Quantum Information Theory

Min-Hsiu Hsieh

University of Technology Sydney, Australia

October 30, 2015

Contents

1 Quantum States and Channels

	Classical	Quantum
Source	Probability vector $P_X(x)$	Density operators $\rho_A \geq 0 \in \mathbb{C}^{d_A \times d_A}$, and $\operatorname{Tr} \rho_A = 1$
Channel	$p_{Y\|X}$	CPTP map Measurement $\Lambda : \mathcal{Q} \to \mathcal{C}$
Entropy	$H(X) = -\sum p_X(x) \log p_X(x)$	$H(A) = -\operatorname{Tr} \rho_A \log \rho_A$
Conditional Entropy	$H(Y\|X) = H(XY) - H(X)$	$H(A\|B)_\rho = H(AB) - H(B)$
Mutual Information	$I(X:Y) = H(X) - H(X\|Y)$	$I(A:B)_\rho = H(A) - H(A\|B)$
Conditional MI	$I(X:Y\|Z) = I(X:YZ) - I(X\|Z)$	$I(A:B\|C)_\rho = I(A:BC) - I(A\|C)$

Quantum State:

For a d-dimensional Hilbert Space $\mathcal{H}_d$, we fix the computational basis $\{|1\rangle, \ldots, |d\rangle\}$ in it, where $|i\rangle = (0, \cdots, 0, 1, 0, \cdots, 0)^T$. A d-dimensional *pure* quantum system can be mathematically described

by a unit-length vector

$$|\psi\rangle = \sum_{i=1}^{d} \alpha_i |i\rangle,$$

where $\alpha_i \in \mathbb{C}$ and $\sum_{i=1}^{d} |\alpha_i|^2 = 1$. We use the convention that the ket notation $|\psi\rangle$ is a column vector, and the bra notation $\langle\psi| := |\psi\rangle^\dagger$ is a row vector (from complex conjugate and transpose of $|\psi\rangle$). However, a quantum system can be *mixed*. A mixed quantum state ρ can be mathematically described as a density operator in $\mathbb{C}^{d\times d}$, namely, it is a Hermitian (self-adjoint) matrix so that $\rho \geq 0$ and $\operatorname{Tr}\rho = 1$. It is easy to verify that when ρ is rank one, $\rho = |\psi\rangle\langle\psi|$ for some $|\psi\rangle \in \mathcal{H}_d$. Therefore, a quantum system ρ is pure if and only if its $\operatorname{rank}(\rho) = 1$. Otherwise, it is mixed. We denote by $\mathcal{D}(\mathcal{H}_d) = \{\rho : \rho \geq 0, \operatorname{Tr}\rho = 1\}$ the set of density operators defined on $\mathcal{H}_d$.

Exercise 1 *Show that* $\operatorname{Tr}[\rho^2] = 1$ *if and only if* ρ *is pure. If* ρ *is mixed, then* $\operatorname{Tr}[\rho^2] < 1$.

For any quantum system $\rho \in \mathcal{D}(\mathcal{H}_d)$ with rank κ, we can always decompose (by spectral decomposition)

$$\rho = \sum_{j=1}^{\kappa} \mu_j |E_j\rangle\langle E_j|$$

where $\mu_j \geq 0$ are eigenvalues of ρ with eigenvectors $|E_j\rangle$. Since $\operatorname{Tr}\rho = 1$, we have $\sum_{j=1}^{\kappa} \mu_j = 1$.

Entanglement:

We define a *bipartite* pure quantum system $|\psi\rangle\langle\psi|_{AB} \in \mathcal{D}(\mathcal{H}_{d_A} \otimes \mathcal{H}_{d_B})$, where $\otimes$ denotes the tensor product. We can think of the quantum state $|\psi\rangle_{AB}$ held by two parties (we often call them Alice (A) and Bob(B) in the quantum regime) whose local spaces are $\mathcal{H}_{d_A}$ and $\mathcal{H}_{d_B}$, respectively. Since $\{|i\rangle_A \otimes |j\rangle_B\}$ forms an othornormal basis for the Hilbert space $\mathcal{H}_{d_A} \otimes \mathcal{H}_{d_B}$, we can write

$$|\psi\rangle_{AB} = \sum_{i=1}^{d_A} \sum_{j=1}^{d_B} \alpha_{ij} |i\rangle_A \otimes |j\rangle_B$$

where $\alpha_{ij} \in \mathbb{C}$ and $\sum_{i=1}^{d_A} \sum_{j=1}^{d_B} |\alpha_{ij}|^2 = 1$.

Exercise 2 *The quantum state* ρ_A *held by Alice when ignoring the other party Bob is defined*

$$\rho_A := \operatorname{Tr}_B |\psi\rangle\langle\psi|_{AB} = \sum_{i,i'=1}^{d_A} \beta_{ii'} |i\rangle\langle i'|,$$

where Tr_B *is the partial trace on* $\mathcal{H}_{d_B}$*. Compute the exact expression of* $\beta_{ii'}$ *in terms of* α_{ij}*.*

A pure state $|\psi\rangle_{AB}$ is *entangled* if and only if it cannot be written as tensor product of two pure states: $|\psi\rangle_{AB} \neq |\phi\rangle_A \otimes |\phi\rangle_B$ for any $|\phi\rangle_A \in \mathcal{H}_{d_A}$ and $|\phi\rangle_B \in \mathcal{H}_{d_B}$.

Exercise 3 *Define the maximally entangled state (or an ebit) to be*

$$|\Phi_+\rangle_{AB} = \frac{1}{\sqrt{2}} \left(|00\rangle_{AB} + |11\rangle_{AB}\right). \tag{1}$$

Verify that $|\Phi_+\rangle_{AB}$ *is entangled.*

For a general quantum state $\rho_{AB} \in \mathcal{D}(\mathcal{H}_{d_A} \otimes \mathcal{H}_{d_B})$, we say it is a *separable* state if and only if

$$\rho_{AB} = \sum_k p_k \rho_A^{(k)} \otimes \rho_B^{(k)}, \tag{2}$$

for some $\rho_A^{(k)} \in \mathcal{D}(\mathcal{H}_{d_A}), \rho_B^{(k)} \in \mathcal{D}(\mathcal{H}_{d_B})$, $\alpha_k \in \mathbb{R}$, and $\sum_k p_k = 1$. If ρ_{AB} is not separable, then it is entangled.

Quantum Ensemble:

A quantum ensemble is a collection of n quantum states $\rho_B^x \in \mathcal{D}(\mathcal{H}_B)$ with probability p_x, where $\sum_{x=1}^n p_x = 1$, and is denoted by $\mathcal{E} = \{p_x, \rho_B^x\}_{x=1}^n$. Equivalently, we can relate $\mathcal{E}$ to a classical-quantum state σ_{XB}:

$$\sigma_{XB} := \sum_{x=1}^n p_x |x\rangle\langle x|_X \otimes \rho_B^x, \tag{3}$$

where $\{|x\rangle_X\}_{x=1}^n$ denotes the computational basis in the auxiliary system X. The system X can be viewed as the classical labels of the corresponding quantum states.

Exercise 4 *Verify that the classical-quantum state σ_{XB} in Eq. (3) is not an entangled state between systems X and B.*

Exercise 5 *Denote $|+\rangle := \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|-\rangle := \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. Let $\mathcal{E} = \{(1/2, |+\rangle), (1/2, |-\rangle)\}$. Write down the classical-quantum state σ_{XB} of the quantum ensemble $\mathcal{E}$ in the matrix form.*

Measurement:

A measurement is a device that reads out classical messages from a quantum system. It can be mathematically described by $\mathbf{A} := \{A_i\}_{i=1}^n$ of measurement operators (i.e. linear operators in $\mathcal{H}$) so that

$$\sum_{i=1}^n A_i^\dagger A_i = I,$$

where I is the identity operator in $\mathcal{H}$. The outcome j after measuring the quantum state ρ with $\mathbf{A}$ happens with probability

$$p_j = \operatorname{Tr} A_j \rho A_j^\dagger,$$

and the resulting quantum state is

$$\rho' = \frac{1}{p_j} A_j \rho A_j^\dagger.$$

Exercise 6 *Show that $\sum_{j=1}^n p_j = 1$.*

If we do not care about the post-measurement quantum state, we can use the *positive operator-valued measures* (POVM) formalism. A POVM Λ with n measurement outcomes consists of $\{\Lambda_i\}_{i=1}^n$ where each $0 \le \Lambda_i \le I$ and $\sum_i \Lambda_i = I$. Applying the measurement Λ on a quantum state ρ will yield outcome k with probability

$$p_k = \operatorname{Tr}[\Lambda_k \rho].$$

Note that the set of projectors $\{\Pi_i := |i\rangle\langle i|\}_{i=1}^d$ is a special case of a POVM measurement.

Exercise 7 *The POVM measurement Λ and general measurement $\mathbf{A}$ can be related as follows. For a measurement $\mathbf{A}$, we can construct elements of POVM measurement*

$$\Lambda_i = A_i^\dagger A_i.$$

For a POVM measurement Λ, there exists a unitary U so that

$$A_i = U\sqrt{\Lambda_i}.$$

For a quantum ensemble $\mathcal{E} = \{p_i, \rho_i\}_{i=1}^n$ and a POVM $\Lambda = \{\Lambda_i\}_{i=1}^n$, define the *successful probability* of identifying the classical messages in $\mathcal{E}$ by

$$P_c(\mathcal{E}, \Lambda) := \sum_{i=1}^n p_i \operatorname{Tr}[\rho_i \Lambda_i].$$

Exercise 8 *Let $\mathcal{E} = \{(1/2, |+\rangle), (1/2, |-\rangle)\}$. Design a POVM Λ so that $P_c(\mathcal{E}, \Lambda) = 1$.*

Quantum Channels:

A most general quantum channel (or operation) $\mathcal{N} : \mathcal{D}(\mathcal{H}_A) \mapsto \mathcal{D}(\mathcal{H}_B)$ is a completely positive and trace-preserving (CPTP) map:

$$\mathcal{N}_{A\to B} \otimes \mathrm{id}_R(|\psi_\rho\rangle\langle\psi_\rho|_{AR}) = \sigma_{BR} \in \mathcal{D}(\mathcal{H}_B \otimes \mathcal{H}_R)$$

for any quantum state $\rho_A \in \mathcal{D}(\mathcal{H}_A)$ and any auxiliary purification system R with purification $|\psi_\rho\rangle_{AR}$ (i.e. $\operatorname{Tr}_R |\psi_\rho\rangle\langle\psi_\rho|_{AR} = \rho_A$).

Exercise 9 *For a quantum system $\rho_A \in \mathcal{D}(\mathcal{H}_A)$ with rank κ: $\rho = \sum_{j=1}^\kappa \mu_j |E_j\rangle\langle E_j|$, we can always purify ρ as follows:*

$$|\psi_\rho\rangle_{AR} = \sum_{j=1}^\kappa \sqrt{\mu_j} |E_j\rangle_A \otimes |j\rangle_R \tag{4}$$

where $\{|i\rangle_R\}$ are orthonormal vectors in $\mathcal{H}_R$. We call such a purification canonical. Verify $\operatorname{Tr}_R \psi_{\rho_{AR}} = \rho_A$.

Show that the purification is not unique in the sense that there exists other pure state $|\phi_\rho\rangle_{AR}$ so that $\operatorname{Tr}_R \phi_{\rho_{AR}} = \rho_A$.

A quantum channel $\mathcal{N}$ can be equivalently represented by the Kraus representation:

$$\mathcal{N}(\rho) = \sum_{j=1} A_j \rho A_j^\dagger,$$

where $\{A_j\}$ are the Kraus operators of the channel $\mathcal{H}$ satisfying $\sum_j A_j^\dagger A_j = I$.

Exercise 10 *Show that a classical channel $p_{Y|X}(y|x)$ acting on a classical input $p_X(x)$ with $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ can be described as a special case of a quantum channel $\mathcal{N}$ on a density operator ρ. Express the Kraus operators of $\mathcal{N}$ in terms of $p_{Y|X}(y|x)$ and the density operator ρ in terms of $p_X(x)$.*

A closed quantum system evolves according to a unitary. Hence a noisy quantum evolution (a quantum channel) $\mathcal{N}$ on $\rho \in \mathcal{D}(\mathcal{H}_A)$ can be considered as:

$$\mathcal{N}(\rho) = \mathrm{Tr}_E[U(\rho \otimes |0\rangle\langle 0|_E)U^\dagger]$$

where U is a unitary evolution on system $\mathcal{H}_A \otimes \mathcal{H}_E$. This relation allows us to construct Kraus operators $\{A_j := \langle j|_E U|0\rangle_E\}$.

Exercise 11 *Define a quantum erasure channel with probability ε:*

$$\mathcal{N}_e(\rho) = (1-\varepsilon)\rho + \varepsilon|e\rangle\langle e|$$

where $|e\rangle\langle e| \perp \rho$. Construct Kraus operators $\{A_j\}$ for $\mathcal{N}_e$.

We can construct a measurement map $\mathcal{E}_\Lambda : A \to AX$ associated with a measurement $\{\Lambda_i\}_{i=1}^n$ as follows:

$$\mathcal{E}_\Lambda(\rho_A) = \sum_{i=1}^n \Lambda_i \rho \Lambda_i \otimes |i\rangle\langle i|_X.$$

Entropic Measures:

Define the *von Neumann entropy* of a quantum state $\rho_A \in \mathcal{D}(\mathcal{H}_A)$ to be

$$H(\rho_A) = H(A)_\rho := -\operatorname{Tr} \rho_A \log \rho_A.$$

Let the spectral decomposition of ρ be

$$\rho_A = \sum_{x \in \mathcal{X}} p_x |x\rangle\langle x|_A.$$

Then $H(A)_\rho = \sum_{x\in\mathcal{X}} -p_x \log p_x := H(X)$, where $H(X)$ is the Shannon entropy of a random variable X whose distribution is $\Pr(X = x) = p_x$.

Exercise 12 *Show that $H(\rho) = 0$ if and only if ρ is pure.*

Exercise 13 *Show that $H(\rho) = \log d$ if and only if ρ is a completely mixed state I/d in $\mathcal{H}_d$.*

The quantum conditional entropy of a bipartite quantum state ρ_{AB} is defined as

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

where $H(B)_\rho$ is the von Neumann entropy of the reduced density operator $\rho_B = \mathrm{Tr}_A \rho_{AB}$.

Exercise 14 *Show that the quantum conditional entropy of a pure entangled state $|\psi\rangle_{AB}$ is negative.*

The quantum mutual information $I(A:B)_\rho$ of a quantum state ρ_{AB} is defined as

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

Lemma 15 (Data Processing Inequality) *Let $\sigma_{BR} = \mathcal{N}_{A\to B}(\rho_{AR})$. Then*

$$I(B:R)_\sigma \leq I(A:R)_\rho.$$

The conditional quantum mutual information $I(A:B|C)_\rho$ of a quantum state ρ_{ABC} is defined as

$$I(A:B|C)_\rho := H(A|C)_\rho - H(A|BC)_\rho.$$

Exercise 16 (very hard) *Show that $I(A:B|C)_\rho \geq 0$ for any ρ_{ABC}. This is the so-called* strong subadditivity.

2 Toolbox

Quantum Typicality

In this section, we will fix the distribution p_x on $\mathcal{X}$. Let $x^n := x_1 x_2 \cdots x_n$, where $x_i \in \mathcal{X}$ for each i. Let $N(x|x^n)$ denote the number of occurrences of the symbol x in $\mathcal{X}$ in the sequence x^n. The type t_{x^n} of a sequence x^n is a probability vector whose element

$$t_{x^n}(a) = \frac{1}{n} N(a|x^n) \qquad \forall a \in \mathcal{X}.$$

Define the set of sequences of type t by

$$\mathcal{T}_t^n = \{x^n : t_{x^n} = t\}.$$

Let

$$\tau_\delta = \{t : \forall a \in \mathcal{X}, |t_a - p_a| \leq \delta\}.$$

Define the δ-typical set $T_{\delta,X}^n$ be

$$\begin{aligned} T_{\delta,X}^n &= \left\{x^n : \forall a \in \mathcal{X}, \left|\frac{1}{n} N(a|x^n) - p_a\right| \leq \delta\right\} \\ &= \bigcup_{t \in \tau_\delta} \mathcal{T}_t^n. \end{aligned}$$

Lemma 17 *For any $\epsilon, \delta > 0$ and n sufficiently large,*

- $\Pr\{X^n \in \mathcal{T}_{\delta,X}^n\} \geq 1 - \epsilon$.
- $|\mathcal{T}_{\delta,X}^n| \leq 2^{n[H(X)+c\delta]}$ *for some contant c.*
- $2^{-n[H(X)+c\delta]} \leq \Pr(x^n) \leq 2^{-n[H(X)-c\delta]}$, $\forall x^n \in \mathcal{T}_{\delta,X}^n$.

Exercise 18 *Prove Lemma 17.*

Recall that a density operator can be written in terms of

$$\rho = \sum_{x \in \mathcal{X}} p_x |x\rangle\langle x|.$$

The eigenvalues $\{p_x\}$ form a probability distribution (of a random variable X) so that we can define typical sequences and so on. Moreover

$$H(\rho) = H(p) = H(X).$$

Thus, we can define the type projector

$$\Pi_t^n = \sum_{x^n \in \mathcal{T}_t^n} |x^n\rangle\langle x^n|,$$

and the δ-typical projector

$$\Pi_{\delta,\rho}^n = \sum_{x^n \in \mathcal{T}_{\delta,X}^n} |x^n\rangle\langle x^n| = \sum_{t \in \tau_\delta} \Pi_t^n.$$

Exercise 19 *For $t \in \tau_\delta$, prove that*

$$|\Pi_t^n| \geq 2^{n[H(\rho)-\eta(\delta)]}$$

where $\eta(\delta) \to 0$ as $\delta \to 0$.

Lemma 20 *For any $\epsilon, \delta > 0$ and n sufficiently large,*

- $\operatorname{Tr} \rho^{\otimes n} \Pi_{\delta,\rho}^n \geq 1 - \epsilon$.
- $|\Pi_{\delta,\rho}^n| = \operatorname{Tr} \Pi_{\delta,\rho}^n \leq 2^{n[H(\rho)+c\delta]}$ *for some contant c.*
- $2^{-n[H(\rho)+c\delta]} \Pi_{\delta,\rho}^n \leq \Pi_{\delta,\rho}^n \rho^{\otimes n} \Pi_{\delta,\rho}^n \leq 2^{-n[H(\rho)-c\delta]} \Pi_{\delta,\rho}^n$.

Exercise 21 *Prove Lemma 20.*

For any sequence $x^n \in \mathcal{T}_{\delta,X}^n$, we can permute x^n into

$$\pi(x^n) := x_\uparrow = (1, \cdots, 1, \cdots, |\mathcal{X}|, \cdots, |\mathcal{X}|).$$

where the number of occurrences of symbol a is $m_a := N(a|x^n)$. We can then define the conditional typical projector $\Pi_{\delta,\rho}^n(x_\uparrow)$

$$\Pi_{\delta,\rho}^n(x_\uparrow) = \Pi_{\delta,\rho_1}^{m_1} \otimes \Pi_{\delta,\rho_2}^{m_2} \otimes \cdots \otimes \Pi_{\delta,\rho_{|\mathcal{X}|}}^{m_{|\mathcal{X}|}},$$

where each typical projector $\Pi_{\delta,\rho_i}^{m_i}$ of ρ_i satisfies $\operatorname{Tr} \Pi_{\delta,\rho_i}^{m_i} \rho_i^{\otimes m_i} \geq 1 - |\mathcal{X}|^{-1}\epsilon$. Since $x^n \in \mathcal{T}_{\delta,X}^n$, $m_i \approx np_i$. Therefore, there exists n large enough so that $\Pi_{\delta,\rho_i}^{m_i}$, $\forall i$, are typical projectors.

We can then define the conditional typical projector for $\rho_{x^n} := \rho_{x_1} \otimes \cdots \otimes \rho_{x_n}$ as follows:

$$\Pi_{\delta,\rho_{x^n}}^n = U_\pi \Pi_{\delta,\rho}^n(x_\uparrow) U_\pi^\dagger,$$

where U_π is the unitary permuting the corresponding Hilbert spaces:

$$U_\pi \rho_{x_\uparrow} U_\pi^\dagger = \rho_{x^n}.$$

Lemma 22 *For any $\epsilon, \delta > 0$ and n sufficiently large,*

- $\operatorname{Tr} \rho_{x^n} \Pi_{\delta,\rho_{x^n}}^n \geq 1 - \epsilon$.
- $|\Pi_{\delta,\rho_{x^n}}^n| \leq 2^{n[H(B|X)+c\delta]}$ *for some contant c.*
- $2^{-n[H(B|X)+c\delta]} \Pi_{\delta,\rho_{x^n}}^n \leq \Pi_{\delta,\rho_{x^n}}^n \rho_{x^n} \Pi_{\delta,\rho_{x^n}}^n \leq 2^{-n[H(B|X)-c\delta]} \Pi_{\delta,\rho_{x^n}}^n$.

Exercise 23 *Prove Lemma 22.*

Distant Measures

I will only introduce one distant measure in this short course. You can find a few others in the literature.

Define the *trace norm* (or the ℓ_1-norm) of an Hermitian operator A to be:

$$\|A\|_1 := \operatorname{Tr} \sqrt{A^\dagger A}.$$

Exercise 24 *Let*

$$X = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix}.$$

Compute $\|X\|_1$.

Proposition 25 *The trace norm satisfies*

- *Faithfulness:* $\|A\|_1 = 0$ *if and only if A=0.*
- *Homogeneity:* $\|cA\|_1 = |c|\|A\|_1$ *for any* $c \in \mathbb{C}$.
- *Triangle Inequality:* $\|A + B\|_1 \leq \|A\|_1 + \|B\|_1$.

Exercise 26 *Let A be any Hermitian operator. Show that*

$$\|A\|_1 = \max_{-I \leq \Lambda \leq I} \operatorname{Tr} \Lambda A.$$

One of the most commonly used distant measures is called the *trace distance*. The trace distance between two density operators ρ and σ is $\|\rho - \sigma\|_1$.

Lemma 27 (Monotonicity) *The trace distance is monotone under cptp maps* $\mathcal{N}$ *:*

$$\|\mathcal{N}(\rho - \sigma)\|_1 \leq \|\rho - \sigma\|_1 \tag{5}$$

Exercise 28 *Show that if the two states* ρ *and* σ *commute, then the trace distance is equivalent to the variational distance of two probability distributions.*

Exercise 29 *Fix a quantum ensemble* $\mathcal{E} = \{(p_0, \rho_0), (p_1, \rho_1)\}$. *Show that the success probability* $P_c(\mathcal{E}) := \max_\Lambda P_c(\mathcal{E}, \Lambda)$ *is*

$$P_c(\mathcal{E}) = \frac{1}{2} + \frac{1}{2}\|p_0\rho_0 - p_1\rho_1\|_1.$$

Lemma 30 (gentle measurement) *Fix a density operator* ρ *and an operator* $0 \leq \Lambda \leq I$ *so that*

$$\operatorname{Tr} \Lambda\rho \geq 1 - \epsilon.$$

Then

$$\|\rho - \sqrt{\Lambda}\rho\sqrt{\Lambda}\|_1 \leq 2\sqrt{\epsilon}.$$

Exercise 31 *Prove Lemma 30.*

Lemma 32 *If* $\|\rho - \sigma\|_1 \leq \epsilon$, *then*

$$|H(\rho) - H(\sigma)| \leq 2\epsilon \log d + 2h(\epsilon),$$

where $h(x) = -x \log x - (1 - x) \log(1 - x)$.

Lemma 33 *If* $\|\rho_{AB} - \sigma_{AB}\|_1 \leq \epsilon$, *then*

$$|I(A:B)_\rho - I(A:B)_\sigma| \leq 6\epsilon \log d_A + 4h(\epsilon),$$

where $h(x) = -x\log x - (1-x)\log(1-x)$.

The set of generalized Pauli matrices $\{U_m\}_{m\in[d^2]}$ is defined by $U_{l\cdot d+k} = \hat{Z}_d(l)\hat{X}_d(k)$ for $k, l = 0, 1, \cdots, d-1$ and

$$\begin{aligned} \hat{X}_d(k) &= \sum_s |s\rangle\langle s+k| = \hat{X}_d(1)^k, \\ \hat{Z}_d(l) &= \sum_s e^{i2\pi sl/d}|s\rangle\langle s| = \hat{Z}_d(1)^l. \end{aligned} \tag{6}$$

The $+$ sign denotes addition modulo d.

We will always use $|\Phi_d\rangle$ to represent the d-dimensional maximally entangled state (subscript will be omitted when the dimension is clear from the context):

$$|\Phi_d\rangle^{AB} = \frac{1}{\sqrt{d}}\sum_{i=1}^{d} |i\rangle^A |i\rangle^B. \tag{7}$$

We have the following result:

$$\frac{1}{d^2}\sum_{m=1}^{d^2} (U_m \otimes I)\Phi^{AB}(U_m^\dagger \otimes I) = \pi^A \otimes \pi^B, \tag{8}$$

where $\pi^A = \pi^B = \frac{I}{d}$.

Exercise 34 *Prove Eq. (8).*

We will also need the following equality:

$$(I \otimes U)|\Phi\rangle = (U^{tr} \otimes I)|\Phi\rangle \tag{9}$$

for any operator U, and U^{tr} denotes transposition of U.

Exercise 35 *Prove Eq. (9).*

3 Source Coding: Schumacher Compression

For a quantum source $\rho_A \in \mathcal{H}_A$ with purification $|\psi^\rho\rangle_{AR}$, we define an (n, R, ϵ) source code by

- compression operation $\mathcal{E} : \mathcal{H}_d^{\otimes n} \mapsto \mathcal{H}_{2^{nR}}$;
- decompression operation $\mathcal{D} : \mathcal{H}_{2^{nR}} \mapsto \mathcal{H}_d^{\otimes n}$

so that

$$\left\|(\psi_{AR}^\rho)^{\otimes n} - \mathcal{D}\circ\mathcal{E}\left((\psi_{AR}^\rho)^{\otimes n}\right)\right\|_1 \leq \epsilon.$$

We call R is *achievable* if for any $\delta, \epsilon > 0$, there exists an $(n, R+\delta, \epsilon)$ source code. Define $\mathcal{C}(\rho) = \inf\{R : R \text{ is achievable}\}$.

Theorem 36 (Quantum Data Compression [Sch95])

$$\mathcal{C}(\rho) = H(\rho).$$

Direct Coding Theorem. Let the spectral decomposition of $\rho = \sum_{x\in\mathcal{X}} p_x |x\rangle\langle x|$. Shorthand $\psi^n_{AR} \equiv (\psi^\rho_{AR})^{\otimes n}$, $\Pi_0 \equiv \Pi^n_{\delta,\rho}$ and $\Pi_1 \equiv I - \Pi^n_{\delta,\rho}$. Note that $\{\Pi_0, \Pi_1\}$ forms a projective measurement. We can construct the compression operator $\mathcal{E}$ as the composition of the following operations:

$$\mathcal{E}_1(\rho^{\otimes n}) := \sigma_1 = \Pi_0\rho^{\otimes n}\Pi_0 \otimes |0\rangle\langle 0|_X + \Pi_1\rho^{\otimes n}\Pi_1 \otimes |1\rangle\langle 1|_X$$

$$\mathcal{E}_2(\sigma_1) := \sigma_2 = \Pi_0\rho^{\otimes n}\Pi_0 \otimes |0\rangle\langle 0|_X + \mathrm{Tr}(\Pi_1\rho^{\otimes n})|\perp\rangle\langle\perp| \otimes |1\rangle\langle 1|_X.$$

The decompression operation $\mathcal{D}$ simply discards the classical system X:

$$\mathcal{D}(\sigma_2) := \sigma = \Pi_0\rho^{\otimes n}\Pi_0 + \mathrm{Tr}(\Pi_1\rho^{\otimes n})|\perp\rangle\langle\perp|.$$

We can verify that

$$\begin{aligned}
\|\psi^n_{AR} - \mathcal{D}\circ\mathcal{E}(\psi^n_{AR})\|_1 &\leq \|\psi^\rho_{AR} \otimes |0\rangle\langle 0|_X - \mathcal{E}(\psi^n_{AR})\|_1 \\
&\leq \|(\psi^n_{AR} - \Pi_0\psi^n_{AR}\Pi_0) \otimes |0\rangle\langle 0|_X\|_1 + \|\,\mathrm{Tr}(\Pi_1\rho^{\otimes n})|\perp\rangle\langle\perp|\,\|_1 \\
&\leq 2\sqrt{\epsilon} + \epsilon
\end{aligned}$$

where the first inequality follows from monotonicity of trace distance (Lemma 27); the second inequality follows from the triangle inequality for trace distance; the third inequality uses the gentle measurement lemma (Lemma 30) and quantum typicality $\mathrm{Tr}\,\Pi_1\rho^{\otimes n} \leq \epsilon$.

■

Converse. For any $(n, R+\delta, \epsilon)$ source code with $\mathcal{E} : A^n \to W$ and $\mathcal{D} : W \to A$ with $|W| = 2^{nR}$, let

$$\omega_{\hat{A}^n R^n} = \mathcal{D}(\sigma^n_{WA}),$$

where

$$\sigma_{WR^n} := \mathcal{E}(\psi^n_{AR}),$$

so that

$$\|\psi^n_{AR} - \omega^n_{AR}\|_1 \leq \epsilon.$$

Then

$$\begin{aligned}
2nR &\geq |H(W)_\sigma| + |H(W|R^n)_\sigma| \\
&\geq |H(W)_\sigma - H(W|R^n)_\sigma| \\
&= I(W : R^n)_\sigma \\
&\geq I(\hat{A}^n : R^n)_\omega \\
&\geq I(A^n : R^n)_\psi - n\epsilon' \\
&= 2H(A^n)_\phi - n\epsilon' \\
&= 2nH(\rho) - n\epsilon'.
\end{aligned}$$

The fourth line follows from data processing inequality (Lemma 15). The fifth line follows from the continuity of the mutual information (Lemma 33).

■

4 Channel Coding: HSW Theorem

The packing lemma below will prove to be a powerful tool in quantum information theory. The technique used here is simple, directly analogous to the classical coding theorem.

Lemma 37 (Packing [HDW08]) *We are given an ensemble $\{\lambda_m, \sigma_m\}_{m\in\mathcal{S}}$ with average density operator*

$$\sigma = \sum_{m\in\mathcal{S}} \lambda_m \sigma_m.$$

Assume the existence of projectors Π and $\{\Pi_m\}_{m\in\mathcal{S}}$ with the following properties:

$$\begin{aligned} \operatorname{Tr} \sigma_m \Pi_m &\geq 1-\epsilon, && (10)\\ \operatorname{Tr} \sigma_m \Pi &\geq 1-\epsilon, && (11)\\ \operatorname{Tr} \Pi_m &\leq d, && (12)\\ \Pi\sigma\Pi &\leq D^{-1}\Pi && (13) \end{aligned}$$

for all $m \in \mathcal{S}$ and some positive integers D and d. Let $N = \lfloor \gamma D/d \rfloor$ for some $0 < \gamma < 1$ where $\lfloor r \rfloor$ represents the largest integer less than r. Then there exists a map $f : [N] \to \mathcal{S}$, and a corresponding set of POVMs $\{\Lambda_k\}_{k\in[N]}$ which reliably distinguishes between the states $\{\sigma_{f(k)}\}_{k\in[N]}$ in the sense that

$$\operatorname{Tr} \sigma_{f(k)} \Lambda_k \geq 1 - 4(\varepsilon + \sqrt{8\varepsilon}) - 8\gamma$$

for all $k \in [N]$.

Proof. Let X^N denote a sequence of random variables $X_1, X_2, \ldots, X_N$, where each random variable X_k takes values from $\mathcal{S}$ and is distributed according to λ. Set $f(k) = X_k$. Each random code $C = \{\sigma_{x_k}\}_{k\in[N]}$ is generated according to $X_k = x_k$. Define $p_e(k)$ to be the probability of error for a single codeword σ_{x_k}:

$$p_e(k) = \operatorname{Tr} \sigma_{x_k}(I - \Lambda_k),$$

where the POVM elements $\{\Lambda_k\}$ are constructed by the so-called *square root measurement* [Hol98, SW97]

$$\Lambda_k = \left(\sum_{l=1}^{N} \Upsilon_{x_l}\right)^{-\frac{1}{2}} \Upsilon_{x_k} \left(\sum_{l=1}^{N} \Upsilon_{x_l}\right)^{-\frac{1}{2}}$$

with

$$\Upsilon_m = \Pi\Pi_m\Pi.$$

Define $p_e(C)$ to be the average probability of error, averaged over all codewords in C:

$$p_e(C) = \frac{1}{N}\sum_{k=1}^{N} p_e(k).$$

Define $\bar{p}_e$ to be the average probability of error, averaged over all possible random codes C to be:

$$\bar{p}_e = \mathbb{E}_{X^N}\left[p_e(C)\right].$$

The idea here is that if the average probability of error $\bar{p}_e$ is small enough, we can then show the existence of at least one good code. In what follows, we will first show that $\bar{p}_e \leq \epsilon'$ for some $\epsilon' \to 0$

when $n \to \infty$.

Invoking Lemma 38, we can now place an upper bound on $p_e(C)$:

$$p_e(C) \leq \frac{1}{N} \sum_{k=1}^{N} \left[2(1 - \operatorname{Tr} \sigma_{x_k} \Upsilon_{x_k}) + 4 \sum_{l \neq k} \operatorname{Tr} \sigma_{x_k} \Upsilon_{x_l} \right]. \tag{14}$$

The gentle operator lemma (Lemma 30) and property (11) give

$$\|\Pi \sigma_m \Pi - \sigma_m\| \leq \sqrt{8\epsilon}. \tag{15}$$

By property (10) and (15)

$$\begin{aligned} \operatorname{Tr} \sigma_m \Upsilon_m &\geq \operatorname{Tr} \sigma_m \Pi_m - \|\Pi \sigma_m \Pi - \sigma_m\| \\ &\geq 1 - \epsilon - \sqrt{8\epsilon}. \end{aligned} \tag{16}$$

For $k \neq l$, the random variables X_k and X_l are independent. Thus

$$\begin{aligned} \mathbb{E}_{X^N} \left[\operatorname{Tr} \sigma_{X_k} \Upsilon_{X_l} \right] &= \operatorname{Tr} \left(\Pi \mathbb{E} \sigma_{X_k} \Pi \; \mathbb{E} \Pi_{X_l} \right) \\ &\leq D^{-1} \mathbb{E} \operatorname{Tr} \Pi \Pi_{X_l} \\ &\leq d/D. \end{aligned} \tag{17}$$

The first inequality follows from $\mathbb{E}\, \sigma_{X_k} = \sigma$ and property (12). The second follows from $\Pi \leq \mathbf{1}$ and property (13). Taking the expectation of (14), and incorporating (16) and (17) gives

$$\begin{aligned} \bar{p}_e &\leq 2(\varepsilon + \sqrt{8\varepsilon}) + 4(N-1)d/D, \\ &\leq 2(\varepsilon + \sqrt{8\varepsilon}) + 4Nd/D \\ &= 2(\varepsilon + \sqrt{8\varepsilon}) + 4\gamma =: \epsilon'. \end{aligned} \tag{18}$$

Two more standard steps are needed.

i) Derandomization. There exists at least one particular value x^N of the string X^N such that this code $C^* = \{\sigma_{x^N}\}$ for which $p_e(C^*)$ is at least as small as the expectation value. Thus

$$p_e(C^*) \leq \varepsilon'. \tag{19}$$

ii) Average to maximal error probability. Since

$$p_e(C^*) = \frac{1}{N} \sum_{k \in N} p_e(k) \leq \varepsilon',$$

then $p_e(k) \leq 2\varepsilon'$ for at least half the indices k. Throw the others away and redefine f, N and γ accordingly. This further changes the error estimate to $4(\varepsilon + \sqrt{8\epsilon}) + 8\gamma$.

■

Lemma 38 (Hayashi, Nagaoka [HN03]) *For any operators $0 \leq S \leq \mathbf{1}$ and $T \geq 0$, we have*

$$\mathbf{1} - \sqrt{S+T}^{-1} S \sqrt{S+T}^{-1} \leq 2(\mathbf{1} - S) + 4T.$$

Classical Communication

For a quantum channel $\mathcal{N} : A \to B$, we define an (n, R, ϵ) channel code by

- an encoding operation $\mathcal{E} : X \equiv \{1, 2, \cdots, 2^{nR}\} \to A$;
- a decoding POVM $\Lambda : B \to \{1, 2, \cdots, 2^{nR}\} \equiv \widehat{X}$

so that

$$\Pr\{X \neq \widehat{X}\} \leq \epsilon.$$

We say that the rate R is *achievable* if for any $\epsilon, \delta > 0$ there exists an $(n, R - \delta, \epsilon)$ channel code. We define the classical capacity over the quantum channel $\mathcal{N}$:

$$\mathcal{C}(\mathcal{N}) = \sup\{R : R \text{ is achievable}\}.$$

Define the Holevo quantity of a quantum channel $\mathcal{N} : A \to B$:

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

where

$$\rho_{XB} = \sum_x p_x |x\rangle\langle x|_X \otimes \mathcal{N}_{A\to B}(\psi_x^A).$$

Denote

$$\chi_r(\mathcal{N}) = \lim_{n\to\infty} \frac{1}{n} \chi(\mathcal{N}^{\otimes n}).$$

Theorem 39 (HSW theorem [Hol98, SW97])

$$\mathcal{C}(\mathcal{N}) = \chi_r(\mathcal{N}).$$

Direct Coding Theorem. Fix any ensemble $\{p_x, \rho_x\}$. We construct a new ensemble $\{p'_{x^n}, \rho_{x^n}\}$, where

$$p'_{x^n} = \begin{cases} \frac{p_{x^n}}{\Pr[\mathcal{T}^n_{\delta,X}]}, & \text{if } x^n \in \mathcal{T}^n_{\delta,X} \\ 0, & otherwise \end{cases}.$$

It is not hard to verify that

$$\|p' - p\|_1 \leq 2\epsilon.$$

We can now apply packing lemma on the ensemble $\mathcal{S} = \{p'_{x^n}, \sigma_{x^n}\}$ to prove the direct coding theorem, where

$$\sigma_{x^n} = \mathcal{N}(\rho_{x_1}) \otimes \mathcal{N}(\rho_{x_2}) \otimes \cdots \otimes \mathcal{N}(\rho_{x_n}).$$

Note that

$$\mathbb{E}[\mathcal{S}] := \bar{\sigma} = \sum_{x^n} p'_{x^n} \sigma_{x^n}.$$

We also have

$$\|\bar{\sigma} - \sigma^{\otimes n}\|_1 \leq 2\epsilon$$

where $\sigma := \mathcal{N}(\sum_x p_x \rho_x)$. The projectors of Π and $\{\Pi_m\}$ in the packing lemma are chosen as follows:

$$\begin{aligned} \Pi &\equiv \Pi^n_{\delta,\sigma} \\ \Pi_m &\equiv \Pi^n_{\delta,\sigma_{x^n}}. \end{aligned}$$

Then by the properties of (conditional) typical projectors

$$\begin{aligned}
\operatorname{Tr}\Pi^n_{\delta,\sigma_{x^n}}\sigma_{x^n} &\geq 1-\epsilon \\
\operatorname{Tr}\Pi^n_{\delta,\sigma}\sigma^{\otimes n} &\geq 1-\epsilon \\
\operatorname{Tr}\Pi^n_{\delta,\sigma_{x^n}} &\leq 2^{n[H(B|X)_\sigma+c\delta]} \\
\Pi^n_{\delta,\sigma}\bar{\sigma}\Pi^n_{\delta,\sigma} &\leq (1-\epsilon)^{-1}2^{-n[H(B)_\sigma-c\delta]}\Pi^n_{\delta,\sigma},
\end{aligned}$$

where

$$\begin{aligned}
d &= 2^{n[H(B|X)_\sigma+c\delta]} \\
D &= (1-\epsilon)^{-1}2^{-n[H(B)_\sigma-c\delta]}.
\end{aligned}$$

Choosing $N = 2^{n[I(X:B)_\sigma-3c\delta]}$ and $\gamma = 2^{-nc\delta}$. The error probability is

$$p_e \leq 2(\epsilon+\sqrt{8\epsilon}) + 4\times 2^{-nc\delta} \to_{n\to\infty} 0. \tag{20}$$

■

Converse. Here we can use a simple trick. Instead of proving the converse for classical capacity, we prove a converse for common randomness generation. Since classical communication can be used to generate common randomness, hence the capacity of common randomness generation can only be larger than the classical capacity $\mathcal{C}(\mathcal{N})$.

The general protocol for common randomness generation begins with Alice preparing a maximally correlated state

$$\bar{\Phi}_{MM'} = \frac{1}{2^{nR}}\sum_{i=1}^{2^{nR}} |ii\rangle\langle ii|.$$

After her encoding and sending through $\mathcal{N}$, Bob performs his decoding measurement on the channel output B^n of the sate σ_{MB^n} to recover a state $\omega_{MM'}$ that is ϵ-close to $\bar{\Phi}_{MM'}$:

$$\|\omega_{MM'} - \bar{\Phi}_{MM'}\|_1 \leq \epsilon.$$

Then

$$\begin{aligned}
nR &= I(M:M')_{\bar{\Phi}} \\
&\leq I(M:M')_\omega + n\epsilon' \\
&\leq I(M:B^n)_\sigma + n\epsilon' \\
&\leq \chi(\mathcal{N}^{\otimes n}) + n\epsilon'.
\end{aligned}$$

The first inequality follows from continuity of mutual information (Lemma 33). The second inequality follows from data processing inequality (Lemma 15). The final inequality follows from the definition of Holevo χ quantity.

■

5 Entanglement-assisted Classical Coding

We first describe a general entanglement-assisted classical communication protocol. Alice and Bob are connected by a large number n uses of the quantum channel $\mathcal{N} : A' \to B$. Alice controls the channel input system A' and Bob has access to the channel output B. They also have entanglement in the form of n copies of some pure bipartite state $\varphi_{A'B'}$. Any such state is determined upto a local unitary transformation by the local density operator $\rho^{A'} = \mathrm{Tr}_{B'} \varphi_{A'B'}$. Alice and Bob use these resources to communicate, in analogy to superdense coding. Based on her message Alice performs a quantum operation on her share of the entanglement. She then sends it through the quantum channel. Bob performs a decoding measurement on the channel output plus his share of the entanglement. They endeavor to maximize the communication rate.

We define an (n, R, ε) entanglement-assisted code by

- a set of unitary encoding maps $\{\mathcal{E}_k\}_{k\in[2^{nR}]}$ acting on $A'^n := A'_1 \dots A'_n$ for Alice;
- Bob's decoding measurement $\Lambda = \{\Lambda_k\}_{k\in[2^{nR}]}$ acting on $B^n B'^n$.

such that for all $k \in [2^{nR}]$

$$\mathrm{Tr}[\Lambda_k((\mathcal{N}^{\otimes n} \circ \mathcal{E}_k) \otimes I)(\varphi^{\otimes n})] \geq 1 - \epsilon.$$

We say that the rate R is achievable if for any $\epsilon, \delta > 0$ there exists an $(n, R-\delta, \epsilon)$ entanglement-assisted code. Define the entanglement-assisted classical capacity of a channel $\mathcal{N}$

$$\mathcal{C}_{\mathrm{ea}}(\mathcal{N}) = \sup\{R : R \text{ is achievable.}\}.$$

Define

$$I(\mathcal{N}_{A'\to B}) = \max_{\varphi_{AA'}} I(A : B)_\sigma$$

where $\sigma_{AB} = \mathcal{N}(\varphi_{AA'})$.

Theorem 40 (Entanglement-assisted Capacity [BSST02])

$$\mathcal{C}_{\mathrm{ea}}(\mathcal{N}) = I(\mathcal{N}).$$

Direct Coding Theorem. The proof can be done using the packing lemma. However, it requires further manipulation. The following proof comes from [HDW08].

Let the size of distinct types be T, and $t_1, \dots, t_T$ be an ordering of the types t_{x^n}. For each type t_α, we denote the size of its type class $d_\alpha = |\mathcal{T}^n_{t_\alpha}|$, and denote its type projector $\Pi^n_{t_\alpha}$. Define $|\Phi_\alpha\rangle$ to be the maximally entangled state on a pair of d_α-dimensional quantum systems A'^n and B'^n

$$|\Phi_\alpha\rangle_{A'^n B'^n} = \frac{1}{\sqrt{d_\alpha}} \sum_{x^n \in \mathcal{T}^n_{t_\alpha}} |x^n\rangle_{A'^n} |x^n\rangle_{B'^n}. \tag{21}$$

The maximally mixed state

$$\pi_\alpha = \frac{1}{d_\alpha} \Pi^n_{t_\alpha}.$$

Note that Alice and Bob' s preshared entangled state admits the following decomposition:

$$|\varphi\rangle^{\otimes n} := |\Psi\rangle_{A'^n B'^n} = \sum_\alpha \sqrt{p_\alpha} |\Phi_\alpha\rangle, \tag{22}$$

where $p_\alpha = \sum_{x^n \in \mathcal{T}^n_{t_\alpha}} p_{x^n}$. The distinct types induce a decomposition of the Hilbert space $\mathcal{H}^{\otimes n}$ of A'^n (correspondingly of B'^n) into a direct sum

$$\mathcal{H}^{\otimes n} = \bigoplus_{\alpha=1}^{T} \mathcal{H}_{t_\alpha}.$$

Let $\mathcal{G} = \{(g_1, g_2, \cdots, g_T) : g_\alpha \in \{1, \cdots, d_\alpha^2\}, \alpha \in \{1, \cdots, T\}\}$, $\mathcal{B} = \{(b_1, b_2, \cdots, b_T) : b_\alpha \in \{0, 1\}\}$, and $\mathcal{S} = \mathcal{G} \times \mathcal{B}$. Every element $\vec{s} \in \mathcal{S}$ is uniquely determined by $\vec{g} \in \mathcal{G}$ and $\vec{b} \in \mathcal{B}$. Define a unitary operation $U_{\vec{s}}$ for each $\vec{s} \in \mathcal{S}$ to be

$$U_{\vec{s}} \equiv U_{\vec{g},\vec{b}} = \bigoplus_{\alpha=1}^{T} (-1)^{b_\alpha} U_{g_\alpha} \tag{23}$$

where $\{U_{g_\alpha}\}$ are the d_α^2 generalized Pauli operators (6) defined on $\mathcal{H}_{t_\alpha}$. Define

$$\begin{aligned} \sigma_{\vec{s}}^{B^n B'^n} &:= (\mathcal{N}^{\otimes n} \otimes I) \left[(U_{\vec{s}} \otimes I) \Psi_{A'^n B'^n} (U_{\vec{s}}^\dagger \otimes I) \right] \\ &= (I \otimes U_{\vec{s}}^{tr}) \theta^{\otimes n} (I \otimes U_{\vec{s}}^*), \end{aligned} \tag{24}$$

where

$$\theta = \mathcal{N}(\varphi_{A'B'}).$$

The last equality follows from (9).

Consider the ensemble $\{1/|\mathcal{S}|, \sigma_{\vec{s}}\}_{\vec{s} \in \mathcal{S}}$. Let σ be the average state of the ensemble, then

$$\begin{aligned} \sigma &= \frac{1}{|\mathcal{S}|} \sum_{\vec{s} \in \mathcal{S}} \sigma_{\vec{s}} \\ &= \frac{1}{|\mathcal{B}||\mathcal{G}|} \sum_{\vec{g} \in \mathcal{G}} \sum_{\vec{b} \in \mathcal{B}} \sum_{\alpha, \alpha'} \sqrt{p_\alpha p_{\alpha'}} (\mathcal{N}^{\otimes n} \otimes I) \left[(U_{\vec{g},\vec{b}} \otimes I) |\Phi_\alpha\rangle\langle\Phi_{\alpha'}| (U_{\vec{g},\vec{b}}^\dagger \otimes I) \right]. \\ &= \sum_\alpha p_\alpha \left(\mathcal{N}^{\otimes n}(\pi_\alpha^n) \otimes \pi_\alpha^n \right). \end{aligned} \tag{25}$$

The last equality comes from (26) and (27) below. When $\alpha = \alpha'$,

$$\begin{aligned} &\frac{1}{|\mathcal{B}||\mathcal{G}|} \sum_{\vec{g} \in \mathcal{G}} \sum_{\vec{b} \in \mathcal{B}} p_\alpha (\mathcal{N}^{\otimes n} \otimes I) \left[(U_{\vec{g},\vec{b}} \otimes I) \Phi_\alpha (U_{\vec{g},\vec{b}}^\dagger \otimes I) \right] \\ &= (\mathcal{N}^{\otimes n} \otimes I) \frac{1}{|\mathcal{G}|} \sum_{g_1} \cdots \sum_{g_T} p_\alpha (U_{g_\alpha} \otimes I) \Phi_\alpha (U_{g_\alpha}^\dagger \otimes I) \\ &= (\mathcal{N}^{\otimes n} \otimes I) p_\alpha (\pi_\alpha^n \otimes \pi_\alpha^n). \end{aligned} \tag{26}$$

The last equality follows from (8). When $\alpha \neq \alpha'$, we get (27):

$$\begin{aligned} &\frac{1}{|\mathcal{B}||\mathcal{G}|} \sum_{\vec{g} \in \mathcal{G}} \sum_{\vec{b} \in \mathcal{B}} \sqrt{p_\alpha p_{\alpha'}} (\mathcal{N}^{\otimes n} \otimes I) \left[(U_{\vec{g},\vec{b}} \otimes I) |\Phi_\alpha\rangle\langle\Phi_{\alpha'}| (U_{\vec{g},\vec{b}}^\dagger \otimes I) \right] \\ &= \frac{1}{d_\alpha^2 d_{\alpha'}^2} \sqrt{p_\alpha p_{\alpha'}} \sum_{b_\alpha b_{\alpha'}} \frac{(-1)^{b_\alpha + b_{\alpha'}}}{4} \left\{ \sum_{g_\alpha g_{\alpha'}} (\mathcal{N}^{\otimes n} \otimes I) \left[(U_{g_\alpha} \otimes I) |\Phi_\alpha\rangle\langle\Phi_{\alpha'}| (U_{g_{\alpha'}}^\dagger \otimes I) \right] \right\} \\ &= 0. \end{aligned} \tag{27}$$

Define the projectors on $B'^n B^n$

$$\begin{aligned} \Pi_{\vec{s}} &:= (I \otimes U_{\vec{s}}^{tr})\, \Pi_{\delta,\theta}^n\, (I \otimes U_{\vec{s}}^*), && (28)\\ \Pi &:= \Pi_{\delta,\mathcal{N}(\rho)}^n \otimes \Pi_{\delta,\rho}^n. && (29) \end{aligned}$$

For all $\epsilon > 0, \delta > 0$ and all sufficiently large n,

$$\begin{aligned} \operatorname{Tr} \sigma_{\vec{s}} \Pi_{\vec{s}} &\geq 1-\epsilon && (30)\\ \operatorname{Tr} \sigma_{\vec{s}} \Pi &\geq 1-\epsilon && (31)\\ \operatorname{Tr} \Pi_{\vec{s}} &\leq 2^{n[H(AB)_\theta + c\delta]} && (32)\\ \Pi\sigma\Pi &\leq 2^{-n[H(A)_\theta + H(B)_\theta - c\delta]}\Pi. && (33) \end{aligned}$$

Let $\lambda_{\vec{s}} = \frac{1}{|\mathcal{S}|}$ and $R = I(A:B)_\theta - (2c+1)\delta$. We now apply the packing lemma to the ensemble $\{\lambda_{\vec{s}}, \sigma_{\vec{s}}\}_{\vec{s}\in\mathcal{S}}$ and projectors Π and $\Pi_{\vec{s}}$. Thus there exist a map $f : [2^{nR}] \to \mathcal{S}$ and a POVM $\{\Lambda_k\}_{k\in[2^{nR}]}$ such that

$$\operatorname{Tr} \sigma_{f(k)} \Lambda_k \geq 1 - \epsilon', \tag{34}$$

with

$$\epsilon' = 4(\epsilon + \sqrt{8\epsilon}) + 16 \times 2^{-n\delta}.$$

Proofs of properties (30)-(33).

I. Eq. (30): By (24) and (28),

$$\begin{aligned} \operatorname{Tr} \sigma_{\vec{s}} \Pi_{\vec{s}} &= \operatorname{Tr} \theta^{\otimes n} \Pi_{\delta,\theta}^n \\ &\geq 1 - \epsilon. \end{aligned} \tag{35}$$

The last line follows since $\Pi_{\delta,\theta}^n$ is the δ-typical projector of θ.

II. Eq. (31): Shorthand $\check{P} = I - P$. Then

$$\begin{aligned} \Pi &= \Pi_{\delta,\mathcal{N}(\rho)}^n \otimes \Pi_{\delta,\rho}^n \\ &\geq I \otimes I - I \otimes \check{\Pi}_{\delta,\rho}^n - \check{\Pi}_{\delta,\mathcal{N}(\rho)}^n \otimes I. \end{aligned} \tag{36}$$

We have

$$\begin{aligned} &\operatorname{Tr} \sigma_{\vec{s}} \Pi \\ &\geq \operatorname{Tr} \sigma_{\vec{s}} - \operatorname{Tr} \sigma_{\vec{s}} (I \otimes \check{\Pi}_{\delta,\rho}^n) - \operatorname{Tr} \sigma_{\vec{s}} (\check{\Pi}_{\delta,\mathcal{N}(\rho)}^n \otimes I) \\ &= 1 - \operatorname{Tr}[\rho^{\otimes n} \check{\Pi}_{\delta,\rho}^n] - \operatorname{Tr}[\mathcal{N}(\rho)^{\otimes n} \check{\Pi}_{\delta,\mathcal{N}(\rho)}^n] \\ &\geq 1 - 2\epsilon. \end{aligned} \tag{37}$$

III. Eq. (32): This follows directly from the property of quantum typicality.

$$\operatorname{Tr} \Pi_{\vec{s}} = \operatorname{Tr} \Pi_{\delta,\theta}^n \leq 2^{n[H(AB)_\theta + c\delta]}. \tag{38}$$

IV. Eq. (33): From Exercise (19), we can bound the density operator π_α by

$$\pi_\alpha = \frac{\Pi_{t_\alpha}^n}{\operatorname{Tr} \Pi_{t_\alpha}^n} \leq 2^{-n[H(\rho) - \eta(\delta)]} \Pi_{\delta,\rho}^n. \tag{39}$$

Then

$$\begin{aligned}
\Pi\sigma\Pi &= \left(\Pi^n_{\delta,\mathcal{N}(\rho)} \otimes \Pi^n_{\delta,\rho}\right) \left[\sum_\alpha p_\alpha (\mathcal{N}^{\otimes n}(\pi_\alpha) \otimes \pi_\alpha)\right] \left(\Pi^n_{\delta,\mathcal{N}(\rho)} \otimes \Pi^n_{\delta,\rho}\right) \\
&= \sum_\alpha p_\alpha \left[\left(\Pi^n_{\delta,\mathcal{N}(\rho)} \mathcal{N}^{\otimes n}(\pi_\alpha) \Pi^n_{\delta,\mathcal{N}(\rho)}\right) \otimes \left(\Pi^n_{\delta,\rho} \pi_\alpha \Pi^n_{\delta,\rho}\right)\right] \\
&\leq \left(\Pi^n_{\delta,\mathcal{N}(\rho)} \mathcal{N}^{\otimes n} \left(\sum_\alpha p_\alpha \pi_\alpha\right) \Pi^n_{\delta,\mathcal{N}(\rho)}\right) \otimes \left(2^{-n[H(\rho)-\eta(\delta)]} \Pi^n_{\delta,\rho}\right) \\
&\leq \left(2^{-n[H(\mathcal{N}(\rho))-c\delta]} \Pi^n_{\delta,\mathcal{N}(\rho)}\right) \otimes \left(2^{-n[H(\rho)-\eta(\delta)]} \Pi^n_{\delta,\rho}\right) \\
&= 2^{-n[H(\rho)+H(\mathcal{N}(\rho))-c\delta-\eta(\delta)]}\, \Pi \\
&= 2^{-n[H(A)_\theta+H(B)_\theta-c\delta-\eta(\delta)]}\, \Pi
\end{aligned}$$

where the first inequality follows from (39) and the second inequality follows since $\sum_\alpha p_\alpha \pi_\alpha = \rho^{\otimes n}$.

■

Converse. It suffices to prove a converse for the entanglement-assisted common randomness generation. In this protocol, Alice prepares a common randomness state $\bar{\Phi}_{MM'}$ of size 2^{nR}, and performs an encoding operation before sending through the channel $\mathcal{N}$. Bob then performs a decoding POVM on the channel output B and his half preshared entangled system T_B of $\sigma_{MT_BB^n}$ to generate $\omega_{MM'}$ so that

$$\|\omega_{MM'} - \bar{\Phi}_{MM'}\|_1 \leq \epsilon.$$

■

Then

$$\begin{aligned}
nR &= I(M:M')_{\bar{\Phi}} \\
&\leq I(M:M')_\omega + n\epsilon' \\
&\leq I(M:B^nT_B)_\sigma + n\epsilon' \\
&= I(T_BM:B^n)_\sigma + I(M:T_B)_\sigma - I(B^n:T_B)_\sigma + n\epsilon' \\
&\leq I(T_BM:B^n)_\sigma + n\epsilon' \\
&\leq \max_\sigma I(T_BM:B^n)_\sigma + n\epsilon' \\
&= I(\mathcal{N}^{\otimes n}) + n\epsilon' \\
&= nI(\mathcal{N}) + n\epsilon'.
\end{aligned}$$

The first inequality follows from the continuity of mutual information (Lemma 33). The second inequality uses data processing inequality (Lemma 15). The third inequality follows since $I(M : T_B)_\sigma = 0$ and $I(B^n : T_B)_\sigma \geq 0$. The second last line uses the result in Exercise 41. The last line follows since the quantity $I(\mathcal{N}^{\otimes n}) = nI(\mathcal{N})$ is additive.

Exercise 41 *Denote* $\sigma_{XAB} = \sum_x p_x |x\rangle\langle x|_X \otimes \mathcal{N}_{A'\to B}(\varphi^{AA'}_x)$. *Show that*

$$\max_\sigma I(XA:B)_\sigma = I(\mathcal{N}).$$

Exercise 42 *Show that*

$$I(\mathcal{N}_1 \otimes \mathcal{N}_2) = I(\mathcal{N}_1) + I(\mathcal{N}_2).$$

6 Private Coding

The core technical tool for proving the private capacity is the following covering lemma. The following explicit form of covering lemma first appeared in Ref. [DHW06].

Covering Lemma

We first prove a quantum generalization of the covering lemma.

Lemma 43 (Covering lemma) *We are given an ensemble $\{p_x, \sigma_x\}_{x\in\mathcal{X}}$ with average density operator $\sigma = \sum_{x\in\mathcal{X}} p_x\sigma_x$. Assume the existence of projectors Π and $\{\Pi_x\}_{x\in\mathcal{X}}$ with the following properties ($\forall\ x \in \mathcal{X}$):*

$$\begin{aligned} \operatorname{Tr}\sigma_x\Pi_x &\geq 1-\epsilon, && (40)\\ \operatorname{Tr}\sigma_x\Pi &\geq 1-\epsilon, && (41)\\ \operatorname{Tr}\Pi &\leq D, && (42)\\ \Pi_x\sigma_x\Pi_x &\leq d^{-1}\Pi_x. && (43) \end{aligned}$$

In addition, we require Π_x and σ_x to commute for all x. The obfuscation error of a set $\mathcal{S} \subseteq \mathcal{X}$ is defined as

$$oe(\mathcal{S}) = \left\| \frac{1}{|\mathcal{S}|}\sum_{x\in\mathcal{S}}\sigma_x - \sigma \right\|_1.$$

Define the set $\mathcal{C} = \{X_s\}_{s\in[N]}$, where X_s is a random variable chosen independently according to the distribution p on $\mathcal{X}$, and $N = \lceil\gamma^{-1}D/d\rceil$ for some $0 < \gamma < 1$. Then

$$\Pr\{oe(\mathcal{C}) \geq 3\epsilon + 19\sqrt{\epsilon}\} \leq 2D\exp\left(-\frac{\epsilon^3}{2\ \ln 2\gamma}\right). \quad (44)$$

Proof. The proof of the covering lemma involves the following steps.

1. Define $\sigma'_x = \Pi_x\sigma_x\Pi_x$. Since σ_x and Π_x commute, (40) implies

$$\|\sigma_x - \sigma'_x\|_1 \leq \epsilon.$$

2. Define $\omega'_x = \Pi\sigma'_x\Pi$. Then (41) and Exercise 26 give

$$\begin{aligned} \operatorname{Tr}\omega'_x &= \operatorname{Tr}\Pi\sigma'_x \\ &\geq \operatorname{Tr}\Pi\sigma_x - \|\sigma_x - \sigma'_x\|_1 \\ &\geq 1 - 2\epsilon. \end{aligned} \quad (45)$$

   Furthermore, the gentle measurement lemma (Lemma 30) gives

$$\|\omega'_x - \sigma'_x\|_1 \leq \sqrt{16\epsilon}.$$

   Applying the triangle inequality, we have

$$\begin{aligned} \|\omega'_x - \sigma_x\|_1 &\leq \|\omega'_x - \sigma'_x\|_1 + \|\sigma'_x - \sigma_x\|_1 \\ &\leq \epsilon + \sqrt{16\epsilon}. \end{aligned} \quad (46)$$

3. Define $\omega' = \sum_{x\in\mathcal{X}} p(x)\omega'_x$. Let $\hat{\Pi}$ be the projector onto the subspace spanned by the eigenvectors of ω' with eigenvalue $\geq \epsilon D^{-1}$. Define $\omega_x = \hat{\Pi}\omega'_x\hat{\Pi}$ and $\omega = \hat{\Pi}\omega'\hat{\Pi}$. Since (42) implies that the support of ω' has dimension $\leq D$, eigenvalues smaller than ϵD^{-1} contribute at most ϵ to $\operatorname{Tr}\omega'$. Together with (45) thus gives

$$\operatorname{Tr}\omega \geq \operatorname{Tr}\omega' - \epsilon \geq 1 - 3\epsilon. \tag{47}$$

Furthermore, the gentle measurement lemma (Lemma 30) gives

$$\|\omega - \omega'\|_1 \leq \sqrt{24\epsilon}. \tag{48}$$

4. Consider the operator ensemble $\{p_x, d\omega_x\}_{x\in\mathcal{X}}$. The expectation value of this ensemble is

$$\begin{aligned}\sum_{x\in\mathcal{X}} p_x d\omega_x &= d\left(\hat{\Pi}\sum_{x\in\mathcal{X}} p_x\omega'_x\hat{\Pi}\right)\\ &= d\hat{\Pi}\omega'\hat{\Pi}\\ &\geq tI,\end{aligned}$$

where $t = \epsilon d/D$. Now application of the operator Chernoff bound (Lemma 44) gives

$$\Pr\left\{\frac{1}{N}\sum_{s=1}^{N}\omega_{M_s} \notin [(1\pm\epsilon)\omega]\right\} \leq 2D\ \exp\left(-N\frac{\epsilon^2 t}{2\ \ln 2}\right). \tag{49}$$

5. The last step is to translate (49) into a statement about σ_{M_s}. Assume that for some set $\mathcal{S}\in\mathcal{X}$ the following condition holds:

$$\frac{1}{|\mathcal{S}|}\sum_{m\in\mathcal{S}}\omega_m \in [(1\pm\epsilon)\omega].$$

This implies that

$$\left\|\frac{1}{|\mathcal{S}|}\sum_{m\in\mathcal{S}}\omega_m - \omega\right\|_1 \leq \epsilon. \tag{50}$$

Together with (47) thus gives

$$\operatorname{Tr}\left(\frac{1}{|\mathcal{S}|}\sum_{m\in\mathcal{S}}\omega_m\right) \geq 1 - 4\epsilon. \tag{51}$$

Application of the gentle measurement lemma (Lemma 30) to (51) gives

$$\left\|\frac{1}{|\mathcal{S}|}\sum_{m\in\mathcal{S}}\omega'_m - \frac{1}{|\mathcal{S}|}\sum_{m\in\mathcal{S}}\omega_m\right\|_1 \leq \sqrt{32\epsilon}. \tag{52}$$

Application of the triangle inequality together with (46) gives

$$\begin{aligned}\left\|\frac{1}{|\mathcal{S}|}\sum_{m\in\mathcal{S}}\omega'_m - \frac{1}{|\mathcal{S}|}\sum_{m\in\mathcal{S}}\sigma_m\right\|_1 &\leq \frac{1}{|\mathcal{S}|}\sum_{m\in\mathcal{S}}\|\omega'_m - \sigma_m\|_1\\ &\leq \epsilon + \sqrt{16\epsilon},\end{aligned} \tag{53}$$

and analogously

$$\|\omega' - \sigma\|_1 \leq \epsilon + \sqrt{16\epsilon}. \tag{54}$$

Finally, combination of (48), (50), and (52)-(54) with the triangle inequality gives

$$oe(\mathcal{S}) = \left\| \frac{1}{|\mathcal{S}|} \sum_m \sigma_m - \sigma \right\|_1 \quad \leq \quad 3\epsilon + 19\sqrt{\epsilon}.$$

The statement of the lemma follows immediately from (49).

■

Lemma 44 *(Operator Chernoff Bound [AW02]). Let $\xi_1, \cdots, \xi_\nu$ be independent and identically distributed random variables with values in the algebra $B(\mathcal{H})$ of bounded linear operators on some Hilbert space $\mathcal{H}$, which are bounded between 0 and the identity operator I. Assume that the expectation value $\mathbb{E}\xi_s = \theta \geq tI$. Then for every $0 < \eta < 1/2$*

$$\Pr\left\{ \frac{1}{\nu} \sum_{s=1}^{\nu} \xi_s \neq [(1 \pm \eta)\theta] \right\} \leq 2 \ \dim \mathcal{H} \ \exp\left(-\nu \frac{\eta^2 t}{2 \ \ln 2} \right),$$

where $[(1 \pm \eta)\theta] \equiv [(1-\eta)\theta; (1+\eta)\theta]$ is an interval in the operator order: $[A; B] \equiv \{\xi \in B(\mathcal{H}) : A \leq \xi \leq B\}$.

Consider an ensemble $\{p_{x^n}, \sigma^E_{x^n}\}_{x^n \in \mathcal{X}^n}$ with average density operator $\sigma^E = \sum_{x^n} p_{x^n} \sigma^E_{x^n}$. We can define a *covering code* $\mathcal{C}$ as follows.

Corollary 45 (Covering Code) *There exists a covering code $\mathcal{C} = \{X_s\}_{s \in [\mathbf{S}]}$ of size $\mathbf{S} = 2^{n[I(X:E)_\sigma + 3c\delta]}$ so that for all $\epsilon, \delta > 0$ and sufficiently large n,*

$$\Pr\{oe(\mathcal{C}) \geq 3\epsilon + 19\sqrt{\epsilon}\} \leq 2|d_E|^n \ \exp\left(-\frac{\epsilon^3}{4 \ln 2} 2^{nc\delta} \right). \tag{55}$$

Proof. We can relate to Lemma 43 through the identifications $\mathcal{X} \to \mathcal{X}^n$, $\sigma_x \to \sigma_{x^n}$, $p \to p^n$, $\sigma \to \sigma^E$, $\Pi \to \Pi^n_{\delta(|\mathcal{X}|+1),\sigma}$, and $\Pi_x \to \hat{\Pi}^n_{\delta,\sigma_{x^n}}$ with

$$\hat{\Pi}^n_{E|X,\delta}(x^n) = \begin{cases} \Pi^n_{E|X,\delta}(x^n), & x^n \in \mathcal{T}^n_{X,\delta}, \\ 0, & \text{otherwise.} \end{cases}$$

Thus, we see that

$$\begin{aligned} D &= 2^{n[H(E)_\sigma + c\delta]} \\ d &= 2^{n[H(E|X)_\sigma - c\delta]}. \end{aligned}$$

These follow from the properties of typical subspaces and conditionally typical subspaces mentioned before. ■

Private Communication

A quantum channel $\mathcal{N}$ with the classical encoding map $\mathcal{E} : \mathcal{X} \to A$ can always be viewed as a classical-quantum channel $W : \mathcal{X} \to A$ so that

$$W(x) = \mathcal{N}(\mathcal{E}(x)) := \sigma^{BE}_{E(x)}.$$

Moreover, in the private setting, a classical-quantum channel $W : \mathcal{X} \to BE$ will generate two output quantum systems, where B is for the legitimate receiver while E goes to the eavesdropper.

We can thus define an (n, R, ϵ) *private code* as follows.

1. An encoding map $E : \{0,1\}^{nR} \to \mathcal{X}^n$ by Alice; Alice encodes the index m as $E(m)$ and sends it through the channel $W^{\otimes n}$, generating the state

$$\Upsilon^{MBE} = \frac{1}{2^{nR}} \sum_{m \in \{0,1\}^{nR}} |m\rangle\langle m|^M \otimes \sigma^{BE}_{E(m)}. \tag{56}$$

2. A decoding POVM $\{\Lambda_{m'}\}_{m' \in \{0,1\}^{nR}}$;

so that

$$\left\| \tilde{\Upsilon}^{BE} - \tau^B \otimes \sigma^E \right\|_1 \leq \epsilon, \tag{57}$$

where $\tilde{\Upsilon}^{BE}$ is the quantum system after Bob's decoding operation, and

$$\tau^B = \frac{1}{2^{nR}} \sum_m |m\rangle\langle m|^B$$

contains the private classical information that is decoupled from Eve's state σ^E.

We say the rate R is *achievable* if for any $\epsilon, \delta > 0$ and sufficiently large n there exists an $(n, R - \delta, \epsilon)$ private code. The private capacity $\mathcal{P}(\mathcal{N})$ is defined as

$$\mathcal{P}(\mathcal{N}) = \sup\{R : R \text{ is achievable}\}.$$

Let

$$I_p(\mathcal{N}) = \max_\rho I(X : B)_\sigma - I(X : E)_\sigma$$

where

$$\rho_{XA} = \sum_x p_x |x\rangle\langle x|_X \otimes \rho^A_x.$$

is the input to the channel $\mathcal{N}$ generating $\sigma_{XBE} = \sum_x p_x |x\rangle\langle x|_X \otimes \mathcal{N}_{A \to BE}(\rho^A_x)$.

Theorem 46 (Private Capacity [Dev05])

$$\mathcal{P}(\mathcal{N}) = \lim_{n \to \infty} \frac{1}{n} I_p(\mathcal{N}^{\otimes n}).$$

Direct Coding Theorem.

Fix $\epsilon, \delta > 0$ and a sufficiently large n. Consider the ensemble $\{p_{x^n}, \sigma^{BE}_{x^n}\}$ of the channel output $W^{\otimes n}$. There exists an encoding map $E : (M, S) \to \mathcal{X}^n$ for Alice, where $X^n \in \mathcal{X}^n$ is i.i.d. with distribution p, M represents the private classical message taken values from $\{0,1\}^{nR}$, and S represents the bits with value taken from $\{0,1\}^{nR_s}$ that needs to be sacrificed in order to blind eavesdropper's information about the private messages.

First, we invoke the HSW theorem (direct coding theorem of Theorem 39). There exists a code $\mathcal{C} = \{X_{E(m,s)}\}_{(m,s)\in 2^{nr}}$, where $r := R + R_s = I(X:B) - 2(c + c'\delta)\delta$ so that

$$\mathbb{E}[p_e(\mathcal{C})] \leq \epsilon.$$

For each $m \in \{0,1\}^{nR}$, define $\mathcal{C}_m = \{X_{E(m,s)}\}_{s\in[2^{nR_s}]}$, where $\mathcal{C}_m$ works as a covering code. Choose $R_s = I(X:E) + 3c\delta$. For any $m \in \{0,1\}^{nR}$, define the logic statement ℓ_m by $oe(\mathcal{C}_m) \leq 3\epsilon + 19\sqrt{\epsilon}$, where

$$oe(\mathcal{C}_m) = \left\| \frac{1}{2^{nR_s}} \sum_s \sigma^E_{X_{E(m,s)}} - \sigma^E \right\|_1 ,$$

where

$$\sigma^E = \sum_{x^n} p_{x^n} \sigma^E_{x^n}$$

and $\sigma^E_{x^n} = \mathrm{Tr}_B \, \sigma^{BE}_{x^n}$. By Corollary 45, $\forall m$,

$$\Pr\{\text{not } \ell_m\} \leq 2|d_E|^n \, \exp\left(-\frac{\epsilon^3}{4\ln 2} 2^{nc\delta}\right). \tag{58}$$

The probability of (58) can be made $\leq \epsilon 2^{-nR}$ for some R when n is sufficient large since the right-hand side is a double exponential in n.

Define the logic statement ℓ_0 by $\{p_e(\mathcal{C}) \leq \sqrt{\epsilon}\}$. By the Markov inequality, $\Pr\{\text{not } \ell_0\} \leq \sqrt{\epsilon}$. By the union bound,

$$\Pr\{\text{not } (\ell_0 \wedge \ell_1 \wedge \cdots \wedge \ell_{|m|})\} \leq \sum_{i=0}^{2^{nR}} \Pr\{\text{not } \ell_i\} \leq \epsilon + \sqrt{\epsilon},$$

where $\wedge$ means the logic operator "AND". Hence there exists a specific choice of $\{X_{E(m,s)}\}$, say $\{x_{E(m,s)}\}$, for which all these conditions are satisfied.

Denote by $\tilde{\Upsilon}^{BE}$ the state after Bob's POVM measurement and

$$\tilde{\Upsilon}^{BE}_0 = \frac{1}{2^{nR}} \sum_m |m\rangle\langle m|^B \otimes \frac{1}{2^{nR_s}} \sum_s \sigma^E_{X_{f(m,s)}} .$$

Consequently,

$$\begin{aligned} \|\tilde{\Upsilon}^{BE} - \tau^B \otimes \sigma^E\|_1 &\leq \|\tilde{\Upsilon}^{BE} - \tilde{\Upsilon}^{BE}_0\|_1 + \|\tilde{\Upsilon}^{BE}_0 - \tau^B \otimes \sigma^E\|_1 \\ &\leq 4\epsilon + 20\sqrt{\epsilon}, \end{aligned}$$

as claimed in (57).

■

Converse. We use the same trick. We consider the task of secret-key generation, where Alice prepares $\bar{\Phi}_{MM'}$ of size 2^{nR}. She then encodes M' before sending through the channel $\mathcal{N}$. Bob performs his POVM on the channel output

$$\sigma_{MBE} = 2^{-nR} \sum_m |m\rangle\langle m|_M \otimes \sigma^{B^n E^n}_m$$

yielding the state $\omega_{M\hat{M}E}$ so that

$$\|\omega_{M\hat{M}E} - \bar{\Phi}_{MM'} \otimes \sigma_{E^n}\|_1 \leq \epsilon.$$

The above condition guarantees

$$I(M : E^n)_\omega \leq n\epsilon'. \tag{59}$$

We have

$$\begin{aligned} nR &= I(M : M')_{\bar{\Phi}} \\ &\leq I(M : \hat{M})_\omega + n\epsilon' \\ &\leq I(M : B^n)_\sigma + n\epsilon' \\ &\leq I(M : B^n)_\sigma - I(M : E^n)_\sigma + 2n\epsilon' \\ &\leq I_p(\mathcal{N}^{\otimes n}) + n\epsilon, \end{aligned}$$

where the second line uses continuity of mutual information; the third line uses data processing inequality; the fourth lines follows from Eq. (59); the last line follows from the definition of I_p.

■

References

[AW02] R. Ahlswede and A. Winter. Strong converse for identification via quantum channels. *IEEE Transactions on Information Theory*, 48(3):569–579, Mar 2002. doi:10.1109/18.985947.

[BSST02] C.H. Bennett, P.W. Shor, J.A. Smolin, and A.V. Thapliyal. Entanglement-assisted capacity of a quantum channel and the reverse shannon theorem. *IEEE Transactions on Information Theory*, 48(10):2637–2655, Oct 2002. doi:10.1109/TIT.2002.802612.

[Dev05] I. Devetak. The private classical capacity and quantum capacity of a quantum channel. *IEEE Transactions on Information Theory*, 51(1):44–55, Jan 2005. doi:10.1109/TIT.2004.839515.

[DHW06] Igor Devetak, Patrick Hayden, and Andreas Winter. Principles of quantum information theory. (unpublished), 2006.

[HDW08] Min-Hsiu Hsieh, Igor Devetak, and Andreas Winter. Entanglement-Assisted Capacity of Quantum Multiple-Access Channels. *IEEE Trans. Inf. Theory*, 54:3078, 2008.

[HN03] M. Hayashi and H. Nagaoka. General formulas for capacity of classical-quantum channels. *IEEE Transactions on Information Theory*, 49(7):1753–1768, July 2003. doi:10.1109/TIT.2003.813556.

[Hol98] A. S. Holevo. The capacity of the quantum channel with general signal states. *IEEE Trans. Inf. Theory*, 44:269, 1998. doi:10.1109/18.651037.

[Sch95] Benjamin Schumacher. Quantum coding. *Phys. Rev. A*, 51:2738–2747, 1995.

[SW97] Benjamin Schumacher and Michael D. Westmoreland. Sending classical information via noisy quantum channels. *Phys. Rev. A*, 56:131–138, 1997. doi:10.1103/PhysRevA.56.131.