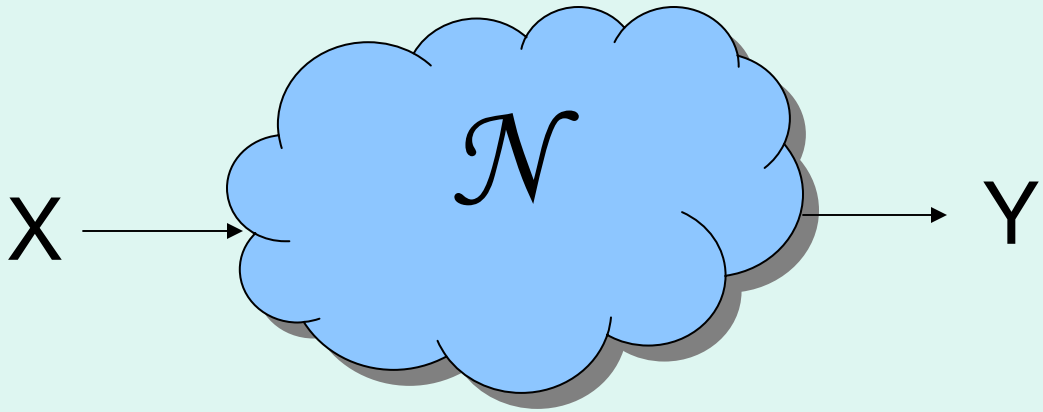
Quantum Channels and Their Capacities: Lectures 3+4

The many capacities of a quantum channel

Channel Capacity



p(y|x)

Capacity: bits per channel use in the limit of many channels

$$C = \max_X I(X;Y)$$

I(X;Y) is the mutual information

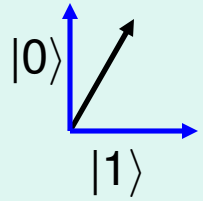
Overview

- Quantum States and Channels
- Definitions of Capacities and statements of Coding Theorems
- Sketch of Quantum Coding Theorem
- Introduction to Additivity

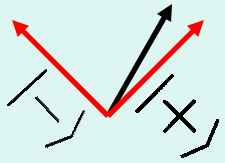
Pure Quantum States

- Qubit: $|\psi\rangle = \alpha\,|0\rangle + \beta|1\rangle$, α,β complex and $|\alpha|^2 + |\beta|^2 = 1$.

- If you measure $|\psi\rangle$ in the $|0\rangle,|1\rangle$ basis, you get 0 with prob. $|\alpha|^2$ and 1 with prob. $|\beta|^2$



- You could use some other basis, though. Like $|+\rangle,|-\rangle$, with $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$



For a d-level system $|\psi\rangle$ is a unit vector in C^d

Mixed Quantum States

- Pure states are the minimum uncertainty states in quantum mechanics.
- We can also have mixed states:
  $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ with $p_i \geq 0$ and $\sum_i p_i = 1$
- Can think of it as a bipartite pure state with one part traced out: $\rho_B = Tr_A |\psi_{AB}\rangle\langle\psi_{AB}|$
- A pure whole can have mixed parts

Entropy and Typical Spaces

- Any $\rho_B = \mathrm{Tr}_A\, |\psi\rangle\langle\psi|_{AB}$
- $S(\rho_B) = -\mathrm{Tr}\, \rho_B \log \rho_B$ is the entropy
- It measures the uncertainty in B
- Given n copies of $|\psi\rangle_{AB}$, we can reversibly map B to a space of dimension $2^{n\, S(\rho_B)}$. This is the “typical space”.
- Analogous to “typical sets” of classical information theory. $2^{nH(p)}$ strings

Properties of von Neumann entropy

$$\text{Definition: } S(\rho) = -\text{Tr}\rho \log \rho$$

Properties of von Neumann entropy

Definition: $S(\rho) = -\mathrm{Tr}\rho\log\rho$

$S(\rho) \geq 0$, 0 iff ρ is pure. $\quad S(\rho) \leq \log D$

Concave: $S(p_1\rho_1 + p_2\rho_2) \geq p_1 S(\rho_1) + p_2 S(\rho_2)$

Properties of von Neumann entropy

$$\text{Definition: } S(\rho) = -\text{Tr}\rho\log\rho$$

$S(\rho) \geq 0$, 0 iff ρ is pure. $\qquad S(\rho) \leq \log D$

$$\text{Concave: } S(p_1\rho_1 + p_2\rho_2) \geq p_1 S(\rho_1) + p_2 S(\rho_2)$$

$$\text{Continuous: } |S(\sigma) - S(\rho)| \leq \epsilon \log d + \epsilon \log \tfrac{1}{\epsilon}$$

$$\text{for } ||\sigma - \rho||_1 = \epsilon \leq 1/e.$$

Properties of von Neumann entropy

Definition: $S(\rho) = -\mathrm{Tr}\rho\log\rho$

$S(\rho) \geq 0$, 0 iff ρ is pure. $\quad S(\rho) \leq \log D$

Concave: $S(p_1\rho_1 + p_2\rho_2) \geq p_1 S(\rho_1) + p_2 S(\rho_2)$

Continuous: $|S(\sigma) - S(\rho)| \leq \epsilon \log d + \epsilon \log \frac{1}{\epsilon}$

for $||\sigma - \rho||_1 = \epsilon \leq 1/e$.

Subadditivity: $S(AB) \leq S(A) + S(B)$

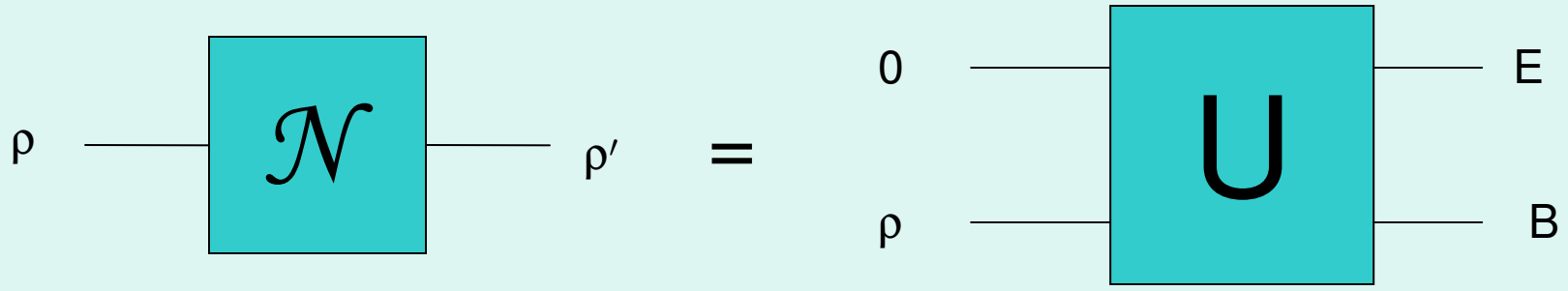
Strong Subadditivity:

$S(ABC) + S(C) \leq S(AC) + S(BC)$

Noisy Quantum Channels

- Noiseless quantum evolution: $\rho \rightarrow U\rho U^{\dagger}$

  Unitary satisfies $U^{\dagger}U = I$
- Noisy quantum evolution: unitary interaction with inaccessible environment



$$\rho \rightarrow \mathrm{Tr}_E\, U(\, \rho \otimes |0\rangle\langle 0|\,) U^{\dagger}$$

Overview

- ~~Quantum States and Channels~~
- Definitions of Capacities and statements of Coding Theorems
- Sketch of Quantum Coding Theorem
- Introduction to Additivity

Classical Capacity of Quantum Channel

We can understand coding schemes for classical information in terms of the Holevo Information:
$\chi(\mathcal{N}) = \max_{\{p_x, \rho_x\}} I(X;B)$
where $I(X;B) = H(X) + H(B) - H(XB)$ uses von Neumann entropy and is evaluated on the state $\sum_x p_x |x\rangle\langle x| \otimes \mathcal{N}(\rho_x)$

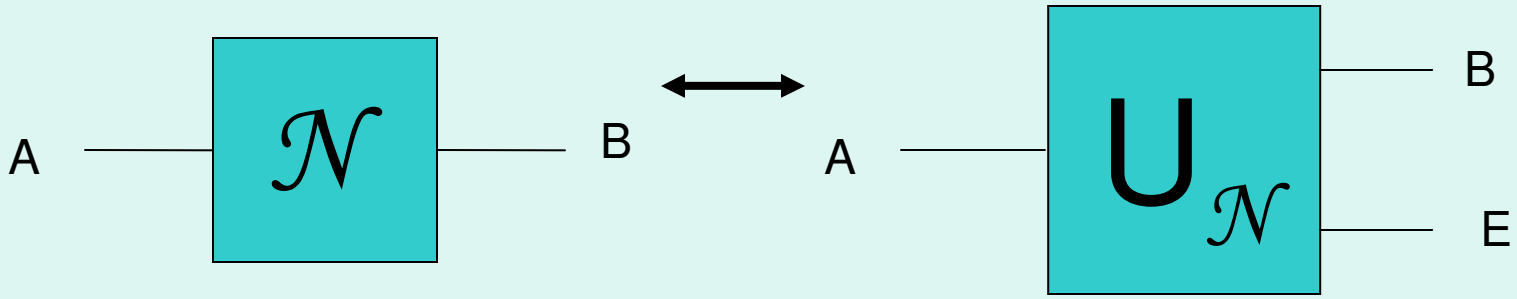
Random coding arguments show that $\chi(\mathcal{N})$ is an achievable rate, so $C(\mathcal{N}) \geq \chi(\mathcal{N})$. Furthermore,

n uses

$$C(\mathcal{N}) = \lim_{n \to \infty} (1/n)\, \chi(\mathcal{N} \otimes \ldots \otimes \mathcal{N})$$

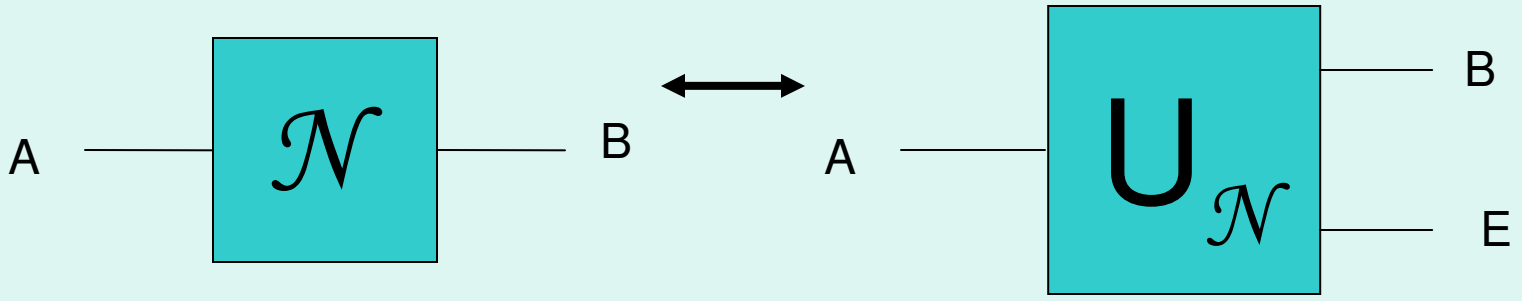
(see Holevo 73, 79, 98, Schumacher-Westmoreland 97)

Private Classical Capacity

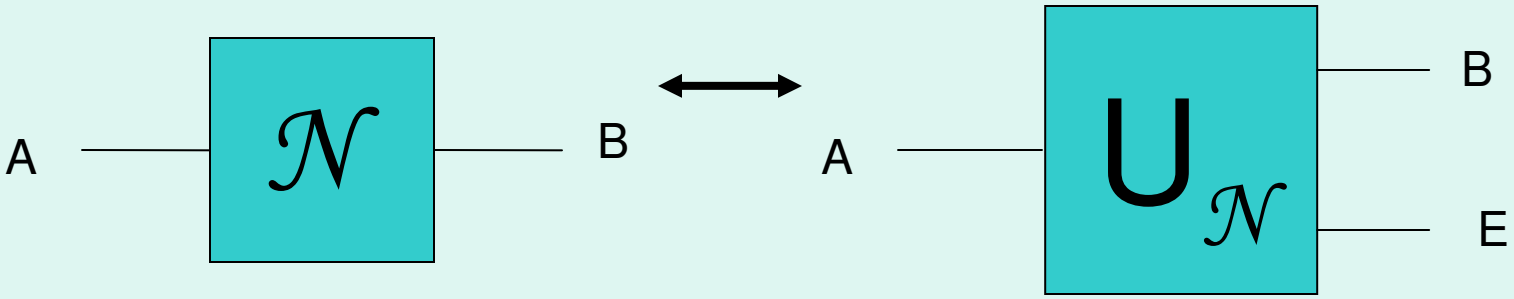


- Quantum channel has one sender, two receivers.
- Best rate for classical messages from A to B while E learns nothing = private capacity. Call it P($\mathcal{N}$).
- Related to quantum key distribution---the fact than by analysing the map from A to B we can infer the map from A to E allows unconditional security that is impossible classically.

Private Classical Capacity



Private Classical Capacity



- Let $P^1(\mathcal{N}) = \max_{p_v,\phi_v} I(V;B) - I(V;E)$, with mutual informations evaluated on $\sum_v p_v |v\rangle\langle v| \otimes U\phi_v U^\dagger$

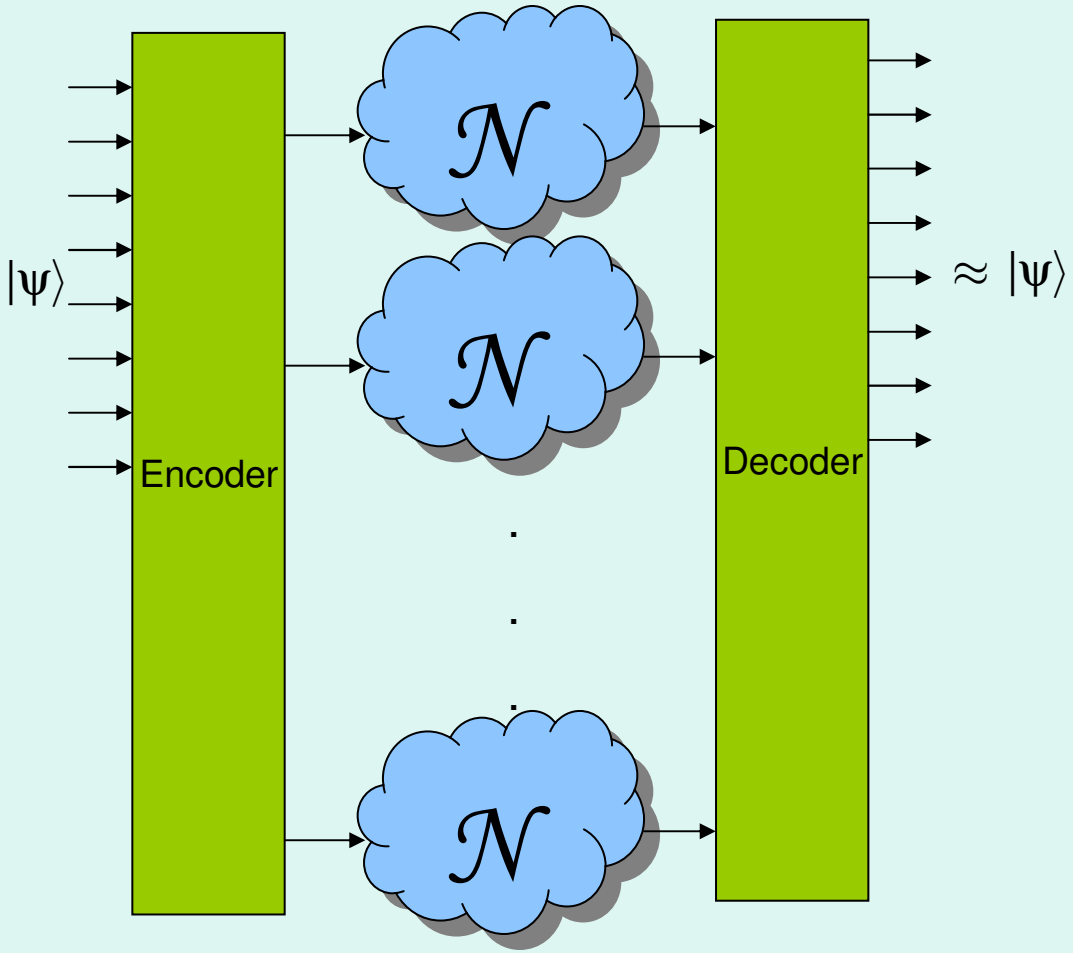
- Random coding and privacy amplification shows $P(\mathcal{N}) \geq P^1(\mathcal{N})$ and, in fact we can get

$$P(\mathcal{N}) = \lim_{n\to\infty} (1/n)\, P^1(\mathcal{N} \otimes \ldots \otimes \mathcal{N})$$

n uses

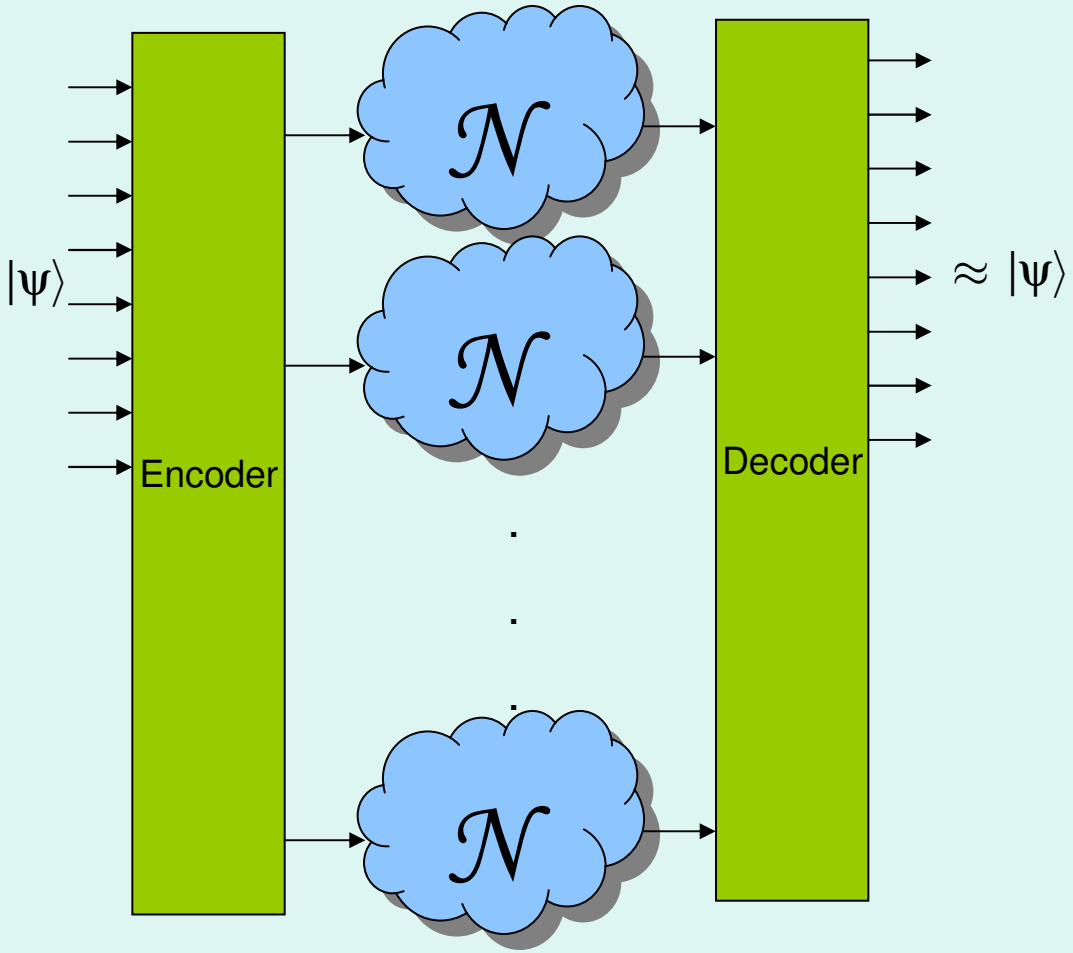
See Devetak 03

Quantum Capacity

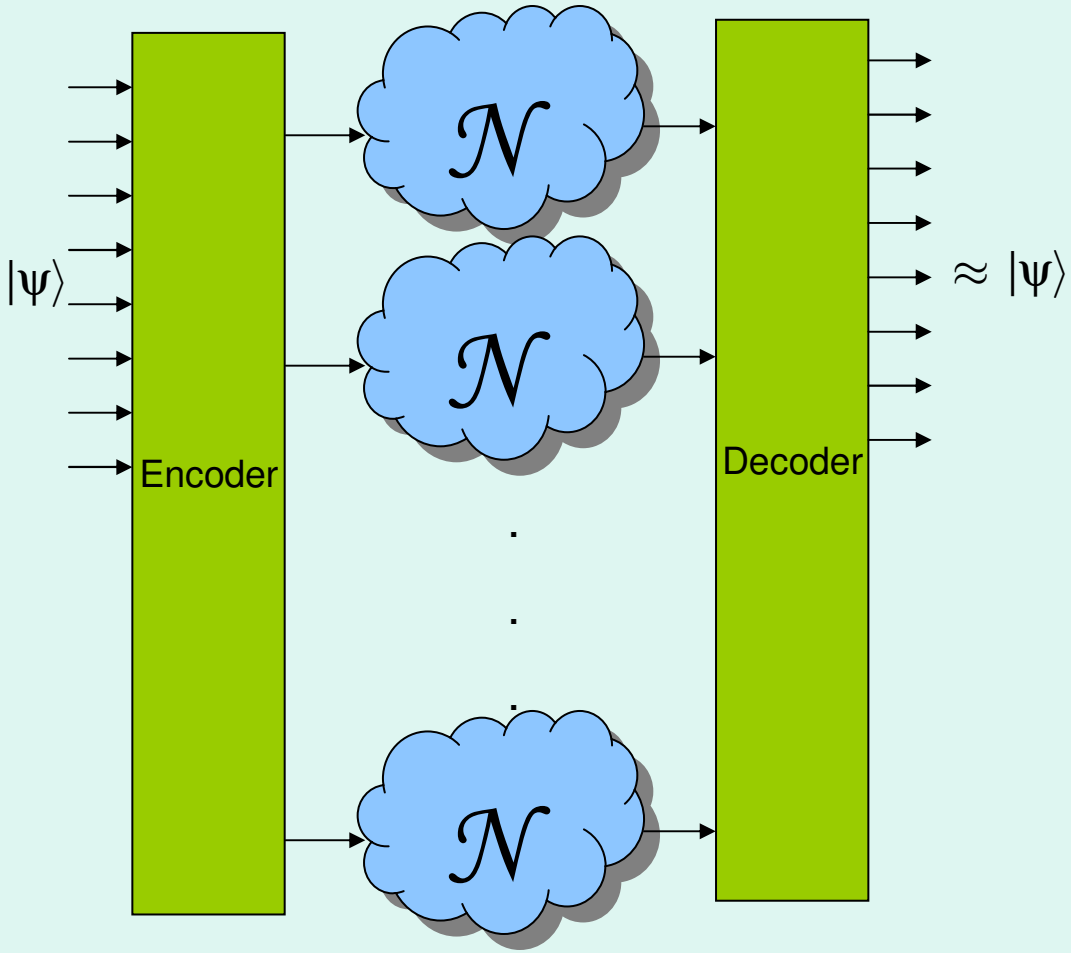


- If we try to transmit an arbitrary quantum state, we arrive at the quantum capacity, $Q(\mathcal{N})$.
- The quantum capacity, measured in qubits per channel use, characterizes the ultimate limit on quantum error correction.

Quantum Capacity



Quantum Capacity



Define the coherent information:

$$Q^1(\mathcal{N}) = \max_\phi H(B) - H(E),$$

with entropies evaluated on $U\phi U^\dagger$. Then, we can show that $Q^1(\mathcal{N})$ is an achievable rate for quantum communication, so

$$Q(\mathcal{N}) \geq Q^1(\mathcal{N})$$

Furthermore,

$$Q(\mathcal{N}) = \lim_{n\to\infty}(1/n)Q^1(\mathcal{N}\otimes \ldots\otimes\mathcal{N})$$

See Lloyd 97, Shor 02, Devetak 03

Coherent Information and no-cloning

- **No cloning:** there is no physical operation that copies an unknown quantum state.
- Basically, because $|\psi\rangle \rightarrow |\psi\rangle|\psi\rangle$ isn't linear

S(B) is how much information B has

S(E) is how much information E has

$Q^1 = S(B) - S(E)$ is how much more Bob knows than Eve.

$\approx$ how much secret information we can send to Bob

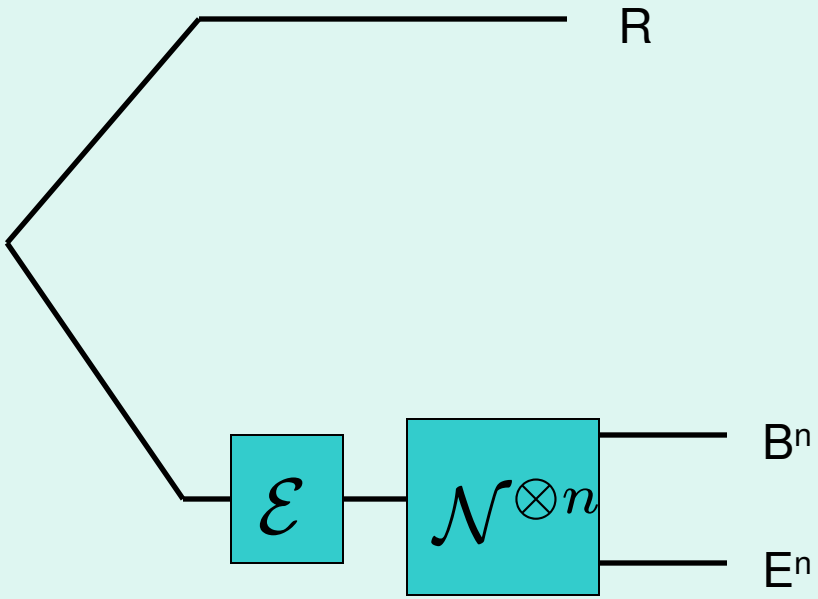
Overview

- Quantum States and Channels
- Definitions of Capacities and statements of Coding Theorems
- Sketch of Quantum Coding Theorem
- Introduction to Additivity

Sketch of Achievability of Coherent Information

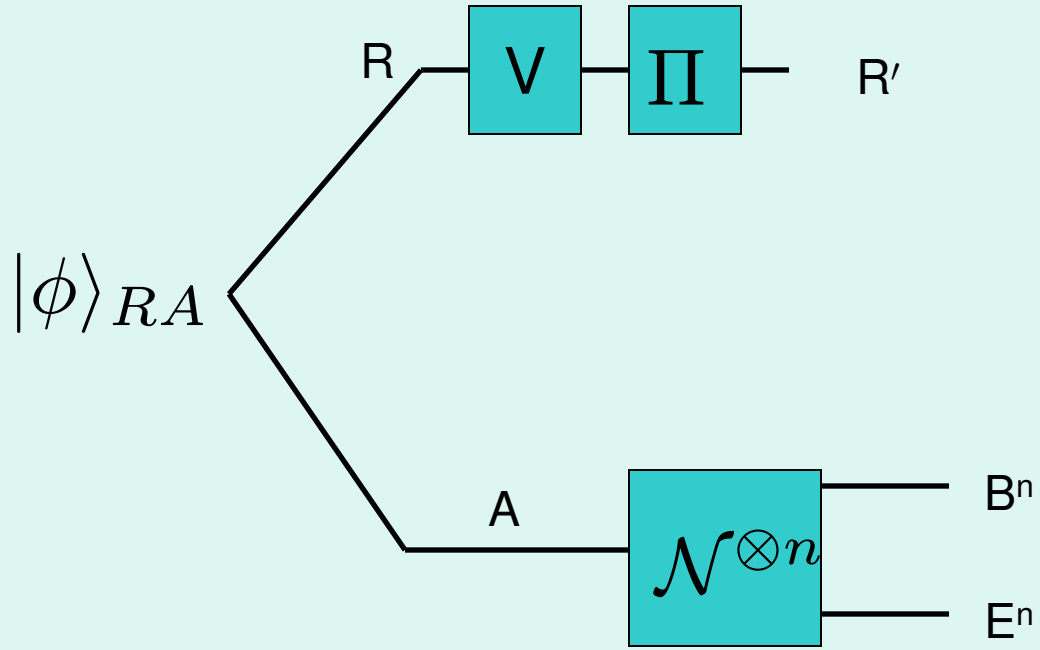
- Step 1: If you can transmit half of a maximally entangled state reliably, then you can transmit a quantum state.
- Step 2: If you can decouple your reference system from the environment, then you have a pure entangled state.

Decoupling



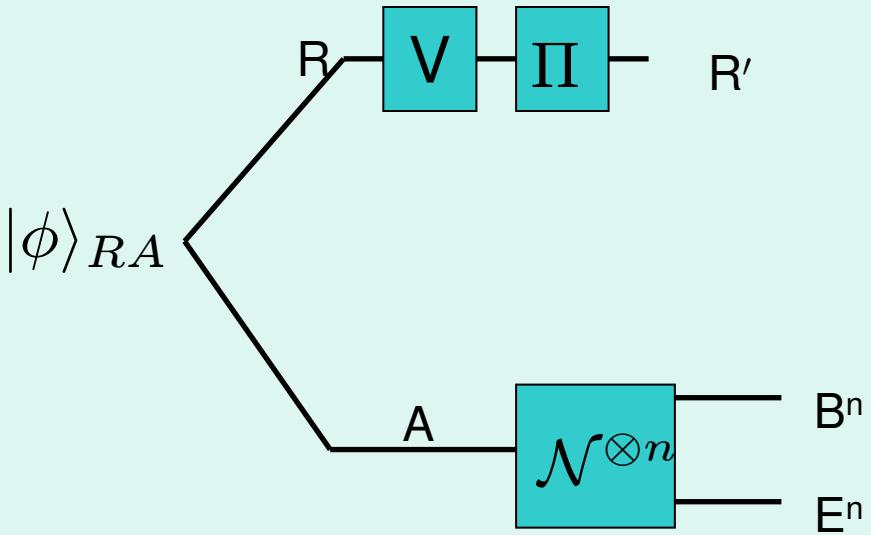
If $\rho_{RE^n} \approx \rho_R \otimes \rho_{E^n}$, then $\rho_{RB_1B_2E^n} = |\varphi\rangle_{RB_1} |\psi\rangle_{B_2E^n}$

Decoupling with Random encoding



$$|\phi\rangle_{RA} = \frac{1}{\sqrt{|R|}} \sum_{i=1}^{|R|} |i\rangle|i\rangle$$

Decoupling with Random encoding



Choose V randomly.

Consider the state this circuit generates on R'Eⁿ.

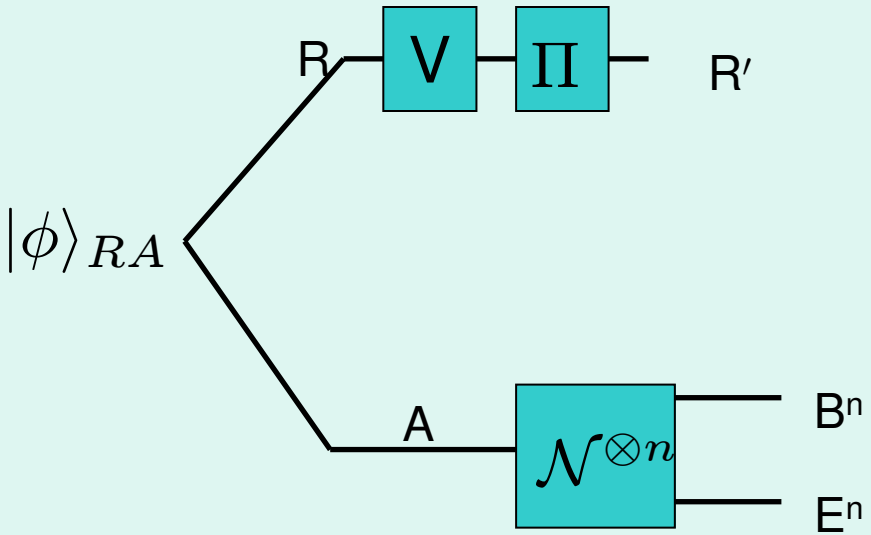
$$\left(\int dV ||\sigma_{R'E^n}(V) - \sigma_{R'}^{\max} \otimes \sigma_{E^n}||_1\right)^2 \leq |R'E^n|(\mathrm{Tr}(\sigma_{RE^n}))^2$$

Estimate: |R'| = $2^{n(\text{rate})}$

$|E^n| \approx 2^{n\, S(E)}$

Spectrum of RE is same as spectrum of B. When iid, this is maximally mixed with dimension $2^{nS(B)}$. This gives $(\mathrm{Tr}(\sigma_{RE^n}))^2 \approx 2^{nS(B)}$

Decoupling with Random encoding



So, as long as rate < S(B) – S(E), the deviation from a product state between R' and E^n becomes arbitrarily small. Which enables transmission.

Choose V randomly.

Consider the state this circuit generates on R'E^n.

$$\left(\int dV ||\sigma_{R'E^n}(V) - \sigma_{R'}^{\max} \otimes \sigma_{E^n}||_1\right)^2 \leq |R'E^n|(\mathrm{Tr}(\sigma_{RE^n}))^2$$

Estimate: $|R'| = 2^{n(\text{rate})}$

$|E^n| \approx 2^{n\,S(E)}$

Spectrum of RE is same as spectrum of B. When iid, this is maximally mixed with dimension $2^{nS(B)}$. This gives $(\mathrm{Tr}(\sigma_{RE^n}))^2 \approx 2^{nS(B)}$

Overview

- Quantum States and Channels
- Definitions of Capacities and statements of Coding Theorems
- Sketch of Quantum Coding Theorem
- Introduction to Additivity

Additivity: definition and motivation

- A function on channels is called additive if $f(\mathcal{N} \otimes \mathcal{M}) = f(\mathcal{N}) + f(\mathcal{M})$
- Recall that $Q(\mathcal{N}) = \lim_{n \to \infty}(1/n)Q^1(\mathcal{N}^{\otimes n})$. If we could show that Q^1 was additive, we'd have $Q(\mathcal{N}) = Q^1(\mathcal{N})$.
- Similarly, $C(N) = \lim_{n \to \infty}(1/n)\, \chi(\mathcal{N}^{\otimes n})$ and $P(\mathcal{N}) = \lim_{n \to \infty}(1/n)P^1(\mathcal{N}^{\otimes n})$, so if χ and P^1 were additive, we'd have single-letter capacities for classical and private communication.

Most things aren't additive

- Q^1 is not additive for the very noisy depolarizing channel (Shor-Smolin '96)
- P^1 isn't additive for BB84 channel (Smith-Renes-Smolin, '08)
- χ is nonadditive for high-dimensional random channel (Hastings '09)
- Q^1 and P^1 can both be extremely nonadditive (Smith-Smolin 08, 09)

But sometimes they are

- χ is additive for depolarizing, erasure, and entanglement breaking channels.
- Q^1 and P^1 are additive for degradable channels*, Q^1 is for PPT channels.

* Just like a degraded broadcast channel when you take the less noisy user to be the channel output and the more noisy user to be the environment

See King, Shor, Devetak-Shor, Horodecki, …

A different kind of (non)additivity

Already saw that Q^1 wasn't additive, but what about $Q(\mathcal{N}) = \lim_{n\to\infty}(1/n)Q^1(\mathcal{N}^{\otimes n})$?

Since $Q(\mathcal{N}\otimes\mathcal{N}) = 2\,Q(\mathcal{N})$, this is actually a question about how different channels interact: Can $Q(\mathcal{N}\otimes\mathcal{M}) > Q(\mathcal{N}) + Q(\mathcal{M})$?

Yes

A different kind of (non)additivity

The only channels with zero classical capacity have no correlation between input and output. However, quantum information is more delicate, so there are nontrivial quantum channels with $Q(\mathcal{N}) = 0$. A good example is the 50% quantum erasure channel ($\rho \rightarrow \frac{1}{2}\rho + \frac{1}{2}|e\rangle\langle e|$).

There's a more complicated kind of channel with $Q(\mathcal{M}) = 0$, called a private PPT channel. These have $P(\mathcal{M}) > 0$.

You can show that for any such PPT channel, $Q(\mathcal{N} \otimes \mathcal{M}) \geq \frac{1}{2} P(\mathcal{M}) > 0$, so in the end, we get

$Q(\mathcal{N}) = 0$ and $Q(\mathcal{M}) = 0$, but $Q(\mathcal{N} \otimes \mathcal{M}) > 0$.

This is for two qubit channels, but with larger channels you can make the additivity violation very large (1/8 log d). Get similar nonadditivity for the private classical capacity.

Additivity Questions

Information \ Quantity	Capacity	Correlation Measure
Classical	Classical Capacity **?**	Holevo Information χ = max I(X;B) **No** (Hastings '09)
Private	Private Capacity **No** (Li-Winter-Zou-Guo '09 Smith-Smolin-08/09)	Private Information max I(X;B)-I(X;E) **No** (Smith-Renes-Smolin '08)
Quantum	Quantum Capacity **No** (Smith-Yard '08)	Coherent Information max S(B)-S(E) **No** (Div-Shor-Smolin '98)
Entanglement assisted	Entanglement assisted classical capacity **Yes** (Bennett-Shor-Smolin-Thapliyal '99)	Quantum Mutual Information **Yes** (Bennett-Shor-Smolin-Thapliyal '99)

Summary

- Quantum channels are unitary interactions of an input with an inaccessible environment
- They have many capacities: classical, private, quantum (and more!)
- Each type of information has a coding theorem, which tells you an acheivable communication rate (“direct”).
- These involve random coding (e.g., quantum).
- Additivity relates to finding upper bounds and computable characterizations of quantum capacities. Usually things aren’t additive (and this is cool!).

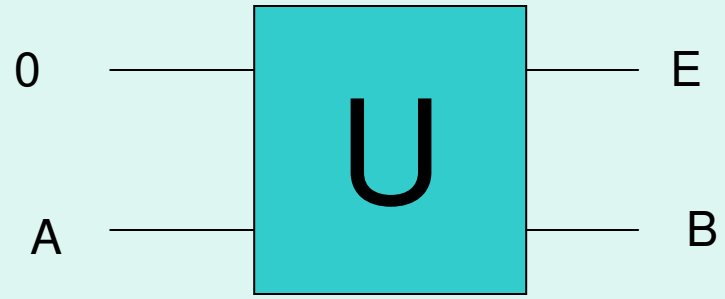
Where to learn more

- Decoupling proof of quantum channel capacity: Hayden, Horodecki, Winter, Yard ’07
- Strong Subadditivity: Ed Effros, PNAS 08

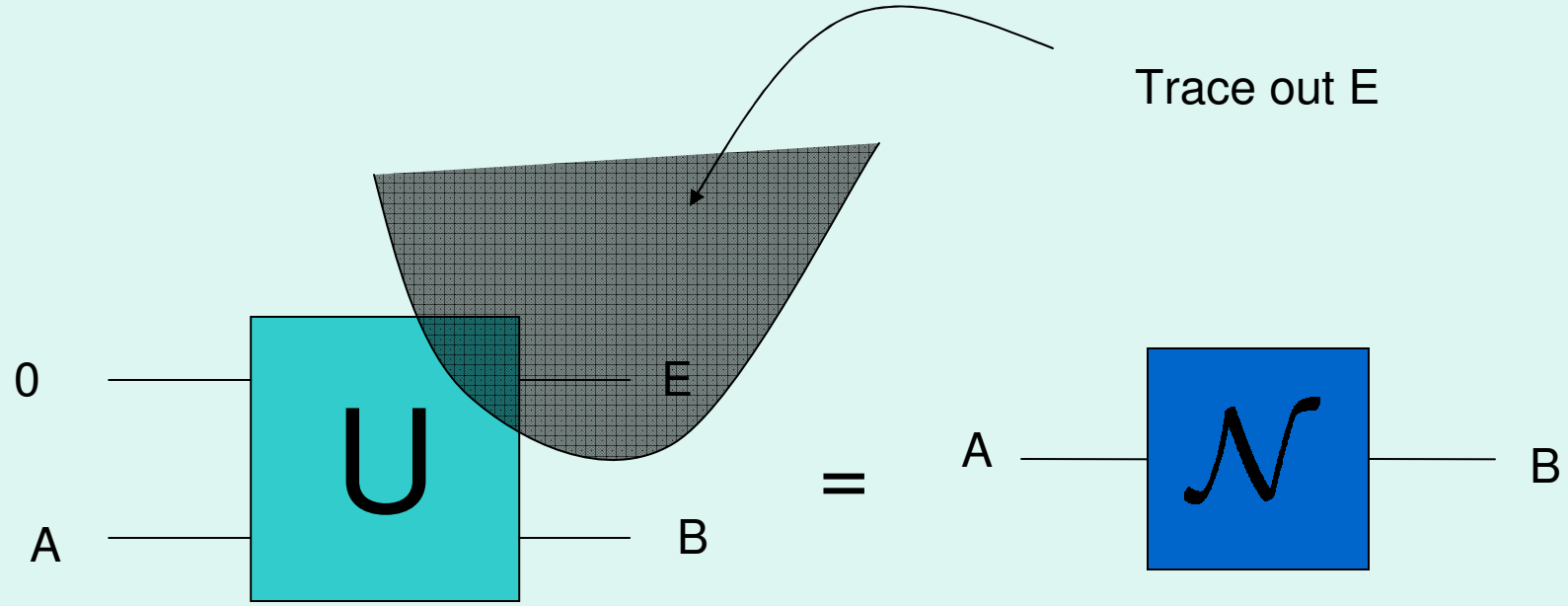
- I am here till Wednesday morning. Come talk to me!

Bonus Material

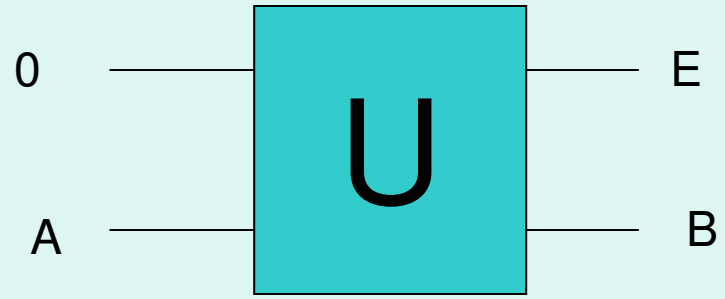
Degradable Channels



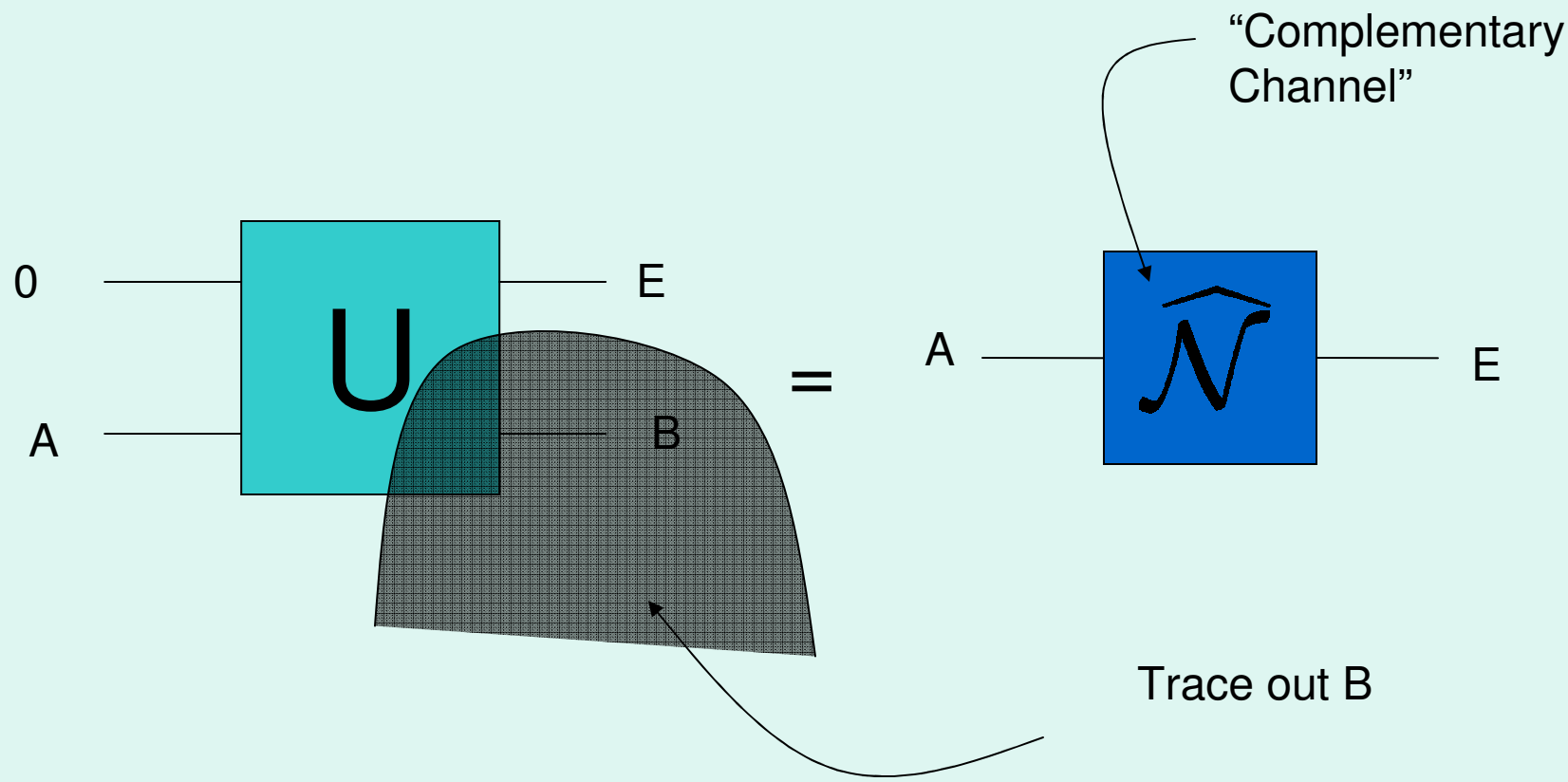
Degradable Channels



Degradable Channels

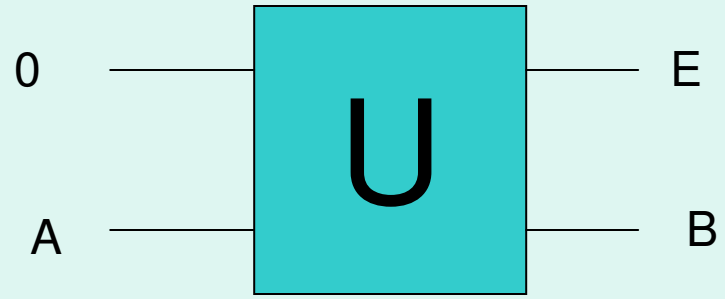


Degradable Channels



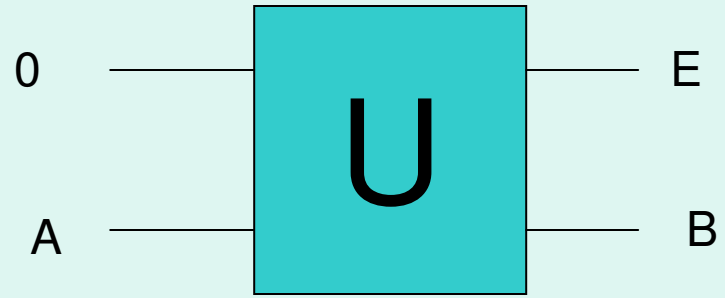
Degradable Channels

Bob can simulate Eve



Degradable Channels

Bob can simulate Eve



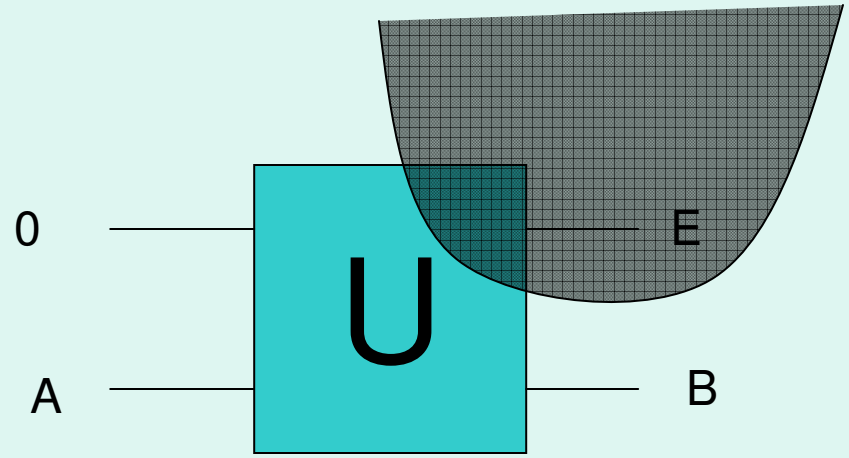
degradable if $\widehat{\mathcal{N}}$ is noisier than $\mathcal{N}$:

$$\widehat{\mathcal{N}} = \mathcal{D} \circ \mathcal{N}$$

For some “degrading channel” $\mathcal{D}$

Degradable Channels

Bob can simulate Eve



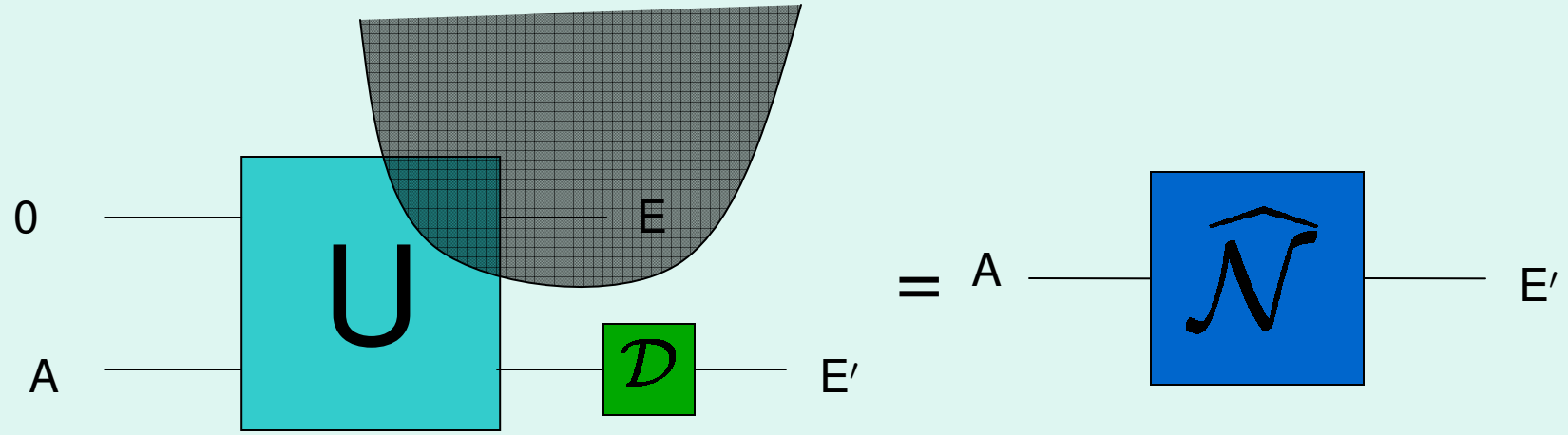
degradable if $\widehat{\mathcal{N}}$ is noisier than $\mathcal{N}$:

$$\widehat{\mathcal{N}} = \mathcal{D} \circ \mathcal{N}$$

For some “degrading channel” $\mathcal{D}$

Degradable Channels

Bob can simulate Eve



degradable if $\widehat{\mathcal{N}}$ is noisier than $\mathcal{N}$:

$$\widehat{\mathcal{N}} = \mathcal{D} \circ \mathcal{N}$$

For some “degrading channel” $\mathcal{D}$

Degradable Channels: Erasure Channel

- Erases with probability p:

$$\mathcal{E}_p(\rho) = (1-p)\rho + p|e\rangle\langle e|$$

Orthogonal erasure flag

Degradable Channels: Erasure Channel

- Erases with probability p:

Orthogonal erasure flag

$$\mathcal{E}_p(\rho) = (1-p)\rho + p|e\rangle\langle e|$$

- Envnment gets the state when its erased:

$$\hat{\mathcal{E}}_p(\rho) = p\rho + (1-p)|e\rangle\langle e|$$

Degradable Channels: Erasure Channel

- Erases with probability p:

Orthogonal erasure flag

$$\mathcal{E}_p(\rho) = (1-p)\rho + p|e\rangle\langle e|$$

- Envnment gets the state when its erased:

$$\widehat{\mathcal{E}}_p(\rho) = p\rho + (1-p)|e\rangle\langle e|$$

- As long as p $\leq$ ½ , we can degrade:

$$\widehat{\mathcal{E}}_p = \mathcal{D}_p \circ \mathcal{E}_p$$

- $\mathcal{D}_p$ just throws away ρ with prob $\frac{1-2p}{1-p}$

Degradable Channels: Amplitude Damping

- Acts like $\mathcal{A}_\gamma(\rho) = A_0 \rho A_0^\dagger + A_1 \rho A_1^\dagger$ with $A_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix}$ and $A_1 = \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix}$. Models relaxation from excited to ground state.
- $\widehat{\mathcal{A}}_\gamma = \mathcal{A}_{1-\gamma}$ (Up to a unitary)
- So, if $\gamma \leq$ ½, can just damp more and simulate complementary channel

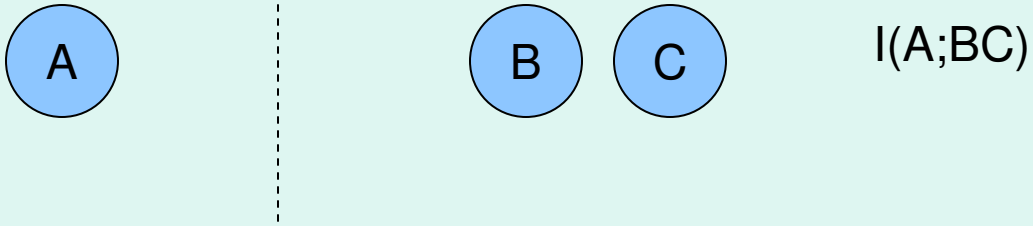
Degradable Channels: A few more

- Half of the qubit channels with two Kraus operators (the other half are reverse-degradable)
- Pure loss bosonic gaussian channel
- Channels whose complement is entanglement breaking (“Hadamard channels”)
- Unruh channel??

How to check if a channel is degradable

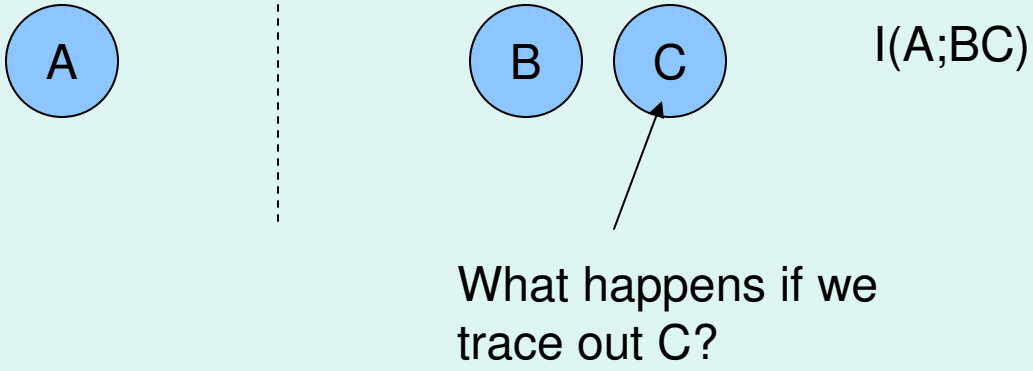
- Let's say we have a channel $\mathcal{N}$ with complementary channel $\widehat{\mathcal{N}}$. How do we know if it's degradable?
- Well, $\mathcal{N}$ is a linear map on density matrices, so it has an inverse, $\mathcal{N}^{-1}$. Generally, this is not a CPTP map.
- If there's a degrading map, $\mathcal{D}$, it'll have to be equal to $\widehat{\mathcal{N}} \circ \mathcal{N}^{-1}$, so check if this is CPTP.

Monotonicity of Mutual Information and Strong Subadditivity



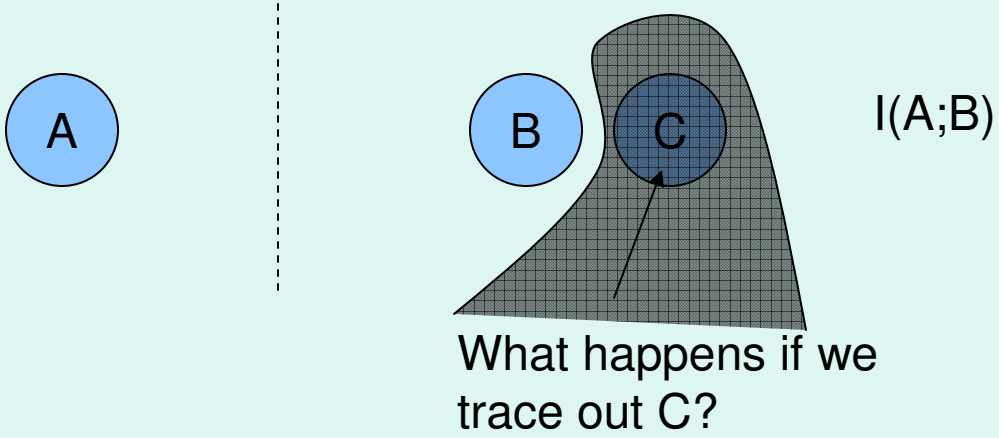
- Mutual Information:
  $I(A;B) = S(A) + S(B) - S(AB)$

Monotonicity of Mutual Information and Strong Subadditivity



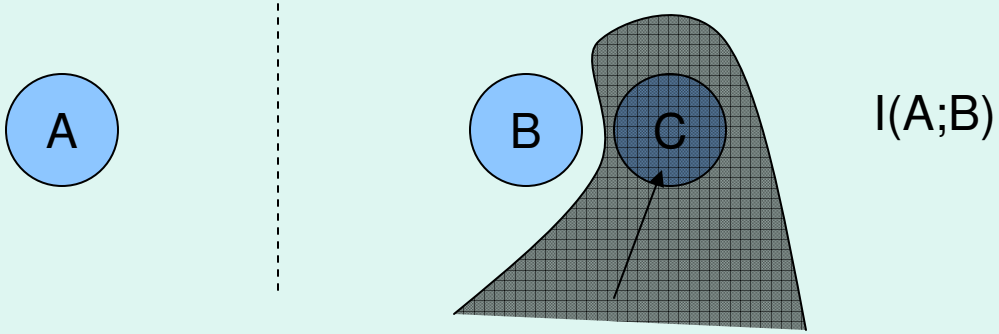
- Mutual Information:
  $I(A;B) = S(A) + S(B) - S(AB)$

Monotonicity of Mutual Information and Strong Subadditivity



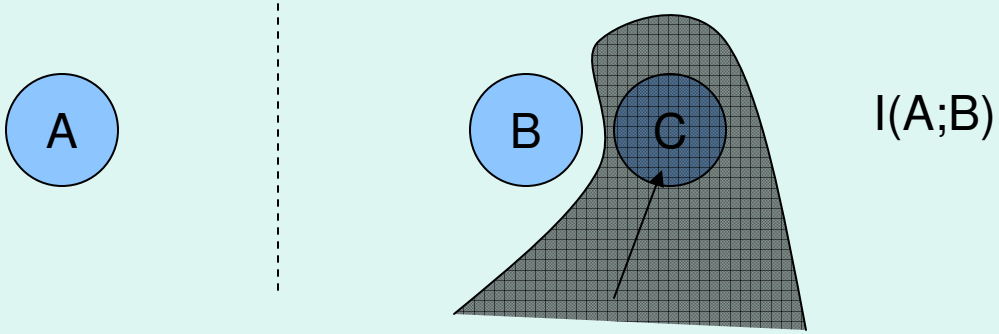
- Mutual Information:
  $I(A;B) = S(A) + S(B) - S(AB)$
- Monotonicity: $I(A;BC) \geq I(A;B)$

Monotonicity of Mutual Information and Strong Subadditivity



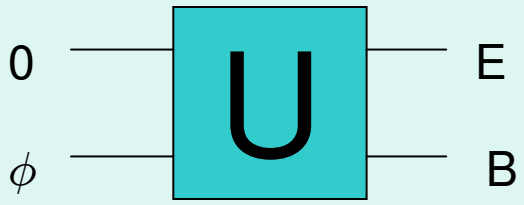
- Mutual Information:
  $I(A;B) = S(A) + S(B) - S(AB)$
- Monotonicity: $I(A;BC) \geq I(A;B)$
- Equivalently: $S(BC) + S(AB) \geq S(ABC) + S(B)$

Monotonicity of Mutual Information and Strong Subadditivity



- Mutual Information:
  $I(A;B) = S(A) + S(B) - S(AB)$
- Monotonicity: $I(A;BC) \geq I(A;B)$
- Equivalently: $S(BC) + S(AB) \geq S(ABC) + S(B)$
- This tells us mutual information can only decrease under local processing

Additivity of Coherent Information for degradable channels



Recall that $Q^1 = \max_\phi S(B) - S(E)$
ϕ is a mixed state on the input.

- The capacity is given by
  $Q(\mathcal{N}) = \lim_{n \to \infty}(1/n)\, Q^1(\mathcal{N} \otimes \ldots \otimes \mathcal{N})$
- If we could show $Q^1(\mathcal{N} \otimes \mathcal{N}) \leq 2Q^1(\mathcal{N})$ for any degradable channel, we'd have $Q(\mathcal{N}) = Q^1(\mathcal{N})$ for degradable channels

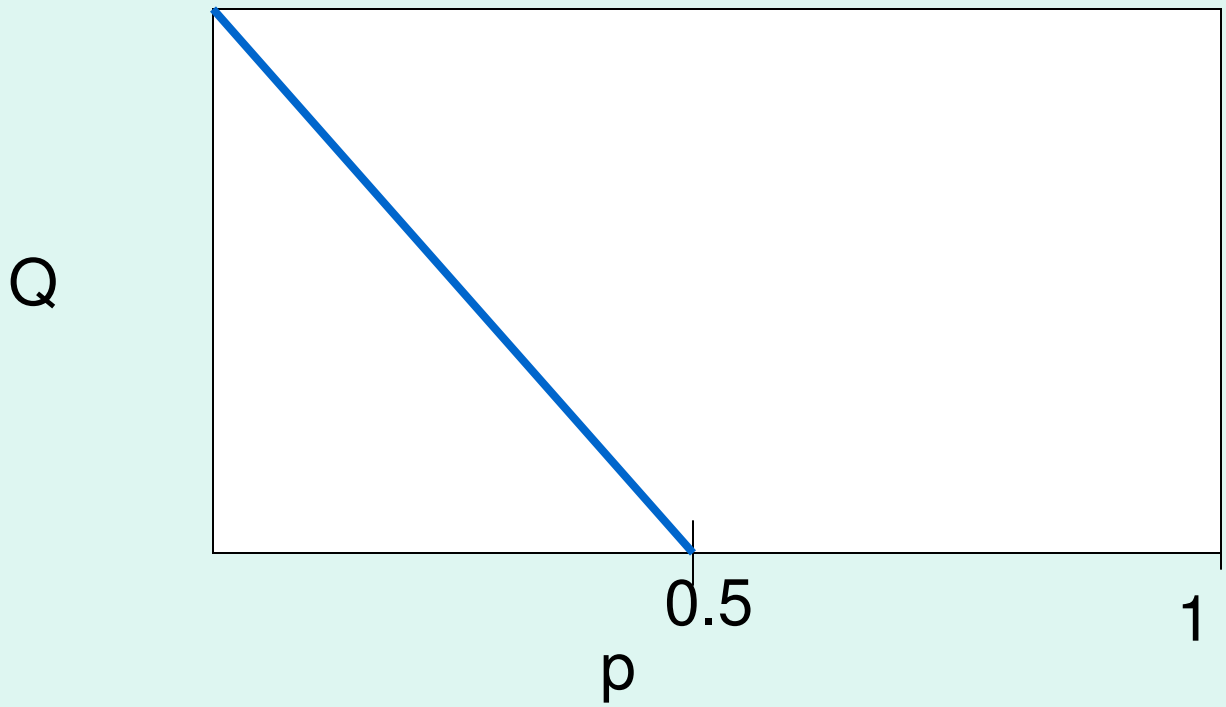
Additivity of Coherent Information for degradable channels

- Say we have ϕ_{12}, a mixed state on A_1A_2, the input of $\mathcal{N} \otimes \mathcal{N}$, with $Q^1(\mathcal{N} \otimes \mathcal{N}) = S(B_1B_2)\text{-}S(E_1E_2)$ (entropies evaluated on $\rho_{B_1B_2E_1E_2} = U \otimes U\, \phi_{12}\, U^\dagger \otimes U^\dagger$, where U is the isometric extension of $\mathcal{N}$.
- This gives us two input states to try on a single use of the channel: $\phi_1 = Tr_2\, \phi_{12}$, and $\phi_2 = Tr_1 \phi_{12}$. Let's evaluate how much coherent information the two of these give us.
- ϕ_1 gives us $S(B_1)\text{-}S(E_1)$, evaluated on $\rho_{B_1E_1} = U\phi_1U^\dagger = Tr_{B_2E_2}\rho_{B_1B_2E_1E_2}$, and similarly ϕ_2 gives $S(B_2)\text{-}S(E_2)$. Note that these entropies are evaluated on $\rho_{B_1B_2E_1E_2}$!
- Since $S(B_1)\text{-}S(E_1) \leq Q^1(\mathcal{N})$, and similarly for 2, now we just have to show that $S(B_1B_2) - S(E_1E_2) \leq S(B_1)\text{-}S(E_1) + S(B_2)\text{-}S(E_2)$.
- This is equivalent to showing $S(E_1)+S(E_2)-S(E_1E_2) = I(E_1; E_2) \leq I(B_1;B_2) = S(B_1)+S(B_2)\text{-}S(B_1B_2)$
- We can degrade B_1 to E_1, and B_2 to E_2, so this follows by monotonicity of mutual information!

The Capacity of Some Channels

- Erasure channel: $\mathcal{E}_p(\rho) = (1-p)\rho + p|e\rangle\langle e|$

$$Q(\mathcal{E}_p) = 1 - 2p$$



The Capacity of Some Channels

- Amplitude Damping Channel

$$\mathcal{A}_\gamma(\rho) = A_0 \rho A_0^\dagger + A_1 \rho A_1^\dagger$$

- $Q = \max_t (H(t\gamma) - H(t(1-\gamma)))$