zuzana.gavorova@uab.cat

# 1. Classical communication

$A, B$ - Hilbert spaces

$\mathcal{L}(A) \supseteq S(A) = \{\rho \in \mathcal{L}(A) \mid \operatorname{tr}\rho = 1, \rho \geq 0\}$

**Def. 1.1.** A quantum channel $\eta : \mathcal{L}(A) \to \mathcal{L}(B)$ is a linear map that is completely positive (CP) and trace preserving (TP).

TP: $\forall \rho$, $\operatorname{tr}[\eta(\rho)] = \operatorname{tr}[\rho]$ $(\operatorname{tr}_B \circ \eta = \operatorname{tr}_A)$

Positive: if $\rho \geq 0$, then $\eta(\rho) \geq 0$.

CP: $\mathrm{id}_C \otimes \eta$ is positive for any auxiliary system $C$.

**Example 1.2.** Examples of channels:

$\langle Vu, Vu \rangle = \langle u, u \rangle$, $\forall |u\rangle \in A$
$V^\dagger V = \mathbb{1}_A$ (half of unitary)
Unitary iff $|A| = |B|$.

1. The identity channel: $\mathrm{id}_A(\rho) = \rho$.
2. Isometry channel: given an isometry $V : A \to B$, the map $\mathcal{V}(\rho) = V\rho V^\dagger$ is CPTP.
3. Partial trace: let $A = B \otimes E$, $\operatorname{tr}_E : \mathcal{L}(A) \to \mathcal{L}(B)$ is CPTP.
4. Composition and $\otimes$ of channels: if $\eta : A \to B$, $m : B \to C$ are CPTP, $m \circ \eta : A \to C$ is CPTP.
   If $\eta_i : A_i \to B_i$, for $i \in \{1, 2\}$ are CPTP, $\eta_1 \otimes \eta_2 : A_1 \otimes A_2 \to B_1 \otimes B_2$ is CPTP.
5. Convex combinations of channels: if $\eta_i : A \to B$ are CPTP for $i \in I$, and $(p_i)_{i \in I}$ is a prob. vector $(p_i \geq 0, \sum_i p_i = 1)$, then $\sum_i p_i \eta_i$ is a CPTP map.

**Lemma 1.3.** For a linear map $\eta : A \to B$, the following are equivalent:

1. $\eta$ is CPTP.
2. $\eta$ is TP and $\mathrm{id}_{A'} \otimes \eta$ is positive for $A' \underset{\text{isomorphic}}{\cong} A$.
3. The Choi-Jamiołkowski matrix of $\eta$, $J(\eta)^{A'B}$, is positive semidef. and satisfies $\operatorname{tr}_B J(\eta) = \frac{1}{|A|}\mathbb{1}_{A'}$.

$J(\eta)^{A'B} := (\mathrm{id}_{A'} \otimes \eta)(\phi^{A'A})$, $\phi^{A'A} = |\phi\rangle\langle\phi|$, $|\phi\rangle = \frac{1}{\sqrt{|A|}}\sum_i |i\rangle \otimes |i\rangle$

4. $\eta$ has a Stinespring dilation: there exists an isometry $V: A \to B \otimes E$ with a suitable Hilbert space $E$, such that $\eta(\rho) = \mathrm{tr}_E[V\rho V^\dagger]$.

5. $\eta$ has a Kraus representation: there exist operators $K_\alpha : A \to B$ such that $\eta(\rho) = \sum_\alpha K_\alpha \rho K_\alpha^\dagger$ and $\sum_\alpha K_\alpha^\dagger K_\alpha = \mathbb{1}$.

__Proof.__ (1) $\to$ (2) $\to$ (3) $\to$ (4) $\to$ (5) $\to$ (1)

1. $\to$ 2. : just take $C = A'$.

2. $\to$ 3. : $\mathrm{id}_{A'} \otimes \eta$ is positive, $\phi^{A'A} \in S(A'A) \to \underbrace{(\mathrm{id}_{A'} \otimes \eta)\,\phi^{A'A}}_{J(\eta)^{A'A}} \geq 0$

$$\mathrm{tr}_B[J(\eta)] = \mathrm{tr}_B[(\mathrm{id}_A \otimes \eta)\,\phi^{A'A}] = \mathrm{tr}_A\, \phi^{A'A} = \frac{1}{|A|}\mathbb{1}_{A'}.$$

3. $\to$ 4. : $J(V) = (\mathbb{1}_{A'} \otimes V)\,|\phi\rangle\langle\phi|\,(\mathbb{1}_{A'} \otimes V^\dagger)$

$$\begin{tikzcd}
\eta \arrow[r] & J(\eta) \\
V \arrow[u, "\mathrm{tr}_E"] \arrow[r, "\text{Choi-J.}"] & J(V) \arrow[u, "\text{purification}"'] \arrow[ul, bend left]
\end{tikzcd}$$

Claim: $\eta(\rho) = |A|\,\mathrm{tr}_{A'}[(\rho_{A'}^T \otimes \mathbb{1}_B)\,J(\eta)^{A'B}]$.

To show the claim, we will use that $(\mathbb{1}_{A'} \otimes \rho)|\phi\rangle^{A'A} \overset{*}{=} (\rho^T \otimes \mathbb{1}_B)|\phi\rangle^{A'A}$ (exercise). Then, we get

$\rho = |A|\,\mathrm{tr}_{A'}[(\mathbb{1}_{A'} \otimes \rho)|\phi\rangle\langle\phi|] \overset{*}{=} |A|\,\mathrm{tr}_{A'}[(\rho^T \otimes \mathbb{1}_A)\,\phi]$. Thus:

$$\eta(\rho) = |A|\,\mathrm{tr}_{A'}[(\mathbb{1}_{A'} \otimes \eta)(\rho^T \otimes \mathbb{1}_A)\,\phi] = |A|\,\mathrm{tr}_{A'}[(\rho^T \otimes \mathbb{1}_A)\underbrace{(\mathbb{1}_{A'} \otimes \eta)\,\phi}_{J(\eta)}] \quad \checkmark$$

All purifications of the same state are unique, up to isometries on the purifying Hilbert space

Take some $\psi^{A'BE}$ purification of $J(\eta)$ for some appropriate $E$
$$\begin{cases} \mathrm{tr}_E\, \psi^{A'BE} = J(\eta) \\ \mathrm{tr}_{EB}\, \psi^{A'BE} = \mathrm{tr}_B\, J(\eta) = \frac{1}{|A|}\mathbb{1}_{A'} = \mathrm{tr}_A\, \phi^{A'A} \end{cases}$$

All purifications of the same state are unique up to the isometry on the purifying Hilbert space.

$\exists$ isometry $V: A \to B \otimes E$ s.t. $|\psi\rangle^{A'BE} = (\mathbb{1}_A \otimes V)|\phi\rangle^{A'A}$. The channel $\mathcal{V}(\rho) = V\rho V^\dagger$, where $\mathcal{V}: A \to B \otimes E$ is an isometry channel.

Observe $\psi = (\mathbb{1}_{A'} \otimes \mathcal{V})\, \phi^{A'A} = J(\mathcal{V}) \to J(\eta) = \mathrm{tr}_E\, J(\mathcal{V})$.

$$\eta(\rho) = |A|\, \mathrm{tr}_{A'}[(\rho^T \otimes \mathbb{1}_B)\, \mathrm{tr}_E\, J(\mathcal{V})^{A'BE}] \to \eta(\rho) = \mathrm{tr}_E [\underbrace{|A|\, \mathrm{tr}_{A'}[(\rho^T \otimes \mathbb{1}_{BE})\, J(\mathcal{V})^{A'BE}]}_{\mathcal{V}(\rho)}]$$

Then, $\eta(\rho) = \mathrm{tr}_E [V\rho V^\dagger]$.

4. $\to$ 5.: Choose any basis for $E$: $\{|\alpha\rangle\}$.

$$\eta(\rho) = \mathrm{tr}_E[V\rho V^\dagger] = \sum_\alpha \underbrace{(\mathbb{1}_B \otimes \langle\alpha|)V}_{K_\alpha}\rho \underbrace{V^\dagger(\mathbb{1}_B \otimes |\alpha\rangle)}_{K_\alpha^\dagger} = \sum_\alpha K_\alpha \rho K_\alpha^\dagger$$

Also, $\sum_\alpha K_\alpha^\dagger K_\alpha = \sum_\alpha V^\dagger(\mathbb{1}_B \otimes |\alpha\rangle)(\mathbb{1}_B \otimes \langle\alpha|)V = V^\dagger V = \mathbb{1}_A$.

5. $\to$ 1.: Any map of the form $\rho \mapsto K\rho K^\dagger$ is positive. Then, $\langle u|K\rho K^\dagger|u\rangle \overset{?}{\geq} 0$. As $|v\rangle = K^\dagger|u\rangle$, $\langle v|\rho|v\rangle \geq 0$, $\forall |u\rangle \in A$.

$\rho \mapsto (\mathbb{1}_C \otimes K_\alpha)\rho(\mathbb{1}_C \otimes K_\alpha)^\dagger$ for any $C$ are of this form $\to$ positive

$\rho \mapsto K_\alpha \rho K_\alpha^\dagger$ is CP, $\eta$ is CP (as it is a sum of those)

$\mathrm{tr}[\eta(\rho)] = \sum_\alpha \mathrm{tr}[K_\alpha \rho K_\alpha^\dagger] = \mathrm{tr}\rho$ ✓

Example 1.4. More examples on CPTP channels (with special physical meaning)

1. Identity channel: $\mathrm{id}_A: A \to A$
2. Constant channel: $P_\sigma: A \to B$, $P_\sigma(\rho) = \sigma\, \mathrm{tr}\rho$ loss of purity
3. Qubit depolarizing channel: $D_p(\rho) = (1-p)\rho + \frac{p}{3}(X\rho X^\dagger + Y\rho Y^\dagger + Z\rho Z^\dagger) = (1-q)\rho + q\, \mathrm{tr}\rho\, \frac{\mathbb{1}}{2}$
4. Qubit dephasing channel: $\mathcal{L}_p(\rho) = (1-p)\rho + pZ\rho Z^\dagger$
5. Qubit Pauli error channel: $\eta_{\vec{p}}(\rho) = p_0\rho + p_1 X\rho X^\dagger + p_2 Y\rho Y^\dagger + p_3 Z\rho Z^\dagger$ generalizes 3 and 4

6. Qubit amplitude damping channel : $a_\delta(\rho) = A_0 \rho A_0^\dagger + A_1 \rho A_1^\dagger$ ; $A_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\delta} \end{pmatrix}$, $A_1 = \begin{pmatrix} 0 & \sqrt{\delta} \\ 0 & 0 \end{pmatrix}$

7. Erasure channel : $\mathcal{E}_q : A \to B$ where $B = A \oplus |\perp\rangle\langle\perp|$, $\mathcal{E}_q(\rho) = (1-q)\rho + q|\perp\rangle\langle\perp|$.

- State preparations $X \to B$
- Measurements, i.e. POVMs $(M_y, y \in Y)$ $A \to Y$

Definition 1.5. A classical to quantum channel (cq-channel) is a CPTP map $\eta : X \to B$ such that $\eta(\xi) = \sum_x \langle x|\xi|x\rangle \rho_x$, for an orthonormal basis $\{|x\rangle\}$ of $X$ and states $\rho_x \in S(B)$.

- Prepare $\rho_x$ if the input is $|x\rangle\langle x|$.
- Prepares a mixture of $\rho_x$'s in general.

A quantum classical channel (qc-channel) is a CPTP map $m : A \to Y$ such that $m(\alpha) = \sum_y |y\rangle\langle y| \operatorname{tr}(\alpha M_y)$ for an orthonormal basis $\{|y\rangle\}$ and a POVM $(M_y : y \in Y)$ on $A$.

- Implements the POVM and records the classical outcome.

---

Channels map states to states

We can use Born's rule : $p(i|\rho) = \operatorname{tr}[\rho M_i]$ for POVM $(M_i : i \in I)$ to equivalently use maps taking observables to observables : adjoint CP maps

adjoint CP map

Definition 1.6. For a linear map $T : \mathcal{L}(A) \to \mathcal{L}(B)$, define $T^* : \mathcal{L}(B) \to \mathcal{L}(A)$ as the unique linear map such that $\forall \xi \in \mathcal{L}(A), Y \in \mathcal{L}(B)$, $\operatorname{tr}[T(\xi)^\dagger Y] = \operatorname{tr}[\xi^\dagger T^*(Y)]$.

- Note we have $T^{**} = T$.
- A CPTP map is Hermitian preserving $\to \eta(\xi)^\dagger = \eta(\xi^\dagger)$ we can omit the $\dagger$ in
- Moreover, for $\eta$ CPTP, $\eta^*$ is CP and unit-preserving $\leftrightarrow \eta^*(\mathbb{1}_B) = \mathbb{1}_A$. → Prove this with Kraus

↳ transforms POVMs into POVMs (sum to the identity!)

**Example 1.7.** (adjoint maps of some specific channels)

1. $\mathrm{Id}_A$ $\quad \mathrm{Id}_A$
2. $\mathcal{K}(\rho) = K\rho K^\dagger$ $\quad \mathcal{K}^*(Y) = K^\dagger Y K$
3. $\eta(\rho) = \sum_\alpha K_\alpha \rho K_\alpha^\dagger$ $\quad \eta^*(Y) = \sum_\alpha K_\alpha^\dagger Y K_\alpha$
4. $\mathrm{tr}_E : \mathcal{L}(BE) \to \mathcal{L}(B)$

   $\mathrm{tr}_E^*(Y) = Y \otimes \mathbb{I}_E$
5. Composed channel $m \circ \eta$ $\quad (m \circ \eta)^* = \eta^* \circ m^*$
6. Tensor product $\eta_1 \otimes \eta_2$ $\quad (\eta_1 \otimes \eta_2)^* = \eta_1^* \otimes \eta_2^*$
7. Stinespring form $\eta(\rho) = \mathrm{tr}_E[V\rho V^\dagger]$ $\quad \eta^*(Y) = V^* \circ \mathrm{tr}_E^* = V^\dagger (Y \otimes \mathbb{I}_E) V$
8. Constant channel $P_\sigma = \sigma \,\mathrm{tr}$ $\quad P_\sigma^*(Y) = (\mathrm{tr}[\sigma Y])\, \mathbb{I}_A$

## 1.2. One-shot communication, plain & entanglement assisted

$$[K] \ni m \to \boxed{E} \xrightarrow{A} \boxed{\eta} \xrightarrow{B} \boxed{(\Delta_{m'})} \to m' \in [K]$$

$[K] = \{1, 2, \dots, K\}$

$E(m) = \rho_m$

$$\underbrace{\text{encoder} \qquad\qquad \text{decoder}}_{\text{code: what we are optimizing}}$$

**Definition 1.8.** A code for $\eta : A \to B$ is a pair $(E, D)$ consisting of a map $E : [K] \to S(A)$ (encoder) from the set of messages $[K] = \{1, 2, \dots, K\}$ into the states $S(A)$, and a measurement, i.e. a POVM $D = (\Delta_{m'} : m' \in [K])$ on $B$ (the decoder).

A code has two performance criteria: the (one-shot) rate of the code $R = \log K$, and the average (over messages) error probability, $P_e = 1 - \frac{1}{K} \sum_{m \in [K]} \mathrm{tr}[\eta(E(m))\Delta_m]$

The largest $R$ such that the code with $P_e \leq \varepsilon$ exists is called $\varepsilon$-one-shot capacity of the channel $\eta$,

$$C_\varepsilon(\eta) = \sup_{\text{codes with } P_e \leq \varepsilon} R.$$

**Remark 1.9.** 1. Capacity increases (or stays the same) under the increase of error tolerance for $0 \leq \varepsilon \leq \varepsilon' \leq 1$:

$C_{\varepsilon'}(\eta) \geq C_\varepsilon(\eta)$

2. Capacity decreases (or stays the same) under pre- and post-processing:

$C_\varepsilon(\mathcal{N}) \geq C_\varepsilon(\mathcal{N}\circ a) \geq C_\varepsilon(B\circ\mathcal{N}\circ a)$

Proof. 1. Any code for $\mathcal{N}$ with rate $R$ and error $\leq \varepsilon$ is also a code with rate $R$ and error $\leq \varepsilon'$.

$R^* \to C_\varepsilon(\mathcal{N})$. $C_{\varepsilon'}(\mathcal{N}) = \sup_{\text{codes } P_e \leq \varepsilon} R \geq R^*$   ↳ best code in $\varepsilon$, for sure $\varepsilon'$ is, at least, that.   $S_\varepsilon \subseteq \tilde{S}_{\varepsilon'}$

2. Any code for $B\circ\mathcal{N}\circ a$ with rate $R$ and error $\varepsilon$ can be used to build a code for $\mathcal{N}$ with rate $R$ and error $\leq \varepsilon$.

$$[K] \ni m \to \underbrace{\boxed{E}\ \boxed{a}}_{\tilde{E} = a\circ E}\ \boxed{\mathcal{N}}\ \underbrace{\boxed{B}\ \boxed{(D_{m'})}}_{\tilde{D}_{m'} = B^*(D_{m'})} \to m' \in [K]$$

suggested new code

you can eliminate (hermitian-preserving)

$tr[T(\xi)^\dagger Y] = tr[\xi^\dagger T^*(Y)]$

new code

$$\tilde{Pr}[M'=m' \mid M=m] = tr[\mathcal{N}(\tilde{E}(m))\tilde{D}_{m'}] = tr[\mathcal{N}(a(\overset{\rho_m}{\overset{\|}{E(m)}}))B^*(D_{m'})] = tr[B\circ\mathcal{N}\circ a(\rho_m)D_{m'}] =$$
$$= Pr[M'=m' \mid M=m]$$

If we want to prove (2), $\sup_{x\in S} R(x) \leq \sup_{y\in\tilde{S}} \tilde{R}(y)$, it is enough to show $\sup_{x\in S} R(x) = \sup_{f(x)\in f(S)\cap\tilde{S}} \tilde{R}(f(x))$.

Then, enough to show
- $R(x) = \tilde{R}(f(x)), \forall x$
- $f(S) \subseteq \tilde{S}$ the error in the new guy is small enough

Example 1.10. (capacities for simple channels)

1. $\log|A| \leq C_\varepsilon(id_A) \leq \log|A| - \log(1-\varepsilon)$

Lower bound: will show $\log|A| \leq C_0(id_A) \leq C_\varepsilon(id_A)$   ↳ show this   ↳ Remark 1.9.

$E_m = |m\rangle\langle m|$ ← $K = |A|$

$D_{m'} = |m'\rangle\langle m'|$, $\sum D_{m'} = \mathbb{1}$

$Pr[M'=m' \mid M=m] = tr[E(m)D_{m'}] = \delta_{mm'} \to P_e = 1 - \frac{1}{K}\sum_{m\in[K]} Pr[M'=m \mid M=m] = 0 \to R = \log|A|$

$C_0(id_A) \geq \log|A|$

Upper bound: assume $K$ and see how the bound $P_e \leq \varepsilon$ bounds $K$: $P_{succ} \geq 1-\varepsilon$:

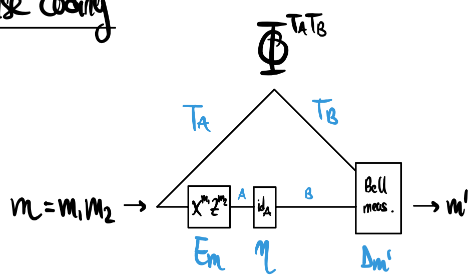
$$1-\varepsilon \leq P_{succ} = \frac{1}{K}\sum_{m=1}^{K} tr[id_A(\rho_m)D_m] \leq \frac{1}{K}\sum_{m=1}^{K}\left(tr[\rho_m D_m] + tr[(\mathbb{1}-\rho_m)D_m]\right) = \frac{1}{K}\sum_m tr[D_m] = \frac{|A|}{K}$$

Then, $1-\varepsilon \le \frac{|A|}{K} \longrightarrow K \le \frac{|A|}{1-\varepsilon} \longrightarrow R \le \log \frac{|A|}{1-\varepsilon}$ for any code $P_e \le \varepsilon$.

2. $C_\varepsilon(P_\sigma) \le \log \left\lfloor \frac{1}{1-\varepsilon} \right\rfloor$, $P_\sigma(\rho_m) = (\mathrm{tr}\,\rho_m)\sigma$

$$1-\varepsilon \le P_{succ} = \frac{1}{K}\sum_m \mathrm{tr}(\sigma D_m) = \frac{1}{K}\mathrm{tr}\,\sigma = \frac{1}{K} \longrightarrow K \le \frac{1}{1-\varepsilon}$$

## Dense coding




2 bits can be communicated using 1 qbit and 1 ebit

**Def. 1.11.** An entanglement assisted code for $\eta: A \to B$ is a triplet $(\omega, E, D)$:

- a state $\omega$ on $T_A \otimes T_B$
- a map $E: [K] \to CPTP(T_A \to A)$ from the set of messages $[K] = \{1, 2, \dots, K\}$ to the set of channels from $T_A$ to $A$

$E(m) = \mathcal{E}_m$

↑ modulation operators (depends on $m$)

- a measurement (i.e. a POVM) $D = (D_{m'} : m' \in [K])$ on $B \otimes T_B$.

Same two performance criteria:

- The one-shot rate of the code $R = \log K$
- The average (over messages) error probability $P_e = 1 - \frac{1}{K}\sum_{m\in[K]} \mathrm{tr}\left[\{(\eta \circ \mathcal{E}_m) \otimes 1_B\}(\omega^{T_AT_B})\, D_m\right]$

The largest rate $R$ s.t. an entanglement assisted code with $P_e \le \varepsilon$ exists is called the $\varepsilon$-one-shot entanglement assisted capacity, $C_\varepsilon^{(ea)}(\eta) = \sup_{\substack{\text{e.a. codes} \\ \text{with } P_e \le \varepsilon}} R$.

**Remark 1.12.** $C_0^{(e.a.)}(id_2) \ge 2$, because of dense coding

**Remark 1.13.** Same as plain capacity, the e.a. capacity is non-increasing under the decrease of error tolerance, and under pre- and post-processing.

*Proof.* Same as for plain capacity.

**Remark 1.12. (continued).** $C_\varepsilon^{(ea)}(\mathrm{id}_2) \geq C_0^{(ea)}(\mathrm{id}_2) \geq 2$

**Remark 1.14.** E.A. code is a generalisation of plain code: $C_\varepsilon^{(ea)}(\eta) \geq C_\varepsilon(\eta)$.

*Proof.* Any plain code $(E, D)$ with rate $R$ and $P_e \leq \varepsilon$ can be used to define an entanglement assisted code $(\omega, \mathcal{E}_m, \tilde{D}_m)$ as:

- $\omega^{T_A T_B}$ anything
- for the modulation $\mathcal{E}_m = P_{\rho_m}(\rho_m = E(m))$, the constant channel.
  - this gives $(\mathcal{E}_m \otimes \mathrm{id}_{T_B})(\omega^{T_A T_B}) = \rho_m \otimes \omega^{T_B}$ after encoding
  - it gives $\eta(\rho_m)^B \otimes \omega^{T_B}$ on $B \otimes T_B$
- for the measurement $\tilde{D}_{m'} = D_{m'} \otimes \mathbb{1}_B$

The rate of the constructed e.a. code is also $R = \log K$, the error?

$$P_r^{(ea)}[M' = m' \mid M' = m] = \mathrm{tr}[\{(\eta \circ \mathcal{E}_m) \otimes \mathrm{id}_{T_B}\}(\omega)\tilde{D}_{m'}] = \mathrm{tr}[(\eta \otimes \mathrm{id}_{T_B})(\rho_m^A \otimes \omega^{T_B})\tilde{D}_{m'}]$$

$$= \mathrm{tr}[(\eta(\rho_m) \otimes \omega^{T_B})(D_{m'} \otimes \mathbb{1}_{T_B})] = \mathrm{tr}[\eta(\rho_m) D_{m'}] = P_r^{(plain)}[M' = m' \mid M = m]$$

$$P_e^{(ea)} = P_e^{(plain)} \leq \varepsilon \rightarrow C_\varepsilon^{(ea)}(\eta) \geq \underbrace{R^{*(plain)}}_{\text{best rate for plain codes}} = C_\varepsilon(\eta) \quad \square$$

## 1.3. Coding theorems via hypothesis testing

**Def. 1.15.** For two states, $\rho$ and $\sigma$, on the same system and $0 \leq \varepsilon \leq 1$, the <u>hypothesis testing relative entropy</u> is $D_H^\varepsilon(\rho\|\sigma) = -\log \min_{\substack{0 \leq M \leq \mathbb{1} \\ \mathrm{tr}[\rho M] \geq 1-\varepsilon}} \mathrm{tr}[\rho M]$.

2 types of errors:

- The error of announcing $\rho$ when the state is $\sigma \longrightarrow$ is being minimised (objective function)
- The error of announcing $\sigma$ when the state is $\rho \longrightarrow$ is upper bounded by $\varepsilon$

**Lemma 1.16.** $D_h^\varepsilon$ has the following properties:

1. For any states $\rho, \sigma$: $D_h^\varepsilon(\rho\|\sigma) \geq D_h^\varepsilon(\rho\|\rho) = -\log(1-\varepsilon) \geq 0$.
2. For any states $\rho, \sigma$ on $A$ and any channel $\eta: A \to B$: $D_h^\varepsilon(\rho\|\sigma) \geq D_h^\varepsilon(\eta(\rho)\|\eta(\sigma))$.
3. For the maximally mixed state, $\tau_A = \frac{1}{|A|}\mathbb{1}_A$ and pure $\rho$, $D_h^\varepsilon(\rho\|\tau) = \log|A| - \log(1-\varepsilon)$.

Proof. First we prove (2): take $M$ that optimises $D_h^\varepsilon(\eta(\rho)\|\eta(\sigma))$. It satisfies $1-\varepsilon \leq tr[\eta(\rho)M] =$

$$= tr[\rho\underbrace{\eta^*(M)}_{\substack{\tilde{M} \text{ satisfies the error tolerance} \\ \text{for } D_h^\varepsilon(\rho\|\sigma) \quad \text{"feasible"}}}].$$

$$D_h^\varepsilon(\rho\|\sigma) = \max_{M' \text{ feasible}} -\log tr[\sigma M'] \geq -\log tr[\sigma\tilde{M}] = -\log tr[\sigma\eta^*(M)] = -\log tr[\eta(\sigma)M] = D_h^\varepsilon(\eta(\rho)\|\eta(\sigma))$$

New proof of 1. Left neq. set $\eta = P_\sigma$.

$$D_h^\varepsilon(\rho\|\rho) = -\log \min_{\substack{X \geq 1-\varepsilon \\ \text{another condition} \\ \text{on } X}} X \leq -\log \min_{X \geq 1-\varepsilon} X = -\log(1-\varepsilon)$$

It remains to show lower bound for equality. Suggest a specific distinguishing measurement: $M_1 = (1-\varepsilon)\mathbb{1}$.

$tr[\rho M_1] = 1-\varepsilon \to M_1$ is feasible

$$D_h^\varepsilon(\rho\|\rho) \geq -\log tr[\rho M_1] = -\log(1-\varepsilon) \quad \square$$

$M$ feasible $\to 1-\varepsilon \leq tr[\rho M] \to 1-\varepsilon \leq tr[M]$

3. $$D_h^\varepsilon(\rho\|\tau) = -\log \min_{M \text{ feasible}} tr[\tau M] = \log|A| - \log \min_{M \text{ feas.}} tr[M] \leq \log|A| - \log(1-\varepsilon)$$

Remaining $\geq$: suggested specific $M_2 = (1-\varepsilon)\rho$, $tr[\rho M_2] \geq 1-\varepsilon$ ↑ satisfied since $\rho$ pure

$$D_h^\varepsilon(\rho\|\tau) \geq \log|A| - \log tr[(1-\varepsilon)\rho] = \log|A| - \log(1-\varepsilon) \quad \square$$

Recall: $C(\eta^{class.}) := \sup_{\substack{\text{families of codes} \\ \text{s.t. } \lim_{n\to\infty} \frac{1}{n} P_e^{(n)} = 0}} R = \max_{p(x)} I(X;Y)$ Shannon channel coding

$$= \max_{p(x)} D(p(x,y)\|p(x)p(y))$$

$$C(\mathcal{N}_{A\to B}) := \sup_{\substack{\text{families of codes} \\ \text{s.t. } \lim_{n\to\infty} \frac{1}{n} P_e^{(n)} = 0}} R$$

**Holevo - Schumacher - Westmoreland theorem**

$$C(\mathcal{N}) = \lim_{\ell\to\infty} \frac{1}{\ell} \sup_{\substack{p(x^\ell) \\ \rho_{x^\ell}}} I(X^{\otimes \ell} ; B^{\otimes \ell})$$

$$= \lim_{\ell\to\infty} \frac{1}{\ell} \sup \Delta(\omega^{X^{\otimes\ell} B^{\otimes\ell}} \| \omega^{X^{\otimes\ell}} \omega^{B^{\otimes\ell}})$$

where $\omega = \sum p(x^\ell) |x^\ell \rangle\langle x^\ell| \otimes \mathcal{N}^{\otimes \ell}(\rho_{x^\ell})$

**Theorem 1.17.** For any channel $\mathcal{N} : A \to B$ and any $0 \le \varepsilon \le 1$, we have for any $\varepsilon' < \varepsilon$:

$$2^{C_\varepsilon(\mathcal{N})} \geq \sup_{p(x), \rho_x} \left\lfloor 2^{\Delta_h^{\varepsilon'}(\omega^{XB} \| \omega^X \otimes \omega^B)) - \log \frac{4\varepsilon}{(\varepsilon-\varepsilon')^2}} \right\rfloor$$

where $\rho_x \in S(A)$ and $p(x)$ is a prob. density function $\omega^{XB} = \sum_x p(x) |x\rangle\langle x|^X \otimes \mathcal{N}(\rho_x)$.

**Theorem 1.18.** $C_\varepsilon(\mathcal{N}) \le \sup_{p(x), \rho_x} \Delta_h^\varepsilon(\omega^{XB} \| \omega^X \otimes \omega^B)$.