A walk through the zoo of quantum information entropies

Francesco Buscemi* (Nagoya University)

21 March 2022

YITP Intl Workshop *Quantum Information Entropy in Physics*

*www.quantumquia.com

Notation and conventions

- Hilbert spaces $(\mathcal{H}_A, \mathcal{H}_B, \dots)$ are all finite-dimensional
- set of non-negative linear operators: $\mathcal{P}(\mathcal{H}_A), \mathcal{P}(\mathcal{H}_B), \dots$
- the word "state" denotes normalized density matrices $(\rho_A, \sigma_A, \dots)$
- set of density matrices: $\mathcal{D}(\mathcal{H}_A), \mathcal{D}(\mathcal{H}_B), \dots$
- support: $\mathsf{supp}\rho_A := (\mathsf{ker}\rho_A)^\perp$
- quantum channels $\mathcal{E} : A \to B$ are completely positive trace-preserving (CPTP) linear maps from operators on $\mathcal{H}_A$ to operators on $\mathcal{H}_B$
- trace-dual map $\mathcal{E}^\dagger : B \to A$ is defined by $\mathrm{Tr}[\mathcal{E}(X_A)\ Y_B] = \mathrm{Tr}[X_A\ \mathcal{E}^\dagger(Y_B)]$ for all linear operators X_A and Y_B
- identity operator: $\mathbb{1}_A$; identity channel: id; maximally mixed state: $\omega_A := d_A^{-1}\mathbb{1}_A$

Recommended references

- F. Dupuis, L. Kraemer, P. Faist, J. M. Renes, and R. Renner: *Generalized Entropies*. Freely available on the arXiv at `https://doi.org/10.48550/arXiv.1211.3141`
- M. Tomamichel: *Quantum Information Processing with Finite Resources – Mathematical Foundations*. Freely available on the arXiv at `https://doi.org/10.48550/arXiv.1504.00233`
- S. Khatri and M. M. Wilde: *Principles of Quantum Communication Theory: A Modern Approach*. Freely available on the arXiv at `https://doi.org/10.48550/arXiv.2011.04672`
- please refer to the above references' bibliography also for the right credits to give for each definition and result introduced here

In a zoo there are many animals:
they all have their role
and they are all beautiful

Basic information-theoretic quantities

- von Neumann entropy:

$$\rho_A \in \mathcal{D}(\mathcal{H}_A) \rightsquigarrow H(A)_\rho := -\operatorname{Tr}\{\rho_A \log_2 \rho_A\}$$

- quantum conditional entropy:

$$\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B) \rightsquigarrow H(A|B)_\rho := H(AB)_\rho - H(B)_\rho$$

- quantum mutual entropy:

$$\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B) \rightsquigarrow I(A;B)_\rho := H(A)_\rho + H(B)_\rho - H(AB)_\rho$$

- quantum conditional mutual information:

$$\rho_{ABC} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C) \rightsquigarrow I(A;C|B)_\rho := H(A|B)_\rho + H(C|B)_\rho - H(AC|B)_\rho$$

In fact, they all arise from one single quantity

ザ quantum relative entropy: Kullback–Leibler–Umegaki relative entropy

Definition

Let $\rho \in \mathcal{D}(\mathcal{H})$, $Q \in \mathcal{P}(\mathcal{H})$, and $Q_\epsilon := Q + \epsilon \mathbb{1}$; the Kullback–Leibler–Umegaki (KLU) relative entropy is given by

$$D(\rho \| Q) := \lim_{\epsilon \to 0^+} \mathrm{Tr}\left\{ \rho \log_2 \rho - \rho \log_2 Q_\epsilon \right\} .$$

Useful properties:

- Klein's inequality: $\mathrm{Tr}\{\rho\} \geq \mathrm{Tr}\{Q\} \implies D(\rho \| Q) \geq 0$
- $\rho \leq Q \implies D(\rho \| Q) \leq 0$
- $Q \leq Q' \implies D(\rho \| Q) \geq D(\rho \| Q')$
- $c > 0 \implies D(\rho \| cQ) + \log_2 c = D(\rho \| Q)$

How other quantities arise from KLU relative entropy

- von Neumann entropy: $H(A)_\rho = -D(\rho_A \| \mathbb{1}_A)$
- quantum conditional entropy:
  $H(A|B)_\rho = -D(\rho_{AB} \| \mathbb{1}_A \otimes \rho_B) = -\inf_{\sigma_B \in \mathcal{D}(\mathcal{H}_B)} D(\rho_{AB} \| \mathbb{1}_A \otimes \sigma_B)$
- quantum mutual information:
  $I(A;B)_\rho = D(\rho_{AB} \| \rho_A \otimes \rho_B) = \inf_{\sigma_A \in \mathcal{D}(\mathcal{H}_A), \tau_B \in \mathcal{D}(\mathcal{H}_B)} D(\rho_{AB} \| \sigma_A \otimes \tau_B)$
- quantum conditional mutual information:
  $I(A;C|B)_\rho = D(\rho_{ABC} \| \sigma_{ABC})$ where

$$\sigma_{ABC} := 2^{\log_2(\rho_{AB} \otimes \mathbb{1}_C) + \log_2(\mathbb{1}_A \otimes \rho_{BC}) - \log_2(\mathbb{1}_A \otimes \rho_B \otimes \mathbb{1}_C)}$$

Natural properties of KLU relative entropy

- positivity on states: for any $\rho, \sigma \in \mathcal{D}(\mathcal{H})$, $D(\rho\|\sigma) \geq 0$
- faithfulness: for any $\rho, \sigma \in \mathcal{D}(\mathcal{H})$, $D(\rho\|\sigma) = 0 \iff \rho = \sigma$
- data-processing (DP) property: for any $\rho \in \mathcal{D}(\mathcal{H})$, any $Q \in \mathcal{P}(\mathcal{H})$, and any quantum channel $\mathcal{E}$, $D(\rho\|Q) \geq D(\mathcal{E}(\rho)\|\mathcal{E}(Q))$

The role of the data-processing property

In particular, the DP property of KLU relative entropy is very useful to prove inequalities.

- example (mathematics): $I(A;C|B)_\rho \geq 0$ (strong subadditivity). Proof:

$$\begin{aligned} I(A;C|B)_\rho &= H(A|B)_\rho + H(C|B)_\rho - H(AC|B)_\rho \\ &= H(C|B)_\rho - H(C|AB)_\rho \\ &= D(\rho_{ABC}\|\rho_{AB}\otimes \mathbb{1}_C) - D(\rho_{BC}\|\rho_B \otimes \mathbb{1}_C) \\ &= D(\rho_{ABC}\|\rho_{AB}\otimes \omega_C) - D(\rho_{BC}\|\rho_B \otimes \omega_C) \\ &\geq 0\,, \end{aligned}$$

  where the last inequality comes from taking the partial trace over A
- example (physics): $\Gamma_A = Z^{-1}e^{-\beta\mathscr{H}_A}$, $D(\rho_A\|\Gamma_A) = \beta(F - \langle\mathscr{H}_A\rangle_\rho) - H(A)_\rho$, hence, for a channel such that $\mathcal{E}(\Gamma_A) = \Gamma_A$,

$$\Delta H(A) \geq -\beta\Delta\langle\mathscr{H}_A\rangle$$

Is the KLU relative entropy “unique”?

$\rightsquigarrow$ **Rényi’s axiomatic approach**

Petz–Rényi relative entropies

Definition

Let $\rho \in \mathcal{D}(\mathcal{H})$, $Q \in \mathcal{P}(\mathcal{H})$, and $Q_\epsilon := Q + \epsilon \mathbb{1}$; for all $\alpha \in (0,1) \cup (1,+\infty)$, the Petz–Rényi α-relative entropy is defined as

$$D_\alpha(\rho \| Q) := \lim_{\epsilon \to 0^+} \frac{1}{\alpha - 1} \log_2 \mathrm{Tr}\{\rho^\alpha \, Q_\epsilon^{1-\alpha}\} \, .$$

- faithfulness: for all $\rho, \sigma \in \mathcal{D}(\mathcal{H})$, $D_\alpha(\rho \| \sigma) = 0 \iff \rho = \sigma$
- monotonicity: $\alpha \le \alpha' \implies D_\alpha(\rho \| Q) \le D_{\alpha'}(\rho \| Q)$
- DP property: for any $\alpha \in (0,1) \cup (1,2]$, any $\rho \in \mathcal{D}(\mathcal{H})$, any $Q \in \mathcal{P}(\mathcal{H})$, and any channel $\mathcal{E}$, $D_\alpha(\rho \| Q) \ge D_\alpha(\mathcal{E}(\rho) \| \mathcal{E}(Q))$
- for any $\rho \in \mathcal{D}(\mathcal{H})$ and any $Q \in \mathcal{P}(\mathcal{H})$, $\lim_{\alpha \to 1} D_\alpha(\rho \| Q) = D(\rho \| Q)$

"Sandwiched" Rényi relative entropies

Definition

Let $\rho \in \mathcal{D}(\mathcal{H})$, $Q \in \mathcal{P}(\mathcal{H})$, and $Q_\epsilon := Q + \epsilon \mathbb{1}$; for all $\alpha \in (0,1) \cup (1,+\infty)$, the sandwiched Rényi α-relative entropy is defined as

$$\begin{aligned}\widetilde{D}_\alpha(\rho\|Q) :&= \lim_{\epsilon\to 0^+} \frac{1}{\alpha-1} \log_2 \mathrm{Tr}\left\{\left(Q_\epsilon^{\frac{1-\alpha}{2\alpha}}\ \rho\ Q_\epsilon^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right\} \\ &= \lim_{\epsilon\to 0^+} \frac{1}{\alpha-1} \log_2 \mathrm{Tr}\left\{\left(\rho^{\frac{1}{2}}\ Q_\epsilon^{\frac{1-\alpha}{\alpha}}\ \rho^{\frac{1}{2}}\right)^\alpha\right\}.\end{aligned}$$

- faithfulness: for all $\rho, \sigma \in \mathcal{D}(\mathcal{H})$, $\widetilde{D}_\alpha(\rho\|\sigma) = 0 \iff \rho = \sigma$
- monotonicity: $\alpha \le \alpha' \implies \widetilde{D}_\alpha(\rho\|Q) \le \widetilde{D}_{\alpha'}(\rho\|Q)$
- DP property: for any $\alpha \in [\frac{1}{2}, 1) \cup (1,+\infty)$, any $\rho \in \mathcal{D}(\mathcal{H})$, any $Q \in \mathcal{P}(\mathcal{H})$, and any channel $\mathcal{E}$, $\widetilde{D}_\alpha(\rho\|Q) \ge \widetilde{D}_\alpha(\mathcal{E}(\rho)\|\mathcal{E}(Q))$
- for any $\rho \in \mathcal{D}(\mathcal{H})$ and any $Q \in \mathcal{P}(\mathcal{H})$, $\lim_{\alpha\to 1} \widetilde{D}_\alpha(\rho\|Q) = D(\rho\|Q)$
- for any $\rho, \sigma \in \mathcal{D}(\mathcal{H})$, $\widetilde{D}_{\frac{1}{2}}(\rho\|\sigma) = -\log_2 \left\|\sqrt{\rho}\sqrt{\sigma}\right\|_1^2 := -\log_2 F(\rho,\sigma)$

Geometric Rényi relative entropies

Definition

Let $\rho \in \mathcal{D}(\mathcal{H})$, $Q \in \mathcal{P}(\mathcal{H})$, and $Q_\epsilon := Q + \epsilon \mathbb{1}$; for all $\alpha \in (0,1) \cup (1,+\infty)$, the geometric Rényi α-relative entropy is defined as

$$\widehat{D}_\alpha(\rho\|Q) := \lim_{\epsilon\to 0^+} \frac{1}{\alpha-1} \log_2 \mathrm{Tr}\left\{\sqrt{Q_\epsilon}\left(\frac{1}{\sqrt{Q_\epsilon}}\rho\frac{1}{\sqrt{Q_\epsilon}}\right)^\alpha \sqrt{Q_\epsilon}\right\} .$$

- $\widehat{D}_2(\rho\|Q) = D_2(\rho\|Q)$
- faithfulness: for all $\rho, \sigma \in \mathcal{D}(\mathcal{H})$, $\widehat{D}_\alpha(\rho\|\sigma) = 0 \iff \rho = \sigma$
- monotonicity: $\alpha \le \alpha' \implies \widehat{D}_\alpha(\rho\|Q) \le \widehat{D}_{\alpha'}(\rho\|Q)$
- DP property: for any $\alpha \in (0,1) \cup (1,2]$, any $\rho \in \mathcal{D}(\mathcal{H})$, any $Q \in \mathcal{P}(\mathcal{H})$, and any channel $\mathcal{E}$, $\widehat{D}_\alpha(\rho\|Q) \ge \widehat{D}_\alpha(\mathcal{E}(\rho)\|\mathcal{E}(Q))$
- for $\alpha \to 1$ does not converge to KLU relative entropy, but to Belavkin–Staszewski's: $\lim_{\alpha\to 1} \widehat{D}_\alpha(\rho\|Q) = \mathrm{Tr}\{\rho \log_2(\sqrt{\rho} Q^{-1} \sqrt{\rho})\}$

Fundamental relation between generalized quantum Rényi relative entropies

Theorem

For any $\rho \in \mathcal{D}(\mathcal{H})$, any $Q \in \mathcal{P}(\mathcal{H})$, and any $\alpha \in (0,1) \cup (1,+\infty)$, if $[\rho, Q] = 0$,

$$\widetilde{D}_\alpha(\rho\|Q) = D_\alpha(\rho\|Q) = \widehat{D}_\alpha(\rho\|Q) \; .$$

If $[\rho, Q] \neq 0$, then for any $\alpha \in (0,1) \cup (1,2]$:

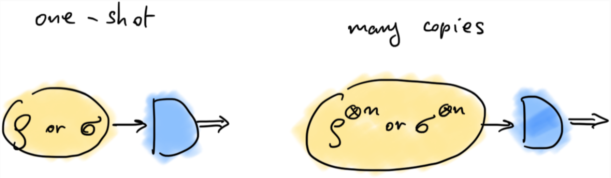
$$\widetilde{D}_\alpha(\rho\|Q) \leq D_\alpha(\rho\|Q) \leq \widehat{D}_\alpha(\rho\|Q) \; .$$

Moreover: any "natural" quantum generalization of the classical Rényi relative entropies must lie in between $\widetilde{D}_\alpha$ and $\widehat{D}_\alpha$.

With so many entropies, which is the “right” one to use?

It depends...

Example: the quantum Chernoff bound



- guessing probability:
  $P_{\text{guess}}(p, \rho, \sigma) := \max_{0 \leq E \leq \mathbb{1}} \operatorname{Tr}\{pE\rho + (1-p)(\mathbb{1} - E)\sigma\}$
- intuition: the more copies are given, the more distinguishable the states become, that is, $P_{\text{guess}}(p, \rho^{\otimes n}, \sigma^{\otimes n}) \xrightarrow{n \to \infty} 1$
- question: how fast?
- answer (quantum Chernoff bound):
  $\lim_{n\to\infty} -\frac{1}{n} \log_2 [1 - P_{\text{guess}}(p, \rho^{\otimes n}, \sigma^{\otimes n})] = C(\rho, \sigma)$ where

$$C(\rho\|\sigma) := \sup_{\alpha \in (0,1)} -\log_2 \operatorname{Tr}\{\rho^{\alpha}\, \sigma^{1-\alpha}\}\,.$$

Instead, the cases $\alpha \to 0$ and $\alpha \to \infty$ become particularly useful in the finite block-length regime

"min" and "max" quantum relative entropies

Definition

Let $\rho \in \mathcal{D}(\mathcal{H})$ and $Q \in \mathcal{P}(\mathcal{H})$ the min- and max- quantum relative entropy are defined as

$$D_{\min}(\rho\|Q) := \lim_{\alpha\to 0} D_\alpha(\rho\|Q) , \qquad (\equiv D_0(\rho\|Q)) ,$$
$$D_{\max}(\rho\|Q) := \lim_{\alpha\to\infty} \widetilde{D}_\alpha(\rho\|Q) .$$

The max-relative entropy

Theorem

For $\rho \in \mathcal{D}(\mathcal{H})$ and $Q \in \mathcal{P}(\mathcal{H})$, if $\mathrm{supp}\rho \subseteq \mathrm{supp}Q$, *we have*

$$\begin{aligned} D_{\max}(\rho\|Q) &= \log_2 \lambda_{\max}(\sqrt{\rho}Q^{-1}\sqrt{\rho}) \\ &= \log_2 \inf\{\lambda : \rho \leq \lambda Q\} \\ &= \log_2 \sup_{M\geq 0\,:\,\mathrm{Tr}\{MQ\}\leq 1} \mathrm{Tr}\{M\rho\}\ ; \end{aligned}$$

otherwise $D_{\max}(\rho\|Q) := +\infty$.

- for any $\rho \in \mathcal{D}(\mathcal{H})$ and any $Q \in \mathcal{P}(\mathcal{H})$, $D_{\max}(\rho\|Q) = \lim_{\alpha\to\infty} \widehat{D}_\alpha(\rho\|Q)$
- faithfulness: for all $\rho, \sigma \in \mathcal{D}(\mathcal{H})$, $D_{\max}(\rho\|\sigma) = 0 \iff \rho = \sigma$
- DP property: for any $\rho \in \mathcal{D}(\mathcal{H})$, any $Q \in \mathcal{P}(\mathcal{H})$, and any channel $\mathcal{E}$, $D_{\max}(\rho\|Q) \geq D_{\max}(\mathcal{E}(\rho)\|\mathcal{E}(Q))$

The min-relative entropy

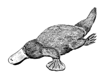
Theorem

For $\rho \in \mathcal{D}(\mathcal{H})$ and $Q \in \mathcal{P}(\mathcal{H})$, we have

$$D_{\min}(\rho \| Q) = -\log_2 \mathrm{Tr}\{\Pi_\rho \, Q\} \,.$$

- DP property: for any $\rho \in \mathcal{D}(\mathcal{H})$, any $Q \in \mathcal{P}(\mathcal{H})$, and any channel $\mathcal{E}$, $D_{\min}(\rho \| Q) \geq D_{\min}(\mathcal{E}(\rho) \| \mathcal{E}(Q))$
- it is not faithful: $D_{\min}(\rho \| \sigma) = 0 \;\not\Longrightarrow\; \rho = \sigma$

カモノハシ: Hilbert α-divergences



Definition (F.B. and G.Gour, PRA, 2017)

For any $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ and any $\alpha \geq 1$, the Hilbert α-divergence is defined as

$$H_\alpha(\rho\|\sigma) := \frac{\alpha}{\alpha - 1} \log_2 \sup_{\alpha^{-1}\mathbb{1} \leq E \leq \mathbb{1}} \frac{\mathrm{Tr}\{E\rho\}}{\mathrm{Tr}\{E\sigma\}} .$$

- faithfulness: for all $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ and all $\alpha \geq 1$,
  $H_\alpha(\rho\|\sigma) = 0 \iff \rho = \sigma$
- DP property: for any $\alpha \geq 1$, any $\rho, \sigma \in \mathcal{D}(\mathcal{H})$, and any channel $\mathcal{E}$,
  $H_\alpha(\rho\|\sigma) \geq H_\alpha(\mathcal{E}(\rho)\|\mathcal{E}(\sigma))$
- $\lim_{\alpha\to\infty} H_\alpha(\rho\|\sigma) = D_{\max}(\rho\|\sigma)$
- $\lim_{\alpha\to 1^+} H_\alpha(\rho\|\sigma) = \frac{1}{2\ln 2}\|\rho - \sigma\|_1$ (i.e., a Brègman divergence!)

Smoothing and conditioning

What is it?

- "**smoothing**", in the context of quantum information entropies, means to make the quantity continuous in its left argument
  - think in particular about $D_{\min}(\rho\|\sigma) = -\log_2 \mathrm{Tr}\{\Pi_\rho \sigma\}$, which is not continuous in ρ
- this is important both mathematically (to deal with limits etc.) and technologically (because there is always error in the characterization of states)
- "**conditioning**" instead is the process of updating information-theoretic quantities when side-information is available
- both smoothing and conditioning are treated as optimizations over some sets (i.e., as entropic projections)

How to smooth?

Definition (State-smoothing and operator-smoothing)

For any $\rho \in \mathcal{D}(\mathcal{H})$, let us define

$$B_\epsilon(\rho) := \{\bar{\rho} \in \mathcal{D}(\mathcal{H}) : \bar{\rho} \approx_\epsilon \rho\} ,$$
$$B^*_\epsilon(\rho) := \{0 \le P \le \mathbb{1} : \mathrm{Tr}\{P\rho\} \ge 1 - \epsilon\} .$$

Then, for any $Q \in \mathcal{P}(\mathcal{H})$,

$$D^\epsilon_{\max}(\rho\|\sigma) := \inf_{\bar{\rho} \in B_\epsilon(\rho)} D_{\max}(\bar{\rho}\|Q) ,$$
$$D^\epsilon_{\min}(\rho\|Q) \equiv D^\epsilon_H(\rho\|Q) := -\log_2 \inf_{P \in B^*_\epsilon(\rho)} \mathrm{Tr}\{P\, Q\} .$$

The quantity $D^\epsilon_H(\rho\|Q)$, now known as hypothesis-testing relative entropy, was originally introduced in [F.B. and N. Datta. IEEE Trans. Inf. Th. (2010)] as the operator-smoothed Petz relative quasi-entropy of order 0.

The asymptotic equipartition property

Theorem

For any $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ *and any* $\epsilon \in (0,1)$

$$\lim_{n\to\infty} \frac{1}{n} D_H^{\epsilon}(\rho^{\otimes n} \| \sigma^{\otimes n}) = D(\rho\|\sigma)\ ,$$

and

$$\lim_{n\to\infty} \frac{1}{n} D_{\max}^{\epsilon}(\rho^{\otimes n} \| \sigma^{\otimes n}) = D(\rho\|\sigma)\ .$$

How to condition?

There are various ways to express $H(A|B)_\rho$, but we focus on the following two:

$$H(A|B)_\rho = -D(\rho_{AB} \| \mathbb{1}_A \otimes \rho_B) \ ,$$
$$H(A|B)_\rho = -\inf_{\sigma_B \in \mathcal{D}(\mathcal{H}_B)} D(\rho_{AB} \| \mathbb{1}_A \otimes \sigma_B) \ .$$

Definition

For all $\epsilon \geq 0$ we define

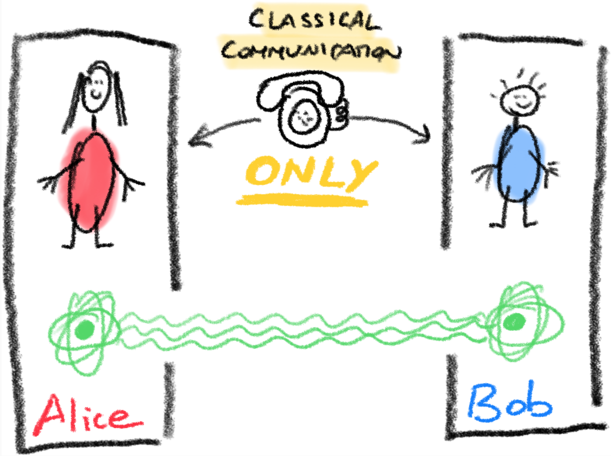
$$H^\epsilon_{\min}(A|\sigma_B)_\rho := -D^\epsilon_{\max}(\rho_{AB} \| \mathbb{1}_A \otimes \sigma_B)$$
$$H^\epsilon_{\min}(A|B)_\rho := -\inf_{\sigma_B \in \mathcal{D}(\mathcal{H}_B)} D^\epsilon_{\max}(\rho_{AB} \| \mathbb{1}_A \otimes \sigma_B) \ .$$

Applications in quantum information theory: one-shot entanglement theory

References

- F.B. and N. Datta: The quantum capacity of channels with arbitrarily correlated noise. IEEE TIT (2010). `http://arxiv.org/abs/0902.0158`
- F.B. and N. Datta: Distilling entanglement from arbitrary resources. J. Math. Phys. (2010). `http://arxiv.org/abs/1006.1896`
- F.B. and N. Datta: Entanglement cost in practical scenarios. Phys. Rev. Lett. (2011). `http://arxiv.org/abs/0906.3698`
- F.B. and N. Datta: General theory of assisted entanglement distillation. IEEE TIT (2013). `http://arxiv.org/abs/1009.4464`

Quantum entanglement as a resource: the LOCC paradigm



One-shot distillation and dilution

- One-shot entanglement distillation:

$$\rho_{AB} \xmapsto{\mathcal{L}\in\mathrm{LOCC}} \underbrace{\Psi^+_{A'B'} \otimes \cdots \Psi^+_{A'B'}}_{N_{\max}(\rho_{AB})}.$$

- One-shot entanglement dilution:

$$\underbrace{\Psi^+_{AB} \otimes \cdots \otimes \Psi^+_{AB}}_{M_{\min}(\sigma_{A'B'})} \xmapsto{\mathcal{L}\in\mathrm{LOCC}} \sigma_{A'B'}.$$

- Correspondingly,
  - one-shot distillable entanglement: $E_D^{(1)}(\rho_{AB}) = N_{\max}(\rho_{AB})$;
  - one-shot entanglement cost: $E_C^{(1)}(\sigma_{A'B'}) = M_{\min}(\sigma_{A'B'})$

Allowing for finite accuracy

With an eye to practical implementations we define:

One-shot entanglement ϵ-distillation:

$$\rho_{AB} \xmapsto{\mathcal{L}\in\mathrm{LOCC}} \tilde{\rho}_{A'B'} \overset{\varepsilon}{\approx} \underbrace{\Psi^-_{A'B'} \otimes \cdots \Psi^-_{A'B'}}_{N_{\max}(\rho_{AB};\varepsilon)}.$$

One-shot entanglement ϵ-dilution

$$\underbrace{\Psi^-_{AB} \otimes \cdots \otimes \Psi^-_{AB}}_{M_{\min}(\sigma_{A'B'};\varepsilon)} \xmapsto{\mathcal{L}\in\mathrm{LOCC}} \tilde{\sigma}_{A'B'} \overset{\varepsilon}{\approx} \sigma_{A'B'}.$$

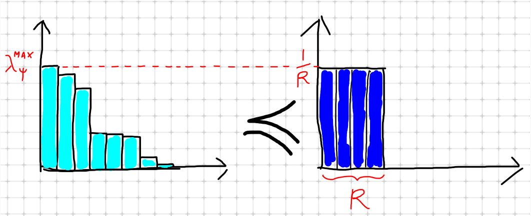
Correspondingly,

- one-shot ϵ-distillable entanglement: $E_D^{(1)}(\rho_{AB};\epsilon) = N_{\max}(\rho_{AB};\epsilon)$;
- one-shot entanglement ϵ-cost: $E_C^{(1)}(\sigma_{A'B'};\epsilon) = M_{\min}(\sigma_{A'B'};\epsilon)$

Pure states: one-shot zero-error distillable entanglement

☞ Given an initial pure state ψ_{AB}, denote by $\lambda_\psi^{\downarrow i}$ the i-th largest eigenvalue of $\psi_A := \mathrm{Tr}_B\{\psi_{AB}\}$. **Nielsen** (1999) showed that a maximally entangled state of rank R, i.e. $R^{-1/2}\sum_{i=1}^R |i\rangle|i\rangle$, can be distilled if and only if

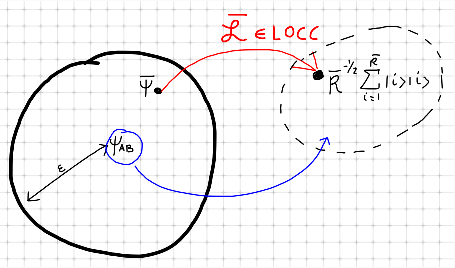
1. $\lambda_\psi^{\max} \equiv \lambda_\psi^{\downarrow 1} \le R^{-1}$;
2. $\lambda_\psi^{\downarrow 1} + \lambda_\psi^{\downarrow 2} \le 2R^{-1}$;
3. and so on.



☞ $E_D^{(1)}(\psi_{AB};0) \ge \log_2 \left\lfloor \frac{1}{\lambda_\psi^{\max}} \right\rfloor = -D_{\max}(\psi_A \| \mathbb{1}_A) =: H_{\min}(A)_\psi$

Considering finite accuracy

Consider the set of pure states $\mathscr{P}_\epsilon(\psi_{AB}) := \left\{ |\bar{\psi}_{AB}\rangle : \bar{\psi}_{AB} \overset{\epsilon}{\approx} \psi_{AB} \right\}$



☞ A maximally entangled state of rank $\bar{R} = \left\lfloor \frac{1}{\lambda_{\bar{\psi}}^{\max}} \right\rfloor$ can always be distilled up to an ϵ-error, i.e.,

$$E_D^{(1)}(\psi_{AB};\epsilon) \geq \max_{\bar{\psi}\in\mathscr{P}_\epsilon(\psi)} \log_2 \left\lfloor \frac{1}{\lambda_{\bar{\psi}}^{\max}} \right\rfloor = -D_{\max}^{\epsilon}(\psi_A \| \mathbb{1}_A) =: H_{\min}^{\epsilon}(A)_\psi$$

The min-entropy is the one-shot distillable entanglement

Theorem

For any pure bipartite state ψ_{AB} and any $\epsilon \geq 0$, the min-entropy of the reduced state is the one-shot LOCC-distillable entanglement:

$$H_{\min}^{\epsilon}(A)_{\psi} \leq E_D^{(1)}(\psi_{AB};\epsilon) \leq H_{\min}^{\epsilon'}(A)_{\psi} - \log_2(1 - 2\sqrt{\epsilon}) \ ,$$

where $\epsilon' := 2^{\frac{5}{4}}\epsilon^{\frac{1}{8}}$.

The max-entropy is the one-shot entanglement cost

- **Vidal, Jonathan, and Nielsen** (2000): a pure bipartite state ψ_{AB} can be obtained by LOCC from a maximally entangled state of rank R with a minimum error of $\epsilon = 1 - \sum_{i=1}^{R} \lambda_\psi^{\downarrow i}$.
- as a consequence, $E_C^{(1)}(\psi_{AB}; 0) = \log_2 \operatorname{rank} \psi_A = -D_{\min}(\psi_A \| \mathbb{1}_A) =: H_{\max}(A)_\psi$.
- also with finite accuracy:

$$E_C^{(1)}(\psi_{AB}; \epsilon) \simeq H_{\max}^{\epsilon}(A)_\psi,$$

where $H_{\max}^{\epsilon}(A)_\rho := D_{\min}^{\epsilon}(\psi_A \| \mathbb{1}_A)$.

Summary of the pure state case

$$\boxed{\begin{array}{ccccccc} E_D^{(1)}(\psi_{AB};\epsilon) & \simeq & H_{\min}^{\epsilon}(A)_\psi & \leq & H_{\max}^{\epsilon}(A)_\psi & \simeq & E_C^{(1)}(\psi_{AB};\epsilon) \\ \downarrow & & \searrow & & \swarrow & & \downarrow \\ E_D^{\infty}(\psi_{AB}) & = & & H(A)_\psi & & = & E_C^{\infty}(\psi_{AB}) \end{array}}$$

where "$F(\rho;\epsilon) \to G(\rho)$" means $\lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{1}{n}F(\rho^{\otimes n};\epsilon) = G(\rho)$

☞ well-known phenomenon of "asymptotic reversibility" of pure-state entanglement

Reverse data-processing theorems
and the theory of statistical comparison

Extending the data-processing inequality

The data-processing property is so useful that any improvement can be very important.

There are two main directions of investigation:

- theory of approximate reversibility:

$$D(\rho\|\sigma) - D(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)) \geq 0 + \delta$$

- theory of statistical sufficiency, deficiency, and comparison:

$$D(\rho_1\|\rho_2) \geq D(\sigma_1\|\sigma_2) \overset{?}{\Longleftrightarrow} \sigma_i = \mathcal{E}(\rho_i)$$

Starting point

Question. For which triples $(\rho, \sigma, \mathcal{E})$, $D(\rho\|\sigma) = D(\mathcal{E}(\rho)\|\mathcal{E}(\sigma))$?

Petz (1986,1988)

If and only if $\tilde{\mathcal{E}}_\sigma(\bullet) := \sqrt{\sigma}\mathcal{E}^\dagger \left[\frac{1}{\sqrt{\mathcal{E}(\sigma)}} \bullet \frac{1}{\sqrt{\mathcal{E}(\sigma)}}\right] \sqrt{\sigma}$ satisfies

$$\tilde{\mathcal{E}}_\sigma \circ \mathcal{E}(\rho) = \rho \ .$$

(The other equality $\tilde{\mathcal{E}}_\sigma \circ \mathcal{E}(\sigma) = \sigma$ is satisfied *by construction*.)

Question

More generally, suppose that $D(\rho\|\sigma) = D(\rho'\|\sigma')$. Does there exist a quantum channel $\mathcal{E}$ such that $\mathcal{E}(\rho) = \rho'$ and $\mathcal{E}(\sigma) = \sigma'$?

No: the answer is not so simple!

Quantum statistical comparison

The problem

Given two families of quantum states, $E = \{\rho_i : i \in \mathbb{I}\}$ and $F = \{\sigma_i : i \in \mathbb{I}\}$, express the condition:

there exists a channel $\mathcal{E}$ such that $\mathcal{E}(\rho_i) = \sigma_i$ for all $i \in \mathbb{I}$

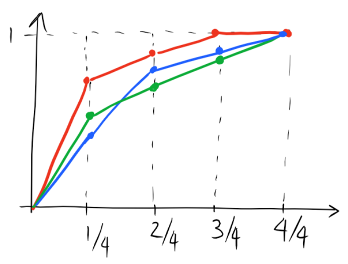
as a collection of inequalities of the form

$$g(E) \geq g(F) \ ,$$

for all g in a suitable family of real-valued functions.

The prototype of statistical comparison: Lorenz curves and majorization

- two probability distributions, $\mathbf{p} = (p_1, \ldots, p_n)$ and $\mathbf{q} = (q_1, \ldots, q_n)$
- truncated sums $P(k) = \sum_{i=1}^{k} p_i^{\downarrow}$ and $Q(k) = \sum_{i=1}^{k} q_i^{\downarrow}$, for all $k = 1, \ldots, n$
- **p** majorizes **q**, i.e., $\mathbf{p} \succeq \mathbf{q}$, whenever $P(k) \geq Q(k)$, for all k
- minimal element: uniform distribution $\mathbf{e} = n^{-1}(1, 1, \cdots, 1)$



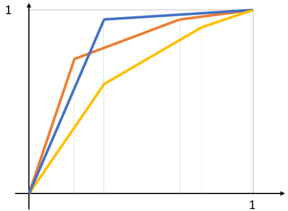
$(x_k, y_k) = (k/n, P(k)), \quad 1 \leq k \leq n$

Hardy, Littlewood, and Pólya (1929)

$\mathbf{p} \succeq \mathbf{q} \iff \mathbf{q} = M\mathbf{p}$, for some bistochastic matrix M (i.e., $M\mathbf{e} = \mathbf{e}$)

Generalization: dichotomies and relative majorization

- two *pairs* of probability distributions, i.e., two *dichotomies*, $(\mathbf{p}_1, \mathbf{p}_2)$ and $(\mathbf{q}_1, \mathbf{q}_2)$, of dimension m and n, respectively
- relabel entries such that ratios p_1^i / p_2^i and q_1^j / q_2^j are nonincreasing
- construct the truncated sums $P_1(k) = \sum_{i=1}^k p_1^i$ and $P_2(k) = \sum_{i=1}^k p_2^i$
- do the same for $Q_1(k)$ and $Q_2(k)$
- $(\mathbf{p}_1, \mathbf{p}_2) \succeq (\mathbf{q}_1, \mathbf{q}_2)$ iff the relative Lorenz curve of the former is never below that of the latter



Relative Lorenz curves:

$(x_k, y_k) = (P_2(k), P_1(k))$

Blackwell's Theorem for Dichotomies (1953)

$(\mathbf{p}_1, \mathbf{p}_2) \succeq (\mathbf{q}_1, \mathbf{q}_2) \iff \mathbf{q}_i = M\mathbf{p}_i$, for some stochastic matrix M.

The case of quantum dichotomies

Question

When is one quantum dichotomy (ρ_1, ρ_2) sufficient for another one (σ_1, σ_2)? That is, when does there exist a quantum channel $\mathcal{E}$ such that $\sigma_i = \mathcal{E}(\rho_i)$?

- qubit case: very similar to the classical case (Alberti and Uhlmann, 1983) otherwise counterexamples (Matsumoto, 2014)
- finite dimensional case: needs an extended comparison (F.B., arXiv:1505.00535; G. Gour, D. Jennings, F.B., R. Duan, I. Marvian, Nat. Comm., 2018)
- quantum relative Lorenz curves: the following are equivalent (F.B. and G. Gour, PRA, 2017)
    - $(\rho_1, \rho_2) \succeq (\sigma_1, \sigma_2)$
    - $H_\alpha(\rho_1 \| \rho_2) \geq H_\alpha(\sigma_1 \| \sigma_2)$ and $H_\alpha(\rho_2 \| \rho_1) \geq H_\alpha(\sigma_2 \| \sigma_1)$ for all $\alpha \geq 1$
    - $D_H^\epsilon(\rho_1 \| \rho_2) \geq D_H^\epsilon(\sigma_1 \| \sigma_2)$ for all $\epsilon \in [0, 1]$

Conclusions

Not even mentioned in this introduction

- Brègman divergences and entropic projections
- infinite dimensional case
- theory of approximate reversibility and Petz's transpose map
- Bayesian inference and learning
- Fisher information and information geometry
- large deviation theory
- DP property under positive maps and statistical morphisms
- additivity properties
- channel entropies
- many applications (information and communication theory, complexity theory, cryptography, statistical mechanics, etc.)
- ...

Take-home messages

- entropies are statistical concepts: there is no entropy without a stochastic process (perhaps hidden)
- entropies are measures of statistic distinguishability
- we need many different entropies because there are many inequivalent notions of "distinguishability" (discrimination, guesswork, estimation, etc.)
- KLU relative entropy is very special within all statistical sciences, but not by any means the only one to learn and use

Presenter's work on the subject

- Buscemi, F; Sutter, D; Tomamichel, M: *An information-theoretic treatment of quantum dichotomies*. Quantum 3, 209 (2019). arXiv link
- Gour, G; Jennings, D; Buscemi, F; Duan, R; Marvian, I: *Quantum majorization and a complete set of entropic conditions for quantum thermodynamics*. Nature Communications 9, 5352 (2018). arXiv link
- Buscemi, F: *Reverse Data-Processing Theorems and Computational Second Laws*. Springer Proceedings in Mathematics & Statistics, vol 261, 135-159 (2018). arXiv link
- Buscemi, F; Gour, G: *Quantum relative Lorenz curves*. Phys. Rev. A 95, 012110 (2017). arXiv link
- Buscemi, F: *Fully quantum second-law–like statements from the theory of statistical comparisons*. Preprint arXiv:1505.00535 [quant-ph].
- Buscemi, F; Datta, N: *General theory of assisted entanglement distillation*. IEEE Trans. Inf. Th., vol. 59 (3), pp.1940-1954 (2013). arXiv link
- Buscemi, F: *Comparison of quantum statistical models: equivalent conditions for sufficiency*. Comm. Math. Phys., vol. 310, pp.625-647 (2012). arXiv link
- Buscemi, F; Datta, N: *Entanglement cost in practical scenarios*. Phys. Rev. Lett., vol. 106, 130503 (2011). arXiv link
- Buscemi, F; Datta, N: *Distilling entanglement from arbitrary resources*. J. Math. Phys., vol. 51 (10), 102201 (2010). arXiv link
- Buscemi, F; Datta, N: *The quantum capacity of channels with arbitrarily correlated noise*. IEEE Tran. Inf. Th., vol. 56 (3), pp.1447-1460 (2010). arXiv link