$$Q \longrightarrow \boxed{\mathcal{E}} \xrightarrow{A} \boxed{\mathcal{W}} \xrightarrow{B} \boxed{\mathcal{D}} \longrightarrow \hat{Q}$$

noisy quantum channel.

Now want to transmit quantum states

Rk: $\mathcal{W}$ is not necessarily communication link, it can be the memory of a quantum computer.

Def: An (M, ε) quantum code $(\mathcal{E}, \mathcal{D})$ for a quantum channel $\mathcal{W} : L(A) \to L(B)$ is given by:

- $\mathcal{E} : L(Q) \to L(A)$
- $\mathcal{D} : L(B) \to L(\hat{Q})$

$\Big\}$ quantum channels

with $\dim(Q) = \dim(\hat{Q}) = M$

and $1 - \varepsilon \leq \int F\left(|\psi\rangle\langle\psi|, \mathcal{D} \circ \mathcal{W} \circ \mathcal{E}(|\psi\rangle\langle\psi|)\right)^2 d\psi$ — normalized unitarily invariant measure unit vectors in Q

Rk: * For classical information, replace measure $d\psi$ with uniform distribution over a fixed basis. In that case, we may assume $F \in \{0, 1\}$

* Other measures are also possible: $F \geq 1 - \varepsilon \quad \forall |\psi\rangle$
All can be related.

* It is not a priori clear that "interesting" encodings exist. In fact, because of "no-cloning" unclear how to add redundancy to improve error.

Def: The quantum capacity of $\mathcal{W}$ is defined as

$$Q(\mathcal{W}) = \lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{\max\{\log M : \exists\ (M,\varepsilon) \text{ code for } \mathcal{W}^{\otimes n}\}}{n}$$

Ex:

- $\mathcal{W}(S) = \mathrm{Tr}(S)\,\sigma$ ← fixed state.

If $(\mathcal{E}, \mathcal{D})$ is an (M,ε) code then for

$$\int \langle\psi|\mathcal{D}\circ \underbrace{\mathcal{W}\circ\mathcal{E}(|\psi\rangle\langle\psi|)}_{\sigma}|\psi\rangle d\psi \geq (1-\varepsilon)$$

$$\int \langle\psi|\mathcal{D}(\sigma)|\psi\rangle d\psi \geq (1-\varepsilon)$$

$$\mathrm{Tr}\left(\int |\psi\rangle\langle\psi| d\psi \cdot \mathcal{D}(\sigma)\right) \geq (1-\varepsilon)$$

By unitary invariance, $\int |\psi\rangle\langle\psi| d\psi = \frac{I_Q}{M}$

So $M \leq \frac{1}{(1-\varepsilon)}$

and $Q(\mathcal{W}) = 0$ as expected.

The definition of average fidelity might seem complicated to compute but it is actually simple.

Will use instead the channel fidelity F_c

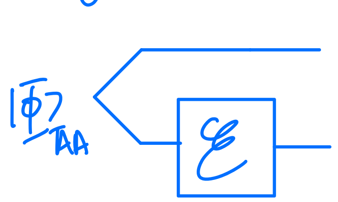
Prop: For a quantum channel $\mathcal{E}: L(A) \to L(A)$ we have

$$\int \langle\psi|\mathcal{E}(|\psi\rangle\langle\psi|)|\psi\rangle \, d\psi = \frac{d \cdot F_c(\mathcal{E}) + 1}{d+1} \qquad d = \dim A.$$

where $F_c(\mathcal{E}) := \langle\Phi|_{\bar{A}A} (I_{\bar{A}} \otimes \mathcal{E})(|\Phi\rangle\langle\Phi|_{\bar{A}A})|\Phi\rangle_{\bar{A}A} = \frac{1}{d^2}\sum_{a a'} \langle a|\mathcal{E}(|a\rangle\langle a'|)|a'\rangle$

where $|\Phi\rangle_{\bar{A}A} = \frac{1}{\sqrt{d}} \sum_a |a\,a\rangle_{\bar{A}A}$.

Average $\simeq$ evaluating fidelity for a single state $|\Phi\rangle_{\bar{A}A}$



$\leftarrow$ reference system we added.

$\mathrm{Tr}(F_{\bar{A}A}(S_{\bar{A}} \otimes T_A)) = \mathrm{Tr}(ST)$. $F|\psi\rangle\otimes|\varphi\rangle = |\varphi\rangle\otimes|\psi\rangle$

Flip operator on $\bar{A}\otimes A$: $F = \sum_{a,a'} |a\,a'\rangle\langle a'\,a|$

Proof:

$$\int \mathrm{Tr}(|\psi\rangle\langle\psi| \cdot \mathcal{E}(|\psi\rangle\langle\psi|))\, d\psi = \int \mathrm{Tr}\left(F_{\bar{A}A}\, |\psi\rangle\langle\psi|_{\bar{A}} \otimes \mathcal{E}(|\psi\rangle\langle\psi|)_A\right) d\psi$$

Haar measure

$\mathbb{E} f(U) = \mathbb{E} f(UV)$ for any V

$$= \int \mathrm{Tr}\left((I_{\bar{A}} \otimes \mathcal{E}^*)(F)\; |\psi\rangle\langle\psi| \otimes |\psi\rangle\langle\psi|\right) d\psi.$$

$$= \mathrm{Tr}\left((I_{\bar{A}} \otimes \mathcal{E}^*)(F_{\bar{A}A}) \int U|0\rangle\langle 0|U^* \otimes U|0\rangle\langle 0|U^*\, dU\right)$$

arbitrary fixed state

U chosen from Haar measure on unitary group

Compute the integral

$$\int (U\otimes U)(|0\rangle\langle 0| \otimes |0\rangle\langle 0|)\, U^* \otimes U^*\, dU = \alpha I_{\bar{A}A} + \beta F_{\bar{A}A}$$

(from representation theory) Schur–Weyl duality.

$\uparrow$ commutes with any $V\otimes V$

α, β can be found by computing $\mathrm{Tr}(\cdot)$ and $\mathrm{Tr}(F\cdot)$

We get $\int (U\otimes U)(|0\rangle\langle 0|\otimes|0\rangle\langle 0|) U^*\otimes U^* dU = \frac{1}{d(d+1)}(I_{\bar{A}A} + F_{\bar{A}A})$

Back to our calculation

$$= \mathrm{Tr}\left((I_{\bar{A}}\otimes\mathcal{E}^*)(F_{\bar{A}A}) \cdot \frac{1}{d(d+1)} (I_{\bar{A}A} + F_{A\bar{A}}) \right)$$

$$= \frac{1}{d(d+1)} \underbrace{\mathrm{Tr}(F_{\bar{A}A}\, I_{\bar{A}} \otimes \mathcal{E}(I_A))}_{\mathrm{Tr}(\mathcal{E}(I_A)) = d}$$

$$+ \frac{1}{d(d+1)} \underbrace{\sum_{a,a'} \mathrm{Tr}\big(|a\rangle\langle a'|\otimes|a'\rangle\langle a| \,(I_A\otimes\mathcal{E})(F_{\bar{A}A})\big)}_{\substack{= \sum_{a a'} \mathrm{Tr}(|a\rangle\langle a'|\otimes|a'\rangle\langle a| \; |a'\rangle\langle a|\otimes\mathcal{E}(|a\rangle\langle a'|)) \\ = d^2 \langle\Phi|\mathcal{E}(|\Phi\rangle\langle\Phi|)|\Phi\rangle}}$$

$$= \frac{1}{d+1} + \frac{d}{d+1} F_c(\mathcal{E}) \quad \blacksquare$$

Rk: If $\mathcal{E}(S) = \sum_x K_x S K_x^*$ then

$$F_c(\mathcal{E}) = \frac{1}{d^2} \sum_{a,a'} \sum_x \langle a|K_x|a\rangle\langle a'|K_x^*|a'\rangle$$

$$= \frac{1}{d^2}\sum_x \mathrm{Tr}(K_x)\,\mathrm{Tr}(K_x^*) = \frac{1}{d^2}\sum_x |\mathrm{Tr}(K_x)|^2 .$$

Ex: Perfect classical channel: $\{|x\rangle\}_{x\in X}$ basis

$$\mathcal{W}(S) = \sum_x |x\rangle\langle x|S|x\rangle\langle x| .$$

measures and saves outcome "x".

Easy: classical capacity $C(\mathcal{W}) = \log|X|$.

Expect this channel is not useful for transmitting quantum information.

Compute $F_c(\mathcal{D}\circ\mathcal{W}\circ\mathcal{E})$ and show it cannot be close to 1 for $M\geq 2$.

Can write $\mathcal{D}\circ\mathcal{W}\circ\mathcal{E}(S) = \sum_{k,\ell,x} D_\ell |x\rangle\langle x| E_k S E_k^* |x\rangle\langle x| D_\ell^*$

So

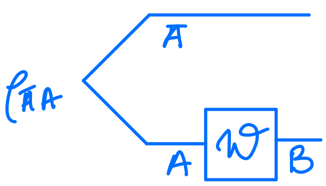
$$
\begin{aligned}
M^2\cdot F_c(\mathcal{D}\circ\mathcal{W}\circ\mathcal{E}) &= \sum_{k,\ell,x} \left|\mathrm{Tr}(D_\ell |x\rangle\langle x| E_k)\right|^2 \\
&= \sum_{k,\ell,x} \left|\langle x|E_k D_\ell|x\rangle\right|^2 \\
&\leq \sum_{k,\ell,x} \langle x|E_k E_k^*|x\rangle \langle x|D_\ell^* D_\ell|x\rangle \qquad \text{(Cauchy-Schwarz)} \\
&= \sum_{k} \mathrm{Tr}(E_k E_k^*) \qquad \text{Using } \sum_\ell D_\ell^* D_\ell = I \text{ and } \sum_k E_k^* E_k = I \\
&= \dim Q = M.
\end{aligned}
$$

So $F_c(\mathcal{D}\circ\mathcal{W}\circ\mathcal{E}) \leq \frac{1}{M}$.

Def (Coherent information).

For a channel $\mathcal{W}: L(A)\to L(B)$, the coherent information is defined as

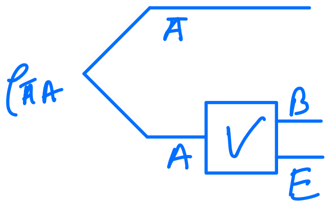
$$I_c(\mathcal{W}) = \max_{\rho_{\bar{A}A}\in S(\bar{A}\otimes A)} -H(\bar{A}|B)_{(I_{\bar{A}}\otimes\mathcal{W})(\rho_{\bar{A}A})}.$$



Rks:

- $\omega \mapsto -H(\bar{A}|B)_\omega$ is convex so maximum is achieved on extreme points i.e. pure states. Can restrict to $\rho_{\bar{A}A}$ pure.
- As all purifications of ρ_A in $\bar{A}$ are related by a unitary on $\bar{A}$ and $-H(\bar{A}|B)$ is invariant under unitaries on $\bar{A}$ we can choose a fixed purification of ρ_A.

- By considering a Stinespring representation, there exists an isometry $V: A \to B \otimes E$ (ie $V^* V = I_A$)



For $\rho_{\bar{A}A} = |\psi\rangle\langle\psi|_{\bar{A}A}$ pure

After applying the channel:

$$\omega_{\bar{A}BE} := (I_{\bar{A}} \otimes V) \rho_{\bar{A}A} (I_{\bar{A}} \otimes V^*) \text{ is also pure}$$
$$= (I_{\bar{A}} \otimes V)|\psi\rangle\langle\psi|(I_{\bar{A}} \otimes V^*).$$

$$-H(\bar{A}|B)_\omega = -H(\bar{A}B)_\omega + H(B)_\omega = \underbrace{-H(E)_\omega}_{\omega_{\bar{A}BE} \text{ is pure}} + H(B)_\omega.$$

So $$I_c(\mathcal{N}) = \max_{\rho_A} H(B)_{V\rho_A V^*} - H(E)_{V\rho_A V^*}$$

$I_c(\mathcal{N})$ large: output of the channel $\mathcal{N}$ has more entropy than environment.

Note that this expression does not depend on choice of dilation V as they are all related by a unitary on E.

<u>Ex</u>: - For any channel $\mathcal{N}$, $\underline{I_c(\mathcal{N}) \geq 0}$.

Choose $\rho_{\bar{A}A} = |0\rangle\langle 0|_{\bar{A}} \otimes |0\rangle\langle 0|_A$ where $|0\rangle \in A$ an arbitrary unit vector.

$$\omega_{\bar{A}B} = |0\rangle\langle 0|_{\bar{A}} \otimes \mathcal{N}(|0\rangle\langle 0|)$$

$$-H(\bar{A}|B)_\omega = -H(\bar{A}B)_\omega + H(B)_\omega = \underbrace{-H(\bar{A})_\omega}_{=0} - H(B)_\omega + H(B)_\omega = 0$$

- <u>Erasure channel</u> $\mathcal{N}: L(\underset{\text{span}\{|0\rangle,|1\rangle\}}{\mathbb{C}^2}) \to L(\underset{\text{span}\{|0\rangle,|1\rangle,|e\rangle\}}{\mathbb{C}^3})$ erasure symbol

$$\mathcal{N}(S) = (1-p)S + p\,\mathrm{Tr}(S)|e\rangle\langle e|$$

$$\rho_{\bar{A}A} = |\psi\rangle\langle\psi|_{\bar{A}A} \quad \text{with} \quad |\psi\rangle_{\bar{A}A} = \sqrt{\lambda}\,|00\rangle_{\bar{A}A} + \sqrt{1-\lambda}\,|11\rangle_{\bar{A}A}.$$

$$\begin{aligned}
\omega_{AB} = (I_{\bar{A}} \otimes \mathcal{W})(\rho_{\bar{A}A}) &= \lambda\,|0\rangle\langle 0|_{\bar{A}} \otimes \mathcal{W}(|0\rangle\langle 0|) + (1-\lambda)\,|1\rangle\langle 1|_{\bar{A}} \otimes \mathcal{W}(|1\rangle\langle 1|) \\
&\quad + \sqrt{\lambda(1-\lambda)}\,|0\rangle\langle 1|_{\bar{A}} \otimes \mathcal{W}(|0\rangle\langle 1|) + \sqrt{\lambda(1-\lambda)}\,|1\rangle\langle 0|_{\bar{A}} \otimes \mathcal{W}(|1\rangle\langle 0|) \\
&= (1-p)\,|\psi\rangle\langle\psi|_{\bar{A}B} + p\left(\lambda|0\rangle\langle 0|_{\bar{A}} + (1-\lambda)|1\rangle\langle 1|_{\bar{A}}\right) \otimes |e\rangle\langle e|_B
\end{aligned}$$

$$\begin{aligned}
-H(\bar{A}B) &= (1-p)\log(1-p) + p\lambda\log(p\lambda) + p(1-\lambda)\log(p(1-\lambda)) \\
&= (1-p)\log(1-p) + p\log p + p\left(\lambda\log\lambda + (1-\lambda)\log(1-\lambda)\right)
\end{aligned}$$

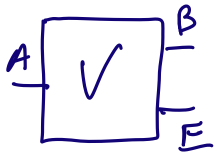
$$\begin{aligned}
H(B) &= -(1-p)\lambda\log((1-p)\lambda) - (1-p)(1-\lambda)\log(1-p)(1-\lambda) - p\log p \\
&= -(1-p)\log(1-p) - (1-p)\left(\lambda\log\lambda + (1-\lambda)\log(1-\lambda)\right) - p\log p
\end{aligned}$$

$$\begin{aligned}
I_c(\mathcal{W}) &= \max_{\lambda\in[0,1]} (1-p)h_2(\lambda) - p\,h_2(\lambda) \\
&= \begin{cases} 1-2p. & \text{for } p \le 1/2 \\ 0 & \text{for } p \ge 1/2 \end{cases}
\end{aligned}$$

$h_2(\lambda) = -\lambda\log\lambda - (1-\lambda)\log(1-\lambda)$
binary entropy function.

Rk: The classical capacity of this channel is $1-p$. The fact that $I_c(\mathcal{W}) = 0$ is for $p = \frac{1}{2}$ is consistent with no-cloning.

Stinespring dilation



Channel $A \to B$ and $A \to E$ are both erasure channels with $p = 1/2$

If $I_c(\mathcal{W}) > 0$ could transmit information to B and E simultaneously

Th: For any channel $\mathcal{W}$:

$$Q(\mathcal{W}) = \sup_n \frac{1}{n} I_c(\mathcal{W}^{\otimes n})$$

Rks: * Easy to see I_c is superadditive. But it is not additive even for very simple channels e.g. depolarizing $\mathcal{W}(\rho)=(1-p)\rho+p\frac{I}{d}$

$Q(\mathcal{W}) \geq I_c(\mathcal{W})$ but inequality can be strict.

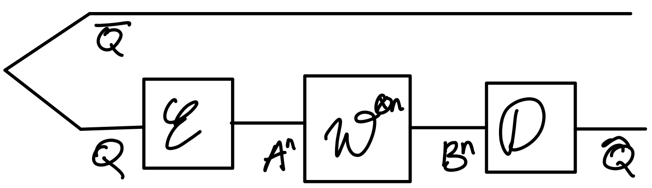
* Unlike for classical capacity where $C(\mathcal{W})=0 \Leftrightarrow$ input & output independent.

For Q we do not have a good characterization of $Q(\mathcal{W})=0$

Proof sketch:

- Converse bound. Assume $\varepsilon=0$

Consider an $(M,\varepsilon=0)$ code for $\mathcal{W}^{\otimes n}$



We have $\mathcal{I}_{\bar{Q}}\otimes\mathcal{D}\circ\mathcal{W}^{\otimes n}\circ\mathcal{E}(|\Phi\rangle\langle\Phi|_{\bar{Q}Q}) = |\Phi\rangle\langle\Phi|_{\bar{Q}\hat{Q}}$

Note that

$$H(\bar{Q}|\hat{Q})_{|\Phi\rangle\langle\Phi|} = H(\bar{Q}\hat{Q}) - H(\hat{Q}) = -\log M.$$

Data processing inequality

$$H(\bar{Q}|B^n)_{(\mathcal{I}_{\bar{Q}}\otimes(\mathcal{W}^{\otimes n}\circ\mathcal{E}))(|\Phi\rangle\langle\Phi|)} \leq H(\bar{Q}|\hat{Q})_{\mathcal{I}_{\bar{Q}}\otimes\mathcal{D}\circ\mathcal{W}^{\otimes n}\circ\mathcal{E}(|\Phi\rangle\langle\Phi|)}$$

$$= -\log M.$$

But $I_c(\mathcal{W}^{\otimes n}) \geq -H(\bar{Q}|B^n)_{(\mathcal{I}_{\bar{Q}}\otimes(\mathcal{W}^{\otimes n}\circ\mathcal{E}))(|\Phi\rangle\langle\Phi|)} \geq \log M.$

$\mathcal{I}_{\bar{Q}}\otimes\mathcal{E}(|\Phi\rangle\langle\Phi|)$ is a specific state.

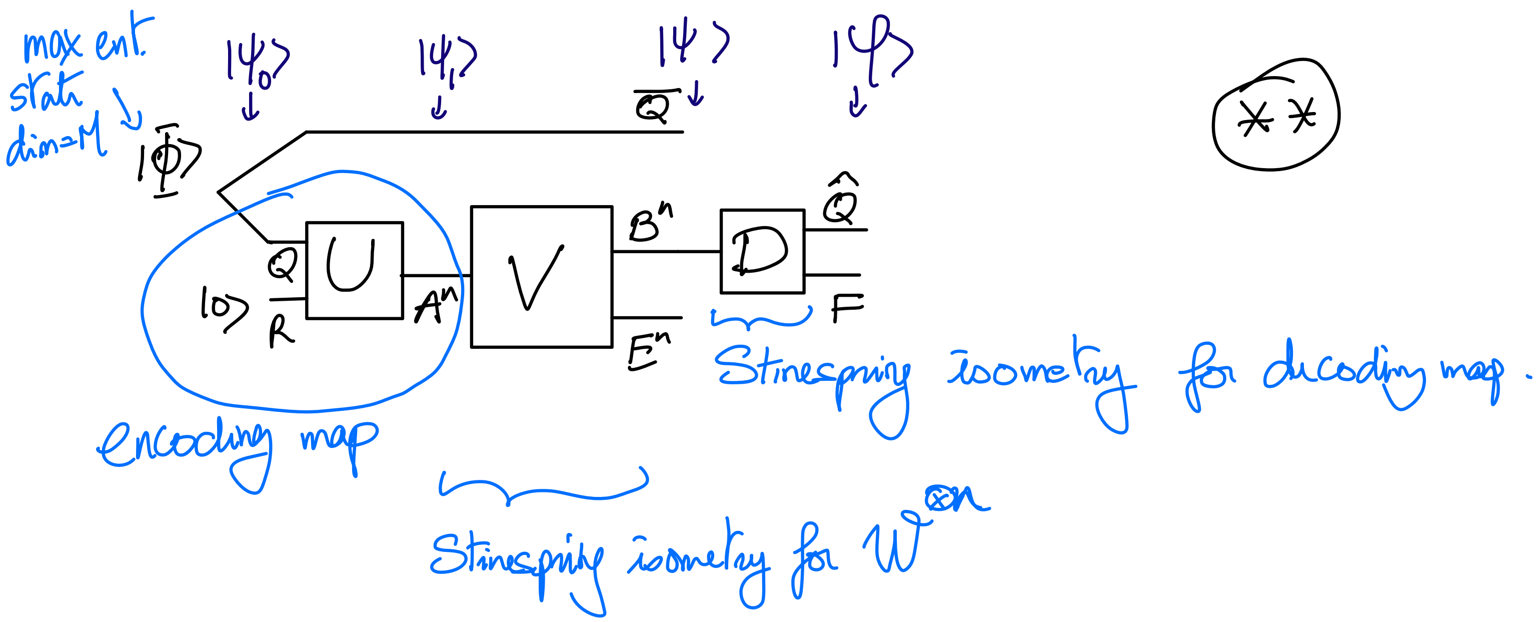
$$\frac{\log M}{n} \leq \frac{I_c(\mathcal{W}^{\otimes n})}{n}$$

For $\varepsilon > 0$, show that $H(\bar{Q}|\hat{Q})_{(\mathcal{I}_{\bar{Q}} \otimes \mathcal{D} \circ \mathcal{W}^{\otimes n} \circ \mathcal{E})(|\Phi\rangle\langle\Phi|)} \underset{\sim}{\leq} -(1-\varepsilon)\log M$

Fano's inequality

• **Achievability**.

Have to construct an encoding & decoding.



$(**)$

Use the idea of decoupling

Assume we managed to construct U such that

$$F\left(\psi_{\bar{Q}E^n},\ \psi_{\bar{Q}} \otimes \underbrace{\gamma_{E^n}}_{\text{arbitrary state}}\right)^2 \geq 1-\varepsilon \qquad (*)$$

where $\psi = |\psi\rangle\langle\psi|$

* A purification of $\psi_{\bar{Q}E^n}$ is $|\psi_{\bar{Q}B^nE^n}\rangle$
* A purification of $\psi_{\bar{Q}} \otimes \gamma_{E^n}$ is $|\Phi\rangle_{\bar{Q}\hat{Q}} \otimes |\gamma\rangle_{E^n\hat{E}^n}$.

Uhlmann's theorem: There exists an isometry

$D : B^n \to \underbrace{\hat{Q}\hat{E}^n}_{\text{purifying system for state } \psi_{\bar{Q}} \otimes \gamma_{E^n}}$ such that

$\underbrace{B^n}_{\text{purifying system for state } \psi_{\bar{Q}E^n}}$

$$\underbrace{F(\psi_{\bar{Q}E^n}, \psi_{\bar{Q}} \otimes \psi_{E^n})^2}_{1-\varepsilon \le} = |\langle \Phi|_{\bar{Q}\hat{Q}} \otimes \langle\gamma|_{E^n\hat{E}^n} D |\psi_{\bar{Q}B^nE^n}\rangle|^2$$

$$= F(|\Phi\rangle\langle\Phi|_{\bar{Q}\hat{Q}} \otimes |\gamma\rangle\langle\gamma|, D\psi_{\bar{Q}B^nE^n}D^*)^2$$

$$\le F(|\Phi\rangle\langle\Phi|_{\bar{Q}\hat{Q}}, \mathrm{Tr}_{\hat{E}^nE^n}(D\psi D^*))^2$$

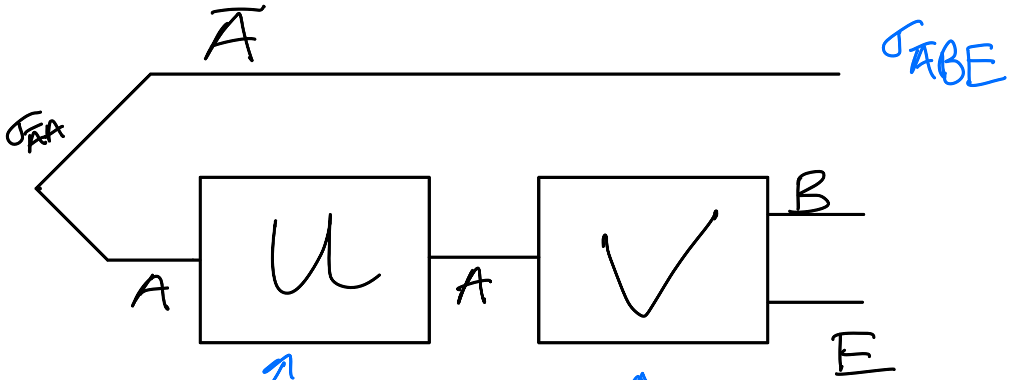
Conclusion: $\mathcal{D}(S) = \mathrm{Tr}_{\hat{E}^n}(DSD^*)$ is a decoder achieving fidelity $1-\varepsilon$ provided decoupling property $(*)$ is satisfied.

It remains to find U so that $(*)$ is satisfied when $\frac{\log M}{n} \le \frac{I_c(\mathcal{W}^{\otimes n})}{n} - \delta$.

Idea: choose U at random from Haar measure.

Recall $I_c(\mathcal{W}^{\otimes n}) = \max_{\rho_{A^n}} H(B^n)_{\mathcal{V}^{\otimes n}(\rho_{A^n})\mathcal{V}^{\otimes n}} - H(E^n)_{\mathcal{V}^{\otimes n}(\rho_{A^n})\mathcal{V}^{\otimes n}}$

Will only show that when $\frac{\log M}{n} \le (\log \dim B - \log \dim E) - \delta$ good codes exist.



unitary chosen at random   fixed isometry. $V^\dagger V = I$

Th (Decoupling)

replace by marginal

$$\mathbb{E}_U \left\{ \Delta\left(\sigma_{\bar A E}\,,\ \sigma_{\bar A} \otimes \frac{\mathrm{Tr}_B VV^\dagger}{\dim(A)}\right)^2 \right\} \le \frac{1}{2}\, \frac{(\dim \bar A)(\dim E)}{\dim B}\, \mathrm{Tr}\left(\sigma_{\bar A A}^2\right)$$

[See Dupuis Thesis for a more general statement].

Use this with for diagram (**) with

$$\bar A \to \bar Q\,,\quad A \to Q \otimes R\,,\quad B \to B^n\,,\quad E \to E^n$$

$$\sigma_{\bar A A} = \underbrace{|\Phi\rangle\langle\Phi|_{\bar Q Q}}_{M\text{-dim max entangled state}} \otimes \underbrace{|0\rangle\langle 0|_R}_{\text{fixed state}}$$

$$\text{Choose} \quad M = \varepsilon^2 \cdot \frac{\dim B^n}{\dim E^n} = 2^{\,n(\log\dim B - \log\dim E) - 2\log 1/\varepsilon}$$

corresponds to $H(B) - H(E)$ for a state with maximally mixed marginal.

$\exists\ \mathcal{U}$ s.t.

$$\Delta\left(\psi_{\bar{A}E^n}, \psi_{\bar{A}} \otimes \gamma_{E^n}\right)^2 \leq \varepsilon^2$$

$$F\left(\psi_{\bar{A}E^n}, \psi_{\bar{A}} \otimes \gamma_{E^n}\right) \geq 1 - \Delta \geq 1 - \varepsilon.$$

$\Rightarrow$ Can find a decoder achieving a channel fidelity $(1-\varepsilon)^2$ ▪

Rk: Description of the code not explicit: U random and decoder obtained from Uhlmann theorem.
In practice, want $\mathcal{E}$ and $\mathcal{D}$ explicit, efficient, "geometrically friendly", fault tolerant, ...

Proof of decoupling inequality:

Recall $\Delta(\rho, \sigma) = \frac{1}{2}\|\rho - \sigma\|_1$. Will assume $V = I$, $A = BE$

- Trace norm $\rightarrow$ Frobenius norm (Cauchy-Schwarz)

$\| \quad \|_1^2 \leq (\dim) \| \quad \|_2^2$

$$\left\| \sigma_{\bar{A}E} - \sigma_{\bar{A}} \otimes \frac{I_E}{\dim E} \right\|_1^2 \leq (\dim \bar{A})(\dim E) \left\| \sigma_{\bar{A}E} - \sigma_{\bar{A}} \otimes \frac{I_E}{\dim E} \right\|_2^2$$

$$= \dim \bar{A} \dim E \left(\mathrm{Tr}\left(\sigma_{\bar{A}E}^2\right) - 2\,\mathrm{Tr}\left(\sigma_{\bar{A}E} \cdot \sigma_{\bar{A}} \otimes \frac{I_E}{\dim E}\right) + \mathrm{Tr}\left(\sigma_{\bar{A}}^2\right) \cdot \frac{1}{\dim E} \right)$$

$$\mathrm{Tr}\left(\sigma_{\bar{A}E} \cdot \sigma_{\bar{A}} \otimes I_E\right) = \mathrm{Tr}\left(\sigma_{\bar{A}}^2\right)$$

$$= (\dim \bar{A})(\dim E)\left(\mathrm{Tr}\left(\sigma_{\bar{A}E}^2\right) - \mathrm{Tr}\left(\sigma_{\bar{A}}^2\right) \cdot \frac{1}{\dim E}\right)$$

- Compute expectations

* $$\mathbb{E}_{\mathcal{U}}\left\{ \underbrace{\mathrm{Tr}(\sigma_{\bar{A}}^2)\cdot\frac{1}{\dim E}}_{\text{independent of } U} \right\} = \mathrm{Tr}(\sigma_{\bar{A}}^2)\cdot\frac{1}{\dim E}$$

dropped some $I_{\bar{A}}$.

* $$\mathbb{E}_{\mathcal{U}}\left\{\mathrm{Tr}\left(\sigma_{\bar{A}E}^2\right)\right\} = \mathbb{E}_{\mathcal{U}}\left\{\mathrm{Tr}\left[\left(\mathrm{Tr}_B\left(U\sigma_{\bar{A}A}U^* \right)\right)^2\right]\right\}$$

Expectation of a polynomial in U $\Rightarrow$ can in principle compute it via Weingarten calculus.

$$= \mathbb{E}_{\mathcal{U}}\left\{\mathrm{Tr}\left[\mathrm{Tr}_{B_1}\left(U\sigma_{\bar{A}_1A_1}U^* \right)\otimes\mathrm{Tr}_{B_2}\left(U\sigma_{\bar{A}_2A_2}U^* \right) F_{\bar{A}_1E_1,\bar{A}_2E_2}\right]\right\}$$

flip operator

$$= \mathbb{E}_{\mathcal{U}}\left\{\mathrm{Tr}\left[U\sigma_{\bar{A}_1A_1}U^* \otimes U\sigma_{\bar{A}_2A_2}U^* \; I_{B_1B_2}\otimes F_{\bar{A}_1\bar{A}_2}\otimes F_{E_1E_2}\right]\right\}$$

$$= \mathrm{Tr}\left(\sigma_{\bar{A}_1A_1}\otimes\sigma_{\bar{A}_2A_2}\cdot F_{\bar{A}_1\bar{A}_2}\otimes \underbrace{\mathbb{E}_{\mathcal{U}}\left\{U^*_{A_1}\otimes U^*_{A_2}\left(I_{B_1B_2}\otimes F_{E_1E_2}\right) U\otimes U\right\}}_{=\alpha I_{A_1A_2}+\beta F_{A_1A_2}}\right)$$

$$(\dim B)^2\dim E = \mathrm{Tr}\left(I_{B_1B_2}\otimes F_{E_1E_2}\right) = \alpha(\dim A)^2 + \beta(\dim A)$$

$$\dim B(\dim E)^2 = \mathrm{Tr}\left(F_{B_1B_2}\otimes I_{E_1E_2}\right) = \alpha(\dim A) + \beta(\dim A)^2$$

Solving this system:

$$\alpha = \frac{\dim B\dim A - \dim E}{(\dim A)^2 - 1} = \frac{1}{(\dim E)}\cdot\left(\frac{1-\frac{\dim E}{(\dim A)^2}}{1-1/(\dim A)^2}\right) \simeq \frac{1}{\dim E}.$$

$$\beta \simeq \frac{1}{\dim B}$$

$$= \mathrm{Tr}\left(\rho_{\bar{A}_1 A_1} \otimes \rho_{\bar{A}_2 A_2} \cdot F_{\bar{A}_1 \bar{A}_2} \otimes \left(\alpha I_{A_1 A_2} + \beta F_{A_1 A_2}\right)\right)$$

$$= \alpha \, \mathrm{Tr}\left(\rho_{\bar{A}}^2\right) + \beta \, \mathrm{Tr}\left(\rho_{\bar{A}A}^2\right) \quad \blacksquare$$