0|)$ and $\mathcal{N}(|1\rangle\langle 1|)$ denote the Unruh channel outputs corresponding to the respective input states $|0\rangle\langle 0|^{A'}$ and $|1\rangle\langle 1|^{A'}$

$$\begin{aligned}
\mathcal{N}(|0\rangle\langle 0|) &\equiv \bigoplus_{l=2}^{\infty} p_l(z) S_l(|0\rangle\langle 0|) \\
&= \bigoplus_{l=2}^{\infty} \frac{p_l(z)}{\Delta_{l-1}} \sum_{i=0}^{l-2} (l-1-i)|i\rangle\langle i|^B, \\
\mathcal{N}(|1\rangle\langle 1|) &\equiv \bigoplus_{l=2}^{\infty} p_l(z) S_l(|1\rangle\langle 1|) \\
&= \bigoplus_{l=2}^{\infty} \frac{p_l(z)}{\Delta_{l-1}} \sum_{i=0}^{l-2} (i+1)|i+1\rangle\langle i+1|^B.
\end{aligned}$$

The following states are useful for calculating the Holevo information:

$$\begin{aligned}
\rho^{XB} &= \tfrac{1}{2}\left[|0\rangle\langle 0|^X \otimes \psi_0^B + |1\rangle\langle 1|^X \otimes \psi_1^B\right], \\
\psi_0^B &= \mu\mathcal{N}(|0\rangle\langle 0|) + (1-\mu)\mathcal{N}(|1\rangle\langle 1|) \\
&= \bigoplus_{l=2}^{\infty} \frac{p_l(z)}{\Delta_{l-1}} \sum_{i=0}^{l-1} \lambda_i^{l-1}(\mu)|i\rangle\langle i|^B, \\
\psi_1^B &= (1-\mu)\mathcal{N}(|0\rangle\langle 0|) + \mu\mathcal{N}(|1\rangle\langle 1|) \\
&= \bigoplus_{l=2}^{\infty} \frac{p_l(z)}{\Delta_{l-1}} \sum_{i=0}^{l-1} \lambda_i^{l-1}(1-\mu)|i\rangle\langle i|^B,
\end{aligned}$$

$$\begin{aligned}
\rho^B &= \frac{1}{2}[\mathcal{N}(|0\rangle\langle 0|) + \mathcal{N}(|1\rangle\langle 1|)] \\
&= \bigoplus_{l=2}^{\infty} \frac{p_l(z)}{l} \sum_{i=0}^{l-1} |i\rangle\langle i|^B,
\end{aligned}$$

where $\lambda_i^N(\mu) \equiv (N-2i)\mu + i$. We then compute the entropies $H(B)$ and $H(B|X)$

$$\begin{aligned}
H(B)_\rho &= -\sum_{l=2}^{\infty}\sum_{i=0}^{l-1} \frac{p_l(z)}{l} \log_2\left(\frac{p_l(z)}{l}\right) \\
&= -\sum_{l=2}^{\infty} p_l(z) \log_2\left(\frac{p_l(z)}{l}\right), \\
H(B|X)_\rho &= \tfrac{1}{2}[H(B)_{\psi_0} + H(B)_{\psi_1}] \\
&= -\sum_{l=2}^{\infty}\sum_{i=0}^{l-1} \frac{p_l(z)\lambda_i^{l-1}(\mu)}{\Delta_{l-1}} \log_2\left[\frac{p_l(z)\lambda_i^{l-1}(\mu)}{\Delta_{l-1}}\right].
\end{aligned} \tag{F1}$$

Observe that the following relationships hold:

$$\begin{aligned}
\sum_{i=0}^{l-1} \lambda_i^{(l-1)}(\mu) &= \Delta_{l-1}, \\
\sum_{l=2}^{\infty} p_l(z) &= 1.
\end{aligned}$$

These relationships allow us to rewrite the expression in (F1) for $H(B|X)$ as follows:

$$\begin{aligned}
H(B|X) = H(B) - 1 &+ \sum_{l=2}^{\infty} p_l(z) \log_2(l-1) \\
&- \sum_{l=2}^{\infty}\sum_{i=0}^{l-1} \frac{p_l(z)}{\Delta_{l-1}} \lambda_i^{(l-1)}(\mu) \log_2\left[\lambda_i^{(l-1)}(\mu)\right].
\end{aligned} \tag{F2}$$

and we get the Holevo information

$$\begin{aligned}
I(X;B) = 1 &- \sum_{l=2}^{\infty} p_l(z) \log_2(l-1) \\
&+ \sum_{l=2}^{\infty} \frac{p_l(z)}{\Delta_{l-1}} \sum_{i=0}^{l-1} \lambda_i^{(l-1)}(\mu) \log_2\left[\lambda_i^{(l-1)}(\mu)\right].
\end{aligned}$$

The Holevo information $I(X;B)$ coincides with the expression for the classical capacity of an Unruh channel when $\mu = 0$ (Corollary 3, Eq. (18) in Ref. [30], though note the slightly different definition of $\Delta_l$ in that paper)

$$I(X;B)_{\mu=0} = 1 - \sum_{l=2}^{\infty} p_l(z) \log_2(l-1) + \sum_{l=2}^{\infty} \frac{p_l(z)}{\Delta_{l-1}} \sum_{i=0}^{l-1} i \log_2 i.$$

The Holevo information vanishes when $\mu = 1/2$.

We now compute the coherent information. Let $\mathcal{N}^c(|0\rangle\langle 0|)$ and $\mathcal{N}^c(|1\rangle\langle 1|)$ denote the outputs of the complementary

channel of the Unruh channel corresponding to the respective input states $|0\rangle\langle 0|^{A'}$ and $|1\rangle\langle 1|^{A'}$

$$\begin{aligned}\mathcal{N}^c(|0\rangle\langle 0|) &\equiv \bigoplus_{l=2}^{\infty} p_l(z) S_l^c(|0\rangle\langle 0|)\\ &= \bigoplus_{l=2}^{\infty} \frac{p_l(z)}{\Delta_{l-1}} \sum_{i=0}^{l-2} (l-1-i)|i\rangle\langle i|^E\\ \mathcal{N}^c(|1\rangle\langle 1|) &\equiv \bigoplus_{l=2}^{\infty} p_l(z) S_l^c(|1\rangle\langle 1|)\\ &= \bigoplus_{l=2}^{\infty} \frac{p_l(z)}{\Delta_{l-1}} \sum_{i=0}^{l-2} (i+1)|i\rangle\langle i|^E.\end{aligned}$$

The following states are important in this calculation:

$$\begin{aligned}\rho^{XE} &= \tfrac{1}{2}\left[|0\rangle\langle 0|^X \otimes \psi_0^E + |1\rangle\langle 1|^X \otimes \psi_1^E\right],\\ \psi_0^E &= \mu\mathcal{N}^c(|0\rangle\langle 0|) + (1-\mu)\mathcal{N}^c(|1\rangle\langle 1|)\\ &= \bigoplus_{l=2}^{\infty} \frac{p_l(z)}{\Delta_{l-1}} \sum_{i=0}^{l-2} \eta_i^{(l-1)}(\mu)|i\rangle\langle i|^E,\\ \psi_1^E &= (1-\mu)\mathcal{N}^c(|0\rangle\langle 0|) + \mu\mathcal{N}^c(|1\rangle\langle 1|)\\ &= \bigoplus_{l=2}^{\infty} \frac{p_l(z)}{\Delta_{l-1}} \sum_{i=0}^{l-2} \eta_i^{(l-1)}(1-\mu)|i\rangle\langle i|^E,\end{aligned}$$

where $\eta_i^{(l-1)}(\mu) \equiv (l-2-2i)\mu + i + 1$. Then the conditional entropy $H(E|X)$ is as follows:

$$H(E|X) = -\sum_{l=2}^{\infty} \frac{p_l(z)}{\Delta_{l-1}} \sum_{i=0}^{l-2} \eta_i^{(l-1)(\mu)} \log_2 \left[\frac{p_l(z)\eta_i^{(l-1)}(\mu)}{\Delta_{l-1}}\right].$$

The relation $\sum_{i=0}^{l-2} \eta_i^{(l-1)}(\mu) = \Delta_{l-1}$ allows us to simplify the above expression for the conditional entropy $H(E|X)$

$$\begin{aligned}H(E|X) = H(B) - 1 + \sum_{l=2}^{\infty} p_l(z)\log_2(l-1)\\ - \sum_{l=2}^{\infty}\sum_{i=0}^{l-2} \frac{p_l(z)}{\Delta_{l-1}} \eta_i^{(l-1)}(\mu)\log_2\left[\eta_i^{(l-1)}(\mu)\right]. \quad \text{(F3)}\end{aligned}$$

We finally obtain the coherent information as the difference of (F2) and (F3)

$$\begin{aligned}I(A\rangle BX) = -\sum_{l=2}^{\infty} \frac{p_l(z)}{\Delta_{l-1}} \sum_{i=0}^{l-1} \lambda_i^{(l-1)}(\mu)\log_2\left[\lambda_i^{(l-1)}(\mu)\right]\\ + \sum_{l=2}^{\infty} \frac{p_l(z)}{\Delta_{l-1}} \sum_{i=0}^{l-2} \eta_i^{(l-1)}(\mu)\log_2\left[\eta_i^{(l-1)}(\mu)\right].\end{aligned}$$

The above expression vanishes when $\mu = 0$, and it coincides with the expression for the quantum capacity of the Unruh channel when $\mu = 1/2$ (in Sec. III B of Ref. [33])

$$I(A\rangle BX)_{\mu=1/2} = \sum_{k=0}^{\infty} p_{k+2}(z)\log_2\left(\frac{k+2}{k+1}\right).$$

We now trace out the CE trade-off curve for a single use of the Unruh channel. We use the same input state as in Lemma 6 because that lemma proves that this input state traces out both the CQ curve and the CE curve. We then have here that $H(A|X) = H_2(\mu)$, and we obtain the expression in the statement of the theorem by noting that $I(AX;B) = H(A|X) + H(B) - H(E|X)$. ∎


[1] A. S. Holevo, IEEE Trans. Inf. Theory **44**, 269 (1998).
[2] B. Schumacher and M. D. Westmoreland, Phys. Rev. A **56**, 131 (1997).
[3] S. Lloyd, Phys. Rev. A **55**, 1613 (1997).
[4] P. W. Shor, "The quantum channel capacity and coherent information," Lecture Notes, MSRI Workshop on Quantum Computation, San Francisco, 2002 (unpublished); available at [http://www.msri.org/publications/ln/msri/2002/quantumcrypto/shor/1].
[5] I. Devetak, IEEE Trans. Inf. Theory **51**, 44 (2005).
[6] I. Devetak and A. Winter, Phys. Rev. Lett. **93**, 080501 (2004).
[7] P. Hayden, M. Horodecki, A. Winter, and J. Yard, Open Syst. Inf. Dyn. **15**, 7 (2008).
[8] R. Klesse, Open Syst. Inf. Dyn. **15**, 21 (2008).
[9] M. Horodecki, S. Lloyd, and A. Winter, Open Syst. Inf. Dyn. **15**, 47 (2008).
[10] P. Hayden, P. W. Shor, and A. Winter, Open Syst. Inf. Dyn. **15**, 71 (2008).
[11] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996).
[12] H. Barnum, E. Knill, and M. A. Nielsen, IEEE Trans. Inf. Theory **46**, 1317 (2000).
[13] C. E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948).
[14] M. B. Hastings, Nature Phys. **5**, 255 (2009).
[15] P. Hayden and A. Winter, Commun. Math. Phys. **284**, 263 (2008).
[16] M. Fukuda, C. King, and D. K. Moser, Commun. Math. Phys. **296**, 111 (2010).
[17] F. G. S. L. Brandao and M. Horodecki, Open Syst. Inf. Dyn. **17**, 31 (2010).
[18] D. P. DiVincenzo, P. W. Shor, and J. A. Smolin, Phys. Rev. A **57**, 830 (1998).
[19] G. Smith and J. A. Smolin, Phys. Rev. Lett. **98**, 030501 (2007).
[20] P. W. Shor, Current Developments in Mathematics **2005**, 173 (2007).
[21] P. W. Shor, Commun. Math. Phys. **246**, 453 (2004).
[22] G. G. Amosov, A. S. Holevo, and R. F. Werner, Problems of Information Transmission **36**, 25 (2000).
[23] C. King, J. Math. Phys. **43**, 4641 (2002).
[24] C. H. Bennett, D. P. DiVincenzo, and J. A. Smolin, Phys. Rev. Lett. **78**, 3217 (1997).
[25] C. King, K. Matsumoto, M. Nathanson, and M. B. Ruskai, Markov Processes and Related Fields **13**, 391 (2007), J. T. Lewis memorial issue.

[26] P. W. Shor, J. Math. Phys. **43**, 4334 (2002).
[27] C. King, IEEE Trans. Inf. Theory **49**, 221 (2003).
[28] N. Datta, A. S. Holevo, and Y. Suhov, Int. J. Quantum. Inform. **4**, 85 (2006).
[29] M. Fukuda, J. Phys. A, **38**, L753 (2005).
[30] K. Brádler, e-print arXiv:0903.1638.
[31] I. Devetak and P. W. Shor, Commun. Math. Phys. **256**, 287 (2005).
[32] T. S. Cubitt, M. B. Ruskai, and G. Smith, J. Math. Phys. **49**, 102104 (2008).
[33] K. Brádler, N. Dutil, P. Hayden, and A. Muhammad, e-print arXiv:0909.3297.
[34] V. Giovannetti and R. Fazio, Phys. Rev. A **71**, 032314 (2005).
[35] K. Brádler, P. Hayden, and P. Panangaden, J. High Energy Phys. 08 (2009) 074.
[36] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, Phys. Rev. Lett. **83**, 3081 (1999).
[37] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, IEEE Trans. Inf. Theory **48**, 2637 (2002).
[38] C. H. Bennett and S. J. Wiesner, Phys. Rev. Lett. **69**, 2881 (1992).
[39] P. W. Shor, in *Quantum Information, Statistics, Probability*, edited by O. Hirota (Rinton Press, Princeton, NJ, 2004), pp. 144–152, e-print arXiv:quant-ph/0402129.
[40] M.-H. Hsieh and M. M. Wilde, IEEE Trans. Inf. Theory, (2010), e-print arXiv:0811.4227.
[41] M.-H. Hsieh and M. M. Wilde, e-print arXiv:0901.3038.
[42] I. Kremsky, M.-H. Hsieh, and T. A. Brun, Phys. Rev. A **78**, 012341 (2008).
[43] M. M. Wilde and T. A. Brun, in *Proceedings of the IEEE International Symposium on Information Theory, Toronto, Ontario, Canada, July 2008* (IEEE, New York, 2008), pp. 359–363.
[44] I. Devetak, A. W. Harrow, and A. Winter, IEEE Trans. Inf. Theory **54**, 4587 (2008).
[45] I. Devetak, A. W. Harrow, and A. Winter, Phys. Rev. Lett. **93**, 230504 (2004).
[46] H. Barnum, P. Hayden, R. Jozsa, and A. Winter, Proceedings of the Royal Society A **457**, 2019 (2001).
[47] P. Hayden, R. Jozsa, and A. Winter, J. Math. Phys. **43**, 4404 (2002).
[48] M. Koashi and N. Imoto, e-print arXiv:quant-ph/0104001.
[49] C. H. Bennett, P. Hayden, D. W. Leung, P. W. Shor, and A. Winter, IEEE Trans. Inf. Theory **51**, 56 (2005).
[50] A. Abeyesinghe and P. Hayden, Phys. Rev. A **68**, 062319 (2003).
[51] G. Kuperberg, IEEE Trans. Inf. Theory **49**, 1465 (2003).
[52] M. M. Wilde and M.-Hsiu. Hsieh, e-print arXiv:1004.0458.
[53] T. A. Brun, I. Devetak, and M.-H. Hsieh, Science **314**, 436 (2006).
[54] I. Devetak, T. A. Brun, and M.-H. Hsieh, *New Trends in Mathematical Physics* (Springer, The Netherlands, 2009), pp. 161–172.
[55] M.-H. Hsieh, I. Devetak, and T. Brun, Phys. Rev. A **76**, 062313 (2007).
[56] M. M. Wilde and T. A. Brun, Phys. Rev. A **81**, 042333 (2010).
[57] F. Brito, D. P. DiVincenzo, R. H. Koch, and M. Steffen, New J. Phys. **10**, 033027 (2008).
[58] P. W. Milonni and M. L. Hardies, Phys. Lett. A **92**, 321 (1982).
[59] A. Lamas-Linares, C. Simon, J. C. Howell, and D. Bouwmeester, Science **296**, 712 (2002).
[60] C. Simon, G. Weihs, and A. Zeilinger, Phys. Rev. Lett. **84**, 2993 (2000).
[61] W. G. Unruh, Phys. Rev. D **14**, 870 (1976).
[62] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).
[63] P. Bergmans, IEEE Trans. Inf. Theory **19**, 197 (1973).
[64] M. Horodecki, P. W. Shor, and M. Beth. Ruskai, Rev. Math. Phys. **15**, 629 (2003).
[65] W. K. Wootters and W. H. Zurek, Nature (London) **299**, 802 (1982).
[66] N. Gisin and S. Massar, Phys. Rev. Lett. **79**, 2153 (1997).
[67] A. S. Holevo, e-print arXiv:quant-ph/0212025.
[68] K. Brádler, P. Hayden, and P. Panangaden (unpublished).
[69] M.-H. Hsieh, Z. Luo, and T. Brun, Phys. Rev. A **78**, 042306 (2008).
[70] M.-H. Hsieh and M. M. Wilde, Phys. Rev. A **80**, 022306 (2009).
[71] A. W. Roberts and D. E. Varberg, *Convex functions*, volume 57 of *Pure and applied mathematics* (Academic Press, New York, 1973).
[72] M.-H. Hsieh, I. Devetak, and A. Winter, IEEE Trans. Inf. Theory **54**, 3078 (2008).