One-Shot Capacity Bounds on the Simultaneous Transmission of Classical and Quantum Information

Farzin Salek®, Student Member, IEEE, Anurag Anshu®, Min-Hsiu Hsieh®, Senior Member, IEEE, Rahul Jain®, and Javier Rodríguez Fonollosa®, Senior Member, IEEE

Abstract—We study the communication capabilities of a quantum channel under the most general channel model known as the one-shot model. Unlike classical channels that can only be used to transmit classical information (bits), a quantum channel can be used for transmission of classical information, quantum information (qubits) and simultaneous transmission of classical and quantum information. In this work, we investigate the one-shot capabilities of a quantum channel for simultaneously transmitting bits and qubits. This problem was studied in the asymptotic regime for a memoryless channel where a regularized characterization of the capacity region was reported. It is known that the transmission of private classical information is closely related to the problem of quantum information transmission. We resort to this idea and find achievable and converse bounds on the simultaneous transmission of the public and private classical information. Then shifting the classical private rate to the quantum information rate leads to a rate region for simultaneous transmission of classical and quantum information. In the case of asymptotic i.i.d. setting, our one-shot result is evaluated to the known results in the literature. Our main tools used in the achievability proofs are position-based decoding and convex-split lemma.

Index Terms—One-shot coding, channel coding, private capacity, quantum capacity.

I. INTRODUCTION

SHANNON modeled a noisy (classical) channel as a stochastic map \( W_{X \rightarrow Y} \) taking classical inputs to classical outputs according to some probability distribution, \( p_{Y|X}(y|x) \) [23]. In his paper, he defined and computed the fundamental feature of a channel, its capacity: the amount of classical information, i.e., bits, that can be reliably transmitted from a sender to a remote receiver over a classical channel. In the limit of many independent uses of a stationary memoryless channel, Shannon showed that its capacity in bits per use of the channel is equal to the mutual information between the input and output.

Since nature is fundamentally quantum, it seemed necessary to enhance or replace Shannon’s channel model with a quantum channel model that takes quantum mechanics into account. Many years after Shannon in the context of quantum information theory, a quantum channel is modelled by a completely-positive trace-preserving map (CPTP) with possibly different input and output Hilbert spaces. Denoted by \( N_{A \rightarrow B} \), a quantum channel with input and output systems \( A \) and \( B \) respectively, can now be used to accomplish a variety of information-processing tasks and accordingly different capacities can be defined. In the next two subsections, we review some concepts in the asymptotic and one-shot regimes.

A. Memoryless and Stationary Channels, Asymptotic Regime

Perhaps the most direct analogue of the capacity of a classical channel, \( C(V) \), is the classical capacity of a quantum channel, \( C(N) \), i.e., the highest rate (in bits per use of the channel) at which a sender can transmit classical information faithfully to a remote receiver through a quantum channel with general quantum inputs and quantum outputs. The classical capacity1 was independently studied in [24] and [25] where

1hereafter, we talk about quantum channels unless otherwise specified, hence we drop the term quantum.
an achievability bound, i.e., $C(N) \geq \chi(N)$, known as HSW theorem was reported, where $\chi(N)$ is the celebrated Holevo Information [26] defined as follows:

$$\chi(N) := \max_{I(X;B)_\rho} \{ I(X;B)_\rho - I(X;E)_\rho \},$$

where $p(x)$ is a probability distribution, $\rho_{XRB} = \sum_x p(x) \rho_{X|x} \otimes N_{A\rightarrow B}(\rho_A^x)$ is a bipartite quantum state and $I(X;B)_\rho$ is the quantum mutual information (see Definition 6). The classical capacity equals the regularized Holevo information, taking a limit over many copies of the channel. So unlike the classical channel, we don’t fully know the capabilities of a quantum channel for transmitting classical information.

In certain scenarios, a sender may wish to communicate classical information to a receiver by means of a quantum channel such that the information must remain secret from some third party surrounding the legitimate receiver. This information-processing task gives rise to the notion of private capacity of a quantum channel. Cai-Winter-Yeung [28] and Devetak [27] showed that the achievable rates for classical private capacity can be formulated as the difference between the Holevo information of the sender and the legitimate receiver and that of the sender and the eavesdropper(s) as given below:

$$\mathcal{P}(N) := \max_\rho \{ I(X;B)_\rho - I(X;E)_\rho \},$$

where $\rho_{XBE} = \sum_x p(x) \rho_{X|x} \otimes U_{A\rightarrow BE}(\rho_A^x)$ and $U_{A\rightarrow BE}$ is an isometric extension of the channel $N_{A\rightarrow B}$. They also showed that the private capacity equals the regularized form of $\mathcal{P}(N)$ meaning that this ability of the quantum channel is still not fully understood.

The capacity of a quantum channel to transmit quantum information is called the quantum capacity of the channel and we represent it by $Q_{reg}(N)$. For a given quantum channel, one would like to understand the best rates (in terms of qubits per use of the channel) at which quantum information can be transmitted over the channel. The quantum capacity theorem was first considered in [29] and later in [30]. Subsequently, by taking advantage of the properties of the private classical codes, Devetak [27] showed that the quantum capacity is given by the regularized coherent information of the channel:

$$Q_{reg}(N) := \lim_{k \to \infty} \frac{1}{k} Q(N^{\otimes k})$$

where the coherent information is defined as $Q(N) := \max_{\phi_{RA}} I(R;B)_\sigma$ (see Definition 7) and the optimization is with respect to all pure, bipartite states $\phi_{RA}$ and $\sigma_{RB} = N_{A\rightarrow B}(\phi_{RA})$.

Devetak and Shor [19] unified the classical and quantum capacities and introduced a new information-processing task studying the simultaneously achievable rates for transmission of classical and quantum information over a quantum channel. Since we will follow the results of [19] closely in this paper, we mention its main theorem:

**Theorem 1 ([19]):** The capacity region of $N$ for simultaneous transmission of classical and quantum information is as follows:

$$S_{reg}(N) := \lim_{k \to \infty} \frac{1}{k} S(N^{\otimes k}),$$

where $S(N)$ is the union, over all states $\rho_{XRB} = \sum_x p(x) \rho_{X|x} \otimes \rho_{RB}^{x}$ arising from the channel $N_{A\rightarrow B}$, i.e., for $x \in \text{supp}(p(x))$, $\rho_{RB}^{x} = \mathcal{N}(\phi_{RA}^x)$ for pure states $\{|\phi_{RA}^x\rangle\}$, of the $(r, R)$ pairs obeying

$$0 \leq r \leq I(X;B)_\rho,$$

$$0 \leq R \leq I(R;BX)_\rho,$$

where $r$ and $R$ are the rates of the classical and quantum information, respectively.

The result of Devetak and Shor is generalized in [20] such that the rate of a secret key that used to achieve noiseless private capacity, enters the tradeoff. It is known that the interplay between public classical communication, private classical communication and secret key is rather analogous to how classical communication, quantum communication and entanglement interact with one another. This interaction was studied in [21] from an information-theoretic perspective and the corresponding rate regions for several realistic channels were computed.

**B. General Channels, One-Shot Regime**

All the aforementioned capacities are originally evaluated under the assumptions that the channels are memoryless and stationary and they are available to be used arbitrarily many times. However, in many real-world scenarios, we encounter channels which are neither stationary nor memoryless. Therefore, it is of fundamental importance to think of coding schemes for the channels which fail to satisfy these assumptions. The independent channel uses are relaxed in [31] and [32] and general channels with memory are studied in [48] and [49], albeit these results are derived in the form of a limit such that the error probability vanishes as the number of channel uses goes to infinity. Later researchers considered single-serving scenarios where a given channel is used only once. This approach gives rise to a high level of generality that no assumptions are made on the structure of the channel and the associated capacity is usually referred to as one-shot capacity.

The one-shot capacity of a classical channel was characterized in terms of min- and max-entropies in [33]. The one-shot classical capacity of a quantum channel is addressed by a hypothesis testing approach in [34] and [1], yielding expressions in terms of the generalized (Rényi) relative entropies and a smooth relative entropy quantity, respectively. By taking advantage of two primitive information-theoretic protocols, privacy amplification and information reconciliation, authors of [35] constructed coding schemes for one-shot transmission of public and private classical information. Their results come in terms of the min- and max-entropies. Two new tools namely position-based decoding [2] and convex-split lemma [41], are employed in [3] where one-shot achievability bounds on the public and private transmission rates are reported (note that

\(^2\)It is the same for various information-processing tasks: subspace transmission, entanglement transmission or entanglement generation.
prior to this work, one-shot bounds on the public transmission rates on both assisted and unassisted cases were reported in [2] and [1], respectively. Recently, [36] reported tight upper and lower bounds for the one-shot capacity of the wiretap channel. This was done by proving a one-shot version of the quantum covering lemma (see [42]) along with an operator Chernoff bound for non-square matrices. Inner and outer bounds on the one-shot quantum capacity of an arbitrary channel are studied in [15]. The general scenario of [15] leads to the evaluation of the quantum capacity of a channel with arbitrary correlated noise in the repeated uses of the channel.

This work is the continuation of the authors’ former work on the simultaneous transmission of private and public information [16]. In this work, we aim to study the problem of simultaneous transmission of classical and quantum information over a single use of a quantum channel. In other words, we are interested in the one-shot tradeoff between the number of bits and qubits that are simultaneously achievable. The root of our approach is the well-known quantum capacity theorem via private classical communication [27]. The basic intuition underlying the quantum capacity is the no-cloning theorem which states that it is impossible to create an identical copy of an arbitrary unknown quantum state. We know well that associated to every quantum channel there is an environment (Eve). If Eve can learn anything about the quantum information that Alice is trying to send to Bob, Bob will not be able to retrieve this information, otherwise the no-cloning theorem would be violated. Hence, to transmit quantum information, Alice needs to store her quantum information in such subspaces of her input space that Eve does not have access to. By using this idea, Devetak [27] proves that a code for private classical communication can be readily translated into a code for quantum communication. Note that Devetak’s proof shows the aforementioned translation in the asymptotic regime, however, one can easily check that the same holds true in the one-shot regime and the proof follows along the same lines. We provide a proof sketch in appendix B. Therefore, if we can come up with a protocol for simultaneously transmitting public and private classical information, we are able to adapt it for the simultaneous transmission of classical and quantum information.

C. Techniques and Tools

Main tools in our achievability bounds are position-based decoding and convex-split lemma. Our technique is a simple application of superposition coding in classical information theory (not to be confused with the concept of superposition in the quantum mechanics), along with convex-split lemma and position-based decoding. In this manner, we significantly differ from the technique of Devetak and Shor [19], whose method was inherently asymptotic i.i.d. and could not have been adapted in the one-shot setting.

We briefly review position-based decoding and the convex-split lemma. Assume Alice and Bob have a way of creating the following state shared between them (in other words, they have this resource at their disposal before any communication takes place):

$$\rho_{XA}^{\otimes |M|} = \rho_{XA}^1 \otimes \ldots \otimes \rho_{XA}^m \otimes \ldots \otimes \rho_{XA}^{|M|},$$

where Alice possesses $A$ systems and Bob has $X$ systems. Here, the positions of states is denoted by superscripts. Alice wishes to transmit the $m$-th copy of the state above through the channel $\mathcal{N}_{A \rightarrow B}$ to Bob. This induces the following state on Bob’s side:

$$\rho_{X|M|B}^m = \rho_X^1 \otimes \ldots \otimes \rho_X^m \otimes \ldots \otimes \rho_X^{|M|}.$$

If Bob has a means by which he can distinguish between the induced states for different values of $m$ (hypotheses), which happens to be reduced to the problem of distinguishing between states $\rho_{XB}$ and $\rho_{X} \otimes \rho_B$, he is able to learn about the transmitted message $m$. Position-based decoding in fact, relates the communication problem to a problem in binary hypothesis testing. On the other hand, once Alice chooses the $m$-th system uniformly and sends it over the channel, the induced state on receiver side can generally be considered as:

$$\frac{1}{|M|} \sum_{m=1}^{|M|} \rho_X^1 \otimes \ldots \otimes \rho_X^m \otimes \ldots \otimes \rho_X^{|M|} \otimes \rho_B,$$

convex-split lemma argues that if the number of systems, $|M|$, is almost equal to a quantity known as max-mutual information, the induced state is close to the following state:

$$\rho_{X}^1 \otimes \ldots \otimes \rho_X^m \otimes \ldots \otimes \rho_X^{|M|} \otimes \rho_B,$$

meaning that the receiver will not be able to distinguish between the induced states and the product state above, resulting in its ignorance about the chosen message $m$.

A subset of the current authors have generalized the problem studied in this paper and considered the simultaneous transmission of private, public and confidential messages, see [51] and [52]. In that work, the proofs are based on conditional versions of the convex-split lemma and position based decoding.

The rest of the paper is organized as follows. In Section II, we give preliminaries and definitions. A code for simultaneous transmission of public and private information is formally discussed in Section III. This section also includes our main results. Section IV is devoted to the description of the protocol as well as our achievability proof. Converse bounds are proven in section V. In Section VI, we argue how the well-known asymptotic bounds can be quickly recovered by many independent uses of a memoryless channel. We conclude the paper by a discussion in Section VII.

II. Preliminaries

We denote (quantum) systems by capital letters, and we use subscripts to denote the systems on which the mathematical objects are defined (we may drop the subscript if it does not lead to ambiguity). The Hilbert space corresponding to a quantum system $A$ is denoted by $\mathcal{H}_A$ and its dimension is shown by $|\mathcal{H}_A|$. Conventionally, a random variable $X$ taking on its values from some finite alphabet $\mathcal{X}$ with cardinality $|\mathcal{X}|$ can be associated with a (classical) system (which we also referred to as $X$) whose Hilbert space has orthonormal basis labeled by $x$, i.e., $\{|x\rangle\}_{x \in \mathcal{X}}$ and dimension $|\mathcal{H}_X| = |\mathcal{X}|$. This notation is adopted throughout the paper. The set of linear operators on $\mathcal{H}_A$ is denoted by $\mathcal{L}(\mathcal{H}_A)$ and the set of
between the density operator $\rho$, a positive-semidefinite operator, i.e., $\rho_A \in \mathcal{P}(\mathcal{H}_A)$, with trace equal to one. We denote the set of quantum states in $\mathcal{H}_A$ by $\mathcal{D}(\mathcal{H}_A)$. The identity operator acting on $\mathcal{H}_A$ is shown by $\mathbb{1}$. The trace norm of the linear operator $\rho_A \in \mathcal{L}(\mathcal{H}_A)$ is defined as $\|\rho_A\|_1 = \text{Tr} \sqrt{\rho_A^\dagger \rho_A}$ where $\rho_A^\dagger$ is the conjugate transpose of $\rho_A$. The support of an operator $\rho$, supp($\rho$), is defined to be the subspace orthogonal to its kernel. If the support of $\rho$ is contained in that of $\sigma$, we write $\rho \subseteq \sigma$. Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be Hilbert spaces associated to systems $A$ and $B$, respectively. We can consider the composite system of $A$ and $B$ as a single system with Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. For a bipartite state $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, marginal systems are defined via partial trace as $\text{Tr}_B(\rho_{AB}) = \rho_A$ and $\text{Tr}_A(\rho_{AB}) = \rho_B$. For a pair of integers $i \leq j$, we define the discrete interval $[i : j] := \{i, i+1, ..., j\}$. For Hermitian operators $M$ and $N$, $M \leq N$ means that $(N - M) \in \mathcal{P}(\mathcal{H})$.

Let us consider a binary hypothesis test discriminating between the density operator $\rho_A$ (null hypothesis) and $\sigma_A$ (alternative hypothesis) where $\rho_A, \sigma_A \in \mathcal{D}(\mathcal{H}_A)$. The task is to distinguish between the two hypotheses by means of some quantum measurement $(T_A, \mathbb{1} - T_A)$ such that $0 \leq T_A \leq \mathbb{1}$. The test decides in favor of $\rho_A$ (resp. $\sigma_A$) when the outcome corresponding to $T_A$ (resp. $\mathbb{1} - T_A$) occurs. Two kinds of errors can be defined here: Type I error occurs when the true hypothesis was $\rho_A$ but $\sigma_A$ is decided and Type II error is the opposite kind of error. The error probabilities corresponding to Type I and Type II errors are respectively as follows:

$$
\alpha(T_A, \rho_A) := \text{Tr}(\mathbb{1} - T_A)\rho_A), \quad \beta(T_A, \sigma_A) := \text{Tr}(T_A\sigma_A).
$$

In the setting of asymmetric hypothesis testing, the aim is to minimize $\beta(T_A, \sigma_A)$ under a constraint on $\alpha(T_A, \rho_A)$. This task gives rise to the definition of the hypothesis testing relative entropy defined as follows:

**Definition 1 (Hypothesis testing relative entropy [1], [15]):**

$$
D_H^\beta(\rho_A || \sigma_A) := -\inf_{\alpha(T_A, \rho_A) \leq \epsilon} \beta(T_A, \sigma_A).
$$

In quantum information theory, one often needs to measure the distance between two quantum states. Let us again consider the task of distinguishing between two quantum states $\rho_A$ and $\sigma_A$ by means of a binary test operator $0 \leq T_A \leq \mathbb{1}$. Intuitively, the closer the states are, the harder they can be distinguished. We further assume that $\rho_A$ and $\sigma_A$ are prepared with equal probabilities. It can be easily shown that the optimal success probability in distinguishing the states equals

$$
\frac{1}{2}(1 + \text{max}_{0 \leq T_A \leq \mathbb{1}} \text{Tr}(T_A(\rho_A - \sigma_A))).
$$

The optimization problem is evaluated as follows:

$$
\max_{0 \leq T_A \leq \mathbb{1}} \text{Tr}(T_A(\rho_A - \sigma_A)) = \{[\rho_A - \sigma_A]_+, (\rho_A - \sigma_A)\} - \{[\rho_A - \sigma_A]_-(\rho_A - \sigma_A)\} \\
\geq \frac{1}{2}\|\rho_A - \sigma_A\|_1.
$$

where $\{\rho_A - \sigma_A\}_+$ denotes the projector onto the subspace where the operator $(\rho_A - \sigma_A)$ is non-negative, and $\{\rho_A - \sigma_A\}_-$ denotes the projector onto the subspace where $(\rho_A - \sigma_A)$ is non-positive. This operational interpretation leads to a distance measure called trace distance defined below.

**Definition 2 (Trace Distance [12]):** The trace distance between two quantum states $\rho_A, \sigma_A$ is given by:

$$
D(\rho_A, \sigma_A) := \frac{1}{2}\|\rho_A - \sigma_A\|_1.
$$

We frequently use the following properties of the trace distance:

- Trace distance is convex. For two ensembles $\{\rho(x), \rho_A^x\}$ and $\{\sigma(x), \sigma_A^x\}$, where $\rho_A^x, \sigma_A^x \in \mathcal{D}(\mathcal{H}_A)$ for all $x$, let $\rho_{XA} := \sum_x p(x)\rho(x)\otimes \rho_A^x$ and $\sigma_{XA} := \sum_x p(x)\sigma(x)\otimes \sigma_A^x$ be the associated classical-quantum (CQ) states, respectively. Then,

$$
\|\rho_{XA} - \sigma_{XA}\|_1 = \sum_x p(x)\|\rho_A^x - \sigma_A^x\|_1.
$$

Moreover, the following property can be easily checked:

$$
\|\rho_{XA} - \sigma_{XA}\|_1 \leq \|\rho - \sigma\|_1.
$$

- Trace distance is monotone non-increasing with respect to CPTP maps. That is, for quantum states $\rho$ and $\sigma$ and the map $\mathcal{N}$, the following inequality holds:

$$
\|\mathcal{N}(\rho) - \mathcal{N}(\sigma)\|_1 \leq \|\rho - \sigma\|_1.
$$

- Trace distance is invariant with respect to tensor-product states, meaning that for quantum states $\rho, \sigma$ and $\tau$, we have that:

$$
\|\rho \otimes \tau - \sigma \otimes \tau\|_1 = \|\rho - \sigma\|_1.
$$

- Trace distance fulfills the triangle inequality; That is, for any three quantum states $\rho, \sigma$ and $\tau$, the following inequality holds:

$$
\|\rho - \sigma\|_1 \leq \|\rho - \tau\|_1 + \|\tau - \sigma\|_1.
$$

**Definition 3 (Fidelity [45], [12]):** The fidelity between two states $\rho, \sigma \in \mathcal{D}(\mathcal{H}_A)$ is defined as:

$$
F(\rho, \sigma) := \sqrt{\text{Tr}(\sqrt{\sqrt{\rho} \sigma \sqrt{\rho}})}.
$$

**Definition 4 (Purified Distance [44], [38]):** Let $\rho_A, \sigma_A \in \mathcal{D}(\mathcal{H}_A)$. The purified distance between $\rho_A$ and $\sigma_A$ is defined as:

$$
P(\rho, \sigma) := \sqrt{1 - F(\rho, \sigma)^2},
$$

where $F(\rho, \sigma)$ is the fidelity. The purified distance is a metric on $\mathcal{D}(\mathcal{H})$. We use the purified distance to specify an $\epsilon$-ball around $\rho_A \in \mathcal{D}(\mathcal{H}_A)$, that is $B^\epsilon(\rho_A) := \{\rho_A' \in \mathcal{D}(\mathcal{H}_A) : P(\rho_A', \rho_A) \leq \epsilon\}$.

The purified distance is also monotone non-increasing with respect to quantum channels, obeys the triangle inequality and is invariant with respect to tensor product states. Moreover, the following expression shows its relation to the trace distance [38]:

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

\[3\] In general, $[\omega]_+$ denotes the projector onto the positive eigenspace of $\omega$ and $[\omega]_- = 1 - [\omega]_+$.\]
In addition to the hypothesis testing relative entropy, several different relative entropies and variances appear in our results and we shall consider their definitions here.

**Definition 5 (Conditional von Neumann entropy):** For a bipartite state \( \rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B) \), we define the conditional von Neumann entropy of \( A \) given \( B \) as follows:

\[
H(A|B) := H(AB) - H(B),
\]
where

\[
H(A) := -\text{Tr}(\rho_A \log_2 \rho_A),
\]

\( H(A)_\rho \) is the von Neumann entropy [17], corresponding to the Shannon entropy of the eigenvalues of \( \rho_A \).

**Definition 6 (Quantum Mutual Information):** The quantum mutual information of a bipartite state \( \rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B) \) is defined as follows:

\[
I(A;B)_\rho := H(A)_\rho + H(B)_\rho - H(AB)_\rho.
\]

The conditional quantum mutual information of a tripartite state \( \rho_{ABC} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C) \) is defined in an analogous way to its classical counterpart as follows:

\[
I(A;B|C)_\rho := H(A|C)_\rho + H(B|C)_\rho - H(AB|C)_\rho.
\]

**Definition 7 (Coherent Information):** The coherent information of a bipartite state \( \rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B) \) is defined as follows:

\[
I(A)(\rho) := H(B)_\rho - H(AB)_\rho.
\]

The conditional coherent information of a tripartite state \( \rho_{ABC} \) is defined as \( I(A)(B|C)_\rho := H(B|C)_\rho - H(AB|C)_\rho \) and it can be shown that \( I(A)(B) = I(A)B) \). In particular, for the CQ state \( \rho_{XAB} = \sum_x p(x) |x\rangle \langle x| \otimes \rho^A_B \) we have \( I(A)(B) = \sum_x p(x) I(A)(B) \).

**Definition 8 (Quantum Relative Entropy) [4]:** The quantum relative entropy for \( \rho_A, \sigma_A \in \mathcal{D}(\mathcal{H}_A) \) is defined as

\[
D(\rho_A\|\sigma_A) := \text{Tr}(\rho_A \log_2 \rho_A - \log_2 \sigma_A),
\]
whenever supp(\( \rho_A \)) \subseteq supp(\( \sigma_A \)) and otherwise it equals \( +\infty \).

**Fact 1 (Relation between the quantum relative entropy and the hypothesis testing relative entropy [11]):** For all states \( \rho_A \) and \( \sigma_A \) and \( \epsilon \in [0,1) \), the following inequality holds

\[
D_H(\rho_A\|\sigma_A) \leq \frac{1}{1 - \epsilon} \left[ D(\rho_A\|\sigma_A) + h_B(\epsilon) \right],
\]
where \( h_B(\epsilon) := -\log_2 \epsilon - (1 - \epsilon) \log_2(1 - \epsilon) \) is the binary entropy function.

**Definition 9 (Quantum relative entropy variance [13]):** The quantum relative entropy variance for \( \rho_A, \sigma_A \in \mathcal{D}(\mathcal{H}_A) \) is given by:

\[
V(\rho_A\|\sigma_A) := \text{Tr}[\rho_A \log_2 \rho_A - \log_2 \sigma_A - D(\rho_A\|\sigma_A)^2],
\]
whenever supp(\( \rho_A \)) \subseteq supp(\( \sigma_A \)) and \( D(\rho_A\|\sigma_A) \) is the quantum relative entropy.

**Definition 10 (Max-relative entropy [40]):** Max-relative entropy for \( \rho_A, \sigma_A \in \mathcal{D}(\mathcal{H}_A) \) is defined as:

\[
D_{\text{max}}(\rho_A\|\sigma_A) := \inf \left\{ \lambda \in \mathbb{R} : \rho_A \leq 2^\lambda \sigma_A \right\},
\]
where it is well-defined if supp(\( \rho_A \)) \subseteq supp(\( \sigma_A \)).

An important property of the max-relative entropy is its monotonicity under quantum operations.

**Fact 2 (Monotonicity of max-relative entropy [40]):** For quantum states \( \rho_A, \sigma_A \in \mathcal{D}(\mathcal{H}_A) \) and any CPTP map \( \mathcal{E} : \mathcal{L}(\mathcal{H}_A) \rightarrow \mathcal{L}(\mathcal{H}_B) \), it holds that

\[
D_{\text{max}}(\mathcal{E}(\rho_A)\|\mathcal{E}(\sigma_A)) \leq D_{\text{max}}(\rho_A\|\sigma_A).
\]

It can be shown that the monotonicity property also holds for the hypothesis testing relative entropy in the same direction.

**Fact 3 (Relation between quantum relative entropy and max-relative entropy [40]):** For quantum states \( \rho_A, \sigma_A \in \mathcal{D}_\leq(\mathcal{H}_A) \), it holds that

\[
D(\rho_A\|\sigma_A) \leq D_{\text{max}}(\rho_A\|\sigma_A).
\]

**Definition 11 (Smooth max-relative entropy [40]):** For a parameter \( \epsilon \in (0,1) \), smooth max-relative entropy for \( \rho_A, \sigma_A \in \mathcal{D}(\mathcal{H}_A) \) is defined as:

\[
D^\epsilon_{\text{max}}(\rho_A\|\sigma_A) := \inf_{\rho_A'} \left( D_{\text{max}}(\rho_A'\|\sigma_A) + \epsilon \right)
\]

or

\[
D^\epsilon_{\text{max}}(\rho_A\|\sigma_A) := \inf_{\rho_A'} \left( D_{\text{max}}(\rho_A'\|\sigma_A) + \epsilon \right).
\]

Fact 4 ([13] and [14]): Let \( \epsilon \in (0,1) \) and \( n \) be an integer. For any pair of states \( \rho_A \) and \( \sigma_A \) and their \( n \)-fold products, i.e., \( \rho_A^{\otimes n} \) and \( \sigma_A^{\otimes n} \), the following equations hold:

\[
D^\epsilon_{\text{max}}(\rho_A^{\otimes n}\|\sigma_A^{\otimes n}) = n D(\rho_A\|\sigma_A) + \epsilon n \log_2 n + O(\log_2 n),
\]

\[
D^\epsilon_{\text{max}}(\rho_A^{\otimes n}\|\sigma_A^{\otimes n}) = n D(\rho_A\|\sigma_A) + \epsilon n \log_2 n + O(n),
\]

where \( \Phi(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-\frac{x^2}{2}) dx \) is the cumulative distribution function for a standard Gaussian random variable and its inverse is defined as \( \Phi^{-1}(\epsilon) := \sup\{\alpha \in \mathbb{R} | \Phi(\alpha) \leq \epsilon\} \).

In the following we will define new entropic quantities that the analysis of their asymptotic behaviour requires Fact 4 as well as a useful result in information theory known as the asymptotic equipartition property (AEP). Let \( X^n = (X_1, X_2, ..., X_n) \) be a sequence of independent and identically distributed (i.i.d.) random variables. The AEP states that for any \( 0 < \epsilon < 1 \), any \( \delta > 0 \) and for large enough \( n \), a randomly chosen i.i.d. sequence \( x^n \) is with probability more than \( 1 - \epsilon \) in a \( \delta \)-typical set of sequences that satisfy

\[
\frac{1}{n} \sum_{i=1}^{n} I(x_i|\mathcal{X}^{i-1}) - I(x_i) \leq \delta,
\]

where \( N(x_i|\mathcal{X}^{i-1}) \) is the number of occurrences of \( x_i \) in the sequence \( x^n \). To use these concepts in quantum information, the notion of typical subspace is defined. Consider the state \( \rho_X = \sum_{x} p(x)|x\rangle \langle x| \). The \( \delta \)-typical subspace is a subspace of the full Hilbert space \( \mathcal{H}_X \otimes \mathcal{H}_Y \), associated with many copies of the density operator, i.e., \( \rho_X^{\otimes n} = \sum_{x^n} p(x^n)|x^n\rangle \langle x^n| \), that is spanned by states \( |x^n\rangle \) whose corresponding classical sequences are \( \delta \)-typical. For an introduction to the quantum typicality and more on the properties of the typical subspace, we refer the reader to [12].

We will present our results in terms of mutual information-like quantities defined below. We note that quantum mutual information (Definition 6) of a bipartite state \( \rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B) \).
$\mathcal{H}_B$ can be defined alternatively by quantum relative entropy (Definition 8) as follows:

$$I(A; B)_\rho := D(\rho_{AB} \| \rho_A \otimes \rho_B).$$

**Definition 12** (Hypothesis testing-mutual information [1]): For a bipartite state $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and a parameter $\epsilon \in (0, 1)$, from the hypothesis testing-relative entropy (Definition 11), the hypothesis testing-mutual information is defined as follows:

$$I_{H}(A; B)_\rho := D(\rho_{AB} \| \rho_A \otimes \rho_B).$$

**Definition 13** (Max-mutual information [10]): For a bipartite state $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and a parameter $\epsilon \in (0, 1)$, from the max-relative entropy (Definition 10), the max-mutual information can be defined as follow:

$$I_{\text{max}}(A; B)_\rho := D_{\text{max}}(\rho_{AB} \| \rho_A \otimes \rho_B).$$

**Definition 14** (Smooth max-mutual information [10]): For a bipartite state $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and a parameter $\epsilon \in (0, 1)$, from the max-mutual information (Definition 13), we define smooth max-mutual information as follows:

$$I_{\text{max}}^\epsilon(A; B)_\rho := \inf_{\rho'_{AB} \in \mathcal{B}(\rho_{AB})} D_{\text{max}}(\rho'_{AB} \| \rho_A \otimes \rho_B).$$

The following quantity is similar to smooth max-mutual information.

**Definition 15** (smooth max-mutual information, alternate definition [2]): For a bipartite state $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and a parameter $\epsilon \in (0, 1)$, the smooth max-mutual information alternately can be defined as follows:

$$I_{\text{max}}^\epsilon(B; A)_\rho := \inf_{\rho'_{AB} \in \mathcal{B}(\rho_{AB})} D_{\text{max}}(\rho'_{AB} \| \rho_A \otimes \rho_B).$$

**Fact 5** (Relation between two definitions of the smooth max-mutual information, [11] and see lemma 2 in [2]): Let $\epsilon \in (0, 1)$ and $\gamma \in (0, \epsilon)$. For a bipartite state $\rho_{AB}$, it holds that:

$$I_{\text{max}}^\epsilon(B; A)_\rho \leq I_{\text{max}}^{\epsilon - \gamma}(A; B)_\rho + \log_2 \left( \frac{3}{\gamma^2} \right).$$

**Definition 16** (Conditional smooth hypothesis testing-mutual information): Let $\rho_{ABX} := \sum_x p_X(x) |x\rangle \langle x| \otimes \rho_{AB}$ be a CQ state and $\epsilon \in (0, 1)$. We define

$$I_{H}^\epsilon(A; B|X)_\rho := \min_{\rho'_{X} \in \text{supp}(\rho_{X}^*)} I_{H}^\epsilon(A; B|\rho'_{AB}),$$

where maximization is over all $\rho'_{X} = \sum_x p_X(x) |x\rangle \langle x| \otimes \rho_{AB}^*$ satisfying $P(\rho'_{X}, \rho_X) \leq \epsilon$.

**Definition 17** (Conditional smooth max-mutual information): Let $\rho_{ABX} := \sum_x p_X(x) |x\rangle \langle x| \otimes \rho_{AB}$ be a CQ state and $\epsilon \in (0, 1)$. The conditional smooth max-mutual information is defined as follows:

$$I_{\text{max}}^\epsilon(A; B|X)_\rho := \min_{\rho'_{X} \in \text{supp}(\rho_{X}^*)} I_{\text{max}}^\epsilon(A; B|\rho'_{AB}),$$

where minimization is over all $\rho'_{X} = \sum_x p_X(x) |x\rangle \langle x| \otimes \rho_{AB}$ satisfying $P(\rho'_{X}, \rho_X) \leq \epsilon$.

---

### Lemma 1

Let $\rho_{XAB} := \sum_x p(x) |x\rangle \langle x| \otimes \rho_{AB}$. Then the following holds:

$$\lim_{n \to \infty} \frac{1}{n} I_{H}(A \otimes^n; B \otimes^n|X^n)_{\rho \otimes^n} = I(A; B|X)_\rho.$$

**Proof:** The following is easily seen from the definition,

$$I_{H}^{\epsilon}(A \otimes^n; B \otimes^n|X^n)_{\rho \otimes^n} := \max_{\rho_{X}^n \in \text{supp}(\rho_{X}^*)} D_{H}(\rho_{X^n|B^n} \| \rho_{X^n}\otimes \rho_{B^n}).$$

In order to apply the asymptotic results given in Fact 4, we first produce $\rho_{X^n}$ by projecting $\rho_{X^n}$ onto its typical subspace and properly normalize it. We know that the resulting state is close to the initial product state. Conditioned on a particular typical sequence $x^n$, the state $\rho_{X^n} \otimes \cdots \otimes \rho_{x^n}$ is in fact a tensor-product state that can be written as $\rho_{AB} \otimes \cdots \otimes \rho_{AB} \otimes \cdots \otimes \rho_{AB}$, in which $x(i), i \in [1 : n]$ indicates the $i$-th index in the sequence $x^n$. From the definition of the typical sequences, we know that for $n$ large enough, each realization $x$ appears almost $np(x)$ times in each sequence. Hence, for any $\delta > 0$, as $n \to \infty$, by using Fact 4 for each chosen sequence, the multi-letter formula above can be written as shown by (3), as shown at the top of the next page, where $x_i, i \in [1 : |X'|]$ denotes an element of the alphabet $X'$ and the second equality follows from Fact 4 and the fully quantum AEP [14].

**Lemma 2:** Let $\rho_{XAB} := \sum_x p(x) |x\rangle \langle x| \otimes \rho_{AB}$. Then the following holds.

$$\lim_{n \to \infty} \frac{1}{n} I_{\text{max}}^{\epsilon}(A \otimes^n; B \otimes^n|X^n)_{\rho \otimes^n} = I(A; B|X)_\rho.$$ 

**Proof:** The proof is very similar to that of Lemma 1. It employs the properties of the typical sequences as well as the fully quantum asymptotic equipartition property (AEP) for smooth max-mutual information [14].

**Lemma 3:** For a CQ state $\rho_{XAB} := \sum_x p(x) |x\rangle \langle x| \otimes \rho_{AB}$, the following inequality is true.

$$I_{H}^{\epsilon}(A; B|X)_\rho \leq \frac{1}{1 - \epsilon} \left( I(A; B|X) + h_b(\epsilon) \right).$$

**Proof:** Considering the definition of the conditional hypothesis testing-mutual information and the fact that

$$\min_{x} D_{H}(\rho_{AB}^x \| \rho_{B}^x) \leq \sum_x p(x) D_{H}(\rho_{AB}^x \| \rho_{AB}^x),$$

and also from Fact 1 for all $x$, we have:

$$D_{H}(\rho_{AB}^x \| \rho_{B}^x) \leq \frac{1}{1 - \epsilon} (D(\rho_{AB}^x \| \rho_{AB}^x) + h_b(\epsilon)),$$

by plugging into the aforementioned inequality, we can get the result. We note than in order for the above to be true, we should have $\rho_X^x \leq \rho_X$. However, in case $\rho_X^x$ goes beyond the support of $\rho_X$, it can be projected onto the support of $\rho_X$. Since $P(\rho_X^x, \rho_X) \leq \epsilon$, from the monotonicity of the purified distance, it can be seen that the state after being projected will remain $\epsilon$-close to the initial state.

**Lemma 4:** Let $\rho_{XAB} := \sum_x p(x) |x\rangle \langle x| \otimes \rho_{AB}$. The following inequality holds.

$$I_{\text{max}}^{\epsilon}(A; B|X)_\rho \geq I(A; B|X)_\rho - 2\epsilon \log |\mathcal{H}_A| - 2(1 + \epsilon) h_b(\epsilon).$$
where $P(\rho_{AB}^x, \tilde{\rho}_B^x) \leq \epsilon$. From Fact 3 we further know that $D_{max}(\rho_{AB}^x, \rho_{AB}^x) \geq D_{max}(\tilde{\rho}_A^x, \tilde{\rho}_B^x)$. Now we deploy Alicki-Fannes-Winter (AFW) inequality [47] (an improvement over [46]) for the quantum mutual information saying that: (from the relation between the purified and trace distances, we know that $\frac{1}{2}\|\rho_{AB} - \tilde{\rho}_A^x\| \leq \epsilon$)

$$D_{max}(\rho_{AB}^x, \rho_{AB}^x) \geq D_{max}(\tilde{\rho}_A^x, \tilde{\rho}_B^x) \geq -2\epsilon \log |H_A| - 2(1 + \epsilon)h_2\left(\frac{\epsilon}{1 + \epsilon}\right).$$

Therefore,

$$D_{max}(\rho_{AB}^x, \rho_{AB}^x) \geq D_{max}(\tilde{\rho}_A^x, \tilde{\rho}_B^x) \geq -2\epsilon \log |H_A| - 2(1 + \epsilon)h_2\left(\frac{\epsilon}{1 + \epsilon}\right),$$

and plugging back into the right-hand side of (4), we well get the desired result.

Lemma 5 (Convex-split lemma [41]): Fix $\epsilon \in (0, 1)$ and $\delta \in (0, \epsilon)$. Let $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and define the state $\tau_{A_1, \ldots, A_k}^{\otimes}$ as follows:

$$\tau_{A_1, \ldots, A_k}^{\otimes} \coloneqq \frac{1}{|K|} \sum_{k=1}^{|K|} \rho_{A_1} \otimes \cdots \otimes \rho_{A_{k-1}} \otimes \rho_{A_k} \otimes \rho_{A_{k+1}} \otimes \cdots \otimes \rho_{A_{|K|}}.$$

If

$$\log_2 |K| \geq \tilde{I}_{max}(B; A)_\rho + 2 \log_2 \left(\frac{1}{\delta}\right),$$

then

$$P(\tau_{A_1, \ldots, A_k}^{\otimes}, \rho_{AB}^x, \tilde{\rho}_B^x) \leq \sqrt{\epsilon},$$

where $\tilde{\rho}_B$ is the marginal of some state $\rho_{AB} \in B^{\sqrt{\epsilon}(\rho_{AB})}$. The above smooth version of convex-split lemma is taken from [3], which improved the error parameters in the smooth version given in [2].

Lemma 6 (Hayashi-Nagaoka operator inequality [32]): Let $T, S \in \mathcal{P}(\mathcal{H}_A)$ such that $(I - S) \in \mathcal{P}(\mathcal{H}_A)$. Then for all constants $c > 0$, the following inequality holds:

$$I - (S + T)^{-\frac{c}{2}} S (S + T)^{\frac{c}{2}} \leq (1 + c)(I - S) + (2 + c + c^{-1})T.$$
where for each message \( m \), \( \rho_{m}^{\text{LE}} \) and \( \rho_{L}^{m} \) are appropriate marginals of the state \( \rho_{L}^{\text{BE}} = \frac{1}{|L|} \sum_{l \in L} \rho(N(M(m), t)) \) and \( \rho_{L}^{m} \) can be any arbitrary state.

- A decoding operation by Bob \( D : D(H_B) \rightarrow \hat{M} \hat{L} \) such that

\[
P_{F} \left( (\hat{M}, \hat{L}) \neq (M, L) \right) \leq \epsilon, \tag{6}
\]

where \( \hat{L} \) and \( \hat{M} \) denote the estimates of the public and private messages, respectively.

A rate pair \( (r, R) \) is said to be \((\epsilon, \epsilon')\)-achievable if there exist encoding and decoding maps \( (E, D) \) such that (5) and (6) are fulfilled. For a given \((\epsilon, \epsilon')\), the one-shot capacity region for the simultaneous transmission of public and private information of the channel \( N, C_{\epsilon, \epsilon'}(N) \), is the closure of all achievable rate pairs in a \((r, R, \epsilon, \epsilon')\) coding scheme. In this work, our aim is to find upper and lower bounds on \( C_{\epsilon, \epsilon'}(N) \).

In the following, we first have Theorem 2 that establishes a lower bound on \( C_{\epsilon, \epsilon'}(N) \) referred to as achievability and then Theorem 3 that states an upper bound on \( C_{\epsilon, \epsilon'}(N) \), i.e., the converse. This section ends with a discussion about the translation of the private classical capacity to the quantum capacity in one-shot regime.

**Theorem 2 (Achievability):** For any fixed \( \epsilon \in (0, 1), \epsilon' \in (0, 1) \), and \( \delta, \delta' \) such that \( \delta \in (0, \epsilon) \), \( \delta' \in (0, \epsilon') \), there exists a one-shot \((r, R, \epsilon, \epsilon', \delta, \delta')\) code for the channel \( N_{A \rightarrow BE} \) if the twin \((r, R)\) satisfies the following bounds:

\[
r \leq I_{H}^{-\delta}(X; B)_{\rho} - \log_{2} \left( \frac{4\epsilon}{\delta'} \right),
\]

\[
R \leq I_{H}^{-\delta}(Y; B|X)_{\rho} - \hat{I}_{\max}^{\epsilon' + \delta'}(Y; E|X)_{\rho}
- \log_{2} \left( \frac{4\epsilon}{\delta'} \right) - 2 \log_{2} \left( \frac{1}{\delta'} \right),
\]

for some quantum state \( \rho \) arising from the channel. We call the region above \( C_{\epsilon}(N) \), therefore, we have

\[
C_{\epsilon}(N) \subseteq C^{\delta + 2\epsilon + \sqrt{\epsilon'} + 2(\epsilon + \sqrt{\epsilon'}) + \sqrt{\epsilon'}}(N).
\]

**Theorem 3 (Converse):** For any fixed \( \epsilon \in (0, 1), \epsilon' \in (0, 1) \), every one-shot \((r, R, \epsilon, \epsilon')\) public-private code for the channel \( N_{A \rightarrow BE} \), must satisfy the following inequalities:

\[
r \leq I_{H}^{-\delta}(X; B)_{\rho},
\]

\[
R \leq I_{H}^{\epsilon'}(Y; B|X)_{\rho} - I_{\max}^{\epsilon'}(Y; E|X)_{\rho},
\]

for some state \( \rho_{XYBE} = \sum_{x,y} p(x,y) |x\rangle \langle x| \otimes |y\rangle \langle y| \otimes \rho_{BE}^{x,y} \). We refer to this region as \( C_{\epsilon}(N) \). In fact, we have \( C_{\epsilon, \epsilon'}(N) \subseteq C_{\epsilon}(N) \).

Once there is a code for simultaneous transmission of public and private classical information, this code can be translated into a coherent code that is capable of transmitting classical and quantum information simultaneously. In other words, the rate pair (public classical, private classical) can be shifted to the rate pair (public classical, quantum) (or simply (classical, quantum)). We can then translate our one-shot (public, private) code to a one-shot (classical, quantum) code. Note that the proof is implicit in findings of Devetak [27] such that one can mimic his procedure to see the result in one-shot setting. Henceforth, we have a one-shot code for simultaneous transmission of classical and quantum information.

By evaluating the asymptotic behaviour of the rate region given by Theorem 2 and Theorem 3 (Section VI), we recover Theorem 1 of [19], the well-known result of Devetak and Shor, as a corollary.

**IV. Achievability**

We consider a general quantum channel which is preprended by an encoder (modulator) that associates a particular input state to every classical input pair. In this sense, Alice can be thought of as being in possession of an ensemble \( \{p_{X,Y}(x,y), \omega_{A,x}^{y} \} \) such that the input distribution \( p(x,y) \) and the encoder need to be optimized over to get our capacity results. In our protocol, Bob runs two successive decodings, his first decoder has \(|M|\) possible classical outputs as well as a post-measurement quantum state. His second decoder takes the resulted states of the first decoder and its output is a classical system of dimension \(|L|\). Before we get into achievability proof, we describe our protocol.

**A. Protocol Description**

Fix a joint probability distribution \( p_{X,Y}(x,y) \) over the finite alphabets \( \{X \times Y\} \), \( \epsilon, \epsilon' \in (0, 1) \), \( \delta \in (0, \epsilon) \), \( \delta' \in (0, \epsilon') \) and \( \rho_{XYBE} = \sum_{x,y} p(x,y) |x\rangle \langle x| \otimes |y\rangle \langle y| \otimes \rho_{BE}^{x,y} \). Let

\[
r \leq I_{H}^{-\delta}(X; B)_{\rho} - \log_{2} \left( \frac{4\epsilon}{\delta'} \right),
\]

\[
R + \tilde{R} \leq I_{H}^{-\delta}(Y; B|X)_{\rho} - \log_{2} \left( \frac{4\epsilon}{\delta'} \right),
\]

\[
\tilde{R} \geq I_{\max}^{\epsilon'}(E; Y|X)_{\rho} + 2 \log_{2} \left( \frac{1}{\delta'} \right).
\]

We choose \(|M| = 2^{R}, |L| = 2^{R} \) and \(|K| = 2^{\tilde{R}} \) implying that \( r \) and \( R \) denote our public and private rates, respectively and \(|K| \) stands for the size of a local key, a uniformly distributed random variable \( K \), used by Alice for obfuscation purpose. Let the sender Alice, legitimate receiver Bob and Eve be connected by means of a quantum (wiretap) channel \( N^{A \rightarrow BE} \).

Alice wants to convey to Bob, in a single use of a quantum channel, a classical message \( m \in M \) and simultaneously, a private classical message \( \ell \in L \) where both messages are uniformly distributed on their corresponding sets. The message \( m \) is public, meaning that Bob has to be able to decode it correctly with small probability of error. On the other hand, message \( \ell \) is private and while Bob has to receive it with negligible error probability, it must be kept secret from Eve. We clarify that our definition of public and private messages is the same as in [19] and these correspond respectively to common and confidential messages defined in [18].

The position-based decoding is employed in order to accomplish this information-processing task, therefore before communication begins, Alice, Bob and Eve share the state given in (7), where Alice controls the system \( A \), Bob has systems \((X, Y)\) and Eve is in possession of \((X', Y')\) systems. Our coding scheme is, in spirit, inspired from the well-known superposition coding in classical information theory [43]. Our coding scheme is, in spirit, inspired from the well-known superposition coding in classical information theory [43].
each of which is use to accomplish a certain part of the task. There are \(|\mathcal{M}|\) bins in the first place, inside each of them, there are \(|\mathcal{L}|\mathcal{K}|\) states that are divided into \(|\mathcal{L}|\) bins, again inside each one there are \(|\mathcal{K}|\) states.

Upon receiving the message pair \((m, \ell)\), Alice goes to the \(m\)-th copy of \(\rho_{XX'(AYY')|L|K}|\). There she runs the protocol for the private capacity, by considering \(|\mathcal{L}|\mathcal{K}|\) copies and choosing a system A uniformly at random from the \(\ell\)-th bin. Upon receiving \(B\), Bob performs a position-based decoding to obtain the public message \(m\) (and hence the correct copy of \(\rho_{XX'(AYY')|L|K}|\)). The choice of the rate for public message \(r\) ensures that this is possible and gentle measurement lemma ensures that the quantum state of the correct copy of \(\rho_{XX'(AYY')|L|K}|\) is almost unchanged after Bob’s decoding.

To decode \(\ell\), Bob performs another position-based decoding conditioned on \(X\), meaning that having found the correct copy of \(\rho_{XX'(AYY')|L|K}|\) used in the transmission, Bob applies a decoder that depends on \(X\), and it works for all \(x \in \mathcal{X}\). For this strategy, Bob first appeals to the definition of the conditional smooth hypothesis testing-mutual information, to assume that the distribution over \(X\) was \(p'(x)\) (achieving the infimum in the definition) with negligible error. Then for \(x \in \text{supp}(\rho_{XX'})\), he performs position-based decoding. The choice of \(R + \tilde{R}\) guarantees the successful decoding for every \(x\) and at the same time, the security criterion is ensured from the fact that even if Eve is aware of the correct copy of \(\rho_{XX'(AYY')|L|K}|\), the condition that convex-split lemma imposes on \(|\mathcal{K}|\), gives her very small information about \(\ell\) for every \(x \in \text{supp}(\rho_{XX'})\) (where here \(\rho_{XX'} = \sum x p(x|x)x|XX'\) and \(p_{XX'}(x)\) is the distribution achieving the infimum in the alternate definition of conditional smooth max-mutual information). Now we can derandomize the protocol by fixing the values in corresponding systems. Upon derandomization, the code is publicly available.

Before we proceed to the error analysis of the direct part, we make the following remarks. The state that is fed into the second decoder differs from the original state although negligibly, this adds to the error probability of the private message. Moreover, since successive cancellation decoding is being performed, in the event of a failure of the first decoder, the second decoder will fail as well. We also take the contribution of this event into account. Moreover, note that there is just one decoding map in general, Bob’s (two) separate decoding are just a property of our protocol.

\[ \rho_{XX'(AYY')|L|K}| := \left( \sum_x p(x|x)\langle x|XX'| xx' \otimes \left( \sum_y p(y|x)\langle y|YY'| yy' \otimes \rho_{XX'}^{x,y} \right) \right)^{\otimes |\mathcal{L}|\mathcal{K}|}. \] (7)

### B. Achievability Proof

As is learned in the preceding subsection, we start with a randomness assisted protocol and derandomize it later. We get started on our proof by introducing the encoder and the decoders. We then analyze the average error probability of the public message. Likewise, we inspect the second decoder and analyze the average error probability of the private message. Finally, we study the secrecy requirement.

In the achievability part of our randomness assisted code, for the private message, we stick to a single criterion known as privacy error introduced in [7], [8] and [3]. The general idea is to merge the secrecy of the private message \(5\) as well as its error probability \(6\) into one single criterion. While this idea was used in [7] and [8] in understanding upper bounds for private communication protocols, it had not been used in an achievability proof prior to [3]. We should note that the main advantage of dealing with single criterion reveals when the code is to be derandomized. Our procedure is that we analyze the error probability of Bob in detecting the private message separately from keeping Eve ignorant. This leads to two separate criteria and then the separate criteria are merged into one single criterion. It is clear that if the joined criterion is satisfied, each of the single criteria is also fulfilled. After we prove the correctness of these criteria for the randomness assisted code, we immediately proceed to derandomize the code in the succeeding step that the unassisted criteria set out by Definition 18 can be inferred. The derandomization involves some procedures that appeared in [3] and [22].

Alice, Bob and Eve are allowed to share some quantum state among themselves. Moreover, Alice has access to a source of uniform dummy randomness given in random variable \(K\). Further, let \(\tilde{R} = \log_2 |K|\). The state initially shared between three parties is given by equation (7), as shown at the top of this page, where Alice possesses the quantum systems \(A\), Bob possesses the classical systems \(X, Y\) and Eve has the classical systems \(X', Y'\). For ease of notation, we further define \(\Upsilon_T : XX'YX'Y | L, K | := \rho_{XX'(AYY')|L|K}|^{\otimes |\mathcal{M}|} |^{\otimes |\mathcal{L}|\mathcal{K}|},\) with it being clear that for example \(\Upsilon_T A = \rho_{A|L, K}|^{\otimes |\mathcal{M}|} |^{\otimes |\mathcal{L}|\mathcal{K}|}\).

The encoding and decoding pairs are as follows:

- Alice performs some encoding operation \(\mathcal{E} : \mathcal{M} \rightarrow \mathcal{A}\).
- Let us denote the state in (7) after channel transmission as:

\[ \left( \rho_{XX'(AYY')|L|K}|^{\otimes |\mathcal{M}|} |^{\otimes |\mathcal{L}|\mathcal{K}|}, \right) \otimes \rho_{XX'(YY')|L|K}|^{\otimes |\mathcal{M}|} |^{\otimes |\mathcal{L}|\mathcal{K}|}^{-1} BE, \]

(8)

where \((m, \ell, k) \in [1 : 2^m] \times [1 : 2^\ell] \times [1 : 2^\theta]\) are the public message, the private message and a dummy index drawn uniformly at random by the encoder and \(p_{XX'(YY')|L|K}|^{\otimes |\mathcal{M}|} |^{\otimes |\mathcal{L}|\mathcal{K}|}^{-1} BE\) is given by equation (9), as shown at the top of the next page.

- After the channel action, Bob performs a decoding operation (quantum instrument) \(D^1 : BX \rightarrow MB\) on his \(\Upsilon_T\)
Having defined the decoders, it is seen that the phrase in (12),

\[ D^2 : MBY \rightarrow L \] (which is input the classical output of the first decoder, the disturbed quantum output, Bob's \( \mathcal{T}_Y \) systems and outputs a classical system \( L \).

\[ D^2_{MBY \rightarrow L} (\sigma^{m,(\ell,k)}_{XY} |\mathcal{C}| |\mathcal{B}|) := \sum_{\ell=1}^{|\mathcal{C}|} \rho^{\ell} \{ \ell \} |\ell \rangle \langle \ell |, \]

where \( \{ |\ell \rangle \}_{\ell=1}^{|\mathcal{C}|} \) is some orthonormal basis and \( \sigma^{m,(\ell,k)}_{XY} |\mathcal{C}| |\mathcal{B}| \) comes about by tracing out uninvolved systems in the disturbed state.

Having defined the decoders, it is seen that the phrase in (12), as shown at the top of the next page, indicates the probability of an erroneous detection of the public message, while the expression in (13), as shown at the top of the next page, captures the notions of an erroneous detection of the private message as well as the secrecy condition of the eavesdropper (the latter is clarified below). After we derandomize the code, we see that the criteria mentioned in Definition 18 can be set out from these criteria by using the monotonicity of the trace distance and properly adjusting the constants.

\[ D^1_{BX \rightarrow MB} (\rho^{m,(\ell,k)}_{XBM}) := \sum_{m=1}^{M} \text{Tr} [\Lambda^m_{XBM} \rho^{m,(\ell,k)}_{XBM}] |m\rangle \langle m| \tilde{M} , \]

where \( \Lambda^m_{XBM} \) is given in (15), as shown at the top of the next page, and for \( m \in [1 : |\mathcal{M}|] : \]

\[ \Gamma^m_{XBM} = \mathbb{1}^1_X \otimes \mathbb{1}^2_X \otimes \ldots \otimes T^m_{XBM} \otimes \ldots \otimes \mathbb{1}^{|\mathcal{M}|}_X , \]

in which, \( T^m_{XBM} \) is a test operator distinguishing between two hypotheses, \( \rho_X \) and \( \rho_Y \otimes \rho_B \) and \( \rho^{m,(\ell,k)}_{XBM} \) can be seen from (7). In fact, Bob needs to discriminate between the following states for different values of \( m \in \mathcal{M} \).

\[ \rho^{m,(\ell,k)}_{XBM} := \rho^{m,(\ell,k)}_{XBM} - \rho^{m,(\ell,k)}_{XBM} \].

Note that to decode the public message \( m \), Bob’s decoder does not care about the copy selected by Alice among \( |\mathcal{C}| \) copies (no matter which one is selected). In other words, to accomplish the protocol for transmitting the public message, it suffices to consider \( |\mathcal{M}| \) copies of \( \rho_{XY} = \sum_x p_X(x) |x\rangle \langle x| \otimes \omega^x_M \) shared between Alice and Bob, where \( \omega^x_M = \sum_{x'} p(y|x) \omega^x_A \). Besides, as is clear from the former discussion, Bob’s first decoder faces an \( |\mathcal{M}| \) -ary hypothesis testing problem. This \( |\mathcal{M}| \)-ary hypothesis testing problem can be reduced to a binary hypothesis testing problem, in which a binary test operator discriminates between two hypotheses. However, it should not be confused with the fact that in general we deal with an \( |\mathcal{M}| \)-ary problem.

Let \( T_{XBM} \) be a test operator in a binary hypothesis testing scenario with null and alternative hypotheses being \( \rho_{X} \) and \( \rho_{Y} \otimes \rho_{B} \), respectively. Discriminator employed by Bob succeeds in guessing null and alternative hypotheses with probabilities \( \text{Tr}(T_{XBM} \rho_{X}) \) and \( \text{Tr}((\mathbb{1}^1_X - T_{XBM}) (\rho_{Y} \otimes \rho_{B})) \), respectively. And accordingly, the error probabilities associated to the type I and II errors are \( \text{Tr}((\mathbb{1}^1_X - T_{XBM}) \rho_{X}) \) and \( \text{Tr}(T_{XBM} (\rho_{Y} \otimes \rho_{B})) \), respectively.

It is notation-wise useful to assume that the error probability of the hypothesis tester is \( \epsilon - \delta \) where \( \delta \in (0, \epsilon) \) implying that overall probability of error \( \epsilon \) is greater than or equal to that of the hypothesis tester. Having introduced the test operator, we can define the following measurement operator for all \( m \in [1 : |\mathcal{M}|] : \]

\[ \Gamma^m_{XBM} = \mathbb{1}^1_X \otimes \ldots \otimes T^m_{XBM} \otimes \ldots \otimes \mathbb{1}^{|\mathcal{M}|}_X . \]

If Alice sends the \( m \)-th message (copy), the probability of producing the correct message at the output equals:

\[ \text{Tr}(\Gamma^m_{XBM} \rho^{m,(\ell,k)}_{XBM}) \]

\[ = \text{Tr}((\mathbb{1}^1_X \otimes \ldots \otimes T^m_{XBM} \otimes \ldots \otimes \mathbb{1}^{|\mathcal{M}|}_X) (\rho^m_{X} \otimes \ldots \otimes \rho^{m,(\ell,k)}_{XBM} \otimes \ldots \otimes \rho^{|\mathcal{M}|}_X) \]

\[ = \text{Tr}(T^m_{XBM} \rho_{X}) , \]

(16)
\[ P_e = |\hat{M} \neq M| := \frac{1}{M} \sum_{m=1}^{\lfloor |M|/2 \rfloor} \| D^{1}_{B_{X \rightarrow \hat{M}}}(\rho^{m,(l,k)}_{X|\hat{M}|B}) - |m\rangle \langle m|_{\hat{M}} \|_1 \leq \epsilon, \] (12)

\[ P_{\text{prior}} := \frac{1}{|\mathcal{L}|} \sum_{l=1}^{\lfloor |\mathcal{L}|/2 \rfloor} \| D^{2}_{M_{B_{Y \rightarrow L}}(\sigma^{m,(l,k)}_{X|Y|\mathcal{L}|E}) - |l\rangle \langle l|_{L} \otimes \hat{\sigma}^{m,(l,k)}_{E} \|_1 \leq 2(\epsilon + \sqrt{\epsilon}) + \sqrt{\epsilon'}, \] (13)

where in the last equality we drop the dependence on \( m \) since it is the same for all messages. And probability of deciding in favor of \( m' \neq m \) when \( m \) was sent is equal to:

\[
\begin{align*}
\text{Tr}[\Gamma^{m'}_{X|\mathcal{M}|B} \rho^{m,(l,k)}_{X|\mathcal{M}|B}] &= \text{Tr}[\Gamma^{m'}_{X|\mathcal{M}|B} \rho^{m,(l,k)}_{X} \rho^{m'}_{X}]
\quad = \text{Tr}[\rho^{m,(l,k)}_{X} \rho^{m'}_{X}]
\quad = \text{Tr}[\rho^{m,(l,k)}_{X} \rho^{m'}_{X}]
\quad = \text{Tr}[\rho^{m,(l,k)}_{X} \rho^{m'}_{X}]
\end{align*}
\]

\[\Gamma^{m}_{X|\mathcal{M}|B} := \left( \sum_{m'=1}^{\lfloor |\mathcal{M}|/2 \rfloor} \Gamma^{m'}_{X|\mathcal{M}|B} \right)^{-\frac{1}{2}} \Gamma^{m}_{X|\mathcal{M}|B} \left( \sum_{m'=1}^{\lfloor |\mathcal{M}|/2 \rfloor} \Gamma^{m'}_{X|\mathcal{M}|B} \right)^{-\frac{1}{2}} - (2 + c + c^{-1}) \text{Tr}[\Gamma^{m}_{X|\mathcal{M}|B} \rho^{1,(l,k)}_{X|\mathcal{M}|B}]
\]

\[\Gamma^{m}_{X|\mathcal{M}|B} := \left( \sum_{m'=1}^{\lfloor |\mathcal{M}|/2 \rfloor} \Gamma^{m'}_{X|\mathcal{M}|B} \right)^{-\frac{1}{2}} \Gamma^{m}_{X|\mathcal{M}|B} \left( \sum_{m'=1}^{\lfloor |\mathcal{M}|/2 \rfloor} \Gamma^{m'}_{X|\mathcal{M}|B} \right)^{-\frac{1}{2}} - (2 + c + c^{-1}) \text{Tr}[\Gamma^{m}_{X|\mathcal{M}|B} \rho^{1,(l,k)}_{X|\mathcal{M}|B}]
\]

where in the last equality we remove the index \( m' \) because this quantity is the same for all \( m' \neq m \). This endorses our claim saying that we are facing a binary hypothesis testing problem. From the aforementioned measurement operators, the square-root measurement given in (15) is formed acting as Bob’s POVM to detect the public message \( m \). The mentioned POVM construction and the coding scheme, known as position-based coding, first appeared in [41] and [42].

We now focus on the analysis of the test probability of the position-based decoder. The POVM elements above are unitary permutations of one another. In particular, it can be easily shown that all of the elements can be reached by a unitary permutation of the first one, i.e., \( \Lambda^{m}_{X|\mathcal{M}|B} := U^{m}(\pi_{m}) U_{X|\mathcal{M}|B}^{\dagger} U_{X|\mathcal{M}|B}^{\dagger} U^{m}(\pi_{m}) \) in which \( \pi_{m} \) denotes the permutation operator [3]. Having said this, we find the probability of error for the first message, i.e., Alice received \( m = 1 \) and has chosen and sent one of the \( |\mathcal{L}| \) subsystems of the first copy over the channel. We emphasize again that although Alice selects a particular \( A \) subsystem out of \( |\mathcal{L}| \) copies based on reliability and security of the private message, at this point, when Bob aims to estimate the public message, no matter which \( A \) was chosen by Alice, it does not affect Bob’s decision about the public message.

We begin by applying the Hayashi-Nagaoka operator inequality (Lemma 6) with \( S = \Lambda^{1}_{X|\mathcal{M}|B} \) and \( T = \sum_{m=1}^{\lfloor |\mathcal{M}|/2 \rfloor} \Gamma^{m}_{X|\mathcal{M}|B} \) (This \( T \) should not be confused with the test operator \( T^{m}_{X|\mathcal{M}|B} \)):

\[ P_{r}(\hat{M} \neq 1|\mathcal{M}=1) \leq \text{Tr}[1\Lambda^{1}_{X|\mathcal{M}|B} - \Lambda^{1}_{X|\mathcal{M}|B}] \rho^{1,(l,k)}_{X|\mathcal{M}|B} \leq \text{Tr}[(1 + c)(\Lambda^{1}_{X|\mathcal{M}|B} - \Lambda^{1}_{X|\mathcal{M}|B}) \rho^{1,(l,k)}_{X|\mathcal{M}|B}] \]

\[\text{log}_{2} |\mathcal{M}| \leq I_{H}^{\epsilon}(X;\mathcal{B})_{\rho} + \text{log}_{2} \left( \frac{\epsilon - (1 + c)(\epsilon - \delta)}{2 + c + c^{-1}} \right), \] (18)
and average probability of error of the public message for the one-shot assisted code is
\[ \frac{1}{|M|} \sum_{m=1}^{|M|} \text{Tr}((I_{X|m} \otimes I_B - \Lambda_{X|m}^m)(P_{X|m}^{m,(\ell,k)}) \leq \epsilon. \] (19)

In the following, we deal with the private message and the second decoder. Before we move on to the privacy analysis, we make a couple of remarks. If the first decoder fails, the second decoder breaks down completely since as intuitively clear, it ends up with a state having zero information about the position of the sent message. We precisely evaluate the contribution of the first decoder to the error of the second decoder. Moreover, Bob’s first decoder acts on his X systems as well as the output of the channel. The Y systems remain intact and in fact, when Bob applies the first decoder, one can assume that the uninvolved systems are being multiplied by the identity operators. Considering this point and the action of the POVM, the resulting state on systems X and B are (up to normalization) \[ \frac{\sqrt{\Lambda_{X|m}^m}}{\sqrt{\Lambda_{X|m}^m}} P_{X|m}^{m,(\ell,k)} \Lambda_{X|m}^m \] and by taking uninvolved systems into account, we define the state that passes to the second decoder as in (20), as shown at the top of the next page.

2) Correctness and Secrecy of Private Message, (Privacy error Eq. (13): Reconsider the state in (20) showing the state resulted from transmitting the (\ell, k)-th A subsystem through the channel (for a given m) after Bob applies his first decoder. Remember that in the first part of the protocol it did not matter which copy out of [L][K] copies was chosen but now it does matter as Bob and Eve try to decode the private message. Bob’s decoder for the private message \( \ell \) is constructed as follows:
\[ D_{MBY}^{\ell}(\sigma_{XY|L[K]|B}^{m,(\ell,k)}) \]
\[ = \sum_{\ell=1}^{X} \text{Tr}(P_{XY|L[K]|B}^{m,(\ell,k)} \sigma_{XY|L[K]|B}^{m,(\ell,k)} |\ell\rangle |\ell\rangle \),
where for all \( x \in X \)
\[ P_{XY|L[K]|B}^{\ell} = \sum_{k=1}^{K} P_{XY|L[K]|B}^{m,(\ell,k)} \]
and \( P_{XY|L[K]|B}^{m,(\ell,k)} \) is given in (22), as shown at the top of the next page, in which, for all \( x \in X \), \( Z_{YB} \) is a binary test operator distinguishing between two hypotheses \( \sigma_{YB}^x \) and \( \sigma_{YB}^\delta \) with an error of \( \epsilon - \delta \), i.e.,
\[ \text{Tr}(Z_{YB} \sigma_{YB}^\delta) \geq 1 - (\epsilon - \delta), \]
where \( \epsilon \in (0, 1) \) and \( \delta \in (0, \epsilon) \). Note that the variable \( x \) appearing in the operator indicates the fact that the decoding works for all \( x \in X \).

Bob has to be able to distinguish between states \( \sigma_{XY|L[K]|B}, \sigma_{XY|L[K]|B}^{x}, \ldots, \sigma_{XY|L[K]|B}^{m,(\ell,k)} \). We will see that this amounts to Bob being able to distinguish between the following states:
\[ \sum_{x} p_X(x) |x\rangle \langle x| \otimes \sigma_{YB}^x, \]
\[ \sum_{x} p_X(x) |x\rangle \langle x| \otimes \sigma_{YB}^\delta \otimes \sigma_{YB}^\delta, \]
or more precisely, between state \( \sigma_{YB}^x \) and \( \sigma_{YB}^\delta \otimes \sigma_{YB}^\delta \) for all \( x \in X \). We importantly note that after detecting the public message \( m \), Bob is faced a \( [L][K] \)-ary hypothesis testing problem. This scenario should not be confused by the binary hypothesis testing above, i.e., Alice distinguishes between \( \sigma_{YB}^x \) and \( \sigma_{YB}^\delta \otimes \sigma_{YB}^\delta \) for all \( x \in X \), the latter happens to be a byproduct of the general scenario once we go into the error analysis. Now see that if the pair \( (\ell, k) \) was chosen, the action of the operator \( N_{XY|L[K]|B}^{m,(\ell,k)} \) would be as follows:
\[ \text{Tr}(N_{XY|L[K]|B}^{m,(\ell,k)} \sigma_{XY|L[K]|B}^{m,(\ell,k)}) = \sum_{x} p_X(x) \text{Tr}(Z_{YB} \sigma_{YB}^x), \]
and for any other operator, i.e., the private message-local key pair \( (\ell, k) \) is confused by \( (\ell', k') \), either \( k \neq k' \) or \( \ell \neq \ell' \):
\[ \text{Tr}(N_{XY|L[K]|B}^{m,(\ell',k')} \sigma_{XY|L[K]|B}^{m,(\ell',k')}) = \text{Tr}(|\ell\rangle \langle \ell| \otimes (\sigma_{YB}^{x,m,(\ell',k')} \otimes \sigma_{YB}^{x,m,(\ell',k')})) \]
\[ = \sum_{x} p_X(x) \text{Tr}(Z_{YB} \sigma_{YB}^{x,m,(\ell',k')} \otimes \sigma_{YB}^{x,m,(\ell',k')}). \]

We can think of the states \( \sigma_{YB}^{x,m} \) as the null and alternative hypotheses, respectively. As a typical procedure in quantum error analysis, Bob forms the square-root measurement operators given in (22) acting as his POVMs to detect the private message-local key pair \( (\ell, k) \). It can be shown that each measurement operator \( P_{Y|L[K]|B}^{\ell,(k)} \) is related to the first one \( P_{Y|L[K]|B}^{(1,1)} \) by a unitary permutation of \( Y|L[K]|B \) systems for all \( x \in X \). This fact gives rise to the following identity, for all \( \ell \in [1 : |L|] \) and \( k \in [1 : |K|] \):
\[ \text{Tr}(I_{XY|L[K]|B} - P_{Y|L[K]|B}^{(1,1)} \sigma_{XY|L[K]|B}^{m,(1,1)}) = \text{Tr}(I_{XY|L[K]|B} - P_{Y|L[K]|B}^{\ell,(k)} \sigma_{XY|L[K]|B}^{m,(\ell,k)}) \]
meaning that the error probability is the same for all private messages, in other words, it is independent from a particular chosen twin \( (\ell, k) \). And again this implies that average error probability equals individual error probabilities. In what follows, we deploy Hayashi-Nagaoka operator inequality (Lemma 6) to analyze the error probability. Let’s assume \( (\ell = 1, k = 1) \) was sent. Moreover, let’s choose \( S = N_{Y|L[K]|B}^{(1,1)} \) and \( T = \sum_{\ell' \neq 1} \sum_{k' \neq 1} N_{Y|L[K]|B}^{(\ell',k')} \) in Hayashi-Nagaoka inequality. We have
\[ \text{Tr}(I_{XY|L[K]|B} - P_{Y|L[K]|B}^{(1,1)} \sigma_{XY|L[K]|B}^{m,(1,1)}) \leq (1 + c) \text{Tr}(I_{XY|L[K]|B} - N_{Y|L[K]|B}^{(1,1)} \sigma_{XY|L[K]|B}^{m,(1,1)}) \]
\[ + (2 + c + c^{-1}) \sum_{\ell' \neq 1} \sum_{k' \neq 1} \text{Tr}(N_{XY|L[K]|B}^{m,(\ell',k')} \sigma_{XY|L[K]|B}^{m,(\ell',k')}) \]
\[ = (1 + c) \sum_{x} p_X(x) \text{Tr}(I_{YB} - Z_{YB}^{(1,1)} \sigma_{YB}^{x,m,(1,1)}) \]
\[ + (2 + c + c^{-1}) \sum_{x} p_X(x) \text{Tr}(I_{YB} - Z_{YB}^{(1,1)} \sigma_{YB}^{x,m,(1,1)}) \]
\[
\sigma_{X^m, Y^m}^{x,m,(\ell,k)}^{(X,Y\otimes\mathcal{E}[\mathcal{K}])BE} := \sum_x p_X(x) \langle x | x \otimes | x \rangle \langle x | x' \otimes \sigma_{Y, Y'}^{x,m,(1,1)} \otimes \cdots \otimes \sigma_{Y, Y'}^{x,m,(\ell-1)} \otimes \sigma_{Y, Y'}^{x,m,(\ell,k)} \otimes \cdots \otimes \sigma_{Y, Y'}^{x,m,([L],[K])} \rangle.
\]

we end up having:

\[
\log_2 |\mathcal{L}| |\mathcal{K}| \leq \mathcal{I}^{\ell,0}_{\mathcal{H}}(Y; B|X)_{\sigma_{XYB}} - \log_2 \left( \frac{4C}{\delta^2} \right).
\]

The derivation above ensures that in the privacy error in (13), Bob’s error in detecting private message is satisfied (note that each separate criterion comes about by tracing out the other one).

We now turn our attention to Eve’s state and security criterion which is merged into (13). We also assume that Eve has detected the public message. From (20), for a fixed \((\ell, k)\), Eve’s state is

\[
\sigma_{X^m, Y^m}^{x,m,(\ell,k)} \sigma_{Y, Y'}^{x,m,([L],[K])} := \sum_x p_X(x) \langle x | x \otimes 1_{Y_1}^{(1,1)} \otimes \cdots \otimes 1_{Y_1}^{([L],[K])} \otimes \cdots \otimes 1_{Y_1}^{(\ell-1)} \otimes Z_{Y_1}^{(\ell,k)} \otimes 1_{Y_1}^{(\ell,k+1)} \cdots \otimes 1_{Y_1}^{([L],[K])} \rangle.
\]

As we discussed before, \(k\) is a local key exclusively in possession of Alice and for a given private message \(\ell\), it is chosen uniformly at random; Hence, for a given message \(\ell\), the state of Eve can be written as equation (25), as shown at the top of the next page. We would like her to learn almost nothing about the sent private message. In other words, her state becomes independent from the chosen index \(\ell\):

\[
\forall \ell, \ell' : \frac{1}{2} \|\sigma_{X^m, Y^m}^{x,m,\ell} \otimes \sigma_{Y, Y'}^{x,m,\ell'} - \tilde{\sigma}_{X^m, Y^m}^{x,m,\ell} \otimes \tilde{\sigma}_{Y, Y'}^{x,m,\ell'} \|_1 \leq \sqrt{\epsilon'},
\]

where

\[
\sigma_{X^m, Y^m}^{x,m,\ell} := \sum_x p_X(x) \langle x | x \otimes \sigma_{Y, Y'}^{x,m,(1,1)} \otimes \cdots \otimes \sigma_{Y, Y'}^{x,m,(\ell,k)} \rangle.
\]

The right-hand side of the expression above should be maximized with respect to \(c\). Since it is a down-side parabola when it comes to maximization, we pick its global maximum which occurs at \(c = \frac{\epsilon - (1 + c)(\epsilon - \delta)}{2c + c^{-1}}\). By plugging it back into the expression \(\sigma_{X^m, Y^m}^{x,m,\ell} \sigma_{Y, Y'}^{x,m,\ell'} \), we get:

\[
\log_2 |\mathcal{L}| |\mathcal{K}| = I_{\mathcal{H}}^{\ell,0}(Y; B|X)_{\sigma_{XYB}} + \log_2 \left( \frac{\epsilon - (1 + c)(\epsilon - \delta)}{2c + c^{-1}} \right).
\]

\(\sigma_{X^m, Y^m}^{x,m,\ell} \sigma_{Y, Y'}^{x,m,\ell'} \) is the marginal of \(\sigma_{X^m, Y^m}^{x,m,\ell} \) and \(P(\sigma_{X^m, Y^m}^{x,m,\ell}, \tilde{\sigma}_{Y, Y'}^{x,m,\ell'}) \leq \sqrt{\epsilon'} - \delta'\) in which \(\delta' \in (0, \sqrt{\epsilon'})\). From the invariance of trace distance with respect to tensor-product states, we can expand the security constraint (26) as given by (28), as shown at the top of the next page.

From the convex-split lemma and the definition of the conditional smooth max-mutual information (see Definition 17),

\(\sigma_{X^m, Y^m}^{x,m,\ell} \sigma_{Y, Y'}^{x,m,\ell'} \) denotes the disturbed state after Bob finds the public message, without loss of generality, we also assume Eve affects the initial state in the same way.

\footnote{Note that the state in (20) denotes the disturbed state after Bob finds the public message, without loss of generality, we also assume Eve affects the initial state in the same way.}
\[ I_{H^{-3}}(Y; B | X)_{\sigma_{XYB}} := \max_{\sigma_X^Y} \min_{x \in \text{supp}(\sigma_X^Y)} \left\{ -\log_2 \inf_{0 \leq \rho_{YB} \leq 1, \alpha(Z_{YB}^\alpha, \sigma_Y^x, \sigma_B^x) \leq \epsilon - \delta} \beta(Z_{YB}^\alpha, \sigma_Y^x, \sigma_B^x) \right\}, \quad (24) \]

where \( \sigma_{XYB} = \sum_x p_X(x | x) x^Y \otimes \sigma_Y^x \otimes \sigma_B^x \) and \( P(\sigma_X^Y, \sigma_Y^x) \leq \epsilon'. \)

\[ \sigma_{XY'}^{m,\ell}(1|K)_{E} := \frac{1}{|K|} \sum_{k=1}^{|K|} \sigma_{XY' \otimes 1|K}^{m,\ell}(k) = \sum_x p_X(x | x) x_{X'}^{Y'} \otimes \sigma_{Y'}^{x,m,\ell},(1,1) \otimes \ldots \otimes \sigma_{Y'}^{x,m,\ell}(1,|K|) \quad \sigma_{YB}^{\ell} \] \[
\otimes \left[ \frac{1}{|K|} \sum_{k=1}^{|K|} \sigma_{Y'}^{x,m,\ell}(1,1) \otimes \ldots \otimes \sigma_{Y'}^{x,m,\ell}(1,|K|) \right] \otimes \ldots \otimes \sigma_{Y'}^{x,m,\ell}(1,|K|). \quad (25) \]

\[
\frac{1}{2} \left\| \sigma_{XY' \otimes 1|K}^{m,\ell} - \sum_x p_X(x | x) x_{X'}^{Y'} \otimes \sigma_{Y'}^{x,m,\ell} \otimes \tilde{\sigma}_E \right\|_1 = \frac{1}{2} \left\| \sum_x p_X(x | x) x_{X'}^{Y'} \otimes \left( \sigma_{Y'}^{x,m,\ell} - \sigma_{Y'}^{x,m,\ell} \otimes \tilde{\sigma}_E \right) \right\|_1 \leq \frac{1}{2} \sum_x p_X(x) \frac{1}{2} \left\| \sigma_{Y'}^{x,m,\ell} - \sigma_{Y'}^{x,m,\ell} \otimes \tilde{\sigma}_E \right\|_1 \quad (28) \]

if the following condition holds,\(^9\)
\[ \log_2 |K| = \tilde{\gamma}_{\max} - \delta - \rho_{\sigma_{YB}^\alpha} + 2 \log_2 \left( \frac{1}{\delta} \right), \quad (29) \]
then
\[ P(\sigma_{XY' \otimes 1|K}^{m,\ell} | E, X) \leq \sqrt{\epsilon'}. \]

We continue by expanding \( \sigma_{XY' \otimes 1|K}^{m,\ell} = \sum_x p_X(x | x) x_{X'}^{Y'} \otimes \sigma_{Y'}^{x,m,\ell} \) as in equation (31), as shown at the top of the next page.

Reconsider the optimal test operator \( \Theta_{YB} \), we can write the following equation:
\[ \text{Tr} \{ \Theta_{YB} \sigma_{YB}^y \} = \text{Tr} \{ \Theta_{YB} \sum_y p(y | x) | y \rangle \langle y | \otimes \sigma_{B}^{x,y} \} \]
\[ = \sum_y p(y | x) \text{Tr} \{ | y \rangle \langle y | \sigma_{B}^{x,y} \} \]
\[ = \sum_y p(y | x) \text{Tr} \{ G_{x,y} \sigma_{B}^{x,y} \}, \]
where \( G_{x,y} := \langle y | \sigma_{YB}^y \rangle \). In an analogous way:
\[ \text{Tr} \{ \Theta_{YB} \sigma_{YB}^x \} = \sum_y p(y | x) \text{Tr} \{ G_{x,y} \sigma_{B}^{x,y} \}. \]

The above derivations lead the test operator to be considered as \( \Theta_{YB} = \sum_y | y \rangle \langle y | \otimes \sigma_{B}^{x,y} \), i.e., the operator classical on \( Y \) achieves the same optimal values as any general operator. Next we try to embed the test operator in the \( N_{\gamma^I_{X'Y'|K}|B} \) as given in (23), as shown at the top of the previous page, and expanded as in (32), as shown at the top of the next page. Observe the structure of the POVM given in (33), as shown at the top of the next page. And finally our POVM has the form given in equation (34), as shown at the top of the next page. To build a POVM on the full space, we add \( \Omega_B^0 = \mathbb{1} - \sum \Omega_B^{x,k} \) to the set \( \{ \Omega_B^{x,k} \}_{x=1, k=1} \). By

\(^9\)To maintain consistency, in the following expression, we show Eve’s \( X' \) and \( Y' \) systems with \( X \) and \( Y \), respectively.
\[ \sigma_{Y}^{x,m,\ell} = \sigma_{Y}^{x,m,\ell,0} \otimes \cdots \otimes \sigma_{Y}^{x,m,\ell,|K|} \]

\[ = \sum_{y_{11}, \ldots, y_{11}} p(y_{11}|x) p(y_{11}|x) \cdots \sum_{y_{11}} p(y_{11}|x) \right| y_{11} \rangle \langle y_{11} | \otimes \cdots \sum_{y_{11}} p(y_{11}|x) \right| y_{11} \rangle \langle y_{11} | \right| y_{11} \rangle \langle y_{11} | \otimes \cdots \sum_{y_{11}} p(y_{11}|x) \right| y_{11} \rangle \langle y_{11} | \right| y_{11} \rangle \langle y_{11} | \]

\[ = \sum_{y_{11}, y_{11}, \ldots, y_{11}, y_{11}} p(y_{11}|x) p(y_{11}|x) \cdots \sum_{y_{11}} p(y_{11}|x) \right| y_{11} \rangle \langle y_{11} | \otimes \cdots \sum_{y_{11}} p(y_{11}|x) \right| y_{11} \rangle \langle y_{11} | \right| y_{11} \rangle \langle y_{11} | \otimes \cdots \sum_{y_{11}} p(y_{11}|x) \right| y_{11} \rangle \langle y_{11} | \right| y_{11} \rangle \langle y_{11} | \]

\[ \text{hence} \]

\[ \sigma_{X,Y}^{m,\ell,k} = \sum_{x, y_{11}, y_{11}, \ldots, y_{11}, y_{11}} p_X(y|x) p(y_{11}|x) \cdots \sum_{y_{11}} p(y_{11}|x) \right| x \rangle \langle x | \otimes \cdots \sum_{y_{11}} p(y_{11}|x) \right| x \rangle \langle x | \otimes \cdots \sum_{y_{11}} p(y_{11}|x) \right| x \rangle \langle x | \otimes \cdots \sum_{y_{11}} p(y_{11}|x) \right| x \rangle \langle x | \right| x \rangle \langle x | \]

\[ \text{(31)} \]

\[ N_{X,Y}^{(\ell,k)}(\ell) = \left( \cdots \sum_{y_{11}, \ldots, y_{11}} p(y_{11}|x) \right| x \rangle \langle x | \otimes \cdots \sum_{y_{11}, \ldots, y_{11}} p(y_{11}|x) \right| x \rangle \langle x | \otimes \cdots \sum_{y_{11}, \ldots, y_{11}} p(y_{11}|x) \right| x \rangle \langle x | \right| x \rangle \langle x | \right| x \rangle \langle x | \]

\[ \text{(32)} \]

\[ P_{X,Y}^{(\ell,k)}(\ell) = \left( \cdots \sum_{y_{11}, \ldots, y_{11}} p(y_{11}|x) \right| x \rangle \langle x | \otimes \cdots \sum_{y_{11}, \ldots, y_{11}} p(y_{11}|x) \right| x \rangle \langle x | \otimes \cdots \sum_{y_{11}, \ldots, y_{11}} p(y_{11}|x) \right| x \rangle \langle x | \right| x \rangle \langle x | \right| x \rangle \langle x | \]

\[ \text{(33)} \]

\[ \text{combining (31) and (33), we find that} \]

\[ \text{Tr}(\mathbb{I}_{Y}^{(\ell)} \otimes |k \rangle \langle k |) = \sum_{\ell=1}^{|K|} \text{Tr}(\mathbb{I}_{Y}^{(\ell)}) \sigma_{Y}^{x,m,\ell,k} \]

\[ \text{and from (30), the equality of the average and the individual error probabilities, yields equation (35), as shown at the top of the next page.} \]

\[ \text{By taking advantage of the POVMs} \left( \Omega_{B}^{x,y,\ell,k} \right)_{\ell=1,1,k=1}^{|K|,|K|}, \]

\[ \text{the following measurement channels are defined} \]

\[ D_{B \rightarrow \ell,k}^{2}(\omega_{B}) := \sum_{\ell=1}^{|K|} \sum_{k=1}^{|K|} \text{Tr}(\Omega_{B}^{x,y,\ell,k} \otimes |\ell \rangle \langle \ell | \otimes |k \rangle \langle k |) \]

\[ \text{(36)} \]

\[ \text{averaging over} \ \ell, \ k \ \text{and} \ (x, y_{11}, \ldots, y_{11}, |K|) \text{and using (35), we get equation (38), as shown at the top of the next page.} \]
\[
\frac{1}{|\mathcal{L}|} \sum_{\ell=1}^{|\mathcal{L}|} \sum_{k=1}^{|\mathcal{K}|} \text{Tr}(I_{y|z}|k) B - P^{(\ell,k)}_{X,Y|z|k} \sigma^{m,(\ell,k)}_{X,Y|z|k} B) \]
\[
= \frac{1}{|\mathcal{L}|} \sum_{\ell=1}^{|\mathcal{L}|} \sum_{k=1}^{|\mathcal{K}|} p_{x}(x) p(y_{11}|x) \ldots p(y_{\mathcal{L}|k}|x) \text{Tr}(I_{y} - \Omega^{x,y}_{B} \sigma^{x,y}_{B} ) \leq \epsilon. \quad (35)
\]
\[
= \sum_{x,y_{11}, \ldots, y_{|\mathcal{L}|}|k} p_{x}(x) p(y_{11}|x) \ldots p(y_{|\mathcal{L}|}|k|x) \left[ \frac{1}{|\mathcal{L}|} \sum_{\ell=1}^{|\mathcal{L}|} \sum_{k=1}^{|\mathcal{K}|} \left\| D^{2}_{B \rightarrow L \hat{k}}(\sigma^{x,y}_{B}) - \frac{1}{|\mathcal{L}|} \langle \ell | L \rangle \right\|_{1} \right] \leq \epsilon. \quad (38)
\]
\[
= \sum_{x,y_{11}, \ldots, y_{|\mathcal{L}|}|k} p_{x}(x) p(y_{11}|x) \ldots p(y_{|\mathcal{L}|}|k|x) \left[ \frac{1}{|\mathcal{L}|} \sum_{\ell=1}^{|\mathcal{L}|} \sum_{k=1}^{|\mathcal{K}|} \left\| D^{2}_{B \rightarrow L \hat{k}}(\sigma^{x,y}_{B}) - \frac{1}{|\mathcal{L}|} \langle \ell | L \rangle \right\|_{1} \right] \leq \epsilon. \quad (39)
\]

In equation (38), if we take the average over \(k\) inside the trace distance and trace out \(\hat{k}\) system, by the convexity and monotonicity of the trace distance, we obtain the equations in (39), as shown at the top of this page.

Considering the POVM \(\{\Omega^{x,y}_{B} \}_{|\mathcal{L}|,|\mathcal{K}|,|l|=1,k=1}\), the probability of getting \(\ell'\) conditioned on the fact that \((\ell, k)\) was sent is equal to \(\sum_{k'=1}^{k} \text{Tr}(\Omega^{x,y}_{B} \sigma^{x,y}_{B})\) and it is clear from the uniformity of the local key that the probability of getting \(\ell'\) given that \(I\) was sent, equals \(P_{r}(\ell'|I) = \frac{1}{|\mathcal{L}|} \sum_{l=1}^{|\mathcal{L}|} \sum_{k=1}^{|\mathcal{K}|} \text{Tr}(\Omega^{x,y}_{B} \sigma^{x,y}_{B}) = \frac{1}{|\mathcal{L}|} \sum_{k'=1}^{|\mathcal{K}|} \sum_{k'=1}^{|\mathcal{K}|} \text{Tr}(\Omega^{x,y}_{B} \sigma^{x,y}_{B})\). Note that evidently \(\sum_{|\mathcal{L}|} P_{r}(\ell'|I) = 1\). If the trace above was only applied to the \(B\) system, we have:

\[
\frac{1}{|\mathcal{L}|} \sum_{\ell=1}^{|\mathcal{L}|} \sum_{k=1}^{|\mathcal{K}|} \text{Tr}(\Omega^{x,y}_{B} \sigma^{x,y}_{B}) = P_{r}(\ell'|I) u_{E}^{\ell',\ell},
\]

where

\[
u_{E}^{\ell',\ell} := \frac{1}{|\mathcal{L}|} \sum_{k=1}^{|\mathcal{K}|} \sum_{l=1}^{|\mathcal{L}|} \text{Tr}(\Omega^{x,y}_{B} \sigma^{x,y}_{B})
\]

And by summing up over all \(\ell'\) we get: (see that \(\sum_{k'=1}^{k} \text{Tr}(\Omega^{x,y}_{B} \sigma^{x,y}_{B}) = \sigma_{E}^{x,y}_{B}\) for a given pair \((\ell, k)\))

\[
\frac{1}{|\mathcal{L}|} \sum_{\ell=1}^{|\mathcal{L}|} P_{r}(\ell'|I) u_{E}^{\ell',\ell} \]

Hence, the following equation follows:

\[
D^{2}_{B \rightarrow L} \left( \frac{1}{|\mathcal{K}|} \sum_{k=1}^{|\mathcal{K}|} \sigma^{x,y}_{B} \right) = \sum_{\ell'=1}^{|\mathcal{L}|} P_{r}(\ell'|I) \langle \ell'| L \rangle \otimes u_{E}^{\ell',\ell},
\]

and by tracing out Eve’s system:

\[
D^{2}_{B \rightarrow L} \left( \frac{1}{|\mathcal{K}|} \sum_{k=1}^{|\mathcal{K}|} \sigma^{x,y}_{B} \right) = \sum_{\ell'=1}^{|\mathcal{L}|} P_{r}(\ell'|I) \langle \ell'| L \rangle.
\]

We move forward with the chain of inequalities ending up in (40), as shown at the top of the next page, where the first inequality follows from the convexity of the trace distance and the second equality emerges because of the invariance of the trace distance with respect to tensor-product states. This result together with (39) leads to the inequality given by (41), as shown at the top of the next page. This is equivalent to the criterion dealing with Bob’s error in detecting the private message. We continue by expanding Eve’s security condition as given in (42), as shown at the top of the next page, where the last equality comes about by using the invariance of trace distance with respect to tensor-product states.

We deal with two important expressions in (41) and (42), the former is Bob’s error in detecting the private message and the latter is the security of Eve. Now it is time to unify two criteria into the so-called privacy error. To this end, let’s consider (41) and (42) together with their imposed bounds on the cardinalities of \(|\mathcal{L}|\) and \(|\mathcal{K}|\). We employ triangle inequality for the trace distance to merge them into the privacy error as given in (43), as shown at the top of the next page, (remember that in the assisted code, there is no difference between average and individual error probabilities). This immediately implies the privacy criterion given in (13) in the sense that if this holds, the single criterion in (13) also holds.

We are now done with the assisted code. As we proceed to derandomize the code, it will be clear that the procedure employed to unify two error criteria is helpful. Before we proceed to derandomize the code, we would like to consider
\[
\frac{1}{|\mathcal{L}|} \sum_{l=1}^{|\mathcal{L}|} D^2_{B \rightarrow L} \left( \frac{1}{|\mathcal{K}|} \sum_{k=1}^{|\mathcal{K}|} \sigma_{BE}^{x,m,y_l,k} \right) - |\ell\rangle \langle \ell| \otimes \frac{1}{|\mathcal{K}|} \sum_{k=1}^{|\mathcal{K}|} \sigma_{E}^{x,m,y_{\ell,k}} \right) \right) \right|_1
\]

\[
= \frac{1}{|\mathcal{L}|} \sum_{\ell=1}^{|\mathcal{L}|} \left[ \frac{1}{2} \left\| |\ell\rangle \langle \ell| \otimes u_{E}^{\ell,\ell} - |\ell\rangle \langle \ell| \otimes u_{E}^{\ell,\ell} \right\|_1 \right] = \frac{1}{|\mathcal{L}|} \sum_{\ell=1}^{|\mathcal{L}|} \sum_{\ell' \neq \ell} \left[ \frac{1}{2} \left\| |\ell\rangle \langle \ell| \otimes u_{E}^{\ell,\ell'} - |\ell\rangle \langle \ell| \otimes u_{E}^{\ell,\ell'} \right\|_1 \right]
\]

\[
\leq \frac{1}{|\mathcal{L}|} \sum_{\ell=1}^{|\mathcal{L}|} \sum_{\ell' \neq \ell} \frac{1}{2} D^2_{B \rightarrow L} \left( \frac{1}{|\mathcal{K}|} \sum_{k=1}^{|\mathcal{K}|} \sigma_{BE}^{x,m,y_{\ell,k}} \right) - |\ell\rangle \langle \ell| \otimes \frac{1}{|\mathcal{K}|} \sum_{k=1}^{|\mathcal{K}|} \sigma_{E}^{x,m,y_{\ell,k}} \right) \right|_1, \quad (40)
\]

\[
\sum_{x,y_{11},...,y_{|\mathcal{L}|,|\mathcal{K}|}} p_X(x)p_Y(y_{11}|x)...p_Y(y_{|\mathcal{L}|,|\mathcal{K}|}|x) \left[ \frac{1}{|\mathcal{L}|} \sum_{\ell=1}^{|\mathcal{L}|} \frac{1}{2} D^2_{B \rightarrow L} \left( \frac{1}{|\mathcal{K}|} \sum_{k=1}^{|\mathcal{K}|} \sigma_{BE}^{x,m,y_{\ell,k}} \right) - |\ell\rangle \langle \ell| \otimes \frac{1}{|\mathcal{K}|} \sum_{k=1}^{|\mathcal{K}|} \sigma_{E}^{x,m,y_{\ell,k}} \right) \right|_1 \leq \epsilon. \quad (41)
\]

\[
\frac{1}{2} \left\| \sigma_{XY}^{m,\ell} - \sum_{x} p_X(x) |x\rangle \langle x| \otimes \sigma_{Y|\mathcal{L}|,|\mathcal{K}|}^{x,m} \otimes \sigma_{E}^{x,m} \right\|_1
\]

\[
= \frac{1}{2} \left[ \left\| \frac{1}{|\mathcal{K}|} \sum_{k=1}^{|\mathcal{K}|} p_X(x)p_Y(x|y_{11}|x)...p_Y(x|y_{|\mathcal{L}|,|\mathcal{K}|}|x) |x\rangle \langle x| \otimes |y_{11},...,y_{|\mathcal{L}|,|\mathcal{K}|}\rangle \langle y_{11},...,y_{|\mathcal{L}|,|\mathcal{K}|}| \otimes \sigma_{E}^{y_{\ell},y_{k}} - \sigma_{E}^{x,m} \right\|_1
\]

\[
= \frac{1}{2} \left[ \left\| \frac{1}{|\mathcal{L}|} \sum_{\ell=1}^{|\mathcal{L}|} \sum_{\ell' \neq \ell} \frac{1}{2} D^2_{B \rightarrow L} \left( \frac{1}{|\mathcal{K}|} \sum_{k=1}^{|\mathcal{K}|} \sigma_{BE}^{x,m,y_{\ell,k}} \right) - |\ell\rangle \langle \ell| \otimes \sigma_{E}^{x,m} \right) \right|_1 \leq \sqrt{\epsilon}, \quad (42)
\]

\[
\sum_{x,y_{11},...,y_{|\mathcal{L}|,|\mathcal{K}|}} p_X(x)p_Y(x|y_{11},x)|p_Y(x|y_{|\mathcal{L}|,|\mathcal{K}|}|x) \left[ \frac{1}{2} \left\| |\ell\rangle \langle \ell| \otimes \frac{1}{|\mathcal{K}|} \sum_{k=1}^{|\mathcal{K}|} \sigma_{E}^{x,m,y_{\ell,k}} - |\ell\rangle \langle \ell| \otimes \sigma_{E}^{x,m} \right\|_1 \right] \leq \epsilon + \sqrt{\epsilon}, \quad (43)
\]

Two extra error terms. The error probability of the second decoder depends on the error probability of the first decoder in two directions, first, the second decoder is fed a state close to the actual received state and second, the second decoder applies a quantum instrument depending on the estimate of the transmitted message \( m \). This can, without losing the generality, be written as follows:

\[
D^2_{MBY \rightarrow L} := \sum_{m} |m\rangle \langle m|_M \otimes D^2_{BY \rightarrow L}.
\]

In the following we show how this fact contributes to the error probability. First one is the difference between the received state and the disturbed state being fed into the second decoder. Since the probability of error of the first decoder is at most \( \epsilon \), we know from gentle measurement lemma that:

\[
\left\| \sigma_{XY}^{m,\ell,k} - \rho_{XY}^{m,\ell,k} \right\|_1 \leq 2\sqrt{\epsilon},
\]

and for the second term we have the chain of inequalities given by (44), as shown at the top of the next page; where the equality follows from the the observation in (10), the first and second inequalities follow the convexity and monotonicity of trace distance, respectively. Adding these two terms to (43) will result in \( P_{prf} \leq 2(\epsilon + \sqrt{\epsilon}) + \sqrt{\epsilon}. \n
3) Derandomization: We can now fix the classical registers and obtain a protocol without shared randomness, i.e.,
derandomize the code. The derandomization is a standard technique and its mathematical details are given in the appendix.

V. CONVERSE

In this section we give upper bounds for the capacity region $C_r^{e'}(N)$. Proof: [Proof of theorem (3)] Two messages $m \in \mathcal{M}$ and $\ell \in \mathcal{L}$ are sent through the channel $\mathcal{N}_{A \rightarrow BE}$ and their estimates are $\hat{M}$ and $\hat{L}$, respectively. From definition (18), an $(\epsilon, e')$-code satisfies $Pr(M \neq \hat{M}) \leq \epsilon$. A hypothesis testing problem can be associated to the problem of detecting $m$ leading to an expression for the error probability of the public message. To see how it works out, consider a binary hypothesis testing problem in which null and alternative hypothesis are $H_0: \rho = \rho_0$ and $H_1: \rho = \rho_1$. It is easily seen that type I error, i.e., deciding in favor of $H_1$ while the true state was $H_0$, is exactly equal to the error probability $Pr(M \neq \hat{M})$ which is less than or equal to $\epsilon$ by assumption. On the other hand, type II error, deciding $H_0$ on $H_1$, equals $\frac{1}{|\mathcal{M}|}$ (the distribution over message set is uniform). Then from the definition of the hypothesis testing mutual information, we have the following:

$$r \leq I_H^\epsilon(M; \hat{M})_\rho,$$

where $r = \log |\mathcal{M}|$ is the rate of the public message. Furthermore, from the quantum DPI, we have:

$$I_H^\epsilon(M; \hat{M})_\rho \leq I_H^\epsilon(M; B)_\rho.$$

Finally, using the injectivity of the encoder, we define a random variable $X$ whose distribution is built by projecting the distribution of $M$ on its image on $X$ and zero otherwise. Setting $X = M$, we get the following:

$$r \leq I_H^\epsilon(X; B)_\rho.$$

In regards to the private rate $R = \log |\mathcal{L}|$, consider the following chain of inequalities:

$$\epsilon \geq Pr\{M \neq \hat{M}\} \geq \sum_{m,\ell} p(m) p(\ell) \sum_{(\hat{m}, \hat{\ell}) \neq (m, \ell)} p(\hat{m}, \hat{\ell}|m, \ell) \geq \sum_{m,\ell} p(m) p(\ell) \sum_{(\hat{m}, \hat{\ell}) \neq (m, \ell)} p(\hat{m}|m, \ell) \geq \sum_m p(m) Pr(\hat{L} \neq L|M = m),$$

where the first line is due to the assumption. From Markov’s inequality, we know that with probability at least $1 - \sqrt{\epsilon}$, the following holds for a randomly generated $m \in \mathcal{M}$:

$$Pr(\hat{L} \neq L|M = m) \leq \sqrt{\epsilon}.$$

Then following the same strategy as for the public rate, we consider a binary hypothesis testing problem distinguishing between $\rho_0^m$ and $\rho_1^m \otimes \rho_0^\ell$ conditioned on previously specified $m$, we will have:

$$R \leq I_H^\epsilon(L; \hat{L} | M = m)_\rho.$$

Then, to get $I_H^\epsilon(L; \hat{L} | M)_\rho$, according to Definition 16, we can optimize the expression with respect to $\rho_M^\ell$ where $P(\rho_M^\ell, \rho_M^\ell) \leq \sqrt{\epsilon}$. Then from the monotonicity of the hypothesis-testing relative entropy applied to $\hat{L}$ system, we have:

$$I_H^\epsilon(L; \hat{L} | M)_\rho \leq I_H^\epsilon(L; B | M)_\rho.$$

By the same argument that we defined $X := M$, we also define $Y := L$ and so the following results:

$$R \leq I_H^\epsilon(Y; B | X)_\rho.$$

On the other hand, from the secrecy condition (5), we know that for every $m$, the following is true:

$$\frac{1}{2} \|\rho_{LE}^m - \rho_L^m \otimes \rho_E^m\|_1 \leq \epsilon',$$

and from the relation between the purified distance and the trace distance it holds that:

$$P(\rho_{LE}^m, \rho_L^m \otimes \rho_E^m) \leq \sqrt{2\epsilon'}.$$

From the definition of the smooth-max-relative entropy we see that $D_{\max}^{\epsilon'}/\sqrt{\epsilon'}(\rho_{LE}^m, \rho_L^m \otimes \rho_E^m) = 0$. And by considering the optimization in Definition 17 over $\rho_M^\ell$ such that $P(\rho_M^\ell, \rho_M^\ell) \leq \sqrt{\epsilon'}$, we have $I_{\max}^{\epsilon'}/\sqrt{\epsilon'}(L; E | M) = 0$. Setting $M := X$ and $L := Y$ as before and plugging into (45), the following bound on the private rate holds:

$$R \leq I_H^\epsilon(Y; B | X) - I_{\max}^{\epsilon'}/\sqrt{\epsilon'}(Y; E | X).$$

□
VI. ASYMPTOTIC ANALYSIS

We evaluate our rate region in the asymptotic limit of many uses of a memoryless channel. The capacity theorem for simultaneous transmission of classical and quantum information was proved by Devetak and Shor [19]. In this section, we recover their result from our theorems. We define the rate region of the simultaneous transmission of the classical and quantum information as follows:

\[ C_\infty(N) := \lim_{\epsilon, \epsilon' \to 0} \lim_{n \to \infty} \frac{1}{n} C^{\epsilon, \epsilon'}(N^{\otimes n}). \]

Let \( C(N) \) be the set of rate pairs \((r', R')\) such that

\[
\begin{align*}
  r' &\leq I(X; B)_p, \\
  R' &\leq I(Y; B|X)_p - I(Y; E|X)_p
\end{align*}
\]

where all the entropic quantities are computed over all \( \rho_{X,Y,B,E} := \sum_{x,y} P(x,y) (x|y) (y|\emptyset) (\emptyset|B) (B|A) \) arising from the channel. Then the capacity region \( C_\infty(N) \) is the (normalized) union over \( \ell \) uses of the channel \( N \) as below:

\[ C_\infty(N) = \bigcup_{\ell=1}^\infty C^{\epsilon, \epsilon'}(N^{\otimes \ell}). \quad (47) \]

In the rest of this section, our aim is to prove the capacity region above. Before doing so, we slightly modify the expression for the private rate in Theorem 2 by using Fact 5. Note that Fact 5 deals with unconditional expressions, however, conditional expressions are trivial noting their definitions. Therefore, the following holds:

\[
\begin{align*}
  I_\max^{\ell-\delta}(Y; E|X) &\leq I_\max^{\ell-\delta-\gamma}(Y; E|X) \\
  &\quad + \log_2 \left( \frac{3}{\gamma^2} \right),
\end{align*}
\]

where \( \gamma \in (0, \sqrt{\ell-\delta}) \). And so the achievability of the private rate appears as follows:

\[
R \geq I_H^{\ell-\delta}(Y; B|X) - I_\max^{\ell-\delta-\gamma}(Y; E|X) \\
- \log_2 \left( \frac{4\epsilon}{\delta^2} - 2 \log_2 \left( \frac{1}{\delta} \right) \right) - \log_2 \left( \frac{3}{\gamma^2} \right).
\]

Like all capacity theorems, the proof of the aforementioned capacity region is accomplished in two steps, direct part that we show all such rates are achievable, i.e., the right-hand side of the equation (47) is contained \((\subseteq)\) inside \( C_\infty(N) \) and the converse part that goes in the opposite direction saying that those rates cannot be exceeded, i.e., \( C_\infty(N) \) is contained inside the union on the right-hand side of (47).

For the direct part, we use our one-shot lower bounds on the capacity region and apply quantum AEP for the (conditional) smooth hypothesis testing- and max-mutual information. From Theorem 2, for \( m \) uses of the channel \( N \) (or as one may like to think of it, one use of the superchannel \( N^{\otimes m} \)), the following lower bound on the capacity region \( C^{\epsilon, \epsilon'}(N^{\otimes m}) \) can be seen:

\[
\bigcup_{\ell=1}^m C_\alpha(N^{\otimes \ell}) \subseteq C^{3\epsilon + 2\sqrt{\epsilon} + \sqrt{\epsilon^2 + 2(\epsilon + \sqrt{\epsilon}) + \sqrt{\epsilon^2}}}(N^{\otimes m}),
\]

where \( C_\alpha(N^{\otimes \ell}) \) is the set of all twins \((r', R')\) satisfying:

\[
\begin{align*}
  r' &\leq I_H^{\ell-\delta}(X'; B^{\otimes \ell}) - \log_2 \left( \frac{4\epsilon}{\delta^2} \right), \\
  R' &\leq I_H^{\ell-\delta}(Y'; B^{\otimes \ell}|X') - I_\max^{\ell-\delta-\gamma}(Y'; E^{\otimes \ell}|X') \\
  &\quad - \log_2 \left( \frac{4\epsilon}{\delta^2} \right) - 2 \log_2 \left( \frac{1}{\delta} \right) - \log_2 \left( \frac{3}{\gamma^2} \right).
\end{align*}
\]

Since the region above is basically a lower bound on the capacity region, we are free to assume that the sequences of the random variables are generated in an i.i.d. fashion according to the corresponding distributions. This empowers us to make use of quantum AEP as described below.

From Fact 4 we have

\[
\lim_{\epsilon \to 0} \lim_{m \to \infty} \frac{1}{m} I_H^{\ell-\delta}(X^m; B^{\otimes m}) = I(Y; B).
\]

Likewise, applying Lemma 1 and Lemma 2 give rise respectively to the following identities:

\[
\begin{align*}
  \lim_{\epsilon \to 0} \lim_{m \to \infty} \frac{1}{m} I_H^{\ell-\delta}(Y^m; B^{\otimes m}|X^m) &= I(Y; B|X), \\
  \lim_{\epsilon \to 0} \lim_{m \to \infty} \frac{1}{m} I_H^{\ell-\delta}(E^{\otimes m}|X^m) &= I(Y; E|X).
\end{align*}
\]

Plugging back into the respective equations, we obtain

\[
C(N) \subseteq \lim_{\epsilon, \epsilon' \to 0} \lim_{m \to \infty} \frac{1}{m} C^{\epsilon, \epsilon'}(N^{\otimes m}),
\]

where \( C(N) \), as defined before, consists of rate pairs \((r', R')\) satisfying

\[
\begin{align*}
  r' &\leq I(X; B)_p, \\
  R' &\leq I(Y; B|X)_p - I(Y; E|X)_p
\end{align*}
\]

Last step of the direct part is to consider a superchannel \( N^{\otimes \ell} \) \((\ell \) independent uses of the channel \( N \) \) and let \( n = m\ell \) and repeat the above argument, i.e., use the superchannel \( m \) times. Finally by letting \( n \to \infty \) and evaluating the union of the regions, we obtain the desired result.

To prove the converse, we consider our upper bounds given in Theorem 3 in the case of \( n \) uses of the channel \( N \) and we have:

\[
C^{\epsilon, \epsilon'}(N^{\otimes n}) \subseteq \bigcup_{n=1}^{\infty} C_\alpha(N^{\otimes n}),
\]

where \( C_\alpha(N^{\otimes n}) \) includes all ordered twins \((r', R')\) satisfying

\[
\begin{align*}
  r' &\leq I_\ell^\epsilon(X^n; B^{\otimes n}), \\
  R' &\leq I_\ell^\epsilon(Y^n; B^{\otimes n}|X^n) - I_\max^\ell(E^{\otimes n}; Y^n|X^n).
\end{align*}
\]

To upper bound right-hand side of (48) we apply Fact 1. The first term on the right-hand side of (49) can be upper bounded by making use of lemma (3) and for the second term, we use Lemma 4 replacing \(|\mathcal{H}_A|\) with \(|\gamma|^n\). The inequalities are as follows:

\[
\begin{align*}
  r' &\leq \frac{1}{1-\epsilon} (I(X^n; B^{\otimes n}) + h_b(\epsilon)), \\
  R' &\leq \frac{1}{1-\epsilon} (I(Y^n; B^{\otimes n}|X^n) + h_b(\epsilon)) - I(E^{\otimes n}; Y^n|X^n) \\
  &\quad + 2n \sqrt{2\epsilon^2} \log |\gamma| + 2(1 + \sqrt{2\epsilon}) h_b \left( \frac{\sqrt{2\epsilon^2}}{1 + \sqrt{2\epsilon}} \right).
\end{align*}
\]
Multiplying by \( \frac{1}{n} \) and taking the limits \( n \to \infty \) and \( \epsilon, \epsilon' \to 0 \), (changing \( n \) with \( \ell \)) the desired result is achieved.

A. Private Information to Coherent Information

Here we argue that the private rate that has been given in terms of the difference between two mutual-information-like quantities, is in principle, the coherent information appearing in [19]. To see how this plays out, consider an ensemble\( E = \{p_X(x), (\phi^x)^{RA} \}_{x \in \mathcal{X}} \) where \( X \) is a random variable with alphabet \( \mathcal{X} \) and distribution \( p_X(x) \) and A and R are quantum systems such that \( R \) plays the role of a reference system. Assuming an auxiliary classical system \( \sigma_X = \sum_x p_X(x)\langle x|X \rangle \), the following state can be associated to the ensemble:

\[
\sigma_{XRA} = \sum_x p_X(x)\langle x|X \rangle \otimes |\phi^x \rangle \langle \phi^x |_{RA}.
\]  

(50)

If channel \( N_{A \rightarrow BE} \) acts on this state, we get the following coherent state:

\[
N_{A \rightarrow BE}(\sigma_{XRA}) = \sum_x p_X(x)\langle x|X \rangle \otimes |\phi^x \rangle \langle \phi^x |_{RBE},
\]

and the conditional coherent information on it, is evaluated as follows:

\[
I(R;BX) \coloneqq -H(R|BX) = H(B|X) - H(R|B|X)
\]

\[
\overset{(a)}{=} H(B|X) - H(E|X),
\]

where \((a)\) follows from the fact that the state \( |\phi^x \rangle \langle \phi^x |_{RBE} \) is a pure state (conditioned on \( X \)).

We proceed with applying the Schmidt decomposition to the pure states \( |\phi^x \rangle \rangle_{RBE} \) with respect to the cut \( R|BE \). Let \( \{|y^x\rangle\} \) and \( \{|y^{x,y}\rangle\} \rangle_{BE} \) be orthonormal bases for \( R \) and \( BE \) systems. Then from Schmidt decomposition we have that

\[
|\phi^x \rangle \rangle_{RBE} = \sum_y \sqrt{p_Y|X(y|x)}|y^x\rangle \otimes |y^{x,y}\rangle \rangle_{BE}.
\]

We want to get a decoherent version of the state \( |\phi^x \rangle \rangle_{RBE} \) by measuring the \( R \) system in the basis \( \{y^x\} \rangle \). Since after the measurement, \( R \) system becomes a classical system, hereafter we show it by \( Y \). Let \( \tilde{\phi}^x \rangle \rangle_{YBE} \) denote the decoherent state resulting from the measurement, then

\[
\tilde{\phi}^x \rangle \rangle_{YBE} = \sum_y p_Y|X(y|x)|y^x\rangle \langle y^x | \otimes |y^{x,y}\rangle \langle y^{x,y} |_{BE},
\]

and let the decoherent state \( \tilde{\sigma}_{XBE} \) be as follow:

\[
\tilde{\sigma}_{XYBE} = \sum_x p_X(x)\langle x|X \rangle \otimes \sum_y p_Y|X(y|x)|y^x\rangle \langle y^x | \otimes |y^{x,y}\rangle \langle y^{x,y} |_{BE}.
\]

This state is the same as was held by Bob and Eve after decoding for the public message. If the correctness of the following equality can be proven, which turns out to be straightforward, we can argue about the correctness of our claim.

\[
I(R;BX)_{\tilde{\sigma}} = I(Y;B|X)_{\tilde{\sigma}} - I(Y;E|X)_{\tilde{\sigma}}.
\]

(51)

The right-hand side of (51) can be expanded as follow:

\[
I(Y;B|X)_{\tilde{\sigma}} = I(Y;E|X)_{\tilde{\sigma}}
\]

\[
\overset{(a)}{=} H(B|X) - H(B|X,Y) - H(E|X) + H(E|X,Y)
\]

\[
\overset{(b)}{=} H(B|X) - H(E|X),
\]

where \((a)\) follows by the definition of the conditional mutual information and \((b)\) is due to the fact that conditioned on \( X \) and \( Y \), the state on \( BE \) is a pure state. Observe the last expression is a function solely of the density operator given in (50). It is evident that for the regularized formula, we consider \( n \)-fold states in our proof instead. This proves our claim.

VII. CONCLUSION

We studied the one-shot capacity of a quantum channel for simultaneous transmission of classical and quantum information. Our main tools are position-based decoding and convex-split lemma. We first consider the problem of simultaneous transmission of public and private classical information and then we discussed that the private rate can be translated into quantum capacity. We also provided converse bounds. By evaluating our achievability and converse bounds in asymptotic \( n \)-fold regime, we recovered the well-known results in the literature.

APPENDIX A

DERANDOMIZATION OF THE CODE

We aim to derandomize the assisted code. As mentioned in the introductory section, this development follows the procedure used in [3] and [22]. We start with the public message. We saw that the optimal operator \( \Pi_{XB} \) is such that

\[
\text{Tr}[\Pi_{XB} \rho_{XB}] \geq 1 - (\epsilon - \delta) \text{ and } \text{Tr}[\Pi_{XB}(\rho_{X} \otimes \rho_{B})] = 2^{-I_{n}^{\text{opt}}(x:b)_{\mu}},
\]

we rewrite the two error types with slightly different notations as follows:

\[
\text{Tr}[\Pi_{XB} \rho_{XB}] = \text{Tr}\left\{ \Pi_{XB} \sum_{x} p_{X}(x)\langle x|X \rangle \otimes \rho_{B}^{x} \right\}
\]

\[
= \sum_{x} p_{X}(x)\text{Tr}\{\langle x| \Pi_{XB} |x \rangle \rho_{B}^{x}\}
\]

\[
= \sum_{x} p_{X}(x)\text{Tr}\{W^{x}_{B} \rho_{B}\},
\]

in which the operator \( W^{x}_{B} \) is defined as \( W^{x}_{B} := \langle x| \Pi_{XB} |x \rangle \). In an analogous way, we have that

\[
\text{Tr}[\Pi_{XB}(\rho_{X} \otimes \rho_{X})] = \text{Tr}\left\{ \Pi_{XB} \sum_{x} p_{X}(x)\langle x|X \rangle \otimes \rho_{B}\right\}
\]

\[
= \sum_{x} p_{X}(x)\text{Tr}\{\langle x| \Pi_{XB} |x \rangle \rho_{B}\}
\]

\[
= \sum_{x} p_{X}(x)\text{Tr}\{W^{x}_{B} \rho_{B}\}.
\]

These expressions imply that it is sufficient to take the optimal test to be \( \Pi_{XB} = \sum_{x} \langle x|\otimes W^{x}_{B} \) with aforementioned \( W^{x}_{B} \). In other words, the test \( \Pi_{XB} \) can achieve the same error probability as any other \( \Pi_{XB} \) would do. We proceed with dissecting
that we just introduced, the operator assuming the particular structure for the optimal test operator \( \sum_{x_1, ..., x_{|\mathcal{M}|}} p_X(x_1) ... p_X(x_{|\mathcal{M}|}) |x_{|\mathcal{M}|} \rangle \langle x_{|\mathcal{M}|} | x_1, ..., x_{|\mathcal{M}|} \otimes \rho_{X_B}^{m_{(l,k)}}. \) (52)

\[
\Gamma_{X_{|\mathcal{M}|}}^{m_{(l,k)}} = \mathbb{I}_X \otimes \ldots \otimes T_{X_B}^{m_{(l,k)}} \otimes \ldots \otimes \mathbb{I}_X = \sum_{x_1} |x_1 \rangle \langle x_1 | x \otimes \ldots \otimes \left( \sum_{x_m} |x_m \rangle \langle x_m | W_B^{x_m} \right) \otimes \ldots \otimes \sum_{x_{|\mathcal{M}|}} |x_{|\mathcal{M}|} \rangle \langle x_{|\mathcal{M}|} | x \otimes W_B^{x_{|\mathcal{M}|}}. \] (53)

It turns out that each term involved in \( \text{Tr}((\mathbb{I}_X X_{|\mathcal{M}|}^{m_{(l,k)}} - \Lambda_{X_{|\mathcal{M}|}}^{m_{(l,k)}}) \rho_{X_{|\mathcal{M}|}}^{m_{(l,k)}}) \) where \( \rho_{X_{|\mathcal{M}|}}^{m_{(l,k)}} \) is given in (52), as shown at the top of this page. By assuming the particular structure for the optimal test operator that we just introduced, the operator \( \Gamma_{X_{|\mathcal{M}|}}^{m_{(l,k)}} \) appears as given in (53), as shown at the top of this page. And

\[
\left( \sum_{m' = 1}^{M_{|\mathcal{M}|}} \Gamma_{X_{|\mathcal{M}|}}^{m'} \right)^{-\frac{1}{2}} = \left( \sum_{m' = 1}^{M_{|\mathcal{M}|}} \sum_{x_1, ..., x_{|\mathcal{M}|}} |x_1 \rangle \langle x_1 | x \otimes W_B^{x_{|\mathcal{M}|}} \right)^{-\frac{1}{2}} = \left( \sum_{x_1, ..., x_{|\mathcal{M}|}} |x_1 \rangle \langle x_1 | x \otimes \left( \sum_{m' = 1}^{M_{|\mathcal{M}|}} W_B^{x_{|\mathcal{M}|}} \right)^{-\frac{1}{2}} \right)^{-\frac{1}{2}} = \frac{1}{|\mathcal{M}|} \sum_{x_1, ..., x_{|\mathcal{M}|}} p_X(x_1) \ldots p_X(x_{|\mathcal{M}|}) \cdot \text{Tr}((\mathbb{I}_B - \Delta_B^{m_{(l,k)}}) \rho_B^{x_{|\mathcal{M}|}}) \leq \epsilon, \] (54)

\[\text{This conclusion is known as the Shannon trick. The sequence } \{x_1, ..., x_{|\mathcal{M}|}\} \text{ serves as the codebook used to transmit the public message.} \]

\[\text{As for the second part, we take (43), and average over all private messages as given in (55), as shown at the top of the next page. And we again employ Shannon trick to conclude that there exists at least one sequence of values } \{y_1, ..., y_{|\mathcal{L}|_{|\mathcal{K}|}} \} \text{ such that equation (56), as shown at the top of the next page, holds.}\]

\[\text{We can now argue that there exist values } (x_1, ..., x_{|\mathcal{M}|}) \text{ serving as public codebook for the transmission of the public message and conditioned on a particular codeword of the public codebook, there exist values } (y_1, ..., y_{|\mathcal{L}|_{|\mathcal{K}|}}) \text{ serving as private codebook ensuring that the privacy criterion holds. Now we have a codebook of size } |\mathcal{M}| |\mathcal{L}| |\mathcal{K}|, (x_1, ..., x_{|\mathcal{M}|}, y_1, ..., y_{|\mathcal{L}|_{|\mathcal{K}|}}), \text{ that is publicly available serving as our deterministic codebook for simultaneous transmission of public and private messages.}\]

\[\text{APPENDIX B}\]

\[\text{ONE-SHOT QUANTUM CAPACITY: IMITATING DEVetak’S ASYMPTOTIC PROOF}\]

\[\text{As we mentioned in the introduction, a one-shot version of Devetak’s asymptotic proof of quantum capacity follows along}\]
...become a basis can take a classical protocol working on letters and turn it into The last step is to use the local decodings, which exist by the general idea. We shall freely use the notation introduced so far. We have now seen that there exists a good codebook \( \{x(\ell, k)\}_{\ell, k} \) selected from a distribution \( p(x) \) and a corresponding POVM \( \{\Omega_B^{(\ell, k)}\} \) such that once Alice transmits a state corresponding to \( x(\ell, k) \) over the channel \( N_{A \rightarrow BE} \), Bob is able to reliably work out both Alice’s message \( \ell \) and the local key \( k \) and at the same time, Eve happens to learn very little about Alice’s message \( \ell \). This holds true for the rates of the private message \( \log_2 |\mathcal{L}| \) and the local randomness \( \log_2 |\mathcal{K}| \) as are specified. Now a quantum code can be obtained by a “making coherent” of this code (see [50] for coherifying general protocols).

The first idea of making protocols coherent is that classical words/letters \( x \) become basis states \( |x\rangle \) of the Hilbert space. Functions \( f : x \rightarrow f(x) \) thus induce linear operators on Hilbert space, but only permutations (one-to-one functions) are really interesting, since they give rise to unitaries (isometries, resp.). The second idea is thus to make classical computations first reversible, by extending them to one-to-one functions. The last step is to use the local decodings, which exist by the classical theorem. In summary, “making coherent” means we can take a classical protocol working on letters and turn it into a bunch of unitaries acting as permutations on the basis states, and that we can run perfectly well on superpositions.

From the recipe outlined above, Alice’s messages \( \ell \in \mathcal{L} \) become a basis \( \{|\ell\rangle\}_{\ell \in \mathcal{L}} \) of the Hilbert space. Suppose that Alice shares a state \( |\phi\rangle_{RA_1} \) with a reference system \( R \):

\[
|\phi\rangle_{RA_1} := \sum_{\ell, k \in \mathcal{L}} a_{\ell, k} |\ell\rangle_R |\ell\rangle_{A_1},
\]

where \( |\ell\rangle_R \) and \( |\ell\rangle_{A_1} \) are some orthonormal bases for \( R \) and \( A_1 \), respectively. A number of different information-processing tasks can be considered as quantum communications. The strongest definition of quantum capacity, however, corresponds to a task known as entanglement transmission. According to this task, Alice aims to transfer her share of entanglement with a reference system to Bob with Alice no longer entangled with the reference, i.e., the following state:

\[
|\phi\rangle_{RB} := \sum_{\ell, k \in \mathcal{L}} a_{\ell, k} |\ell\rangle_R |\ell\rangle_B,
\]

where now Bob holds the \( B \) system. Suppose there is an ensemble of quantum states \( \{p(x), |y^x\rangle_A\} \). Alice uses her classical code to create a quantum codebook whose codewords are as follows:

\[
|\phi^\ell\rangle_A := \frac{1}{\sqrt{|\mathcal{K}|}} \sum_{k \in \mathcal{K}} a^{\ell, k} |y^x(\ell, k)\rangle_A,
\]

where the states \( |y^x(\ell, k)\rangle_A \) are from the aforementioned ensemble and \( x(\ell, k) \) belong to the (classical) private codebook. Alice’s action would be to coherently copy the value of \( \ell \) in register \( A_1 \) to another register \( A_2 \). She then applies some isometric encoding from \( A_2 \) register to \( A \). Two steps are performed with the following map:

\[
\left( \sum_\ell |\ell\rangle_{A_1} \otimes |\phi^\ell\rangle_A \right) \left( \sum_\ell |\ell\rangle_{A_2} \otimes |\phi^\ell\rangle_{A_1} \right),
\]

Alice then transmits the codeword over the channel giving rise to the following state:

\[
\sum_{\ell, k} a_{\ell, k} |\ell\rangle_R |\ell\rangle_{A_1} |\phi^\ell\rangle_{BE}.
\]

From the classical protocol we know that Bob can detect both the message \( \ell \) and the local key \( k \) with high probability. Bob constructs a coherent version of his POVM as follows:

\[
\sum_{\ell, k} \sqrt{\Omega_B^{(\ell, k)}} \otimes |\ell\rangle_{B_1} |k\rangle_{B_2}.
\]

From gentle measurement lemma, the state after Bob’s decoding will be close to the following state:

\[
\sum_{\ell, k} \frac{1}{\sqrt{|\mathcal{K}|}} \sum_{i, \ell, k} a_{i, \ell} |\ell\rangle_R |\ell\rangle_{A_1} e^{\gamma_{\ell, k}} |\phi^x(\ell, k)\rangle_{BE} |\ell\rangle_{B_1} |k\rangle_{B_2}.
\]

On the other hand, from secrecy requirement, it can be seen that there exists some isometry on Bob’s \( B \) and \( B_2 \) systems such that after its application, Eve’s system will be decoupled from the rest and the following state will result:

\[
\sum_{\ell, k} a_{i, \ell} |\ell\rangle_R |\ell\rangle_{A_1} |\ell\rangle_{B_1}.
\]

So far they have successfully implemented an approximate coherent channel from systems \( A_1 \) to \( A_1 B_1 \). Alice is allowed...
to use a forward classical channel to communicate with Bob in order to turn the above coherent channel to a quantum channel. Alice’s strategy is to first perform a Fourier transform on the register $A_1$ then measure the register in the computational basis and communicate the classical output to Bob. Bob will perform a controlled unitary based on the classical letter he received and the desired state will be achieved. Note that it can be shown that there exists a scheme that does not require the use of this forward classical channel.

ACKNOWLEDGMENT

The first author is indebted to Andreas Winter for his constant advice; In regards to this work, the part on making protocols coherent is due to his private communications with Prof. Winter. He also sincerely thanks the centre for quantum software and information at the University of Technology Sydney for their hospitality and support while part of this work was done. The authors are grateful to M. M. Wilde for his comments on the first version of this work. The paper benefited considerably from the valuable comments by the anonymous reviewers.

REFERENCES

[1] L. Wang and R. Renner, “One-shot classical-quantum capacity and hypothesis testing,” Phys. Rev. Lett., vol. 108, May 2012, Art. no. 200501.
[2] A. Anshu, R. Jain, and N. A. Warsi, “One shot entanglement assisted classical and quantum communication over noisy quantum channels: A hypothesis testing and convex split approach,” Nov. 2017, arXiv:1702.01940. [Online]. Available: https://arxiv.org/abs/1702.01940
[3] M. Wilde, “Coherence-based coding and convex splitting for private communication over quantum channels,” Quantum Inf. Process., vol. 16, p. 264, Oct. 2017.
[4] H. Umegaki, “Conditional expectation in an operator algebra, IV (entropy and information),” Kodai Math. Seminar Rep., vol. 14, no. 2, pp. 59–85, 1962.
[5] H. Nagaoka and M. Hayashi, “An information-spectrum approach to classical and quantum hypothesis testing for simple hypotheses,” IEEE Trans. Inf. Theory, vol. 53, no. 2, pp. 534–549, Feb. 2007.
[6] A. Winter, “Coding theorem and strong converse for quantum channels,” IEEE Trans. Inf. Theory, vol. 45, no. 7, pp. 2481–2485, Nov. 1999.
[7] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, “General paradigm for distilling classical key from quantum states,” IEEE Trans. Inf. Theory, vol. 55, no. 4, pp. 1898–1929, Apr. 2009.
[8] M. M. Wilde, M. Tomamichel, and M. Berta, “Converse bounds for private communication over quantum channels,” IEEE Trans. Inf. Theory, vol. 63, no. 3, pp. 1792–1817, Mar. 2017.
[9] T. Ogawa and H. Nagaoka, “Making good codes for classical-quantum channel coding via quantum hypothesis testing,” IEEE Trans. Inf. Theory, vol. 53, no. 6, pp. 2261–2266, Jun. 2007.
[10] M. Berta, M. Christandl, and R. Renner, “The quantum reverse Shannon theorem based on one-shot information theory,” Commun. Math. Phys., vol. 306, no. 3, pp. 579–615, Sep. 2011.
[11] N. Ciganović, N. J. Beaudry, and R. Renner, “Smooth max-information as one-shot generalization for mutual information,” IEEE Trans. Inf. Theory, vol. 53, no. 3, pp. 1573–1581, Mar. 2014.
[12] M. M. Wilde, Quantum Information Theory, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, Feb. 2017.
[13] K. Li, “Second-order asymptotics for quantum hypothesis testing,” Ann. Statist., vol. 42, no. 1, pp. 171–189, 2014.
[14] M. Tomamichel and M. Hayashi, “A hierarchy of information quantities for finitely block length analysis of quantum tasks,” IEEE Trans. Inf. Theory, vol. 59, no. 11, pp. 7693–7710, Nov. 2013.
[15] F. Buscemi and N. Datta, “The quantum capacity of channels with arbitrarily correlated noise,” IEEE Trans. Inf. Theory, vol. 56, no. 3, pp. 1447–1460, Mar. 2010.
[16] F. Salek, A. Anshu, M.-H. Hsieh, R. Jain, and J. R. Fonollosa, “One-shot capacity bounds on the simultaneous transmission of public and private information over quantum channels,” in Proc. IEEE Int. Symp. Inf. Theory, Vail, CO, USA, Jun. 2018, pp. 296–300.
Farzin Salek is a doctoral researcher at the Quantum Information Group (GIQ), Department of Physics at the Universitat Autònoma de Barcelona, Barcelona, Spain and at the Department of Signal Theory and Communications, Universitat Politècnica de Catalunya, Barcelona, Spain. He received his M.Sc. degree from Tarbiat Modares University, Tehran, Iran in electrical engineering-communication systems in 2014. He was a visiting scholar at the department of information and communication Technologies, Pompeu Fabra University, Barcelona, Spain in 2015. During the summer of 2018, he was an honorary appointee with the Centre for Quantum Software and Information (QSI) at University of Technology Sydney (UTS), Sydney, Australia. His research interests include classical and quantum information theory and quantum machine learning.

Anurag Anshu is a joint postdoctoral researcher at the Institute for Quantum Computing, University of Waterloo and the Perimeter Institute for Theoretical Physics. He finished his PhD from Centre for Quantum Technologies, NUS, Singapore. He is interested in quantum information and quantum complexity.

Min-Hsiu Hsieh (M’09–SM’16) received his PhD degree in electrical engineering from the University of Southern California, Los Angeles, in 2008. From 2008-2010, he was a Researcher at the ERATO-SORST Quantum Computation and Information Project, Japan Science and Technology Agency, Tokyo, Japan. From 2010-2012, he was a Postdoctoral Researcher at the Statistical Laboratory, the Centre for Mathematical Sciences, the University of Cambridge, UK. He is now an Australian Research Council (ARC) Future Fellow and an Associate Professor at the Centre for Quantum Software and Information (UTS-QSI), Faculty of Engineering and Information Technology (FEIT), University of Technology Sydney (UTS). His scientific interests include quantum Shannon theory, entanglement theory, and quantum coding theory.

Rahul Jain obtained his Ph.D. degree in computer science from the Tata Institute of Fundamental Research, Mumbai, India, in 2003. He completed two postdoctoral fellowships: two years at the University of California, Berkeley, CA, USA, followed by two years at the Institute for Quantum Computing at the University of Waterloo, Waterloo, ON, Canada. He joined Centre for Quantum Technologies (CQT), Singapore, as a Principal Investigator and the National University of Singapore (NUS), Singapore, as an Assistant Professor in 2008. He is presently an Associate Professor (starting July 2013) at NUS and Principal Investigator at CQT. His research interests are in the areas of information theory, quantum computation, cryptography, communication complexity, and computational complexity theory.

Javier Rodríguez Fonollosa (S’90–M’92–SM’98) received the Ph.D. degree in electrical and computer engineering at Northeastern University, Boston, MA, in 1992. In 1993 he joined the Department of Signal Theory and Communications of UPC where he became Associate Professor in 1996, Professor in 2003 and Department Head from 2006 until 2010. In 1995 he led UPC’s participation in the European Commission funded ACTS Mobile projects TSUNAMI (II) and SUNBEAM that included the analysis of adaptive antennas in 2G and 3rd generation cellular mobile communication systems. Since January 2000 until 2003 he was technical and project coordinator of the IST projects METRA and IMETRA dedicated to the introduction of multi-antenna terminals in UMTS and Systems beyond 3G. Since January 2006 to December 2008 he coordinated the Sixth Framework Programme IST project SURFACE which evaluated the performance of a generalized air interface with self-configuration capabilities. In November 2006 he initiated the coordination of the 5 year Type C-Consolider project Fundamental bounds in Network Information Theory of the National Research Plan of Spain. Since December 2008 to December 2014 he was the Coordinator of the CONSOLIDER-INGENIO 2010 Foundations and Methodologies for Future Communication and Sensor Networks (COMONSENS), a 6 year, 3.5 Million effort of 135 researchers belonging to 10 universities and research centers in Spain. The project continued under his coordination as a research network Red COMONSENS for three years starting December 2015. He was department head of the Signal Theory and Communications department of UPC from October 2006 to January 2010. In 2009 he coordinated the ERASMUS MUNDUS 2009-2013 Master of Science in Research on Information and Communication Technologies (MERIT) EMMC Joint Master Program. Since May 2005 he is member of the Editorial Board of the EURASIP Signal Processing Journal. In June 1995 and September 2001 he was co-chairman and organizer of the IEEE Signal Processing/ATHOS Workshop on Higher-Order Statistics held in Begur, Girona, Spain and of the IST Mobile Communications Summit 2001 held in Sitges, Barcelona, Spain. He was elected IEEE Senior Member and member of the Signal Processing for Communications (SPCOM) Technical Committee of the IEEE Signal Processing Society in February 1998 and January 1999 respectively. Since February 2010 until July 2014 he was manager of the Communications and Electronic Technologies (TEC) area of the National Research Plan of Spain.