Strong Converse and Finite Resource Trade-Offs for Quantum Channels

Marco Tomamichel†, Mario Berta, Joseph M. Renes, Mark M. Wilde, Andreas Winter

†School of Physics, The University of Sydney

Finite Resource Information Theory

- We are on the verge of engineering small, reliable quantum information processors.

![Image of quantum device](Source: Kelly et al., Nature 519, 66-69 (2015))

- It is important to understand the fundamental limits for information processing with such small quantum devices.
- We are interested in analytic and easy to evaluate formulas that characterize the trade-off between
  1. the information processing rate (in qubits per use of a resource)
  2. the tolerated error / infidelity
  3. the size of quantum devices / coding block length
Quantum Coding: Channels

- **Quantum channel**: completely positive trace-preserving linear map $\mathcal{N} \equiv \mathcal{N}_{A \rightarrow B}$ from (states on) $A$ to (states on) $B$.

  $A \xrightarrow{\mathcal{N}} B$

Assume $A$ and $B$ are finite-dimensional.

- The channel is memoryless:

  $A_1 \xrightarrow{\mathcal{N}} B_1$
  $A_2 \xrightarrow{\mathcal{N}} B_2$
  $\vdots$
  $A_n \xrightarrow{\mathcal{N}} B_n$

  $\equiv A^n \xrightarrow{\mathcal{N} \otimes n} B^n$
Quantum Coding: Encoder and Decoder

• **Entanglement transmission code** (for $\mathcal{N} \otimes n$):

$$\mathcal{C}_n = \{d_n, \mathcal{E}_n, \mathcal{D}_n\}.$$

1. **code size** $d_n$:
   - Hilbert spaces $M$, $M'$, $M''$ of dimension $d_n$.
   - maximally entangled state

$$|\phi\rangle_{MM'} = \frac{1}{\sqrt{d_n}} \sum_{i=1}^{d_n} |i\rangle_M \otimes |i\rangle_{M'}.$$

2. **encoder** $\mathcal{E}_n$: quantum channel from $M'$ to $A^n$.

3. **decoder** $\mathcal{D}_n$: quantum channel from $B^n$ to $M''$. 


Quantum Coding: Encoder and Decoder

- **Entanglement transmission code** (for $\mathcal{N}^\otimes n$):

$$C_n = \{d_n, E_n, D_n\}.$$

1. **code size** $d_n$:
   - Hilbert spaces $M, M', M''$ of dimension $d_n$.
   - maximally entangled state

$$|\phi\rangle_{MM'} = \frac{1}{\sqrt{d_n}} \sum_{i=1}^{d_n} |i\rangle_M \otimes |i\rangle_{M'}.$$

2. **encoder** $E_n$: quantum channel from $M'$ to $A^n$.

3. **decoder** $D_n$: quantum channel from $B^n$ to $M''$. 
Quantum Coding: Encoder and Decoder

- **Entanglement transmission code** (for $\mathcal{N}^\otimes n$):
  \[ C_n = \{ d_n, E_n, D_n \} \cdot \]

1. **Code size $d_n$:**
   - Hilbert spaces $M, M', M''$ of dimension $d_n$.
   - Maximally entangled state
     \[ |\phi\rangle_{MM'} = \frac{1}{\sqrt{d_n}} \sum_{i=1}^{d_n} |i\rangle_M \otimes |i\rangle_{M'} \cdot \]

2. **Encoder $E_n$:** quantum channel from $M'$ to $A^n$.

3. **Decoder $D_n$:** quantum channel from $B^n$ to $M''$. 
Quantum Coding: Entanglement Fidelity

- **Fidelity** with maximally entangled state:

\[ F(C_n, \mathcal{N}^\otimes n) = \text{tr} \left( (\mathcal{D}_n \circ \mathcal{N}^\otimes n \circ \mathcal{E}_n)(\phi_{MM'})\phi_{MM''} \right). \]
Achievable Region and Capacity

- A triple $(R, n, \varepsilon)$ is achievable on $\mathcal{N}$ if $\exists C_n$ with

\[
\frac{1}{n} \log d_n \geq R, \quad \text{and} \quad F(C_n \mathcal{N}^{\otimes n}) \geq 1 - \varepsilon.
\]

- Boundary of (non-asymptotic) achievable region:

\[
\hat{R}(n; \varepsilon, \mathcal{N}) := \max \{ R : (R, n, \varepsilon) \text{ is achievable on } \mathcal{N} \}.
\]

- The *quantum capacity*, $Q(\mathcal{N})$, is the rate at which qubits can be transmitted with fidelity approaching one asymptotically.

\[
Q_{\varepsilon}(\mathcal{N}) := \lim_{n \to \infty} \hat{R}(n; \varepsilon, \mathcal{N}), \quad \varepsilon \in (0, 1)
\]

\[
Q(\mathcal{N}) := \lim_{\varepsilon \to 0} Q_{\varepsilon}(\mathcal{N}).
\]

- These are operational quantities: the task of information theory is to relate them to (easy to evaluate) information quantities.
Quantum Capacity Theorem


\[
Q(\mathcal{N}) = \lim_{\ell \to \infty} \frac{1}{\ell} I_c(\mathcal{N} \otimes \ell), \quad \text{where}
\]
\[
I_c(\mathcal{N}) := \max_{\rho_A} \left\{ -H(A|B)_{\omega} \right\},
\]

and \( \omega_{AB} = \mathcal{N}_{A' \rightarrow B}(\psi_{A'A}^\rho) \) for the purification \( \psi_{A'A}^\rho \) of \( \rho_A \).

- This result is unsatisfactory for several reasons:
  - It is not a single-letter formula, i.e. not easier to compute than the original optimization problem.
  - We need to consider arbitrarily large \( \ell \) in general (Cubitt+’14).
  - The formula simplifies for channels which satisfy \( I_c(\mathcal{N} \otimes \ell) = \ell I_c(\mathcal{N}) \), e.g. for degradable channels like dephasing channels.
  - But even so, this does not tell us about \( \varepsilon > 0 \) and finite \( n \).
Capacity and Strong Converse

- Before we consider finite resource trade-offs, we need to fully understand the asymptotic limit $n \to \infty$.
- The first thing we would like to know:

\[
\hat{R}(n; \varepsilon, \mathcal{N}) = Q(\mathcal{N}) + o(1)
\]
State of the Art

- Prior to this work, the strong converse property could only be established for some channels with trivial capacity.
- Morgan and Winter showed that degradable quantum channels satisfy a “pretty strong” converse:

\[ Q_{\varepsilon}(\mathcal{N}) = Q(\mathcal{N}) \quad \text{for all } \varepsilon \in \left(0, \frac{1}{2}\right) \]

(Extending their proof to all \( \varepsilon \in (0, 1) \) appears difficult.)
- Strong converse rates are known, for example the entanglement-assisted capacity established via channel simulation (Bennett+’02), or the entanglement cost of a channel (Berta+’13).
- However, they are not tight except for trivial channels.
Result 1: Rains Entropy is Strong Converse Rate

- The *Rains relative entropy* of the channel is defined as
  \[ R(\mathcal{N}) := \max_{\rho_A} \min_{\sigma_{AB} \in \text{Rains}(A:B)} D(\mathcal{N}_{A' \to B}(\psi^A_{A'}A) \parallel \sigma_{AB}). \]

**Theorem**

*For any channel \( \mathcal{N} \), communication at a rate exceeding \( R(\mathcal{N}) \) implies (exponentially) vanishing fidelity.*

- Key Idea: Consider correlations \( \sigma_{AB} \) that are useless for quantum communication. Classically:
  \[ C(W) = \max_{P_X} \min_{Q_X, Q_Y} D(P_X \times W_{Y|X} \parallel Q_X \times Q_Y) = \max_{P_X} I(X : Y). \]

- A state \( \sigma_{AB} \in \text{Rains}(A : B) \) satisfies
  \[ \text{tr} \left( \phi_{AB} \sigma_{AB} \right) \leq \frac{1}{d} \quad \forall \text{ maximally entangled } \phi_{AB}. \]

- Rains used this set in the context of entanglement distillation (Rains’99).
Result 1: Covariant Channels

- The Rains relative entropy of symmetric channels simplifies.
- Covariance group of the channel $\mathcal{N}$: Group $G$ with unitary representations $U_A$ and $V_B$ such that

$$\mathcal{N}_{A\rightarrow B}(U_A(g)(\cdot)U_A^\dagger(g)) = V_B(g)\mathcal{N}_{A\rightarrow B}(\cdot)V_B^\dagger(g) \quad \forall g \in G$$

Lemma (Channel Covariance)

Let $G$ be a covariance group of $\mathcal{N}$. Then,

$$R(\mathcal{N}) = \max_{\bar{\rho}_A} \min_{\sigma_{AB}} D(\mathcal{N}_{A'\rightarrow B}(\psi_{AA'}^{\bar{\rho}}) \parallel \sigma_{AB})$$

where $\bar{\rho}_A = U_A(g)\bar{\rho}_A U_A^\dagger(g)$, i.e. $\bar{\rho}_A$ is invariant under $G$.

- Covariance group of $\mathcal{N}^\otimes n$ always contains permutations $S_n$. Thus, we can restrict to permutation invariant states $\bar{\rho}_{A^n}$.
- If the channel is covariant with regards to a one-design on $A$, the optimal state is the maximally entangled state.
Result 1: Assisted Codes

- Remains valid for codes with classical post-processing assistance.
  - Includes forward classical communication assistance (all channels).
  - Includes two-way communication assistance (covariant channels).
    - Proof via teleportation (Bennett+’96, see also Pirandola+’15).
Example: Dephasing Channels Satisfy Strong Converse

• For all quantum channels we thus have

$$I_c(N) \leq Q(N) \leq Q_\epsilon(N) \leq R(N).$$

Theorem

For generalized dephasing channels $\mathcal{Z}$, we have $I_c(\mathcal{Z}) = R(\mathcal{Z})$.

• The inequalities collapse and $Q_\epsilon(\mathcal{Z}) = Q(\mathcal{Z})$.

• Includes qubit dephasing channel:

$$\mathcal{Z}_\lambda : \rho \mapsto (1 - \lambda)\rho + \lambda \mathcal{Z} \rho \mathcal{Z},$$

with $Q_\epsilon(\mathcal{Z}_\lambda) = 1 - h(\lambda)$ for all $\epsilon \in (0, 1)$.

• One- or two-way classical assistance does not help.
Result 2: Outer Bounds on Achievable Region

**Theorem**

If the covariance group of \( \mathcal{N} \) is a one-design on \( A \), then

\[
\hat{R}(n; \varepsilon, \mathcal{N}) \leq R(\mathcal{N}) + \sqrt{\frac{V_R(\mathcal{N})}{n}} \Phi^{-1}(\varepsilon) + O\left(\frac{\log n}{n}\right)
\]

- \( V_R(\mathcal{N}) \) is (Rains) quantum channel dispersion.
- \( R(\mathcal{N}) = \min_{\sigma_{AB} \in \text{Rains}(A:B)} D(\mathcal{N}_{A' \rightarrow B}(\phi_{A'A}) \parallel \sigma_{AB}) \),
- \( V_R(\mathcal{N}) := V(\mathcal{N}_{A' \rightarrow B}(\phi_{A'A}) \parallel \sigma^*_{AB}) \),
- \( \sigma^*_{AB} \) is the minimizer of the channel Rains information,
- \( V(\rho \parallel \sigma) := \text{tr} \left( \rho (\log \rho - \log \sigma)^2 \right) - D(\rho \parallel \sigma)^2 \),
- \( \Phi^{-1}(\cdot) \) is inverse of cumulative normal distribution function.
Result 2: Outer Bounds on Achievable Region

**Theorem**

If the covariance group of $\mathcal{N}$ is a one-design on $A$, then

$$\hat{R}(n; \varepsilon, \mathcal{N}) \leq R(\mathcal{N}) + \sqrt{\frac{V_R(\mathcal{N})}{n}} \Phi^{-1}(\varepsilon) + O\left(\frac{\log n}{n}\right)$$

- $V_R(\mathcal{N})$ is (Rains) quantum channel dispersion.

$$R(\mathcal{N}) = \min_{\sigma_{AB} \in \text{Rains}(A:B)} D(\mathcal{N}_{A' \rightarrow B}(\phi_{A' A}) \parallel \sigma_{AB})$$

$$V_R(\mathcal{N}) := V(\mathcal{N}_{A' \rightarrow B}(\phi_{A' A}))$$

- $\sigma_{AB}^*$ is the minimizer of the channel Rains.
- $V(\rho \parallel \sigma) := \text{tr} \left( \rho (\log \rho - \log \sigma)^2 \right) - D(\rho \parallel \sigma)$
- $\Phi^{-1}(\cdot)$ is inverse of cumulative normal distribution.
Result 2: Inner Bound on Achievable Region

**Theorem**

For any quantum channel $\mathcal{N}$, we have

$$\hat{R}(n; \varepsilon, \mathcal{N}) \geq I_c(\mathcal{N}) + \sqrt{\frac{V_c(\mathcal{N})}{n}} \Phi^{-1}(\varepsilon) + O\left(\frac{\log n}{n}\right)$$

- $V_c(\mathcal{N})$ is (Hashing) quantum channel dispersion.
- $I_c(\mathcal{N}) = \max_{\rho_A} \left\{ D(\mathcal{N}_{A' \rightarrow B}(\psi^\rho_{AA'})) \| 1_A \otimes \mathcal{N}_{A \rightarrow B}(\rho_A) \right\}$,

$$V_c(\mathcal{N}) := V(\mathcal{N}_{A' \rightarrow B}(\psi^\rho_{A' A}) \| 1_A \otimes \mathcal{N}_{A \rightarrow B}(\rho^*_A)),$$

- $\rho^*_A$ is optimal input state for coherent information.
- This inner bound was independently established by Beigi+’15.
- Sometimes the upper and lower bounds agree...
Example: Qubit Dephasing Channel

- Bounds agree, classical assistance does not help:

$$\hat{R}(n; \varepsilon, Z_\gamma) = 1 - h(\gamma) + \sqrt{\frac{v(\gamma)}{n}} \Phi^{-1}(\varepsilon) + \frac{\log n}{2n} + O\left(\frac{1}{n}\right).$$

- Dephasing channel: $\gamma = 0.1$ and fixed fidelity $1 - \varepsilon = 95%$.
- Corresponds to binary symmetric channel (e.g. Polyanskiy+’10).
Example: Qubit Dephasing Channel

- Bounds agree, classical assistance does not help:

\[
\hat{R}(n; \varepsilon, Z_\gamma) = 1 - h(\gamma) + \sqrt{\frac{v(\gamma)}{n}} \Phi^{-1}(\varepsilon) + \log \frac{n}{2n} + O\left(\frac{1}{n}\right).
\]

- Dephasing channel: \( \gamma = 0.1 \) and fixed fidelity \( 1 - \varepsilon = 95\% \).
- Corresponds to binary symmetric channel (e.g. Polyanskiy+’10).
Example: Qubit Erasure Channel

- Erasure channel $\mathcal{E}_\beta : \rho \mapsto (1 - \beta)\rho + \beta|k\rangle\langle k|$. 
- Bounds agree if we allow two-way classical assistance:

$$\hat{R}(n; \varepsilon, \mathcal{E}_\beta) = 1 - \beta + \sqrt{\frac{\beta(1 - \beta)}{n}} \Phi^{-1}(\varepsilon) + O\left(\frac{1}{n}\right).$$

- Erasure channel: $\beta = 0.25$ and $1 - \varepsilon = 99\%$. 
Example: Qubit Depolarizing Channel

- Depolarizing channel: \( \rho \mapsto (1 - \alpha)\rho + \frac{\alpha}{3}(X\rho X + Y\rho Y + Z\rho Z) \).

- Exact outer bound for \( \alpha = 0.0825 \) and \( \epsilon = 5.5\% \).
- Inner bounds: unassisted, outer bounds: two-way assisted
Step 1: Arimoto-Type (One-Shot) Converse Bounds

- Consider \( \mathcal{C} = \{d, \mathcal{E}, \mathcal{D}\} \) for \( \mathcal{N} \) with \( F(\mathcal{C}, \mathcal{N}) \geq 1 - \varepsilon \).
- Test if a state is \( \phi_{MM''} \), or not:
  \[
  T(\cdot) = p|0\rangle\langle 0| + (1 - p)|1\rangle\langle 1|, \quad p = \text{tr} \left( \phi_{MM''} \cdot \right).
  \]
- Let \( \rho_{AM} = \mathcal{E}(\phi_{MM'}) \). Due to data-processing, we have
  \[
  \min_{\sigma_{BM}} D(\mathcal{N}(\rho_{AM})\|\sigma_{BM}) \geq \min_{\sigma_{BM}} D(T \circ \mathcal{D} \circ \mathcal{N}(\rho_{AM})\|T \circ \mathcal{D}(\sigma_{BM}))
  \]
  for any divergence satisfying data-processing.
- The latter quantity can be bounded using
  \[
  \langle 0|T \circ \mathcal{D} \circ \mathcal{N}(\rho_{AM})|0\rangle \geq 1 - \varepsilon, \quad \langle 0|T \circ \mathcal{D}(\sigma_{RB})|0\rangle \leq \frac{1}{d}.
  \]
- Second order: use divergence related to hypothesis testing, \( D^\varepsilon_H \), and its asymptotic expansion (T+Hayashi’13,Li’14).
- Strong converse: use sandwiched Rényi divergence, \( \tilde{D}_\alpha \).
Step 2: Asymptotics for Strong Converse

Lemma

Optimizing over codes we have the following one-shot converse:

\[ \hat{R}(1; \varepsilon, \mathcal{N}) \leq \max_{\rho_A} \min_{\sigma_{AB}} \tilde{D}_\alpha (\mathcal{N}_{A' \rightarrow B} (\psi^{\rho}_{AA'}) \| \sigma_{AB}) + \frac{\alpha \log \frac{1}{1-\varepsilon}}{\alpha - 1} \]

- This yields an upper bound on the $\varepsilon$-capacity:

\[ Q_\varepsilon (\mathcal{N}) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \max_{\rho_{A^n}} \min_{\sigma_{A^n B^n}} \tilde{D}_\alpha (\mathcal{N}^{\otimes n} (\psi^{\rho}_{A^n A'^n}) \| \sigma_{A^n B^n}) \leq \hat{R}_\alpha (\mathcal{N}^{\otimes n}) \]

- We can restrict optimization to permutation invariant $\rho_{A^n}$.
- It remains to show that $\hat{R}_\alpha (\mathcal{N})$ satisfies an asymptotic sub-additivity property, i.e. $\hat{R}_\alpha (\mathcal{N}^{\otimes n}) \leq n\hat{R}_\alpha (\mathcal{N}) + o(n)$. 
Step 3: Asymptotic Sub-Additivity

• Employ the fact that \( \psi_{A^n A'}^\rho \) is in the symmetric subspace:

\[
\psi_{A A'}^\rho \leq P_{A^n R^n}^{\text{symm}} \leq n|A|^2 \int d\mu(\theta) \theta^{\otimes n}_{A R}.
\]

• The quantum way to restrict to product states in the converse.

• This allows us to show (skipping a few technical steps) that

\[
\tilde{R}_\alpha(\mathcal{N}^{\otimes n}) \leq n\tilde{R}_\alpha(\mathcal{N}) + O(\log(n)).
\]

• Hence, \( Q_\varepsilon(\mathcal{N}) \leq \tilde{R}_\alpha(\mathcal{N}) \) for all \( \alpha > 1 \).

• And, thus, by continuity as \( \alpha \to 1 \), we find \( Q_\varepsilon(\mathcal{N}) \leq R(\mathcal{N}) \).

• A more detailed analysis reveals that the fidelity converges exponentially fast to 0 for any \( d > R(\mathcal{N}) \).
Conclusion

- The (asymptotic) capacity is insufficient to characterize information transmission over quantum channels in realistic settings.
- However, using the channel dispersion, we can characterize the achievable region using only two parameters.
- These approximations agree very well with numerical results already for small instances.

Open Questions:
- Strong converse for all degradable channels.
- Find second order outer bound for general (not only covariant) channels.
- Find better inner bounds for two-way assisted achievable region.
- Consider other important qubit channels, e.g., amplitude damping.