Introduction to Quantum Channel Capacities

Graeme Smith

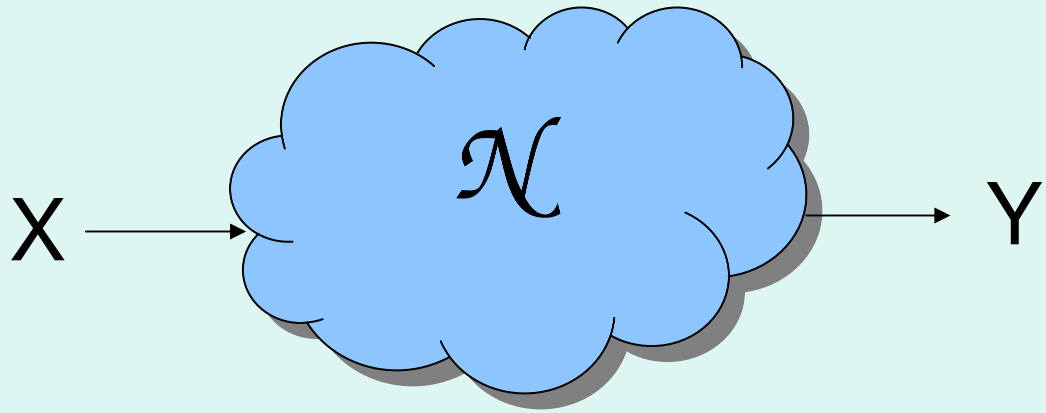
IBM Research

Information Theory

“The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point”

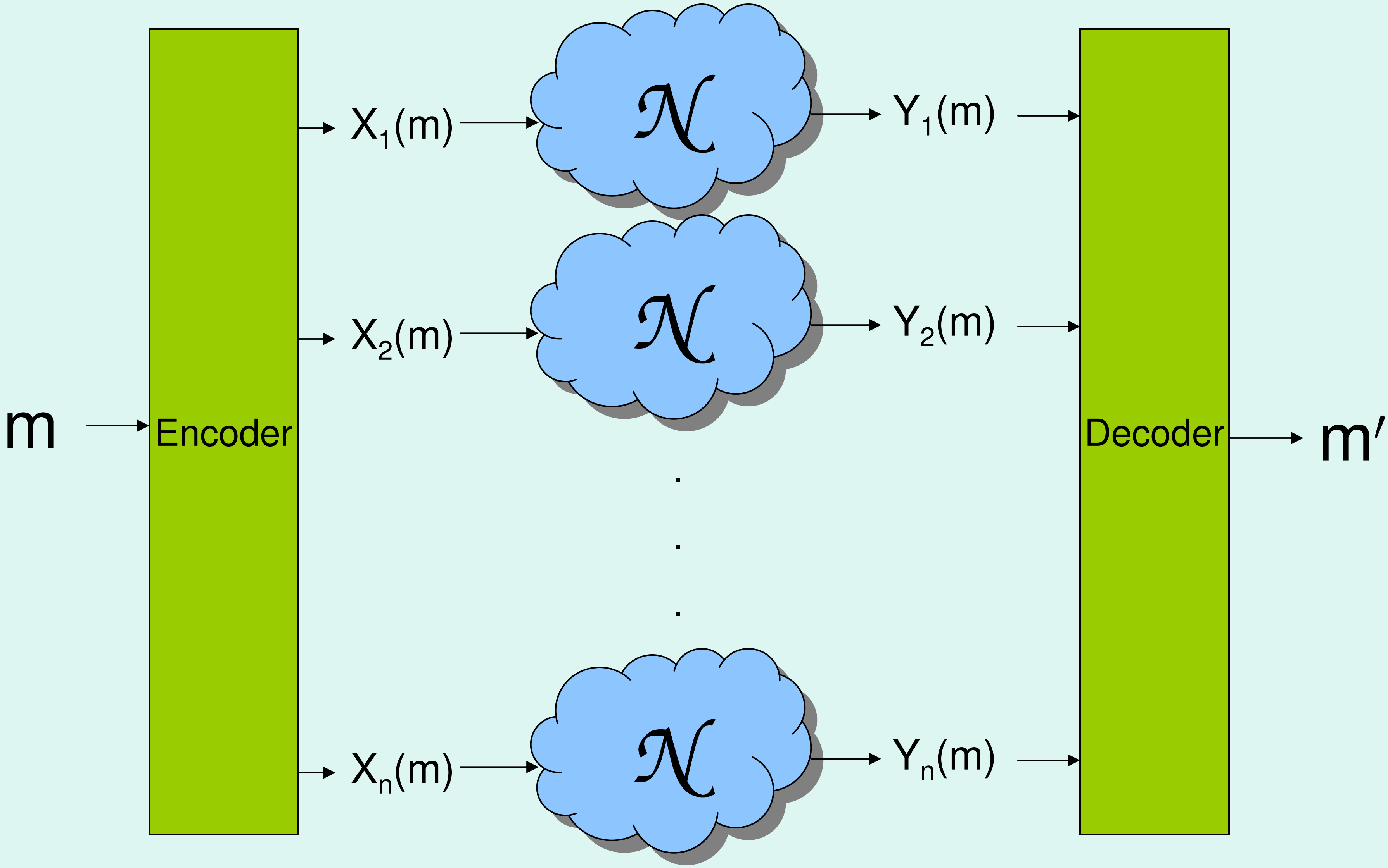
Source coding, channel coding, detection, cryptography…

Channel Coding

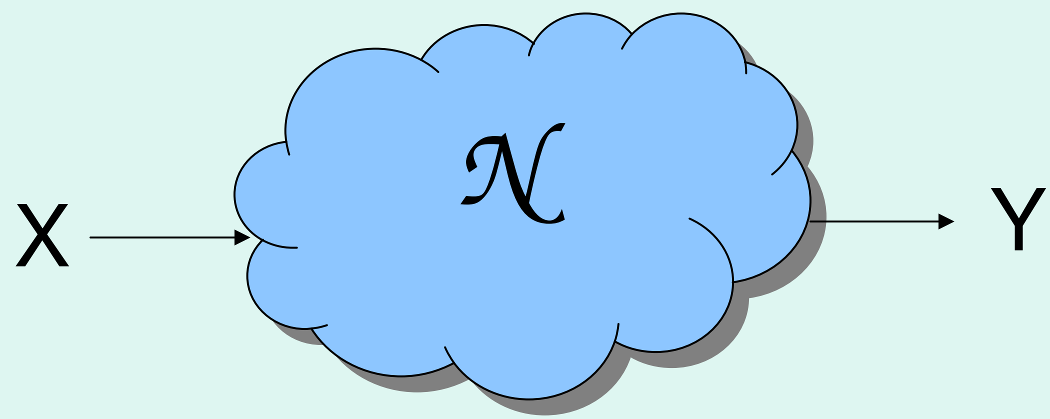


p(y|x)

Channel Coding



Channel Coding



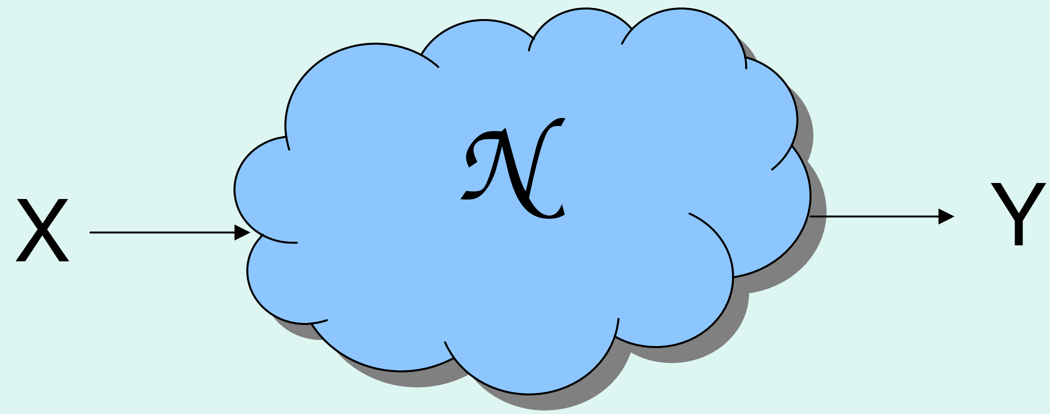
p(y|x)

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

Goal 1) understand capacity as a function of p(y|x)

Goal 2) find practical constructions for approaching capacity

Channel Coding



p(y|x)

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

Goal 1) understand capacity as a function of p(y|x)

Goal 2) find practical constructions for approaching capacity

Outline

- Classical coding theorem with converse
- Quantum states and channels
- Three quantum coding theorems
- Questions of additivity
  (where's the converse?)

Classical Coding Theorem: Typical Sequences

- Let $X_1, \ldots, X_n$ be i.i.d. (independent identically distributed) 0/1 r.v.s with $\Pr(X = 1) = p$, $\Pr(X=0)=1-p$

- If we look at the string $(X_1, \ldots, X_n)$ then w.h.p. it'll have $\approx pn$ 1's and $\approx (1-p)n$ 0's.

- Call these typical sequences.

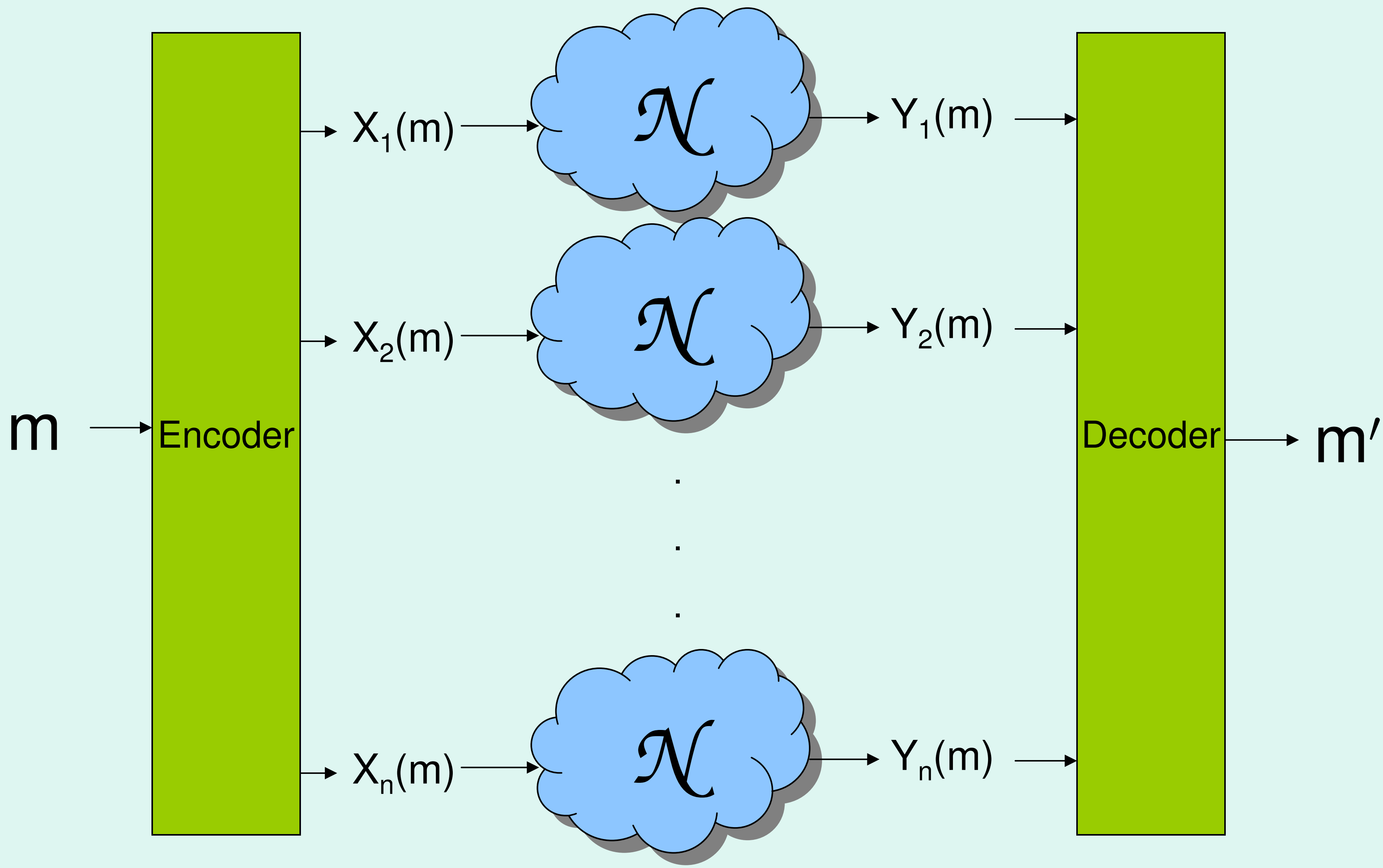
- How many? n choose $pn \approx 2^{n\,H(p)}$, where $H(p) = -p \log_2 p - (1-p) \log_2 (1-p)$

- Similar story for nonbinary variables.

Classical Coding Theorem: Conditionally Typical Sequences

- Now we have $(X_1,Y_1),\ldots (X_n,Y_n)$
- Let's say I tell you $(X_1,\ldots,X_n)$. For a typical $(X_1,\ldots,X_n)$, is there a high probability set that $(Y_1,\ldots,Y_n)$ almost certainly lives in?
- Yes! Call these Y's "conditionally typical".
- How many are there? $\approx 2^{n\, H(Y|X)}$, where $H(Y|X) = \sum p_x\, H(Y|x) = H(YX)-H(X)$

Channel Coding



Classical Coding Theorem

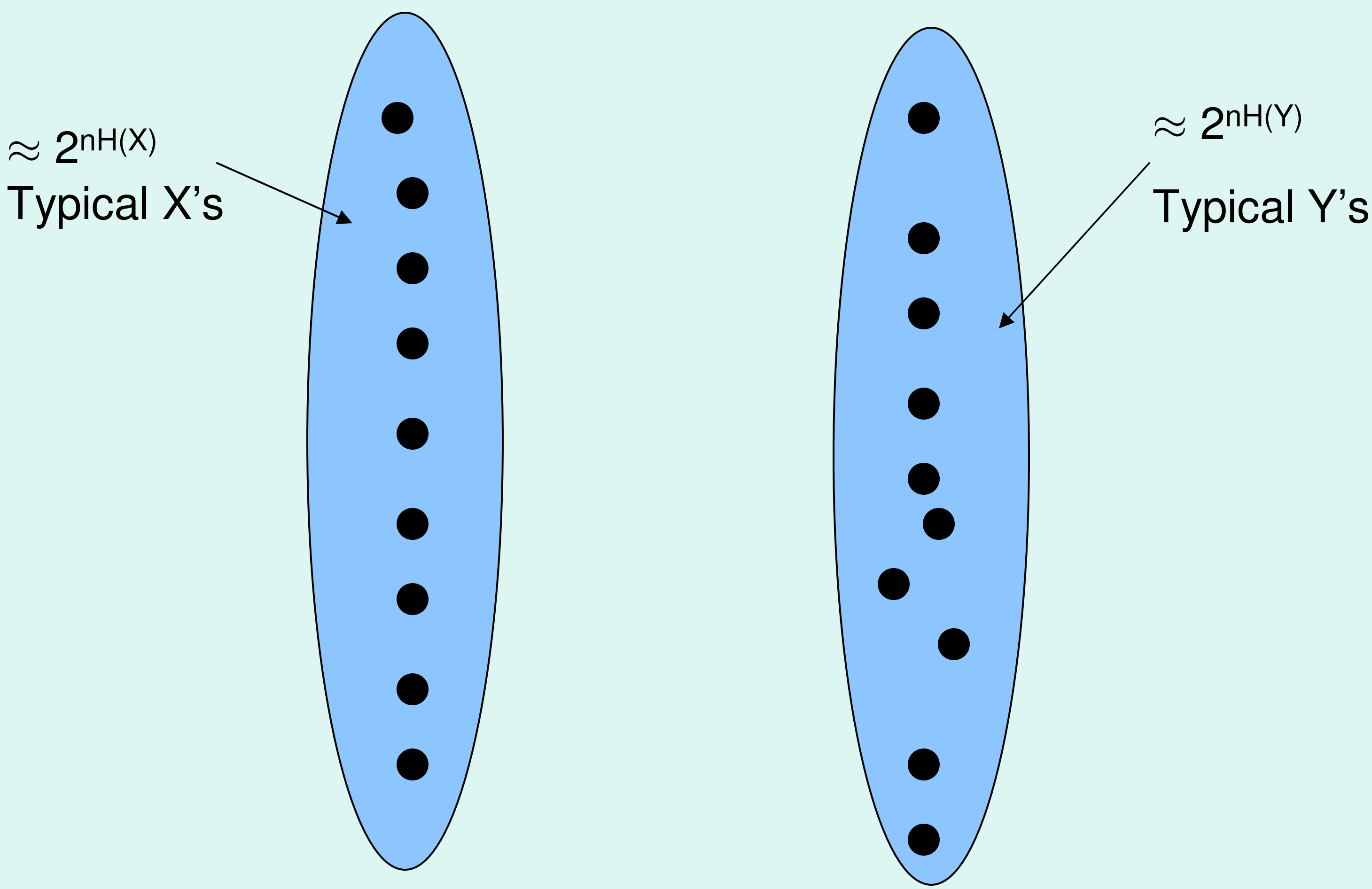
- We have a channel $\mathcal{N}$ with probs p(y|x)
- The capacity of $\mathcal{N}$ is the maximal number of bits per channel use we can send given a large number of uses.
- Shannon's Theorem:

$$\mathbf{C(\mathcal{N}) = max_X\ I(X;Y),}$$

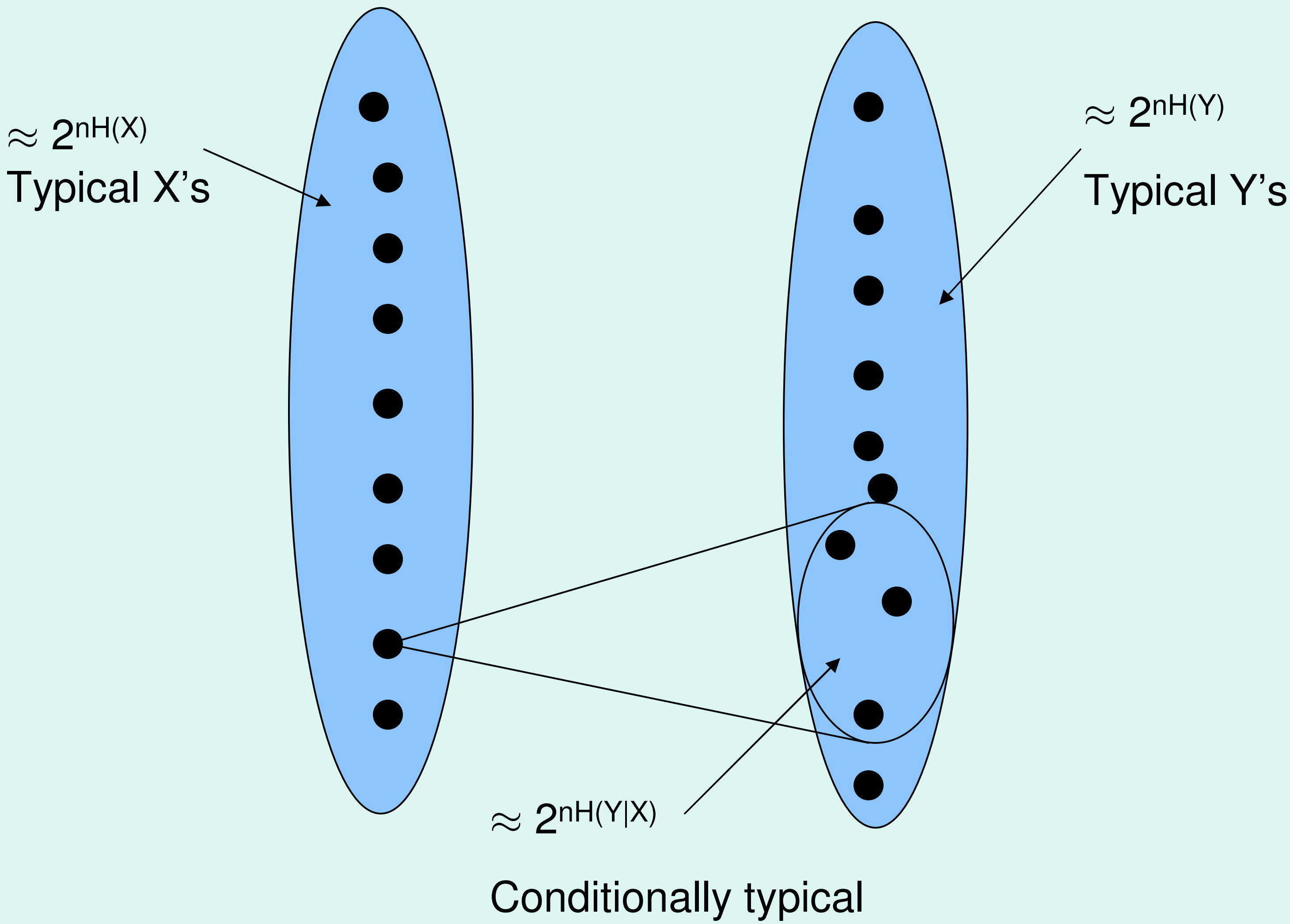
where $I(X;Y) = H(Y) - H(Y|X)$

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

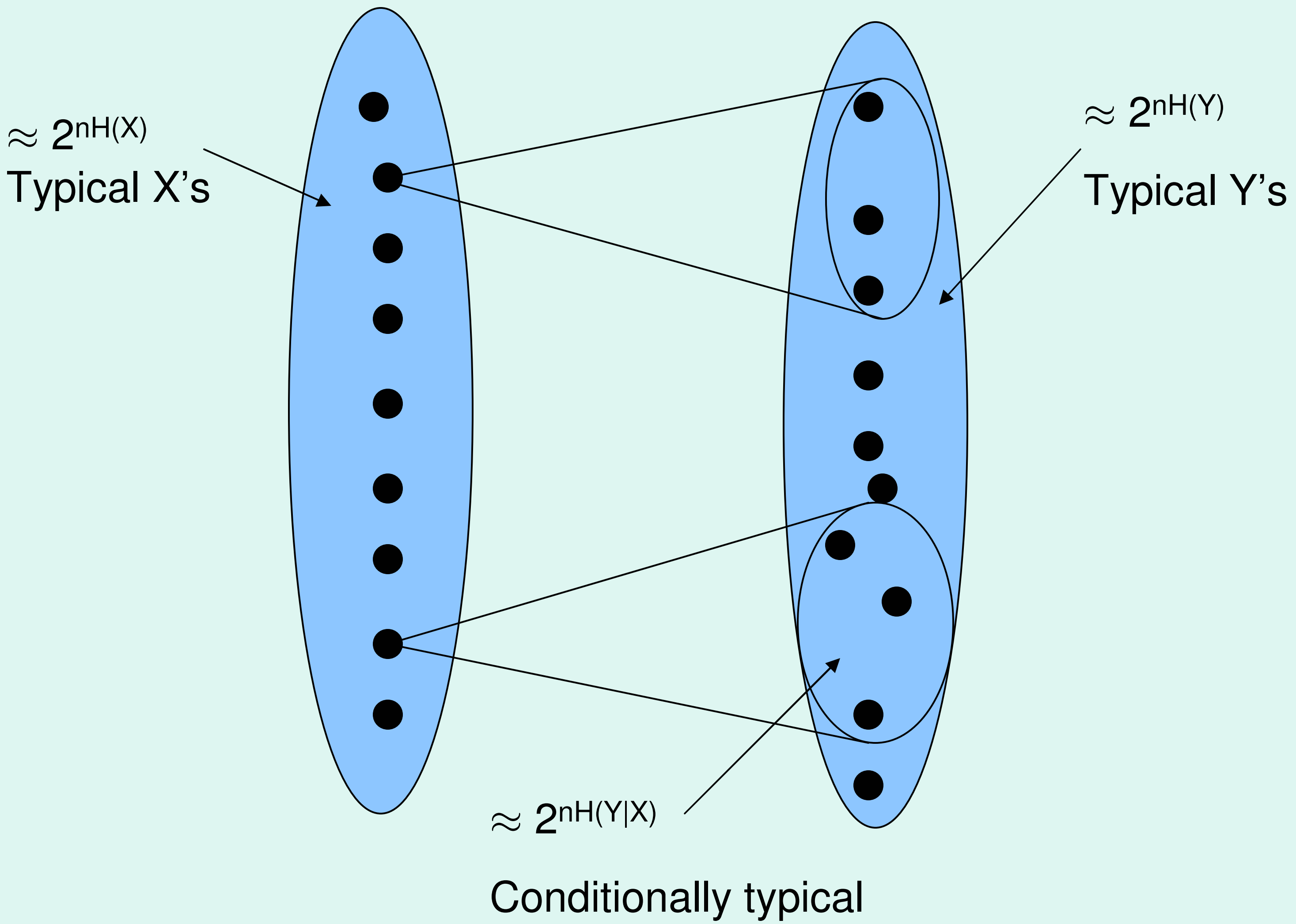
Classical Coding Theorem: How to pick a code



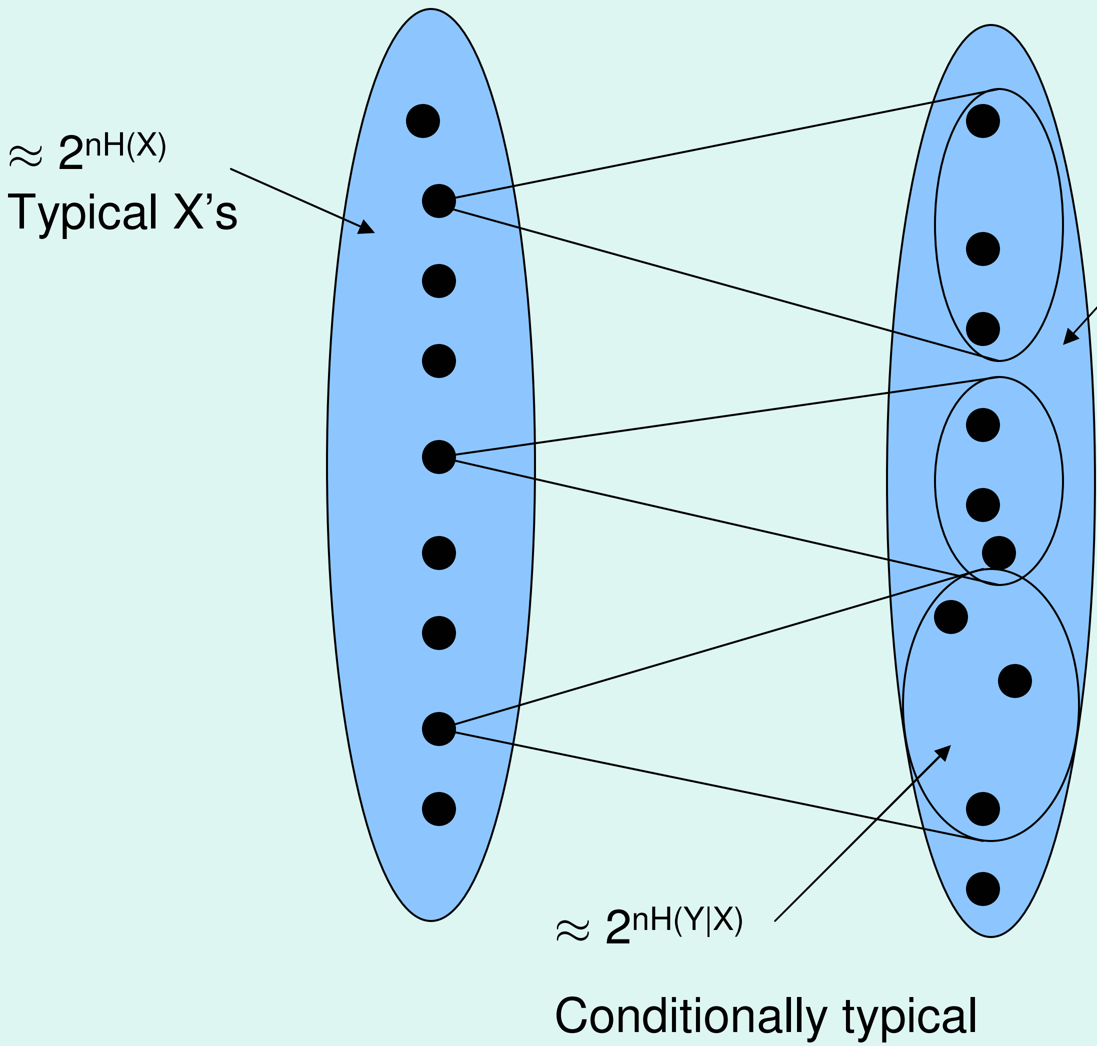
Classical Coding Theorem: How to pick a code



Classical Coding Theorem: How to pick a code



Classical Coding Theorem: How to pick a code



Keep adding typical X's to your code until you just barely fill up the set of typical Y's with their images.

How many X's can you take?

$\approx 2^{nH(Y)}/2^{nH(Y|X)}$

$= 2^{n\,I(X;Y)}$

Classical Converse

- We showed that any $R \leq \max_X I(X;Y)$ is achievable.
- Now we want that any $R > \max_X I(X;Y)$ is not.

Classical Converse

- Step one:
  show $C(\mathcal{N}) \leq \lim_{n \to \infty}(1/n)\max I(X^{(n)}; Y^{(n)})$
  $X^{(n)}$ is a r.v. on n inputs to the channel

- Step two:
  show $\max I(X^{(n)},Y^{(n)}) \leq n \max I(X;Y)$

- Step one involves continuity of entropy
  (see, e.g., Debbie Leung's talk)

- Step two is about additivity. It fails in most quantum cases (see Chris King's & B. Collins' talks, also Toby Cubitt's).

Quantum States and Channels

- Pure state: $|\psi\rangle \in$ C^d
- Mixed state: $\rho \in$ B(C^d) $\rho \geq$ 0, Tr ρ = 1
- In General: ρ = $\sum_\iota \lambda_i \, |\psi_i\rangle\langle\psi_i|$

- $\mathcal{N}$: B(C^d) $\to$ B(C^d), completely positive trace preserving.
- $\mathcal{N}(\rho)$ = $\sum_\kappa A_k \rho A_k^\dagger$ with $\sum A_k^\dagger A_k$ = I
- $\mathcal{N}(\rho)$ = Tr_E $U\rho U^\dagger$ with U : A$\to$ BE isometry.

Entropy and Typical Spaces

- Any ρ_B = $\mathrm{Tr}_A\, |\psi\rangle\langle\psi|_{AB}$
- $H(\rho_B)$ = -Tr ρ log ρ is the entropy
- It measures the uncertainty in B
- Given n copies of $|\psi\rangle$, can reversibly map B to a space of dimension $2^{n\,H(\rho_B)}$. This is the "typical space".

Quantum Coding Theorems

There are several kinds of information you can try to send with a quantum channel:

- Classical Information
- Private Classical Information
- Quantum Information

There are different capacities for each of these.

Actually, there are even more: I might give you free entanglement or free two-way classical communication to help.

Quantum Coding Theorems: Classical Information

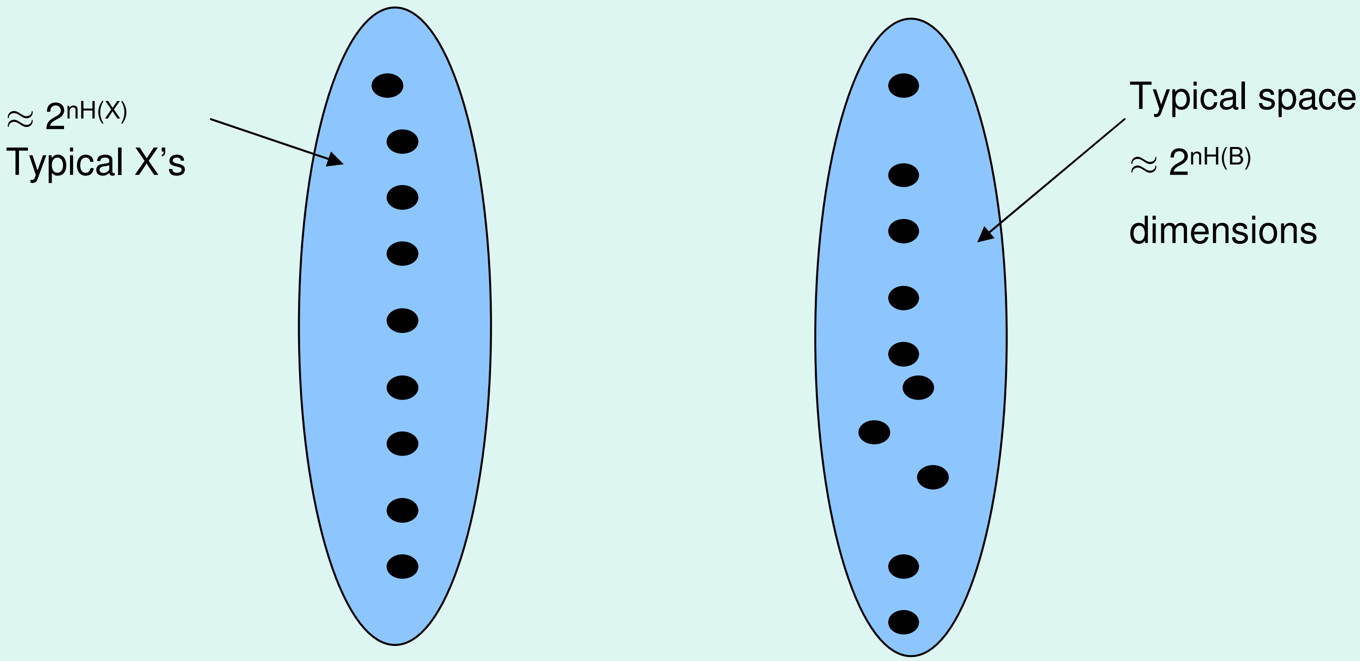
- Our channel maps $\mathcal{N}$: A$'$ $\rightarrow$ B
- We want a code {1,…,M} $\rightarrow$ $\rho_m \in$ B((A$'$)^n)
- Want large log M and $\mathcal{N}^{\otimes n}(\rho_m)$ distinguishable ($\approx$ orthogonal)

Quantum Coding Theorems: Achievable Classical Rate

- Our channel maps $\mathcal{N}$: A$'\to$ B
- Let X $\leftrightarrow$ p_x, and $\phi_x \in$ B(A$'$).
  We call $\mathcal{E}$ = {p_x, ϕ_x} an ensemble.
- $\chi(\mathcal{N}, \mathcal{E})$ = I(X;B), eval. on $\sum p_x|x\rangle\langle x|\otimes\mathcal{N}(\phi_x)$
- $\chi(\mathcal{N}, \mathcal{E})$ is achievable by random coding and "square-root measurement" (HSW 98)

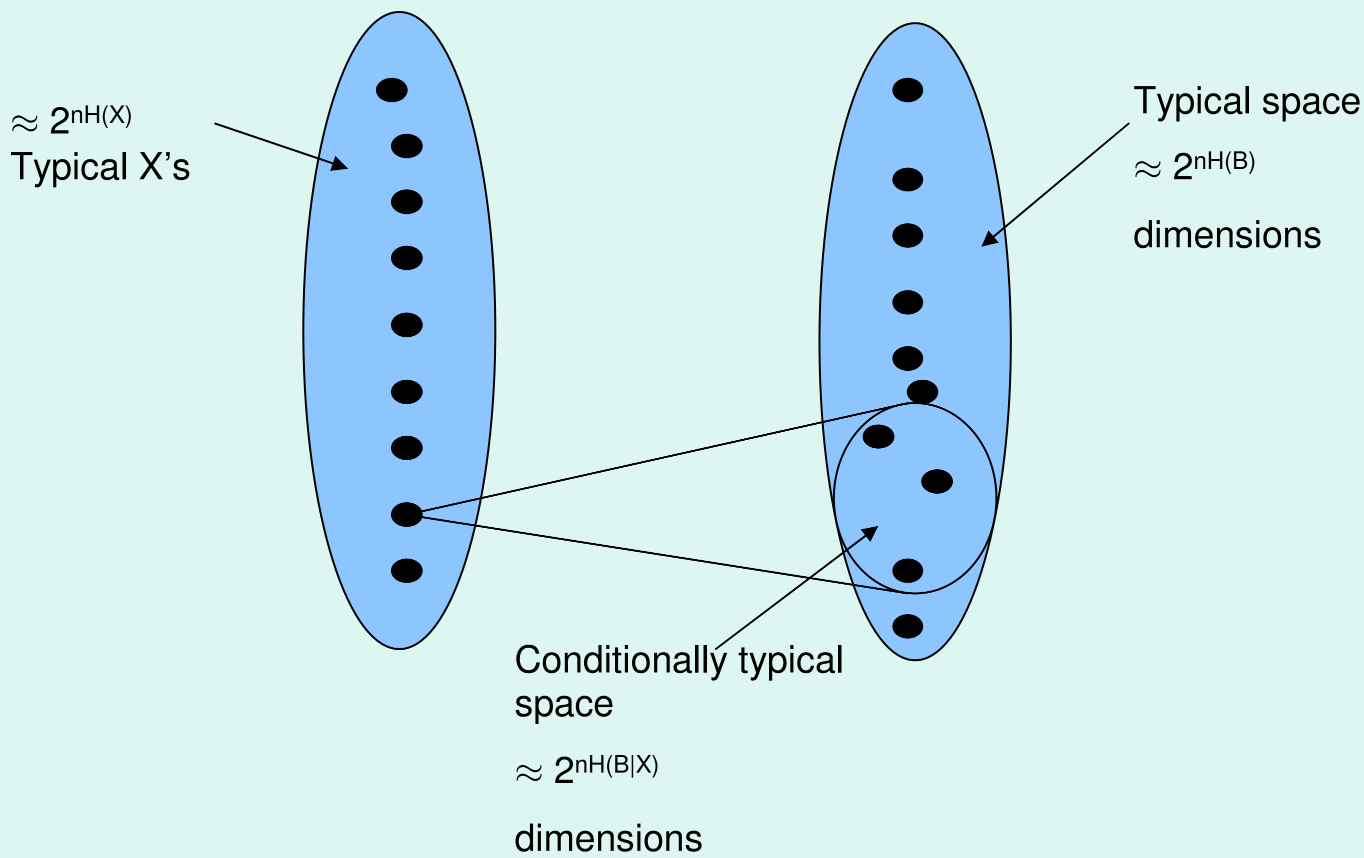
Quantum Coding Theorems: Achievable Classical Rate

$\mathcal{N}$: A $\rightarrow$ B, X $\rightarrow \phi_x \rightarrow \mathcal{N}(\phi_x) = \rho_x$ and let $\rho = \sum_\xi p_x \rho_x$



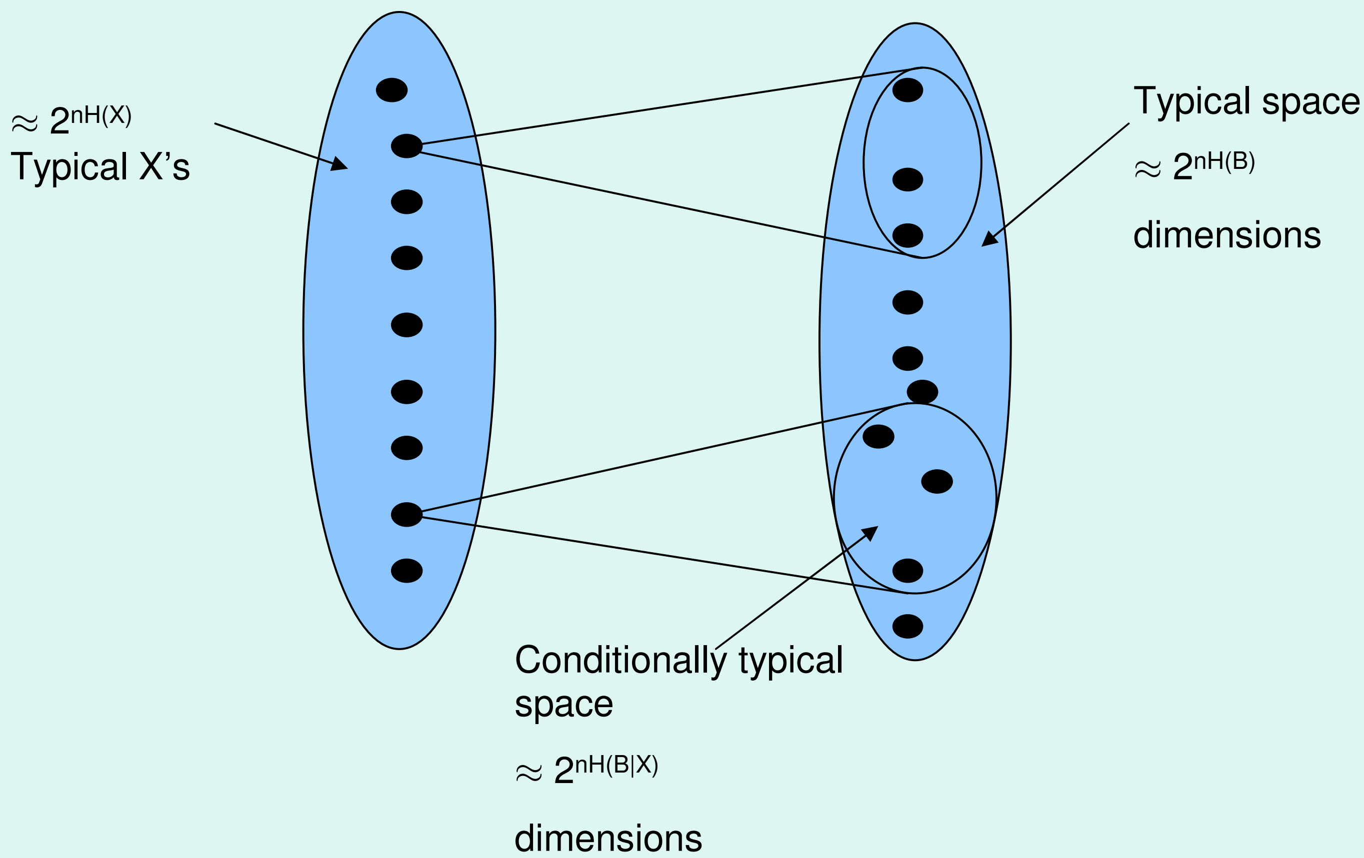
Quantum Coding Theorems: Achievable Classical Rate

$\mathcal{N}$: A $\rightarrow$ B, X $\rightarrow \phi_x \rightarrow \mathcal{N}(\phi_x) = \rho_x$ and let $\rho = \sum_\xi p_x \rho_x$



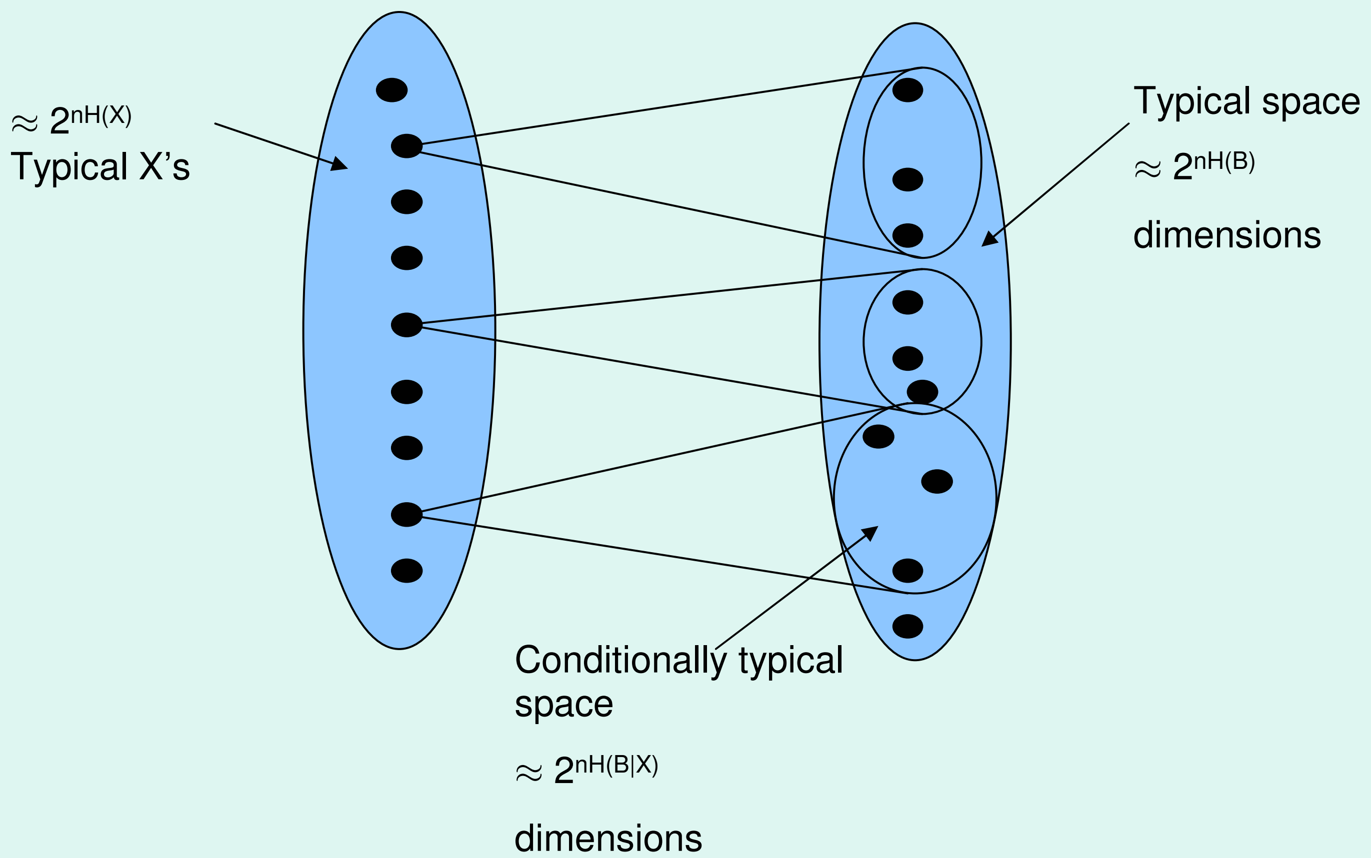
Quantum Coding Theorems: Achievable Classical Rate

$\mathcal{N} : A \rightarrow B$, $X \rightarrow \phi_x \rightarrow \mathcal{N}(\phi_x) = \rho_x$ and let $\rho = \sum_\xi p_x \rho_x$



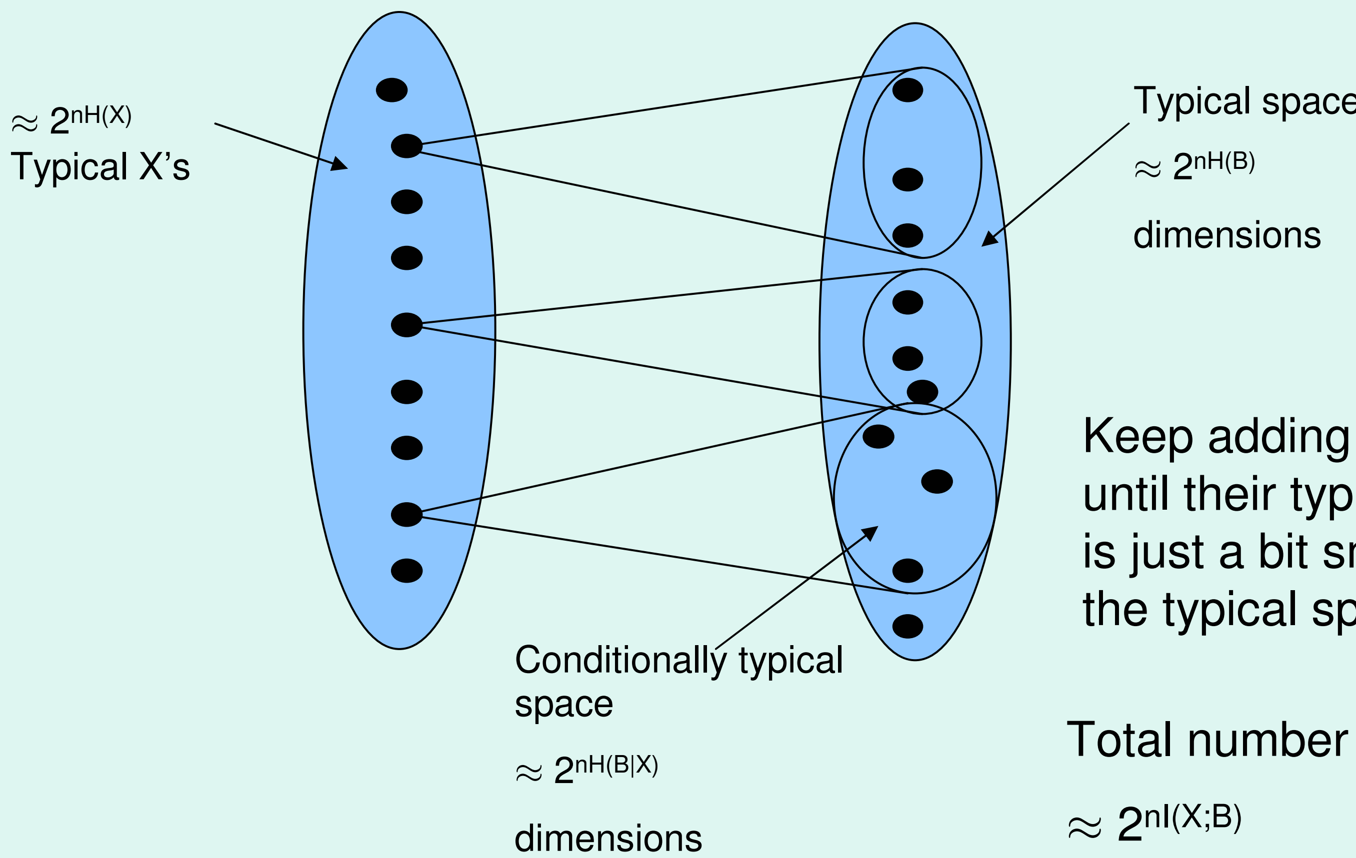
Quantum Coding Theorems: Achievable Classical Rate

$\mathcal{N}$: A $\rightarrow$ B, X $\rightarrow \phi_x \rightarrow \mathcal{N}(\phi_x) = \rho_x$ and let $\rho = \sum_\xi p_x \rho_x$



Quantum Coding Theorems: Achievable Classical Rate

$\mathcal{N}: A \to B$, $X \to \phi_x \to \mathcal{N}(\phi_x) = \rho_x$ and let $\rho = \sum_\xi p_x \rho_x$



Keep adding typical X's until their typical image is just a bit smaller than the typical space of B.

Total number of X's:

$\approx 2^{nI(X;B)}$

Quantum Coding Theorems: Partial Converse

- We saw that $\chi(\mathcal{N}, \mathcal{E})$ is achievable.
- Let $\chi(\mathcal{N}) = \max_{\mathcal{E}} \chi(\mathcal{N}, \mathcal{E})$.
  Then $(1/n)\ \chi(\mathcal{N}^{\otimes n})$ is achievable too.
- In fact, $C(\mathcal{N}) = \lim_{n \to \infty} (1/n)\ \chi(\mathcal{N}^{\otimes n})$

- For some channels $C(\mathcal{N}) = \chi(\mathcal{N})$, but not for others.
- Even better: for some $\mathcal{N}$ any code with rate $R > \chi(\mathcal{N})$ has exponentially bad fidelity ("Strong Converse" cf Stephanie Wehner's talk).

Quantum Coding Theorems: Private Classical Information

- Recall $\mathcal{N}(\rho) = \mathrm{Tr}_E U\rho U^\dagger$ and let $\mathcal{N}'(\rho)=\mathrm{Tr}_B\, U\rho U^\dagger$
- Want to send classical information to B but ensure E learns nothing of the message.
- Let $P^1(\mathcal{N}) = \max_{\mathcal{E}} \chi(\mathcal{N}, \mathcal{E}) - \chi(\mathcal{N}', \mathcal{E})$.

  Then $P^1(\mathcal{N})$ is acheivable.

- This is proved in two steps. 1) you can communicate to B at rate $\chi(\mathcal{N}, \mathcal{E})$. By averaging over these messages at a rate of $\chi(\mathcal{N}', \mathcal{E})$ we can smear out any information E gets.

Quantum Coding Theorems: Private Classical Information

- Like the classical capacity, we have $P(\mathcal{N}) = \lim_{n \to \infty}(1/n)P^1(\mathcal{N}^{\otimes n})$

- $P(\mathcal{N}) = P^1(\mathcal{N})$ for degradable channels. That means B can simulate E.

- However, even for qubit channels we can find instances of $P(\mathcal{N}) \neq P^1(\mathcal{N})$. The difference can be large.

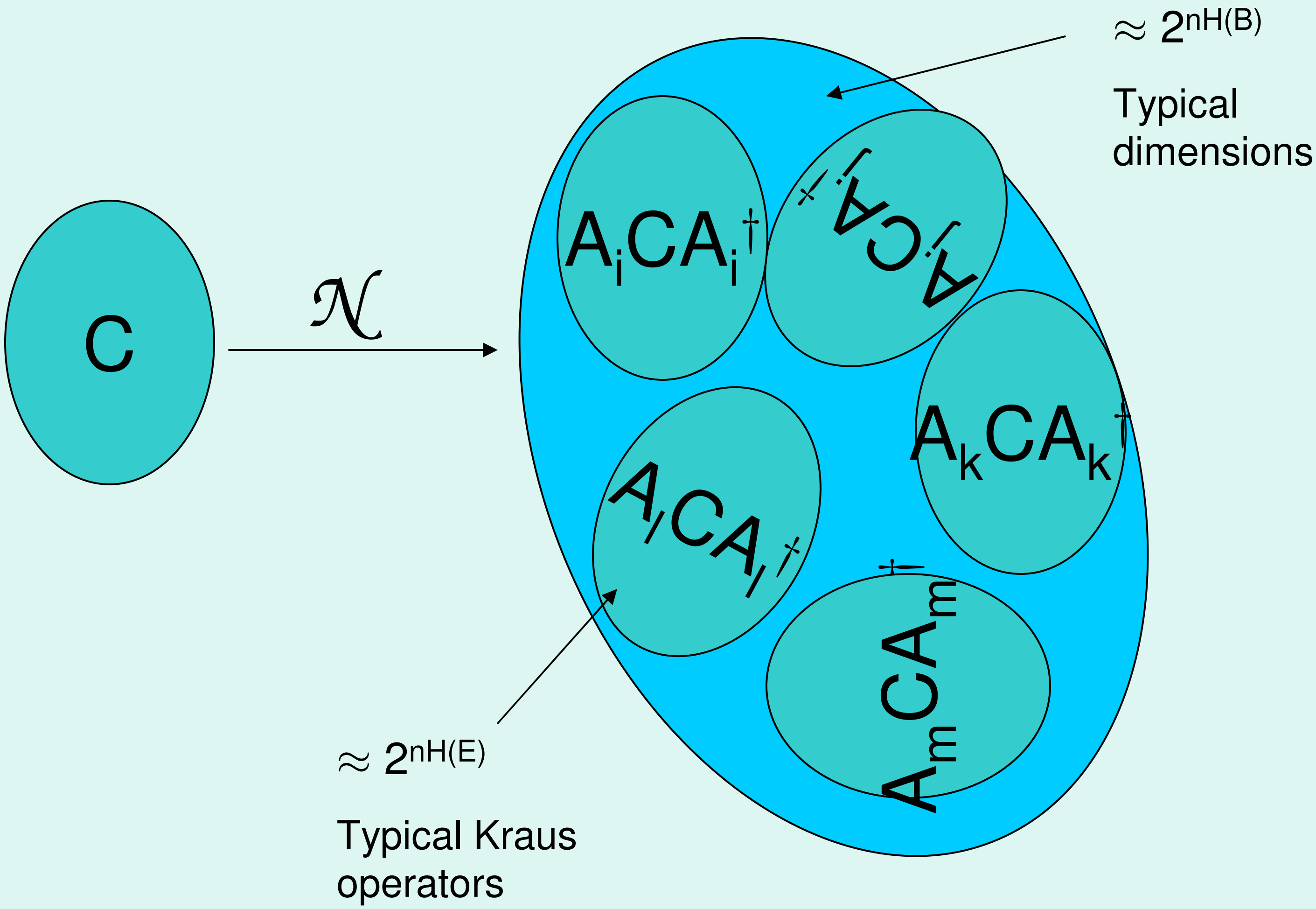
Quantum Coding Theorems: Quantum Information

- Our channel maps $\mathcal{N} : \mathrm{A}' \to \mathrm{B}$
- We want to find a subspace $\mathrm{C} \subset (\mathrm{A}')^{\otimes \mathrm{n}}$ and a decoding operation $\mathcal{D}$ such that

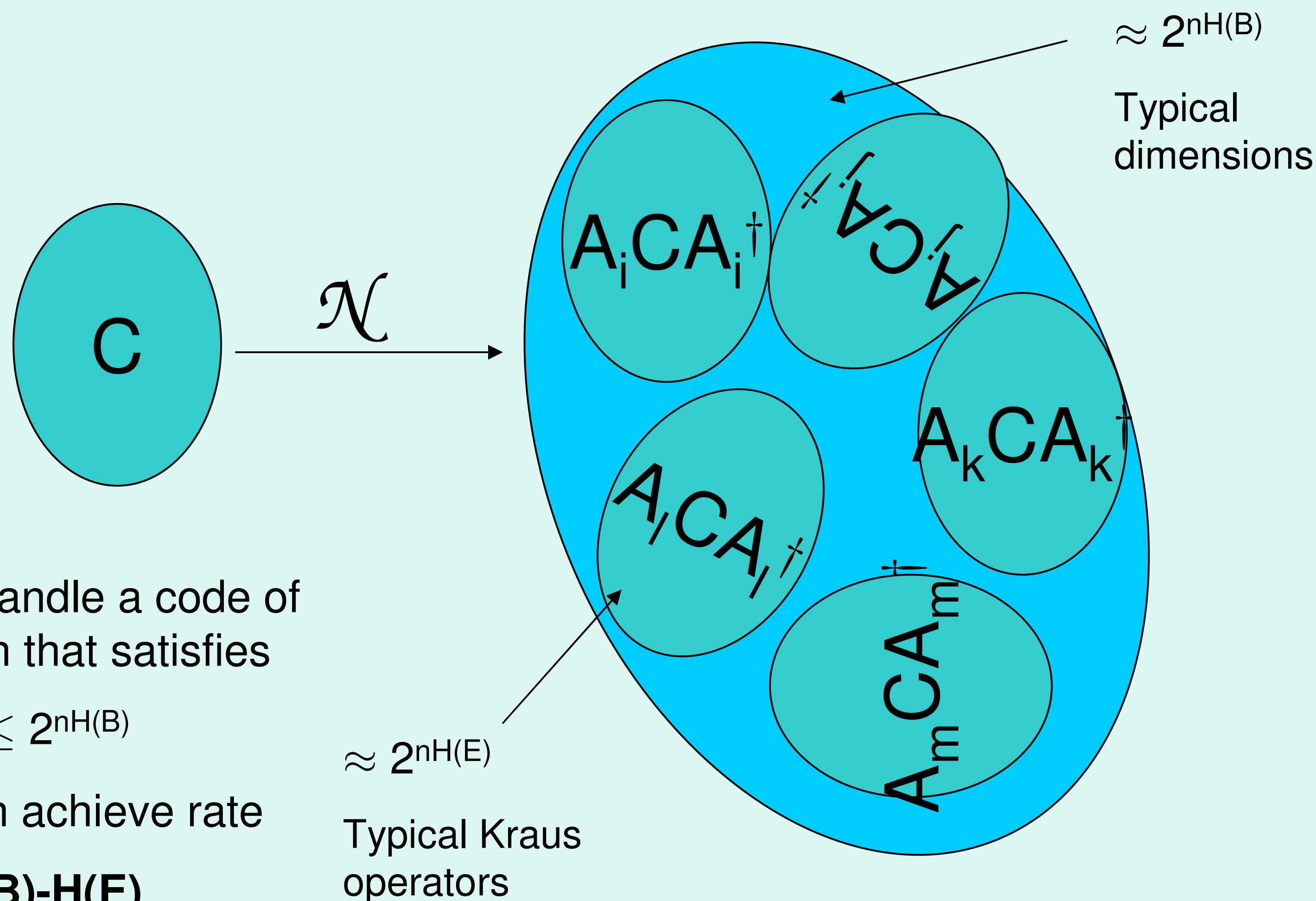
$$\mathcal{D} \circ \mathcal{N}^{\otimes \mathrm{n}}|_{\mathrm{C}} \approx \mathrm{id}_{\mathrm{C}}$$

- log dim C will be the number of qubits we can send.

Quantum Coding Theorems: Quantum Information



Quantum Coding Theorems: Quantum Information



We can handle a code of dimension that satisfies

$d_C \; 2^{nH(E)} \leq 2^{nH(B)}$

so we can achieve rate

R \leq H(B)-H(E)

Quantum Coding Theorems: Quantum Information

- Recall $\mathcal{N}(\rho)$ = $\mathrm{Tr}_E U\rho U^\dagger$ and let $\mathcal{N}'(\rho)$= $\mathrm{Tr}_B\ U\rho U^\dagger$
- Let $Q^1(\mathcal{N})$ = $\max_\rho$ H(B)-H(E), where the entropies are on $\mathcal{N}(\rho)$ and $\mathcal{N}'(\rho)$, respectively.
- $Q(\mathcal{N}) = \lim_{n \to \infty}(1/n)Q^1(\mathcal{N}^{\otimes n})$
- $Q(\mathcal{N}) = Q^1(\mathcal{N})$ for some channels but not for others. They can be very different.
- Get similar answer even if you're uncertain of the channel (see Igor Bjelakovic's talk)

Entanglement assisted capacity: The only one that's totally solved

- Lets say in addition to $\mathcal{N}$ I give Alice and Bob an arbitrary $|\psi\rangle_{\mathcal{AB}}$ to use. Note: $|\psi\rangle_{\mathcal{AB}}$ is no good for communication alone.

- In this setting the classical capacity of $\mathcal{N}$ is

$$C_E(\mathcal{N}) = \max I(A;B)$$

- No regularization needed!

Additivity Primer

A function on channels is additive if $f(\mathcal{N} \otimes \mathcal{M}) = f(\mathcal{N}) + f(\mathcal{M})$.

An additive information measure may give simple capacity formulas.

An additive capacity uniquely quantifies communication capability.

Information \ Quantity	Capacity	Information
Classical	Classical Capacity **?**	Holevo Information: χ = max I(X;B) **No** (Hastings '09)
Private	Private Capacity **No**(Li-Winter-Zou-Guo '09)	Private Information: max I(X;B)-I(X;E) **No** (S-Renes-Smolin '08)
Quantum	Quantum Capacity **No** (S-Yard '08)	Coherent Information: max S(B)-S(E) **No** (Div-Shor-Smolin '98)
Entanglement Assisted Classical	E.A. Capacity **Yes** (Bennett-Shor-Smolin-Thaplyial '01)	Mutual Information: max I(A;B) **Yes** (Bennett-Shor-Smolin-Thaplyial '01)

Additivity Primer

Even though most natural information measures and capacities are nonadditive in general, there are nontrivial examples where additivity holds. We want more!

- Entanglement Breaking Channels: If $\mathcal{N}$ is EB and $\mathcal{M}$ is arbitrary, $\chi(\mathcal{N} \otimes \mathcal{M})=\chi(\mathcal{N})+\chi(\mathcal{M})$. Even better, $C(\mathcal{N} \otimes \mathcal{M}) = C(\mathcal{N})+C(\mathcal{M})$.
- Degradable Channels: B can simulate E. Coherent information is additive. Any two such channels have additive capacity. The private capacity equals the quantum, so this behaves too.
- Unital qubit channels, depolarizing channels, bosonic gaussian channels and others have additive χ.

Known unknowns: Some things I haven't mentioned

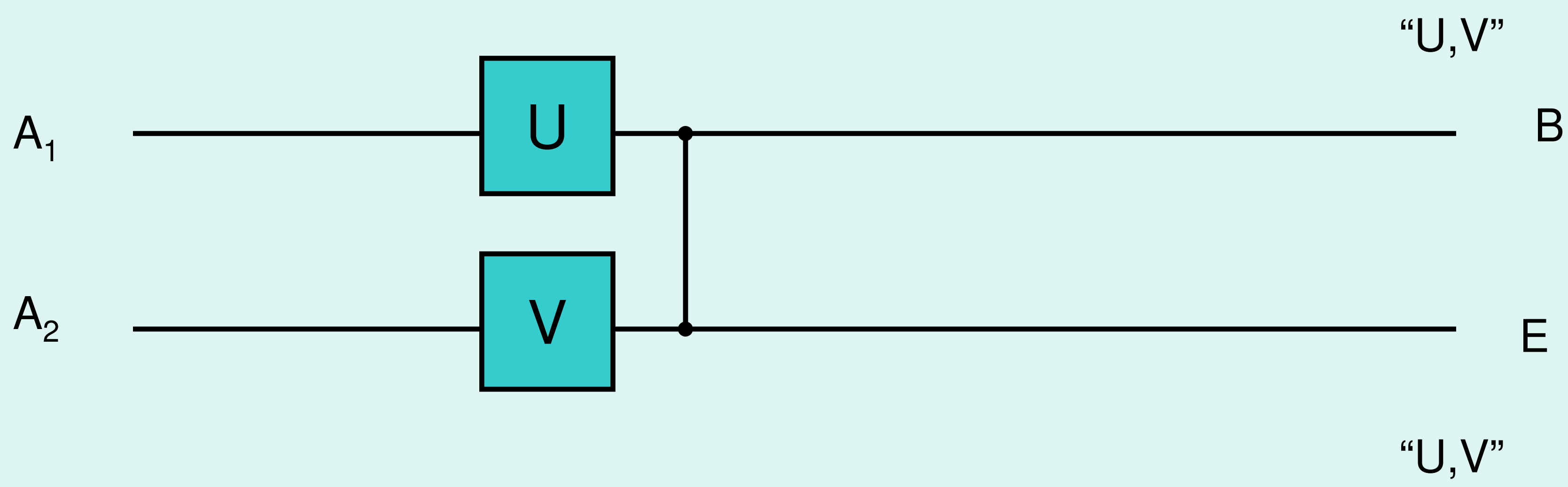
- Multiple access channels, broadcast channels, and multi-user information theory (Careful, though: some of these are hard classically).
- How do we actually achieve these rates?
  (slightly unsatisfying answer: Forney construction.)
- Coding theory, fault-tolerance, etc.
- Pure-state source coding (aka "data compression") is actually solvable.
- Two-way capacities and relationship to entanglement and LOCC.
- PPT criterion and NPT bound entanglement?
- $P \neq Q$
- Connections between Quantum Key Distribution and private capacities (tomography, non-iid, etc.).
- Beyond i.i.d. (symmetrization and de Finetti arguments)
- Identification capacity, environment assisted capacity, capacity of unitary interactions, symmetric side channels, commitment capacity, reverse Shannon theorem, embezzling states, entanglement measures, zero-error…

Rocket Channels

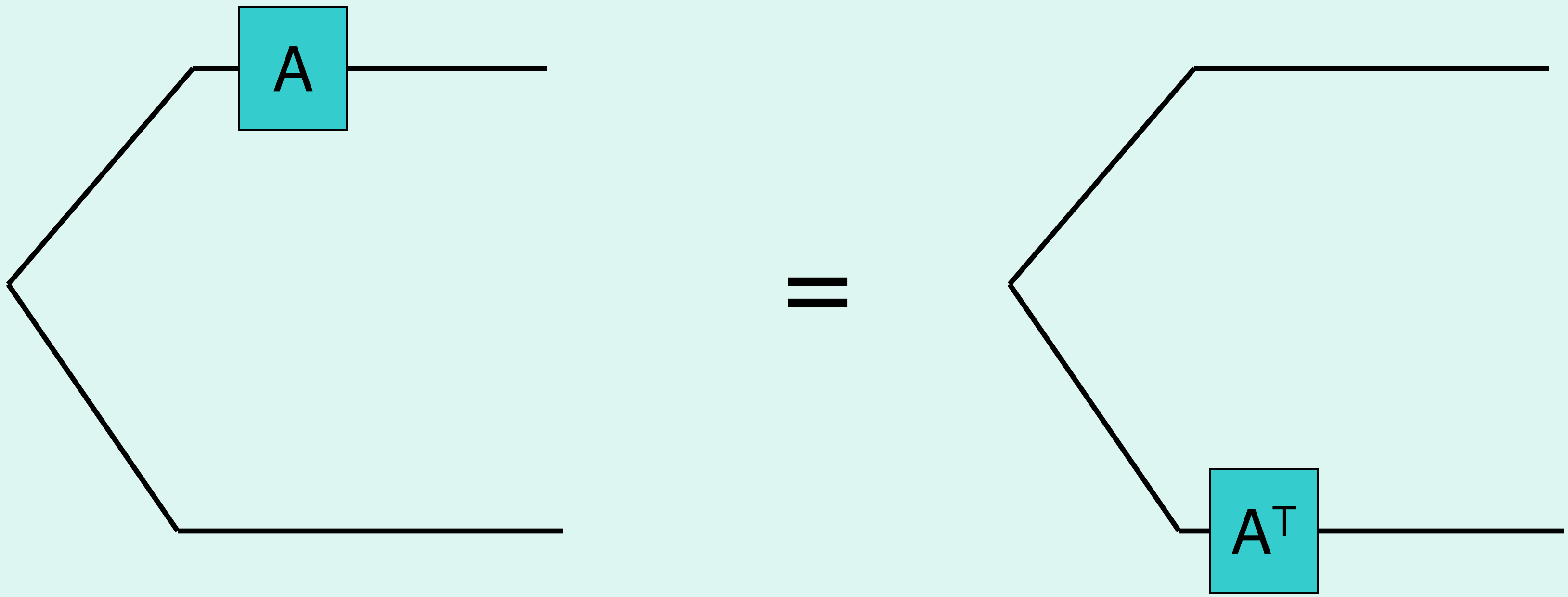
- Simple channel displaying extensive nonadditivity of private capacity when used with a 50% erasure.
- Actually, even small *classical* capacity and they have large joint *quantum* capacity.
- Circuit diagram looks like a rocket!

Rocket Channels

$$\mathcal{R}_d = E(\mathcal{R}^{U,V}{}_d \otimes |UV\rangle\langle UV|)$$



A nice identity



$A \otimes I\,|\phi_d\rangle = \sum A_{i,j}\,|i\rangle\,|j\rangle$ $\qquad$ $I \otimes B\,|\phi_d\rangle = \sum B_{j,i}\,|i\rangle\,|j\rangle$

$$|\phi_d\rangle = \sum |i\rangle\,|i\rangle$$

Bob can undo interaction