3-1-2015

Multiplicativity of Completely Bounded p-Norms Implies a Strong Converse for Entanglement-Assisted Capacity

Manish K. Gupta
Louisiana State University

Mark M. Wilde
Louisiana State University

Multiplicativity of completely bounded p-norms implies a strong converse for entanglement-assisted capacity

Manish K. Gupta
Hearne Institute for Theoretical Physics,
Department of Physics and Astronomy,
Louisiana State University,
Baton Rouge, Louisiana 70803, USA

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

July 4, 2018

Abstract

The fully quantum reverse Shannon theorem establishes the optimal rate of noiseless classical communication required for simulating the action of many instances of a noisy quantum channel on an arbitrary input state, while also allowing for an arbitrary amount of shared entanglement of an arbitrary form. Turning this theorem around establishes a strong converse for the entanglement-assisted classical capacity of any quantum channel. This paper proves the strong converse for entanglement-assisted capacity by a completely different approach and identifies a bound on the strong converse exponent for this task. Namely, we exploit the recent entanglement-assisted "meta-converse" theorem of Matthews and Wehner, several properties of the recently established sandwiched Rényi relative entropy (also referred to as the quantum Rényi divergence), and the multiplicativity of completely bounded p-norms due to Devetak *et al*. The proof here demonstrates the extent to which the Arimoto approach can be helpful in proving strong converse theorems, it provides an operational relevance for the multiplicativity result of Devetak *et al*., and it adds to the growing body of evidence that the sandwiched Rényi relative entropy is the correct quantum generalization of the classical concept for all $\alpha > 1$.

1 Introduction

An important lesson learned in theoretical quantum information science is that entanglement assistance tends to simplify problems of interest and, perhaps surprisingly, makes such problems more like their classical counterparts. For example, in quantum computational complexity theory, the canonical QIP-complete problem is distinguishing two quantum channels specified by quantum circuits [41]. In order to solve this problem, two parties, traditionally called a prover and verifier, engage in an entanglement-assisted discrimination strategy, and furthermore, it is known that the computational complexity of this task does not increase if these parties engage in a completely classical strategy to solve this problem [25]. In the theory of quantum error correction, the entanglement-assisted stabilizer formalism allows for producing a quantum error-correcting code from an arbitrary classical error-correcting code, while retaining the desirable properties of the imported classical code [11, 12]. However, importing arbitrary classical codes is not possible if entanglement assistance is not available. Furthermore, entanglement assistance helps to resolve technical problems that arise in the construction of quantum LDPC [23, 24] and turbo codes [48]. In quantum rate distortion theory (the theory of lossy quantum data compression), the most well understood setting is again the entanglement-assisted setting, with there being a simple formula that characterizes optimal compression rates [15].

Perhaps the earliest observation in this spirit is due to Bennett *et al.* [7, 8] and Holevo [22], who established that a simple formula characterizes the capacity of a quantum channel for classical communication when unlimited entanglement assistance is available. This result is one of the strongest in quantum Shannon theory and provides a "fully quantum" generalization of Shannon's well known formula for the classical capacity of a classical channel [43]. Furthermore, this formula is robust under the presence of a noiseless quantum feedback channel from receiver to sender—Bowen established that the entanglement-assisted capacity does not increase in the presence of such a quantum feedback channel [10].

In later work, Bennett *et al.* [6] and Berta *et al.* [9] strengthened the interpretation of the formula for entanglement-assisted capacity, by demonstrating that a so-called strong converse theorem holds in this setting. A strong converse theorem establishes that, if the rate of communication in any given coding scheme exceeds the capacity, then the error probability of this scheme tends to one in the limit of many channel uses. Coupled with the achievability part of a coding theorem (that there always exist a coding scheme with error probability tending to zero in the limit of many channel uses if the rate of communication is less than capacity), a strong converse theorem establishes the capacity as a very sharp line dividing achievable communication rates from unachievable ones. Furthermore, strong converse theorems find applications in establishing security in particular models of cryptography [29].

Bennett *et al.* and Berta *et al.* established a strong converse theorem for entanglement-assisted capacity by proving what is known as the entanglement-assisted quantum reverse Shannon theorem. Such a theorem corresponds to a compression-like quantum information processing task and characterizes the optimal rates of communication at which it is possible to simulate a quantum channel. In entanglement-assisted channel simulation, the goal is for a sender and receiver, who share an unlimited amount of entanglement before the protocol begins, to use as few noiseless classical bit channels as possible to simulate the action of many independent instances of the channel on any quantum input (this input can be entangled with another system not fed into the channel), in such a way that any third party should not be able to distinguish between the original channels and the simulation. Interestingly, the rate of communication required for channel simulation corresponds to a strong converse rate for capacity, because, if it were possible to use the channel to communicate at a rate larger than its channel simulation rate, then a sender and receiver could "get out" more communication than they invested originally (essentially getting "something for nothing, bits for free"). Carrying this *reductio ad absurdum* argument out in more detail, one can show that the error probability in fact increases exponentially fast to one if the rate of communication exceeds the channel simulation rate. The main result of Bennett *et al.* and Berta *et al.* is that the optimal channel simulation rate in the presence of shared entanglement is equal to the entanglement-assisted capacity of the channel, and by the above argument, this establishes a strong converse theorem for the entanglement-assisted capacity.

2 Summary of results

In this paper, we establish a strong converse theorem for the entanglement-assisted capacity by a route completely different from that of Bennett *et al.* and Berta *et al.* Furthermore, we identify a bound on the strong converse exponent for this task. Our motivation is two-fold: first, for a theorem as important as this one, it is certainly reasonable to have multiple proofs to shed further light on the topic. More importantly, our approach here might illuminate alternate ways for establishing a strong converse theorem for other capacities, the most pressing of which is the quantum capacity of degradable channels [33]. Our approach taken here is in the line of Arimoto [2], a route by which several strong converse theorems have now been established [39, 28, 49] and for which the general framework has been extended significantly [40, 44]. Our proof makes use of the sandwiched Rényi relative entropy [36, 49], known also by the name of "quantum Rényi divergence" [36], and it exploits several properties of this entropy in order to establish the strong converse theorem.

A pleasing aspect of the present paper is that it provides an operational relevance for the main result of Devetak, Junge, King, and Ruskai [17] (see related follow up work in [26]). Indeed, in the present paper, for simplicity, we will *define*[1] the completely bounded $1 \to \alpha$ norm of a completely positive map $\mathcal{M}$ as

$$\|\mathcal{M}\|_{\mathrm{CB},1\to\alpha} \equiv \max_{\rho_A} \left\| \left(\rho_A^{1/2\alpha} \otimes I_B\right) \Gamma_{AB}^{\mathcal{M}} \left(\rho_A^{1/2\alpha} \otimes I_B\right) \right\|_\alpha, \tag{2.1}$$

where $\Gamma_{AB}^{\mathcal{M}}$ is the Choi matrix of the channel $\mathcal{M}$, $\|\cdot\|_\alpha$ denotes the Schatten α-norm for $\alpha \geq 1$, and the optimization is over density operators ρ_A (see the next section for formal definitions of these objects). A special case of the main result of [17] is that these norms are multiplicative for all completely positive maps $\mathcal{M}_1$ and $\mathcal{M}_2$, in the sense that

$$\|\mathcal{M}_1 \otimes \mathcal{M}_2\|_{\mathrm{CB},1\to\alpha} = \|\mathcal{M}_1\|_{\mathrm{CB},1\to\alpha} \|\mathcal{M}_2\|_{\mathrm{CB},1\to\alpha}. \tag{2.2}$$

Due to the connection between the sandwiched Rényi relative entropy and α-norms, we can apply the above multiplicativity result to our setting in order to establish a strong converse for entanglement-assisted capacity. Furthermore, it is interesting to observe that the completely bounded α-norm in (2.1) converges to the so-called "diamond norm" [27] in the limit as $\alpha \to 1$. The diamond norm is used all throughout quantum information theory as a measure of distance between quantum channels [1, 20, 41, 42, 46] because it is the operationally relevant distance measure in the setting of entanglement-assisted discrimination of quantum channels (the most general strategy that one could use to distinguish quantum channels). Of course, this is the setting with which one would be dealing if trying to determine the value of a single bit encoded with an entanglement-assisted communication strategy, so it appears that our connection of the main result of [17] to entanglement-assisted capacity is the natural one to make.[2]

We now outline our proof for the strong converse of the entanglement-assisted classical capacity, and the following sections give detailed arguments.

[1] Note that [17] defined the completely bounded norms $\|\cdot\|_{\mathrm{CB},p\to q}$ in a very general way and proved that these norms reduce to the expression in (2.1) when $p = 1$ and $q = \alpha$.

[2] The authors of [17] connected their main technical result to additivity of a quantity now known as reverse coherent information [19]. However, in spite of the statements made in [19], we are not convinced that the reverse coherent information possesses a compelling operational interpretation.

1. We say that a quantity is a generalized divergence [40, 44] (a generalization of von Neumann relative entropy) if it satisfies the following monotonicity inequality for all density operators ρ and σ and channels $\mathcal{N}$:
$$\mathbf{D}(\rho\|\sigma) \geq \mathbf{D}(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)). \tag{2.3}$$
From such a divergence, we can derive a generalized mutual information of a quantum channel according to the following recipe:
$$I_{\mathbf{D}}(\mathcal{N}) \equiv \max_{\rho_A} I_{\mathbf{D}}(A;B)_\omega \tag{2.4}$$
$$\omega_{AB} \equiv \rho_A^{1/2}\Gamma_{AB}^{\mathcal{N}}\rho_A^{1/2}, \tag{2.5}$$
$$\Gamma_{AB}^{\mathcal{N}} \equiv \mathcal{N}_{A'\to B}(\Gamma_{AA'}), \tag{2.6}$$
$$\Gamma_{AA'} \equiv |\Gamma\rangle\langle\Gamma|_{AA'}, \tag{2.7}$$
$$|\Gamma\rangle_{AA'} \equiv \sum_i |i\rangle_A|i\rangle_{A'}, \tag{2.8}$$
$$I_{\mathbf{D}}(A;B)_\tau \equiv \min_{\sigma_B} \mathbf{D}(\tau_{AB}\|\tau_A \otimes \sigma_B). \tag{2.9}$$
(We explain all of these quantities in further detail in the main text.) Our first step then is to recall [32, Propositions 20 and 21], which establish a relationship between Type I and II errors in hypothesis testing and the rate R and success probability for any (n, R, ε) entanglement-assisted code (a code that uses n instances of a channel $\mathcal{N}$ at a fixed rate R and has an error probability no larger than ε). We use this "meta converse" theorem in order to obtain an upper bound on any entanglement-assisted code's success probability in terms of its rate, blocklength, and any generalized mutual information derived from a generalized divergence $\mathbf{D}(\rho\|\sigma)$ as above.

2. Next we recall the definition of the sandwiched Rényi relative entropy [36, 49], also referred to as the quantum Rényi divergence [36], which is a particular divergence between two density operators ρ and σ. It is defined [36, 49] for $\alpha \in (1, \infty)$ as
$$\widetilde{D}_\alpha(\rho\|\sigma) \equiv \frac{1}{\alpha-1}\log \mathrm{Tr}\left\{\left(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right\}, \tag{2.10}$$
when the support of ρ is contained in the support of σ and it is equal to $+\infty$ otherwise (in this work, we focus exclusively on the regime $\alpha > 1$). In particular, $\widetilde{D}_\alpha(\rho\|\sigma)$ was shown to obey the monotonicity inequality mentioned above for all $\alpha \in (1, 2]$ [49, 36] and later work proved that this monotonicity holds for all $\alpha \in [1/2, 1) \cup (1, \infty)$ [18, 5]. Furthermore, this quantity converges to the von Neumann relative entropy in the limit as $\alpha \searrow 1$. This is one reason why the sandwiched Rényi relative entropy is relevant for us in establishing a strong converse for entanglement-assisted capacity.

Remark: In a few recent works, it has been said, somewhat ambiguously, that this new Rényi relative entropy is useful because it captures the "non-commutativity of quantum states." However, the previous notion of Rényi relative entropy, $D_\alpha(\rho\|\sigma) \equiv \frac{1}{\alpha-1}\log\mathrm{Tr}\{\rho^\alpha\sigma^{1-\alpha}\}$ is perfectly well defined for non-commutative quantum states and proves to be useful in the regime when $\alpha \in [0, 1)$ [3, 21, 37, 4, 38, 34, 35]. In our opinion, the sandwiched Rényi relative entropy has proved useful when $\alpha > 1$ because it can be related to a Schatten α-norm in the following way:

$$\widetilde{D}_\alpha(\rho\|\sigma) = \frac{\alpha}{\alpha-1}\log\left\|\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}\right\|_\alpha, \tag{2.11}$$

and there are many properties of these α-norms and prior results established for them that come into play when establishing properties of information measures derived from the sandwiched Rényi relative entropy (the present paper being no exception). Thus, there is a growing consensus [49, 36, 35, 16] that the sandwiched Rényi relative entropy is the correct generalization of the classical Rényi relative entropy at least for the regime $\alpha > 1$.

3. We evaluate the upper bound on success probability mentioned in the first step above, by using the sandwiched Rényi relative entropy as the divergence. This yields the following upper bound on the success probability of any rate R entanglement-assisted scheme that uses a channel n times:

$$p_{\text{succ}} \leq 2^{-n\sup_{\alpha>1}\left(\frac{\alpha-1}{\alpha}\right)\left(R-\frac{1}{n}\widetilde{I}_\alpha(\mathcal{N}^{\otimes n})\right)}, \tag{2.12}$$

where $\widetilde{I}_\alpha(\mathcal{M})$ is the sandwiched Rényi mutual information of a quantum channel $\mathcal{M}$, derived by the same recipe in (2.4)-(2.9), taking $\mathbf{D} = \widetilde{D}_\alpha$. By inspecting the above formula, we can observe that if additivity of $\widetilde{I}_\alpha$ holds for $\alpha \in (1,\infty)$, i.e.,

$$\widetilde{I}_\alpha(\mathcal{N}^{\otimes n}) = n\widetilde{I}_\alpha(\mathcal{N}), \tag{2.13}$$

then by a standard argument [39, 28], which we elaborate for our case here, the strong converse follows.

4. As a precursor to proving additivity of $\widetilde{I}_\alpha(\mathcal{N}^{\otimes n})$, we relate the sandwiched Rényi mutual information of a channel $\mathcal{N}$ to an α-norm of the states involved

$$\begin{aligned}\widetilde{I}_\alpha(\mathcal{N}) &= \max_{\rho_A}\min_{\sigma_B}\widetilde{D}_\alpha\left(\rho_A^{1/2}\Gamma^{\mathcal{N}}_{AB}\rho_A^{1/2}\middle\|\rho_A\otimes\sigma_B\right) &(2.14)\\ &= \max_{\rho_A}\min_{\sigma_B}\frac{\alpha}{\alpha-1}\log\left\|\left(\left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2}\left(\rho_A^{\frac{1}{\alpha}}\otimes\sigma_B^{\frac{1-\alpha}{\alpha}}\right)\left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2}\right)\right\|_\alpha &(2.15)\end{aligned}$$

Using Hölder duality of norms and the Lieb concavity theorem [30], we then show that

$$\left\|\left(\left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2}\left(\rho_A^{\frac{1}{\alpha}}\otimes\sigma_B^{\frac{1-\alpha}{\alpha}}\right)\left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2}\right)\right\|_\alpha \tag{2.16}$$

is concave in ρ_A for $\alpha \in (1,\infty)$. Convexity of the α-norm and operator convexity of $x^{(1-\alpha)/\alpha}$ for $\alpha \in (1,\infty)$ implies that the above function is convex in σ_B for $\alpha \in (1,\infty)$. These properties are sufficient for us to apply the Sion minimax theorem [45] in order to exchange the minimum with the maximum for $\alpha \in (1,\infty)$:

$$\begin{aligned}\widetilde{I}_\alpha(\mathcal{N}) &= \max_{\rho_A}\min_{\sigma_B}\widetilde{D}_\alpha\left(\rho_A^{1/2}\Gamma^{\mathcal{N}}_{AB}\rho_A^{1/2}\middle\|\rho_A\otimes\sigma_B\right) &(2.17)\\ &= \min_{\sigma_B}\max_{\rho_A}\widetilde{D}_\alpha\left(\rho_A^{1/2}\Gamma^{\mathcal{N}}_{AB}\rho_A^{1/2}\middle\|\rho_A\otimes\sigma_B\right). &(2.18)\end{aligned}$$

5. From here, we can exploit the multiplicativity of completely bounded α-norms [17] and [5, Theorem 11] to establish that the sandwiched Rényi mutual information is additive as a function of quantum channels, in the sense that

$$\widetilde{I}_\alpha(\mathcal{N}_1 \otimes \mathcal{N}_2) = \widetilde{I}_\alpha(\mathcal{N}_1) + \widetilde{I}_\alpha(\mathcal{N}_2), \tag{2.19}$$

for all quantum channels $\mathcal{N}_1$ and $\mathcal{N}_2$ and all $\alpha \in (1, \infty)$. This additivity result along with an inductive argument gives us the additivity relation in (2.13).

6. Combining the above results, we obtain the following bound on the success probability for any (n, R, ε) entanglement-assisted coding scheme for a channel $\mathcal{N}$:

$$p_{\text{succ}} \leq 2^{-n \sup_{\alpha>1}\left(\frac{\alpha-1}{\alpha}\right)\left(R-\widetilde{I}_\alpha(\mathcal{N})\right)}. \tag{2.20}$$

Finally, by a standard argument [39, 44] (which we elaborate for our case here), we can choose $\varepsilon > 0$ such that $\widetilde{I}_\alpha(\mathcal{N}) < I(\mathcal{N}) + \varepsilon$ for all $\alpha > 1$ in some neighborhood of 1, so that the success probability decays exponentially fast to zero with n if $R > I(\mathcal{N})$, where $I(\mathcal{N})$ is the entanglement-assisted capacity of the channel $\mathcal{N}$ (in this case, $I(\mathcal{N})$ can be constructed according to the recipe in (2.4)-(2.9) with the generalized divergence taken as the von Neumann relative entropy). The strong converse theorem for the entanglement-assisted capacity then follows.

The next section reviews some notations and definitions, and the rest of the paper proceeds in the order above, giving detailed proofs for each step. We then conclude with a brief summary.

3 Notation and Definitions

Let $\mathcal{B}(\mathcal{H})$ denote the algebra of bounded linear operators acting on a Hilbert space $\mathcal{H}$. We restrict ourselves to finite-dimensional Hilbert spaces throughout this paper. The α-norm of an operator X is defined as

$$\|X\|_\alpha \equiv \operatorname{Tr}\{(\sqrt{X^\dagger X})^\alpha\}^{1/\alpha}, \tag{3.1}$$

where $\alpha \geq 1$. Let $\mathcal{B}(\mathcal{H})_+$ denote the subset of positive semi-definite operators (we often simply say that an operator is "positive" if it is positive semi-definite). We also write $X \geq 0$ if $X \in \mathcal{B}(\mathcal{H})_+$. An operator ρ is in the set $\mathcal{S}(\mathcal{H})$ of density operators if $\rho \in \mathcal{B}(\mathcal{H})_+$ and $\operatorname{Tr}\{\rho\} = 1$. The tensor product of two Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ is denoted by $\mathcal{H}_A \otimes \mathcal{H}_B$. Given a multipartite density operator $\rho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$, we unambiguously write $\rho_A = \operatorname{Tr}_B\{\rho_{AB}\}$ for the reduced density operator on system A. A linear map $\mathcal{N}_{A\to B} : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ is positive if $\mathcal{N}_{A\to B}(\sigma_A) \in \mathcal{B}(\mathcal{H}_B)_+$ whenever $\sigma_A \in \mathcal{B}(\mathcal{H}_A)_+$. Let id_A denote the identity map acting on a system A. A linear map $\mathcal{N}_{A\to B}$ is completely positive if the map $\text{id}_R \otimes \mathcal{N}_{A\to B}$ is positive for a reference system R of arbitrary size. A linear map $\mathcal{N}_{A\to B}$ is trace-preserving if $\operatorname{Tr}\{\mathcal{N}_{A\to B}(\tau_A)\} = \operatorname{Tr}\{\tau_A\}$ for all input operators $\tau_A \in \mathcal{B}(\mathcal{H}_A)$. If a linear map is completely positive and trace-preserving (CPTP), we say that it is a quantum channel or quantum operation. A positive operator-valued measure (POVM) is a set $\{\Lambda^m\}$ of positive operators such that $\sum_m \Lambda^m = I$.

The sandwiched Rényi relative entropy [36, 49], also referred to as the quantum Rényi divergence [36], between two density operators ρ and σ is defined for $\alpha \in (1, \infty)$ as follows:

$$\widetilde{D}_\alpha(\rho\|\sigma) \equiv \frac{1}{\alpha - 1} \log \operatorname{Tr}\left\{\left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho \sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right\}, \tag{3.2}$$

whenever the support of ρ is contained in the support of σ and it is equal to $+\infty$ otherwise. Throughout this work, we will be considering only the range $\alpha \in (1, \infty)$. For such choices, the Hölder conjugate of α is α' such that $\frac{1}{\alpha} + \frac{1}{\alpha'} = 1$, so that $\alpha' = \alpha/(\alpha - 1) \in (1, \infty)$. We can define a sandwiched Rényi mutual information of a bipartite state ρ_{AB} as

$$\widetilde{I}_\alpha(A;B)_\rho \equiv \min_{\sigma_B} \widetilde{D}_\alpha(\rho_{AB} \| \rho_A \otimes \sigma_B). \tag{3.3}$$

Let $|\Gamma\rangle_{AA'}$ denote the "maximally-entangled-like" vector:

$$|\Gamma\rangle_{AA'} \equiv \sum_i |i\rangle_A |i\rangle_{A'}. \tag{3.4}$$

We can then define the sandwiched Rényi mutual information of a channel as

$$\widetilde{I}_\alpha(\mathcal{N}) \equiv \max_{\rho_A} \widetilde{I}_\alpha(A;B)_\omega, \tag{3.5}$$

where

$$\omega_{AB} \equiv \rho_A^{1/2} \mathcal{N}_{A' \to B}(\Gamma_{AA'}) \rho_A^{1/2}. \tag{3.6}$$

In what follows, we will use the following abbreviation for the Choi matrix:

$$\Gamma^{\mathcal{N}}_{AB} \equiv \mathcal{N}_{A' \to B}(\Gamma_{AA'}), \tag{3.7}$$

so that

$$\omega_{AB} = \rho_A^{1/2} \Gamma^{\mathcal{N}}_{AB} \rho_A^{1/2}. \tag{3.8}$$

The quantum relative entropy $D(\rho \| \sigma)$ is defined as

$$D(\rho \| \sigma) \equiv \operatorname{Tr}\{\rho[\log \rho - \log \sigma]\}, \tag{3.9}$$

whenever the support of ρ is contained in the support of σ and it is equal to $+\infty$ otherwise.

4 Bounding the success probability of any entanglement-assisted code with a generalized divergence

We first review the steps in any general (n, R, ε) protocol for entanglement-assisted classical communication over n uses of a quantum channel. Such a protocol begins with a sender Alice and a receiver Bob sharing an arbitrary bipartite entangled state $\Psi_{T_A T_B}$, where Alice possesses the system T_A and Bob the system T_B. Their goal is to use the entangled state $\Psi_{T_A T_B}$ and n instances of a noisy channel $\mathcal{N}_{A' \to B}$ in order for Alice to transmit a message M to Bob. The receiver Bob combines his share T_B of the entanglement and the n output systems of the noisy channel in order to decode the message. This scheme is an (n, R, ε) protocol if the error probability is no larger than $\varepsilon > 0$ and the rate $R = \frac{1}{n} \log_2 |M|$, where $|M|$ denotes the size of the message.

For the purposes of proving a strong converse theorem, we can assume that Alice selects the message M according to a uniform distribution. (The rate at which they can communicate when Alice uses a particular message distribution can only be larger than that for a scheme that should work for all message distributions.) Thus, the protocol begins with Alice preparing a classically-correlated state of the following form:

$$\overline{\Phi}_{MM'} = \frac{1}{|M|} \sum_m |m\rangle\langle m|_M \otimes |m\rangle\langle m|_{M'}, \tag{4.1}$$

and their goal will be for her and Bob to share a state close to this one at the end of the protocol. Alice appends the registers MM' to her share T_A of the entanglement, so that the global state is

$$\overline{\Phi}_{MM'} \otimes \Psi_{T_A T_B}. \tag{4.2}$$

The most general encoding that she can perform is a CPTP map $\mathcal{E}_{M'T_A \to A'^n}$, taking the $M'T_A$ registers to a register A'^n that can be transmitted through n instances of the channel:

$$\begin{aligned}
\rho_{MA'^nT_B} &= \mathcal{E}_{M'T_A \to A'^n}(\overline{\Phi}_{MM'} \otimes \Psi_{T_A T_B}) && (4.3)\\
&= \frac{1}{|M|} \sum_m |m\rangle\langle m|_M \otimes \mathcal{E}_{M'T_A \to A'^n}(|m\rangle\langle m|_{M'} \otimes \Psi_{T_A T_B}) && (4.4)\\
&= \frac{1}{|M|} \sum_m |m\rangle\langle m|_M \otimes \mathcal{E}^m_{T_A \to A'^n}(\Psi_{T_A T_B}). && (4.5)
\end{aligned}$$

In the above, we are exploiting the fact that a single CPTP map $\mathcal{E}_{M'T_A \to A'^n}$ acting on the registers $M'T_A$ can be written as $|M|$ CPTP maps $\{\mathcal{E}^m_{T_A \to A'^n}\}$ acting on the register T_A (since M is a classical register) [47]. Following [32], we define the average code density operator as follows:

$$\rho_{A'^n} \equiv \frac{1}{|M|} \sum_m \mathcal{E}^m_{T_A \to A'^n}(\Psi_{T_A}). \tag{4.6}$$

Alice then transmits the systems A'^n over n uses of the noisy channel $\mathcal{N}_{A' \to B}$, with the overall state becoming

$$\begin{aligned}
\omega_{MB^nT_B} &\equiv \mathcal{N}_{A'^n \to B^n}(\rho_{MA'^nT_B}) && (4.7)\\
&= \frac{1}{|M|} \sum_m |m\rangle\langle m|_M \otimes \mathcal{N}_{A'^n \to B^n}(\mathcal{E}^m_{T_A \to A'^n}(\Psi_{T_A T_B})), && (4.8)
\end{aligned}$$

where $\mathcal{N}_{A'^n \to B^n} \equiv (\mathcal{N}_{A' \to B})^{\otimes n}$. Bob performs a POVM $\{\Lambda^m_{B^nT_B}\}$ on the registers B^nT_B in order to estimate the message m sent by Alice. The overall state after this step is

$$\omega_{M\hat{M}} \equiv \frac{1}{|M|} \sum_{m,m'} |m\rangle\langle m|_M \otimes \mathrm{Tr}\Big\{\Lambda^{m'}_{B^nT_B} \mathcal{N}_{A'^n \to B^n}(\mathcal{E}^m_{T_A \to A'^n}(\Psi_{T_A T_B}))\Big\} \big|m'\big\rangle\big\langle m'\big|_{\hat{M}}. \tag{4.9}$$

Thus, the average success probability for Bob detecting the message correctly is given by

$$\begin{aligned}
p_{\mathrm{succ}} &= \frac{1}{|M|} \sum_m \Pr\Big\{\hat{M} = m | M = m\Big\} && (4.10)\\
&= \frac{1}{|M|} \sum_m \mathrm{Tr}\{\Lambda^m_{B^nT_B} \mathcal{N}_{A'^n \to B^n}(\mathcal{E}^m_{T_A \to A'^n}(\Psi_{T_A T_B}))\} && (4.11)
\end{aligned}$$

For any (n, R, ϵ) protocol, the above success probability is bounded from below by $1 - \varepsilon$.

At this point, we recall the "entanglement-assisted meta-converse" [32, Propositions 20 and 21]:

Proposition 1 ([32, Propositions 20 and 21]) *For any entanglement-assisted code of the above form, with average code density operator $\rho_{A'^n}$, there exists a two-outcome POVM $\{T_{AB^n}, I - T_{AB^n}\}$ such that*

$$p_{\text{succ}} = \text{Tr}\{T_{AB^n}\mathcal{N}_{A'^n\to B^n}(\psi^\rho_{AA'^n})\}, \tag{4.12}$$

$$\frac{1}{|M|} = \text{Tr}\{T_{AB^n}(\psi^\rho_A \otimes \sigma_{B^n})\}, \tag{4.13}$$

where $\psi^\rho_{AA'^n}$ is a purification of the average code density operator $\rho_{A'^n}$ and σ_{B^n} is any density operator.

This proposition allows us to relate the rate and success probability of an entanglement-assisted code to any generalized divergence, extending the framework of [40, 44] to the entanglement-assisted case. Let $I_{\mathbf{D}}(\mathcal{M})$ denote the generalized mutual information of a quantum channel $\mathcal{M}$, constructed from any generalized divergence according to the recipe in (2.4)-(2.9). Let $\delta(p\|q)$ be equal to the generalized divergence $\mathbf{D}(\rho_p\|\rho_q)$ evaluated for the commuting qubit states

$$\rho_p = p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1|, \tag{4.14}$$

$$\rho_q = q|0\rangle\langle 0| + (1-q)|1\rangle\langle 1|. \tag{4.15}$$

(Note that monotonicity of the generalized divergence implies its unitary invariance, which in turn implies that it is independent of the basis when evaluated for classical, commuting states.)

Proposition 2 *The following bound holds for any (n, R, ε) entanglement-assisted code:*

$$I_{\mathbf{D}}(\mathcal{N}^{\otimes n}) \geq \delta(\varepsilon\|1 - 2^{-nR}). \tag{4.16}$$

Proof. Our starting point for a proof is to exploit the test from Proposition 1 and define the classical states $\rho_{p_{\text{succ}}}$ and $\rho_{1/|M|}$, which arise from applying the measurement $\{T_{AB^n}, I - T_{AB^n}\}$ to the following states:

$$\mathcal{N}_{A'^n\to B^n}(\psi^\rho_{AA'^n}), \tag{4.17}$$

$$\psi^\rho_A \otimes \sigma_{B^n}, \tag{4.18}$$

respectively. Without loss of generality, we can assume that $\varepsilon \leq 1-2^{-nR}$ (otherwise, there would be no need to prove the strong converse since the error probability would obey the bound $\varepsilon > 1-2^{-nR}$). Then we have the following inequalities:

$$\begin{aligned} \delta(\varepsilon\|1 - 2^{-nR}) &\leq \delta(1 - p_{\text{succ}}\|1 - 2^{-nR}) &(4.19)\\ &\leq \mathbf{D}(\mathcal{N}_{A'^n\to B^n}(\psi^\rho_{AA'^n})\|\psi^\rho_A \otimes \sigma_{B^n}). &(4.20)\end{aligned}$$

The first inequality follows from the monotonicity $\delta(p'\|q) \leq \delta(p\|q)$ whenever $p \leq p' \leq q$ [40] (recall that we have $1 - p_{\text{succ}} \leq \varepsilon \leq 1 - 2^{-nR}$). The second inequality follows from monotonicity of the generalized divergence under the test $\{T_{AB^n}, I - T_{AB^n}\}$. Since the inequality holds for all states σ_{B^n}, we can find the tightest upper bound on $\delta(\varepsilon\|1 - 2^{-nR})$ for a code with average code density operator $\rho_{A'^n}$ by taking a minimum

$$\delta(\varepsilon\|1 - 2^{-nR}) \leq \min_{\sigma_{B^n}} \mathbf{D}(\mathcal{N}_{A'^n\to B^n}(\psi^\rho_{AA'^n})\|\psi^\rho_A \otimes \sigma_{B^n}). \tag{4.21}$$

Finally, we can remove the dependence of the upper bound on any particular code by maximizing over all code density operators $\rho_{A'^n}$:

$$\delta\big(\varepsilon\|1-2^{-nR}\big) \leq \max_{\rho_{A'^n}} \min_{\sigma_{B^n}} \mathbf{D}\big(\mathcal{N}_{A'^n\to B^n}(\psi^{\rho}_{AA'^n})\big\|\psi^{\rho}_{A}\otimes\sigma_{B^n}\big). \tag{4.22}$$

This is then equivalent to the inequality in the statement of the proposition. ■

For our purposes here, we can evaluate the bound from Proposition 2 by setting the divergence to be the sandwiched Rényi relative entropy (however, note that there is no need for the assumption $\varepsilon \leq 1-2^{-nR}$ when employing the sandwiched Rényi relative entropy). Following steps identical to those in [49, Section 6], we arrive at the following bound on the success probability of any entanglement-assisted code:

$$p_{\text{succ}} \leq 2^{-n\left(\frac{\alpha-1}{\alpha}\right)\left(R-\frac{1}{n}\widetilde{I}_{\alpha}(\mathcal{N}^{\otimes n})\right)}. \tag{4.23}$$

We stress that this bound holds for all $\alpha > 1$ and $n \geq 1$. In order to arrive at the strong converse, we should now prove that the sandwiched Rényi mutual information is additive as a function of quantum channels for $\alpha \in (1,\infty)$.

5 Additivity of the sandwiched Rényi mutual information of a quantum channel

In this section, we show that the sandwiched Rényi mutual information $\widetilde{I}_{\alpha}(\mathcal{N}^{\otimes n})$ is additive as a function of quantum channels for $\alpha \in (1,\infty)$. Before doing so, we require a few supplementary lemmas.

Lemma 3 *The following equality holds for $\alpha \in (1,\infty)$*

$$\widetilde{D}_{\alpha}\Big(\rho_A^{1/2}\Gamma^{\mathcal{N}}_{AB}\rho_A^{1/2}\Big\|\rho_A\otimes\sigma_B\Big) = \frac{\alpha}{\alpha-1}\log\left\|\left(\left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2}\left(\rho_A^{\frac{1}{\alpha}}\otimes\sigma_B^{\frac{1-\alpha}{\alpha}}\right)\left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2}\right)\right\|_{\alpha} \tag{5.1}$$

Proof. This follows from

$$\begin{aligned}
&\widetilde{D}_{\alpha}\Big(\rho_A^{1/2}\Gamma^{\mathcal{N}}_{AB}\rho_A^{1/2}\Big\|\rho_A\otimes\sigma_B\Big) \\
&= \frac{1}{\alpha-1}\log\operatorname{Tr}\left\{\left((\rho_A\otimes\sigma_B)^{\frac{1-\alpha}{2\alpha}}\rho_A^{1/2}\Gamma^{\mathcal{N}}_{AB}\rho_A^{1/2}(\rho_A\otimes\sigma_B)^{\frac{1-\alpha}{2\alpha}}\right)^{\alpha}\right\} && (5.2)\\
&= \frac{1}{\alpha-1}\log\operatorname{Tr}\left\{\left(\left(\rho_A^{\frac{1-\alpha}{2\alpha}}\otimes\sigma_B^{\frac{1-\alpha}{2\alpha}}\right)\rho_A^{1/2}\Gamma^{\mathcal{N}}_{AB}\rho_A^{1/2}\left(\rho_A^{\frac{1-\alpha}{2\alpha}}\otimes\sigma_B^{\frac{1-\alpha}{2\alpha}}\right)\right)^{\alpha}\right\} && (5.3)\\
&= \frac{1}{\alpha-1}\log\operatorname{Tr}\left\{\left(\left(\rho_A^{\frac{1-\alpha}{2\alpha}}\rho_A^{1/2}\otimes\sigma_B^{\frac{1-\alpha}{2\alpha}}\right)\Gamma^{\mathcal{N}}_{AB}\left(\rho_A^{1/2}\rho_A^{\frac{1-\alpha}{2\alpha}}\otimes\sigma_B^{\frac{1-\alpha}{2\alpha}}\right)\right)^{\alpha}\right\} && (5.4)\\
&= \frac{1}{\alpha-1}\log\operatorname{Tr}\left\{\left(\left(\rho_A^{\frac{1}{2\alpha}}\otimes\sigma_B^{\frac{1-\alpha}{2\alpha}}\right)\Gamma^{\mathcal{N}}_{AB}\left(\rho_A^{\frac{1}{2\alpha}}\otimes\sigma_B^{\frac{1-\alpha}{2\alpha}}\right)\right)^{\alpha}\right\} && (5.5)\\
&= \frac{1}{\alpha-1}\log\operatorname{Tr}\left\{\left(\left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2}\left(\rho_A^{\frac{1}{\alpha}}\otimes\sigma_B^{\frac{1-\alpha}{\alpha}}\right)\left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2}\right)^{\alpha}\right\} && (5.6)\\
&= \frac{\alpha}{\alpha-1}\log\left\|\left(\left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2}\left(\rho_A^{\frac{1}{\alpha}}\otimes\sigma_B^{\frac{1-\alpha}{\alpha}}\right)\left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2}\right)\right\|_{\alpha} && (5.7)
\end{aligned}$$

■

Lemma 4 *The following function is concave in $\rho_A \in \mathcal{S}(\mathcal{H}_A)$ and convex in $\sigma_B \in \mathcal{S}(\mathcal{H}_B)$ for $\alpha \in (1,\infty)$:*

$$\left\| \left(\left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2} \left(\rho_A^{\frac{1}{\alpha}} \otimes \sigma_B^{\frac{1-\alpha}{\alpha}} \right) \left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2} \right) \right\|_{\alpha}. \tag{5.8}$$

Proof. Convexity in σ_B follows immediately from operator convexity of $x^{(1-\alpha)/\alpha}$ for $\alpha \in (1,\infty)$ and convexity of the a-norm. The other statement follows in a few steps. We first reexpress the α-norm as the following optimization:

$$\left\| \left(\left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2} \left(\rho_A^{\frac{1}{\alpha}} \otimes \sigma_B^{\frac{1-\alpha}{\alpha}} \right) \left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2} \right) \right\|_{\alpha} = \max_{\|X\|_{\frac{\alpha}{\alpha-1}} \leq 1} \operatorname{Tr}\left\{ X \left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2} \left(\rho_A^{\frac{1}{\alpha}} \otimes \sigma_B^{\frac{1-\alpha}{\alpha}} \right) \left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2} \right\}. \tag{5.9}$$

Due to the operator $\left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2} \left(\rho_A^{\frac{1}{\alpha}} \otimes \sigma_B^{\frac{1-\alpha}{\alpha}} \right) \left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2}$ being positive, it suffices to restrict the optimization to be over positive X operators such that $\|X\|_{\frac{\alpha}{\alpha-1}} \leq 1$. However, this restriction is equivalent to

$$\operatorname{Tr}\left\{ X^{\frac{\alpha}{\alpha-1}} \right\} \leq 1, \ \operatorname{Tr}\{X\} \geq 0. \tag{5.10}$$

Thus, with the substitution $Y = X^{\alpha/(\alpha-1)}$, we can rewrite the above as

$$\max_{Y \geq 0, \operatorname{Tr}\{Y\} \leq 1} \operatorname{Tr}\left\{ Y^{\frac{\alpha-1}{\alpha}} \left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2} \left(\rho_A^{\frac{1}{\alpha}} \otimes \sigma_B^{\frac{1-\alpha}{\alpha}} \right) \left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2} \right\}. \tag{5.11}$$

To prove concavity in ρ, we require Lieb's concavity theorem [30], a special case of which is the statement that the following function

$$\operatorname{Tr}\left\{ X R^{1-t} X^{\dagger} S^{t} \right\}, \tag{5.12}$$

for $R, S \geq 0$ and $t \in [0,1]$, is jointly concave in R and S. Indeed, for $i \in \{0,1\}$, consider any Y_i such that $Y_i \geq 0$, $\operatorname{Tr}\{Y_i\} \leq 1$, density operators ρ_i, and $\lambda \in [0,1]$. We begin with

$$\begin{aligned}
&\lambda \operatorname{Tr}\left\{ Y_0^{\frac{\alpha-1}{\alpha}} \left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2} \left(\rho_0^{\frac{1}{\alpha}} \otimes \sigma_B^{\frac{1-\alpha}{\alpha}} \right) \left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2} \right\} \\
&\qquad + (1-\lambda) \operatorname{Tr}\left\{ Y_1^{\frac{\alpha-1}{\alpha}} \left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2} \left(\rho_1^{\frac{1}{\alpha}} \otimes \sigma_B^{\frac{1-\alpha}{\alpha}} \right) \left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2} \right\} \\
&\leq \operatorname{Tr}\left\{ (\lambda Y_0 + (1-\lambda) Y_1)^{\frac{\alpha-1}{\alpha}} \left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2} \left((\lambda \rho_0 + (1-\lambda)\rho_1)^{\frac{1}{\alpha}} \otimes \sigma_B^{\frac{1-\alpha}{\alpha}} \right) \left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2} \right\} &(5.13) \\
&\leq \max_{Y \geq 0, \operatorname{Tr}\{Y\} \leq 1} \operatorname{Tr}\left\{ Y^{\frac{\alpha-1}{\alpha}} \left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2} \left((\lambda \rho_0 + (1-\lambda)\rho_1)^{\frac{1}{\alpha}} \otimes \sigma_B^{\frac{1-\alpha}{\alpha}} \right) \left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2} \right\} &(5.14) \\
&= \left\| \left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2} \left((\lambda \rho_0 + (1-\lambda)\rho_1)^{\frac{1}{\alpha}} \otimes \sigma_B^{\frac{1-\alpha}{\alpha}} \right) \left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2} \right\|_{\alpha} &(5.15)
\end{aligned}$$

Since the calculation is independent of which Y_0 and Y_1 we started with, concavity of (5.8) in ρ follows. ■

Lemma 5 *The following equality holds for $\alpha \in (1,\infty)$*

$$\widetilde{I}_{\alpha}(\mathcal{N}) = \min_{\sigma_B} \max_{\rho_A} \widetilde{D}_{\alpha}\left(\rho_A^{1/2} \Gamma^{\mathcal{N}}_{AB} \rho_A^{1/2} \middle\| \rho_A \otimes \sigma_B \right). \tag{5.16}$$

Proof. Consider that

$$\widetilde{I}_\alpha(\mathcal{N}) = \max_{\rho_A}\min_{\sigma_B} \widetilde{D}_\alpha\left(\rho_A^{1/2}\Gamma^{\mathcal{N}}_{AB}\rho_A^{1/2}\middle\|\rho_A\otimes\sigma_B\right) \tag{5.17}$$

$$= \max_{\rho_A}\min_{\sigma_B} \frac{\alpha}{\alpha-1}\log\left\|\left(\left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2}\left(\rho_A^{\frac{1}{\alpha}}\otimes\sigma_B^{\frac{1-\alpha}{\alpha}}\right)\left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2}\right)\right\|_\alpha \tag{5.18}$$

$$= \frac{\alpha}{\alpha-1}\log\max_{\rho_A}\min_{\sigma_B}\left\|\left(\left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2}\left(\rho_A^{\frac{1}{\alpha}}\otimes\sigma_B^{\frac{1-\alpha}{\alpha}}\right)\left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2}\right)\right\|_\alpha \tag{5.19}$$

$$= \frac{\alpha}{\alpha-1}\log\min_{\sigma_B}\max_{\rho_A}\left\|\left(\left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2}\left(\rho_A^{\frac{1}{\alpha}}\otimes\sigma_B^{\frac{1-\alpha}{\alpha}}\right)\left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2}\right)\right\|_\alpha \tag{5.20}$$

$$= \min_{\sigma_B}\max_{\rho_A} \frac{\alpha}{\alpha-1}\log\left\|\left(\left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2}\left(\rho_A^{\frac{1}{\alpha}}\otimes\sigma_B^{\frac{1-\alpha}{\alpha}}\right)\left[\Gamma^{\mathcal{N}}_{AB}\right]^{1/2}\right)\right\|_\alpha \tag{5.21}$$

$$= \min_{\sigma_B}\max_{\rho_A} \widetilde{D}_\alpha\left(\rho_A^{1/2}\Gamma^{\mathcal{N}}_{AB}\rho_A^{1/2}\middle\|\rho_A\otimes\sigma_B\right), \tag{5.22}$$

where we applied the Sion minimax theorem [45]. ∎

We now prove additivity by exploiting the above lemmas and some results in [5, 17]:

Lemma 6 *The sandwiched Rényi mutual information is additive as a function of channels for* $\alpha\in(1,\infty)$, *in the sense that*

$$\widetilde{I}_\alpha(\mathcal{N}_1\otimes\mathcal{N}_2) = \widetilde{I}_\alpha(\mathcal{N}_1) + \widetilde{I}_\alpha(\mathcal{N}_2). \tag{5.23}$$

Proof. The inequality below is straightforward

$$\widetilde{I}_\alpha(\mathcal{N}_1\otimes\mathcal{N}_2) \geq \widetilde{I}_\alpha(\mathcal{N}_1) + \widetilde{I}_\alpha(\mathcal{N}_2), \tag{5.24}$$

following from [5, Theorem 11] and the fact that we can choose tensor-product states as a special case of the optimization on the left hand side.

We now prove the other inequality:

$$\widetilde{I}_\alpha(\mathcal{N}_1\otimes\mathcal{N}_2) \leq \widetilde{I}_\alpha(\mathcal{N}_1) + \widetilde{I}_\alpha(\mathcal{N}_2). \tag{5.25}$$

From the above lemmas, we can reexpress $\widetilde{I}_\alpha(\mathcal{N})$ as

$$\min_{\sigma_B}\max_{\rho_A}\frac{\alpha}{\alpha-1}\log\left\|\sigma_B^{\frac{1-\alpha}{2\alpha}}\mathcal{N}_{A'\to B}\left(\rho_A^{\frac{1}{2\alpha}}\Gamma_{AA'}\rho_A^{\frac{1}{2\alpha}}\right)\sigma_B^{\frac{1-\alpha}{2\alpha}}\right\|_\alpha. \tag{5.26}$$

Defining the CP map $\Theta_\sigma(X) = \sigma^{1/2}X\sigma^{1/2}$, we can write the above as

$$\min_{\sigma_B}\frac{\alpha}{\alpha-1}\log\max_{\rho_A}\left\|\left(\Theta_{\sigma^{\frac{1-\alpha}{\alpha}}}\circ\mathcal{N}_{A'\to B}\right)\left(\rho_A^{\frac{1}{2\alpha}}\Gamma_{AA'}\rho_A^{\frac{1}{2\alpha}}\right)\right\|_\alpha$$
$$= \min_{\sigma_B}\frac{\alpha}{\alpha-1}\log\left\|\Theta_{\sigma^{\frac{1-\alpha}{\alpha}}}\circ\mathcal{N}_{A'\to B}\right\|_{\mathrm{CB},1\to\alpha}. \tag{5.27}$$

The equality follows from [17, Theorem 10], in which these authors showed that the $(\mathrm{CB}, 1\to\alpha)$ norm of a CP map $\mathcal{M}_{A'\to B}$ is equal to

$$\|\mathcal{M}_{A'\to B}\|_{\mathrm{CB},1\to\alpha} \equiv \sup_{X>0}\frac{\|(X\otimes I)\mathcal{M}_{A'\to B}(\Gamma_{AA'})(X\otimes I)\|_\alpha}{\|X^2\|_\alpha} \tag{5.28}$$

$$= \sup_{Y>0,\mathrm{Tr}\{Y\}\leq 1}\left\|\left(Y^{\frac{1}{2\alpha}}\otimes I\right)\mathcal{M}_{A'\to B}(\Gamma_{AA'})\left(Y^{\frac{1}{2\alpha}}\otimes I\right)\right\|_\alpha \tag{5.29}$$

With this result in hand, defining

$$\Gamma^{\mathcal{N}_1 \otimes \mathcal{N}_1}_{A_1 A_2 B_1 B_2} \equiv (\mathcal{N}_1 \otimes \mathcal{N}_2)\Big(\Gamma_{A_1 A_2 A_1' A_2'}\Big), \tag{5.30}$$

we can now prove the other inequality:

$$\widetilde{I}_\alpha(\mathcal{N}_1 \otimes \mathcal{N}_2) = \max_{\rho_{A_1 A_2}} \min_{\sigma_{B_1 B_2}} \widetilde{D}_\alpha\Big(\rho^{1/2}_{A_1 A_2} \Gamma^{\mathcal{N}_1 \otimes \mathcal{N}_1}_{A_1 A_2 B_1 B_2} \rho^{1/2}_{A_1 A_2} \Big\| \rho_{A_1 A_2} \otimes \sigma_{B_1 B_2}\Big) \tag{5.31}$$

$$= \min_{\sigma_{B_1 B_2}} \max_{\rho_{A_1 A_2}} \widetilde{D}_\alpha\Big(\rho^{1/2}_{A_1 A_2} \Gamma^{\mathcal{N}_1 \otimes \mathcal{N}_1}_{A_1 A_2 B_1 B_2} \rho^{1/2}_{A_1 A_2} \Big\| \rho_{A_1 A_2} \otimes \sigma_{B_1 B_2}\Big) \tag{5.32}$$

$$\leq \min_{\sigma_{B_1} \otimes \sigma_{B_2}} \max_{\rho_{A_1 A_2}} \widetilde{D}_\alpha\Big(\rho^{1/2}_{A_1 A_2} \Gamma^{\mathcal{N}_1 \otimes \mathcal{N}_1}_{A_1 A_2 B_1 B_2} \rho^{1/2}_{A_1 A_2} \Big\| \rho_{A_1 A_2} \otimes \sigma_{B_1} \otimes \sigma_{B_2}\Big) \tag{5.33}$$

$$= \frac{\alpha}{\alpha - 1} \log \min_{\sigma_{B_1} \otimes \sigma_{B_2}} \left\| \left(\Theta_{\sigma_{B_1}^{\frac{1-\alpha}{\alpha}}} \circ \mathcal{N}_1\right) \otimes \left(\Theta_{\sigma_{B_2}^{\frac{1-\alpha}{\alpha}}} \circ \mathcal{N}_2\right) \right\|_{\mathrm{CB},1\to\alpha} \tag{5.34}$$

Using [17, Theorem 11] (the main result there), we find that the above is equal to

$$= \frac{\alpha}{\alpha - 1} \log \min_{\sigma_{B_1} \otimes \sigma_{B_2}} \left\| \left(\Theta_{\sigma_{B_1}^{\frac{1-\alpha}{\alpha}}} \circ \mathcal{N}_1\right) \right\|_{\mathrm{CB},1\to\alpha} \left\| \left(\Theta_{\sigma_{B_2}^{\frac{1-\alpha}{\alpha}}} \circ \mathcal{N}_2\right) \right\|_{\mathrm{CB},1\to\alpha} \tag{5.35}$$

$$= \frac{\alpha}{\alpha - 1} \log \min_{\sigma_{B_1}} \left\| \left(\Theta_{\sigma_{B_1}^{\frac{1-\alpha}{\alpha}}} \circ \mathcal{N}_1\right) \right\|_{\mathrm{CB},1\to\alpha} \min_{\sigma_{B_2}} \left\| \left(\Theta_{\sigma_{B_2}^{\frac{1-\alpha}{\alpha}}} \circ \mathcal{N}_2\right) \right\|_{\mathrm{CB},1\to\alpha} \tag{5.36}$$

$$= \frac{\alpha}{\alpha - 1} \log \min_{\sigma_{B_1}} \left\| \left(\Theta_{\sigma_{B_1}^{\frac{1-\alpha}{\alpha}}} \circ \mathcal{N}_1\right) \right\|_{\mathrm{CB},1\to\alpha} + \frac{\alpha}{\alpha - 1} \log \min_{\sigma_{B_2}} \left\| \left(\Theta_{\sigma_{B_2}^{\frac{1-\alpha}{\alpha}}} \circ \mathcal{N}_2\right) \right\|_{\mathrm{CB},1\to\alpha} \tag{5.37}$$

$$= \widetilde{I}_\alpha(\mathcal{N}_1) + \widetilde{I}_\alpha(\mathcal{N}_2). \tag{5.38}$$

■

6 Final steps for the strong converse

In this section, we outline the remaining steps to prove the strong converse theorem. Returning to (4.23), the additivity relation from Lemma 6 (along with an inductive argument) allows us to conclude the following upper bound on the success probability when using an entanglement-assisted code to communicate over a quantum channel $\mathcal{N}$:

$$p_{\text{succ}} \leq 2^{-n \sup_{\alpha>1}\left(\frac{\alpha-1}{\alpha}\right)\left(R - \widetilde{I}_\alpha(\mathcal{N})\right)}. \tag{6.1}$$

The quantity $\sup_{\alpha>1}\left(\frac{\alpha-1}{\alpha}\right)\Big(R - \widetilde{I}_\alpha(\mathcal{N})\Big)$ is thus our bound on the strong converse exponent, which holds for all $n \geq 1$.

Recall the definition of the (traditional) quantum Rényi relative entropy for $\alpha \in (1, \infty)$:

$$D_\alpha(\rho \| \sigma) \equiv \frac{1}{\alpha - 1} \log \mathrm{Tr}\big\{\rho^\alpha \sigma^{1-\alpha}\big\}, \tag{6.2}$$

whenever the support of ρ is in the support of σ and it is equal to $+\infty$ otherwise. We define the Rényi quantum mutual information of a bipartite state ρ_{AB} as follows:

$$I_\alpha(A;B)_\rho = \min_{\sigma_B} D_\alpha(\rho_{AB} \| \rho_A \otimes \sigma_B). \tag{6.3}$$

By applying the following Lieb-Thirring trace inequality [31, 13], which holds for $B \geq 0$, any operator C, and for $\alpha \geq 1$:

$$\mathrm{Tr}\{(CBC^\dagger)^\alpha\} \leq \mathrm{Tr}\{(C^\dagger C)^\alpha B^\alpha\}, \tag{6.4}$$

we find that the following inequality holds for $\alpha > 1$ [49]:

$$\widetilde{D}_\alpha(\rho\|\sigma) \leq D_\alpha(\rho\|\sigma). \tag{6.5}$$

This in turn implies the following upper bound on the success probability of any entanglement-assisted code for all $\alpha \in (1, \infty)$ by combining (6.5) with (6.1):

$$p_{\mathrm{succ}} \leq 2^{-n\left(\frac{\alpha-1}{\alpha}\right)(R - I_\alpha(\mathcal{N}))}, \tag{6.6}$$

where we define $I_\alpha(\mathcal{N})$ according to the recipe in (2.4)-(2.9) with the divergence set to (6.2).

We next prove the following "quantum Sibson identity," which will be helpful in obtaining an explicit form for the Rényi quantum mutual information (see [44] for a variant which is relevant for Rényi coherent information):

Lemma 7 (Quantum Sibson identity) *The following quantum Sibson identity holds for $\alpha \in (1, \infty)$*

$$D_\alpha(\rho_{AB} \| \rho_A \otimes \sigma_B) = D_\alpha(\sigma_B^* \| \sigma_B) + D_\alpha(\rho_{AB} \| \rho_A \otimes \sigma_B^*), \tag{6.7}$$

where σ_B^* is defined as

$$\sigma_B^* \equiv \frac{\left(\mathrm{Tr}_A\{\rho_A^{1-\alpha}\rho_{AB}^\alpha\}\right)^{1/\alpha}}{\mathrm{Tr}\left\{\left(\mathrm{Tr}_A\{\rho_A^{1-\alpha}\rho_{AB}^\alpha\}\right)^{1/\alpha}\right\}}. \tag{6.8}$$

Proof. It is clear that σ_B^* is a positive operator because $\rho_A^{\frac{1-\alpha}{2}} \rho_{AB}^\alpha \rho_A^{\frac{1-\alpha}{2}}$ is positive, and the partial trace maintains positivity while being equal to $\mathrm{Tr}_A\{\rho_A^{1-\alpha}\rho_{AB}^\alpha\}$ from cyclicity. The above relation then implies that

$$\left(\sigma_B^* \mathrm{Tr}\left\{\left(\mathrm{Tr}_A\{\rho_A^{1-\alpha}\rho_{AB}^\alpha\}\right)^{1/\alpha}\right\}\right)^\alpha = \mathrm{Tr}_A\{\rho_A^{1-\alpha}\rho_{AB}^\alpha\}. \tag{6.9}$$

We can then expand $D_\alpha(\rho_{AB} \| \rho_A \otimes \sigma_B)$ as follows:

$$\begin{aligned}
D_\alpha(\rho_{AB} \| \rho_A \otimes \sigma_B) &= \frac{1}{\alpha - 1} \log \mathrm{Tr}\{\rho_{AB}^\alpha(\rho_A^{1-\alpha} \otimes \sigma_B^{1-\alpha})\} && (6.10)\\
&= \frac{1}{\alpha - 1} \log \mathrm{Tr}\{\mathrm{Tr}_A\{\rho_A^{1-\alpha}\rho_{AB}^\alpha\}\ \sigma_B^{1-\alpha}\} && (6.11)\\
&= \frac{1}{\alpha - 1} \log \mathrm{Tr}\left\{\left(\sigma_B^*\ \mathrm{Tr}\left\{\left(\mathrm{Tr}_A\{\rho_A^{1-\alpha}\rho_{AB}^\alpha\}\right)^{1/\alpha}\right\}\right)^\alpha \sigma_B^{1-\alpha}\right\} && (6.12)\\
&= \frac{1}{\alpha - 1}\left[\log \mathrm{Tr}\{(\sigma_B^*)^\alpha\ \sigma_B^{1-\alpha}\} + \alpha \log \mathrm{Tr}\left\{\left(\mathrm{Tr}_A\{\rho_A^{1-\alpha}\rho_{AB}^\alpha\}\right)^{1/\alpha}\right\}\right] && (6.13)\\
&= D_\alpha(\sigma_B^* \| \sigma_B) + \frac{\alpha}{\alpha - 1} \log \mathrm{Tr}\left\{\left(\mathrm{Tr}_A\{\rho_A^{1-\alpha}\rho_{AB}^\alpha\}\right)^{1/\alpha}\right\}. && (6.14)
\end{aligned}$$

Now consider expanding $D_\alpha(\rho_{AB} \| \rho_A \otimes \sigma_B^*)$:

$$
\begin{aligned}
& D_\alpha(\rho_{AB} \| \rho_A \otimes \sigma_B^*) \\
& = \frac{1}{\alpha - 1} \log \mathrm{Tr}\Big\{ \mathrm{Tr}_A\{\rho_A^{1-\alpha}\rho_{AB}^\alpha\} \; (\sigma_B^*)^{1-\alpha} \Big\} && (6.15) \\
& = \frac{1}{\alpha - 1} \log \mathrm{Tr}\Big\{ \mathrm{Tr}_A\{\rho_A^{1-\alpha}\rho_{AB}^\alpha\} \; \Big(\big(\mathrm{Tr}_A\{\rho_A^{1-\alpha}\rho_{AB}^\alpha\}\big)^{1/\alpha} \Big)^{1-\alpha} \Big\} + \log\Big(\mathrm{Tr}\Big\{ \big(\mathrm{Tr}_A\{\rho_A^{1-\alpha}\rho_{AB}^\alpha\}\big)^{1/\alpha} \Big\} \Big) && (6.16) \\
& = \frac{1}{\alpha - 1} \log \mathrm{Tr}\Big\{ \mathrm{Tr}_A\{\rho_A^{1-\alpha}\rho_{AB}^\alpha\} \; \big(\mathrm{Tr}_A\{\rho_A^{1-\alpha}\rho_{AB}^\alpha\}\big)^{\frac{1-\alpha}{\alpha}} \Big\} + \log\Big(\mathrm{Tr}\Big\{ \big(\mathrm{Tr}_A\{\rho_A^{1-\alpha}\rho_{AB}^\alpha\}\big)^{1/\alpha} \Big\} \Big) && (6.17) \\
& = \frac{1}{\alpha - 1} \log \mathrm{Tr}\Big\{ \big(\mathrm{Tr}_A\{\rho_A^{1-\alpha}\rho_{AB}^\alpha\}\big)^{\frac{1}{\alpha}} \Big\} + \log\Big(\mathrm{Tr}\Big\{ \big(\mathrm{Tr}_A\{\rho_A^{1-\alpha}\rho_{AB}^\alpha\}\big)^{1/\alpha} \Big\} \Big) && (6.18) \\
& = \frac{\alpha}{\alpha - 1} \log \mathrm{Tr}\Big\{ \big(\mathrm{Tr}_A\{\rho_A^{1-\alpha}\rho_{AB}^\alpha\}\big)^{\frac{1}{\alpha}} \Big\}. && (6.19)
\end{aligned}
$$

Combining (6.14) and (6.19) gives (6.7). ■

By exploiting Lemma 7, we obtain an explicit form for it:

Corollary 8 *The Rényi quantum mutual information has an explicit form for $\alpha \in (1, \infty)$ given by*

$$
I_\alpha(A;B)_\rho = \frac{\alpha}{\alpha - 1} \log \mathrm{Tr}\Big\{ \big(\mathrm{Tr}_A\{\rho_A^{1-\alpha}\rho_{AB}^\alpha\}\big)^{\frac{1}{\alpha}} \Big\} \tag{6.20}
$$

Proof. From the identity in Lemma 7, we can conclude that

$$
\begin{aligned}
I_\alpha(A;B)_\rho &= \min_{\sigma_B} D_\alpha(\rho_{AB} \| \rho_A \otimes \sigma_B) && (6.21) \\
&= \min_{\sigma_B} [D_\alpha(\sigma_B^* \| \sigma_B) + D_\alpha(\rho_{AB} \| \rho_A \otimes \sigma_B^*)] && (6.22) \\
&= D_\alpha(\rho_{AB} \| \rho_A \otimes \sigma_B^*) && (6.23) \\
&= \frac{\alpha}{\alpha - 1} \log \mathrm{Tr}\Big\{ \big(\mathrm{Tr}_A\{\rho_A^{1-\alpha}\rho_{AB}^\alpha\}\big)^{\frac{1}{\alpha}} \Big\}. && (6.24)
\end{aligned}
$$

■

The following lemma will be helpful for us in relating Rényi quantum mutual information to the von Neumann quantum mutual information:

Lemma 9 *The following identity holds for a bipartite state ρ_{AB}:*

$$
\lim_{\alpha \searrow 1} \left[\frac{\partial}{\partial \alpha} \log \mathrm{Tr}\Big\{ \big(\mathrm{Tr}_A\{\rho_A^{1-\alpha}\rho_{AB}^\alpha\}\big)^{\frac{1}{\alpha}} \Big\} \right] = I(A;B)_\rho \, . \tag{6.25}
$$

Proof. A proof follows by exploiting some ideas from [14] and [39]. It suffices to show that

$$
\lim_{\alpha \searrow 1} \left[\frac{\partial}{\partial \alpha} \log \mathrm{Tr}\Big\{ \big(\mathrm{Tr}_A\{\rho_A^{1-\alpha}\rho_{AB}^\alpha\}\big)^{\frac{1}{\alpha}} \Big\} \right] = -\mathrm{Tr}\{\rho_A \log \rho_A\} - \mathrm{Tr}\{\rho_B \log \rho_B\} + \mathrm{Tr}\{\rho_{AB} \log \rho_{AB}\}. \tag{6.26}
$$

(In this proof, we will take log to denote the natural logarithm, but note that the result follows simply by replacing the natural logarithm in both definitions with the binary logarithm.)

Let us rewrite the expression inside the trace, using $\alpha = 1 + \beta$ where $\beta > 0$, as

$$\operatorname{Tr}\left\{ \left(\operatorname{Tr}_A \left\{ \rho_A^{-\beta} \rho_{AB}^{1+\beta} \right\} \right)^{\frac{1}{1+\beta}} \right\}. \tag{6.27}$$

Furthermore, we can introduce two parameters $\beta_1 > 0$ and $\beta_2 > 0$, so that the above expression is a special case of

$$f(\beta_1, \beta_2) \equiv \operatorname{Tr}\left\{ \left(\operatorname{Tr}_A \left\{ \rho_A^{-\beta_1} \rho_{AB}^{1+\beta_1} \right\} \right)^{\frac{1}{1+\beta_2}} \right\}. \tag{6.28}$$

We then have that

$$\lim_{\alpha \searrow 1} \left[\frac{\partial}{\partial \alpha} \log \operatorname{Tr}\left\{ \left(\operatorname{Tr}_A \{ \rho_A^{1-\alpha} \rho_{AB}^{\alpha} \} \right)^{\frac{1}{\alpha}} \right\} \right] = \frac{\lim_{\beta \searrow 0} \left[\frac{\partial}{\partial \beta} f(\beta, \beta) \right]}{f(0,0)} \tag{6.29}$$

$$= \lim_{\beta_1 \searrow 0} \left[\frac{\partial}{\partial \beta_1} f(\beta_1, 0) \right] + \lim_{\beta_2 \searrow 0} \left[\frac{\partial}{\partial \beta_2} f(0, \beta_2) \right], \tag{6.30}$$

where the second equality follows in part because $f(0,0) = 1$. Now consider the following Taylor expansions around $\beta = 0$:

$$X^{-\beta} = I - \beta \log X + O\big(\beta^2\big), \tag{6.31}$$

$$X^{1+\beta} = X + \beta X \log X + O\big(\beta^2\big), \tag{6.32}$$

$$X^{\frac{1}{1+\beta}} = X - \beta X \log X + O\big(\beta^2\big). \tag{6.33}$$

From these, we calculate $f(\beta_1, 0)$ as

$$f(\beta_1, 0) = \operatorname{Tr}\left\{ \rho_A^{-\beta_1} \rho_{AB}^{1+\beta_1} \right\} \tag{6.34}$$

$$= \operatorname{Tr}\{ \rho_{AB} - \beta_1 \rho_{AB} \log \rho_A + \beta_1 \rho_{AB} \log \rho_{AB} \} + O\big(\beta_1^2\big) \tag{6.35}$$

$$= \operatorname{Tr}\{\rho_{AB}\} - \beta_1 \operatorname{Tr}\{\rho_{AB} \log \rho_A\} + \beta_1 \operatorname{Tr}\{\rho_{AB} \log \rho_{AB}\} + O\big(\beta_1^2\big). \tag{6.36}$$

It then follows that

$$\lim_{\beta_1 \searrow 0} \left[\frac{\partial}{\partial \beta_1} f(\beta_1, 0) \right] = -\operatorname{Tr}\{\rho_A \log \rho_A\} + \operatorname{Tr}\{\rho_{AB} \log \rho_{AB}\}. \tag{6.37}$$

We then calculate $f(0, \beta_2)$ as

$$f(0, \beta_2) = \operatorname{Tr}\left\{ (\operatorname{Tr}_A\{\rho_{AB}\})^{\frac{1}{1+\beta_2}} \right\} \tag{6.38}$$

$$= \operatorname{Tr}\left\{ (\rho_B)^{\frac{1}{1+\beta_2}} \right\} \tag{6.39}$$

$$= \operatorname{Tr}\{\rho_B\} - \beta_2 \operatorname{Tr}\{\rho_B \log \rho_B\} + O\big(\beta_2^2\big). \tag{6.40}$$

It then follows that

$$\lim_{\beta_2 \searrow 0} \left[\frac{\partial}{\partial \beta_2} f(0, \beta_2) \right] = -\operatorname{Tr}\{\rho_B \log \rho_B\}. \tag{6.41}$$

Putting these together, we find that

$$\lim_{\beta \searrow 0} \left[\frac{\partial}{\partial \beta} f(\beta, \beta) \right] = -\operatorname{Tr}\{\rho_A \log \rho_A\} - \operatorname{Tr}\{\rho_B \log \rho_B\} + \operatorname{Tr}\{\rho_{AB} \log \rho_{AB}\} = I(A;B)_\rho. \tag{6.42}$$

■

Let $I(\mathcal{N})$ denote the entanglement-assisted capacity of a quantum channel $\mathcal{N}$, which [8] proved is a function of $\mathcal{N}$ and constructed according to the recipe in (2.4)-(2.9) with the generalized divergence taken as the quantum relative entropy $D(\rho\|\sigma)$.

Lemma 10 *If $R > I(\mathcal{N})$ then*

$$\exists\, \beta > 1, \forall\alpha \in (1,\beta), \qquad \left(\frac{\alpha-1}{\alpha}\right)(R - I_\alpha(\mathcal{N})) > 0. \tag{6.43}$$

Proof. The argument here is very similar to the proof of [39, Lemma 3] and that of [44, Lemma 8]. We include it here for completeness. Let

$$g(\alpha,\rho) \equiv \left(\frac{\alpha-1}{\alpha}\right)(R - I_\alpha(A;B)_\omega), \tag{6.44}$$

where $\omega_{AB} \equiv \rho_A^{1/2}\Gamma_{AB}^{\mathcal{N}}\rho_A^{1/2}$. We can expand $g(\alpha,\rho)$ using Lemma 8 as

$$\begin{aligned} g(\alpha,\rho) &= \left(\frac{\alpha-1}{\alpha}\right)\left(R - \frac{\alpha}{\alpha-1}\log\mathrm{Tr}\left\{\left(\mathrm{Tr}_A\{\rho_A^{1-\alpha}\rho_{AB}^{\alpha}\}\right)^{\frac{1}{\alpha}}\right\}\right) &(6.45)\\ &= \left(1-\frac{1}{\alpha}\right)R - \log\mathrm{Tr}\left\{\left(\mathrm{Tr}_A\{\rho_A^{1-\alpha}\rho_{AB}^{\alpha}\}\right)^{\frac{1}{\alpha}}\right\}. &(6.46)\end{aligned}$$

Suppose now that $R > I(\mathcal{N})$, as in the statement of the lemma. Then for $\forall\rho$ we have that $g(1,\rho) = 0$ and by Lemma 9,

$$g'(1,\rho) = R - I(A;B)_\omega > 0. \tag{6.47}$$

Now, suppose for a contradiction that (6.43) does not hold, or equivalently, that

$$\forall\beta > 1, \exists\alpha \in (1,\beta), \qquad \min_\rho g(\alpha,\rho) \le 0. \tag{6.48}$$

Then there exists a real sequence $\{\alpha_n\}$ and a sequence $\{\rho_n\}$ of states in $\mathcal{S}(\mathcal{H}_A)$ such that

$$\alpha_n \in \left(1, 1+\frac{1}{n}\right) \quad \text{and} \quad g(\alpha_n,\rho_n) \le 0. \tag{6.49}$$

Since $\mathcal{S}(\mathcal{H}_A)$ is a compact space, there exists a subsequence of ρ_n that converges to some state $\rho_\infty \in \mathcal{S}(\mathcal{H}_A)$ as $n \to \infty$. Relabeling the subsequence to be $\{\rho_n\}$, without loss of generality we can assume that $\rho_n \to \rho_\infty$ as $n \to \infty$. By the mean value theorem, we have that

$$\forall n, \exists\gamma_n \in (1,\alpha_n), \qquad g'(\gamma_n,\rho_n) = \frac{g(\alpha_n,\rho_n) - g(1,\rho_n)}{\alpha_n - 1} \le 0. \tag{6.50}$$

Since $g'(\alpha,\rho)$ is a continuous function of (α,ρ), (6.50) yields

$$g'(1,\rho_\infty) \le 0, \tag{6.51}$$

which contradicts (6.47). ■

Now, Lemma 10 and (6.6) yield our main theorem:

Theorem 11 (Strong converse for EA capacity) *For any sequence of entanglement-assisted codes for a channel $\mathcal{N}$ and with rate $R > I(\mathcal{N})$, the success probability decays exponentially to zero as $n \to \infty$.*

7 Conclusion

This paper provides an alternate path for establishing a strong converse theorem for the entanglement-assisted capacity of any quantum channel. The strong converse theorem, along with the coding theorem from [7, 8, 22], refines our understanding of the entanglement-assisted capacity as a sharp dividing line between what rates of communication are possible or impossible. The approach taken here is to exploit the entanglement-assisted "meta-converse" from [32], several properties of the sandwiched Rényi relative entropy (especially its relation to α-norms) [49, 36, 5, 18], and the main result from [17]. The appeal of the present paper is that it demonstrates the extent to which the powerful Arimoto approach [2, 40, 44] can be helpful in establishing strong converses, and furthermore, we provide an operational relevance for the main result in [17]. The present paper also adds to the existing body of evidence [49, 36, 35, 16] that the sandwiched Rényi relative entropy is the correct quantum generalization of the classical concept for all $\alpha > 1$. Finally, some of the ideas in this work might be helpful in solving the open question from [33] (i.e., establishing a strong converse theorem for the quantum capacity of degradable quantum channels).

Acknowledgements—We are grateful to Naresh Sharma and Andreas Winter for insightful discussions and to Milán Mosonyi for some helpful comments on our paper. We also thank the Department of Physics and Astronomy and the Center for Computation and Technology at Louisiana State University for providing startup funds to support this research.

References

[1] Dorit Aharonov, Alexei Kitaev, and Noam Nisan. Quantum circuits with mixed states. *In Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, pages 20–30, 1998. arXiv:quant-ph/9806029.

[2] Suguru Arimoto. On the converse to the coding theorem for discrete memoryless channels. *IEEE Transactions on Information Theory*, 19:357–359, May 1973.

[3] K. M. R. Audenaert, J. Calsamiglia, R. Muñoz Tapia, E. Bagan, Ll. Masanes, A. Acin, and F. Verstraete. Discriminating states: The quantum Chernoff bound. *Physical Review Letters*, 98:160501, April 2007. arXiv:quant-ph/0610027.

[4] K. M. R. Audenaert, Michael Nussbaum, A. Szkoła, and Frank Verstraete. Asymptotic error rates in quantum hypothesis testing. *Communications in Mathematical Physics*, 279(1):251–283, 2008. arXiv:0708.4282.

[5] Salman Beigi. Sandwiched Rényi divergence satisfies data processing inequality. *Journal of Mathematical Physics*, 54(12):122202, December 2013. arXiv:1306.5920.

[6] Charles H. Bennett, Igor Devetak, Aram W. Harrow, Peter W. Shor, and Andreas Winter. Quantum reverse Shannon theorem. *IEEE Transactions on Information Theory*, 60(5):2926–2959, May 2014. arXiv:0912.5537.

[7] Charles H. Bennett, Peter W. Shor, John A. Smolin, and Ashish V. Thapliyal. Entanglement-assisted classical capacity of noisy quantum channels. *Physical Review Letters*, 83:3081–3084, October 1999. arXiv:quant-ph/9904023.

[8] Charles H. Bennett, Peter W. Shor, John A. Smolin, and Ashish V. Thapliyal. Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem. *IEEE Transactions on Information Theory*, 48(10):2637–2655, 2002. arXiv:quant-ph/0106052.

[9] Mario Berta, Matthias Christandl, and Renato Renner. The quantum reverse Shannon theorem based on one-shot information theory. *Communications in Mathematical Physics*, 306(3):579–615, August 2011. arXiv:0912.3805.

[10] Garry Bowen. Quantum feedback channels. *IEEE Transactions on Information Theory*, 50:2429–2433, October 2004. arXiv:quant-ph/0209076.

[11] Todd A. Brun, Igor Devetak, and Min-Hsiu Hsieh. Correcting quantum errors with entanglement. *Science*, 314(5798):436–439, October 2006. arXiv:quant-ph/0610092.

[12] Todd A. Brun, Igor Devetak, and Min-Hsiu Hsieh. Catalytic quantum error correction. *IEEE Transactions on Information Theory*, 60(6):3073–3089, June 2014. arXiv:quant-ph/0608027.

[13] Eric A. Carlen. Trace inequalities and quantum entropy: An introductory course. *Contemporary Mathematics*, 529:73–140, 2010.

[14] Eric A. Carlen and Elliott H. Lieb. A Minkowski type trace inequality and strong subadditivity of the quantum entropy II. *Letters in Mathematical Physics*, 83(2):107–126, 2008. arXiv:0710.4167.

[15] Nilanjana Datta, Min-Hsiu Hsieh, and Mark M. Wilde. Quantum rate distortion, reverse Shannon theorems, and source-channel separation. *IEEE Transactions on Information Theory*, 59:615–630, January 2013. arXiv:1108.4940.

[16] Nilanjana Datta and Felix Leditzky. A limit of the quantum Rényi divergence. *Journal of Physics A: Mathematical and Theoretical*, 47(4):045304, January 2014. arXiv:1308.5961.

[17] Igor Devetak, Marius Junge, Christopher King, and Mary Beth Ruskai. Multiplicativity of completely bounded p-norms implies a new additivity result. *Communications in Mathematical Physics*, 266:37–63, August 2006. arXiv:quant-ph/0506196.

[18] Rupert L. Frank and Elliott H. Lieb. Monotonicity of a relative Rényi entropy. June 2013. arXiv:1306.5358.

[19] Raúl García-Patrón, Stefano Pirandola, Seth Lloyd, and Jeffrey H. Shapiro. Reverse coherent information. *Physical Review Letters*, 102(21):210501, May 2009. arXiv:0808.0210.

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

[21] Masahito Hayashi. Error exponent in asymmetric quantum hypothesis testing and its application to classical-quantum channel coding. *Physical Review A*, 76:062301, December 2007. arXiv:quant-ph/0611013.

[22] Alexander S. Holevo. On entanglement assisted classical capacity. *Journal of Mathematical Physics*, 43(9):4326–4333, 2002. arXiv:quant-ph/0106075.

[23] Min-Hsiu Hsieh, Todd A. Brun, and Igor Devetak. Entanglement-assisted quantum quasicyclic low-density parity-check codes. *Physical Review A*, 79(3):032340, March 2009. arXiv:0803.0100.

[24] Min-Hsiu Hsieh, Wen-Tai Yen, and Li-Yi Hsu. High performance entanglement-assisted quantum LDPC codes need little entanglement. *IEEE Transactions on Information Theory*, 57(3):1761–1769, March 2011. arXiv:0906.5532.

[25] Rahul Jain, Zhengfeng Ji, Sarvagya Upadhyay, and John Watrous. QIP = PSPACE. *Communications of the ACM*, 53(12):102–109, December 2010. arXiv:0905.1300.

[26] Anna Jencova. A relation between completely bounded norms and conjugate channels. *Communications in Mathematical Physics*, 266(1):65–70, August 2006. arXiv:quant-ph/0601071.

[27] Alexei Kitaev. Quantum computations: algorithms and error correction. *Russian Mathematical Surveys*, 52(6):1191–1249, 1997.

[28] Robert Koenig and Stephanie Wehner. A strong converse for classical channel coding using entangled inputs. *Physical Review Letters*, 103:070504, August 2009. arXiv:0903.2838.

[29] Robert Koenig, Stephanie Wehner, and Jürg Wullschleger. Unconditional security from noisy quantum storage. *IEEE Transactions on Information Theory*, 58(3):1962–1984, March 2012. arXiv:0906.1030.

[30] Elliot H. Lieb. Convex trace functions and the Wigner-Yanase-Dyson conjecture. *Advances in Mathematics*, 11:267–288, December 1973.

[31] Elliott H. Lieb and Walter Thirring. *Studies in mathematical physics*, chapter Inequalities for the moments of the eigenvalues of the Schroedinger Hamiltonian and their relation to Sobolev inequalities, pages 269–297. Princeton University Press, Princeton, 1976.

[32] William Matthews and Stephanie Wehner. Finite blocklength converse bounds for quantum channels. October 2012. arXiv:1210.4722.

[33] Ciara Morgan and Andreas Winter. "Pretty strong" converse for the quantum capacity of degradable channels. *IEEE Transactions on Information Theory*, 60(1):317–333, January 2014. arXiv:1301.4927.

[34] Milán Mosonyi and Fumio Hiai. On the quantum Rényi relative entropies and related capacity formulas. *IEEE Transactions on Information Theory*, 57(4):2474–2487, April 2011. arXiv:0912.1286.

[35] Milán Mosonyi and Tomohiro Ogawa. Quantum hypothesis testing and the operational interpretation of the quantum Rényi relative entropies. September 2013. arXiv:1309.3228.

[36] Martin Müller-Lennert, Frédéric Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel. On quantum Rényi entropies: A new generalization and some properties. *Journal of Mathematical Physics*, 54(12):122203, December 2013. arXiv:1306.3142.

[37] Hiroshi Nagaoka. The converse part of the theorem for quantum Hoeffding bound. November 2006. arXiv:quant-ph/0611289.

[38] Michael Nussbaum and Arleta Szkoła. The Chernoff lower bound for symmetric quantum hypothesis testing. *The Annals of Statistics*, 37(2):1040–1057, 2009. arXiv:quant-ph/0607216.

[39] Tomohiro Ogawa and Hiroshi Nagaoka. Strong converse to the quantum channel coding theorem. *IEEE Transactions on Information Theory*, 45(7):2486–2489, November 1999. arXiv:quant-ph/9808063.

[40] Yury Polyanskiy and Sergio Verdú. Arimoto channel coding converse and Rényi divergence. In *Proceedings of the 48th Annual Allerton Conference on Communication, Control, and Computation*, pages 1327–1333, September 2010.

[41] Bill Rosgen and John Watrous. On the hardness of distinguishing mixed-state quantum computations. *Proceedings of the 20th IEEE Conference on Computational Complexity*, pages 344–354, June 2005. arXiv:cs/0407056.

[42] Massimiliano F. Sacchi. Entanglement can enhance the distinguishability of entanglement-breaking channels. *Physical Review A*, 72:014305, July 2005. arXiv:quant-ph/0505174.

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

[44] Naresh Sharma and Naqueeb Ahmad Warsi. On the strong converses for the quantum channel capacity theorems. June 2012. arXiv:1205.1712.

[45] Maurice Sion. On general minimax theorems. *Pacific Journal of Mathematics*, 8(1):171–176, 1958.

[46] John Watrous. Simpler semidefinite programs for completely bounded norms. *Chicago Journal of Theoretical Computer Science*, 2013(8):1–19, July 2013. arXiv:1207.5726.

[47] Mark M. Wilde. *From Classical to Quantum Shannon Theory*. June 2011. arXiv:1106.1445.

[48] Mark M. Wilde, Min-Hsiu Hsieh, and Zunaira Babar. Entanglement-assisted quantum turbo codes. *IEEE Transactions on Information Theory*, 60(2):1203–1222, February 2013. arXiv:1010.1256v3.

[49] Mark M. Wilde, Andreas Winter, and Dong Yang. Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy. *Communications in Mathematical Physics*, 331(2):593–622, October 2014. arXiv:1306.1586.