Quantum Channel Capacities Per Unit Cost

Dawei Ding∗ Dmitri S. Pavlichin† Mark M. Wilde‡

June 14, 2017

Abstract

Communication over a noisy channel is often conducted in a setting in which different input symbols to the channel incur a certain cost. For example, for the additive white Gaussian noise channel, the cost associated with a real number input symbol is the square of its magnitude. In such a setting, it is often useful to know the maximum amount of information that can be reliably transmitted per cost incurred. This is known as the capacity per unit cost. In this paper, we generalize the capacity per unit cost to various communication tasks involving a quantum channel; in particular, we consider classical communication, entanglement-assisted classical communication, private communication, and quantum communication. For each task, we define the corresponding capacity per unit cost and derive a formula for it analogous to that of the usual capacity. Furthermore, for the special case in which there is a zero-cost quantum state, we obtain expressions for the various capacities per unit cost in terms of an optimized relative entropy involving the zero-cost state. For each communication task, we construct an explicit pulse-position-modulation coding scheme that achieves the capacity per unit cost. Finally, we compute capacities per unit cost for various quantum Gaussian channels and introduce the notion of a blocklength-constrained capacity per unit cost.

1 Introduction

The main concerns of information theory are determining limitations on information processing and how to attain them [CT06]. In the task of communication over a noisy channel, for example, the usual goal is to compute the capacity of the channel, which is informally defined as the maximum number of bits that can be reliably transmitted over the channel divided by the total number of channel uses. As the number of channel uses is often directly proportional to the overall transmission time, the capacity measures the maximum rate of information transmission per unit time. Hence, the capacity is a limit to communication when given a certain time constraint.

However, we need not restrict ourselves to time constraints. We can seek limits to communication with respect to other types of constraints, and in some practical settings these other constraints are more relevant. For example, we can imagine a setting in which a satellite deep in space is transmitting information back to Earth [PPR81]. In this case, the amount of time taken to transmit the information is not as much of a concern as is the finite amount of battery energy. The receivers on Earth can afford to wait an arbitrary amount of time, but once the batteries are consumed, the

∗Stanford Institute for Theoretical Physics, Stanford University, Stanford, California 94305, USA
†Department of Applied Physics, Stanford University, Stanford, California 94305, USA
‡Hearne Institute for Theoretical Physics, Department of Physics and Astronomy, Center for Computation and Technology, Louisiana State University, Baton Rouge, Louisiana 70803, USA
satellite is no longer useful if there is no external energy source available. Hence, in this case, the relevant constraint is an energy constraint.

In general, we can consider constraints with respect to a certain cost associated with transmission. Then, the corresponding communication limit should informally be the maximum amount of information that can be reliably transmitted divided by the total cost incurred in transmission. This is the notion of capacity per unit cost, which was introduced and extensively studied for classical channels in [Ver90]. After the development of quantum information theory, capacity per unit cost was extended to channels with classical binary inputs and quantum outputs [CHP07].

In this paper, we generalize the capacity per unit cost to various communication tasks involving quantum channels, including classical communication, entanglement-assisted classical communication, private communication, and quantum communication. To do so, we first recall from classical information theory the notion of a cost function, which associates to each input symbol a non-negative real number. In the quantum case, we employ a cost observable $G \geq 0$ [Hol03, Hol04] in order to quantify the cost of transmitting a quantum state. We then generalize the results in [Ver90] to classical communication over quantum channels by giving a formula for the capacity per unit cost of the form

$$\sup_{\{p_X, \rho_A(x)\}} \frac{I(X; B)_{\rho}}{\text{tr}[G \bar{\rho}_A]},$$

where $A, B$ denote quantum systems, $\bar{\rho}_A \equiv \sum_x p_X(x) \rho_A(x)$, and the mutual information $I(X; B)$ is evaluated with respect to a classical–quantum state $\rho_{XB}$ associated with the channel. For simplicity, we focus on channels with a single-letter cost-constrained classical capacity, including entanglement-breaking channels for example [Sho02a, Shi06, Hol08]. The formula in eq. (1.1) is derived via the known formula [Hol04] for the cost-constrained classical capacity, also known as the classical capacity cost function, of a quantum channel and reduces to it when $G$ is the identity operator. This is intuitive since in this case the cost is the number of channel uses, and so every quantum state incurs unit cost.

Now, more interesting results come about in the special case in which there is a zero-cost quantum state, that is, some state $\psi^0$ such that $\text{tr}[G \psi^0] = 0$. This is a natural setting to consider, given that transmitting a zero-cost state often physically corresponds to not actively sending anything through the channel. For example, for a bosonic channel the zero-cost state is the vacuum. Now, by the positive semi-definiteness of $G$, without loss of generality we can take $\psi^0$ to be pure. In this case, we find that the capacity per unit cost reduces to the following expression:

$$\sup_{\psi \neq \psi^0} \frac{D(\mathcal{N}(\psi) \| \mathcal{N}(\psi^0))}{\langle \psi | G | \psi \rangle},$$

where $D$ denotes the quantum relative entropy [Ume62] and the supremum is with respect to pure states $\psi$. The above expression is intuitive given the fact that pulse-position-modulation (PPM) protocols achieve the capacity per unit cost when there is a zero-cost state [Ver90]. Such PPM protocols encode information into the position of a $\psi$-pulse amidst a baseline of zero-cost $\psi^0$ states. Hence, for these protocols, we expect the relevant variables for computing the capacity per unit cost to be the distinguishability of the states $\mathcal{N}(\psi)$ and $\mathcal{N}(\psi^0)$ in addition to the cost of $\psi$. For other communication tasks such as entanglement-assisted, private, or quantum communication, we find related results and point to the respective Sections 4, 5, and 6 for details. We apply these formulae to various quantum Gaussian channels in Section 7. In Section 8 we introduce the notion of a blocklength-constrained capacity per unit cost, analogous to that of a cost-constrained capacity,
and derive a formula for it. We find that a blocklength constraint can ensure that the capacity per unit cost is finite and thus can play a similar role to a cost constraint for the usual capacity.

In classical information theory, communication limits with respect to costs other than time were first considered in [Rez61, Sha48, Pie78, Gol49, Gal87]. Such ideas have appeared in quantum settings as well, for instance for quantum bosonic channels [GHT11, Guh11, GDS11, DBEM11] where the relevant cost is the photon number.

Related Work: After deriving many of the results in this paper and while drafting this manuscript, a related work appeared on the quant-ph arXiv [Jar17]. In [Jar17], the author considers classical capacity per unit cost for particular channels that accept a general classical input symbol and output a quantum state (known as classical–quantum channels in the literature), thus generalizing the approach in [CHP07]. In particular, a cost function is considered to quantify the cost of classical input symbols. We note here that our paper generalizes this setup to the fully quantum case in which there is a cost observable and the channels considered have quantum inputs and quantum outputs.

2 Preliminaries

For simplicity, we restrict our developments to finite-dimensional Hilbert spaces, with the exception of Section 7, which applies to quantum Gaussian channels. Let $\mathcal{H}_A$ and $\mathcal{H}_B$ denote finite-dimensional Hilbert spaces, and let $\mathcal{L}(\mathcal{H}_A)$ and $\mathcal{L}(\mathcal{H}_B)$ denote spaces of linear operators acting on those respective Hilbert spaces. We denote by $\mathcal{N}_{A\rightarrow B}:\mathcal{L}(\mathcal{H}_A)\rightarrow\mathcal{L}(\mathcal{H}_B)$ a quantum channel, defined to be a completely positive and trace-preserving map. By Stinespring’s dilation theorem [Sti55, Pau02], $\mathcal{N}_{A\rightarrow B}$ can be extended to an isometric channel $\mathcal{U}_{A\rightarrow BE}:\mathcal{L}(\mathcal{H}_A)\rightarrow\mathcal{L}(\mathcal{H}_B \otimes \mathcal{H}_E)$, where $\mathcal{H}_E$ is some other finite-dimensional Hilbert space and $\mathcal{N}_{A\rightarrow B} = \text{tr}_E \circ \mathcal{U}_{A\rightarrow BE}$.

Now, let $G \in \mathcal{L}(\mathcal{H}_A)$ be a positive semi-definite operator acting on $\mathcal{H}_A$. Throughout, we refer to $G$ as the cost observable. This is the standard cost constraint used in applications of quantum Shannon theory (see, for instance, [Hol03, Hol04, Hol08, Hol12]). This is also a quantum generalization of a classical cost function [Ver90], which is a map from the input alphabet to the non-negative reals. The cost function can be embedded into a cost observable by letting the spectrum of the latter be the image of the former. However, note that unlike in the classical case, we can use quantum codewords that are not eigenstates of the cost observable. This might even be necessary to achieve the capacity of a quantum channel. For example, for single-mode phase-insensitive bosonic Gaussian channels, the relevant cost observable is the photon number operator, but it is known that coherent states, not number states, achieve the classical capacity [GGL+04, GHGP15, GGPCH14].

Lastly, given two quantum states $\rho, \sigma \in \mathcal{S}(\mathcal{H}_A)$, where $\mathcal{S}(\mathcal{H}_A) \subset \mathcal{L}(\mathcal{H}_A)$ denotes the set of positive semi-definite operators with unit trace, a quantum hypothesis test with $N$ copies is a binary positive-operator valued measure (POVM) $\{\Lambda_N, I - \Lambda_N\}$ that distinguishes between $N$ copies of the two states. The two states to be distinguished are called the null and alternative hypotheses, respectively. Now, there are two possible errors that can occur. Taking $\Lambda_N$ to be the measurement result that declares the state to be $\rho^\otimes N$, the error probabilities are given by

$$\alpha_N(\Lambda_N) \equiv \text{tr}[(I - \Lambda_N)\rho^\otimes N],$$

$$\beta_N(\Lambda_N) \equiv \text{tr}[\Lambda_N\sigma^\otimes N].$$

The errors are called the Type I and Type II errors, respectively. Then, for some $\epsilon \in (0, 1)$, we can
define the following quantity:
\[
\beta_N^*(\varepsilon) \equiv \inf_{\Lambda_N} \{ \beta_N(\Lambda_N) | \sigma_N(\Lambda_N) \leq \varepsilon \}.
\] (2.3)

That is, it is the lowest Type II error possible, given that the Type I error does not exceed \(\varepsilon\). By Quantum Stein’s Lemma [HP91, ON00], for all \(\varepsilon \in (0,1)\),
\[
\lim_{N \to \infty} -\frac{1}{N} \log_2 \beta_N^*(\varepsilon) = D(\rho\|\sigma),
\] (2.4)
where the quantum relative entropy \(D(\rho\|\sigma)\) is defined as
\[
D(\rho\|\sigma) \equiv \text{tr}[\rho(\log_2 \rho - \log_2 \sigma)]
\] (2.5)
whenever \(\text{supp}(\rho) \subseteq \text{supp}(\sigma)\) and it is equal to \(+\infty\) otherwise.

### 3 Classical Communication

We first consider the case of unassisted classical communication over a quantum channel \(\mathcal{N}\). Let \(n, M \in \mathbb{N}, \nu \in \mathbb{R}_{>0}\), and \(\varepsilon \in [0,1]\). We denote an \((n,M,\nu,\varepsilon)\) code as one with blocklength \(n\) and number of messages \(M\). Furthermore, denoting the quantum codewords as the density operators \(\rho_{A^n}(1), \ldots, \rho_{A^n}(M) \in \mathcal{S}(\mathcal{H}_A^n)\), each quantum codeword satisfies
\[
\text{tr}[G_n \rho_{A^n}(x)] \leq \nu,
\] (3.1)
where
\[
G_n \equiv \sum_{j=1}^{n} I^{j-1} \otimes G \otimes I^{n-j}.
\] (3.2)
The operator \(G_n\) is the total cost observable for the \(n\) channel uses, and by inspection, it is equal to the sum of the individual costs incurred for each channel use. Finally, given that the decoder uses a POVM \(\{\Pi_1, \ldots, \Pi_M\}\) to guess the message, the average error probability over the possible messages cannot exceed \(\varepsilon\), i.e.,
\[
\frac{1}{M} \sum_{m=1}^{M} \text{tr}[(I^{\otimes n} - \Pi_m)^{\otimes n}(\rho_{A^n}(m))] \leq \varepsilon.
\] (3.3)

We recall the definition for the cost-constrained classical capacity of a quantum channel [Hol03, Hol04]:

**Definition 1.** Given \(\varepsilon \in [0,1]\), and \(\beta > 0\), a non-negative number \(R\) is an \(\varepsilon\)-achievable rate with average cost not exceeding \(\beta\) if for all \(\delta > 0\), \(\exists n_0 \in \mathbb{N}\) such that if \(n \geq n_0\), then \(\exists\) an \((n,M,n\beta,\varepsilon)\) code such that \(\frac{\log_2 M}{n} > R - \delta\). Then, \(R\) is called achievable if it is \(\varepsilon\)-achievable for all \(\varepsilon \in (0,1)\). The supremum of all achievable rates with average cost not exceeding \(\beta\) is denoted \(C(\mathcal{N},\beta)\), the classical capacity cost function.
For simplicity, for the rest of the section we will consider channels with additive Holevo information at all cost constraints, i.e.,
\[ \forall \beta \geq 0, n \in \mathbb{N}, \quad \chi(\mathcal{N}^\otimes n, n\beta) = n\chi(\mathcal{N}, \beta), \tag{3.4} \]
where
\[ \chi(\mathcal{N}^\otimes n, n\beta) \equiv \sup_{\{p_X, \rho_A^n(x)\}} I(X; B^n)_{\rho} \tag{3.5} \]
and
\[ \rho_{XB^n} = \sum_x p_X(x) |x \rangle \langle x | \otimes \mathcal{N}^\otimes n(\rho_{A^n}(x)), \tag{3.6} \]
\[ \tilde{\rho}_A^n = \sum_x p_X(x) \rho_{A^n}(x). \tag{3.7} \]

Similar to the classical case [Ver90], the classical capacity cost function can be computed as an optimization of a mutual information with respect to input ensembles that satisfy the cost constraint:

**Theorem 2** ([SW97, Hol98, Hol03, Hol04]). For a channel with additive Holevo information at all cost constraints, the classical capacity cost function is given by
\[ C(\mathcal{N}, \beta) = \chi(\mathcal{N}, \beta) \equiv \sup_{\{p_X, \rho_A^n(x)\}} I(X; B)_{\rho}. \tag{3.8} \]

We now give the definition of the classical capacity per unit cost.

**Definition 3.** Given \( \varepsilon \in [0, 1) \), a non-negative number \( R \) is an \( \varepsilon \)-achievable rate per unit cost if for every \( \delta > 0 \), \( \exists \nu_0 > 0 \) such that if \( \nu \geq \nu_0 \), \( \exists \) an \( (n, M, \nu, \varepsilon) \) code such that \( \log_2 M > \nu(R - \delta) \). \( R \) is achievable if it is achievable for all \( \varepsilon \in (0, 1) \) and the capacity per unit cost is the supremum of all achievable rates per unit cost, denoted as \( C(\mathcal{N}) \).

Observe that the above definition is similar to that of the usual capacity, except we replace the blocklength \( n \) by the cost \( \nu \). In fact, we can also give an expression for the classical capacity per unit cost in terms of an optimized mutual information:

**Theorem 4.** The capacity per unit cost for a channel with additive Holevo information at all cost constraints is given by
\[ C(\mathcal{N}) = \sup_{\beta > 0} \frac{C(\mathcal{N}, \beta)}{\beta} = \sup_{\{p_X, \rho_A^n(x)\}} \frac{I(X; B)_{\rho}}{\text{tr}[G \rho_A^\beta]}. \tag{3.9} \]

The proof of Theorem 4 is based on [Ver90] and follows from the achievability and converse for the cost-constrained classical capacity. In general it is sufficient to prove the coding theorem for the cost-constrained capacity in order to establish a coding theorem for the capacity per unit cost.

**Proof of Theorem 4.** We first show the achievability statement \( C(\mathcal{N}) \geq \sup_{\beta > 0} \frac{C(\mathcal{N}, \beta)}{\beta} \). Let \( \beta > 0 \). Let \( R \) be an achievable rate per channel use with average cost not exceeding \( \beta \). We claim that \( R/\beta \) is an achievable rate per unit cost. This is clear for \( R = 0 \), so we can assume \( R > 0 \). To see the
claim, let \( \varepsilon \in (0, 1) \) and fix some \( \delta > 0 \). Then, by definition \( \exists n_0 \) such that for all \( n \geq n_0 \) there is an \((n, M, n\beta, \varepsilon)\) code such that
\[
\frac{\log_2 M}{n} > R - \frac{\beta \delta}{2}.
\] (3.10)
This same code is a \((n, M, n\beta, \varepsilon)\) code such that
\[
\frac{\log_2 M}{n\beta} > \frac{R}{\beta} - \frac{\delta}{2}.
\] (3.11)
Now, let \( \nu_0 = \max \left\{ (n_0 + 1)\beta, \frac{2R}{\delta} \right\} \) and \( \nu \geq \nu_0 \). If \( \nu = n\beta \) for some \( n \in \mathbb{N} \), then \( n \geq n_0 \), and so the above \((n, M, n\beta, \varepsilon)\) code satisfies the necessary requirements. If instead \( n\beta < \nu < (n + 1)\beta \) for some \( n \in \mathbb{N} \), then we note
\[
n + 1 > \frac{\nu}{\beta} \geq \frac{\nu_0}{\beta} \geq \frac{2R}{\delta \beta},
\] (3.12)
and so
\[
\left( \frac{R}{\beta} - \frac{\delta}{2} \right) \frac{n\beta}{\nu} > \left( \frac{R}{\beta} - \frac{\delta}{2} \right) \frac{n}{n + 1}
= \left( \frac{R}{\beta} - \frac{\delta}{2} \right) \left( 1 - \frac{1}{n + 1} \right)
> \left( \frac{R}{\beta} - \frac{\delta}{2} \right) \left( 1 - \frac{\delta \beta}{2R} \right)
> \frac{R}{\beta} - \delta.
\] (3.13)
Now, \((n + 1)\beta > \nu \geq \nu_0 \), so \( n \geq n_0 \). Hence, the above \((n, M, n\beta, \varepsilon)\) code is a \((n, M, \nu, \varepsilon)\) code such that
\[
\frac{\log_2 M}{\nu} > \left( \frac{R}{\beta} - \frac{\delta}{2} \right) \frac{n\beta}{\nu} \geq \frac{R}{\beta} - \delta.
\] (3.17)
Hence we have shown achievability.

We next show the converse statement \( C(\mathcal{N}) \leq \sup_{\beta > 0} \frac{C(\mathcal{N}, \beta)}{\beta} \). Suppose \( \mathcal{N} \) has an \((n, M, \nu, \varepsilon)\) code. By a standard data-processing argument and entropy continuity bound [Wil17, Hol12, Win16], we have
\[
\log_2 M \leq \chi(\mathcal{N}, \nu) + f(M, \varepsilon),
\] (3.18)
where \( f(M, \varepsilon) = \varepsilon \log_2 M + (\varepsilon + 1) \log_2 (\varepsilon + 1) - \varepsilon \log_2 \varepsilon \), so that \( \lim_{\varepsilon \to 0} f(M, \varepsilon) = 0 \). Thus,
\[
\frac{\log_2 M}{\nu} \leq \frac{\chi(\mathcal{N}, \nu)}{\nu} + \frac{f(M, \varepsilon)}{\nu}
= \frac{1}{\nu} \sup_{\{p_X, \rho_A(x)\}} \text{tr}[\overline{G}_\rho A] \frac{I(X; B)_\rho + \frac{f(M, \varepsilon)}{\nu}}{\nu}
\leq \sup_{\beta > 0} \frac{1}{\beta} \sup_{\{p_X, \rho_A(x)\}} \text{tr}[\overline{G}_\rho A] \leq \beta \frac{I(X; B)_\rho + \frac{f(M, \varepsilon)}{\nu}}{\nu}
= \sup_{\beta > 0} \frac{C(\mathcal{N}, \beta)}{\beta} + \frac{f(M, \varepsilon)}{\nu}.
\] (3.21)
Thus, for any \( \varepsilon \)-achievable rate per unit cost \( R \), for any \( \delta > 0 \) there exists \( \nu_0 \) such that for \( \nu \geq \nu_0 \),

\[
R - \delta < \sup_{\beta > 0} \frac{C(N, \beta)}{\beta} + \frac{f(M, \varepsilon)}{\nu}.
\]

(3.23)

Hence, for all \( \delta > 0 \),

\[
R - \delta < \liminf_{\nu \to \infty} \left( \sup_{\beta > 0} \frac{C(N, \beta)}{\beta} + \frac{f(M, \varepsilon)}{\nu} \right).
\]

(3.24)

Hence, if \( R \) is an achievable rate per unit cost, then

\[
R \leq \sup_{\beta > 0} \frac{C(N, \beta)}{\beta}.
\]

(3.25)

This establishes the first equality in Theorem 4.

To show the second equality, we first argue

\[
\sup_{\beta > 0} \frac{C(N, \beta)}{\beta} = \sup_{\beta > 0} \frac{1}{\beta} \sup_{\{p_X, \rho_A(x)\}} \frac{1}{\text{tr}[G\rho_A]} I(X; B)_\rho
\]

\[\leq \sup_{\beta > 0} \frac{1}{\text{tr}[G\rho_A]} I(X; B)_\rho \]

\[= \sup_{\{p_X, \rho_A(x)\}} \frac{I(X; B)_\rho}{\text{tr}[G\rho_A]}. \]

(3.26)

(3.27)

(3.28)

Note that the inequality is trivial if for some ensemble \( \text{tr}[G\rho_A] = 0 \) and \( I(X; B)_\rho > 0 \). Now, this is also an achievable rate per unit cost since for any \( \{p_X, \rho_A(x)\} \), we can achieve a rate per channel use \( I(X; B)_\rho \) using cost-constrained Holevo-Schumacher-Westmoreland (HSW) coding [Hol98, SW97, Hol03, Hol04, Hol12]. The average cost per channel use is then exactly \( \text{tr}[G\rho_A] \), and so we achieve a capacity per unit cost equal to

\[
\frac{I(X; B)_\rho}{\text{tr}[G\rho_A]}.
\]

(3.29)

This concludes the proof.

Now, suppose that we have a state \( \psi^0 \) with zero cost, i.e., \( \text{tr}[G\psi^0] = 0 \). As mentioned above, without loss of generality, \( \psi^0 \) can be taken pure since otherwise we can spectrally decompose it and conclude that all of its eigenstates must have zero cost since \( G \geq 0 \). In this special case, the capacity per unit cost is given by the following simple expression:

**Theorem 5.** If there is a state \( \psi^0 \) with zero cost, then the capacity per unit cost of a channel with additive Holevo information at all cost constraints is

\[
C(N) = \sup_{\psi \neq \psi^0} \frac{D(N(\psi)||N(\psi^0))}{\langle \psi | G | \psi \rangle},
\]

(3.30)

where \( \psi \) is pure.
Just as was found in [Ver90], the expression for the capacity per unit cost is arguably simpler than that for the capacity cost function given in eq. (3.8). The latter requires an optimization over ensembles on the input space while the former only requires an optimization over the input space itself.

We now give a proof of Theorem 5.

**Proof.** Without loss of generality, \( \psi^0 \) is the unique zero-cost state. Otherwise, let \( \phi^0 \neq \psi^0 \) be a zero-cost state. If \( \mathcal{N}(\phi^0) \neq \mathcal{N}(\psi^0) \), the capacity per unit cost is infinite since we can send a binary message with zero cost. If on the other hand \( \mathcal{N}(\phi^0) = \mathcal{N}(\psi^0) \), then \( \phi^0 \) is the same as \( \psi^0 \) for the purposes of communicating over \( \mathcal{N} \).

We first prove the direct part. To begin with, we note that the possibility of time-sharing (interpolation between two different protocols) implies that \( C(N, \beta) \) is concave in \( \beta \). Furthermore, we have a zero-cost state and so \( C(N, \beta)/\beta \) is monotone non-increasing on \( (0, +\infty) \). We conclude that

\[
C(N) = \lim_{\beta \searrow 0} \frac{C(N, \beta)}{\beta}.
\]  

(3.31)

Now, let \( \beta \in (0, \langle \psi | G | \psi \rangle) \) and consider the following classical-quantum state:

\[
\rho_{XA}^\beta = \left( 1 - \frac{\beta}{\langle \psi | G | \psi \rangle} \right) |0\rangle \langle 0| \otimes \psi^0 + \frac{\beta}{\langle \psi | G | \psi \rangle} |1\rangle \langle 1| \otimes \psi,
\]

(3.32)

where \( \psi \neq \psi^0 \). By Theorem 4, we can achieve the following rate per unit cost:

\[
\frac{I(X; B)_{\rho^\beta}}{\beta}.
\]

(3.33)

Now, recall the following identity, which holds for a classical-quantum state \( \rho_{XB} = \sum_x p_X(x) |x\rangle \langle x|_X \otimes \rho^x_B \) [YO93, Hol12] where for all \( x \), \( p_X(x) > 0 \):

\[
I(X; B)_\rho = D(\rho_{XB} \| \rho_X \otimes \rho_B)
= \sum_x p_X(x) D(\rho^x_B \| \rho_B).
\]

(3.34)

(3.35)

This expression is well defined because \( \text{supp}(\rho^x_B) \subseteq \text{supp}(\rho_B) \) for all \( x \). Hence by eq. (3.35) and non-negativity of quantum relative entropy when evaluated on quantum states, we obtain

\[
\frac{I(X; B)_{\rho^\beta}}{\beta} \geq \frac{D(\mathcal{N}(\psi) \| \rho^\beta_B)}{\langle \psi | G | \psi \rangle}.
\]

(3.36)

So by the lower semicontinuity of the relative entropy [Hol12],

\[
C(N) = \lim_{\beta \searrow 0} \frac{C(N, \beta)}{\beta}
\]

(3.37)

\[
\geq \lim_{\beta \searrow 0} \frac{I(X; B)_{\rho^\beta}}{\beta}
\]

(3.38)

\[
\geq \liminf_{\beta \searrow 0} \frac{D(\mathcal{N}(\psi) \| \rho^\beta_B)}{\langle \psi | G | \psi \rangle}.
\]

(3.39)

\[
\geq \frac{D(\mathcal{N}(\psi) \| \mathcal{N}(\psi^0))}{\langle \psi | G | \psi \rangle}.
\]

(3.40)
This holds for all $\psi \neq \psi^0$, and so we obtain the direct part
\[
C(N) \geq \sup_{\psi \neq \psi^0} \frac{D(N(\psi) \| N(\psi^0))}{\langle \psi | G | \psi \rangle}. \tag{3.41}
\]

For the converse, we start with
\[
I(X; B)_\rho = \inf_{\sigma_B \in S(H_B)} D(\rho_{XB} \| \rho_X \otimes \sigma_B) \tag{3.42}
\leq D(\rho_{XB} \| \rho_X \otimes N(\psi^0)) \tag{3.43}
= \sum_x p_X(x) D(\rho^x \| N(\psi^0)). \tag{3.44}
\]

The first equality is a well known identity [Wil17, Exercise 11.8.2]. Note that if any of the relative entropies are infinite, then the bound is trivial. Therefore, $\forall \beta > 0$,
\[
\frac{C(N, \beta)}{\beta} = \frac{1}{\beta} \sup_{\{p_X, \psi^x\}} \frac{I(X; B)}{\text{tr}[G \rho] \leq \beta} \tag{3.45}
= \frac{1}{\beta} \sup_{\{p_X, \psi^x\}} \frac{I(X; B)}{\text{tr}[G \psi] \leq \beta} \tag{3.46}
\leq \frac{1}{\beta} \sup_{\{p_X, \psi^x\}, \psi^x \neq \psi^0} \frac{\sum_x p_X(x) D(\rho^x \| N(\psi^x))}{\langle \psi^x | G | \psi^x \rangle} \tag{3.47}
= \frac{1}{\beta} \sup_{\{p_X, \psi^x\}, \psi^x \neq \psi^0} \frac{\sum_x p_X(x) D(\rho^x \| N(\psi^0))}{\langle \psi^x | G | \psi^x \rangle} \tag{3.48}
\leq \sup_{\psi \neq \psi^0} \frac{D(N(\psi) \| N(\psi^0))}{\langle \psi | G | \psi \rangle} \sup_{\{p_X, \psi^x\}, \psi^x \neq \psi^0} \frac{\sum_x p_X(x) \langle \psi^x | G | \psi^x \rangle}{\beta} \tag{3.49}
\leq \sup_{\psi \neq \psi^0} \frac{D(N(\psi) \| N(\psi^0))}{\langle \psi | G | \psi \rangle}, \tag{3.50}
\]

where eq. (3.46) follows since pure state ensembles maximize the Holevo information (even with a cost constraint), and we can divide by $\langle \psi^x | G | \psi^x \rangle$ in eq. (3.49) since we assumed $\psi^0$ is the unique zero-cost state. \hfill \Box

3.1 Pulse-Position-Modulation Scheme for Classical Communication

We can also directly prove the achievability part of Theorem 5 without going through the cost-constrained capacity, as was done in [Ver90] for the classical case. This follows by using a PPM scheme along the following lines.
**Encoding:** Let \( \psi \neq \psi^0 \) be a pure state and fix \( M, N \in \mathbb{N} \). For a message \( m \in [1 : M] \), the sender transmits the following length-\( MN \) sequence of states:

\[
\left((\psi^0)^\otimes N\right)^{\otimes m-1} \otimes \psi^0 \otimes \left((\psi^0)^\otimes N\right)^{\otimes M-m}.
\]

That is, the message is encoded in the position of a \( \psi \)-“pulse” amidst a baseline of zero-cost states. Note that the cost of each codeword is \( N \langle \psi | G | \psi \rangle \).

**Decoding:** Let \( \varepsilon \in (0, 1) \). The receiver obtains the state

\[
\left(\mathcal{N}(\psi^0)^\otimes N\right)^{\otimes m-1} \otimes \mathcal{N}(\psi) \otimes \left(\mathcal{N}(\psi^0)^\otimes N\right)^{\otimes M-m}.
\]

Then, the receiver uses a quantum hypothesis test to deduce the position of the pulse. Specifically, he performs \( M \) independent binary hypothesis tests with \( N \) copies where the null hypothesis is \( \mathcal{N}(\psi) \) and the alternative hypothesis is \( \mathcal{N}(\psi^0) \). If the receiver obtains a test result of the form eq. (3.53) for some \( \hat{m} \), then \( \hat{m} \) is declared. Otherwise an error is declared.

**Error Analysis:** Let \( A_{iN} \), for \( i \in [1 : M] \), denote the POVM of the \( i \)th hypothesis test, and let \( \alpha_{iN}(A_{iN}) \) and \( \beta_{iN}(A_{iN}) \) denote the Type I and Type II errors, respectively. Now, the error probability \( p_e \) is independent of the message by symmetry, so we can fix some message index \( i \). Furthermore, since each POVM acts on independent size-\( N \) blocks, we can apply the classical union bound as follows:

\[
p_e \leq \alpha_{iN} + (M - 1)\beta_{iN}.
\]

By eq. (2.4), for \( \varepsilon \in (0, 1) \),

\[
\lim_{N \to \infty} -\frac{1}{N} \log_2 \beta_{iN}^*((\varepsilon/2)) = D(\mathcal{N}(\psi)\|\mathcal{N}(\psi^0)).
\]

Using the test \( A_{iN} \) that achieves \( \beta_{iN}^*((\varepsilon/2)) \) and given \( \delta > 0 \), for sufficiently large \( N \), the probability of error is bounded by

\[
p_e \leq \frac{\varepsilon}{2} + (M - 1)2^{-ND(\mathcal{N}(\psi)\|\mathcal{N}(\psi^0)) + N\delta}.
\]

Hence, if

\[
\frac{\log_2 M}{N \langle \psi | G | \psi \rangle} < \frac{D(\mathcal{N}(\psi)\|\mathcal{N}(\psi^0))}{\langle \psi | G | \psi \rangle} - \frac{2\delta}{\langle \psi | G | \psi \rangle},
\]

then \( p_e < \varepsilon \) for sufficiently large \( N \). Therefore, \( \frac{D(\mathcal{N}(\psi)\|\mathcal{N}(\psi^0))}{\langle \psi | G | \psi \rangle} \) is an achievable rate per unit cost.

4 Entanglement-Assisted Communication

We now consider the case of communication with unlimited entanglement assistance. We define an \((n, M, \nu, \varepsilon) \) code in the same way as in the unassisted case, with the exception that the sender and receiver are allowed to share an arbitrary quantum state of arbitrary dimension before communication begins and they can use this resource in the encoding and decoding. The entanglement-assisted
capacity cost function \( C_{EA}(\mathcal{N}, \beta) \) is defined similarly but again takes into account the entanglement assistance.

Let \( A \) and \( A' \) denote quantum systems with isomorphic Hilbert spaces. Define for a bipartite state \( \varphi_{AA'} \)
\[
\varphi_{AB} \equiv (\text{id}_A \otimes \mathcal{N}_{A' \to B})(\varphi_{AA'}).
\] (4.1)

We recall the following theorem:

**Theorem 6** ([Hol03]). The entanglement-assisted capacity cost function for a quantum channel \( \mathcal{N}_{A' \to B} \) is given by
\[
C_{EA}(\mathcal{N}, \beta) = \max_{\varphi_{AA'}} \frac{I(A; B)}{\text{tr}[G\varphi_{A'}] \leq \beta},
\] (4.2)

where \( \varphi_{AA'} \) is a pure bipartite state.

We define the entanglement-assisted capacity per unit cost \( C_{EA}(\mathcal{N}) \) in the same manner and obtain an expression for it in the same way as in the unassisted case (i.e., as done in Theorem 4).

**Theorem 7.** The entanglement-assisted capacity per unit cost for a quantum channel \( \mathcal{N}_{A' \to B} \) is given by
\[
C_{EA}(\mathcal{N}) = \sup_{\beta > 0} \frac{C_{EA}(\mathcal{N}, \beta)}{\beta} = \sup_{\varphi_{AA'}} \frac{I(A; B)}{\text{tr}[G\varphi_{A'}]},
\] (4.3)

Now suppose that we have a zero-cost pure state \( \psi^0 \). Similar to the unassisted case, we obtain the expression for the entanglement-assisted capacity per unit cost given in Theorem 8. Note that, like the mutual information, the quantity to be optimized is only a function of the input state \( \varphi_{A'} \) and not of the specific purification. Also note that, unlike the unassisted case, \( C_{EA} \) is ostensibly as difficult to calculate as \( C_{EA} \).

**Theorem 8.** If there is a state \( \psi^0 \) with zero cost, then the entanglement-assisted capacity per unit cost of a channel \( \mathcal{N}_{A' \to B} \) is given by
\[
C_{EA}(\mathcal{N}) = \sup_{\varphi_{AA'}} C_{EA,\psi^0}(\mathcal{N}, \varphi),
\] (4.4)

where
\[
C_{EA,\psi^0}(\mathcal{N}, \varphi) = \begin{cases} 
D(\varphi_{AB} \| \varphi_A \otimes \mathcal{N}(\psi^0_{A'})) / \text{tr}[G\varphi_{A'}] & D(\varphi_{AB} \| \varphi_A \otimes \mathcal{N}(\psi^0_{A'})) > 0, \\
0 & \text{otherwise}
\end{cases}
\] (4.5)

**Proof.** The proof proceeds much like in the unassisted case. Suppose that \( \varphi^0 \neq \psi^0 \) has zero cost. Now, if \( D(\varphi^0_{AB} \| \varphi^0_A \otimes \mathcal{N}(\psi^0_{A'})) > 0 \), we clearly have infinite \( C_{EA}(\mathcal{N}) \) since we can send a distinguishable binary message (using the entangled state \( \varphi^0_{AA'} \)) with zero cost. Hence it is sufficient to assume that \( D(\varphi^0_{AB} \| \varphi^0_A \otimes \mathcal{N}(\psi^0_{A'})) = 0 \).

We now prove achievability. By concavity of \( C_{EA}(\mathcal{N}, \beta) \) with respect to \( \beta \) and the existence of a zero-cost state, \( C_{EA} \) is
\[
C_{EA}(\mathcal{N}) = \lim_{\beta \to 0} \frac{C_{EA}(\mathcal{N}, \beta)}{\beta}.
\] (4.6)

Consider some \( \varphi_{AA'} \). Since we can trivially achieve zero rate, suppose that \( \varphi_{A'} \) has positive cost. Then, define the following state:
\[
\rho_{XAA'}^\beta \equiv \left(1 - \frac{\beta}{\text{tr}[G\varphi_{A'}]}\right) |0\rangle\langle 0|_X \otimes \varphi_A \otimes \psi^0_{A'} + \frac{\beta}{\text{tr}[G\varphi_{A'}]} |1\rangle\langle 1|_X \otimes \varphi_{AA'},
\] (4.7)
where $\beta \in (0, \text{tr}[G\varphi_{A'}])$. By Theorem 7 and the data-processing inequality for mutual information, we obtain the following entanglement-assisted rate per unit cost for the mixed state $\rho_{XAB}^\beta \equiv (\text{id}_X \otimes N'_{A' \rightarrow B})(\rho_{XAA'}^\beta)$:

$$\frac{I(XA; B)_{\rho^\beta}}{\beta} \quad (4.8)$$

Now, we can write the mutual information of any classical-quantum state $\rho_{XAB} = \sum_x p_X(x) |x\rangle\langle x|_X \otimes \rho_{AB}^x$ as the following convex sum of relative entropies:

$$I(XA; B)_{\rho} = \text{tr} [\rho_{XAB} (\log_2 \rho_{XAB} - \log_2 \rho_X \otimes \rho_B)] \quad (4.9)$$

$$= \sum_x p_X(x) \text{tr}_{AB} [\rho_{AB}^x (\log_2 \rho_{AB}^x - \log_2 \rho_A^x \otimes \rho_B)] \quad (4.10)$$

$$= \sum_x p_X(x) D(\rho_{AB}^x \| \rho_A^x \otimes \rho_B). \quad (4.11)$$

Thus, by the non-negativity of quantum relative entropy when evaluated on quantum states,

$$\frac{I(XA; B)_{\rho^\beta}}{\beta} \geq \frac{D(\varphi_{AB} \| \varphi_A \otimes \rho_B^\beta)}{\text{tr}[G\varphi_{A'}]} \quad (4.12)$$

Hence, again by the lower semicontinuity of the relative entropy,

$$C_{EA}(N) = \lim_{\beta \searrow 0} \frac{C_{EA}(N, \beta)}{\beta} \quad (4.14)$$

$$\geq \lim_{\beta \searrow 0} \frac{I(XA; B)_{\rho^\beta}}{\beta} \quad (4.15)$$

$$\geq \liminf_{\beta \searrow 0} \frac{D(\varphi_{AB} \| \varphi_A \otimes \rho_B^\beta)}{\text{tr}[G\varphi_{A'}]} \quad (4.16)$$

$$\geq \frac{D(\varphi_{AB} \| \varphi_A \otimes N(\psi_0^A'))}{\text{tr}[G\varphi_{A'}]} \quad (4.17)$$

For the converse, we have for any pure input state $\varphi_{AA'}$,

$$I(A; B)_\varphi = \inf_{\sigma_B} D(\varphi_{AB} \| \varphi_A \otimes \sigma_B) \quad (4.18)$$

$$\leq D(\varphi_{AB} \| \varphi_A \otimes N(\psi_0^A')). \quad (4.19)$$

Note again that if the relative entropy is infinite, then the bound is trivial. Hence,

$$\frac{C_{EA}(N, \beta)}{\beta} \leq \sup_{\varphi_{AA'} \subseteq \text{tr}[G\varphi_{A'}] \leq \beta} \frac{D(\varphi_{AB} \| \varphi_A \otimes N(\psi_0^A'))}{\beta}. \quad (4.20)$$
Now, we assumed that for any zero-cost state \( \varphi^0 \), \( D(\varphi_{AB} || \varphi^0_A \otimes N(\psi^0_{A'})) = 0 \). Thus we can take the supremum over non-zero cost states. If there are not any, then eq. (4.20) implies that the upper bound is 0, which would conclude the converse. Otherwise, we can argue

\[
\frac{C_{EA}(N, \beta)}{\beta} \leq \sup_{\varphi_{AA'}} \frac{D(\varphi_{AB} || \varphi_A \otimes N(\psi^0_{A'})) \operatorname{tr}[G\varphi_{A'}]}{\beta} \leq \sup_{\varphi_{AA'}} \frac{D(\varphi_{AB} || \varphi_A \otimes N(\psi^0_{A'})) \operatorname{tr}[G\varphi_{A'}]}{\operatorname{tr}[G\varphi_{A'}]}.
\]

This concludes the proof.

### 4.1 Pulse-Position-Modulation Scheme for Entanglement-Assisted Classical Communication

We propose a PPM scheme that achieves the rate given in eq. (4.4), thereby providing an alternative proof of the direct part of Theorem 8. This will be much like the scheme in the unassisted case except with the greater discriminatory power that entanglement assistance provides.

**Encoding**: Let \( \varphi_{A'} \) be a positive-cost state and fix \( M, N \in \mathbb{N} \). The sender and receiver share \( MN \) copies of a pure state \( \varphi_{AA'} \), where \( A' \) is at the sender and \( A \) is at the receiver, where we have \( N \) copies for each message in \([1 : M]\). Hence the overall shared state is

\[
\bigotimes_{i=1}^M \left( \varphi_{A_iA'_{i'}} \right)^{\otimes N}.
\]

(4.23)

For a message \( m \in [1 : M] \), the sender transmits a \( \varphi_{A'} \)-pulse amidst a zero-cost state baseline by using the following sequence of states:

\[
\bigotimes_{i=1}^{m-1} \left( \psi^0_{A'_{i'}} \right)^{\otimes N} \otimes \left( \varphi_{A_m} \right)^{\otimes N} \otimes \bigotimes_{j=m+1}^M \left( \psi^0_{A'_{j'}} \right)^{\otimes N}.
\]

(4.24)

That is, the sender transmits the zero-cost state, but at every \( m \)th block of length \( N \), he sends his shares of the corresponding copies of \( \varphi_{AA'} \).

**Decoding**: Let \( \varepsilon \in (0,1) \). Now, since \( \varphi_A \) purifies \( \varphi_{A'} \), whenever the sender transmits \( \psi^0 \), the receiver obtains a product state \( \varphi_A \otimes N(\psi^0) \). Hence, the receiver now has the state

\[
\bigotimes_{i=1}^{m-1} \left( \varphi_{A_i} \otimes N(\psi^0_{A'_{i'}}) \right)^{\otimes N} \otimes \left( \text{id}_A \otimes N \right)(\varphi_{A_mA'_m})^{\otimes N} \otimes \bigotimes_{j=m+1}^M \left( \varphi_{A_j} \otimes N(\psi^0_{A'_{j'}}) \right)^{\otimes N}.
\]

(4.25)

Then, the receiver uses quantum hypothesis testing along with his shares of the entangled states to deduce the position of the pulse. He performs \( M \) independent binary hypothesis tests with \( N \) copies where the null hypothesis is \( \varphi_{AB} \equiv (\text{id}_A \otimes N_{A'\rightarrow B})(\varphi_{AA'}) \) and the alternative hypothesis is \( \varphi_A \otimes N(\psi^0_{A'}) \). If the receiver obtains a test result of the form eq. (4.25) for some \( \hat{m} \), then \( \hat{m} \) is declared. Otherwise an error is declared.
Error Analysis: The error analysis follows in exactly the same way as the unassisted case. We conclude that we can obtain vanishing error in transmission provided that, for some \( \delta > 0 \),

\[
\frac{\log_2 M}{N \text{tr}[G\varphi_{A'}]} < \frac{D(\varphi_{AB}||\varphi_A \otimes N(\psi^0_A))}{\text{tr}[G\varphi_{A'}]} - \frac{\delta}{\text{tr}[G\varphi_{A'}]}.
\] (4.26)

Hence we achieve the entanglement-assisted rate per unit cost \( \frac{D(\varphi_{AB}||\varphi_A \otimes N(\psi^0_A))}{\text{tr}[G\varphi_{A'}]} \).

Similar to the position-based decoding scheme for entanglement-assisted classical communication [AJW17], this scheme does not consume all the entanglement needed to implement the encoding. This follows from the gentle-measurement lemma [Win99a, Win99b]: the entangled states that were not transmitted but measured by the decoder will only be negligibly disturbed, given that the decoding measurement succeeds with high probability. Now, the natural measure of rate of entanglement consumption in this setting is entanglement consumed per unit cost, and in this scheme it can be expressed in terms of the entanglement entropy of \( \varphi_{AA'} \) as follows:

\[
\frac{NS(A)\varphi}{N \text{tr}[G\varphi_{A'}]} = \frac{S(A)\varphi}{\text{tr}[G\varphi_{A'}]}.
\] (4.27)

5 Private Communication

We now consider private communication over a quantum channel. This was first studied in [Dev05, CWY04] when there is no cost constraint and recently in [WQ16] when there is a cost constraint. Given a noisy channel \( N_{A \rightarrow B} \), let \( U_{A \rightarrow BE} \) denote an isometric channel extending it and let \( N^c_{A \rightarrow E} = \text{tr}_{B} \circ U_{A \rightarrow BE} \) denote the induced complementary channel.

The formulation here is based on [EHLG13], but note that here we use a definition of a private code with the privacy based on trace distance [KRBM07, Wil17, WQ16]. Namely, we define an \((n,M,\nu,\varepsilon,\zeta)\) private code as having blocklength \( n \in \mathbb{N} \), number of messages \( M \in \mathbb{N} \), total cost at most \( \nu \in \mathbb{R}_{>0} \), and probability of error of the receiver's decoding at most \( \varepsilon \in [0,1] \). The quantity \( \zeta \in [0,1] \) bounds the privacy error for each message: for each message, we demand that the eavesdropper's state is approximately independent of the message. Specifically, for all \( m \in [1 : M] \), we require that

\[
\frac{1}{2}\left\| (N^c_{A \rightarrow E})^\otimes n (\rho_{A^n}(m)) - \sigma_{E^n} \right\|_1 \leq \zeta,
\] (5.1)

where \( \rho_{A^n}(m) \) are the codewords and \( \sigma_{E^n} \) is some fixed state independent of \( m \).

We can now establish some definitions.

**Definition 9 ([WQ16]).** Given \( \beta > 0 \), \( R_p \) is an achievable private communication rate with average cost not exceeding \( \beta \) if for all \( \varepsilon,\zeta \in (0,1) \) and \( \delta > 0 \), \( \exists n_0 \) such that \( \forall n \geq n_0 \), there is an \((n,M,n\beta,\varepsilon,\zeta)\) code for which

\[
\frac{\log_2 M}{n} > R_p - \delta.
\] (5.2)

The supremum of all achievable rates with average cost not exceeding \( \beta \) as a function of \( \beta \) is the private capacity cost function \( P(N,\beta) \).

We recall the formula for \( P(N,\beta) \) when \( N \) is a degradable channel:
Theorem 10 (\cite{wq16}). The private capacity cost function for a degradable channel \( N_{A \to B} \) is given by
\[
P(N, \beta) = \sup_{\{p_X, \psi_A(x)\}} I(X; B)_\rho - I(X; E)_\rho, \tag{5.3}
\]
where each state \( \psi_A(x) \) is pure,
\[
\rho_{XBE} = \sum_x |x⟩⟨x|_X \otimes U_{A \to BE}(\psi_A(x)), \tag{5.4}
\]
and \( \bar{\psi}_A = \sum_x p_X(x)\psi_A(x) \) is the average input state.

Now we give the definition for the private capacity per unit cost.

Definition 11. \( R_p \) is an achievable private communication rate per unit cost if for all \( \varepsilon, \zeta \in (0, 1) \) and \( \delta > 0 \), there \( \exists \nu_0 > 0 \) such that \( \forall \nu \geq \nu_0 \), there is an \((n, M, \nu, \varepsilon, \zeta)\) code for which
\[
\log_2 M > \nu (R_p - \delta). \tag{5.5}
\]
The private capacity per unit cost is equal to the supremum of all achievable private communication rates per unit cost, denoted by \( P(N) \).

We can obtain an expression for the private capacity per unit cost via the private capacity cost function. A proof of this follows from the achievability and converse of the cost-constrained private capacity per channel use [WQ16], just as in the proof of Theorem 4.

Theorem 12. The private capacity per unit cost of a degradable channel \( N_{A \to B} \) is given by
\[
P(N) = \sup_{\beta > 0} \frac{P(N, \beta)}{\beta} = \sup_{\{p_X, \psi_A(x)\}} \frac{I(X; B)_\rho - I(X; E)_\rho}{\text{tr}[G\bar{\psi}_A]}, \tag{5.6}
\]
where \( \bar{\psi}_A \) is the average input state.

Now again suppose that we have a zero-cost state \( \psi^0 \). We then obtain the following simpler expression for the private capacity per unit cost.

Theorem 13. If there is a state \( \psi^0 \) with zero cost, then the private capacity per unit cost of a degradable channel \( N_{A \to B} \) is given by
\[
P(N) = \sup_{\psi} P_{\psi^0}(N, \psi), \tag{5.7}
\]
where \( \psi \) is pure,
\[
P_{\psi^0}(N, \psi) \equiv \begin{cases} N_X(\psi, \psi^0) & N_X(\psi, \psi^0) > 0 \\ 0 & \text{otherwise} \end{cases}, \tag{5.8}
\]
\[
N_X(\psi, \psi^0) \equiv D(N(\psi)||N(\psi^0)) - D(N^c(\psi)||N^c(\psi^0)), \tag{5.9}
\]
and \( N^c_{A \to E} \) denotes the complementary channel of \( N \) corresponding to an isometric channel \( U_{A \to BE} \) extending \( N \).
Proof. Suppose that the state $\varphi^0$ has zero cost. If $N_N(\varphi^0, \psi^0) \neq 0$, we have a zero-cost binary alphabet over which we can form ensembles for which $I(X; B) - I(X; E) > 0$. Hence, by Theorem 12, we can attain infinite private capacity per unit cost. Thus, it suffices to assume $N_N(\varphi^0, \psi^0) = 0$.

Now, once again by concavity and the existence of a zero-cost state,

$$P(N) = \lim_{\beta \downarrow 0} \frac{P(N, \beta)}{\beta}. \quad (5.10)$$

Let $\beta \in (0, \langle \psi | G | \psi \rangle)$, and let $\psi$ be a pure state. We can assume that $\psi$ has positive cost since it is trivial to attain zero rate. Consider the following classical-quantum state:

$$\rho^\beta_{XBE} = \left(1 - \frac{\beta}{\langle \psi | G | \psi \rangle}\right)|0\rangle\langle 0|_X \otimes U_{A \rightarrow BE}(\psi^0_A) + \frac{\beta}{\langle \psi | G | \psi \rangle}|1\rangle\langle 1|_X \otimes U_{A \rightarrow BE}(\psi_A). \quad (5.11)$$

By a similar argument as in the unassisted case applied to each relative entropy in $N_N$, in the limit $\beta \downarrow 0$, this ensemble achieves the desired quantity:

$$P(N) \geq \lim_{\beta \downarrow 0} \frac{I(X; B)_{\rho^\beta} - I(X; E)_{\rho^\beta}}{\beta} \geq \frac{N_N(\psi, \psi^0)}{\langle \psi | G | \psi \rangle}. \quad (5.12)$$

In arriving at the above result, we need to make use of the lower semi-continuity of the private information as a function of the input ensemble. This is proven for bounded cost ensembles in certain settings in Corollary 3 of [SH16] and in particular applies to our case here.

For the converse, we note that for any ensemble $\{p_X, \rho^x\}$,

$$I(X; B)_{\rho} - I(X; E)_{\rho} = \sum_x p_X(x)N_N(\rho^x, \bar{\rho}) \quad (5.13)$$

$$= \sum_x p_X(x)N_N(\rho^x, \psi^0) - N_N(\bar{\rho}, \psi^0) \quad (5.14)$$

$$\leq \sum_x p_X(x)N_N(\rho^x, \psi^0). \quad (5.15)$$

The inequality follows since by degradability of $\mathcal{N}$ and monotonicity of relative entropy [Lin75], $N_N(\rho, \sigma) \geq 0$ for all states $\rho$ and $\sigma$. If any of the $N_N$ quantities are infinite, the bound is trivial.

We can then argue for all $\beta > 0$,

$$\frac{P(N, \beta)}{\beta} = \frac{1}{\beta} \sup_{\{p_X, \rho^x\}} \frac{I(X; B) - I(X; E)}{\text{tr}[\psi^0]} \geq \frac{1}{\beta} \sup_{\{p_X, \rho^x\}} \sum_x p_X(x)N_N(\rho^x, \psi^0). \quad (5.16)$$

$$\leq \frac{1}{\beta} \sup_{\{p_X, \rho^x\}} \sum_x p_X(x)N_N(\psi^x, \psi^0). \quad (5.17)$$

Now, just as in the entanglement-assisted case, we can restrict the supremum to be taken over
positive-cost states $\psi^x$:

$$\frac{P(N,\beta)}{\beta} \leq \frac{1}{\beta} \sup_{\{p_X,\psi^x\}} \sum_x p_X(x) N_N(\psi^x,\psi^0)$$

$$= \sup_{\{p_X,\psi^x\}} \sum_x p_X(x) \frac{N_N(\psi^x,\psi^0) \langle \psi^x | G | \psi^x \rangle}{\beta}$$

$$\leq \sup_{\psi} \frac{N_N(\psi,\psi^0)}{\langle \psi | G | \psi \rangle}. \quad (5.18)$$

This concludes the proof. \qed

### 5.1 Pulse-Position-Modulation Scheme for Private Communication

We now give an alternative proof of the achievability part of Theorem 13 via a PPM scheme that achieves the private capacity per unit cost for a degradable channel $N_{A \to B}$.

**Codebook**: As discussed above, without loss of generality, we can restrict the discussion to positive-cost pure states $\psi$ such that $N_N(\psi,\psi^0) > 0$. Let $\psi$ be such a state. Then, fix $M, L, N \in \mathbb{N}$. We have $ML$ codewords labeled by $(m,l)$ where $m \in [1 : M]$ and $l \in [1 : L]$. For codeword $(m,l)$, the corresponding input quantum state is

$$\left( (\psi^0 \otimes N) \otimes L \right)^{\otimes m-1} \otimes \left( (\psi^0 \otimes N) \otimes L \otimes \psi \otimes N \otimes (\psi^0 \otimes N) \otimes L - l \right) \otimes \left( (\psi^0 \otimes N) \otimes L \right)^{\otimes M - m}. \quad (5.21)$$

This can be understood as a $\psi$-pulse amidst a baseline of $\psi^0$ states, which is itself a pulse amidst a $(\psi^0 \otimes L)$ baseline. Note that the cost of this codeword is $N \langle \psi | G | \psi \rangle$.

**Encoding**: The sender transmits the message $m$ to the receiver and uses $l$ to obscure the message on the eavesdropper’s side. Given message $m \in [1 : M]$, he uniformly chooses at random $l \in [1 : L]$ and transmits $N$ times the quantum state corresponding to $(m,l)$.

**Decoding**: The receiver performs $ML$ binary quantum hypothesis tests, using $N$ copies for each test, in order to determine the pulse position. Again this can be done with vanishing error provided that, for some $\delta > 0$,

$$\frac{\log_2(ML)}{N} < D(N(\psi)||N(\psi^0)) - \delta. \quad (5.22)$$

**Privacy**: Given the randomization over $l$, the state that the eavesdropper obtains is

$$\left( (N^c(\psi^0) \otimes N) \otimes L \right)^{\otimes m-1} \otimes \xi_L \otimes \left( (N^c(\psi^0) \otimes N) \otimes L \right)^{\otimes M - m}, \quad (5.23)$$

where

$$\xi_L \equiv \left[ \frac{1}{L} \sum_{l=1}^{L} (N^c(\psi^0) \otimes N)^{\otimes l-1} \otimes N^c(\psi) \otimes N \otimes (N^c(\psi^0) \otimes N)^{\otimes L - l} \right]. \quad (5.24)$$

17
Now, by a corollary to the convex-split lemma \[AJW17\], the state in eq. (5.23) is approximately $\mathcal{N}^{c}(\psi^{0}) \otimes NLM$ if $L$ is chosen large enough. More precisely, given $\delta', \varepsilon > 0$,

$$
\frac{1}{2} \left\| \mathcal{N}^{c}(\psi^{0}) \otimes NLM - \mathcal{N}^{c}(\psi^{0}) \right\|_{1} = \frac{1}{2} \left\| \mathcal{N}^{c}(\psi^{0}) \otimes NLM \right\|_{1} \leq 2\varepsilon + \delta' \tag{5.25}
$$

if

$$
L > 2^{D_{\max}^{\varepsilon}(\mathcal{N}^{c}(\psi) \otimes N^{c}(\psi^{0})) - 2\delta'}, \tag{5.27}
$$

where $D_{\max}^{\varepsilon}$ denotes the smooth max-relative entropy \[Dat09\]:

$$
D_{\max}^{\varepsilon}(\rho \parallel \sigma) \equiv \inf_{\tilde{\rho} \in B^{\varepsilon}(\rho)} D_{\max}(\tilde{\rho} \parallel \sigma). \tag{5.28}
$$

Here, $B^{\varepsilon}(\rho)$ denotes the $\varepsilon$-ball around $\rho$: 

$$
B^{\varepsilon}(\rho) \equiv \left\{ \tilde{\rho} \geq 0 : \sqrt{1 - F^{2}(\tilde{\rho}, \rho)} \leq \varepsilon, \text{tr}[\tilde{\rho}] = 1 \right\}, \tag{5.29}
$$

where $F$ is the quantum fidelity \[Uhl76\], and

$$
D_{\max}(\rho \parallel \sigma) \equiv \log \inf \{ \lambda \geq 0 : \rho \leq \lambda \sigma \}. \tag{5.30}
$$

Hence, using the quantum asymptotic equipartition property \[Tom16\] for smooth max-relative entropy, for small enough $\varepsilon$, all $\delta'' > 0$, and sufficiently large $N$, eq. (5.27) is satisfied if

$$
L > 2^{N(D(\mathcal{N}^{c}(\psi) \parallel \mathcal{N}^{c}(\psi^{0})) + \delta'')} - 2\delta'. \tag{5.31}
$$

Taking the logarithm on both sides and dividing by $N$, we obtain

$$
\frac{\log_{2} L}{N} > D(\mathcal{N}^{c}(\psi) \parallel \mathcal{N}^{c}(\psi^{0})) + \delta'' - \frac{2\log_{2} \delta'}{N}. \tag{5.32}
$$

For large enough $N$, the condition becomes

$$
\frac{\log_{2} L}{N} > D(\mathcal{N}^{c}(\psi) \parallel \mathcal{N}^{c}(\psi^{0})) + \delta'' \tag{5.33}
$$

for some $\delta''' > 0$.

We conclude that we can attain arbitrarily low decoding and privacy error if

$$
\frac{\log_{2} (ML)}{N} < D(\mathcal{N}(\psi) \parallel \mathcal{N}(\psi^{0})) - \delta \tag{5.34}
$$

and

$$
\frac{\log_{2} L}{N} > D(\mathcal{N}^{c}(\psi) \parallel \mathcal{N}^{c}(\psi^{0})) + \delta'''. \tag{5.35}
$$

\[\text{1}\] The original definition differs slightly from the definition in \[AJW17\], which is the one we use here.
Combining the two inequalities, we obtain
\[
\log_2 \frac{M}{N} < \frac{N_{\mathcal{N}}(\psi, \psi^0)}{\langle \psi | G | \psi \rangle} - \frac{\delta + \delta''}{\langle \psi | G | \psi \rangle}.
\] (5.36)

Hence, we can achieve a private rate per unit cost of \( \frac{N_{\mathcal{N}}(\psi, \psi^0)}{\langle \psi | G | \psi \rangle} \).

Note that the above protocol does not use the fact that the channel is degradable. Hence, this gives an achievability result for non-degradable quantum channels as well. Furthermore, we can achieve this with a mixed state \( \rho \) instead of a pure state \( \psi \), which might be necessary for non-degradable channels. We conclude for general quantum channels \( \mathcal{N} \),

\[
P(\mathcal{N}) \geq \sup_{\rho \in S(\mathcal{H}_A)} P_{\psi^0}(\mathcal{N}, \rho).
\] (5.37)

6 Quantum Communication

We now formulate the quantum capacity per unit cost. Quantum capacity was first studied in [SN96, Llo97, BKN00, BNS98, Sho02b, Dev05] when there is no cost constraint and recently in [WQ16] when there is a cost constraint. Let \( \mathcal{N}_{A \rightarrow B} \) be a quantum channel, and let \( n, Q \in \mathbb{N} \), \( \nu \in \mathbb{R}_{>0} \), and \( \varepsilon \in [0,1] \). An \((n, Q, \nu, \varepsilon)\) quantum code has blocklength \( n \), dimension \( Q \) for the total input space, total cost at most \( \nu \), and quantum decoding error at most \( \varepsilon \). In more detail, an \((n, Q, \nu, \varepsilon)\) code for quantum communication consists of an encoding channel \( E_n : S(\mathcal{H}_S) \rightarrow S(\mathcal{H}_A \otimes^n) \) and a decoding channel \( D_n : S(\mathcal{H}_B \otimes^n) \rightarrow S(\mathcal{H}_S) \), where \( \dim(\mathcal{H}_S) = Q \). The cost constraint imposes the following bound on all states resulting from the output of the encoding channel \( E_n \):

\[
\text{Tr} \{ G_n E_n(\rho_S) \} \leq \nu,
\] (6.1)

where \( \rho_S \in S(\mathcal{H}_S) \) and \( G_n \) is defined in eq. (3.2). Finally, we have the error bounded by \( \varepsilon \), in the sense that for all pure states \( \phi_{RS} \in S(\mathcal{H}_R \otimes \mathcal{H}_S) \), where \( \mathcal{H}_R \) is isomorphic to \( \mathcal{H}_S \), the following trace distance bound holds:

\[
\frac{1}{2} \| \phi_{RS} - (\text{id}_R \otimes N_A \rightarrow B \otimes^n \circ \mathcal{E}^n)(\phi_{RS}) \|_1 \leq \varepsilon.
\] (6.2)

**Definition 14.** Given \( \beta > 0 \), \( R_q \) is an achievable quantum communication rate with average cost not exceeding \( \beta \) if for all \( \varepsilon \in (0,1) \) and \( \delta > 0 \), \( \exists n_0 \) such that for \( n \geq n_0 \), \( \exists \) an \((n, Q, n\beta, \varepsilon)\) code for which

\[
\frac{\log_2 Q}{n} > R_q - \delta.
\] (6.3)

The quantum capacity cost function \( Q(\mathcal{N}, \beta) \) is equal to the supremum of all achievable quantum communication rates with average cost not exceeding \( \beta \).

Building on [DS05], the expression for \( Q(\mathcal{N}, \beta) \) when \( \mathcal{N} \) is degradable was shown in [WQ16] to be the following.

**Theorem 15 ([WQ16]).** The quantum capacity cost function for a degradable channel \( \mathcal{N}_{A \rightarrow B} \) is given by

\[
Q(\mathcal{N}, \beta) = \sup_{\varphi_{RA} \leq \beta} I(R|B)_\varphi,
\] (6.4)

where \( I(R|B) \equiv S(B) - S(RB) \) is the coherent information [SN96] and

\[
\varphi_{RB} \equiv (\text{id}_R \otimes N_{A \rightarrow B})(\varphi_{RA}).
\] (6.5)
We then define the quantum capacity per unit cost.

**Definition 16.** $R_q$ is an achievable quantum communication rate per unit cost if for any $\delta, \varepsilon > 0$, $\exists \nu_0 > 0$ such that for $\nu \geq \nu_0$ there is an $(n, Q, \nu, \varepsilon)$ code for which

$$\log_2 Q > \nu (R_q - \delta).$$  \hspace{1cm} (6.6)

The quantum capacity per unit cost is then defined to be the supremum of all achievable quantum communication rates per unit cost and is denoted as $Q(N)$.

Using the achievability and converse proofs for $Q(N, \beta)$, and reasoning similar to that in the proof of Theorem 4, we obtain the following expressions for $Q(N)$.

**Theorem 17.** The quantum capacity per unit cost for a degradable channel $N_{A \rightarrow B}$ is given by

$$Q(N) = \sup_{\beta > 0} \frac{Q(N, \beta)}{\beta} = \sup_{\varphi_{RA}} \frac{I(R)B|_\varphi}{\text{tr}[G\varphi_A]}.$$  \hspace{1cm} (6.7)

However, note that since $Q(N, \beta) = P(N, \beta)$ \cite{Smi08, WQ16} for degradable channels, in this case the quantum capacity per unit cost is equal to the private capacity per unit cost. In particular, if a zero-cost state exists, then the quantum capacity per unit cost of a degradable channel is given by the expression in Theorem 13.

### 6.1 Pulse-Position-Modulation Scheme for Quantum Communication

We propose a PPM scheme for achieving the quantum capacity per unit cost for a degradable channel $N_{A \rightarrow B}$. We do this by operating the PPM scheme for private communication in a coherent fashion analogous to that of [Dev05]. Since the approach is so similar (yet tailored to a PPM coding scheme), we merely sketch the proof for simplicity and point to [Dev05] for more details (see also [DW04] in this context). The task we consider is entanglement generation, in which the goal is to establish a maximally entangled state between the sender and receiver. To generalize this to arbitrary quantum states, we again point to [Dev05]. Let $U_{A \rightarrow BE}$ be an isometric channel extending $N_{A \rightarrow B}$ and let $U_{A \rightarrow BE}$ denote the corresponding isometry. Let $\psi^0$ be a zero-cost pure state and $\psi$ a positive-cost pure state for which $N_{N}(\psi, \psi^0) > 0$.

We first consider the case in which $\psi$ is orthogonal to $\psi^0$.

**Codebook:** As in the private PPM scheme, we fix $M, L, N \in \mathbb{N}$. For each ordered pair $(m, l)$, where $m \in [1 : M]$ and $l \in [1 : L]$, consider the following pure quantum state:

$$\psi^{m,l} \equiv \left[\left(\left(\psi^0 \otimes N\right) \otimes L\right)^{\otimes m-1} \otimes \left(\left((\psi^0 \otimes N) \otimes N \otimes \left((\psi^0 \otimes N) \otimes L\right)^{-1}\right) \otimes \left((\psi^0 \otimes N) \otimes L\right)^{\otimes M-m}\right) \right] \hspace{1cm} (6.8)$$

Note that since $\psi$ is orthogonal to $\psi^0$, $\psi^{m,l}$ are orthogonal for different $m, l$. Observe also that since $G|\psi^0\rangle = 0$,

$$G_K|\psi^{m,l}\rangle_{AK} = \sum_{j=n(m,l)}^{n(m,l)+N-1} \left(I^{\otimes j-1} \otimes G \otimes I^{\otimes K-j}\right)|\psi^{m,l}\rangle_{AK},$$  \hspace{1cm} (6.9)

where $K \equiv NML$ and $n(m,l) \equiv NL(m-1) + N(l-1) + 1$ is the position of the first $\psi$ state in $\psi^{m,l}$.
**Encoding:** Let $|\Phi\rangle_{RA}$ denote the maximally entangled state to be established with the receiver, where the dimension of $\hat{A}$, the system the sender is to transmit, is at most $M$. We can decompose the state with respect to some orthonormal bases $\{|m\rangle_R\}$ and $\{|m\rangle_{\hat{A}}\}$:

$$|\Phi\rangle_{RA} = \frac{1}{\sqrt{M}} \sum_m |m\rangle_R |m\rangle_{\hat{A}}. \quad (6.10)$$

Depending on the value of $m$, the sender coherently prepares a uniform superposition of $|\psi^{m,l}\rangle$ over the $l$ variable. That is, the mapping is given by

$$|m\rangle_{\hat{A}} \mapsto \frac{1}{\sqrt{L}} \sum_{l=1}^{L} |\psi^{m,l}\rangle_{AK}. \quad (6.11)$$

Since $\psi^{m,l}$ are orthogonal for different $m,l$, the above mapping is an isometry. The overall state after the encoding is

$$|\Psi\rangle_{RAK} = \frac{1}{\sqrt{ML}} \sum_{m,l} |m\rangle_R |\psi^{m,l}\rangle_{AK}. \quad (6.12)$$

Note that the reduced state on $A^K$ is

$$\Psi_{AK} = \frac{1}{ML} \sum_{m,l,l'} |\psi^{m,l}\rangle \langle \psi^{m,l'}|_{AK}$$

and therefore has cost $N \langle \psi | G | \psi \rangle$ by eq. (6.9).

**Decoding:** After $K$ uses of the isometric extension, the overall state is

$$|\Psi\rangle_{RB^KE^K} = \frac{1}{\sqrt{ML}} \sum_{m,l} |m\rangle_R |\psi^{m,l}\rangle_{BK^KE^K}, \quad (6.14)$$

where $|\psi^{m,l}\rangle_{BK^KE^K} \equiv U_{A \rightarrow BE}^{\otimes K} |\psi^{m,l}\rangle_{AK}$. Let $\{\Lambda_{m,l}^{m,l'}\}_{m,l}$ denote the POVM used as a decoder in the private communication protocol. A coherent version of this measurement is given by the isometry $V_{B^K \rightarrow BK^K \hat{B}_0 \hat{B}_1} = \sum_{m,l} \sqrt{\Lambda_{m,l}^{m,l}} |m\rangle_{B_0} \otimes |l\rangle_{B_1}$, and after performing it, the resulting state is approximately equal to the following one:

$$|\Psi\rangle_{RB^KE^K \hat{B}_0 \hat{B}_1} = \frac{1}{\sqrt{ML}} \sum_{m,l} |m\rangle_R |\psi^{m,l}\rangle_{BK^KE^K} |m\rangle_{\hat{B}_0} |l\rangle_{\hat{B}_1}. \quad (6.15)$$

At this point, we know from the privacy condition for the private code, that for each $m$, the following approximation holds: $\frac{1}{L} \sum_l |\psi^{m,l}_{E^K}\rangle \approx N^c |(\psi^0)^{\otimes K}\rangle$. Thus, given $m$, the eavesdropper’s system is approximately independent of $m$. By Uhlmann’s theorem [Uhl76], for each $m$, there exists an isometry $W_{B^K \hat{B}_1 \rightarrow \hat{B}_2}^m$ such that

$$W_{B^K \hat{B}_1 \rightarrow \hat{B}_2}^m \left[ \frac{1}{\sqrt{L}} \sum_l |\psi^{m,l}_{B^KE^K}\rangle |l\rangle_{\hat{B}_1} \right] \approx |\varsigma\rangle_{E^K \hat{B}_2}, \quad (6.16)$$
where $|\psi\rangle_{E^K B^2}$ is some state independent of $m$. So this means that the receiver can perform the controlled isometry $\sum_m |m\rangle_B \otimes W^m_{B^K B^1 \rightarrow B^2}$, and the resulting state is approximately close to the following state:

$$\frac{1}{\sqrt{\lambda}} \sum_m |m\rangle_R \otimes |\psi\rangle_{E^K B^2}. \quad (6.17)$$

By the properties of the PPM scheme for private communication, the quantum rate per unit cost of this scheme is equal to $N_X(\psi, \psi^0) / \langle \psi | G | \psi \rangle$.

Now, if $\psi$ is not orthogonal to $\psi^0$, we implement the above protocol but replacing $\psi \otimes N$ with its normalized rejection from $(\psi^0) \otimes N$. That is, we take the component of $|\psi\rangle \otimes N$ orthogonal to $|\psi^0\rangle \otimes N$ and normalize it. Calling this $\psi^{1}_A N$, we have

$$|\psi^{1}_A\rangle = \frac{1}{\sqrt{1 - |\langle \psi^0 | \psi\rangle|^2 N}} \left(|\psi\rangle \otimes N - \langle \psi^0 | \psi\rangle^N |\psi^0\rangle \otimes N\right), \quad (6.18)$$

and so

$$\frac{1}{2} \norm{\psi^{1}_A - (\psi \otimes N)}_1 = \norm{\langle \psi^0 | \psi\rangle^N \equiv \delta_N. \quad (6.19)$$

By monotonicity, the trace distance is at most $\delta_N$ after $N$ uses of $\mathcal{N}$ or $\mathcal{N}^c$. Hence, since $\psi$ is positive-cost and so $|\langle \psi^0 | \psi\rangle| < 1$, by using $\psi^{1}_A$ we expect to obtain the desired rate in the limit of large $N$. Indeed, since $\psi^{1}_A$ is orthogonal to $(\psi^0) \otimes N$, we can implement the above protocol and achieve a quantum rate per unit cost

$$D_{H}^\epsilon \left(\mathcal{N}^{\otimes N}(\psi^{1}_A) \| \mathcal{N}(\psi^0) \otimes N\right) - D_{\text{max}}^\epsilon \left(\mathcal{N}^c(\psi^{1}_A) \| \mathcal{N}^{c}(\psi^0) \otimes N\right) \quad (6.20)$$

for any $\epsilon, \epsilon' > 0$, where $D_{H}^\epsilon$ is the hypothesis testing relative entropy [WR12, BD10]:

$$D_{H}^\epsilon (\rho \| \sigma) = -\log_2 \inf_{0 \leq \Lambda \leq I} \tr[\Lambda \sigma]. \quad (6.21)$$

Note that this is simply the negative logarithm of the quantity defined in eq. (2.3). For sufficiently large $N$, we have that $\delta_N < \epsilon$. Then, by Lemma 7 in [DTW16],

$$D_{H}^\epsilon \left(\mathcal{N}^{\otimes N}(\psi^{1}_A) \| \mathcal{N}(\psi^0) \otimes N\right) \geq D_{H}^{-\delta_N} \left(\mathcal{N}(\psi^0) \otimes N \| \mathcal{N}(\psi^0) \otimes N\right) \geq D_{H}^\epsilon \left(\mathcal{N}(\psi^0) \otimes N \| \mathcal{N}^{c}(\psi^0) \otimes N\right), \quad (6.22)$$

where $\delta(\epsilon) = \epsilon - \delta_N(\epsilon)$ and $N(\epsilon)$ is the smallest integer such that $\delta_N(\epsilon) < \epsilon$. Furthermore, by the definition of smooth max-relative entropy in eq. (5.28) and the triangle inequality, for sufficiently large $N$ such that $\delta_N < \epsilon'$, we have

$$D_{\text{max}}^\epsilon(\mathcal{N}^c(\psi^0) \otimes N) \leq D_{\text{max}}^\epsilon(\mathcal{N}^c(\psi^0) \otimes N) \leq D_{\text{max}}^{\delta(\epsilon')} \left(\mathcal{N}^{c}(\psi^0) \otimes N \| \mathcal{N}^{c}(\psi^0) \otimes N\right). \quad (6.23)$$

Lastly,

$$\langle \psi^{1}_A | G_N | \psi^{1}_A \rangle = \frac{1}{1 - |\langle \psi^0 | \psi\rangle|^2 N} \langle \psi^{1}_A \otimes N | G_N | \psi^{1}_A \otimes N \rangle = \frac{N \langle \psi | G | \psi \rangle}{1 - |\langle \psi^0 | \psi\rangle|^2 N}. \quad (6.24)$$

We conclude that we can achieve

$$\left(1 - |\langle \psi^0 | \psi\rangle|^2 N\right) \frac{D_{H}^\epsilon \left(\mathcal{N}(\psi) \otimes N \| \mathcal{N}^{c}(\psi) \otimes N\right) - D_{\text{max}}^{\delta(\epsilon')} \left(\mathcal{N}^{c}(\psi) \otimes N \| \mathcal{N}^{c}(\psi^0) \otimes N\right)}{N \langle \psi | G | \psi \rangle}. \quad (6.25)$$
By Quantum Stein’s Lemma (eq. (2.4)) and the quantum asymptotic equipartition property for smooth max-relative entropy, in the limit of large $N$ we can therefore achieve a quantum rate per unit cost of $N_{N(\psi,\psi^0)}^{(\psi|\mathcal{G}|\psi)}$.

Note that just like in the private case, this scheme gives achievability for non-degradable channels as well.

### 7 Capacities per Unit Cost of Quantum Gaussian Channels

By the methods developed and used in [Hol03, Hol04, Hol12, WQ16], we can generalize the above results to infinite dimensions, in particular for quantum Gaussian channels. We will use the formulas derived above to compute various capacities per unit cost for specific Gaussian channels, where the cost observable is the photon number operator $\hat{n} = \sum_{n=0}^{\infty} n|n\rangle\langle n|$, with $|n\rangle$ being a photon number state. Since the channels we consider already have known capacity cost functions, it is easiest for us to compute the capacity per unit cost via the following formula:

$$
C(\mathcal{N}) = \lim_{\bar{n} \to 0} C(\mathcal{N}, \bar{n}) / \bar{n}.
$$

Note that we could also compute the capacities per unit cost using the optimized relative entropy formulas that we obtained in the previous sections.

The quantum Gaussian channels we consider are the following [Hol02, HW01, GGPCH14]. The thermal channel $\mathcal{E}_{n}^{n_{th}}$ with transmissivity $\eta \in (0, 1)$ and thermal photon number $n_{th} \in \mathbb{R}_{>0}$ is a Gaussian channel which mixes the input signal with a thermal Gaussian state. This is summarized by the following Heisenberg evolution:

$$
a_{in} \mapsto \sqrt{\eta}a_{in} + \sqrt{1-\eta}a_{th},
$$

where $a_{in}$, $a_{th}$ are the annihilation operators for the input and thermal modes, respectively. We also consider the additive classical noise channel $\mathcal{N}_N$ with variance $\bar{N} \in \mathbb{R}_{>0}$, which describes classical noise which displaces the signal in phase space according to a Gaussian distribution. The Heisenberg evolution is

$$
a_{in} \mapsto a_{in} + \xi,
$$

where $\xi$ is a complex normal random variable with mean zero and variance $\bar{N}$. Next, the amplifier channel $\mathcal{A}_{n}^{n_{th}}$ with gain parameter $\kappa > 1$ and thermal photon number $n_{th} \in \mathbb{R}_{>0}$ describes the effect of a two-mode squeezing Hamiltonian which acts on the input mode and a thermal mode. This effectively amplifies the input signal but at the cost of adding noise. The resulting Heisenberg evolution is given by

$$
a_{in} \mapsto \sqrt{\kappa}a_{in} + \sqrt{\kappa-\kappa}a_{th}^\dagger.
$$

Lastly, we also look at the weak conjugate of the amplifier channel, the contravariant amplifier $\tilde{\mathcal{A}}_{n}^{n_{th}}$ with gain parameter $\kappa > 1$, and thermal photon number $n_{th} \in \mathbb{R}_{>0}$.

### 7.1 Classical Communication over Gaussian Channels

We compute the unassisted classical capacities per unit cost for these four channels. For the thermal channel, the classical capacity cost function is given by [GHGP15, GGPCH14]

$$
C(\mathcal{E}_{n}^{n_{th}}, \bar{n}) = g(\eta \bar{n} + (1-\eta)n_{th}) - g((1-\eta)n_{th}),
$$

23
where
\[ g(x) \equiv (x + 1) \log_2(x + 1) - x \log_2 x. \] (7.6)

Hence, by applying eq. (7.1), we find the following:
\[ C(\mathcal{E}_{\eta}^{n_{th}}) = \eta \log_2 \left( 1 + \frac{1}{n_{th}(1 - \eta)} \right). \] (7.7)

Now, in the limit \( n_{th} \to 0 \), the thermal channel reduces to the pure-loss bosonic channel \( \mathcal{L}_\eta [GGL+04] \), which has classical capacity
\[ C(\mathcal{L}_\eta, \bar{n}) = g(\eta \bar{n}). \] (7.8)

With a simple calculation we find that \( C(\mathcal{L}_\eta) \) is infinite, an observation made in [GDS11] and earlier in the context of infinite bandwidth [LL66, CD94, YKL75]\(^2\). One way to make sense of this infinity is to look at \( C(\mathcal{E}_{\eta}^{n_{th}}) \) for small \( n_{th} \):
\[ C(\mathcal{E}_{\eta}^{n_{th}}) = -\eta \log_2(n_{th}(1 - \eta)) + O(n_{th}). \] (7.9)

This is finite for any \( n_{th} > 0 \), but diverges as \(-\log_2(n_{th})\) as \( n_{th} \to 0 \). Since \( n_{th} \) physically corresponds to temperature, this suggests that the infinity comes about since zero temperature is unphysical. It also implies that the capacity per unit cost can be increased arbitrarily by going to lower temperatures. We find that a toy model of a classical binary channel reproduces this qualitative difference between the zero and non-zero temperature cases and the logarithmic scaling of the capacity per unit cost in the noise parameter of the channel:

\[ C(B_{\epsilon, \delta}) = -(1 - \epsilon) \log_2(\delta) - \epsilon \log_2(1 - \delta) - h(\epsilon), \] (7.11)

where \( h(x) \equiv -x \log_2(x) - (1 - x) \log_2(1 - x) \) is the binary entropy function. Thus the capacity per unit cost of the classical binary channel \( B_{\epsilon, \delta} \) diverges as \(-\log_2(\delta)\) as \( \delta \to 0 \), analogous to the behavior of (7.9) as \( n_{th} \to 0 \).

\(^2\)For a connection between infinite bandwidth capacity and capacity per unit cost, see [Ver90].
Now, consider the additive-noise channel, which has the following classical capacity cost function [GHGP15, GGPCH14]

\[ C(\mathcal{N}, \bar{n}) = g(\bar{n} + N) - g(N). \] (7.12)

Hence,

\[ C(\mathcal{N}_N) = \log_2 \left( 1 + \frac{1}{N} \right). \] (7.13)

Next, we consider the amplifier channel, which has a classical capacity cost function [GHGP15, GGPCH14]

\[ C(\mathcal{A}_N^{n_{th}}, \bar{n}) = g(\kappa \bar{n} + (\kappa - 1)(n_{th} + 1)) - g((\kappa - 1)(n_{th} + 1)). \] (7.14)

We find

\[ C(\mathcal{A}_N^{n_{th}}) = \kappa \log_2 \left( 1 + \frac{1}{(\kappa - 1)(n_{th} + 1)} \right). \] (7.15)

Finally, for the contravariant amplifier channel, which has the capacity cost function [GGPCH14]

\[ C(\tilde{\mathcal{A}}_N^{n_{th}}, \bar{n}) = g(\kappa n_{th} + (\kappa - 1)(\bar{n} + 1)) - g(\kappa(n_{th} + 1) - 1), \] (7.16)

we have

\[ C(\tilde{\mathcal{A}}_N^{n_{th}}) = (\kappa - 1) \log_2 \left( 1 + \frac{1}{\kappa(n_{th} + 1) - 1} \right). \] (7.17)

Note that given eq. (3.31) and the achievability result in Theorem 4, when a zero-cost state exists, we can achieve the capacity per unit cost with any code that achieves the cost-constrained capacity in the limit of zero cost. For example, for the pure-loss bosonic channel, single-photon-detection achieves the classical capacity in the limit \( \bar{n} \rightarrow 0 \) [TG14]. In general, we can achieve the capacity per unit cost with any code that achieves \( C(\beta_{\text{max}}) \) where \( \beta_{\text{max}} = \arg \sup_{\beta > 0} C(\beta)/\beta \).

### 7.2 Entanglement-Assisted Communication over Gaussian Channels

We now compute the entanglement-assisted capacity per unit cost for the first three channels. For the thermal channel, the entanglement-assisted capacity cost function is given by [HW01, HG12]

\[
C_{\text{EA}}(\mathcal{E}_\eta^{n_{th}}, \bar{n}) = g(\bar{n}) + g(\eta \bar{n} + (1 - \eta)n_{th})
- g\left(\frac{1}{2} \left( \sqrt{((1 + \eta)\bar{n} + (1 - \eta)n_{th} + 1)^2 - 4\eta\bar{n}(\bar{n} + 1) - (1 - \eta)(\bar{n} - n_{th} - 1)} \right) \right)
- g\left(\frac{1}{2} \left( \sqrt{((1 + \eta)\bar{n} + (1 - \eta)n_{th} + 1)^2 - 4\eta\bar{n}(\bar{n} + 1) + (1 - \eta)(\bar{n} - n_{th} - 1)} \right) \right). \] (7.18)

We compute the limit as per eq. (7.1) and find that the entanglement-assisted capacity per unit cost of the thermal channel diverges\(^3\). For the additive noise channel, the entanglement-assisted capacity cost function is [HW01, HG12]

\[
C_{\text{EA}}(\mathcal{N}, \bar{n}) = g(\bar{n}) + g(\bar{n} + N) - g\left(\frac{1}{2} \left( \sqrt{(N + 1)^2 + 4N\bar{n} - N - 1} \right) \right)
- g\left(\frac{1}{2} \left( \sqrt{(N + 1)^2 + 4N\bar{n} + N - 1} \right) \right), \] (7.19)

\(^3\)For many of these calculations, see the Mathematica file included in arXiv posting.
Thermal Classical Noise Amplifier

Figure 1: The figure illustrates the divergence of $C_{EA}(\mathcal{N}, \bar{n})/\bar{n}$ as $\bar{n} \to 0$. For the thermal channel we set $n_{th} = 10$, $\eta = 0.7$, for the classical noise channel $N = 10$, and for the amplifier channel $n_{th} = 10$, $\kappa = 1.3$.

which also gives an infinite capacity per unit cost. Lastly, for the amplifier channel,

$$C_{EA}(A_{\kappa}^{n_{th}}, \bar{n}) = g(\bar{n}) + g(\kappa\bar{n} + (\kappa - 1)(n_{th} + 1))$$

$$- g\left(\frac{1}{2}\left(\sqrt{((\kappa + 1)\bar{n} + (\kappa - 1)(n_{th} + 1) + 1)^2 - 4\kappa\bar{n}(\bar{n} + 1)} - (\kappa - 1)(\bar{n} + n_{th} + 1) - 1\right)\right)$$

$$- g\left(\frac{1}{2}\left(\sqrt{((\kappa + 1)\bar{n} + (\kappa - 1)(n_{th} + 1) + 1)^2 - 4\kappa\bar{n}(\bar{n} + 1) + (\kappa - 1)(\bar{n} + n_{th} + 1) - 1}\right)\right).$$ (7.20)

Computing the ratio again gives an infinite value as $\bar{n} \to 0$.

We show the divergence of $C_{EA}$ by plotting the bits per photon against the number of photons for these three Gaussian channels in fig. 1. It is likely that the divergences for these Gaussian channels appear since we are allowing for an unbounded amount of entanglement assistance per unit cost. However, it is not true that this potentially infinite entanglement assistance always leads to divergences. Trivial examples include channels where $C_{EA}(\mathcal{N}, \beta) = 0$ for all $\beta$ and when the cost observable is positive definite. We can also find examples using the fact that entanglement assistance sometimes does not help. Take for instance the state preparation qubit channel

$$\mathcal{N}_s : \rho \mapsto \langle 0| \rho |0\rangle \rho^0 + \langle 1| \rho |1\rangle \rho^1$$ (7.21)

with cost observable $G = |1\rangle\langle 1|$, where $\rho^0, \rho^1 \in \mathcal{S}(\mathcal{H}_A)$ such that $0 < D(\rho^1||\rho^0) < \infty$. By Proposition 4 in [Shi12], $C_{EA}(\mathcal{N}_s, \beta) = C(\mathcal{N}_s, \beta)$ for all $\beta$. Now, we have a zero-cost state $|0\rangle\langle 0|$, so we can use
Figure 2: Plot of $P(N, \bar{n})/\bar{n} = Q(N, \bar{n})/\bar{n}$ versus $\bar{n}$. For the ideal amplifier we set $\kappa = 3$ and for the pure-loss channel $\eta = 0.7$.

Theorem 5:

$$C(N_s) = \sup_{\psi \neq |0\rangle \langle 0|} \frac{D(N_s(\psi)||N_s(|0\rangle \langle 0|))}{\langle \psi | G | \psi \rangle}$$

$$= \sup_{\psi \neq |0\rangle \langle 0|} \frac{\left(|c_0|^2 \rho^0 + |c_1|^2 \rho^1\right)}{|c_1|^2}$$

$$\leq \sup_{\psi \neq |0\rangle \langle 0|} \frac{1}{|c_1|^2} \left(|c_0|^2 D(\rho^0 || \rho^0) + |c_1|^2 D(\rho^1 || \rho^0)\right)$$

$$= D(\rho^1 || \rho^0)$$

$$< \infty,$$

where $|\psi\rangle \equiv c_0 |0\rangle + c_1 |1\rangle$ and we can divide by $|c_1|^2$ since $\psi \neq |0\rangle \langle 0|$. The inequality follows from the convexity of the relative entropy in the first argument. We conclude $C_{EA} (N_s) = C(N_s) < \infty$. However, since $D(\rho^1 || \rho^0) > 0$, this channel clearly has a non-zero $C(N_s, \beta)$ and so non-zero $C_{EA} (N_s, \beta)$. It is an interesting open question to find an explicit nontrivial channel for which there is a gain from entanglement assistance but $C_{EA}$ is finite.

7.3 Private and Quantum Communication over Gaussian Channels

We next compute the private and quantum capacities per unit cost for degradable Gaussian channels. In particular, we consider ideal amplifiers $A_0^0$ and pure-loss channels $E_0^0$ for $\eta \leq 1/2$. For the former, we have the capacity cost function [WQ16]

$$P(A_0^0, \bar{n}) = Q(A_0^0, \bar{n}) = g(\kappa(\bar{n} + 1) - 1) - g((\kappa - 1)(\bar{n} + 1)).$$

Hence,

$$P(A_0^0) = Q(A_0^0) = \log_2 \left( \frac{\kappa}{\kappa - 1} \right),$$

where $g(x) = x \log_2 e - (1/x)$. For pure-loss channels,

$$P(E_0^0, \bar{n}) = Q(E_0^0, \bar{n}) = g(\eta(\bar{n} + 1) - 1) - g((\eta - 1)(\bar{n} + 1)).$$

Hence,

$$P(E_0^0) = Q(E_0^0) = \log_2 \left( \frac{\eta}{\eta - 1} \right),$$

where $\eta = \frac{2}{\gamma + 1}$.

27
which demonstrates that the quantum and private capacity per unit cost are equal to the unconstrained quantum and private capacity of the ideal amplifier channel [WPGG07, WQ16]. Finally, for the pure-loss channel, we have that [WQ16]

\[ P(\mathcal{E}_\eta^0, \bar{n}) = Q(\mathcal{E}_\eta^0, \bar{n}) = g(\eta \bar{n}) - g((1 - \eta)\bar{n}). \]  

(7.29)

We divide by \( \bar{n} \) and find the following divergent term for \( \eta > 1/2 \) as \( \bar{n} \to 0 \):

\[ (2\eta - 1) \log_2(1/\bar{n}). \]  

(7.30)

Hence, the pure-loss channel for \( \eta > 1/2 \) has an infinite private and quantum capacity per unit cost, while for \( \eta \leq 1/2 \) the capacities per unit cost are zero since the channel is antidegradable in this regime. In fig. 2 we plot the private bits or qubits per photon against the number of photons for these two channels.

8 Blocklength-Constrained Capacity Per Unit Cost

In Section 7, we find that there are several channels with infinite capacities per unit cost. Although we are able to ascribe these infinities to unphysical assumptions such as zero temperature and unbounded entanglement assistance, we can alternatively ascribe almost any infinite capacity per unit cost to the absence of a time constraint. It is important to notice that although Definition 3 is very similar to Definition 1, there is a significant asymmetry: while the latter has a constraint on the cost divided by blocklength, the former does not have an analogous constraint. Hence, just as cost constraints tamed infinite capacities, we likewise expect that blocklength constraints would tame infinite capacities per unit cost.

Indeed, we can show that the former implies the latter. We first define the blocklength-constrained capacity per unit cost. For concreteness, we state everything in the unassisted classical communication setting.

**Definition 18.** A non-negative number \( R \) is an achievable rate per unit cost with blocklength constraint \( \alpha \in \mathbb{R}_{>0} \) if for every \( \varepsilon \in (0,1), \delta > 0, \exists \nu_0 > 0 \) such that if \( \nu \geq \nu_0 \), \( \exists \) an \((n, M, \nu, \varepsilon)\) code such that \( \log_2 M > \nu(R - \delta) \) and \( n \leq \nu \alpha \). The capacity per unit cost with blocklength constraint \( \alpha \) is the supremum of all achievable rates per unit cost with constraint \( \alpha \), denoted as \( C(\mathcal{N}, \alpha) \).

Conveniently, we can show that \( C(\mathcal{N}, \alpha) \) has a characterization similar to that of \( C(\mathcal{N}) \). The proof of this theorem uses results from the proof of Theorem 4 and that of the corollary in [Ver90].

**Theorem 19.** For blocklength constraint \( \alpha > 0 \), the blocklength-constrained capacity per unit cost of a quantum channel \( \mathcal{N} \) is given by

\[ C(\mathcal{N}, \alpha) = \sup_{\beta \geq \frac{1}{\alpha}} \frac{C(\mathcal{N}, \beta)}{\beta}. \]  

(8.1)

**Proof.** We first show achievability. Let \( \beta > 0 \). By definition, we can achieve a rate \( C(\mathcal{N}, \beta) \) with cost \( \beta \). Let \( \varepsilon \in (0,1), \delta > 0 \). Then, using the direct part in the proof of Theorem 4, \( \exists \nu_0 \) such that for \( \nu \geq \nu_0 \) we can find an \((n, M, \nu, \varepsilon)\) code where

\[ \frac{\log_2 M}{\nu} > \frac{C(\mathcal{N}, \beta)}{\beta} - \delta \]  

(8.2)
and
\[ n \leq \nu \frac{\beta}{\beta}. \quad (8.3) \]
Hence, for \( \beta \geq 1/\alpha \), we achieve a rate per unit cost \( C(\mathcal{N}, \beta) / \beta \) with blocklength constraint \( \frac{1}{\beta} \leq \alpha \).

We now show the converse. We claim that if we can achieve a rate per unit cost \( \bar{R} \) with blocklength constraint \( \alpha \), then we can achieve a rate \( R / \alpha \) at cost \( 1/\alpha \).

To see this, let \( \varepsilon \in (0, 1), \delta > 0 \). Then, by Definition 18, there exists \( \nu_0 \) such that for all \( \nu \geq \nu_0 \), there is an \((n, M, \nu, \varepsilon)\) code where \( n \leq \nu\alpha \) and
\[ \log_2 M > R - \nu \alpha - \delta. \quad (8.4) \]

Thus, this is an \((n, M, \nu, \varepsilon)\) code that satisfies
\[ \frac{\log_2 M}{\nu} \geq \frac{R}{\alpha} - \delta. \quad (8.5) \]

Let \( n_0 \equiv \nu_0 \alpha \) and \( n \geq n_0 \). Then, let \( \nu = n / \alpha \). This implies \( \nu \geq n_0 / \alpha = \nu_0 \), so the above code with this \( \nu \) achieves a rate per channel use \( \bar{R}/\alpha \) and also has an average cost \( 1/\alpha \). This establishes the claim.

We now proceed by contradiction. Suppose we can achieve a rate per unit cost \( \bar{R} > \sup_{\beta \geq 1/\alpha} \frac{C(\mathcal{N}, \beta)}{\beta} \) with constraint \( \alpha \). Then, by the claim we can achieve the following rate with cost \( 1/\alpha \):
\[ \frac{R}{\alpha} > \sup_{\beta \geq 1/\alpha} \frac{C(\mathcal{N}, \beta)}{\beta \alpha} \geq C(\mathcal{N}, 1/\alpha). \quad (8.6) \]

Since this contradicts the definition of the capacity cost function, the converse follows.

In particular, when there is a zero-cost state, by concavity \( C(\mathcal{N}', \beta) / \beta \) is monotone non-increasing on \((0, +\infty)\) and so
\[ C(\mathcal{N}', \alpha) = \alpha C(\mathcal{N}', 1/\alpha). \quad (8.7) \]

This therefore removes the infinities, assuming that the capacity cost function does not diverge at finite cost. In particular, the classical capacity per unit cost for the pure-loss channel and the entanglement-assisted capacities per unit cost for the Gaussian channels we consider are finite if we impose a blocklength constraint. Also note that eq. (8.7) gives an operational interpretation of the ratio of the cost-constrained capacity to the cost: When there is a zero-cost state, for any \( \beta > 0 \), \( C(\mathcal{N}', \beta) / \beta \) is the capacity per unit cost with blocklength constraint \( 1/\beta \). Furthermore, by inverting eq. (8.7), we obtain an operational interpretation of the ratio of the blocklength-constrained capacity per unit cost to the blocklength constraint.

Interestingly, even in the general case where there might not be a zero-cost state, we can express the cost-constrained capacity in terms of the blocklength-constrained capacity per unit cost. Indeed, by the same proof of Theorem 19 except simply switching the achievability and the converse, we can show that
\[ C(\mathcal{N}, \beta) = \sup_{\alpha \geq 1/\beta} \frac{C(\mathcal{N}, \alpha)}{\alpha}. \quad (8.8) \]

This expression and Theorem 19 therefore establish a duality between the two quantities.
9 Discussion

In this paper we generalized the notion of capacity per unit cost to quantum channels for the tasks of classical communication, entanglement-assisted classical communication, private communication, and quantum communication. There are some simple extensions of our results that hold but for which we do not provide details. First is to consider non-additive quantum channels, for which we will find regularized expressions of the formulas given in Theorem 4, Theorem 5, Theorem 12, and Theorem 17. For example, for the regularized classical capacity per unit cost we obtain

\[ C(N) = \lim_{n\to\infty} \sup_{\beta>0} \frac{C(N^{\otimes n}, n\beta)}{n\beta}. \]  

(9.1)

Next, for private capacity per unit cost, we could also consider, as in [CWY04], the more general case of a quantum wiretap channel \( N_{A\to BE} \) where \( N \) is not necessarily an isometry. The results we found in Section 5 apply directly to the more general case of a degraded quantum wiretap channel.

We think there are a number of interesting directions for future work. One direction is to consider non-degradable quantum channels with a zero-cost state. As discussed earlier, a lower bound for the private capacity per unit cost in this case is given by our private PPM scheme and for classical channels in [EHLG13], but the converse remains open. It would also be interesting to give an expression for the quantum capacity per unit cost in terms of an optimized relative entropy. Another possible direction is to prove a strong converse for the capacity per unit cost. However, as was shown in [BGPWW15], the strong converse holds for Gaussian channels with an approximate peak cost constraint but not an average cost constraint. Hence, for Gaussian channels we cannot directly prove a strong converse for the classical capacity per unit cost in the same way as in Theorem 4. Lastly, we can consider the regime of finite blocklength or finite-sized measurement blocks and see how much information can be sent per cost in this context.

In general, we can find new directions by extending concepts associated with blocklength to those of cost. For instance, instead of one-shot capacities we could consider one-photon capacities, or more generally, “unit-cost” capacities. Instead of finite blocklength analyses of information processing tasks, we could perform finite cost analyses. The reverse substitution is also interesting. As we saw above, rather than cost-constrained capacities, we can consider blocklength-constrained capacities per unit cost. For probabilistic protocols or settings with feedback, we could imagine having an expected blocklength constraint instead of an expected cost constraint. In essence, there are many possible questions one could ask by looking at more general costs, as long as one has the time to consider them.

Acknowledgements. We are grateful to Emmanuel Abbe, Saikat Guha, Patrick Hayden, Alexander Holevo, Hideo Mabuchi, Sergio Verdu, and Tsachy Weissman for helpful conversations about this work. DD acknowledges support from the Stanford Graduate Fellowship program. DSP acknowledges support from the Army Research Office under grant number W911NF-16-1-0086. MMW acknowledges support from the US National Science Foundation and the Office of Naval Research.

References

[AJW17] Anurag Anshu, Rahul Jain, and Naqueeb Ahmad Warsi. One shot entanglement assisted classical and quantum communication over noisy quantum channels: A
hypothesis testing and convex split approach. February 2017. arXiv:1702.01940.

[BD10] Francesco Buscemi and Nilanjana Datta. The quantum capacity of channels with arbitrarily correlated noise. *IEEE Transactions on Information theory*, 56(3):1447–1460, March 2010. arXiv:0902.0158.

[BGPWW15] Bhaskar Roy Bardhan, Raul Garcia-Patron, Mark M. Wilde, and Andreas Winter. Strong converse for the classical capacity of all phase-insensitive bosonic Gaussian channels. *IEEE Transactions on Information Theory*, 61(4):1842–1850, April 2015. arXiv:1401.4161.

[BKN00] Howard Barnum, Emanuel Knill, and Michael Nielsen. On quantum fidelities and channel capacities. *IEEE Transactions on Information Theory*, 46(4):1317–1329, July 2000. arXiv:quant-ph/9809010.

[BNS98] Howard Barnum, M. A. Nielsen, and Benjamin Schumacher. Information transmission through a noisy quantum channel. *Physical Review A*, 57(6):4153–4175, June 1998. arXiv:quant-ph/9702049.

[CD94] Carlton M. Caves and Peter D. Drummond. Quantum limits on bosonic communication rates. *Reviews of Modern Physics*, 66(2):481, April 1994.

[CHP07] Imre Csiszár, Fumio Hiai, and Dénes Petz. Limit relation for quantum entropy and channel capacity per unit cost. *Journal of Mathematical Physics*, 48(9):092102, September 2007. arXiv:0704.0046.

[CT06] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley-Interscience, second edition, 2006.

[CWY04] N. Cai, Andreas Winter, and Raymond W. Yeung. Quantum privacy and quantum wiretap channels. *Problems of Information Transmission*, 40(4):318–336, October 2004.

[Dat09] Nilanjana Datta. Min- and max-relative entropies and a new entanglement monotone. *IEEE Transactions on Information Theory*, 55(6):2816–2826, June 2009. arXiv:0803.2770.

[DBEM11] S. Dolinar, K. M. Birnbaum, B. I. Erkmen, and B. Moision. On approaching the ultimate limits of photon-efficient and bandwidth-efficient optical communication. In *2011 International Conference on Space Optical Systems and Applications (ICSOS)*, pages 269–278, May 2011. arXiv:1104.2643.

[Dev05] Igor Devetak. The private classical capacity and quantum capacity of a quantum channel. *IEEE Transactions on Information Theory*, 51(1):44–55, January 2005. arXiv:quant-ph/0304127.

[DS05] Igor Devetak and Peter W. Shor. The capacity of a quantum channel for simultaneous transmission of classical and quantum information. *Communications in Mathematical Physics*, 256(2):287–303, June 2005. arXiv:quant-ph/0311131.
Nilanjana Datta, Marco Tomamichel, and Mark M Wilde. On the second-order asymptotics for entanglement-assisted communication. Quantum Information Processing, 15(6):2569–2591, June 2016. arXiv:1405.1797.

Igor Devetak and Andreas Winter. Relating quantum privacy and quantum coherence: An operational approach. Physical Review Letters, 93(8):080501, August 2004. arXiv:quant-ph/0307053.

Mustafa El-Halabi, Tie Liu, and Costas N. Georghiades. Secrecy capacity per unit cost. IEEE Journal on Selected Areas in Communications, 31(9):1909–1920, September 2013.

Robert G. Gallager. Energy limited channels: Coding, multiaccess, and spread spectrum. Technical report, 1987.

Saikat Guha, Zachary Dutton, and Jeffrey H. Shapiro. On quantum limit of optical communications: concatenated codes and joint-detection receivers. In Proceedings of the 2011 IEEE International Symposium on Information Theory, pages 274–278. IEEE, 2011. arXiv:1102.1963.

Vittorio Giovannetti, Saikat Guha, Seth Lloyd, Lorenzo Maccone, Jeffrey H Shapiro, and Horace P Yuen. Classical capacity of the lossy bosonic channel: The exact solution. Physical Review Letters, 92(2):027902, January 2004. arXiv:quant-ph/0308012.

Vittorio Giovannetti, Raul Garcia-Patron, Nicholas J. Cerf, and Alexander S. Holevo. Ultimate classical communication rates of quantum optical channels. Nature Photonics, 8(10):796–800, September 2014. arXiv:1312.6225.

Vittorio Giovannetti, Alexander S. Holevo, and Raul Garcia-Patron. A solution of gaussian optimizer conjecture for quantum channels. Communications in Mathematical Physics, 334(3):1553–1571, March 2015. arXiv:1312.2251.

Saikat Guha, Jonathan L. Habif, and M. Takeoka. Approaching Helstrom limits to optical pulse-position demodulation using single photon detection and optical feedback. Journal of Modern Optics, 58(3-4):257–265, December 2011. arXiv:1001.2447.

M. J. E. Golay. Note on the theoretical efficiency of information reception with PPM. Proceedings of the Institute of Radio Engineers, 37(9):1031–1031, 1949.

Saikat Guha. Structured optical receivers to attain superadditive capacity and the Holevo limit. Physical Review Letters, 106(24):240502, June 2011. arXiv:1101.1550.

Alexander S. Holevo and Vittorio Giovannetti. Quantum channels and their entropic characteristics. Reports on progress in physics, 75(4):046001, April 2012. arXiv:1202.6480.

Alexander S. Holevo. The capacity of the quantum channel with general signal states. IEEE Transactions on Information Theory, 44(1):269–273, January 1998. arXiv:quant-ph/9611023.
Alexander S. Holevo. Sending quantum information with Gaussian states. In *Quantum Communication, Computing, and Measurement*, pages 75–82. Springer, 2002. arXiv preprint quant-ph/9809022.

Alexander S. Holevo. Entanglement-assisted capacity of constrained channels. In *First International Symposium on Quantum Informatics*, pages 62–69. International Society for Optics and Photonics, 2003. arXiv:quant-ph/0211170.

Alexander S. Holevo. Entanglement-assisted capacities of constrained quantum channels. *Theory of Probability & Its Applications*, 48(2):243–255, July 2004. arXiv:quant-ph/0211170.

Alexander S. Holevo. Entanglement-breaking channels in infinite dimensions. *Problems of Information Transmission*, 44(3):171–184, 2008. arXiv:0802.0235.

Alexander S. Holevo. *Quantum systems, channels, information: a mathematical introduction*, volume 16. Walter de Gruyter, 2012.

Fumio Hiai and Dénes Petz. The proper formula for relative entropy and its asymptotics in quantum probability. *Communications in Mathematical Physics*, 143(1):99–114, December 1991.

Alexander S. Holevo and Reinhard F. Werner. Evaluating capacities of bosonic Gaussian channels. *Physical Review A*, 63(3):032312, February 2001. arXiv:quant-ph/9912067.

Marcin Jarzyna. Classical capacity per unit cost for quantum channels. April 2017. arXiv:1704.06696.

Robert König, Renato Renner, Andor Bariska, and Ueli Maurer. Small accessible quantum information does not imply security. *Physical Review Letters*, 98:140502, April 2007. arXiv:quant-ph/0512021.

Göran Lindblad. Completely positive maps and entropy inequalities. *Communications in Mathematical Physics*, 40(2):147–151, June 1975.

Dmitry S. Lebedev and Lev B. Levitin. Information transmission by electromagnetic field. *Information and Control*, 9(1):1–22, 1966.

Seth Lloyd. Capacity of the noisy quantum channel. *Physical Review A*, 55(3):1613–1622, March 1997. arXiv:quant-ph/9604015.

Tomohiro Ogawa and Hiroshi Nagaoka. Strong converse and Stein’s lemma in quantum hypothesis testing. *IEEE Transactions on Information Theory*, 46(7):2428–2433, November 2000. arXiv:quant-ph/9906090.

Vern Paulsen. *Completely bounded maps and operator algebras*, volume 78. Cambridge University Press, 2002.

J. Pierce. Optical channels: Practical limits with photon counting. *IEEE Transactions on Communications*, 26(12):1819–1821, December 1978.
[PPR81] J. Pierce, E. Posner, and E. Rodemich. The capacity of the photon counting channel. *IEEE Transactions on Information Theory*, 27(1):61–77, January 1981.

[Rez61] Fazlollah M. Reza. *An introduction to information theory*. Courier Corporation, 1961.

[SH16] M. E. Shirokov and Alexander S. Holevo. Lower semicontinuity of the entropic disturbance and its applications in quantum information theory. August 2016. arXiv:1608.02203.

[Sha48] Claude E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, and 623–656, 1948.

[Shi06] Maksim E. Shirokov. The Holevo capacity of infinite dimensional channels and the additivity problem. *Communications in Mathematical Physics*, 262(1):137–159, February 2006. arXiv:quant-ph/0408009.

[Shi12] Maxim E. Shirokov. Conditions for coincidence of the classical capacity and entanglement-assisted capacity of a quantum channel. *Problems of Information Transmission*, 48(2):85–101, 2012.

[Sho02a] Peter W. Shor. Additivity of the classical capacity of entanglement-breaking quantum channels. *Journal of Mathematical Physics*, 43(9):4334–4340, 2002. arXiv:quant-ph/0201149.

[Sho02b] Peter W. Shor. The quantum channel capacity and coherent information. In *Lecture Notes, MSRI Workshop on Quantum Computation*, 2002.

[Smi08] Graeme Smith. Private classical capacity with a symmetric side channel and its application to quantum cryptography. *Physical Review A*, 78(2):022306, August 2008. arXiv:0705.3838.

[SN96] Benjamin Schumacher and Michael A. Nielsen. Quantum data processing and error correction. *Physical Review A*, 54(4):2629–2635, October 1996. arXiv:quant-ph/9604022.

[Sti55] William Forrest Stinespring. Positive functions on C*-algebras. *Proceedings of the American Mathematical Society*, 6(2):211–216, 1955.

[SW97] Benjamin Schumacher and Michael D. Westmoreland. Sending classical information via noisy quantum channels. *Physical Review A*, 56(1):131, July 1997.

[TG14] Masahiro Takeoka and Saikat Guha. Capacity of optical communication in loss and noise with general quantum Gaussian receivers. *Physical Review A*, 89(4):042309, April 2014. arXiv:1401.5132.

[Tom16] Marco Tomamichel. *Quantum Information Processing with Finite Resources: Mathematical Foundations*, volume 5 of *Springer Briefs in Mathematical Physics*. Springer, 2016. arXiv:1504.00233.

[Uhl76] Armin Uhlmann. The “transition probability” in the state space of a *-algebra. *Reports on Mathematical Physics*, 9(2):273–279, 1976.
References

[Ume62] Hisaharu Umegaki. Conditional expectations in an operator algebra IV (entropy and information). *Kodai Mathematical Seminar Reports*, 14(2):59–85, 1962.

[Ver90] Sergio Verdu. On channel capacity per unit cost. *IEEE Transactions on Information Theory*, 36(5):1019–1030, September 1990.

[Wil17] Mark M. Wilde. *Quantum information theory*. Cambridge University Press, second edition, 2017. arXiv:1106.1445.

[Win99a] Andreas Winter. Coding theorem and strong converse for quantum channels. *IEEE Transactions on Information Theory*, 45(7):2481–2485, November 1999. arXiv:1409.2536.

[Win99b] Andreas Winter. *Coding Theorems of Quantum Information Theory*. PhD thesis, Universität Bielefeld, July 1999. arXiv:quant-ph/9907077.

[Win16] Andreas Winter. Tight uniform continuity bounds for quantum entropies: Conditional entropy, relative entropy distance and energy constraints. *Communications in Mathematical Physics*, 1(347):291–313, October 2016. arXiv:1507.07775.

[WPGG07] Michael M. Wolf, David Pérez-García, and Geza Giedke. Quantum capacities of bosonic channels. *Physical Review Letters*, 98(13):130501, March 2007. arXiv:quant-ph/0606132.

[WQ16] Mark M. Wilde and Haoyu Qi. Energy-constrained private and quantum capacities of quantum channels. September 2016. arXiv:1609.01997.

[WR12] Ligong Wang and Renato Renner. One-shot classical-quantum capacity and hypothesis testing. *Physical Review Letters*, 108(20):200501, May 2012. arXiv:1007.5456.

[YKL75] H. Yuen, R. Kennedy, and Melvin Lax. Optimum testing of multiple hypotheses in quantum detection theory. *IEEE Transactions on Information Theory*, 21(2):125–134, March 1975.

[YO93] Horace P. Yuen and Masanao Ozawa. Ultimate information carrying limit of quantum systems. *Physical Review Letters*, 70(4):363, January 1993.