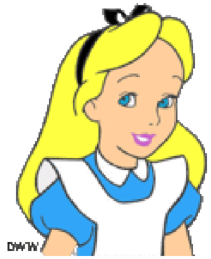
Additivity
in Quantum Shannon Theory

Mark M. Wilde

School of Computer Science
McGill University

2010 International Workshop on Quantum Information Science
ERATO-SORST Project, Tokyo, Japan
Monday, March 8, 2010

Tutorial Overview



Classical Tasks

Transmission of classical information

Transmission of private classical information

Quantum Tasks

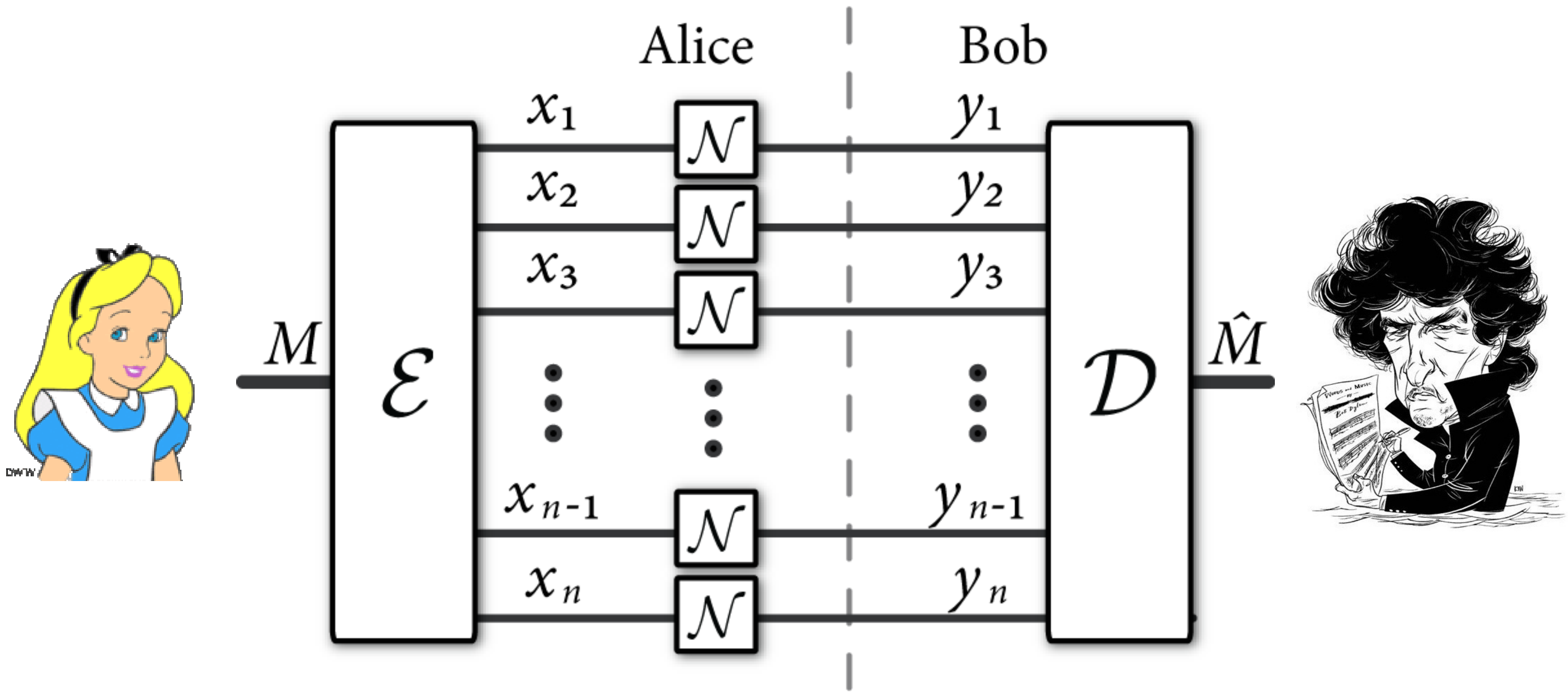
Transmission of classical information

Trans. of classical info. with help of unlimited entanglement

Transmission of quantum information

Transmission of private classical information
Trade-off coding

"Dynamic" Shannon Theory



Classical channel is the stochastic map $\mathcal{N} \equiv p_{Y|X}(y|x)$

Given a large number *n* of uses of a classical channel, what is the **largest rate** of reliable communication?

(where rate is $\dfrac{\log(M)}{n}$)

Shannon's Capacity Theorem

Largest reliable rate is the **capacity**

$$I(\mathcal{N}) \equiv \max_{p_X(x)} I(X;Y)$$

Might call this measure
the mutual information of the classical channel

Proof follows from three important steps:

1) Direct Coding Theorem (construction of random code)

2) Converse Theorem (bounding the channel information throughput)

3) **Additivity** of the proposed channel information measure

The Importance of Additivity

Implies a **complete understanding**
of a channel's transmission capabilities

Implies the **proposed** capacity formula is the correct one

Without additivity, the best characterization is
an **intractable "regularization"**

$$I_{\mathrm{reg}}(\mathcal{N}) \equiv \lim_{n\to\infty} \frac{1}{n} I(\mathcal{N}^{\otimes n})$$

(pretty much useless)

Justification of postdoc salary:

"Probably every quantum information theorist
worth his salt has had a go on that one."

-Werner 2005

Additivity of Classical Channels

Given two classical channels:

$$\mathcal{N}_1 \equiv p_{Y_1|X_1}(y_1|x_1)$$

$$\mathcal{N}_2 \equiv p_{Y_2|X_2}(y_2|x_2)$$

Does **additivity** of channel mutual information hold?

$$I(\mathcal{N}_1 \otimes \mathcal{N}_2) = I(\mathcal{N}_1) + I(\mathcal{N}_2)$$

"Easy direction" always holds:

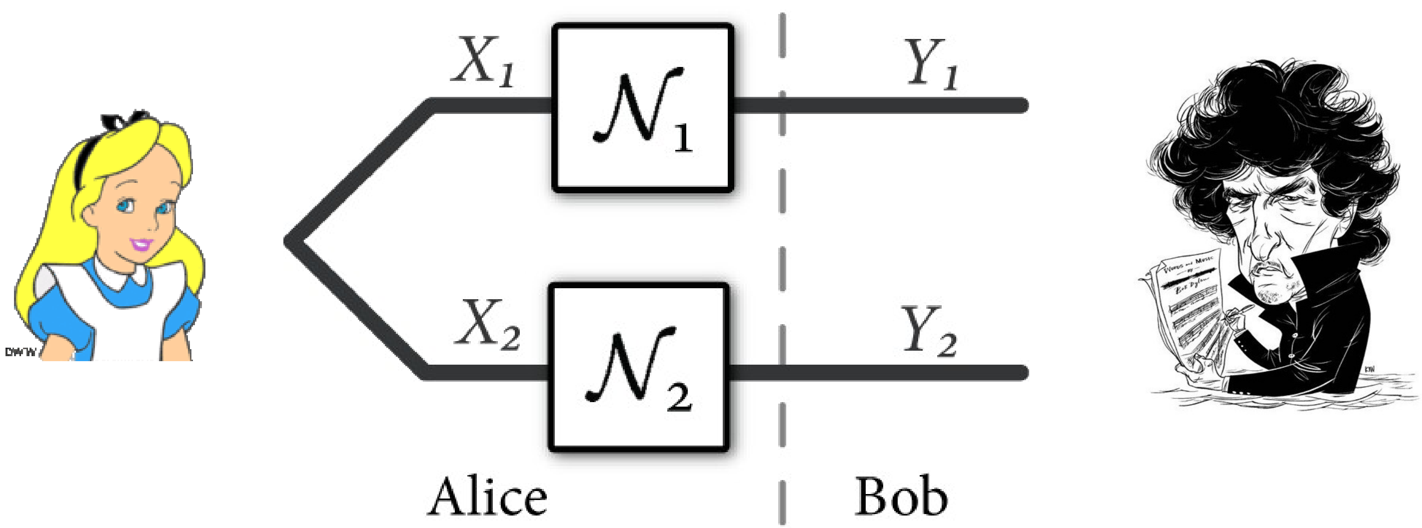
$$I(\mathcal{N}_1 \otimes \mathcal{N}_2) \geq I(\mathcal{N}_1) + I(\mathcal{N}_2)$$

Choose $p_{X_1,X_2}(x_1, x_2) = p^*_{X_1}(x_1)p^*_{X_2}(x_2)$

Additivity of Classical Channels (Ctd)

Does **"hard direction"** hold?

$$I(\mathcal{N}_1 \otimes \mathcal{N}_2) \leq I(\mathcal{N}_1) + I(\mathcal{N}_2)$$



Correlations between inputs **do not increase** information throughput?

Yes!

(and holds for all classical channels)

Follows because Y_1 independent of X_2
and Y_2 is conditionally independent of X_1 and Y_1 given X_2

Additivity of Classical Channels (Ctd)

Additivity of classical channel mutual information holds:

$$I(\mathcal{N}_1 \otimes \mathcal{N}_2) = I(\mathcal{N}_1) + I(\mathcal{N}_2)$$

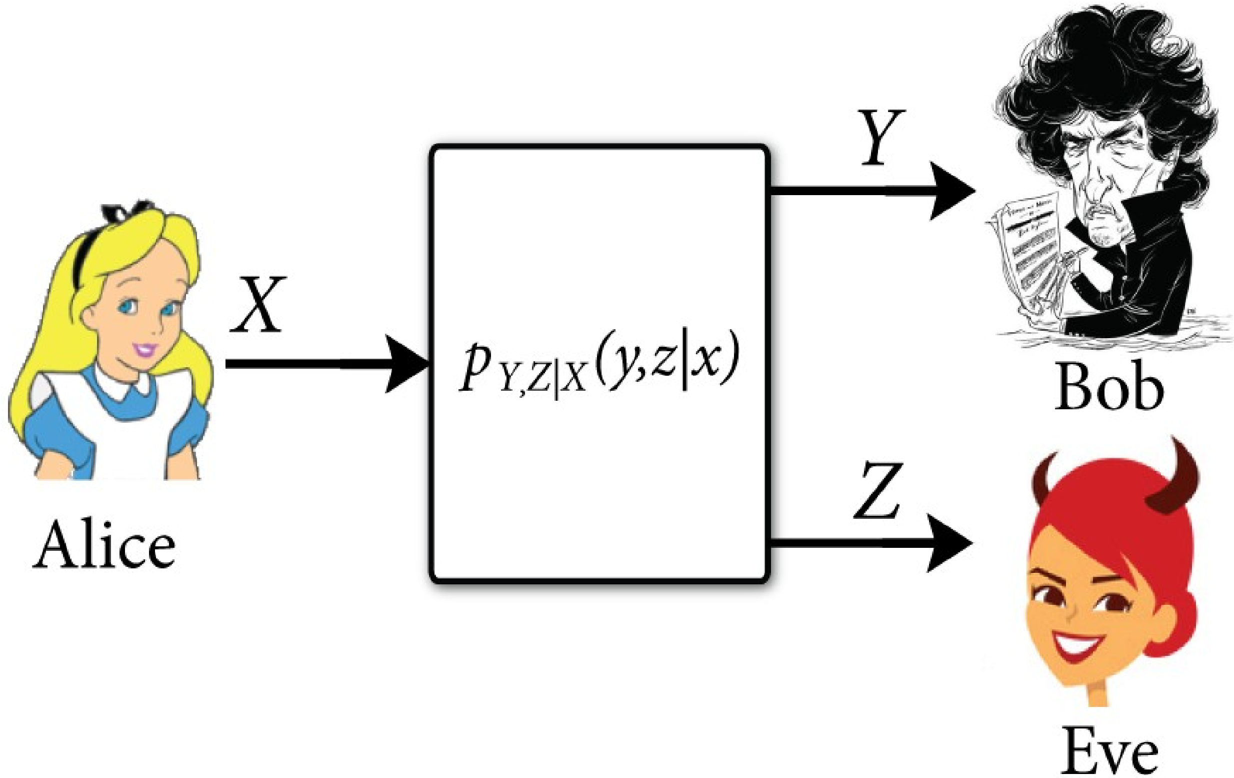
By **induction**, it holds that

$$I_{\mathrm{reg}}(\mathcal{N}) = I(\mathcal{N})$$

(No need for regularization)

Implies a **complete understanding** of the transmission capabilities of classical memoryless channels

Classical Wiretap Channel



Wiretap channel is the stochastic map $\mathcal{N} \equiv p_{Y,Z|X}(y,z|x)$

Private information of the wiretap channel:

$$P(\mathcal{N}) \equiv \max_{p_X(x)} I(X;Y) - I(X;Z)$$

Aaron D. Wyner, "The wire-tap channel", Bell. Sys. Tech. Jour., vol. 54, pp. 1355–1387, 1975.

Additivity of Classical Wiretap Channels

Given two classical wiretap channels:

$$\mathcal{N}_1 \equiv p_{Y_1,Z_1|X_1}(y_1, z_1|x_1)$$

$$\mathcal{N}_2 \equiv p_{Y_2,Z_2|X_2}(y_2, z_2|x_2)$$

Does **additivity** of channel private information hold?

$$P(\mathcal{N}_1 \otimes \mathcal{N}_2) = P(\mathcal{N}_1) + P(\mathcal{N}_2)$$

"Easy direction" again always holds:

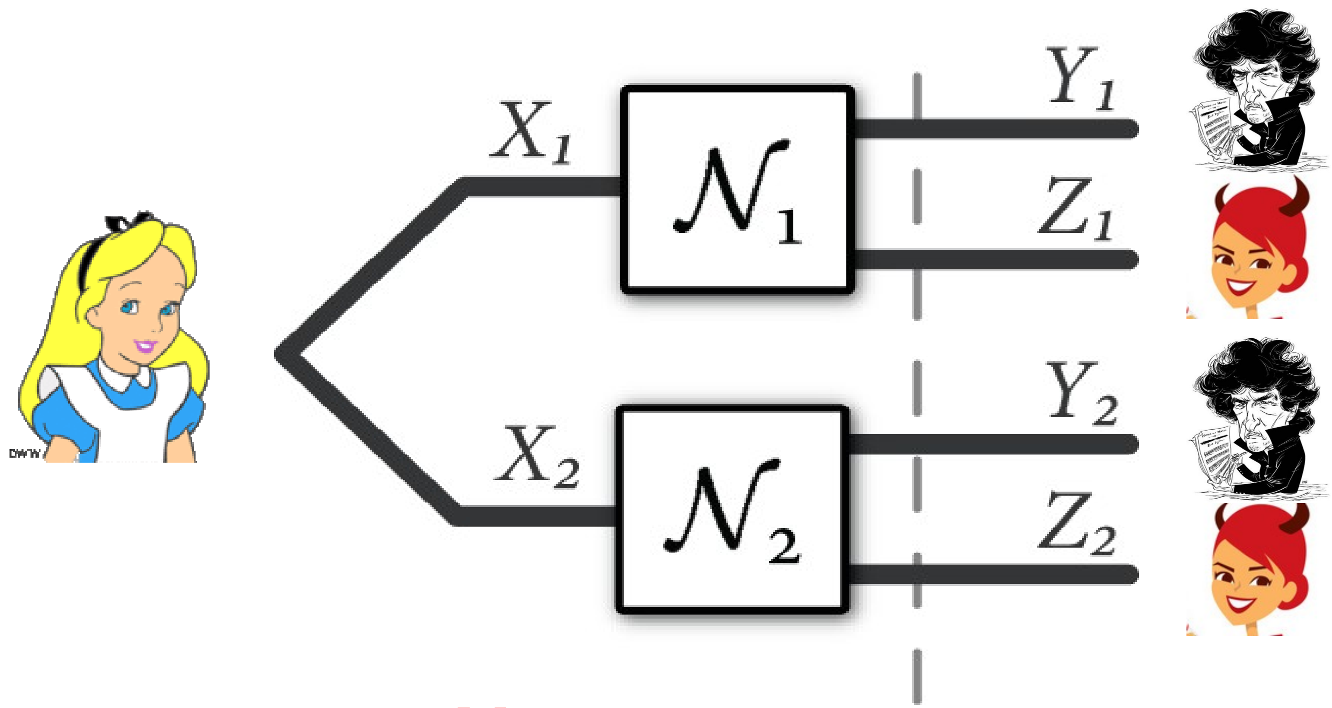
$$P(\mathcal{N}_1 \otimes \mathcal{N}_2) \geq P(\mathcal{N}_1) + P(\mathcal{N}_2)$$

Choose $p_{X_1,X_2}(x_1, x_2) = p^*_{X_1}(x_1)p^*_{X_2}(x_2)$

Additivity of Classical Wiretap Channels

Does **"hard direction"** hold?

$$P(\mathcal{N}_1 \otimes \mathcal{N}_2) \leq P(\mathcal{N}_1) + P(\mathcal{N}_2)$$



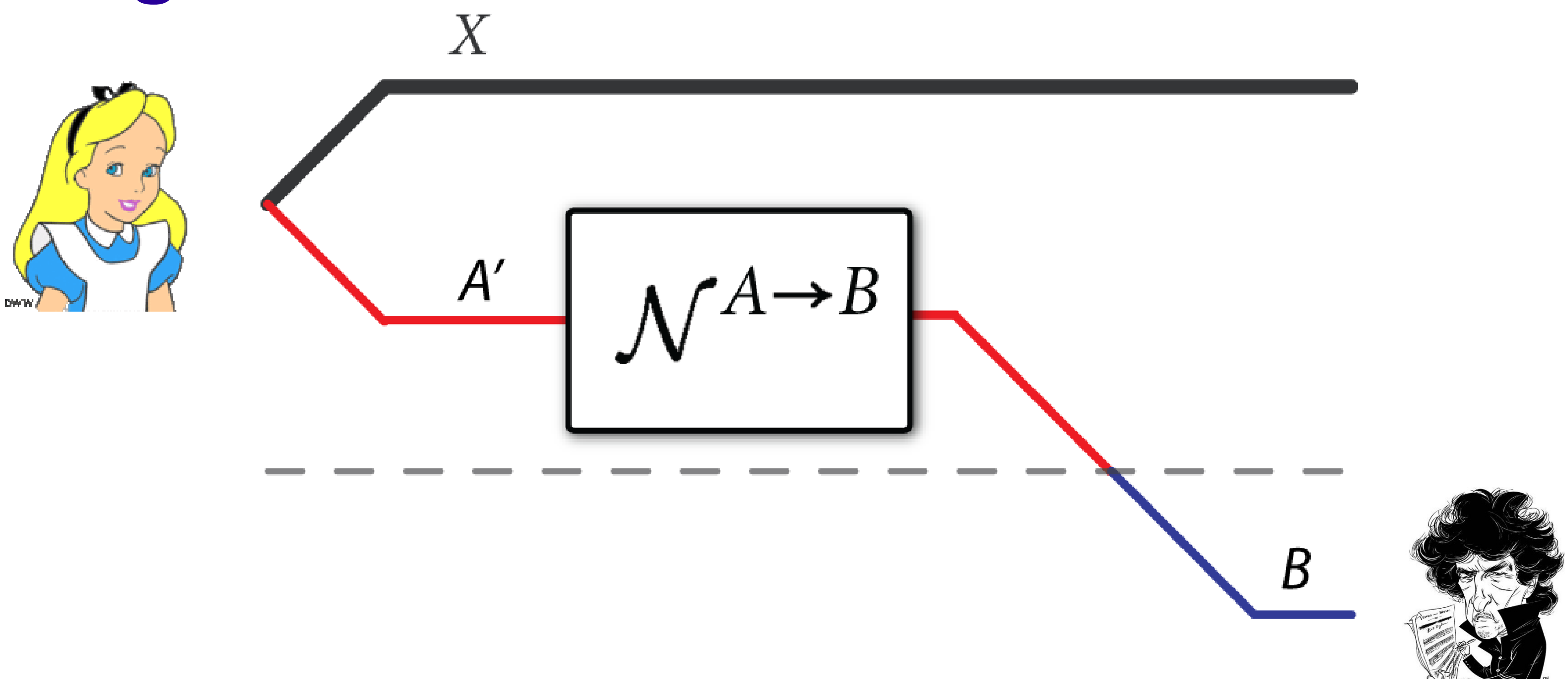
Not in general,

but **does** if correlations in Bob's outputs are greater than Eve's:

$$I(Y_1; Y_2) \geq I(Z_1; Z_2)$$

Concept of **degradability** useful in quantum setting as well

Sending Classical Data over Quantum Channels



Correlate classical data with quantum states:

$$\sum_x p_X(x)|x\rangle\langle x|^X \otimes \mathcal{N}^{A'\to B}(\phi_x^{A'})$$

Holevo information of a quantum channel:

$$\chi(\mathcal{N}) \equiv \max_{\{p_X(x),\phi_x\}} I(X;B)$$

Holevo (1998), Schumacher and Westmoreland (1997)

Additivity of Holevo Information?

Given two quantum channels (CPTP maps),

does **additivity** of channel Holevo information hold?

$$\chi(\mathcal{N}_1 \otimes \mathcal{N}_2) = \chi(\mathcal{N}_1) + \chi(\mathcal{N}_2)$$

"Easy direction" always holds:

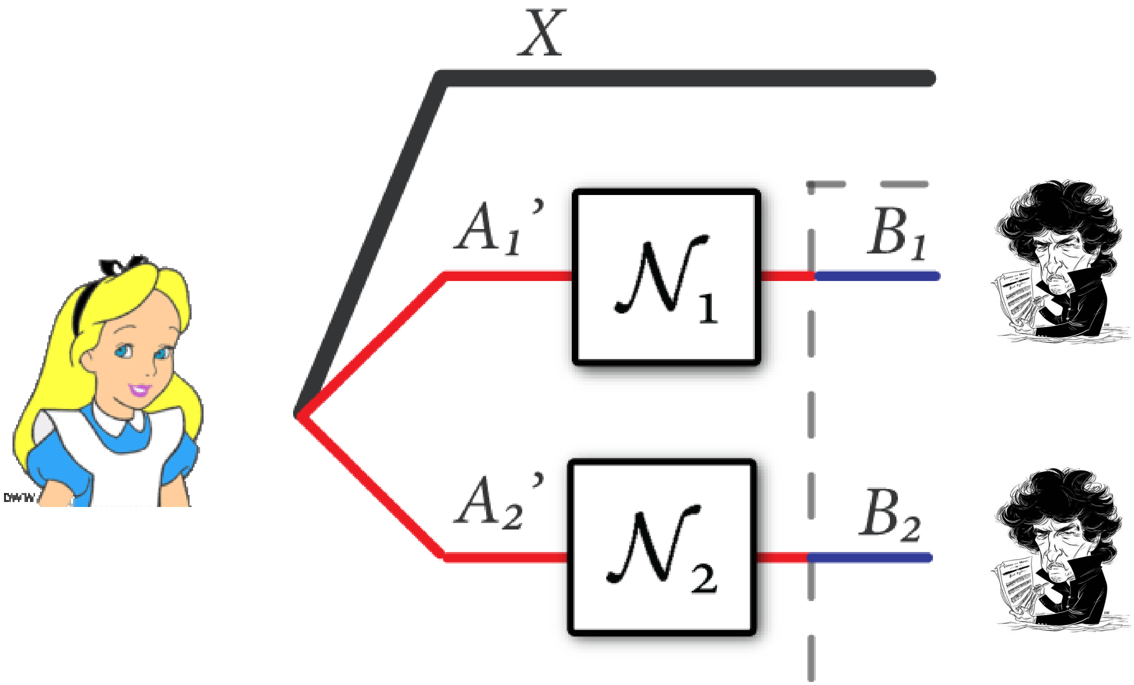
$$\chi(\mathcal{N}_1 \otimes \mathcal{N}_2) \geq \chi(\mathcal{N}_1) + \chi(\mathcal{N}_2)$$

Can choose ensemble on **LHS** to be a **tensor product** of the ones that individually maximize **RHS**

Additivity of Holevo Information?

Does **"hard direction"** hold?

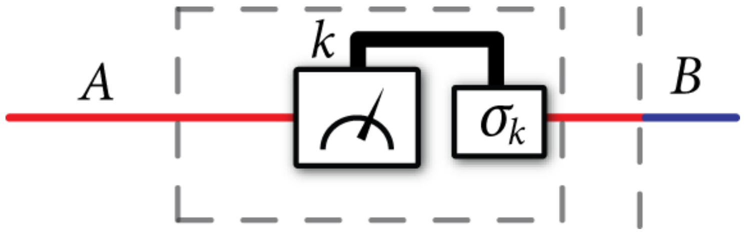
$$\chi(\mathcal{N}_1 \otimes \mathcal{N}_2) \leq \chi(\mathcal{N}_1) + \chi(\mathcal{N}_2)$$



If **true** for a given channel,
then entanglement **does not boost** information throughput
according to the Holevo measure

Simplest Example for Holevo Additivity

Suppose one channel is **entanglement-breaking**:



Then additivity holds:

$$\chi(\mathcal{N}_1 \otimes \mathcal{N}_2) = \chi(\mathcal{N}_1) + \chi(\mathcal{N}_2)$$

Proof: State on Bob's systems is **separable**

$$\sum_y p_{Y|X}(y|x)\rho_{x,y}^{B_1} \otimes \sigma_{x,y}^{B_2}$$

Give classical variable Y to Alice and separable state becomes **product** when conditioned on Y

Shor, arXiv:quant-ph/0201149 (2002)

Random Counterexample to Holevo Additivity

Consider random unitary channels:

$$\mathcal{N}(\rho) \equiv \sum_i p_I(i) U_i \rho U_i^\dagger$$

$$\overline{\mathcal{N}}(\rho) \equiv \sum_i p_I(i) U_i^* \rho U_i^T$$

where unitaries selected according to Haar measure

Then **additivity fails** according to Hastings' probabilistic argument (*and Shor's equivalence of additivity conjectures*):

$$\chi(\mathcal{N} \otimes \overline{\mathcal{N}}) > \chi(\mathcal{N}) + \chi(\overline{\mathcal{N}})$$

Open problem to find **explicit counterexamples** to additivity *(rather than a random construction)*

Shor (2004), Hastings 2008, Hayden and Winter (2008)

What all this means...

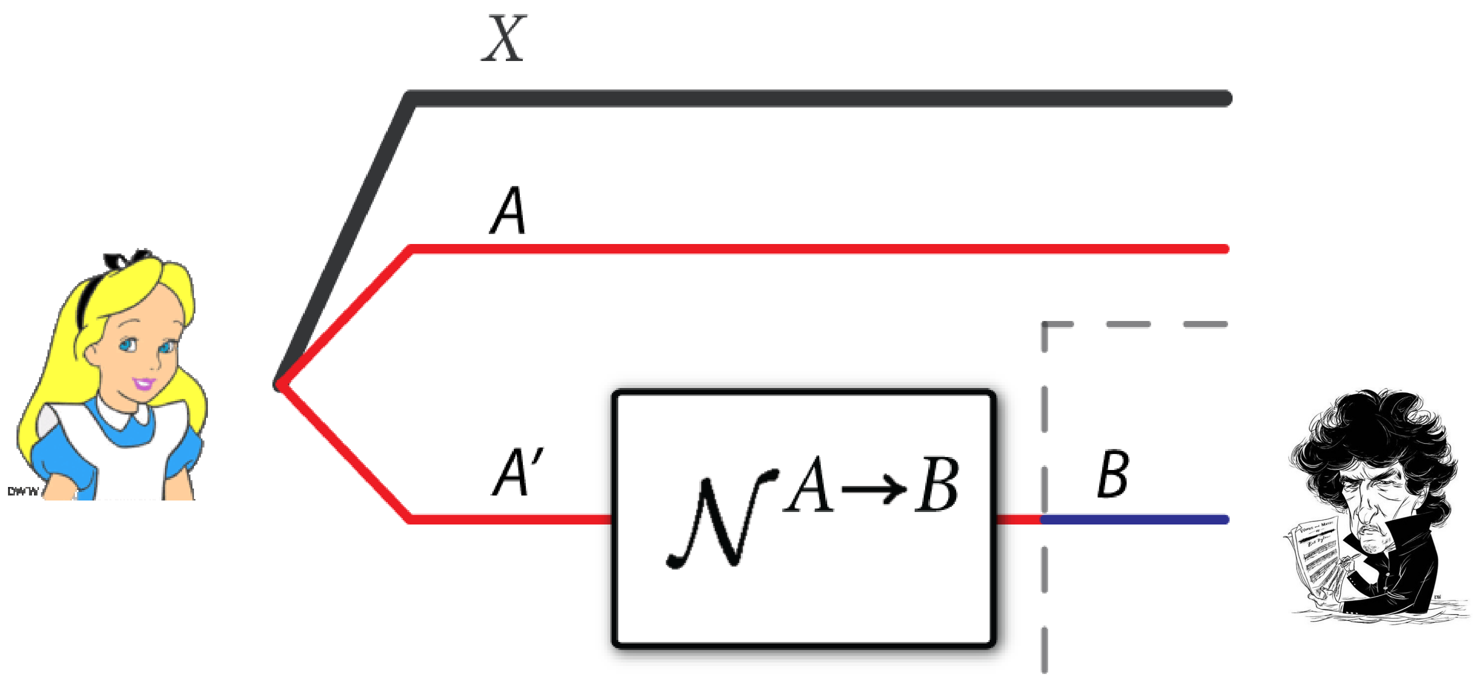
The **HSW formula** is **unsatisfactory** as a measure of a quantum channel's ability to transmit classical information

Regularization is necessary (for now):

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

Classical capacity **could still be additive**
(we just don't know the right formula)

Sending Classical Data over EA Quantum Channels



Correlate classical data with entangled quantum states:

$$\sum_x p_X(x)|x\rangle\langle x|^X \otimes \mathcal{N}^{A'\to B}(\phi_x^{AA'})$$

Mutual information of a quantum channel:

$$I(\mathcal{N}) \equiv \max_{\{p_X(x),\phi_x\}} I(AX;B)$$

Bennett et al. (2002), Shor (2004)

Additivity of Channel Mutual Information

First, can simplify expression for channel mutual info.

$$I(\mathcal{N}) \equiv \max_{\phi} I(A;B)$$

(follows from concavity of entropy and a few other arguments...)

Given two quantum channels,
does **additivity** of channel mutual information hold?

$$I(\mathcal{N}_1 \otimes \mathcal{N}_2) = I(\mathcal{N}_1) + I(\mathcal{N}_2)$$

"Easy direction" always holds:

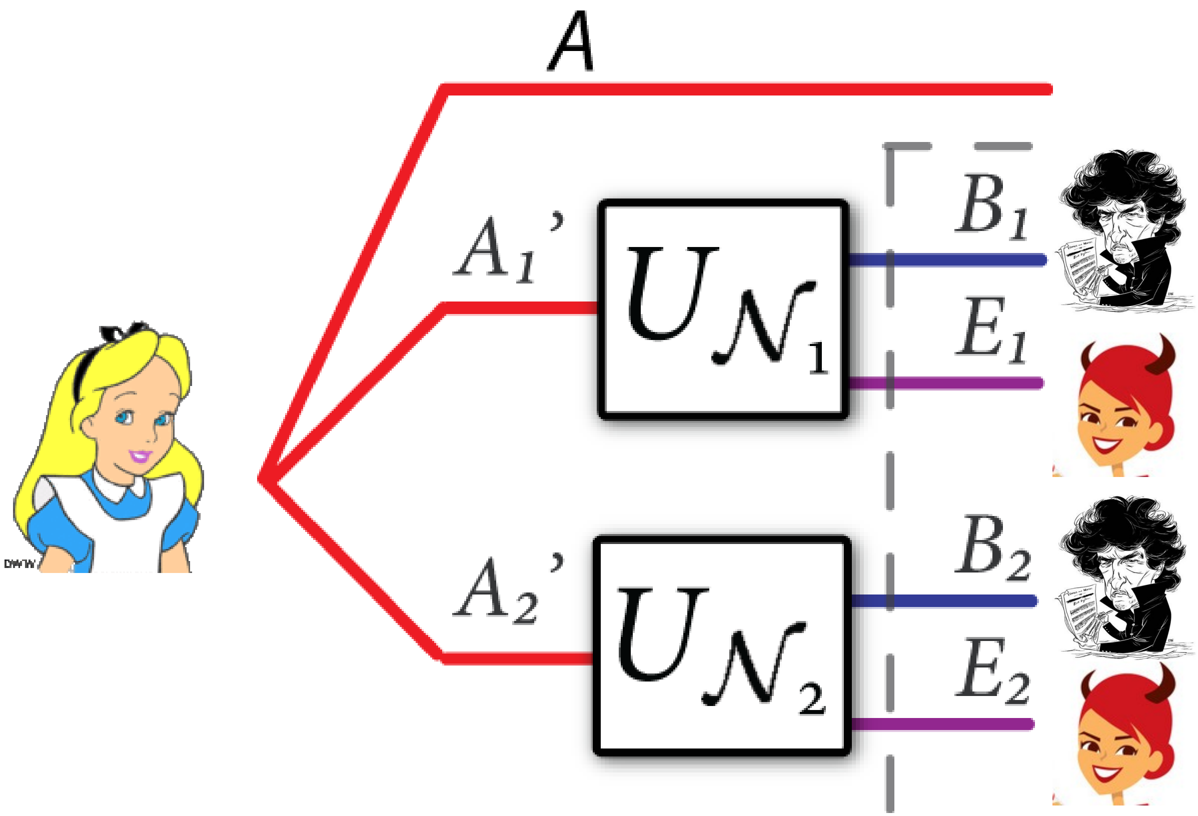
$$I(\mathcal{N}_1 \otimes \mathcal{N}_2) \geq I(\mathcal{N}_1) + I(\mathcal{N}_2)$$

Can choose ensemble on **LHS** to be a **tensor product**
of the ones that individually maximize **RHS**

Additivity of Channel Mutual Information

Does **"hard direction"** hold?

$$I(\mathcal{N}_1 \otimes \mathcal{N}_2) \leq I(\mathcal{N}_1) + I(\mathcal{N}_2)$$



Yes!

(follows from "one part subadditivity" and "three parts strong subadditivity")

$$\begin{aligned} I(A;B_1B_2) &= H(B_1B_2) + H(B_1B_2|E_1E_2) \\ &\leq H(B_1) + H(B_1|E_1) + H(B_2) + H(B_2|E_2) \\ &= I(AA_2';B_1) + I(AA_1';B_2) \end{aligned}$$

Additivity of Channel Mutual Information

Additivity of quantum channel mutual information holds for all quantum channels!

$$I(\mathcal{N}_1 \otimes \mathcal{N}_2) = I(\mathcal{N}_1) + I(\mathcal{N}_2)$$

By **induction**, it holds that

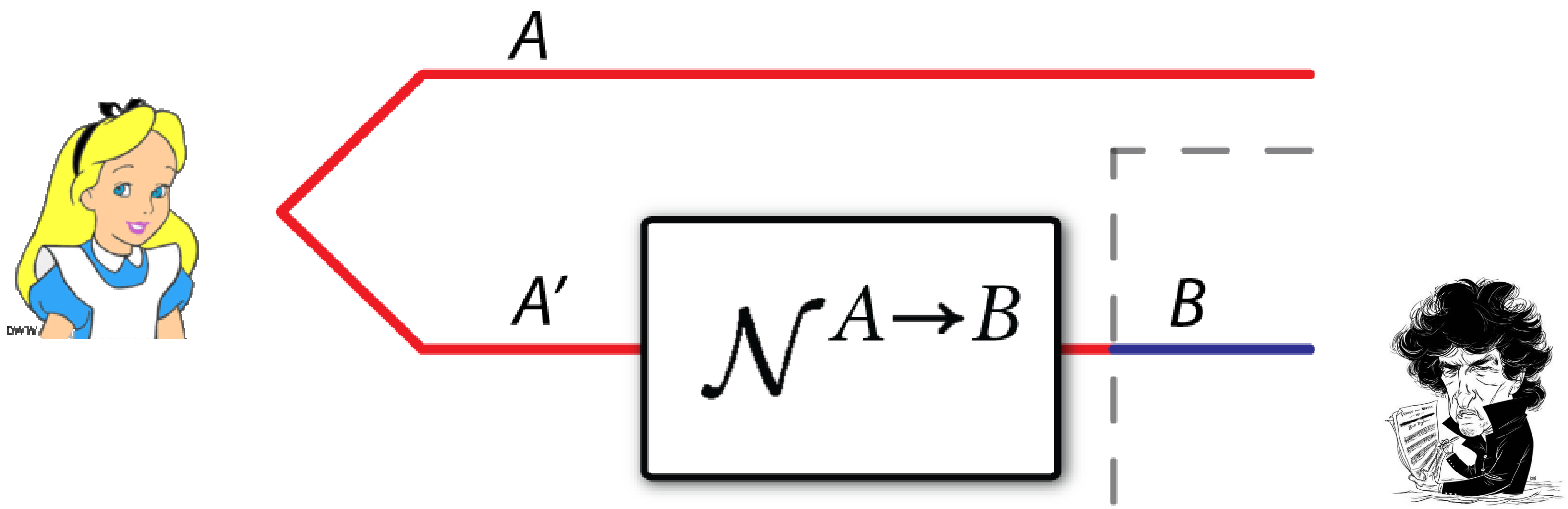
$$I_{\text{reg}}(\mathcal{N}) = I(\mathcal{N})$$

(No need for regularization)

Implies a **complete understanding** of the transmission capabilities of a quantum channel assisted with unlimited entanglement

Hayden's Musing:
What's so special about **entanglement assistance**?
It makes quantum Shannon theory
and quantum coding theory both "look" classical
(c.f., talk of Min-Hsiu Hsieh)

Sending Quantum Data over Quantum Channels



Preserving entanglement is the same as transmitting quantum data

$$\mathcal{N}^{A'\to B}(\phi^{AA'})$$

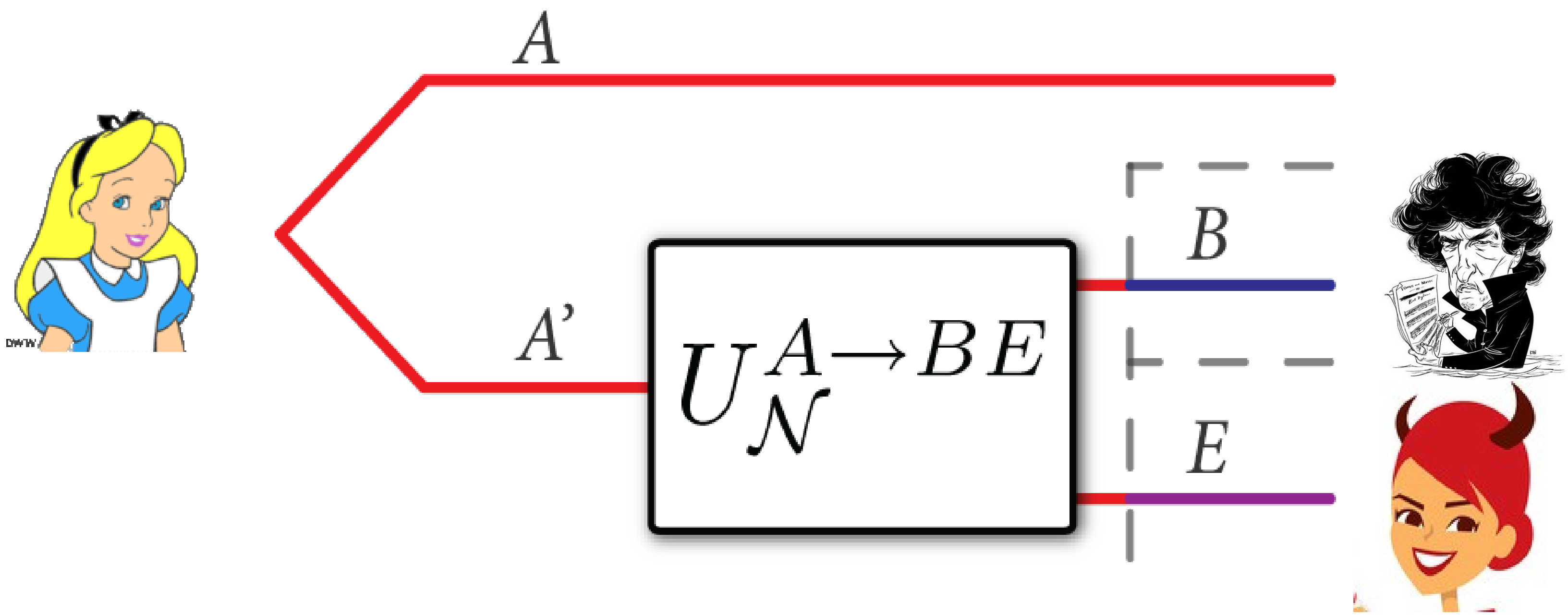
Coherent information of a quantum channel:

$$Q(\mathcal{N}) \equiv \max_{\phi} I(A\rangle B)$$

where $$I(A\rangle B) \equiv H(B) - H(AB)$$

Lloyd (1997), Shor (2002), Devetak (2005)

A Useful Alternate Viewpoint



Coherent information of a quantum channel:

$$Q(\mathcal{N}) \equiv \max_{\phi} H(B) - H(E)$$

Qualitatively “looks like” classical wiretap setting

Devetak (2005)

Additivity of Channel Coherent Information

Given two quantum channels,
does **additivity** of channel coherent information hold?

$$Q(\mathcal{N}_1 \otimes \mathcal{N}_2) = Q(\mathcal{N}_1) + Q(\mathcal{N}_2)$$

"Easy direction" always holds:

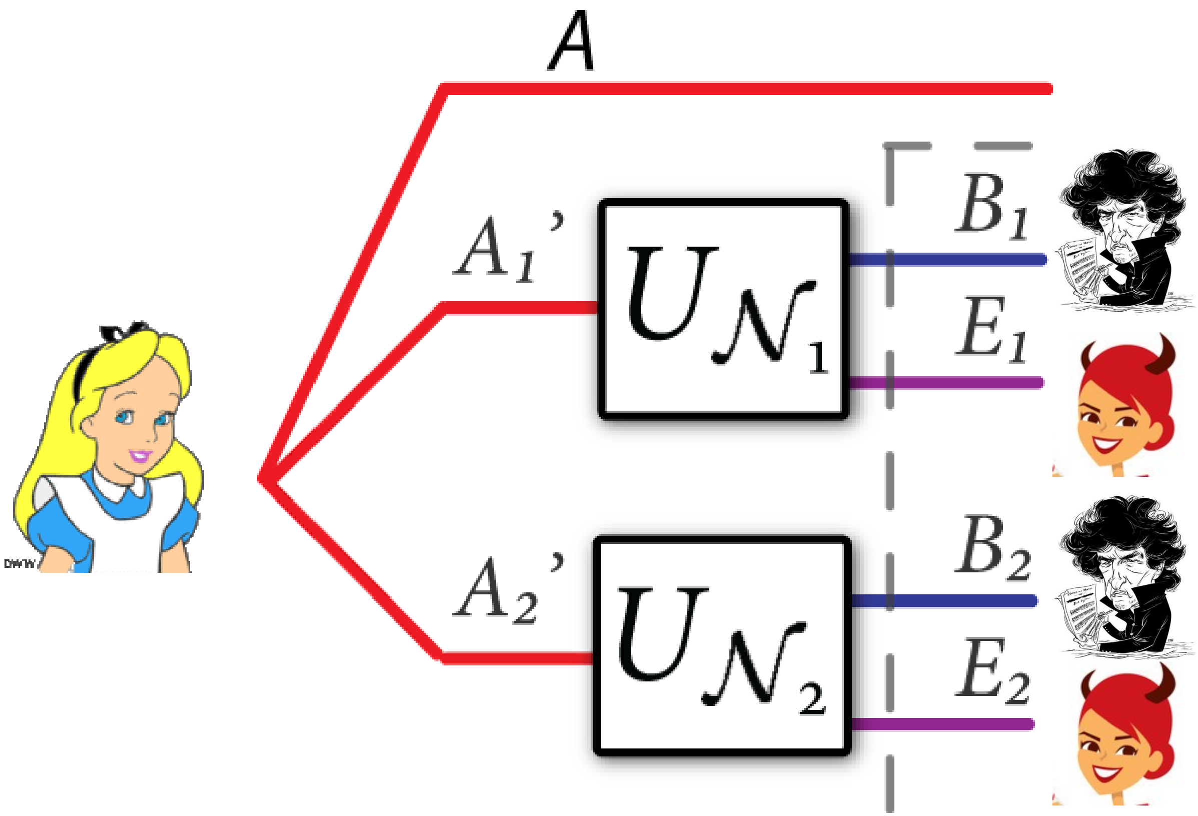
$$Q(\mathcal{N}_1 \otimes \mathcal{N}_2) \geq Q(\mathcal{N}_1) + Q(\mathcal{N}_2)$$

Can choose ensemble on **LHS** to be a **tensor product** of the ones that individually maximize **RHS**

Additivity of Channel Coherent Information

Does **"hard direction"** hold?

$$Q(\mathcal{N}_1 \otimes \mathcal{N}_2) \leq Q(\mathcal{N}_1) + Q(\mathcal{N}_2)$$



Not Always!

But does if

$$I(B_1;B_2) \geq I(E_1;E_2)$$

(holds for degradable channels)

$$\begin{aligned}
I(A\rangle B_1 B_2) &= H(B_1 B_2) - H(E_1 E_2) \\
&= H(B_1) - H(E_1) + H(B_2) - H(E_2) - [I(B_1;B_2) - I(E_1;E_2)] \\
&\leq H(B_1) - H(E_1) + H(B_2) - H(E_2) \\
&= I(AA_2'\rangle B_1) + I(AA_1'\rangle B_2)
\end{aligned}$$

Counterexample to Coherent Info. Additivity

Noisy quantum channel is the depolarizing channel
(*lets the qubit through or replaces it with the maximally mixed state*)

$$\mathcal{N}(\rho) = (1-p)\rho + p\frac{I}{2}$$

Concatenating a random code with a five-qubit repitition code outperforms a random code

Implies that $Q(\mathcal{N}^{\otimes 5}) > 5Q(\mathcal{N})$

Technique essentially exploits that we don't need to correct all quantum errors (degeneracy of quantum codes)

The **LSD formula** is **unsatisfactory** as a measure of a quantum channel's ability to transmit quantum information

DiVincenzo, Shor, Smolin (1996)

Even More Suprising...

Quantum capacity itself cannot be an additive function on two different quantum channels

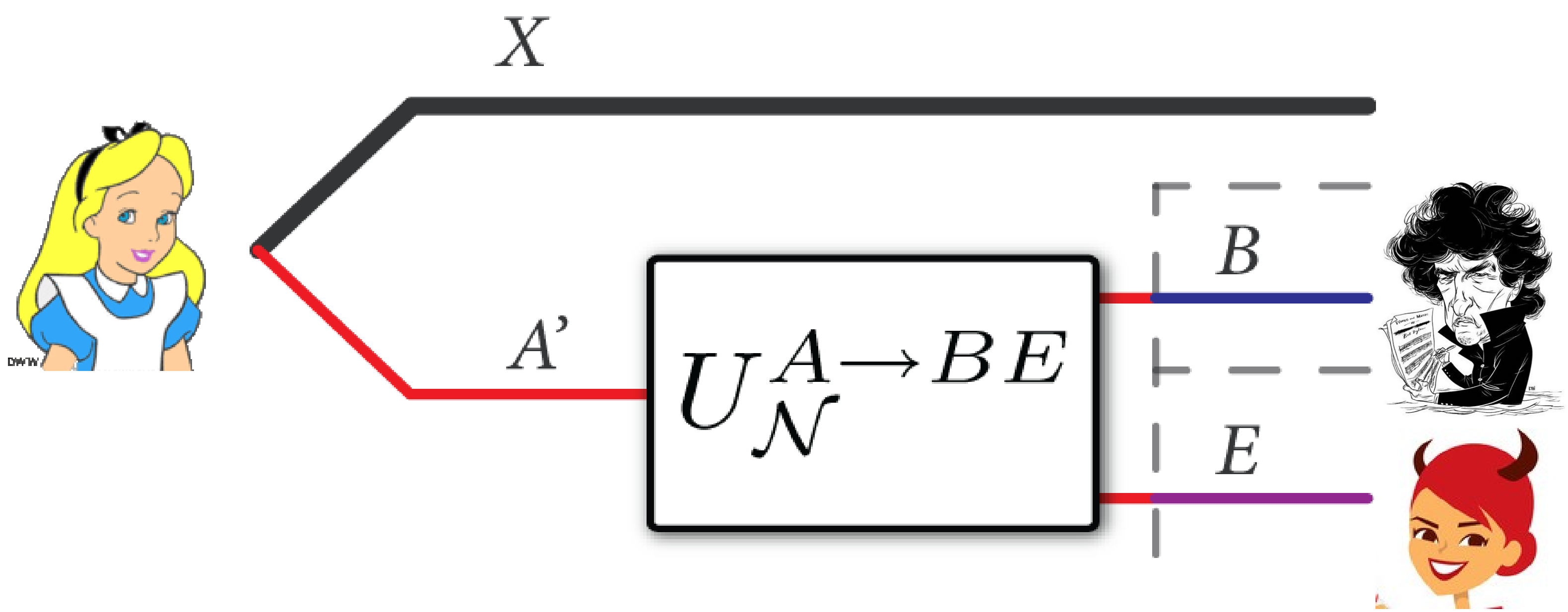
Horodecki channel with Zero Quantum Capacity (*can only create bound entangled states*)

50% erasure channel with Zero Quantum Capacity (*by the no-cloning theorem*)

But the joint channel has Nonzero Quantum Capacity!

Smith and Yard (2008)

Sending Private Data over Quantum Channels



Correlate classical data with channel input

$$\sum_x p_X(x)|x\rangle\langle x|^X \otimes U_{\mathcal{N}}^{A'\to BE}(\rho_x^{A'})$$

Private information of a quantum channel:

$$P(\mathcal{N}) \equiv \max_{\{p_X(x),\rho_x\}} I(X;B) - I(X;E)$$

Devetak (2005), Cai, Winter, Yeung (2004)

Additivity of Channel Private Information

Additivity **does not always** hold,
But **does** for the class of **degradable** channels
(*proof similar to quantum case but slightly different*)

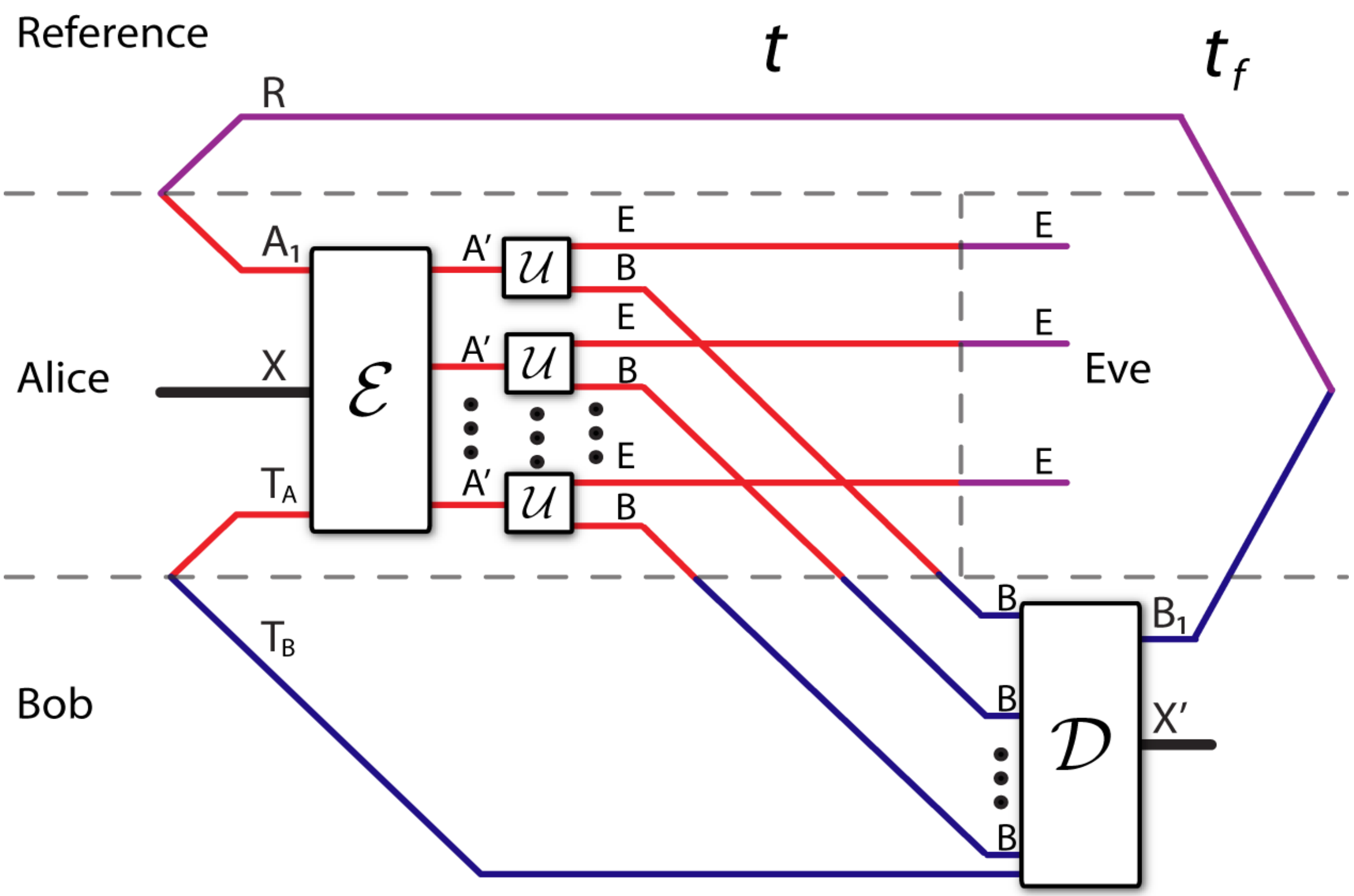
In fact, **quantum capacity** is the same as **private capacity** for the class of degradable channels

In general, the private information is **unsatisfactory** as a formula to characterize private information transmission
(*does not give a tractable optimization problem*)

Smith (2007), Smith, Renes, Smolin (2006)

Trade-off Coding

Suppose Alice wants to send classical and quantum data
With the help of shared entanglement
(*generalizes many of the above settings*)



Hsieh and Wilde (2009)

Trade-off Coding (Ctd.)

Let ***C*** be classical data rate,
Q quantum data rate, and
E entanglement consumption rate.

Three-dimensional capacity region is union of

$$\begin{aligned} C + 2Q &\leq I(AX;B) \\ Q &\leq I(A\rangle BX) + E \\ C + Q &\leq I(X;B) + I(A\rangle BX) + E \end{aligned}$$

over all states of the form:

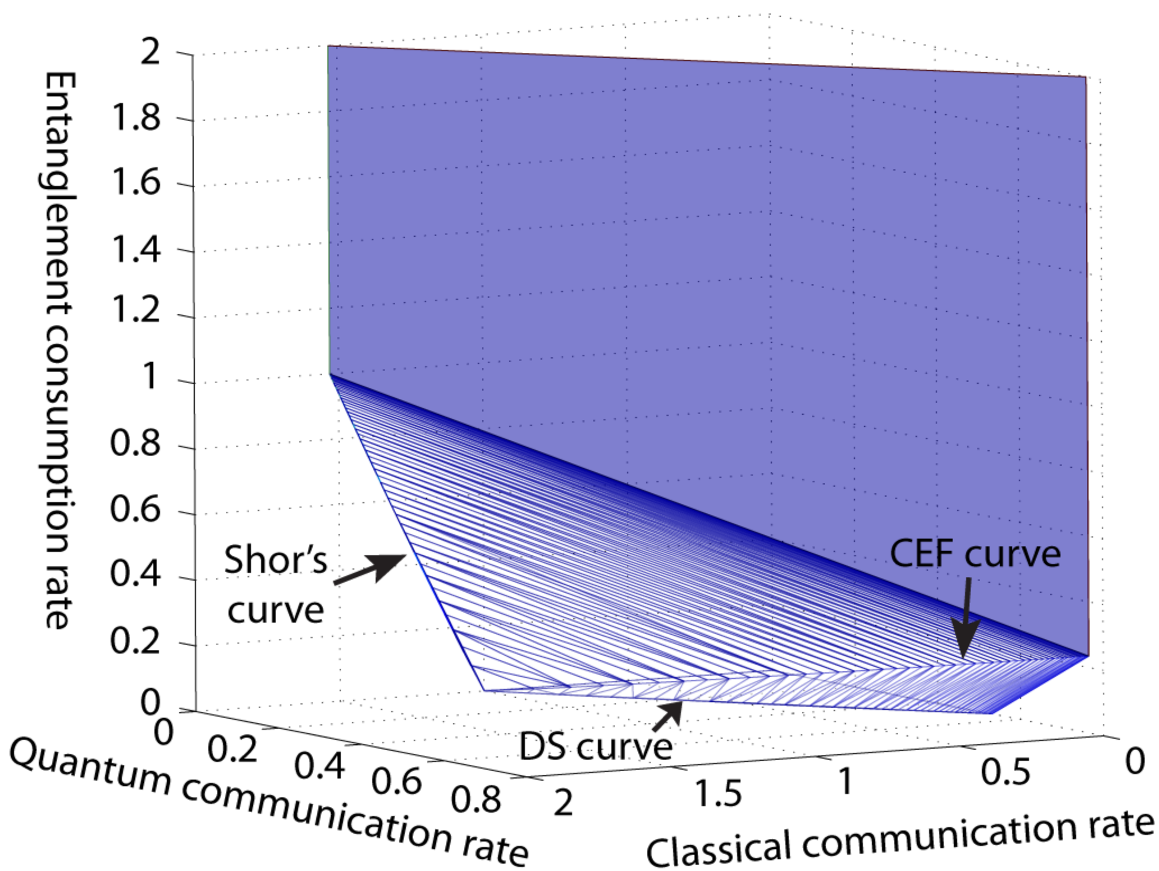
$$\sum_x p_X(x) |x\rangle\langle x|^X \otimes \mathcal{N}^{A' \to B}(\phi_x^{AA'})$$

Hsieh and Wilde (2009)

Trade-off Coding (Ctd.)

Full region is additive for the class of "Hadamard" channels
(*channels whose complements are entanglement-breaking*)

Means that we can actually plot it!



Hsieh and Wilde (2009), Bradler, Hayden, Touchette, Wilde (2010)

Conclusion

Additivity is at the heart of our understanding of classical information theory

Additivity does not hold in many cases for quantum channels (*but does for entanglement-assisted capacities*)

Open problem: Find a better formula for the classical capacity

Open problem: Find explicit counterexample to Holevo additivity

Open problem: Determine if the classical capacity is an additive function on quantum channels

Open problem: Find a better formula for the quantum capacity

Open problem: Find a better characterization for the triple trade-off capacity region other than the multi-letter one