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,
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 $\mathcal{W}_{X\to Y}$ 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 $\mathcal{N}_{A\to B}$, 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, $C(\mathcal{W})$, is the classical capacity of a quantum channel, $C(\mathcal{N})$, 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 capacity[1] was independently studied in [24] and [25] where an achievability bound, i.e., $C(\mathcal{N}) \geq \chi(\mathcal{N})$, known as HSW theorem was reported, where $\mathcal{X}(\mathcal{N})$ is the celebrated Holevo Information [26] defined as follows:

$$\chi(\mathcal{N}) := \max_{p(x),\rho} I(X;B)_\rho,$$

where $p(x)$ is a probability distribution, $\rho_{XB} = \sum_x p_X(x)|x\rangle\langle x|_X \otimes \mathcal{N}_{A\to B}(\rho_A^x)$ 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

This work was presented in part in ISIT 2018, see [16].

F. Salek is jointly affiliated with the Quantum Information Group (GIQ), Department of Physics at the Universitat Autònoma de Barcelona, Barcelona, Spain and with the Department of Signal Theory and Communications, Universitat Politècnica de Catalunya, Barcelona, Spain, e-mail: (farzin.salek@gmail.com).

J. R. Fonollosa is with the Department of Signal Theory and Communications, Universitat Politècnica de Catalunya, Barcelona, Spain, e-mail: (farzin.salek@upc.edu, javier.fonollosa@upc.edu).

Min-Hsiu Hsieh is with the Center for Quantum Software and Information, Sydney University of Technology, Sydney, Australia, e-mail: (min-hsiu.hsieh@uts.edu.au).

Anurag Anshu is with the Institute for Quantum Computing, University of Waterloo and the Perimeter Institute for Theoretical Physics, Canada, e-mail: (a0109169@u.nus.edu).

Rahul Jain is with Centre for Quantum Technologies, National University of Singapore; MajuLab, UMI 3654, Singapore and is VAJRA Adjunct Faculty, TIFR, Mumbai, India, e-mail: (rahul@comp.nus.edu.sg).

[1] hereafter, we talk about quantum channels unless otherwise specified, hence we drop the term quantum.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2019.2945800, IEEE Transactions on Information Theory

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:

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

where $\rho_{XBE} = \sum_x p_X(x)|x\rangle\langle x|_X \otimes \mathcal{U}^{\mathcal{N}}_{A\to BE}(\rho^x_A)$ and $\mathcal{U}^{\mathcal{N}}_{A\to BE}$ is an isometric extension of the channel $\mathcal{N}_{A\to B}$. They also showed that the private capacity equals the regularized form of $\mathcal{P}(\mathcal{N})$ 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 $Q_{reg}(\mathcal{N})$. 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:

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

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

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 $\mathcal{N}$ for simultaneous transmission of classical and quantum information is as follows:

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

where $S(\mathcal{N})$ is the union, over all states $\rho_{XRB} = \sum_x p(x)\,|x\rangle\langle x| \otimes \rho^x_{BR}$ arising from the channel $\mathcal{N}_{A\to B}$, i.e., for $x \in \mathrm{supp}(p(x))$, $\rho^x_{RB} = \mathcal{N}(\phi^x_{RA})$ for pure states $|\phi^x\rangle_{RA}$, of the (r, R) pairs obeying

$$\begin{aligned} 0 &\le r \le I(X;B)_\rho, \\ 0 &\le R \le I(R\rangle BX)_\rho. \end{aligned}$$

where r and R are the rates of the classical and quantum[2] 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 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.

In this paper, 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 classcial communication [27]. The basic intuition underlying the quantum capacity is the no-cloning theorem

[2] It is the same for various information-processing tasks: subspace transmission, entanglement transmission or entanglement generation

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2019.2945800, IEEE Transactions on Information Theory

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):

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

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 $\mathcal{N}_{A\to B}$ to Bob. This induces the following state on Bob's side :

$$\rho_{X^{|\mathcal{M}|}B}^{m}=\rho_{X}^{1}\otimes...\otimes\rho_{XB}^{m}\otimes...\otimes\rho_{X}^{|\mathcal{M}|}.$$

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 ρ_{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:

$$\frac{1}{|\mathcal{M}|}\sum_{m=1}^{|\mathcal{M}|}\rho_{X}^{1}\otimes...\otimes\rho_{XB}^{m}\otimes...\otimes\rho_{X}^{|\mathcal{M}|},$$

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

$$\rho_{X}^{1}\otimes...\otimes\rho_{X}^{m}\otimes...\otimes\rho_{X}^{|\mathcal{M}|}\otimes\rho_B,$$

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 public, private 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 non-negative operators by $\mathcal{P}(\mathcal{H}_A)$. A state of system A is 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 $\mathbb{1}_A$. 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 ρ_A. The support of an operator ρ, supp(ρ), is defined to be the subspace orthogonal to its kernel. If the support of ρ in contained in that of σ, 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 $[i:j]\coloneqq\{i,i+1,...,j\}$. For Hermitain 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 ρ_A (null hypothesis) and σ_A (alternative hypothesis) where $\rho_A,\sigma_A\in\mathcal{D}(\mathcal{H}_A)$. The task is

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2019.2945800, IEEE Transactions on Information Theory

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 ρ_A (resp. σ_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 ρ_A but σ_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:

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

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]):

$$D_H^\epsilon(\rho_A \| \sigma_A) := -\log_2 \inf_{\substack{0 \leq T_A \leq \mathbb{1}, \\ \alpha(T_A, \rho_A) \leq \epsilon}} \beta(T_A, \sigma_A).$$

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 ρ_A and σ_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 ρ_A and σ_A are prepared with equal probabilities. It can be easily shown that the optimal success probability in distinguishing the states equals $\frac{1}{2}(1 + \max_{0 \leq T \leq \mathbb{1}} \mathrm{Tr}\{T_A(\rho_A - \sigma_A)\})$. The optimization problem is evaluated as follows:

$$\begin{aligned}\max_{0 \leq T_A \leq \mathbb{1}} \mathrm{Tr}\{T_A(\rho_A - \sigma_A)\} &= [\{\rho_A - \sigma_A\}_+(\rho_A - \sigma_A)] \\ &\quad - [\{\rho_A - \sigma_A\}_-(\rho_A - \sigma_A)] \\ &:= \frac{1}{2}\|\rho_A - \sigma_A\|_1,\end{aligned}$$

where $\{\rho_A - \sigma_A\}_+$ denotes the projector onto the subspace where the operator $(\rho_A - \sigma_A)$ is non-negative, and $\{\rho_A - \sigma_A\}_- = \mathbb{1} - \{\rho_A - \sigma_A\}_+$[3]. 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 ρ_A, σ_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 $\{p(x), \rho_A^x\}$ and $\{p(x), \sigma_A^x\}$, where $\rho_A^x, \sigma_A^x \in \mathcal{D}(\mathcal{H}_A)$ for all x, let $\rho_{XA} := \sum_x p(x)\,|x\rangle\langle x| \otimes \rho_A^x$ and $\sigma_{XA} := \sum_x p(x)\,|x\rangle\langle x| \otimes \sigma_A^x$ be the associated classical-quantum (CQ) states, respectively. Then,
$$\left\|\sum_x p(x)\rho_A^x - \sum_x p(x)\sigma_A^x\right\|_1 \leq \sum_x p(x)\|\rho_A^x - \sigma_A^x\|_1.$$
Moreover, the following property can be easily checked:
$$\|\rho_{XA} - \sigma_{XA}\|_1 = \sum_x p(x)\|\rho_A^x - \sigma_A^x\|_1.$$
- Trace distance is monotone non-increasing with respect to CPTP maps. That is, for quantum states ρ and σ 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 ρ, σ and τ, 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 ρ, σ and τ, 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:

$$F(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1.$$

Definition 4 (Purified Distance [44], [38]): Let $\rho_A, \sigma_A \in \mathcal{D}(\mathcal{H}_A)$. The purified distance between ρ_A and σ_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 ϵ-ball around $\rho_A \in \mathcal{D}(\mathcal{H}_A)$, that is $\mathcal{B}^\epsilon(\rho_A) := \{\rho'_A \in \mathcal{D}(\mathcal{H}_A) : P(\rho'_A, \rho_A) \leq \epsilon\}$.

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]:

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

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:

$$H(A|B)_\rho := H(AB)_\rho - H(B)_\rho,$$

where

$$H(A)_\rho := -\mathrm{Tr}\{\rho_A \log_2 \rho_A\},$$

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

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

[3] In general, $\{\omega\}_+$ denotes the projector onto the positive eigenspace of ω and $\{\omega\}_- = \mathbb{1} - \{\omega\}_+$.

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:

$$I(A\rangle B)_\rho := H(B)_\rho - H(AB)_\rho.$$

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

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

$$D(\rho_A\|\sigma_A) := \mathrm{Tr}\{\rho_A[\log_2 \rho_A - \log_2 \sigma_A]\},$$

whenever $\mathrm{supp}(\rho_A) \subseteq \mathrm{supp}(\sigma_A)$ and otherwise it equals $+\infty$.

Fact 1 (Relation between the quantum relative entropy and the hypothesis testing relative entropy [1]): For all state ρ_A and σ_A and $\epsilon \in [0,1)$, the following inequality holds

$$D^\epsilon_H(\rho_A\|\sigma_A) \leq \frac{1}{1-\epsilon}\left[D(\rho_A\|\sigma_A) + h_b(\epsilon)\right],$$

where $h_b(\epsilon) := -\epsilon \log_2 \epsilon - (1-\epsilon)\log_2(1-\epsilon)$ 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:

$$V(\rho_A\|\sigma_A) := \mathrm{Tr}\{\rho_A[\log_2 \rho_A - \log_2 \sigma_A - D(\rho_A\|\sigma_A)]^2\},$$

whenever $\mathrm{supp}(\rho_A) \subseteq \mathrm{supp}(\sigma_A)$ 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:

$$D_{max}(\rho_A\|\sigma_A) := \inf\left\{\lambda \in \mathbb{R} : \rho_A \leq 2^\lambda \sigma_A\right\}, \qquad (2)$$

where it is well-defined if $\mathrm{supp}(\rho_A) \subseteq \mathrm{supp}(\sigma_A)$.

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 $\mathcal{E} : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B)$, it holds that

$$D_{max}(\mathcal{E}(\rho_A)\|\mathcal{E}(\sigma_A)) \leq D_{max}(\rho_A\|\sigma_A).$$

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

$$D(\rho_A\|\sigma_A) \leq D_{max}(\rho_A\|\sigma_A).$$

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:

$$D^\epsilon_{max}(\rho_A\|\sigma_A) := \inf_{\rho'_A \in \mathcal{B}^\epsilon(\rho_A)} D_{max}(\rho'_A\|\sigma_A).$$

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

$$\begin{aligned} D^\epsilon_H\left(\rho_A^{\otimes n}\|\sigma_A^{\otimes n}\right) &= nD\left(\rho_A\|\sigma_A\right) \\ &+ \sqrt{nV\left(\rho_A\|\sigma_A\right)}\Phi^{-1}(\epsilon) + O(\log n), \end{aligned}$$

$$\begin{aligned} D^\epsilon_{max}\left(\rho_A^{\otimes n}\|\sigma_A^{\otimes n}\right) &= nD\left(\rho_A\|\sigma_A\right) \\ &- \sqrt{nV\left(\rho_A\|\sigma_A\right)}\Phi^{-1}(\epsilon^2) + O(\log n), \end{aligned}$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^x exp(-\frac{x^2}{2})dx$ 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 δ-*typical set* of sequences that satisfy

$$\left|\frac{1}{n}N(x_i|x^n) - p(x_i)\right| \leq \delta,$$

where $N(x_i|x^n)$ 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 δ-typical subspace is a subspace of the full Hilbert space $\mathcal{H}_{X_1} \otimes ... \otimes \mathcal{H}_{X_n}$, 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 δ-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 $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ 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 1), the hypothesis testing-mutual information is defined as follows:

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

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:

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

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:

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

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:

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

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 ρ_{AB}, it holds that:

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

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

$$I^{\epsilon}_{H}(A;B|X)_{\rho} := \max_{\rho'} \min_{x\in\text{supp}(\rho'_X)} I^{\epsilon}_{H}(A;B)_{\rho^x_{AB}},$$

where maximization is over all $\rho'_X = \sum_x p'_X(x)|x\rangle\langle x|_X$ satisfying $P(\rho'_X,\rho_X)\leq\epsilon$.

Definition 17 (Conditional smooth max-mutual information[4]): Let $\rho_{ABX} := \sum_x p_X(x)|x\rangle\langle x|_X\otimes\rho^x_{AB}$ be a CQ state and $\epsilon\in[0,1)$. The conditional smooth max-mutual information is defined as follows:

$$I^{\epsilon}_{max}(A;B|X)_{\rho} := \min_{\rho'} \max_{x\in\text{supp}(\rho'_X)} I^{\epsilon}_{max}(A;B)_{\rho^x_{AB}},$$

where minimization is over all $\rho'_X = \sum_x p'_X(x)|x\rangle\langle x|_X$ satisfying $P(\rho'_X,\rho_X)\leq\epsilon$.

Lemma 1: Let $\rho_{XAB} = \sum_x p(x)|x\rangle\langle x|_X\otimes\rho^x_{AB}$. Then the following holds:

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

Proof: The following is easily seen from the definition,

$$\begin{aligned}&I^{\epsilon}_{H}(A^{\otimes n};B^{\otimes n}|X^n)_{\rho^{\otimes n}}\\ &\quad:= \max_{\rho'_{X^n}} \min_{x^n\in supp(\rho'_{X^n})} D^{\epsilon}_{H}(\rho^{x^n}_{A^nB^n}\|\rho^{x^n}_{A^n}\otimes\rho^{x^n}_{B^n}).\end{aligned}$$

In order to be able to apply the asymptotic results given in Fact 4, we first produce ρ'_{X^n} by projecting $\rho^{\otimes n}_X$ 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 $\rho^{x^n}_{A^nB^n}$ is in fact a tensor-product state that can be written as $\rho^{x(1)}_{AB}\otimes...\otimes\rho^{x(i)}_{AB}\otimes...\otimes\rho^{x(n)}_{AB}$ 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 $\delta\geq 0$, as $n\to\infty$, by using Fact 4 for each chosen sequence, the multi-letter formula above can be written as shown by (3) where x_i, $i\in[1:|\mathcal{X}|]$ denotes an element of the alphabet $\mathcal{X}$ and the second equality follows from Fact 4 and the fully quantum AEP [14]. □

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

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

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 $\rho_{XAB} = \sum_x p(x)|x\rangle\langle x|_X\otimes\rho^x_{AB}$, the following inequality is true.

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

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

$$\min_x D^{\epsilon}_{H}(\rho^x_{AB}\|\rho^x_A\otimes\rho^x_B) \leq \sum_x p(x)D^{\epsilon}_{H}(\rho^x_{AB}\|\rho^x_A\otimes\rho^x_B),$$

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

$$D^{\epsilon}_{H}(\rho^x_{AB}\|\rho^x_A\otimes\rho^x_B) \leq \frac{1}{1-\epsilon}\left(D(\rho^x_{AB}\|\rho^x_A\otimes\rho^x_B)+h_b(\epsilon)\right),$$

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

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

$$\begin{aligned}I^{\epsilon}_{max}(A;B|X)_{\rho} \geq\;& I(A;B|X)_{\rho}\\ &-2\epsilon\log|\mathcal{H}_A| - 2(1+\epsilon)h_b(\frac{\epsilon}{1+\epsilon}).\end{aligned}$$

Proof: In the the following simple inequality:

$$\max_x D^{\epsilon}_{max}(\rho^x_{AB}\|\rho^x_A\otimes\rho^x_B) \geq \sum_x p(x)D^{\epsilon}_{max}(\rho^x_{AB}\|\rho^x_A\otimes\rho^x_B), \tag{4}$$

we have to deal with $D^{\epsilon}_{max}(\rho^x_{AB}\|\rho^x_A\otimes\rho^x_B)$ and try to bound it from below. Let $\bar{\rho}^x_{AB}$ be the state achieving the minimum in the definition of $D^{\epsilon}_{max}(\rho^x_{AB}\|\rho^x_A\otimes\rho^x_B)$, hence

$$D^{\epsilon}_{max}(\rho^x_{AB}\|\rho^x_A\otimes\rho^x_B) \geq D_{max}(\bar{\rho}^x_{AB}\|\bar{\rho}^x_A\otimes\bar{\rho}^x_B)$$

where $P(\rho^x_{AB},\bar{\rho}^x_{AB})\leq\epsilon$. From Fact 3 we further know that $D_{max}(\bar{\rho}^x_{AB}\|\bar{\rho}^x_A\otimes\bar{\rho}^x_B)\geq D(\bar{\rho}^x_{AB}\|\bar{\rho}^x_A\otimes\bar{\rho}^x_B)$. 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 $\frac{1}{2}\|\rho^x_{AB}-\bar{\rho}^x_{AB}\|\leq\epsilon$)

$$\begin{aligned}D(\bar{\rho}^x_{AB}\|\bar{\rho}^x_A\otimes\bar{\rho}^x_B) \geq\;& D(\rho^x_{AB}\|\rho^x_A\otimes\rho^x_B)\\ &-2\epsilon\log|\mathcal{H}_A| - 2(1+\epsilon)h_2(\frac{\epsilon}{1+\epsilon}).\end{aligned}$$

[4] Conditional alternate smooth max-information can be defined in the same way.

$$\lim_{n\to\infty}\frac{1}{n}D_H^{\epsilon}(\rho_{AB}^{x^n}\|\rho_A^{x^n}\otimes\rho_B^{x^n}) = \lim_{n\to\infty}\frac{1}{n}D_H^{\epsilon}\left(\rho_{AB}^{np(x_1)\pm\delta}\otimes ... \otimes\rho_{AB}^{np(x_{|\mathcal{X}|})\pm\delta}\|(\rho_A^{x_1}\otimes\rho_B^{x_1})^{\otimes np(x_1)\pm\delta}\otimes ... \otimes(\rho_A^{x_{|\mathcal{X}|}}\otimes\rho_B^{x_{|\mathcal{X}|}})^{\otimes np(x_{|\mathcal{X}|})\pm\delta}\right)$$
$$= \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{|\mathcal{X}|}\big(np(x_i)\pm\delta\big)D(\rho_{AB}^{x_i}\|\rho_A^{x_i}\otimes\rho_B^{x_i}) = \sum_{x=1}^{|\mathcal{X}|}p(x)D(\rho_{AB}^{x}\|\rho_A^{x}\otimes\rho_B^{x}) \coloneqq I(A;B|X)_\rho. \tag{3}$$

Therefore,

$$D_{max}^{\epsilon}(\rho_{AB}^x\|\rho_A^x\otimes\rho_B^x) \geq D(\rho_{AB}^x\|\rho_A^x\otimes\rho_B^x) - 2\epsilon\log|\mathcal{H}_A| - 2(1+\epsilon)h_2(\frac{\epsilon}{1+\epsilon}),$$

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 $\tau_{A_1...A_K B}$ as follows:

$$\tau_{A_1...A_{|\mathcal{K}|}B} = \frac{1}{|\mathcal{K}|}\sum_{k=1}^{|\mathcal{K}|}\rho_{A_1}\otimes ... \otimes\rho_{A_{k-1}}\otimes\rho_{A_kB}\otimes\rho_{A_{k+1}}\otimes ... \otimes\rho_{A_{|\mathcal{K}|}}.$$

If

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

then

$$P(\tau_{A_1...A_{|\mathcal{K}|}B}, \rho_{A_1}\otimes ... \otimes\rho_{A_k}\otimes ... \otimes\rho_{A_{|\mathcal{K}|}}\otimes\tilde{\rho}_B) \leq \sqrt{\epsilon},$$

where $\tilde{\rho}_B$ is the marginal of some state $\tilde{\rho}_{AB} \in \mathcal{B}^{\sqrt{\epsilon}-\delta}(\rho_{AB})$. 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 $(\mathbb{1}-S) \in \mathcal{P}(\mathcal{H}_A)$. Then for all constants $c > 0$, the following inequality holds:

$$\mathbb{1} - (S+T)^{-\frac{1}{2}}S(S+T)^{-\frac{1}{2}} \leq (1+c)(\mathbb{1}-S) + (2+c+c^{-1})T.$$

Lemma 7 (Gentle measurement lemma [6]): Let $\rho_A \in \mathcal{D}(\mathcal{H}_A)$ and $0 \leq \Lambda_A \leq \mathbb{1}$ be a measurement operator. If the measurement operator decides in favor of ρ_A with high probability, $\text{Tr}\{\Lambda_A\rho_A\} \geq 1-\epsilon$ for $\epsilon \in [0,1]$, then

$$\left\|\rho_A - \sqrt{\Lambda_A}\rho_A\sqrt{\Lambda_A}\right\|_1 \leq 2\sqrt{\epsilon}.$$

We note that here we consider quantum communication channels with quantum input and outputs. One may consider channels with classical inputs and quantum outputs, i.e., CQ channels. In this case, an encoder has to be prepended to the CQ channel such that it associates a particular input state to every classical input.

III. Problem Statement And Main Results

In this section, we first define a simultaneous public-private one-shot code, then we present our main results. Latter, we discuss the translation of the public-private code to a classical-quantum code. Two classical messages $(m,\ell) \in \mathcal{M}\times\mathcal{L}$ are to be transmitted from a sender to a receiver in the presence of an eavesdropper by using a quantum channel only once, i.e., one-shot communication is considered. The sender Alice, wishes to reliably communicate a public message m and (simultaneously) a private message ℓ to the legitimate receiver Bob such that ℓ must not be leaked to the eavesdropper Eve. The quantum (wiretap) channel to be used by three parties is denoted by $\mathcal{N}_{A\to BE}$ and it takes quantum states from $\mathcal{H}_A$ to $\mathcal{H}_B\otimes\mathcal{H}_E$ where Alice is assumed to control the input system A and systems B and E are outputs received by Bob and Eve, respectively. Let M and L be the random variables[5] corresponding to Alice's choices of the public and private messages, respectively[6]. We formally define a one-shot simultaneous public-private code in the following.

Definition 18: Fix ϵ, $\epsilon' \in (0,1)$ and let r and R be the rates of the public and private messages, respectively (i.e., $|\mathcal{M}| = 2^r$ and $|\mathcal{L}| = 2^R$). A one-shot (r,R,ϵ,ϵ')-simultaneous public-private code for the channel $\mathcal{N}_{A\to BE}$ consists of

- An encoding operation by Alice $\mathcal{E}: ML \to \mathcal{D}(\mathcal{H}_A)$ such that
$$\forall m\in\mathcal{M}, \qquad \frac{1}{2}\|\rho_{LE}^m - \rho_L\otimes\tilde{\rho}_E^m\|_1 \leq \epsilon', \tag{5}$$
where for each message m, ρ_{LE}^m and ρ_L are appropriate marginals of the state $\rho_{LBE}^m = \frac{1}{|\mathcal{L}|}\sum_{\ell=1}^{|\mathcal{L}|}|\ell\rangle\langle\ell|\otimes\mathcal{N}(\mathcal{E}(m,\ell))$ and $\tilde{\rho}_E^m$ can be any arbitrary state.
- A decoding operation by Bob $\mathcal{D}: \mathcal{D}(\mathcal{H}_B) \to \hat{M}\hat{L}$ such that
$$Pr\Big((\hat{M},\hat{L}) \neq (M,L)\Big) \leq \epsilon, \tag{6}$$
where $\hat{M}$ and $\hat{L}$ denote the estimates of the public and private messages, respectively.

A rate pair (r,R) is said to be (ϵ,ϵ')-achievable if there exist encoding and decoding maps $(\mathcal{E},\mathcal{D})$ such that (5) and (6) are fulfilled. For a given (ϵ,ϵ'), the one-shot capacity region for the simultaneous transmission of public and private

[5] M and L basically are registers which hold the public and private messages, respectively. Here with slightly abuse of notation, we refer to them as random variables to which, corresponding classical states can be tied.

[6] In the literature, for example [18], the public and private messages are referred to as the common and confidential messages, respectively. If Eve were to receive the common message, it could have been considered without jeopardizing the confidential message. Indeed, as we will see, the secrecy analysis is guaranteed assuming Eve has detected the common (or the public) message.

information of the channel $\mathcal{N}$, $\mathcal{C}^{\epsilon,\epsilon'}(\mathcal{N})$, 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 $\mathcal{C}^{\epsilon,\epsilon'}(\mathcal{N})$.

In the following, we first have Theorem 2 that establishes a lower bound on $\mathcal{C}^{\epsilon,\epsilon'}(\mathcal{N})$ referred to as achievability and then Theorem 3 that states an upper bound on $\mathcal{C}^{\epsilon,\epsilon'}(\mathcal{N})$, 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 δ, δ' such that $\delta \in (0,\epsilon), \delta' \in (0,\epsilon')$, there exists a one-shot $(r, R, 3\epsilon + 2\sqrt{\epsilon} + \sqrt{\epsilon'}, 2(\epsilon + \sqrt{\epsilon}) + \sqrt{\epsilon'})$ code for the channel $\mathcal{N}_{A\to BE}$ if the twin (r, R) satisfies the following bounds:

$$\begin{aligned} r &\leq I_H^{\epsilon-\delta}(X;B)_\rho - \log_2(\frac{4\epsilon}{\delta^2}), \\ R &\leq I_H^{\epsilon-\delta}(Y;B|X)_\rho - \tilde{I}_{max}^{\sqrt{\epsilon'}-\delta'}(Y;E|X)_\rho \\ &\qquad\qquad - \log_2(\frac{4\epsilon}{\delta^2}) - 2\log_2(\frac{1}{\delta'}), \end{aligned}$$

for some quantum state ρ arising from the channel. We call the region above $\mathcal{C}_a(\mathcal{N})$, therefore, we have

$$\mathcal{C}_a(\mathcal{N}) \subseteq \mathcal{C}^{3\epsilon+2\sqrt{\epsilon}+\sqrt{\epsilon'},2(\epsilon+\sqrt{\epsilon})+\sqrt{\epsilon'}}(\mathcal{N}).$$

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 $\mathcal{N}_{A\to BE}$, must satisfy the following inequalities:

$$\begin{aligned} r \leq& I_H^{\epsilon}(X;B)_\rho, \\ R \leq& I_H^{\sqrt{\epsilon}}(Y;B|X)_\rho - I_{max}^{\sqrt{2\epsilon'}}(Y;E|X)_\rho, \end{aligned}$$

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 $\mathcal{C}_c(\mathcal{N})$. In fact, we have $\mathcal{C}^{\epsilon,\epsilon'}(\mathcal{N}) \subseteq \mathcal{C}_c(\mathcal{N})$.

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 (classcial, 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 prepended 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 $|\mathcal{M}|$ 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 $|\mathcal{L}|$. 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 $\{\mathcal{X} \times \mathcal{Y}\}$, $\epsilon, \epsilon' \in (0,1)$, $\delta \in (0,\epsilon)$, $\delta' \in (0, \sqrt{\epsilon'})$ and $\rho_{XYBE} = \sum_{x,y} p(x,y)|x\rangle\langle x| \otimes |y\rangle\langle y| \otimes \rho_{BE}^{x,y}$. Let

$$\begin{aligned} r &\leq I_H^{\epsilon-\delta}(X;B)_\rho - \log_2(\frac{4\epsilon}{\delta^2}), \\ R + \tilde{R} &\leq I_H^{\epsilon-\delta}(Y;B|X)_\rho - \log_2(\frac{4\epsilon}{\delta^2}), \\ \tilde{R} &\geq \tilde{I}_{max}^{\sqrt{\epsilon'}-\delta'}(E;Y|X)_\rho + 2\log_2(\frac{1}{\delta'}). \end{aligned}$$

We choose $|\mathcal{M}| = 2^r$, $|\mathcal{L}| = 2^R$ and $|\mathcal{K}| = 2^{\tilde{R}}$ implying that r and R denote our public and private rates, respectively and $|\mathcal{K}|$ 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 $\mathcal{N}^{A\to BE}$.

Alice wants to convey to Bob, in a single use of a quantum channel, a classical message $m \in \mathcal{M}$ and simultaneously, a private classical message $\ell \in \mathcal{L}$ 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 ℓ 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, inferred from the well-known superposition coding in classical information theory [43]. We can think of the state (7) as the superposition of two states, 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, ℓ), Alice goes to the m-th copy of $\rho_{XX'(AYY')^{\otimes|\mathcal{L}||\mathcal{K}|}}^{\otimes|\mathcal{M}|}$. 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 ℓ-th bin. Upon receiving B, Bob performs a position-based decoding to obtain the public message m (and hence the correct copy of $\rho_{XX'(AYY')^{\otimes|\mathcal{L}||\mathcal{K}|}}$). 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 $\rho_{XX'(AYY')^{\otimes|\mathcal{L}||\mathcal{K}|}}$ is almost unchanged after Bob's decoding.

$$\rho^{\otimes|\mathcal{M}|}_{XX'(AYY')^{\otimes|\mathcal{L}||\mathcal{K}|}} := \left(\sum_x p(x)|x\rangle\langle x|_X \otimes |x\rangle\langle x|_{X'} \otimes \left(\sum_y p(y|x)|y\rangle\langle y|_Y \otimes |y\rangle\langle y|_{Y'} \otimes \rho^{x,y}_A\right)^{\otimes|\mathcal{L}||\mathcal{K}|}\right)^{\otimes|\mathcal{M}|} . \tag{7}$$

To decode ℓ, Bob performs another position-based decoding conditioned on X, meaning that having found the correct copy of $\rho_{XX'(AYY')^{\otimes|\mathcal{L}||\mathcal{K}|}}$ 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 $x \in \mathrm{supp}(\rho_{X'})$, 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 $\rho_{XX'(AYY')^{\otimes|\mathcal{L}||\mathcal{K}|}}$, the condition that convex-split lemma imposes on $|\mathcal{K}|$, gives her very small information about ℓ for every $x \in \mathrm{supp}(\rho_{X'})$ (where here $\rho_{X'} = \sum_x p_{X'}(x)|x\rangle\langle x|_{X'}$ and $p_{X'}(x)$ 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 decodings are just a property of our protocol.

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 $\tilde{R} = \log_2|\mathcal{K}|$. The state initially shared between three parties is given by equation (7), 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 $\Upsilon_{T_X T_{X'} T_A T_Y T_{Y'}} := \rho^{\otimes|\mathcal{M}|}_{XX'(AYY')^{\otimes|\mathcal{L}||\mathcal{K}|}}$ with it being clear that for example $\Upsilon_{T_A} = \rho^{\otimes|\mathcal{M}|}_{A^{\otimes|\mathcal{L}||\mathcal{K}|}}$.[7]

The encoding and decoding pairs are as follows:

- Alice performs some encoding operation $\mathcal{E} : MLA \to A$. Let us denote the state in (7) after channel transmission as:
$$\left(\rho_{XX'(AYY')^{\otimes|\mathcal{L}||\mathcal{K}|}}\right)^{\otimes|\mathcal{M}|-1} \otimes \rho^{m,(\ell,k)}_{XX'(YY')^{\otimes|\mathcal{L}||\mathcal{K}|}(A)^{\otimes|\mathcal{L}||\mathcal{K}|-1}BE}, \tag{8}$$
where $(m, \ell, k) \in [1 : 2^r] \times [1 : 2^R] \times [1 : 2^{\tilde{R}}]$ are the public message, the private message and a dummy index drawn uniformly at random by the encoder and $\rho^{m,(\ell,k)}_{XX'(YY')^{\otimes|\mathcal{L}||\mathcal{K}|}(A)^{\otimes|\mathcal{L}||\mathcal{K}|-1}BE}$ is given by equation (9).
- After the channel action, Bob performs a decoding operation (quantum instrument) $\mathcal{D}^1 : BX \to \hat{M}B$ on his Υ_{T_X} systems as well as the received system, whose outputs are a classical system $\hat{M}$ and a quantum system in $\mathcal{D}(\mathcal{H}_B)$ (the decoder will be defined formally later, see (14)). The action of the quantum decoder $\mathcal{D}^1_{BX\to\hat{M}B}$ on Bob's corresponding systems is as follows:
$$\mathcal{D}^1_{BX\to\hat{M}B}(\rho^{m,(\ell,k)}_{X^{\otimes|\mathcal{M}|}B}) := \sum_{m'=1}^{|\mathcal{M}|} |m'\rangle\langle m'|_{\hat{M}} \otimes \mathcal{D}^{1,m'}_{BX\to B}(\rho^{m,(\ell,k)}_{X^{\otimes|\mathcal{M}|}B}), \tag{10}$$
where $\{|m\rangle\}_{m=1}^{|\mathcal{M}|}$ is some orthonormal basis and $\rho^{m,(\ell,k)}_{X^{\otimes|\mathcal{M}|}B}$ can be seen from (8) by tracing out uninvolved systems. Moreover, tracing out the classical system $\hat{M}$ gives the induced quantum operation $\mathcal{D}^1_{BX\to B} =$

[7] Due to the cumbersome notations we face, the tensor product states are shown for example as either $\rho^{\otimes|\mathcal{M}|}_X$ or $\rho_{X^{\otimes|\mathcal{M}|}}$.

$$\rho^{m,(\ell,k)}_{XX'(YY')^{\otimes|\mathcal{L}||\mathcal{K}|}(A)^{\otimes|\mathcal{L}||\mathcal{K}|-1}BE} :=$$
$$\sum_x p_X(x)|x\rangle\langle x|_X \otimes |x\rangle\langle x|_{X'} \otimes \rho^{x,m,(1,1)}_{YY'A} \otimes ... \otimes \rho^{x,m,(\ell,k-1)}_{YY'A} \otimes \mathcal{N}_{A\to BE}\left(\rho^{x,m,(\ell,k)}_{YY'A}\right) \otimes \rho^{x,m,(\ell,k+1)}_{YY'A} ... \otimes \rho^{x,m,(|\mathcal{L}|,|\mathcal{K}|)}_{YY'A}. \quad (9)$$

$\sum_m \mathcal{D}^{1,m}_{BX\to B}$ such that its sum is trace preserving, i.e., $\text{Tr}\left\{\sum_{m'=1}^{|\mathcal{M}|} \mathcal{D}^{1,m'}_{BX\to B}(\rho^{m,(\ell,k)}_{X^{\otimes|\mathcal{M}|}B})\right\} = 1.$

Let $\sigma^{m,(\ell,k)}_{XX'(YY')^{\otimes|\mathcal{L}||\mathcal{K}|}BE}$ denote the *disturbed* state after Bob applied his first decoder (this state will be defined formally later, see (20)).

- Bob's second decoder is another quantum map $\mathcal{D}^2 : \hat{M}BY \to \hat{L}$ which is input the classical output of the first decoder, the disturbed quantum output, Bob's Υ_{T_Y} systems and outputs a classical system $\hat{L}$.

$$\mathcal{D}^2_{\hat{M}BY\to\hat{L}}(\sigma^{m,(\ell,k)}_{XY^{\otimes|\mathcal{L}||\mathcal{K}|}B}) := \sum_{\ell=1}^{|\mathcal{L}|} p_{\hat{L}}(\ell)|\ell\rangle\langle\ell|_{\hat{L}}, \quad (11)$$

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

Having defined the decoders, it is seen that the phrase in (12) indicates the probability of an erroneous detection of the public message, while the expression in (13) 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.

1) Correctness of Public Message: Eq. (12): All systems are assumed to be traced out except those used by Bob's first decoder (we could have considered multiplying those systems by identity operator as well). To decode the public message m, Bob employs the following decoding instrument:

$$\mathcal{D}^1_{BX\to\hat{M}B}(\rho^{m,(\ell,k)}_{X^{\otimes|\mathcal{M}|}B}) \quad (14)$$
$$:= \sum_{m=1}^{|\mathcal{M}|} \text{Tr}\{\Lambda^m_{X^{|\mathcal{M}|}B}\rho^{m,(\ell,k)}_{X^{\otimes|\mathcal{M}|}B}\}|m\rangle\langle m|_{\hat{M}}$$
$$\otimes \frac{\sqrt{\Lambda^m_{X^{|\mathcal{M}|}B}}\rho^{m,(\ell,k)}_{X^{\otimes|\mathcal{M}|}B}\sqrt{\Lambda^m_{X^{|\mathcal{M}|}B}}}{\text{Tr}\{\Lambda^m_{X^{|\mathcal{M}|}B}\rho^{m,(\ell,k)}_{X^{\otimes|\mathcal{M}|}B}\}},$$

where $\Lambda^m_{X^{|\mathcal{M}|}B}$ is given in (15), and for $m \in [1 : |\mathcal{M}|]$:

$$\Gamma^m_{X^{|\mathcal{M}|}B} = \mathbb{1}^1_X \otimes \mathbb{1}^2_X \otimes ... \otimes T^m_{XB} \otimes ... \otimes \mathbb{1}^{|\mathcal{M}|}_X,$$

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

$$\rho^{m,(\ell,k)}_{X^{\otimes|\mathcal{M}|}B} := \rho_X^{\otimes|\mathcal{M}|-1} \otimes \rho^{m,(\ell,k)}_{XB}.$$

Note that to decode the public message m, Bob's decoder does not care about the copy selected by Alice among $|\mathcal{L}||\mathcal{K}|$ 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 $\rho_{XA} = \sum_x P_X(x)|x\rangle\langle x| \otimes \omega^x_A$ shared between Alice and Bob, where $\omega^x_A = \sum_y p(y|x)\omega^{x,y}_A$. 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_{XB} be a test operator in a binary hypothesis testing scenario with null and alternative hypotheses being ρ_{XB} and $\rho_X \otimes \rho_B$, respectively. Discriminator employed by Bob succeeds in guessing null and alternative hypotheses with probabilities $\text{Tr}\{T_{XB}\rho_{XB}\}$ and $\text{Tr}\{(\mathbb{1}_{XB} - T_{XB})(\rho_X \otimes \rho_B)\}$, respectively. And accordingly, the error probabilities associated to the type I and II errors are $\text{Tr}\{(\mathbb{1}_{XB} - T_{XB})\rho_{XB}\}$ and $\text{Tr}\{T_{XB}(\rho_X \otimes \rho_B)\}$, 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 (ϵ) 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 $m \in [1 : |\mathcal{M}|]$:

$$\Gamma^m_{X^{|\mathcal{M}|}B} = \mathbb{1}^1_X \otimes ... \otimes T^m_{XB} \otimes ... \otimes \mathbb{1}^{|\mathcal{M}|}_X.$$

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

$$\text{Tr}\{\Gamma^m_{X^{|\mathcal{M}|}B}\rho^{m,(\ell,k)}_{X^{\otimes|\mathcal{M}|}B}\}$$
$$= \text{Tr}\{(\mathbb{1}_X{}^1 \otimes \mathbb{1}^2_X \otimes ... \otimes T^m_{XB} \otimes ... \otimes \mathbb{1}^{|\mathcal{M}|}_X)$$
$$(\rho^1_X \otimes ... \otimes \rho^{m,(\ell,k)}_{XB} \otimes ... \otimes \rho^{|\mathcal{M}|}_X)\}$$
$$= \text{Tr}\{T^m_{XB}\rho^m_{XB}\} = \text{Tr}\{T_{XB}\rho_{XB}\}, \quad (16)$$

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:

$$\text{Tr}\{\Gamma^{m'}_{X^{|\mathcal{M}|}B}\rho^{m,(\ell,k)}_{X^{\otimes|\mathcal{M}|}B}\}$$
$$= \text{Tr}\{(\mathbb{1}^m_X \otimes T^{m'}_{XB})(\rho^{m,(\ell,k)}_{XB} \otimes \rho^{m'}_X)\}$$
$$= \text{Tr}\{T^{m'}_{XB}(\rho^{m,(\ell,k)}_B \otimes \rho^{m'}_X)\} = \text{Tr}\{T_{XB}(\rho_B \otimes \rho_X)\}, \quad (17)$$

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 [2].

$$P_e =\{\hat{M} \neq M\} := \frac{1}{M}\sum_{m=1}^{|\mathcal{M}|}\frac{1}{2}\left\|\mathcal{D}^1_{BX\to\hat{M}}(\rho^{m,(\ell,k)}_{X^{\otimes|\mathcal{M}|}B}) - |m\rangle\langle m|_{\hat{M}}\right\|_1 \leq \epsilon, \tag{12}$$

$$P_{priv} :=\frac{1}{|\mathcal{L}|}\sum_{l=1}^{|\mathcal{L}|}\frac{1}{2}\left\|\mathcal{D}^2_{\hat{M}BY\to\hat{L}}(\sigma^{m,(\ell,k)}_{XX'(YY')^{\otimes|\mathcal{L}||\mathcal{K}|}BE}) - |l\rangle\langle l|_{\hat{L}}\otimes\hat{\sigma}_{X'Y'^{\otimes|\mathcal{L}||\mathcal{K}|}E}\right\|_1 \leq 2(\epsilon+\sqrt{\epsilon})+\sqrt{\epsilon'}, \tag{13}$$

where

$$\hat{\sigma}_{X'Y'^{\otimes|\mathcal{L}||\mathcal{K}|}E} := \sum_x p_X(x)|x\rangle\langle x|_{X'}\otimes\sigma^{x,m,(\ell,k)}_{Y'^{\otimes|\mathcal{L}||\mathcal{K}|}}\otimes\tilde{\sigma}^{x,m}_E \quad \text{and} \quad P(\sigma^x_{YE},\tilde{\sigma}^x_{YE})\leq\sqrt{\epsilon'}.$$

$$\Lambda^m_{X^{|\mathcal{M}|}B} := \left(\sum_{m'=1}^{|\mathcal{M}|}\Gamma^{m'}_{X^{|\mathcal{M}|}B}\right)^{-\frac{1}{2}}\Gamma^m_{X^{|\mathcal{M}|}B}\left(\sum_{m'=1}^{|\mathcal{M}|}\Gamma^{m'}_{X^{|\mathcal{M}|}B}\right)^{-\frac{1}{2}}. \tag{15}$$

We now focus on the analysis of the error 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., $\Lambda^m_{X^{|\mathcal{M}|}B} = U^{\pi(m)}_{X^{|\mathcal{M}|}B}\Lambda^1_{X^{|\mathcal{M}|}B}U^{\pi(m)\dagger}_{X^{|\mathcal{M}|}B}$ in which $\pi(.)$ denotes the permutatin 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 $|\mathcal{L}||\mathcal{K}|$ A subsystems of the first copy over the channel. We emphasize again that although Alice selects a particular A subsystem out of $|\mathcal{L}||\mathcal{K}|$ 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 $S = \Gamma^1_{X^{|\mathcal{M}|}B}$ and $T = \sum_{m\neq 1}\Gamma^m_{X^{|\mathcal{M}|}B}$ (This T should not be confused with the test operator T^m_{XB}):

$$\begin{aligned}
&Pr(\hat{M}\neq 1|M=1)\\
&\quad= \mathrm{Tr}\{(\mathbb{1}_{X^{|\mathcal{M}|}B}-\Lambda^1_{X^{|\mathcal{M}|}B})\rho^{1,(\ell,k)}_{X^{\otimes|\mathcal{M}|}B}\}\\
&\quad\leq \mathrm{Tr}\{((1+c)(\mathbb{1}_{X^{|\mathcal{M}|}B}-\Gamma^1_{X^{|\mathcal{M}|}B})\rho^{1,(\ell,k)}_{X^{\otimes|\mathcal{M}|}B}\}\\
&\qquad\qquad+(2+c+c^{-1})\mathrm{Tr}\{(\sum_{m\neq 1}\Gamma^m_{X^{|\mathcal{M}|}B})\rho^{1,(\ell,k)}_{X^{\otimes|\mathcal{M}|}B}\}\\
&\quad= (1+c)\mathrm{Tr}\{(\mathbb{1}_{X^{|\mathcal{M}|}B}-\Gamma^1_{X^{|\mathcal{M}|}B})\rho^{1,(\ell,k)}_{X^{\otimes|\mathcal{M}|}B}\}\\
&\qquad\qquad+(2+c+c^{-1})\mathrm{Tr}\{(\sum_{m\neq 1}\Gamma^m_{X^{|\mathcal{M}|}B})\rho^{1,(\ell,k)}_{X^{\otimes|\mathcal{M}|}B}\}
\end{aligned}$$

$$\begin{aligned}
&= (1+c)\mathrm{Tr}\{(\mathbb{1}^1_X - T^1_{XB})\rho^{1,(\ell,k)}_{XB}\}\\
&\qquad+(2+c+c^{-1})\sum_{m\neq 1}\mathrm{Tr}\{T^m_{XB}(\rho^{m,(\ell,k)}_B\otimes\rho^m_X)\}\\
&= (1+c)\mathrm{Tr}\{(\mathbb{1}_{XB}-T_{XB})\rho_{XB}\}\\
&\qquad+(2+c+c^{-1})(|\mathcal{M}|-1)\mathrm{Tr}\{T_{XB}(\rho_B\otimes\rho_X)\},
\end{aligned}$$

where in the second last equality, the first and second terms follow from (16) and (17), respectively. Let Π_{XB} be the optimal test operator in the following optimization: (see Definition 1 and Definition 12)

$$I_H^{\epsilon-\delta}(X;B)_{\rho_{XB}} := -\log_2 \inf_{\substack{0\leq T_{XB}\leq\mathbb{1},\\ \alpha(T_{XB},\rho_{XB})\leq\epsilon-\delta}} \beta(T_{XB},\rho_X\otimes\rho_B),$$

then,

$$\begin{aligned}
&\mathrm{Tr}\{(\mathbb{1}_{X^{|\mathcal{M}|}B}-\Lambda^1_{X^{|\mathcal{M}|}B})\rho^{1,(\ell,k)}_{X^{\otimes|\mathcal{M}|}B}\}\\
&\quad\leq(1+c)\mathrm{Tr}\{(\mathbb{1}_{XB}-\Pi_{XB})\rho_{XB}\}\\
&\qquad\qquad+(2+c+c^{-1})(|\mathcal{M}|-1)\mathrm{Tr}\{\Pi_{XB}(\rho_B\otimes\rho_X)\}\\
&\quad\leq(1+c)(\epsilon-\delta)+(2+c+c^{-1})|\mathcal{M}|2^{-I_H^{\epsilon-\delta}(X;B)_\rho}.
\end{aligned}$$

The last term above is set equal to ϵ, if we solve for $|\mathcal{M}|$, we end up with the following term

$$\log_2|\mathcal{M}| = I_H^{\epsilon-\delta}(X;B)_\rho + \log_2\left(\frac{\epsilon-(1+c)(\epsilon-\delta)}{2+c+c^{-1}}\right),$$

the expression inside the $\log$ has a global maximum with respect to c, i.e., the parabola is down-side. We put first derivative equal to zero and pick $c=\frac{\delta}{2\epsilon-\delta}$ and by doing so finally the following bound holds:

$$\log_2|\mathcal{M}| \leq I_H^{\epsilon-\delta}(X;B)_\rho - \log_2(\frac{4\epsilon}{\delta^2}), \tag{18}$$

and average probability of error of the public message for the one-shot assisted code is

$$\frac{1}{|\mathcal{M}|}\sum_{m=1}^{|\mathcal{M}|}\mathrm{Tr}\{(\mathbb{1}_{X^{|\mathcal{M}|}B}-\Lambda^m_{X^{|\mathcal{M}|}B})\rho^{m,(\ell,k)}_{X^{\otimes|\mathcal{M}|}B}\}\leq\epsilon. \tag{19}$$

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 is 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) $\sqrt{\Lambda^{m}_{X^{|\mathcal{M}|}B}}\rho^{m,(\ell,k)}_{X^{\otimes|\mathcal{M}|}B}\sqrt{\Lambda^{m}_{X^{|\mathcal{M}|}B}}$, and by taking uninvolved systems into account, we define the state that passes to the second decoder as in (20).

2) Correctness and secrecy of Private Message, (Privacy error) Eq. (13): Reconsider the state in (20) showing the state resulted from transmitting the (ℓ,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 $|\mathcal{L}||\mathcal{K}|$ 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 ℓ is constructed as follows:

$$\begin{aligned}\mathcal{D}^{2}_{\hat{M}BY\to\hat{L}}(\sigma^{m,(\ell,k)}_{XY^{\otimes|\mathcal{L}||\mathcal{K}|}B}) \qquad\qquad\qquad\qquad (21)\\ :=\sum_{l=1}^{|\mathcal{L}|}\mathrm{Tr}\{P^{\ell}_{XY^{|\mathcal{L}||\mathcal{K}|}B}\sigma^{m,(\ell,k)}_{XY^{\otimes|\mathcal{L}||\mathcal{K}|}B}\}|\ell\rangle\langle\ell|_{\hat{L}},\end{aligned}$$

where for all $x\in\mathcal{X}$

$$P^{\ell}_{XY^{|\mathcal{L}||\mathcal{K}|}B}=\sum_{k=1}^{|\mathcal{K}|}P^{(\ell,k)}_{XY^{|\mathcal{L}||\mathcal{K}|}B},$$

and $P^{x,(\ell,k)}_{XY^{|\mathcal{L}||\mathcal{K}|}B}$ is given in (22), in which, for all $x\in\mathcal{X}$, Z_{YB} is a binary test operator distinguishing between two hypotheses σ^{x}_{YB} and $\sigma^{x}_{Y}\otimes\sigma^{x}_{B}$ with an error of $\epsilon-\delta$, i.e.,

$$\mathrm{Tr}\{Z_{YB}\sigma^{x}_{YB}\}\geq 1-(\epsilon-\delta),$$

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 $x\in\mathcal{X}$.

Bob has to be able to distinguish between states $\sigma^{m,(1,1)}_{XY^{\otimes|\mathcal{L}||\mathcal{K}|}B},\sigma^{m,(1,2)}_{XY^{\otimes|\mathcal{L}||\mathcal{K}|}B},\dots,\sigma^{m,(l,k)}_{XY^{\otimes|\mathcal{L}||\mathcal{K}|}B},\sigma^{m,(|\mathcal{L}|,|\mathcal{K}|)}_{XY^{\otimes|\mathcal{L}||\mathcal{K}|}B}$; We will see that this amounts to Bob being able to distinguish between the following states:

$$\begin{aligned}&\sum_{x}p_{X}(x)|x\rangle\langle x|_{X}\otimes\sigma^{x}_{YB},\\&\sum_{x}p_{X}(x)|x\rangle\langle x|_{X}\otimes\sigma^{x}_{Y}\otimes\sigma^{x}_{B},\end{aligned}$$

or more precisely, between state σ^{x}_{YB} and $\sigma^{x}_{Y}\otimes\sigma^{x}_{B}$ for all $x\in\mathcal{X}$. We importantly note that after detecting the public message m, Bob is faced a $|\mathcal{L}||\mathcal{K}|$-ary hypothesis testing problem. This scenario should not be confused by the binary hypothesis testing above, i.e., Alice distinguishes between σ^{x}_{YB} and $\sigma^{x}_{Y}\otimes\sigma^{x}_{B}$ for all $x\in\mathcal{X}$, the latter happens to be a byproduct of the general scenario once we go into the error analysis. Now see that if the pair (ℓ,k) was chosen, the action of the operator $N^{(\ell,k)}_{Y^{|\mathcal{L}||\mathcal{K}|}B}$ would be as follows:

$$\begin{aligned}&\mathrm{Tr}\{N^{(\ell,k)}_{XY^{|\mathcal{L}||\mathcal{K}|}B}\sigma^{m,(\ell,k)}_{XY^{\otimes|\mathcal{L}||\mathcal{K}|}B}\}\\&\qquad=\sum_{x}p_{X}(x)\mathrm{Tr}\{Z^{(\ell,k)}_{YB}\sigma^{x,m,(\ell,k)}_{YB}\},\end{aligned}$$

and for any other operator, i.e., the private message-local key pair (ℓ,k) is confused by (ℓ',k'), either $k\neq k'$, $l\neq\ell'$ or $(k\neq k',\ell\neq\ell')$:

$$\begin{aligned}&\mathrm{Tr}\{N^{(\ell',k')}_{XY^{|\mathcal{L}||\mathcal{K}|}B}\sigma^{m,(\ell,k)}_{XY^{\otimes|\mathcal{L}||\mathcal{K}|}B}\}\\&=\mathrm{Tr}\Big\{|x\rangle\langle x|_{X}\otimes(Z^{(\ell',k')}_{YB}\otimes\mathbb{1}^{(\ell,k)}_{Y})\\&\qquad(\sum_{x}p_{X}(x)|x\rangle\langle x|_{X}\otimes\sigma^{x,m,(\ell',k')}_{Y}\otimes\sigma^{x,m,(\ell,k)}_{YB})\Big\}\\&=\sum_{x}p_{X}(x)\mathrm{Tr}\{Z^{(\ell',k')}_{YB}(\sigma^{x,m,(\ell',k')}_{Y}\otimes\sigma^{x,m,(\ell,k)}_{B})\}.\end{aligned}$$

We can think of the states $\sigma^{x,m}_{YB}$ and $\sigma^{x,m}_{Y}\otimes\sigma^{x,m}_{B}$ 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 (ℓ,k). It can be shown that each measurement operator $P^{(\ell,k)}_{Y^{|\mathcal{L}||\mathcal{K}|}B}$ is related to the first one $P^{(1,1)}_{Y^{|\mathcal{L}||\mathcal{K}|}B}$ by a unitary permutation of $Y^{|\mathcal{L}||\mathcal{K}|}$ systems for all $x\in\mathcal{X}$. This fact gives rise to the following identity, for all $\ell\in[1:|\mathcal{L}|]$ and $k\in[1:|\mathcal{K}|]$:

$$\begin{aligned}&\mathrm{Tr}\{(\mathbb{1}_{XY^{|\mathcal{L}||\mathcal{K}|}B}-P^{(1,1)}_{XY^{|\mathcal{L}||\mathcal{K}|}B})\sigma^{m,(1,1)}_{XY^{\otimes|\mathcal{L}||\mathcal{K}|}B}\}\\&\qquad=\mathrm{Tr}\{(\mathbb{1}_{XY^{|\mathcal{L}||\mathcal{K}|}B}-P^{(\ell,k)}_{XY^{|\mathcal{L}||\mathcal{K}|}B})\sigma^{m,(\ell,k)}_{XY^{\otimes|\mathcal{L}||\mathcal{K}|}B}\},\end{aligned}$$

meaning that the error probability is the same for all private messages, in other words, it is independent from a particular chosen twin (ℓ,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 $S=N^{(1,1)}_{Y^{|\mathcal{L}||\mathcal{K}|}B}$ and $T=\sum_{\ell'\neq1}\sum_{k'\neq1}N^{(\ell',k')}_{Y^{|\mathcal{L}||\mathcal{K}|}B}$ in Hayashi-Nagaoka inequality. We have

$$\begin{aligned}&\mathrm{Tr}\{(\mathbb{1}_{Y^{|\mathcal{L}||\mathcal{K}|}B}-P^{(1,1)}_{XY^{|\mathcal{L}||\mathcal{K}|}B})\sigma^{m,(1,1)}_{XY^{\otimes|\mathcal{L}||\mathcal{K}|}B}\}\\&\leq(1+c)\mathrm{Tr}\{(\mathbb{1}_{Y^{|\mathcal{L}||\mathcal{K}|}B}-N^{(1,1)}_{XY^{|\mathcal{L}||\mathcal{K}|}B})\sigma^{m,(1,1)}_{XY^{\otimes|\mathcal{L}||\mathcal{K}|}B}\}\\&\qquad+(2+c+c^{-1})\sum_{l'\neq1}\sum_{k'\neq1}\mathrm{Tr}\{N^{(\ell',k')}_{XY^{|\mathcal{L}||\mathcal{K}|}B}\sigma^{m,(1,1)}_{XY^{\otimes|\mathcal{L}||\mathcal{K}|}}\}\\&=(1+c)\sum_{x}p_{X}(x)\mathrm{Tr}\{(\mathbb{1}_{YB}-Z^{(1,1)}_{YB})\sigma^{x,m,(1,1)}_{YB}\}\\&+(2+c+c^{-1})\sum_{x}p_{X}(x)\big(\\&\qquad\sum_{l'\neq1}\sum_{k'\neq1}\mathrm{Tr}\{Z^{(\ell',k')}_{YB}(\sigma^{x,m,(\ell',k')}_{Y}\otimes\sigma^{x,m,(1,1)}_{B})\}\big)\\&=(c+1)\sum_{x}p_{X}(x)\mathrm{Tr}\{(\mathbb{1}_{YB}-Z_{YB})\sigma^{x,m}_{YB}\}\\&+(2+c+c^{-1})(|\mathcal{L}||\mathcal{K}|-1)(\\&\qquad\qquad\sum_{x}p_{X}(x)\mathrm{Tr}\{Z_{YB}(\sigma^{x,m}_{Y}\otimes\sigma^{x,m}_{B})\})\end{aligned}$$

For each realization x, let Θ^{x}_{YB} denote the measurement operator that is the answer to the optimization mentioned in Definition 1 with $\alpha(Z_{YB},\sigma^{x}_{YB}):=\mathrm{Tr}\{(\mathbb{1}-Z_{YB})\sigma^{x}_{YB}\}$ and $\beta(Z_{YB},\sigma^{x}_{Y}\otimes\sigma^{x}_{B}):=\mathrm{Tr}\{Z_{YB}(\sigma^{x}_{Y}\otimes\sigma^{x}_{B})\}$ where by assumption it detects the joint state with an error probability

$$\sigma^{m,(\ell,k)}_{XX'(YY')^{\otimes|\mathcal{L}||\mathcal{K}|}BE} \\ \coloneqq \sum_x p_X(x)|x\rangle\langle x|_X \otimes |x\rangle\langle x|_{X'} \otimes \sigma^{x,m,(1,1)}_{YY'} \otimes ... \otimes \sigma^{x,m,(\ell,k-1)}_{YY'} \otimes \sigma^{x,m,(\ell,k)}_{YY'BE} \otimes \sigma^{x,m,(\ell,k+1)}_{YY'} ... \otimes \sigma^{x,m,(|\mathcal{L}|,|\mathcal{K}|)}_{YY'}. \quad (20)$$

$$P^{(\ell,k)}_{XY^{|\mathcal{L}||\mathcal{K}|}B} \coloneqq \left(\sum_{\ell'=1}^{|\mathcal{L}|}\sum_{k'=1}^{|\mathcal{K}|} N^{(\ell',k')}_{XY^{|\mathcal{L}||\mathcal{K}|}B}\right)^{-\frac{1}{2}} N^{(\ell,k)}_{XY^{|\mathcal{L}||\mathcal{K}|}B} \left(\sum_{\ell'=1}^{|\mathcal{L}|}\sum_{k'=1}^{|\mathcal{K}|} N^{(\ell',k')}_{XY^{|\mathcal{L}||\mathcal{K}|}B}\right)^{-\frac{1}{2}}, \quad (22)$$

where for all $\ell \in [1:|\mathcal{L}|]$, and $k \in [1:|\mathcal{K}|]$,

$$N^{(\ell,k)}_{XY^{|\mathcal{L}||\mathcal{K}|}B} \coloneqq |x\rangle\langle x|_X \otimes \mathbb{1}^{(1,1)}_Y \otimes ... \otimes \mathbb{1}^{(1,|\mathcal{K}|)}_Y \otimes ... \otimes \mathbb{1}^{(\ell,k-1)}_Y \otimes Z^{(\ell,k)}_{YB} \otimes \mathbb{1}^{(\ell,k+1)}_Y ... \otimes \mathbb{1}^{(|\mathcal{L}|,|\mathcal{K}|)}_Y. \quad (23)$$

of $\epsilon - \delta$ where $\delta \in (0, \epsilon)$. This optimization can be done for all x, but from the definition of the conditional hypothesis testing mutual information (Definition 16), the x minimizing the expression given in equation (24) over a nearby distribution is of particular interest in error analysis; The error probability simplifies as follows:

$$\begin{aligned}
&\mathrm{Tr}\{(\mathbb{1}_{Y^{|\mathcal{L}||\mathcal{K}|}B} - P^{(1,1)}_{XY^{|\mathcal{L}||\mathcal{K}|}B})\sigma^{m,(1,1)}_{XY^{\otimes|\mathcal{L}||\mathcal{K}|}B}\} \\
&\quad \le (c+1)\sum_x p_X(x)\mathrm{Tr}\{(\mathbb{1}_{YB} - \Theta_{YB})\sigma^{x,m}_{YB}\} \\
&\quad + (2+c+c^{-1})|\mathcal{L}||\mathcal{K}|\sum_x p_X(x)\mathrm{Tr}\{\Theta_{YB}(\sigma^{x,m}_Y \otimes \sigma^{x,m}_B)\} \\
&\quad \le (c+1)(\epsilon - \delta) \\
&\qquad\quad + (2+c+c^{-1})|\mathcal{L}||\mathcal{K}|2^{-I_H^{\epsilon-\delta}(Y;B|X)_{\sigma_{XYB}}},
\end{aligned}$$

where in the last line, the first expression is derived from the assumption that for all x, $\mathrm{Tr}\{\Theta_{YB}\sigma^{x,m}_{YB}\} \ge 1 - (\epsilon - \delta)$, and the second expression follows from (24). By putting the last line above equal to ϵ (Bob's error in detecting private message is ϵ) and solving it for $|\mathcal{L}||\mathcal{K}|$, we get:

$$\begin{aligned}
\log_2 |\mathcal{L}||\mathcal{K}| =& I_H^{\epsilon-\delta}(Y;B|X)_{\sigma_{XYB}} \\
&+ \log_2\left(\frac{\epsilon - (1+c)(\epsilon-\delta)}{2+c+c^{-1}}\right).
\end{aligned}$$

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 $c = \frac{\delta}{2\epsilon - \delta}$. By plugging it back into the expression we end up having:

$$\log_2 |\mathcal{L}||\mathcal{K}| \le I_H^{\epsilon-\delta}(Y;B|X)_{\sigma_{XYB}} - \log_2(\frac{4\epsilon}{\delta^2}).$$

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 (ℓ, k), Eve's state is[8]

$$\begin{aligned}
\sigma^{m,(\ell,k)}_{X'Y'^{\otimes|\mathcal{L}||\mathcal{K}|}E} = \sum_x p_X(x)|x\rangle\langle x|_{X'} \otimes \sigma^{x,m,(1,1)}_{Y'} \otimes ... \\
\otimes \sigma^{x,m,(\ell,k)}_{Y'E} \otimes ... \otimes \sigma^{x,m,(|\mathcal{L}|,|\mathcal{K}|)}_{Y'}.
\end{aligned}$$

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

$$\forall m, \ell: \quad \frac{1}{2}\|\sigma^{m,\ell}_{X'Y'^{\otimes|\mathcal{L}||\mathcal{K}|}E} - \hat{\sigma}_{X'Y'^{\otimes|\mathcal{L}||\mathcal{K}|}E}\|_1 \le \sqrt{\epsilon'}, \quad (26)$$

where

$$\hat{\sigma}_{X'Y'^{\otimes|\mathcal{L}||\mathcal{K}|}E} \coloneqq \sum_x p_X(x)|x\rangle\langle x|_{X'} \otimes \sigma^{x,m}_{Y'^{\otimes|\mathcal{L}||\mathcal{K}|}} \otimes \tilde{\sigma}^{x,m}_E \quad (27)$$

for $\epsilon' \in (0,1)$ and some state $\tilde{\sigma}^{m,x}_E$ that is the marginal of $\tilde{\sigma}^{m,x}_{Y'E}$ and $P(\sigma^{m,x}_{Y'E}, \tilde{\sigma}^{m,x}_{Y'E}) \le \sqrt{\epsilon'} - \delta'$ 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).

From the convex-split lemma and the definition of the conditional smooth max-mutual information (see Definition 17), if the following condition holds[9],

$$\log_2 |\mathcal{K}| = \tilde{I}_{max}^{\sqrt{\epsilon'}-\delta'}(E;Y|X)_\sigma + 2\log_2(\frac{1}{\delta'}), \quad (29)$$

then

$$P(\sigma^{m,\ell}_{X'Y'^{\otimes|\mathcal{L}||\mathcal{K}|}E}, \hat{\sigma}_{X'Y'^{\otimes|\mathcal{L}||\mathcal{K}|}E}) \le \sqrt{\epsilon'}$$

is satisfied with $\hat{\sigma}_{X'Y'^{|\mathcal{L}||\mathcal{K}|}E}$ defined in (27) and from the relation between purified distance and trace distance correctness of (26) is guaranteed. Note also that $P(\sigma^{x,m}_E, \tilde{\sigma}^{x,m}_E) \le$

[8] 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.

[9] To maintain consistency, in the following expression, we show Eve's X' and Y' systems with X and Y, respectively.

$$I_H^{\epsilon-\delta}(Y;B|X)_{\sigma_{XYB}} \coloneqq \max_{\sigma'_X} \min_{x\in\text{supp}\left(\sigma'_X\right)} \left\{ -\log_2 \inf_{\substack{0\leq Z^x_{YB}\leq \mathbb{1},\\ \alpha(Z^x_{YB},\sigma^x_{YB})\leq\epsilon-\delta}} \beta(Z^x_{YB}, \sigma^x_Y \otimes \sigma^x_B) \right\}, \tag{24}$$

$$\text{where} \qquad \sigma_{XYB} = \sum_x p_X(x)|x\rangle\langle x|_X \otimes \sigma^x_{YB} \qquad \text{and} \qquad P(\sigma'_X, \sigma_X) \leq \epsilon''.$$

$$\begin{aligned}\sigma^{m,\ell}_{X'Y'^{\otimes|\mathcal{L}||\mathcal{K}|}E} \coloneqq& \frac{1}{|\mathcal{K}|}\sum_{k=1}^{|\mathcal{K}|} \sigma^{m,(\ell,k)}_{X'Y'^{\otimes|\mathcal{L}||\mathcal{K}|}E} \coloneqq \sum_x p_X(x)|x\rangle\langle x|_{X'} \otimes \sigma^{x,m,(1,1)}_{Y'} \otimes ...\\ &\otimes \left[\frac{1}{|\mathcal{K}|}\sum_{k=1}^{|\mathcal{K}|} \sigma^{x,m,(\ell,1)}_{Y'} \otimes ... \otimes \sigma^{x,m,(\ell,k)}_{Y'E} \otimes ... \otimes \sigma^{x,m,(\ell,|\mathcal{K}|)}_{Y'} \right] \otimes ... \otimes \sigma^{x,m,(|\mathcal{L}|,|\mathcal{K}|)}_{Y'}.\end{aligned} \tag{25}$$

$$\begin{aligned}&\frac{1}{2}\left\| \sigma^{m,\ell}_{X'Y'^{\otimes|\mathcal{L}||\mathcal{K}|}E} - \sum_x p_X(x)|x\rangle\langle x|_{X'} \otimes \sigma^{x,m}_{Y'^{\otimes|\mathcal{L}||\mathcal{K}|}} \otimes \tilde{\sigma}^{x,m}_E \right\|_1\\ &= \frac{1}{2}\left\| \sum_x p_X(x)|x\rangle\langle x|_{X'} \otimes \left(\sigma^{x,m,\ell}_{Y'^{\otimes|\mathcal{L}||\mathcal{K}|}E} - \sigma^{x,m}_{Y'^{\otimes|\mathcal{L}||\mathcal{K}|}} \otimes \tilde{\sigma}^{x,m}_E \right) \right\|_1 = \sum_x p_X(x)\frac{1}{2}\left\| \sigma^{x,m,\ell}_{Y'^{\otimes|\mathcal{L}||\mathcal{K}|}E} - \sigma^{x,m}_{Y'^{\otimes|\mathcal{L}||\mathcal{K}|}} \otimes \tilde{\sigma}^{x,m}_E \right\|_1\\ &= \frac{1}{2}\sum_x p_X(x) \left\| \frac{1}{K}\sum_{k=1}^{|\mathcal{K}|} \sigma^{x,m,(\ell,1)}_{Y'} \otimes ... \otimes \sigma^{x,m,(\ell,k-1)}_{Y'} \otimes \left(\sigma^{x,m,(\ell,k)}_{Y'E} - \sigma^{x,m}_{Y'} \otimes \tilde{\sigma}^{x,m}_E \right) \otimes \sigma^{x,m,(\ell,k+1)}_{Y'} \otimes ... \otimes \sigma^{x,m,(|\mathcal{L}|,|\mathcal{K}|)}_{Y'} \right\|_1 .\end{aligned} \tag{28}$$

$P(\sigma^{x,m}_{YE}, \tilde{\sigma}^{x,m}_{YE}) \leq \sqrt{\epsilon'} - \delta'$. So far, we have shown the correctness of two separate criteria for the assisted code. For our purposes here we would like to have a single condition for the private message encompassing both conditions discussed lately and so in the following, by sticking to the reciepe set out by [3], we try to merge two conditions and deal with a single *privacy error*. We see that the single criterion will be beneficial once we derandomize the code and upon derandomization, the requirements set out in the definition of the unassisted code will be fulfilled.

We saw that the average error probability is equal to the individual error probabilities:

$$\begin{aligned}&\text{Tr}\{(\mathbb{1}_{XY^{|\mathcal{L}||\mathcal{K}|}B} - P^{(\ell,k)}_{XY^{|\mathcal{L}||\mathcal{K}|}B})\sigma^{m,(\ell,k)}_{XY^{\otimes|\mathcal{L}||\mathcal{K}|}B}\} =\\ &\frac{1}{|\mathcal{L}||\mathcal{K}|}\sum_{l=1}^{|\mathcal{L}|}\sum_{k=1}^{|\mathcal{K}|}\text{Tr}\{(\mathbb{1}_{Y^{|\mathcal{L}||\mathcal{K}|}B} - P^{(\ell,k)}_{XY^{|\mathcal{L}||\mathcal{K}|}B})\sigma^{m,(\ell,k)}_{XY^{\otimes|\mathcal{L}||\mathcal{K}|}B}\} \leq \epsilon\end{aligned} \tag{30}$$

We continue by expanding $\sigma^{m,(\ell,k)}_{XY^{\otimes|\mathcal{L}||\mathcal{K}|}B} = \sum_x p_X(x)|x\rangle\langle x|_X \otimes \sigma^{x,m,(\ell,k)}_{Y^{\otimes|\mathcal{L}||\mathcal{K}|}B}$ as in equation (31).

Reconsider the optimal test operator Θ_{YB}, we can write the following equation:

$$\begin{aligned}\text{Tr}\{\Theta_{YB}\sigma^x_{YB}\} &= \text{Tr}\{\Theta_{YB}(\sum_y p(y|x)|y\rangle\langle y| \otimes \sigma^{x,y}_B)\}\\ &= \sum_y p(y|x)\text{Tr}\{\langle y|\Theta^x_{YB}|y\rangle\sigma^{x,y}_B\}\\ &= \sum_y p(y|x)\text{Tr}\{G^{x,y}_B\sigma^{x,y}_B\},\end{aligned}$$

where $G^{x,y}_B \coloneqq \langle y|\Theta^x_{YB}|y\rangle$. In an analogous way:

$$\text{Tr}\{\Theta_{YB}(\sigma^x_Y \otimes \sigma^x_B)\} = \sum_y p(y|x)\text{Tr}\{G^{x,y}_B\sigma^x_B\}.$$

The above derivations lead the test operator to be considered as $\Theta_{YB} = \sum_y |y\rangle\langle y|_Y \otimes G^{x,y}_B$, 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 $N^{(\ell,k)}_{XY^{|\mathcal{L}||\mathcal{K}|}B}$ as given in (23) and expanded as in (32). Observe the structure of the POVM given in (33). And finally our POVM has the form given in equation (34). To build a POVM on the full space, we add $\Omega^0_B = \mathbb{1}_B - \sum_\ell\sum_k \Omega^{x,(\ell,k)}_B$ to the set $\{\Omega^{x,(\ell,k)}_B\}^{|\mathcal{L}|,|\mathcal{K}|}_{\ell=1,k=1}$. By combining (31) and (34), we find that

$$\begin{aligned}&\text{Tr}\{(\mathbb{1}_{Y^{|\mathcal{L}||\mathcal{K}|}B} - P^{(\ell,k)}_{XY^{|\mathcal{L}||\mathcal{K}|}B})\sigma^{m,(\ell,k)}_{XY^{\otimes|\mathcal{L}||\mathcal{K}|}B}\}\\ &\quad= \sum_{x,y_{11}\cdots y_{|\mathcal{L}||\mathcal{K}|}} p_X(x)p(y_{11}|x)...p(y_{|\mathcal{L}||\mathcal{K}|}|x)\\ &\qquad\qquad \times \text{Tr}\{(\mathbb{1}_B - \Omega^{x,y_{\ell k}}_B)\sigma^{x,m,y_{\ell k}}_B\}\end{aligned}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2019.2945800, IEEE Transactions on Information Theory

$$\begin{aligned}\sigma_{Y^{\otimes|\mathcal{L}||\mathcal{K}|}B}^{x,m,(\ell,k)} &= \sigma_{Y}^{x,m,(1,1)} \otimes ... \otimes \sigma_{Y}^{x,m,(1,|\mathcal{K}|)} \otimes ... \otimes \sigma_{Y}^{x,m,(\ell,k-1)} \otimes \sigma_{YB}^{x,m,(\ell,k)} \otimes \sigma_{Y}^{x,m,(\ell,k+1)} ... \otimes \sigma_{Y}^{x,m,(|\mathcal{L}|,|\mathcal{K}|)} \\ &= \sum_{y_{11}} p(y_{11}|x)|y_{11}\rangle\langle y_{11}| \otimes ... \otimes \sum_{y_{1|\mathcal{K}|}} p(y_{1|\mathcal{K}|}|x)|y_{1|\mathcal{K}|}\rangle\langle y_{1|\mathcal{K}|}| \otimes ... \otimes \sum_{y_{\ell k-1}} p(y_{\ell k-1})|y_{\ell k-1}\rangle\langle y_{\ell k-1}| \\ &\quad \otimes ... \otimes \sum_{y_{\ell k}} p(y_{\ell k}|x)|y_{\ell k}\rangle\langle y_{\ell k}| \otimes \sigma_{B}^{x,m,y_{\ell k}} \otimes \sum_{y_{\ell k+1}} p(y_{\ell k+1}|x)|y_{\ell k+1}\rangle\langle y_{\ell k+1}| ... \otimes \sum_{y_{|\mathcal{L}||\mathcal{K}|}} p(y_{|\mathcal{L}||\mathcal{K}|}|x)|y_{|\mathcal{L}||\mathcal{K}|}\rangle\langle y_{|\mathcal{L}||\mathcal{K}|}| \\ &= \sum_{y_{11}y_{12}...y_{lk}...y_{|\mathcal{L}||\mathcal{K}|}} p(y_{11}|x)p(y_{12}|x)...p(y_{\ell k}|x)...p(y_{|\mathcal{L}||\mathcal{K}|}|x)|y_{11}...y_{\ell k}...y_{|\mathcal{L}||\mathcal{K}|}\rangle\langle y_{11}...y_{\ell k}...y_{|\mathcal{L}||\mathcal{K}|}| \otimes \sigma_{B}^{x,m,y_{\ell k}},\end{aligned}$$

hence

$$\sigma_{XY^{\otimes|\mathcal{L}||\mathcal{K}|}B}^{m,(\ell,k)} = \sum_{x,y_{11}y_{12}...y_{lk}...y_{|\mathcal{L}||\mathcal{K}|}} p_{XY^{|\mathcal{L}||\mathcal{K}|}}(x,y_{11}...y_{lk}...y_{|\mathcal{L}||\mathcal{K}|})|x\rangle\langle x|_X \otimes |y_{11}...y_{\ell k}...y_{|\mathcal{L}||\mathcal{K}|}\rangle\langle y_{11}...y_{\ell k}...y_{|\mathcal{L}||\mathcal{K}|}| \otimes \sigma_{B}^{x,m,y_{\ell k}}. \tag{31}$$

$$\begin{aligned}N_{XY^{|\mathcal{L}||\mathcal{K}|}B}^{(\ell,k)} &= |x\rangle\langle x|_X \otimes \mathbb{1}_{Y}^{(1,1)} \otimes ... \otimes \Theta_{YB}^{(\ell,k)} \otimes ... \otimes \mathbb{1}_{Y}^{(|\mathcal{L}|,|\mathcal{K}|)} \\ &= \sum_{y_{11}...y_{lk}...y_{|\mathcal{L}||\mathcal{K}|}} |x\rangle\langle x|_X \otimes |y_{11}...y_{\ell k}...y_{|\mathcal{L}||\mathcal{K}|}\rangle\langle y_{11}...y_{\ell k}...y_{|\mathcal{L}||\mathcal{K}|}| \otimes G_{B}^{x,y_{\ell k}}.\end{aligned} \tag{32}$$

$$\begin{aligned}&\left(\sum_{\ell'}\sum_{k'} N_{XY^{|\mathcal{L}||\mathcal{K}|}B}^{(\ell',k')}\right)^{-\frac{1}{2}} \\ &= \sum_{y_{11}...y_{|\mathcal{L}||\mathcal{K}|}} |x\rangle\langle x|_X \otimes |y_{11}...y_{|\mathcal{L}||\mathcal{K}|}\rangle\langle y_{11}...y_{|\mathcal{L}||\mathcal{K}|}| \otimes \left(\sum_{\ell'}\sum_{k'} G_{B}^{x,y_{\ell'k'}}\right)^{-\frac{1}{2}},\end{aligned} \tag{33}$$

$$\begin{aligned}&P_{XY^{|\mathcal{L}||\mathcal{K}|}B}^{(\ell,k)} \\ &= \left(\sum_{\ell'}\sum_{k'} N_{XY^{|\mathcal{L}||\mathcal{K}|}B}^{(\ell',k')}\right)^{-\frac{1}{2}} N_{XY^{|\mathcal{L}||\mathcal{K}|}B}^{(\ell,k)} \left(\sum_{l'}\sum_{k'} N_{XY^{|\mathcal{L}||\mathcal{K}|}B}^{(\ell',k')}\right)^{-\frac{1}{2}} \\ &= \sum_{y_{11}...y_{|\mathcal{L}||\mathcal{K}|}} |x\rangle\langle x|_X \otimes |y_{11}...y_{|\mathcal{L}||\mathcal{K}|}\rangle\langle y_{11}...y_{|\mathcal{L}||\mathcal{K}|}| \otimes \Omega_{B}^{x,y_{\ell k}},\end{aligned} \tag{34}$$

where

$$\Omega_{B}^{x,(\ell,k)} := \left(\sum_{\ell'}\sum_{k'} G_{B}^{x,y_{\ell'k'}}\right)^{-\frac{1}{2}} G_{B}^{x,y_{\ell k}} \left(\sum_{\ell'}\sum_{k'} G_{B}^{x,y_{\ell'k'}}\right)^{-\frac{1}{2}}.$$

and from (30), the equality of the average and the individual error probabilities, yields equation (35).

By taking advantage of the POVMs $\{\Omega_{B}^{x,(\ell,k)}\}_{\ell=1,k=1}^{|\mathcal{L}|,|\mathcal{K}|}$, the following measurement channels are defined

$$\mathcal{D'}_{B\to\hat{L}}^{2}(\omega_B) := \sum_{\ell=1}^{|\mathcal{L}|}\sum_{k=1}^{|\mathcal{K}|} \mathrm{Tr}\{\Omega_{B}^{x,y_{\ell,k}}\omega_B\}|\ell\rangle\langle\ell|_{\hat{L}}, \tag{36}$$

$$\mathcal{D'}_{B\to\hat{L}\hat{K}}^{2}(\omega_B) := \sum_{\ell=1}^{|\mathcal{L}|}\sum_{k=1}^{|\mathcal{K}|} \mathrm{Tr}\{\Omega_{B}^{x,y_{\ell,k}}\omega_B\}|\ell\rangle\langle\ell|_{\hat{L}} \otimes |k\rangle\langle k|_{\hat{K}}, \tag{37}$$

where ω_B is a general quantum state and $\mathrm{Tr}_{\hat{K}} \circ \mathcal{D'}_{B\to\hat{L}\hat{K}}^{2} = \mathcal{D'}_{B\to\hat{L}}^{2}$. Note that in (36) the probability of getting a particular ℓ equals $\sum_{k=1}^{|\mathcal{K}|} \mathrm{Tr}\{\Omega_{B}^{x,y_{\ell k}}\omega_B\}$. By direct calculations, it can be seen that:

$$\begin{aligned}&\mathrm{Tr}\{(\mathbb{1}_B - \Omega_{B}^{x,y_{\ell,k}})\sigma_{B}^{x,m,y_{\ell,k}}\} \\ &\quad= \frac{1}{2}\left\|\mathcal{D'}_{B\to\hat{L}\hat{K}}^{2}(\sigma_{B}^{x,m,y_{\ell,k}}) - |\ell\rangle\langle\ell|_{\hat{L}} \otimes |k\rangle\langle k|_{\hat{K}}\right\|_1,\end{aligned}$$

averaging over ℓ, k and $(x, y_{1,1}...y_{|\mathcal{L}|,|\mathcal{K}|})$ and using (35), we get equation (38). 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).

Considering the POVM $\{\Omega_{B}^{x,y_{\ell k}}\}_{\ell=1,k=1}^{|\mathcal{L}|,|\mathcal{K}|}$, the probability of getting ℓ' conditioned on the fact that (ℓ, k) was sent is equal to $\sum_{k'=1}^{K} \mathrm{Tr}\{\Omega_{B}^{x,y_{\ell'k'}}\sigma_{B}^{x,y_{\ell k}}\}$ and it

$$\begin{aligned}
&\frac{1}{|\mathcal{L}|}\frac{1}{|\mathcal{K}|}\sum_{\ell=1}^{|\mathcal{L}|}\sum_{k=1}^{|\mathcal{K}|}\text{Tr}\{(\mathbb{1}_{Y^{|\mathcal{L}||\mathcal{K}|}B} - P^{(\ell,k)}_{XY^{|\mathcal{L}||\mathcal{K}|}B})\sigma^{m,(\ell,k)}_{XY^{\otimes|\mathcal{L}||\mathcal{K}|}B}\}\\
&\quad = \frac{1}{|\mathcal{L}|}\frac{1}{|\mathcal{K}|}\sum_{\ell=1}^{|\mathcal{L}|}\sum_{k=1}^{|\mathcal{K}|}\sum_{x,y_{11}\dots y_{|\mathcal{L}||\mathcal{K}|}} p_X(x)p(y_{11}|x)\dots p(y_{|\mathcal{L}||\mathcal{K}|}|x)\text{Tr}\{(\mathbb{1}_B - \Omega_B^{x,y_{\ell k}})\sigma_B^{x,m,y_{\ell k}}\}\\
&\quad = \sum_{x,y_{11}\dots y_{|\mathcal{L}||\mathcal{K}|}} p_X(x)p(y_{11}|x)\dots p(y_{|\mathcal{L}||\mathcal{K}|}|x)\left(\frac{1}{|\mathcal{L}|}\frac{1}{|\mathcal{K}|}\sum_{\ell=1}^{|\mathcal{L}|}\sum_{k=1}^{|\mathcal{K}|}\text{Tr}\{(\mathbb{1}_B - \Omega_B^{x,y_{\ell,k}})\sigma_B^{x,m,y_{\ell k}}\}\right) \leq \epsilon. \qquad (35)
\end{aligned}$$

$$\sum_{x,y_{11}\dots y_{|\mathcal{L}||\mathcal{K}|}} p_X(x)p(y_{11}|x)\dots p(y_{|\mathcal{L}||\mathcal{K}|}|x)\left[\frac{1}{|\mathcal{L}|}\frac{1}{|\mathcal{K}|}\sum_{\ell}\sum_{k}\frac{1}{2}\left\|\mathcal{D'}^2_{B\to\hat{L}\hat{K}}(\sigma_B^{x,m,y_{\ell k}}) - |\ell\rangle\langle\ell|_{\hat{L}}\otimes|k\rangle\langle k|_{\hat{K}}\right\|_1\right] \leq \epsilon. \qquad (38)$$

$$\begin{aligned}
&\sum_{x,y_{11},\dots,y_{|\mathcal{L}||\mathcal{K}|}} p_X(x)p(y_{11}|x)\dots p(y_{|\mathcal{L}||\mathcal{K}|}|x)\left[\frac{1}{|\mathcal{L}|}\sum_{\ell=1}^{|\mathcal{L}|}\frac{1}{2}\left\|(\text{Tr}_{\hat{K}}\circ\mathcal{D'}^2_{B\to\hat{L}\hat{K}})\left(\frac{1}{|\mathcal{K}|}\sum_{k=1}^{|\mathcal{K}|}\sigma_B^{x,m,y_{\ell k}}\right) - |\ell\rangle\langle\ell|_{\hat{L}}\right\|_1\right]\\
&\quad = \sum_{x,y_{11},\dots,y_{|\mathcal{L}||\mathcal{K}|}} p_X(x)p(y_{11}|x)\dots p(y_{|\mathcal{L}||\mathcal{K}|}|x)\left[\frac{1}{|\mathcal{L}|}\sum_{\ell=1}^{|\mathcal{L}|}\frac{1}{2}\left\|\mathcal{D'}^2_{B\to\hat{L}}\left(\frac{1}{|\mathcal{K}|}\sum_{k=1}^{|\mathcal{K}|}\sigma_B^{x,m,y_{\ell k}}\right) - |\ell\rangle\langle\ell|_{\hat{L}}\right\|_1\right] \leq \epsilon.\\
&\qquad (39)
\end{aligned}$$

is clear from the uniformity of the local key that the probability of getting l' given that l was sent, equals $Pr(l'|l) = \frac{1}{|\mathcal{K}|}\sum_{k=1}^{|\mathcal{K}|}\sum_{k'=1}^{|\mathcal{K}|}\text{Tr}\{\Omega_B^{x,y_{\ell' k'}}\sigma_{BE}^{x,m,y_{\ell k}}\} = \frac{1}{|\mathcal{K}|}\sum_{k=1}^{|\mathcal{K}|}\sum_{k'=1}^{|\mathcal{K}|}\text{Tr}\{\Omega_B^{x,y_{\ell' k'}}\sigma_B^{x,m,y_{\ell k}}\}$. Note that evidently $\sum_{\ell'=1}^{|\mathcal{L}|}Pr(\ell'|\ell) = 1$. If the trace above was only applied to the B system, we would have :

$$\frac{1}{|\mathcal{K}|}\sum_{k=1}^{|\mathcal{K}|}\sum_{k'=1}^{|\mathcal{K}|}\text{Tr}_B\{\Omega_B^{x,y_{\ell',k'}}\sigma_{BE}^{x,m,y_{\ell,k}}\} = Pr(\ell'|\ell)u_E^{\ell',\ell},$$

where

$$u_E^{\ell',\ell} := \frac{\frac{1}{|\mathcal{K}|}\sum_{k=1}^{|\mathcal{K}|}\sum_{k'=1}^{|\mathcal{K}|}\text{Tr}_B\{\Omega_B^{x,y_{\ell',k'}}\sigma_{BE}^{x,m,y_{\ell,k}}\}}{Pr(\ell'|\ell)}.$$

And by summing up over all ℓ' we get: (see that $\sum_{k'=1}^{|\mathcal{K}|}\sum_{\ell'=1}^{|\mathcal{L}|}\text{Tr}_B\{\Omega_B^{x,y_{\ell',k'}}\sigma_{BE}^{x,m,y_{\ell,k}}\} = \sigma_E^{x,m,y_{\ell,k}}$ for a given pair (ℓ,k))

$$\sum_{\ell'=1}^{|\mathcal{L}|}Pr(\ell'|\ell)u_E^{\ell',\ell} = \frac{1}{|\mathcal{K}|}\sum_{k=1}^{|\mathcal{K}|}\sigma_E^{x,m,y_{\ell,k}}.$$

Hence, the following equation follows:

$$\mathcal{D'}^2_{B\to\hat{L}}\left(\frac{1}{|\mathcal{K}|}\sum_{k=1}^{|\mathcal{K}|}\sigma_{BE}^{x,m,y_{\ell,k}}\right) = \sum_{\ell'=1}^{|\mathcal{L}|}Pr(\ell'|\ell)|\ell'\rangle\langle\ell'|_{\hat{L}}\otimes u_E^{\ell',\ell},$$

and by tracing out Eve's system:

$$\mathcal{D'}^2_{B\to\hat{L}}\left(\frac{1}{|\mathcal{K}|}\sum_{k=1}^{|\mathcal{K}|}\sigma_B^{x,m,y_{\ell,k}}\right) = \sum_{\ell'=1}^{|\mathcal{L}|}Pr(\ell'|\ell)|\ell'\rangle\langle\ell'|_{\hat{L}}.$$

We move forward with the chain of inequalities ending up in (40), 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). 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), 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 later 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) (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 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

$$\begin{aligned}
&\frac{1}{|\mathcal{L}|}\sum_{l=1}^{|\mathcal{L}|}\frac{1}{2}\left\|\mathcal{D'}^{2}_{B\to\hat{L}}\left(\frac{1}{|\mathcal{K}|}\sum_{k=1}^{|\mathcal{K}|}\sigma_{BE}^{x,m,y_{\ell,k}}\right)-|\ell\rangle\langle\ell|_{\hat{L}}\otimes\frac{1}{|\mathcal{K}|}\sum_{k=1}^{|\mathcal{K}|}\sigma_{E}^{x,m,y_{\ell,k}}\right\|_1\\
&\quad=\frac{1}{|\mathcal{L}|}\sum_{\ell=1}^{|\mathcal{L}|}\frac{1}{2}\left\|\sum_{\ell'=1}^{|\mathcal{L}|}Pr(\ell'|\ell)|\ell'\rangle\langle\ell'|_{\hat{L}}\otimes u_E^{\ell',\ell}-|\ell\rangle\langle\ell|_{\hat{L}}\otimes\sum_{\ell'=1}^{|\mathcal{L}|}Pr(\ell'|\ell)u_E^{\ell',\ell}\right\|_1\\
&\quad\leq\frac{1}{|\mathcal{L}|}\sum_{l=1}^{|\mathcal{L}|}\sum_{\ell'=1}^{|\mathcal{L}|}Pr(\ell'|\ell)\left[\frac{1}{2}\||\ell'\rangle\langle\ell'|_{\hat{L}}\otimes u_E^{\ell',\ell}-|\ell\rangle\langle\ell|_{\hat{L}}\otimes u_E^{\ell',\ell}\|_1\right]=\frac{1}{|\mathcal{L}|}\sum_{\ell=1}^{|\mathcal{L}|}\sum_{\ell'=1}^{|\mathcal{L}|}Pr(\ell'|\ell)\left[\frac{1}{2}\||\ell'\rangle\langle\ell'|_{\hat{L}}-|\ell\rangle\langle\ell|_{\hat{L}}\|_1\right]\\
&\quad=\frac{1}{|\mathcal{L}|}\sum_{\ell=1}^{|\mathcal{L}|}\sum_{\ell'\neq\ell}Pr(\ell'|\ell)=\frac{1}{|\mathcal{L}|}\sum_{\ell=1}^{|\mathcal{L}|}\frac{1}{2}\left\|\mathcal{D'}^{2}_{B\to\hat{L}}\left(\frac{1}{|\mathcal{K}|}\sum_{k=1}^{|\mathcal{K}|}\sigma_{B}^{x,m,y_{\ell,k}}\right)-|\ell\rangle\langle\ell|_{\hat{L}}\right\|_1, \qquad (40)
\end{aligned}$$

$$\sum_{x,y_{1,1},\dots,y_{|\mathcal{L}|,|\mathcal{K}|}}p_X(x)p(y_{1,1}|x)...p(y_{|\mathcal{L}|,|\mathcal{K}|}|x)\left[\frac{1}{|\mathcal{L}|}\sum_{\ell=1}^{|\mathcal{L}|}\frac{1}{2}\left\|\mathcal{D'}^{2}_{B\to\hat{L}}\left(\frac{1}{|\mathcal{K}|}\sum_{k=1}^{|\mathcal{K}|}\sigma_{BE}^{x,m,y_{\ell,k}}\right)-|\ell\rangle\langle\ell|_{\hat{L}}\otimes\frac{1}{|\mathcal{K}|}\sum_{k=1}^{|\mathcal{K}|}\sigma_{E}^{x,m,y_{\ell,k}}\right\|_1\right]\leq\epsilon. \qquad (41)$$

$$\begin{aligned}
&\frac{1}{2}\left\|\sigma^{m,\ell}_{XY^{\otimes|\mathcal{L}||\mathcal{K}|}E}-\sum_x p_X(x)|x\rangle\langle x|_X\otimes\sigma^{x,m}_{Y^{\otimes|\mathcal{L}||\mathcal{K}|}}\otimes\tilde{\sigma}_E^{x,m}\right\|_1\\
&=\frac{1}{2}\left\|\frac{1}{|\mathcal{K}|}\sum_{k=1}^{|\mathcal{K}|}\sum_{x,y_{11}\dots y_{|\mathcal{L}||\mathcal{K}|}}p_X(x)p_{Y|X}(y_{11}|x)...p_{Y|X}(y_{|\mathcal{L}||\mathcal{K}|}|x)|x\rangle\langle x|_X\otimes|y_{11}...y_{|\mathcal{L}||\mathcal{K}|}\rangle\langle y_{11}...y_{|\mathcal{L}||\mathcal{K}|}|_{Y^{|\mathcal{L}||\mathcal{K}|}}\otimes(\sigma_E^{x,m,y_{\ell k}}-\tilde{\sigma}_E^{x,m})\right\|_1\\
&=\frac{1}{2}\left\|\sum_{x,y_{11}\dots y_{\ell k}}p_X(x)p_{Y|X}(y_{11}|x)...p_{Y|X}(y_{|\mathcal{L}|,|\mathcal{K}|}|x)|x\rangle\langle x|_X\otimes|y_{11}...y_{|\mathcal{L}||\mathcal{K}|}\rangle\langle y_{11}...y_{|\mathcal{L}|,|\mathcal{K}|}|_{Y^{|\mathcal{L}||\mathcal{K}|}}\otimes\left(\frac{1}{|\mathcal{K}|}\sum_{k=1}^{|\mathcal{K}|}\sigma_E^{x,m,y_{\ell,k}}-\tilde{\sigma}_E^{x,m}\right)\right\|_1\\
&=\sum_{x,y_{11}\dots y_{|\mathcal{L}||\mathcal{K}|}}p_X(x)p_{Y|X}(y_{11}|x)...p_{Y|X}(y_{|\mathcal{L}||\mathcal{K}|}|x)\left[\frac{1}{2}\left\|\frac{1}{|\mathcal{K}|}\sum_{k=1}^{|\mathcal{K}|}\sigma_E^{x,m,y_{\ell k}}-\tilde{\sigma}_E^{x,m}\right\|_1\right]\\
&=\sum_{x,y_{11}\dots y_{|\mathcal{L}||\mathcal{K}|}}p_X(x)p_{Y|X}(y_{11}|x)...p_{Y|X}(y_{|\mathcal{L}||\mathcal{K}|}|x)\left[\frac{1}{2}\left\||\ell\rangle\langle\ell|_{L'}\otimes\frac{1}{|\mathcal{K}|}\sum_{k=1}^{|\mathcal{K}|}\sigma_E^{x,m,y_{\ell k}}-|\ell\rangle\langle\ell|_{L'}\otimes\tilde{\sigma}_E^{x,m}\right\|_1\right]\leq\sqrt{\epsilon'}, \qquad (42)
\end{aligned}$$

$$\sum_{x,y_{1,1},\dots,y_{|\mathcal{L}|,|\mathcal{K}|}}p_X(x)p_{Y|X}(y_{1,1}|x)...p_{Y|X}(y_{|\mathcal{L}|,|\mathcal{K}|}|x)\left[\frac{1}{2}\left\|\mathcal{D'}^{2}_{B\to\hat{L}}\left(\frac{1}{|\mathcal{K}|}\sum_{k=1}^{|\mathcal{K}|}\sigma_{BE}^{x,m,y_{\ell,k}}\right)-|\ell\rangle\langle\ell|_{\hat{L}}\otimes\tilde{\sigma}_E^{x,m}\right\|_1\right]\leq\epsilon+\sqrt{\epsilon'}, \qquad (43)$$

applies a quantum instrument depending on the estimate of the transmitted message m. This can, without losing the generality, be written as follows:

$$\mathcal{D}^2_{\hat{M}BY\to\hat{L}}:=\sum_m|m\rangle\langle m|_{M'}\otimes\mathcal{D}^{2,m}_{BY\to\hat{L}}.$$

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 ϵ, we know from gentle measurement lemma that:

$$\left\|\sigma^{m,(\ell,k)}_{XY^{\otimes|\mathcal{L}||\mathcal{K}|}B}-\rho^{m,(\ell,k)}_{XY^{\otimes|\mathcal{L}||\mathcal{K}|}B}\right\|_1\leq 2\sqrt{\epsilon},$$

and for the second term we have the chain of inequalities given by (44); 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 $P_{priv}\leq 2(\epsilon+\sqrt{\epsilon})+\sqrt{\epsilon'}$.

3) Derandomization: We can now fix the classical registers and obtain a protocol without shared randomness, i.e., deran-

$$\begin{aligned}
&\left\|(\sum_m |m\rangle\langle m|_{M'} \otimes \mathcal{D}^{2,m}_{BY\to\hat{L}})(\mathcal{D}^{1}_{BX\to\hat{M}B}(\rho^{m,(\ell,k)}_{X^{\otimes|\mathcal{M}|}B})) - (|m\rangle\langle m|_{M'} \otimes \mathcal{D}^{2,m}_{BY\to\hat{L}})(\mathcal{D}^{1}_{BX\to\hat{M}B}(\rho^{m,(\ell,k)}_{X^{\otimes|\mathcal{M}|}B}))\right\|_1 \\
&\quad= \left\|\sum_{m'\neq m} |m'\rangle\langle m'|_{M'} \otimes \mathcal{D}^{2,m'}_{BY\to\hat{L}}(\mathcal{D}^{1,m'}_{BX\to B}(\rho^{m,(\ell,k)}_{X^{\otimes|\mathcal{M}|}B}))\right\|_1 \\
&\quad\leq \sum_{m'\neq m} \left\||m'\rangle\langle m'|_{M'} \otimes \mathcal{D}^{2,m'}_{BY\to\hat{L}}(\mathcal{D}^{1,m'}_{BX\to B}(\rho^{m,(\ell,k)}_{X^{\otimes|\mathcal{M}|}B}))\right\|_1 \\
&\quad\leq \sum_{m'\neq m} \left\|\mathcal{D}^{1,m'}_{BX\to B}(\rho^{m,(\ell,k)}_{X^{\otimes|\mathcal{M}|}B})\right\|_1 \leq \epsilon. \qquad (44)
\end{aligned}$$

domize 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 $\mathcal{C}^{\epsilon,\epsilon'}(\mathcal{N})$.

Proof of theorem (3): Two messages $m \in \mathcal{M}$ and $\ell \in \mathcal{L}$ are sent through the channel $\mathcal{N}_{A\to BE}$ and their estimates are $\hat{M}$ and $\hat{L}$, respectively. From definition (18), an (ϵ,ϵ')-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

$$\rho_{MM'} = \frac{1}{|\mathcal{M}|}\sum_{m=1}^{|\mathcal{M}|} |m\rangle\langle m|_M \otimes |m\rangle\langle m|_{M'} \quad \text{and}$$

$$\rho_M \otimes \rho_{M'} = \frac{1}{|\mathcal{M}|}\sum_{m=1}^{|\mathcal{M}|} |m\rangle\langle m|_M \otimes \frac{1}{|\mathcal{M}|}\sum_{m=1}^{|\mathcal{M}|} |m\rangle\langle m|_{M'},$$

respectively. It is easily seen that type I error, i.e., deciding in favor of $\rho_M \otimes \rho_{M'}$ while the true state was $\rho_{MM'}$, is exactly equal to the error probability $Pr(M \neq \hat{M})$ which is less than or equal to ϵ by assumption. On the other hand, type II error, deciding $\rho_{MM'}$ on $\rho_M \otimes \rho_{M'}$, 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:

$$r \leq I_H^{\epsilon}(M;M')_\rho.$$

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

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

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:

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

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 $\rho^m_{L\hat{L}}$ and $\rho^m_L \otimes \rho^m_L$ conditioned on previously specified m, we will have:

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

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

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

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

$$R \leq I_H^{\sqrt{\epsilon}}(Y;B|X)_\rho. \qquad (45)$$

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

$$\frac{1}{2}\|\rho^m_{LE} - \rho_L \otimes \tilde{\rho}^m_E\|_1 \leq \epsilon',$$

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

$$P(\rho^m_{LE}, \rho_L \otimes \tilde{\rho}^m_E) \leq \sqrt{2\epsilon'}.$$

From the definition of the smooth max-relative entropy we see that $D_{max}^{\sqrt{2\epsilon'}}(\rho^m_{LE}, \rho_L \otimes \tilde{\rho}^m_E) = 0$. And by considering the optimization in Definition 17 over ρ'_M such that $P(\rho'_M, \rho_M) \le \sqrt{\epsilon'}$, we have $I_{max}^{\sqrt{2\epsilon'}}(L;E|M) = 0$. Setting $M := X$ and $L := Y$ as before and plugging into (45), the following bound on the private rate holds:

$$R \le I_H^{\sqrt{\epsilon}}(Y;B|X) - I_{max}^{\sqrt{2\epsilon'}}(Y;E|X). \qquad (46)$$

□

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:

$$\mathcal{C}_\infty(\mathcal{N}) := \lim_{\epsilon,\epsilon' \to 0} \lim_{n\to\infty} \frac{1}{n} \mathcal{C}^{\epsilon,\epsilon'}(\mathcal{N}^{\otimes n}).$$

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

$$\begin{aligned} r' &\le I(X;B)_\rho, \\ R' &\le I\left(Y;B|X\right)_\rho - I\left(Y;E|X\right)_\rho \end{aligned}$$

where all the entropic quantities are computed over all $\rho_{XYBE} := \sum_{x,y} p(x,y) \left|x\middle\rangle\middle\langle x\right| \otimes \left|y\middle\rangle\middle\langle y\right| \otimes \mathcal{N}_{A\to BE}(\rho_A^{x,y})$ arising from the channel. Then the capacity region $\mathcal{C}_\infty(\mathcal{N})$ is the (normalized) union over ℓ uses of the channel $\mathcal{N}$ as below:

$$\mathcal{C}_\infty(\mathcal{N}) = \bigcup_{\ell=1}^{\infty} \frac{1}{\ell} \mathcal{C}(\mathcal{N}^{\otimes \ell}). \qquad (47)$$

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:

$$\begin{aligned} \tilde{I}_{max}^{\sqrt{\epsilon'}-\delta'}(Y;E|X) \le I_{max}^{\sqrt{\epsilon'}-\delta'-\gamma}&(Y;E|X) \\ &+ \log_2\left(\frac{3}{\gamma^2}\right), \end{aligned}$$

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

$$\begin{aligned} R \ge I_H^{\epsilon-\delta}\left(Y;B|X\right) &- I_{max}^{\sqrt{\epsilon'}-\delta'-\gamma}(Y;E|X) \\ &- \log_2(\frac{4\epsilon}{\delta^2}) - 2\log_2(\frac{1}{\delta'}) - \log_2\left(\frac{3}{\gamma^2}\right). \end{aligned}$$

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 $\mathcal{C}_\infty(\mathcal{N})$ and the converse part that goes in the opposite direction saying that those rates cannot be exceeded, i.e., $\mathcal{C}_\infty(\mathcal{N})$ 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 $\mathcal{N}$ (or as one may like to think of it, one use of the superchannel $\mathcal{N}^{\otimes m}$), the following lower bound on the capacity region $\mathcal{C}^{\epsilon,\epsilon'}(\mathcal{N}^{\otimes m})$ can be seen:

$$\bigcup_{\ell=1}^{m} \mathcal{C}_a(\mathcal{N}^{\otimes \ell}) \subseteq \mathcal{C}^{3\epsilon+2\sqrt{\epsilon}+\sqrt{\epsilon'}, 2(\epsilon+\sqrt{\epsilon})+\sqrt{\epsilon'}}(\mathcal{N}^{\otimes m}),$$

where $\mathcal{C}_a(\mathcal{N}^{\otimes \ell})$ is the set of all twins (r', R') satisfying:

$$\begin{aligned} r' &\le I_H^{\epsilon-\delta}\left(X^\ell; B^{\otimes \ell}\right) - \log_2(\frac{4\epsilon}{\delta^2}), \\ R' &\le I_H^{\epsilon-\delta}\left(Y^\ell; B^{\otimes\ell}|X^\ell\right) - I_{max}^{\sqrt{\epsilon'}-\delta'-\gamma}(Y^\ell; E^{\otimes \ell}|X^\ell) \\ &\qquad - \log_2(\frac{4\epsilon}{\delta^2}) - 2\log_2(\frac{1}{\delta'}) - \log_2\left(\frac{3}{\gamma^2}\right). \end{aligned}$$

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

$$\lim_{\epsilon\to 0} \lim_{m\to\infty} \frac{1}{m} I_H^{\epsilon-\delta}\left(X^m; B^{\otimes m}\right) = I(X;B).$$

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

$$\begin{aligned} \lim_{\epsilon\to 0} \lim_{m\to\infty} \frac{1}{m} I_H^{\epsilon-\delta}\left(Y^m; B^{\otimes m}|X^m\right) &= I(Y;B|X), \\ \lim_{\epsilon'\to 0} \lim_{m\to\infty} \frac{1}{m} I_{max}^{\sqrt{\epsilon'}-\delta'-\gamma}(Y^m; E^{\otimes m}|X^m) &= I(Y;E|X). \end{aligned}$$

Plugging back into the respective equations, we obtain

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

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

$$\begin{aligned} r' &\le I(X;B)_\rho, \\ R' &\le I(Y;B|X)_\rho - I(Y;E|X)_\rho. \end{aligned}$$

Last step of the direct part is to consider a superchannel $\mathcal{N}^{\otimes \ell}$ (ℓ independent uses of the channel $\mathcal{N}$) 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 $\mathcal{N}$ and we have:

$$\mathcal{C}^{\epsilon,\epsilon'}(\mathcal{N}^{\otimes n}) \subseteq \bigcup_{n=1}^{\infty} \mathcal{C}_c(\mathcal{N}^{\otimes n})$$

where $\mathcal{C}_c(\mathcal{N}^{\otimes n})$ includes all ordered twins (r', R') satisfying

$$r' \le I_H^{\epsilon}\left(X^n; B^{\otimes n}\right), \qquad (48)$$

$$R' \le I_H^{\epsilon}(Y^n; B^{\otimes n}|X^n) - I_{max}^{\sqrt{2\epsilon'}}\left(E^{\otimes n}; Y^n|X^n\right). \qquad (49)$$

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 $|\mathcal{Y}|^n$. The inequalities are as follows:

$$
\begin{aligned}
r' &\leq \frac{1}{1-\epsilon}\left(I(X^n;B^{\otimes n})+h_b(\epsilon)\right),\\
R' &\leq \frac{1}{1-\epsilon}\left(I(Y^n;B^{\otimes n}|X^n)+h_b(\epsilon)\right)-I\left(E^{\otimes n};Y^n|X^n\right)\\
&\quad+2n\sqrt{2\epsilon'}\log|\mathcal{Y}|+2(1+\sqrt{2\epsilon'})h_b(\frac{\sqrt{2\epsilon'}}{1+\sqrt{2\epsilon'}}).
\end{aligned}
$$

Multiplying by $\frac{1}{n}$ and taking the limits $n\to\infty$ and $\epsilon,\epsilon'\to 0$, (changing n with ℓ) 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 of quantum states $\mathcal{E}=\{p_X(x),|\phi^x\rangle_{RA}\}_{x\in\mathcal{X}}$ 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 $\sigma_X=\sum_x p_X(x)|x\rangle\langle x|_X$, the following state can be associated to the ensemble:

$$
\sigma_{XRA}=\sum_x p_X(x)|x\rangle\langle x|_X\otimes|\phi^x\rangle\langle\phi^x|_{RA}. \tag{50}
$$

If channel $\mathcal{N}_{A\to BE}$ acts on this state, we get the following *coherent* state:

$$
\mathcal{N}_{A\to BE}(\sigma_{XRA})=\sum_x p_X(x)|x\rangle\langle x|_X\otimes|\phi^x\rangle\langle\phi^x|_{RBE},
$$

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

$$
\begin{aligned}
I(R\rangle BX) &\coloneqq -H(R|BX)=H(B|X)-H(RB|X)\\
&\overset{(a)}{=} H(B|X)-H(E|X),
\end{aligned}
$$

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 $\{|\phi^x\rangle_{RBE}\}_{x\in\mathcal{X}}$ with respect to the cut $R|BE$. Let $\{|y^x\rangle_R\}$ and $|\psi^{x,y}\rangle_{BE}$ be orthonormal bases for R and BE systems. Then from Schmidt decomposition we have that

$$
|\phi^x\rangle_{RBE}=\sum_y\sqrt{p_{Y|X}(y|x)}|y^x\rangle_R\otimes|\psi^{x,y}\rangle_{BE}.
$$

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

$$
\bar{\phi}^x_{YBE}=\sum_y p_{Y|X}(y|x)|y^x\rangle\langle y^x|_Y\otimes|\psi^{x,y}\rangle\langle\psi^{x,y}|_{BE},
$$

and let the decoherent state $\bar{\sigma}_{XRBE}$ be as follow:

$$
\begin{aligned}
\bar{\sigma}_{XYBE}=&\sum_x p_X(x)|x\rangle\langle x|_X\\
&\otimes\sum_y p_{Y|X}(y|x)|y^x\rangle\langle y^x|_Y\otimes|\psi^{x,y}\rangle\langle\psi^{x,y}|_{BE}.
\end{aligned}
$$

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,

$$
I(R\rangle BX)_\sigma=I(Y;B|X)_{\bar{\sigma}}-I(Y;E|X)_{\bar{\sigma}}. \tag{51}
$$

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

$$
\begin{aligned}
&I(Y;B|X)_{\bar{\sigma}}-I(Y;E|X)_{\bar{\sigma}}\\
&\overset{(a)}{=} H(B|X)-H(B|X,Y)-H(E|X)+H(E|X,Y)\\
&\overset{(b)}{=} H(B|X)-H(E|X),
\end{aligned}
$$

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 i.i.d. 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 Π_{XB} is such that $\text{Tr}\{\Pi_{XB}\rho_{XB}\}\geq 1-(\epsilon-\delta)$ and $\text{Tr}\{\Pi_{XB}(\rho_X\otimes\rho_B)\}=2^{-I_H^{\epsilon-\delta}(X;B)_\rho}$, we rewrite the two error types with slightly different notations as follows :

$$
\begin{aligned}
\text{Tr}\{\Pi_{XB}\rho_{XB}\}&=\text{Tr}\left\{\Pi_{XB}\sum_x p_X(x)|x\rangle\langle x|_X\otimes\rho_B^x\right\}\\
&=\sum_x p_X(x)\text{Tr}\{\langle x|\Pi_{XB}|x\rangle_X\rho_B^x\}\\
&=\sum_x p_X(x)\text{Tr}\{W_B^x\rho_B^x\},
\end{aligned}
$$

in which the operator W_B^x is defined as $W_B^x\coloneqq\langle x|\Pi_{XB}|x\rangle_X$. In an analogous way, we have that

$$
\begin{aligned}
\text{Tr}\{\Pi_{XB}(\rho_B\otimes\rho_X)\}&=\text{Tr}\left\{\Pi_{XB}\sum_x p_X(x)|x\rangle\langle x|_X\otimes\rho_B\right\}\\
&=\sum_x p_X(x)\text{Tr}\{\langle x|\Pi_{XB}|x\rangle_X\rho_B\}\\
&=\sum_x p_X(x)\text{Tr}\{W_B^x\rho_B\}.
\end{aligned}
$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2019.2945800, IEEE Transactions on Information Theory

These expressions imply that it is sufficient to take the optimal test to be $\Pi_{XB} = \sum_x |x\rangle\langle x|_X \otimes W_B^x$ with aforementioned W_B^x; In other words, the test Π_{XB} can achieve the same error probability as any other Π_{XB} would do. We proceed with dissecting each term involved in $\mathrm{Tr}\{(\mathbb{1}_{X^{|\mathcal{M}|}B} - \Lambda^m_{X^{|\mathcal{M}|}B})\rho^{m,(l,k)}_{X^{\otimes|\mathcal{M}|}B}\}$ where $\rho^{m,(l,k)}_{X^{\otimes|\mathcal{M}|}B}$ is given in (52). By assuming the particular structure for the optimal test operator that we just introduced, the operator $\Gamma^m_{X^{|\mathcal{M}|}B}$ appears as given in (53). And

$$\begin{aligned}
&\left(\sum_{m'=1}^{|\mathcal{M}|} \Gamma^{m'}_{X^{|\mathcal{M}|}B}\right)^{-\frac{1}{2}} \\
&= \left(\sum_{m'=1}^{|\mathcal{M}|} \sum_{x_1 \ldots x_{|\mathcal{M}|}} |x_1 \ldots x_{|\mathcal{M}|}\rangle\langle x_1 \ldots x_{|\mathcal{M}|}|_X \otimes W_B^{x_{m'}}\right)^{-\frac{1}{2}} \\
&= \left(\sum_{x_1 \ldots x_{|\mathcal{M}|}} |x_1 \ldots x_{|\mathcal{M}|}\rangle\langle x_1 \ldots x_{|\mathcal{M}|}|_X \otimes \sum_{m'=1}^{|\mathcal{M}|} W_B^{x_{m'}}\right)^{-\frac{1}{2}} \\
&= \sum_{x_1 \ldots x_{|\mathcal{M}|}} |x_1 \ldots x_{|\mathcal{M}|}\rangle\langle x_1 \ldots x_{|\mathcal{M}|}|_X \otimes \left(\sum_{m'=1}^{|\mathcal{M}|} W_B^{x_{m'}}\right)^{-\frac{1}{2}},
\end{aligned}$$

and finally

$$\begin{aligned}
\Lambda^m_{X^{|\mathcal{M}|}B} &= \left(\sum_{m'=1}^{|\mathcal{M}|} \Gamma^m_{X^{|\mathcal{M}|}B}\right)^{-\frac{1}{2}} \Gamma^{m'}_{X^{|\mathcal{M}|}B} \left(\sum_{m'=1}^{|\mathcal{M}|} \Gamma^{m'}_{X^{|\mathcal{M}|}B}\right)^{-\frac{1}{2}} \\
&= \sum_{x_1 \ldots x_{|\mathcal{M}|}} |x_1 \ldots x_{|\mathcal{M}|}\rangle\langle x_1 \ldots x_{|\mathcal{M}|}|_X \otimes \Delta^m_B,
\end{aligned}$$

where

$$\Delta^m_B := \left(\sum_{m'=1}^{|\mathcal{M}|} W_B^{x_{m'}}\right)^{-\frac{1}{2}} W_B^{x_m} \left(\sum_{m'=1}^{|\mathcal{M}|} W_B^{x_{m'}}\right)^{-\frac{1}{2}}.$$

Note that the obtained POVM, $\{\Delta^m_B\}_{m=1}^{|\mathcal{M}|}$, can be completed by adding $\Delta^0_B = \mathbb{1} - \sum_{m'=1}^{|\mathcal{M}|} \Delta^{m'}_B$. By putting everything that has derived so far into the error term, we will have:

$$\begin{aligned}
&\mathrm{Tr}\{(\mathbb{1}_{X^{|\mathcal{M}|}B} - \Lambda^m_{X^{|\mathcal{M}|}B})\rho^{m,(l,k)}_{X^{\otimes|\mathcal{M}|}B}\} \\
&\qquad = \sum_{x_1,\ldots,x_{|\mathcal{M}|}} p_X(x_1)\ldots p_{x_{|\mathcal{M}|}} \mathrm{Tr}\{(\mathbb{1}_B - \Delta^m_B)\rho_B^{x_m,(l,k)}\}.
\end{aligned}$$

By assuming a uniform distribution on the message set, averaging it over all messages results in

$$\begin{aligned}
&\frac{1}{|\mathcal{M}|}\sum_{m=1}^{|\mathcal{M}|} \mathrm{Tr}\{(\mathbb{1}_{X^{|\mathcal{M}|}B} - \Lambda^m_{X^{|\mathcal{M}|}B})\rho^{m,(l,k)}_{X^{\otimes|\mathcal{M}|}B}\} \\
&= \frac{1}{|\mathcal{M}|}\sum_{m=1}^{|\mathcal{M}|} \sum_{x_1,\ldots,x_{|\mathcal{M}|}} p_X(x_1)\ldots p_X(x_{|\mathcal{M}|}) \\
&\qquad\qquad \times \mathrm{Tr}\{(\mathbb{1}_B - \Delta^m_B)\rho_B^{x_m,(l,k)}\} \\
&= \sum_{x_1,\ldots,x_{|\mathcal{M}|}} p_X(x_1)\ldots p_X(x_{|\mathcal{M}|}) \\
&\qquad\qquad \times \left[\frac{1}{|\mathcal{M}|}\mathrm{Tr}\{(I_B - \Delta^m_B)\rho_B^{x_m,(l,k)}\}\right],
\end{aligned}$$

the last expression above shows averaging over all codebooks and we know that

$$\begin{aligned}
&\sum_{x_1,\ldots,x_{|\mathcal{M}|}} p_X(x_1)\ldots p_X(x_{|\mathcal{M}|}) \\
&\qquad\qquad \times \left[\frac{1}{|\mathcal{M}|}\mathrm{Tr}\{(\mathbb{1}_B - \Delta^m_B)\rho_B^{x_m,(l,k)}\}\right] \leq \epsilon,
\end{aligned}$$

which in turn, says that there exists at least one particular set of values of $\{x_1, \ldots x_{|\mathcal{M}|}\}$ such that

$$\frac{1}{|\mathcal{M}|}\sum_{m=1}^{|\mathcal{M}|} \mathrm{Tr}\{(I_B - \Delta^m_B)\rho_B^{x_m,(l,k)}\} \leq \epsilon. \tag{54}$$

This conclusion is known as the *Shannon trick*. The sequence $\{x_1 \ldots x_{|\mathcal{M}|}\}$ serves as the codebook used to transmit the public message.

As for the second part, we take (43) and average over all private messages as given in (55). And we again employ Shannon trick to conclude that there exists at least one sequence of values $(y_{1,1} \ldots y_{|\mathcal{L}|,|\mathcal{K}|}|x)$ such that equation (56) holds.

We can now argue that there exist values $(x_1 \ldots x_{|\mathcal{M}|})$ serving as *public codebook* for the transmission of the public message and conditioned on a particular codeword of the public codebook, there exist values $(y_{1,1} \ldots y_{|\mathcal{L}|,|\mathcal{K}|})$ serving as *private codebook* ensuring that the privacy criterion holds. Now we have a codebook of size $|\mathcal{M}||\mathcal{L}||\mathcal{K}|$, $\{x_1, \ldots, x_{|\mathcal{M}|}, y_1, \ldots, y_{|\mathcal{L}||\mathcal{K}|}\}$, that is publicly available serving as our deterministic codebook for simultaneous transmission of public and private messages.

Appendix B
One-shot Quantum Capacity: Imitating Devetak's Asymptotic Proof

As we mentioned in the introduction, a one-shot version of Devetak's asymptotic proof of quantum capacity follows along the same lines [27] . Here we briefly outline the proof and the general idea. We shall freely use the notation introduced so far. We have now seen that there exists a good codebook $\{x(\ell, k)\}_{\ell\in\mathcal{L},k\in\mathcal{K}}$ 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 $\mathcal{N}_{A\to BE}$, Bob is able to reliably work out both Alice's message ℓ and the local key k and at the same time, Eve happens to learned very little about Alice's message ℓ. This holds true for the rates of the private message $\log_2|\mathcal{L}|$ and the local randomness $\log_2|\mathcal{L}|$ 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 $f : x \to f(x)$ 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

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2019.2945800, IEEE Transactions on Information Theory

$$\rho^{m,(l,k)}_{X^{\otimes|\mathcal{M}|}B} \coloneqq \rho^{1}_{X} ... \otimes \rho^{m,(l,k)}_{XB} \otimes ... \otimes \rho^{|\mathcal{M}|}_{X} = \sum_{x_1,\dots,x_{|\mathcal{M}|}} p_X(x_1)...p_X(x_{|\mathcal{M}|}) |x_1....x_{|\mathcal{M}|}\rangle\langle x_1...x_{|\mathcal{M}|}|_{X_1...X_{|\mathcal{M}|}} \otimes \rho^{m,(l,k)}_{B}. \tag{52}$$

$$\Gamma^{m}_{X^{|\mathcal{M}|}B} = \mathbb{1}^{1}_{X} \otimes ... \otimes T^{m}_{XB} \otimes ... \otimes \mathbb{1}^{|\mathcal{M}|}_{X} = \sum_{x_1} |x_1\rangle\langle x_1|_X \otimes ... \otimes \left(\sum_{x_m} |x_m\rangle\langle x_m|_X \otimes W^{x_m}_{B} \right) \otimes ... \otimes \sum_{x_{|\mathcal{M}|}} |x_{|\mathcal{M}|}\rangle\langle x_{|\mathcal{M}|}|_X$$
$$= \sum_{x_1 \dots x_{|\mathcal{M}|}} |x_1 ... x_{|\mathcal{M}|}\rangle\langle x_1...x_{|\mathcal{M}|}|_X \otimes W^{x_m}_{B}. \tag{53}$$

$$\epsilon + \sqrt{\epsilon'} \geq \frac{1}{|\mathcal{L}|} \sum_{\ell=1}^{|\mathcal{L}|} \sum_{x, y_{11}, \dots, y_{|\mathcal{L}||\mathcal{K}|}} p_X(x) p_{Y|X}(y_{11}|x) ... p_{Y|X}(y_{|\mathcal{L}||\mathcal{K}|}|x) \left[\frac{1}{2} \left\| \mathcal{D'}^{2}_{B\to\hat{L}} \left(\frac{1}{|\mathcal{K}|} \sum_{k=1}^{|\mathcal{K}|} \sigma^{x,m,y_{\ell k}}_{BE} \right) - |\ell\rangle\langle\ell|_{\hat{L}} \otimes \tilde{\sigma}^{x,m}_{E} \right\|_1 \right]$$
$$= \sum_{x, y_{11}, \dots, y_{|\mathcal{L}||\mathcal{K}|}} p_X(x) p_{Y|X}(y_{11}|x) ... p_{Y|X}(y_{|\mathcal{L}||\mathcal{K}|}|x) \left(\frac{1}{|\mathcal{L}|} \sum_{l=1}^{|\mathcal{L}|} \left[\frac{1}{2} \left\| \mathcal{D'}^{2}_{B\to\hat{L}} \left(\frac{1}{|\mathcal{K}|} \sum_{k=1}^{|\mathcal{K}|} \sigma^{x,m,y_{\ell,k}}_{BE} \right) - |\ell\rangle\langle\ell|_{\hat{L}} \otimes \tilde{\sigma}^{x,m}_{E} \right\|_1 \right] \right). \tag{55}$$

$$\frac{1}{|\mathcal{L}|} \sum_{l=1}^{|\mathcal{L}|} \left[\frac{1}{2} \left\| \mathcal{D'}^{2}_{B\to\hat{L}} \left(\frac{1}{|\mathcal{K}|} \sum_{k=1}^{|\mathcal{K}|} \sigma^{x,m,y_{\ell,k}}_{BE} \right) - |\ell\rangle\langle\ell|_{\hat{L}} \otimes \tilde{\sigma}^{x,m}_{E} \right\|_1 \right] \leq \epsilon + \sqrt{\epsilon'}. \tag{56}$$

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 $\{|\ell\rangle_{A_1}\}_{\ell\in\mathcal{L}}$ of the Hilbert space. Suppose that Alice shares a state $|\varphi\rangle_{RA_1}$ with a reference system R:

$$|\varphi\rangle_{RA_1} \coloneqq \sum_{i,\ell\in\mathcal{L}} \alpha_{i,\ell} |i\rangle_R |\ell\rangle_{A_1},$$

where $|i\rangle_R$ 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:

$$|\varphi\rangle_{RB} \coloneqq \sum_{i,\ell\in\mathcal{L}} \alpha_{i,\ell} |i\rangle_R |\ell\rangle_{B},$$

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

$$|\phi^\ell\rangle_A \coloneqq \frac{1}{\sqrt{|\mathcal{K}|}} \sum_{k\in\mathcal{K}} e^{\gamma_{\ell,k}} \left|\psi^{x(\ell,k)}\right\rangle_A,$$

where the states $\left|\psi^{x(\ell,k)}\right\rangle_A$ 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 ℓ in register A_1 to another register A_2. She then applies some isometric encoding from A_2 register to A. These two steps are performed with the following map:

$$\left(\sum_{\ell} |\ell\rangle\langle\ell|_{A_2} \otimes |\phi^\ell\rangle_A \right) \left(\sum_{\ell} |\ell\rangle\langle\ell|_{A_1} \otimes |\ell\rangle_{A_2} \right),$$

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

$$\sum_{i,\ell\in\mathcal{L}} \alpha_{i,\ell} |i\rangle_R |\ell\rangle_{A_1} |\phi^\ell\rangle_{BE}.$$

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

$$\sum_{\ell,k} \sqrt{\Omega^{\ell,k}_{B}} \otimes |\ell\rangle_{B_1} |k\rangle_{B_2}.$$

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

$$\sum_{i,\ell} \sum_{k} \frac{1}{\sqrt{|\mathcal{K}|}} \alpha_{i,\ell} |i\rangle_R |\ell\rangle_{A_1} e^{\delta_{\ell,k}} \left|\phi^{x(\ell,k)}\right\rangle_{BE} |\ell\rangle_{B_1} |k\rangle_{B_2}.$$

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:

$$\sum_{i,\ell} \alpha_{i,\ell} |i\rangle_R |\ell\rangle_{A_1} |\ell\rangle_{B_1}.$$

So far they have successfully implemented an approximate coherent channel from systems A_1 to A_1B_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 than 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. FS acknowledges partial financial support by the Baidu-UAB collaborative project 'Learning of Quantum Hidden Markov Models', the Spanish MINECO (project FIS2016-86681-P) with the support of FEDER funds, and the Generalitat de Catalunya (project 2017-SGR-1127). M.-H. Hsieh was supported by an ARC Future Fellowship under Grant FT140100574 and by US Army Research Office for Basic Scientific Research Grant W911NF-17-1-0401. The work of F. Salek and J. R. Fonollosa has been funded by the "Ministerio de Economa, Industria y Competitividad" of the Spanish Government, ERDF funds (TEC2013-41315-R, TEC2015-69648-REDC, TEC2016-75067-C4-2-R) and the Catalan Government (2017 SGR 578 AGAUR). A. Anshu and R. Jain are supported by the Singapore Ministry of Education through the Tier 3 Grant "Random numbers from quantum processes" MOE2012-T3-1-009. R. Jain is also supported by the VAJRA Grant, Department of Science and Technology, Government of India and by the NRF2017-NRF-ANR004 VanQuTe grant. 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, no. 20, p. 200501, 2012.

[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," 2017, arXiv:1702.01940, 2017.

[3] M. M. Wilde, "Position-based coding and convex splitting for private communication over quantum channels," *Quantum Inf. Process*, vol. 16, no. 10, p. 264, 2017.

[4] H. Umegaki, "Conditional expectations in an operator algebra IV (entropy and information)," *Kodai Math. Sem. Rep.*, vol. 14, 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, 2007.

[6] A. Winter, "Coding theorem and strong converse for quantum channels," *IEEE Trans. Inf. Theory,* vol. 45, no. 7, pp. 2481-2485, 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. 18981929, April 2009.

[8] Mark M. Wilde, Marco Tomamichel, and Mario Berta, "Converse bounds for private communication over quantum channels," *IEEE Trans. Inf. Theory,* vol. 63, no. 3, pp. 17921817, March 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, 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, 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. 60, pp. 1537-1581, 2014.

[12] Mark M. Wilde, *Quantum Information Theory.* Cambridge University press, second edition, february 2017.

[13] K. Li, "Second order asymptotics for quantum hypothesis testing," *Ann. Statist.,* vol. 42, pp. 171-189, 2014.

[14] M. Tomamichel and M. Hayashi, "A hierarchy of information quantities for finite block length analysis of quantum tasks," *IEEE Transactions on Information Theory*, 59(11):76937710, November 2013.

[15] F. Buscemi, N. Datta, "The quantum capacity of channels with arbitrary correlated noise," *IEEE Trans. Inf. Theory,* vol. 56, No. 3, March 2010.

[16] F. Salek, A. Anshu, M.-H. Hsieh, R. Jain, J. R. Fonollosa, "One-shot Capacity Bounds on the Simultaneous Transmission of Public and Private Information Over Quantum Channels", *IEEE. Int. Symp. Information Theory,* CO, USA, 2018.

[17] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik,* Springer, Berlin, 1932.

[18] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory,* vol. 24, no. 3, pp. 339-348, 1978.

[19] I. Devetak, P. Shor, "The capaity of a quantum channel for simultanous transmission of classical and quantum information," *Commmun. Math. Phys.* 256, 287-303, 2005.

[20] Min-Hsiu Hsieh and M. M. Wilde, "Public and private communication with a quantum channel and a secret key," *Phys. Rev. A,* vol. 8, Iss. 2, 2009.

[21] Min-Hsiu Hsieh and M. M. Wilde, "Public and private resource tradeoffs for a quantum channel," *Quantum Inf. Process*, vol. 11, Iss. 6, p. 14651501, 2012.

[22] H. Qi, Q. Wang and M. M. Wilde, "Applications of position-based decoding to classical communication over quantum channels," *Journal of Physics A: Mathematical and Theoretical,* vol. 51, no. 44, 2018.

[23] C. E. Shannon, "A mathematical theory of communication," *Bell system Tech. J.*, Vol. 27, pp. 379-656, 1948.

[24] A. S. Holevo, "the capacity of a quantum channel with general signal states," *IEEE. Trans. Inf. Theory,* vol. 44, no. 1, pp. 269-273, Jan. 1998.

[25] B. Schumacher and M. D. Westmoreland, "Sending classical information via noisy quantum channels," *Phys. Rev. Lett.,* vol. 56, no. 1, p. 131, Jul. 1997.

[26] A. S. Holevo, "Statistical problems problems in quantum physics," *In proceedings of the second Japan-USSR symposium on probability theory,* ser. Lecture notes in mathematics, G. Maruyama and J. V. Prokhorov, Eds., vol. 330. berlin: Springer-Verlag, pp. 104-119, 1973.

[27] I. Devetak, "The private classical capacity and quantum capacity of a quantum channel," *IEEE. Trans. Inf. Theory,* vol. 51, pp. 44-55, 2005.

[28] N. Cai, A. Winter and R. Yeung, "Quantum privacy and quantum wiretap channels," *problems of information transmission,* vol. 40, no. 4, pp. 1613-1622, 1997.

[29] S. Lloyd, "Capacity of the noisy quantum channel," *Phys, Rev. A,* vol. 55, p. 1613, 1996.

[30] P. W. Shor, "The quantum channel capacity and coherent information," in *MSRI Seminar,* Nov. 2002, unpublished.

[31] T. S. Han and S, Verdu, "Approximation theory of output statistics," *IEEE. Trans. Inf. Theory.,* vol. 39, no. 3, pp. 752-772, 1993.

[32] M. Hayashi and H. Nagaoka, "General formulas for capacity of classical-quantum channels," *IEEE. Trans. Inf. Theory,* vol. 49, no. 7, pp. 1753-1768, 2003.

[33] R. Renner, S. Wolf, and J. Wullschleger, "The signle-serving channel capacity," in *Proc. IEEE. Int. Symp. Information Theory,* pp. 1424-1427, 2006.

[34] M. Mosonyi and N. Datta, "Generalized relative entropies and the capacity of classical-quantum channel," *J. Mathematical Physics,* vol. 15, no. 7, pp. 072104-14, 2009.

[35] J. M. Renes and R. Renner, "Noisy channel coding via privacy amplification and information reconciliation," *IEEE. Trans. Inf. Theory,* vol. 57, no. 11, 2011.

[36] J. Radhakrishnan, P. Sen and N. A. Warsi, "One-shot private classical capacity of quantum wiretap channel: based on one-shot quantum covering lemma," ,arXiv: 1703.01932v1, 2017.

[37] R. Renner, "Security of quantum key distribution," Ph.D. dissertation, ETH Zürich, Zürich, Switzerland, 2005.

[38] M. Tomamichel, "A framework for non-asymptotic quantum information theory," Ph.D. dissertation, ETH Zürich, Zürich, Switzerland, 2012.

[39] H. Qi, Q. Wang and M. M. Wilde, "Applications of position-based decoding to classical communication over quantum channels", arXiv: 1704.01361, 2017.

[40] N. Datta, "Min- max-relative entropies and a new entanglement monotone", *IEEE. Trans. Inf. Theory,* vol. 59, pp. 2816-2816, 2009.

[41] A. Anshu, V. K. Devabathini and R. Jain, "Quantum message compression with application", *Phys. Rev. Lett.*, vol. 119, p. 120506, 2017.

[42] R. Ahlswede and A. Winter, "Strong converse for identification via quantum channels," *IEEE Trans. Inf. Theory,* vol. 48, pp. 569-579, 2002.

[43] T. Cover, J. Thomas, *Elements of Information Theory*, 2nd, ed. New York: Wiley-Interscience, 2006.

[44] A. Gilchrist, N. Langford, and M. Nielsen, "Distance Measures to Compare Real and Ideal Quantum Processes", *Phys. Rev. A.*, vol. 71(6), pp. 062310, 2005.

[45] A. Uhlmann, "The Transition Probability for States of Star-Algebras", *Ann. Phys.,* vol. 497(4), pp. 524532, 1985.

[46] A. Winter, "Tight uniform continuity bounds for quantum entropies: conditional entropy, relative entropy distance and energy constraints", *Commun. Math. Phys.*, 347(1), 291313, 2016.

[47] M. E. Shirokov, "Tight continuity bounds for the quantum conditional mutual information, for the Holevo quantity and for capacities of quantum channels", *J. Math. Phys.*, 58, 102202, 2017.

[48] N. Datta, T. Dorlas, "The coding theorem for a class of quantum channels with long-term memory", *J. Phys. A: Math. Gen.*, vol. 40, pp. 8147-8164, 2007.

[49] T. Dorlas, C. Morgan, "The invalidity of a strong capacity for a quantum channel with memory,", *Phys. Rev. A*, 84, 042318, 2011.

[50] I. Devetak, A. Harrow and A. Winter, "A Resource Framework for Quantum Shannon Theory ," *IEEE Trans. Inf. Theory,* vol. 54, no. 10, pp. 4587-4618, 2008.

[51] F. Salek, Min-Hsiu Hsieh, and J. R. Fonollosa, "Publicness, Privacy and Confidentiality in the Single-Serving Quantum Broadcast Channel," in *Proc. IEEE. Int. Symp. Information Theory,* pp. 1712-1716, July 2019.

[52] F. Salek, Min-Hsiu Hsieh, and J. R. Fonollosa, "Publicness, Privacy and Confidentiality in the Single-Serving Quantum Broadcast Channel,", arXiv: 1903.04463, 2019.