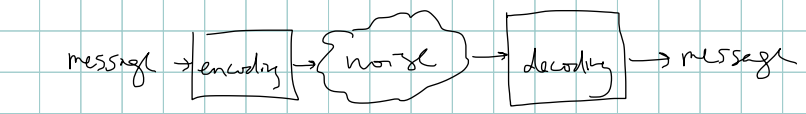
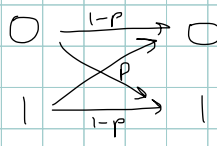
Channel capacities

How efficiently can A send a message to B over a noisy channel?
Consider the classical case first.



how many bits do we need to send so we can decode reliably?
send n bits to convey k bits $\Rightarrow$ rate $\frac{k}{n} = R$

Ex: binary symmetric channel



i.e. $\Pr(y|x) = \begin{cases} 1-p & x=y \\ p & x=\bar{y} \end{cases}$

we should choose input codewords that can be reliably decoded after going through the channel

with n input bits, typically pn bits are flipped
# of strings with $p \cdot n$ bits flipped is about $\frac{1}{p^{pn}(1-p)^{(1-p)n}} \approx 2^{nH_2(p)}$

$\Rightarrow$ the most likely outputs are in a "Hamming ball" of "Hamming radius" $\approx p \cdot n$, with volume $\approx 2^{nH_2(p)}$

to convey nR bits, need $2^{nR} \leq \frac{2^n}{2^{nH_2(p)}} \Rightarrow R \leq 1 - H_2(p)$

in fact this is achievable
to prove this, use a random code (useful for proving bounds on R, though not efficient)
choose 2^{nR} codewords uniformly at random from $\{0,1\}^n$
send through channel
to decode, consider a Hamming sphere of volume $2^{n(H_2(p)+\delta)}$
probability of another codeword falling in the same sphere $\leq \frac{2^{n(H_2(p)+\delta)}}{2^n} \cdot 2^{nR}$

$= 2^{-n[1-H_2(p)-R-\delta]}$

provided $R < 1 - H_2(p)$ (the capacity), we can take δ small and still have a small probability of failure. $\mathbb{E}_x[\Pr(\text{decoding error for } x)] < \varepsilon$ for any $\varepsilon > 0$

by Markov's inequality ($\Pr(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$), fraction of codewords w/ error probability $\geq 2\varepsilon$ is at most $\frac{1}{2}$
throw out the half of the codewords with highest prob. of decoding error. rate $R - \frac{1}{n} \to R$ as $n \to \infty$

Shannon's noisy coding theorem (for the binary symmetric channel):

- If $R < 1 - H_2(p)$, there is a scheme of rate R with $\Pr(\text{correct transmission}) \to 1$ as $n \to \infty$
- If $R > 1 - H_2(p)$, any rate-R scheme has $\Pr(\text{correct transmission}) \to 0$ as $n \to \infty$

What about other channels?

In general, describe a channel by a distribution $\Pr(y \mid x)$ (x: codeword; y: output of channel)

Choose x from a distribution with entropy $H(X) \Rightarrow 2^{nH(X)}$ typical strings

For a typical received value of y, # possible messages sent:

$$2^{nH(X|Y)} \quad \text{where} \quad H(X|Y) = \sum_y \Pr(y)\, H(X|Y=y)$$

($H(X|Y)$: # of additional bits needed to specify x once y is known)

$$= \sum_{x,y} \Pr(y)\, \Pr(x|y) \log \frac{1}{\Pr(x|y)}$$

SKIP

So with a random code,

$$\Pr(\text{decoding error}) \le \frac{2^{n(H(X|Y)+\delta)}}{2^{nH(X)}} 2^{nR} \Rightarrow R \ge \underbrace{H(X) - H(X|Y)}_{= I(X;Y)} - \delta$$

the mutual information between X and Y

note: $I(X;Y) = H(Y) - H(Y|X)$

$= H(X) + H(Y) - H(X,Y)$

Can also show this is the best possible rate

A scheme of rate R takes one of 2^{nR} messages and encodes them into n bits. After the channel acts, we apply a decoding procedure to recover the message. We say the scheme is reliable if its probability of correct decoding goes to 1 as $n \to \infty$.

Def: The mutual information between random variables X & Y is

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$$

$$= H(X) - \sum_y \Pr(y)\, H(X|Y=y) = H(X) + H(Y) - H(X,Y)$$

$$= H(X) - \sum_{x,y} \Pr(y)\, \Pr(x|y) \log \frac{1}{\Pr(x|y)}$$

This measures how much info X & Y have in common (and $H(X|Y)$ measures the uncertainty about X when we know Y)

Shannon's noisy channel coding theorem

Let $C = \max_X I(X;Y)$ where Y is the distribution of channel outputs when the input is distributed according to X (explicitly, $\Pr(y) = \sum_x \Pr(x)\Pr(y|x)$).

If $R < C$ then $\exists$ a reliable scheme of rate R.
If $R > C$ then no scheme of rate R is reliable.
C is the capacity of the channel.

Can prove this using random coding.

Quantum capacities

What if we have a quantum channel $\mathcal{N}$?
Suppose we want to send classical info.

Holevo quantity: for an ensemble where ρ_x occurs with probability p_x,
$\chi(\{p_x, \rho_x\}) = S(\rho) - \sum_x p_x S(\rho_x)$

Holevo bound: Suppose we measure a POVM with operators $\{E_y\}$
Then $I(X;Y) \leq \chi$
↑ how much we learn about X from the measurement outcome

Given a q. operation $\mathcal{N}$, define $\chi(\mathcal{N}) = \max_{\{p_x, \rho_x\}} \chi(\{p_x, \mathcal{N}(\rho_x)\})$

HSW Theorem (Holevo, Schumacher, Westmoreland):
If we are restricted to product-state inputs for the n channel uses (but can do collective decoding), then the capacity is $\chi(\mathcal{N})$.

More generally, the capacity is the regularized Holevo information:

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

Q: Do we ever need entangled inputs?
I.e., is the capacity additive?
No! (Hastings 2008) $\exists$ channels for which $\chi(\mathcal{N}^{\otimes 2}) > 2\,\chi(\mathcal{N})$.

We can also consider other capacities.

Ex: Capacity to send q. info

The <u>coherent information</u> is $I_c(\rho, \mathcal{N}) = S(\mathcal{N}(\rho)) - S(\overbrace{\mathrm{tr}_B(U\rho U^\dagger)}^{\text{state of environment}})$

$\mathcal{N}$: map from A to B

U: isometry from A to BE (environment)

The capacity to send q. states is exactly the optimal (regularized) coherent info: $\lim_{n\to\infty} \frac{1}{n} \max_\rho I_c(\rho, \mathcal{N}^{\otimes n})$

This has long been known to be superadditive.

Ex: Capacity to send classical info in the presence of free shared entanglement:

$$\max_\rho \left[S(\rho) + S(\mathcal{N}(\rho)) - S((I \otimes \mathcal{N})(|\psi\rangle\langle\psi|)) \right]$$

($|\psi\rangle$: purification of ρ)

this is additive! no regularization needed