2-7-2020

Conditional Quantum One-Time Pad

Kunal Sharma
Louisiana State University

Eyuri Wakakuwa
The University of Electro-Communications

Mark M. Wilde
Louisiana State University

Recommended Citation
Sharma, K., Wakakuwa, E., & Wilde, M. (2020). Conditional Quantum One-Time Pad. *Physical Review Letters, 124* (5) https://doi.org/10.1103/PhysRevLett.124.050503

Conditional quantum one-time pad

Kunal Sharma,[1] Eyuri Wakakuwa,[2] and Mark M. Wilde[1]

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

[2]*Graduate School of Informatics and Engineering, University of Electro-Communications, 1-5-1 Chofugaoka, Chofu-shi, Tokyo, 182-8585, Japan*

(Dated: February 11, 2020)

Suppose that Alice and Bob are located in distant laboratories, which are connected by an ideal quantum channel. Suppose further that they share many copies of a quantum state ρ_{ABE}, such that Alice possesses the A systems and Bob the BE systems. In our model, there is an identifiable part of Bob's laboratory that is insecure: a third party named Eve has infiltrated Bob's laboratory and gained control of the E systems. Alice, knowing this, would like to use their shared state and the ideal quantum channel to communicate a message in such a way that Bob, who has access to the whole of his laboratory (BE systems), can decode it, while Eve, who has access only to a sector of Bob's laboratory (E systems) and the ideal quantum channel connecting Alice to Bob, cannot learn anything about Alice's transmitted message. We call this task the conditional one-time pad, and in this paper, we prove that the optimal rate of secret communication for this task is equal to the conditional quantum mutual information $I(A;B|E)$ of their shared state. We thus give the conditional quantum mutual information an operational meaning that is different from those given in prior works, via state redistribution, conditional erasure, or state deconstruction. We also generalize the model and method in several ways, one of which is a secret-sharing task, i.e., the case in which Alice's message should be secure from someone possessing only the AB or AE systems but should be decodable by someone possessing all systems A, B, and E.

Introduction—This paper shows that the optimal rate of a communication task, which we call the conditional one-time pad, is equal to a fundamental information quantity called the conditional quantum mutual information. To prove this statement, we operate in the regime of quantum Shannon theory [1–3], supposing that Alice and Bob possess a large number n of copies of a quantum state ρ_{ABE}. We suppose that one party Alice has access to all of the A systems, and another party Bob has access to all of the BE systems. We suppose that Bob's laboratory is divided into two parts, one of which is secure (the B part) and the other which is insecure (the E part) and accessible to an eavesdropper Eve. We also suppose that Alice and Bob are connected by an ideal quantum channel, but the eavesdropper Eve can observe any quantum system that is transmitted over the ideal channel if she so desires. The goal of a conditional quantum one-time pad protocol is for Alice to encode a message m into her A systems, in such a way that if she sends her A systems over the ideal quantum channel, then

1. Bob can decode the message m reliably by performing a measurement on all of the ABE systems, while
2. an eavesdropper possessing the AE systems has essentially no chance of determining the message m if she tried to figure it out.

We prove that the optimal asymptotic rate at which this task can be accomplished is equal to the conditional quantum mutual information of the state ρ_{ABE}, defined as

$$I(A;B|E)_\rho \equiv I(A;BE)_\rho - I(A;E)_\rho, \qquad (1)$$

where the quantum mutual information of a state σ_{FG} is defined as $I(F;G)_\sigma \equiv H(F)_\sigma + H(G)_\sigma - H(FG)_\sigma$, with $H(F)_\sigma \equiv -\operatorname{Tr}\{\sigma_F \log_2 \sigma_F\}$ denoting the quantum entropy of the reduced state σ_F.

Our main result thus gives an operational meaning to the conditional quantum mutual information (CQMI) that is conceptually different from those appearing in prior works [4–7]. CQMI has previously been interpreted as the optimal rate of quantum communication from a sender to a receiver to accomplish the task of state redistribution [4, 5], in which the goal is for a sender to transmit one of her systems to a receiver who possesses a system correlated with the systems of the sender. CQMI has also been interpreted as the optimal rate of noise needed to accomplish the task of conditional erasure or state deconstruction [6, 7], in which (briefly) the goal is to apply noise to the AE systems of $\rho_{ABE}^{\otimes n}$ such that the resulting A systems are locally recoverable from the E systems alone while the marginal state $\rho_{BE}^{\otimes n}$ is negligibly disturbed. Recently, the dynamic counterpart of CQMI has been interpreted as the optimal rate of entanglement-assisted private communication over quantum broadcast channels [8], which is inspired by the conditional one-time pad protocol presented in this work.

The conditional mutual information is an information quantity that plays a central role in quantum information theory. The fact that it is non-negative for any quantum state is non-trivial and known as the strong subadditivity of quantum entropy [9, 10]. The strong subadditivity inequality is at the core of nearly every coding theorem in quantum information theory (see, e.g., [1–3]). The CQMI is also the information quantity underlying an entanglement measure called squashed entanglement

[11], a quantum correlation measure called quantum discord [12, 13] (as shown in [14]), and a steering quantifier called intrinsic steerability [15]. The CQMI is also a witness of Markovianity in the sense that if $I(A;B|E)$ is small, then the correlations between systems A and B are mediated by the system E via a recovery channel from E to AE [16]. Moreover, the CQMI of three regions with a non-trivial topology leads to the topological entanglement entropy of the system, which essentially characterizes irreducible many-body correlation [17–19]. The CQMI is thus an important information quantity to study quantum correlations in condensed matter systems (see, e.g., [20]). Furthermore, in the context of thermodynamics, the CQMI has been used to establish that the free fermion non-equilibrium steady state is an approximate quantum Markov chain [21]. The CQMI also plays an important role in high energy physics [22–24].

The basic intuition for the achievability of the conditional mutual information for the conditional one-time pad task is obtained by inspecting the expansion in (1) and is as follows: the authors of [25] showed that the quantum mutual information of a bipartite state is equal to the optimal rate of a task they called the (unconditional) quantum one-time pad. In our setting, the result of [25] implies that Alice can communicate a message secure against an eavesdropper, who can observe only the A systems, such that Bob, in possession of the BE systems, can decode it reliably, as long as the number of messages is $\approx nI(A;BE)_\rho$ bits. Here, we show that the message of Alice can be secured against an eavesdropper having access to both the A and E systems if Alice sacrifices $\approx nI(A;E)_\rho$ bits of the message, such that the total number of bits of the message is $\approx nI(A;BE)_\rho - nI(A;E)_\rho = nI(A;B|E)_\rho$, where we have employed (1). The main idea for a code construction to accomplish the above task is the same as that for the classical wiretap channel [26], which has been extended in a certain way to the quantum case in [27, 28]. To prove the achievability part of the main result of our paper, we use a coding technique developed in [29, Section III-A] and which was rediscovered shortly thereafter in [25] and later used in [30]. We also employ tools known as the quantum packing and covering lemmas (see, e.g., [3]). To establish optimality of the CQMI for the conditional one-time pad task, we employ entropy inequalities. We note that the aforementioned methods also lead to a proof of the main result of [31], which concerns a kind of quantum one-time pad protocol different from that developed in [25] or the present paper.

A modification of the coding structure for the conditional one-time pad protocol allows us to establish that the following information quantity

$$I(A;BE)_\rho - \max\{I(A;B)_\rho, I(A;E)_\rho\} \tag{2}$$

of a tripartite state ρ_{ABE} is an optimal achievable rate for a particular secret-sharing task that we call *information scrambling*. In this modified task, we suppose that Alice, Bob, and Eve are three distinct parties. Alice's laboratory is distant from Bob and Eve's, but we imagine that Bob and Eve's laboratories are close together, and an ideal quantum channel connects Alice's laboratory to Bob and Eve's. The goal of the information scrambling task is for Alice to communicate a message in such a way that it can be decoded only by someone who possesses all three ABE systems. If someone possesses only the AB systems or only the AE systems, then such a person can figure out essentially nothing about the encoded message.

Our finding here shows that the quantity in (2) is an optimal achievable rate for information scrambling, such that the message is encoded in the non-local degrees of freedom of $\rho_{ABE}^{\otimes n}$ and cannot be decoded exclusively from the local degrees of freedom, which in this case are constituted by systems AB or systems AE.

The rest of our paper proceeds as follows. We first formally define the conditional one-time pad task. We then sketch a proof for the achievability part of our result. We finally discuss variations of the main task, such as the information scrambling task mentioned above and more general tasks, and then we conclude with a brief summary.

The supplementary material provides a detailed proof of the achievability part of our main result. It also establishes the optimality part of our main result: that Alice cannot communicate at a rate higher than the conditional mutual information $I(A;B|E)$ while still satisfying the joint demands of reliable decoding for Bob (who gets the ABE systems) and security against an eavesdropper who has access to the AE systems. The optimality proof is based on entropy inequalities and identities.

Conditional quantum one-time pad—We use notation and concepts standard in quantum information theory and point the reader to [3] for further background. Let $n, M \in \mathbb{N}$ and let $\varepsilon, \delta \in [0,1]$. An $(n, M, \varepsilon, \delta)$ conditional one-time pad protocol begins with Alice and Bob sharing n copies of the state ρ_{ABE}, so that their state is $\rho_{ABE}^{\otimes n}$. As mentioned previously, Bob has access to the BE systems, but we consider the E systems to be insecure and jointly accessible by an eavesdropper. Alice and Bob are connected by an ideal quantum channel, which Eve has access to as well (later we argue that it suffices for Alice and Bob to use only $\approx nH(A)_\rho$ ideal qubit channels, but for now we suppose that the ideal quantum channel can transmit as many qubits as desired). At the beginning of the protocol, Alice picks a message $m \in \{1, \ldots, M\}$ and applies an encoding channel $\mathcal{E}^m_{A^n \to A'}$ to the A^n systems of $\rho_{ABE}^{\otimes n}$, leading to the state $\omega^m_{A'B^nE^n} \equiv \mathcal{E}^m_{A^n \to A'}(\rho_{ABE}^{\otimes n})$. She transmits the system A' of $\omega^m_{A'B^nE^n}$ over the ideal quantum channel. Bob applies a decoding positive operator-valued measure $\{\Lambda^m_{A'B^nE^n}\}_m$ to the systems $A'B^nE^n$ of $\omega^m_{A'B^nE^n}$ in order to figure out which message was transmitted. The protocol is ε-reliable if Bob can determine the message m with probability not smaller than $1-\varepsilon$:

$$\forall m : \mathrm{Tr}\{\Lambda^m_{A'B^nE^n} \omega^m_{A'B^nE^n}\} \geq 1 - \varepsilon. \tag{3}$$

The protocol is δ-secure if the reduced state $\omega^m_{A'E^n}$ on systems $A'E^n$ is nearly indistinguishable from a constant state $\sigma_{A'E^n}$ independent of the message m:

$$\forall m : \frac{1}{2} \left\| \omega^m_{A'E^n} - \sigma_{A'E^n} \right\|_1 \leq \delta, \tag{4}$$

where we have employed the normalized trace distance.

We say that a rate R is achievable for the conditional quantum one-time pad if for all $\varepsilon, \delta \in (0,1)$, $\gamma > 0$, and sufficiently large n, there exists an $(n, 2^{n[R-\gamma]}, \varepsilon, \delta)$ conditional one-time pad protocol of the above form. The conditional one-time pad capacity of a state ρ_{ABE} is equal to the supremum of all achievable rates.

Achievability of CQMI for conditional one-time pad— Here we mostly sketch an argument that the CQMI $I(A;B|E)_\rho$ is a lower bound on the conditional one-time pad capacity of ρ_{ABE}, while the supplementary material contains a detailed proof. First, consider the reduced state ρ_A and a spectral decomposition for it as $\rho_A = \sum_x p_X(x)|x\rangle\langle x|_A$, where p_X is a probability distribution and $\{|x\rangle_A\}_x$ is an orthonormal basis. Let $|\phi\rangle_{AR} = \sum_x \sqrt{p_X(x)}|x\rangle_A|x\rangle_R$ be a purification of ρ_A. Let $|\psi\rangle_{ABEF}$ denote a purification of ρ_{ABE}, with F playing the role of a purifying system. Since all purifications are related by an isometry acting on the purifying system, there exists an isometry $U_{R\to BEF}$ such that $U_{R\to BEF}|\phi\rangle_{AR} = |\psi\rangle_{ABEF}$. Applying the isometry $U_{R\to BEF}$ followed by a partial trace over F can be thought of as a channel $\mathcal{N}_{R\to BE}$ that realizes the state ρ_{ABE} as $\mathcal{N}_{R\to BE}(\phi_{AR}) = \rho_{ABE}$. Similarly, if we apply the isometry $U_{R\to BEF}$ and trace over FB, then this is a channel $\mathcal{M}_{R\to E}$ that realizes the reduced state ρ_{AE} as $\mathcal{M}_{R\to E}(\phi_{AR}) = \rho_{AE}$.

If we take n copies of ρ_{ABE}, then the state $\rho^{\otimes n}_{ABE}$ can be thought of as the following state $\mathcal{N}^{\otimes n}_{R\to BE}(\phi^{\otimes n}_{AR})$. The pure state $|\phi\rangle^{\otimes n}_{AR}$ admits an information-theoretic type decomposition of the following form: $|\phi\rangle^{\otimes n}_{AR} = \sum_t \sqrt{p(t)}|\Phi_t\rangle_{A^nR^n}$, where the label t indicates a type class and $|\Phi_t\rangle_{A^nR^n}$ is a maximally entangled state of Schmidt rank d_t with support on the type class subspace labeled by t. We can then consider forming encoding unitaries out of the generalized Pauli shift and phase-shift operators

$$V_{A^n}(x_t, z_t) = X_{A^n}(x_t) Z_{A^n}(z_t), \tag{5}$$

which act on a given type class subspace t and where $x_t, z_t \in \{0, \ldots, d_t - 1\}$. The overall encoding unitary allows for an additional phase $(-1)^{b_t}$ for $b_t \in \{0,1\}$ and has the form

$$U_{A^n}(s) = \bigoplus_t (-1)^{b_t} V_{A^n}(x_t, z_t), \tag{6}$$

where s is a vector $[(b_t, x_t, z_t)]_t$.

The coding scheme is based on random coding, as is usually the case in quantum Shannon theory, and works as follows. Let $M, K \in \mathbb{N}$. Alice has a message variable $m \in \{1, \ldots, M\}$ and a local key variable $k \in \{1, \ldots, K\}$. For each pair (m,k), Alice picks a vector s, of the form described previously, uniformly at random and labels it as $s(m,k)$. The set $\mathcal{C} = \{s(m,k)\}_{m,k}$ constitutes the code, and observe that it is initially selected randomly. If Alice wishes to send message m, then she picks k uniformly at random from $k \in \{1, \ldots, K\}$, applies the encoding unitary $U_{A^n}(s(m,k))$ to the state $\rho^{\otimes n}_{ABE}$ and sends the A^n systems to Bob. Bob's goal is to decode both the message variable m and the local key variable k. Based on the packing lemma, it follows that if $\log_2 MK \approx nI(A;BE)_\rho$, then there is a decoding measurement $\{\Lambda^{m,k}_{A^nB^nE^n}\}$ for Bob, constructed from typical projectors and corresponding to a particular selected code $\mathcal{C}$, such that

$$\mathbb{E}_{\mathcal{C}} \left\{ \frac{1}{MK} \sum_{m,k} \mathrm{Tr}\{\Lambda^{m,k}_{A^nB^nE^n} U_{A^n}(S(m,k)) \rho^{\otimes n}_{ABE} \times U^\dagger_{A^n}(S(m,k))\} \right\} \geq 1 - \varepsilon, \tag{7}$$

for all $\varepsilon \in (0,1)$ and sufficiently large n, and where the expectation is with respect to the random choice of code $\mathcal{C}$. On the other hand, from the perspective of someone who does not know the choice of k and who does not have access to the systems B^n, the state has the following form:

$$\tau^m_{A^nE^n} \equiv \frac{1}{K} \sum_{k=1}^{K} U_{A^n}(s(m,k)) \rho^{\otimes n}_{AE} U^\dagger_{A^n}(s(m,k)). \tag{8}$$

The quantum covering lemma and the properties of typical projectors guarantee that

$$\Pr_{\mathcal{C}}\{\|\tau^m_{A^nE^n} - \overline{\tau}_{A^nE^n}\|_1 \leq \delta + 4\sqrt{\delta} + 24\sqrt[4]{\delta}\} \geq 1 - 2D \exp\left(-\frac{\delta^3 K 2^{-n[I(A;E)_\rho + \delta']}}{4} \right), \tag{9}$$

where D is a parameter that is no more than exponential in n, $\delta' > 0$ is a small constant, and

$$\overline{\tau}_{A^nE^n} \equiv \mathbb{E}_S\{U_{A^n}(S) \rho^{\otimes n}_{AE} U^\dagger_{A^n}(S)\}. \tag{10}$$

Thus, as long as we pick $\log_2 K \approx nI(A;E)_\rho$, then there is an extremely good chance that the state $\tau^m_{A^nE^n}$ will be nearly indistinguishable from the average state $\overline{\tau}_{A^nE^n}$. Now, we can define the event E_0 to be the event that Bob's measurement decodes with high average success probability and the event E_m to be the event that $\|\tau^m_{A^nE^n} - \overline{\tau}_{A^nE^n}\|_1$ is small. The union bound of probability theory then guarantees that there is a non-zero probability for there to be a code $\{s(m,k)\}_{m,k}$ such that the average success probability of Bob's decoder is arbitrarily high and $\|\tau^m_{A^nE^n} - \overline{\tau}_{A^nE^n}\|_1$ is arbitrarily small for all m, with these statements holding for sufficiently large n. So this means that such a code $\{s(m,k)\}_{m,k}$

exists. A final "expurgation" argument guarantees that Bob can decode each m and k with arbitrarily high probability and that $\|\tau^m_{A^nE^n} - \overline{\tau}_{A^nE^n}\|_1$ is arbitrarily small for all m. Therefore, the number of bits that Alice can communicate securely is thus

$$\begin{aligned}\log_2 M &= \log_2 MK - \log_2 K \\ &\approx nI(A;BE)_\rho - nI(A;E)_\rho \\ &= nI(A;B|E)_\rho, \end{aligned} \tag{11}$$

so that $I(A;B|E)_\rho$ is an achievable rate. This concludes the achievability proof sketch. As indicated previously, the optimality proof is given in the supplementary material.

We note that it actually suffices to use $\approx nH(A)_\rho$ noiseless qubit channels for the communication of the A systems, rather than $n \log|A|$ noiseless qubit channels. This is because Alice can perform Schumacher compression [32] of her A^n systems before transmitting them, and the structure of the encoding unitaries is such that this can be done regardless of which message is being transmitted (see the discussion at the end of [3, Section 22.3]). The Schumacher compression causes a negligible disturbance to each of the states that is transmitted.

Conditional one-time pad of a quantum message—We note that it is possible to define a conditional quantum one-time pad of a quantum message, in which the goal is to transmit one share $\hat{M}$ of a quantum state $|\varphi\rangle_{M''\hat{M}}$ securely in such a way that Bob, possessing systems $A'B^nE^n$, can decode the quantum message in $\hat{M}$, while someone possessing the systems $A'E^n$ cannot learn anything about the quantum system $\hat{M}$. Our result here is that $I(A;B|E)_\rho/2$ is the optimal rate for this task of conditional one-time pad of a quantum message. The optimality proof is nearly identical to the optimality proof given previously, except that we start with the assumption that the initial state $|\varphi\rangle_{M''\hat{M}}$ is a maximally entangled state $|\Phi\rangle_{M''\hat{M}}$, such that the quantum information in system $\hat{M}$ can be decoded well. Then, the proof starts with the condition that $\log_2 M = I(M'';\hat{M})_\Phi/2$ and proceeds identically from there. For the achievability part, we perform a coherent version of the above protocol, as reviewed in [3, Section 22.4], and we find that it generates coherent bits [33], which are secure from someone possessing the A^nE^n systems, at a rate equal to $I(A;B|E)_\rho$. By the coherent communication identity from [33], it follows that qubits can be transmitted securely at a rate equal to $I(A;B|E)_\rho/2$.

Generalizations—We note that the coding scheme outlined above in the achievability proof can be generalized in several interesting ways. Suppose that Alice shares a state with "many Bobs", i.e., one of the form $\rho_{AB_1\cdots B_\ell}$ for some positive integer $\ell \geq 2$. Then Alice might wish to encode a message m in her A systems of $\rho^{\otimes n}_{AB_1\cdots B_\ell}$ in such a way that only someone possessing all of the systems $AB_1 \cdots B_\ell$ would be able to decode it, but someone possessing system A and some subset $\mathcal{B}_i \in \{B_1,\ldots,B_\ell\}$ would not be able to determine anything about the message m. Alice might wish to protect the message against several different subsets $\mathcal{B}_i$, for $i \in \{1,\ldots,p\}$, as in secret sharing. Then we could structure a coding scheme similar to our achievability proof to have a message variable $m \in \{1,\ldots,M\}$ and a local key variable $k \in \{1,\ldots,K\}$, such that

$$\log_2 MK \approx nI(A;B_1\cdots B_\ell), \tag{12}$$

$$\log_2 K \approx n[\max\{I(A;\mathcal{B}_1),\ldots,I(A;\mathcal{B}_p)\}]. \tag{13}$$

Given that

$$I(A;B_1\cdots B_\ell)_\rho - \max\{I(A;\mathcal{B}_1)_\rho,\ldots,I(A;\mathcal{B}_p)_\rho\} \tag{14}$$

is always non-negative, the coding scheme guarantees that this information difference is an achievable rate that accomplishes the desired task. We note that the secret-sharing task discussed above is different from the previously considered protocols in [34, 35], and references therein.

A particular case of interest is the scenario mentioned earlier in the paper and which we called information scrambling. There, Alice, Bob, and Eve share a state ρ_{ABE}, and the goal is for Alice to encode a message in the A system such that someone possessing the ABE systems can decode it, but someone possessing the AB systems or the AE systems cannot determine anything about the message m (i.e., the message m has been scrambled in the nonlocal degrees of freedom of the state ρ_{ABE} and is not available in ρ_{AB} or ρ_{AE}). According to the above reasoning, an achievable rate for this task is the information quantity $I(A;BE)_\rho - \max\{I(A;B)_\rho, I(A;E)_\rho\}$. This rate is also optimal.

We note also that our methods give a concrete and transparent approach to prove the results of [31], as discussed in the supplementary material. In particular, we have established an information-theoretic converse of that result using entropy identities and inequalities along the lines presented previously, and the achievability part of that result can be accomplished by using the encoding unitaries discussed earlier, along with the quantum packing and covering lemmas.

Our operational interpretation of the conditional mutual information also leads to an interesting operational interpretation of the squashed entanglement of a bipartite state ρ_{AB}: we can consider squashed entanglement to be the optimal rate of secure communication in the conditional one-time pad if an eavesdropper has the E system of the worst possible extension ρ_{ABE} of the state ρ_{AB}, given that squashed entanglement is defined as $\frac{1}{2}\inf_{\rho_{ABE}}\{I(A;B|E)_\rho : \mathrm{Tr}_E\{\rho_{ABE}\} = \rho_{AB}\}$ [11]. This is analogous to the interpretations from [36] and the follow-up one in [7].

Conclusion—In this paper, we proved that the conditional mutual information $I(A;B|E)_\rho$ of a tripartite state ρ_{ABE} is equal to the optimal rate of secure communication for a task that we call the conditional one-time pad. This represents a fundamentally different operational interpretation of conditional mutual information that is conceptually simple at the same time. Furthermore, due to the fact that the optimal rate is given by conditional mutual information, the conditional one-time pad is an example of a communication task in which non-Markov quantum states are used as a resource [37, 38]. In the continuing quest to understand a refined generalization of conditional mutual information, as has been attempted previously in [39–42], the protocol of conditional one-time pad might end up being helpful in this effort.

ACKNOWLEDGMENTS

We thank David Ding, Rahul Jain, and Andreas Winter for discussions related to the topic of this paper. MMW acknowledges support from the Office of Naval Research and the National Science Foundation. KS acknowledges support from the Department of Physics and Astronomy at LSU and the National Science Foundation under Grant No. 1714215.

[1] Masahito Hayashi. *Quantum Information: An Introduction*. Springer, 2006.

[2] Alexander S. Holevo. *Quantum Systems, Channels, Information*. de Gruyter Studies in Mathematical Physics (Book 16). de Gruyter, November 2012.

[3] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, second edition, February 2017. arXiv:1106.1445v8.

[4] Igor Devetak and Jon Yard. Exact cost of redistributing multipartite quantum states. *Physical Review Letters*, 100(23):230501, June 2008.

[5] Jon Yard and Igor Devetak. Optimal quantum source coding with quantum side information at the encoder and decoder. *IEEE Transactions on Information Theory*, 55(11):5339–5351, November 2009. arXiv:0706.2907.

[6] Mario Berta, Fernando G. S. L. Brandao, Christian Majenz, and Mark M. Wilde. Conditional decoupling of quantum information. *Physical Review Letters*, 121(4):040504, July 2018. arXiv:1808.00135.

[7] Mario Berta, Fernando G. S. L. Brandao, Christian Majenz, and Mark M. Wilde. Deconstruction and conditional erasure of quantum correlations. *Physical Review A*, 98(4):042320, October 2018. arXiv:1609.06994.

[8] Haoyu Qi, Kunal Sharma, and Mark M. Wilde. Entanglement-assisted private communication over quantum broadcast channels. *Journal of Physics A*, 51(37):374001, September 2018. arXiv:1803.03976.

[9] Elliott H. Lieb and Mary Beth Ruskai. Proof of the strong subadditivity of quantum-mechanical entropy. *Journal of Mathematical Physics*, 14(12):1938–1941, December 1973.

[10] Elliott H. Lieb and Mary Beth Ruskai. A fundamental property of quantum-mechanical entropy. *Physical Review Letters*, 30(10):434–436, March 1973.

[11] Matthias Christandl and Andreas Winter. "Squashed entanglement" - an additive entanglement measure. *Journal of Mathematical Physics*, 45(3):829–840, March 2004. arXiv:quant-ph/0308088.

[12] Wojciech H. Zurek. Einselection and decoherence from an information theory perspective. *Annalen der Physik*, 9(11-12):855–864, November 2000. arXiv:quant-ph/0011039.

[13] Harold Ollivier and Wojciech H. Zurek. Quantum discord: A measure of the quantumness of correlations. *Physical Review Letters*, 88(1):017901, December 2001. arXiv:quant-ph/0105072.

[14] Marco Piani. Problem with geometric discord. *Physical Review A*, 86(3):034101, September 2012. arXiv:1206.0231.

[15] Eneet Kaur, Xiaoting Wang, and Mark M. Wilde. Conditional mutual information and quantum steering. *Physical Review A*, 96(2):022332, August 2017. arXiv:1612.03875.

[16] Omar Fawzi and Renato Renner. Quantum conditional mutual information and approximate Markov chains. *Communications in Mathematical Physics*, 340(2):575–611, December 2015. arXiv:1410.0664.

[17] Alexei Kitaev and John Preskill. Topological entanglement entropy. *Physical Review Letters*, 96:110404, March 2006. arXiv:hep-th/0510092.

[18] Michael Levin and Xiao-Gang Wen. Detecting topological order in a ground state wave function. *Physical Review Letters*, 96:110405, March 2006. arXiv:cond-mat/0510613.

[19] Isaac H. Kim. Perturbative analysis of topological entanglement entropy from conditional independence. *Physical Review B*, 86:245116, December 2012. arXiv:1210.2360.

[20] Bei Zeng, Xie Chen, Duan-Lu Zhou, and Xiao-Gang Wen. Quantum information meets quantum matter. August 2015. arXiv:1508.02595.

[21] Raghu Mahajan, C. Daniel Freeman, Sam Mumford, Norm Tubman, and Brian Swingle. Entanglement structure of non-equilibrium steady states. August 2016. arXiv:1608.05074.

[22] Bartlomiej Czech, Lampros Lamprou, Samuel McCandlish, and James Sully. Integral geometry and holography. *Journal of High Energy Physics*, 2015:175, October 2015. arXiv:1505.05515.

[23] Dawei Ding, Patrick Hayden, and Michael Walter. Conditional mutual information of bipartite unitaries and scrambling. *Journal of High Energy Physics*, 2016(12):145, December 2016. arXiv:1608.04750.

[24] Fernando Pastawski, Jens Eisert, and Henrik Wilming. Towards holography via quantum source-channel codes. *Physical Review Letters*, 119:020501, July 2017. arXiv:1611.07528.

[25] Benjamin Schumacher and Michael D. Westmoreland. Quantum mutual information and the one-time pad. *Physical Review A*, 74(4):042305, October 2006. arXiv:quant-ph/0604207.

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

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

[28] Ning Cai, Andreas Winter, and Raymond W. Yeung. Quantum privacy and quantum wiretap channels. *Problems of Information Transmission*, 40(4):318–336, October 2004.
[29] Min-Hsiu Hsieh, Igor Devetak, and Andreas Winter. Entanglement-assisted capacity of quantum multiple-access channels. *IEEE Transactions on Information Theory*, 54(7):3078–3090, July 2008. arXiv:quant-ph/0511228.
[30] Nilanjana Datta, Marco Tomamichel, and Mark M. Wilde. On the second-order asymptotics for entanglement-assisted communication. *Quantum Information Processing*, 15(6):2569–2591, June 2016. arXiv:1405.1797.
[31] Fernando G. S. L. Brandao and Jonathan Oppenheim. Quantum one-time pad in the presence of an eavesdropper. *Physical Review Letters*, 108(4):040504, January 2012. arXiv:1004.3328.
[32] Benjamin Schumacher. Quantum coding. *Physical Review A*, 51(4):2738–2747, April 1995.
[33] Aram Harrow. Coherent communication of classical messages. *Physical Review Letters*, 92(9):097902, March 2004. arXiv:quant-ph/0307091.
[34] Mark Hillery, Vladimir Buzek, and Andre Berthiaume. Quantum secret sharing. *Physical Review A*, 59(3):1829, March 1999. arXiv:quant-ph/9806063.
[35] Kaushik Senthoor and Pradeep Kiran Sarvepalli. Communication efficient quantum secret sharing. January 2018. arXiv:1801.09500.
[36] Jonathan Oppenheim. A paradigm for entanglement theory based on quantum communication. January 2008. arXiv:0801.0458.
[37] Patrick Hayden, Richard Jozsa, Denes Petz, and Andreas Winter. Structure of states which satisfy strong subadditivity of quantum entropy with equality. *Communications in Mathematical Physics*, 246(2):359–374, April 2004. arXiv:quant-ph/0304007.
[38] Eyuri Wakakuwa. Operational resource theory of non-Markovianity. September 2017. arXiv:1709.07248.
[39] Mario Berta, Kaushik Seshadreesan, and Mark M. Wilde. Rényi generalizations of the conditional quantum mutual information. *Journal of Mathematical Physics*, 56(2):022205, February 2015. arXiv:1403.6102.
[40] Nilanjana Datta, Min-Hsiu Hsieh, and Jonathan Oppenheim. An upper bound on the second order asymptotic expansion for the quantum communication cost of state redistribution. *Journal of Mathematical Physics*, 57(5):052203, May 2016. arXiv:1409.4352.
[41] Mario Berta, Matthias Christandl, and Dave Touchette. Smooth entropy bounds on one-shot quantum state redistribution. *IEEE Transactions on Information Theory*, 62(3):1425–1439, March 2016. arXiv:1409.4338.
[42] Anurag Anshu, Vamsi Krishna Devabathini, and Rahul Jain. Quantum communication using coherent rejection sampling. *Physical Review Letters*, 119(12):120506, September 2017. arXiv:1410.3031.
[43] Armin Uhlmann. The "transition probability" in the state space of a *-algebra. *Reports on Mathematical Physics*, 9(2):273–279, 1976.
[44] Andreas Winter. Tight uniform continuity bounds for quantum entropies: conditional entropy, relative entropy distance and energy constraints. *Communications in Mathematical Physics*, 347(1):291–313, October 2016. arXiv:1507.07775.
[45] Alexander S. Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problems of Information Transmission*, 9:177–183, 1973.

Supplementary material

Appendix A: Preliminaries

We begin by reviewing some definitions and prior results relevant for the rest of the supplementary material. We point readers to [1–3] for background.

1. Quantum states, quantum fidelity, and trace distance

Throughout this work, we restrict ourselves to finite-dimensional Hilbert spaces. Let $\mathcal{H}$ denote a Hilbert space. Let $\mathcal{D}(\mathcal{H})$ denote the set of density operators (positive semi-definite with unit trace) acting on $\mathcal{H}$. The Hilbert space for a composite system AB is denoted as $\mathcal{H}_{AB}$ where $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$. The density operator corresponding to a state of a composite system AB is denoted as ρ_{AB}, and the reduced state $\rho_A = \operatorname{Tr}_B\{\rho_{AB}\}$, where $\rho_A \in \mathcal{D}(\mathcal{H}_A)$. A purification of a density operator ρ_A is a pure state ψ_{RA} such that $\operatorname{Tr}_R\{\psi_{RA}\} = \rho_A$, where R is called the purifying system. All purifications of a density operator are related by an isometry acting on the purifying system. The maximally mixed state acting on the Hilbert space $\mathcal{H}_A$ is denoted by $\pi_A \equiv I_A/\dim(\mathcal{H}_A)$.

The fidelity of two quantum states $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ is defined as [43] $F(\rho,\sigma) \equiv \left\|\sqrt{\rho}\sqrt{\sigma}\right\|_1^2$, where $\|\cdot\|_1$ denotes the trace norm. The trace distance between two density operators $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ is equal to $\|\rho - \sigma\|_1$. The operational interpretation of trace distance is that it is linearly related to the maximum success probability in distinguishing two quantum states.

2. Typical set, strongly typicality, typical subspace, typical projector, properties of typical subspace

Suppose that a random variable X takes values in an alphabet $\mathcal{X}$ with cardinality $|\mathcal{X}|$. Consider an i.i.d. information source that samples independently from the distribution $p_X(x)$, corresponding to random variable X, and emits n realizations $x_1, \ldots, x_n$.

Let $N(x|x^n)$ be the number of occurrences of the symbol $x \in \mathcal{X}$ in the sequence x^n.

Definition 1 (Type) *The type or empirical distribution t_{x^n} of a sequence x^n is a probability mass function whose elements are $t_{x^n}(x)$ where*

$$t_{x^n}(x) \equiv \frac{1}{n} N(x|x^n). \tag{A1}$$

Definition 2 (Type Class) *Let $T_t^{X^n}$ denote the type class of a particular type t. The type class $T_t^{X^n}$ is the set of all sequences with length n and type t:*

$$T_t^{X^n} \equiv \{x^n \in \mathcal{X}^n : t^{x^n} = t\}. \tag{A2}$$

Definition 3 (Strongly Typical Set) *The δ-strongly typical set $T_\delta^{X^n}$ is the set of all sequences with an empirical distribution $\frac{1}{n}N(x|x^n)$ that has maximum deviation δ from the true distribution $p_X(x)$. Furthermore, the empirical distribution $\frac{1}{n}N(x|x^n)$ of any sequence in $T_\delta^{X^n}$ vanishes for any letter x for which $p_X(x) = 0$:*

$$T_\delta^{X^n} \equiv \left\{ x^n : \forall x \in \mathcal{X},\ \left| \frac{1}{n} N(x|x^n) - p_X(x) \right| \le \delta \ \textit{if}\ p_X(x) > 0,\ \textit{else}\ \frac{1}{n} N(x|x^n) = 0 \right\}. \tag{A3}$$

We now discuss the notion of a quantum information source and recall definitions of a typical subspace and typical projectors. Analogous to the notion of a classical information source, a quantum information source randomly emits pure qudit states in a finite-dimensional Hilbert space $\mathcal{H}_A$. Consider the following spectral decomposition of a density operator ρ:

$$\rho_A = \sum_{x \in \mathcal{X}} p_X(x) |x\rangle\langle x|_A \ . \tag{A4}$$

Now suppose that the quantum information source emits a large number n of random quantum states. The density operator corresponding to the emitted state is given by

$$\rho_{A^n} \equiv \rho_{A_1} \otimes \cdots \otimes \rho_{A^n} = \rho_A^{\otimes n} \ . \tag{A5}$$

A spectral decomposition of the aforementioned state is as follows:

$$\rho_{A^n} = \sum_{x^n \in \mathcal{X}^n} p_{X^n}(x^n) |x^n\rangle\langle x^n|_{A^n}, \tag{A6}$$

where $p_{X^n}(x^n) = \prod_{i=1}^n p_X(x_i)$, and $|x^n\rangle_{A^n} \equiv |x_1\rangle_{A_1} \cdots |x_n\rangle_{A_n}$.

Definition 4 (Typical Subspace) *The δ-typical subspace $T_{A^n}^{\rho,\delta}$ associated with many copies of a density operator, as defined in (A4), is a subspace of the Hilbert space $\mathcal{H}_{A^n} = \mathcal{H}_{A_1} \otimes \cdots \otimes \mathcal{H}_{A_n}$. It is spanned by states $|x^n\rangle_{A^n}$ whose corresponding classical sequences x^n are δ-typical, i.e.,*

$$T_{A^n}^{\rho,\delta} \equiv \mathrm{span}\{|x^n\rangle_{A^n} : x^n \in T_\delta^{X^n}\} \ . \tag{A7}$$

Definition 5 (Typical Projector) *The typical projector $\Pi_{A^n}^{\rho,\delta}$ is a projector onto the typical subspace associated with a density operator, as defined in (A4):*

$$\Pi_{A^n}^{\rho,\delta} \equiv \sum_{x^n \in T_\delta^{X^n}} |x^n\rangle\langle x^n|_{A^n}. \tag{A8}$$

Definition 6 (Type Class Subspace) *The type class subspace associated to type t is the subspace spanned by all states with the same type:*

$$T_{A^n}^t \equiv \mathrm{span}\left\{ |x^n\rangle_{A^n} : x^n \in T_t^{X^n} \right\}. \tag{A9}$$

Definition 7 (Type Class Projector) *Let $\Pi^t_{A^n}$ denote the type class subspace projector associated to type t:*

$$\Pi^t_{A^n} \equiv \sum_{x^n \in T^{X^n}_t} |x^n\rangle\langle x^n|_{A^n}. \tag{A10}$$

Using the aforementioned definitions, we now state three useful properties of the strongly typical subspace $T^{\rho;\delta}_{A^n}$. We point readers to [3] for a review of the proofs of these properties.

Property A.1 (Unit Probability) *The probability that the quantum state ρ_{A^n} is in the strongly typical subspace $T^{\rho;\delta}_{A^n}$ approaches one as n becomes large, i.e.,*

$$\mathrm{Tr}\{\Pi^{\rho;\delta}_{A^n}\rho_{A^n}\} \geq 1-\varepsilon, \tag{A11}$$

for all $\varepsilon \in (0,1)$, $\delta > 0$, and sufficiently large n.

Property A.2 (Exponentially Smaller Dimension) *The dimension of the strongly typical subspace ($T^{\rho;\delta}_{A^n}$) is exponentially smaller than the dimension $|A|^n$ of the entire space of quantum states when the output of the quantum information source is not maximally mixed. The mathematical form of this property is as follows:*

$$\mathrm{Tr}\{\Pi^{\rho;\delta}_{A^n}\} \leq 2^{n(H(A)+c\delta)}, \tag{A12}$$

where c is some positive constant.

Property A.3 (Equipartition) *The action of the strongly typical projector $\Pi^{\rho;\delta}_{A^n}$ on the density operator ρ_{A^n} is to select out all the basis states of ρ_{A^n} that are in the typical subspace and form a sliced operator. The following operator inequality holds for the sliced operator $\Pi^{\rho;\delta}_{A^n}\rho_{A^n}\Pi^{\rho;\delta}_{A^n}$:*

$$2^{-n(H(A)+c\delta)}\Pi^{\rho;\delta}_{A^n} \leq \Pi^{\rho;\delta}_{A^n}\rho_{A^n}\Pi^{\rho;\delta}_{A^n} \leq 2^{-n(H(A)-c\delta)}\Pi^{\rho;\delta}_{A^n}\ . \tag{A13}$$

3. Packing lemma

Definition 8 (Ensemble) *Suppose that a random variable X with probability density function $p_X(x)$ takes values in an alphabet $\mathcal{X}$ with cardinality $|\mathcal{X}|$. Consider an ensemble $\{p_X(x), \sigma_x\}_{x\in\mathcal{X}}$ of quantum states where each realization x can be encoded into a quantum state $\sigma_x \in \mathcal{D}(\mathcal{H})$. The expected density operator of the ensemble is*

$$\sigma \equiv \sum_{x\in\mathcal{X}} p_X(x)\sigma_x\ . \tag{A14}$$

For a reliable transmission of classical information, Alice can select a subset $\mathcal{C}$ of $\mathcal{X}$ for encoding, and Bob's task is to distinguish this subset of states. The subset $\mathcal{C}$ constitutes the code. We now recall the statement of the packing lemma.

Lemma 9 (Packing Lemma) *Suppose that Alice has an ensemble $\{p_X(x), \sigma_x\}_{x\in\mathcal{X}}$, as in Definition 8. Suppose that codeword subspace projectors $\{\Pi_x\}_{x\in\mathcal{X}}$ and a code subspace projector Π exist, and they project onto subspaces of the Hilbert space $\mathcal{H}$, and these projectors and ensemble satisfy the following conditions:*

$$\mathrm{Tr}\{\Pi\sigma_x\} \geq 1-\varepsilon, \tag{A15}$$
$$\mathrm{Tr}\{\Pi_x\sigma_x\} \geq 1-\varepsilon, \tag{A16}$$
$$\mathrm{Tr}\{\Pi_x\} \leq d, \tag{A17}$$
$$\Pi\sigma\Pi \leq \frac{1}{D}\Pi, \tag{A18}$$

where $\varepsilon \in (0,1)$, $D > 0$, and $d \in (0,D)$. Suppose that $\mathcal{M}$ is a message set of size $|\mathcal{M}|$ with elements m. Consider a set $\mathcal{C} = \{C_m\}_{m\in\mathcal{M}}$ of random variables C_m generated independently at random according to $p_X(x)$, so that each random variable C_m takes a value in $\mathcal{X}$ and corresponds to the message m. However, its distribution is independent of the particular message m, and therefore, the set $\mathcal{C}$ constitutes a random code. Then there exists a corresponding

POVM $\{\Lambda_m\}_{m\in\mathcal{M}}$ *that reliably distinguishes the states* $\{\sigma_{C_m}\}_{m\in\mathcal{M}}$*, in the sense that the expectation of the average probability of detecting the correct state is high:*

$$\mathbb{E}_{\mathcal{C}}\left\{\frac{1}{|\mathcal{M}|}\sum_{m\in\mathcal{M}}\mathrm{Tr}\left\{\Lambda_m\sigma_{C_m}\right\}\right\}\geq 1-2\left(\varepsilon+2\sqrt{\varepsilon}\right)-4\left|\mathcal{M}\right|\frac{d}{D}, \tag{A19}$$

when D/d *is large,* $|\mathcal{M}|\ll D/d$*, and* ε *is small.*

We note that Bob can construct POVM $\{\Lambda_m\}_{m\in\mathcal{M}}$ by using the codeword subspace projectors $\{\Pi_x\}_{x\in\mathcal{X}}$ and the code subspace projector Π. In particular, Bob can employ a square-root measurement or a sequential decoding strategy. We point readers to [3, Chapter 16] for a review of an explicit construction of POVM and a complete proof of the packing lemma.

4. Covering lemma

The goal of the covering lemma is to cover Eve's space in such a way that Eve cannot distinguish different classical messages that Alice is sending to Bob. We start by defining two relevant ensembles for the covering lemma. We follow the convention from [3] and refer to the two different ensembles as the "true ensemble" and the "fake ensemble."

Definition 10 (True Ensemble) *For our discussion, the true ensemble is defined in the same way as in Definition 8.*

Definition 11 (Fake Ensemble) *Let* $\mathcal{G}$ *be a set such that* $\mathcal{G}\subseteq\mathcal{X}$*. The fake ensemble is defined as follows:*

$$\left\{1/\left|\mathcal{G}\right|,\sigma_g\right\}_{g\in\mathcal{G}}. \tag{A20}$$

Let $\overline{\sigma}$ *denote the "fake expected density operator" of the fake ensemble:*

$$\overline{\sigma}(\mathcal{G})\equiv\frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}}\sigma_g. \tag{A21}$$

The goal for Alice is to generate a fake ensemble such that the trace distance between $\bar{\sigma}(\mathcal{G})$ in (A21) and σ in (A14) is small. Moreover, in order to achieve a higher private communication rate, it is required to make the size of fake ensembles as small as possible while still having privacy from Eve. We call the trace distance between $\bar{\sigma}(\mathcal{G})$ and σ the obfuscation error $o_e(\mathcal{G})$, i.e.,

$$o_e(\mathcal{G})=\left\|\bar{\sigma}(\mathcal{G})-\sigma\right\|_1 . \tag{A22}$$

We now recall the statement of the covering lemma.

Lemma 12 (Covering Lemma) *Let* $\{p_X(x),\sigma_x\}_{x\in\mathcal{X}}$ *be an ensemble as in Definition 8. Suppose a total subspace projector* Π *and codeword subspace projectors* $\{\Pi_x\}_{x\in\mathcal{X}}$ *are given, they project onto subspaces of* $\mathcal{H}$*, and these projectors and each state* σ_x *satisfy the following conditions:*

$$\mathrm{Tr}\{\Pi\sigma_x\}\geq 1-\varepsilon, \tag{A23}$$

$$\mathrm{Tr}\{\Pi_x\sigma_x\}\geq 1-\varepsilon, \tag{A24}$$

$$\mathrm{Tr}\{\Pi\}\leq D, \tag{A25}$$

$$\Pi_x\sigma_x\Pi_x\leq\frac{1}{d}\Pi_x, \tag{A26}$$

where $\varepsilon\in(0,1)$*,* $D>0$*, and* $d\in(0,D)$*. Suppose that* $\mathcal{G}$ *is a set of size* $|\mathcal{G}|$ *with elements* g*. Let a random covering code* $\mathcal{C}\equiv\{C_g\}_{g\in\mathcal{G}}$ *consist of random codewords* C_g *where the codewords* C_g *are chosen independently according to the distribution* $p_X(x)$ *and give rise to a fake ensemble* $\left\{1/\left|\mathcal{G}\right|,\sigma_{C_g}\right\}_{g\in\mathcal{G}}$*. Then the following bound exists on the probability of having a small obfuscation error* $o_e(\mathcal{C})$ *of the random covering code:*

$$\Pr_{\mathcal{C}}\left\{o_e(\mathcal{C})\leq\varepsilon+4\sqrt{\varepsilon}+24\sqrt[4]{\varepsilon}\right\}\geq 1-2D\exp\left(-\frac{\varepsilon^3}{4}\frac{|\mathcal{G}|\,d}{D}\right), \tag{A27}$$

when ε *is small and* $|\mathcal{G}|\gg\varepsilon^3 d/D$*. Thus, it is highly likely that a given fake ensemble* $\left\{1/\left|\mathcal{G}\right|,\sigma_{C_g}\right\}_{g\in\mathcal{G}}$ *has its expected density operator indistinguishable from the expected density operator of the original ensemble* $\{p_X(x),\sigma_x\}_{x\in\mathcal{X}}$*.*

We point readers to [3, Chapter 17] for a proof of the covering lemma.

5. Properties of encoding unitaries

Consider the reduced state ρ_A and a spectral decomposition for it as $\rho_A = \sum_x p_X(x)|x\rangle\langle x|_A$, where p_X is a probability distribution and $\{|x\rangle_A\}_x$ is an orthonormal basis. Consider the following purification of ρ_A:

$$|\psi\rangle_{AR} = \sum_x \sqrt{p_X(x)}|x\rangle_A|x\rangle_R, \tag{A28}$$

where $p_X(x) > 0$ for all x, $\sum_x p_X(x) = 1$, and $\{|x\rangle_A\}$ and $\{|x\rangle_R\}$ are orthonormal bases for systems A and R, respectively. We now start with n copies of the above state and write in terms of its type decomposition, as given in Definition 2:

$$|\psi^n\rangle_{A^nR^n} = \sum_{x^n} \sqrt{p_{X^n}(x^n)}|x^n\rangle_{A^n}|x^n\rangle_{R^n} \tag{A29}$$

$$= \sum_t \sqrt{p_{X^n}(x_t^n)d_t}\frac{1}{\sqrt{d_t}} \sum_{x^n\in T_t} |x^n\rangle_{A^n}|x^n\rangle_{R^n} \tag{A30}$$

$$= \sum_t \sqrt{p(t)}|\Phi_t\rangle_{A^nR^n}, \tag{A31}$$

where d_t is the size of the type class T_t and

$$p(t) \equiv p_{X^n}(x_t^n)d_t, \tag{A32}$$

$$|\Phi_t\rangle_{A^nR^n} \equiv \frac{1}{\sqrt{d_t}} \sum_{x^n\in T_t} |x^n\rangle_{A^n}|x^n\rangle_{R^n}. \tag{A33}$$

We now consider a Heisenberg-Weyl set of d_t^2 operators that act on all the A^n systems of $|\Phi_t\rangle_{A^nR^n}$. We denote one of these operators by $V(x_t, z_t) \equiv X(x_t)Z(z_t)$ where $x_t, z_t \in \{0, \ldots, d_t - 1\}$. Along with these operators, Alice applies a phase $(-1)^{b_t}$ in each subspace. Therefore, the resulting unitary operator can be expressed as a direct sum of all these unitary operators:

$$U_{A^n}(s) = \bigoplus_t (-1)^{b_t} V_{A^n}(x_t, z_t), \tag{A34}$$

where s is a vector containing all the indices needed to specify the unitary $U(s)$:

$$s \equiv [(x_t, z_t, b_t)]_t. \tag{A35}$$

We now recall that a transpose trick holds for such unitary operators. Consider the following chain of inequalities:

$$(U_{A^n}(s) \otimes I_{R^n})|\psi^n\rangle_{A^nR^n} = \sum_t \sqrt{p(t)}(-1)^{b_t} V_{A^n}(x_t, z_t)|\Phi_t\rangle_{A^nR^n} \tag{A36}$$

$$= \sum_t \sqrt{p(t)}(-1)^{b_t} V_{R^n}^T(x_t, z_t)|\Phi_t\rangle_{A^nR^n} \tag{A37}$$

$$= \left(I_{A^n} \otimes U_{R^n}^T(s)\right)|\psi^n\rangle_{A^nR^n}, \tag{A38}$$

where we have used the direct-sum property of the unitary operator (A34) and a transpose trick associated with the maximally entangled state $|\Phi_t\rangle_{A^nR^n}$.

Let $|\psi\rangle_{ABEF}$ denote a purification of ρ_{ABE}. Since all purifications are related by an isometry acting on the purifying system, there exists an isometry $U_{R\to BEF}$ such that $U_{R\to BEF}|\phi\rangle_{AR} = |\psi\rangle_{ABEF}$. Applying the isometry $U_{R\to BEF}$ followed by a partial trace over F can be thought of as a channel $\mathcal{N}_{R\to BE}$ that realizes the state ρ_{ABE} as $\mathcal{N}_{R\to BE}(\phi_{AR}) = \rho_{ABE}$. Moreover, the state $\rho_{ABE}^{\otimes n}$ can be thought of as the following state $\mathcal{N}_{R\to BE}^{\otimes n}(\phi_{AR}^{\otimes n})$.

Appendix B: Achievability of CQMI for conditional one-time pad

In this section, we provide a proof that the conditional quantum mutual information $I(A;B|E)_\rho$ is a lower bound on the conditional one-time pad capacity of ρ_{ABE}.

In the conditional one-time pad task, the goal for Alice is to encode information in her share of the state ρ_{ABE} in such a way that Bob can reliably decode the information, while maintaining privacy from Eve. We now construct a coding scheme based on random coding for reliable communication between Alice and Bob. Let $M, K \in \mathbb{N}$. Alice has message variable $m \in \{1, \ldots, M\}$ and a local key variable $k \in \{1, \ldots, K\}$. If Alice wishes to send message m, then she picks k uniformly at random from $k \in \{1, \ldots, K\}$. For each pair (m, k), the random code is selected in such a way that the vector s, of the form described in (A35), is chosen uniformly at random and then the encoding unitary $U_{A^n}(s(m,k))$ acting on the state $\rho_{ABE}^{\otimes n}$ is associated to m and k. Therefore, the set $\mathcal{C} = \{s(m,k)\}_{m,k}$ constitutes a random code, given that the s vectors are picked uniformly at random; i.e., the way that they are selected is independent of m and k, but after they are chosen, the association to m and k is made. Let $\mathcal{S}$ denote the set of all possible vectors s. To be clear, the ensemble from which Alice and Bob are selecting their code can be expressed as

$$\left\{\frac{1}{|\mathcal{S}|}, U_{A^n}(s)\rho_{A^nB^nE^n}U_{A^n}^\dagger(s)\right\}_{s\in\mathcal{S}}. \tag{B1}$$

As described in Lemma 9, if the four inequalities corresponding to the codeword subspace projectors, the code subspace projector, and aforementioned ensemble are satisfied, then there exists a decoding POVM that can reliably decode Alice's transmitted message. Consider the following respective codeword subspace projectors and a code subspace projector:

$$U_{A^n}(s)\Pi^{\rho,\delta}_{A^nB^nE^n}U_{A^n}^\dagger(s), \tag{B2}$$

$$\Pi^{\rho,\delta}_{A^n} \otimes \Pi^{\rho,\delta}_{B^nE^n}, \tag{B3}$$

where $\Pi^{\rho,\delta}_{A^nB^nE^n}$, $\Pi^{\rho,\delta}_{A^n}$, and $\Pi^{\rho,\delta}_{B^nE^n}$ are the typical projectors for n copies of the states ρ_{ABE}, ρ_A, and ρ_{BE}, respectively.

Let $\bar{\rho}_{A^nB^nE^n}$ denote the expected density operator of the ensemble in (B1). We now state the four conditions corresponding to the packing lemma for our code:

$$\mathrm{Tr}\left\{\left(\Pi^{\rho,\delta}_{A^n} \otimes \Pi^{\rho,\delta}_{B^nE^n}\right)\left(U_{A^n}(s)\rho_{A^nB^nE^n}U_{A^n}^\dagger(s)\right)\right\} \geq 1-\varepsilon, \tag{B4}$$

$$\mathrm{Tr}\left\{\left(U_{A^n}(s)\Pi^{\rho,\delta}_{A^nB^nE^n}U_{A^n}^\dagger(s)\right)\left(U_{A^n}(s)\rho_{A^nB^nE^n}U_{A^n}^\dagger(s)\right)\right\} \geq 1-\varepsilon, \tag{B5}$$

$$\mathrm{Tr}\left\{U_{A^n}(s)\Pi^{\rho,\delta}_{A^nB^nE^n}U_{A^n}^\dagger(s)\right\} \leq 2^{n[H(ABE)_\rho + c\delta]} \tag{B6}$$

$$\left(\Pi^{\rho,\delta}_{A^n} \otimes \Pi^{\rho,\delta}_{B^nE^n}\right)\overline{\rho}_{A^nB^nE^n}\left(\Pi^{\rho,\delta}_{A^n} \otimes \Pi^{\rho,\delta}_{B^nE^n}\right) \leq 2^{-n[H(A)_\rho + H(BE)_\rho - \nu(n,\delta) - c\delta]}\left(\Pi^{\rho,\delta}_{A^n} \otimes \Pi^{\rho,\delta}_{B^nE^n}\right), \tag{B7}$$

where c is some constant, and $\nu(n,\delta)$ is given by

$$\nu(n,\delta) = \frac{\delta}{2}\dim(\mathcal{H}_B)\log_2\dim(\mathcal{H}_B) + h_2\!\left(\dim(\mathcal{H}_B)\frac{\delta}{2}\right) + \dim(\mathcal{H}_B)\frac{1}{n}\log(n+1), \tag{B8}$$

which approaches zero as $n \to \infty$ and $\delta \to 0$. In order to establish these inequalities, we use the properties of typical projectors described in Section A 2 and encoding unitaries described in Section A 5. We point readers to [3, Chapter 23] for a review of related proofs to establish these inequalities.

We now invoke Lemma 9 to demonstrate the existence of a reliable code. Since the four conditions (B4)–(B7) hold, there exists a POVM $\{\Lambda^{m,k}_{A^nB^nE^n}\}_{m,k}$ that can detect the transmitted states with an arbitrarily low expectation of the average probability of error, as described in Lemma 9. In particular, we get the following upper bound on the expectation of average probability of error:

$$\mathbb{E}_{\mathcal{C}}\left\{\frac{1}{MK}\sum_{m,k}\mathrm{Tr}\{(I - \Lambda^{m,k}_{A^nB^nE^n})U_{A^n}(S(m,k))\rho_{ABE}^{\otimes n}U_{A^n}^\dagger(S(m,k))\}\right\}$$

$$\leq 2\left(\varepsilon + 2\sqrt{\varepsilon}\right) + 4\ MK\ 2^{-n[H(A)_\rho + H(BE)_\rho - \nu(n,\delta) - c\delta]}2^{n[H(ABE)_\rho + c\delta]} \tag{B9}$$

$$\leq 2\left(\varepsilon + 2\sqrt{\varepsilon}\right) + 4\ MK\ 2^{-n[I(A;BE)_\rho - \nu(n,\delta) - 2c\delta]} \tag{B10}$$

$$\leq 2\left(\varepsilon + 2\sqrt{\varepsilon}\right) + 4 \cdot 2^{-nc\delta}, \tag{B11}$$

where we considered the size of the message set and the local key set combined to be

$$MK = 2^{n[I(A;BE) - \nu(n,\delta) - 3c\delta]}\ . \tag{B12}$$

Therefore, the number of bits per channel use encoded by M and K is

$$\frac{1}{n}\log_2 MK = I(A;BE)_\rho - \nu(n,\delta) - 3c\delta \;. \tag{B13}$$

Let $\varepsilon' \in (0,1)$ and $\delta' \in (0,1)$. If we pick n large enough and δ small enough, we can have both $\nu(n,\delta)+3c\delta \leq \delta'$ and $2\left(\varepsilon + 2\sqrt{\varepsilon}\right) + 4\cdot 2^{-nc\delta} \leq \varepsilon'$. Therefore, if $\log_2 MK \approx nI(A;BE)_\rho$, Alice can reliably communicate classical messages to Bob.

We now provide a proof for maintaining privacy from an eavesdropper in the conditional one-time pad task. Consider the following respective codeword subspace projectors and a code subspace projector.

$$U_{A^n}(s)\Pi^{\rho,\delta}_{A^nE^n}U^\dagger_{A^n}(s), \tag{B14}$$

$$\Pi^{\rho,\delta}_{A^n}\otimes\Pi^{\rho,\delta}_{E^n}, \tag{B15}$$

where $\Pi^{\rho,\delta}_{A^nE^n}$, $\Pi^{\rho,\delta}_{A^n}$, and $\Pi^{\rho,\delta}_{E^n}$ are the typical projectors for many copies of the states ρ_{AE}, ρ_A, and ρ_E, respectively.

Furthermore, consider the following ensemble derived from (B1) by tracing over the B^n systems:

$$\left\{\frac{1}{|\mathcal{S}|}, U_{A^n}(s)\rho_{A^nE^n}U^\dagger_{A^n}(s)\right\}_{s\in\mathcal{S}} . \tag{B16}$$

The ensemble average of this ensemble is given by

$$\overline{\tau}_{A^nE^n} \equiv \mathbb{E}_S\left\{U_{A^n}(S)\rho^{\otimes n}_{AE}U^\dagger_{A^n}(S)\right\}. \tag{B17}$$

Since for any message m, Alice picks k uniformly at random from $k\in\{1,\ldots,K\}$, then from the perspective of an Eve who does not know the choice of k, the state has the following form:

$$\tau^m_{A^nE^n} \equiv \frac{1}{K}\sum_{k=1}^{K} U_{A^n}(s)\rho^{\otimes n}_{AE}U^\dagger_{A^n}(s). \tag{B18}$$

As described in Lemma 12, if the four inequalities corresponding to the codeword subspace projectors, the code subspace projector, and the above mentioned ensemble are satisfied, then it is highly likely that $\tau^m_{A^nE^n}$ in (B18) is indistinguishable from $\overline{\tau}_{A^nE^n}$ in (B17).

We now state the four conditions corresponding to the covering lemma for our code:

$$\mathrm{Tr}\{(\Pi^{\rho,\delta}_{A^n}\otimes\Pi^{\rho,\delta}_{E^n})(U_{A^n}(s)\rho_{A^nE^n}U^\dagger_{A^n}(s))\} \geq 1-\varepsilon, \tag{B19}$$

$$\mathrm{Tr}\{(U_{A^n}(s)\Pi^{\rho,\delta}_{A^nE^n}U^\dagger_{A^n}(s))(U_{A^n}(s)\rho_{A^nE^n}U^\dagger_{A^n}(s))\} \geq 1-\varepsilon, \tag{B20}$$

$$\mathrm{Tr}\{\Pi^{\rho,\delta}_{A^n}\otimes\Pi^{\rho,\delta}_{E^n}\} \leq 2^{n(H(A)_\rho+H(E)_\rho+2c\delta)}, \tag{B21}$$

$$\left(U_{A^n}(s)\Pi^{\rho,\delta}_{A^nE^n}U^\dagger_{A^n}(s)\right)U_{A^n}(s)\rho_{A^nE^n}U^\dagger_{A^n}(s)\left(U_{A^n}(s)\Pi^{\rho,\delta}_{A^nE^n}U^\dagger_{A^n}(s)\right) \leq 2^{-n(H(AE)_\rho-c\delta)}\left(U_{A^n}(s)\Pi^{\rho,\delta}_{A^nE^n}U^\dagger_{A^n}(s)\right) , \tag{B22}$$

where c is some constant. Proofs of these properties are available in [3].

We now invoke Lemma 12 and arrive at the following inequality:

$$\Pr_{\mathcal{C}}\{\|\tau^m_{A^nE^n}-\overline{\tau}_{A^nE^n}\|_1 \leq \varepsilon+4\sqrt{\varepsilon}+24\sqrt[4]{\varepsilon}\} \geq 1-2^{n(H(A)_\rho+H(E)_\rho+2c\delta+1/n)}\exp\left(-\frac{\varepsilon^3K2^{-n[I(A;E)_\rho-3c\delta]}}{4}\right) . \tag{B23}$$

Thus, if we choose the size of the key set to be $K=2^{n[I(A;E)_\rho+4c\delta]}$, then $\exp\{-\varepsilon^3 2^{nc\delta}/(4)\}$ is doubly exponentially decreasing in n. Therefore, if $\log_2 K \approx nI(A;E)_\rho$, it is highly likely that the state $\tau^m_{A^nE^n}$ will be nearly indistinguishable from the average state $\overline{\tau}_{A^nE^n}$.

As described earlier, in the conditional one-time pad task, the goal for Alice is to encode information in her share of the state ρ_{ABE} in such a way that Bob can reliably decode the information, while maintaining privacy from Eve. So far, we have shown that Alice can reliably communicate to Bob. Moreover, we have also discussed a strategy that Alice can implement to communicate a classical message m to Bob, such that the quantum state that Eve can access has essentially no dependence on the message m.

Next, we would like to show the existence of a code that is both reliable and secure. Using the union bound of probability theory, it can be shown that there is a non-zero probability for there to be a code $\{s(m,k)\}_{m,k}$ such that the average success probability of Bob's decoder is arbitrarily high and $\|\tau^m_{A^nE^n} - \overline{\tau}_{A^nE^n}\|_1$ is arbitrarily small for all m, with these statements holding for sufficiently large n. Furthermore, a final "expurgation" argument can be applied to show that Bob can decode each m and k with arbitrarily high probability and that $\|\tau^m_{A^nE^n} - \overline{\tau}_{A^nE^n}\|_1$ is arbitrarily small for all m. These techniques have been used to establish a formula for the capacity of quantum channel for transmitting private classical information in [27, 28]. We point readers to [3, Chapter 23] for a review of a related proof to establish the desired result.

Therefore, the number of bits that Alice can communicate securely is

$$\log_2 M = \log_2 MK - \log_2 K \approx n[I(A;BE)_\rho - I(A;E)_\rho] = nI(A;B|E)_\rho \ , \tag{B24}$$

and $I(A;B|E)_\rho$ is an achievable rate. This concludes the achievability proof.

Appendix C: Optimality of CQMI for conditional one-time pad

In this appendix, we establish that the conditional one-time pad capacity of ρ_{ABE} cannot exceed $I(A;B|E)_\rho$. To see this, consider an arbitrary $(n, M, \varepsilon, \delta)$ protocol of the above form, and suppose that the message m is chosen uniformly at random. Then the overall state that describes all systems is

$$\omega_{\hat{M}A'B^nE^n} \equiv \frac{1}{M}\sum_{m=1}^{M} |m\rangle\langle m|_{\hat{M}} \otimes \omega^m_{A'B^nE^n}, \tag{C1}$$

where $\{|m\rangle_{\hat{M}}\}_m$ is an orthonormal basis and the state $\omega^m_{A'B^nE^n}$ is defined in the main text. We can describe Bob's decoding measurement as a measurement channel

$$\mathcal{M}_{A'B^nE^n\to M'}(\theta_{A'B^nE^n}) \equiv \sum_m \mathrm{Tr}\{\Lambda^m_{A'B^nE^n}\theta_{A'B^nE^n}\}|m\rangle\langle m|_{M'}, \tag{C2}$$

so that the final output state is

$$\omega_{\hat{M}M'} = \mathcal{M}_{A'B^nE^n\to M'}(\omega_{\hat{M}A'B^nE^n}). \tag{C3}$$

By the condition in (3) and some further calculations, it follows that

$$\frac{1}{2}\left\|\omega_{\hat{M}M'} - \overline{\Phi}_{\hat{M}M'}\right\|_1 \leq \varepsilon, \tag{C4}$$

where $\overline{\Phi}_{\hat{M}M'} \equiv \frac{1}{M}\sum_{m=1}^{M}|m\rangle\langle m|_{\hat{M}} \otimes |m\rangle\langle m|_{M'}$ is a maximally classically correlated state. A uniform bound for the continuity of mutual information [44] implies that

$$\log_2 M = I(\hat{M};M')_{\overline{\Phi}} \tag{C5}$$

$$\leq I(\hat{M};M')_\omega + \varepsilon \log_2 M + g(\varepsilon), \tag{C6}$$

where $g(\varepsilon) \equiv (\varepsilon+1)\log_2(\varepsilon+1) - \varepsilon\log_2\varepsilon$, with the property that $\lim_{\varepsilon\to 0} g(\varepsilon) = 0$. From the Holevo bound [45] or more generally quantum data processing (see, e.g., [3, Section 11.9.2]), it follows that

$$I(\hat{M};M')_\omega \leq I(\hat{M};A'B^nE^n)_\omega. \tag{C7}$$

By the condition in (4), it follows that

$$\frac{1}{2}\left\|\omega_{\hat{M}A'E^n} - \omega_{\hat{M}} \otimes \sigma_{A'E^n}\right\|_1 \leq \delta, \tag{C8}$$

which in turn implies from [44] that

$$I(\hat{M};A'E^n)_\omega \leq \delta \log_2 M + g(\delta). \tag{C9}$$

Putting everything together leads to the following bound:

$$\log_2 M \leq I(\hat{M};B^n|A'E^n)_\omega + (\varepsilon+\delta)\log_2 M + g(\varepsilon) + g(\delta), \tag{C10}$$

where we used that $I(\hat{M};A'B^nE^n)_\omega - I(\hat{M};A'E^n)_\omega = I(\hat{M};B^n|A'E^n)_\omega$. Now by several applications of the chain rule for conditional mutual information, we find that

$$I(\hat{M};B^n|A'E^n)_\omega = \sum_{i=1}^{n} I(\hat{M};B_i|B^{i-1}A'E^n)_\omega \tag{C11}$$

$$\leq \sum_{i=1}^{n} I(\hat{M}A'B^{i-1}E^{i-1}E^n_{i+1};B_i|E_i)_\omega \tag{C12}$$

$$\leq \sum_{i=1}^{n} I(A_i;B_i|E_i)_{\rho^{\otimes n}} \tag{C13}$$

$$= nI(A;B|E)_\rho. \tag{C14}$$

The second inequality follows because we can consider the sequential action of 1) tensoring in the states $\rho_{ABE}^{\otimes i-1}$ and $\rho_{ABE}^{\otimes n-i}$ to the ith copy of ρ_{ABE}, 2) tensoring in the state $\frac{1}{M}\sum_{m=1}^{M}|m\rangle\langle m|_{\hat{M}}$, 3) applying the encoding $\mathcal{E}^m_{A^n\to A'}$ conditioned on the value m in $\hat{M}$, and 4) tracing over the systems B^n_{i+1} all as a local channel $\mathcal{N}_{A_i\to\hat{M}A'B^{i-1}E^{i-1}E^n_{i+1}}$ acting on the A_i system of $\rho_{A_iB_iE_i}$, so that

$$\omega_{\hat{M}A'B^{i-1}E^{i-1}B_iE_i} = \mathcal{N}_{A_i\to\hat{M}A'B^{i-1}E^{i-1}E^n_{i+1}}(\rho_{A_iB_iE_i}), \tag{C15}$$

and the conditional mutual information does not increase under the action of a local channel on an unconditioned system [11]:

$$I(\hat{M}A'B^{i-1}E^{i-1};B_i|E_i)_\omega \leq I(A_i;B_i|E_i)_{\rho^{\otimes n}}. \tag{C16}$$

Alternatively, the inequality resulting from (C11)–(C14) may be seen by the following steps:

$$I(\hat{M};B^n|A'E^n)_\omega = I(\hat{M}A';B^n|E^n)_\omega - I(A';B^n|E^n)_\omega \tag{C17}$$

$$\leq I(\hat{M}A';B^n|E^n)_\omega \tag{C18}$$

$$\leq I(A^n;B^n|E^n)_{\rho^{\otimes n}} \tag{C19}$$

$$= nI(A;B|E)_\rho. \tag{C20}$$

The first inequality follows from non-negativity of conditional mutual information [9, 10], and the second inequality follows from monotonicity of conditional mutual information [11] with respect to a local channel acting on one of the unconditioned systems [in this case, the local channel is the encoding channel that tensors in the maximally mixed state on system $\hat{M}$ and applies the channel $\sum_m |m\rangle\langle m|_{\hat{M}}(\cdot)|m\rangle\langle m|_{\hat{M}}\otimes\mathcal{E}^m_{A^n\to A'}(\cdot)$].

Putting everything together, we find the following bound for any (n,M,ε,δ) conditional one-time pad protocol:

$$\frac{1-\varepsilon-\delta}{n}\log_2 M \leq I(A;B|E)_\rho + \frac{g(\varepsilon)+g(\delta)}{n}. \tag{C21}$$

Taking the limit as $n\to\infty$ and then as $\varepsilon,\delta\to 0$ allows us to conclude that the conditional mutual information $I(A;B|E)_\rho$ is an upper bound on the conditional one-time pad capacity of ρ_{ABE}.

Appendix D: A proof of the converse theorem for a secret-sharing task

In this section, we provide a proof of the converse theorem for a secret-sharing task that we call *information scrambling*. The goal of the information scrambling task is for Alice to communicate a message in such a way that it can be decoded only by someone who possesses all three ABE systems. If someone possesses only the AB systems or only the AE systems, then such a person can figure out essentially nothing about the encoded message.

By using arguments similar to (C4)–(C10), we find the following two inequalities:

$$\log_2 M \leq nI(A;B|E)_\rho + (\varepsilon+\delta)\log_2 M + g(\varepsilon) + g(\delta)\ . \tag{D1}$$

$$\log_2 M \leq nI(A;E|B)_\rho + (\varepsilon+\delta)\log_2 M + g(\varepsilon) + g(\delta)\ . \tag{D2}$$

Therefore, by combining (D1) and (D2), we arrive at the following bound for any (n,M,ε,δ) secret-sharing protocol:

$$\frac{1-\varepsilon-\delta}{n}\log_2 M \leq I(A;BE)_\rho - \max\{I(A;B)_\rho, I(A;E)_\rho\} + \frac{g(\varepsilon)+g(\delta)}{n}. \tag{D3}$$

Taking the limit as $n \to \infty$ and then as $\varepsilon, \delta \to 0$ allows us to conclude that $I(A;BE)_\rho - \max\{I(A;B)_\rho, I(A;E)_\rho\}$ is an upper bound on the information scrambling capacity of ρ_{ABE}.

The proof for the achievability part follows similarly to the achievability part for the conditional one-time pad task, except that we have extra security conditions that should hold. To handle this, we just invoke the covering lemma again to be sure that the key variable is large enough to protect the message variable against local parties who do not have access to all systems of the full state.

Appendix E: A proof of the converse theorem for the communication protocol in [31]

We begin by recalling the communication protocol described in [31]. Suppose that Alice, Bob, and Eve share n copies of the quantum state ρ_{ABE}, so that their state is $\rho_{ABE}^{\otimes n}$. In this communication protocol, Alice, Bob, and Eve have access to the A, B, and E systems, respectively. Alice and Bob are connected by an ideal quantum channel, which Eve has access to as well. The goal of this protocol is for Alice to encode a message m into her A systems, in such a way that if she sends her A systems over the ideal quantum channel, then Bob can decode the message m reliably by performing a measurement on all of the AB systems, while Eve, possessing the AE systems, has essentially no chance of determining the message m if she tried to figure it out.

At the beginning of the protocol, Alice picks $m \in \{1, \ldots, M\}$ and applies an encoding channel $\mathcal{E}^m_{A^n\to A'}$ to the A^n systems of $\rho_{ABE}^{\otimes n}$, leading to the state $\omega^m_{A'B^nE^n} \equiv \mathcal{E}^m_{A^n\to A'}(\rho_{ABE}^{\otimes n})$. She transmits the system A' of $\omega^m_{A'B^nE^n}$ over the ideal quantum channel. Bob applies a decoding positive operator-valued measure $\{\Lambda^m_{A'B^n}\}_m$ to the systems $A'B^n$ of $\omega^m_{A'B^nE^n}$ in order to figure out which message was transmitted. The protocol is ε-reliable if Bob can determine the message m with probability not smaller than $1-\varepsilon$:

$$\forall m : \mathrm{Tr}\{\Lambda^m_{A'B^n}\omega^m_{A'B^n}\} \geq 1-\varepsilon. \tag{E1}$$

The protocol is δ-secure if the reduced state $\omega^m_{A'E^n}$ on systems $A'E^n$ is nearly indistinguishable from a constant state $\sigma_{A'E^n}$ independent of the message m:

$$\forall m : \frac{1}{2}\left\|\omega^m_{A'E^n} - \sigma_{A'E^n}\right\|_1 \leq \delta \ . \tag{E2}$$

Achievable rates and capacity are defined similarly to the previous cases.

We now provide a proof to establish an upper bound on the rate of communication for an arbitrary protocol of the above form, which is different from the proof given in [31]. Let the message m be chosen uniformly at random, and the overall state that describes all systems is defined in the same way as in (C1). We can describe Bob's decoding measurement as a measurement channel similar to (C2):

$$\mathcal{M}_{A'B^n\to M'}(\theta_{A'B^n}) \equiv \sum_m \mathrm{Tr}\{\Lambda^m_{A'B^n}\theta_{A'B^n}\}|m\rangle\langle m|_{M'}, \tag{E3}$$

so that the final output state is

$$\omega_{\hat{M}M'} = \mathcal{M}_{A'B^n\to M'}(\omega_{\hat{M}A'B^n}), \tag{E4}$$

where

$$\omega_{\hat{M}A'B^nE^n} \equiv \sum_m \frac{1}{M}|m\rangle\langle m|_{\hat{M}} \otimes \omega^m_{A'B^nE^n}. \tag{E5}$$

By using arguments similar to (C4)–(C10), we find the following bound:

$$\log_2 M \leq I(\hat{M};A'B^n)_\omega - I(\hat{M};A'E^n)_\omega + (\varepsilon+\delta)\log_2 M + g(\varepsilon) + g(\delta) \ . \tag{E6}$$

Now by several applications of the chain rule for conditional mutual information, we find that

$$I(\hat{M};A'B^n)_\omega - I(\hat{M};A'E^n)_\omega = I(\hat{M};B^n|A')_\omega - I(\hat{M};E^n|A')_\omega \tag{E7}$$

$$= \sum_{i=1}^n I(\hat{M};B_i|B_1^{i-1}E_{i+1}^nA')_\omega - I(\hat{M};E_i|B_1^{i-1}E_{i+1}^nA')_\omega \tag{E8}$$

$$\leq \sum_{i=1}^n \sup_{\mathcal{N}_{A\to A''A'''}} [I(A'';B_i|A''')_\tau - I(A'';E_i|A''')_\tau] \tag{E9}$$

$$= n \sup_{\mathcal{N}_{A\to A''A'''}} [I(A'';B|A''')_\tau - I(A'';E|A''')_\tau] \ , \tag{E10}$$

where the first inequality follows because the action of 1) tensoring in the states $\rho_{ABE}^{\otimes i-1}$ and $\rho_{ABE}^{\otimes n-i}$ to the ith copy of ρ_{ABE}, 2) tensoring in the state $\frac{1}{M}\sum_{m=1}^{M}|m\rangle\langle m|_{\hat{M}}$, 3) applying the encoding $\mathcal{E}_{A^n\to A'}^{m}$ conditioned on the value m in $\hat{M}$, and 4) tracing over the systems B_{i+1}^{n} and E_1^{i-1} can be understood as a local channel $\mathcal{N}_{A_i\to\hat{M}A'B_1^{i-1}E_{i+1}^{n}}$ acting on the A_i system of $\rho_{A_iB_iE_i}$. Relabeling system $\hat{M}$ as A'' and systems $A'B_1^{i-1}E_{i+1}^{n}$ as A''', and defining

$$\tau_{A''A'''BE} \equiv \mathcal{N}_{A\to A''A'''}(\rho_{ABE}), \tag{E11}$$

we arrive at the inequality, following from the fact that the supremum is taken over quantum channels $\mathcal{N}_{A\to A''A'''}$.

Putting everything together, we find the following bound on the number of bits that Alice can communicate securely:

$$\frac{1-\varepsilon-\delta}{n}\log_2 M \leq \sup_{\mathcal{N}_{A\to A''A'''}} [I(A'';B|A''')_\tau - I(A'';E|A''')_\tau] + \frac{g(\varepsilon)+g(\delta)}{n}. \tag{E12}$$

Taking the limit as $n\to\infty$ and then as $\varepsilon,\delta\to 0$ allows us to conclude that

$$\sup_{\mathcal{N}_{A\to A''A'''}} [I(A'';B|A''')_\tau - I(A'';E|A''')_\tau] \tag{E13}$$

is an upper bound on the capacity of ρ_{ABE} for this communication task.

Appendix F: A proof of the direct coding theorem for the communication protocol in [31]

We now provide a proof of the direct coding theorem for the communication protocol described above, which follows from the coding scheme developed in Section B. Let $M,K\in\mathbb{N}$. Alice has message variable $m\in\{1,\ldots,M\}$ and a local key variable $k\in\{1,\ldots,K\}$. Before communicating to Bob, Alice can apply any local operation to the A^n systems. Suppose that Alice applies a quantum channel $\mathcal{N}_{A\to A''A'''}$ to all n copies of the state ρ_{ABE}. Then the overall state that describes all systems is

$$\hat{\rho}_{A''A'''BE}^{\otimes n} = \mathcal{N}_{A\to A''A'''}^{\otimes n}(\rho_{ABE}^{\otimes n}) \ . \tag{F1}$$

Similar to the coding scheme developed in Section B, if Alice wishes to send message m, then she picks k uniformly at random from $k\in\{1,\ldots,K\}$. For each pair (m,k), the random code is selected in such a way that a vector s, of the form described in (A35), is chosen uniformly at random and associated with the pair (m,k). So if Alice wishes to send the pair (m,k), she applies the encoding unitary $U_{A''^n}(s(m,k))$ to the state $\hat{\rho}_{A''A'''BE}^{\otimes n}$ and sends $A''^nA'''^n$ systems to Bob over the ideal quantum channel. Then we could structure a coding scheme similar to our achievability proof in Section B, such that

$$\log_2 MK \approx nI(A'';BA'''), \tag{F2}$$
$$\log_2 K \approx nI(A'';EA''') \ . \tag{F3}$$

Then if

$$I(A'';BA''') - I(A'';EA''') \tag{F4}$$

is strictly positive, the coding scheme guarantees that this information difference is an achievable rate. Moreover, Alice can further improve the achievable rate of secure communication by optimizing over quantum channels $\mathcal{N}_{A\to A''A'''}$. Therefore, the following rate is an achievable rate:

$$\sup_{\mathcal{N}_{A\to A''A'''}} [I(A'';B|A''') - I(A'';E|A''')] \ . \tag{F5}$$

This concludes the achievability proof.