10-1-2017

Position-based coding and convex splitting for private communication over quantum channels

Mark M. Wilde
Louisiana State University

Recommended Citation

Wilde, M. (2017). Position-based coding and convex splitting for private communication over quantum channels. *Quantum Information Processing, 16* (10) https://doi.org/10.1007/s11128-017-1718-4

Position-based coding and convex splitting for private communication over quantum channels

Mark M. Wilde*

March 5, 2017

Abstract

The classical-input quantum-output (cq) wiretap channel is a communication model involving a classical sender X, a legitimate quantum receiver B, and a quantum eavesdropper E. The goal of a private communication protocol that uses such a channel is for the sender X to transmit a message in such a way that the legitimate receiver B can decode it reliably, while the eavesdropper E learns essentially nothing about which message was transmitted. The ε-one-shot private capacity of a cq wiretap channel is equal to the maximum number of bits that can be transmitted over the channel, such that the privacy error is no larger than $\varepsilon \in (0, 1)$. The present paper provides a lower bound on the ε-one-shot private classical capacity, by exploiting the recently developed techniques of Anshu, Devabathini, Jain, and Warsi, called position-based coding and convex splitting. The lower bound is equal to a difference of the hypothesis testing mutual information between X and B and the "alternate" smooth max-information between X and E. The one-shot lower bound then leads to a non-trivial lower bound on the second-order coding rate for private classical communication over a memoryless cq wiretap channel.

1 Introduction

Among the many results of information theory, the ability to use the noise in a wiretap channel for the purpose of private communication stands out as one of the great conceptual insights [Wyn75]. A classical wiretap channel is modeled as a conditional probability distribution $p_{Y,Z|X}$, in which the sender Alice has access to the input X of the channel, the legitimate receiver Bob has access to the output Y, and the eavesdropper Eve has access to the output Z. The goal of private communication is for Alice and Bob to use the wiretap channel in such a way that Alice communicates a message reliably to Bob, while at the same time, Eve should not be able to determine which message was transmitted. The author of [Wyn75] proved that the mutual information difference

$$\max_{p_X} [I(X;Y) - I(X;Z)] \tag{1.1}$$

is an achievable rate for private communication over the wiretap channel, when Alice and Bob are allowed to use it many independent times. Since then, the interest in the wiretap channel has not waned, and there have been many increasingly refined statements about achievable rates for private communication over wiretap channels [CK78, Hay06, Tan12, Hay13, YAG13, YSP16, TB16].

*Hearne Institute for Theoretical Physics, Department of Physics and Astronomy, Center for Computation and Technology, Louisiana State University, Baton Rouge, Louisiana 70803, USA

Many years after the contribution of [Wyn75], the protocol of quantum key distribution was developed as a proposal for private communication over a quantum channel [BB84]. Quantum information theory started becoming a field in its own right, during which many researchers revisited several of the known results of Shannon's information theory under a quantum lens. This was not merely an academic exercise: doing so revealed that remarkable improvements in communication rates could be attained for physical channels of practical interest if quantum-mechanical strategies are exploited [GGL+04].

One important setting which was revisited is the wiretap channel, and in the quantum case, the simplest extension of the classical model is given by the classical-input quantum-output wiretap channel (abbreviated as *cq wiretap channel*) [Dev05, CWY04]. It is described as the following map:

$$x \to \rho_{BE}^{x}, \tag{1.2}$$

where x is a classical symbol that Alice can input to the channel and ρ_{BE}^{x} is the joint output quantum state of Bob and Eve's system, represented as a density operator acting on the tensor-product Hilbert space of Bob and Eve's quantum systems. The goal of private communication over the cq wiretap channel is similar to that for the classical wiretap channel. However, in this case, Bob is allowed to perform a collective quantum measurement over all of his output quantum systems in order to determine Alice's message, while at the same time, we would like for it be difficult for Eve to figure out anything about the transmitted message, even if she has access to a quantum computer memory that can store all of the quantum systems that she receives from the channel output. The authors of [Dev05, CWY04] independently proved that a quantum generalization of the formula in (1.1) is an achievable rate for private communication over a cq quantum wiretap channel, if Alice and Bob are allowed to use it many independent times. Namely, they proved that the following Holevo information difference is an achievable rate:

$$\max_{p_X} \left[I(X;B) - I(X;E) \right], \tag{1.3}$$

where the information quantities in the above formula are the Holevo information to Bob and Eve, respectively, and will be formally defined later in the present paper.

Since the developments of [Dev05, CWY04], there has been an increasing interest in the quantum information community to determine refined characterizations of communication tasks [TH13, Li14, TT15, DTW16, DHO16, DL15, BDL16, TBR16, WTB17], strongly motivated by the fact that it is experimentally difficult to control a large number of quantum systems, and in practice, one has access only to a finite number of quantum systems anyway. One such scenario of interest, as discussed above, is the quantum wiretap channel. Hitherto, the only work offering achievable one-shot rates for private communication over cq wiretap channels is [RR11]. However, that work did not consider bounding the second-order coding rate for private communication over the cq wiretap channel.

The main contribution of the present paper is a lower bound on the one-shot private capacity of a cq wiretap channel. Namely, I prove that

$$\log_2 M^*_{\text{priv}}(\varepsilon_1 + \sqrt{\varepsilon_2}) \geq I_H^{\varepsilon_1 - \eta_1}(X;B) - \widetilde{I}_{\max}^{\sqrt{\varepsilon_2} - \eta_2}(E;X) - \log_2(4\varepsilon_1/\eta_1^2) - 2\log_2(1/\eta_2). \tag{1.4}$$

In the above, $\log_2 M^*_{\text{priv}}(\varepsilon_1 + \sqrt{\varepsilon_2})$ represents the maximum number of bits that can be sent from Alice to Bob, using a cq wiretap channel once, such that the privacy error (to be defined formally later) does not exceed $\varepsilon_1 + \sqrt{\varepsilon_2} \in (0,1)$, with $\varepsilon_1, \varepsilon_2 \in (0,1)$. The quantities on the right-hand side of

the above inequality are particular one-shot generalizations of the Holevo information to Bob and Eve, which will be defined later. It is worthwhile to note that the one-shot information quantities in (1.4) can be computed using semi-definite programming, and the computational runtime is polynomial in the dimension of the channel. Thus, for channels of reasonable dimension, the quantities can be efficiently estimated numerically. The constants η_1 and η_2 are chosen so that $\eta_1 \in (0, \varepsilon_1)$ and $\eta_2 \in (0, \sqrt{\varepsilon_2})$. By substituting an independent and identically distributed (i.i.d.) cq wiretap channel into the right-hand side of the above inequality, using second-order expansions for the one-shot Holevo informations [TH13, Li14], and picking $\eta_1, \eta_2 = 1/\sqrt{n}$, we find the following lower bound on the second-order coding rate for private classical communication:

$$\log_2 M^*_{\text{priv}}(n, \varepsilon_1 + \sqrt{\varepsilon_2}) \geq n\,[I(X;B) - I(X;E)] \\ + \sqrt{nV(X;B)}\Phi^{-1}(\varepsilon_1) + \sqrt{nV(X;E)}\Phi^{-1}(\varepsilon_2) + O(\log n). \quad (1.5)$$

In the above, $\log_2 M^*_{\text{priv}}(n, \varepsilon_1 + \sqrt{\varepsilon_2})$ represents the maximum number of bits that can be sent from Alice to Bob, using a cq wiretap channel n times, such that the privacy error does not exceed $\varepsilon_1 + \sqrt{\varepsilon_2} \in (0, 1)$. The Holevo informations from (1.3) make an appearance in the first-order term (proportional to the number n of channel uses) on the right-hand side above, while the second order term (proportional to $\sqrt{n}$) consists of the quantum channel dispersion quantities $V(X;B)$ and $V(X;E)$ [TT15], which will be defined later. They additionally feature the inverse Φ^{-1} of the cumulative Gaussian distribution function Φ. Thus, the one-shot bound in (1.4) leads to a lower bound on the second-order coding rate, which is comparable to bounds that have appeared in the classical information theory literature [Tan12, YAG13, YSP16, TB16].

To prove the one-shot bound in (1.4), I use two recent and remarkable techniques: position-based coding [AJW17] and convex splitting [ADJ17]. The main idea of position-based coding [AJW17] is conceptually simple. To communicate a classical message from Alice to Bob, we allow them to share a quantum state $\rho_{RA}^{\otimes M}$ before communication begins, where M is the number of messages, Bob possesses the R systems, and Alice the A systems. If Alice wishes to communicate message m, then she sends the mth A system through the channel. The reduced state of Bob's systems is then

$$\rho_{R_1} \otimes \cdots \otimes \rho_{R_{m-1}} \otimes \rho_{R_m B} \otimes \rho_{R_{m+1}} \otimes \cdots \otimes \rho_{R_M}, \quad (1.6)$$

where $\rho_{R_m B} = \mathcal{N}_{A_m \to B}(\rho_{R_m A_m})$ and $\mathcal{N}_{A_m \to B}$ is the quantum channel. For all $m' \neq m$, the reduced state for systems $R_{m'}$ and B is the product state $\rho_{R_{m'}} \otimes \rho_B$. However, the reduced state of systems $R_m B$ is the (generally) correlated state $\rho_{R_m B}$. So if Bob has a binary measurement which can distinguish the joint state ρ_{RB} from the product state $\rho_R \otimes \rho_B$ sufficiently well, he can base a decoding strategy off of this, and the scheme will be reliable as long as the number of bits $\log_2 M$ to be communicated is chosen to be roughly equal to a one-shot mutual information known as hypothesis testing mutual information (cf., [WR12]). This is exactly what is used in position-based coding, and the authors of [AJW17] thus forged a transparent and intuitive link between quantum hypothesis testing and communication for the case of entanglement-assisted communication.

Convex splitting [ADJ17] is rather intuitive as well and can be thought of as dual to the coding scenario mentioned above. Suppose instead that Alice and Bob have a means of generating the state in (1.6), perhaps by the strategy mentioned above. But now suppose that Alice chooses the variable m uniformly at random, so that the state, from the perspective of someone ignorant of the choice of m, is the following mixture:

$$\frac{1}{M} \sum_{m=1}^{M} \rho_{R_1} \otimes \cdots \otimes \rho_{R_{m-1}} \otimes \rho_{R_m B} \otimes \rho_{R_{m+1}} \otimes \cdots \otimes \rho_{R_M}. \quad (1.7)$$

The convex-split lemma guarantees that as long as $\log_2 M$ is roughly equal to a one-shot mutual information known as the alternate smooth max-mutual information, then the state above is nearly indistinguishable from the product state $\rho_R^{\otimes M} \otimes \rho_B$.

Both position-based coding and convex splitting have been used recently and effectively to establish a variety of results in one-shot quantum information theory [AJW17, ADJ17]. In the present paper, I use the approaches in conjunction to construct codes for the cq wiretap channel. The main underlying idea follows the original approach of [Wyn75], by allowing for a message variable $m \in \{1, \ldots, M\}$ and a local key variable $k \in \{1, \ldots, K\}$ (local randomness), the latter of which is selected uniformly at random and used to confuse the eavesdropper Eve. Before communication begins, Alice, Bob, and Eve are allowed share to MK copies of the common randomness state $\theta_{X_A X_B X_E} \equiv \sum_x p_X(x)|xxx\rangle\langle xxx|_{X_A X_B X_E}$. We can think of the MK copies of $\theta_{X_A X_B X_E}$ as being partitioned into M blocks, each of which contain K copies of the state $\theta_{X_A X_B X_E}$. If Alice wishes to send message m, then she picks k uniformly at random and sends the (m, k) X_A system through the cq wiretap channel in (1.2). As long as $\log_2 MK$ is roughly equal to the hypothesis testing mutual information $I_H^\varepsilon(X; B)$, then Bob can use a position-based decoder to figure out both m and k. As long as $\log_2 K$ is roughly equal to the alternate smooth max-mutual information $\widetilde{I}_{\max}^\varepsilon(E; X)$, then the convex-split lemma guarantees that the overall state of Eve's systems, regardless of which message m was chosen, is nearly indistinguishable from the product state $\rho_{X_E}^{\otimes MK} \otimes \rho_E$. Thus, in such a scheme, Bob can figure out m while Eve cannot figure out anything about m. This is the intuition behind the coding scheme and gives a sense of why $\log_2 M = \log_2 MK - \log_2 K \approx I_H^\varepsilon(X; B) - \widetilde{I}_{\max}^\varepsilon(E; X)$ is an achievable number of bits that can be sent privately from Alice to Bob. The main purpose of the present paper is to develop the details of this argument and furthermore show how the scheme can be derandomized, so that the MK copies of the common randomness state $\theta_{X_A X_B X_E}$ are in fact not necessary.

The rest of the paper proceeds as follows. In Section 2, I review some preliminary material, which includes several metrics for quantum states and pertinent information measures. Section 3 develops the position-based coding approach for classical-input quantum-output communication channels. Position-based coding was developed in [AJW17] to highlight a different approach to entanglement-assisted communication, but I show in Section 3 how the approach can be used for shared randomness-assisted communication; I also show therein how to derandomize codes in this case (i.e., the shared randomness is not actually necessary for classical communication over cq channels). Section 4 represents the main contribution of the present paper, which is a lower bound on the ε-one-shot private classical capacity of a cq wiretap channel. The last development in Section 4 is to show how the one-shot lower bound leads to a lower bound on the second-order coding rate for private classical communication over a memoryless cq wiretap channel. Therein, I also show how these lower bounds simplify for pure-state cq wiretap channels and when using binary phase-shift keying as a coding strategy for private communication over a pure-loss bosonic channel. Section 5 concludes with a summary and some open questions for future work.

2 Preliminaries

I use notation and concepts that are standard in quantum information theory and point the reader to [Wil16] for background. In the rest of this section, I review concepts that are less standard and set some notation that will be used later in the paper.

Trace distance, fidelity, and purified distance. Let $\mathcal{D}(\mathcal{H})$ denote the set of density operators acting on a Hilbert space $\mathcal{H}$, let $\mathcal{D}_{\leq}(\mathcal{H})$ denote the set of subnormalized density operators (with trace not exceeding one) acting on $\mathcal{H}$, and let $\mathcal{L}_{+}(\mathcal{H})$ denote the set of positive semi-definite operators acting on $\mathcal{H}$. The trace distance between two quantum states $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ is equal to $\|\rho - \sigma\|_1$, where $\|C\|_1 \equiv \mathrm{Tr}\{\sqrt{C^\dagger C}\}$ for any operator C. It has a direct operational interpretation in terms of the distinguishability of these states. That is, if ρ or σ are prepared with equal probability and the task is to distinguish them via some quantum measurement, then the optimal success probability in doing so is equal to $(1 + \|\rho - \sigma\|_1 / 2) / 2$. The fidelity is defined as $F(\rho, \sigma) \equiv \left\|\sqrt{\rho}\sqrt{\sigma}\right\|_1^2$ [Uhl76], and more generally we can use the same formula to define $F(P, Q)$ if $P, Q \in \mathcal{L}_{+}(\mathcal{H})$. Uhlmann's theorem states that [Uhl76]

$$F(\rho_A, \sigma_A) = \max_U |\langle \phi^\sigma |_{RA} U_R \otimes I_A | \phi^\rho \rangle_{RA}|^2 , \tag{2.1}$$

where $|\phi^\rho\rangle_{RA}$ and $|\phi^\sigma\rangle_{RA}$ are fixed purifications of ρ_A and σ_A, respectively, and the optimization is with respect to all unitaries U_R. The same statement holds more generally for $P, Q \in \mathcal{L}_{+}(\mathcal{H})$. The fidelity is invariant with respect to isometries and monotone non-decreasing with respect to channels. The sine distance or C-distance between two quantum states $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ was defined as

$$C(\rho, \sigma) \equiv \sqrt{1 - F(\rho, \sigma)} \tag{2.2}$$

and proven to be a metric in [Ras02, Ras03, GLN05, Ras06]. It was later [TCR09] (under the name "purified distance") shown to be a metric on subnormalized states $\rho, \sigma \in \mathcal{D}_{\leq}(\mathcal{H})$ via the embedding

$$P(\rho, \sigma) \equiv C(\rho \oplus [1 - \mathrm{Tr}\{\rho\}] , \sigma \oplus [1 - \mathrm{Tr}\{\sigma\}]) . \tag{2.3}$$

The following inequality relates trace distance and purified distance:

$$\frac{1}{2} \|\rho - \sigma\|_1 \leq P(\rho, \sigma). \tag{2.4}$$

Relative entropies and variances. The quantum relative entropy of two states ω and τ is defined as [Ume62]

$$D(\omega \| \tau) \equiv \mathrm{Tr}\{\omega[\log_2 \omega - \log_2 \tau]\} \tag{2.5}$$

whenever $\mathrm{supp}(\omega) \subseteq \mathrm{supp}(\tau)$ and it is equal to $+\infty$ otherwise. The quantum relative entropy variance is defined as [TH13, Li14]

$$V(\omega \| \tau) \equiv \mathrm{Tr}\{\omega[\log_2 \omega - \log_2 \tau - D(\omega \| \tau)]^2\}, \tag{2.6}$$

whenever $\mathrm{supp}(\omega) \subseteq \mathrm{supp}(\tau)$. The hypothesis testing relative entropy [BD10, WR12] of states ω and τ is defined as

$$D_H^\varepsilon(\omega \| \tau) \equiv -\log_2 \inf_\Lambda \{\mathrm{Tr}\{\Lambda\tau\} : 0 \leq \Lambda \leq I \wedge \mathrm{Tr}\{\Lambda\omega\} \geq 1 - \varepsilon\} . \tag{2.7}$$

The max-relative entropy for states ω and τ is defined as [Dat09]

$$D_{\max}(\omega \| \tau) \equiv \inf \left\{ \lambda \in \mathbb{R} : \omega \leq 2^\lambda \tau \right\} . \tag{2.8}$$

The smooth max-relative entropy for states ω and τ and a parameter $\varepsilon \in (0,1)$ is defined as [Dat09]

$$D^{\varepsilon}_{\max}(\omega\|\tau) \equiv \inf\left\{\lambda \in \mathbb{R} : \widetilde{\omega} \leq 2^{\lambda}\tau \wedge P(\omega,\widetilde{\omega}) \leq \varepsilon\right\}. \tag{2.9}$$

The following second-order expansions hold for D^{ε}_{H} and $D^{\varepsilon}_{\max}$ when evaluated for tensor-power states [TH13, Li14]:

$$D^{\varepsilon}_{H}(\omega^{\otimes n}\|\tau^{\otimes n}) = nD(\omega\|\tau) + \sqrt{nV(\omega\|\tau)}\Phi^{-1}(\varepsilon) + O(\log n), \tag{2.10}$$

$$D^{\sqrt{\varepsilon}}_{\max}(\omega^{\otimes n}\|\tau^{\otimes n}) = nD(\omega\|\tau) - \sqrt{nV(\omega\|\tau)}\Phi^{-1}(\varepsilon) + O(\log n). \tag{2.11}$$

The above expansion features the cumulative distribution function for a standard normal random variable:

$$\Phi(a) \equiv \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a} dx\, \exp\left(-x^2/2\right), \tag{2.12}$$

and its inverse, defined as $\Phi^{-1}(\varepsilon) \equiv \sup\{a \in \mathbb{R} \,|\, \Phi(a) \leq \varepsilon\}$.

Mutual informations and variances. The quantum mutual information $I(X;B)_{\rho}$ and information variance $V(X;B)_{\rho}$ of a bipartite state ρ_{XB} are defined as

$$I(X;B)_{\rho} \equiv D(\rho_{XB}\|\rho_X \otimes \rho_B), \tag{2.13}$$

$$V(X;B)_{\rho} \equiv V(\rho_{XB}\|\rho_X \otimes \rho_B). \tag{2.14}$$

In this paper, we are exclusively interested in the case in which system X of ρ_{XB} is classical, so that ρ_{XB} can be written as

$$\rho_{XB} = \sum_x p_X(x)|x\rangle\langle x|_X \otimes \rho^x_B, \tag{2.15}$$

where p_X is a probability distribution, $\{|x\rangle_X\}_x$ is an orthonormal basis, and $\{\rho^x_B\}_x$ is a set of quantum states. The hypothesis testing mutual information is defined as follows for a bipartite state ρ_{XB} and a parameter $\varepsilon \in (0,1)$:

$$I^{\varepsilon}_{H}(X;B)_{\rho} \equiv D^{\varepsilon}_{H}(\rho_{XB}\|\rho_X \otimes \rho_B). \tag{2.16}$$

From the smooth max-relative entropy, one can define a mutual information-like quantity for a state θ_{AB} as follows:

$$D^{\varepsilon}_{\max}(\theta_{AB}\|\theta_A \otimes \theta_B). \tag{2.17}$$

Note that we have the following expansions, as a direct consequence of (2.10)–(2.11) and definitions:

$$I^{\varepsilon}_{H}(X^n;B^n)_{\rho^{\otimes n}} = nI(X;B)_{\rho} + \sqrt{nV(X;B)_{\rho}}\Phi^{-1}(\varepsilon) + O(\log n), \tag{2.18}$$

$$D^{\sqrt{\varepsilon}}_{\max}(\rho^{\otimes n}_{XB}\|\rho^{\otimes n}_X \otimes \rho^{\otimes n}_B) = nI(X;B)_{\rho} - \sqrt{nV(X;B)_{\rho}}\Phi^{-1}(\varepsilon) + O(\log n). \tag{2.19}$$

Another quantity, related to that in (2.17), is as follows [AJW17]:

$$\widetilde{I}^{\varepsilon}_{\max}(B;A)_{\theta} \equiv \inf_{\theta'_{AB}\,:\,P(\theta'_{AB},\theta_{AB})\leq\varepsilon} D_{\max}(\theta'_{AB}\|\theta_A \otimes \theta'_B). \tag{2.20}$$

We recall a relation [AJW17, Lemma 1] between the quantities in (2.17) and (2.20), giving a very slight modification of it which will be useful for our purposes here:

Lemma 1 *For a state θ_{AB}, $\varepsilon \in (0,1)$, and $\gamma \in (0,\varepsilon)$, the following inequality holds*

$$\widetilde{I}^{\varepsilon}_{\max}(B;A)_{\theta} \leq D^{\varepsilon-\gamma}_{\max}(\theta_{AB}\|\theta_A \otimes \theta_B) + \log_2\left(\frac{3}{\gamma^2}\right). \tag{2.21}$$

Proof. To see this, recall [AJW17, Claim 2]: For states ω_{AB}, ξ_A, κ_B, there exists a state $\overline{\omega}_{AB}$ such that $P(\omega_{AB},\overline{\omega}_{AB}) \leq \delta$ and

$$D_{\max}(\overline{\omega}_{AB}\|\xi_A \otimes \overline{\omega}_B) \leq D_{\max}(\omega_{AB}\|\xi_A \otimes \kappa_B) + \log_2\left(\frac{3}{\delta^2}\right). \tag{2.22}$$

Let θ^*_{AB} denote the optimizer for $D^{\varepsilon-\gamma}_{\max}(\theta_{AB}\|\theta_A\otimes\theta_B)$. Then, in (2.22), taking $\omega_{AB} = \theta^*_{AB}$, $\xi_A = \theta_A$, $\kappa_B = \theta_B$, we find that there exists a state $\overline{\theta}_{AB}$ such that $P(\theta^*_{AB},\overline{\theta}_{AB}) \leq \gamma$ and

$$D_{\max}(\overline{\theta}_{AB}\|\theta_A \otimes \overline{\theta}_B) \leq D^{\varepsilon-\gamma}_{\max}(\theta_{AB}\|\theta_A \otimes \theta_B) + \log_2\left(\frac{3}{\gamma^2}\right). \tag{2.23}$$

By the triangle inequality for the purified distance, we conclude that $P(\theta_{AB},\overline{\theta}_{AB}) \leq P(\theta_{AB},\theta^*_{AB}) + P(\theta^*_{AB},\overline{\theta}_{AB}) \leq (\varepsilon-\gamma)+\gamma = \varepsilon$. Since the quantity on the left-hand side includes an optimization over all states θ'_{AB} satisfying $P(\theta'_{AB},\theta_{AB}) \leq \varepsilon$, we conclude the inequality in (2.21). ∎

Hayashi–Nagaoka operator inequality. A key tool in analyzing error probabilities in communication protocols is the Hayashi–Nagaoka operator inequality [HN03]: given operators S and T such that $0 \leq S \leq I$ and $T \geq 0$, the following inequality holds for all $c > 0$

$$I - (S+T)^{-1/2}S(S+T)^{-1/2} \leq (1+c)(I-S) + (2+c+c^{-1})T. \tag{2.24}$$

Convex-split lemma. The convex-split lemma from [ADJ17] has been a key tool used in recent developments in quantum information theory [AJW17, ADJ17]. We now state a variant of the convex-split lemma, which is helpful for obtaining one-shot bounds for privacy and an ensuing lower bound on the second-order coding rate. Its proof closely follows proofs available in [AJW17, ADJ17] but has some slight differences. For completeness, Appendix A contains a proof of Lemma 2.

Lemma 2 (Convex split) *Let ρ_{AB} be a state, and let $\tau_{A_1\cdots A_K B}$ be the following state:*

$$\tau_{A_1\cdots A_K B} \equiv \frac{1}{K}\sum_{k=1}^{K} \rho_{A_1} \otimes \cdots \otimes \rho_{A_{k-1}} \otimes \rho_{A_k B} \otimes \rho_{A_{k+1}} \otimes \cdots \otimes \rho_{A_K}. \tag{2.25}$$

Let $\varepsilon \in (0,1)$ and $\eta \in (0,\sqrt{\varepsilon})$. If

$$\log_2 K = \widetilde{I}^{\sqrt{\varepsilon}-\eta}_{\max}(B;A)_{\rho} + 2\log_2\left(\frac{1}{\eta}\right), \tag{2.26}$$

then

$$P(\tau_{A_1\cdots A_K B}, \rho_{A_1} \otimes \cdots \otimes \rho_{A_K} \otimes \widetilde{\rho}_B) \leq \sqrt{\varepsilon}, \tag{2.27}$$

for some state $\widetilde{\rho}_B$ such that $P(\rho_B,\widetilde{\rho}_B) \leq \sqrt{\varepsilon}-\eta$.

3 Public classical communication

3.1 Definition of the one-shot classical capacity

We begin by defining the ε-one-shot classical capacity of a cq channel

$$x \to \rho_B^x. \tag{3.1}$$

We can write the classical–quantum channel in fully quantum form as the following quantum channel:

$$\mathcal{N}_{X'\to B}(\sigma_{X'}) = \sum_x \langle x|_{X'}\sigma_{X'}|x\rangle_{X'}\rho_B^x, \tag{3.2}$$

where $\{|x\rangle_{X'}\}_x$ is some orthonormal basis. Let $M \in \mathbb{N}$ and $\varepsilon \in (0,1)$. An (M,ε) classical communication code consists of a collection of probability distributions $\{p_{X|M}(x|m)\}_{m=1}^M$ (one for each message m) and a decoding positive operator-valued measure (POVM) $\{\Lambda_B^m\}_{m=1}^M$,[1] such that

$$\frac{1}{M}\sum_{m=1}^M \operatorname{Tr}\{(I_B - \Lambda_B^m)\rho_B^m\} = \frac{1}{M}\sum_{m=1}^M \frac{1}{2}\left\|\mathcal{M}_{B\to\hat{M}}(\rho_B^m) - |m\rangle\langle m|_{\hat{M}}\right\|_1 \leq \varepsilon. \tag{3.3}$$

We refer to the left-hand side of the above inequality as the *decoding error*. In the above, $\{|m\rangle_{\hat{M}}\}_{m=1}^M$ is an orthonormal basis, we define the state ρ_B^m as

$$\rho_B^m = \sum_x p_{X|M}(x|m)\rho_B^x, \tag{3.4}$$

and the measurement channel $\mathcal{M}_{B\to\hat{M}}$ as

$$\mathcal{M}_{B\to\hat{M}}(\omega_B) \equiv \sum_m \operatorname{Tr}\{\Lambda_B^m\omega_B\}|m\rangle\langle m|_{\hat{M}}. \tag{3.5}$$

The equality in (3.3) follows by direct calculation:

$$\begin{aligned}
&\left\|\mathcal{M}_{B\to\hat{M}}(\rho_B^m) - |m\rangle\langle m|_{\hat{M}}\right\|_1 \\
&= \left\|\sum_{m'} \operatorname{Tr}\{\Lambda_B^{m'}\rho_B^m\}|m'\rangle\langle m'|_{\hat{M}} - |m\rangle\langle m|_{\hat{M}}\right\|_1 &(3.6)\\
&= \left\|\sum_{m'\neq m} \operatorname{Tr}\{\Lambda_B^{m'}\rho_B^m\}|m'\rangle\langle m'|_{\hat{M}} - (1-\operatorname{Tr}\{\Lambda_B^m\rho_B^m\})|m\rangle\langle m|_{\hat{M}}\right\|_1 &(3.7)\\
&= \sum_{m'\neq m} \operatorname{Tr}\{\Lambda_B^{m'}\rho_B^m\} + (1-\operatorname{Tr}\{\Lambda_B^m\rho_B^m\}) &(3.8)\\
&= 2\operatorname{Tr}\{(I_B - \Lambda_B^m)\rho_B^m\}. &(3.9)
\end{aligned}$$

For a given channel $\mathcal{N}_{X'\to B}$ and ε, the one-shot classical capacity is equal to $\log_2 M_{\text{pub}}^*(\varepsilon)$, where $M_{\text{pub}}^*(\varepsilon)$ is the largest M such that (3.3) can be satisfied for a fixed ε.

One can allow for shared randomness between Alice and Bob before communication begins, in which case one obtains the one-shot shared randomness assisted capacity of a cq channel.

[1] We could allow for a decoding POVM to be $\{\Lambda_B^m\}_{m=0}^M$, consisting of an extra operator $\Lambda_B^0 = I_B - \sum_{m=1}^M \Lambda_B^m$, if needed.

3.2 Lower bound on the one-shot, randomness-assisted classical capacity

We first consider a one-shot protocol for randomness assisted, public classical communication in which the goal is for Alice to use the classical-input quantum-output (cq) channel in (3.1) once to send one of M messages with error probability no larger than $\varepsilon \in (0,1)$. The next section shows how to derandomize such that the shared randomness is not needed.

The main result of this section is that

$$I_H^{\varepsilon-\eta}(X;B)_\rho - \log_2(4\varepsilon/\eta^2), \tag{3.10}$$

for all $\eta \in (0,\varepsilon)$, is a lower bound on the ε-one-shot randomness-assisted, classical capacity of the cq channel in (3.1). Although this result is already known from [WR12], the development in this section is an important building block for the wiretap channel result in Section 4.2, and so we go through it in full detail for the sake of completeness. Also, the approach given here uses position-based decoding for the cq channel.

Fix a probability distribution p_X over the channel input alphabet. Consider the following classical–classical state:

$$\rho_{XX'} \equiv \sum_x p_X(x)|x\rangle\langle x|_X \otimes |x\rangle\langle x|_{X'}, \tag{3.11}$$

which we can think of as representing shared randomness. Let ρ_{XB} denote the following state, which results from sending the X' system of $\rho_{XX'}$ through the channel $\mathcal{N}_{X'\to B}$:

$$\rho_{XB} \equiv \mathcal{N}_{X'\to B}(\rho_{XX'}) = \sum_x p_X(x)|x\rangle\langle x|_X \otimes \rho_B^x. \tag{3.12}$$

The coding scheme works as follows. Let Alice and Bob share M copies of the state $\rho_{XX'}$, so that their shared state is

$$\rho_{X^M X'^M} \equiv \rho_{X_1X_1'} \otimes \cdots \otimes \rho_{X_MX_M'} = \rho_{XX'}^{\otimes M}. \tag{3.13}$$

Alice has the systems labeled by X', and Bob has the systems labeled by X. If Alice would like to communicate message m to Bob, then she simply sends system X_m' over the classical–quantum channel. In such a case, the reduced state for Bob is as follows:

$$\rho_{X^M B}^m \equiv \rho_{X_1} \otimes \cdots \otimes \rho_{X_{m-1}} \otimes \rho_{X_{m+1}} \otimes \cdots \otimes \rho_{X_M} \otimes \rho_{X_mB}. \tag{3.14}$$

Observe that each state $\rho_{X^M B}^m$ is related to the first one $\rho_{X^M B}^1$ by a permutation $\pi(m)$ of the X^M systems:

$$W_{X^M}^{\pi(m)} \rho_{X^M B}^1 W_{X^M}^{\pi(m)\dagger} = \rho_{X^M B}^m, \tag{3.15}$$

where $W_{X^M}^{\pi(m)}$ is a unitary representation of the permutation $\pi(m)$.

If Bob has a way of distinguishing the joint state ρ_{XB} from the product state $\rho_X \otimes \rho_B$, then with high probability, he will be able to figure out which message m was communicated. Let T_{XB} denote a test (measurement operator) satisfying $0 \le T_{XB} \le I_{XB}$, which we think of as identifying ρ_{XB} with high probability ($\ge 1-\varepsilon$) and for which the complementary operator $I_{XB} - T_{XB}$ identifies $\rho_X \otimes \rho_B$ with the highest probability subject to the constraint $\mathrm{Tr}\{T_{XB}\rho_{XB}\} \ge 1-\varepsilon$. From such a test, we form the following measurement operator:

$$\Gamma_{X^M B}^m \equiv T_{X_mB} \otimes I_{X_1} \otimes \cdots \otimes I_{X_{m-1}} \otimes I_{X_{m+1}} \otimes \cdots \otimes I_{X_M}, \tag{3.16}$$

which we think of as a test to figure out whether the reduced state on systems $X_m B$ is ρ_{XB} or $\rho_X \otimes \rho_B$. Observe that each message operator $\Gamma^m_{X^M B}$ is related to the first one $\Gamma^1_{X^M B}$ by a permutation $\pi(m)$ of the X^M systems:

$$W^{\pi(m)}_{X^M} \Gamma^1_{X^M B} W^{\pi(m)\dagger}_{X^M} = \Gamma^m_{X^M B}. \tag{3.17}$$

If message m is transmitted and the measurement operator $\Gamma^m_{X^M B}$ acts, then the probability of it accepting is

$$\mathrm{Tr}\{\Gamma^m_{X^M B} \rho^m_{X^M B}\} = \mathrm{Tr}\{T_{XB} \rho_{XB}\}. \tag{3.18}$$

If however the measurement operator $\Gamma^{m'}_{X^M B}$ acts, where $m' \neq m$, then the probability of it accepting is

$$\mathrm{Tr}\{\Gamma^{m'}_{X^M B} \rho^m_{X^M B}\} = \mathrm{Tr}\{T_{XB} [\rho_X \otimes \rho_B]\}. \tag{3.19}$$

From these measurement operators, we then form a square-root measurement as follows:

$$\Lambda^m_{X^M B} \equiv \left(\sum_{m'=1}^M \Gamma^{m'}_{X^M B} \right)^{-1/2} \Gamma^m_{X^M B} \left(\sum_{m'=1}^M \Gamma^{m'}_{X^M B} \right)^{-1/2}. \tag{3.20}$$

Again, each message operator $\Lambda^m_{X^M B}$ is related to the first one $\Lambda^1_{X^M B}$ by a permutation of the X^M systems:

$$\begin{aligned}
&W^{\pi(m)}_{X^M} \Lambda^1_{X^M B} W^{\pi(m)\dagger}_{X^M} \\
&= W^{\pi(m)}_{X^M} \left(\sum_{m'=1}^M \Gamma^{m'}_{X^M B} \right)^{-1/2} \Gamma^1_{X^M B} \left(\sum_{m'=1}^M \Gamma^{m'}_{X^M B} \right)^{-1/2} W^{\pi(m)\dagger}_{X^M} && (3.21) \\
&= W^{\pi(m)}_{X^M} \left(\sum_{m'=1}^M \Gamma^{m'}_{X^M B} \right)^{-1/2} W^{\pi(m)\dagger}_{X^M} W^{\pi(m)}_{X^M} \Gamma^1_{X^M B} W^{\pi(m)\dagger}_{X^M} W^{\pi(m)}_{X^M} \left(\sum_{m'=1}^M \Gamma^{m'}_{X^M B} \right)^{-1/2} W^{\pi(m)\dagger}_{X^M} && (3.22) \\
&= \left(\sum_{m'=1}^M W^{\pi(m)}_{X^M} \Gamma^{m'}_{X^M B} W^{\pi(m)\dagger}_{X^M} \right)^{-1/2} \Gamma^m_{X^M B} \left(\sum_{m'=1}^M W^{\pi(m)}_{X^M} \Gamma^{m'}_{X^M B} W^{\pi(m)\dagger}_{X^M} \right)^{-1/2} && (3.23) \\
&= \left(\sum_{m'=1}^M \Gamma^{m'}_{X^M B} \right)^{-1/2} \Gamma^m_{X^M B} \left(\sum_{m'=1}^M \Gamma^{m'}_{X^M B} \right)^{-1/2}. && (3.24)
\end{aligned}$$

This is called the position-based decoder and was analyzed in [AJW17] for the case of entanglement-assisted communication. The error probability under this coding scheme is as follows for each message m:

$$\mathrm{Tr}\{(I_{R^M B} - \Lambda^m_{R^M B}) \rho^m_{R^M B}\}. \tag{3.25}$$

The error probability is in fact the same for each message, due to the observations in (3.15) and (3.21)–(3.24):

$$\begin{aligned}
\mathrm{Tr}\{(I_{R^M B} - \Lambda^1_{R^M B}) \rho^1_{R^M B}\} &= \mathrm{Tr}\{(I_{R^M B} - \Lambda^1_{R^M B}) W^{\pi(m)\dagger}_{X^M} W^{\pi(m)}_{X^M} \rho^1_{R^M B} W^{\pi(m)\dagger}_{X^M} W^{\pi(m)}_{X^M}\} && (3.26) \\
&= \mathrm{Tr}\{(I_{R^M B} - W^{\pi(m)}_{X^M} \Lambda^1_{R^M B} W^{\pi(m)\dagger}_{X^M}) W^{\pi(m)}_{X^M} \rho^1_{R^M B} W^{\pi(m)\dagger}_{X^M}\} && (3.27) \\
&= \mathrm{Tr}\{(I_{R^M B} - \Lambda^m_{R^M B}) \rho^m_{R^M B}\}. && (3.28)
\end{aligned}$$

So let us analyze the error probability for the first message $m = 1$. Applying the Hayashi-Nagaoka operator inequality in (2.24), with $S = \Gamma^1_{X^M B}$, $T = \sum_{m' \neq 1} \Gamma^{m'}_{X^M B}$, $c_{\mathrm{I}} \equiv 1 + c$, and $c_{\mathrm{II}} = 2 + c + c^{-1}$ for $c > 0$, we find that this error probability can be bounded from above as

$$\begin{aligned}
&\operatorname{Tr}\{(I_{R^M B} - \Lambda^1_{R^M B})\rho^1_{R^M B}\} \\
&\leq c_{\mathrm{I}} \operatorname{Tr}\{(I_{X^M B} - \Gamma^1_{X^M B})\rho^1_{X^M B}\} + c_{\mathrm{II}} \sum_{m' \neq 1} \operatorname{Tr}\{\Gamma^{m'}_{X^M B}\rho^1_{X^M B}\} && (3.29) \\
&= c_{\mathrm{I}} \operatorname{Tr}\{(I_{XB} - T_{XB})\rho_{XB}\} + c_{\mathrm{II}} \sum_{m' \neq 1} \operatorname{Tr}\{T_{RB}\left[\rho_X \otimes \rho_B\right]\} && (3.30) \\
&= c_{\mathrm{I}} \operatorname{Tr}\{(I_{XB} - T_{XB})\rho_{XB}\} + c_{\mathrm{II}}(M-1) \operatorname{Tr}\{T_{XB}\left[\rho_X \otimes \rho_B\right]\}. && (3.31)
\end{aligned}$$

Consider the hypothesis testing mutual information:

$$I^{\varepsilon}_H(X;B)_\rho \equiv D^{\varepsilon}_H(\rho_{XB} \| \rho_X \otimes \rho_B), \tag{3.32}$$

where

$$D^{\varepsilon}_H(\rho\|\sigma) \equiv -\log_2 \inf_{\Lambda} \left\{\operatorname{Tr}\{\Lambda\sigma\} : 0 \leq \Lambda \leq I \wedge \operatorname{Tr}\{\Lambda\rho\} \geq 1 - \varepsilon\right\}. \tag{3.33}$$

Take the test T_{XB} in Bob's decoder to be Υ^*_{XB}, where Υ^*_{XB} is the optimal measurement operator for $I^{\varepsilon-\eta}_H(X;B)_\rho$ for $\eta \in (0, \varepsilon)$. Then the error probability is bounded as

$$\begin{aligned}
&\operatorname{Tr}\{(I_{X^M B} - \Lambda^1_{X^M B})\rho^1_{X^M B}\} \\
&\leq c_{\mathrm{I}} \operatorname{Tr}\{(I_{XB} - \Upsilon^*_{XB})\rho_{XB}\} + c_{\mathrm{II}} M \operatorname{Tr}\{\Upsilon^*_{XB}\left[\rho_X \otimes \rho_B\right]\} && (3.34) \\
&\leq c_{\mathrm{I}}\left(\varepsilon - \eta\right) + c_{\mathrm{II}} M 2^{-I^{\varepsilon-\eta}_H(X;B)_\rho}. && (3.35)
\end{aligned}$$

Now pick $c = \eta/(2\varepsilon - \eta)$ and we get that the last line above $= \varepsilon$, for

$$\log_2 M = I^{\varepsilon-\eta}_H(X;B)_\rho - \log_2(4\varepsilon/\eta^2). \tag{3.36}$$

Indeed, consider that we would like to have c such that

$$\varepsilon = c_{\mathrm{I}}\left(\varepsilon - \eta\right) + c_{\mathrm{II}} M 2^{-I^{\varepsilon-\eta}_H(X;B)_\rho}. \tag{3.37}$$

Rewriting this, we find that M should satisfy

$$\log_2 M = I^{\varepsilon-\eta}_H(X;B)_\rho + \log_2\left(\frac{\varepsilon - c_{\mathrm{I}}\left(\varepsilon - \eta\right)}{c_{\mathrm{II}}}\right). \tag{3.38}$$

Picking $c = \eta/(2\varepsilon - \eta)$ then implies (after some algebra) that

$$\frac{\varepsilon - c_{\mathrm{I}}\left(\varepsilon - \eta\right)}{c_{\mathrm{II}}} = \frac{\eta^2}{4\varepsilon}. \tag{3.39}$$

So the quantity $I^{\varepsilon-\eta}_H(X;B)_\rho - \log_2(4\varepsilon/\eta^2)$ represents a lower bound on the ε-one-shot randomness-assisted, classical capacity of the cq channel in (3.1). The bound holds for both average error probability and maximal error probability, and this coincidence is due to the protocol having the assistance of shared randomness.

3.3 Lower bound on the one-shot classical capacity

Now I show how to derandomize the above randomness-assisted code. The main result of this section is the following lower bound on the ε-one shot classical capacity of the cq channel in (3.1), holding for all $\eta \in (0, \varepsilon)$:

$$\log_2 M^*_{\text{pub}}(\varepsilon) \geq I_H^{\varepsilon-\eta}(X;B)_\rho - \log_2(4\varepsilon/\eta^2). \tag{3.40}$$

Again, note that although this result is already known from [WR12], the development in this section is an important building block for the wiretap channel result in Section 4.2. As stated previously, the approach given here uses position-based decoding for the cq channel.

By the reasoning from the previous section, we have the following bound on the average error probability for a randomness-assisted code:

$$\frac{1}{M}\sum_{m=1}^{M} \text{Tr}\{(I_{X^M B} - \Lambda^m_{X^M B})\rho^m_{X^M B}\} \leq \varepsilon, \tag{3.41}$$

if

$$\log_2 M = I_H^{\varepsilon-\eta}(X;B)_\rho - \log_2(4\varepsilon/\eta^2). \tag{3.42}$$

So let us analyze the expression $\text{Tr}\{(I_{X^M B} - \Lambda^m_{X^M B})\rho^m_{X^M B}\}$. By definition, it follows that

$$\rho^m_{X^M B} = \sum_{x_1,\ldots,x_M} p_X(x_1)\cdots p_X(x_M)|x_1,\ldots,x_M\rangle\langle x_1,\ldots,x_M|_{X_1\cdots X_M} \otimes \rho_B^{x_m}. \tag{3.43}$$

Also, recall that Υ^*_{XB} is optimal for $I_H^{\varepsilon-\eta}(X;B)_\rho$, which implies that

$$\text{Tr}\{\Upsilon^*_{XB}\rho_{XB}\} \geq 1 - (\varepsilon - \eta), \tag{3.44}$$

$$\text{Tr}\{\Upsilon^*_{XB}[\rho_X \otimes \rho_B]\} = 2^{-I_H^{\varepsilon-\eta}(X;B)_\rho}. \tag{3.45}$$

But consider that

$$\text{Tr}\{\Upsilon^*_{XB}\rho_{XB}\} = \text{Tr}\left\{\Upsilon^*_{XB}\sum_x p_X(x)|x\rangle\langle x|_X \otimes \rho_B^x\right\} \tag{3.46}$$

$$= \sum_x p_X(x)\,\text{Tr}\{\langle x|_X \Upsilon^*_{XB}|x\rangle_X \rho_B^x\} \tag{3.47}$$

$$= \sum_x p_X(x)\,\text{Tr}\{Q_B^x \rho_B^x\}, \tag{3.48}$$

where we define

$$Q_B^x \equiv \langle x|_X \Upsilon^*_{XB}|x\rangle_X. \tag{3.49}$$

Similarly, we have that

$$\text{Tr}\{\Upsilon^*_{XB}[\rho_X \otimes \rho_B]\} = \text{Tr}\left\{\Upsilon^*_{XB}\sum_x p_X(x)|x\rangle\langle x|_X \otimes \rho_B\right\} \tag{3.50}$$

$$= \sum_x p_X(x)\,\text{Tr}\{\langle x|_X \Upsilon^*_{XB}|x\rangle_X \rho_B\} \tag{3.51}$$

$$= \sum_x p_X(x)\,\text{Tr}\{Q_B^x \rho_B\}. \tag{3.52}$$

This demonstrates that it suffices to take the optimal measurement operator Υ^*_{XB} to be $\sum_x |x\rangle\langle x|_X \otimes Q^x_B$, with Q^x_B defined as in (3.49), and this will achieve the same optimal value as Υ^*_{XB} does.

Taking Υ^*_{XB} as such, now consider that

$$\Gamma^m_{X^M B} = \Upsilon^*_{X_m B} \otimes I_{X_1} \otimes \cdots \otimes I_{X_{m-1}} \otimes I_{X_{m+1}} \otimes \cdots \otimes I_{X_M} \tag{3.53}$$

$$= \sum_{x_m} |x_m\rangle\langle x_m|_{X_m} \otimes Q^{x_m}_B \otimes I_{X_1} \otimes \cdots \otimes I_{X_{m-1}} \otimes I_{X_{m+1}} \otimes \cdots \otimes I_{X_M} \tag{3.54}$$

$$= \sum_{x_1,\ldots,x_M} |x_1,\ldots,x_M\rangle\langle x_1,\ldots,x_M|_{X^M} \otimes Q^{x_m}_B. \tag{3.55}$$

Then this implies that

$$\left(\sum_{m'=1}^M \Gamma^{m'}_{X^M B}\right)^{-1/2} = \left(\sum_{m'=1}^M \sum_{x_1,\ldots,x_M} |x_1,\ldots,x_M\rangle\langle x_1,\ldots,x_M|_{X^M} \otimes Q^{x_{m'}}_B\right)^{-1/2} \tag{3.56}$$

$$= \left(\sum_{x_1,\ldots,x_M} |x_1,\ldots,x_M\rangle\langle x_1,\ldots,x_M|_{X^M} \otimes \sum_{m'=1}^M Q^{x_{m'}}_B\right)^{-1/2} \tag{3.57}$$

$$= \sum_{x_1,\ldots,x_M} |x_1,\ldots,x_M\rangle\langle x_1,\ldots,x_M|_{X^M} \otimes \left(\sum_{m'=1}^M Q^{x_{m'}}_B\right)^{-1/2}, \tag{3.58}$$

so that

$$\Lambda^m_{X^M B} = \left(\sum_{m'=1}^M \Gamma^{m'}_{X^M B}\right)^{-1/2} \Gamma^m_{X^M B} \left(\sum_{m'=1}^M \Gamma^{m'}_{X^M B}\right)^{-1/2} \tag{3.59}$$

$$= \sum_{x_1,\ldots,x_M} |x_1,\ldots,x_M\rangle\langle x_1,\ldots,x_M|_{X^M} \otimes \Omega^{x_m}_B, \tag{3.60}$$

where

$$\Omega^{x_m}_B \equiv \left(\sum_{m'=1}^M Q^{x_{m'}}_B\right)^{-1/2} Q^{x_m}_B \left(\sum_{m'=1}^M Q^{x_{m'}}_B\right)^{-1/2}. \tag{3.61}$$

Observe that $\{\Omega^{x_m}_B\}_{m=1}^M$ is a POVM on the support of $\sum_{m'=1}^M Q^{x_{m'}}_B$ and can be completed to a POVM on the full space by adding $\Omega^{x_0}_B \equiv I_B - \sum_{m'=1}^M Q^{x_{m'}}_B$. By employing (3.43) and (3.60), we find that

$$\mathrm{Tr}\{(I_{X^M B} - \Lambda^m_{X^M B})\rho^m_{X^M B}\} = \sum_{x_1,\ldots,x_M} p_X(x_1)\cdots p_X(x_M)\,\mathrm{Tr}\{(I_B - \Omega^{x_m}_B)\rho^{x_m}_B\}, \tag{3.62}$$

so that the average error probability is as follows:

$$\frac{1}{M}\sum_{m=1}^M \mathrm{Tr}\{(I_{X^M B} - \Lambda^m_{X^M B})\rho^m_{X^M B}\}$$

$$= \frac{1}{M}\sum_{m=1}^M \sum_{x_1,\ldots,x_M} p_X(x_1)\cdots p_X(x_M)\,\mathrm{Tr}\{(I_B - \Omega^{x_m}_B)\rho^{x_m}_B\} \tag{3.63}$$

$$= \sum_{x_1,\ldots,x_M} p_X(x_1)\cdots p_X(x_M)\left[\frac{1}{M}\sum_{m=1}^M \mathrm{Tr}\{(I_B - \Omega^{x_m}_B)\rho^{x_m}_B\}\right]. \tag{3.64}$$

The last line above is the same as the usual "Shannon trick" of exchanging the average over the messages with the expectation over a random choice of code. By employing the bound in (3.41), we find that

$$\sum_{x_1,\ldots,x_M} p_X(x_1)\cdots p_X(x_M)\left[\frac{1}{M}\sum_{m=1}^{M}\operatorname{Tr}\{(I_B-\Omega_B^{x_m})\rho_B^{x_m}\}\right]\leq\varepsilon. \tag{3.65}$$

Then there exists a particular set of values of $x_1, \ldots, x_M$ such that

$$\frac{1}{M}\sum_{m=1}^{M}\operatorname{Tr}\{(I_B-\Omega_B^{x_m})\rho_B^{x_m}\}\leq\varepsilon. \tag{3.66}$$

This sequence $x_1, \ldots, x_M$ constitutes the codewords and $\{\Omega_B^{x_m}\}_{m=1}^M$ is a corresponding POVM that can be used as a decoder. The number of bits that the code can transmit is equal to $\log_2 M = I_H^{\varepsilon-\eta}(X;B)_\rho - \log_2(4\varepsilon/\eta^2)$. No shared randomness is required for this code (it is now derandomized).

Remark 3 *To achieve maximal error probability 2ε, one can remove the worst half of the codewords, and then a lower bound on the achievable number of bits is*

$$I_H^{\varepsilon-\eta}(X;B)_\rho - \log_2 2 - \log_2(4\varepsilon/\eta^2) = I_H^{\varepsilon-\eta}(X;B)_\rho - \log_2(8\varepsilon/\eta^2). \tag{3.67}$$

4 Private classical communication

4.1 Definition of the one-shot private classical capacity

Now suppose that Alice, Bob, and Eve are connected by a classical-input quantum-quantum-output (cqq) channel of the following form:

$$x\to\rho_{BE}^x, \tag{4.1}$$

where Bob has system B and Eve system E. The fully quantum version of this channel is as follows:

$$\mathcal{N}_{X'\to BE}(\sigma_{X'}) = \sum_x \langle x|_{X'}\sigma_{X'}|x\rangle_{X'}\rho_{BE}^x, \tag{4.2}$$

where $\{|x\rangle_{X'}\}_x$ is some orthonormal basis.

We define the one-shot private classical capacity in the following way. Let $M\in\mathbb{N}$ and $\varepsilon\in(0,1)$. An (M,ε) private communication code consists of a collection of probability distributions $\{p_{X|M}(x|m)\}_{m=1}^M$ (one for each message m) and a decoding POVM $\{\Lambda_B^m\}_{m=1}^M$, such that

$$\frac{1}{M}\sum_{m=1}^{M}\frac{1}{2}\left\|\mathcal{M}_{B\to\hat{M}}(\rho_{BE}^m) - |m\rangle\langle m|_{\hat{M}}\otimes\sigma_E\right\|_1\leq\varepsilon. \tag{4.3}$$

We refer to the left-hand side of the above inequality as the *privacy error*. In the above, $\{|m\rangle_{\hat{M}}\}_{m=1}^M$ is an orthonormal basis, the state σ_E can be any state, we define the state ρ_{BE}^m as

$$\rho_{BE}^m = \sum_x p_{X|M}(x|m)\rho_{BE}^x, \tag{4.4}$$

and the measurement channel $\mathcal{M}_{B\to\hat{M}}$ as

$$\mathcal{M}_{B\to\hat{M}}(\omega_B) \equiv \sum_m \operatorname{Tr}\{\Lambda_B^m \omega_B\}|m\rangle\langle m|_{\hat{M}}. \tag{4.5}$$

For a given channel $\mathcal{N}_{X'\to BE}$ and ε, the one-shot private classical capacity is equal to $\log_2 M^*_{\text{priv}}(\varepsilon)$, where $M^*_{\text{priv}}(\varepsilon)$ is the largest M such that (4.3) can be satisfied for a fixed ε.

The condition in (4.3) combines the reliable decoding and security conditions into a single average error criterion. We can see how it represents a generalization of the error criterion in (3.3), which was for public classical communication over a cq channel. One could have a different definition of one-shot private capacity, in which there are two separate criteria, but the approach above will be beneficial for our purposes. In any case, a code satisfying (4.3) satisfies the two separate criteria as well, as is easily seen by invoking the monotonicity of trace distance.[2] Having a single error criterion for private capacity is the same as the approach taken in [HHHO09] and [WTB17], and in the latter paper, it was shown that notions of asymptotic private capacity are equivalent when using either a single error criterion or two separate error criteria.

4.2 Lower bound on the one-shot private classical capacity

The main result of this section is the following lower bound on the ε-one shot private capacity of a cq wiretap channel, holding for all $\varepsilon_1, \varepsilon_2 \in (0,1)$, such that $\varepsilon_1 + \sqrt{\varepsilon_2} \in (0,1)$, and $\eta_1 \in (0, \varepsilon_1)$ and $\eta_2 \in (0, \sqrt{\varepsilon_2})$:

$$\log_2 M^*_{\text{priv}}(\varepsilon_1 + \sqrt{\varepsilon_2}) \geq I_H^{\varepsilon_1-\eta_1}(X;B)_\rho - \widetilde{I}_{\max}^{\sqrt{\varepsilon_2}-\eta_2}(E;X)_\rho - \log_2(4\varepsilon_1/\eta_1^2) - 2\log_2(1/\eta_2)\,. \tag{4.6}$$

To begin with, we allow Alice, Bob, and Eve shared randomness of the following form:

$$\rho_{XX'X''} \equiv \sum_x p_X(x)|x\rangle\langle x|_X \otimes |x\rangle\langle x|_{X'} \otimes |x\rangle\langle x|_{X''}, \tag{4.7}$$

where Bob has the X system, Alice the X' system, and Eve the X'' system. It is natural here to let Eve share the randomness as well, and this amounts to giving her knowledge of the code to be used. Let $\rho_{XX''BE}$ denote the state resulting from sending the X' system through the channel $\mathcal{N}_{X'\to BE}$ in (4.2):

$$\rho_{XX''BE} \equiv \sum_x p_X(x)|x\rangle\langle x|_X \otimes \rho_{BE}^x \otimes |x\rangle\langle x|_{X''}. \tag{4.8}$$

The coding scheme that Alice and Bob use is as follows. There is the message $m \in \{1,\ldots,M\}$ and a local key $k \in \{1,\ldots,K\}$. The local key k represents local, uniform randomness that Alice has, but which is not accessible to Bob or Eve. We assume that Alice, Bob, and Eve share MK copies of the state in (4.7) before communication begins, and we denote this state as

$$\rho_{X^{MK}X'^{MK}X''^{MK}} = \rho_{X_{1,1}X'_{1,1}X''_{1,1}} \otimes \cdots \otimes \rho_{X_{M,K}X'_{M,K}X''_{M,K}} = \rho_{XX'X''}^{\otimes MK}. \tag{4.9}$$

[2] Indeed, starting with (4.3) and applying monotonicity of trace distance under partial trace of the E system, we get that $\frac{1}{M}\sum_{m=1}^M \frac{1}{2}\|\mathcal{M}_{B\to\hat{M}}(\rho_B^m) - |m\rangle\langle m|_{\hat{M}}\|_1 \leq \varepsilon$. Recalling (3.3), we can interpret this as asserting that the decoding error probability does not exceed ε. Doing the same but considering a partial trace over the B system implies that $\frac{1}{M}\sum_{m=1}^M \frac{1}{2}\|\rho_E^m - \sigma_E\|_1 \leq \varepsilon$, which is a security criterion. So we get that the conventional two separate criteria are satisfied if a code satisfies the single privacy error criterion in (4.3).

To send the message m, Alice picks k uniformly at random from the set $\{1,\ldots,K\}$. She then sends the (m,k)th X' system through the channel $\mathcal{N}_{X'\to BE}$. Thus, when m and k are chosen, the reduced state on Bob and Eve's systems is

$$\rho^{m,k}_{X^{MK}X''^{MK}BE} = \rho_{X_{1,1}X''_{1,1}} \otimes \cdots \otimes \rho_{X_{m,k-1}X''_{m,k-1}} \otimes \rho_{X_{m,k}X''_{m,k}BE} \otimes \rho_{X_{m,k+1}X''_{m,k+1}} \otimes \cdots \otimes \rho_{X_{M,K}X''_{M,K}}, \tag{4.10}$$

and the state of Bob's systems is

$$\rho^{m,k}_{X^{MK}B} = \rho_{X_{1,1}} \otimes \cdots \otimes \rho_{X_{m,k-1}} \otimes \rho_{X_{m,k}B} \otimes \rho_{X_{m,k+1}} \otimes \cdots \otimes \rho_{X_{M,K}}. \tag{4.11}$$

For Bob to decode, he uses the position-based decoder to decode both the message m and the local key k. Let $\{\Lambda^{m,k}_{X^{M,K}B}\}_{m,k}$ denote his decoding POVM. By the reasoning from Section 3.2, as long as

$$\log_2 MK = I_H^{\varepsilon_1-\eta_1}(X;B)_\rho - \log_2(4\varepsilon_1/\eta_1^2), \tag{4.12}$$

where $\varepsilon_1 \in (0,1)$ and $\eta_1 \in (0,\varepsilon_1)$, then we have the following bound holding for all m,k:

$$\operatorname{Tr}\{(I_{X^{M,K}B} - \Lambda^{m,k}_{X^{MK}B})\rho^{m,k}_{X^{MK}B}\} \leq \varepsilon_1, \tag{4.13}$$

where $\Lambda^{m,k}_{X^{MK}B}$ is defined as in Sections 3.2 and 3.3. By the reasoning from Section 3.3, we can also write (4.13) as

$$\sum_{x_{1,1},\ldots,x_{M,K}} p_X(x_{1,1})\cdots p_X(x_{M,K}) \left[\frac{1}{MK}\sum_{m=1}^{M}\sum_{k=1}^{K} \operatorname{Tr}\{(I_B - \Omega_B^{x_{m,k}})\rho_B^{x_{m,k}}\}\right] \leq \varepsilon_1, \tag{4.14}$$

with $\Omega_B^{x_{m,k}}$ defined as in Section 3.3. Define the following measurement channels:

$$\mathcal{M}_{B\to\hat{M}}(\omega_B) \equiv \sum_{m,k} \operatorname{Tr}\{\Omega_B^{x_{m,k}}\omega_B\}|m\rangle\langle m|_{\hat{M}}, \tag{4.15}$$

$$\mathcal{M}'_{B\to\hat{M}\hat{K}}(\omega_B) \equiv \sum_{m,k} \operatorname{Tr}\{\Omega_B^{x_{m,k}}\omega_B\}|m\rangle\langle m|_{\hat{M}} \otimes |k\rangle\langle k|_{\hat{K}}, \tag{4.16}$$

with it being clear that $\operatorname{Tr}_{\hat{K}} \circ \mathcal{M}'_{B\to\hat{M}\hat{K}} = \mathcal{M}_{B\to\hat{M}}$. Consider that

$$\begin{aligned}
&\frac{1}{2}\left\|\mathcal{M}'_{B\to\hat{M}\hat{K}}\left(\rho_B^{x_{m,k}}\right) - |m\rangle\langle m|_{\hat{M}} \otimes |k\rangle\langle k|_{\hat{K}}\right\|_1 \\
&= \frac{1}{2}\left\|\sum_{m',k'} \operatorname{Tr}\{\Omega_B^{x_{m',k'}}\rho_B^{x_{m,k}}\}|m'\rangle\langle m'|_{\hat{M}} \otimes |k'\rangle\langle k'|_{\hat{K}} - |m\rangle\langle m|_{\hat{M}} \otimes |k\rangle\langle k|_{\hat{K}}\right\|_1 && (4.17)\\
&= \frac{1}{2}\left\|\sum_{(m',k')\neq(m,k)} \operatorname{Tr}\{\Omega_B^{x_{m',k'}}\rho_B^{x_{m,k}}\}|m'\rangle\langle m'|_{\hat{M}} \otimes |k'\rangle\langle k'|_{\hat{K}} - (1-\operatorname{Tr}\{\Omega_B^{x_{m,k}}\rho_B^{x_{m,k}}\})|m\rangle\langle m|_{\hat{M}} \otimes |k\rangle\langle k|_{\hat{K}}\right\|_1 && (4.18)\\
&= 1 - \operatorname{Tr}\{\Omega_B^{x_{m,k}}\rho_B^{x_{m,k}}\} && (4.19)\\
&= \operatorname{Tr}\{(I_B - \Omega_B^{x_{m,k}})\rho_B^{x_{m,k}}\}. && (4.20)
\end{aligned}$$

Now averaging the above quantity over m, k, and $x_{1,1}$, $\ldots$, $x_{M,K}$, and applying the condition in (4.14), we get that

$$\sum_{x_{1,1},\ldots,x_{M,K}} p_X(x_{1,1})\cdots p_X(x_{M,K}) \left[\frac{1}{MK} \sum_{m,k} \frac{1}{2} \left\| \mathcal{M}'_{B\to\hat{M}\hat{K}} \left(\rho_B^{x_{m,k}} \right) - |m\rangle\langle m|_{\hat{M}} \otimes |k\rangle\langle k|_{\hat{K}} \right\|_1 \right] \leq \varepsilon_1. \tag{4.21}$$

Applying convexity of the trace distance to bring the average over k inside and monotonicity with respect to partial trace over system $\hat{K}$ to the left-hand side of (4.21), we find that

$$\begin{aligned} &\sum_{x_{1,1},\ldots,x_{M,K}} p_X(x_{1,1})\cdots p_X(x_{M,K}) \left[\frac{1}{M} \sum_{m=1}^{M} \frac{1}{2} \left\| (\mathrm{Tr}_{\hat{K}} \circ \mathcal{M}'_{B\to\hat{M}\hat{K}}) \left(\frac{1}{K} \sum_{k=1}^{K} \rho_B^{x_{m,k}} \right) - |m\rangle\langle m|_{\hat{M}} \right\|_1 \right] \\ &= \sum_{x_{1,1},\ldots,x_{M,K}} p_X(x_{1,1})\cdots p_X(x_{M,K}) \left[\frac{1}{M} \sum_{m=1}^{M} \frac{1}{2} \left\| \mathcal{M}_{B\to\hat{M}} \left(\frac{1}{K} \sum_{k=1}^{K} \rho_B^{x_{m,k}} \right) - |m\rangle\langle m|_{\hat{M}} \right\|_1 \right] \leq \varepsilon_1. \end{aligned} \tag{4.22}$$

Let us define the state

$$\omega_E^{x_{m'},x_m} \equiv \frac{\frac{1}{K}\sum_{k,k'=1}^{K} \mathrm{Tr}_B\{\Omega_B^{x_{m',k'}} \rho_{BE}^{x_{m,k}}\}}{q(x_{m'}|x_m)}, \tag{4.23}$$

$$q(x_{m'}|x_m) \equiv \frac{1}{K} \sum_{k,k'=1}^{K} \mathrm{Tr}\{\Omega_B^{x_{m',k'}} \rho_{BE}^{x_{m,k}}\}. \tag{4.24}$$

Consider that

$$\sum_{m'} q(x_{m'}|x_m) \omega_E^{x_{m'},x_m} = \frac{1}{K} \sum_{k=1}^{K} \rho_E^{x_{m,k}}. \tag{4.25}$$

Then we can write

$$\mathcal{M}_{B\to\hat{M}} \left(\frac{1}{K} \sum_{k=1}^{K} \rho_{BE}^{x_{m,k}} \right) = \sum_{m'} q(x_{m'}|x_m) |m'\rangle\langle m'|_{\hat{M}} \otimes \omega_E^{x_{m'},x_m}, \tag{4.26}$$

so that

$$\mathcal{M}_{B\to\hat{M}} \left(\frac{1}{K} \sum_{k=1}^{K} \rho_B^{x_{m,k}} \right) = \sum_{m'} q(x_{m'}|x_m) |m'\rangle\langle m'|_{\hat{M}}. \tag{4.27}$$

Using these observations, we can finally write

$$\begin{aligned}
&\frac{1}{M}\sum_{m=1}^{M}\frac{1}{2}\left\|\mathcal{M}_{B\to\hat{M}}\left(\frac{1}{K}\sum_{k=1}^{K}\rho_{BE}^{x_{m,k}}\right)-|m\rangle\langle m|_{\hat{M}}\otimes\frac{1}{K}\sum_{k=1}^{K}\rho_{E}^{x_{m,k}}\right\|_1 \\
&=\frac{1}{M}\sum_{m=1}^{M}\frac{1}{2}\left\|\sum_{m'}q(x_{m'}|x_m)|m'\rangle\langle m'|_{\hat{M}}\otimes\omega_E^{x_{m'},x_m}-|m\rangle\langle m|_{\hat{M}}\otimes\sum_{m'}q(x_{m'}|x_m)\omega_E^{x_{m'},x_m}\right\|_1 \qquad (4.28)\\
&\leq\frac{1}{M}\sum_{m=1}^{M}\sum_{m'}q(x_{m'}|x_m)\left[\frac{1}{2}\left\||m'\rangle\langle m'|_{\hat{M}}\otimes\omega_E^{x_{m'},x_m}-|m\rangle\langle m|_{\hat{M}}\otimes\omega_E^{x_{m'},x_m}\right\|_1\right] \qquad (4.29)\\
&=\frac{1}{M}\sum_{m=1}^{M}\sum_{m'}q(x_{m'}|x_m)\left[\frac{1}{2}\left\||m'\rangle\langle m'|_{\hat{M}}-|m\rangle\langle m|_{\hat{M}}\right\|_1\right] \qquad (4.30)\\
&=\frac{1}{M}\sum_{m=1}^{M}\sum_{m'\neq m}q(x_{m'}|x_m) \qquad (4.31)\\
&=\frac{1}{M}\sum_{m=1}^{M}\frac{1}{2}\left\|\mathcal{M}_{B\to\hat{M}}\left(\frac{1}{K}\sum_{k=1}^{K}\rho_{B}^{x_{m,k}}\right)-|m\rangle\langle m|_{\hat{M}}\right\|_1. \qquad (4.32)
\end{aligned}$$

Combining with (4.22), the above development implies that

$$\begin{aligned}
&\sum_{x_{1,1},\ldots,x_{M,K}}p_X(x_{1,1})\cdots p_X(x_{M,K})\\
&\qquad\times\left[\frac{1}{M}\sum_{m=1}^{M}\frac{1}{2}\left\|\mathcal{M}_{B\to\hat{M}}\left(\frac{1}{K}\sum_{k=1}^{K}\rho_{BE}^{x_{m,k}}\right)-|m\rangle\langle m|_{\hat{M}}\otimes\frac{1}{K}\sum_{k=1}^{K}\rho_{E}^{x_{m,k}}\right\|_1\right]\leq\varepsilon_1. \qquad (4.33)
\end{aligned}$$

Now we consider the state on Eve's systems and the analysis of privacy. If m and k are fixed, then her state is

$$\rho_{X^{MK}E}^{m,k}=\rho_{X_{1,1}}\otimes\cdots\otimes\rho_{X_{m,k-1}}\otimes\rho_{X_{m,k}E}\otimes\rho_{X_{m,k+1}}\otimes\cdots\otimes\rho_{X_{M,K}}. \qquad (4.34)$$

(For simplicity of notation, in the above and what follows we are labeling her systems X'' as X.) However, k is chosen uniformly at random, and so conditioned on the message m being fixed, the state of Eve's systems is as follows:

$$\begin{aligned}
\rho_{X^{MK}E}^{m}&\equiv\frac{1}{K}\sum_{k=1}^{K}\rho_{X^{MK}E}^{m,k} \qquad (4.35)\\
&=\rho_{X_{1,1}}\otimes\cdots\otimes\rho_{X_{m-1,K}}\\
&\qquad\otimes\left[\frac{1}{K}\sum_{k=1}^{K}\rho_{X_{m,1}}\otimes\cdots\otimes\rho_{X_{m,k-1}}\otimes\rho_{X_{m,k}E}\otimes\rho_{X_{m,k+1}}\otimes\cdots\otimes\rho_{X_{m,K}}\right]\\
&\qquad\otimes\rho_{X_{m+1,1}}\otimes\cdots\otimes\rho_{X_{M,K}}. \qquad (4.36)
\end{aligned}$$

We would like to show for $\varepsilon_2\in(0,1)$ that

$$\frac{1}{2}\left\|\rho_{X^{MK}E}^{m}-\rho_{X^{MK}}\otimes\widetilde{\rho}_E\right\|_1\leq\varepsilon_2, \qquad (4.37)$$

for some state $\widetilde{\rho}_E$. By the invariance of the trace distance with respect to tensor-product states, i.e.,

$$\left\| \sigma \otimes \tau - \omega \otimes \tau \right\|_1 = \left\| \sigma - \omega \right\|_1, \tag{4.38}$$

we find that

$$\frac{1}{2} \left\| \rho^m_{X^{MK}E} - \rho_{X^{MK}} \otimes \widetilde{\rho}_E \right\|_1 \tag{4.39}$$

$$= \frac{1}{2} \left\| \rho^m_{X_{m,1}\cdots X_{m,K}E} - \rho_{X_{m,1}\cdots X_{m,K}} \otimes \widetilde{\rho}_E \right\|_1 \tag{4.40}$$

$$= \frac{1}{2} \left\| \frac{1}{K} \sum_{k=1}^{K} \rho_{X_{m,1}} \otimes \cdots \otimes \rho_{X_{m,k-1}} \otimes \left(\rho_{X_{m,k}E} - \rho_{X_{m,k}} \otimes \widetilde{\rho}_E \right) \otimes \rho_{X_{m,k+1}} \otimes \cdots \otimes \rho_{X_{m,K}} \right\|_1 . \tag{4.41}$$

From Lemma 2 and the relation in (2.4) between trace distance and purified distance, we find that if we pick K such that

$$\log_2 K = \widetilde{I}^{\sqrt{\varepsilon_2}-\eta_2}_{\max}(E;X)_\rho + 2\log_2(1/\eta_2), \tag{4.42}$$

then we are guaranteed that

$$\frac{1}{2} \left\| \rho^m_{X^{MK}E} - \rho_{X^{MK}} \otimes \widetilde{\rho}_E \right\|_1 \leq \sqrt{\varepsilon_2}, \tag{4.43}$$

where $\widetilde{\rho}_E$ is some state such that $P(\widetilde{\rho}_E, \rho_E) \leq \sqrt{\varepsilon_2} - \eta_2$.

Consider that we can rewrite

$$\frac{1}{2} \left\| \rho^m_{X^{MK}E} - \rho_{X^{MK}} \otimes \widetilde{\rho}_E \right\|_1 \tag{4.44}$$

$$= \frac{1}{2} \left\| \frac{1}{K} \sum_{k=1}^{K} \sum_{x_{1,1}\cdots x_{M,K}} p_X(x_{1,1}) \cdots p_X(x_{M,K}) |x_{1,1}\cdots x_{M,K}\rangle\langle x_{1,1}\cdots x_{M,K}|_{X^{M,K}} \otimes \left(\rho_E^{x_{m,k}} - \widetilde{\rho}_E \right) \right\|_1 \tag{4.45}$$

$$= \frac{1}{2} \left\| \sum_{x_{1,1}\cdots x_{M,K}} p_X(x_{1,1}) \cdots p_X(x_{M,K}) |x_{1,1}\cdots x_{M,K}\rangle\langle x_{1,1}\cdots x_{M,K}|_{X^{M,K}} \otimes \left(\frac{1}{K} \sum_{k=1}^{K} \rho_E^{x_{m,k}} - \widetilde{\rho}_E \right) \right\|_1 \tag{4.46}$$

$$= \sum_{x_{1,1}\cdots x_{M,K}} p_X(x_{1,1}) \cdots p_X(x_{M,K}) \left[\frac{1}{2} \left\| \frac{1}{K} \sum_{k=1}^{K} \rho_E^{x_{m,k}} - \widetilde{\rho}_E \right\|_1 \right] \leq \sqrt{\varepsilon_2}. \tag{4.47}$$

Applying (4.38) to (4.47), we find that

$$\sum_{x_{1,1}\cdots x_{M,K}} p_X(x_{1,1}) \cdots p_X(x_{M,K}) \left[\frac{1}{2} \left\| |m\rangle\langle m|_{\hat{M}} \otimes \frac{1}{K} \sum_{k=1}^{K} \rho_E^{x_{m,k}} - |m\rangle\langle m|_{\hat{M}} \otimes \widetilde{\rho}_E \right\|_1 \right] \leq \sqrt{\varepsilon_2}. \tag{4.48}$$

Putting together (4.12), (4.42), (4.33), and (4.48), we find that if

$$\log_2 M = I_H^{\varepsilon_1-\eta_1}(X;B)_\rho - \log_2(4\varepsilon_1/\eta_1^2) - \left[\widetilde{I}^{\sqrt{\varepsilon_2}-\eta_2}_{\max}(E;X)_\rho + 2\log_2(1/\eta_2) \right] \tag{4.49}$$

$$= I_H^{\varepsilon_1-\eta_1}(X;B)_\rho - \widetilde{I}^{\sqrt{\varepsilon_2}-\eta_2}_{\max}(E;X)_\rho - \log_2(4\varepsilon_1/\eta_1^2) - 2\log_2(1/\eta_2), \tag{4.50}$$

then we have by the triangle inequality that

$$\sum_{x_{1,1},\ldots,x_{M,K}} p_X(x_{1,1})\cdots p_X(x_{M,K}) \left[\frac{1}{2}\left\|\mathcal{M}_{B\to\hat{M}}\left(\frac{1}{K}\sum_{k=1}^{K}\rho_{BE}^{x_{m,k}}\right) - |m\rangle\langle m|_{\hat{M}}\otimes\widetilde{\rho}_E\right\|_1\right] \leq \varepsilon_1 + \sqrt{\varepsilon_2}. \tag{4.51}$$

So this gives what is achievable with shared randomness (again, no difference between average and maximal error if shared randomness is allowed).

We now show how to derandomize the code. We take the above and average over all messages m. We find that

$$\begin{aligned}
&\varepsilon_1 + \sqrt{\varepsilon_2} \\
&\geq \frac{1}{M}\sum_{m=1}^{M}\sum_{x_{1,1},\ldots,x_{M,K}} p_X(x_{1,1})\cdots p_X(x_{M,K}) \left[\frac{1}{2}\left\|\mathcal{M}_{B\to\hat{M}}\left(\frac{1}{K}\sum_{k=1}^{K}\rho_{BE}^{x_{m,k}}\right) - |m\rangle\langle m|_{\hat{M}}\otimes\widetilde{\rho}_E\right\|_1\right] &(4.52)\\
&= \sum_{x_{1,1},\ldots,x_{M,K}} p_X(x_{1,1})\cdots p_X(x_{M,K}) \left(\frac{1}{M}\sum_{m=1}^{M}\left[\frac{1}{2}\left\|\mathcal{M}_{B\to\hat{M}}\left(\frac{1}{K}\sum_{k=1}^{K}\rho_{BE}^{x_{m,k}}\right) - |m\rangle\langle m|_{\hat{M}}\otimes\widetilde{\rho}_E\right\|_1\right]\right). &(4.53)
\end{aligned}$$

So we can conclude that there exist particular values $x_{1,1}, \ldots, x_{M,K}$ such that

$$\frac{1}{M}\sum_{m=1}^{M}\left[\frac{1}{2}\left\|\mathcal{M}_{B\to\hat{M}}\left(\frac{1}{K}\sum_{k=1}^{K}\rho_{BE}^{x_{m,k}}\right) - |m\rangle\langle m|_{\hat{M}}\otimes\widetilde{\rho}_E\right\|_1\right] \leq \varepsilon_1 + \sqrt{\varepsilon_2}. \tag{4.54}$$

Thus, our final conclusion is that the number of achievable bits that can be sent such that the privacy error is no larger than $\varepsilon_1 + \sqrt{\varepsilon_2}$ is equal to

$$\log_2 M = I_H^{\varepsilon_1-\eta_1}(X;B)_\rho - \widetilde{I}_{\max}^{\sqrt{\varepsilon_2}-\eta_2}(E;X)_\rho - \log_2(4\varepsilon_1/\eta_1^2) - 2\log_2(1/\eta_2). \tag{4.55}$$

4.3 Second-order asymptotics for private classical communication

In this section, I show how the lower bound on one-shot private capacity leads to a non-trivial lower bound on the second-order coding rate of private communication over an i.i.d. cq wiretap channel. I also show how the bounds simplify for pure-state cq wiretap channels and when using binary phase-shift keying as a coding strategy for private communication over a pure-loss bosonic channel.

Applying Lemma 1 to (4.55) with $\gamma \in (0, \sqrt{\varepsilon} - \eta)$, we can take

$$\begin{aligned}
\log_2 M &= I_H^{\varepsilon_1-\eta_1}(X;B)_\rho - \widetilde{I}_{\max}^{\sqrt{\varepsilon_2}-\eta_2}(E;X)_\rho - \log_2(4\varepsilon_1/\eta_1^2) - 2\log_2(1/\eta_2) &(4.56)\\
&\geq I_H^{\varepsilon_1-\eta_1}(X;B)_\rho - D_{\max}^{\sqrt{\varepsilon_2}-\eta_2-\gamma}(\rho_{XE}\|\rho_X\otimes\rho_E) - \log_2(4\varepsilon_1/\eta_1^2) - 2\log_2(1/\eta_2) - \log_2(3/\gamma^2). &(4.57)
\end{aligned}$$

while still achieving the performance in (4.54).

Substituting an i.i.d. cq wiretap channel into the one-shot bounds, evaluating for such a case and using the expansions for I_H^ε in (2.18) and $D_{\max}^\varepsilon$ in (2.19), while taking $\eta_1 = \eta_2 = \gamma = 1/\sqrt{n}$, for sufficiently large n, we get that

$$\log_2 M^*_{\text{priv}}(n, \varepsilon_1 + \sqrt{\varepsilon_2}) \geq n\left[I(X;B)_\rho - I(X;E)_\rho\right] + \sqrt{nV(X;B)_\rho}\Phi^{-1}(\varepsilon_1) + \sqrt{nV(X;E)_\rho}\Phi^{-1}(\varepsilon_2) + O(\log n). \quad (4.58)$$

4.3.1 Example: Pure-state cq wiretap channel

Let us consider applying the inequality in (4.58) to a cq pure-state wiretap channel of the following form:

$$x \to |\psi^x\rangle\langle\psi^x|_B \otimes |\varphi^x\rangle\langle\varphi^x|_E, \quad (4.59)$$

in which the classical input x leads to a pure quantum state $|\psi^x\rangle\langle\psi^x|_B$ for Bob and a pure quantum state $|\varphi^x\rangle\langle\varphi^x|_E$ for Eve. This channel may seem a bit particular, but we discuss in the next section how one can induce such a channel from a practically relevant channel, known as the pure-loss bosonic channel. In order to apply the inequality in (4.58) to the channel in (4.59), we fix a distribution $p_X(x)$ over the input symbols, leading to the following classical–quantum state:

$$\rho_{XBE} \equiv \sum_x p_X(x)|x\rangle\langle x|_X \otimes |\psi^x\rangle\langle\psi^x|_B \otimes |\varphi^x\rangle\langle\varphi^x|_E. \quad (4.60)$$

It is well known and straightforward to calculate that the following simplifications occur

$$I(X;B)_\rho = H(B)_\rho = H(\rho_B), \quad (4.61)$$
$$I(X;E)_\rho = H(E)_\rho = H(\rho_E), \quad (4.62)$$

where $H(\sigma) \equiv -\operatorname{Tr}\{\sigma \log_2 \sigma\}$ denotes the quantum entropy of a state σ and

$$\rho_B = \sum_x p_X(x)|\psi^x\rangle\langle\psi^x|_B, \quad (4.63)$$
$$\rho_E = \sum_x p_X(x)|\varphi^x\rangle\langle\varphi^x|_E. \quad (4.64)$$

Proposition 4 below demonstrates that a similar simplification occurs for the information variance quantities in (4.58), in the special case of a pure-state cq wiretap channel. By employing it, we find the following lower bound on the second-order coding rate for a pure-state cq wiretap channel:

$$\log_2 M^*_{\text{priv}}(n, \varepsilon_1 + \sqrt{\varepsilon_2}) \geq n\left[H(\rho_B) - H(\rho_E)\right] + \sqrt{nV(\rho_B)}\Phi^{-1}(\varepsilon_1) + \sqrt{nV(\rho_E)}\Phi^{-1}(\varepsilon_2) + O(\log n), \quad (4.65)$$

where $V(\rho_B)$ and $V(\rho_E)$ are defined from (4.67) below.

Proposition 4 *Let*

$$\rho_{XB} = \sum_x p_X(x)|x\rangle\langle x|_X \otimes |\psi^x\rangle\langle\psi^x|_B \quad (4.66)$$

be a classical–quantum state corresponding to a pure-state ensemble $\{p_X(x), |\psi^x\rangle_B\}_x$. *Then the Holevo information variance* $V(X;B)_\rho = V(\rho_{XB}\|\rho_X \otimes \rho_B)$ *is equal to the entropy variance* $V(\rho_B)$ *of the expected state* $\rho_B = \sum_x p_X(x)|\psi^x\rangle\langle\psi^x|_B$, *where*

$$V(\sigma) = \mathrm{Tr}\{\sigma\,[-\log_2\sigma - H(\sigma)]^2\}. \tag{4.67}$$

That is, when ρ_{XB} *takes the special form in* (4.66), *the following equality holds*

$$V(X;B)_\rho = V(\rho_B). \tag{4.68}$$

Proof. For the cq state in (4.66), consider that $I(X;B)_\rho = H(B)_\rho = H(\rho_B)$. Furthermore, we have that

$$\log_2\rho_{XB} = \log_2\left[\sum_x p_X(x)|x\rangle\langle x|_X \otimes |\psi^x\rangle\langle\psi^x|_B\right] \tag{4.69}$$

$$= \sum_x \log_2\left(p_X(x)\right)|x\rangle\langle x|_X \otimes |\psi^x\rangle\langle\psi^x|_B, \tag{4.70}$$

which holds because the eigenvectors of ρ_{XB} are $\{|x\rangle_X \otimes |\psi^x\rangle_B\}_x$. Then

$$V(X;B) = V(\rho_{XB}\|\rho_X\otimes\rho_B) \tag{4.71}$$

$$= \mathrm{Tr}\{\rho_{XB}\,[\log_2\rho_{XB} - \log_2(\rho_X\otimes\rho_B)]^2\} - [I(X;B)_\rho]^2 \tag{4.72}$$

$$= \mathrm{Tr}\{\rho_{XB}\,[\log_2\rho_{XB} - \log_2\rho_X\otimes I_B - I_X\otimes\log_2\rho_B]^2\} - [H(B)_\rho]^2\,. \tag{4.73}$$

By direct calculation, we have that

$$\log_2\rho_{XB} - \log_2\rho_X\otimes I_B - I_X\otimes\log_2\rho_B$$
$$= \sum_x \log_2\left(p_X(x)\right)|x\rangle\langle x|_X\otimes|\psi^x\rangle\langle\psi^x|_B - \sum_x \log_2\left[p_X(x)\right]|x\rangle\langle x|_X\otimes I_B - \sum_x |x\rangle\langle x|_X\otimes\log_2\rho_B \tag{4.74}$$

$$= -\sum_x |x\rangle\langle x|_X\otimes\left[\log_2\left(p_X(x)\right)(I_B - |\psi^x\rangle\langle\psi^x|_B) + \log_2\rho_B\right]. \tag{4.75}$$

Observe that $I_B - |\psi^x\rangle\langle\psi^x|_B$ is the projection onto the space orthogonal to $|\psi^x\rangle_B$. Then we find that

$$[\log_2\rho_{XB} - \log_2\rho_X\otimes I_B - I_X\otimes\log_2\rho_B]^2$$
$$= \left[-\sum_x |x\rangle\langle x|_X\otimes\left[\log_2\left(p_X(x)\right)(I_B - |\psi^x\rangle\langle\psi^x|_B) + \log_2\rho_B\right]\right]^2 \tag{4.76}$$

$$= \sum_x |x\rangle\langle x|_X\otimes\left[\log_2\left(p_X(x)\right)(I_B - |\psi^x\rangle\langle\psi^x|_B) + \log_2\rho_B\right]^2. \tag{4.77}$$

Furthermore, we have that

$$[\log_2\left(p_X(x)\right)(I_B - |\psi^x\rangle\langle\psi^x|_B) + \log_2\rho_B]^2$$
$$= [\log_2\left(p_X(x)\right)]^2(I_B - |\psi^x\rangle\langle\psi^x|_B) + \log_2\left(p_X(x)\right)(I_B - |\psi^x\rangle\langle\psi^x|_B)(\log_2\rho_B)$$
$$+ \log_2\left(p_X(x)\right)(\log_2\rho_B)(I_B - |\psi^x\rangle\langle\psi^x|_B) + [\log_2\rho_B]^2 \tag{4.78}$$

So then, by direct calculation,

$$\operatorname{Tr}\{\rho_{XB}\left[\log_2 \rho_{XB} - \log_2 \rho_X \otimes I_B - I_X \otimes \log_2 \rho_B\right]^2\} \tag{4.79}$$

$$= \operatorname{Tr}\left\{\left[\sum_{x'} p_X(x')|x'\rangle\langle x'|_X \otimes |\psi^{x'}\rangle\langle\psi^{x'}|_B\right]\left[\sum_x |x\rangle\langle x|_X \otimes \left[\log_2\left(p_X(x)\right)\left(I_B - |\psi^x\rangle\langle\psi^x|_B\right) + \log_2 \rho_B\right]^2\right]\right\} \tag{4.80}$$

$$= \sum_x p_X(x) \operatorname{Tr}\left\{|\psi^x\rangle\langle\psi^x|_B\left[\left[\log_2\left(p_X(x)\right)\left(I_B - |\psi^x\rangle\langle\psi^x|_B\right) + \log_2 \rho_B\right]^2\right]\right\} \tag{4.81}$$

$$= \sum_x p_X(x) \operatorname{Tr}\left\{|\psi^x\rangle\langle\psi^x|_B\left[\log_2 \rho_B\right]^2\right\} \tag{4.82}$$

$$= \operatorname{Tr}\{\rho_B\left[\log_2 \rho_B\right]^2\}. \tag{4.83}$$

In the second-to-last equality, we used the expansion in (4.78) and the fact that $|\psi^x\rangle\langle\psi^x|_B$ and $I_B - |\psi^x\rangle\langle\psi^x|_B$ are orthogonal. Finally, putting together (4.73) and (4.83), we conclude (4.68). ■

4.3.2 Example: Pure-loss bosonic channel

We can induce a pure-state cq wiretap channel from a pure-loss bosonic channel. In what follows, we consider a coding scheme called binary phase-shift keying (BPSK). Let us recall just the basic facts needed from Gaussian quantum information to support the argument that follows (a curious reader can consult [Ser17] for further details). The pure-loss channel of transmissivity $\eta \in (0,1)$ is such that if the sender inputs a coherent state $|\alpha\rangle$ with $\alpha \in \mathbb{C}$, then the outputs for Bob and Eve are the coherent states $|\sqrt{\eta}\alpha\rangle_B$ and $|\sqrt{1-\eta}\alpha\rangle_E$, respectively. Note that the overlap of any two coherent states $|\alpha\rangle$ and $|\beta\rangle$ is equal to $|\langle\alpha|\beta\rangle|^2 = e^{-|\alpha-\beta|^2}$, and this is in fact the main quantity that we need to evaluate the information quantities in (4.65). The average photon number of a coherent state $|\alpha\rangle$ is equal to $|\alpha|^2$. A BPSK-coding scheme induces the following pure-state cq wiretap channel from the pure-loss channel:

$$0 \to |\alpha\rangle_A \to |\sqrt{\eta}\alpha\rangle_B \otimes |\sqrt{1-\eta}\alpha\rangle_E, \tag{4.84}$$

$$1 \to |-\alpha\rangle_A \to |-\sqrt{\eta}\alpha\rangle_B \otimes |-\sqrt{1-\eta}\alpha\rangle_E. \tag{4.85}$$

That is, if the sender would like to transmit the symbol "0," then she prepares the coherent state $|\alpha\rangle_A$ at the input, and the physical channel prepares the coherent state $|\sqrt{\eta}\alpha\rangle_B$ for Bob and $|\sqrt{1-\eta}\alpha\rangle_E$ for Eve. A similar explanation holds for when the sender inputs the symbol "1." A BPSK-coding scheme is such that the distribution $p_X(x)$ is unbiased: there is an equal probability 1/2 to pick "0" or "1" when selecting codewords. Thus, the expected density operators at the output for Bob and Eve are respectively as follows:

$$\rho_B = \frac{1}{2}\left(|\sqrt{\eta}\alpha\rangle\langle\sqrt{\eta}\alpha|_B + |-\sqrt{\eta}\alpha\rangle\langle-\sqrt{\eta}\alpha|_B\right), \tag{4.86}$$

$$\rho_E = \frac{1}{2}\left(|\sqrt{1-\eta}\alpha\rangle\langle\sqrt{1-\eta}\alpha|_B + |-\sqrt{1-\eta}\alpha\rangle\langle-\sqrt{1-\eta}\alpha|_B\right). \tag{4.87}$$

A straightforward computation reveals that the eigenvalues for ρ_B are a function only of the overlap $\left|\langle-\sqrt{\eta}\alpha|\sqrt{\eta}\alpha\rangle\right|^2 = e^{-4\eta|\alpha|^2} \equiv e^{-4\eta\bar{n}}$ and are equal to [GW12]

$$p^B(\eta,\bar{n}) \equiv \frac{1}{2}\left(1 + e^{-2\eta\bar{n}}\right), \qquad 1 - p^B(\eta,\bar{n}) = \frac{1}{2}\left(1 - e^{-2\eta\bar{n}}\right). \tag{4.88}$$

Similarly, the eigenvalues of ρ_E are given by

$$p^E(\eta, \bar{n}) \equiv \frac{1}{2}\left(1 + e^{-2(1-\eta)\bar{n}}\right), \qquad 1 - p^E(\eta, \bar{n}) = \frac{1}{2}\left(1 - e^{-2(1-\eta)\bar{n}}\right). \tag{4.89}$$

We can then immediately plug in to (4.65) to find a lower bound on the second-order coding rate for private communication over the pure-loss bosonic channel:

$$\begin{aligned}\log_2 M^*_{\text{priv}}(n, \varepsilon_1 + \sqrt{\varepsilon_2}) \geq n\left[h_2(p^B(\eta, \bar{n})) - h_2(p^E(\eta, \bar{n}))\right] \\ + \sqrt{n v_2(p^B(\eta, \bar{n}))}\Phi^{-1}(\varepsilon_1) + \sqrt{n v_2(p^E(\eta, \bar{n}))}\Phi^{-1}(\varepsilon_2) + O(\log n),\end{aligned} \tag{4.90}$$

where h_2 and v_2 respectively denote the binary entropy and binary entropy variance:

$$h_2(\gamma) \equiv -\gamma \log_2 \gamma - (1-\gamma)\log_2(1-\gamma), \tag{4.91}$$

$$v_2(\gamma) \equiv \gamma\left[\log_2 \gamma + h_2(\gamma)\right]^2 + (1-\gamma)\left[\log_2(1-\gamma) + h_2(\gamma)\right]^2. \tag{4.92}$$

A benchmark against which we can compare the performance of a BPSK code with $|\alpha|^2 = \bar{n}$ is the energy-constrained private capacity of a pure-loss bosonic channel [WQ16], given by

$$g(\eta\bar{n}) - g((1-\eta)\bar{n}), \tag{4.93}$$

where $g(x) \equiv (x+1)\log_2(x+1) - x\log_2 x$. Figure 1 plots the normal approximation [PPV10] of the lower bound on the second-order coding rate of BPSK coding for various parameter choices for ε_1, ε_2, η, and $\bar{n}$, comparing it against the asymptotic performance of BPSK and the actual energy-constrained private capacity in (4.93). The normal approximation consists of all terms in (4.90) besides the $O(\log n)$ term and typically serves as a good approximation for non-asymptotic capacity even for small values of n (when (4.90) is not necessarily valid), as previously observed in [PPV10, TH13, TBR16].

5 Conclusion

This paper establishes a lower bound on the ε-one-shot private classical capacity of a cq wiretap channel, which in turn leads to a lower bound on the second-order coding rate for private communication over an i.i.d. cq wiretap channel. The main techniques used are position-based decoding [AJW17] in order to guarantee that Bob can decode reliably and convex splitting [ADJ17] to guarantee that Eve cannot determine which message Alice transmitted. It is my opinion that these two methods represent a powerful approach to quantum information theory, having already been used effectively in a variety of contexts in [ADJ17, AJW17].

For future work, it would be good to improve upon the lower bounds given here. Extensions of the methods of [YSP16] and [TB16] might be helpful in this endeavor.

Note: After the completion of the results in the present paper, Naqueeb Warsi informed the author of an unpublished result from [War15], which establishes a lower bound on the ε-one-shot private capacity of a cq wiretap channel in terms of a difference of the hypothesis testing mutual information and a smooth max-mutual information.

Acknowledgements. I am grateful to Anurag Anshu, Saikat Guha, Rahul Jain, Haoyu Qi, Qingle Wang, and Naqueeb Warsi for discussions related to the topic of this paper. I acknowledge support from the Office of Naval Research and the National Science Foundation.

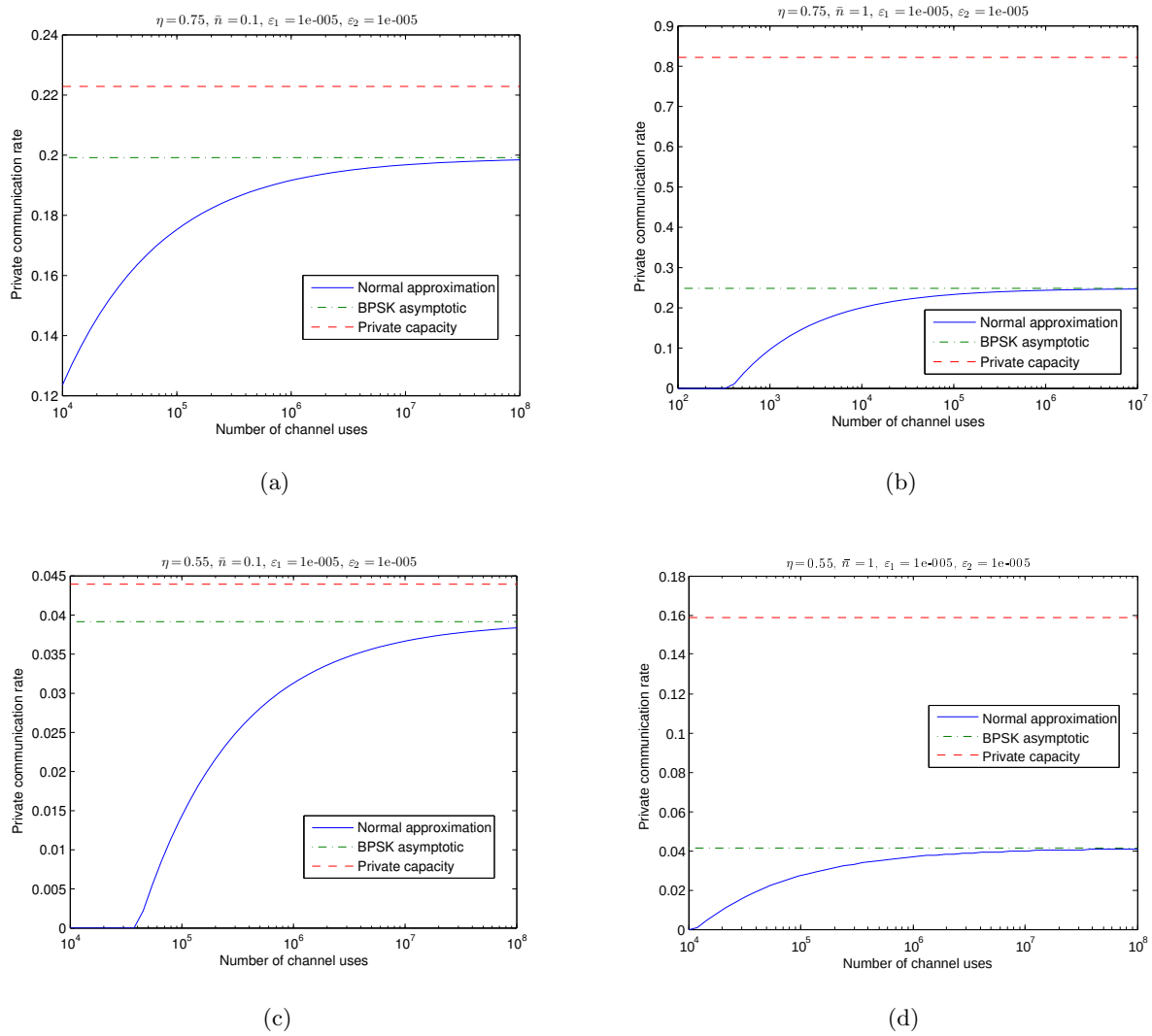

Figure 1: The figures plot the normal approximation for non-asymptotic BPSK private communication (using (4.90)), the asymptotic limit for BPSK, and the asymptotic energy-constrained private capacity for various values of the channel transmissivity η, the mean photon number $\bar{n}$, ε_1, and ε_2.

A Proof of convex-split lemma

For the sake of completeness, this appendix features a proof of Lemma 2. Let $\widetilde{\rho}_{AB}$ be the optimizer for

$$\lambda^* \equiv \widetilde{I}_{\max}^{\sqrt{\varepsilon}-\eta}(B;A)_\rho = \inf_{\rho'_{AB} \,:\, P(\rho'_{AB},\rho_{AB}) \leq \sqrt{\varepsilon}-\eta} D_{\max}(\rho'_{AB} \| \rho_A \otimes \rho'_B). \tag{A.1}$$

We take $\widetilde{\rho}_B$ as the marginal of $\widetilde{\rho}_{AB}$. We define the following state, which we think of as an approximation to $\tau_{A_1 \cdots A_K B}$:

$$\widetilde{\tau}_{A_1 \cdots A_K B} \equiv \frac{1}{K} \sum_{k=1}^{K} \rho_{A_1} \otimes \cdots \otimes \rho_{A_{k-1}} \otimes \widetilde{\rho}_{A_k B} \otimes \rho_{A_{k+1}} \otimes \cdots \otimes \rho_{A_K}. \tag{A.2}$$

In fact, it is a good approximation if $\sqrt{\varepsilon} - \eta$ is small: Consider from joint concavity of the root fidelity that

$$\begin{aligned}
&\sqrt{F}(\widetilde{\tau}_{A_1 \cdots A_K B}, \tau_{A_1 \cdots A_K B}) \\
&\geq \frac{1}{K} \sum_{k=1}^{K} \sqrt{F}(\rho_{A_1} \otimes \cdots \otimes \rho_{A_{k-1}} \otimes \widetilde{\rho}_{A_k B} \otimes \rho_{A_{k+1}} \otimes \cdots \otimes \rho_{A_K}, \\
&\qquad\qquad \rho_{A_1} \otimes \cdots \otimes \rho_{A_{k-1}} \otimes \rho_{A_k B} \otimes \rho_{A_{k+1}} \otimes \cdots \otimes \rho_{A_K}) \qquad \text{(A.3)} \\
&= \frac{1}{K} \sum_{k=1}^{K} \sqrt{F}(\widetilde{\rho}_{A_k B}, \rho_{A_k B}) \qquad \text{(A.4)} \\
&= \sqrt{F}(\widetilde{\rho}_{AB}, \rho_{AB}), \qquad \text{(A.5)}
\end{aligned}$$

which in turn implies that

$$F(\widetilde{\tau}_{A_1 \cdots A_K B}, \tau_{A_1 \cdots A_K B}) \geq F(\widetilde{\rho}_{AB}, \rho_{AB}). \tag{A.6}$$

So the inequality in (A.6), the definition of the purified distance, and the fact that $P(\rho_{AB}, \widetilde{\rho}_{AB}) \leq \sqrt{\varepsilon} - \eta$ imply that

$$P(\widetilde{\tau}_{A_1 \cdots A_K B}, \tau_{A_1 \cdots A_K B}) \leq \sqrt{\varepsilon} - \eta. \tag{A.7}$$

Let $\omega = \sum_y p_Y(y) \omega^y$, for p_Y a probability distribution and $\{\omega^y\}_y$ a set of states. Then the following property holds for quantum relative entropy and a state κ such that $\operatorname{supp}(\omega) \subseteq \operatorname{supp}(\kappa)$:

$$D(\omega \| \kappa) = \sum_y p_Y(y) \left[D(\omega^y \| \kappa) - D(\omega^y \| \omega) \right]. \tag{A.8}$$

Applying (A.8), it follows that

$$\begin{aligned}
&D(\widetilde{\tau}_{A_1 \cdots A_K B} \| \rho_{A_1} \otimes \cdots \otimes \rho_{A_K} \otimes \widetilde{\rho}_B) \\
&\quad = \frac{1}{K} \sum_{k=1}^{K} D(\rho_{A_1} \otimes \cdots \otimes \rho_{A_{k-1}} \otimes \rho_{A_k B} \otimes \rho_{A_{k+1}} \otimes \cdots \otimes \rho_{A_K} \| \rho_{A_1} \otimes \cdots \otimes \rho_{A_K} \otimes \widetilde{\rho}_B) \\
&\qquad\qquad - \frac{1}{K} \sum_{k=1}^{K} D(\rho_{A_1} \otimes \cdots \otimes \rho_{A_{k-1}} \otimes \rho_{A_k B} \otimes \rho_{A_{k+1}} \otimes \cdots \otimes \rho_{A_K} \| \widetilde{\tau}_{A_1 \cdots A_K B}). \qquad \text{(A.9)}
\end{aligned}$$

The first term in (A.9) on the right-hand side of the equality simplifies as

$$\begin{aligned}
&\frac{1}{K}\sum_{k=1}^{K} D(\rho_{A_1}\otimes\cdots\otimes\rho_{A_{k-1}}\otimes\rho_{A_kB}\otimes\rho_{A_{k+1}}\otimes\cdots\otimes\rho_{A_K}\|\rho_{A_1}\otimes\cdots\otimes\rho_{A_K}\otimes\widetilde{\rho}_B)\\
&=\frac{1}{K}\sum_{k=1}^{K} D(\rho_{A_kB}\|\rho_{A_k}\otimes\widetilde{\rho}_B) &\text{(A.10)}\\
&=D(\rho_{AB}\|\rho_A\otimes\widetilde{\rho}_B). &\text{(A.11)}
\end{aligned}$$

We now lower bound the last term in (A.9). Consider that a partial trace over systems A_1, …, A_{k-1}, A_{k+1}, …, A_K gives

$$\begin{aligned}
&D(\rho_{A_1}\otimes\cdots\otimes\rho_{A_{k-1}}\otimes\rho_{A_kB}\otimes\rho_{A_{k+1}}\otimes\cdots\otimes\rho_{A_K}\|\widetilde{\tau}_{A_1\cdots A_KB})\\
&\qquad\geq D(\rho_{A_kB}\|\widetilde{\tau}_{A_kB}) = D(\rho_{AB}\|\,[1/K]\,\widetilde{\rho}_{AB}+[1-1/K]\,\rho_A\otimes\widetilde{\rho}_B), \quad\text{(A.12)}
\end{aligned}$$

where the equality follows because $\widetilde{\tau}_{A_kB}=[1/K]\,\widetilde{\rho}_{AB}+[1-1/K]\,\rho_A\otimes\widetilde{\rho}_B$. Thus, averaging the inequality in (A.12) over k implies that

$$\begin{aligned}
&\frac{1}{K}\sum_{k=1}^{K} D(\rho_{A_1}\otimes\cdots\otimes\rho_{A_{k-1}}\otimes\rho_{A_kB}\otimes\rho_{A_{k+1}}\otimes\cdots\otimes\rho_{A_K}\|\widetilde{\tau}_{A_1\cdots A_KB})\\
&\qquad\geq D(\rho_{AB}\|\,[1/K]\,\widetilde{\rho}_{AB}+[1-1/K]\,\rho_A\otimes\widetilde{\rho}_B), \quad\text{(A.13)}
\end{aligned}$$

Putting together (A.9), (A.11), and (A.13), we find that

$$\begin{aligned}
&D(\widetilde{\tau}_{A_1\cdots A_KB}\|\rho_{A_1}\otimes\cdots\otimes\rho_{A_K}\otimes\rho_B)\\
&\qquad\leq D(\rho_{AB}\|\rho_A\otimes\widetilde{\rho}_B) - D(\rho_{AB}\|\,[1/K]\,\widetilde{\rho}_{AB}+[1-1/K]\,\rho_A\otimes\widetilde{\rho}_B) \quad\text{(A.14)}
\end{aligned}$$

By the definition of λ^* in (A.1), we have that

$$\widetilde{\rho}_{AB}\leq 2^{\lambda^*}\rho_A\otimes\widetilde{\rho}_B, \tag{A.15}$$

which means that

$$[1/K]\,\widetilde{\rho}_{AB}+[1-1/K]\,\rho_A\otimes\widetilde{\rho}_B\leq\left[1+(2^{\lambda^*}-1)/K\right]\rho_A\otimes\widetilde{\rho}_B. \tag{A.16}$$

An important property of quantum relative entropy is that $D(\omega\|\tau)\geq D(\omega\|\tau')$ if $\tau\leq\tau'$. Applying it to (A.16) and the right-hand side of (A.14), we get that

$$\begin{aligned}
&D(\rho_{AB}\|\rho_A\otimes\widetilde{\rho}_B) - D(\rho_{AB}\|\,[1/K]\,\widetilde{\rho}_{AB}+[1-1/K]\,\rho_A\otimes\widetilde{\rho}_B)\\
&\leq D(\rho_{AB}\|\rho_A\otimes\widetilde{\rho}_B) - D(\rho_{AB}\|\left[1+(2^{\lambda^*}-1)/K\right]\rho_A\otimes\widetilde{\rho}_B) &\text{(A.17)}\\
&= D(\rho_{AB}\|\rho_A\otimes\widetilde{\rho}_B) - D(\rho_{AB}\|\rho_A\otimes\widetilde{\rho}_B) + \log_2\left[1+(2^{\lambda^*}-1)/K\right] &\text{(A.18)}\\
&= \log_2\left[1+(2^{\lambda^*}-1)/K\right]. &\text{(A.19)}
\end{aligned}$$

Then the well known inequality $D(\omega\|\tau) \geq -\log_2 F(\omega,\tau)$, (A.14), and (A.17)–(A.19) imply that

$$-\log_2 F(\widetilde{\tau}_{A_1\cdots A_K B}, \rho_{A_1}\otimes\cdots\otimes\rho_{A_K}\otimes\widetilde{\rho}_B) \leq \log_2\left[1+(2^{\lambda^*}-1)/K\right], \tag{A.20}$$

which in turn implies that

$$F(\widetilde{\tau}_{A_1\cdots A_K B}, \rho_{A_1}\otimes\cdots\otimes\rho_{A_K}\otimes\widetilde{\rho}_B) \geq \frac{1}{1+\frac{2^{\lambda^*}-1}{K}} = 1-\frac{2^{\lambda^*}-1}{K+2^{\lambda^*}-1} \geq 1-\frac{2^{\lambda^*}}{K}. \tag{A.21}$$

So if we pick K such that

$$\log_2 K = \inf_{\rho'_{AB}\,:\,P(\rho'_{AB},\rho_{AB})\leq\sqrt{\varepsilon}-\eta} D_{\max}(\rho'_{AB}\|\rho_A\otimes\rho'_B) + 2\log_2\left(\frac{1}{\eta}\right), \tag{A.22}$$

then we are guaranteed that

$$P(\widetilde{\tau}_{A_1\cdots A_K B}, \rho_{A_1}\otimes\cdots\otimes\rho_{A_K}\otimes\widetilde{\rho}_B) \leq \eta. \tag{A.23}$$

By the triangle inequality for the purified distance, we then get that

$$\begin{aligned} &P(\tau_{A_1\cdots A_K B}, \rho_{A_1}\otimes\cdots\otimes\rho_{A_K}\otimes\widetilde{\rho}_B) \\ &\leq P(\tau_{A_1\cdots A_K B}, \widetilde{\tau}_{A_1\cdots A_K B}) + P(\widetilde{\tau}_{A_1\cdots A_K B}, \rho_{A_1}\otimes\cdots\otimes\rho_{A_K}\otimes\widetilde{\rho}_B) &\text{(A.24)}\\ &\leq \left(\sqrt{\varepsilon}-\eta\right)+\eta = \sqrt{\varepsilon}. &\text{(A.25)}\end{aligned}$$

This concludes the proof.

References

[ADJ17] Anurag Anshu, Vamsi Krishna Devabathini, and Rahul Jain. Quantum message compression with applications. February 2017. arXiv:1410.3031v4.

[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.

[BB84] Charles H. Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of IEEE International Conference on Computers Systems and Signal Processing*, pages 175–179, Bangalore, India, December 1984.

[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.

[BDL16] Salman Beigi, Nilanjana Datta, and Felix Leditzky. Decoding quantum information via the Petz recovery map. *Journal of Mathematical Physics*, 57(8):082203, August 2016. arXiv:1504.04449.

[CK78] Imre Csiszár and Janos Körner. Broadcast channels with confidential messages. *IEEE Transactions on Information Theory*, 24(3):339–348, May 1978.

[CWY04] Ning 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.

[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.

[DHO16] Nilanjana Datta, Min-Hsiu Hsieh, and Jonathan Oppenheim. An upper bound on the second order asymptotic expansion for the quantum communication cost of state redistribution. *Journal of Mathematical Physics*, 57(5):052203, May 2016. arXiv:1409.4352.

[DL15] Nilanjana Datta and Felix Leditzky. Second-order asymptotics for source coding, dense coding, and pure-state entanglement conversions. *IEEE Transactions on Information Theory*, 61(1):582–608, January 2015. arXiv:1403.2543.

[DTW16] 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.

[GGL+04] 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.

[GLN05] Alexei Gilchrist, Nathan K. Langford, and Michael A. Nielsen. Distance measures to compare real and ideal quantum processes. *Physical Review A*, 71(6):062310, June 2005. arXiv:quant-ph/0408063.

[GW12] Saikat Guha and Mark M. Wilde. Polar coding to achieve the Holevo capacity of a pure-loss optical channel. *Proceedings of the 2012 IEEE International Symposium on Information Theory*, pages 546–550, 2012. arXiv:1202.0533.

[Hay06] Masahito Hayashi. General nonasymptotic and asymptotic formulas in channel resolvability and identification capacity and their application to the wiretap channel. *IEEE Transactions on Information Theory*, 52(4):1562–1575, April 2006. arXiv:cs/0503088.

[Hay13] Masahito Hayashi. Tight exponential analysis of universally composable privacy amplification and its applications. *IEEE Transactions on Information Theory*, 59(11):7728–7746, November 2013. arXiv:1010.1358.

[HHHO09] Karol Horodecki, Michal Horodecki, Pawel Horodecki, and Jonathan Oppenheim. General paradigm for distilling classical key from quantum states. *IEEE Transactions on Information Theory*, 55(4):1898–1929, April 2009. arXiv:quant-ph/0506189.

[HN03] Masahito Hayashi and Hiroshi Nagaoka. General formulas for capacity of classical-quantum channels. *IEEE Transactions on Information Theory*, 49(7):1753–1768, July 2003. arXiv:quant-ph/0206186.

[Li14] Ke Li. Second order asymptotics for quantum hypothesis testing. *Annals of Statistics*, 42(1):171–189, February 2014. arXiv:1208.1400.

[PPV10] Yury Polyanskiy, H. Vincent Poor, and Sergio Verdú. Channel coding rate in the finite blocklength regime. *IEEE Transactions on Information Theory*, 56(5):2307–2359, May 2010.

[Ras02] Alexey E. Rastegin. Relative error of state-dependent cloning. *Physical Review A*, 66(4):042304, October 2002.

[Ras03] Alexey E. Rastegin. A lower bound on the relative error of mixed-state cloning and related operations. *Journal of Optics B: Quantum and Semiclassical Optics*, 5(6):S647, December 2003. arXiv:quant-ph/0208159.

[Ras06] Alexey E. Rastegin. Sine distance for quantum states. February 2006. arXiv:quant-ph/0602112.

[RR11] Joseph M. Renes and Renato Renner. Noisy channel coding via privacy amplification and information reconciliation. *IEEE Transactions on Information Theory*, 57(11):7377–7385, Nov 2011.

[Ser17] Alessio Serafini. *Quantum Continuous Variables*. CRC Press, 2017.

[Tan12] Vincent Y. F. Tan. Achievable second-order coding rates for the wiretap channel. In *2012 IEEE International Conference on Communication Systems (ICCS)*, pages 65–69, November 2012.

[TB16] Mehrdad Tahmasbi and Matthieu R. Bloch. Second order asymptotics for degraded wiretap channels: How good are existing codes? In *2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 830–837, September 2016.

[TBR16] Marco Tomamichel, Mario Berta, and Joseph M. Renes. Quantum coding with finite resources. *Nature Communications*, 7:11419, May 2016. arXiv:1504.04617.

[TCR09] Marco Tomamichel, Roger Colbeck, and Renato Renner. A fully quantum asymptotic equipartition property. *IEEE Transactions on Information Theory*, 55(12):5840–5847, December 2009. arXiv:0811.1221.

[TH13] Marco Tomamichel and Masahito Hayashi. A hierarchy of information quantities for finite block length analysis of quantum tasks. *IEEE Transactions on Information Theory*, 59(11):7693–7710, November 2013. arXiv:1208.1478.

[TT15] Marco Tomamichel and Vincent Y. F. Tan. Second-order asymptotics for the classical capacity of image-additive quantum channels. *Communications in Mathematical Physics*, 338(1):103–137, August 2015. arXiv:1308.6503.

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

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

[War15] Naqueeb Ahmad Warsi. *One-shot bounds in classical and quantum information theory.* PhD thesis, Tata Institute of Fundamental Research, Mumbai, India, December 2015. Not publicly available; communicated by email on March 5, 2017.

[Wil16] Mark M. Wilde. *From Classical to Quantum Shannon Theory.* March 2016. arXiv:1106.1445v7.

[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.

[WTB17] Mark M. Wilde, Marco Tomamichel, and Mario Berta. Converse bounds for private communication over quantum channels. *IEEE Transactions on Information Theory*, 63(3):1792–1817, March 2017. arXiv:1602.08898.

[Wyn75] Aaron D. Wyner. The wire-tap channel. *Bell System Technical Journal*, 54(8):1355–1387, October 1975.

[YAG13] Mohammad Hossein Yassaee, Mohammad Reza Aref, and Amin Gohari. Non-asymptotic output statistics of random binning and its applications. In *2013 IEEE International Symposium on Information Theory*, pages 1849–1853, July 2013. arXiv:1303.0695.

[YSP16] Wei Yang, Rafael F. Schaefer, and H. Vincent Poor. Finite-blocklength bounds for wire-tap channels. In *2016 IEEE International Symposium on Information Theory (ISIT)*, pages 3087–3091, July 2016. arXiv:1601.06055.