Computational Relative Entropy

ÉTS Montreal

Johannes Jakob Meyer

Freie Universität Berlin

Guiding Question

Can **computational restrictions** be built into **quantum information theory**?

Part I

Why is this even **necessary**?

Relative Entropy

RELATIVE ENTROPY
Measure of dissimilarity of quantum states (a.k.a. KL Divergence)

$$D(\rho \,\|\, \sigma) = \mathrm{Tr}[\rho(\log \rho - \log \sigma)]$$

FOUND ALL OVER INFORMATION THEORY

- Hypothesis testing
- Sanov's Theorem
- Log-Sobolev Inequalities
- Transportation-Cost Inequalities
- ...

PARENT QUANTITY FOR

- Entropy
- Mutual Information
- Coherent Information
- Holevo Chi
- ...

Single-shot Approach to Relative Entropy

HYPOTHESIS TESTING RELATIVE ENTROPY

Operational definition of the best type II error achievable at fixed type I error

$$D_h^\epsilon(\rho \,\|\, \sigma) := -\log \min_{0 \leq \Lambda \leq \mathbb{I}} \left\{ \mathrm{Tr}[\Lambda\sigma] \,|\, \mathrm{Tr}[\Lambda\rho] \geq 1 - \epsilon \right\}$$

QUANTUM STEIN'S LEMMA

Best asymptotic exponent in the limit of many copies

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

Operational Interpretation is Inherited

HYPOTHESIS TESTING ENTROPY

Optimal rank of projector that compresses a quantum state

$$-D_h^\epsilon(\rho \,\|\, \mathbb{I}) := \log \min_{0 \leq \Lambda \leq \mathbb{I}} \left\{ \mathrm{Tr}[\Lambda] \,|\, \mathrm{Tr}[\Lambda\rho] \geq 1 - \epsilon \right\}$$

QUANTUM STEIN'S LEMMA

Implies this is equal to the regular entropy asymptotically

$$\lim_{m \to \infty} -\frac{1}{m} D_h^\epsilon(\rho^{\otimes m} \,\|\, \mathbb{I}) = -D(\rho \,\|\, \mathbb{I}) = S(\rho)$$

Computational Setting

STATE FAMILIES

We consider families of quantum states indexed by a scaling variable n, for example a family of n-qubit states

$$\rho_n \Leftrightarrow (\rho_n)_{n\in\mathbb{N}}$$

COMPUTATIONAL EFFICIENCY

We say that a computational is efficient if it has a complexity scaling **polynomially in n**. Therefore, how the scaling variable enters into the state set is crucial, but this usually emerges naturally.

Optimal Test

LIKELIHOOD RATIO TEST
For classical distributions

$$f(x) = \begin{cases} 1 & \text{if } \log \frac{p(x)}{q(x)} \geq R \\ 0 & \text{else.} \end{cases}$$

I.I.D. SETTING
For m i.i.d. copies of p and q

$$f(x_1, \ldots, x_m) = \begin{cases} 1 & \text{if } \frac{1}{m} \sum_{i=1}^{m} \log \frac{p(x_i)}{q(x_i)} \geq R \\ 0 & \text{else.} \end{cases}$$

Optimal Test

LIKELIHOOD RATIO CONCENTRATES

For true distribution p

$$\frac{1}{m}\sum_{i=1}^{m}\log\frac{p(x_i)}{q(x_i)} \to \sum_{x} p(x)\log\frac{p(x)}{q(x)} = D(p\,\|\,q)$$

OPTIMAL THRESHOLD

To achieve 1-epsilon overlap with p, we need to choose

$$R \approx D(p\,\|\,q) + O(\Phi^{-1}(\epsilon)/\sqrt{m})$$

Optimal Test

IMPLEMENTATION
Measure and predict

$$(x_1, x_2, \ldots, x_m) \to f(x_1, x_2, \ldots, x_m) \in \{0, 1\}$$

SIMPLEST ALGORITHM
Store the log-likelihood for all symbols and compute f in time poly(m,d)

$$f : [d]^m \to \{0, 1\}$$

The dimension d usually scales **exponentially** – the above is only possible if the logic inside of f is algorithmically compressible.

We observe that

The **operational interpretation** of relative entropy requires **unbounded** computational power

Part II

How we add **computational restrictions** to the theory

Adding Computational Restrictions

G-COMPLEXITY HYPOTHESIS TESTING RELATIVE ENTROPY[1]

Operational definition of the best type II error achievable at fixed type I error and fixed gate complexity of the measurement

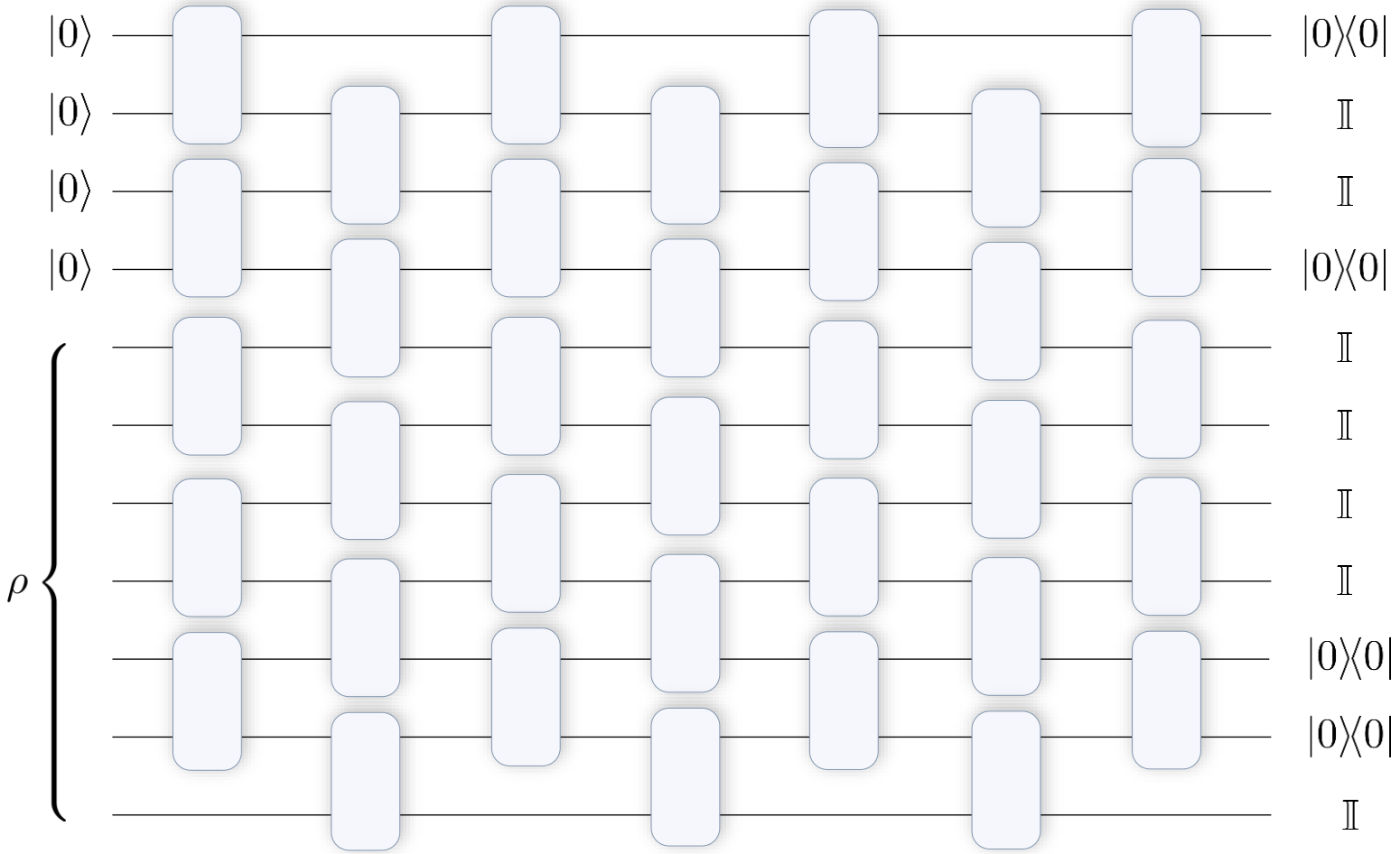
$$D_h^\epsilon(\rho \,\|\, \sigma; G) := -\log \min_{0 \leq \Lambda \leq \mathbb{I}} \left\{ \mathrm{Tr}[\Lambda\sigma] \mid \mathrm{Tr}[\Lambda\rho] \geq 1 - \epsilon, C(\Lambda) \leq G \right\}$$

GATE COMPLEXITY

The gate complexity is the number of elementary 2-local gates needed to implement a measurement with access to auxiliary systems in the zero state and measurement as projection onto the zero state

[1]Munson et al., PRX Quantum 6 010346 (2025)

Computational Model for Measurements



Computational Relative Entropy

COMPUTATIONAL RELATIVE ENTROPY

The G-complexity hypothesis testing relative entropy is a single-shot quantity. To get something like relative entropy, we have to regularize.

We define the computational relative entropy as the best asymptotic exponent for complexity-constrained hypothesis testing in the limit of polynomially many copies

$$\underline{D}(\rho_n \,\|\, \sigma_n) :\simeq \lim_{\epsilon \to 0} \lim_{\ell \to \infty} \liminf_{k \to \infty} \frac{1}{n^k} D_h^\epsilon(\rho_n^{\otimes n^k} \,\|\, \sigma_n^{\otimes n^k}; n^{k\ell})$$

Part III

Unpacking the definition

Computational Relative Entropy

COMPUTATIONAL RELATIVE ENTROPY

$$\underline{D}(\rho_n \,\|\, \sigma_n) :\simeq \lim_{\epsilon \to 0} \lim_{\ell \to \infty} \liminf_{k \to \infty} \frac{1}{n^k} D_h^\epsilon(\rho_n^{\otimes n^k} \,\|\, \sigma_n^{\otimes n^k} ; n^{k\ell})$$

Weak Converse Limit

Polynomial Regularization

Polynomial Gate Complexity

Polynomial Regularization

REGULARIZATION

In unbounded quantum information theory, regularization allows us to get rid of finite-size correction terms. We want something similar, but it should work if we only regularize over a polynomially large number of copies.

Idea: Ignore negligible contributions!

NEGLIGIBLE FUNCTIONS

A function is negligible if it vanishes faster than any polynomial asymptotically

$$f(n) = \mathrm{negl}(n) \quad \Leftrightarrow \quad \text{for all } k \in \mathbb{N} : \lim_{n \to \infty} \frac{|f(n)|}{n^k} = 0$$

Formalization

PSEUDOMETRIC OF ASYMPTOTIC EXPONENT

We define the following pseudometric that allows us to rigorously capture equality up to negligible functions and to define limits involving polynomials

$$d(f,g) := \limsup_{n\to\infty} \begin{cases} \exp\left(\frac{\log|f(n)-g(n)|}{\log n}\right) & \text{if } f(n) \neq g(n), \\ 0 & \text{else.} \end{cases}$$

EQUIVALENCE RELATION

We define functions to be equivalent if they only differ by a negligible function.

$$f \simeq g \Leftrightarrow d(f,g) = 0.$$

NEGLIGIBLE FUNCTIONS

Pseudometric captures negligible functions as

$$f(n) = \mathrm{negl}(n) \Leftrightarrow d(f,0) = 0.$$

The Workhorse

LIMIT OF INVERSE POLYNOMIALS
The limit of inverse polynomials is negligible

$$\lim_{k\to\infty} \frac{1}{n^k} \simeq 0.$$

PROOF
The proof follows directly from the definition.

$$\begin{aligned}\lim_{k\to\infty} d(n^{-k}, 0) &= \lim_{k\to\infty} \lim_{n\to\infty} \exp\left(\frac{\log n^{-k}}{\log n}\right) \\ &= \lim_{k\to\infty} e^{-k} \\ &= 0.\end{aligned}$$

Part IV

Can we **prove** anything **useful** about this quantity?

Properties of the Computational Relative Entropy

- Data-processing under poly-complexity maps

$$C(\Phi_n) \leq O(\mathrm{poly}(n)) \Rightarrow \underline{D}(\Phi_n[\rho_n] \,\|\, \Phi_n[\sigma_n]) \lesssim \underline{D}(\rho_n \,\|\, \sigma_n)$$

- Superadditivity under tensor products (can be strict!)

$$\underline{D}(\rho_n \otimes \tau_n \,\|\, \sigma_n \otimes \upsilon_n) \gtrsim \underline{D}(\rho_n \,\|\, \sigma_n) + \underline{D}(\tau_n \,\|\, \upsilon_n)$$

- Additivity under tensor powers

$$m \leq O(\mathrm{poly}(n)) \Rightarrow \underline{D}(\rho_n \,\|\, \sigma_n) \simeq \frac{1}{m} \underline{D}(\rho_n^{\otimes m} \,\|\, \sigma_n^{\otimes m})$$

Computational Pinsker's Inequality

COMPUTATIONAL TOTAL VARIATION DISTANCE

Apply the same idea to a symmetric hypothesis test between two quantum states

$$d_{\mathrm{TV}}(\rho, \sigma; G) := \max_{C(\Lambda) \leq G} \frac{1}{2} \operatorname{Tr}[\Lambda(\rho - \sigma)]$$

$$\underline{d}_{\mathrm{TV}}(\rho_n, \sigma_n) :\simeq \lim_{\ell \to \infty} d_{\mathrm{TV}}(\rho_n, \sigma_n; n^\ell)$$

COMPUTATIONAL PINSKER

$$\underline{d}_{\mathrm{TV}}(\rho_n, \sigma_n) \lesssim \sqrt{\frac{1}{2} \underline{D}(\rho_n \,\|\, \sigma_n)}$$

Computational Smoothing

COMPUTATIONAL INDISTINGUISHABILITY

We say two quantum states are computationally indistinguishable if for all m = poly(n).

$$\underline{d}_{\mathrm{TV}}(\rho_n^{\otimes m}, \sigma_n^{\otimes m}) \simeq 0$$
$$\Leftrightarrow \ \rho_n \approx_c \sigma_n$$

COMPUTATIONAL SMOOTHING

The computational relative entropy has the computational smoothign property

$$\underline{D}(\rho_n \,\|\, \sigma_n) \simeq \min_{\tilde{\rho}_n \approx_c \rho_n} \underline{D}(\tilde{\rho}_n \,\|\, \sigma_n)$$

PSEUDO RELATIVE ENTROPY BOUND

$$\underline{D}(\rho_n \,\|\, \sigma_n) \lesssim \min_{\tilde{\rho}_n \approx_c \rho_n} D(\tilde{\rho}_n \,\|\, \sigma_n)$$

Computational Stein‘s Lemma

COMPUTATIONALLY MEASURED RELATIVE ENTROPY

Regularizing the relative entropy measured over efficient two-outcome measurements

$$\underline{D}^{\mathcal{D}}(\rho_n \,\|\, \sigma_n) :\simeq \lim_{\ell\to\infty} \liminf_{k\to\infty} \frac{1}{n^k} \max_{C(\mathcal{M})\leq n^{k\ell}} D(\mathcal{M}[\rho_n^{\otimes n^k}] \,\|\, \mathcal{M}[\sigma_n^{\otimes n^k}])$$

COMPUTATIONAL STEIN‘S LEMMA

Under regularity assumptions on rho and sigma, we have

$$\underline{D}(\rho_n \,\|\, \sigma_n) \simeq \underline{D}^{\mathcal{D}}(\rho_n \,\|\, \sigma_n)$$

Part V

Can we **prove** Something **useful** with this quantity?

Computational Separations

- Unconditionally,

$$\underline{D}(\rho_n \,\|\, \sigma_n) \simeq 0, \text{ but } D(\rho_n \,\|\, \sigma_n) = \infty, C(\rho_n, \sigma_n) = O(\exp(n))$$

- Conditioned on quantum-hard one-way functions

$$\underline{D}(\rho_n \,\|\, \sigma_n) \simeq 0, \text{ but } D(\rho_n \,\|\, \sigma_n) = \infty, C(\rho_n, \sigma_n) = O(\text{poly}(n))$$

- Conditioned on classically-hard quantum-easy one-way functions

$$\underline{D}^{\text{cl}}(\rho_n \,\|\, \sigma_n) \simeq 0, \text{ but } \underline{D}(\rho_n \,\|\, \sigma_n) \gtrsim \omega(\log n)$$

State Compression

COMPUTATIONAL ENTROPY
We derive the computational entropy in analogy to the unbounded setting

$$\underline{S}(\rho_n) :\simeq -\underline{D}(\rho_n \,\|\, \mathbb{I})$$

PROPERTIES
The computational entropy inherits many properties of the computational relative entropy, like computational smoothing

$$\underline{S}(\rho_n) \simeq \max_{\tilde{\rho}_n \approx_c \rho_n} \underline{S}(\tilde{\rho}_n)$$

COMPRESSION
We prove that the computational entropy as defined above exactly quantifies the rate at which we can reliably compress a quantum state when compression and expansion have to be computationally efficient

Entanglement Theory

COMPUTATIONAL RAINS BOUND

The best rate at which we can distill entanglement from a quantum state with computationally efficient operations is bounded by

$$\underline{E}_D(\rho_n^{A_n B_n}) \lesssim \underline{R}_D(\rho_n^{A_n B_n}) :\simeq \min_{\sigma_n^{A_n B_n}\ \mathrm{PPT}} \underline{D}(\rho_n^{A_n B_n} \,\|\, \sigma_n^{A_n B_n})$$

ENTANGLEMENT DILUTION VIA COMPRESSION

The best rate at which we can create an entangled state from a shared resource with computationally efficient operations is bounded by

$$\underline{E}_C(\rho_n^{A_n B_n}) \lesssim \min\{\underline{S}(\rho_n^{A_n}), \underline{S}(\rho_n^{B_n})\}$$

Takeaway I

Information-theoretic bounds can be **arbitrarily far** from the actual achievable performance

Takeaway II

We can define **entropic quantities** that have **computational limitations** inbuilt

Takeaway III

We have a **computational quantum information theory** with the same **look and feel** of the unbounded theory

Outlook

Lots of
exciting science
is left to be done

Thank you for your attention.

johannesjakobmeyer.com
jjmeyer@bsky.social

Paper

Slides