Towards a Quantum Reverse Shannon Theorem

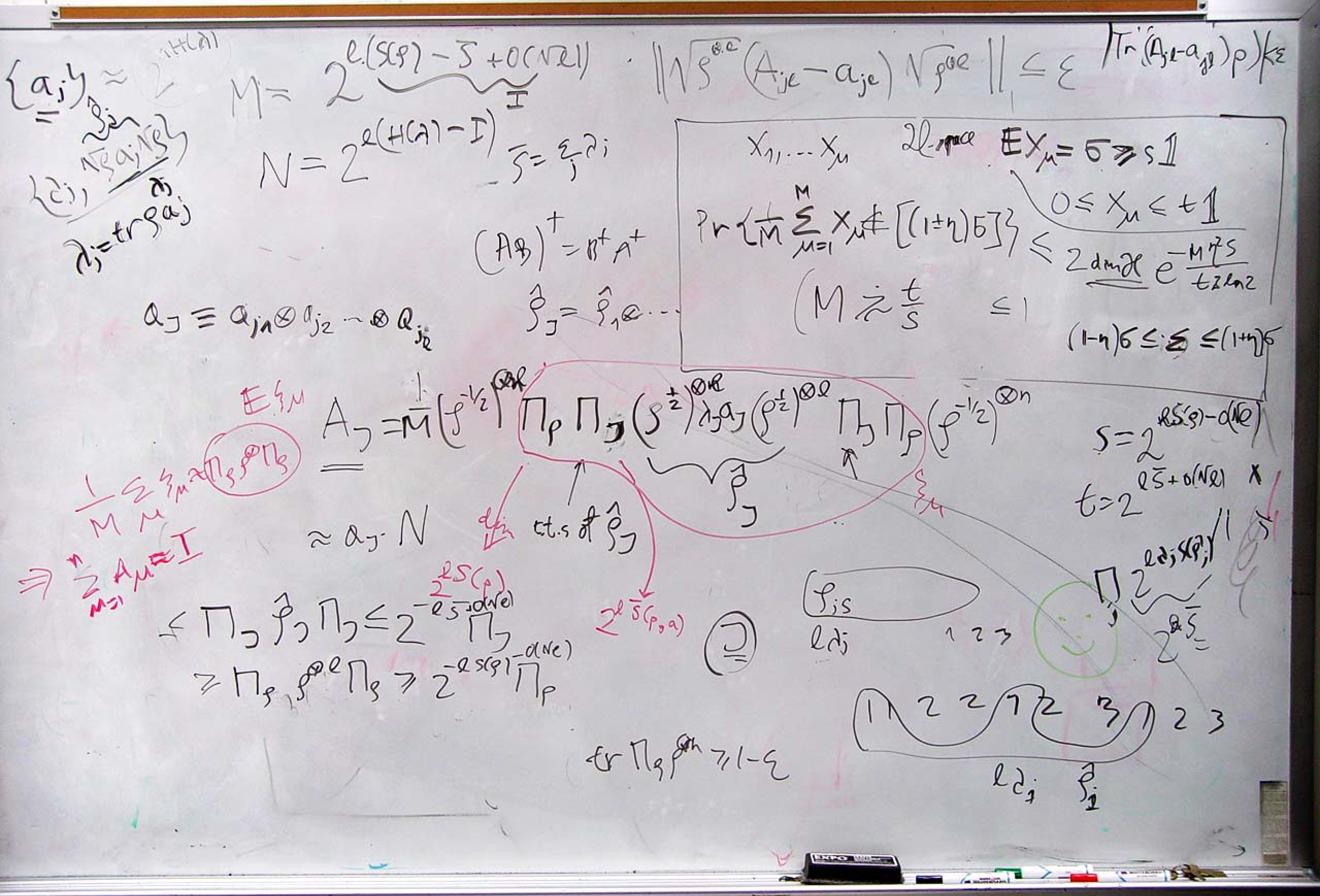
QIP “2003”
MSRI
18 Dec 02

NATIONAL SECURITY AGENCY
UNITED STATES OF AMERICA

SCIENCE TO SHAPE THE FUTURE OF THE ARMY
ARMY RESEARCH OFFICE
1951

ARDA

I. Devetak, A. Harrow,
D. Leung, P. Shor,
J. Smolin, A. Winter,
& CHB, heavily using measurement compression and RSP work described in Andreas’ and Debbie’s talks.



$\{a_j\} \approx 2^{nH(\lambda)}$

$\{\lambda_j, \sqrt{\rho} a_j \sqrt{\rho}\}$ ρ_j

$\lambda_j = \mathrm{tr}\, \rho a_j$

$M = 2^{\ell \cdot (\underbrace{S(\rho) - \bar{S} + O(\sqrt{\ell})}_{I})}$

$N = 2^{\ell (H(\lambda) - I)}$

$\bar{S} = \sum_j \lambda_i$

$\cdot \left\| \sqrt{\rho^{\otimes \ell}} (A_{j\ell} - a_{j\ell}) \sqrt{\rho^{\otimes \ell}} \right\|_1 \leq \varepsilon$

$|\mathrm{Tr}((A_{j\ell} - a_{j\ell})\rho)| < \varepsilon$

$(AB)^\dagger = B^\dagger A^\dagger$

$\hat{\rho}_J = \hat{\rho}_1 \otimes \cdots$

$a_J \equiv a_{j_1} \otimes a_{j_2} \cdots \otimes a_{j_\ell}$

$X_1, \ldots X_\mu$ $\mathcal{H}$-space $\mathbb{E} X_\mu = \sigma \geq s \mathbb{1}$

$0 \leq X_\mu \leq t \mathbb{1}$

$\Pr\left\{ \frac{1}{M} \sum_{\mu=1}^{M} X_\mu \notin [(1 \pm \eta)\sigma] \right\} \leq 2 \dim \mathcal{H}\, e^{-\frac{M \eta^2 s}{t \, 2 \ln 2}}$

$\left(M \gtrsim \frac{t}{s} \right.$ ≤ 1

$(1-\eta)\sigma \leq \Sigma \leq (1+\eta)\sigma$

$\mathbb{E}\, \xi_\mu$

$\underline{\underline{A_J}} = \frac{1}{M} (\rho^{-1/2})^{\otimes \ell}\, \Pi_\rho \Pi_J \underbrace{(\rho^{\frac{1}{2}})^{\otimes \ell} \lambda_J a_J (\rho^{\frac{1}{2}})^{\otimes \ell}}_{\hat{\rho}_J} \Pi_J \Pi_\rho (\rho^{-1/2})^{\otimes n}$

ξ_μ

e.t.s of $\hat{\rho}_J$

d_{J_1}

$\approx a_J \cdot N$

$\frac{1}{M} \sum_\mu \xi_\mu \approx \Pi_\rho \rho^{\otimes} \Pi_\rho$

$\Rightarrow \sum_{\mu=1}^{n} A_\mu \approx \underline{I}$

$2^{\ell S(\rho)}$

$\leq \Pi_J \hat{\rho}_J \Pi_J \leq 2^{-\ell \bar{S} - O(\sqrt{\ell})} \Pi_J$

$2^{\ell \bar{S}(\rho, a)}$

$\geq \Pi_\rho \rho^{\otimes \ell} \Pi_\rho \geq 2^{-\ell S(\rho) - O(\sqrt{\ell})} \Pi_\rho$

$\mathrm{tr}\, \Pi_\rho \rho^{\otimes n} \geq 1 - \varepsilon$

$s = 2^{\ell S(\rho) - O(\sqrt{\ell})}$

$t = 2^{\ell \bar{S} + O(\sqrt{\ell})}$ x

$\rho_{j,s}$

$\ell \lambda_j$

1 2 3

$\prod_j 2^{\ell \lambda_j S(\rho_j)}$

$2^{\ell \bar{S}}$

1 1 2 2 1 2 3 1 2 3

$\ell \lambda_1$ $\hat{\rho}_1$

- Multiple capacities of quantum channels: C,Q…

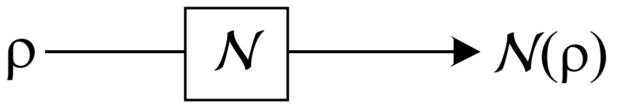
- Trying to simplifying things with free resources:
  - Assisted capacities: Q_2, C_E

- Entanglement assisted cap. C_E as *the* capacity.

- Classical Reverse Shannon Theorem

- Towards a Quantum Reverse Shannon Theorem

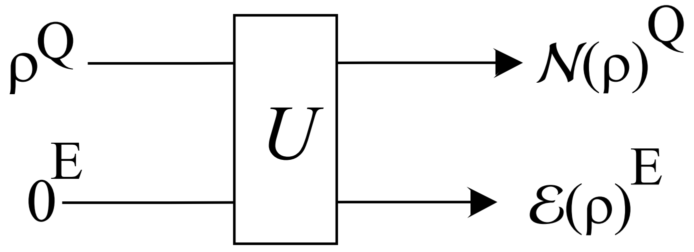
(Discrete Memoryless) Quantum Channel



$$\mathcal{N}(\rho) = \Sigma_k\ A_k\, \rho\, A_k^\dagger$$

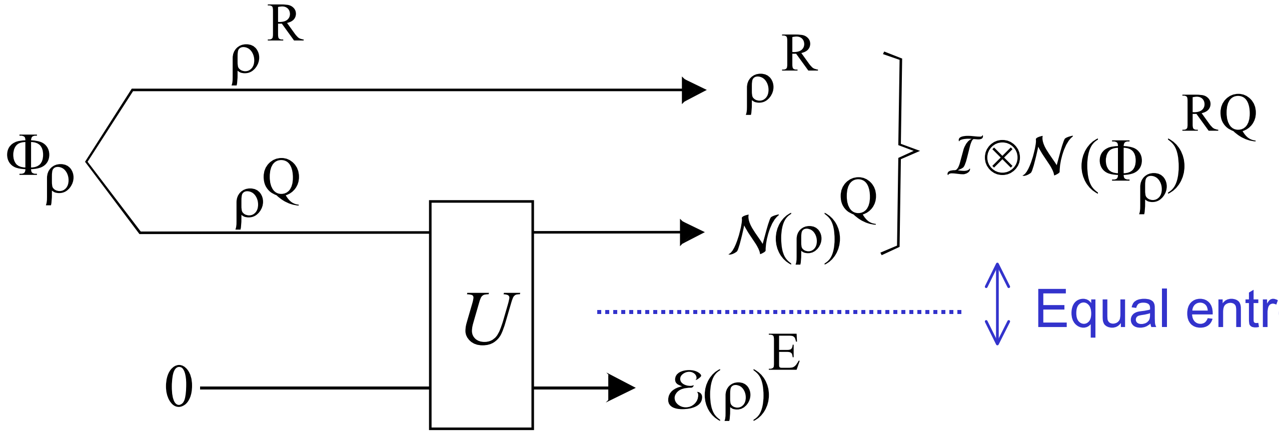
$$\Sigma_k A_k^\dagger A_k = 1$$

Kraus representation.



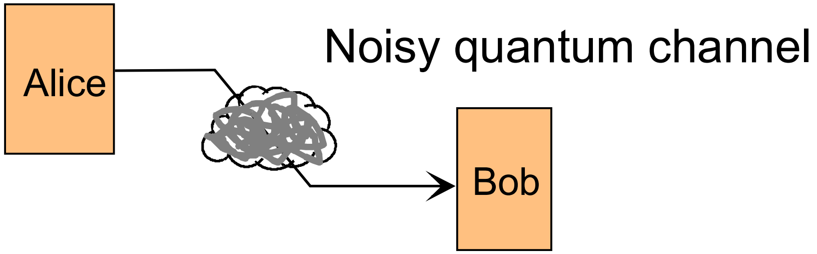
Unitary representation.

Input usually viewed as entangled with a reference system R



Equal entropy

Multiple Capacities of Quantum Channels



Q plain quantum capacity = qubits faithfully trasmitted per channel use, via quantum error correcting codes

C plain classical capacity = bits faithfully trasmitted per channel use

$\boldsymbol{Q_2}$ classically assisted quantum capacity, i.e. qubit capacity in the presence of unlimited 2-way classical communication, (e.g. using entanglement distillation and teleportation)

$\boldsymbol{C_E}$ entanglement assisted classical capacity i.e. bit capacity in the presence of unlimited prior entanglement between sender and receiver.

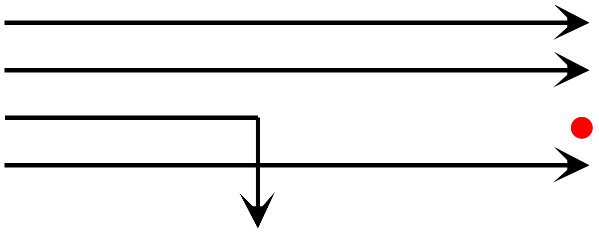
Inequalities among major capacities

All inequalities may be = or < depending on channel

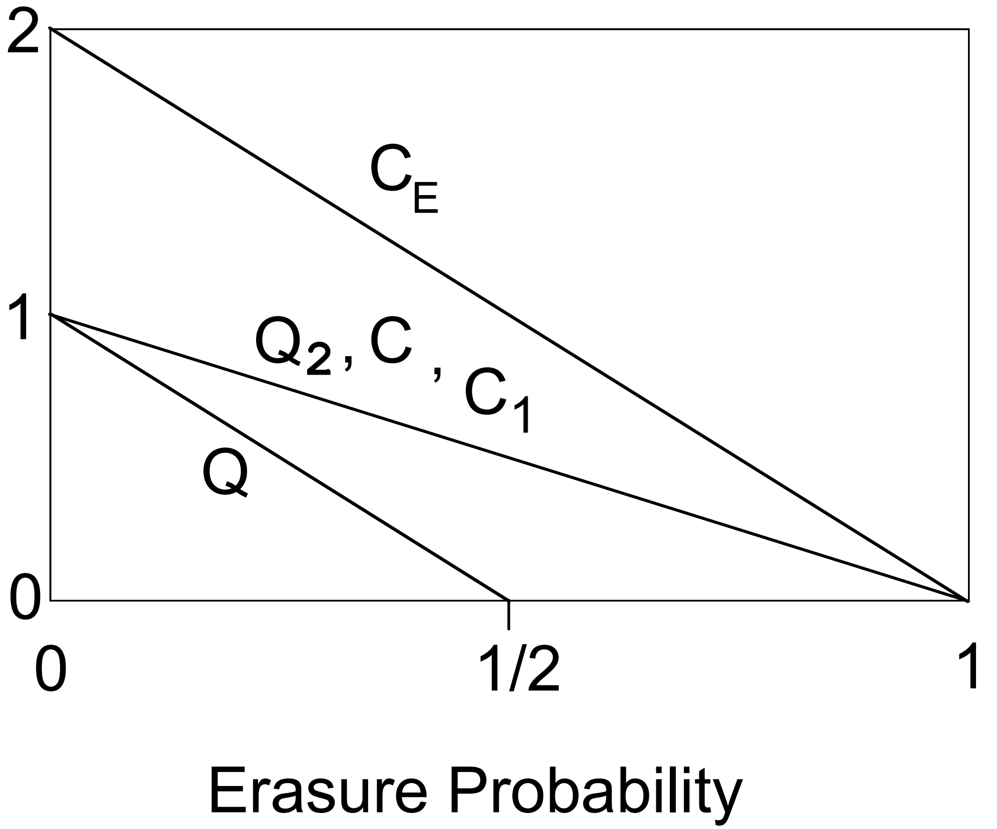
$$Q \underset{\text{by definition}}{\leq} Q_2 \underset{\text{conjectured}}{\leq} C \underset{\text{by definition}}{\leq} C_E \underset{\text{by teleportation and superdense coding}}{=} 2Q_E$$

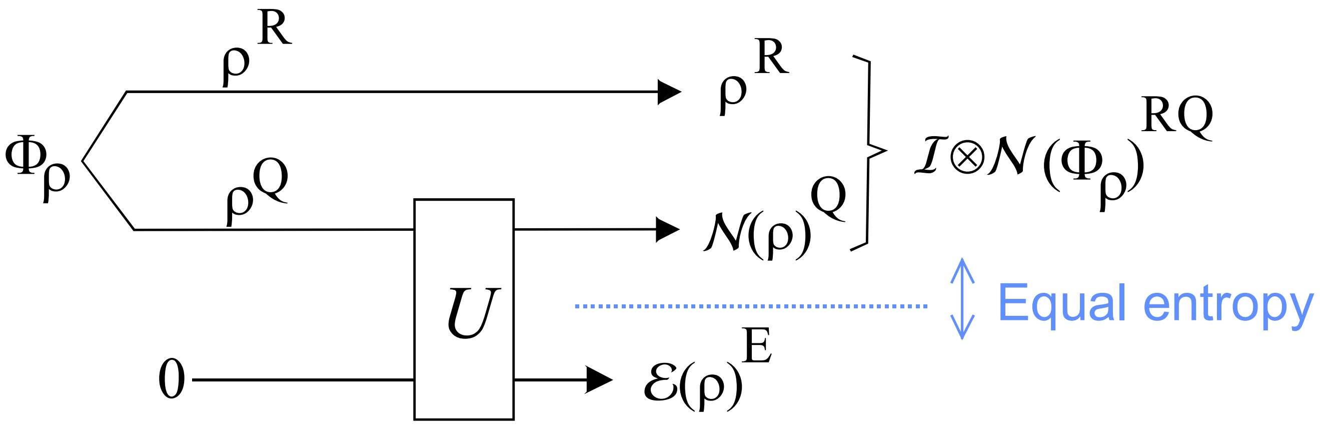
Quantum Erasure Channel

input qubit sometimes lost



Capacities of Quantum Erasure Channel





Entropic quantities related to channel capacities.

Shor ’02

C $=^?$ Holevo capacity = $\max_{\{p_i,\rho_i\}} S(\mathcal{N}(\rho)) - \Sigma p_i S(\mathcal{N}(\rho_i))$

Q = Coherent Information = $\lim_{n\to\infty} \max_{\rho} S(\mathcal{N}(\rho)) - S(\mathcal{E}(\rho))$

C_E = Quantum Mutual Info. = $\max_{\rho} S(\rho) + S(\mathcal{N}(\rho)) - S(\mathcal{E}(\rho))$

Q_2 ≈ Distillable entanglement = $\max_{\rho} D(\mathcal{I}\otimes\mathcal{N}(\Phi\rho))$ = ?

(LOCC-distillable entanglement *D* has no simple expression, may be nonconvex)

Does Free Stuff make the world better?

Robert Owen, Charles Fourier, Edward Bellamy:
Free goods & services will make everything better.

Fourier, Emma Goldman…Haight-Ashbury*
Free Love will make everything better

*same Fourier, (no relation to Fourier transform)

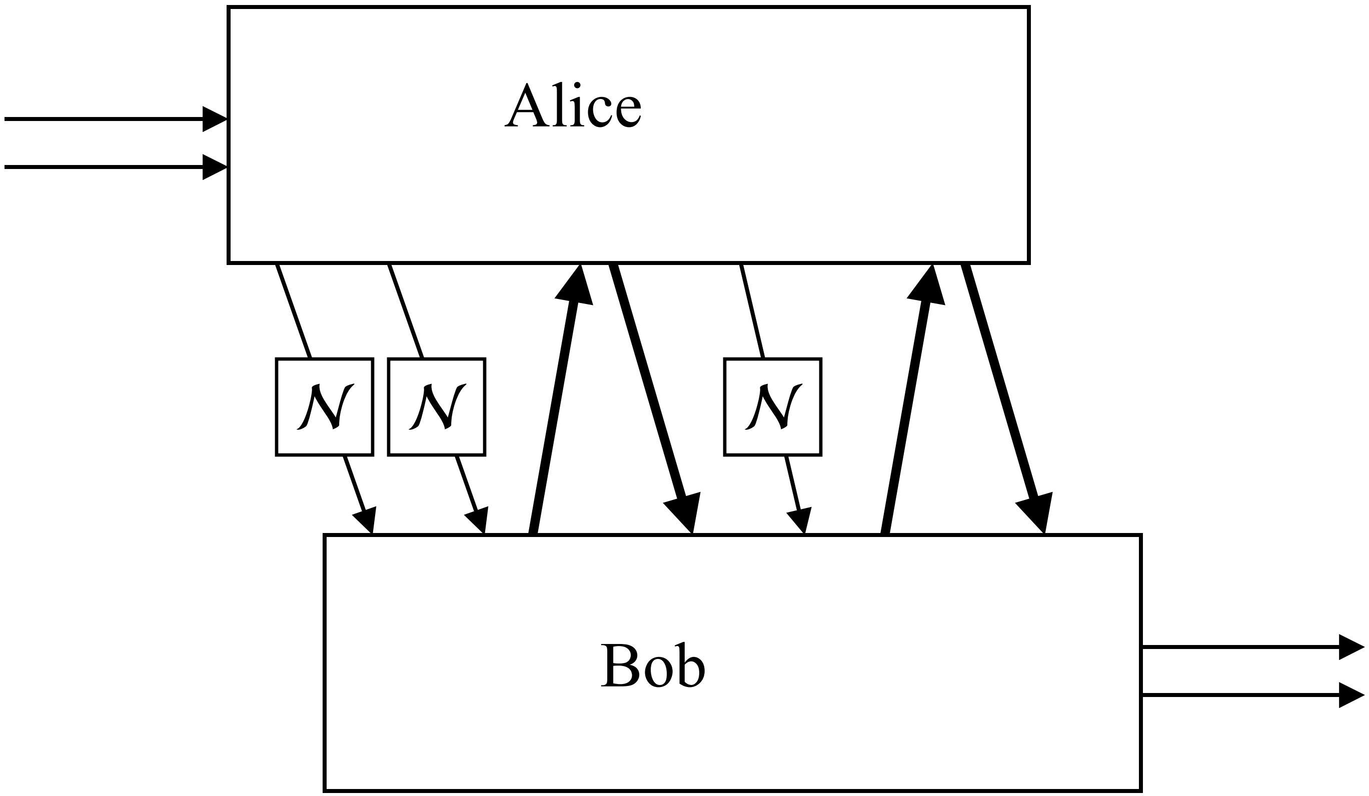
Haight-Ashbury, Timothy Leary:
Free LSD will make everything better

(Gutenberg, FOIA, the Internet, . . . LOCC)
Free Classical Communication

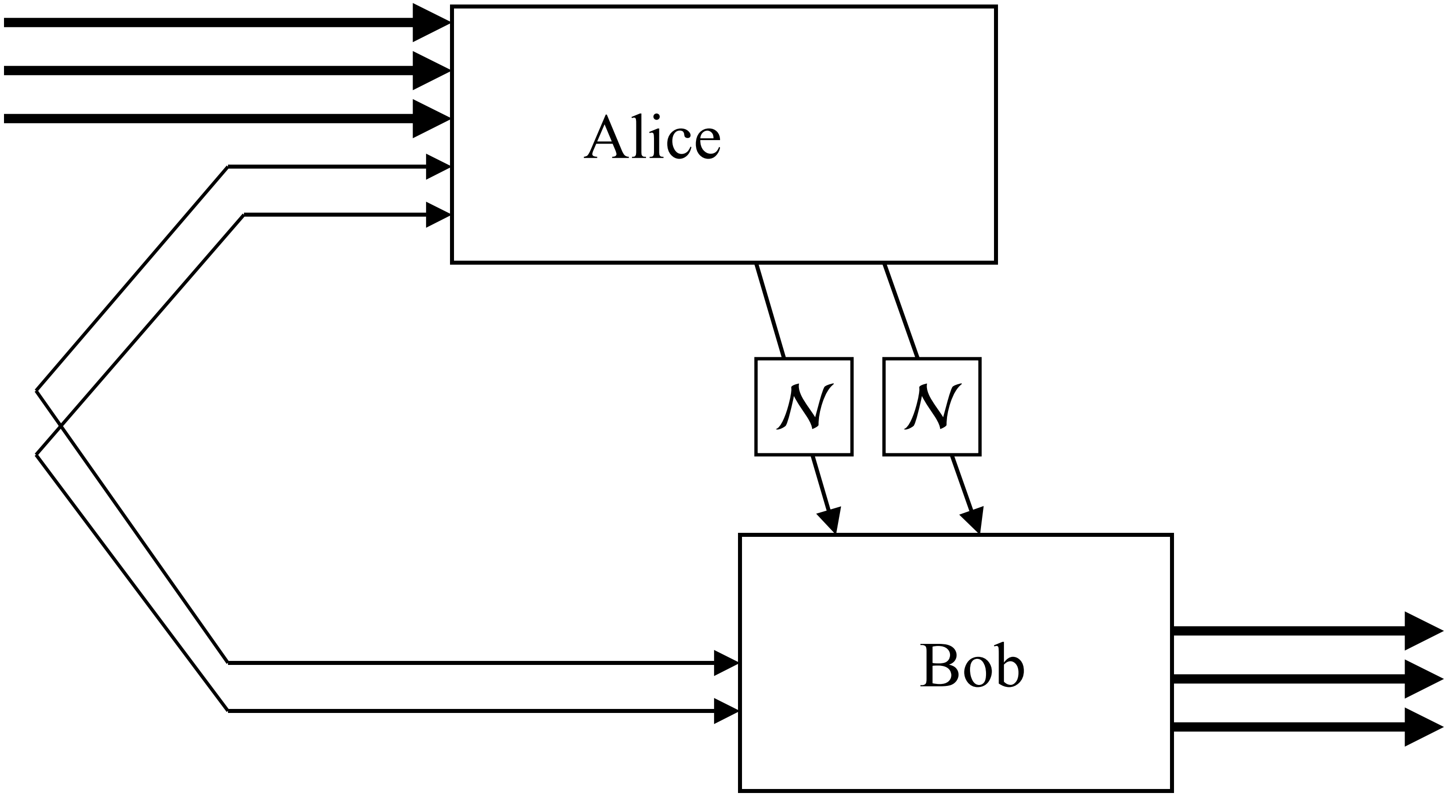
(Aram Harrow, ITP2001 poster session)
Will **Free Entanglement** change the world?

At least it simplifies the theory of quantum interactions & channels.

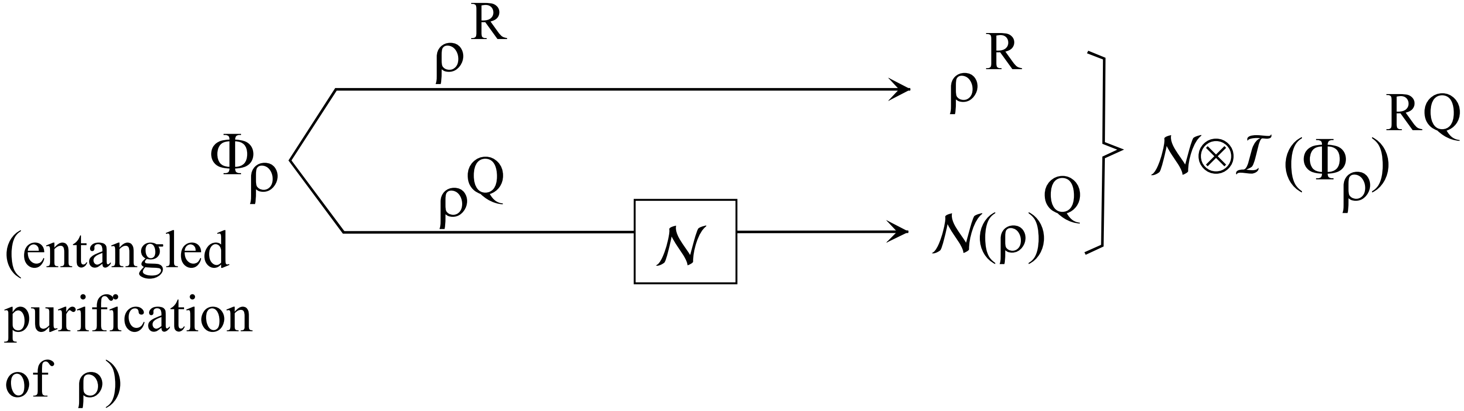
Free classical communication gives Q_2, the classically assisted quantum capacity, e.g. by entanglement sharing, distillation, and teleportation



Free entanglement gives C_E, the entanglement-assisted classical capacity.



For a noiseless channel, $C_E = 2C$ by superdense coding.

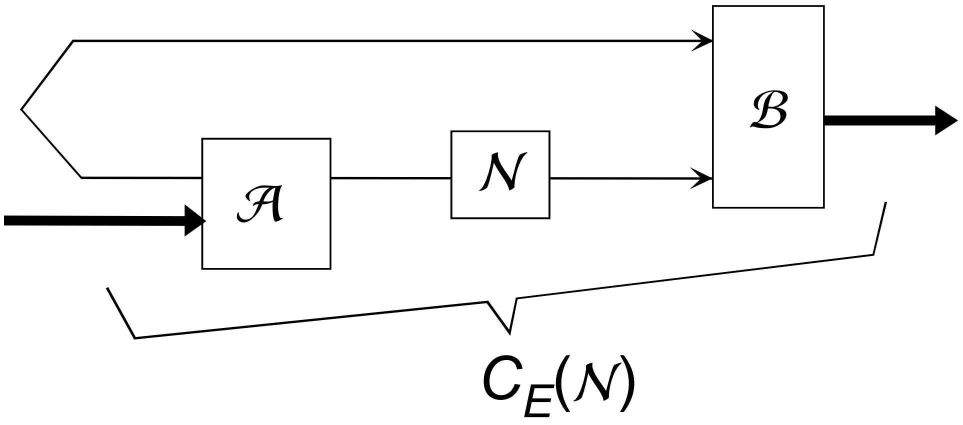

ρ^R
ρ^R
Φ_ρ
ρ^Q
$\mathcal{N}$
$\mathcal{N}(\rho)^Q$
$\mathcal{N}\otimes\mathcal{I}\,(\Phi_\rho)^{RQ}$
(entangled purification of ρ)

$$C_E(\mathcal{N}) = \max_\rho \quad S(\rho) + S(\mathcal{N}(\rho)) - S(\mathcal{N}\otimes\mathcal{I}(\Phi_\rho))$$

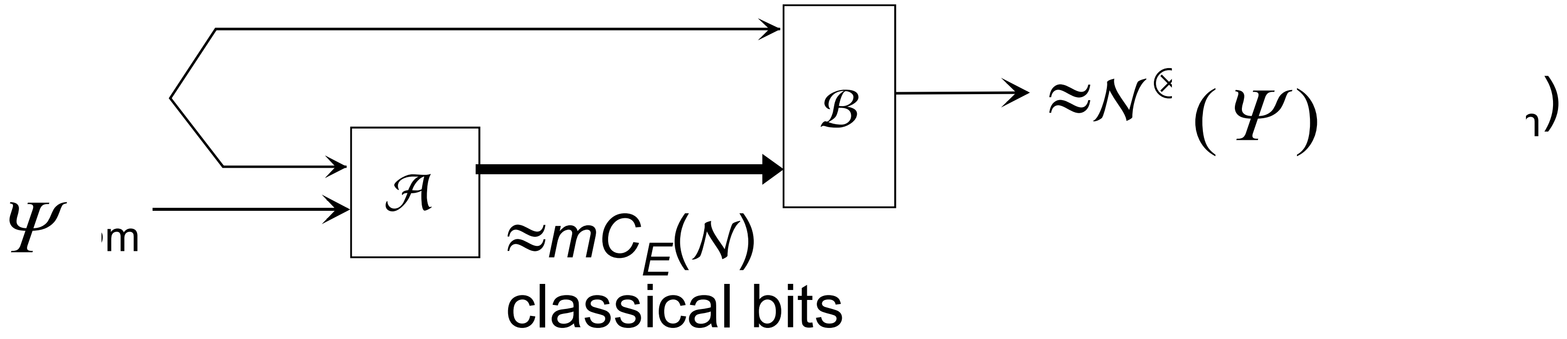
Entanglement-Assisted capacity C_E of a quantum channel $\mathcal{N}$ is equal to the maximum, over channel inputs ρ, of the input (von Neumann) entropy plus the output entropy minus their “joint” entropy (more precisely the joint entropy of the output and a reference system entangled with the late input) (BSST 0106052, Holevo 0106075).

Thus, in retrospect, entanglement-assisted capacity is the natural quantum generalization of the classical capacity of a classical channel.

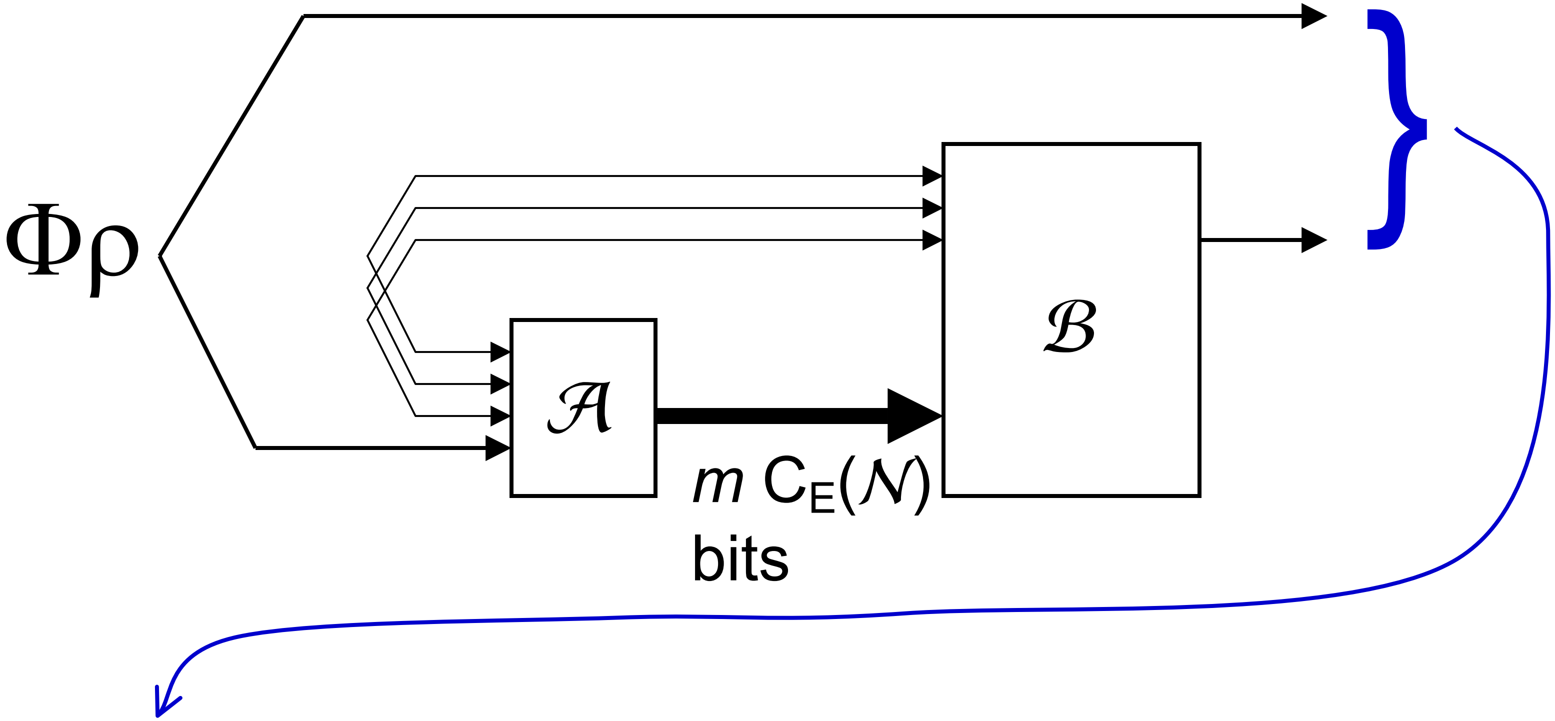
Simplification: $C_E = 2Q_E$ for all channels, by teleportation & superdense coding.



The complicated theory of quantum channel capacity would be greatly simplified if the Quantum Reverse Shannon Theorem (QRST) were true: any quantum channel can be asymptotically simulated by prior entanglement and an amount of classical communication equal to its entanglement assisted capacity. Then, in a world full of entanglement, all quantum channels would be qualitatively equivalent, and quantitatively could be characterized by a single parameter.


$\mathcal{B}$
$\approx\mathcal{N}^{\otimes}(\Psi)$
$)$
$\mathcal{A}$
Ψ m
$\approx mC_E(\mathcal{N})$
classical bits

More generally, we should demand high fidelity on entangled purifications of a mixed state input ρ



Output of simulation, including reference system, should have high fidelity with respect to $(\mathcal{N}\otimes\mathcal{I})^{\otimes m}(\Phi_\rho)$, the output on the same input of m copies of the channel being simulated.

The QRST is known to hold for all (quantum discrete memoryless) channels when their inputs are drawn from a fixed distribution ρ. This is the quantum analog of a classical IID source (Peter's proof).

For many channels, it is known to hold also for arbitrary sources even if the inputs are non-IID and allowed to be entangled across multiple instances of the quantum channel being simulated.

The ability to properly handle non-IID sources is important because, for a channel simulation to be considered faithful, it ought to accurately simulate what the channel would do even on atypical inputs which a malicious adversary might send to expose the weaknesses of the simulation.

Kinds of sources:

Tensor Power (analogous to classical IID): $\rho = \rho^{\otimes m}$

Tensor Product: $\rho = \rho_1 \otimes \rho_2 \otimes \rho_3 \otimes \ldots$

with each factor in H_{in}

(Arbitrary pure: $\psi =$ a general pure state in $H_{\text{in}}^{\otimes \boldsymbol{m}}$)

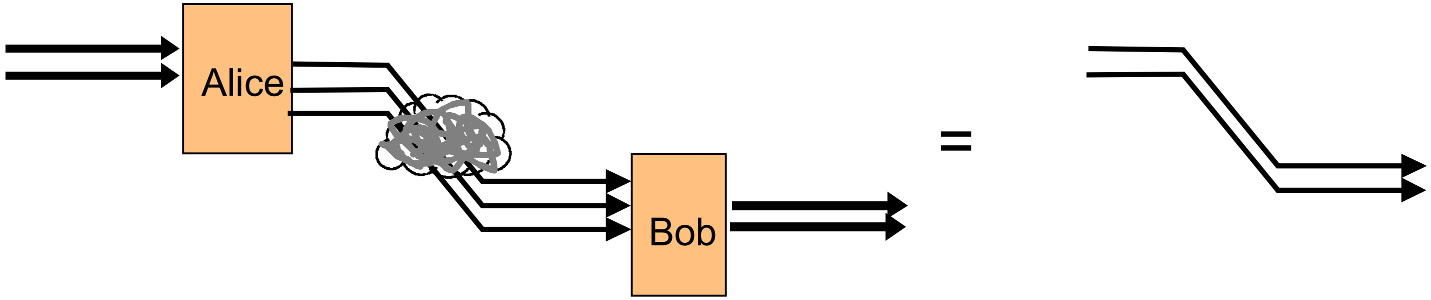
Most general: any pure state Ψ in $\underbrace{H_{\text{in}}^{\otimes \boldsymbol{m}}}_{m \text{ channel inputs}} \otimes \underbrace{H_{\text{in}}^{\otimes \boldsymbol{m}}}_{\text{Purifying reference system}}$

(worst an adversary could send)

Classical Reverse Shannon Theorem (0106052)

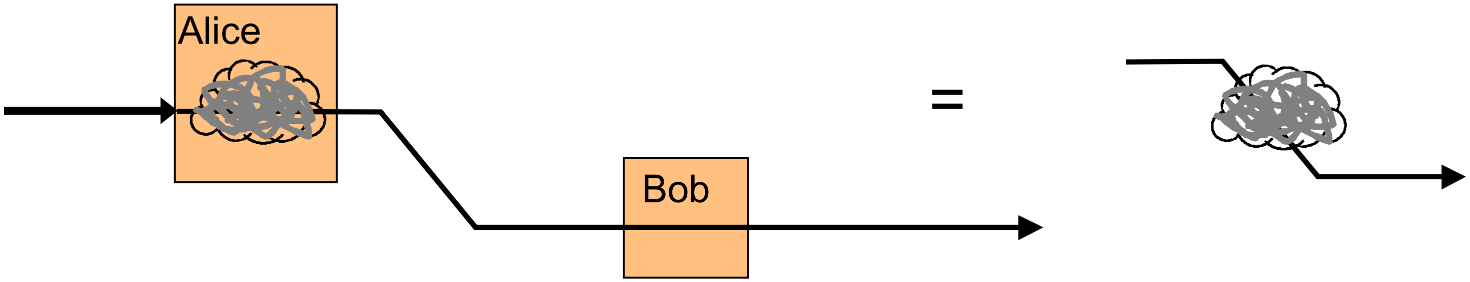
Classical Shannon Theorem:

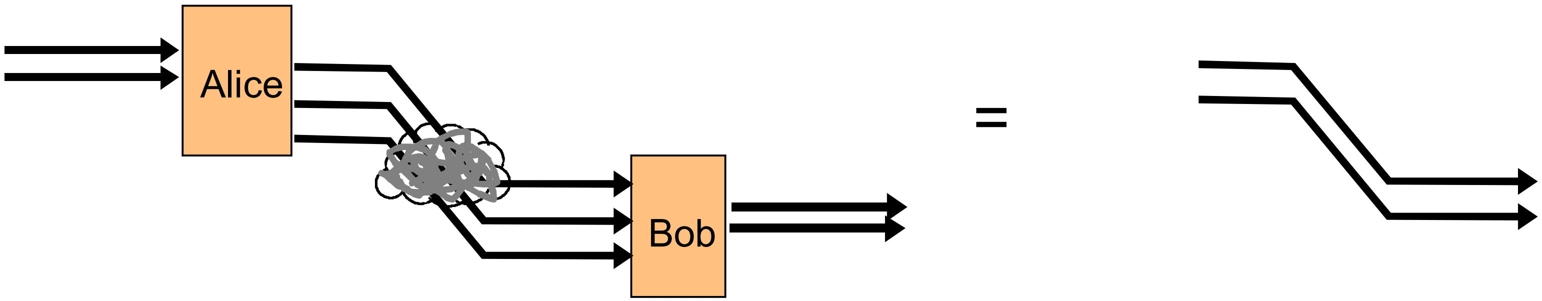
A noisy channel can simulate a noiseless channel



Homer Simpson's Reverse Shanon's Theorem:

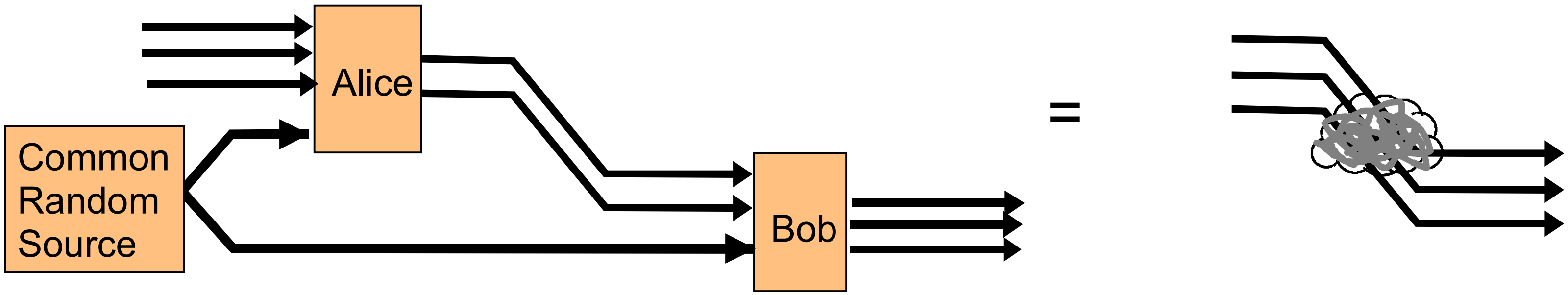
A noiseless channel can simulate a noisy channel.





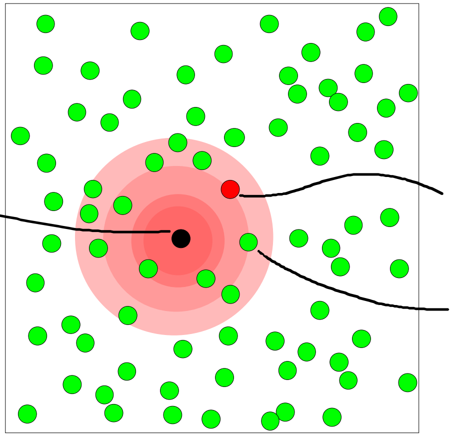
A Better Reverse Shannon Theorem *(quant-ph/0106052)*

In the presence of shared random information between sender and receiver,
a noiseless channel can asymptotically simulate a noisy one ***of equal capacity*****.**



Therefore, in the presence of shared random information,
all classical noisy channels are asymptotically equivalent.

Simulation Method: Alice and Bob first preagree on a sparse set S_R of $2^{n(C+\delta)}$ n-bit strings, using their shared random info R.

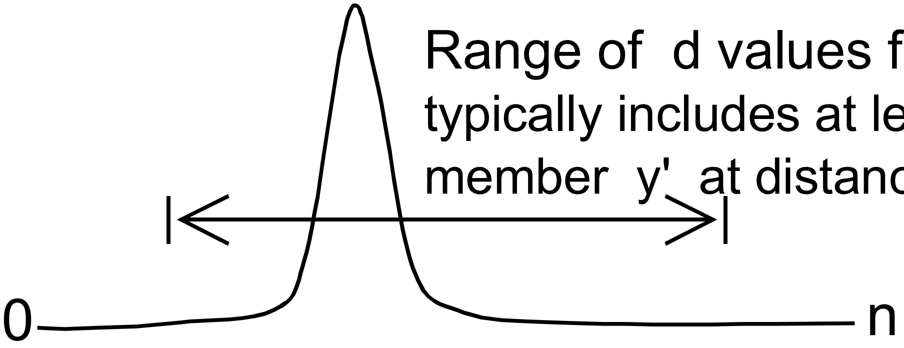


P(y|x)

Alice receives input string x

Next she simulates the channel locally to get a provisional output y

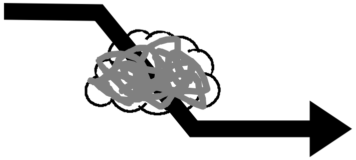
Then she picks y' in S_R at same distance from x as y was, and tells Bob its index using $n(C+\delta)$ bits.



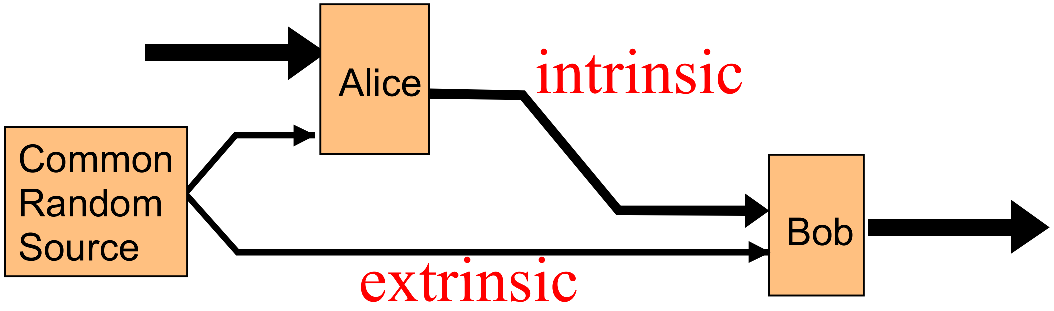
Range of d values for which S_R typically includes at least one member y' at distance d from x.

Distribution of Hamming distances d=|x-y| induced by noisy channel.

In the large *m* limit, sending *m* bits through the noisy channel



can be simulated by sending about *mC* noiseless "intrinsic" bits, which Alice chooses with the help of the input,



and about *m(1-C)* "extrinsic" random bits, which have nothing to do with the channel input, and so can be preagreed before Alice receives the input.

Measurement Compression

Given a density matrix ρ and a POVM $\boldsymbol{a} = \{a_j\}$, define the one-shot output probabilities λ_j=Tr $\rho a_{j.}$, and the square root ensemble $\rho_j = (\sqrt{\rho})\, a_j\, (\sqrt{\rho}) / \lambda_j$ realizing ρ. Then for any tolerance $\varepsilon>0$, there exists a block size l and a POVM $\boldsymbol{B}$, which is a good approximation to $\boldsymbol{A}=\boldsymbol{a}^{\otimes l}$, and where $\boldsymbol{B}$ can be expressed as a convex combination $\boldsymbol{B}=\Sigma_\nu x_\nu \boldsymbol{B}^\nu$ of constituent POVMs $\boldsymbol{B}^\nu$ each having at most M outcomes, where $\log M \approx l\,(S(\rho) - \Sigma_j\, \lambda_j\, S(\rho_j))$ is the Holevo information of the square root ensemble. The approximation is good in the sense that for any entangled purification Φ of $\rho^{\otimes l}$,
$F((\boldsymbol{A}\otimes I)\,\Phi\,, (\boldsymbol{B}\otimes I)\,\Phi\,) > 1-\varepsilon$.

On any tensor power source ρ, the POVM $\boldsymbol{a}$, regarded as QC channel, can be asymptotically simulated by shared randomness and an amount of forward classical communication approaching the quantum mutual information of $\boldsymbol{a}$ on ρ.

QRST for QC channels on known IID sources

$$\text{QMI}\,(\boldsymbol{a},\rho) \;\equiv\; S(\rho) + S(\boldsymbol{a}(\rho)) - S\,(\boldsymbol{a}\otimes\mathcal{I}(\Phi_\rho)).$$

$$\| $$

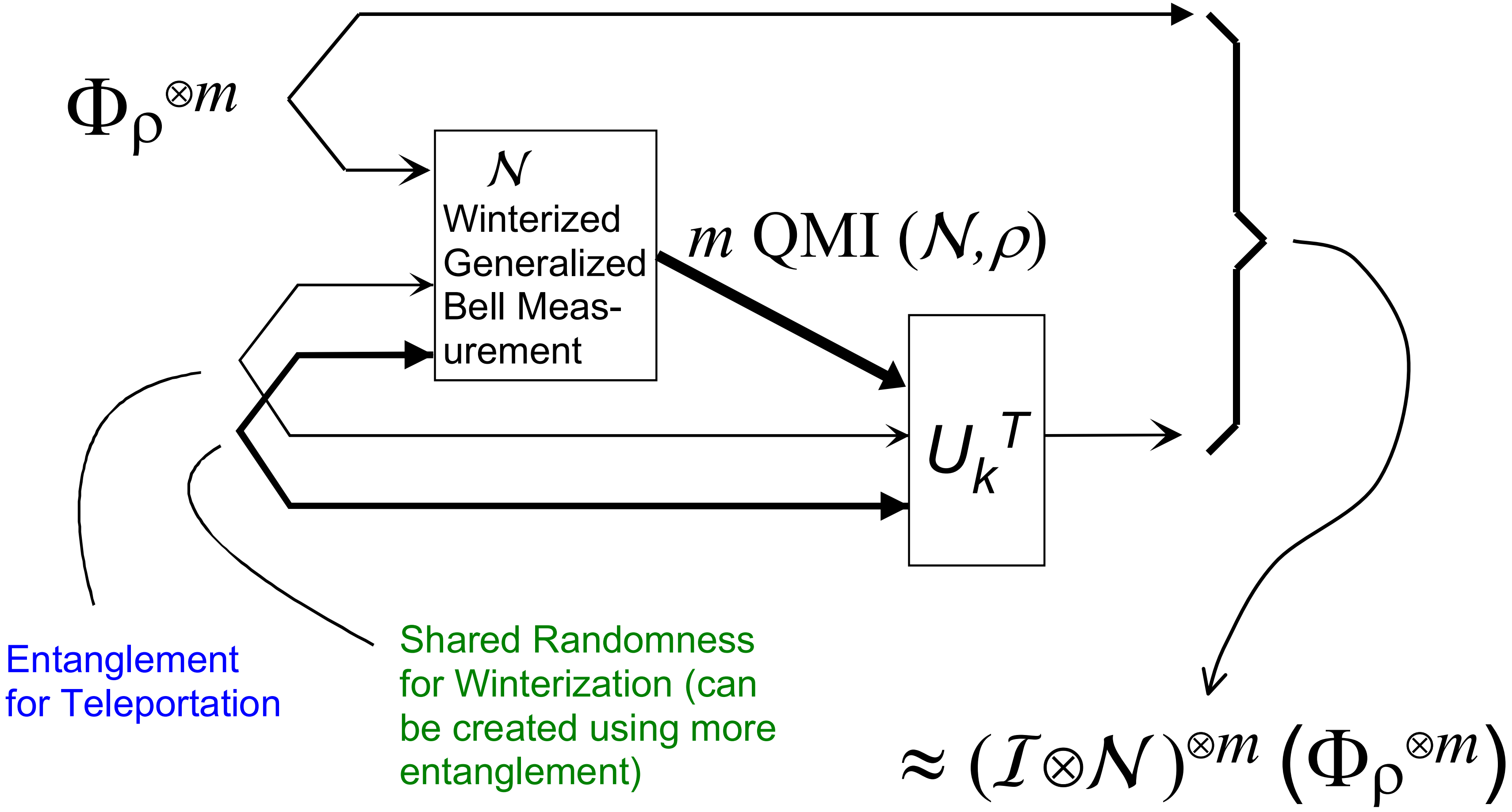
$$S(\boldsymbol{a}(\rho)) + \Sigma_j\, \lambda_j\, S(\rho_j)$$

Sketch of Shor's proof of QRST for tensor power sources, using Winter's compression theorem. Alice's wants to simulate a general noisy channel $\mathcal{N}$, using shared entanglement and as little classical communication to Bob as possible. Let $\mathcal{N}$ be defined by the Kraus operators $\{N_k : k=1\ldots\delta\}$ so on input state ρ the channel output is $\Sigma_k N_k \rho N_k^\dagger$. Let Φ_{in} and Φ_{out} denote projectors onto maximally entangled states sized to the input and output dimensions of $\mathcal{N}$. Let U_j be d_{out} dimensional generalized Pauli matrices.

Generalized Teleportation: Alice performs a POVM with elements $(I\otimes U^*_j N^*_k)\, \Phi_{\mathrm{in}}\, (I\otimes N^T_k U^T_j)$ on the input and her half of a specimen of Φ_{out}, after which she tells Bob only j, the index of which Pauli she performed. He undoes the Pauli, and is left with $\mathcal{N}(\rho)$. This uses 2 log d_{out} bits of classical communication.

Measurement compression: For large block size m, Alice and Bob approximate this POVM by another with an intrinsic cost of m (QMI $(\mathcal{N},\rho)$) + $o(m)$

Overall picture



$\Phi_\rho^{\otimes m}$

$\mathcal{N}$
Winterized
Generalized
Bell Measurement

$m\ \mathrm{QMI}\ (\mathcal{N},\rho)$

U_k^T

Entanglement
for Teleportation

Shared Randomness
for Winterization (can
be created using more
entanglement)

$\approx (\mathcal{I}\otimes\mathcal{N})^{\otimes m}\ (\Phi_\rho^{\otimes m})$

This establishes QRST for a general channel on known IID source.

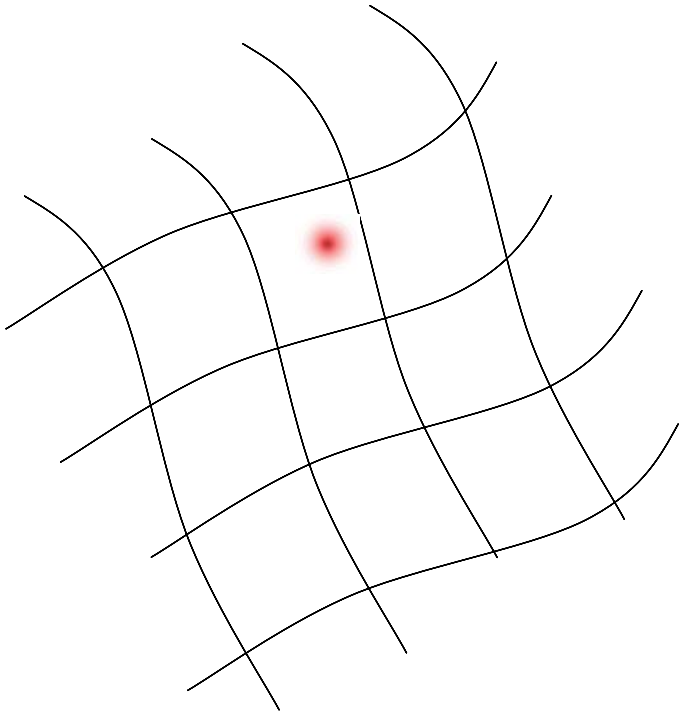
For a general channel on an unknown IID source, use gentle tomography on a large block of inputs to estimate the source, then proceed as with a known IID source.

For a CQ channel on an arbitrary source, Alice performs the initial C part of the channel on a large block of *m* inputs and makes a copy of the results. These results will be unentangled between channel instances, but may not be IID. Using $o(m)$ bits, Alice tells Bob the frequency distribution (type class) of the measured results and they then simulate the full CQ channel on this type class. (Alternatively, this may be viewed as remote state preparation of mixed states which can be done at the cost of the Holevo information of the ensemble, which equals the QMI.)

For a Bell-diagonal channels on arbitrary sources, the noisy quantum channel is directly equivalent to teleportation through a noisy classical channel, which can be simulated using the classical reverse Shannon theorem.

To extend QRST to an unknown tensor power source:
Use gentle tomography to estimate ρ from a large number *m* of copies of ρ without much disturbing the global state.

("Tender measurement" from Keiji Matsumoto's talk)



This may be viewed roughly as choosing a random mesh on the parameter space of ρ coarse enough ($\propto 1/\sqrt{m}$) so for any ρ, a measurement on ρ^m of which cell the average falls in almost always yields the same result. This measurement, when conducted coherently, will therefore scarcely disturb the global state.

(almost) no Information =>

(almost) no disturbance

(sign at Frankfurt Airport)

With gentle tomography, get an estimate of the average density matrix ρ_{est} and its quantum mutual information.

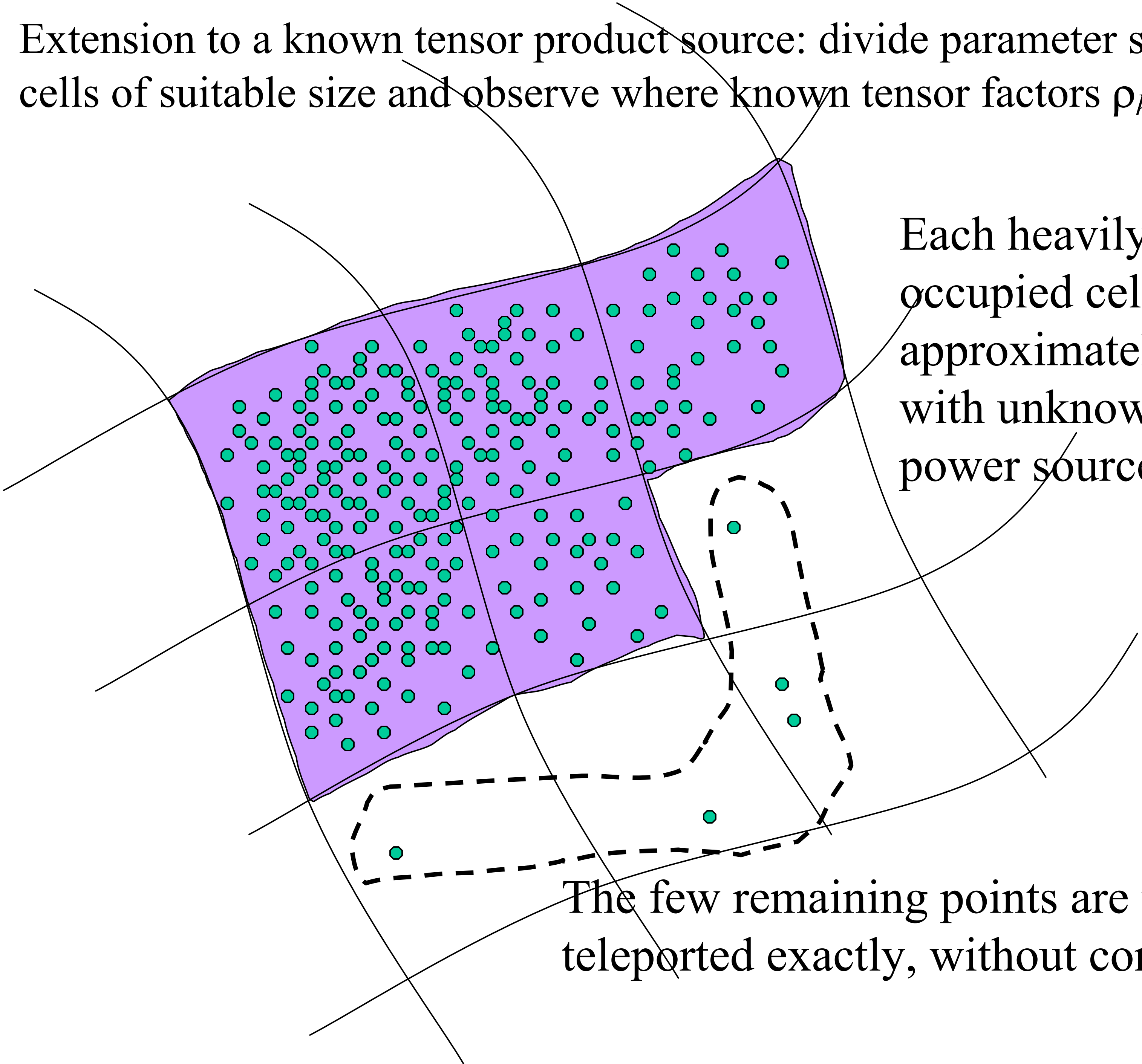
(Crudely speaking): do compressed teleportation using a version of Winter's theorem designed not for the source $\rho_{\text{est}}^{\otimes m}$, but rather for the (non tensor power) source ρ_{union} corresponding to the union of all typical subspaces of density matrices in the same mesh cell as ρ_{est} .

- This union has a dimension only subexponentially greater than the typical subspace of ρ.

- Also we use the fact that the fidelity of measurement compression approaches 1 exponentially with increasing block size, for any forward communication rate R exceeding the QMI.

Therefore, the cost of simulating the channel on an unknown IID source still asymptotically approaches the quantum mutual information of the channel on that source.

Extension to a known tensor product source: divide parameter space into cells of suitable size and observe where known tensor factors ρ_k fall.



Each heavily occupied cell is coded approximately, as with unknown tensor power source

The few remaining points are then teleported exactly, without compressing.

Costs of entanglement assisted channel simulation

Channel / Source	Bell diagonal	Classical or CQ	General Channel
Known tensor power	QMI	QMI	QMI
Unknown tensor power	QMI	QMI	QMI
Known tensor product	ave. QMI	ave. QMI	ave. QMI
General Source	$\leq C_E$	= QMI of collapsed source	$\leq 2 \log \min\{d_{\text{in}}, d_{\text{out}}\}$

QMI = quantum mutual information for the source/channel combination

Open questions:

Prove QRST for most general source model, or find counterexample (a source/channel combination requiring more than CE to simulate).

What other free resource, if any, will similarly simplify the theory of quantum communication in the absence of free entanglement?

- LOCC? probably not
- PPT-preserving operations?
- separability-preserving operations?