Quantum Communication

Lecture Notes

Rahul Jain

These notes can always be improved. Any comments are very much appreciated. I would like to thank all the students of NUS QT5104 in 2024/25 Semester II who have contributed to scribing these notes.

Latest update, April 2026

Contents

$\mathrm{F}(\cdot,\cdot)$	fidelity
$\mathrm{D}(\cdot\|\|\cdot)$	relative-entropy
$\mathrm{D}_{\max}(\cdot\|\|\cdot)$	max relative-entropy
$\mathrm{D}_H(\cdot\|\|\cdot)$	hypothesis testing relative-entropy
$\mathrm{H}(\cdot)$	entropy
$\mathrm{I}_{\max}(\cdot)$	max mutual information
$\rho \approx_\varepsilon \sigma$	$\mathrm{P}(\rho,\sigma) \leq \varepsilon$
$\langle\cdot,\cdot\rangle$	inner product
$\stackrel{\text{def}}{=}$	is defined as
$\mathcal{N}_{A\to B}$	a quantum channel from A to B
$\bar{z}$	complex conjugate of z
$X^\dagger$	adjoint/conjugate transpose of X
X^T	transpose (basis dependent)
$\mathbb{1}$	the identity matrix
span	vector span
Tr	matrix trace
Pr	probability
$\mathbf{1}\{x=y\}$	indicator function, 1 if $x=y$ and 0 otherwise, so that, for example, $\mathbf{1}\{x=y\}+\mathbf{1}\{x\neq y\}=1$
δ_{xy}	shorthand for $\mathbf{1}\{x=y\}$
log	logarithm; to base 2 here, i.e. $\log=\log_2$

Table 0.1: Some basic notation used in these notes.

psd	positive semi-definite
td	trace distance
ONB	orthonormal basis
iff	if and only if

Table 0.2: Some abbreviations used in these notes.

1

Source coding and the convex-split lemma

In this lecture, we study different types of *source-coding* protocols, along with a useful tool in quantum information theory known as the *convex-split lemma* which is helpful in designing such protocols. We begin by introducing some classical scenarios.

1.1 Classical setting

Let Alice receive an input random variable $X \in \{0,1\}^n$ which she intends to send to Bob. Alice sends a message M to Bob and Bob outputs $\hat{X}$. The aim of source coding is to send a small message M while keeping the probability of error ($\varepsilon \geq 0$) small, that is,

$$\Pr\left[X \neq \hat{X}\right] \leq \varepsilon.$$

If Alice and Bob share a random variable S before Alice's input arrives, the protocol becomes a *shared randomness-assisted* protocol.

To ensure that Alice is communicating her input (and not something else), we introduce a Referee R, which keeps a copy of X. The correctness of the protocol can then be rewritten as,

$$XX \approx_\varepsilon X\hat{X}.$$

More generally R and X are two correlated random variables and R need not equal X. The correctness of the protocol can then be rewritten as,

$$RX \approx_\varepsilon R\hat{X}.$$

As a further generalization (Figure 1.1), Alice and Bob could have their own random variables Y and Z, respectively, jointly

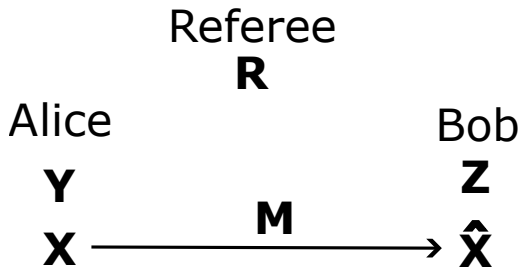

Figure 1.1: Schematics of classical protocol.

correlated with RX. The correctness of the protocol can then be rewritten as,

$$RYZX \approx_\varepsilon RYZ\hat{X}.$$

1.2 Quantum source coding

Quantum source coding is the quantum analogue of the classical case. The most general scenario is known as *quantum state redistribution.* At the beginning Alice (AC), Bob (B), and Referee (R) share a global pure state

$$\psi_{RABC} = |\psi_{RABC}\rangle \langle\psi_{RABC}| .$$

Alice wants to send the register C to Bob by transmitting message M (after some local operations at her end). If Alice and Bob start with some shared state, then we call this protocol an *entanglement-assisted* protocol. For example, Alice and Bob could share several EPR pairs in their registers $\hat{A}$ and $\hat{B}$. After transmission of the message M and Bob performing some local operations, the global state becomes $\varphi_{RAB\tilde{C}}$[1]. The goal of the protocol is to minimize the size (length) of the message M while keeping the error $\varepsilon \geq 0$ small, that is,

$$\psi_{RABC} \approx_\varepsilon \varphi_{RAB\tilde{C}}.$$

The symbol $\approx_\varepsilon$ between two states represents that the two states have the purified distance at most ε. Recall,

[1] Note that the state $\varphi_{RAB\tilde{C}}$ is not necessarily a pure state.

Definition 1.1 (Fidelity and purified distance)**.** *For quantum states ρ, σ, the fidelity between them is defined as,*

$$\mathrm{F}(\rho,\sigma) \stackrel{def}{=} \|\sqrt{\rho}\sqrt{\sigma}\|_1^2.$$

The purified distance between them is defined as

$$\mathrm{P}(\rho,\sigma) \stackrel{def}{=} \sqrt{1-\mathrm{F}(\rho,\sigma)}.$$

Definition 1.2 (ε-ball). *Let ρ be a state.*

$$\mathcal{B}^{\varepsilon}(\rho) := \{\rho' \mid \rho' \text{ is a state and } \mathrm{P}(\rho,\rho') \leq \varepsilon\}.$$

We say $\rho' \approx_{\varepsilon} \rho$ if $\rho' \in \mathcal{B}^{\varepsilon}(\rho)$.

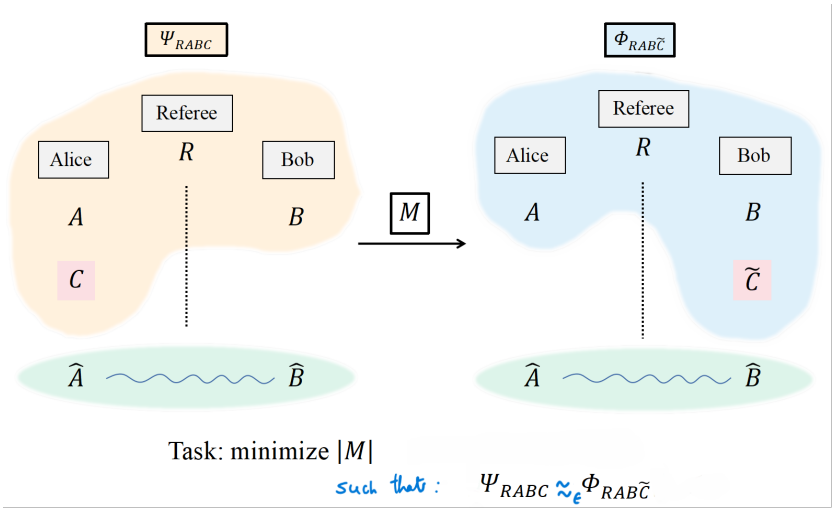

Figure 1.2: Quantum state redistribution.

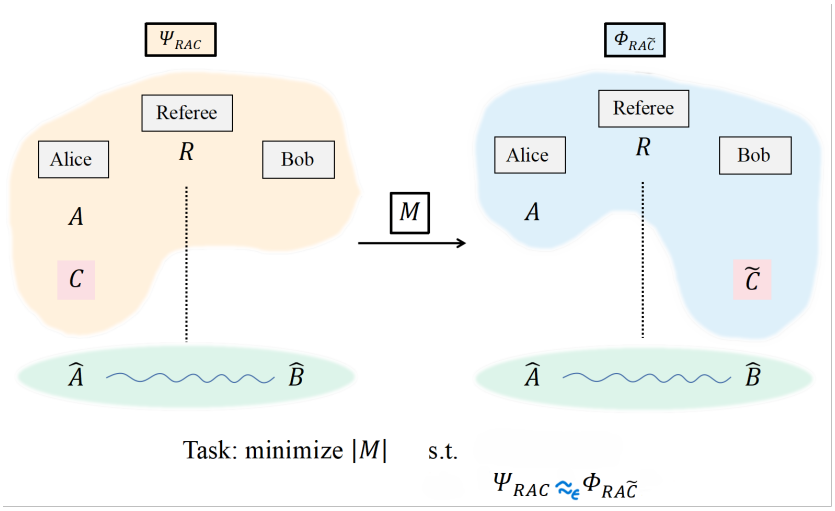

Figure 1.3: Quantum state splitting. A special case of state redistribution when the register B is missing.

An illustrative diagram is shown in Figure 1.2. We consider the following three types of variants.

1. *Quantum state splitting*: No register B at Bob's side (Figure 1.3).

2. *Quantum state merging*: No register A at Alice's side (Figure 1.4).

3. *Quantum state transfer*: Both A and B are absent (Figure 1.5).

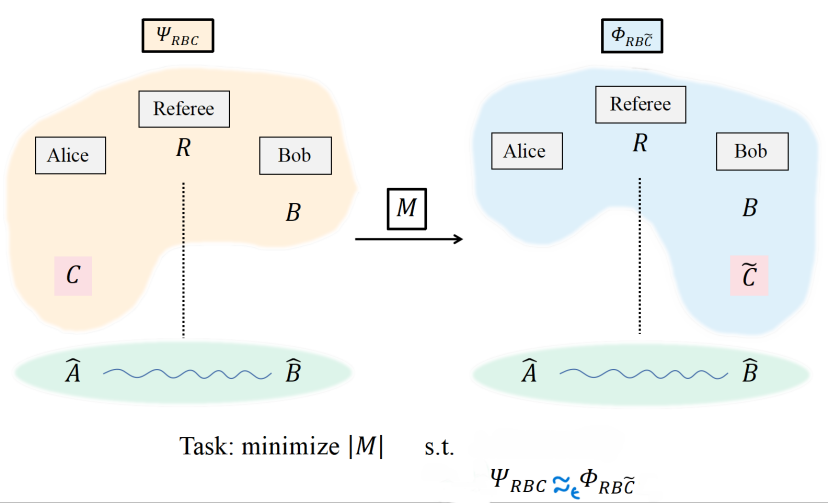

Figure 1.4: Quantum state merging. A special case of state redistribution when the register A is missing.

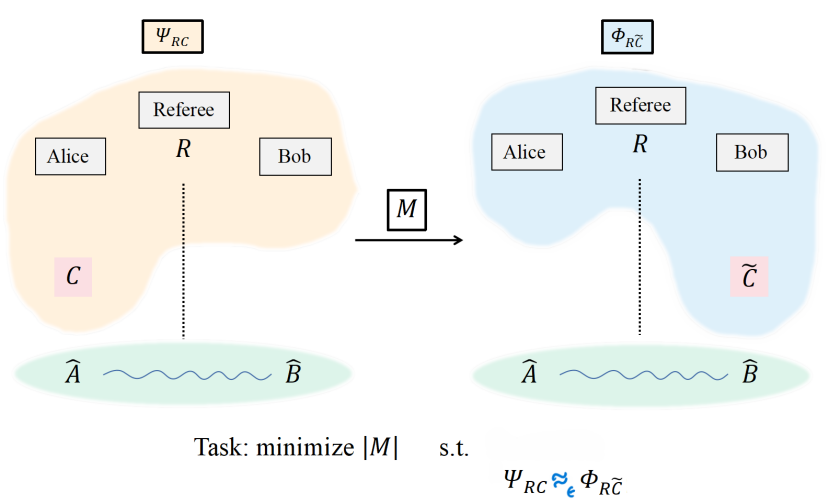

Figure 1.5: Quantum state transfer. A special case of state redistribution when the registers A, B are missing.

Table 1.1 summarizes the presence or absence of the registers and goals.

Protocol	Alice Has	Bob Has	Goal
State Splitting	A, C	*none*	Split AC and send C to Bob
State Merging	C	B	Merge C with B
State Transfer	C	*none*	Transfer C to Bob

Table 1.1: Comparison of quantum communication protocols depending on the registers available to Alice and Bob.

1.3 Convex-split lemma (CSL)

To design some protocols for the above tasks, we will introduce a widely used tool in quantum information theory known as the convex-split lemma [2]. We start with some notation, definitions, and facts.

[2] Anurag Anshu, Vamsi Krishna Devabathini, and Rahul Jain. Quantum communication using coherent rejection sampling. *Physical Review Letters*, 119(12), September 2017. DOI: 10.1103/physrevlett.119.120506. URL http://dx.doi.org/10.1103/PhysRevLett.119.120506

Quantum information primitives. We denote quantum registers by capital alphabetic A, B, etc. We denote density operators by $\rho_{AB} \in \mathcal{D}(AB), \sigma_A \in \mathcal{D}(A), \tau_B \in \mathcal{D}(B)$ and so on.

Definition 1.3 (Relative-entropy)**.** *For quantum states ρ, σ, the relative-entropy between them is defined as,*

$$\mathrm{D}(\rho\|\sigma) \stackrel{def}{=} \mathrm{Tr}\rho\log\rho - \mathrm{Tr}\rho\log\sigma.$$

Definition 1.4 (Max relative-entropy)**.** *Let ρ, σ be quantum states with $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$. The max relative-entropy between them is defined as,*

$$\mathrm{D}_{\max}(\rho\|\sigma) := \inf\left\{\lambda \in \mathbb{R} \mid \rho \leq 2^{\lambda}\sigma\right\}.$$

We have the following facts.

Fact 1.5. *For states ρ, σ, τ,*

$$\mathrm{D}(\rho\otimes\tau\|\sigma\otimes\tau) = \mathrm{D}(\rho\|\sigma).$$

Fact 1.6. *Let $A \geq 0, B \geq C$, then $\mathrm{Tr}AB \geq \mathrm{Tr}AC$.*

Fact 1.7 (Pinsker's inequality)**.** *For states ρ, σ,*

$$\mathrm{F}(\rho,\sigma) \geq 2^{-\mathrm{D}(\rho\|\sigma)}.$$

This implies,

$$1 - \mathrm{F}(\rho,\sigma) \leq (\ln 2)\cdot\mathrm{D}(\rho\|\sigma).$$

Fact 1.8 (Data processing inequality (DPI)). *For states ρ, σ and a CPTP map $\mathcal{E}$, we have*

$$\mathrm{D}(\rho\|\sigma) \geq \mathrm{D}(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)).$$

Consider states $\rho_{AB}, \sigma_{AB} \in \mathcal{D}(AB)$ and let $\mathcal{E}$ be the partial trace of system B. Then,

$$\mathrm{D}(\rho_{AB}\|\sigma_{AB}) \geq \mathrm{D}(\rho_A\|\sigma_A),$$

where $\rho_A = \mathrm{Tr}_B \rho_{AB}$ and $\sigma_A = \mathrm{Tr}_B \sigma_{AB}$.

We begin with the following lemma.

Lemma 1.9. *Let $\mu_1, \mu_2, \ldots, \mu_n, \theta$ be states and $\{p_1, p_2, \ldots p_n\}$ be a probability distribution. Let $\mu = \sum_i p_i \mu_i$ be the average state. Then*

$$\mathrm{D}(\mu\|\theta) = \sum_i p_i \left(\mathrm{D}(\mu_i\|\theta) - \mathrm{D}(\mu_i\|\mu)\right).$$

Proof. Consider,

$$\begin{aligned}
&\sum_i p_i(\mathrm{D}(\mu_i\|\theta) - \mathrm{D}(\mu_i\|\mu)) \\
&= \sum_i p_i(\mathrm{Tr}\mu_i \log \mu_i - \mathrm{Tr}\mu_i \log \theta - \mathrm{Tr}\mu_i \log \mu_i + \mathrm{Tr}\mu_i \log \mu) \\
&= \sum_i p_i(-\mathrm{Tr}\mu_i \log \theta + \mathrm{Tr}\mu_i \log \mu) \\
&= -\mathrm{Tr}\mu \log \theta + \mathrm{Tr}\mu \log \mu \\
&= \mathrm{D}(\mu\|\theta).
\end{aligned}$$

□

Now we are ready to state and prove the convex-split lemma.

Intuition: The convex-split lemma tells us that if a quantum state φ_{RC} is close (in max relative-entropy) to a product state $\varphi_R \otimes \sigma_C$, then by mixing φ_{RC} among n copies of σ_C, the resulting mixture becomes close to $\varphi_R \otimes \sigma_C^{\otimes n}$. The more copies n, the better the approximation. Please refer to Figure 1.6
The convex-split lemma guarantees that the state on the LHS (a convex combination where each term has the actual state φ_{RC} in position Referee and C_j and the state σ_C elsewhere) is approximately equal to the product state on the RHS, with fidelity at least $1 - \delta$.

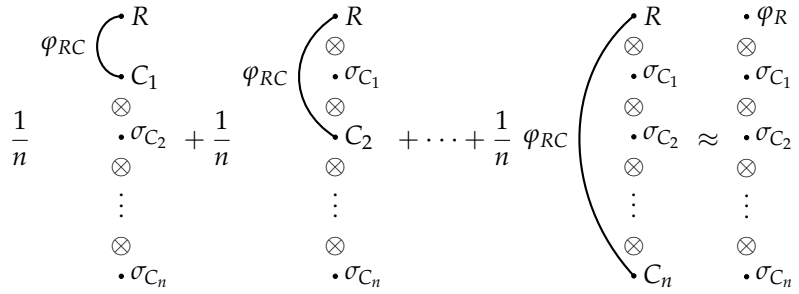

Figure 1.6: State equivalence for the convex-split lemma.

Lemma 1.10 (Convex-split lemma (CSL)). *Let $\varphi_{RC} \in \mathcal{D}(RC)$ and $\sigma_C \in \mathcal{D}(C)$ be states such that $\operatorname{supp}(\varphi_C) \subset \operatorname{supp}(\sigma_C)$. Let*

$$k := \mathrm{D}_{\max}(\varphi_{RC} \| \varphi_R \otimes \sigma_C).$$

Consider the following $(n+1)$-partite state,

$$\tau_{RC_1C_2\ldots C_n} := \frac{1}{n}\sum_{j=1}^{n} \varphi_{RC_j} \otimes \sigma_{C_1} \otimes \cdots \sigma_{C_{j-1}} \otimes \sigma_{C_{j+1}} \otimes \cdots \otimes \sigma_{C_n},$$

where $\varphi_{RC_j} = \varphi_{RC}$ and $\sigma_{C_j} = \sigma_C$ for all $j \in \{1, 2, \ldots, n\}$. Then,

$$\mathrm{D}(\tau_{RC_1C_2\ldots C_n} \| \varphi_C \otimes \sigma_{C_1} \otimes \cdots \otimes \sigma_{C_n}) \leq \log\left(1 + \frac{2^k}{n}\right).$$

Using Pinsker's inequality (Fact 1.7) we get,

$$\mathrm{F}(\tau_{RC_1C_2\ldots C_n}, \varphi_C \otimes \sigma_{C_1} \otimes \cdots \otimes \sigma_{C_n}) \geq \frac{1}{1 + \frac{2^k}{n}}.$$

In particular for $\delta > 0$ and $n = \left\lceil \frac{2^k}{\delta} \right\rceil$,

$$\mathrm{D}(\tau_{RC_1C_2\ldots C_n} \| \varphi_C \otimes \sigma_{C_1} \otimes \cdots \otimes \sigma_{C_n}) \leq \log(1+\delta),$$

and

$$\mathrm{F}(\tau_{RC_1C_2\ldots C_n}, \varphi_C \otimes \sigma_{C_1} \otimes \cdots \otimes \sigma_{C_n}) \geq 1 - \delta.$$

Proof. Denote the n copies of σ_C state by

$$\bar{\sigma} := \sigma_{C_1} \otimes \sigma_{C_2} \otimes \cdots \otimes \sigma_{C_n}.$$

Define a set of quantum states $\left\{\bar{\sigma}^{(-j)}\right\}$ for all $j \in [n]$ as

$$\bar{\sigma}^{(-j)} := \sigma_{C_1} \otimes \sigma_{C_2} \otimes \cdots \otimes \sigma_{C_{j-1}} \otimes \sigma_{C_{j+1}} \otimes \cdots \otimes \sigma_{C_n}.$$

Then we have,

$$\tau_{RC_1C_2\ldots C_n} = \frac{1}{n}\sum_{j=1}^{n} \varphi_{RC_j} \otimes \bar{\sigma}^{(-j)}.$$

Consider

$$\begin{aligned}
\tau_{RC_j} &= \frac{1}{n}\varphi_{RC_j} + \frac{n-1}{n}\left(\varphi_R \otimes \sigma_{C_j}\right) \\
&\leq \frac{2^k}{n}\left(\varphi_R \otimes \sigma_{C_j}\right) + \frac{n-1}{n}\left(\varphi_R \otimes \sigma_{C_j}\right) \qquad (k := \mathrm{D}_{\max}(\varphi_{RC}\|\varphi_R \otimes \sigma_C)) \\
&= \left(1 + \frac{2^k - 1}{n}\right)\left(\varphi_R \otimes \sigma_{C_j}\right).
\end{aligned}$$

Taking logarithm on both the sides (since log is operator monotonic),

$$\log \tau_{RC_j} \leq \log\left(1 + \frac{2^k - 1}{n}\right)\mathbb{I} + \log\left(\varphi_R \otimes \sigma_{C_j}\right). \qquad (1.1)$$

Consider

$$\begin{aligned}
&\mathrm{D}(\tau_{RC_1C_2\ldots C_n}\|\varphi_C \otimes \bar{\sigma}) \\
&= \frac{1}{n}\sum_{j=1}^{n}\left[\mathrm{D}(\varphi_{RC_j} \otimes \bar{\sigma}^{(-j)}\|\varphi_R \otimes \bar{\sigma}) - \mathrm{D}(\varphi_{RC_j} \otimes \bar{\sigma}^{(-j)}\|\tau_{RC_1C_2\ldots C_n})\right] && \text{(Lemma 1.9)} \\
&= \frac{1}{n}\sum_{j=1}^{n}\left[\mathrm{D}(\varphi_{RC_j}\|\varphi_R \otimes \sigma_{C_j}) - \mathrm{D}(\varphi_{RC_j} \otimes \bar{\sigma}^{(-j)}\|\tau_{RC_1C_2\ldots C_n})\right] && \text{(Fact 1.5)} \\
&\leq \frac{1}{n}\sum_{j=1}^{n}\left[\mathrm{D}(\varphi_{RC_j}\|\varphi_R \otimes \sigma_{C_j}) - \mathrm{D}(\varphi_{RC_j}\|\tau_{RC_j})\right] && \text{(Fact 1.8)} \\
&= \frac{1}{n}\sum_{j=1}^{n}\Big[\mathrm{Tr}\varphi_{RC_j}\log\varphi_{RC_j} - \mathrm{Tr}\varphi_{RC_j}\log(\varphi_R \otimes \sigma_{C_j}) \\
&\qquad\qquad - \mathrm{Tr}\varphi_{RC_j}\log\varphi_{RC_j} + \mathrm{Tr}\varphi_{RC_j}\log\tau_{RC_j}\Big] \\
&= \frac{1}{n}\sum_{j=1}^{n}\left[-\mathrm{Tr}\varphi_{RC_j}\log(\varphi_R \otimes \sigma_{C_j}) + \mathrm{Tr}\varphi_{RC_j}\log\tau_{RC_j}\right] \\
&\leq \frac{1}{n}\sum_{j=1}^{n}\Big[-\mathrm{Tr}\varphi_{RC_j}\log(\varphi_R \otimes \sigma_{C_j}) \\
&\qquad\qquad + \mathrm{Tr}\varphi_{RC_j}\left(\log\left(1 + \tfrac{2^k-1}{n}\right)\mathbb{I} + \log(\varphi_R \otimes \sigma_{C_j})\right)\Big] && \text{(Eq. (1.1), Fact 1.6)} \\
&= \log\left(1 + \frac{2^k - 1}{n}\right) \quad \leq \quad \log\left(1 + \frac{2^k}{n}\right).
\end{aligned}$$

Using Pinsker's inequality (Fact 1.7) we get,

$$\mathrm{F}^2(\tau_{RC_1C_2\ldots C_n}, \varphi_C \otimes \bar{\sigma}) \geq 2^{-\mathrm{D}(\tau_{RC_1C_2\ldots C_n}\|\varphi_C\otimes\bar{\sigma})} \geq 2^{-\log\left(1+\frac{2^k}{n}\right)} = \frac{1}{1+\frac{2^k}{n}}.$$

In particular for $\delta > 0$ and $n = \left\lceil \frac{2^k}{\delta} \right\rceil$,

$$\mathrm{D}(\tau_{RC_1C_2\ldots C_n}\|\varphi_C \otimes \sigma_{C_1} \otimes \cdots \otimes \sigma_{C_n}) \leq \log(1+\delta),$$

and

$$\mathrm{F}(\tau_{RC_1C_2\ldots C_n}, \varphi_C \otimes \sigma_{C_1} \otimes \cdots \otimes \sigma_{C_n}) \geq (1+\delta)^{-1} \geq 1-\delta.$$

□

2
Quantum state splitting

2.0.1 *Problem setting*

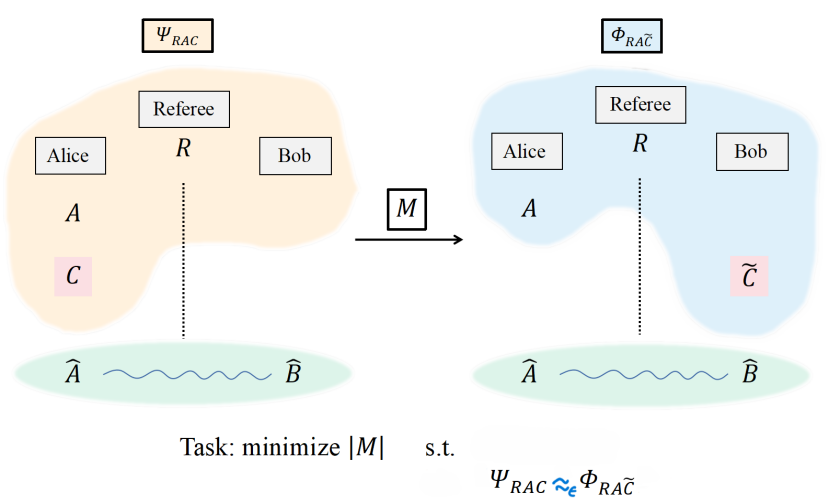

Figure 2.1: Quantum state splitting. The special case of state redistribution when the register B is missing.

2.1 *Protocol for state splitting*

Here we discuss a protocol for state splitting [1]. Let $\varepsilon \geq 0, \delta > 0$ be the error parameters. Let n be the smallest number such that,

$$\log n \geq \mathrm{I}^{\varepsilon}_{\max}(\dot{R} : C)_{\psi_{RC}} + 2\log\frac{1}{\delta}$$

where,

[1] Anurag Anshu, Vamsi Krishna Devabathini, and Rahul Jain. Quantum communication using coherent rejection sampling. *Physical Review Letters*, 119(12), September 2017. DOI: 10.1103/physrevlett.119.120506. URL http://dx.doi.org/10.1103/PhysRevLett.119.120506

Definition 2.1 (Max mutual information). *For a quantum*

state ψ_{RC}, the max mutual information is defined as,

$$\mathrm{I}_{\max}(R:C)_{\psi_{RC}} \stackrel{def}{=} \inf_{\sigma_C} \mathrm{D}_{\max}(\psi_{RC} \| \psi_R \otimes \sigma_C).$$

Definition 2.2 (Smooth max mutual information)**.** *Let $\varepsilon \geq 0$. For a quantum state ψ_{RC}, the ε-smooth max mutual information is defined as,*

$$\begin{aligned}\mathrm{I}^{\varepsilon}_{\max}(R:C)_{\psi_{RC}} &\stackrel{def}{=} \inf_{\psi'_{RC}:\mathrm{P}(\psi_{RC},\psi'_{RC})\leq\varepsilon} \mathrm{I}_{\max}(R:C)_{\psi'_{RC}} \\ &= \inf_{\psi'_{RC}:\mathrm{P}(\psi_{RC},\psi'_{RC})\leq\varepsilon} \inf_{\sigma_C} \mathrm{D}_{\max}(\psi'_{RC} \| \psi'_R \otimes \sigma_C).\end{aligned}$$

The ε-partially smooth max mutual information is defined as,

$$\mathrm{I}^{\varepsilon}_{\max}(\overset{\bullet}{R}:C)_{\psi_{RC}} \stackrel{def}{=} \inf_{\psi'_{RC}:\mathrm{P}(\psi_{RC},\psi'_{RC})\leq\varepsilon \,;\, \psi_R=\psi'_R} \inf_{\sigma_C} \mathrm{D}_{\max}(\psi'_{RC} \| \psi_R \otimes \sigma_C).$$

Let $\psi'_{RC} \approx_{\varepsilon} \psi_{RC}$ (with $\psi_R = \psi'_R$) and σ_C be such that,

$$\mathrm{I}^{\varepsilon}_{\max}(\overset{\bullet}{R}:C)_{\psi_{RC}} = \mathrm{D}_{\max}(\psi'_{RC} \| \psi_R \otimes \sigma_C).$$

We will need the following fundamental result in quantum information theory.

Fact 2.3 (Uhlmann's Theorem)**.** *Let $|\theta\rangle_{A'B}, |\gamma\rangle_{AB}$ be pure states. There exists an isometry $V : A' \to A$, such that:*

$$\mathrm{F}\big(V|\theta\rangle\langle\theta|V^{\dagger}, |\gamma\rangle\langle\gamma|\big) = \mathrm{F}(\theta_B, \gamma_B).$$

V is unitary if $A' \equiv A$.

Let $|\psi'\rangle_{RAC}$ be a purification ψ'_{RC} (guaranteed by Uhlmann's theorem) such that

$$\mathrm{F}(|\psi'\rangle\langle\psi'|_{RAC}, |\psi\rangle\langle\psi|_{RAC}) = \mathrm{F}(\psi'_{RC}, \psi_{RC}).$$

By the CSL (Lemma 1.10), we know that the purified distance between the reduced states (with Referee and Bob) between the LHS and RHS of Figure 2.3 is at most $\sqrt{\delta^2} = \delta$. Uhlmann's theorem (in the context of our protocol, $|\theta\rangle$ is the initial state on LHS and $|\gamma\rangle$ is the target state on RHS of Figure 2.2) guarantees the existence of an isometry V that Alice can apply to achieve $\mathrm{P}(V\theta V^{\dagger}, \gamma) \leq \delta$.

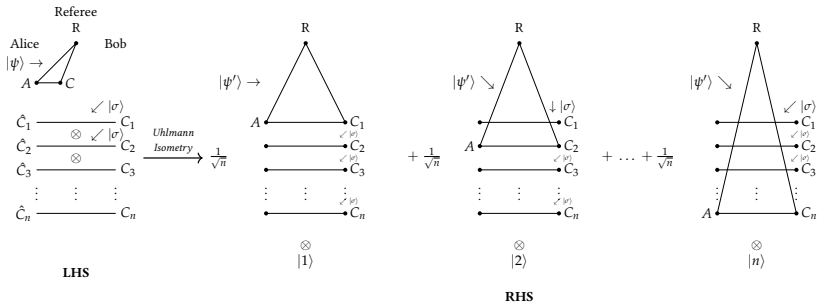

Figure 2.2: Quantum state splitting protocol: The left-hand side (LHS) shows the initial state with Alice holding registers A and C, and Alice and Bob sharing entangled pairs. The right-hand side (RHS) shows the state (that is close to the state after applying Uhlmann's isometry) which is a superposition with coefficients $1/\sqrt{n}$.

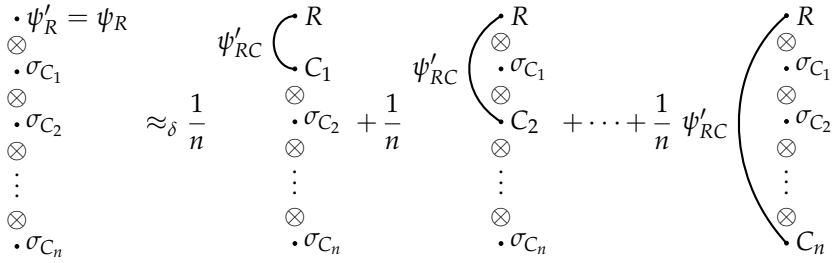

Figure 2.3: State Equivalence: The convex-split lemma guarantees that the state on the RHS (a convex combination where each term has the actual state ψ'_{RC} in position Referee and C_j and the state σ_C elsewhere) is close to the product state on the LHS, with purified distance at most δ.

2.1.1 *Protocol steps*

1. The protocol starts with the initial state $|\psi\rangle_{RAC}$ shared between Referee (R) and Alice (AC).
2. Alice and Bob share n i.i.d (independent and identically distributed) copies of a purification of σ_C, denoted as $|\sigma\rangle$ in registers $\hat{C}_i$ and C_i for $i \in \{1, 2, ..., n\}$.
3. Alice performs the Uhlmann isometry V on her registers based on the CSL.
4. Alice measures the index register j which collapses the superposition.
5. Alice sends the measurement outcome j to Bob using $\log n$ bits of communication.
6. Alice and Bob swap registers according to the value of j.

2.1.2 *Communication cost*

The classical communication cost is given by:

$$\log n = \mathrm{I}^{\varepsilon}_{\max}(\overset{\bullet}{R} : C)_{\psi_{RC}} + 2\log\frac{1}{\delta}.$$

2.1.3 Error analysis

We need the following property about partial traces in quantum systems.

Fact 2.4. *Let*

$$|\gamma\rangle_{A_1A_2B} = \sum_{i=1}^{n} \alpha_i |i\rangle_{A_1} |\gamma^i\rangle_{A_2B}$$

be a state (where $\sum_{i=1}^{n} |\alpha_i|^2 = 1$*). The reduced state on system B is given by:*

$$\gamma_B = \sum_{i=1}^{n} |\alpha_i|^2 \gamma_B^i.$$

This explains why the convex combination on the right side of Figure 2.3 appears when we trace out Alice's systems. Each term in the superposition contributes a component in the RHS of Figure 2.3 with probability $\frac{1}{n}$. The CSL guarantees that this mixture is (δ close to) the desired product state shown on the LHS of Figure 2.3.

At Step 4. in the protocol, when Alice measures the index register, Referee and Bob's combined state becomes (δ close to) a probabilistic mixture weighted by $\frac{1}{n}$.

When Alice sends her measurement outcome j to Bob, Bob knows which register contains the target state, allowing him to effectively "undo" the mixing effect and recover a state $\varphi_{RA\hat{C}} \approx_\delta \psi'_{RAC}$.

Since $\psi_{RAC} \approx_\varepsilon \psi'_{RAC}$,

$$\psi_{RAC} \approx_\varepsilon \psi'_{RAC} \approx_\delta \varphi_{RA\hat{C}}.$$

Using the triangle inequality for the purified distance, we get

$$\psi_{RAC} \approx_{\varepsilon+\delta} \varphi_{RA\hat{C}}.$$

3

State splitting - alternate protocol

3.1 Alternate protocol for state splitting

Here we discuss an alternate protocol for state splitting [1].

Let $\varepsilon \geq 0, \delta > 0$ be the error parameters. Let n be the smallest number such that,

$$\log n \geq \mathrm{I}^{\varepsilon}_{\max}(R:C)_{\psi_{RC}} + 2\log\frac{1}{\delta}.$$

Let $\psi'_{RC} \approx_{\varepsilon} \psi_{RC}$ and σ_C be such that,

$$\mathrm{I}^{\varepsilon}_{\max}(R:C)_{\psi_{RC}} = \mathrm{D}_{\max}(\psi'_{RC} \| \psi'_R \otimes \sigma_C).$$

Let $|\psi'\rangle_{RAC}$ be a purification ψ'_{RC} (guaranteed by Uhlmann's theorem) such that

$$\mathrm{F}(|\psi'\rangle\langle\psi'|_{RAC}, |\psi\rangle\langle\psi|_{RAC}) = \mathrm{F}(\psi'_{RC}, \psi_{RC}).$$

By the CSL (Lemma 1.10), we know that the purified distance between the reduced states (with Referee and Bob) is at most $\sqrt{\delta^2} = \delta$ (see Figure 3.2). Uhlmann's theorem (in the context of our protocol (Figure 3.1), $|\theta\rangle$ is the initial state on LHS and $|\gamma\rangle$ is the target state on RHS) guarantees the existence of an isometry V that Alice can apply to achieve $\mathrm{P}(V\theta V^{\dagger}, \gamma) \leq \delta$.

[1] Anurag Anshu, Vamsi Krishna Devabathini, and Rahul Jain. Quantum communication using coherent rejection sampling. *Physical Review Letters*, 119(12), September 2017. DOI: 10.1103/physrevlett.119.120506. URL http://dx.doi.org/10.1103/PhysRevLett.119.120506

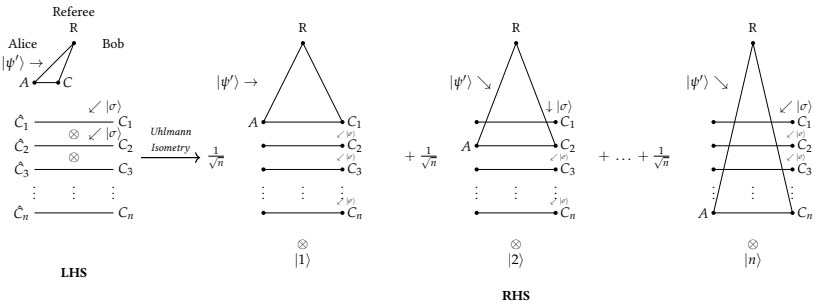

Figure 3.1: Quantum state splitting protocol: The left-hand side (LHS) shows the initial state with Alice holding registers A and C, and Alice and Bob sharing entangled pairs. The right-hand side (RHS) shows the state (that is close to the state after applying Uhlmann's isometry) which is a superposition with coefficients $1/\sqrt{n}$.

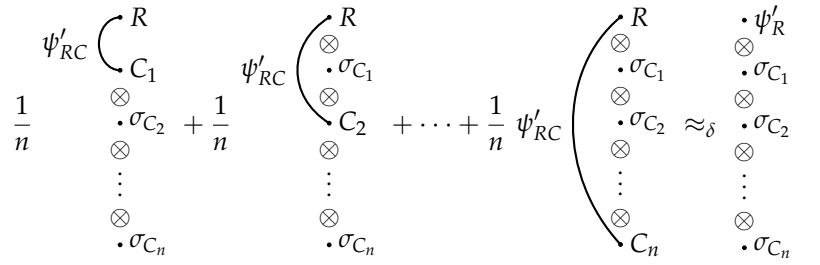

Figure 3.2: State Equivalence: The convex-split lemma guarantees that the state on the LHS (a convex combination where each term has the actual state ψ'_{RC} in position Referee and C_j and the state σ_C elsewhere) is approximately equal to the product state on the RHS, with purified distance at most δ.

3.1.1 *Protocol steps*

1. The protocol starts with the initial state $|\psi\rangle_{RAC}$ shared between Referee (R) and Alice (AC).
2. Alice and Bob share n i.i.d (independent and identically distributed) copies of a purification of σ_C, denoted as $|\sigma\rangle$ in registers $\hat{C}_i$ and C_i for $i \in \{1, 2, ..., n\}$.
3. Alice performs the Uhlmann isometry V on her registers based on the CSL.
4. Alice measures the index register j which collapses the superposition.
5. Alice sends the measurement outcome j to Bob using $\log n$ bits of communication.
6. Alice and Bob swap registers according to the value of j.

3.1.2 *Communication cost*

The classical communication cost is given by:

$$\log n = \mathrm{I}^{\varepsilon}_{\max}(R : C)_{\psi_{RC}} + 2 \log \frac{1}{\delta}.$$

3.1.3 Error analysis

Assume first that the protocol starts with the initial state $|\psi'\rangle_{RAC}$ shared between Referee (R) and Alice (AC). Fact 2.4 explains why the convex combination on the right side of Figure 3.2 appears when we trace out Alice's systems. Each term in the superposition contributes a component in the LHS of Figure 3.2 with probability $\frac{1}{n}$. The CSL guarantees that this mixture is (δ close to) the desired product state shown on the RHS of Figure 3.2.

At Step 4. in the protocol, when Alice measures the index register, Referee and Bob's combined state becomes (δ close to) a probabilistic mixture weighted by $\frac{1}{n}$.

When Alice sends her measurement outcome j to Bob, Bob knows which register contains the target state, allowing him to effectively "undo" the mixing effect and recover a state $\varphi'_{RA\hat{C}} \approx_\delta \psi'_{RAC}$.

Now assume that the protocol starts with the initial state ψ_{RAC} instead. Since $\psi_{RAC} \approx_\varepsilon \psi'_{RAC}$,

$$\psi_{RAC} \approx_\varepsilon \psi'_{RAC} \approx_\delta \varphi'_{RA\hat{C}} \approx_\varepsilon \varphi_{RA\hat{C}}.$$

The last approximation above follows using DPI for fidelity (and hence the purified distance) and noting that the entire communication protocol can be thought of as a CPTP map from the input state on the registers RAC to the output state on the registers $RA\hat{C}$.

Using the triangle inequality for the purified distance, we get

$$\psi_{RAC} \approx_{2\varepsilon+\delta} \varphi_{RA\hat{C}}.$$

Open question: Is it possible to get, for all $\delta > 0$, a protocol where

$$\psi_{RAC} \approx_{\varepsilon+\delta} \varphi_{RA\hat{C}}$$

with communication cost

$$\mathrm{I}^{\varepsilon}_{\max}(R:C)_{\psi_{RC}} + f(\delta)?$$

where f is any function of δ.

4

Converse bound for quantum state splitting

4.1 State splitting

Here, we want to prove the converse (lower bound) of state splitting [1]. The diagram for state splitting is shown in Figure 4.1 and the circuit diagram for the protocol for state splitting is shown in Figure 4.2.

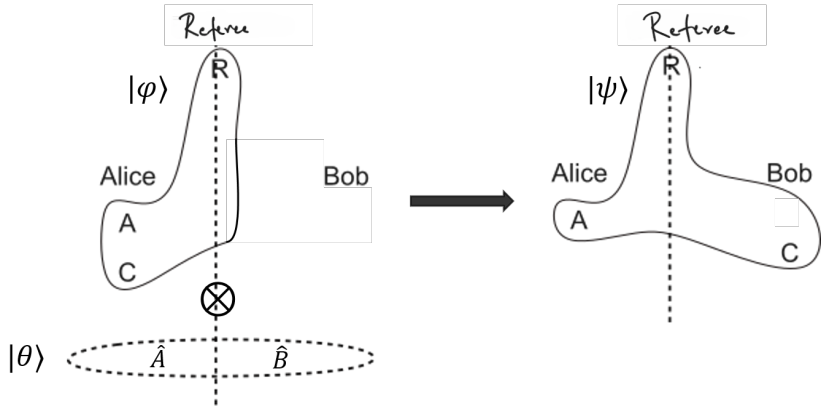

Figure 4.1: Diagram of state splitting with entanglement assistance. Alice splits the states in the system AC and Bob receives the state in the system C.

We need the following facts.

Fact 4.1. *For state* ρ_{AB},

$$\mathrm{I}_{\max}(A:B)_\rho \leq 2\min\{\log|A|, \log|B|\}$$

where $|A| = \dim(\mathrm{supp}(\rho_A))$.

[1] Mario Berta, Matthias Christandl, and Dave Touchette. Smooth entropy bounds on one-shot quantum state redistribution. *IEEE Transactions on Information Theory*, 62(3): 1425–1439, 2016

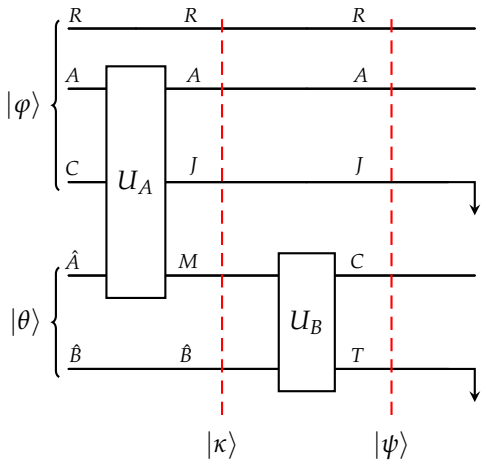

Figure 4.2: Circuit diagram of state splitting with entanglement assistance. Alice and Bob start with the state $|\varphi\rangle_{RAC}$ and share an entangled state $|\theta\rangle_{\hat{A}\hat{B}}$. U_A and U_B are unitary operators. $|\kappa\rangle$ is the state after U_A is applied and $|\psi\rangle$ is the final state after U_B is applied.

Fact 4.2. *For states ρ, σ and unitary U,*

$$\mathrm{D}_{\max}(\rho\|\sigma) = \mathrm{D}_{\max}(U\rho U^\dagger\|U\sigma U^\dagger).$$

4.1.1 Proof of the converse bound

Using Fact 4.1 (with $A \leftarrow R\hat{B}$ and $B \leftarrow M$),

$$\begin{aligned} 2\log|M| &\geq \mathrm{I}_{\max}(R\hat{B}:M)_\kappa \\ &= \min_{\sigma_M} \mathrm{D}_{\max}(\kappa_{R\hat{B}M}\|\kappa_{R\hat{B}}\otimes\sigma_M). \end{aligned}$$

Let ω_M be the state that minimizes the above. Notice that in Figure 4.2, the systems R and $\hat{B}$ remain unchanged after U_A, so $\kappa_{R\hat{B}} = \varphi_R \otimes \theta_{\hat{B}}$. From correctness of the protocol $\varphi_{RAC} \approx_\varepsilon \psi_{RAC}$. This implies (using DPI for purified distance) $\varphi_{RC} \approx_\varepsilon \psi_{RC}$. Hence,

$$\begin{aligned} 2\log|M| &\geq \mathrm{D}_{\max}(\kappa_{R\hat{B}M}\|\kappa_{R\hat{B}}\otimes\omega_M) \\ &= \mathrm{D}_{\max}(\kappa_{R\hat{B}M}\|\varphi_R\otimes\theta_{\hat{B}}\otimes\omega_M) \\ &= \mathrm{D}_{\max}(U_B\kappa_{R\hat{B}M}U_B^\dagger\|\varphi_R\otimes U_B(\theta_{\hat{B}}\otimes\omega_M)U_B^\dagger) && \text{(Fact 4.2)} \\ &= \mathrm{D}_{\max}(\psi_{RCT}\|\varphi_R\otimes\tau_{CT}) && (\tau_{CT} := U_B(\theta_{\hat{B}}\otimes\omega_M)U_B^\dagger) \\ &\geq \mathrm{D}_{\max}(\psi_{RC}\|\varphi_R\otimes\tau_C) && \text{(DPI for } \mathrm{D}_{\max}(\cdot\|\cdot)) \\ &\geq \mathrm{I}^\varepsilon_{\max}(\overset{\bullet}{R}:C)_\varphi. && \text{(Definition 2.2 and } \psi_R = \varphi_R \text{ and } \varphi_{RC} \approx_\varepsilon \psi_{RC}) \end{aligned}$$

4.1.2 Classical communication v/s quantum communication

The protocol for state-splitting presented in the previous lecture used entanglement-assisted classical communication.

Recall that using *superdense coding*, classical communication can be converted to quantum communication by reducing the communication cost by a factor of 2 (by sending one qubit per 2 classical bits). This justifies the factor 2 difference between the achievability and the converse bounds.

Using superdense coding, we get that if the communication is classical then,

$$\log |M| \geq \mathrm{I}^{\varepsilon}_{\max}(\overset{\bullet}{R} : C)_{\varphi}.$$

4.1.3 Putting together

Let cc-ss(ε) denote the optimal entanglement-assisted classical communication cost for state splitting with the purified distance between the input pure state (φ_{RC}) and the output state (ψ_{RC}) being at most ε. Then from the achievability (from the two protocols) and the converse bounds we get $\forall \varepsilon \geq 0, \delta > 0$:

$$\mathrm{I}^{\varepsilon+\delta}_{\max}(\overset{\bullet}{R} : C)_{\varphi} \leq \text{cc-ss}(\varepsilon+\delta) \leq \mathrm{I}^{\varepsilon}_{\max}(\overset{\bullet}{R} : C)_{\varphi} + 2\log\frac{1}{\delta},$$

$$\mathrm{I}^{2\varepsilon+\delta}_{\max}(R : C)_{\varphi} \leq \mathrm{I}^{2\varepsilon+\delta}_{\max}(\overset{\bullet}{R} : C)_{\varphi} \leq \text{cc-ss}(2\varepsilon+\delta) \leq \mathrm{I}^{\varepsilon}_{\max}(R : C)_{\varphi} + 2\log\frac{1}{\delta}.$$

The first inequality in the second expression above follows from the definitions.

4.2 Quantum state merging

State merging is basically the time reversal of state splitting. The diagram for state merging is shown in Figure 4.3 and the circuit diagram for the protocol for state merging is shown in Figure 4.4. Due to this, the achievability and the converse bounds are the same.

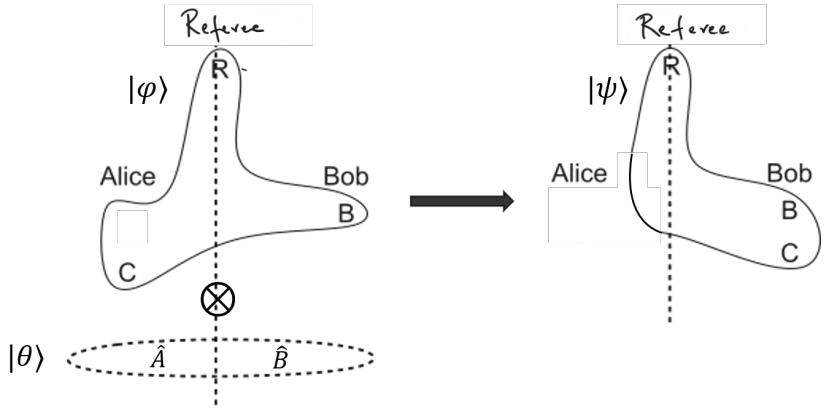

Figure 4.3: Diagram of state merging with entanglement assistance. Alice sends the state in the system C and Bob merges the state with his state in the system B.

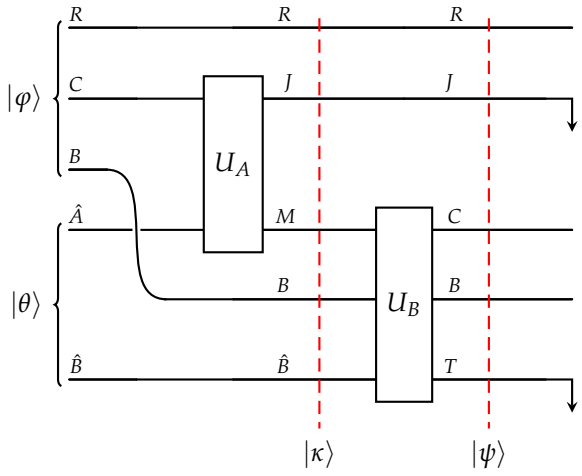

Figure 4.4: Circuit diagram of state merging with entanglement assistance. Alice and Bob start with the state $|\varphi\rangle_{RCB}$ and share an entangled state $|\theta\rangle_{\hat{A}\hat{B}}$. U_A and U_B are unitary gates. $|\kappa\rangle$ is the state after U_A is applied and $|\psi\rangle$ is the final state after U_B is applied.

5
Channel coding and position-based decoding

5.1 Introduction

In this lecture, we introduce the concept of channel coding and explore the *position-based decoding* (PBD) strategy [1].

[1] Anurag Anshu, Rahul Jain, and Naqueeb Ahmad Warsi. Building blocks for communication over noisy quantum networks. *IEEE Transactions on Information Theory*, 65(2):1287–1306, February 2019b. DOI: 10.1109/TIT.2018.2851297

5.2 Point-to-point quantum communication protocol

A communication protocol using a *point-to-point* quantum channel, involves two parties, Alice (the sender) and Bob (the receiver). Alice is allowed to use the quantum channel $\mathcal{N}_{A\to B}$ once to transmit a message m drawn uniformly from $[2^R]$. Bob decodes $\hat{m}$ and we want $\Pr[m = \hat{m}] \geq 1 - \varepsilon$, where ε is a small error tolerance. Alice and Bob are allowed to use a prior entangled state between them.

This communication strategy is known as an (R, ε) *entanglement-assisted code* for the quantum channel $\mathcal{N}_{A\to B}$. The primary objective is to determine the maximum achievable value of R, representing the number of reliable bits transmitted between Alice and Bob.

5.2.1 Position-based decoding (PBD)

We start with the following definition.

Definition 5.1 (Smooth hypothesis testing relative-entropy)**.** *Let ρ, σ be states and $\varepsilon \geq 0$. The* smooth hypothesis

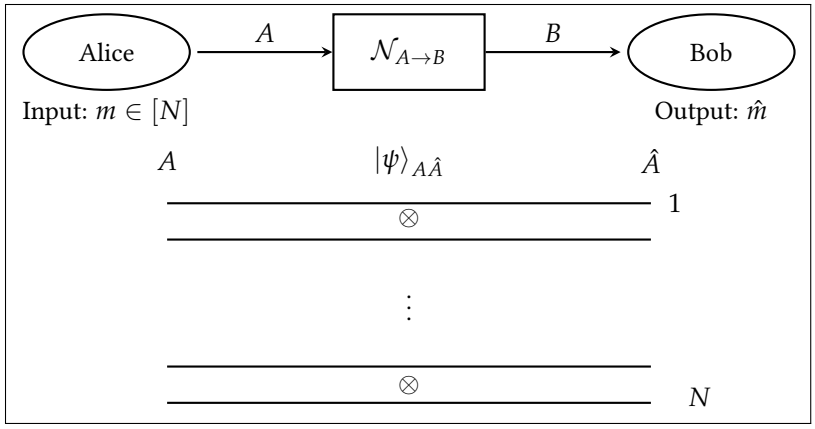

Figure 5.1: Illustration of the position-based decoding protocol.

testing relative-entropy *is defined as follows:*

$$\mathrm{D}_H^{\varepsilon}(\rho\|\sigma) \stackrel{def}{=} \max_{\substack{0\le T\le \mathbb{I} \\ \mathrm{Tr}(T\rho)\ge 1-\varepsilon}} \log\left(\frac{1}{\mathrm{Tr}\, T\sigma}\right).$$

Let $\varepsilon > 0$. Define,

$$R := \max_{|\varphi\rangle_{A\hat{A}}} \left\{ \mathrm{D}_H^{\varepsilon}(\theta_{B\hat{A}}\|\theta_B \otimes \varphi_{\hat{A}}) - \log\frac{1}{\varepsilon} \quad \middle| \quad \theta_{B\hat{A}} := \mathcal{N}_{A\to B}(\varphi_{A\hat{A}}) \right\}.$$

Let $|\psi\rangle_{A\hat{A}}$ be the state that achieves the maximum above. Let,

$$k := \mathrm{D}_H^{\varepsilon}(\theta_{B\hat{A}}\|\theta_B \otimes \psi_{\hat{A}}) \quad ; \quad N := \lfloor 2^R \rfloor = \lfloor \varepsilon \cdot 2^k \rfloor. \tag{5.1}$$

The communication protocol is described as follows (refer to Figure 5.1).

1. Alice and Bob start with $[N]$ i.i.d. copies of $|\psi\rangle_{A\hat{A}}$ shared between them.

2. On receiving input $m \in [N]$, Alice inserts the register A of the mth copy of $|\psi\rangle_{A\hat{A}}$ into the channel.

3. After receiving channel's output, Bob measures the state at his end $\tau^m_{B\hat{A}_1\ldots\hat{A}_N}$ according to the POVM

$$\{\Omega_j \mid j \in [N+1]\},$$

   as described below.

4. Bob outputs $\hat{m}$, which is the outcome of the measurement above.

5.2.2 Bob's decoding measurement

Let $T_{B\hat{A}}$ be the operator that achieves the maximum in the definition of $\mathrm{D}_H^{\varepsilon}(\theta_{B\hat{A}} \| \theta_B \otimes \psi_{\hat{A}})$. For all $j \in [N]$, define the following operators,

$$\Lambda(j) := \mathbb{I}_{\hat{A}_1} \otimes \cdots \otimes T_{B\hat{A}_j} \otimes \cdots \otimes \mathbb{I}_{\hat{A}_N} \quad ; \quad \Omega(j) := \Lambda^{-1/2} \Lambda(j) \Lambda^{-1/2},$$

where,

$$\Lambda := \sum_j \Lambda(j).$$

We can note that for all $j \in [N] : \ 0 \leq \Lambda(j) \leq \mathbb{I}$, $0 \leq \Omega(j)$, and

$$\mathrm{Tr}\ \tau^j_{B\hat{A}_1 \ldots \hat{A}_N} \Lambda(j) = \mathrm{Tr}\ \theta_{B\hat{A}} T_{B\hat{A}} \geq 1 - \varepsilon, \tag{5.2}$$

$$\forall i \neq j : \mathrm{Tr}\ \tau^i_{B\hat{A}_1 \ldots \hat{A}_N} \Lambda(j) = \mathrm{Tr}\ (\theta_B \otimes \psi_{\hat{A}}) T_{B\hat{A}} \leq 2^{-k}, \tag{5.3}$$

and

$$\begin{aligned} \sum_{j=1}^N \Omega(j) &= \sum_{j=1}^N \Lambda^{-1/2} \Lambda(j) \Lambda^{-1/2} \\ &= \Lambda^{-1/2} \left(\sum_{j=1}^N \Lambda(j) \right) \Lambda^{-1/2} \\ &= \Lambda^{-1/2} \Lambda \Lambda^{-1/2} \\ &= \Pi_\Lambda, \end{aligned}$$

where Π_Λ is the projection onto the support of Λ. Define,

$$\Omega(N+1) := \mathbb{I} - \Pi_\Lambda.$$

Hence we have a POVM $\{\Omega(j) \mid j \in [N+1]\}$ since

$$\sum_{j=1}^{N+1} \Omega(j) = \mathbb{I}.$$

5.2.3 Error analysis

We have the following useful fact.

Fact 5.2 (Hayashi-Nagaoka). *Let* $0 \leq S \leq \mathbb{I}$ *and* $T \geq 0$.

Then

$$\mathbb{I} - (S+T)^{-1/2} S (S+T)^{-1/2} \leq 2(\mathbb{I} - S) + 4T.$$

Define $S := \Lambda(m)$ and $T := \Lambda - \Lambda(m)$. Note that

$$\Omega(m) = (S+T)^{-1/2} S (S+T)^{-1/2}.$$

Let $\hat{M}$ denote the output random variable. The probability of incorrect decoding is given by,

$$\begin{aligned}
\Pr\left[\hat{M} \neq m\right] &= \mathrm{Tr}\left((\mathbb{I} - \Omega(m))\, \tau^m_{B\hat{A}_1 \ldots \hat{A}_N} \right) \\
&= \mathrm{Tr}\left((\mathbb{I} - (S+T)^{-1/2} S (S+T)^{-1/2})\, \tau^m_{B\hat{A}_1 \ldots \hat{A}_N} \right) \\
&\leq \mathrm{Tr}\left((2(\mathbb{I} - S) + 4T)\, \tau^m_{B\hat{A}_1 \ldots \hat{A}_N} \right) && \text{(Fact 5.2 and Fact 1.6)} \\
&= 2\mathrm{Tr}\left((\mathbb{I} - \Lambda(m))\, \tau^m_{B\hat{A}_1 \ldots \hat{A}_N} \right) + 4 \sum_{i \neq m} \mathrm{Tr}\left(\Lambda(i)\, \tau^m_{B\hat{A}_1 \ldots \hat{A}_N} \right) \\
&\leq 2\varepsilon + 4 \cdot N \cdot 2^{-k} && \text{(Eq. (5.2) and Eq. (5.3))} \\
&\leq 6\varepsilon. && \text{(Eq. (5.1))}
\end{aligned}$$

5.3 Classical-quantum channel

We consider a *shared-randomness* assisted encoding where information is transmitted through a *classical-quantum* (c-q) channel. Let $\varepsilon > 0$ and define,

$$R := \max_{p_{A\hat{A}}} \left\{ \mathrm{D}^{\varepsilon}_H(\theta_{B\hat{A}} \| \theta_B \otimes p_{\hat{A}}) - \log \frac{1}{\varepsilon} \quad \middle| \quad \theta_{B\hat{A}} := \mathcal{N}_{A \to B}(p_{A\hat{A}}) \right\}.$$

Let $p_{A\hat{A}}$ be the probability distribution that achieves the maximum above. Let,

$$k := \mathrm{D}^{\varepsilon}_H(\theta_{B\hat{A}} \| \theta_B \otimes p_{\hat{A}}) \quad ; \quad N := \lfloor 2^R \rfloor = \lfloor \varepsilon \cdot 2^k \rfloor.$$

The communication protocol is described as follows.

1. Alice and Bob start with $[N]$ i.i.d. copies of $p_{A\hat{A}}$ shared between them.
2. On receiving input $m \in [N]$, Alice inserts the register A of the mth copy of $p_{A\hat{A}}$ into the channel.
3. After receiving channel's output, Bob measures the state at his end $\tau^m_{B\hat{A}_1 \ldots \hat{A}_N}$ according to the POVM

   $$\{\Omega_j \mid j \in [N+1]\},$$

   as described below.

4. Bob outputs $\hat{m}$, which is the outcome of the measurement above.

5.3.1 Bob's decoding measurement and error analysis

Let $T_{B\hat{A}}$ be the operator that achieves the maximum in the definition of $\mathrm{D}_H^\varepsilon(\theta_{B\hat{A}} \| \theta_B \otimes p_{\hat{A}})$. The rest of the description of the measurement is the same as before. Also the error analysis is the same as before.

5.4 Classical-classical channel

We consider a shared-randomness assisted encoding where information is transmitted through a *classical-classical* (c-c) channel. Let $\varepsilon > 0$ and define,

$$R := \max_{p_{A\hat{A}}} \left\{ \mathrm{D}_H^\varepsilon(q_{B\hat{A}} \| q_B \otimes p_{\hat{A}}) - \log \frac{1}{\varepsilon} \quad \middle| \quad q_{B\hat{A}} := \mathcal{N}_{A\to B}(p_{A\hat{A}}) \right\}.$$

Let $p_{A\hat{A}}$ be the probability distribution that achieves the maximum above. Let,

$$k := \mathrm{D}_H^\varepsilon(q_{B\hat{A}} \| q_B \otimes p_{\hat{A}}) \quad ; \quad N := \lfloor 2^R \rfloor = \lfloor \varepsilon \cdot 2^k \rfloor.$$

The communication protocol is described as follows.

1. Alice and Bob start with $[N]$ i.i.d. copies of $p_{A\hat{A}}$ shared between them.
2. On receiving input $m \in [N]$, Alice inserts the register A of the mth copy of $p_{A\hat{A}}$ into the channel.
3. After receiving channel's output, Bob measures the random variable at his end $\tau^m_{B\hat{A}_1 \ldots \hat{A}_N}$ according to the POVM
$$\{\Omega_j \mid j \in [N+1]\},$$
as described below.
4. Bob outputs $\hat{m}$, which is the outcome of the measurement above.

5.4.1 Bob's decoding measurement and error analysis

Let $T_{B\hat{A}}$ be the operator that achieves the maximum in the definition of $\mathrm{D}_H^\varepsilon(q_{B\hat{A}} \| q_B \otimes p_{\hat{A}})$. The rest of the description of the measurement is the same as before. Also the error analysis is the same as before.

5.5 *Randomness unassisted coding for c-q and c-c channels for uniformly random input*

Let M, drawn uniformly from $[N]$, denote the input random variable. Let S denote the randomness used in the protocol. Let $\mathrm{err}(m, s)$ denote the error of the protocol on input m and randomness s.

From the arguments in previous sections, we have for all m:

$$\mathbb{E}_{s \leftarrow S}[\mathrm{err}(m, s)] \leq 6\varepsilon.$$

This implies,

$$\mathbb{E}_{s \leftarrow S}[\mathbb{E}_{m \leftarrow M}[\mathrm{err}(m, s)]] = \mathbb{E}_{m \leftarrow M}[\mathbb{E}_{s \leftarrow S}[\mathrm{err}(m, s)]] \leq 6\varepsilon.$$

Therefore,

$$\exists s : \quad \mathbb{E}_{m \leftarrow M}[\mathrm{err}(m, s)] \leq 6\varepsilon.$$

On fixing the randomness to s we get a randomness unassisted protocol with average error (for uniform input) at most ε.

6

Channel coding converse bound and asymptotic achievability

In the previous lecture, we looked at the achievability of quantum-quantum (q-q), classical-quantum (c-q), and classical-classical (c-c) channel coding in the one-shot setting. In this lecture, we first look at the converse bound of channel coding in the one-shot setting before moving on to the asymptotic analysis of channel coding [1].

[1] Anurag Anshu, Rahul Jain, and Naqueeb Ahmad Warsi. On the near-optimality of one-shot classical communication over quantum channels. *Journal of Mathematical Physics*, 60(1):012204, 01 2019a. DOI: 10.1063/1.5039796. URL https://doi.org/10.1063/1.5039796

6.1 Converse of channel coding

First, let's look at a general entanglement-assisted coding protocol as seen below in Figure 6.1 for message M uniformly drawn from $[N]$. Note that the Encode, $\mathcal{N}$ and Decode blocks are CPTP but not necessarily unitary.

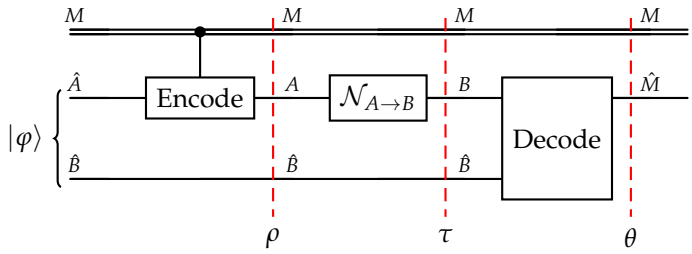

Figure 6.1: A general entanglement-assisted coding protocol.

With figure as a reference, let $|\rho\rangle_{MA\hat{B}T}$ be a purification of $\rho_{MA\hat{B}}$. Define,

$$\Lambda_{M\hat{M}} \stackrel{\text{def}}{=} \sum_{m=1}^{N} |m\rangle\langle m|_M \otimes |m\rangle\langle m|_{\hat{M}}.$$

We can see that this is effectively checking the probability of

$M = \hat{M}$. From the correctness of the protocol,

$$\mathrm{Tr}\Lambda_{M\hat{M}}\theta_{M\hat{M}} = \Pr\left[M = \hat{M}\right]_{\theta_{M\hat{M}}} \geq 1 - \varepsilon.$$

Define $\gamma_{\hat{M}} := \mathrm{Decode}(\tau_{\hat{B}} \otimes \tau_B)$. Note,

$$\mathrm{Tr}\Lambda_{M\hat{M}}\left(\theta_M \otimes \gamma_{\hat{M}}\right) = \Pr\left[M = \hat{M}\right]_{\theta_M \otimes \gamma_{\hat{M}}} = \frac{1}{N}.$$

From the definition of the hypothesis testing relative-entropy (Definition 5.1) we get,

$$\mathrm{D}_H^\varepsilon(\theta_{M\hat{M}} \| \theta_M \otimes \gamma_{\hat{M}}) \geq \log N. \tag{6.1}$$

Consider,

$$\begin{aligned}
&\max_{|\psi\rangle_{\tilde{A}A}} \mathrm{D}_H^\varepsilon(\mathcal{N}_{A\to B}\left(\psi_{\tilde{A}A}\right) \| \psi_{\tilde{A}} \otimes \mathcal{N}_{A\to B}\left(\psi_A\right)) \\
&\geq \mathrm{D}_H^\varepsilon(\mathcal{N}_{A\to B}(\rho_{M\hat{B}TA}) \| \rho_{M\hat{B}T} \otimes \mathcal{N}_{A\to B}\left(\rho_A\right)) && (\text{taking } \tilde{A} = M\hat{B}T) \\
&= \mathrm{D}_H^\varepsilon(\tau_{M\hat{B}TB} \| \rho_{M\hat{B}T} \otimes \tau_B) \\
&\geq \mathrm{D}_H^\varepsilon(\tau_{M\hat{B}B} \| \rho_{M\hat{B}} \otimes \tau_B) && (\text{DPI}) \\
&= \mathrm{D}_H^\varepsilon(\tau_{M\hat{B}B} \| \rho_M \otimes \rho_{\hat{B}} \otimes \tau_B) && (\rho_{M\hat{B}} = \rho_M \otimes \rho_{\hat{B}}) \\
&= \mathrm{D}_H^\varepsilon(\tau_{M\hat{B}B} \| \rho_M \otimes \tau_{\hat{B}} \otimes \tau_B) && (\rho_{\hat{B}} = \tau_{\hat{B}}) \\
&\geq \mathrm{D}_H^\varepsilon(\theta_{M\hat{M}} \| \theta_M \otimes \gamma_{\hat{M}}) && (\text{DPI and } \theta_M = \rho_M)) \\
&\geq \log N. && (\textbf{??})
\end{aligned}$$

Let's compare this with the achievability bound that we found in the previous lecture.

$$\log N \geq \max_{|\psi\rangle_{\tilde{A}A}} \mathrm{D}_H^{\varepsilon/6}(\mathcal{N}_{A\to B}\left(\psi_{\hat{A}A}\right) \| \psi_{\tilde{A}} \otimes \mathcal{N}_{A\to B}\left(\psi_A\right)) - \log\left(\frac{6}{\varepsilon}\right).$$

6.2 Asymptotic achievability

Suppose we have n uses of the channel available, that is $\mathcal{N}_{A^n\to B^n}^{\otimes n}$, and we wish to transmit message $M \in [N]$ using these channels with error at most ε. We use the following definition and Facts.

Definition 6.1 (Smooth max relative-entropy). *Let ρ, σ be states and $\varepsilon > 0$. The ε-smooth max relative-entropy between ρ and σ is defined as,*

$$\mathrm{D}_{\max}^\varepsilon(\rho\|\sigma) \stackrel{def}{=} \inf_{\rho' \approx_\varepsilon \rho} \mathrm{D}_{\max}(\rho'\|\sigma).$$

Fact 6.2 (Asymptotic convergence). *Let ρ, σ be quantum states. Then*

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n}\mathrm{D}_H^{\varepsilon}(\rho^{\otimes n}\|\sigma^{\otimes n}) = \lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n}\mathrm{D}_{\max}^{\varepsilon}(\rho^{\otimes n}\|\sigma^{\otimes n}) = \mathrm{D}(\rho\|\sigma).$$

Fact 6.3. *Let ρ_{AB} be a state. Then,*

$$\min_{\sigma_A,\sigma_B}\mathrm{D}(\rho_{AB}\|\sigma_A\otimes\sigma_B) = \mathrm{D}(\rho_{AB}\|\rho_A\otimes\rho_B) = \mathrm{I}(A:B)_{\rho_{AB}}.$$

Using the achievability protocol using PBD as seen before we have,

$$\begin{aligned}
&\frac{\log N}{n}\\
&\geq \frac{1}{n}\left[\max_{|\psi\rangle_{\hat{A}^n A^n}} \mathrm{D}_H^{\varepsilon/6}(\mathcal{N}_{A^n\to B^n}^{\otimes n}\left(\psi_{\hat{A}^n A^n}\right)\|\psi_{\hat{A}^n}\otimes\mathcal{N}_{A^n\to B^n}^{\otimes n}\left(\psi_{A^n}\right)) - \log\left(\frac{6}{\varepsilon}\right)\right]\\
&\geq \frac{1}{n}\left[\max_{|\psi\rangle_{\hat{A}A}^{\otimes n}} \mathrm{D}_H^{\varepsilon/6}(\mathcal{N}_{A^n\to B^n}^{\otimes n}\left(\psi_{\hat{A}A}^{\otimes n}\right)\|\psi_{\hat{A}}^{\otimes n}\otimes\mathcal{N}_{A^n\to B^n}^{\otimes n}\left(\psi_A^{\otimes n}\right)) - \log\left(\frac{6}{\varepsilon}\right)\right]\\
&= \max_{|\psi\rangle_{\hat{A}A}}\frac{1}{n}\left[\mathrm{D}_H^{\varepsilon/6}(\left(\mathcal{N}_{A\to B}\left(\psi_{\hat{A}A}\right)\right)^{\otimes n}\|\psi_{\hat{A}}^{\otimes n}\otimes\left(\mathcal{N}_{A\to B}\left(\psi_A\right)\right)^{\otimes n}) - \log\left(\frac{6}{\varepsilon}\right)\right]
\end{aligned}$$

Taking $\lim_{\varepsilon\to 0}$ and $\lim_{n\to\infty}$ above we get (note the second term goes to 0),

$$\begin{aligned}
&\max_{|\psi\rangle_{\hat{A}A}}\mathrm{D}(\mathcal{N}_{A\to B}\left(\psi_{\hat{A}A}\right)\|\psi_{\hat{A}}\otimes\mathcal{N}_{A\to B}\left(\psi_A\right)) && \text{(Fact 6.2)}\\
&= \max_{|\psi\rangle_{A\hat{A}}}\mathrm{I}(\hat{A}:B)_{\mathcal{N}_{A\to B}(\psi_{A\hat{A}})} && \text{(Fact 6.3)}\\
&\stackrel{\text{def}}{=} \mathrm{cap}(\mathcal{N}_{A\to B}).
\end{aligned}$$

Above, $\mathrm{cap}(\mathcal{N}_{A\to B})$ is the *entanglement-assisted classical capacity* of the channel $\mathcal{N}_{A\to B}$.

In the c-q and c-c case, we get the classical channel capacity,

$$\mathrm{cap}(\mathcal{N}_{A\to B}) \stackrel{\text{def}}{=} \max_{p_{A\hat{A}}}\mathrm{I}(\hat{A}:B)_{\mathcal{N}_{A\to B}(p_{A\hat{A}})}.$$

In the c-c case the above becomes (will be discussed in a weekly assignment) the more familiar form of the *Shannon capacity* of classical channels,

$$\mathrm{cap}(\mathcal{N}_{A\to B}) \stackrel{\text{def}}{=} \max_{X\leftarrow p_A\,;\,Y\leftarrow\mathcal{N}_{A\to B}(X)}\mathrm{I}(X:Y).$$

7
Channel-coding converse bound in the asymptotic limit

We study the converse bound for channel-coding in the asymptotic limit [1].

[1] Sumeet Khatri, Ludovico Lami, and Mark M. Wilde. *Principles of Quantum Communication Theory: A Modern Approach*. 2025. URL https://www.markwilde.com/PQCT-khatri-lami-wilde.pdf

7.1 Main theorem and proof

The setting is described in Figure 7.1. The channel $\mathcal{N}_{A^n\to B^n}^{\otimes n}$ is n independent uses of the channel $\mathcal{N}_{A\to B}$. Note that the output of Alice's encoding need not be a tensor product state.

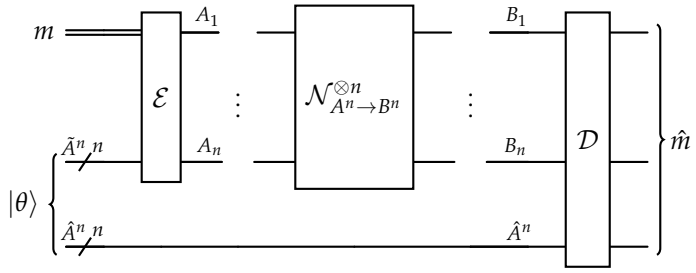

Figure 7.1: An entanglement-assisted protocol over multiple uses of a quantum channel $\mathcal{N}_{A\to B}$. Alice and Bob possess entanglement as a resource in the form of a quantum state $|\theta\rangle$ on the systems $\tilde{A}^n\hat{A}^n$. Alice encodes ($\mathcal{E}$) the message $m \in \{0,1\}^k$ into a quantum state on n quantum systems $A_1 \dots A_n$. Each system A_i is sent through the (same) channel $\mathcal{N}_{A\to B}$ to Bob's side. Bob performs the relevant measurements (decodes, $\mathcal{D}$) and obtains an estimate $\hat{m}$ of the message m sent by Alice.

The following theorem captures the asymptotic limit of the rate of classical information transfer using $\mathcal{N}_{A^n\to B^n}^{\otimes n}$.

Theorem 7.1.

$$\lim_{\substack{\varepsilon\to 0\\ n\to\infty}} \frac{1}{n} \sup_{|\psi\rangle_{\hat{A}^n A^n}} \mathrm{D}_H^{\varepsilon}(\mathcal{N}_{A^n\to B^n}^{\otimes n}(\psi_{\hat{A}^n A^n}) \| \psi_{\hat{A}^n} \otimes \mathcal{N}_{A^n\to B^n}^{\otimes n}(\psi_{A^n})) = \mathrm{cap}(\mathcal{N}_{A\to B}),$$

where

$$\mathrm{cap}(\mathcal{N}_{A\to B}) \stackrel{def}{=} \sup_{|\psi\rangle_{A\hat{A}}} \mathrm{I}(B:\hat{A})_{\mathcal{N}_{A\to B}(\psi_{A\hat{A}})}$$

is the entanglement-assisted classical capacity of the channel $\mathcal{N}_{A\to B}$.

Proof. We show the equality by showing the lower and upper bounds.

1. **Lower Bound.**

   Consider,

$$\begin{aligned}
&\lim_{\substack{\varepsilon\to 0\\ n\to\infty}} \frac{1}{n} \sup_{|\psi\rangle_{\hat{A}^n A^n}} \mathrm{D}^{\varepsilon}_{H}(\mathcal{N}^{\otimes n}_{A^n\to B^n}(\psi_{\hat{A}^n A^n}) \| \psi_{\hat{A}^n} \otimes \mathcal{N}^{\otimes n}_{A^n\to B^n}(\psi_{A^n})) \\
&\geq \lim_{\substack{\varepsilon\to 0\\ n\to\infty}} \frac{1}{n} \left[\sup_{|\psi\rangle_{\hat{A}A}} \mathrm{D}^{\varepsilon}_{H}(\mathcal{N}^{\otimes n}_{A^n\to B^n}\left(\psi^{\otimes n}_{\hat{A}A}\right) \| \psi^{\otimes n}_{\hat{A}} \otimes \mathcal{N}^{\otimes n}_{A^n\to B^n}\left(\psi^{\otimes n}_{A}\right)) \right] \\
&= \sup_{|\psi\rangle_{\hat{A}A}} \lim_{\substack{\varepsilon\to 0\\ n\to\infty}} \frac{1}{n} \left[\mathrm{D}^{\varepsilon}_{H}(\left(\mathcal{N}_{A\to B}\left(\psi_{\hat{A}A}\right)\right)^{\otimes n} \| \psi^{\otimes n}_{\hat{A}} \otimes \left(\mathcal{N}_{A\to B}\left(\psi_{A}\right)\right)^{\otimes n}) \right] \\
&= \sup_{|\psi\rangle_{\hat{A}A}} \mathrm{D}(\mathcal{N}_{A\to B}\left(\psi_{\hat{A}A}\right) \| \psi_{\hat{A}} \otimes \mathcal{N}_{A\to B}\left(\psi_{A}\right)) \qquad \text{(Fact 6.2)} \\
&= \sup_{|\psi\rangle_{A\hat{A}}} \mathrm{I}(\hat{A}:B)_{\mathcal{N}_{A\to B}(\psi_{A\hat{A}})} \qquad \text{(Fact 6.3)} \\
&= \mathrm{cap}(\mathcal{N}_{A\to B}).
\end{aligned}$$

2. **Upper Bound.**

   We need the following Fact relating the hypothesis-testing relative entropy and the relative entropy.

Fact 7.2. *Let* ρ, σ *be quantum states and* $\varepsilon > 0$. *We have*

$$(1-\varepsilon)\mathrm{D}^{\varepsilon}_{H}(\rho\|\sigma) \leq \mathrm{D}(\rho\|\sigma) + \mathrm{H}(\varepsilon).$$

Here

$$\mathrm{H}(x) \stackrel{def}{=} x\log(1/x) + (1-x)\log(1/(1-x))$$

is the binary entropy function.

Consider,

$$
\begin{aligned}
&(1-\varepsilon)\sup_{|\psi\rangle_{\hat{A}^nA^n}} \mathrm{D}_H^{\varepsilon}(\mathcal{N}^{\otimes n}_{A^n\to B^n}(\psi_{\hat{A}^nA^n})\|\psi_{\hat{A}^n}\otimes\mathcal{N}^{\otimes n}_{A^n\to B^n}(\psi_{A^n})) \\
&\leq \sup_{|\psi\rangle_{\hat{A}^nA^n}} \mathrm{D}(\mathcal{N}^{\otimes n}_{A^n\to B^n}(\psi_{\hat{A}^nA^n})\|\psi_{\hat{A}^n}\otimes\mathcal{N}^{\otimes n}_{A^n\to B^n}(\psi_{A^n}))+\mathrm{H}(\varepsilon) && \text{(Fact 7.2)} \\
&= \sup_{|\psi\rangle_{\hat{A}^nA^n}} \mathrm{D}(\tau_{\hat{A}^nB^n}\|\tau_{\hat{A}^n}\otimes\tau_{B^n})+\mathrm{H}(\varepsilon) && (\tau_{\hat{A}^nB^n}\overset{\text{def}}{=}\mathcal{N}^{\otimes n}_{A^n\to B^n}(\psi_{\hat{A}^nA^n})) \\
&= \sup_{\psi_{\hat{A}^nA^n}} \mathrm{I}(\hat{A}^n:B^n)_{\tau}+\mathrm{H}(\varepsilon) && \text{(Fact 6.3)} \\
&\leq n\cdot\mathrm{cap}(\mathcal{N}_{A\to B})+\mathrm{H}(\varepsilon). && \text{(Claim 1)}
\end{aligned}
$$

This implies,

$$
\begin{aligned}
&\lim_{\substack{\varepsilon\to 0\\ n\to\infty}}\frac{1}{n}\cdot\sup_{|\psi\rangle_{\hat{A}^nA^n}} \mathrm{D}_H^{\varepsilon}(\mathcal{N}^{\otimes n}_{A^n\to B^n}(\psi_{\hat{A}^nA^n})\|\psi_{\hat{A}^n}\otimes\mathcal{N}^{\otimes n}_{A^n\to B^n}(\psi_{A^n})) \\
&\leq \lim_{\substack{\varepsilon\to 0\\ n\to\infty}}\frac{1}{(1-\varepsilon)}\cdot\mathrm{cap}(\mathcal{N}_{A\to B})+\frac{1}{n(1-\varepsilon)}\cdot\mathrm{H}(\varepsilon) \\
&= \mathrm{cap}(\mathcal{N}_{A\to B}).
\end{aligned}
$$

□

Claim 1.

$$\sup_{\psi_{\hat{A}^nA^n}} \mathrm{I}(\hat{A}^n:B^n)_{\mathcal{N}^{\otimes n}_{A^n\to B^n}(\psi_{\hat{A}^nA^n})}\leq n\cdot\mathrm{cap}(\mathcal{N}_{A\to B}).$$

Proof. We show the result for $n=2$ case. The general case follows similarly. Refer to Figure 7.2.

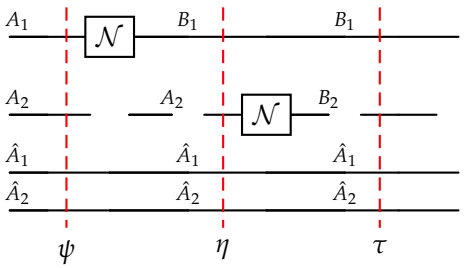

Figure 7.2: $\mathcal{N}\otimes\mathcal{N}=(\mathcal{I}\otimes\mathcal{N})(\mathcal{N}\otimes\mathcal{I})$.

We need the following Facts.

Fact 7.3 (Chain-rule for mutual information). *For a state* ρ_{ABC},

$$\mathrm{I}(A:BC)_{\rho}=\mathrm{I}(A:B)_{\rho}+\mathrm{I}(A:C\mid B)_{\rho}.$$

Fact 7.4. *For a state ρ_{AB},*

$$\mathrm{I}(A:B)_\rho \geq 0.$$

Fact 7.5.

$$\mathrm{cap}(\mathcal{N}_{A\to B}) = \sup_{\rho_{A\hat{A}}} \mathrm{I}(B:\hat{A})_{\mathcal{N}_{A\to B}(\rho_{A\hat{A}})}$$

We have,

$$\begin{aligned}
&\sup_{\psi} \mathrm{I}(\hat{A}_1\hat{A}_2 : B_1B_2)_\tau \\
&= \sup_{\psi} \left(\mathrm{I}(\hat{A}_1\hat{A}_2 : B_1)_\tau + \mathrm{I}(\hat{A}_1\hat{A}_2 : B_2 \mid B_1)_\tau\right) && \text{(Fact 7.3)} \\
&= \sup_{\psi} \left(\mathrm{I}(\hat{A}_1\hat{A}_2 : B_1)_\tau + \mathrm{I}(\hat{A}_1\hat{A}_2B_1 : B_2)_\tau - \mathrm{I}(B_1 : B_2)_\tau\right) && \text{(Fact 7.3)} \\
&\leq \sup_{\psi} \left(\mathrm{I}(\hat{A}_1\hat{A}_2 : B_1)_\tau + \mathrm{I}(\hat{A}_1\hat{A}_2B_1 : B_2)_\tau\right) && \text{(Fact 7.4)} \\
&\leq \sup_{\psi} \mathrm{I}(\hat{A}_1\hat{A}_2 : B_1)_\eta + \sup_{\eta} \mathrm{I}(\hat{A}_1\hat{A}_2B_1 : B_2)_\tau \\
&\leq \sup_{\theta} \mathrm{I}(\tilde{A}_1 : B_1)_{\mathcal{N}_{A_1\to B_1}(\theta)} + \sup_{\theta} \mathrm{I}(\tilde{A}_2 : B_2)_{\mathcal{N}_{A_2\to B_2}(\theta)} \\
&= 2\cdot \mathrm{cap}(\mathcal{N}_{A\to B}). && \text{(Fact 7.5)}
\end{aligned}$$

□

8

Quantum state redistribution

Quantum state redistribution is the generalization of state splitting and state merging that we have seen so far in the one-shot setting. The protocol starts with pure state $|\varphi\rangle_{RABC}$ shared between three parties where the Referee holds the register R, Alice holds the register A and C, and Bob holds the register B.

The goal of the protocol is to transfer the register C to Bob using the prior entangled state $|\theta\rangle$ in the registers $\hat{A}$ and $\hat{B}$ such that the output state ψ_{RABC} satisfies $\psi_{RABC} \approx_\varepsilon \varphi_{RABC}$ (as shown in Figure 8.1).

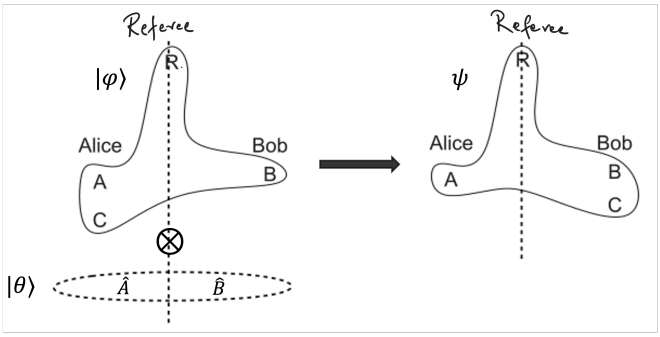

Figure 8.1: State redistribution problem statement where the initial state φ_{RABC} is shared between Alice and Bob. They follow an entanglement-assisted protocol to transfer the register C to Bob. The output state ψ_{RABC} is close to the initial state within purified distance ε.

8.1 Protocol for quantum state redistribution

Here we present a protocol for state redistribution. [1] The initial state and setup are very similar to that of the state splitting protocol. However, now Bob can also make use of the register B which helps Alice compress the communication required from her for Bob to correctly identify the state.

[1] Anurag Anshu, Rahul Jain, and Naqueeb Ahmad Warsi. A one-shot achievability result for quantum state redistribution. *IEEE Transactions on Information Theory*, 64(3):1425–1435, 2018. DOI: 10.1109/TIT.2017.2776112

8.1.1 Initial state and setup

Let $\varepsilon \geq 0, \delta > 0$. The initial state $|\varphi\rangle_{RABC}$ is shared between Referee (R), Alice (AC) and Bob (B). Let $\varphi'_{RBC} \approx_\varepsilon \varphi_{RBC}$ (with $\varphi_{RB} = \varphi'_{RB}$) and σ_C be states. Let

$$\log n := \mathrm{D}_{\max}(\varphi'_{RBC} \| \varphi_{RB} \otimes \sigma_C) + 2 \log \frac{1}{\delta}.$$

Alice and Bob shares n copies of a purification of the quantum state σ_C, in the registers $\hat{C}_i$ and C_i for $i \in \{1, 2, ..., n\}$.

Let $|\varphi'\rangle_{RABC}$ be a purification φ'_{RBC} (guaranteed by Uhlmann's theorem) such that

$$\mathrm{F}(|\varphi'\rangle\langle\varphi'|_{RABC}, |\varphi\rangle\langle\varphi|_{RABC}) = \mathrm{F}(\varphi'_{RBC}, \varphi_{RBC}).$$

8.1.2 Protocol steps

The protocol consists of the following steps.

1. Alice performs the Uhlmann isometry on her registers based on the convex split lemma (CSL). This creates a state (close to RHS of the figure) where each term in the superposition has amplitude $\frac{1}{\sqrt{n}}$ as shown in Figure 8.2.

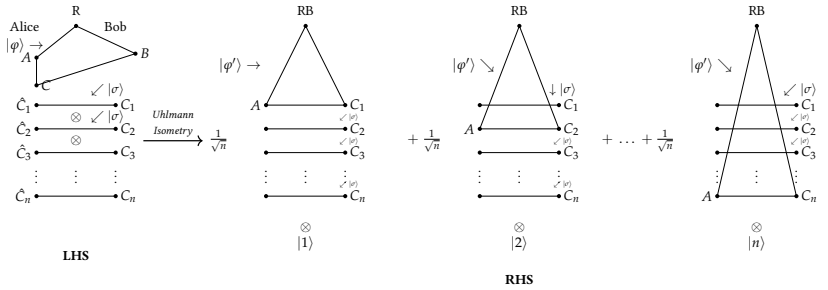

Figure 8.2: State redistribution protocol - Alice's operations: The left-hand side (LHS) shows the initial state with Alice holding registers A and C, and Alice and Bob sharing entangled pairs. The right-hand side (RHS) shows the state (that is close to the state after applying Uhlmann's isometry) which is a superposition with coefficients $1/\sqrt{n}$.

2. Alice measures her index register to find some j.

   If she could communicate j to Bob, Bob would be able to pick up the register C_j obtaining the desired state. The problem is that the number of bits required to communicate j is large (around $\mathrm{D}_{\max}(\varphi'_{RBC} \| \varphi_{RB} \otimes \sigma_C)$).

3. Alice divides $[n]$ into n/k blocks of size k and communicates the block number of j to Bob, where,

$$\log k := \mathrm{D}_H^{2\varepsilon}(\varphi'_{BC} \| \varphi_B \otimes \sigma_C) - \log \frac{1}{\delta}.$$

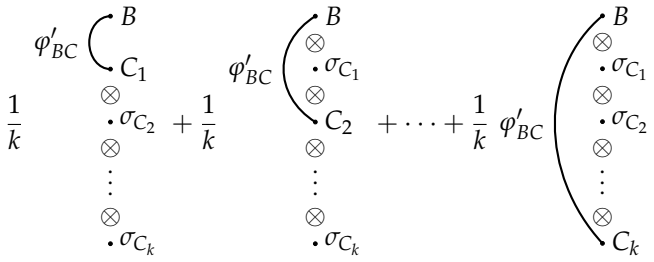

Figure 8.3: State redistribution protocol - Bob's operations: The state with Bob is (close to) a convex combination of states once Alice communicates the block number to Bob. Bob detects j inside the block of size k using PBD.

4. Upon receiving the block number from Alice, the quantum state left with Bob is (close to) a convex mixture as shown in Figure 8.3.

   Bob makes use of his quantum side information (B), employing position-based decoding (PBD) strategy to find j inside the block.

8.1.3 Communication cost

We use the following Fact.

Fact 8.1. *Let $\varepsilon \geq \delta \geq 0$. Let $\rho \approx_\delta \rho'$ and σ be state. Then,*

$$\mathrm{D}_H^{\varepsilon-\delta}(\rho\|\sigma) \leq \mathrm{D}_H^{\varepsilon}(\rho'\|\sigma).$$

The communication cost is:

$$\begin{aligned}
&\log \frac{n}{k} = \log n - \log k \\
&= \min_{\sigma_C, \varphi'_{RBC} \approx_\varepsilon \varphi_{RBC}, \varphi_{RB} = \varphi'_{RB}} \left[\mathrm{D}_{\max}(\varphi'_{RBC}\|\varphi_{RB}\otimes\sigma_C) - \mathrm{D}_H^{2\varepsilon}(\varphi'_{BC}\|\varphi_B\otimes\sigma_C)\right] + 3\log\frac{1}{\delta} \\
&\leq \min_{\sigma_C, \varphi'_{RBC} \approx_\varepsilon \varphi_{RBC}, \varphi_{RB} = \varphi'_{RB}} \left[\mathrm{D}_{\max}(\varphi'_{RBC}\|\varphi_{RB}\otimes\sigma_C) - \mathrm{D}_H^{\varepsilon}(\varphi_{BC}\|\varphi_B\otimes\sigma_C)\right] + 3\log\frac{1}{\delta}.
\end{aligned}$$

The inequality above follows from Fact 8.1 and $\varphi_{BC} \approx_\varepsilon \varphi'_{BC}$.

8.1.4 Error analysis

Using CSL and Uhlmann's theorem, we can argue as before that at Step 2. in the protocol, when Alice measures the index register, Referee and Bob's combined state becomes (δ close to) a probabilistic mixture weighted by $\frac{1}{n}$ (refer to RHS of Figure 8.2).

When Alice sends the block number of her measurement outcome j to Bob, Bob knows which block contains the target state. The quantum state left with Bob is a convex mixture as shown in Figure 8.3. Bob employs PBD to find out j inside the block.

Let $\hat{\gamma}$ be the overall state produced in the protocol after Alice applies Uhlmann's isometry and measures the index register. Let γ be the state on the RHS of Figure 8.2, on measuring the index register. From CSL and Uhlmann's theorem, we know $\gamma \approx_\delta \hat{\gamma}$.

We use the following Fact [2].

Fact 8.2 (Accurate measurement lemma). *Let*

$$\rho_{AB} = \sum_i p_i |i\rangle\langle i|_A \otimes \rho_B^i$$

be a c-q state. Let $\mathcal{M}$ be a measurement performed on the register B to produce an outcome in the register $\hat{A}$. Let $\sigma_{AB\hat{A}}$ be the resulting state after the measurement. Then,

$$\mathrm{F}(\rho_{AB}, \sigma_{AB}) \geq (\Pr\left[\hat{A} = A\right]_\sigma)^2.$$

From the arguments used in analysis of PBD we know that there is measurement that Bob can perform on the state γ in Figure 8.3 to output candidate $\hat{j}$ such that

$$\Pr\left[J = \hat{J}\right] \geq 1 - (4\varepsilon + 4\delta).$$

Let γ' be the resulting state after the measurement. From Fact 8.2 we get,

$$\mathrm{F}(\gamma', \gamma) \geq (1 - (4\varepsilon + 4\delta))^2 \geq 1 - (8\varepsilon + 8\delta).$$

This implies $\gamma' \approx_{\sqrt{8\varepsilon+8\delta}} \gamma$.

If Alice and Bob were using the state γ (and index J to output), the output state would have been exactly $|\varphi'\rangle_{RABC}$. If the output state $\hat{\varphi}_{RABC}$ was produced using $\hat{J}$ in γ', then using DPI we would have,

$$\hat{\varphi}_{RABC} \approx_{\sqrt{8\varepsilon+8\delta}} \varphi'_{RABC}$$

In the protocol the actual output ψ_{RABC} is produced using the PBD measurement on $\hat{\gamma}$ instead of γ. Since $\hat{\gamma} \approx_\delta \gamma$, using DPI we get,

$$\hat{\varphi}_{RABC} \approx_\delta \psi_{RABC}.$$

[2] Anurag Anshu, Rahul Jain, and Naqueeb Ahmad Warsi. A one-shot achievability result for quantum state redistribution. *IEEE Transactions on Information Theory*, 64(3):1425–1435, 2018. DOI: 10.1109/TIT.2017.2776112

Using the triangle inequality for the purified distance we get,

$$\psi_{RABC} \approx_{\delta+\sqrt{8\varepsilon+8\delta}} \varphi'_{RABC}.$$

Since $\varphi_{RABC} \approx_{\varepsilon} \varphi'_{RABC}$, again using the triangle inequality we finally get (since $\varepsilon + \delta + \sqrt{8\varepsilon + 8\delta} \leq 4\sqrt{\varepsilon + \delta}$),

$$\psi_{RABC} \approx_{4\sqrt{\varepsilon+\delta}} \varphi_{RABC}.$$

Open question: Is it possible to get a protocol for state redistribution with error ε and communication that is a polynomial in $\mathrm{I}(R : C \mid B)$ and $\frac{1}{\varepsilon}$?

This is known in the classical case for the analogous task.

9

Converse bound and asymptotics for state redistribution

We will first establish the converse bound, both in the one shot case and the asymptotic limit [1]. Then we will recall the one shot achievability bound and look at its asymptotic limit.

[1] Mario Berta, Matthias Christandl, and Dave Touchette. Smooth entropy bounds on one-shot quantum state redistribution. *IEEE Transactions on Information Theory*, 62(3): 1425–1439, 2016

9.1 Circuit for state redistribution

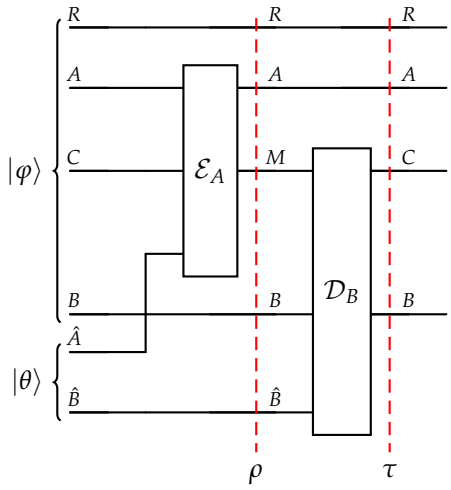

Figure 9.1: State redistribution circuit. Alice's registers are $A, C, \hat{A}$ and Bob's registers are $B, \hat{B}$. Both φ, θ are pure states.

The circuit diagram for a protocol for state redistribution is given above. Alice holds the registers $A, C, \hat{A}$ and Bob holds the registers $B, \hat{B}$ and there is the referee register R. The state $|\varphi\rangle_{RABC}$ is pure. The state $|\theta\rangle$ describing $\hat{A}, \hat{B}$ is also pure and it is taken to be pre-shared entanglement between the two parties. An encoding channel $\mathcal{E}_A$ works on $A, C, \hat{A}$ i.e. on Alice's registers and the outcome are the registers A, M. The quantum register M is then sent to Bob and Bob applies

a decoding channel $\mathcal{D}_B$ on $M, B, \hat{B}$. The goal is that the state after the decoding, $\tau_{RACB} \approx_\varepsilon \varphi_{RACB}$, whilst minimizing the size of the register M to be communicated to Bob. We now intend to find a lower bound on $\log |M|$.

We need the following Fact.

Fact 9.1.

$$\mathrm{D}_{\max}(\rho_A \| \sigma_A) + 2\log|B| \geq \mathrm{D}_{\max}(\rho_{AB} \| \sigma_A \otimes \mu_B).$$

where $\mu_B := \mathbb{1}_B / |B|$.

9.2 A first bound

Let σ_B be the state such that,

$$\mathrm{I}_{\max}(R:B)_\varphi = \mathrm{D}_{\max}(\varphi_{RB} \| \varphi_R \otimes \sigma_B).$$

Consider,

$$\begin{aligned}
&\mathrm{I}_{\max}(R:B)_\varphi + 2\log|M| \\
&= \mathrm{D}_{\max}(\varphi_{RB} \| \varphi_R \otimes \sigma_B) + 2\log|M| \\
&= \mathrm{D}_{\max}(\varphi_{RB} \otimes \theta_{\hat{B}} \| \varphi_R \otimes \sigma_B \otimes \theta_{\hat{B}}) + 2\log|M| \\
&= \mathrm{D}_{\max}(\rho_{RB\hat{B}} \| \varphi_R \otimes \sigma_B \otimes \theta_{\hat{B}}) + 2\log|M| && (\rho_{RB\hat{B}} = \varphi_{RB} \otimes \theta_{\hat{B}}) \\
&\geq \mathrm{D}_{\max}(\rho_{RB\hat{B}M} \| \varphi_R \otimes \sigma_B \otimes \theta_{\hat{B}} \otimes \mu_M) && (\text{Fact 9.1 and } \mu_M := \mathbb{I}_M / |M|) \\
&\geq \mathrm{D}_{\max}(\tau_{RBC} \| \varphi_R \otimes \eta_{BC}) && (\text{DPI and } \eta_{BC} := \mathcal{D}_B(\sigma_B \otimes \theta_{\hat{B}} \otimes \mu_M)) \\
&\geq \mathrm{I}^\varepsilon_{\max}(\overset{\bullet}{R} : BC)_\varphi. && (\tau_{RBC} \approx_\varepsilon \varphi_{RBC} \,;\, \tau_R = \varphi_R)
\end{aligned}$$

This implies,

$$\log|M| \geq \frac{\mathrm{I}^\varepsilon_{\max}(\overset{\bullet}{R} : BC)_\varphi - \mathrm{I}_{\max}(R:B)_\varphi}{2}.$$

9.3 Strengthening the bound

One of the terms in the bound is a smoothed max-information and the other a non-smoothed version. It would be good to have only smoothened versions - and this would also mean strengthening the inequality in this case.

Claim 2.

$$2\log|M| \geq \mathrm{I}^{2\varepsilon}_{\max}(\overset{\bullet}{R} : BC)_\varphi - \mathrm{I}^{\varepsilon}_{\max}(\overset{\bullet}{R} : B)_\varphi.$$

Proof. Let σ_B and $\tilde{\varphi}_{RB} \approx_\varepsilon \varphi_{RB}$ (with $\tilde{\varphi}_R = \varphi_R$) be the states such that,

$$\mathrm{I}^{\varepsilon}_{\max}(\overset{\bullet}{R} : B)_\varphi = \mathrm{D}_{\max}(\tilde{\varphi}_{RB} \| \varphi_R \otimes \sigma_B).$$

Let $\tilde{\varphi}_{RBAC}$ be a purification of $\tilde{\varphi}_{RB}$ guaranteed by Uhlmann's theorem such that $\tilde{\varphi}_{RBAC} \approx_\varepsilon \varphi_{RBAC}$. Imagine that the circuit in Figure 9.1 starts with $\tilde{\varphi}_{RBAC}$; that is Figure 9.2.

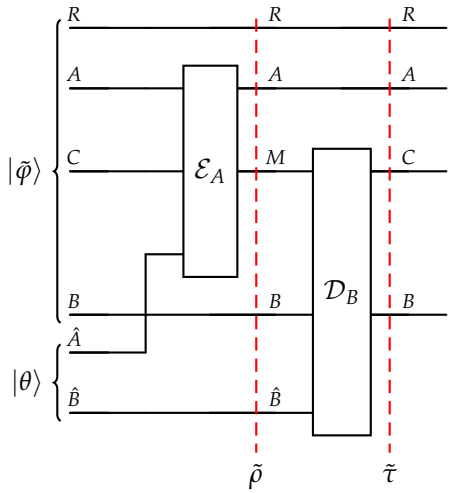

Figure 9.2: Circuit for state redistribution with a close-by starting state.

Consider,

$$\begin{aligned}
&\mathrm{I}^{\varepsilon}_{\max}(\overset{\bullet}{R} : B)_\varphi + 2\log|M| \\
&= \mathrm{D}_{\max}(\tilde{\varphi}_{RB} \| \varphi_R \otimes \sigma_B) + 2\log|M| \\
&= \mathrm{D}_{\max}(\tilde{\varphi}_{RB} \otimes \theta_{\hat{B}} \| \varphi_R \otimes \sigma_B \otimes \theta_{\hat{B}}) + 2\log|M| \\
&= \mathrm{D}_{\max}(\tilde{\rho}_{RB\hat{B}} \| \varphi_R \otimes \sigma_B \otimes \theta_{\hat{B}}) + 2\log|M| && (\tilde{\rho}_{RB\hat{B}} = \tilde{\varphi}_{RB} \otimes \theta_{\hat{B}}) \\
&\geq \mathrm{D}_{\max}(\tilde{\rho}_{RB\hat{B}M} \| \varphi_R \otimes \sigma_B \otimes \theta_{\hat{B}} \otimes \mu_M) && (\text{Fact 9.1 and } \mu_M := \mathbb{I}_M/|M|) \\
&\geq \mathrm{D}_{\max}(\tilde{\tau}_{RBC} \| \varphi_R \otimes \eta_{BC}) && (\text{DPI and } \eta_{BC} := \mathcal{D}_B(\sigma_B \otimes \theta_{\hat{B}} \otimes \mu_M)) \\
&\geq \mathrm{I}^{2\varepsilon}_{\max}(\overset{\bullet}{R} : BC)_\varphi.
\end{aligned}$$

The last inequality follows since $\tilde{\varphi}_{RBC} \approx_\varepsilon \varphi_{RBC}$, we have using DPI, $\tilde{\tau}_{RBC} \approx_\varepsilon \tau_{RBC}$. Since $\varphi_{RBC} \approx_\varepsilon \tau_{RBC}$, from the triangle inequality we have $\tilde{\tau}_{RBC} \approx_{2\varepsilon} \varphi_{RBC}$. Also $\tilde{\tau}_R = \tilde{\varphi}_R = \varphi_R$. □

Open question: Can we get tight achievability and converse bounds for state redistribution in the one-shot setting?

9.4 Asymptotic limit of the converse bound

Here we show the asymptotic limit of the one shot converse bound derived above [2]. This means we want to find a lower bound on the rate

$$R \stackrel{\text{def}}{=} \lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{\log|M_n|}{n}$$

of communication, where $\log|M_n|$ is the number of qubits that are needed to be communicated for a protocol with input $|\varphi\rangle_{RABC}^{\otimes n}$ and overall error ε.

We need the following definition and Facts.

[2] Anurag Anshu, Rahul Jain, and Naqueeb Ahmad Warsi. A one-shot achievability result for quantum state redistribution. *IEEE Transactions on Information Theory*, 64(3):1425–1435, 2018. DOI: 10.1109/TIT.2017.2776112

Fact 9.2.

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n}\mathrm{I}^{\varepsilon}_{\max}(\overset{\bullet}{A}^n : B^n)_{\rho^{\otimes n}} = \mathrm{I}(A:B)_\rho.$$

Consider,

$$\begin{aligned}
R &\stackrel{\text{def}}{=} \lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{\log|M_n|}{n}\\
&\geq \lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{2n}\left(\mathrm{I}^{2\varepsilon}_{\max}(\overset{\bullet}{R}^n : B^nC^n)_{\varphi^{\otimes n}} - \mathrm{I}^{\varepsilon}_{\max}(\overset{\bullet}{R}^n : B^n)_{\varphi^{\otimes n}}\right) && \text{(Claim 2)}\\
&= \frac{1}{2}\left(\mathrm{I}(R:BC)_\varphi - \mathrm{I}(R:B)_\varphi\right) && \text{(Fact 9.2)}\\
&\geq \frac{\mathrm{I}(R:C\mid B)_\varphi}{2}.
\end{aligned}$$

9.5 Asymptotic limit of the achievability bound

We need the following definition and Fact.

Definition 9.3 (Partially smooth max relative-entropy)**.** *Let ρ_{AB}, σ_{AB} be states and $\varepsilon > 0$. The ε-partially smooth max*

relative-entropy between ρ_{AB} and σ_{AB} is defined as,

$$\mathrm{D}^{\varepsilon}_{\max}(\overset{\bullet}{\rho}_{AB}\|\sigma_{AB}) \overset{def}{=} \inf_{\substack{\rho'_{AB}\approx_{\varepsilon}\rho_{AB},\\ \rho_A=\rho'_A}} \mathrm{D}_{\max}(\rho'_{AB}\|\sigma_{AB}).$$

Fact 9.4 (Asymptotic convergence). *Let ρ_{AB}, σ_{AB} be quantum states. Then*

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n}\mathrm{D}^{\varepsilon}_{\max}(\overset{\bullet}{\rho}{}^{\otimes n}_{AB}\|\sigma^{\otimes n}_{AB}) = \mathrm{D}(\rho_{AB}\|\sigma_{AB}).$$

Recall the one shot achievebility bound (with error $4\sqrt{2\varepsilon}$ by setting $\delta = \varepsilon$),

$$\begin{aligned}
&\log|M| \\
&\leq \min_{\sigma_C,\tilde{\varphi}_{RBC}\approx_{\varepsilon}\varphi_{RBC},\varphi_{RB}=\tilde{\varphi}_{RB}} \frac{1}{2}\left[\mathrm{D}_{\max}(\tilde{\varphi}_{RBC}\|\tilde{\varphi}_{RB}\otimes\sigma_C) - \mathrm{D}^{\varepsilon}_H(\varphi_{BC}\|\varphi_B\otimes\sigma_C) + 3\log\frac{1}{\varepsilon}\right].
\end{aligned}$$

In the asymptotic case with input $|\varphi\rangle^{\otimes n}_{RABC}$ and overall error $4\sqrt{2\varepsilon}$ we get,

$$\begin{aligned}
&\log|M_n| \\
&\leq \min_{\sigma_{C^n},\tilde{\varphi}_{R^nB^nC^n}\approx_{\varepsilon}\varphi^{\otimes n}_{RBC},\varphi^{\otimes n}_{RB}=\tilde{\varphi}_{R^nB^n}} \frac{1}{2}\left[\mathrm{D}_{\max}(\tilde{\varphi}_{R^nB^nC^n}\|\varphi^{\otimes n}_{RB}\otimes\sigma_{C^n}) - \mathrm{D}^{\varepsilon}_H(\varphi^{\otimes n}_{BC}\|\varphi^{\otimes n}_B\otimes\sigma_{C^n}) + 3\log\frac{1}{\varepsilon}\right] \\
&\leq \min_{\tilde{\varphi}_{R^nB^nC^n}\approx_{\varepsilon}\varphi^{\otimes n}_{RBC},\varphi^{\otimes n}_{RB}=\tilde{\varphi}_{R^nB^n}} \frac{1}{2}\left[\mathrm{D}_{\max}(\tilde{\varphi}_{R^nB^nC^n}\|\varphi^{\otimes n}_{RB}\otimes\varphi^{\otimes n}_C) - \mathrm{D}^{\varepsilon}_H(\varphi^{\otimes n}_{BC}\|\varphi^{\otimes n}_B\otimes\varphi^{\otimes n}_C) + 3\log\frac{1}{\varepsilon}\right] \\
&= \frac{1}{2}\left[\mathrm{D}^{\varepsilon}_{\max}(\overset{\bullet}{\varphi}{}^{\otimes n}_{RBC}\|\varphi^{\otimes n}_{RB}\otimes\varphi^{\otimes n}_C) - \mathrm{D}^{\varepsilon}_H(\varphi^{\otimes n}_{BC}\|\varphi^{\otimes n}_B\otimes\varphi^{\otimes n}_C) + 3\log\frac{1}{\varepsilon}\right].
\end{aligned}$$

Using Fact 9.4 and Fact 6.2, we get (note that the last term above goes to 0 in the limit),

$$\begin{aligned}
R &\overset{\text{def}}{=} \lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{\log|M_n|}{n} \\
&\leq \frac{1}{2}\left(\mathrm{D}(\varphi_{RBC}\|\varphi_{RB}\otimes\varphi_C) - \mathrm{D}(\varphi_{BC}\|\varphi_B\otimes\varphi_C)\right) \\
&= \frac{1}{2}\left(\mathrm{I}(RB:C)_{\varphi} - \mathrm{I}(B:C)_{\varphi}\right) \\
&= \frac{\mathrm{I}(R:C\mid B)_{\varphi}}{2}.
\end{aligned}$$

Remark: From here we also get the asymptotic achievability and converse bounds for state splitting (and hence also state merging) by removing B, that is $\frac{\mathrm{I}(R:C)_{\varphi}}{2}$.

10
Recap

10.1 Source coding

Here we list the quantum communication cost. Let $\varepsilon \geq 0, \delta > 0$.

10.1.1 State splitting, state merging

- Achievability
  - One-shot:

$$\text{with error } \varepsilon+\delta: \ \frac{1}{2}\left(\mathrm{I}_{\max}^{\varepsilon}(\overset{\bullet}{R}:C)_{\varphi}+2\log\left(\frac{1}{\delta}\right)\right),$$

$$\text{with error } 2\varepsilon+\delta: \ \frac{1}{2}\left(\mathrm{I}_{\max}^{\varepsilon}(R:C)_{\varphi}+2\log\left(\frac{1}{\delta}\right)\right).$$

  - Asymptotic:

$$\frac{1}{2}\mathrm{I}(R:C)_{\varphi}.$$

- Converse
  - One-shot (with error ε):

$$\frac{1}{2}\mathrm{I}_{\max}^{\varepsilon}(\overset{\bullet}{R}:C)_{\varphi}.$$

  - Asymptotic:

$$\frac{1}{2}\mathrm{I}(R:C)_{\varphi}.$$

10.1.2 State redistribution

- Achievability

  - One-shot (with error $4\sqrt{\varepsilon+\delta}$):

$$\frac{1}{2}\min_{\sigma_C,\varphi'_{RBC}\approx_\varepsilon\varphi_{RBC},\varphi'_{RB}=\varphi_{RB}}\left(\mathrm{D}_{\max}(\varphi'_{RBC}\|\varphi_{RB}\otimes\sigma_C)-\mathrm{D}_H^\varepsilon(\varphi_{BC}\|\varphi_B\otimes\sigma_C)+3\log\left(\frac{1}{\delta}\right)\right).$$

  - Asymptotic:

$$\frac{1}{2}\mathrm{I}(R:C\mid B)_\varphi.$$

- Converse
  - One-shot (with error ε):

$$\frac{1}{2}\left(\mathrm{I}_{\max}^{2\varepsilon}(\overset{\bullet}{R}:BC)_\varphi-\mathrm{I}_{\max}^{\varepsilon}(\overset{\bullet}{R}:B)_\varphi\right).$$

  - Asymptotic:

$$\frac{1}{2}\mathrm{I}(R:C\mid B)_\varphi.$$

10.2 Channel-coding

Let $\varepsilon>0$. Let $\mathcal{N}_{A\to B}$ be a q-q channel.

- Achievability
  - One-shot (with error 6ε):

$$\sup_{|\psi\rangle_{A\hat{A}}}\mathrm{D}_H^\varepsilon(\mathcal{N}_{A\to B}(\psi_{A\hat{A}})\|\mathcal{N}_{A\to B}(\psi_A)\otimes\psi_{\hat{A}})-\log\left(\frac{1}{\varepsilon}\right).$$

  - Asymptotic:

$$\mathrm{cap}(\mathcal{N}_{A\to B}):=\sup_{|\psi\rangle_{A\hat{A}}}\mathrm{I}(B:\hat{A})_{\mathcal{N}_{A\to B}(\psi_{A\hat{A}})},$$

    $\mathrm{cap}(\mathcal{N}_{A\to B})$ is called the entanglement-assisted classical capacity.

- Converse
  - One-shot (with error ε):

$$\sup_{|\psi\rangle_{A\hat{A}}}\mathrm{D}_H^\varepsilon(\mathcal{N}_{A\to B}(\psi_{A\hat{A}})\|\mathcal{N}_{A\to B}(\psi_A)\otimes\psi_{\hat{A}}).$$

  - Asymptotic:

$$\mathrm{cap}(\mathcal{N}_{A\to B}).$$

Similar for the c-q and c-c channels where quantum state $|\psi\rangle_{A\hat{A}}$ is replaced with probability distribution $p_{A\hat{A}}$ in sup.

11
The quantum substate theorem

11.1 Theorem statement

The *quantum substate theorem* [1] [2] upper bounds the smooth max relative-entropy in terms of the relative-entropy between two quantum states.

Theorem 11.1 (Substate theorem)**.** *Let ρ, σ be states such that* $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$ *and* $\varepsilon > 0$*. Then*

$$\mathrm{D}_{\max}^{\sqrt{\varepsilon}}(\rho\|\sigma) \leq \frac{1}{\varepsilon}\left(\mathrm{D}(\rho\|\sigma) + 1\right) + \log\left(\frac{1}{1-\varepsilon}\right).$$

[1] Rahul Jain and Ashwin Nayak. Short proofs of the quantum substate theorem. *IEEE Transactions on Information Theory*, 58(6):3664–3669, 2012. DOI: 10.1109/TIT.2012.2184522

[2] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. A property of quantum relative entropy with an application to privacy in quantum communication. *J. ACM*, 56(6), September 2009. ISSN 0004-5411. DOI: 10.1145/1568318.1568323. URL https://doi.org/10.1145/1568318.1568323

11.2 Observational divergence

Before getting into the proof, let us introduce a new information theoretic quantity, the *observational divergence* [3].

[3] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. A property of quantum relative entropy with an application to privacy in quantum communication. *J. ACM*, 56(6), September 2009. ISSN 0004-5411. DOI: 10.1145/1568318.1568323. URL https://doi.org/10.1145/1568318.1568323

Definition 11.2 (Observational divergence)**.** *Let ρ, σ be states, such that* $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$*. The observational divergence between ρ and σ is defined as*

$$\mathrm{D}^{\mathrm{o}}(\rho\|\sigma) \stackrel{def}{=} \max\left\{ \mathrm{Tr}(T\rho)\log\frac{\mathrm{Tr}(T\rho)}{\mathrm{Tr}(T\sigma)} \;\middle|\; \begin{array}{l} 0 \leq T \leq \mathbb{1}, \\ \mathrm{Tr}(T\sigma) \neq 0 \end{array} \right\}.$$

The following claim shows that the observational divergence is upper bounded by the relative-entropy plus one.

Claim 3. *Let ρ, σ be states such that* $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$*.*

Then,

$$\mathrm{D}^{\mathrm{o}}(\rho\|\sigma) \leq \mathrm{D}(\rho\|\sigma) + 1.$$

Proof. Denote T^* as the operator achieving the max in $\mathrm{D}^{\mathrm{o}}(\rho\|\sigma)$. Define the random variable P, Q to be the outcomes on measuring ρ and σ by the POVM $\{\Pi_0 = T^*, \Pi_1 = \mathbb{1} - T^*\}$, respectively. Let $p \stackrel{\text{def}}{=} \mathrm{Tr}(T^*\rho)$ and $q \stackrel{\text{def}}{=} \mathrm{Tr}(T^*\sigma)$. Then,

$$\begin{aligned}\Pr[P = 0] = p, \quad \Pr[P = 1] = 1 - p,\\ \Pr[Q = 0] = q, \quad \Pr[Q = 1] = 1 - q,\end{aligned}$$

and

$$\mathrm{D}^{\mathrm{o}}(\rho\|\sigma) = p \log \frac{p}{q}.$$

By the DPI for relative-entropy we have,

$$p \log \frac{p}{q} + (1-p) \log \frac{1-p}{1-q} = \mathrm{D}(P\|Q) \stackrel{\text{DPI}}{\leq} \mathrm{D}(\rho\|\sigma). \quad (11.1)$$

We know that the binary entropy $\mathrm{H}(p) \leq 1$, and hence,

$$(1-p) \log \frac{1}{1-p} \leq p \log \frac{1}{p} + (1-p) \log \frac{1}{1-p} = \mathrm{H}(p) \leq 1.$$

Using above it follows,

$$(1-p) \log \frac{1-p}{1-q} \geq (1-p) \log(1-p) \geq -1. \quad (11.2)$$

Consider,

$$\begin{aligned}\mathrm{D}^{\mathrm{o}}(\rho\|\sigma) - 1 &= p \log \frac{p}{q} - 1 \\ &\leq p \log \frac{p}{q} + (1-p) \log \frac{1-p}{1-q} && \text{(Eq. (11.2))} \\ &\leq \mathrm{D}(\rho\|\sigma). && \text{(Eq. (11.1))}\end{aligned}$$

Rearranging the above inequality gives us the desired result. □

11.3 Proof of the substate theorem

We need the following Facts.

Fact 11.3. *Let τ, σ be states. Then,*

$$2^{\mathrm{D}_{\max}(\tau\|\sigma)} := \min_{\kappa:\tau\leq\kappa\sigma} \kappa = \max_{M\geq 0, \mathrm{Tr}(M\sigma)\leq 1} \mathrm{Tr}(M\tau).$$

Fact 11.4 (Minimax theorem)**.** *Let A_1, A_2 be non-empty convex, compact subsets of $\mathbb{R}^n$ for some $n \geq 1$. Let $u : A_1 \times A_2 \to \mathbb{R}$ be a continuous function such that,*

1. *$\forall a_2 \in A_2, u(\cdot, a_2)$ is quasi-concave, that is,*
   $\{a_1 \in A_1 : \forall a_1' \in A_1, u(a_1, a_2) \geq u(a_1', a_2)\}$ is convex.
2. *$\forall a_1 \in A_1, u(a_1, \cdot)$ is quasi-convex, that is,*
   $\{a_2 \in A_2 : \forall a_2' \in A_2, u(a_1, a_2) \leq u(a_1, a_2')\}$ is convex.

There exists $(a_1^*, a_2^*) \in A_1 \times A_2$ such that,

$$\max_{a_1\in A_1} \min_{a_2\in A_2} u(a_1, a_2) = \min_{a_2\in A_2} \max_{a_1\in A_1} u(a_1, a_2) = u(a_1^*, a_2^*).$$

Fact 11.5 (Gentle measurement lemma)**.** *Let ρ be a state and $0 \leq A^\dagger A \leq \mathbb{1}$. Then,*

$$\mathrm{F}\Big(\rho, \frac{A\rho A^\dagger}{\mathrm{Tr}(A\rho A^\dagger)}\Big) \geq \mathrm{Tr}(A\rho A^\dagger).$$

Informally: The fidelity between the initial state ρ and the state conditioned on the success of POVM element $(A^\dagger A)$ is large if the success is close to 1.

Fact 11.6 (Joint concavity of square-root fidelity)**.** *Let $\rho_0, \rho_1, \sigma_0, \sigma_1$ be states and $p \in [0, 1]$. Then,*

$$\begin{aligned}&\mathrm{F}^{1/2}(p\rho_0 + (1-p)\rho_1, p\sigma_0 + (1-p)\sigma_1)\\ &\geq p\mathrm{F}^{1/2}(\rho_0, \sigma_0) + (1-p)\mathrm{F}^{1/2}(\rho_1, \sigma_1).\end{aligned}$$

Fact 11.7. *Let $\rho_0, \rho_1, \sigma_0, \sigma_1$ be states, $p \in [0, 1]$ and*

$$\rho = p\,|0\rangle\langle 0| \otimes \rho_0 + (1-p)\,|1\rangle\langle 1| \otimes \rho_1,$$

$$\sigma = p\,|0\rangle\langle 0| \otimes \sigma_0 + (1-p)\,|1\rangle\langle 1| \otimes \sigma_1.$$

Then,

$$\mathrm{F}^{1/2}(\rho, \sigma) = p\mathrm{F}^{1/2}(\rho_0, \sigma_0) + (1-p)\mathrm{F}^{1/2}(\rho_1, \sigma_1).$$

Proof of Theorem 11.1. Since $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$, we can assume without loss of generality that the support of σ is the full support. Denote the dimension of the full support by n. We start with the following claim.

Claim 4.

$$2^{\mathrm{D}_{\max}^{\sqrt{\varepsilon}}(\rho\|\sigma)} = \max_{\substack{M\geq 0,\\ \mathrm{Tr}(M\sigma)\leq 1}} \min_{\substack{\tilde{\rho}\geq 0,\\ \mathrm{Tr}(\tilde{\rho})=1,\\ \mathrm{F}(\tilde{\rho},\rho)\geq 1-\varepsilon}} \mathrm{Tr}(M\tilde{\rho}).$$

Proof. By the definition of the ε-smooth max mutual information, we have

$$\begin{aligned} 2^{\mathrm{D}_{\max}^{\sqrt{\varepsilon}}(\rho\|\sigma)} &\stackrel{\text{def}}{=} \min_{\substack{\tilde{\rho}\geq 0,\\ \mathrm{Tr}(\tilde{\rho})=1,\\ \mathrm{F}(\tilde{\rho},\rho)\geq 1-\varepsilon}} \min_{\kappa:\tilde{\rho}\leq\kappa\sigma} \kappa \\ &= \min_{\substack{\tilde{\rho}\geq 0,\\ \mathrm{Tr}(\tilde{\rho})=1,\\ \mathrm{F}(\tilde{\rho},\rho)\geq 1-\varepsilon}} \max_{\substack{M\geq 0,\\ \mathrm{Tr}(M\sigma)\leq 1}} \mathrm{Tr}(M\tilde{\rho}). \qquad \text{(Fact 11.3)} \end{aligned}$$

Now we want to swap the order of the minimization and maximization using the minimax theorem (Fact 11.4). To apply the minimax theorem, we first need to show that all conditions for the theorem are satisfied. Let

$$A_\rho \stackrel{\text{def}}{=} \{\tilde{\rho} \geq 0 \mid \mathrm{Tr}(\tilde{\rho}) = 1, \mathrm{F}(\tilde{\rho},\rho) \geq 1-\varepsilon\}$$

and

$$A_\sigma \stackrel{\text{def}}{=} \{M \geq 0 \mid \mathrm{Tr}(M\sigma) \leq 1\}.$$

Recall that if a set is closed and bounded, it is compact, so it suffices to prove that A_ρ and A_σ are closed and bounded.

1. The conditions $\mathrm{Tr}(\tilde{\rho}) = 1$ and $\mathrm{F}(\tilde{\rho},\rho) \geq 1-\varepsilon$ imply that A_ρ is closed and bounded. We show that A_ρ is convex by the concavity of square-root fidelity. Let $\tilde{\rho}_1, \tilde{\rho}_2 \in A_\rho$, and $p \in [0,1]$. Then we have

$$\begin{aligned} &\mathrm{F}^{1/2}(p\tilde{\rho}_1 + (1-p)\tilde{\rho}_2, \rho) \\ &\geq p\mathrm{F}^{1/2}(\tilde{\rho}_1,\rho) + (1-p)\mathrm{F}^{1/2}(\tilde{\rho}_2,\rho) \qquad \text{(Fact 11.6)} \\ &\geq \sqrt{1-\varepsilon}. \end{aligned}$$

   The other conditions

$$\mathrm{Tr}(p\tilde{\rho}_1 + (1-p)\tilde{\rho}_2) = 1$$

and $p\tilde{\rho}_1 + (1-p)\tilde{\rho}_2 \geq 0$ are naturally satisfied, and hence, A_ρ is convex.

2. For A_σ, verifying that A_σ is closed and convex is straightforward. Let $\lambda_{\min}(\sigma)$ be the minimum eigenvalue of σ. Since we assume σ has the full support, $\sigma - \lambda_{\min}(\sigma)\mathbb{I} \geq 0$ follows. It implies that

$$\mathrm{Tr}(M(\sigma - \lambda_{\min}(\sigma)\mathbb{1})) \geq 0,$$

and so

$$\mathrm{Tr}(M) \leq \frac{\mathrm{Tr}M\sigma}{\lambda_{\min}(\sigma)} \leq \frac{1}{\lambda_{\min}(\sigma)}.$$

Thus, A_σ is also bounded.

3. Notice that our target function $\mathrm{Tr}(M\tilde{\rho})$ is linear in both $\tilde{\rho}$ and M, so it's also convex and concave on both of them.

Now we can apply the minimax theorem to swap the order of the minimization and maximization, and get the desired. □

To bound the RHS in Claim 4, it is sufficient to show that, for any $M \in A_\sigma$, we can construct a density operator $\rho'_M \in A_\rho$ such that $\mathrm{Tr}(M\rho'_M)$ is bounded. Let the spectral decomposition of M be

$$M = \sum_{i=1}^{n} p_i \left|v_i\right\rangle \left\langle v_i\right|.$$

For $i \in [n]$ we let

$$\lambda_i \stackrel{\text{def}}{=} \left\langle v_i\right| \rho \left|v_i\right\rangle \text{ and } \gamma_i \stackrel{\text{def}}{=} \left\langle v_i\right| \sigma \left|v_i\right\rangle.$$

To better understand this, we can view M, ρ and σ in the eigen basis of M, i.e., $|v_i\rangle$'s, then λ_i's and γ_i's are diagonal entries of ρ and σ in this basis, respectively, i.e.,

$$M = \begin{bmatrix} p_1 & & & \\ & p_2 & 0 & \\ & 0 & \ddots & \\ & & & p_n \end{bmatrix}$$

$$\rho = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & * & \\ & * & \ddots & \\ & & & \lambda_n \end{bmatrix} \quad \sigma = \begin{bmatrix} \gamma_1 & & & \\ & \gamma_2 & * & \\ & * & \ddots & \\ & & & \gamma_n \end{bmatrix}$$

Let $d \stackrel{\text{def}}{=} \mathrm{D}^{\mathrm{o}}(\rho\|\sigma)$. We define a set of 'bad' indices,

$$B \stackrel{\text{def}}{=} \{i \in [n] \mid \lambda_i > 2^{\frac{d}{\varepsilon}}\gamma_i\}.$$

Let $\Pi_B = \sum_{i\in B} |v_i\rangle \langle v_i|$ be the projector onto the subspace spanned by $\{|v_i\rangle\}_{i\in B}$. It follows that

$$\mathrm{Tr}(\Pi_B\rho) > 2^{\frac{d}{\varepsilon}}\mathrm{Tr}(\Pi_B\sigma). \tag{11.3}$$

By the definition of the observational divergence, we have

$$\begin{aligned} d &\geq \mathrm{Tr}(\Pi_B\rho)\log\frac{\mathrm{Tr}(\Pi_B\rho)}{\mathrm{Tr}(\Pi_B\sigma)} \\ &> \frac{d}{\varepsilon}\mathrm{Tr}(\Pi_B\rho). \qquad \text{(Eq. (11.3))} \end{aligned}$$

Rearranging gives us

$$\mathrm{Tr}(\Pi_B\rho) < \varepsilon. \tag{11.4}$$

We can construct a density operator ρ'_M by projecting ρ onto the orthogonal subspace of the subspace specified by B.

$$\begin{aligned} \rho''_M &\stackrel{\text{def}}{=} (\mathbb{I} - \Pi_B)\rho(\mathbb{I} - \Pi_B), \\ \rho'_M &\stackrel{\text{def}}{=} \frac{\rho''_M}{\mathrm{Tr}(\rho''_M)}. \end{aligned}$$

From Eq. (11.4), we have

$$\begin{aligned} \mathrm{Tr}(\rho''_M) &= 1 - \mathrm{Tr}(\Pi_B\rho) \\ &> 1 - \varepsilon. \end{aligned} \tag{11.5}$$

Together with the gentle measurement lemma (Fact 11.5), we have

$$\mathrm{F}(\rho'_M, \rho) \geq \mathrm{Tr}(\rho''_M) > 1 - \varepsilon. \tag{11.6}$$

It implies that the constructed density operator ρ'_M is in A_ρ.

Consider,

$$\begin{aligned}
(1-\varepsilon)\mathrm{Tr}(M\rho'_M) &\leq \mathrm{Tr}(\rho''_M)\mathrm{Tr}(M\rho'_M) && \text{(Eq. (11.5))}\\
&= \mathrm{Tr}(M\rho''_M)\\
&= \sum_{i\notin B} p_i\lambda_i\\
&\leq 2^{\frac{d}{\varepsilon}} \sum_{i\notin B} p_i\gamma_i && \text{(definition of } B\text{)}\\
&\leq 2^{\frac{d}{\varepsilon}} \sum_{i=1}^{n} p_i\gamma_i\\
&= 2^{\frac{d}{\varepsilon}}\mathrm{Tr}(M\sigma)\\
&\leq 2^{\frac{d}{\varepsilon}}. && \text{(definition of } A_\sigma\text{)}
\end{aligned}$$

Therefore, for any $M \in A_\sigma$, we can always find some density operator $\rho'_M \in A_\rho$ such that $\mathrm{Tr}(M\rho'_M)$ is upper bounded by $\frac{2^{d/\varepsilon}}{1-\varepsilon}$. Therefore,

$$\begin{aligned}
2^{\mathrm{D}_{\max}^{\sqrt{\varepsilon}}(\rho\|\sigma)} &= \max_{\substack{M\geq 0,\\ \mathrm{Tr}(M\sigma)\leq 1}} \min_{\substack{\tilde{\rho}\geq 0,\\ \mathrm{Tr}(\tilde{\rho})=1,\\ \mathrm{F}(\tilde{\rho},\rho)\geq 1-\varepsilon}} \mathrm{Tr}(M\tilde{\rho})\\
&\leq \frac{2^{\frac{d}{\varepsilon}}}{1-\varepsilon}\\
&\leq \frac{2^{\frac{\mathrm{D}(\rho\|\sigma)+1}{\varepsilon}}}{1-\varepsilon}. && \text{(Claim 3)}
\end{aligned}$$

Taking log on both sides gives us the desired.

$$\mathrm{D}_{\max}^{\sqrt{\varepsilon}}(\rho\|\sigma) \leq \frac{1}{\varepsilon}(\mathrm{D}(\rho\|\sigma)+1) + \log\frac{1}{1-\varepsilon}. \qquad \square$$

Proof idea: By the minimax theorem, we can swap the order of the minimization and maximization, and then the proof of the substate theorem becomes simpler. Instead of picking a suitable state and optimizing over all POVM operators, the minimax theorem allows us to first fix a POVM operator and construct a corresponding state to bound the ε-smooth maximum relative-entropy.

12
The reverse Shannon theorem

We have seen what it looks like to send information reliably through a noisy channel, that is, simulating a noiseless channel using a noisy channel. Today we see how to do the reverse: try to simulate a noisy channel using a noiseless (ideal) channel. The reason for performing this reverse is to minimize communication cost needed when simulating a noisy channel, using shared resources like randomness or entanglement.

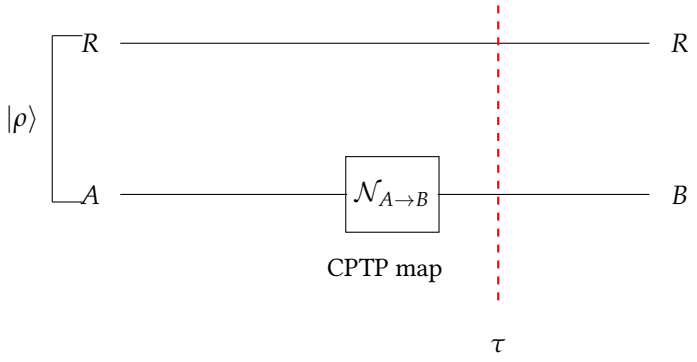

Figure 12.1: Channel $\mathcal{N}_{A\to B}$.

As shown in Figure 12.1, we start with the state $|\rho\rangle_{RA}$ sent through the channel $\mathcal{N}_{A\to B}$ to obtain the state τ.

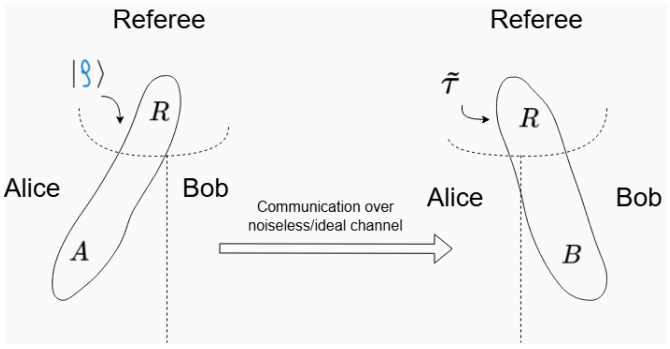

Figure 12.2: Simulating channel $\mathcal{N}_{A\to B}$ using an ideal channel.

As shown in Figure 12.2, Alice and Bob simulate the action of channel $\mathcal{N}_{A\to B}$ using a communication protocol with a noiseless channel. They both know the description of $\mathcal{N}_{A\to B}$ but they don't know the starting state $|\rho\rangle_{RA}$. The requirement is,

$$\forall\, |\rho\rangle_{RA} : \tilde{\tau}_{RB} \approx_\varepsilon \tau_{RB}.$$

The goal is to minimize communication cost while faithfully simulating the channel for every input state $|\rho\rangle_{RA}$.

Let $\varepsilon > 0$. We present a protocol (see Figure 12) with communication cost [1]

$$c \stackrel{\text{def}}{=} \max_{|\rho\rangle_{RA}} \mathrm{I}^{\varepsilon/2}_{\max}(\overset{\bullet}{R} : B)_{\mathcal{N}_{A\to B}(\rho)} + 2\log\left(\frac{2}{\varepsilon}\right).$$

We assume $|\rho\rangle_{RA}$ is the canonical purification of ρ_A. Since we can always perform local isometries on the system R, which commute with the operations of the communication protocol, so the particular purification that exists is not important.

[1] Michael X. Cao, Rahul Jain, and Marco Tomamichel. Quantum channel simulation in fidelity is no more difficult than state splitting. In *2024 IEEE International Symposium on Information Theory (ISIT)*, pages 1421–1425, 2024. DOI: 10.1109/ISIT57864.2024.10619461

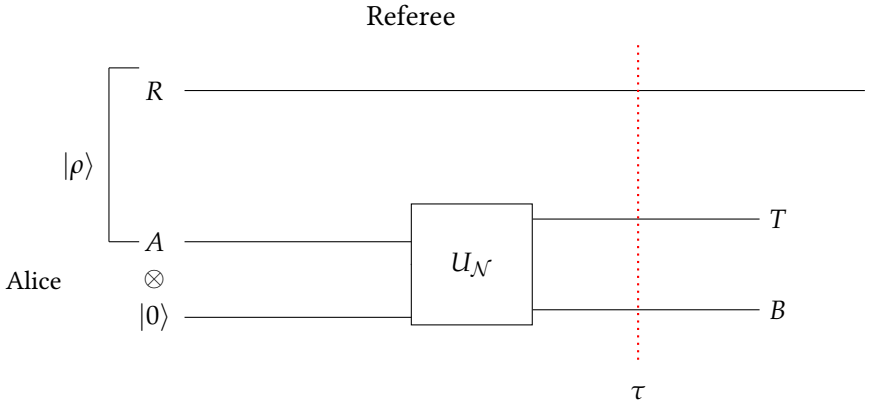

Figure 12.3: Alice's initial operation in communication protocol. After this Alice and Bob would want to implement state splitting protocol to transfer the register B to Bob.

Alice first performs a unitary $U_{\mathcal{N}}$ that effectively applies the channel $\mathcal{N}_{A\to B}$ and generates τ at her end. Alice and Bob then would want to run the state-splitting protocol to transfer the register B across. In contrast to the state-splitting protocol where Alice and Bob need to know τ, here they don't know the starting state $|\rho\rangle_{RA}$ and therefore they don't know τ. Because of this issue, we will again use the powerful minimax theorem.

For a starting state ρ_A and a protocol $\mathcal{P}$, define

$$f(\rho_A, \mathcal{P}) \stackrel{\text{def}}{=} \sqrt{\mathrm{F}(\tilde{\tau}_{RB}, \tau_{RB})},$$

where $\tilde{\tau}_{RB}$ is the output state of $\mathcal{P}$. Let $\varepsilon > 0$. We know that

for every state ρ_A, there exists some protocol $\mathcal{P}_\rho$ with communication c bits (and bounded entanglement), such that

$$f(\rho_A, \mathcal{P}_\rho) \geq \sqrt{1-\varepsilon^2}.$$

Let $\mathcal{S}$ be the closure of the convex hull of $\{\mathcal{P}_\rho\}_\rho$. Then,

$$\sqrt{1-\varepsilon^2} \leq \min_{\rho_A} \max_{\mathcal{P}\in\mathcal{S}} f(\rho_A, \mathcal{P}) = \max_{\mathcal{P}\in\mathcal{S}} \min_{\rho_A} f(\rho_A, \mathcal{P}).$$

We will show that the equality above follows from the minimax theorem (Fact 11.4). For this we need to verify that the conditions needed for the minimax theorem hold. This would imply that there exists a protocol, with communication at most c bits, such that for every input state the purified distance of the output state with τ_{RB} is at most ε.

Below we verify all the conditions needed for the minimax theorem.

1. The set of states $\{\rho_A\}$ is convex and compact.

2. The set $\mathcal{S}$ is convex and closed by definition. Since, for each ρ, the protocol $\mathcal{P}_\rho$ is a bounded-communication, bounded-entanglement protocol, the set $\mathcal{S}$ is also bounded (since the closure of a convex hull of a bounded set is bounded).

3. Fix ρ_A. We want to show that the set (call it $\mathcal{S}(\rho_A)$) of protocols in $\mathcal{S}$ maximizing $f(\rho_A, \cdot)$ is convex.

   Let $p \in [0,1]$ and $\tilde{\tau}^0_{RB}$, $\tilde{\tau}^1_{RB}$ be the output states of protocols $\mathcal{P}_0, \mathcal{P}_1 \in \mathcal{S}(\rho_A)$, respectively, on input $|\rho\rangle_{RA}$ (the canonical purification of ρ_A). Let $\mathcal{P} \stackrel{\text{def}}{=} p\mathcal{P}_0 + (1-p)\mathcal{P}_1$. The protocol $\mathcal{P}$ can be implemented as follows: Alice and Bob choose (using public coins) to implement $\mathcal{P}_0$ with probability p and $\mathcal{P}_1$ with probability $1-p$. Then,

   $$\tilde{\tau}_{RB} \stackrel{\text{def}}{=} p\tilde{\tau}^0_{RB} + (1-p)\tilde{\tau}^1_{RB}$$

   would be the output state of the protocol $\mathcal{P}$. Let $\tau_{RB} = \mathcal{N}_{A\to B}(\rho_{RA})$. From Fact 11.6 we have,

   $$\begin{aligned} f(\rho_A, \mathcal{P}) &= \mathrm{F}^{1/2}(\tilde{\tau}_{RB}, \tau_{RB}) \\ &\geq p \cdot \mathrm{F}^{1/2}(\tilde{\tau}^0_{RB}, \tau_{RB}) + (1-p) \cdot \mathrm{F}^{1/2}(\tilde{\tau}^1_{RB}, \tau_{RB}) \\ &= \max f(\rho_A, \cdot). \end{aligned}$$

   This implies $\mathcal{P} \in \mathcal{S}(\rho_A)$.

4. Fix $\mathcal{P}$. We want to show that the set (call it $\mathcal{S}(\mathcal{P})$) of minimizers of the function $f(\cdot, \mathcal{P})$ is convex.

   Let $p \in [0,1]$ and $\rho_A^0, \rho_A^1 \in \mathcal{S}(\mathcal{P})$. Their convex combination is

   $$\rho_A \stackrel{\text{def}}{=} p\rho_A^0 + (1-p)\rho_A^1.$$

   Define

   $$|\psi\rangle_{DRA} \stackrel{\text{def}}{=} \sqrt{p} \cdot |0\rangle_D |\rho^0\rangle_{RA} + \sqrt{1-p} \cdot |1\rangle_D |\rho^1\rangle_{RA}.$$

   Note that $|\psi\rangle_{DRA}$ is a purification of ρ_A.

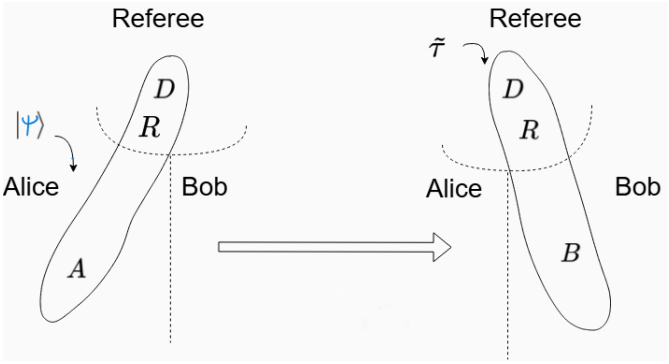

Figure 12.4: Protocol $\mathcal{P}$ with input state $|\psi\rangle_{RDA}$.

   (a) Let $\tilde{\tau}_{DRB}, \tilde{\tau}^0_{RB}, \tilde{\tau}^1_{RB}, \tilde{\gamma}_{RB}$ be the output states of $\mathcal{P}$ on the input states $|\psi\rangle_{DRA}$ (the registers DR are with the Referee, see Figure 12.4), $|\rho^0\rangle_{RA}$, $|\rho^1\rangle_{RA}$ and $|\rho\rangle_{RA}$ respectively.

   (b) Let $\tau_{DRB} \stackrel{\text{def}}{=} \mathcal{N}_{A\to B}(\psi_{DRA})$ and $\gamma_{RB} \stackrel{\text{def}}{=} \mathcal{N}_{A\to B}(\rho_{RA})$.

   (c) Let $\theta_{DRB}, \tilde{\theta}_{DRB}$ be the states obtained by measuring the register D (in the computational basis) in the states $\tau_{DRB}, \tilde{\tau}_{DRB}$ respectively.

   Note that the Uhlmann isometry V that takes $|\psi\rangle_{DRA}$ to $|\rho\rangle_{RA}$, commutes with all the operations in $\mathcal{P}$ and also with

$\mathcal{N}_{A\to B}$. Consider,

$$\begin{aligned}
f(\rho_A, \mathcal{P}) &= \mathrm{F}^{1/2}(\tilde{\gamma}_{RB}, \gamma_{RB}) \\
&= \mathrm{F}^{1/2}(\mathcal{P}(\rho_{RA}), \mathcal{N}_{A\to B}(\rho_{RA})) \\
&= \mathrm{F}^{1/2}(\mathcal{P}(V\psi_{DRB}V^\dagger), \mathcal{N}_{A\to B}(V\psi_{DRB}V^\dagger)) \\
&= \mathrm{F}^{1/2}(V\mathcal{P}(\psi_{DRB})V^\dagger, V\mathcal{N}_{A\to B}(\psi_{DRB})V^\dagger) \\
&= \mathrm{F}^{1/2}(\mathcal{P}(\psi_{DRB}), \mathcal{N}_{A\to B}(\psi_{DRB})) \\
&= \mathrm{F}^{1/2}(\tilde{\tau}_{DRB}, \tau_{DRB}) \\
&\leq \mathrm{F}^{1/2}(\tilde{\theta}_{DRB}, \theta_{DRB}) && \text{(DPI)} \\
&= p \cdot \mathrm{F}^{1/2}(\tilde{\tau}^0_{RB}, \tau^0_{RB}) + (1-p) \cdot \mathrm{F}^{1/2}(\tilde{\tau}^1_{RB}, \tau^1_{RB}) && \text{(Fact 11.7)} \\
&= \min f(\cdot, \mathcal{P}).
\end{aligned}$$

Hence $\rho_A \in \mathcal{S}(\mathcal{P})$.

13
Bibliography

Anurag Anshu, Vamsi Krishna Devabathini, and Rahul Jain. Quantum communication using coherent rejection sampling. *Physical Review Letters*, 119(12), September 2017. DOI: 10.1103/physrevlett.119.120506. URL http://dx.doi.org/10.1103/PhysRevLett.119.120506.

Anurag Anshu, Rahul Jain, and Naqueeb Ahmad Warsi. A one-shot achievability result for quantum state redistribution. *IEEE Transactions on Information Theory*, 64(3):1425–1435, 2018. DOI: 10.1109/TIT.2017.2776112.

Anurag Anshu, Rahul Jain, and Naqueeb Ahmad Warsi. On the near-optimality of one-shot classical communication over quantum channels. *Journal of Mathematical Physics*, 60(1): 012204, 01 2019a. DOI: 10.1063/1.5039796. URL https://doi.org/10.1063/1.5039796.

Anurag Anshu, Rahul Jain, and Naqueeb Ahmad Warsi. Building blocks for communication over noisy quantum networks. *IEEE Transactions on Information Theory*, 65(2):1287–1306, February 2019b. DOI: 10.1109/TIT.2018.2851297.

Mario Berta, Matthias Christandl, and Dave Touchette. Smooth entropy bounds on one-shot quantum state redistribution. *IEEE Transactions on Information Theory*, 62(3):1425–1439, 2016.

Michael X. Cao, Rahul Jain, and Marco Tomamichel. Quantum channel simulation in fidelity is no more difficult than state splitting. In *2024 IEEE International Symposium on Information Theory (ISIT)*, pages 1421–1425, 2024. DOI: 10.1109/ISIT57864.2024.10619461.

Rahul Jain and Ashwin Nayak. Short proofs of the quantum substate theorem. *IEEE Transactions on Information Theory*, 58 (6):3664–3669, 2012. DOI: 10.1109/TIT.2012.2184522.

Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. A property of quantum relative entropy with an application to privacy in quantum communication. *J. ACM*, 56(6), September 2009. ISSN 0004-5411. DOI: 10.1145/1568318.1568323. URL https://doi.org/10.1145/1568318.1568323.

Sumeet Khatri, Ludovico Lami, and Mark M. Wilde. *Principles of Quantum Communication Theory: A Modern Approach*. 2025. URL https://www.markwilde.com/PQCT-khatri-lami-wilde.pdf.