Error Exponents and Strong Converse for Quantum Soft Covering

(arXiv:2202.10995)

Hao-Chung Cheng with Li Gao

haochung@ntu.edu.tw

Department of Electrical Engineering
National Taiwan University

lgao12@uh.edu

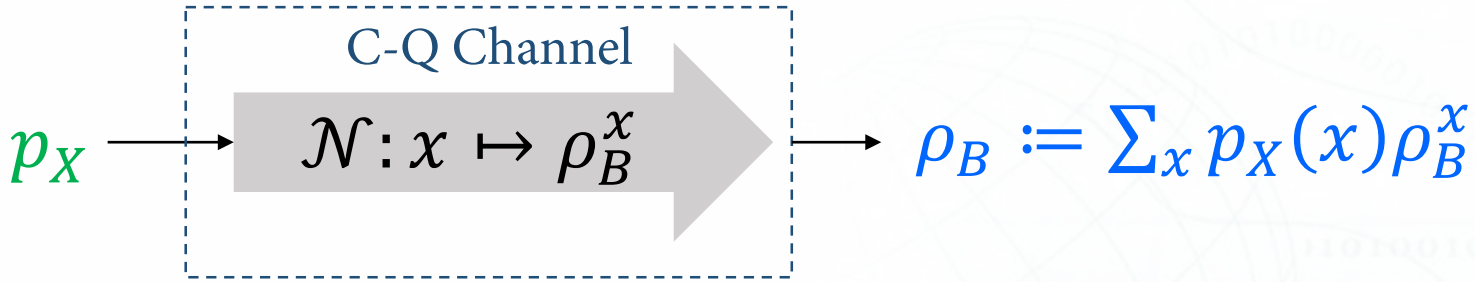
Department of Mathematics
University of Houston

Beyond IID 2022 (Sept. 26 – Sept. 30, 2022, SUSTech)

Problem Formulation

Problem Setting (1/2)

- Fix a classical-quantum (c-q) channel $\mathcal{N}: x \mapsto \rho_B^x$.
  - Input: a classical symbol x in an alphabet $\mathcal{X}$
  - Output: a density operator ρ_B^x on $\mathcal{H}_B$, i.e. $\rho_B^x \geq 0$, $\mathrm{Tr}[\rho_B^x] = 1$
- Given an input distribution p_X on $\mathcal{X}$, the output is:



Question: How well can we approximate ρ_B by querying to the channel?

Problem Setting (2/2)

- Random codebook: $\mathcal{C} := \{x(1), x(2), \ldots, x(M)\}$
  - Each $m \in [1{:}M]$ is equiprobable.
  - Each codeword is *pairwise independent*: $x(m) \sim p_X$.
  - Querying the channel: $x(m) \mapsto \rho_B^{x(m)}$.

Codebook-induced output density

$m \sim$ unif. on $[1{:}M]$ → Codebook $\mathcal{C}$ $m \mapsto x(m) \sim p_X$ → C-Q Channel $\mathcal{N}: x \mapsto \rho_B^x$ → $\rho_B^{\mathcal{C}} := \frac{1}{M}\sum_{x\in\mathcal{C}} \rho_B^x$

Question: For each M, how close is $\rho_B^{\mathcal{C}}$ to the desired marginal ρ_B?

Error Criterion

Schatten 1-norm
$\|A\|_1 := \mathrm{Tr}[|A|]$

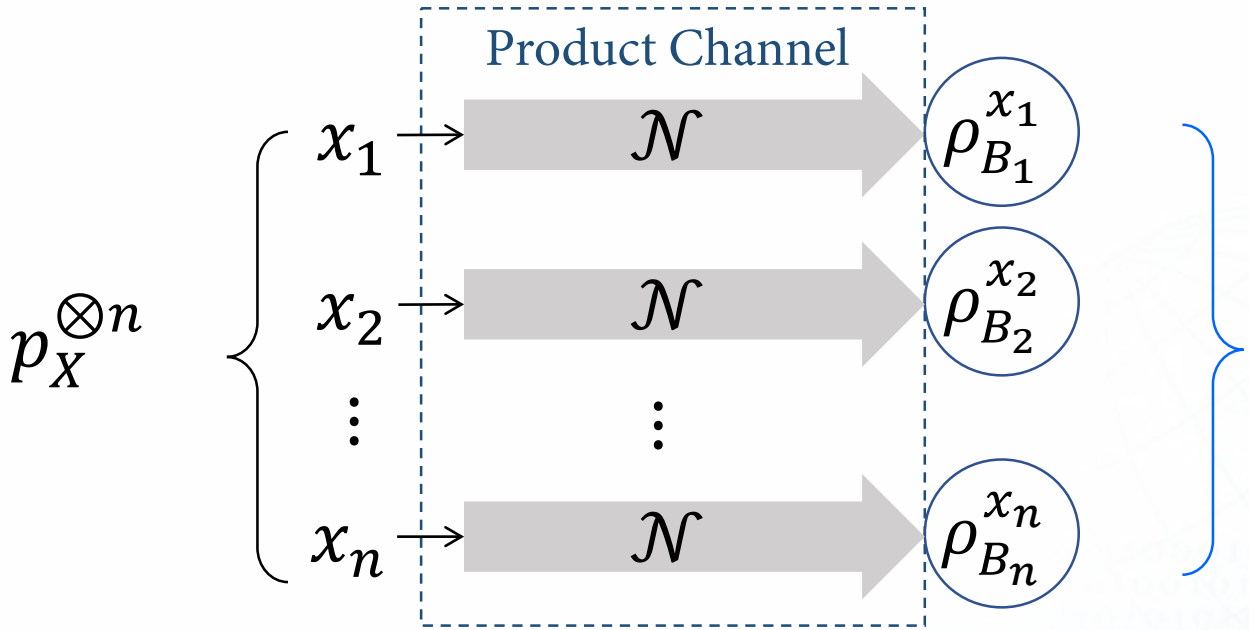
- Trace distance: $\Delta\left(\rho_B^{\mathcal{C}}, \rho_B\right) := \frac{1}{2}\left\|\rho_B^{\mathcal{C}} - \rho_B\right\|_1$.

Question:
For each a codebook size M, how small/large is $\mathbb{E}_{\mathcal{C}}\Delta\left(\rho_B^{\mathcal{C}}, \rho_B\right)$?

- We call this the "**one-shot quantum soft covering**."
  - By "one-shot", we do not impose specific structures on the channel.

Memoryless Asymptotic Setting

- Product (memoryless) channel: $\mathcal{N}^{\otimes n}: x^n \mapsto \rho_{B^n}^{x^n} := \rho_{B_1}^{x_1} \otimes \cdots \otimes \rho_{B_n}^{x_n}$.
- Case I: IID input distribution $p_{X^n} = p_X^{\otimes n}$ on $\mathcal{X}^n$.

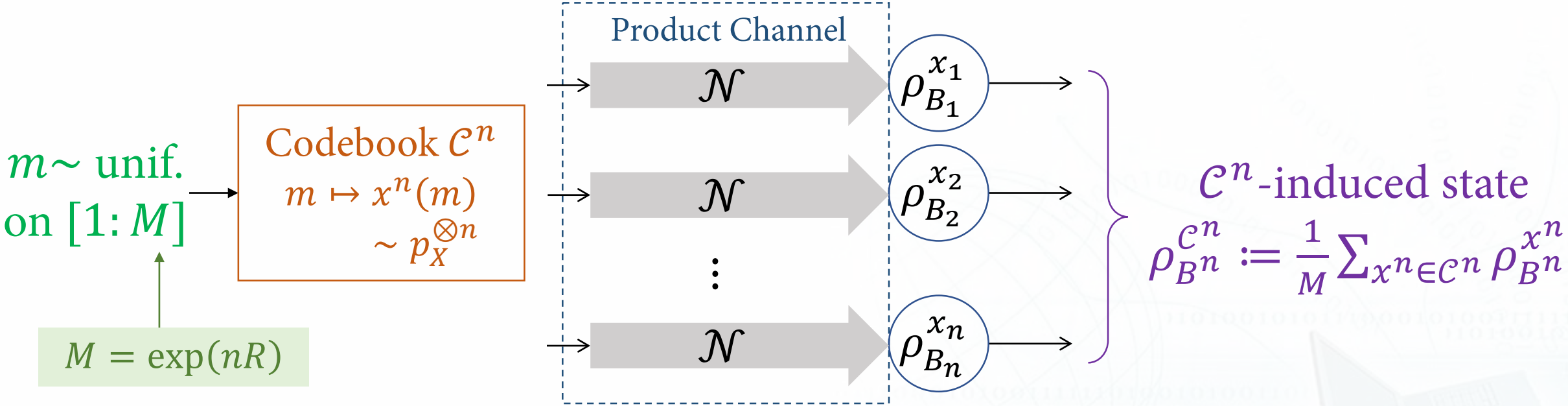


True marginal density

$$\mathcal{N}^{\otimes n}\left(p_X^{\otimes n}\right) = \rho_B^{\otimes n}$$

Memoryless Asymptotic Setting (IID Random Codebook)

- Product (memoryless) channel: $\mathcal{N}^{\otimes n}: x^n \mapsto \rho_{B^n}^{x^n} := \rho_{B_1}^{x_1} \otimes \cdots \otimes \rho_{B_n}^{x_n}$.
- Case I: IID input distribution $p_{X^n} = p_X^{\otimes n}$.

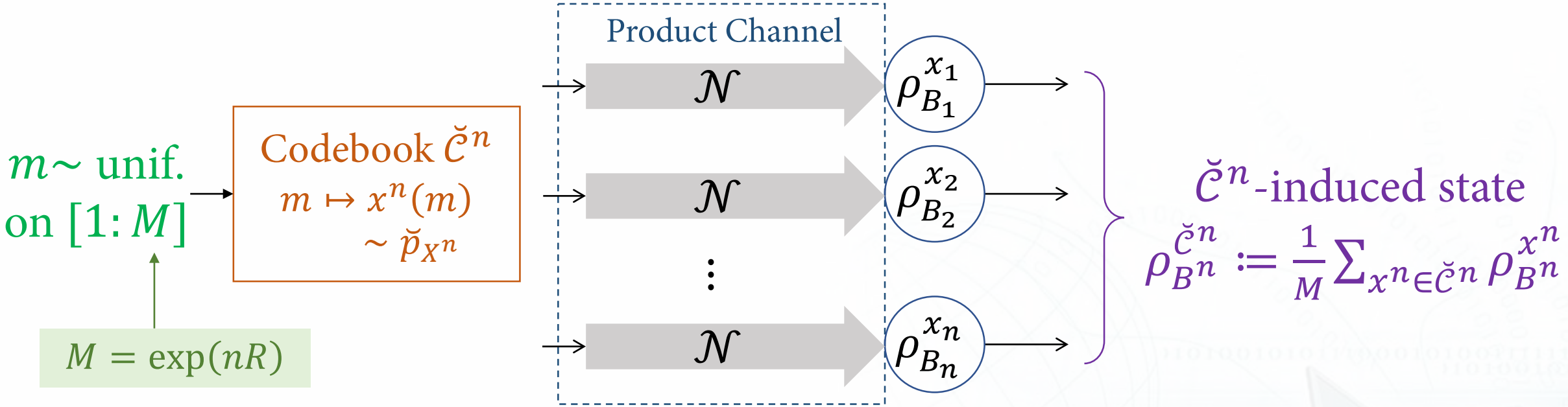


Question: How fast does $\rho_{B^n}^{\mathcal{C}^n}$ converge to $\rho_B^{\otimes n}$ (in n)?

Memoryless Asymptotic Setting (Constant Composition)

- Case II: constant composition input distribution $\breve{p}_{X^n} \overset{\mathcal{N}^{\otimes n}}{\longmapsto} \breve{\rho}_{B^n}$

  (uniform on sequences x^n with empirical distribution p_X)



Question: How fast does $\rho_{B^n}^{\breve{\mathcal{C}}^n}$ converge to $\breve{\rho}_{B^n}$ (in n)?

Literature Review

Historical Remarks (1/4)

- Studying the common information [Wyner'75]
  - It is termed *soft covering* in classical setting

- Approximating output statistics [Han–Verdú'93]
  - Via optimal codebook → *channel resolvability*
  - $R > I(X:Y)$: expected total variation distance vanishes asymptotically
  - $R < I(X:Y)$: weak converse (for full-rank channels)

Historical Remarks (2/4)

- Applications:
  - Identification via channels [Ahlswede–Dueck'89]
  - Random process simulation [Steinberg–Verdú'96]
  - Empirical distribution of good codes [Shamai–Verdú'97], [Polyanskiy–Verdú'97]
  - Channel synthesis [Cuff'13]
  - Lossy source coding [Song–Cuff–Poor'16]
  - Secrecy [Csiszár'96], [Hayashi'06], [Bloch–Laneman'13], [Yu–Tan'19], [Goldfeld–Cuff–Permuter'16], [Parizi–Telatar–Merhav'16], etc…
  - Coding with side information [Watanabe–Kuzuoka–Tan'15]
  - …

Historical Remarks (3/4)

- A *one-shot achievability bound* for soft covering [Hayashi'06]
  - Exponential convergence as $R > I(X:Y)$

- [Cuff'13] improved the achievable error exponent (based on ↑).
- [Yagli–Cuff'19]: optimal error exponents.
  - Asymptotic results for both IID & constant composition codebooks
  - Method of types & [Parizi–Telatar–Merhav'16]

- [Yassaee'19, ISIT'20]: exact polynomial prefactors.
  - A one-shot exponential achievability bound based on Hayashi's result

Historical Remarks (4/4)

- Various distance measures
  - Relative entropy: [Hayashi'06], [Hayashi–Matsumoto'16], [Hayashi'15] [Parizi–Telatar–Merhav'16]: optimal error exponent (for IID & CC)
  - Rényi divergences [Yu–Tan'19]: optimal error exponents
  - Wasserstein distance [Steinberg–Verdú'96]
  - E_γ metric [Liu–Cuff–Verdú'15]
- Strong Converse, i.e. $R < I(X:Y)$
  - [Watanabe–Hayashi: ISIT'14]: strong converse property for *resolvability*; second-order rate & exponential strong converse (with symmetry assump.)
  - [Cuff: ISIT'16]: second-order rate with a probabilistic argument

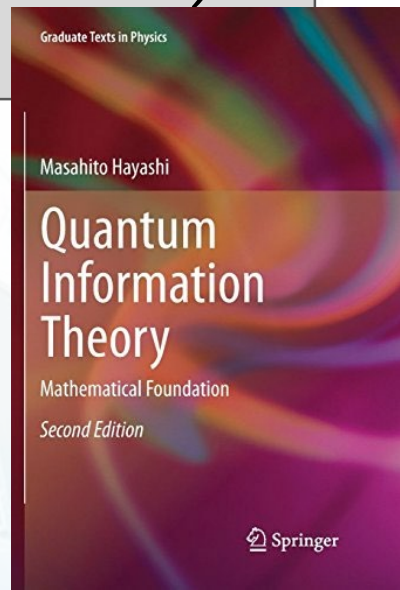
Historical Remarks (Quantum Setting)

- A probability deviation bound

Operator Chernoff Bound [Ahlswede–Winter'02]

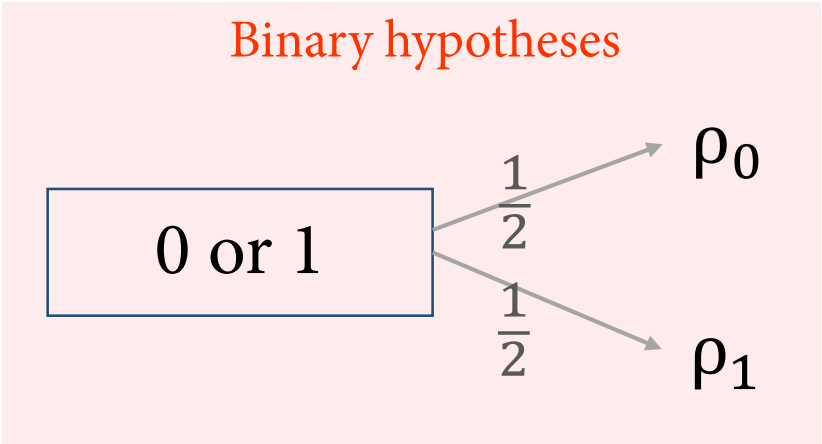
$$\Pr_{\mathcal{C}}\left\{\left|\rho_B^{\mathcal{C}} - \rho_B\right| \preccurlyeq \epsilon\rho_B\right\} \geq 1 - 2\dim(\mathcal{H}_B)\exp\left(-\frac{M\epsilon^2\lambda_{\min}(\rho_B)}{4\ln 2}\right)$$

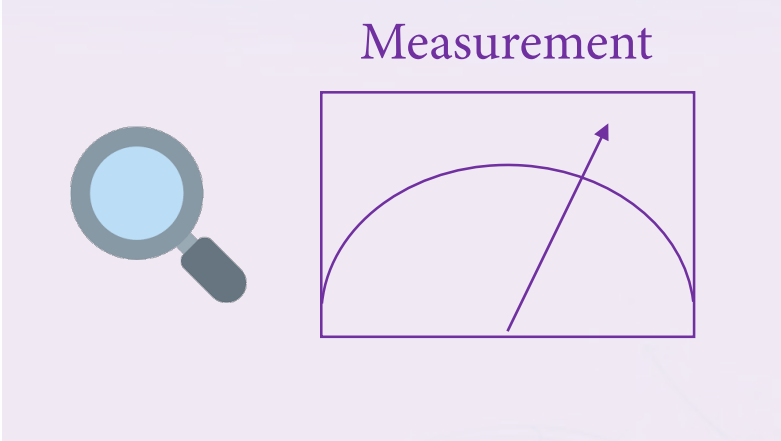
- [Hayashi, *IEEE Trans. Inf. Theory*, 61(10), 2015]
- [Hayashi'17, §9.4]



Why Using Trace Distance?

- Holevo'72–Helström'67 Theorem on quantum state discrimination





0 or 1?

$$\Rightarrow P_{\mathrm{e}}^{\star} = \frac{1}{2}[1 - \Delta(\rho_0, \rho_1)]$$

- We would like $\rho_B^{\mathcal{C}}$ to be *physically indistinguishable* from ρ_B.
- Bounds via other distance measures might not be tight for $\Delta(\cdot,\cdot)$.

Results

Quantum Information Quantities (1/2)

- Two quantum Rényi divergences:

Schatten α-norm:
$\|A\|_{S_\alpha(\mathcal{H})} \equiv \|A\|_\alpha := (\mathrm{Tr}[|A|^\alpha])^{1/\alpha}$

Sandwiched Rényi divergence $D_\alpha^*(\rho\|\sigma) := \frac{1}{\alpha-1}\log \mathrm{Tr}\left[\left(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right]$

$$= \frac{\alpha}{\alpha-1}\log\left\|\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}\right\|_\alpha$$

Petz–Rényi divergence $D_\alpha(\rho\|\sigma) := \frac{1}{\alpha-1}\log \mathrm{Tr}[\rho^\alpha\sigma^{1-\alpha}]$

Quantum Information Quantities (2/2)

- Given p_X, $x \mapsto \rho_B^x$, define joint state $\rho_{XB} := \sum_x p_X(x)|x\rangle\langle x| \otimes \rho_B^x$.

Sandwiched **Rényi** information: $I_\alpha^*(X:B)_\rho := \inf_{\sigma_B} D_\alpha^*(\rho_{XB} \| \rho_X \otimes \sigma_B)$

Sandwiched **Augustin** information: $\breve{I}_\alpha^*(X:B)_\rho := \inf_{\sigma_B} \mathbb{E}_{p_X} D_\alpha^*(\rho_B^x \| \sigma_B)$

A variant Petz–**Rényi** information: $I_\alpha^\downarrow(X:B)_\rho := D_\alpha(\rho_{XB} \| \rho_X \otimes \rho_B)$

A variant Petz–**Augustin** information $\breve{I}_\alpha^\downarrow(X:B)_\rho := \mathbb{E}_{p_X} D_\alpha(\rho_B^x \| \rho_B)$

Published: 04 February 2022

Properties of Noncommutative Rényi and Augustin Information

Hao-Chung Cheng, Li Gao & Min-Hsiu Hsieh

Communications in Mathematical Physics **390**, 501–544 (2022) | Cite this article

1668 IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. 67, NO. 3, MARCH 2021

Divergence Radii and the Strong Converse Exponent of Classical-Quantum Channel Coding With Constant Compositions

Milán Mosonyi and Tomohiro Ogawa

A One-Shot Error Exponent

- For every p_X, $x \mapsto \rho_B^x$, i.e. $\rho_{XB} := \sum_x p_X(x)|x\rangle\langle x| \otimes \rho_B^x$,

Theorem. $\mathbb{E}_{\mathcal{C}}\Delta\left(\rho_B^{\mathcal{C}}, \rho_B\right) \leq \mathrm{e}^{-E^*(\log M)}$

- Error exponent: $E^*(\log M) := \sup_{\alpha\in[1,2]} \frac{\alpha-1}{\alpha}\left(\log M - I_\alpha^*(X:B)_\rho\right)$
  - $E^*(\log M) > 0 \Leftrightarrow \log M > I_1^*(X:B)_\rho = I(X:B)_\rho$
  - **No** closed-form expression for I_α^* in the noncommutative case.
    See [arXiv:2109.06054 [cs.IT] an optimization algorithm with convergence guarantees.

A One-Shot Strong Converse

- For every p_X, $x \mapsto \rho_B^x$ (i.e. $\rho_{XB} := \sum_x p_X(x)|x\rangle\langle x| \otimes \rho_B^x$),

Theorem. $\mathbb{E}_{\mathcal{C}}\Delta\left(\rho_B^{\mathcal{C}}, \rho_B\right) \geq 1 - 4\mathrm{e}^{-E^{\downarrow}(\log M)}$

- Exponent: $E^{\downarrow}(\log M) := \sup_{\alpha\in[1/2,1]} \frac{1-\alpha}{\alpha}\left(I^{\downarrow}_{2-1/\alpha}(X:B)_\rho - \log M\right)$
  - $E^{\downarrow}(\log M) > 0 \Leftrightarrow \log M < I^{\downarrow}_1(X:B)_\rho = I(X:B)_\rho$

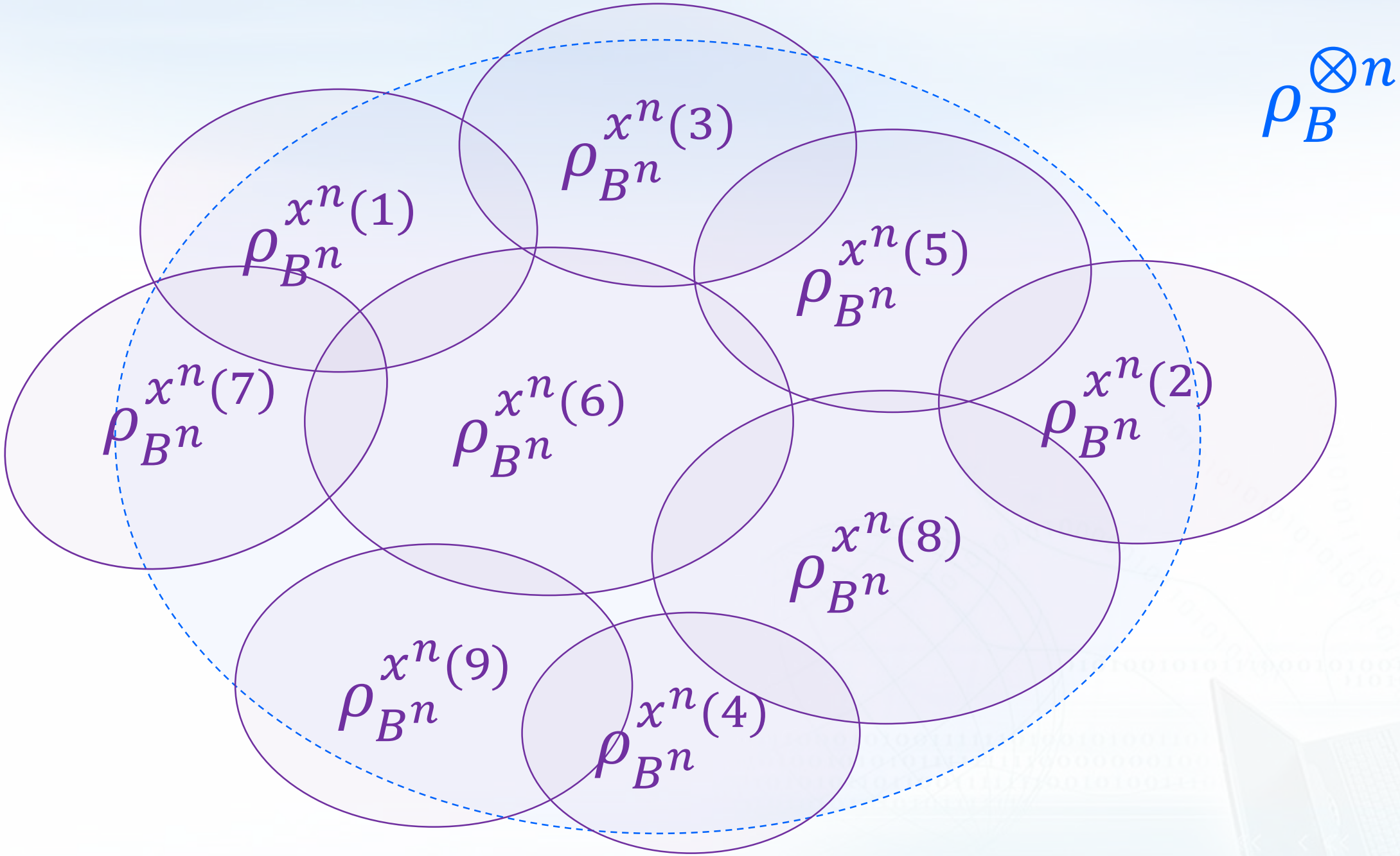
Asymptotic Setting (IID Random Codebook)

- IID input distribution: $p_{X^n} = p_X^{\otimes n}$; $\rho_{X^n B^n} = \rho_{XB}^{\otimes n}$
- Additivity of $\rho \mapsto I_\alpha^*(X:B)_\rho$ and $D_\alpha(\cdot \,\|\, \cdot)$

Rate: $R = \frac{1}{n}\log M$

$$\begin{cases} \mathbb{E}_{\mathcal{C}^n}\Delta\left(\rho_{B^n}^{\mathcal{C}^n}, \rho_B^{\otimes n}\right) \leq \mathrm{e}^{-nE^*(R)}, & R > I(X:B)_\rho \\ \mathbb{E}_{\mathcal{C}^n}\Delta\left(\rho_{B^n}^{\mathcal{C}^n}, \rho_B^{\otimes n}\right) \geq 1 - 4\mathrm{e}^{-nE^{\downarrow}(R)}, & R < I(X:B)_\rho \end{cases}$$

- The above results hold for **any** finite blocklength n.
- $I(X:B)_\rho$ is the minimum achievable rate & strong converse rate.


$\rho_{B^n}^{x^n(1)}$
$\rho_{B^n}^{x^n(3)}$
$\rho_{B^n}^{x^n(5)}$
$\rho_{B^n}^{x^n(7)}$
$\rho_{B^n}^{x^n(6)}$
$\rho_{B^n}^{x^n(2)}$
$\rho_{B^n}^{x^n(8)}$
$\rho_{B^n}^{x^n(9)}$
$\rho_{B^n}^{x^n(4)}$
$\rho_B^{\otimes n}$

Asymptotic Setting (Constant Composition)

- Constant composition input distribution $\breve{p}_{X^n} \overset{\mathcal{N}^{\otimes n}}{\longmapsto} \breve{\rho}_{B^n}$
  (uniform on sequences x^n with empirical distribution p_X)

$$\begin{cases} \mathbb{E}_{\breve{\mathcal{C}}^n} \Delta\left(\rho_{B^n}^{\breve{\mathcal{C}}^n}, \breve{\rho}_{B^n}\right) \leq \mathrm{e}^{-n\breve{E}^*(R)}, & R > I(X:B)_\rho \\ \mathbb{E}_{\breve{\mathcal{C}}^n} \Delta\left(\rho_{B^n}^{\breve{\mathcal{C}}^n}, \breve{\rho}_{B^n}\right) \geq 1 - n^{k_p} \mathrm{e}^{-n\breve{E}^\downarrow(R)}, & R < I(X:B)_\rho \end{cases}$$

- $\breve{E}^*(R) := \sup_{\alpha \in [1,2]} \frac{\alpha-1}{\alpha}\left(R - \breve{I}_\alpha^*(X:B)_\rho\right)$
- $\breve{E}^\downarrow(R) := \sup_{\alpha \in [1/2,1]} \frac{1-\alpha}{\alpha}\left(\breve{I}_{2-1/\alpha}^\downarrow(X:B)_\rho - R\right)$

Augustin information

Remarks

- Comparisons to classical results:
  - Our one-shot achievability (using IID random codebooks) recovers the one-shot bound by [Yassaee'19, ISIT'20] in the commuting setting.
  - Our results (using both random codebooks) imply the achievability in [Yagli–Cuff'19].
- We employ different proof techniques, i.e. **norm interpolation**.
- By Jensen's inequality and concavity of log,

$$\breve{E}^*(R) = \sup_{\alpha\in[1,2]} \frac{\alpha-1}{\alpha}\left(R - \breve{I}^*_\alpha(X:B)_\rho\right) \geq E^*(R) = \sup_{\alpha\in[1,2]} \frac{\alpha-1}{\alpha}\left(R - I^*_\alpha(X:B)_\rho\right)$$

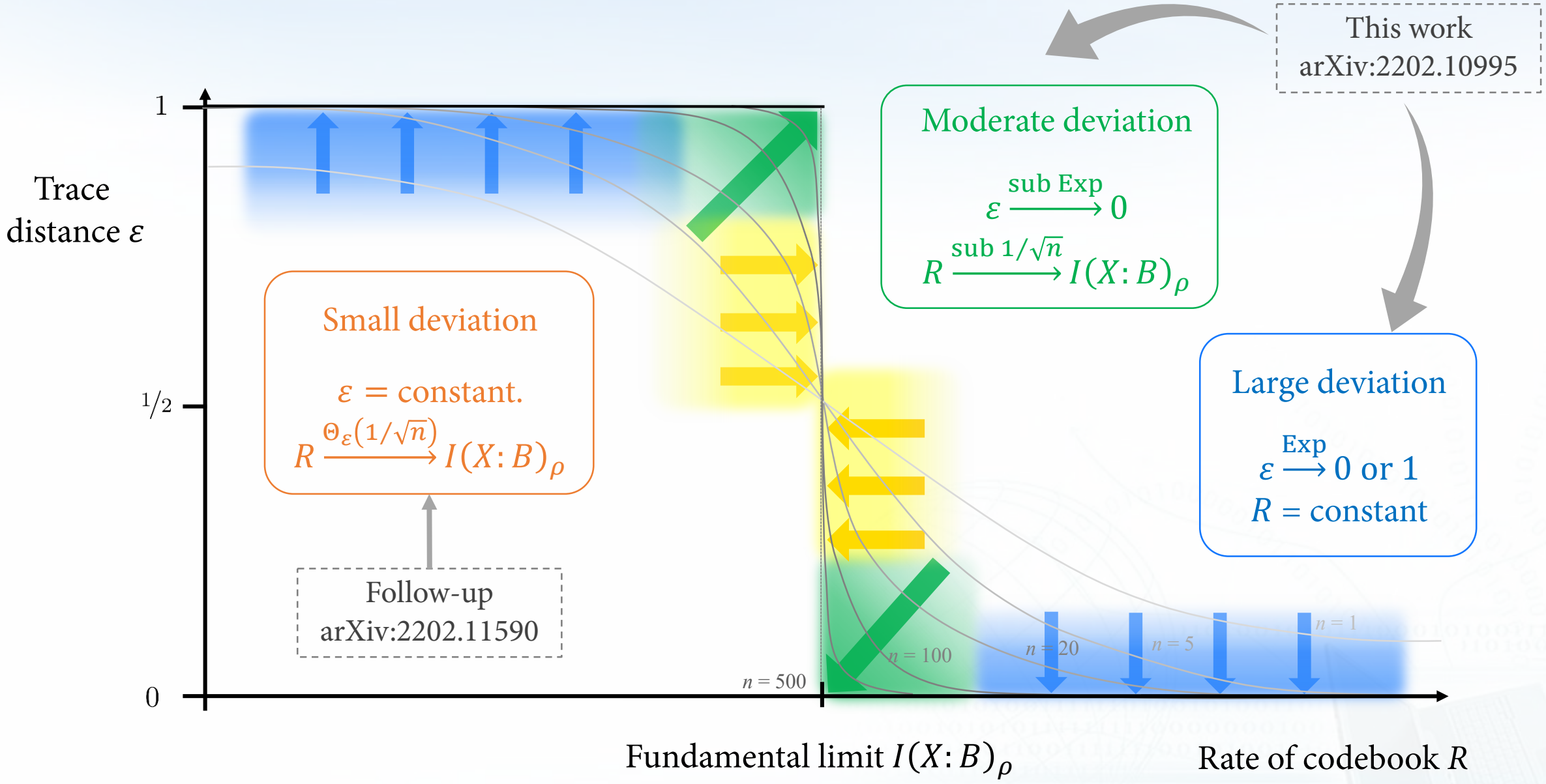
The error exponent is larger when using constant composition codebook.

Moderate Deviation Asymptotics

- (Large deviation) $R > I(X:B)_\rho$: Exponential convergence.
- (Moderate deviation) $R_n = I(X:B)_\rho + a_n$
  - Moderate deviation sequence: (i) $a_n \downarrow 0$; (ii) $\sqrt{n}a_n \uparrow \infty$.

$$\text{IID codebook: } \mathbb{E}_{\mathcal{C}^n}\Delta\left(\rho_{B^n}^{\mathcal{C}^n}, \rho_B^{\otimes n}\right) \lesssim \mathrm{e}^{-\frac{na_n^2}{2V(X:B)_\rho}} \to 0,$$

$$\text{CC codebook: } \mathbb{E}_{\check{\mathcal{C}}^n}\Delta\left(\rho_{B^n}^{\check{\mathcal{C}}^n}, \check{\rho}_{B^n}\right) \lesssim \mathrm{e}^{-\frac{na_n^2}{2\check{V}(X:B)_\rho}} \to 0.$$


This work
arXiv:2202.10995
Moderate deviation
$\varepsilon \xrightarrow{\text{sub Exp}} 0$
$R \xrightarrow{\text{sub } 1/\sqrt{n}} I(X:B)_\rho$
Large deviation
$\varepsilon \xrightarrow{\text{Exp}} 0$ or 1
R = constant
Small deviation
ε = constant.
$R \xrightarrow{\Theta_\varepsilon(1/\sqrt{n})} I(X:B)_\rho$
Follow-up
arXiv:2202.11590
Trace distance ε
1
1/2
0
$n = 500$
$n = 100$
$n = 20$
$n = 5$
$n = 1$
Fundamental limit $I(X:B)_\rho$
Rate of codebook R

Proof Sketch and Ideas

Proof Sketch (1/3)

- Idea: A functional perspective from the noncommutative L_α space

For an operator-valued function $f: \Omega \to S_\infty(\mathcal{H})$, we define

$$\|f\|_{L_\alpha(\Omega,\, S_\alpha(\mathcal{H}))} := \left(\int_\Omega \|f(\omega)\|^\alpha_{S_\alpha(\mathcal{H})}\, \mathrm{d}p(\omega)\right)^{1/\alpha}.$$

For every $m \in [1{:}M]$, $\pi_m: L_\infty(\Omega, p) \to L_\infty\left(\Omega^M, p\right)$

$$\pi_m(f)(\omega_1, \dots, \omega_M) := f(\omega_m)$$

$$\mathrm{E}: L_\infty(\Omega, p) \to L_\infty\left(\Omega^M, p\right)$$

$$\mathrm{E}(f)(\omega_1, \dots, \omega_M) := \mathbb{E}_p[f]$$

$$\Theta(f) := \frac{1}{M}\sum_{m=1}^{M} \pi_m(f) - \mathrm{E}(f)$$

Proof Sketch (2/3)

- C-Q channel with random codes: $\Omega \ni \omega \mapsto \rho_B^{x(\omega)}$

$$\mathbb{E}_{\Omega^M} \left\| \frac{1}{M} \sum_{m=1}^{M} \rho_B^{x(\omega_m)} - \rho_B \right\|_1 = \left\| \Theta\left(\rho_B^{x(\cdot)}\right) \right\|_{L_1\left(\Omega^M, S_1(\mathcal{H})\right)}$$

$$\leq \left\| \Theta\left(\sigma_B^{\frac{1-a}{2\alpha}} \rho_B^{x(\cdot)} \sigma_B^{\frac{1-a}{2\alpha}}\right) \right\|_{L_\alpha\left(\Omega^M, S_\alpha(\mathcal{H})\right)}$$

Holder's inequality

$$\leq \left\| \Theta\colon L_\alpha\left(\Omega, S_\alpha(\mathcal{H})\right) \to L_\alpha\left(\Omega^M, S_\alpha(\mathcal{H})\right) \right\|$$

$$\cdot \left\| \sigma_B^{\frac{1-a}{2\alpha}} \rho_B^{x(\cdot)} \sigma_B^{\frac{1-a}{2\alpha}} \right\|_{L_\alpha\left(\Omega, S_\alpha(\mathcal{H})\right)}$$

$$\to \mathrm{e}^{\frac{\alpha-1}{\alpha} I_\alpha^*(X:B)_\rho}$$

Proof Sketch (2/3)

1. $(\alpha = 1)$ By triangle inequality:

$$\Rightarrow \left\|\Theta\colon L_1\left(\Omega, S_1(\mathcal{H})\right) \to L_1\left(\Omega^M, S_1(\mathcal{H})\right)\right\| \leq 2$$

2. $(\alpha = 2)$ Pairwise independence in $L_2\left(\Omega^M, S_2(\mathcal{H})\right)$:

$$\Rightarrow \left\|\Theta\colon L_2\left(\Omega, S_2(\mathcal{H})\right) \to L_2\left(\Omega^M, S_2(\mathcal{H})\right)\right\| \leq \frac{1}{\sqrt{M}}$$

3. $(\alpha \in [1,2])$ **Norm interpolation**:

$$\Rightarrow \left\|\Theta\colon L_\alpha\left(\Omega, S_\alpha(\mathcal{H})\right) \to L_\alpha\left(\Omega^M, S_\alpha(\mathcal{H})\right)\right\| \leq 2^{\frac{2}{\alpha}-1} M^{\frac{1-\alpha}{\alpha}}$$

$\square$

Conclusions

Conclusions

- One-shot **achievability** and **strong converse** for quantum soft covering are obtained using IID & CC random codebooks.
- Technique: **norm interpolation**.
  - Error exponent of privacy amplification & decoupling [Dupuis'21]
- Applications: wiretap channels & covering-type problems.
- Results also hold for **infinite-dimensional** Hilbert space.

arXiv:2202.10995