Strong Converse Theorems in Quantum Information Theory

Mark M. Wilde

Hearne Institute for Theoretical Physics
Department of Physics and Astronomy
Center for Computation and Technology
Louisiana State University, Baton Rouge

In collaboration with

Bhaskar Roy Bardhan (LSU),
Manish K. Gupta (LSU),
Naresh Sharma (TIFR Mumbai),
Andreas Winter (UAB Barcelona),
Dong Yang (UAB Barcelona)

arXiv:1306.1586, arXiv:1308.6732, arXiv:1310.7028

What is capacity?

Given is a description of a **quantum channel** (perhaps in terms of a Kraus representation):

$$\mathcal{N}(\rho) = \sum_i A_i \rho A_i^\dagger$$

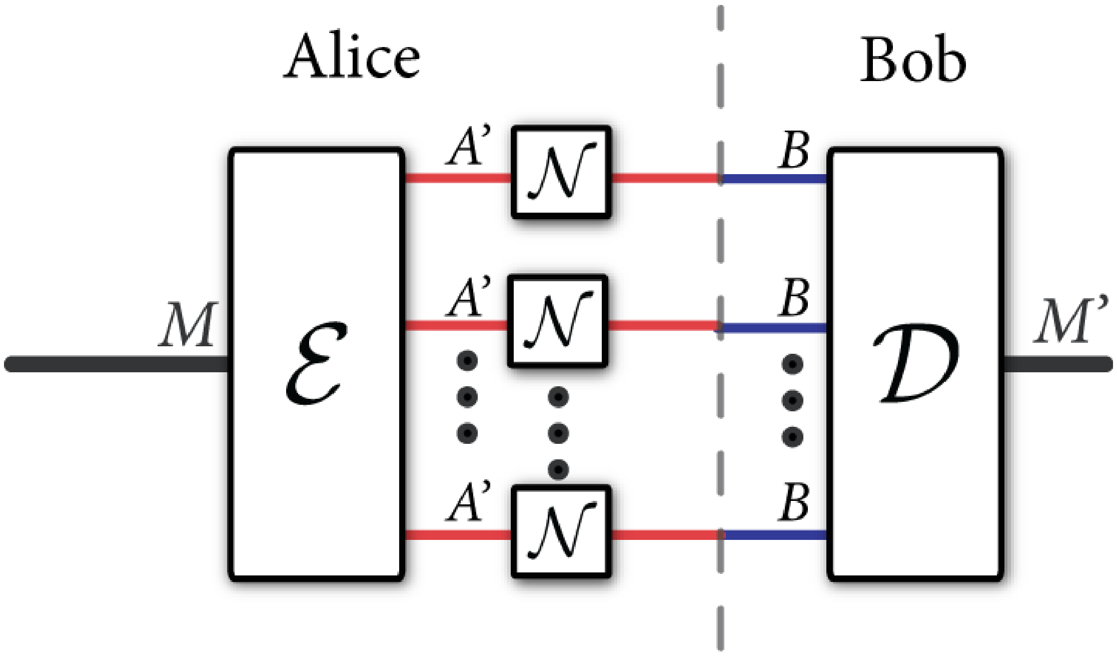
where

$$\sum_i A_i^\dagger A_i = I$$

Capacity is the maximum rate at which a sender can transmit data reliably to a receiver using this channel many times.
(*We will focus on classical data transmission here.*)

What is capacity? (ctd.)

As Shannon did, allow for a slight error in the transmission, but demand that it vanishes in the limit of many channel uses while operating at a fixed rate of communication



An **(*n*,*R*,*ε*) protocol** uses the channel *n* times at a rate *R* while having error no larger than *ε*.

What is capacity? (ctd.)

A rate R is **achievable** if there exists a sequence of $(n, R - \delta, \varepsilon)$ protocols for all $\varepsilon, \delta > 0$ and sufficiently large n.

The **capacity** of a channel is defined as the supremum of all achievable rates.

The capacity is defined operationally and is a function of a channel.

(*Observe that the definitions above effectively coarse grain away all parameters except for rate.*)

A major goal of quantum information theory is to identify a **tractable formula** for capacity.

How to prove a capacity theorem?

First part: Demonstrate the **existence** of (n,$R - \delta$,ε) protocols (codes) operating at some rate R.

Usual approach is to **pick codes at random** à la Shannon and show that some code in the ensemble meets the performance criteria.

This identifies a **lower bound** on capacity

How to prove cap. theorem? (ctd.)

Second part: Identify a rate above which it is impossible to communicate error free.
(*known as a weak converse rate*)

This serves as an **upper bound** on capacity.

If the lower and upper bounds coincide, then you have characterized capacity.

There is an important practical question of whether the formulas characterizing capacity are simple to compute and many of the open questions in QIT revolve around this.

Holevo-Schumacher-Westmoreland Theorem

Holevo, Schumacher, and Westmoreland found that a quantum generalization of Shannon's formula characterizes capacity.

Define the **Holevo information of a channel** as

$$\chi(\mathcal{N}) = \max_{\{p(x),\rho_x\}} I(X;B)_\rho$$

where

$$I(X;B)_\rho = H(X)_\rho + H(B)_\rho - H(XB)_\rho$$

$$\rho_{XB} \equiv \sum_x p(x)|x\rangle\langle x|_X \otimes \mathcal{N}_{A\to B}(\rho_x)$$

HSW Theorem (ctd.)

Holevo, Schumacher, and Westmoreland characterized capacity as a regularization of the Holevo information:

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

If the Holevo information is **additive** for a channel for which we are trying to determine capacity, then capacity is given by $\chi(\mathcal{N})$

However, Hastings proved that it does not have to be additive. (*He proved the existence of a channel for which it is not additive.*)

Weak vs. Strong Converse

The **converse part** of the HSW theorem due to Holevo (1973) only establishes what is called a “weak converse.”

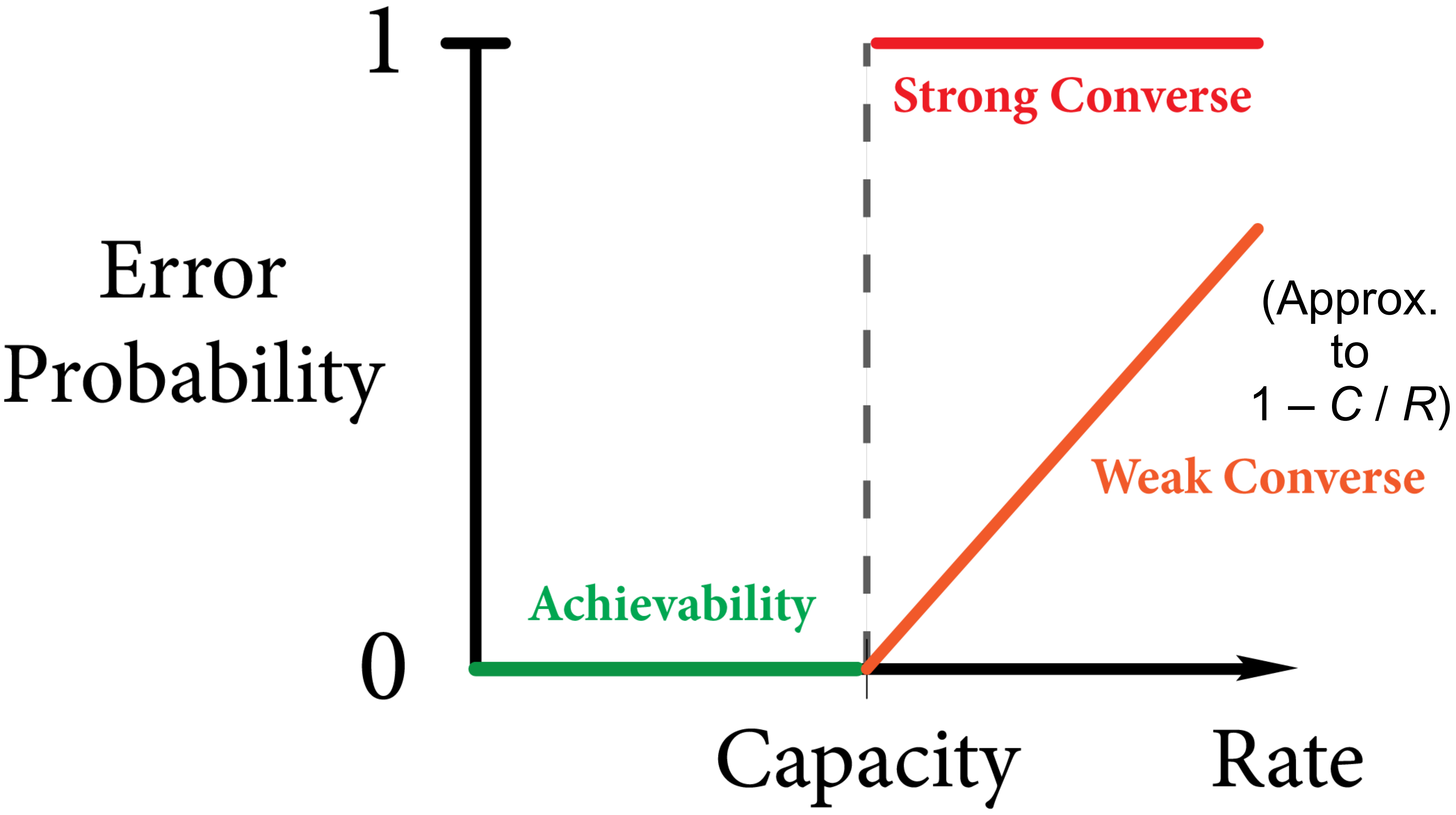
That is, it states that there cannot be an error-free communication scheme if rate exceeds capacity.

$$R \leq \frac{1}{1-\varepsilon} C$$

This suggests that there might be room for a trade-off between rate and error.

A **strong converse theorem** (if it holds) rules out such a possibility. It states that the error probability converges to one in the limit as *n* becomes large if rate exceeds capacity

Weak vs. Strong Converse



Plot should be interpreted in the large *n* limit

Why do we want strong converses?

Strengthens the interpretation of capacity as a **very sharp dividing line**

(*analogous to a phase transition*)

Helpful in proving the **security** of particular models of cryptography in which the eavesdropper is limited to having **noisy storage**

(*see Koenig, Wehner, Wullschleger arXiv:0906.1030*)

What was already known?

Nayak proved a strong converse for classical capacity of noiseless qubit channels (1999)

(*he did not call it this, but the result follows*)

Winter proved a strong converse holds for channels with classical inputs and quantum outputs using a “Wolfowitz” approach (1999)

Ogawa and Nagaoka proved the same theorem with a different approach using Rényi entropies --- the “Arimoto” approach (1999)

Koenig and Wehner proved it for the classical capacity of several covariant channels (2006)

New Results

1) Strong converse for the classical capacity of entanglement-breaking and Hadamard channels (W, Winter, Yang 2013)

2) Strong converse for classical capacity of pure-loss bosonic channel (W and Winter 2013)

3) Strong converse for entanglement-assisted capacity (Gupta and W 2013)

4) Strong converse rates for classical comm. over thermal bosonic channels (Bardhan and W 2013)

For this talk, we focus on 2), 4), & then 1) if time

Review of SC for noiseless channel

Consider any code for communication. It consists of density operators ρ_m, depending on the message *m*, and a decoding POVM $\{\Lambda_m\}$.

Its **average success probability** is

$$\frac{1}{M}\sum_m \mathrm{Tr}\{\Lambda_m \rho_m\}$$

Review of SC for noiseless channel

$$\frac{1}{M}\sum_{m}\mathrm{Tr}\{\Lambda_m\rho_m\}$$

If *R* > 1, the success probability decays exponentially fast to zero.

Review of SC for noiseless channel

Albeit simple, the proof of the strong converse for the noiseless qubit channel highlights a fundamental interplay between

1) **success probability**

2) **number of messages** (related to rate)

3) **dimension** of the space for encoding

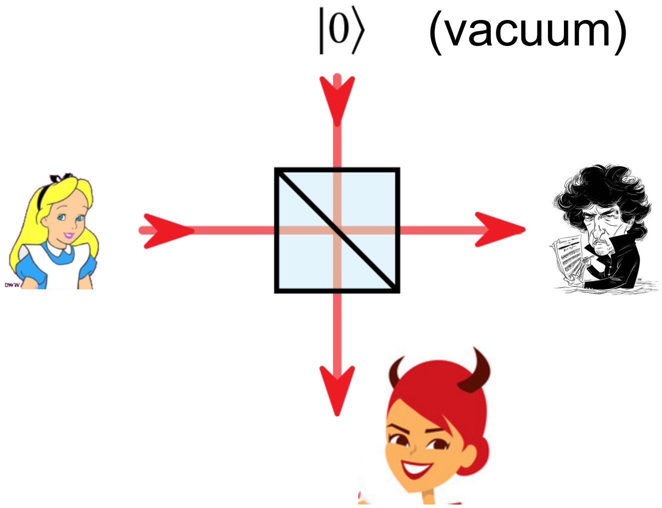
4) **purity** of the channel
(infinity norm of the output states)

Reasoning in a similar way for bosonic channels helps in establishing strong converse rates for them

What is a pure-loss bosonic channel?

Pure-Loss Bosonic Channel

(*models fiber optic or free space transmission*)



Heisenberg input-output relation for channel:

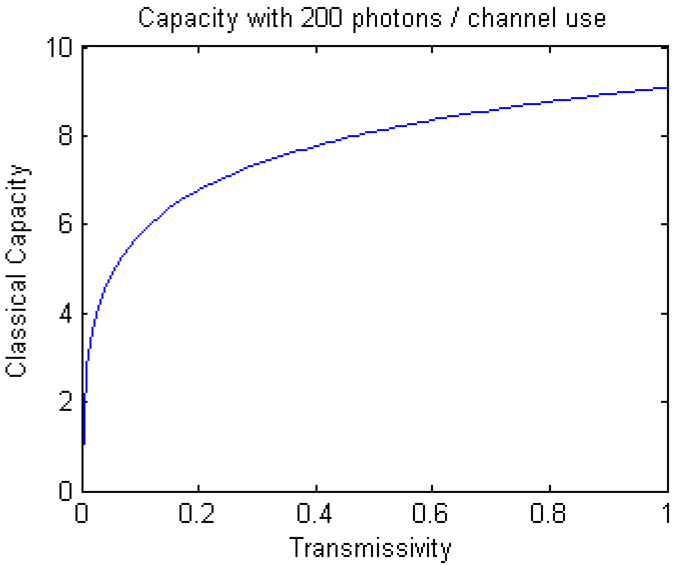
$$\hat{b} = \sqrt{\eta}\hat{a} + \sqrt{1-\eta}\hat{e}$$

Pure-loss bosonic channel capacity

Classical capacity of **pure-loss channel** is exactly

$$g(\eta N_S)$$

where η is **transmissivity** of channel, N_S is the **mean input photon number**, and $g(x) = (x+1)\log(x+1) - x\log x$ is the **entropy** of a **thermal state** with photon number x



Can **achieve** this capacity by selecting **coherent states** randomly according to a complex, isotropic Gaussian prior with variance N_S

Holevo and Werner, arXiv:quant-ph/9912067
Giovannetti *et al.*, *Physical Review Letters* 92, 027902 (2004)

SC for pure-loss bosonic channel

Can prove that a strong converse does not hold under a **mean photon number constraint**

However, instead impose a **maximum photon number constraint**, i.e.,

$$\frac{1}{M}\sum_{m}\mathrm{Tr}\{\Pi_{\lceil nN_S\rceil}\rho_m\}\geq 1-\delta(n)$$

where

$$\Pi_L = \sum_{a^n:\sum_i a_i\leq L} |a_1\rangle\langle a_1|\otimes\cdots\otimes|a_n\rangle\langle a_n|$$

Joint work with Andreas Winter, arXiv:1308.6732

SC for pure-loss bosonic channel

If the input is in the subspace with photon number no larger than ***nN_S***, then the output is with high probability in a subspace of photon number no larger than ***$n(\eta N_S)$***.

(*Physically intuitive, appeal to LLN for a proof*)

An upper bound on the dimension of the subspace with photon number no larger than $n(\eta N_S)$ is

$$2^{n[g(\eta N_S)+\delta]}$$

where δ can be chosen arbitrarily small as n gets large.

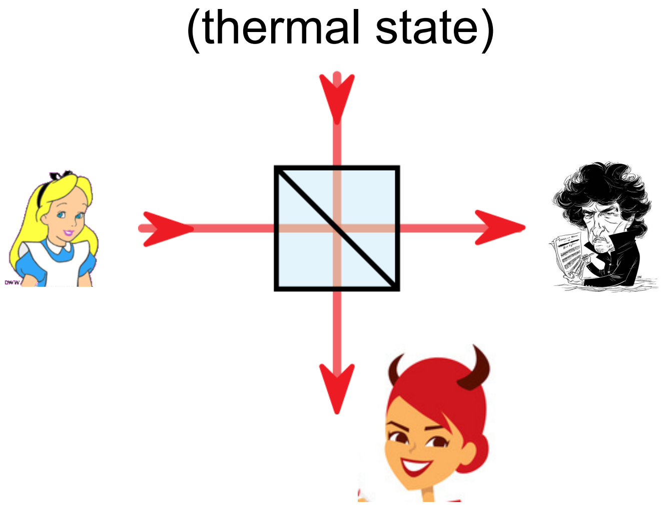
SC for pure-loss bosonic channel

The **purity** of the channel is maximal because we can always send in a coherent state, which retains its purity for the pure-loss bosonic channel.

By arguing along lines very similar to the proof for the noiseless channel, we arrive at an upper bound on the success probability, that is essentially

$$2^{-n[R-g(\eta N_S)]}$$

Thermal bosonic channels



Known lower bound on capacity:

$$g(\eta N_S + (1-\eta)N_B) - g((1-\eta)N_B)$$

Known (weak) upper bounds on capacity:

$$g(\eta N_S/((1-\eta)N_B + 1))$$

Koenig and Smith 2012

$$g(\eta N_S + (1-\eta)N_B) - \log_2(1 + 2(1-\eta)N_B)$$

Giovannetti *et al*. 2004

SC rates for thermal bosonic channels

Recent work with Bardhan shows that both of these upper bounds are **strong converse rates**

To get the Koenig-Smith bound, use the fact that a thermal channel is equivalent to a pure-loss channel with transmissivity ηN_S / $((1-\eta)N_B + 1)$ followed by an amplifier channel

Then apply strong converse for pure-loss channel

Joint work with Bhaskar Roy Bardhan (LSU)

SC rates for thermal bosonic channels

To recover the Giovannetti *et al*. bound, first prove that if the input is in the subspace with photon number no larger than $\boldsymbol{nN_S}$, then the output of the thermal channel is with high probability in the subspace with photon number no larger than $\boldsymbol{n\,(\eta\, N_S + (1\text{-}\eta)\, N_B)}$ (*again with LLN*)

Upper bound the **purity** of the channel with the **smooth min-entropy**, which we can then relate to the collision entropy (Renyi entropy of order 2)

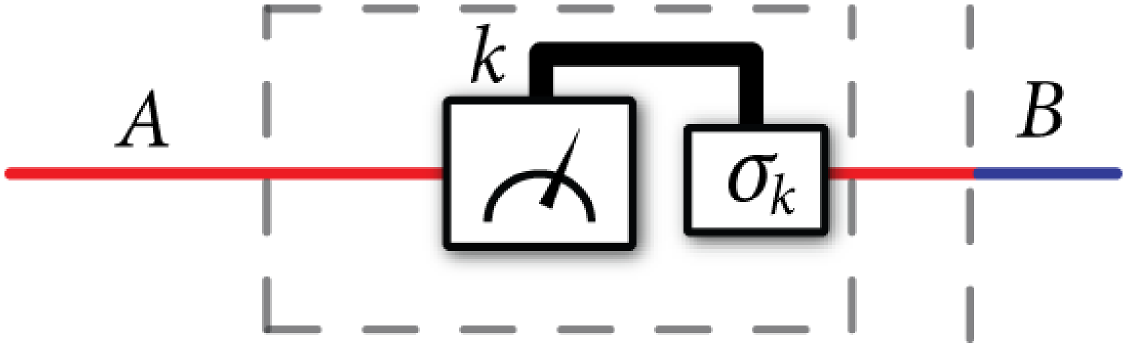
Giovannetti *et al*. have shown that the state leading to the minimum output collision entropy is the vacuum.

We then get the following upper bound:

$$g(\eta N_S + (1-\eta)N_B) - \log_2(1 + 2(1-\eta)N_B)$$

SC for classical capacity of entanglement-breaking channels

An **entanglement-breaking channel** is equivalent to a measurement of the input followed by a preparation of a state conditioned on the outcome of the measurement:



We have shown that a strong converse holds for the classical capacity of these channels

Joint work with Andreas Winter and Dong Yang, arXiv:1306.1586

SC for classical cap. of EB channels

Begin with the "generalized divergence" framework of Sharma and Warsi (arXiv:1205.1712).

A measure is called a **generalized divergence** if it satisfies monotonicity under quantum operations:

$$\mathcal{D}(\rho||\sigma) \geq \mathcal{D}(\mathcal{N}(\rho)||\mathcal{N}(\sigma))$$

Generalized Holevo information of a channel:

$$\chi_{\mathcal{D}}(\mathcal{N}) = \max_{p(x),\rho_x} \min_{\sigma_B} \mathcal{D}(\rho_{XB}||\rho_X \otimes \sigma_B)$$

$$\rho_{XB} \equiv \sum_x p(x)|x\rangle\langle x|_X \otimes \mathcal{N}_{A\to B}(\rho_x)$$

SC for classical cap. of EB channels

Using **monotonicity**, we find the following relation between success probability, rate, and generalized Holevo information for any code:

$$\chi_{\mathcal{D}}(\mathcal{N}^{\otimes n}) \geq \delta(\varepsilon || 1 - 2^{-nR})$$

where δ is the generalized divergence for classical states

$$\delta(p||q) \equiv \mathcal{D}(\rho_p||\rho_q)$$

$$\rho_p \equiv p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1|$$

Sandwiched Rényi relative entropy

We can take the divergence to be the **sandwiched Rényi relative entropy**, defined as

$$\widetilde{D}_\alpha(\rho\|\sigma) \equiv \frac{1}{\alpha-1}\log_2 \mathrm{Tr}\{(\sigma^{(1-\alpha)/2\alpha}\,\rho\,\sigma^{(1-\alpha)/2\alpha})^\alpha\}$$

This divergence already has a bit of history:

V. Jaksic, Y. Ogata, Y. Pautrat, C.-A. Pillet, arXiv:1106.3786

M. Tomamichel. Smooth Entropies: A Tutorial. QCRYPT 2012

S. Fehr. Presentation at Beyond IID Workshop. January 2013

Muller-Lennert. ETH Zurich Master's thesis. April 2013

W, Winter, Yang. arXiv:1306.1586

Müller-Lennert, Dupuis, Szehr, Fehr, Tomamichel. arXiv:1306.3142

Monotonicity of sandwiched entropy

Monotonicity holds for all *α* in [1/2, ∞]

$$\widetilde{D}_\alpha(\rho||\sigma) \geq \widetilde{D}_\alpha(\mathcal{N}(\rho)||\mathcal{N}(\sigma))$$

Frank and Lieb. arXiv:1306.5358

(*see also arXiv:1306.1586, arXiv:1306.3142, Beigi arXiv:1306.5920 for less general ranges of α*)

Evaluating the following bound

$$\chi_{\mathcal{D}}(\mathcal{N}^{\otimes n}) \geq \delta(\varepsilon||1 - 2^{-nR})$$

for the sandwiched relative entropy gives

$$p_{\rm succ} \leq 2^{-n\left(\frac{\alpha-1}{\alpha}\right)\left(R - \frac{1}{n}\widetilde{\chi}_\alpha(\mathcal{N}^{\otimes n})\right)}$$

relevant regime is α in (1,2]

SC for classical cap. of EB channels

If the sandwiched Renyi-Holevo information is **additive**, in the sense that

$$\frac{1}{n}\widetilde{\chi}_\alpha(\mathcal{N}^{\otimes n}) = \widetilde{\chi}_\alpha(\mathcal{N})$$

Then upper bound on success probability becomes

$$p_{\text{succ}} \leq 2^{-n\left(\frac{\alpha-1}{\alpha}\right)(R-\widetilde{\chi}_\alpha(\mathcal{N}))}$$

From here, we can recover the strong converse by a **now standard argument** due to Ogawa and Nagaoka (1999). (Pick *α* close enough to 1)

Goal: Show additivity!

Additivity of Rényi-Holevo information for EB channels

First observation:

Sandwiched Renyi entropy is related to a norm

$$\widetilde{D}_\alpha(\rho||\sigma) \equiv \frac{\alpha}{\alpha-1}\log_2\left\|\sigma^{(1-\alpha)/2\alpha}\,\rho\,\sigma^{(1-\alpha)/2\alpha}\right\|_\alpha$$

Next observation:

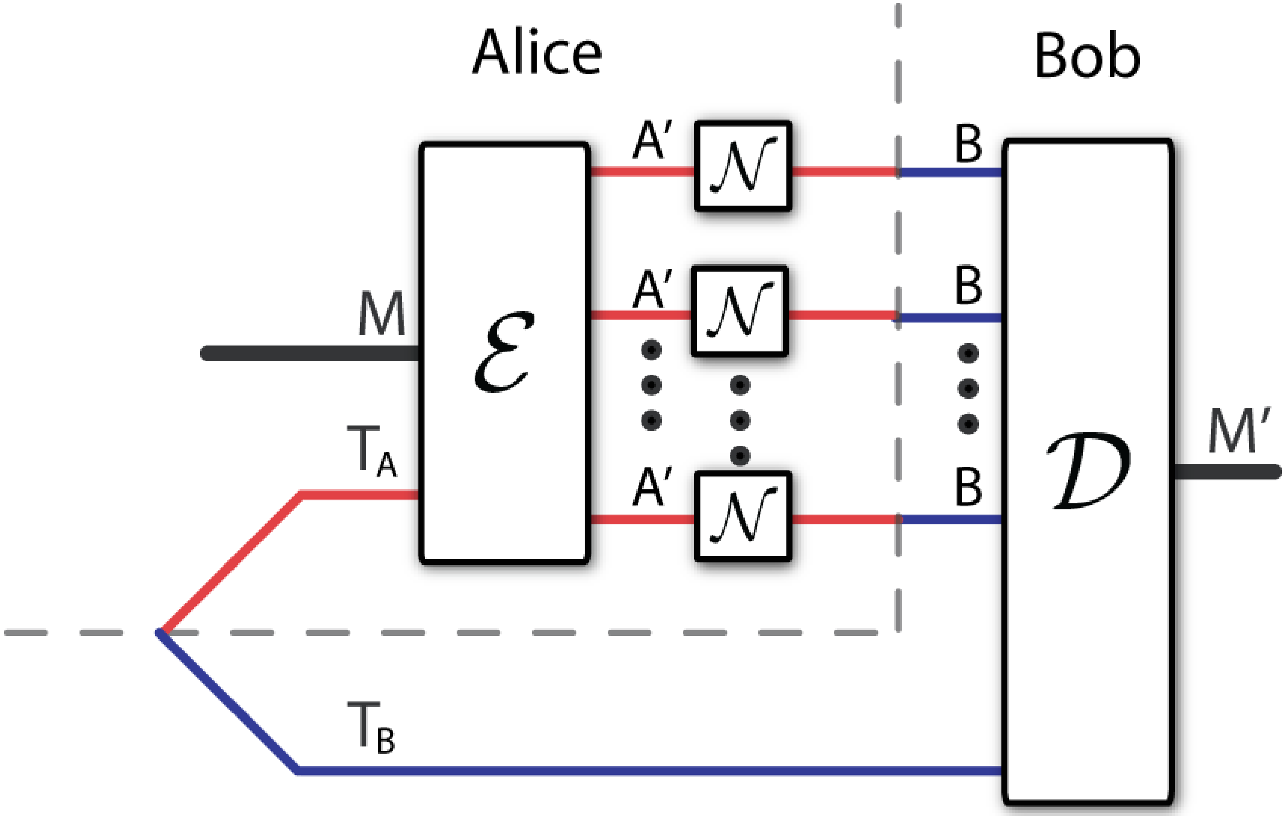
King proved that the maximum output *α*-norm is multiplicative for an EB map and any other CP map

$$\max_{\rho_{A_1A_2}}\left\|(\mathcal{M}_{\mathrm{EB}}\otimes\mathcal{M})(\rho_{A_1A_2})\right\|_\alpha = \max_\omega\left\|\mathcal{M}_{\mathrm{EB}}(\omega)\right\|_\alpha \times \max_\sigma\left\|\mathcal{M}(\sigma)\right\|_\alpha$$

Putting these two observations together (*along with a little more*) gives additivity of the Rényi-Holevo information, from which we get the **strong converse**

SC for entanglement-assisted capacity

Entanglement-assisted setting for communication



Suppose finite-dimensional channel here...

Joint work with Manish K. Gupta (LSU), arXiv:1310.7028

SC for entanglement-assisted capacity

Establish a **generalized divergence framework** by appealing to Propositions 17 and 18 of Matthews-Wehner arXiv:1210.4722

Find an **upper bound** on the success probability of any entanglement-assisted code:

$$p_{\text{succ}} \leq 2^{-n\left(\frac{\alpha-1}{\alpha}\right)\left(R-\frac{1}{n}\widetilde{I}_{\alpha}\left(\mathcal{N}^{\otimes n}\right)\right)}$$

Prove **additivity** of sandwiched quantum mutual information by appealing to multiplicativity result of Devetak, Junge, King, Ruskai in arXiv:quant-ph/0506196

Conclusions

Establishing strong converse theorems is an important step forward for QIT b/c they strengthen the interpretation of capacity and have applications in certain models of cryptography

Main open questions:

Determine whether strong converse theorems hold in other settings

What is the relation to (non-)additivity?

Characterize second-order behavior