Quantum conditional mutual information and approximate Markov chains

Omar Fawzi

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

ENS DE LYON

Caltech

Banff, July 7th, 2015

Joint works with Renato Renner and David Sutter and Renato Renner
arXiv:1410.0664 and arXiv:1504.07251

Entropy and conditioning

- State ρ_A acting on A

$$0 \leq H(A)_\rho \leq \log |A|$$

Entropy and conditioning

- State ρ_A acting on A

$$0 \leq H(A)_\rho \leq \log |A|$$

- State ρ_{AB} acting on $A \otimes B$
  Conditional entropy of A from B's viewpoint

$$H(A|B)_\rho = H(AB)_\rho - H(B)_\rho$$

  Interpretation:
  - Classical B: $\rho_{AB} = \sum_b \rho_A(b) \otimes p(b)|b\rangle\langle b|$

$$H(A|B)_\rho = \sum_b p(b) H(A)_{\rho(b)}$$

  - Quantum B: More subtle
    $H(A|B)_\rho$ can be negative when ρ entangled

$$-\log |A| \leq H(A|B)_\rho \leq \log |A|$$

Mutual information and conditioning

- State ρ_{AB} acting on $A \otimes B$
  Mutual Information:

$$I(A:B)_\rho = H(A)_\rho - H(A|B)_\rho$$

  - Classical ρ: $0 \leq I(A:B)_\rho \leq \min\{\log|A|, \log|B|\}$
  - Quantum ρ: $0 \leq I(A:B)_\rho \leq 2\min\{\log|A|, \log|B|\}$

- State ρ_{ABC} acting on $A \otimes B \otimes C$
  Conditional Mutual Information:

$$I(A:C|B)_\rho = H(C|B)_\rho - H(C|AB)_\rho$$

  - Classical B: $\rho_{ABC} = \sum_b \rho_{AC}(b) \otimes p(b)|b\rangle\langle b|_B$

$$I(A:C|B)_\rho = \sum_b p(b) I(A:C)_{\rho(b)} \in [0, \min\{\log|A|, \log|C|\}]$$

  - Quantum B: More subtle

$$0 \leq I(A:C|B)_\rho \leq 2\min\{\log|A|, \log|C|\}$$

Motivation 1: Operational significance

Optimal rates for information processing tasks

- **Compression** of source A:

$$A \overset{encode}{\Longrightarrow} M \overset{decode}{\Longrightarrow} A$$

Min. rate for compressing $A^{\otimes n}$: $\frac{\log|M|}{n} \to H(A)$

- **Reliable communication** over channel $\mathcal{N} : A \to B$:

$$M \overset{encode}{\Longrightarrow} A \to \mathcal{N} \to B \overset{decode}{\Longrightarrow} M$$

Max. rate of transmission over $\mathcal{N}^{\otimes n}$: $\frac{\log|M|}{n} \to \max_A H(A) - H(A|B)$

- **Randomness extraction** from source A:

$$A \overset{extract}{\Longrightarrow} U$$

Max. rate of extractable randomness from $A^{\otimes n}$: $\frac{\log|U|}{n} \to H(A)$

- Entanglement manipulation, State merging, etc...

Motivation 2: Entropy as proof tool

Very useful as a proof tool

- Bounds on Random Access Codes
- Direct sum results in communication complexity
- de Finetti theorems
- Entanglement measures (squashed entanglement)
- Entanglement in many-body systems
- ...

Motivation 2: Entropy as proof tool

Very useful as a proof tool

- Bounds on Random Access Codes
- Direct sum results in communication complexity
- de Finetti theorems
- Entanglement measures (squashed entanglement)
- Entanglement in many-body systems
- ...

Von Neumann entropy very useful as a proof tool because of

Chain rule

$$I(A_1 \dots A_n : C|B)$$
$$= I(A_1 : C|B) + I(A_2 : C|BA_1) + \cdots + I(A_n : C|BA_1 \dots A_{n-1})$$

Can decompose correlations into parts

Application

Intuition: loosing one bit cannot harm too much

Theorem

For any joint distribution $P_{X_1 \ldots X_n YZ}$ where Y is a bit. There exists a subset $L \subset \{1, \ldots, n\}$ such that $|L| \leq 1/\epsilon$

$$\textit{for all } i, \quad \mathrm{P}_{\mathrm{guess}}(X_i|ZX_L) \geq \mathrm{P}_{\mathrm{guess}}(X_i|ZY) - \sqrt{(2 \ln 2)\epsilon}$$

Interpretation: Y can be replaced by a small number of bits of $X_1 \ldots X_n$

Application

Intuition: loosing one bit cannot harm too much

Theorem

For any joint distribution $P_{X_1 \ldots X_n YZ}$ where Y is a bit. There exists a subset $L \subset \{1, \ldots, n\}$ such that $|L| \leq 1/\epsilon$

$$\textit{for all } i, \quad \mathrm{P}_{\mathrm{guess}}(X_i|ZX_L) \geq \mathrm{P}_{\mathrm{guess}}(X_i|ZY) - \sqrt{(2 \ln 2)\epsilon}$$

Interpretation: Y can be replaced by a small number of bits of $X_1 \ldots X_n$

Algorithm to construct L

$L \leftarrow \emptyset$

while $\exists i \in \{1, \ldots, n\}$ st $I(X_i : Y|ZX_L) > \epsilon$

$\quad L \leftarrow L \cup \{i\}$

Application

Intuition: loosing one bit cannot harm too much

Theorem

For any joint distribution $P_{X_1 \ldots X_n YZ}$ where Y is a bit. There exists a subset $L \subset \{1, \ldots, n\}$ such that $|L| \leq 1/\epsilon$

$$\textit{for all } i, \quad \mathrm{P}_{\mathrm{guess}}(X_i|ZX_L) \geq \mathrm{P}_{\mathrm{guess}}(X_i|ZY) - \sqrt{(2 \ln 2)\epsilon}$$

Interpretation: Y can be replaced by a small number of bits of $X_1 \ldots X_n$

Algorithm to construct L

$L \leftarrow \emptyset$

while $\exists i \in \{1, \ldots, n\}$ st $I(X_i : Y|ZX_L) > \epsilon$

  $L \leftarrow L \cup \{i\}$

Claim 1: The algorithm terminates in $< 1/\epsilon$ steps

Claim 2: When algorithm terminates, $\mathrm{P}_{\mathrm{guess}}(X_i|ZX_L) \gtrapprox_\epsilon \mathrm{P}_{\mathrm{guess}}(X_i|ZY)$

Application (Proof of claim 1)

Theorem

For any joint distribution $P_{X_1 \ldots X_n YZ}$ where Y is a bit. There exists a subset $L \subset \{1, \ldots, n\}$ such that $|L| \leq 1/\epsilon$

$$\textit{for all } i, \quad \mathrm{P}_{\mathrm{guess}}(X_i|ZX_L) \geq \mathrm{P}_{\mathrm{guess}}(X_i|ZY) - \sqrt{(2 \ln 2)\epsilon}$$

Algorithm to construct L
$L \leftarrow \emptyset$
while $\exists i \in \{1, \ldots, n\}$ st $I(X_i : Y|ZX_L) > \epsilon$
  $L \leftarrow L \cup \{i\}$

Claim 1: The algorithm terminates in at most $1/\epsilon$ steps
$L = \{i_1, \ldots, i_\ell\}$

$$I(X_L : Y|Z) = \sum_{p=1}^{\ell} I(X_{i_p} : Y|ZX_{i_1 \ldots i_{p-1}}) \geq \ell \cdot \epsilon$$

But $I(X_L : Y|Z) \leq 1$ because Y is one bit

So $\ell \leq \frac{1}{\epsilon}$ □

Application (Proof of claim 2)

Theorem

For any joint distribution $P_{X_1 \ldots X_n YZ}$ where Y is a bit. There exists a subset $L \subset \{1, \ldots, n\}$ such that $|L| \leq 1/\epsilon$

$$\textit{for all } i, \quad \mathrm{P}_{\mathrm{guess}}(X_i|ZX_L) \geq \mathrm{P}_{\mathrm{guess}}(X_i|ZY) - \sqrt{(2 \ln 2)\epsilon}$$

Algorithm to construct L
$L \leftarrow \emptyset$
while $\exists i \in \{1, \ldots, n\}$ st $I(X_i : Y|ZX_L) > \epsilon$
  $L \leftarrow L \cup \{i\}$

Claim 2: When algorithm terminates, $\mathrm{P}_{\mathrm{guess}}(X_i|ZX_L) \gtrapprox_\epsilon \mathrm{P}_{\mathrm{guess}}(X_i|ZY)$
For all i,

$$\begin{aligned}
\epsilon \geq I(X_i : Y|ZX_L) &= \mathop{\mathbb{E}}_{zx_L} \left\{ I(X_i : Y)_{P_{X_i Y|zx_L}} \right\} \\
&\geq \mathop{\mathbb{E}}_{zx_L} \left\{ \frac{1}{2 \ln 2} \left\| P_{X_i Y|zx_L} - P_{X_i|zx_L} \times P_{Y|zx_L} \right\|_1^2 \right\} \\
&\geq \frac{1}{2 \ln 2} \left(\mathop{\mathbb{E}}_{zx_L} \left\{ \left\| P_{X_i|zx_L y} - P_{X_i|zx_L} \right\|_1 \right\} \right)^2 \\
&\geq \frac{1}{2 \ln 2} \left(\mathrm{P}_{\mathrm{guess}}(X_i|ZX_L Y) - \mathrm{P}_{\mathrm{guess}}(X_i|ZX_L) \right)^2
\end{aligned}$$

$$\mathrm{P}_{\mathrm{guess}}(X_i|ZX_L) \geq \mathrm{P}_{\mathrm{guess}}(X_i|ZX_L Y) - \sqrt{(2 \ln 2)\epsilon} \geq \mathrm{P}_{\mathrm{guess}}(X_i|ZY) - \sqrt{(2 \ln 2)\epsilon} \quad \square$$

What if the systems are quantum

Wanted

For any **quantum** *density operator* $\rho_{X_1 \ldots X_n C B}$ *where* C *is a qubit. There exists a subset* $L \subset \{1, \ldots, n\}$ *such that* $|L| \leq 1/\epsilon$

$$\textit{for all } i, \quad \mathrm{P}_{\mathrm{guess}}(X_i|BX_L) \geq \mathrm{P}_{\mathrm{guess}}(X_i|BC) - \sqrt{(2 \ln 2)\epsilon}$$

Algorithm to construct L
$L \leftarrow \emptyset$
while $\exists i \in \{1, \ldots, n\}$ st $I(X_i : C|BX_L) > \epsilon$
  $L \leftarrow L \cup \{i\}$

Claim 1: The algorithm terminates in at most $1/\epsilon$ steps
Only used **chain rule** and $I(X : C|B) \leq \log |C|$, which still holds quantum ✓

What if the systems are quantum

Wanted

For any **quantum** *density operator* $\rho_{X_1 \ldots X_n C B}$ *where* C *is a qubit. There exists a subset* $L \subset \{1, \ldots, n\}$ *such that* $|L| \leq 1/\epsilon$

$$\textit{for all } i, \quad \mathrm{P}_{\mathrm{guess}}(X_i|BX_L) \geq \mathrm{P}_{\mathrm{guess}}(X_i|BC) - \sqrt{(2 \ln 2)\epsilon}$$

Algorithm to construct L
$L \leftarrow \emptyset$
while $\exists i \in \{1, \ldots, n\}$ st $I(X_i : C|BX_L) > \epsilon$
  $L \leftarrow L \cup \{i\}$

Claim 1: The algorithm terminates in at most $1/\epsilon$ steps

Only used **chain rule** and $I(X : C|B) \leq \log |C|$, which still holds quantum ✓

Claim 2: When algorithm terminates, $\mathrm{P}_{\mathrm{guess}}(X_i|BX_L) \gtrapprox_\epsilon \mathrm{P}_{\mathrm{guess}}(X_i|BC)$

For all i,

$$\epsilon \geq I(X_i : Y|ZX_L) = \mathop{\mathbb{E}}_{zx_L} \left\{ I(X_i : Y)_{P_{X_i Y|zx_L}} \right\} \quad \text{? quantum ?}$$

$$\geq \mathop{\mathbb{E}}_{zx_L} \left\{ \frac{1}{2 \ln 2} \left\| P_{X_i Y|zx_L} - P_{X_i|zx_L} \times P_{Y|zx_L} \right\|_1^2 \right\} \quad \text{? quantum ?}$$

$$\geq \frac{1}{2 \ln 2} \left(\mathrm{P}_{\mathrm{guess}}(X_i|ZX_L Y) - \mathrm{P}_{\mathrm{guess}}(X_i|ZX_L) \right)^2$$

Quantum conditional mutual information

$$\epsilon \geq I(X_i : Y|ZX_L) = \mathop{\mathbb{E}}_{zx_L} \left\{ I(X_i : Y)_{P_{X_i Y|zx_L}} \right\}$$
$$\geq \mathop{\mathbb{E}}_{zx_L} \left\{ \frac{1}{2 \ln 2} \left\| P_{X_i Y|zx_L} - P_{X_i|zx_L} \times P_{Y|zx_L} \right\|_1^2 \right\}$$
$$\geq \frac{1}{2 \ln 2} \left(\mathrm{P}_{\mathrm{guess}}(X_i|ZX_L Y) - \mathrm{P}_{\mathrm{guess}}(X_i|ZX_L) \right)^2$$

Quantum version: Do we still have

$$I(X_i : C|BX_L) \geq c \left(\mathrm{P}_{\mathrm{guess}}(X_i|BX_L C) - \mathrm{P}_{\mathrm{guess}}(X_i|BX_L) \right)^2 \quad ?$$

Quantum conditional mutual information

$$\begin{aligned}\epsilon \geq I(X_i : Y|ZX_L) &= \mathop{\mathbb{E}}_{zx_L} \left\{ I(X_i : Y)_{P_{X_i Y|zx_L}} \right\} \\ &\geq \mathop{\mathbb{E}}_{zx_L} \left\{ \frac{1}{2 \ln 2} \left\| P_{X_i Y|zx_L} - P_{X_i|zx_L} \times P_{Y|zx_L} \right\|_1^2 \right\} \\ &\geq \frac{1}{2 \ln 2} \left(\mathrm{P}_{\mathrm{guess}}(X_i|ZX_L Y) - \mathrm{P}_{\mathrm{guess}}(X_i|ZX_L) \right)^2\end{aligned}$$

Quantum version: Do we still have

$$I(X_i : C|BX_L) \geq c \left(\mathrm{P}_{\mathrm{guess}}(X_i|BX_L C) - \mathrm{P}_{\mathrm{guess}}(X_i|BX_L) \right)^2 \quad ?$$

Objective:

Structure of states ρ_{ABC} for which $I(A : C\|B)_\rho \leq \epsilon$

Wanted: A and C are approx. independent from B's point of view

Small QCMI: $\epsilon = 0$ case

Theorem (Strong subadditivity [Lieb, Ruskai, 1973])

For all quantum states ρ, $I(A:C|B)_\rho \geq 0$

Theorem (QCMI and Markov chains [Petz, 1988])

$$I(A:C|B)_\rho = 0 \quad \Leftrightarrow \quad \exists \mathcal{T} : B \to BC,\ (\mathcal{I}_A \otimes \mathcal{T})(\rho_{AB}) = \rho_{ABC}$$

C can be generated by acting on B

Structure of $\mathcal{T}$: $\mathcal{T}(\gamma) = \rho_{BC}^{1/2} \rho_B^{-1/2} (\gamma \otimes \mathrm{id}_C) \rho_B^{-1/2} \rho_{BC}^{1/2}$

Theorem (QCMI and Markov chains [Hayden, Jozsa, Petz, Winter, 2003])

$$I(A:C|B)_\rho = 0 \quad \Leftrightarrow \quad \rho_{ABC} = \bigoplus_j q_j \rho_{Ab_j^L} \otimes \rho_{b_j^R C}$$

Corollary: $I(A:C|B)_\rho = 0 \Rightarrow$ the state ρ_{AC} is separable

Small QCMI: $\epsilon > 0$ case

Wanted: State-dependent strengthening of $I(A:C|B)_\rho \geq 0$

Theorem (Remainder term for SSA [Carlen, Lieb, 2014] see also [Zhang, Wu 2014])

$$I(A:C|B)_\rho \geq \operatorname{tr}\left[\sqrt{\rho_{ABC}} - \exp\left(\frac{1}{2}\log\rho_{AB} - \frac{1}{2}\log\rho_B + \frac{1}{2}\log\rho_{BC}\right)\right]^2$$

Problem: Term $\exp(\log + \log)$ difficult to interpret **operationally**

Small QCMI: $\epsilon > 0$ case

Wanted: State-dependent strengthening of $I(A:C|B)_\rho \geq 0$

Theorem (Remainder term for SSA [Carlen, Lieb, 2014] see also [Zhang, Wu 2014])

$$I(A:C|B)_\rho \geq \operatorname{tr}\left[\sqrt{\rho_{ABC}} - \exp\left(\frac{1}{2}\log\rho_{AB} - \frac{1}{2}\log\rho_B + \frac{1}{2}\log\rho_{BC}\right)\right]^2$$

Problem: Term $\exp(\log + \log)$ difficult to interpret **operationally**

Theorem (Faithful squashed entanglement [Brandao, Christandl, Yard, 2010])

$$I(A:C|B)_\rho \geq \min_{\sigma_{AC}\ \textit{separable}} \frac{1}{8\ln 2}\|\rho_{AC} - \sigma_{AC}\|^2_{LOCC}$$

Problem: Bound is independent of B, value 0 when A or C classical

Small QCMI: $\epsilon > 0$ case

ρ_{ABC} is a **quantum Markov chain:** $\exists \mathcal{T} : B \to BC,\ (\mathcal{I}_A \otimes \mathcal{T})(\rho_{AB}) = \rho_{ABC}$

Candidate conjecture 1:

$$I(A:C|B)_\rho \leq \epsilon \quad \Rightarrow \quad \rho_{ABC} \approx_{f(\epsilon)} \omega_{ABC}, \text{ with } \omega_{ABC} \text{ Markov chain}$$

Counterexamples [Ibinson, Linden, Winter, 2006] and [Christandl, Schuch, Winter, 2012]
$\to f$ has to depend on dimensions

Small QCMI: $\epsilon > 0$ case

ρ_{ABC} is a **quantum Markov chain:** $\exists \mathcal{T} : B \to BC,\ (\mathcal{I}_A \otimes \mathcal{T})(\rho_{AB}) = \rho_{ABC}$

Candidate conjecture 1:

$$I(A:C|B)_\rho \leq \epsilon \quad \Rightarrow \quad \rho_{ABC} \approx_{f(\epsilon)} \omega_{ABC}, \text{ with } \omega_{ABC} \text{ Markov chain}$$

Counterexamples [Ibinson, Linden, Winter, 2006] and [Christandl, Schuch, Winter, 2012]
$\to f$ has to depend on dimensions

Candidate conjecture 2:

[Li, Winter, 2012], [Kim, 2013], [Zhang, 2013], [Berta, Seshadreesan, Wilde, 2014]

$$I(A:C|B)_\rho \leq \epsilon \quad \Rightarrow \quad \exists \mathcal{T} : B \to BC,\ (\mathcal{I}_A \otimes \mathcal{T})(\rho_{AB}) \approx_\epsilon \rho_{ABC}$$

with $\mathcal{T}(\gamma) = \rho_{BC}^{1/2} \rho_B^{-1/2} (\gamma \otimes \mathrm{id}_C) \rho_B^{-1/2} \rho_{BC}^{1/2}$

Remarks:

- **Conj. 1** and **Conj. 2** are true for **classical** states
- General **quantum** case: **Conj. 2** does **not** imply **Conj. 1**

Main result

A proof of a variant of **Conj. 2**

Theorem

For any ρ_{ABC}, there exists $\mathcal{T} : B \to BC$ such that

$$I(A:C|B)_\rho \geq -2 \log F(\rho_{ABC}, \mathcal{T}(\rho_{AB}))$$

Remarks:

- $F(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1$ is the fidelity
- Implies $I(A:C|B)_\rho \geq \frac{1}{4 \ln 2} \|\rho_{ABC} - \mathcal{T}(\rho_{AB})\|_1^2$
- Structure of the map $\mathcal{T}$

$$\mathcal{T}(\gamma) = V_{BC} \rho_{BC}^{1/2} \rho_B^{-1/2} U_B (\gamma \otimes \mathrm{id}_C) U_B^\dagger \rho_B^{-1/2} \rho_{BC}^{1/2} V_{BC}^\dagger$$

Back to our application

Theorem

For any ρ_{ABC}, there exists $\mathcal{T} : B \to BC$ such that

$$I(A:C|B)_\rho \geq \frac{1}{4 \ln 2} \|\rho_{ABC} - \mathcal{T}(\rho_{AB})\|_1^2$$

- Question was structure of states ρ_{ABC} with $I(A:C|B)_\rho \leq \epsilon$
  $\Longrightarrow$ states for which C can be approximately reconstructed from B

Back to our application

Theorem

For any ρ_{ABC}, there exists $\mathcal{T} : B \to BC$ such that

$$I(A:C|B)_\rho \geq \frac{1}{4 \ln 2} \|\rho_{ABC} - \mathcal{T}(\rho_{AB})\|_1^2$$

- Question was structure of states ρ_{ABC} with $I(A:C|B)_\rho \leq \epsilon$
  $\implies$ states for which C can be approximately reconstructed from B
- Recall desired inequality of the form

$$I(X:C|B)_\rho \leq \epsilon \quad \implies \quad \mathrm{P}_{\mathrm{guess}}(X|BC)_\rho - \mathrm{P}_{\mathrm{guess}}(X|B)_\rho \leq f(\epsilon)$$

Back to our application

Theorem

For any ρ_{ABC}, there exists $\mathcal{T} : B \to BC$ such that

$$I(A:C|B)_\rho \geq \frac{1}{4 \ln 2} \|\rho_{ABC} - \mathcal{T}(\rho_{AB})\|_1^2$$

- Question was structure of states ρ_{ABC} with $I(A:C|B)_\rho \leq \epsilon$
  $\implies$ states for which C can be approximately reconstructed from B
- Recall desired inequality of the form

$$I(X:C|B)_\rho \leq \epsilon \quad \implies \quad \mathrm{P}_{\mathrm{guess}}(X|BC)_\rho - \mathrm{P}_{\mathrm{guess}}(X|B)_\rho \leq f(\epsilon)$$

Strategy for guessing X from B

1. Apply $\mathcal{T}$ getting $\sigma_{XBC} \approx_\delta \rho_{XBC}$ with $\delta = \sqrt{(4 \ln 2)\epsilon}$
2. Pretend the state was ρ_{XBC} and use its optimal strategy

$$\mathrm{P}_{\mathrm{guess}}(X|B)_\rho \geq \mathrm{P}_{\mathrm{guess}}(X|BC)_\rho - \delta$$

Main result: proof sketch

Statement to prove:

$$\exists \mathcal{T} : B \to BC, \quad F(\rho_{ABC}, \mathcal{T}(\rho_{AB})) \geq 2^{-\frac{1}{2} I(A:C|B)}$$

1. Easy special case: **flat marginals** $\rho_B = \frac{\Pi_B}{r_B}$ and $\rho_{BC} = \frac{\Pi_{BC}}{r_{BC}}$

$$F(\rho_{ABC}, \rho_{BC}^{1/2} \rho_B^{-1/2} \rho_{AB} \rho_B^{-1/2} \rho_{BC}^{1/2}) = \sqrt{\frac{r_{BC}}{r_B}} F(\rho_{ABC}, \Pi_{BC} \Pi_B \rho_{AB} \Pi_B \Pi_{BC})$$
$$\geq 2^{-\frac{1}{2}(H(BC)_\rho - H(B)_\rho)} 2^{-\frac{1}{2} D(\rho_{ABC} || \rho_{AB} \otimes \mathrm{id}_C)} = 2^{-\frac{1}{2} I(A:C|B)_\rho}$$

2. General case $\to \approx$ flat marginals: **study** $\rho^{\otimes n}$ and consider types

$$I(A:C|B)_\rho = \frac{I(A^n : C^n | B^n)_{\rho^{\otimes n}}}{n}$$

   Obtain $\mathcal{T}^n_{B^n \to B^n C^n}$ such that $F(\rho_{ABC}^{\otimes n}, \mathcal{T}^n(\rho_{AB}^{\otimes n})) \geq 2^{-\frac{1}{2} I(A^n:C^n|B^n)_{\rho^{\otimes n}}}$

3. If $\mathcal{T}^n_{B^n \to B^n C^n} = \mathcal{T}^{\otimes n}_{B \to BC}$, done.
   For that, **de Finetti reduction**: $\mathcal{T}^n \leq \mathrm{poly}(n) \int \mathcal{T}^{\otimes n} d\mathcal{T}$

Main result: proof sketch

Statement to prove:

$$\exists \mathcal{T} : B \to BC, \quad F(\rho_{ABC}, \mathcal{T}(\rho_{AB})) \geq 2^{-\frac{1}{2} I(A:C|B)}$$

1. Easy special case: **flat marginals** $\rho_B = \frac{\Pi_B}{r_B}$ and $\rho_{BC} = \frac{\Pi_{BC}}{r_{BC}}$

$$F(\rho_{ABC}, \rho_{BC}^{1/2} \rho_B^{-1/2} \rho_{AB} \rho_B^{-1/2} \rho_{BC}^{1/2}) = \sqrt{\frac{r_{BC}}{r_B}} F(\rho_{ABC}, \Pi_{BC} \Pi_B \rho_{AB} \Pi_B \Pi_{BC})$$
$$\geq 2^{-\frac{1}{2}(H(BC)_\rho - H(B)_\rho)} 2^{-\frac{1}{2} D(\rho_{ABC} || \rho_{AB} \otimes \mathrm{id}_C)} = 2^{-\frac{1}{2} I(A:C|B)_\rho}$$

2. General case $\to \approx$ flat marginals: **study** $\rho^{\otimes n}$ and consider types

$$I(A:C|B)_\rho = \frac{I(A^n : C^n | B^n)_{\rho^{\otimes n}}}{n}$$

   Obtain $\mathcal{T}^n_{B^n \to B^n C^n}$ such that $F(\rho_{ABC}^{\otimes n}, \mathcal{T}^n(\rho_{AB}^{\otimes n})) \geq 2^{-\frac{1}{2} I(A^n:C^n|B^n)_{\rho^{\otimes n}}}$

3. If $\mathcal{T}^n_{B^n \to B^n C^n} = \mathcal{T}^{\otimes n}_{B \to BC}$, done.
   For that, **de Finetti reduction**: $\mathcal{T}^n \leq \mathrm{poly}(n) \int \mathcal{T}^{\otimes n} d\mathcal{T}$

Graph of known proofs $\to$ board

The recovery map $\mathcal{T}$

Recovery maps that satisfy inequality:

- $\mathcal{T}(\gamma) = V_{BC}\rho_{BC}^{1/2}\rho_B^{-1/2}U_B(\gamma \otimes \mathrm{id}_C)U_B^\dagger\rho_B^{-1/2}\rho_{BC}^{1/2}V_{BC}^\dagger$
- Optimal map $\mathcal{T}^*$: maximizes $F(\rho_{ABC}, \mathcal{T}(\rho_{AB}))$ (SDP)

The recovery map $\mathcal{T}$

Recovery maps that satisfy inequality:

- $\mathcal{T}(\gamma) = V_{BC}\rho_{BC}^{1/2}\rho_B^{-1/2}U_B(\gamma \otimes \mathrm{id}_C)U_B^\dagger\rho_B^{-1/2}\rho_{BC}^{1/2}V_{BC}^\dagger$
- Optimal map $\mathcal{T}^*$: maximizes $F(\rho_{ABC}, \mathcal{T}(\rho_{AB}))$ (SDP)

Question: Petz map $\mathcal{T}_{Petz}(\gamma) = \rho_{BC}^{1/2}\rho_B^{-1/2}(\gamma \otimes \mathrm{id}_C)\rho_B^{-1/2}\rho_{BC}^{1/2}$?

- Petz map is square-root optimal if ρ_{ABC} pure [Barnum, Knill, 2000]

The recovery map $\mathcal{T}$

Recovery maps that satisfy inequality:

- $\mathcal{T}(\gamma) = V_{BC}\rho_{BC}^{1/2}\rho_B^{-1/2}U_B(\gamma \otimes \mathrm{id}_C)U_B^\dagger\rho_B^{-1/2}\rho_{BC}^{1/2}V_{BC}^\dagger$
- Optimal map $\mathcal{T}^*$: maximizes $F(\rho_{ABC}, \mathcal{T}(\rho_{AB}))$ (SDP)

Question: Petz map $\mathcal{T}_{Petz}(\gamma) = \rho_{BC}^{1/2}\rho_B^{-1/2}(\gamma \otimes \mathrm{id}_C)\rho_B^{-1/2}\rho_{BC}^{1/2}$?

- Petz map is square-root optimal if ρ_{ABC} pure [Barnum, Knill, 2000]
- ... but not in general

The recovery map $\mathcal{T}$

Recovery maps that satisfy inequality:

- $\mathcal{T}(\gamma) = V_{BC}\rho_{BC}^{1/2}\rho_B^{-1/2}U_B(\gamma \otimes \mathrm{id}_C)U_B^\dagger\rho_B^{-1/2}\rho_{BC}^{1/2}V_{BC}^\dagger$
- Optimal map $\mathcal{T}^*$: maximizes $F(\rho_{ABC}, \mathcal{T}(\rho_{AB}))$ (SDP)

Question: Petz map $\mathcal{T}_{Petz}(\gamma) = \rho_{BC}^{1/2}\rho_B^{-1/2}(\gamma \otimes \mathrm{id}_C)\rho_B^{-1/2}\rho_{BC}^{1/2}$?

- Petz map is square-root optimal if ρ_{ABC} pure [Barnum, Knill, 2000]
- ... but not in general

Observe that $\mathcal{T}_{Petz}$ only depends on ρ_{BC}

Theorem

For any ρ_{BC}, there exists a map $\mathcal{T}_{B\to BC}$ such that for any extension ρ_{ABC},

$$I(A:C|B)_\rho \geq -2\log F(\rho_{ABC}, \mathcal{T}(\rho_{AB}))$$

The recovery map $\mathcal{T}$: proving universality

We know that

$$\min_{\rho_{ABC}} \max_{\mathcal{T}:B\to BC} I(A:C|B)_\rho - \frac{1}{4\ln 2}\|\rho_{ABC} - \mathcal{T}(\rho_{AB})\|_1 \geq 0$$

Wanted: Exchange the min and max

To apply **minmax theorem**:

- Concave in $\mathcal{T}$
- Not convex in ρ, but "can make it linear" by adding a label with ρ

$$I(A:C|B)_\rho - \frac{1}{4\ln 2}\|\rho_{ABC} - \mathcal{T}(\rho_{AB})\|_1$$
$$= I(\hat{A}A:C|B)_{|\rho\rangle\langle\rho|_{\hat{A}}\otimes\rho} - \frac{1}{4\ln 2}\||\rho\rangle\langle\rho|_{\hat{A}} \otimes (\rho_{ABC} - \mathcal{T}(\rho_{AB}))\|_1$$

Conclusion

- Conditional mutual information useful for its additivity properties
- **Main result:**

$$I(A:C|B)_\rho \leq \epsilon$$

$$\Rightarrow \quad \rho_{ABC} \text{ approximately satisfies Markov chain condition}$$

Conclusion

- Conditional mutual information useful for its additivity properties
- **Main result:**

$$I(A:C|B)_\rho \leq \epsilon$$
$$\Rightarrow \ \rho_{ABC} \text{ approximately satisfies Markov chain condition}$$

- **Open questions:**
  - Improved lower bounds? Replace log fidelity with relative entropy
  - Explicit structure of recovery map?