Quantum Channel Capacities

Peter Shor
MIT
Cambridge, MA

Purdue University
April 8, 2019

Outline

- Brief intro to quantum mechanics
- Overview of quantum information theory
- Some results on feedback channels

Shannon's theorem (1948)

The entropy of a random variable X is

$$H(X) = \sum_i -p_i \log p_i \, .$$

Channel Coding

A noisy channel $N : X \to Y$ has capacity

$$\max_{p(X)} \quad I(X;Y),$$

where

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y|X) \\ &\quad [H(\text{output}) - H(\text{output given input})] \\ &= H(X) + H(Y) - H(X,Y) \\ &\quad [(\text{output}) + H(\text{input}) - H(\text{joint distribution})] \end{aligned}$$

How Do We Prove Shannon's Theorem?

- Use a random codebook.
- Show that we can (theoretically) decode it

Shannon's construction was not practical, since it takes exponentially long to decode.

John Pierce, 1973

I think that I have never met a physicist who understood information theory. I wish that physicists would stop talking about reformulating information theory and would give us a general expression for the capacity of a channel with quantum effects taken into account rather than a number of special cases.

Quantum Mechanics: The Superposition Principle:

If a quantum system can be in one of two mutually distinguishable states $|\,\mathbf{A}\rangle$ and $|\,\mathbf{B}\rangle$, it can be both these states at once. Namely, it can be in the *superposition* of states

$$\alpha\,|\,\mathbf{A}\rangle + \beta\,|\,\mathbf{B}\rangle$$

where α and β are both complex numbers and $|\alpha|^2 + |\beta|^2 = 1$.

If you look at the system, the chance of seeing it in state $\mathbf{A}$ is $|\alpha|^2$ and in state $\mathbf{B}$ is $|\beta|^2$.

The state of a quantum system is a unit vector in a complex vector space.

We call a two-dimensional quantum system a *qubit*.

Example: If you have a polarized photon, there can only be two distinguishable states, for example, vertical $|\updownarrow\rangle$ and horizontal $|\leftrightarrow\rangle$ polarizations.

All other states can be made from these two.

$$|\nearrow\rangle = \frac{1}{\sqrt{2}}|\leftrightarrow\rangle + \frac{1}{\sqrt{2}}|\updownarrow\rangle$$

$$|\nwarrow\searrow\rangle = \frac{1}{\sqrt{2}}|\leftrightarrow\rangle - \frac{1}{\sqrt{2}}|\updownarrow\rangle$$

$$\left|\circlearrowleft\right\rangle = \frac{1}{\sqrt{2}}|\leftrightarrow\rangle - \frac{i}{\sqrt{2}}|\updownarrow\rangle$$

Quantum Mechanics: Joint Systems:

If you have two qubits, their joint state space is the tensor product of their individual state spaces (e.g., $\mathbb{C}^4$).

Two qubits can be in any superposition of the four states

$$|\updownarrow\updownarrow\rangle \quad |\updownarrow\leftrightarrow\rangle \quad |\leftrightarrow\updownarrow\rangle \quad |\leftrightarrow\leftrightarrow\rangle$$

This includes states such as an EPR pair of photons,

$$\frac{1}{\sqrt{2}}(|\updownarrow\leftrightarrow\rangle - |\leftrightarrow\updownarrow\rangle) = \frac{1}{\sqrt{2}}(|\swarrow\nwarrow\rangle - |\nwarrow\swarrow\rangle),$$

where neither qubit alone has a definite state. These are called *entangled* states.

Quantum Mechanics: Joint Systems:

If you have n qubits, their joint state is described by a 2^n dimensional vector. We now label basis vectors for each qubit by $|0\rangle$ and $|1\rangle$.

The basis states of this vector space are:

$$|000\ldots00\rangle \quad |000\ldots01\rangle \quad \cdots \quad |111\ldots11\rangle$$

This high dimensional tensor product space is where quantum information theory (as well as quantum computation) lives.

Density Matrices

In quantum mechanics, the fundamental objects are often taken to be pure quantum states (unit vectors in $\mathbb{C}^n$).

These are analogous to deterministic objects in classical systems.

For quantum information theory, we need to work with probabilistic ensembles of quantum states. These are represented by *density matrices.*

Density matrix ρ:
Hermitian trace 1 positive semi-definite matrix over $\mathbb{C}^n \times \mathbb{C}^n$.

A rank one density matrix corresponds to a pure state (i.e., vector in $\mathbb{C}^n$).

Density Matrices

Suppose a quantum system is in state v_i with probability p_i.

The density matrix is

$$\rho = \sum_i p_i v_i v_i^\dagger$$

The density matrix ρ gives as much information as possible about the outcomes of experiments performed on the system.

ρ is trace 1, positive semi-definite.

Two systems with the same density matrix ρ are experimentally indistinguisable.

Density Matrices II

Suppose you have a joint quantum system on $\mathbb{C}^a \otimes \mathbb{C}^b$ in the state ρ_{AB}. If you can only do experiments on the second part of the system, it is effectively in the state

$$\rho_B = \mathrm{Tr}_A \ \rho_{AB}$$

Here, Tr_A is the partial trace over the first quantum space. If we have a tensor product state $\rho_A \otimes \rho_B$, then

$$\mathrm{Tr}_A (\rho_A \otimes \rho_B) = (\mathrm{Tr}\rho_A)\rho_B,$$

and we extend this linearly to define the partial trace on entangled states.

Entropy of quantum states

Classical Case

Given n photons, each in state $|\updownarrow\rangle$ or $|\leftrightarrow\rangle$, with probability $\frac{1}{2}$. Any two of these states are completely distinguishable. The entropy is n bits.

Quantum Case

Given n photons, each in state $|\updownarrow\rangle$ or $|\swarrow\!\!\nearrow\rangle$, with probability $\frac{1}{2}$. If the angle between the polarizations is small, any two of these states are barely distinguishable. Intuitively, the entropy should be much less than n bits.

Entropy of density matrices

By thermodynamic arguments, von Neumann deduced the entropy of a quantum system with density matrix ρ as the Shannon entropy of the eigenvalues λ_i of ρ:

$$H_{\mathsf{vN}}(\rho) = H_{\mathsf{Shan}}(\lambda_i).$$

Equivalently,

$$H_{\mathsf{vN}}(\rho) = -\mathsf{Tr}(\rho \log \rho)\,.$$

(Recall $\sum_i \lambda_i = \mathsf{Tr}\rho = 1$.)

Von Neumann Measurements

Suppose you have a quantum state space $\mathbb{C}^n$. A von Neumann measurement corresponds to a complete set of orthogonal subspaces S_1, S_2, ... S_k. (Complete means the S_i span $\mathbb{C}^n$.)

Let Π_{S_i} be the projection onto the i'th subspace S_i.

The corresponding von Neumann measurement operating on the density matrix $\rho \in \mathbb{C}^n \times \mathbb{C}^n$ takes ρ to $\Pi_{S_i} \rho \Pi_{S_i}$ with probability $\mathrm{Tr}(\Pi_{S_i} \rho)$.

The simplest situation is if each of the S_i is one-dimensional, and then $S_i = v_i v_i^\dagger$, where the v_i form a basis of $\mathbb{C}^n$.

Quantum Shannon's Theorem: Holevo Bound χ

Suppose we have a source emitting ρ_i with probability p_i.

$$\chi = H_{\mathsf{vN}}(\sum_i p_i \rho_i) - \sum_i p_i H_{\mathsf{vN}}(\rho_i)$$

How much information I_{acc} (*accessible information*) can we learn about the sequence $\{i\}$? Theorem (Holevo, 1973)

$$I_{\mathsf{acc}} \leq \chi$$

If all the ρ_i commute, the situation is essentially classical, and we get $I_{\mathsf{acc}} = \chi$. Otherwise $I_{\mathsf{acc}} < \chi$.

Theorem (Holevo, Schumacher-Westmoreland, 1996)

The classical-information capacity obtainable using codewords composed of signal states ρ_i, where ρ_i has marginal probability p_i, is

$$\chi(\{\rho_i\};\{p_i\}) = H_{\mathsf{vN}}(\sum_i p_i\rho_i) - \sum_i p_i H_{\mathsf{vN}}(\rho_i)$$

(The entropy of the average output less the average entropy of the output.)

This can be larger than the accessible information.

Quantum Channels

So far what we've dealt with is a sender able to send one out of some set of quantum states.

To ask for the quantum analog of Shannon's theorem, we need to talk about quantum channels.

Input: quantum state $\longrightarrow$ output: quantum state.

Quantum mechanics says that a memoryless quantum channel is a trace-preserving completely positive operator.

If you input a state entangled across two channel inputs, you get an output possibly entangled across two channel outputs.

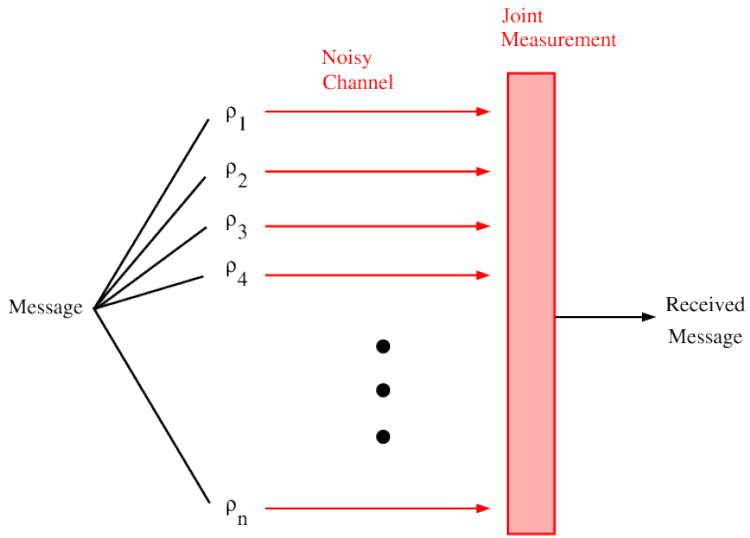
Capacity of Quantum Channels

Does the HSW theorem give the capacity of a quantum channel $\mathcal{N}$?

Possible capacity formula: Maximize $\chi(\{\mathcal{N}(\rho_i)\}; \{p_i\})$ over all output states $\mathcal{N}(\rho)$ of the channel.

This is called $\chi(\mathcal{N})$.

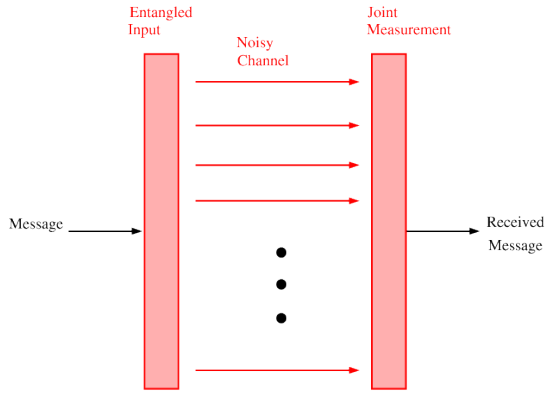
Unentangled Inputs, Joint Measurements



HSW Capacity: Maximize over probability distributions on inputs to the channel ρ_i, p_i:

$$\chi(\{\mathcal{N}(\rho_i)\};\{p_i\})$$

Entangled Inputs, Joint Measurements



Capacity: Maximize over probability distributions on inputs to the channel ρ_i, p_i where ρ_i is in the tensor product space of n inputs:

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

Can be larger.

Measurement of Entanglement

Suppose we have a pure state on systems A and B: $|\psi\rangle_{AB}$.

Its entropy of entanglement is

$$S(\mathrm{Tr}_A |\psi\rangle_{AB}) = S(\mathrm{Tr}_B |\psi\rangle_{AB}).$$

E.g., the state

$$\frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$$

has one bit of entanglement.

An n-qubit state is said to be maximally entangled if it has n bits of entanglement.

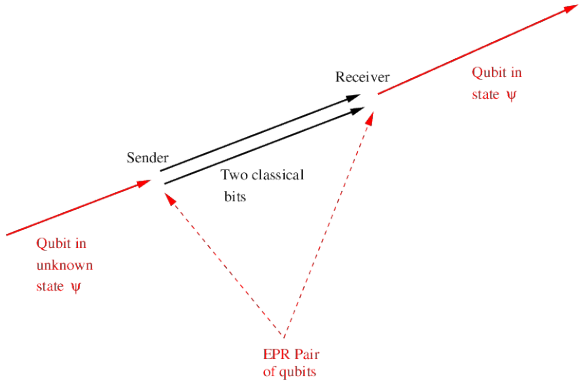
Monogamy of Entanglement

If an n-qubit system A has n bits of entanglement with another system B, then A can't be entangled with anything but B.

And in general, the more an n-qubit system is entangled with one system, the less it can be entangled with any other system.

Teleportation:

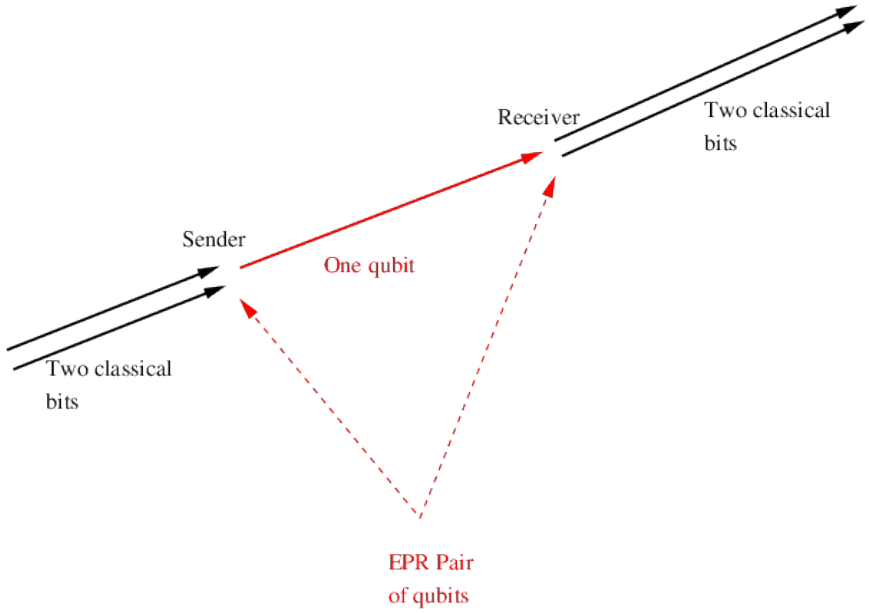
Using an EPR pair of qubits — a maximally entangled pair of qubits — one qubit can be teleported by sending two classical bits.



The sender makes a measurement on the joint state of the unknown qubit and her half of the EPR pair, and the receiver applies to his half of the EPR pair a unitary transformation which depends on the result of the measurement.

Superdense Coding:

Using an EPR pair of qubits, two classical bits can be encoded in one quantum bit.



This is a converse process to teleportation. Now the sender applies the transformation and the receiver makes the measurement.

What does this mean for channel capacity?

Holevo's bound says that a noiseless qubit channel can send at most one bit per signal.

If the sender and the receiver share entangled pairs of qubits, a noiseless qubit channel can send *two* bits per signal.

There is an entanglement-assisted capacity of a quantum channel, and it is larger than the classical capacity of a quantum channel.

Generalizations of Shannon's Formula: entanglement-assisted capacity

Essentially: C_E is the entropy of the input plus the entropy of the output less their joint entropy.

For classical channels, this is the same as the other formula for capacity. For quantum channels, they are different.

The above formulation isn't technically right. There's no such thing as a joint entropy since the input and the output of a quantum channel can never simultaneously exist. We can fix this detail.

Formula for entanglement-assisted capacity

Essentially: C_E is the Entropy of the input plus the entropy of the output less their joint entropy.

The above formulation isn't quite right.

We get around this by using an entangled state as input. The sender keeps half of it, and sends the other half through the channel. Now, we can define an input entropy (entropy of sender's half of the entangled state), an output entropy (entropy of receiver's state), and a joint entropy.

C_E is additive.

Quantum Capacity:

We can also ask for the *quantum capacity* of a channel.

Suppose we have a channel. Then if the *quantum capacity* is Q, we can use n uses of the channel to send $Qn - o(n)$ qubits from the sender to the receiver, and have the receiver decode the output and obtain a quantum state that is nearly the same as the sender encoded.

Quantum Capacity:

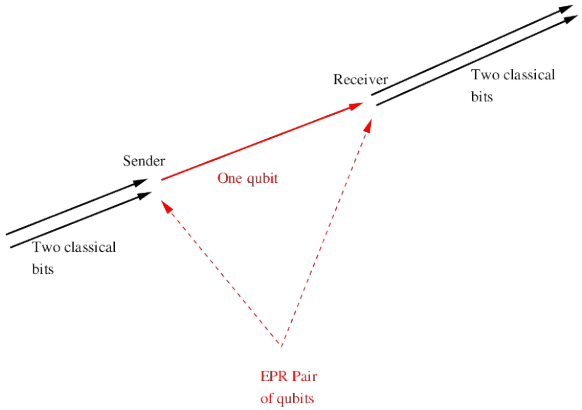
Like entanglement-assisted capacity, you assume the sender inputs an entangled state.

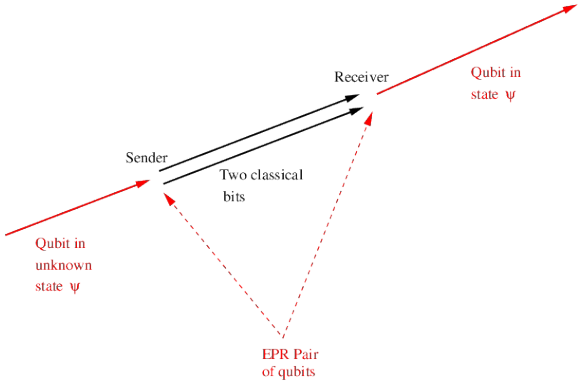
The quantum capacity is essentially the maximum of the entropy of the received state less the entropy of the joint state.

For classical channels (and many quantum channels, that aren't quantum enough), this is 0.

Entanglement-Assisted Quantum Capacity:

Superdense coding and teleportation shows that if the sender and receiver share entanglement, sending two classical bits is quivalent to sending one quantum bit.





Thus, $Q_E = \frac{1}{2} C_E$.

Feedback for Classical Channels

Feedback: the receiver can send information to the sender.

Theorem: Feedback cannot increase the capacity of a memoryless classical channel.

Feedback for Quantum Channels

Feedback can increase the quantum capacity of the erasure channel.

Erasure channel: with probability p, receiver is told the signal is erased, and with probability $1-p$, he gets the input.

Without feedback, a quantum erasure channel with error probability $\geq \frac{1}{2}$ cannot send quantum information.

We use the quantum no-cloning theorem to prove this.

Theorem: you cannot duplicate the state of an unknown qubit.

Feedback for Quantum Channels

Suppose we could encode quantum information in a channel with erasure probability $\geq \frac{1}{2}$.

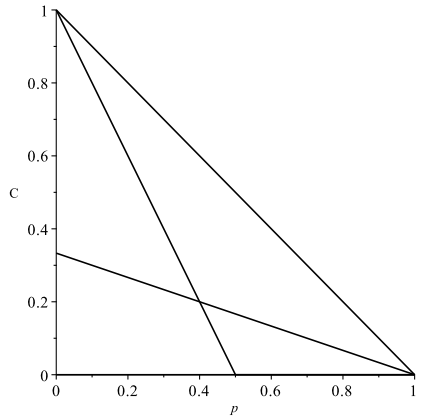
Consider the channel that with probability $\frac{1}{2}$, gives the input to receiver 1 and an erasure state to receiver 2. And with probability $\frac{1}{2}$, does the opposite.

This looks like an erasure channel with probability $\frac{1}{2}$ to each of the two receivers.

Thus, if we could send quantum information over a channel with erasure probability $p = \frac{1}{2}$, both receivers would be able to decode it, and you would have cloned the message qubits, a contradiction.

Capacity of Qubit Erasure Channel with Feedback

Formula for quantum capacity gives $C = 1 - 2p$.

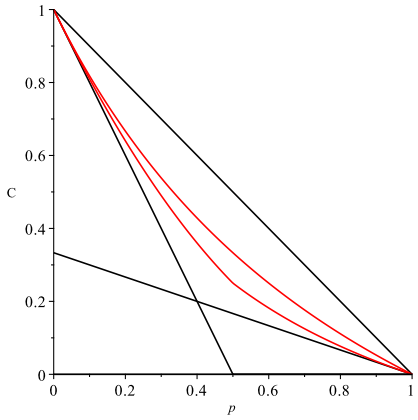


But with feedback, sender and receiver can share an entangled state with one channel use, and send two classical bits with two more channel uses. This protocol gives

$$C = \frac{1}{3}(1 - p)\,.$$

Capacity of Erasure Channel with Feedback

You can combine this with superdense coding and other tricks to do better. While we don't have an exact formula for the capacity, it is between the two red lines on the graph.
(D. Leung, J. Lim, P. Shor, 2007)



Can Feedback Increase Classical Capacity?

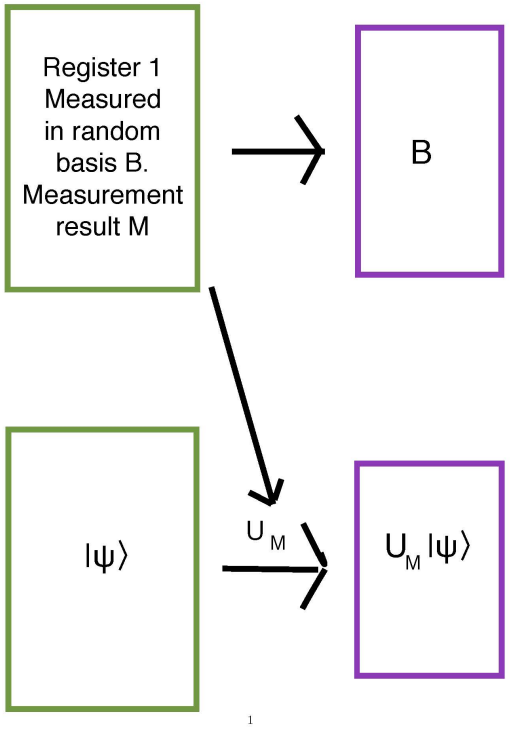
Yes, it can.

Retrocorrectible channels. (Bennett et al ????)

Rather contrived-looking channel with two registers.

Can Feedback Increase Classical Capacity?

The channel takes input $|\phi\rangle$ and $|\psi\rangle$, measures $|\phi\rangle$ in a random basis B, gets the result M, and gives the receiver B and $U_M|\psi\rangle$ Protocol: the sender puts half an entangled state in register 1, the receiver sends her the basis B, and she measures the other half of the entangled state in basis B to find M.



How about simple channels?

For example, can feedback increase the classical capacity of the noiseless qubit channel?

Until recently, this was unknown. (Although it's not clear how many people realized it was an open problem.)

It cannot.

How about simple channels?

For example, can feedback increase the classical capacity of the noiseless qubit channel?

Until recently, this was unknown. (Although it's not clear how many people realized it was an open problem.)

$$\boxed{M \text{ (message)}} : S \text{ (sender)} \longrightarrow R \text{ (receiver)}$$

Proof: We show that the quantity

$$I(M, R) + \text{entanglement}(S, R)$$

cannot increase by more than 1 for each channel transmission.

So after n transmissions $I(M, R) \leq n$, showing that $C \leq n$.

Simple Channels and Feedback (continued)

In fact, with a little more work the proof shows that the capacity with feedback cannot be larger than the maximum entropy of the output.

This shows feedback cannot increase the capacity of the:

- noiseless quantum channel,
- noiseless bosonic channel,
- quantum erasure channel.

Open Questions:

How general is the phenomenon that feedback increases classical capacity of quantum channels?
For example, does feedback increase the capacity of the noisy bosonic channel?

How general is the phenomenon that entangled inputs increase the classical capacity of quantum channels?
For example, does this happen for the depolarizing channel?