Coherent Distributed Source Simulation as Multipartite Quantum State Splitting

Ian George[1] and Hao-Chung Cheng[2-6]

[1]Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA

[2]Department of Electrical Engineering National Taiwan University, Taipei 10617, Taiwan (R.O.C.)

[3]Department of Mathematics, National Taiwan University, Taipei 10617, Taiwan (R.O.C.)

[4]Center for Quantum Science and Engineering, National Taiwan University

[5]Physics Division, National Center for Theoretical Sciences, Taipei 10617, Taiwan (R.O.C.)

[6]Hon Hai Quantum Computing Research Center, Taipei, Taiwan

***Abstract*—Recently, Cheng et al. generalized quantum state splitting (QSS) to the multipartite setting for applications in quantum network information theory. Furthermore, Cheng and Gao recently noted that unipartite convex splitting reduces to soft covering for classical-quantum states. Soft covering is a common primitive in (classical, network) information theory that does not seem to need an extension to the multipartite setting, e.g. when establishing rates for many party distributed source simulation (DSS). This suggests a gap in our understanding of how these tasks are related. In this work we clarify this relation. We first identify "incoherent" QSS without entanglement assistance with DSS in the bipartite setting. Motivated by this, we identify "coherent" DSS as a special case of multipartite QSS. We then establish a one-shot rate region for multipartite QSS in terms of smooth max multipartite mutual information quantities, which in turn implies a one-shot rate region for coherent DSS. Lastly, we establish a strong asymptotic equipartition property for mutual information quantities to establish a strong converse rate region for multipartite QSS and coherent DSS.**

There was at least one mistake in the proof of the strong asymptotic equipartition property, which breaks the claim of a strong converse rate region for the time being. This full version has been updated to only state the weaker results. Significant updates are provided in blue. We have not edited the abstract to make it clear this is the same work.

I. Introduction & Summary

The soft-covering lemma [1] and quantum state splitting [2] are two fundamental techniques for establishing achievable rates for (quantum) information processing tasks in terms of mutual information-induced quantities. In particular, soft-covering has been used to establish achievability rates for distributed source simulation (DSS), channel simulation, channel resolvability (or, relatedly, identification) [3]–[12] and quantum state splitting (QSS) has been used to establish rates for the transfer of quantum messages and channel simulation [2], [13]–[17]. These works have relied on increasingly optimal characterizations of the soft-covering [7], [10], [18], [19] and QSS [2], [20], [21]. These results have culminated in a multipartite version of QSS, which was used to establish a rate for quantum broadcast channel simulation [16] and, along with a related result for multipartite decoupling [22], was a major advancement in tools for establishing quantum multi-user rate regions.

Along with these refinements of soft-covering and QSS, Cheng and Gao also observed that when one considers a classical-quantum state, the error measure for convex splitting, which is used to prove QSS rates, reduces to the measure of error for soft-covering [16]. This formal insight calls into question if there is an operational (rather than formal) correspondence between QSS and some application of soft-covering. In this proceeding, we answer this question, which consists of two parts. In the first part, we establish the rate of distributed source simulation (DSS) as determining the optimal strategy for QSS when correlations with a reference system do not need to be preserved and there is no entanglement assistance. This establishes a correspondence between DSS and a modified version of QSS. This also motivates the notion of distributed source simulation where we require preserving correlations with the reference system, which we call 'coherent' DSS. We establish that for coherent DSS, entanglement assistance from the sender to each party is in general necessary to achieve arbitrarily small error. This allows us to identify 'coherent distributed source simulation' (coherent DSS) as a special case of *multipartite* QSS. Having made this identification, we establish a one-shot rate region for multipartite QSS in terms of smooth max multipartite mutual information and thus a rate region for coherent DSS as well. Finally, we establish a strong asymptotic equipartition property for our smooth multipartite max mutual information quantities. This establishes a strong converse rate region for multipartite QSS and coherent DSS.

a) Conventions: We follow standard notation [23]–[25]. The (subnormalized) density matrices on finite-dimensional joint Hilbert space $A \otimes B$ is denoted $\mathrm{D}(A \otimes B) := \{\rho_{AB} \in \mathrm{Pos}(A \otimes B) : \mathrm{Tr}[\rho_{AB}] = 1\}$, where we use subscripts to denote the relevant spaces. Normally we supress the Kronecker product in our notation, e.g. we may write $A_1^L B_1^2$ to denote $A_1 \otimes \ldots \otimes A_L \otimes B_1 \otimes B_2$. The subnormalized density matrices on AB are denoted $\mathrm{D}_{\leq}(AB)$. We denote the trace distance between two (normalized) states ρ, σ as $\Delta(\rho, \sigma)$ and the purified distance between two (subnormalized) states ρ, σ as $P(\rho, \sigma)$. We denote the space of completely positive, trace-preserving (CPTP) maps from endomorphisms on A to endomorphisms on B as $\mathrm{C}(A, B)$. The Hartley entropy of a register A of a state ρ_{AB} is $H_0(A)_\rho := \log(\mathrm{rank}(\rho_A))$. When we have $L \in \mathbb{N}$, spaces $\{A_i\}_{i\in[L]}$, integers $\{M_i\}_{i\in[L]} \subset \mathbb{N}$, and $S \subset [L]$, we define $M_S := \Pi_{\ell\in S} M_\ell$ and $A_S := \otimes_{\ell\in S} A_\ell$. We also denote any $A_1 - X - A_2$ (short) quantum Markov chains via a subscript, e.g. $\rho_{A_1-X-A_2}$. Finally, throughout, we define any minimization over an empty set to be positive infinity.

II. Unassisted Incoherent Quantum State Splitting as Distributed Source Simulation

In this section we relate QSS modified to not require preservation of correlations with a reference system and without any entanglement assistance to DSS in the two party setting. We first define one-shot distributed source simulation as in [12] except, for consistency within this proceeding, our error measure is the trace distance rather than the L_1-distance.

Definition 1. *Let $\rho_{A_1^2} \in \mathrm{D}(A_1^2)$. Let $\varepsilon \in (0,1)$. The one-shot rate of distributed source simulation (ε-DSS) is the correlation of formation defined as*

$$\begin{aligned} &C_F^\varepsilon(A_1 : A_2)_\rho \\ &\quad = \min\left\{ H_0(X)_{\widetilde{\rho}} : \widetilde{\rho}_{A_1-X-A_2} \,\&\, \Delta(\widetilde{\rho}_{A_1^2}, \rho_{A_1^2}) \leq \varepsilon \right\} . \end{aligned}$$

With this defined, we define our modified bipartite QSS protocol. It only differs from QSS in not considering the reference register in the error criterion, (1), and not providing entanglement assistance (see [17] for comparison).

Definition 2. *Let $\rho_{A_1^2 R}$ be a purification of $\rho_{A_1^2} \in \mathrm{D}(A_1^2)$. Consider a sender Alice and two receivers, Bobs i for $i \in \{1,2\}$, who implement the following protocol:*

1) *Alice holds quantum registers A_1^2 and R is an inaccessible reference system.*
2) *Alice applies a local operation to generate messages $\vec{r} = (r_1, r_2)$, where her local registers update to A'.*
3) *Alice sends message r_i to Bob i via noiseless one-way classical communication.*
4) *Upon receiving their message r_i, Bob i applies a local operation resulting in a joint state $\widehat{\rho}_{RA'B_1^2}$.*

A $(\vec{r},\varepsilon)$-2 party incoherent quantum state splitting protocol without entanglement-assistance (IQSS) for $\rho_{A_1^2 R}$ satisfies

$$\frac{1}{2}\|\widehat{\rho}_{B_1^2} - \rho_{B_1^2}\|_1 \leq \varepsilon \ , \tag{1}$$

where $\rho_{B_1^2 R} := (id_{A_1^2 \to B_1^2} \otimes id_R)(\rho_{A_1^2 R})$. Moreover, we define the sum rate as

$$\mathbf{C}_F^\varepsilon(A_1 : A_2)_\rho := \{\log(M_1^2) : \frac{1}{2}\|\widehat{\rho}_{B_1^2} - \rho_{B_1^2}\|_1 \leq \varepsilon\} \ .$$

We now reduce the rate region to being determined by one-shot DSS.

Proposition 3. *For any bipartite state $\rho_{A_1^2} \in \mathrm{D}(A_1^2)$, the sum rate $(\vec{r},\varepsilon)$-2 party IQSS is*

$$\mathbf{C}_F^\varepsilon(A_1 : A_2)_\rho = 2C_F^\varepsilon(A_1 : A_2)_\rho \ ,$$

and thus optimized by the one-shot DSS scheme.

Proof. First, assume a strategy exists for IQSS for the given target state $\rho_{A_1^2}$ and tolerated error $\varepsilon \in (0,1)$. Without loss of generality denote it

$$\widehat{\rho}_{B_1^2} = \sum_{r_1^2} p(r_1^2) \widehat{\rho}_{B_1}^{r_1} \otimes \widehat{\rho}_{B_2}^{r_2} \ , \tag{2}$$

where we have used that the local operations are solely conditioned on the received messages by Definition 2.

We now show (2) is always a (short) quantum Markov chain. Define the preparation channels $\mathcal{E}_1 : r_1 \mapsto \widehat{\rho}_{B_1}^{r_1}$, $\mathcal{E}_2 : r_1 \mapsto \sum_{r_2} p(r_2|r_1)\widehat{\rho}_{B_2}^{r_2}$, where the latter is a preparation channel by the normalization of conditional probability. By definition of these preparation channels,

$$\begin{aligned} &(\mathcal{E}_1 \otimes \mathcal{E}_2)(\sum_{r_1} p(r_1)|r_1\rangle\langle r_1|_{M_1} \otimes |r_1\rangle\langle r_1|_{M_2}) \\ &= \sum_{r_1} \left[p(r_1)\widehat{\rho}_{B_1}^{r_1} \otimes \sum_{r_2} p(r_2|r_1)\widehat{\rho}_{B_2}^{r_2} \right] \\ &= \sum_{r_1^2} p(r_1)p(r_2|r_1)\widehat{\rho}_{B_1}^{r_1} \otimes \widehat{\rho}_{B_2}^{r_2} \\ &= \widehat{\rho}_{B_1^2} \ , \end{aligned}$$

where the last equality is using the chain rule for conditional probabilities. Thus, we have $B_1 - M_1 - B_2$. By symmetry of this argument, we may also establish $B_1 - M_2 - B_2$. Thus the optimal strategy is for both parties to use the smallest Markov chain that approximates $\rho_{A_1^2}$ to the tolerated error. That is,

$$\begin{aligned} \mathbf{C}_F^\varepsilon(A_1 : A_2)_\rho &= 2 \cdot \min_{\widehat{\rho}_{A_1-X-A_2} : \Delta(\rho_{A_1^2}, \widehat{\rho}_{A_1^2}) \leq \varepsilon} H_0(X)_{\widehat{\rho}} \\ &= 2C_F^\varepsilon(A_1 : A_2)_\rho \ . \end{aligned}$$

This proves the optimal strategy is the optimal $\varepsilon-$DSS strategy. We note if there does not exist a feasible strategy as we had assumed, then our convention implies both rates are infinite and thus the result still holds. □

III. Coherent Distributed Source Simulation as Multipartite Quantum State Splitting

While Proposition 3 does not seem to generalize to the multipartite scenario,[1] it motivates the question of what changes when one considers distributed source simulation (DSS) where we require preserving the correlations with the reference system. We begin by noting such a task requires entanglement assistance in general.

Proposition 4. *Consider the task in Definition 2, but with the error criterion replaced with the new criterion*

$$\frac{1}{2}\|\widehat{\rho}_{B_1^2 R} - \rho_{B_1^2 R}\|_1 \leq \varepsilon \ , \tag{3}$$

where $\rho_{B_1^2 R} := (id_{A_1^2 \to B_1^2} \otimes id_R)(\rho_{A_1^2 R})$. Then the task cannot be performed to arbitrary error $\varepsilon \in (0,1)$ unless $\rho_{A_1^2}$ is a product state.

[1] This is because the construction in the proof of Proposition 3 does not work for more parties.

The result follows immediately from the fact the resulting joint state $\rho_{B_1^2 R}$ is necessarily fully separable but $\rho_{A_1^2 R}$ is entangled between the three registers unless the original state was product.

As we have established that preserving correlations between the reference system and the target to distribute is impossible without assistance, we now define a version of the task which has entanglement assistance.

Definition 5. *Let $\rho_{A_1^L R}$ be a purification of $\rho_{A_1^L} \in \mathrm{D}(A_1^L)$. Consider sender Alice and L receiver, Bobs 1 through L, who implement the following protocol:*

1) *Alice holds quantum registers A_1^L and R is an inaccessible reference system.*
2) *For each $i \in [L]$, an entangled state resource $|\tau\rangle_{\overline{A}_i \overline{B}_i}$ is shared between Alice (holding the $\overline{A}_i$ register) and the receiver Bob i.*
3) *Alice applies a local operation on the A_1^L registers and her portion of the shared entanglement to obtain $\vec{r} = (r_1, \cdots, r_L)$ (her local registers update to A').*
4) *Alice sends r_i to Bob i via noiseless one-way classical communication.*
5) *Upon receiving the messages, each Bob i applies a local operation to $\overline{B}_i$ conditioned on their r_i, resulting in a joint state $\widehat{\rho}_{RA'B_1^L}$.*

A $(\vec{r}, \varepsilon)$-L party coherent distributed source simulation (QDSS) protocol for $\rho_{A_1^L}$ using $\tau_{\overline{A}_1^L \overline{B}_1^L} := \otimes_{i \in [L]} \tau_{\overline{A}_i \overline{B}_i}$ satisfies

$$P_{\mathsf{QDSS}}(\rho_{A_1^L} \| \tau_{A_1^L}) := P(\widehat{\rho}_{B_1^L R}, \rho_{B_1^L R}) \leq \varepsilon \ , \tag{4}$$

where $\rho_{B_1^L R} := (id_{A_1^L \to B_1^L} \otimes id_R)(\rho_{A_1^L} R)$. We define the rate region of QDSS as

$$\left\{ \bigtimes_{\emptyset \neq S \subset [L]} \log(M_S) : P(\widehat{\rho}_{B_1^L R}, \rho_{B_1^L R}) \leq \varepsilon \right\} \ .$$

The above definition is justified as in standard DSS no two receivers share a correlation or channel between them and the sender only sends classical messages. Moreover, the only difference between QDSS and multipartite QSS is that in Definition 5 Alice does not have an extra register A that she wishes to keep (See [17] for formal comparison). Given its importance, we state this observation as a proposition.

Proposition 6. *A $(\vec{r}, \varepsilon)$-m party coherent DSS protocol is a $(\vec{r}, \varepsilon)$-m party state splitting protocol where Alice's A register is trivial.*

a) Coherent DSS Cannot Be Identified with Unipartite Convex Splitting/QSS: Proposition 6 identifies QDSS with *multipartite* DSS. This means that one really needs multipartite convex splitting to achieve QDSS even in the 2-party case. This may be surprising as it is known that *unipartite* convex splitting reduces to soft-covering when one of the registers is classical [16]. However, note that to apply soft-covering to establish multipartite DSS one uses that classical information may be cloned to in some sense ignore the network structure [12], [26]. With quantum states one would not expect this to be the case as teleporting a state to one party is not the same as teleporting a state to a different party.

A. Main Theorems and Multipartite Max Mutual Information Quantities

Our main result is a one-shot rate region for multipartite QSS and QDSS, which relies on the following families of smooth multipartite mutual information quantities, which make use of the smoothing ball $\mathcal{B}^{\varepsilon}(\rho_A) := \{\widetilde{\rho} \in \mathrm{D}_{\leq}(A) : P(\widetilde{\rho}, \rho) \leq \varepsilon\}$.

Definition 7. *Let $\rho_{A_1^L B} \in \mathrm{D}_{\leq}(A_1^L B)$ and $\tau_{A_1^L} \in \mathrm{D}_{\leq}(A_1^L)$. We define the modified smooth max multipartite mutual information as*

$$\mathcal{I}^{\varepsilon}_{\max}(:A_1^L : B)_{\rho|\tau} = \min_{\sigma \in \mathrm{D}(B)} D^{\varepsilon}_{\max}(\rho \| \otimes_i \tau_{A_i} \otimes \sigma_B) \ ,$$

and the 'down arrow' smooth max multipartite mutual information as

$$\begin{aligned} &I^{\downarrow,\varepsilon}_{\max}(:A_1^L : B)_{\rho} \\ &:= \min_{\widetilde{\rho} \in \mathcal{B}^{\varepsilon}(\rho)} \min_{\mathcal{Q}} D_{\max}(\widetilde{\rho} \| \tau_{A_1} \otimes \tau_{A_2} \ldots \otimes \tau_{A_L} \otimes \sigma_B) \ , \end{aligned}$$

where $\mathcal{Q} := \{\bigtimes_{i \in [L]} \tau_{A_i} \times \sigma_B \in \bigtimes_{i \in [L]} \mathrm{D}(A_i) \times \mathrm{D}(B)\}$.

We note that we use the symbol $\downarrow$ as a superscript as in the case there is only one A register, this reduces to $I^{\downarrow}_{\max}$ in the notation of [12]. We also remark that whenever we consider the marginal $\rho_{A_S B}$ where $A_S = \otimes_{i \in S} A_i$, we will simplify the mutual information quantity $\mathbb{I}$ we are considering evaluating over to the notation $\mathbb{I}(:A_S : B)$.

With the multipartite mutual information quantities and simplified notation defined, we state our main theorem.

Theorem 8. *For any $L \in \mathbb{N}$, the one-shot rate region for multipartite quantum state splitting of a target state $\rho_{AA_1^L} \in \mathrm{D}(AA_1^L)$ with error $\delta \in (0,1)$ and entanglement assistance induced by the state $\tau_{A_1^L} \in \mathrm{D}(A_1^L)$, satisfies for all $\emptyset \neq S \subset [L]$,*

$$\begin{aligned} \mathcal{I}^{\varepsilon_S}_{\max}(: A_S : R)_{\rho_{A_S R}|\tau_{A_S}} &+ 2\log(1/\delta'_S) \\ &\geq \log(M_S) \\ &\quad \geq I^{\downarrow,\delta}_{\max}(: C_{i_1} : \cdots C_{i_{|S|}} : R)_{\rho} \ . \end{aligned}$$

where $\rho_{A_1^L R}$ is an arbitrary purification of $\rho_{A_1^L}$, $\delta' \in (0,\delta)$, $T := |\{S : \emptyset \neq S \subset [L]\}|$, $\varepsilon_S := \frac{\delta - \delta'}{2^{|S|}T}$, and $\delta'_S := \frac{\delta'}{2^{|S|-1}T}$ for each $\emptyset \neq S \subset [L]$.

Using Proposition 6, we may conclude the following.

Corollary 9. *For any $L \in \mathbb{N}$ and target state $\rho_{A_1^L} \in \mathrm{D}(A_1^L)$, Theorem 8 provides inner and outer bounds on the rate region for $(\vec{r}, \varepsilon)$-L party coherent distributed source simulation.*

By establishing a weak asymptotic equipartition property for $I^{\downarrow,\varepsilon}_{\max}(:A_L : R)_{\rho}$, we recover the asymptotic rate region.

Theorem 10. *The vanishing-error asymptotic rate region of multipartite quantum state splitting is the polytope*

$$\bigtimes_{\emptyset \neq S \subset [L]} [I(:A_S:R)_\rho), +\infty) \ ,$$

where $\rho_{A_1^L R}$ is an arbitrary purification of $\rho_{A_1^L}$.

Theorem 8 is established by combining Corollary 13 and Lemma 16. The achievability is established in three steps. The first step is to convert the error exponent statement for multipartite convex splitting in terms of Rényi multipartite mutual information quantities from [17] into an error exponent in terms of our $\mathcal{I}^{\varepsilon}_{\max}(:A_S:B)_{\rho|\tau}$ using roughly the same method as [22] uses to convert their Rényi entropy error exponents into smooth entropy error exponents. The second step is to establish error exponents for multipartite QSS from multipartite convex splitting. The last step is to convert our error exponent form into a small deviations achievability result. The converse follows the same proof method as the converse for channel simulation given in [17], [27], which relies on a few properties that we establish for our relevant quantity. The rest of the proceeding explains these proofs in as much detail as the space allows, and the rest is in the appendix for the reviewers.

B. Achievability Proof

The main technical step is to establish error exponents for multipartite convex splitting in terms of smoothed quantities. We do this by modifying the proof of [17, Theorem 3.12] to be more similar to that of the proof of [22, Theorem 3.2].

Lemma 11. *Consider $\rho_{A_1^L E}$ and $\tau_{A_1^L}$. Let $\{M_i\}_{i\in[L]} \subset \mathbb{N}$. Define*

$$\omega_{A_1^{M_1}\ldots A_L^{M_L} E} := (M_L)^{-1} \sum_{i\in[L], m_i \in M_i} \rho_{A_{1,m_1}\ldots A_{L,m_l}E} \cdot \left(\otimes_{\ell\in[L],\overline{m}_\ell \in M_\ell \setminus \{m_\ell\}} \tau_{A_\ell,\overline{m}_\ell}\right) \ ,$$

where $A_\ell^{M_\ell} := A_{\ell,1}\ldots A_{\ell,M_\ell} \cong A_\ell^{\otimes M_\ell}$ for each $\ell \in [L]$. Consider smoothing parameters $\{\varepsilon_S\}_{\emptyset\neq S\subset[L]} \in [0,1]$, then we have the following error exponents for convex splitting:

$$\Delta_{M_S}(\omega, \otimes_{i\in[L]} \tau_{A_i} \otimes \rho_R) \leq \sum_{\emptyset\neq S\subset[L]} 2^{|S|} \cdot \left(2\varepsilon_S + 2^{-E^{\varepsilon_S}_{\log M_S}(\rho_{A_S R}\|\tau_{A_S})}\right) \ , \tag{5}$$

where the error exponent function is defined as

$$E^{\varepsilon_S}_r(\rho_{A_S}) := \sup_{p\in[1,2]} \frac{p-1}{p}(r - \mathcal{I}^{\varepsilon_S}_p(:A_1^L:E)_{\rho_{A_S}|\tau_{A_S}}) \ ,$$

and $I^{\varepsilon_S}_p(:A_1^L:R)$ is the same as the definition $I^{\downarrow,\varepsilon_S}_{\max}$, but with the max divergence replaced with a Sandwiched Rényi divergence with parameter p [25].

Proof. As explained in [17, Top of Page 18], by triangle inequality, we are interested in bounding the term

$$\|\Theta_{[L]} \circ T_S(\frac{\rho_{\mathbf{A}_S E}}{\tau_{A_S}\otimes E})\|_{L_1(\tau_{\mathbf{A}_S}\otimes\sigma_E)}$$

for each S where the norm is Kosaki's weight L_p-norm with respect to the density operator $\tau_{\mathbf{A}_S}\otimes\sigma_E$) (See [16], [17] for background) and $\tau_{\mathbf{A}_S} := \otimes_{i\in S}\tau^{\otimes M_i}_{A_i}$ and $\rho_{\mathbf{A}_S}$ is on the same space. The choice of $\tau_{A_1^L}$ must be consistent across every S. σ_E may be varied for each S, and we will make use of this.

Imagine we choose any $p\in[1,2]$. Now let $\widetilde{\rho}_{A_S}, \sigma_E$ be the optimizer for the modified smooth Rényi mutual information $I^{\varepsilon_S}_p(:A_S:E)_{\rho_{A_S}|\tau_{A_S}}$. Note this means $\frac{1}{2}\|\widetilde{\rho}_{A_S} - \rho_{A_S}\|_1 \leq \varepsilon_S$ since purified distance upper bounds trace distance. Now,

$$\begin{aligned}
&\|\Theta_{[L]}\circ T_S(\frac{\rho_{A_S E}}{\tau_{A_S}\otimes\sigma_E} - \frac{\widetilde{\rho}_{A_S E}}{\tau_{A_S}\otimes\sigma_E})\|_{L_1(\tau_{\mathbf{A}_S}\otimes\sigma_E)} \\
\leq&\|\Theta_{[L]}\circ T_S\| \cdot \|\frac{\rho_{A_S E}}{\tau_{A_S}\otimes\sigma_E} - \frac{\widetilde{\rho}_{A_S E}}{\tau_{A_S}\otimes\sigma_E}\|_{L_1(\tau_{\mathbf{A}_S}\otimes\sigma_E)} \\
\leq&2^{|S|} \cdot \|\frac{\rho_{A_S E}}{\tau_{A_S}\otimes\sigma_E} - \frac{\widetilde{\rho}_{A_S E}}{\tau_{A_S}\otimes\sigma_E}\|_{L_1(\tau_{\mathbf{A}_S}\otimes\sigma_E)} \\
=&2^{|S|} \cdot \|\rho_{A_S E} - \widetilde{\rho}_{A_S E}\|_1 \\
\leq&2^{|S|} \cdot 2\varepsilon_S \ ,
\end{aligned}$$

where the second inequality is [17, Lemmas 3.7 and 3.9], the equality is by definition of Kosaki's weighted L_p-norm for the weight $\gamma = 1/2$, and the final inequality is by our smoothing ball bound. Therefore, by triangle inequality, the inequality we just established, and the map norm,

$$\begin{aligned}
&\|\Theta_{[L]}\circ T_S(\frac{\rho_{A_S E}}{\tau_{A_S}\otimes\sigma_E})\|_{L_1(\tau_{\mathbf{A}_S}\otimes\sigma_E)} \\
\leq&2^{|S|}2\varepsilon_S + \|\Theta_{[L]}\circ T_S(\frac{\widetilde{\rho}_{A_S E}}{\tau_{A_S}\otimes\sigma_E})\|_{L_1(\tau_{\mathbf{A}_S}\otimes\sigma_E)} \\
=&2^{|S|}\left(2\varepsilon_S + \|\frac{\widetilde{\rho}_{A_S E}}{\tau_{A_S}\otimes\sigma_E}\|_{L_1(\tau_{\mathbf{A}_S}\otimes\sigma_E)}\right) \ .
\end{aligned}$$

By complex interpolation provided in [17, Lemma 3.10], we can replace $\|\frac{\widetilde{\rho}_{A_S E}}{\tau_{A_S}\otimes\sigma_E}\|_{L_1(\tau_{A_S}\otimes\sigma_E)}$ with $M_{[L]}^{\frac{1-p}{p}}\|\frac{\widetilde{\rho}_{A_S E}}{\tau_{A_S}\otimes\sigma_E}\|_{L_p(\tau_{A_S}\otimes\sigma_E)}$ for any choice of $p\in[1,2]$.

Now note by definition of Kosaki's weighted L_p-norm for the choice of $\gamma = 1/2$, one can verify [16, Eqs. 113-115]

$$\left\|\frac{\widetilde{\rho}_{A_S E}}{\tau_{A_S}\otimes\sigma_E}\right\|_{L_p(\widetilde{\rho}_{A_S E})} = \exp\left(\frac{p-1}{p}D_p(\widetilde{\rho}_{A_S E}\|\tau_{A_S}\otimes\sigma_E)\right) \ .$$

Finally, note that for every S we chose $\widetilde{\rho}_{A_S E}$ to optimize $\mathcal{I}^{\varepsilon}_p(:A_S:E)_{\rho|\tau}$. Noting there is actually a factor of a half in front of each norm, re-organizing everything in the exponents, and summing over the $\emptyset\neq S\subset[L]$ completes the proof. □

We can now convert the previous lemma to establish an error exponent result for multipartite QSS, where error is measured in purified distance and we use the same error exponents as in the previous lemma. This makes use of Uhlmann's theorem and is similar to proofs in [16], [17].

Lemma 12. *Given a pure state $\rho_{AA_1^L R}$, there exists a $(\vec{r},\varepsilon)$-L-receiver quantum state splitting protocol using entangled resources $|\tau\rangle^{\otimes[M_i]}_{A_i'B_i}$ for $i\in[L]$ such that*

$$P_{\mathsf{Split}}(\rho\|\tau) \leq \sum_{\emptyset\neq S\subset[L]} 2^{|S|-1}\cdot(2\varepsilon_S + 2^{-E^{\varepsilon_S}_{\log M_S}(\rho_{A_S E}\|\tau_{A_S})}) \ ,$$

where

$$P_{\mathsf{Split}}(\rho_{AA_1^L R} \| \tau_{A_1^L}) := P(\widehat{\rho}_{RAB_1^L}, \rho_{RAB_1B_2}) \le \varepsilon \ , \tag{6}$$

and $\widehat{\rho}_{AB_1^L R}$ is the output of the protocol where A is a register Alice holds onto and $B_i \cong A_i$ for all $i \in [L]$.

Finally, we can convert the error exponents into a small deviations achievable rate region.

Corollary 13. *Let $\delta \in (0,1)$. Then there exists a $\delta-$approximate state splitting scheme in purified distance satisfying $\forall \emptyset \neq S \subset [L]$,*

$$\log(M_S) \le \mathcal{I}_{\max}^{\varepsilon_S}(: A_S : E)_{\rho_{A_S E} | \tau_{A_S}} + 2\log\left(1/\delta'_S\right) \ ,$$

where $\delta' \in (0,\delta)$, $T := |\{S : \emptyset \neq S \subset [L]\}|$, $\varepsilon_S := \frac{\delta - \delta'}{2^{|S|}T}$, and $\delta'_S := \frac{\delta'}{2^{|S|-1}T}$ for each $\emptyset \neq S \subset [L]$.

Proof. It suffices to choose parameters that guarantee the RHS of Lemma 12 is upper bounded by δ. A direct calculation will verify that our choices of ε_S and δ'_S will satisfy this so long as $2^{-E^{\varepsilon_S}_{\log M_S}(\rho_{A_S E} \| \tau_{A_S})} \le \delta'_S$ for each $\emptyset \neq S \subset [L]$. We note the following implications and equivalence:

$$\begin{aligned} &\log(M_S) = \mathcal{I}_{\max}^{\varepsilon_S}(: A_S : E)_{\rho_{A_S E} | \tau_{A_S}} + 2\log\left(1/\delta'_S\right) \quad (7) \\ \Rightarrow &\log(M_S) - \mathcal{I}_2^{\varepsilon_S}(: A_S : E)_{\rho_{A_S E} | \tau_{A_S}} \ge -\frac{2}{2-1}\log(\delta'_S) \\ \Leftrightarrow &2^{-\frac{2-1}{2}\left(\log(M_S) - \mathcal{I}_2^{\varepsilon_S}(: A_S : E)_{\rho_{A_S E} | \tau_{A_S}}\right)} \le \delta'_S \\ \Rightarrow &2^{-E^{\varepsilon_S}_{\log M_S}(\rho_{A_S E} \| \tau_{A_S})} \le \delta'_S \ , \end{aligned}$$

where in the first implication we used that the Sandwiched Rényi divergence monotonically increases in parameter p, and the second implication is by the definition of the error exponent function. As this holds for all $\emptyset \neq S \subset [L]$, Eq. (7) defines an achievable rate point, and thus an upper bound on the needed rate. This completes the proof. ☐

C. Converse Proof

The converse follows from the structure of the entanglement assistance, monotonicity under local maps for $I_{\max}^{\downarrow,\varepsilon}(: A_1^L : B)_\rho$, and the following chain rule.

Proposition 14 (Classical-Specific Chain Rule)**.** *Let $\varepsilon \in [0,1)$. Consider $\rho_{A_1^L X_1^L B}$, where the register sizes may vary across index i. Then*

$$\begin{aligned} &I_{\max}^{\downarrow,\varepsilon}(A_1X_1 : A_2X_2 : \cdots : A_LX_L : B)_\rho \\ &\qquad \le I_{\max}^{\downarrow,\varepsilon}(: A_1^L : B)_\rho + \sum_i \log|X_i| \end{aligned}$$

This chain rule may be proven by building a feasible point of the optimization for $I_{\max}^{\downarrow,\varepsilon}(A_1X_1 : A_2X_2 : \cdots : A_LX_L : B)_\rho$ using the optimizer of $I_{\max}^{\downarrow,\varepsilon}(: A_1^L : B)_\rho$.

General multipartite mutual information quantities induced by divergences satisfying data processing are monotonic under local maps via a straightforward proof. We state the specific case we will use.

Proposition 15. *For $\rho \in \mathrm{D}_{\le}(A_1^L B)$, $\varepsilon \in [0,1)$, and CPTP maps $\{\mathcal{E}_i \in \mathrm{C}(A_i, A'_i)\}_{i \in [L]}$, $\mathcal{F} \in \mathrm{C}(B, B')$,*

$$I_{\max}^{\downarrow,\varepsilon}(: A_1^L : B)_\rho \ge I_{\max}^{\downarrow,\varepsilon}(: {A'_1}^L : B')_{\otimes_i \mathcal{E}_i \otimes \mathcal{F}(\rho)}$$

We now establish our converse.

Lemma 16 (Converse Rate Region)**.** *Let $\delta \in (0,1)$ be the error of the state splitting in purified distance. Then,*

$$\log(M_S) \ge I_{\max}^{\downarrow,\delta}(: C_{i_1} : \cdots C_{i_{|S|}} : R)_\rho \quad \forall \emptyset \neq S \subset [L] \ .$$

Proof. This follows the converse for broadcast channel simulation in [17], which follows [27]. Let $\mathcal{E}$ be Alice's encoder and $\mathcal{D}_i$ be party i's local decoder. Define the final state as $\widehat{\rho}$. Note that by definition of state splitting, the original entangled resource is assumed to be such that each party *individually* shares entangled states with each of the encoders just as in Definition 5. This means the registers $\overline{B}_i$ and $\overline{B}_{i'}$ are independent for $i \neq i'$ and each of these registers is also independent of the R register. Thus, any mutual information quantity between the $\overline{B}_i$ registers and R is zero. Also note that by partial trace $\tau_{\overline{B}_1^L R} = (\mathcal{E}(\rho \otimes \tau))_{\overline{B}_1^L} = \otimes_{i \in [L]} \tau_{\overline{B}_i} \otimes \rho_R$. That is, this marginal does not change under the action of the encoder. In this case, for any $\emptyset \neq S \subset [L]$,

$$\begin{aligned} &\sum_{i \in [S]} \log(M_i) \\ =&\sum_{i \in [S]} \log(M_i) + I_{\max}^{\downarrow}(: \overline{B}_{i_1} : \cdots \overline{B}_{i_{|S|}} : R)_\tau \\ =&\sum_{i \in [S]} \log(M_i) + I_{\max}^{\downarrow}(: \overline{B}_{i_1} : \cdots \overline{B}_{i_{|S|}} : R)_{\mathcal{E}(\rho \otimes \tau)} \\ \ge& I_{\max}^{\downarrow}(: M_{i_1}\overline{B}_{i_1} : \ldots : M_{i_{|S|}}\overline{B}_{i_{|S|}} : R)_{\mathcal{E}(\rho \otimes \tau)} \\ \ge& I_{\max}^{\downarrow}(: B_{i_1} : \cdots B_{i_{|S|}} : R)_{\otimes_{i \in \mathcal{D}_i} \circ \mathcal{E}(\rho \otimes \tau)} \\ =& I_{\max}^{\downarrow}(: B_{i_1} : \cdots B_{i_{|S|}} : R)_{\widehat{\rho}} \\ \ge& I_{\max}^{\downarrow,\delta}(: A_{i_1} : \cdots A_{i_{|S|}} : R)_\rho \\ =& I_{\max}^{\downarrow,\delta}(: A_S : R)_\rho \ , \end{aligned}$$

where the first equality is the independence, the second equality is that the encoder does not change this marginal as previously discussed, the first inequality is Proposition 14, the second is the local data-processing property (Proposition 15) and that B_i is the register after the decoder, and the third is noting we require the output state be δ close to the target state ρ in purified distance, so we may minimize over that smoothing ball to get a lower bound and we re-label the B_i's as A_i's. ☐

Acknowledgment

I.G. is supported by NSF Grant No. 2112890. H.-C. Cheng is supported by Grants No. NSTC 112-2636-E-002-009, No. NSTC 112-2119-M-007-006, No. NSTC 112-2119-M-001-006, No. NSTC 112-2124-M-002-003, No. NTU-112V1904-4, NTU-CC-113L891605, and NTU-113L900702.

References

[1] A. Wyner, "The common information of two dependent random variables," *IEEE Transactions on Information Theory*, vol. 21, no. 2, pp. 163–179, 1975.

[2] A. Abeyesinghe, I. Devetak, P. Hayden, and A. Winter, "The mother of all protocols: Restructuring quantum information's family tree," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 465, no. 2108, pp. 2537–2563, 2009.

[3] R. Ahlswede and A. Winter, "Strong converse for identification via quantum channels," *IEEE Transactions on Information Theory*, vol. 48, no. 3, pp. 569–579, 2002.

[4] I. Devetak and A. Winter, "Classical data compression with quantum side information," *Physical Review A*, vol. 68, no. 4, p. 042301, 2003.

[5] N. Cai, A. Winter, and R. W. Yeung, "Quantum privacy and quantum wiretap channels," *problems of information transmission*, vol. 40, pp. 318–336, 2004.

[6] A. Winter, "Secret, public and quantum correlation cost of triples of random variables," in *Proceedings. International Symposium on Information Theory, 2005. ISIT 2005.* IEEE, 2005, pp. 2270–2274.

[7] M. Hayashi, *Quantum Information: An Introduction*. Springer, 2006.

[8] P. Cuff, "Communication requirements for generating correlated random variables," in *2008 IEEE International Symposium on Information Theory*, 2008, pp. 1393–1397.

[9] Z. Luo and I. Devetak, "Channel simulation with quantum side information," *IEEE Transactions on Information Theory*, vol. 55, no. 3, pp. 1331–1342, 2009.

[10] P. Cuff, "Distributed channel synthesis," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7071–7096, 2013.

[11] C. H. Bennett, I. Devetak, A. W. Harrow, P. W. Shor, and A. Winter, "The quantum reverse shannon theorem and resource tradeoffs for simulating quantum channels," *IEEE Transactions on Information Theory*, vol. 60, no. 5, pp. 2926–2959, 2014.

[12] I. George, M.-H. Hsieh, and E. Chitambar, "One-shot distributed source simulation: As quantum as it can get," 2023.

[13] M. Berta, M. Christandl, and R. Renner, "The quantum reverse shannon theorem based on one-shot information theory," *Communications in Mathematical Physics*, vol. 306, pp. 579–615, 2011.

[14] A. Anshu, V. K. Devabathini, and R. Jain, "Quantum communication using coherent rejection sampling," *Physical review letters*, vol. 119, no. 12, p. 120506, 2017.

[15] A. Anshu, M. Berta, R. Jain, and M. Tomamichel, "Partially smoothed information measures," *IEEE Transactions on Information Theory*, vol. 66, no. 8, pp. 5022–5036, 2020.

[16] H.-C. Cheng and L. Gao, "Tight one-shot analysis for convex splitting with applications in quantum information theory," *arXiv preprint arXiv:2304.12055*, 2023.

[17] H.-C. Cheng, L. Gao, and M. Berta, "Quantum broadcast channel simulation via multipartite convex splitting," 2023.

[18] H.-C. Cheng and L. Gao, "Error exponent and strong converse for quantum soft covering," *IEEE Transactions on Information Theory*, 2023.

[19] Y.-C. Shen, L. Gao, and H.-C. Cheng, "Optimal second-order rates for quantum soft covering and privacy amplification," *IEEE Transactions on Information Theory*, pp. 1–1, 2024.

[20] F. Leditzky, M. M. Wilde, and N. Datta, "Strong converse theorems using rényi entropies," *Journal of Mathematical Physics*, vol. 57, no. 8, 2016.

[21] N. Ramakrishnan, M. Tomamichel, and M. Berta, "Moderate deviation expansion for fully quantum tasks," *IEEE Transactions on Information Theory*, 2023.

[22] P. Colomer Saus and A. Winter, "Decoupling by local random unitaries without simultaneous smoothing, and applications to multi-user quantum information tasks," *arXiv e-prints*, pp. arXiv–2304, 2023.

[23] J. Watrous, *The Theory of Quantum Information*. Cambridge University Press, 2018.

[24] M. M. Wilde, *Quantum information theory*. Cambridge university press, 2013.

[25] M. Tomamichel, *Quantum information processing with finite resources: mathematical foundations*. Springer, 2015, vol. 5.

[26] W. Liu, G. Xu, and B. Chen, "The common information of n dependent random variables," in *2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2010, pp. 836–843.

[27] M. X. Cao, N. Ramakrishnan, M. Berta, and M. Tomamichel, "Channel simulation: Finite blocklengths and broadcast channels," *arXiv preprint arXiv:2212.11666*, 2022.

[28] J. Watrous, *The Theory of Quantum Information*. Cambridge University Press, 2018.

APPENDIX

Here we provide the extra proofs for an interested reviewer. We first provide the proof that entanglement assistance is necessary for coherent disitributed source simulation (Proposition 4). We then provide omitted proofs for establishing Theorem 8. The rest of the appendix is proving Theorem **??**.

Proof of Proposition 4. We prove the bipartite case for clarity, but a multipartite extension follows the same proof. The result follows immediately from the fact the resulting joint state $\rho_{B_1^2 R}$ is necessarily fully separable but $\rho_{A_1^2 R}$ is entangled between the three registers in general. To formally show this, note that Alice performs a local CPTP map $\mathcal{E}_{A_1^2 \to A'^2_1 M_1^2}$, resulting in the state

$$\begin{aligned}\rho_{RA'^2_1 M_1^2} \\ = \sum_{m_1,m_2} p(m_1,m_2)|m_1\rangle\langle m_1| \otimes |m_2\rangle\langle m_2| \otimes \rho^{(m_1,m_2)}_{RA'^2_1} \ .\end{aligned}$$

After the messages are transmitted, the Bobs apply local maps $\mathcal{D}_{M_i \to B_i}$, resulting in the state

$$\rho_{RA'^2_1 B_1^2} = \sum_{m_1,m_2} p(m_1,m_2)\rho^{m_1}_{B_1} \otimes \rho^{m_2}_{B_2} \otimes \rho^{(m_1,m_2)}_{RA'^2_1} \ .$$

By definition, this state is fully separable across the subsystems B_1, B_2, and R. As the set of fully separable states is closed,[2] if $\rho_{RA_1^2}$ was entangled between any two registers, one cannot approach the target state to arbitrary precision under the restrictions of the protocol. The only case where the original state is fully separable is when $\rho_{A_1^2}$ is a product state, i.e. $\rho_{A_1^2} = \otimes_{i \in \{1,2\}} |\psi_i\rangle\langle\psi_i|_{A_i}$. This completes the proof. ☐

A. Proofs for Theorem 8

Proof of Lemma 12. We prove the bipartite case, the multipartite argument is an immediate extension. Let $M_1 = 2^{r_1}$, $M_2 = 2^{r_2}$. Let $\tau_{A'_i} = \rho_{A_i}$ for $i \in [2]$. This means $|\tau\rangle_{A'_i B'_i}$ is a purification of ρ_{A_i}. Let Alice and receiver i (Bob i) share M_i copies of $|\tau\rangle_{A'_i B'_i}$, denoted $\otimes_{m \in [M_i]} |\tau\rangle_{A'_{i,m} B'_{i,m}}$, where Alice has the $A'_{i,m}$ systems and Bob has the $B'_{i,m}$ systems. Thus, the initial global state is

$$|\overline{\omega}\rangle = |\rho\rangle_{AA'_1A'_2R} \otimes_{m \in [M_1]} |\tau\rangle_{A'_{1,m}B'_{1,m}} \otimes_{m \in [M_2]} |\tau\rangle_{A'_{2,m}B'_{2,m}} \ . \tag{8}$$

Now the goal is to perform a process such that one ends with the final pure state

$$\begin{aligned}|\omega\rangle := \frac{1}{\sqrt{M_{[2]}}} \sum_{\substack{m_1 \in M_1 \\ m_2 \in M_2}} |m_1\rangle_{M_1} |m_2\rangle_{M_2} |\rho\rangle_{AB_{1,m_1}B_{2,m_2}R} \\ \cdot |0\rangle_{A'_{1,m_1}A'_{2,m_2}} \otimes_{\overline{m_1} \in [M_1] \setminus \{m_1\}} |\tau\rangle_{A'_{1,\overline{m}_1}B_{1,\overline{m}_1}} \\ \otimes_{\overline{m_2} \in [M_2] \setminus \{m_2\}} |\tau\rangle_{A'_{2,\overline{m}_2}B_{2,\overline{m}_2}} \ ,\end{aligned} \tag{9}$$

where $M_{[2]} := \pi_{\ell \in [2]} M_\ell$, Alice holds registers $M_1, M_2, A, A'_{1,[M_1]} := A'_{1,1} \otimes \dots A'_{1,M_1}$, and $A'_{2,[M_2]} := A'_{2,1} \otimes \dots A'_{2,M_2}$, and Bob holds registers $B_{1,[M_1]} := B_{1,1} \otimes \dots B_{1,M_1}$, and $B_{2,[M_2]} := B_{2,1} \otimes \dots B_{2,M_2}$. If this is the case, then Alice can measure registers M_1, M_2 and send outcome m_i to receiver i (Bob i) using r_i bits. Then the two receivers can simply keep registers B_{i,m_i} so that the global state is $|\rho\rangle_{AB_{1,m_1}B_{2,m_2}R} \cong |\rho\rangle_{AB_1B_2R}$.

The rest of the proof is showing there is a process that approximates $|\omega\rangle$ (9) to an error given in the theorem statement. First note that a marginal of $|\overline{\omega}\rangle$ (8) is

$$\overline{\omega}_{A'_{1,[M_1]}A'_{2,[M_2]}R} = \otimes_{m_1 \in [M_1]} \tau_{A'_{1,m_1}} \otimes_{m_2 \in [M_2]} \tau_{A'_{2,m_2}} \otimes \rho_R \ . \tag{10}$$

Lemma 11 tells us that we can approximate the above marginal by the density matrix

$$\begin{aligned}\omega_{A_1,M_1{}'A'_{2,[M_2]}R} \\ = \frac{1}{M_{[2]}} \sum_{(m_1,m_2) \in [M_1] \times [M_2]} \rho_{A_1,m_1{}'A'_{2,m_2}R} \\ \otimes_{\overline{m_1} \in [M_1] \setminus \{m_1\}} \tau_{A'_{1,\overline{m}_1}} \otimes_{\overline{m_2} \in [M_2] \setminus \{m_2\}} \tau_{A'_{2,\overline{m}_2}}\end{aligned} \tag{11}$$

up to trace distance error ε' that is upper bounded by the right hand side of (5) via the replacements $A \to A'_1$, $B \to A'_2$, $E \to R$. Note that $\omega_{A_1,M_1{}'A'_{2,[M_2]}R}$ is a reduced state of ideal final state (9). As already noted, we have $|\rho\rangle_{AB_{1,m_1}B_{2,m_2}R} \cong |\rho\rangle_{AB_1B_2R}$, which implies $\rho_{AB_1B_2R} \cong \rho_{AB_1B_2R}$. Combining these points, $\omega_{AA'_1A'_2R}$ is the marginal of $|\omega\rangle$ such that $\omega_{AB_1B_2} \cong \overline{\omega}_{AA'_1A'_2}$. It follows by Uhlmann's theorem, there is an isometry $V : AA'_1A'_2A'_{1,[M_1]}A'_{2,[M_2]} \to M_1M_2AA'_{1,[M_1]}A'_{2,[M_2]}$ such that $\frac{1}{2}\|V|\overline{\omega}\rangle\langle\overline{\omega}|V^* - |\omega\rangle\langle\omega|\|_1 \leq \sqrt{2\varepsilon'}$. Replacing ε' with the error exponent in Lemma 11 gets the trace distance bound. Alternatively, ε' bounds the purified distance directly, which is what is stated. ☐

Proof of Proposition 14. We prove the two party case as it's notationally simpler and the method immediately generalizes. By definition, we have

$$\begin{aligned}&\exp(I^{\downarrow,\varepsilon}_{\max}(A_1X_1 : A_2X_2 : B)_\rho) \\ =& \min_{\substack{\widetilde{\rho} \in \mathcal{B}^\varepsilon(\rho) \\ \{\tau_{A_iX_i} \in D(A_iX_i)\}_{i \in [2]} \\ \sigma_B \in \mathrm{D}(B)}} \left\{ \lambda : \lambda \tau_{A_1X_1} \otimes \tau_{A_1^2X_1^2} \otimes \sigma_B \succeq \widetilde{\rho}_{A_1^2X_1^2B} \right\} \\ =& \min_{\substack{\widetilde{\rho} \in \mathcal{B}^\varepsilon(\rho) \\ \{q_{X_i} \in \mathcal{P}(\mathcal{X}_i)\}_{i \in [2]} \\ \{\tau_{A_i} \in D(A_i)\}_{i \in [2]} \\ \sigma_B \in \mathrm{D}(B)}} \left\{ \lambda : \lambda q_{X_1}(x_1) q_{X_2}(x_2) \tau^{x_1}_{A_1} \otimes \tau^{x_2}_{A_2} \otimes \sigma_B \right. \\ & \left. \qquad \succeq \widetilde{p}_{X_1^2}(x_1^2) \widetilde{\rho}^{x_1^2}_{A_1^2B} \ \forall x \in \mathcal{X} \right\} \ ,\end{aligned} \tag{12}$$

where we have used the block diagonality and we have labeled the distributions' spaces for clarity.

Let $\log(\mu) = I^{\downarrow,\varepsilon}_{\max}(A_1 : A_2)_\rho$. That is there exists $\widehat{\rho}_{A_1^2} \in \mathcal{B}^\varepsilon(\rho_{A_1^2})$, $\{\sigma_{A_i} \in D(A_i)\}_{i \in [2]}$, $\sigma_B \in \mathrm{D}(B)$ such that $\mu\sigma_{A_1} \otimes \sigma_{A_2} \otimes \sigma_B \succeq \widehat{\rho}_{A_1^2B}$. Now note that by a standard property of purified distance [25], there exists an extension of $\widetilde{\rho}_{A_1^2}$, $\widehat{\rho}_{A_1^2X_1^2B}$, such that $P(\widehat{\rho}_{A_1^2X_1^2B}, \rho_{A_1^2X_1^2B}) \leq \varepsilon$. Moreover, by

[2]This follows from a direct extension of the proof of bipartite separable states being closed in [28].

DPI of purified distance, this $\widehat{\rho}$ has the same QCQ structure as ρ. Then we may write

$$\mu\sigma_{A_1} \otimes \sigma_{A_2} \otimes \sigma_B \succeq \widehat{\rho}_{A_1^2 B} = \sum_{x_1^2 \in \mathcal{X}_1 \times \mathcal{X}_2} \widehat{p}_{X_1^2}(x_1^2)\widehat{\rho}^{x_1^2}_{A_1^2 B} \ , \tag{13}$$

where we just used the structure of our CQ extension. Next note that for any distributions $\{\widehat{q}_{X_1} \in \mathcal{P}(\mathcal{X}_i)\}_{i\in[2]}$, we have

$$\begin{aligned}
&\mu\sigma_{A_1} \otimes \sigma_{A_2} \otimes \sigma_B \\
=&\mu \left(\sum_{x_1 \in \mathcal{X}_1} q_{X_1}(x_1) \right) \sigma_{A_1} \otimes \left(\sum_{x_1 \in \mathcal{X}_2} q_{X_2}(x_2) \right) \sigma_{A_2} \otimes \sigma_B \\
\preceq& \frac{\mu}{\widehat{q}_{\min,1} \cdot \widehat{q}_{\min,2}} \sum_{x_1^2 \in \mathcal{X}^2} q_{X_1}(x_1) q_{X_2}(x_2) \sigma_{A_1} \otimes \sigma_{A_2} \otimes \sigma_B \\
=:& \frac{\mu}{\widehat{q}_{\min,1} \cdot \widehat{q}_{\min,2}} q_{X_1}(x_1) q_{X_2}(x_2) \otimes \widehat{\tau}^{x_1}_{A_1} \otimes \widehat{\tau}^{x_2}_{A_2} \otimes \sigma_B \ ,
\end{aligned} \tag{14}$$

where the first line is by normalization, the second is using $q_{X_1}(x_i)/\widehat{q}_{\min,1} \geq 1$ where $\widehat{q}_{\min,1} := \min_x q_{X_1}(x_1)$ and similarly indexed for 2, and the third is just defining $\widehat{\tau}^{x_1}_{A_1} := \sigma_{A_1}$ for all x_1 and again similarly or the second space.

Now combining the relations (13) and (14) defines a feasible solution $\left(\widehat{q}_{X_1},\, \widehat{q}_{X_2},\, \{\widehat{\tau}^{x_1}_{A_1}\}_{x_1 \in \mathcal{X}_1},\, \{\widehat{\tau}^{x_2}_{A_2}\}_{x_2 \in \mathcal{X}_2}, \sigma_B,\, \widehat{\rho}_{A_1^2 X_1^2 B}\right)$ for (12). Using that $I^{\downarrow,\varepsilon}_{\max}$ is a minimization problem,

$$\begin{aligned}
&I^{\downarrow,\varepsilon}_{\max}(:A_1^L X_1^L : B)_\rho \\
\leq& \log\left(\frac{\mu}{\widehat{q}_{\min,1} \cdot \widehat{q}_{\min,2}} \right) \\
\leq& I^{\downarrow,\varepsilon}_{\max}(:A_1^2 : B)_\rho + \log(\widehat{q}_{\min,1}) + \log(\widehat{q}_{\min,2}) \ .
\end{aligned}$$

Finally, since we had freedom in our choice of $\widehat{q}_{X_1}, \widehat{q}_{X_2}$, we may choose them to be the uniform distribution so that the minimum probability is $|\mathcal{X}_i|^{-1}$ for each i. □

Here we establish a weak converse for $I^{\downarrow,\varepsilon}_{\max}(:A_1^L : B)_\rho$.

a) Weak Asymptotic Equipartition Property:

Proposition 17. *Let $\rho_{A_1^L B} \in \mathrm{D}(A_1^L B)$. Then*

$$\lim_{\varepsilon\to 0} \lim_{n\to\infty} \left[\frac{1}{n} I^{\downarrow,\varepsilon}_{\max}(:A_1^L : B)_\rho \right] \ .$$

Proof. We provide the case where $L = 2$ for notational simplicity, but it generalizes straightforwardly. The proof is similar to that of the weak converse of bipartite max mutual information in [13].

$$\begin{aligned}
&\min_{\widetilde{\rho} \in \mathcal{B}^\varepsilon(\rho_{A_1^2 B})} \min_{\tau_{A_i} \in D_{\max}(A_i)} \min_{\sigma_B} D_{\max}(\widetilde{\rho}_{A_1^2 B} \| \tau_{A_1} \otimes \tau_{A_2} \otimes \sigma_B) \\
\geq& \min_{\widetilde{\rho} \in \mathcal{B}^\varepsilon(\rho_{A_1^2 B})} \min_{\tau_{A_i} \in D(A_i)} \min_{\sigma_B} D(\widetilde{\rho}_{A_1^2 B} \| \tau_{A_1} \otimes \tau_{A_2} \otimes \sigma_B) \\
=& \min_{\widetilde{\rho} \in \mathcal{B}^\varepsilon(\rho_{A_1^2 B})} \min_{\tau_{A_i} \in D(A_i)} \min_{\sigma_B} \Big[D(\widetilde{\rho}_{A_1^2 B} \| \widetilde{\rho}_{A_1} \otimes \widetilde{\rho}_{A_2} \otimes \widetilde{\rho}_B) \\
&+ D(\widetilde{\rho}_{A_1} \| \tau_{A_1}) + D(\widetilde{\rho}_{A_2} \| \tau_{A_2}) + D(\widetilde{\rho}_B \| \tau_B) \Big] \\
=& \min_{\widetilde{\rho} \in \mathcal{B}^\varepsilon(\rho_{A_1^2 B})} D(\widetilde{\rho}_{A_1^2 B} \| \widetilde{\rho}_{A_1} \otimes \widetilde{\rho}_{A_2} \otimes \widetilde{\rho}_B) \\
=& \min_{\widetilde{\rho} \in \mathcal{B}^\varepsilon(\rho_{A_1^2 B})} I(A_1 : A_2 : B)_{\widetilde{\rho}} \\
=& \min_{\widetilde{\rho} \in \mathcal{B}^\varepsilon(\rho_{A_1^2 B})} H(A_1)_{\widetilde{\rho}} + H(A_2)_{\widetilde{\rho}} + H(B)_{\widetilde{\rho}} - H(A_1^2 B)_{\widetilde{\rho}} \ ,
\end{aligned}$$

where we used that $D_{\max} \geq D$ in the inequality, the chain rule for the Umegaki divergence for the first inequality, that $D(\rho\|\sigma) \geq 0$ with equality only if the two states are the same in the second equality and the chain rule for multipartite mutual information in the final equality.

Now we can just consider the n-copy version where we regularize and then take the error to zero:

$$\begin{aligned}
&\lim_{\varepsilon\to 0} \lim_{n\to\infty} \frac{1}{n} \Big[\min_{\widetilde{\rho}^n \mathcal{B}^\varepsilon(\rho^{\otimes n}_{A_1^2 B})} \min_{\tau_{A_i} \in D_{\max}(A_i)} \min_{\sigma_B} \\
&\qquad D_{\max}(\widetilde{\rho}_{A_1^2 B} \| \tau_{A_1} \otimes \tau_{A_2} \otimes \sigma_B) \Big] \\
\geq& \lim_{\varepsilon\to 0} \lim_{n\to\infty} \frac{1}{n} \min_{\widetilde{\rho}^n \mathcal{B}^\varepsilon(\rho^{\otimes n}_{A_1^2 B})} \Big[H(A_1^{\otimes n})_{\widetilde{\rho}^n} + H(A_2)_{\widetilde{\rho}^n} \\
&\qquad + H(B)_{\widetilde{\rho}^n} - H((A_1^2 B)^{\otimes n})_{\widetilde{\rho}^n} \Big] \\
=& \lim_{\varepsilon\to 0} \lim_{n\to\infty} \frac{1}{n} \min_{\widetilde{\rho}^n \mathcal{B}^\varepsilon(\rho^{\otimes n}_{A_1^2 B})} \Big[H(A_1^{\otimes n})_{\widetilde{\rho}^n} + H(A_2)_{\widetilde{\rho}^n} \\
&\qquad + H(B)_{\widetilde{\rho}^n} - H((A_1^2 B)^{\otimes n})_{\widetilde{\rho}^n} \Big] \\
=& \lim_{\varepsilon\to 0} \lim_{n\to\infty} \frac{1}{n} \min_{\widetilde{\rho}^n \in \mathcal{B}^\varepsilon(\rho^{\otimes n}_{A_1^2 B})} \Big[H(A_1^{\otimes n})_{\rho^{\otimes n}} + H(A_2)_{\rho^{\otimes n}} + H(B)_{\rho^{\otimes n}} \\
&\quad - H((A_1^2 B)^{\otimes n})_{\rho^{\otimes n}} \\
&\quad - 2n\varepsilon \left\{ \log|A_1| + \log|A_2| + \log|B| + \log|A_1^2 B| \right\} - 4g_2(\varepsilon) \Big] \\
=& \lim_{\varepsilon\to 0} \Big[I(A_1 : A_2 : B)_\rho \\
&\quad - 2\varepsilon \left\{ \log|A_1| + \log|A_2| + \log|B| + \log|A_1^2 B| \right\} \Big] \\
=& I(A_1 : A_2 : B)_\rho \ ,
\end{aligned}$$

where the inequality is the AFW inequality. This completes the proof. □