Continuity of entropies via integral representations

Mario Berta, Ludovico Lami, Marco Tomamichel

IEEE Transaction on Information Theory (2025), arXiv:2408.15226

RWTH AACHEN UNIVERSITY
SCUOLA NORMALE SUPERIORE
NUS National University of Singapore

Entropy

- Shannon entropy of probability distribution $\mathrm{P} = \{p_x\}_{x \in X}$ is

$$H(X)_P \coloneqq -\sum_{x \in X} p_x \log p_x$$

- Continuity in variational distance for $\frac{1}{2}||P - Q||_1 \leq \epsilon$ as [1]:

$$|H(X)_P - H(X)_Q| \leq \epsilon \cdot \log(|X| - 1) + h_2(\epsilon)$$

with binary entropy $h_2(\epsilon) = -\epsilon \log \epsilon - (1 - \epsilon) \log(1 - \epsilon)$

- Extension to conditional entropy $H(X|Y)_P \coloneqq H(XY)_P - H(Y)_P$ as [2]:

$$|H(X|Y)_P - H(X|Y)_Q| \leq \epsilon \cdot \log(|X| - 1) + h_2(\epsilon)$$

Quantum generalizations

- Von Neumann entropy of quantum state $\rho_A \geq 0$ is

$$H(A)_\rho := -\mathrm{Tr}[\rho_A \log \rho_A]$$

- Continuity in trace distance for $\frac{1}{2}||\rho - \sigma||_1 \leq \epsilon$ as [3,4,5]:

$$|H(A)_\rho - H(A)_\sigma| \leq \epsilon \cdot \log(|A| - 1) + h_2(\epsilon)$$

- Extension to conditional entropy $H(A|B)_\rho := H(AB)_\rho - H(B)_\rho$ as [6,7]:

$$|H(A|B)_\rho - H(A|B)_\sigma| \leq \epsilon \cdot \log|A|^2 + (1+\epsilon) \cdot h_2\left(\frac{\epsilon}{1+\epsilon}\right)$$

- Not exactly tight anymore, improvements / general perspective [7,8]?

Why would we care?

Applications in quantum information theory

- Quantum Shannon theory → channel capacities
- Entanglement theory → entanglement cost, squashed entanglement
- Quantum resource theories → transformation rates

- Example: Quantum capacity $Q(\mathcal{N})$ of ϵ-degrading quantum channel $\mathcal{N}$ of Choi rank r bounded as [9,10]

$$\mathrm{I_c}(\mathcal{N}) \leq Q(\mathcal{N}) \leq \mathrm{I_c}(\mathcal{N}) + f(\epsilon, r)$$

- Extra: Continuity of more general quantum information measures

Main result

Meta semi-continuity relation

- Quantum relative entropy is $D(\rho||\sigma) \coloneqq \mathrm{Tr}[\rho\,(\log\rho - \log\sigma)]$

- Main result (simplified): For $\frac{1}{2}||\rho - \sigma||_1 \leq \epsilon$ and ω, we find

$$D(\rho||\omega) - D(\sigma||\omega) \leq \epsilon \cdot \log\left(e^{D_{\max}(\rho||\omega)} - 1\right) + h_2(\epsilon)$$

with $D_{\max}(\rho||\omega) := \log\inf\,\{\lambda \in \mathbb{R} : \rho \leq e^{\lambda}\omega\}$

- Generates von Neumann entropy: $H(A)_\rho = -D(\rho_A||1_A)$
- Generates conditional entropy: $H(A|B)_\rho = -D(\rho_{AB}||1_A \otimes \rho_B)$

Immediate consequences

$$D(\rho||\omega) - D(\sigma||\omega) \leq \epsilon \cdot \log\left(e^{D_{\max}(\rho||\omega)} - 1\right) + h_2(\epsilon)$$

- Von Neumann entropy, parameter $\beta := \max\{\, ||\rho||_\infty, ||\sigma||_\infty \,\}$ as:

$$|H(A)_\rho - H(A)_\sigma| \leq \epsilon \cdot \log(\beta|A| - 1) + h_2(\epsilon)$$

- Conditional entropy for ρ_{AB}, σ_{AB} with $\rho_B = \sigma_B$ as:

$$|H(A|B)_\rho - H(A|B)_\sigma| \leq \epsilon \cdot \log(|A|^2 - 1) + h_2(\epsilon)$$

versus previous bound $\epsilon \cdot \log|A|^2 + (1+\epsilon) \cdot h_2\left(\frac{\epsilon}{1+\epsilon}\right)$

→ exactly tight in every dimension, conjecture [7,8] for case $\rho_B = \sigma_B$!

Integral representations

Frenkel's integral form

- Motivated from the classical case [11], one finds [12] (see also [13,14])

$$D(\rho||\sigma) = \int_1^\infty d\gamma \left(\frac{1}{\gamma} \cdot E_\gamma(\rho||\sigma) + \frac{1}{\gamma^2} \cdot E_\gamma\,(\sigma||\rho) \right)$$

with hockey-stick divergences

$$E_\gamma(\rho||\sigma) \coloneqq \mathrm{Tr}[\rho - \gamma\sigma]_+$$

- Cleverly employ mathematical properties of hockey-stick divergences:
  - variational representation
  - monotonicity in γ
  - convexity in γ
  - triangle inequalities, etc.

Proof sketch

Intuition in Frenkel form

$$D(\rho||\omega) - D(\sigma||\omega) = \int_1^\infty \frac{d\gamma}{\gamma}\Big(E_\gamma(\rho||\omega) - E_\gamma(\sigma||\omega)\Big) + \int_1^\infty \frac{d\gamma}{\gamma^2}\Big(E_\gamma(\omega||\rho) - E_\gamma(\omega||\sigma)\Big)$$

- Exactly tight for: $\rho_{AB} = \Phi_{AB}$ MES, $\sigma_{AB} = (1-\epsilon)\cdot\Phi_{AB} + \frac{\epsilon}{|A|^2-1}\cdot(1_{AB} - \Phi_{AB})$
- Split up both integrals into three intervals each, via $\mathrm{M} = e^{D_{\max}(\rho||\omega)}$ as:

$$[1,(1-\epsilon)\cdot M], \qquad [(1-\epsilon)\cdot M, M], \qquad [M,\infty]$$

- Bound all terms separately using properties of hockey-stick divergences, e.g., for interval $1 \leq \gamma \leq (1-\epsilon)\cdot M$ we have:

$$E_\gamma(\rho||\omega) = \mathrm{Tr}[\rho - \gamma\omega]_+ \leq \mathrm{Tr}[\rho - \sigma]_+ + \mathrm{Tr}[\sigma - \gamma\omega]_+ \leq \epsilon + E_\gamma(\sigma||\omega)$$

Conclusion & Handover

- Main result: $D(\rho||\omega) - D(\sigma||\omega) \leq \epsilon \cdot \log\left(e^{D_{\max}(\rho||\omega)} - 1\right) + h_2(\epsilon)$
- For $\rho_B = \sigma_B$ we get: $|H(A|B)_\rho - H(A|B)_\sigma| \leq \epsilon \cdot \log(|A|^2 - 1) + h_2(\epsilon)$
- Many applications throughout quantum information theory
- Many questions remain open, e.g., case $\rho_B \neq \sigma_B$, or for mutual information: $I(A:B)_\rho = H(A)_\rho + H(B)_\rho - H(AB)_\rho$ even classically!

- Thank you! I am hiring at RWTH Aachen University: berta@physik.rwth-aachen.de

References

[1] Csiszár, as credited by Petz in his textbook *Quantum Information Theory and Quantum Statistics.*

[2] Alhejji & Smith, *A tight uniform continuity bound for equivocation*, Proc. IEEE ISIT (2020).

[3] Fannes, *A continuity property of the entropy density for spin lattice systems*, Comm. Math. Phys. (1973).

[4] Audenaert, *A sharp continuity estimate for the von Neumann entropy*, J. Phys. A (2007).

[5] Petz, *Quantum Information Theory and Quantum Statistics*, Springer (2008).

[6] Alicki & Fannes, *Continuity of quantum conditional information*, J. Phys. A (2004).

[7] Winter, *Tight uniform continuity bounds for quantum entropies,* Comm. Math. Phys. (2016).

[8] Wilde, *Optimal uniform continuity bound for conditional entropy of classical–quantum states*, QIP (2020).

[9] Sutter et al., *Approximate degradable quantum channels*, IEEE Trans. Inf. Theory (2017).

[10] Leditzky et al., *Quantum and private capacities of low-noise channels*, Phys. Rev. Lett. (2018).

[11] Sason & Verdu, *f-divergence Inequalities*, IEEE Trans. Inf. Theory (2016).

[12] Frenkel, *Integral formula for quantum relative entropy implies data processing inequality*, Quantum (2023).

[13] A. Jencová, *Recoverability of quantum channels via hypothesis testing*, Lett. Math. Phys. (2024).

[14] Hirche & Tomamichel*, Quantum Rényi and f-divergences from integral representations*, Comm. Math. Phys. (2024).