Asymptotic classical communication capacity of quantum states & channels.

Recall given an ensemble $\mathcal{E} = \{p_x, f_x\}$

Treating the $x \rightarrow y$ process as a classical channel, the capacity is $I_{acc}$ of the ensemble.

But we can do better given the Q box -- not because of Alice but because of Bob!

We can do even better given a channel --

$$I_{acc}(\mathcal{E}) = \max_m I(X; Y) \leq C(\mathcal{E}) = \sum_x p_x f_x + \sum_x p_x S(f_x)$$

Q1 How much can Alice comm to Bob if

1. She decides what $x$ instead of drawing $x \sim p(x)$
2. She can use Q box many times? (Bob measures collectively)

Q2 Same question but a channel $N$ (instead of $Q$) is available instead?
Results:
\[ C(N) = \sup_{n} \frac{1}{n} X(\eta^n) \]
where \( X(\eta) = \max_{\epsilon \in \{0,1\}} X(\eta|\epsilon) \)

Direct coding uses the fact:
Asymptotic comm rate of \( \Theta \):
\[ = \max_{p(x)} X(f(p_x, p_x^3)) \]

Converse uses:
- upper bound on # outcomes in optimal measurements
- classical Fano's inequality

Uses random quantum codewords + pretty good measurements

Unconditional typicality \Rightarrow Conditions for gentle meas lemma & packing lemma
Suppose a Q system is prepared in one of the possible states $\rho_1, \rho_2, \ldots, \rho_k$.

The PHM is given by the POVM:

$$M_i = \Lambda^\frac{1}{2} \rho_i \Lambda^{-\frac{1}{2}} \quad \text{for} \quad i = 1, \ldots, k$$

$$M_{K+1} = I - \sum_{i=1}^{K} M_i$$

$$\Lambda = \sum_{i=1}^{K} \rho_i$$, and $\Lambda^\frac{1}{2}$ performed only on supp $(\Lambda)$.

Note that the PHM is still well defined if $\rho_i > 0$ but otherwise un constraint (OK if tr $\rho_i \neq 1$ or $\rho_i \notin SU$).
Elaborating the notations:

\[ \Lambda = \frac{\kappa}{\sum_{i=1}^{k} p_i} \] is positive semidefinite.

Let \( \Lambda = \sum_{j} \lambda_j I_{\Phi_j^c} \) be its spectral decomposition:

\[ \Lambda = \sum_{j} \lambda_j^{-1} I_{\Phi_j^c} \Phi_j, \quad \text{supp}(\Lambda) = \sum_{j} I_{\Phi_j^c} \Phi_j \]

Thus:

\[ M_{\pi} \geq 0, \quad \sum_{i=1}^{K} M_{\pi_i} = \Lambda^{-\frac{1}{2}} \Lambda^{-\frac{1}{2}} = I_{\text{supp}(\Lambda)} \]

\[ M_{K+1} = I - I_{\text{supp}(\Lambda)} = I_{\text{complement}(\text{supp}(\Lambda))}. \]

\[ \forall \pi_i, \quad 0 \leq \text{Tr}(p_i M_{K+1}) \leq \text{Tr}(\Lambda M_{K+1}) = 0 \] (\( i = 1 \).)
Intuitively, the measurement has an error if the state is $\psi_i$ but the outcome is $\psi_j$.

How good is the "pretty good" meas?

* If $\psi_i$'s are orthogonal, it is perfect.

* Upon "googling" many hits on $P_{6,M}$ being optimal in specific applications.

* For pure states $\psi_i = \chi_i \chi_i^* \frac{1}{\sqrt{2}}$ given equally probably, are prob error $\leq \frac{1}{2} \sum_{i\neq j} |\langle \psi_i | \psi_j \rangle|^2$

We'll use packing lemma to bound error prob.
Gentle measurement lemma (Winter ...):

Let $\rho > 0$, $\text{tr}(\rho) \leq 1$, $0 \leq E \leq I$

If $\text{Tr}(\rho M) \geq 1 - 3^{-1}$

then $\exists U$ s.t. $\| U E^2 f E^2 U^+ - \rho \|_{\text{tr}} \leq \sqrt{38}$

(best possible post-meas state for the outcome corresponding to $E$)

Interpretations:

Given a state $\rho$, if meas yields 1 outcome $\text{whp}$

then the state is hardly changed by the meas.
Def: for a set of states $S = \{ \phi_1, \phi_2, \ldots, \phi_k \}$ let distinguishability error of $S$ be
\[
\text{de}(S) = \min_{i} \max_{j \neq i} \Pr(\text{outcome } = j \mid \text{state } = \phi_i)
\]

NB can upper bound $\text{de}(S)$ by considering specific measurements.
The packing lemma:

Let $p_x = \text{prob}(x)$, $S_x$ states, $S = \sum_x p_x S_x$

If projectors $\Pi$, $\Pi_x$ exist s.t. $\forall x$:

1. $\text{Tr} (S_x \Pi) \geq 1 - \varepsilon$
2. $\text{Tr} (S_x \Pi_x) \geq 1 - \varepsilon$
3. $\text{Tr}(\Pi_x) \leq d_1$
4. $\Pi S \Pi \leq \frac{d_0}{d_1}$  \[ \text{for fraction} \]

5. $S = \{ \gamma_1, \ldots, \gamma_K \}$, each $\gamma_x = S_x \text{ wp } p_x$ (iid), $K = \frac{d_0 f}{d_1}$

Then $\mathbb{E}_{S} \text{de}(S) \leq 2 \left( 3 + \sqrt{8 \varepsilon} \right) + 4 f$ (achieved w/ "P6M").
What does this lemma mean?

Conditions ② & ③ say each $S_x$ lives in some $d_1$-dim space (up to $\varepsilon$ approx) defined by $S_x$. Condition ③ says all $S_x$ lives in a space defined by $T_1$. 

---

The diagram illustrates the relationship between $T_1$, $T_2$, $T_3$, and $T_4$, indicating their dimensions and interconnections.
Condition 4 says $\pi \lambda \Sigma \leq \frac{12}{d_0}$

Since $\text{tr}(\Sigma x \Pi) \geq 1 - \lambda$, $\Pi \Sigma x \Pi \approx \Sigma x$

& $\Pi \Sigma \Pi \approx \Sigma$

Condition 4 bounds the max eigenvalue of $\Sigma$

It is no more than $\frac{1}{d_0}$, or $d_0 = \frac{1}{\lambda_{\text{max}}(\Sigma)}$

Where $\lambda_{\text{max}}(\Pi)$ denotes the max eigenvalue of $\Pi$.\n
In general: for 2 rank pure states \( |\Psi\rangle \text{ & } |\Phi\rangle \),

\[ \lambda_{\text{max}} (|\Psi \rangle \langle \Psi | + |\Phi \rangle \langle \Phi |) \text{ is large when } |\Psi \rangle \text{ & } |\Phi \rangle \text{ have high overlap.} \]

In the current problem: \( S = \sum_{x} \lambda_{x} S_{x} \)

We expect \( \lambda_{\text{max}} (S) \) to be small if the \( S_{x} \)'s are distinguishable but having mixed state \( S_{x} \) complicates things.
\[ S_1 = \left(\frac{1}{4} 10 \times 01 \right) + \ldots \]

\[ S_2 = \left(\frac{1}{4} 10 \times 01 \right) + \ldots \]

Here \ldots are smaller terms that do not enter the calculation of \( \text{Max}(S_1, S_2) \).

None the less, \ldots still affect their distinguishability.

This problem is worse if \( \text{rank}(S_1, S_2) \) are large.
The packing lemma tells us, if we're communicating using quantum states $\tilde{S}_x$ and we know little about them except each lives under $T(x)$ of dim $d_i$, then we can send $k = \left(\frac{d_0}{d_1}\right)^{\frac{1}{2}}$ of message w/o much error.

The fraction of space we're willing to learn blank.

This comes from how distinguishable $\tilde{S}_x$'s are.

(Also $d_0 \geq \text{rank}(\tilde{S})$, $d_i \approx$ size of $\tilde{S}_x \Rightarrow \frac{d_0}{d_1}$ sounds right.)
\[ S_0 = (4_0 \times 4_0) \quad P_0 = \frac{1}{2} \]

\[ S_1 = (4_1 \times 4_1) \quad P_1 = \frac{1}{2} \]

\[ \Xi = \begin{bmatrix} 1 & 0 \\ 0 & P \end{bmatrix} \]

Choose \( \bar{1}_0 = (4_0 \times 4_0) \quad \bar{1}_1 = (4_1 \times 4_1) \quad \rho = 1 \)

If we choose \( \bar{1}_1 = 10 \times 10 \) \( \text{when } p \text{ large} \)

\[ \cancel{\bar{1}_1} \times \Xi \bar{1}_1 \leq \frac{\bar{4}_1}{d_0} \quad d_0 = 1 \quad \text{Then } k = 1 \]

If we choose \( \bar{1}_1 = \bar{1}_1 \) \( \text{when } p = \frac{1}{2} \)

\[ \Xi \bar{1}_1 \leq \frac{\bar{4}_1}{d_0} \quad \text{for } d_0 = 2 \quad \text{Then } k = 2 \]
Proof: Let \( y_i = S x_i \). Define a "PGM" using \( \Pi_i \Pi_{x_i} \Pi \) & lower the expected prob of error.

\[
\Lambda_i = \Pi_i \Pi_{x_i} \Pi, \quad \Lambda = \sum_{i=1}^{k} \Lambda_i,
\]

\[
M_i = \Lambda_i^{\frac{1}{2}} \Lambda_i^{-\frac{1}{2}}, \quad M_{k+1} = I - \sum_{i=1}^{k} M_i.
\]

\[
de(S) \leq \frac{1}{k} \sum_{i} \text{Tr} \ Y_i (I - M_i)
\]

The \( \Lambda_i^{\frac{1}{2}} \) in \( M_i \) is nasty...

Use \( I - M_i = I - \Lambda_i^{\frac{1}{2}} \Lambda_i^{-\frac{1}{2}} = I - \left( \Lambda_i + \sum_{j \neq i} \Lambda_j^{\frac{1}{2}} \right) \Lambda_i \left( \Lambda_i + \sum_{j \neq i} \Lambda_j^{\frac{1}{2}} \right)^{-\frac{1}{2}} \)

Optimizes:

\[
I - \left( x + y \right)^{\frac{1}{2}} \left( x + y \right)^{-\frac{1}{2}} \leq 2 \left( I - \Lambda_i \right) + 4 \sum_{j \neq i} \Lambda_j
\]

Sensible approx to \( I - M_i \) if \( \Lambda_i^{\frac{1}{2}} \approx I \).
Since $Y_i > 0$,

\[
\text{de}(S) \leq \frac{1}{K} \sum_{i} \text{Tr} \, Y_i \left( I - \Lambda_i \right)
\]

\[
\leq \frac{2}{K} \sum_{i} \text{Tr} \, Y_i \left( I - \Lambda_i \right) + \frac{1}{K} \sum_{i} \left[ \sum_{j \neq i} \text{Tr} \left( Y_i \Lambda_j \right) \right]
\]

\[
= 1 - \text{Tr} \left( \prod_{X_i} \prod_{X_i} \prod_{X_i} \right)
\]

if $\|X-Y\|_{tr} \leq C$

then $\forall 0 \leq p \leq 1$

$|\text{tr} p(X-Y)| \leq C$

+ gentle meas + (1)

$\leq 3 + \sqrt{83}$ (indeed if $\xi_i \leq \frac{1}{k} \leq \frac{3}{k}$ cancel out)

by (2)
from previous stage, seeking a bound for \( \sum_j \text{Tr}(X_i \Lambda_j) \) for \( j \neq i \).

Here use the fact \( Y_{-i} = S_x w_p x \), drawn iid.

\[
\mathbb{E} \sum_j \text{Tr}(X_i \Lambda_j) = \mathbb{E} \text{Tr} \left[ \left( \mathbb{E} S_{X_i} \right) \Pi \left( \mathbb{E} \Pi X_j \right) \Pi \right]
\]

\[
\text{independent}
\]

\[
\leq \sum_j \text{Tr}(\Pi S \Pi) \left( \mathbb{E} \Pi X_j \right)
\]

\[
\leq \sum_j \text{Tr} \left( \frac{\Pi}{d_0} \right) \left( \mathbb{E} \Pi X_j \right)
\]

\[
\leq \mathbb{E} \sum_j \text{Tr} \left( \frac{\Pi}{d_0} \right) \Pi X_j \mathbb{I} \left( \text{rank} \leq d_1 \right) \left( \text{eigenvalue=0,1} \right)
\]

\[
\leq \mathbb{E} \sum_j \frac{d_1}{d_0} = \frac{\mathbb{E}}{d_0} \leq f.
\]
Back to earlier Qn:

Let $Q$ be a box that emits $P_x$ to Bob if Alice inputs $x$.

For any $n$, consider an $(M, n)$ code transmitting $\log M$ bits to Bob by $n$ uses of $Q$.

Let $P_e = \text{worst prob error}$, $E_P$ expected prob error.

$R$ achievable if there are $(2^n, n)$ codes w/ $P_e \to 0$

Call capacity of $Q$, $C(Q) = \sup R$ achievable.
Theorem: $C(Q) = \max_{pX} \chi(pX, fX)$

Again, need a direct coding proof & a converse proof.
Converse:

Consider the state \( \sum_{x_1, x_2, \ldots, x_n} p(x_1, x_2, \ldots, x_n) \prod_{i=1}^{n} x_i = \cdots = x_1 \times x_1 \cdots x_n \otimes p_{x_1} \otimes \cdots \otimes p_{x_n} \).

System labels \( x_1, x_2, \ldots, x_n \quad B_1, \ldots, B_n \)

\[ n R \leq I(X_1 : \ldots : X_n; B_1, \ldots, B_n) + \text{Holevo inf n-shot arbitrary probs} \]

\[ = S(B_1, \ldots, B_n) - S(B_1, \ldots, B_n | X_1, \ldots, X_n) \]

\[ \leq \sum_{i=1}^{n} S(B_i) - \sum_{i=1}^{n} p(x_i, \ldots, x_n) S(p_{x_i} \otimes \cdots \otimes p_{x_n}) \]

Need \( B_1, \ldots, B_n \) in product state \( \sum_{i=1}^{n} S(B_i) - \sum_{i=1}^{n} p(x_i, \ldots, x_n) \sum_{i=1}^{n} S(p_{x_i}) \geq n \ I(X_i : B_i) \) arbitrary probs 1-shot Holevo inf

\[ \leq \sum_{i=1}^{n} S(B_i) - \sum_{i=1}^{n} p(x_i) S(p_{x_i}) \]

Marginal
Direct coding:

Recall in Shannon's noisy coding theorem

\[ C_1 = X_1, X_2, \ldots, X_n \quad \text{where} \quad X_{ij} \sim p_X \text{ i.i.d.} \]

\[ C_2 = X_{i1}, X_{i2}, \ldots, X_{in} \]

\[ C_M = X_{M1}, X_{M2}, \ldots, X_{Mn} \]

Whp these are typical sequences (prob of outcome \( \approx 2^{-nH(X)} \))

Here, we demand \( C_i \)'s be drawn randomly among strongly typical sequences (next page) and to send message \( i \)

Alice inputs \( X_{i1}, X_{i2}, \ldots, X_{in} \) into the \( n \) uses of \( C_i \).

States to be distinguished by Bob:

\[ Y_i = p_{X_{i1}} \otimes p_{X_{i2}} \otimes \cdots \otimes p_{X_{in}} \]
Strongly typical sequence:

For r.v. $X$, a sequence $C = x_1, x_2, \ldots, x_n$ is $\varepsilon$-strongly typical if the empirical distribution observed in $C$ is $\varepsilon$-close to $\{p(x)\}$.

Say for $x = 1$, $\hat{q}(1) = \frac{1}{n} (\# x_i's\ equal\ to\ 1)$

for $x = 2$, $\hat{q}(2) = \frac{1}{n} (\# x_i's = 2)$

and $\| \hat{q} - p \|_1 = \sum_x |\hat{q}(x) - p(x)| \leq \varepsilon$

Hence, $\varepsilon$-strongly typical sequences are $\varepsilon'$ typical.
Eq. let $\Omega = \{1, 2, 3, 4\}$, $p(a) = \frac{a}{10}$, $n = 20$

$C = 3 \ 3 \ 3 \ 4 \ 4 \ 21 \ 3 \ 2 \ 2 \ 4 \ 3 \ 4 \ 3 \ 4 \ 2 \ 4 \ 4 \ 3$

\[
\begin{array}{c|c}
\text{a} & \frac{g_a}{n} \\
\hline
1 & \frac{1}{20} \\
2 & \frac{5}{20} \\
3 & \frac{7}{20} \\
4 & \frac{7}{20} \\
\end{array}
\quad \quad
\begin{array}{c|c}
\text{p} & \frac{g_p}{n} \\
\hline
1 & \frac{1}{20} \\
2 & \frac{2}{10} \\
3 & \frac{3}{10} \\
4 & \frac{4}{10} \\
\end{array}
\]

$||p - \frac{g}{n}|| = 0.2$, so $C$ is 0.2-Strongly typical.

NB we focus on $n \gg 181$
Given \( C = X_1 \ldots X_n \) is strongly typical.

What do we know about \( \mathcal{E} = \mathcal{E}_1 \otimes \ldots \otimes \mathcal{E}_n \)?

Up to reordering the \( n \) systems, it's \( \bigotimes_x n q_x(x) \).

For each \( x \), by discussion leading to quantum data compression, there's a projector \( \Pi_x \) s.t.

\[
\text{Tr} \left( \Pi_x f_x \otimes n q_x(x) \right) \geq 1 - \delta_1,
\]

\[
\dim \Pi_x \leq 2^{-n q_x(x) \left( S(f_x) + \epsilon_1 \right)}.
\]

Given \( C \), let \( \Pi_C = \bigotimes_x \Pi_x \) acting on the correct systems:

\[
\text{Tr} \left[ \Pi_C f_1 \otimes \ldots \otimes f_n \right] \geq 1 - \delta_2,
\]

\[
\dim \Pi_C \leq 2^{-n \left[ \sum_x q_x(x) S(f_x) + \epsilon_2 \right]}.
\]
e.g. when \( C = 3334423222433432443 \),

7 systems compressed by \( \mathcal{T}_4 \)

state \( Y_C = \underbrace{f_3 \otimes f_3 \otimes f_3 \otimes f_4 \otimes f_4 \otimes \cdots} \quad \otimes f_4 \otimes f_3 \)

for \( x = 1 \), \( \mathcal{T}_1 \) acts on sys 7

\( x = 2 \), \( \mathcal{T}_2 \) acts on sys 6, 9, 10, 11, 17

\( x = 3 \), \( \mathcal{T}_3 \) acts on sys 1, 2, 3, 8, 13, 15, 20

\( x = 4 \), \( \mathcal{T}_4 \) acts on sys 4, 5, 12, 14, 16, 18, 19
More precisely,
for each $n$, for $C$, $\gamma$, $\Pi C$ defined earlier,

& let $f = \sum_x p_x p_x$, $\Pi$ projector onto typical space of $f^\otimes n$

\[ \exists \, \Sigma_n, \delta_n \quad \text{s.t.} \]

1. $Tr(\gamma C \Pi C) \geq 1 - \delta_n$
2. $Tr(\Pi C) \leq 2^{-n \left[ \sum_x p_x S(p_x) + \Sigma_n \right]}$

\text{proof outlined already}

\text{proof}

\[ \exists \gamma \]
3. $Tr(\gamma C \Pi) \geq 1 - \delta_n$
4. $\Pi^{\otimes n} \Pi \leq 2^{-n \left[ S(p) - \Sigma_n \right]} \Pi$

(and $\Sigma_n, \delta_n \to 0$ as $n \to \infty$)
Finally, back to direct coding proof for $C(a_2)$:

The $M$ messages are encoded as: $x_1, x_2, \ldots, x_m$

Where $x_i = p_{x_1} \otimes p_{x_2} \otimes \cdots \otimes p_{x_{in}} = : P_c$

For $C_i = x_{i1} x_{i2} \cdots x_{in}$ a randomly chosen strongly typical sequence.

The hacking lemma now applies.

$S(x)$ being $P_c$, $C$ strongly typical.

$T_1, x$ being $T_1 C$

$d_1, \ldots, 2^n [X(p_x S(p_x) + \delta_n)]$

$T_1, \ldots, T_i$

$\cdots$

$T_1, \ldots, T_i$

$d_1, \ldots, 2^n [S(\frac{m}{n} p_x p_x^t) + \delta_n]$

$K, \ldots, M$

$\vdash M$ can be done.

$\approx 2^n [X(p_x, p_x^t) - \delta_3]$

$\approx 2^n$ takes care of $f & \delta_n$.
Consider the states

\[ p_1 = \{(\pi/4, 10, 0.9, 0.1) \} \text{ for } 14_1 = 10 \]
\[ p_2 = \{(\pi/2, 10, 0.9, 0.1) \} \text{ for } 14_2 = \cos \theta 10 \}
\[ p_3 = \{(\pi/3, 10, 0.9, 0.1) \} \text{ for } 14_3 = \sin \theta 10 \}

\text{such that } 0.4p_1 + 0.3p_2 + 0.3p_3 = \frac{\pi}{2}

So, the box emits one of \( p_1, p_2, p_3 \) on demand.

Applying what we proved, \( C(\Omega) = \max_{p_1, p_2, p_3} \chi(p_1, p_2, p_3) \)

Since \( S(p_x) \) same, \( \chi(p_1, p_2, p_3) = S(\xi_{p_1, p_2, p_3}) - H(\xi_{p_1, p_2, p_3}). \)

By max \( \xi_{p_1, p_2, p_3} = \frac{\pi}{2} \), or \( p_1 = 0.4, \{ p_2, p_3 = 0.3 \} \).
What should Alice send to Bob?

Consider all strings of 1 2 3's of length $n$ with $\approx 40\%$ 1's, 30\% 2's, 30\% 3's.

$H(0.4, 0.3, 0.3) = 1.571$

There are $\approx 2^{n(1.571 + 2)} \approx 2.971^n$ such strings.

Alice is drawing 2 strings from this pool (with replacement). If $d = 0.02$, then are $120 = m$ messages.
Say \( C_1 = 2 \ldots 2 \ldots 3 \ldots 1 \ldots 2 \ldots 1 \ldots 3 \ldots \) 
\( C_2 = 1 \ldots 3 \ldots 2 \ldots 1 \ldots 2 \ldots 1 \ldots 3 \ldots \) 
\( \vdots \) 
\( C_{120} = 2 \ldots 1 \ldots 3 \ldots 1 \ldots 2 \ldots 3 \ldots \)

\( \delta_1 = \beta_2 \beta_1 \beta_1 \beta_2 \beta_2 \beta_2 \beta_2 \beta_2 \beta_3 \beta_3 \beta_3 \beta_3 \) 
\( \delta_2 = \beta_1 \beta_3 \beta_2 \beta_1 \beta_3 \beta_2 \beta_2 \beta_2 \beta_2 \beta_1 \beta_3 \beta_3 \) 
\( \delta_{120} = \beta_2 \beta_1 \beta_3 \beta_1 \beta_2 \beta_3 \beta_1 \beta_1 \beta_2 \beta_3 \)  

Product states, but now classically correlated once \( C_1, C_2, \ldots, C_{120} \) fixed.
Bob's measurement:

\[ \Lambda_i = \prod_{k=1}^{K} \prod_{x_i} \prod_{i} \Lambda^2 \Lambda_i \Lambda^2, \quad \Lambda = \sum_{i=1}^{K} \Lambda_i, \]

\[ M_i = \Lambda^2 \Lambda_i \Lambda^2, \quad M_{k+1} = I - \sum_{i=1}^{K} M_i \]

e.g. \( Y_{1000} = \rho_2 \rho_1 \rho_3 \rho_1 \rho_2 \rho_2 \rho_1 \rho_1 \rho_2 \rho_3 \)

\[(S(\rho_1) = H(0.05) \approx 0.29)\]

Thus 4

down to 2 \( S(\rho_{1/2}) \approx 3 \) dim

Compress to 2 \( 3(0.29 + 0.1) \approx 2.25 \) dims
\[ T_{120} = (\otimes) \text{ of the 3 projections} \]

Tr or rank of \( V_{120} \approx \| 2.25 \|^{2} \times 3 \approx 14.5 \]
\[ \approx 15 \text{ dims and} \]

Theoretically, should be \( 2^{10} (0.29 + 0.1) \approx 15 \) also.

\[ \text{projection onto typical space of} \left( \Psi_{\text{p.i.p.}} \right)^{\otimes 10} \]

\[ \text{which is the full space.} \quad d_{0} \approx 2^{10} \]

\[ \text{Can say} \quad 2^{10 (1 - 0.29 - 0.2)} \]
\[ \approx 120 \text{ messages}. \]
So \( \prod_{i=1}^{120} \tilde{a}_i \approx \tilde{a}_{120} \).

Same for \( \prod_{i=1}^{120} \tilde{a}_i \).

\[
\Lambda = \sum_{i=1}^{120} \prod_{i} \tilde{a}_i \xrightarrow{\text{collective}} \text{highly collective}
\]

\[
\left\{ M_i = \Lambda^{-\frac{1}{2}} \prod_{i} \tilde{a}_i \Lambda^{-\frac{1}{2}} \right\} \text{ defines a highly collective measurement}
\]
What is the $\text{Jacc}$ of the ensemble $E = \{s_i, f_i\}$ in the above example?

That of the $\bigcirc$ time step $\approx 0.585$

$\text{Jacc}$ for the current ensemble should be even lower.
Now, classical capacity of a quantum channel.

Basic use:

\[
\begin{align*}
\text{any } f & \xrightarrow{N} N(f) \\
\text{Alice} & \text{ Bob}
\end{align*}
\]

This is like our "Q" box:

\[
\begin{align*}
\text{"f"} & \xrightarrow{Q_N} N(f)
\end{align*}
\]

Alice can also use \(N^{\text{Nor}}\): Corresponding Q box:

\[
\begin{align*}
\text{\(f\)} & \xrightarrow{N} N^{\text{Nor}}(f) \\
\text{\(f\)} & \xrightarrow{Q_{N^{\text{Nor}}}} N^{\text{Nor}}(f)
\end{align*}
\]
It follows from the capacity discussion of the Do box that

\[ C(N) \geq \sup_r \frac{1}{r} X(N^r) \]

and

\[ \leq \max \chi (\{p_x, N(p_x)\}) \]

\[ \uparrow \]

Input to n channels.

This is called the HSW theorem, after Holmgren, Schumacher & Westmoreland.

Protocol 1:

\[ x_1 \xrightarrow{Q} f(x_1) \xrightarrow{y_1} \]
\[ x_2 \xrightarrow{Q} f(x_2) \xrightarrow{y_2} \]
\[ x_n \xrightarrow{Q} f(x_n) \xrightarrow{y_n} \]

Alice chooses \( p(x) \) to maximize \( \text{I}(x; y) \) (the channel capacity of \( N_0 \)).

Protocol 2: Shannon's direct coding method for \( N_0 \)
Protocol 2:

\[ x_1 \rightarrow Q \rightarrow f_{x_1} \rightarrow y \]
\[ x_2 \rightarrow Q \rightarrow f_{x_2} \]
\[ x_n \rightarrow Q \rightarrow f_{x_n} \]

Alice chooses \( p(x) \) to max \( X \{ fp(x), f x_3 \} \) and this is the capacity CQA.

Protocol: codewords randomly chosen from strongly typical sequences.
Protocol 3: \[ p_{x_1} \xrightarrow{Z} N^{\otimes n}(f_{x_1}) \]

\[ p_{x_2} \xrightarrow{Z} N^{\otimes n}(f_{x_2}) \]

\[ \vdots \]

\[ p_{x_m} \xrightarrow{Z} N^{\otimes n}(f_{x_m}) \]

Max over \( \{ p_x, f_x \} \) \& use input, use output

Capacity: \( C(N) = \sup_{\mu \in \chi} \frac{1}{n} \chi(f_x, N^{\otimes n}(f_x)) \)
Note that the capacity expression has an optimization over — called a “regularized” expression in contrast to the capacity expression for the classical channel (or the Q-box which involve only 1 copy of the resource (they are called single lettered expression).

The step $S(B_1, \ldots, B_n) = S(B_1) + S(B_2) + \cdots + S(B_n)$ in the converse of capacity of Q-box becomes true when only 1 state $\mathcal{N}_{\psi^+}$ is entangled in the channel setting.
To do:

- Some examples
- Additivity results vs non-additivity (this really says it so)
- Brief discussion on the non-additivity graph

Problems like not having a successful add to capacity of a DP channel