Non asymptotic and moderate deviation results in quantum hypothesis testing

arXiv:1612.01464v2

Cambyse Rouzé (University of Cambridge)
Joint with Nilanjana Datta

Beyond I.I.D. in Information Theory
24-28 July 2017, Singapore

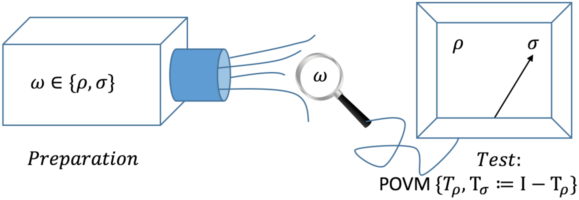
Overview

1. Quantum hypothesis testing in the i.i.d. setting

2. Quantum hypothesis testing beyond i.i.d.

3. Application to the determination of the capacity of classical-quantum channels

Framework



- Let $\mathcal{H}$ finite dimensional Hilbert space, $\rho, \sigma \in \mathcal{D}_+(\mathcal{H})$, set of states on $\mathcal{H}$

  ρ (null hypothesis) *and* σ (alternative hypothesis).
- A test is a POVM $\{T, \mathbb{I} - T\}$, $T \in \mathcal{B}(\mathcal{H})$, $0 \leq T \leq \mathbb{I}$
- Possible errors in inference with associated error probabilities:

$$\alpha(T) = \mathrm{Tr}(\rho(\mathbb{I} - T)) \qquad \text{type I error}$$
$$\beta(T) = \mathrm{Tr}(\sigma T) \qquad \text{type II error}$$

- Trade-off between these two errors, various ways of optimizing them.
- In the asymmetric case, minimize the type II error while controlling the type I error:

$$\beta(\epsilon) = \inf_{0 \leq T \leq 1} \{\beta(T) | \, \alpha(T) \leq \epsilon\}.$$

Size-dependent quantum hypothesis testing

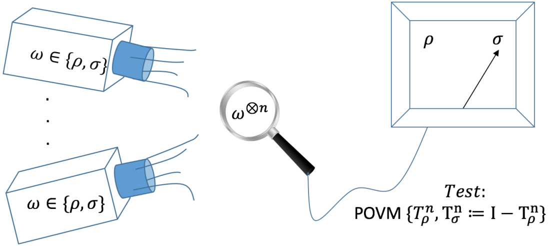
- For states $\rho_n, \sigma_n \in \mathcal{D}_+(\mathcal{H}^{\otimes n})$, define the *type II error exponent:*

$$R_n := -\frac{1}{n} \log \beta(T_n), \qquad \Leftrightarrow \qquad \beta(T_n) = e^{-nR_n},$$

  as well as the *type I threshold*

$$\varepsilon_n \text{ such that } \alpha(T_n) \leq \varepsilon_n.$$

- Example: the i.i.d. scenario:

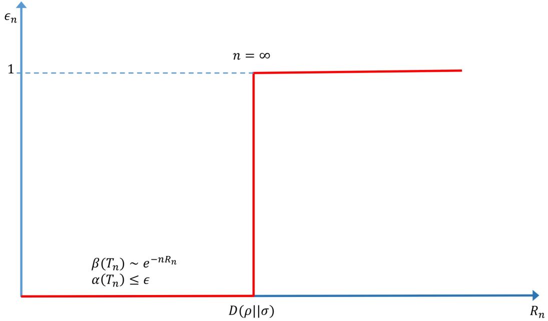


Quantum Stein's lemma

Theorem (Hiai&Petz91, Ogawa&Nagaoka00)

$$\max_{\alpha(T_n)\leq\varepsilon} R_n \underset{n\to\infty}{\to} D(\rho\|\sigma), \quad \forall\epsilon\in(0,1),$$

$$D(\rho\|\sigma) = \mathrm{Tr}(\rho(\log\rho - \log\sigma)) \quad \textit{Umegaki's quantum relative entropy.}$$

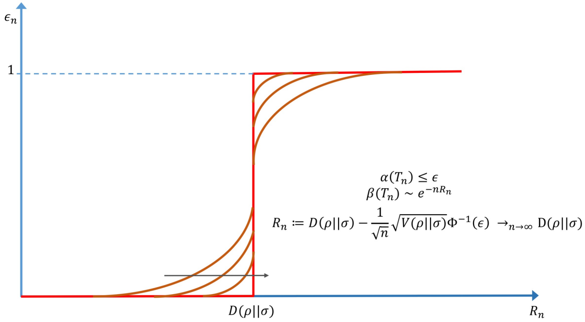


Small deviations and second order asymptotics

Theorem (Tomamichel&Hayashi13, Li14)

$$\max_{\alpha(T_n)\leq\varepsilon} R_n = D(\rho\|\sigma) + \frac{1}{\sqrt{n}}s_1(\epsilon) + \mathcal{O}\left(\frac{\log n}{n}\right), \quad s_1(\epsilon) := \sqrt{V(\rho\|\sigma)}\Phi^{-1}(\epsilon),$$

$$V(\rho\|\sigma) = \mathrm{Tr}(\rho(\log\rho - \log\sigma)^2) - D(\rho\|\sigma)^2 \quad \textit{quantum information variance}, \quad \Phi \sim \mathcal{N}(0,1).$$

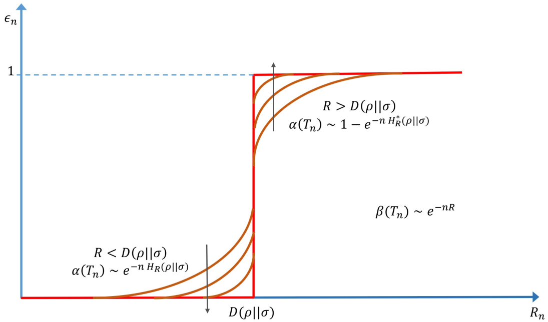


Large deviations and the Hoeffding bound

Theorem (Hayashi06, Nagaoka06, Mosonyi&Ogawa14)

$$R < D(\rho\|\sigma) : -\frac{1}{n}\log\left\{\min_{\beta(T_n)\leq e^{-nR}} \alpha(T_n)\right\} \underset{n\to\infty}{\to} \sup_{0<\alpha<1} \frac{\alpha-1}{\alpha}[R - D_\alpha(\rho\|\sigma)]$$

$$R > D(\rho\|\sigma) : -\frac{1}{n}\log\left\{\max_{\beta(T_n)\leq e^{-nR}} (1-\alpha(T_n))\right\} \underset{n\to\infty}{\to} \sup_{1<\alpha} \frac{\alpha-1}{\alpha}[R - D^*_\alpha(\rho\|\sigma)]$$

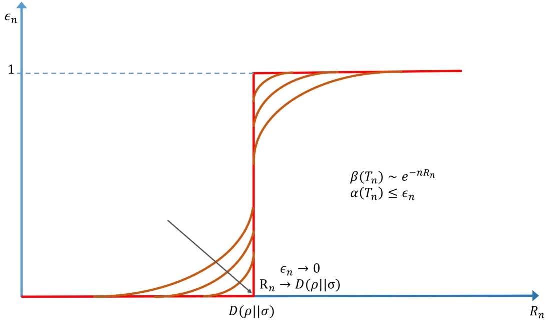


Moderate deviations

Theorem (Chubb&Tan&Tomamichel17, Chen&Hsieh17)

Moderate sequence $\{a_n\}_{n\in\mathbb{N}}$: $a_n \to 0$, $\sqrt{n}a_n \to \infty$; $\varepsilon_n = e^{-na_n^2}$:

$$\max_{\alpha(T_n)\leq\varepsilon_n} R_n = D(\rho\|\sigma) - \sqrt{2V(\rho\|\sigma)}a_n + o(a_n),$$



Finite sample size bounds

Asymptotic regime: $\{\rho^{\otimes n}, \sigma^{\otimes n}\}, \; n \to \infty$	$\rightsquigarrow$	Finite sample size: $\{\rho^{\otimes n}, \sigma^{\otimes n}\}, \; n < \infty$

Theorem (Audenaert&Mosonyi&Verstraete12)

$$D(\rho\|\sigma) - \frac{g(\epsilon)}{\sqrt{n}} \leq \max_{\alpha(T_n)\leq\varepsilon} R_n \leq D(\rho\|\sigma) + \frac{f(\epsilon)}{\sqrt{n}}$$

$$f(\epsilon) \propto \log(1-\epsilon)^{-1}, \qquad g(\epsilon) \propto \log \epsilon^{-1},$$

What is known

	i.i.d.	Beyond i.i.d.
Stein's lemma	$\alpha(T_n) \leq \varepsilon$ $\beta(T_n) \sim e^{-nD(\rho\|\sigma)}$	✓[1]
Small deviations	$\alpha(T_n) \leq \varepsilon$ $\beta(T_n) \sim e^{-nD(\rho\|\sigma)-\sqrt{n}\, s_1(\varepsilon)}$	✓[2]
Large deviations $(R < D(\rho\|\sigma))$	$\beta(T_n) \leq e^{-nR}$ $\alpha(T_n) \sim e^{-nH_R(\rho\|\sigma)}$	✓[3]
Moderate deviations	$\alpha(T_n) \leq e^{-na_n^2}$ $\beta(T_n) \sim e^{-nD(\rho\|\sigma)+\sqrt{2V(\rho\|\sigma)}a_n\sqrt{n}}$	?
Finite sample size	$\alpha(T_n) \leq \varepsilon$ $\beta(T_n) \leq e^{-nD(\rho\|\sigma)+\sqrt{n}g(\epsilon)}$	?

[1] Bjelakovic&Siegmund-Schultze04, Bjelakovic&Deuschel&Krueger&Seiler&Siegmund-Schultze&Szkola08, Hiai&Mosonyi&Ogawa08, Mosonyi&Hiai&Ogawa&Fannes08, Brandao&Plenio10, Jaksic&Ogata&Pillet&Seiringer12, etc.

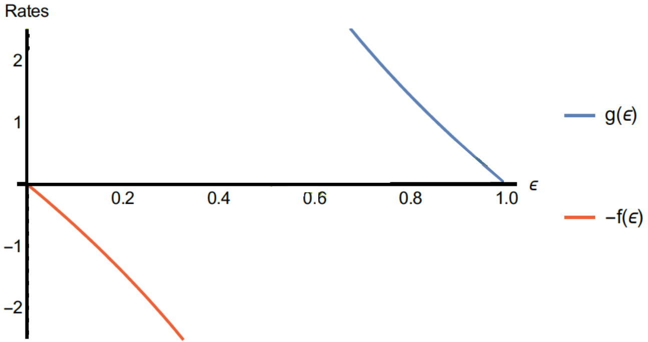
[2] Datta&Pautrat&R16

[3] Mosonyi&Hiai&Ogawa&Fannes08, Jaksic&Ogata&Pillet&Seiringer12

A tighter lower bound on the error exponent for finite sample size

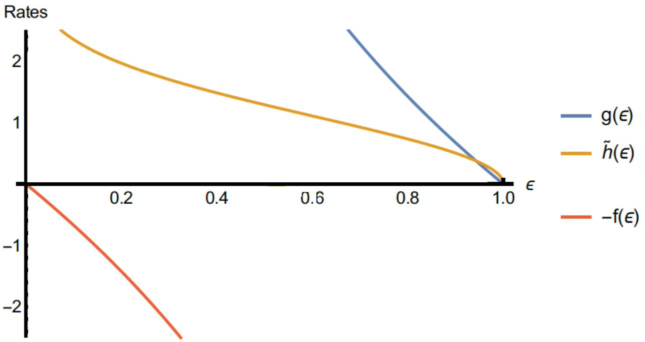
Theorem (Audenaert&Mosonyi&Verstraete12)

$$D(\rho\|\sigma) - \frac{g(\epsilon)}{\sqrt{n}} \le \max_{\alpha(T_n)\le\varepsilon} R_n, \qquad g(\epsilon) \propto \log \epsilon^{-1}.$$



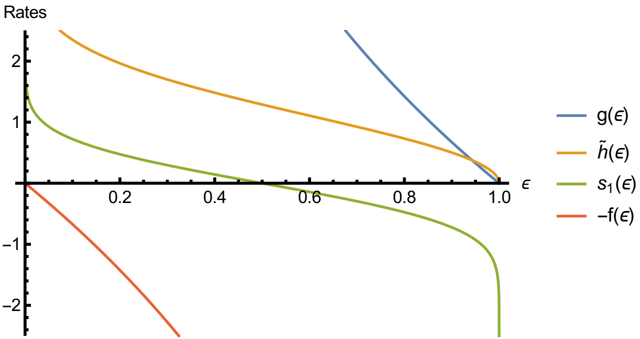
Theorem (R&Datta16)

$$D(\rho\|\sigma) - \frac{1}{\sqrt{n}}\tilde{h}(\varepsilon) \leq \max_{\alpha(T_n)\leq\varepsilon} R_n, \quad \tilde{h}(\varepsilon) \propto \sqrt{\log \varepsilon^{-1}}.$$



Theorem (R&Datta16)

$$D(\rho\|\sigma) - \frac{1}{\sqrt{n}}\tilde{h}(\varepsilon) \leq \max_{\alpha(T_n)\leq\varepsilon} R_n, \quad \tilde{h}(\varepsilon) \propto \sqrt{\log \varepsilon^{-1}}.$$



$$\max_{\alpha(T_n)\leq\varepsilon} R_n = D(\rho\|\sigma) + \frac{1}{\sqrt{n}} s_1(\epsilon) + \mathcal{O}\left(\frac{\log n}{n}\right)$$

Key tool 1: martingale concentration inequalities

- Concentration inequalities provide upper bounds on tail probabilities of the type

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq r).$$

- Different methods exist to derive concentration inequalities (martingales, logarithmic Sobolev inequalities, transportation cost, Talagrand's induction method, etc...)
- A *martingale* is a sequence $\{X_n\}_{n\in\mathbb{N}}$ of integrable random variables for which:

$$\mathbb{E}[X_{n+1}|X_1, \ldots, X_n] = X_n \quad \text{almost surely.}$$

- Example: sum of independent integrable centered random variables $S_n := \sum_{k=1}^n Y_k$:

$$\mathbb{E}[S_{n+1}|Y_1, \ldots, Y_n] = \sum_{k=1}^{n+1} \mathbb{E}[Y_k|Y_1, \ldots, Y_n] = \sum_{k=1}^{n} \mathbb{E}[Y_k|Y_1, \ldots, Y_n] + \mathbb{E}[Y_{n+1}|Y_1, \ldots, Y_n]$$
$$= \sum_{k=1}^{n} Y_k + \mathbb{E}[Y_{n+1}] = S_n + 0 = S_n,$$

- Azuma-Hoeffding's martingale concentration inequality: if $\forall k \in \{1, \ldots, n\}$, $|X_k - X_{k-1}| \leq d$, then:

$$\boxed{\mathbb{P}(X_n - X_0 \geq r) \leq \exp\left(-\frac{r^2}{2nd^2}\right).}$$

Key tool 2: the relative modular operator

$$\Delta_{\rho|\sigma} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}), \qquad A \mapsto \rho A \sigma^{-1}.$$

- Noncommutative generalization of the Radon Nikodym derivative between two measures.
- $\Delta_{\rho|\sigma}$ is a positive operator, admits a spectral decomposition: given

$$\rho := \sum_{\lambda \in \mathrm{sp}\, \rho} \lambda P_\lambda(\rho), \qquad \sigma := \sum_{\mu \in \mathrm{sp}\, \sigma} \mu P_\mu(\sigma),$$

$$\log \Delta_{\rho|\sigma} := \sum_{x \in \mathcal{X}} x \Pi_x(\log \Delta_{\rho|\sigma}), \qquad \Pi_x(\log \Delta_{\rho|\sigma}) := \sum_{\lambda \in \mathrm{sp}(\rho), \mu \in \mathrm{sp}(\sigma), \log(\lambda/\mu) = x} L_{P_\lambda(\rho)} R_{P_\mu(\sigma)},$$

  where $\mathcal{X}$ denotes the set of real numbers $\log(\lambda/\mu)$, $L_A(B) = AB$, $R_A(B) = BA$.
- The spectral theorem allows us to go from the quantum setting to classical probability theory: given a Hilbert space $\mathcal{V}$, a pure state $\Psi \in \mathcal{V}$ and a self adjoint operator $O = \sum_{x \in \mathrm{sp}(O)} x \Pi_x$ on $\mathcal{V}$,

$$\mathbb{P}(X \in A) := \langle \Psi, \Pi_A(O) \Psi \rangle, \qquad \Pi_A := \sum_{x \in A \subset \mathrm{sp}(O)} \Pi_x$$

  defines a random variable on $\mathrm{sp}(O)$, so that $\forall f : \mathrm{sp}(O) \to \mathbb{R}$:

$$\mathbb{E}[f(X)] = \langle \Psi, f(O) \Psi \rangle$$

From error exponents to concentration

Lemma (Li14, Datta&Pautrat&R16)

For all $L > 0$ there exists a test T such that

$$\mathrm{Tr}(\rho^{\otimes n}(\mathbf{1} - T)) \leq \langle (\rho^{\otimes n})^{1/2}, \Pi_{(-\infty, \log L]}(\log \Delta_{\rho^{\otimes n}|\sigma^{\otimes n}})((\rho^{\otimes n})^{1/2})\rangle \quad \textit{and} \quad \mathrm{Tr}(\sigma^{\otimes n} T) \leq L^{-1}. \tag{1}$$

From error exponents to concentration

Lemma (Li14, Datta&Pautrat&R16)

For all $L > 0$ there exists a test T such that

$$\mathrm{Tr}(\rho^{\otimes n}(\mathbf{1} - T)) \le \langle (\rho^{\otimes n})^{1/2}, \Pi_{(-\infty, \log L]}(\log \Delta_{\rho^{\otimes n}|\sigma^{\otimes n}})((\rho^{\otimes n})^{1/2})\rangle \quad \textit{and} \quad \mathrm{Tr}(\sigma^{\otimes n} T) \le L^{-1}. \tag{1}$$

- Via spectral theorem, define the random variable X_n of law ($\Omega_\tau = \tau^{1/2}$)

$$\mathbb{P}(X_n = x) := \langle \Omega_{\rho^{\otimes n}}, \Pi_x(\log \Delta_{\rho^{\otimes n}|\sigma^{\otimes n}})(\Omega_{\rho^{\otimes n}})\rangle$$

i.e. $\forall f : \mathrm{sp}(\log(\Delta_{\rho^{\otimes n}|\sigma^{\otimes n}})) \to \mathbb{R}$

$$\langle \Omega_{\rho^{\otimes n}}, f(\log \Delta_{\rho^{\otimes n}|\sigma^{\otimes n}})(\Omega_{\rho^{\otimes n}})\rangle = \sum_{x \in \mathcal{X}} f(x) \langle \Omega_{\rho^{\otimes n}}, \Pi_x(\log \Delta_{\rho^{\otimes n}|\sigma^{\otimes n}})(\Omega_{\rho^{\otimes n}})\rangle \equiv \mathbb{E}[f(X_n)].$$

From error exponents to concentration

Lemma (Li14, Datta&Pautrat&R16)

For all $L > 0$ there exists a test T such that

$$\mathrm{Tr}(\rho^{\otimes n}(\mathbf{1} - T)) \leq \langle (\rho^{\otimes n})^{1/2}, \Pi_{(-\infty, \log L]}(\log \Delta_{\rho^{\otimes n}|\sigma^{\otimes n}})((\rho^{\otimes n})^{1/2}) \rangle \quad \text{and} \quad \mathrm{Tr}(\sigma^{\otimes n} T) \leq L^{-1}. \tag{1}$$

- Via spectral theorem, define the random variable X_n of law $(\Omega_\tau = \tau^{1/2})$

$$\mathbb{P}(X_n = x) := \langle \Omega_{\rho^{\otimes n}}, \Pi_x(\log \Delta_{\rho^{\otimes n}|\sigma^{\otimes n}})(\Omega_{\rho^{\otimes n}}) \rangle$$

  i.e. $\forall f : \mathrm{sp}(\log(\Delta_{\rho^{\otimes n}|\sigma^{\otimes n}})) \to \mathbb{R}$

$$\langle \Omega_{\rho^{\otimes n}}, f(\log \Delta_{\rho^{\otimes n}|\sigma^{\otimes n}})(\Omega_{\rho^{\otimes n}}) \rangle = \sum_{x \in \mathcal{X}} f(x) \langle \Omega_{\rho^{\otimes n}}, \Pi_x(\log \Delta_{\rho^{\otimes n}|\sigma^{\otimes n}})(\Omega_{\rho^{\otimes n}}) \rangle \equiv \mathbb{E}[f(X_n)].$$

- X_n has characteristic function

$$\mathbb{E}\left[e^{iuX_n}\right] = \langle \Omega_{\rho^{\otimes n}}, e^{iu \log \Delta_{\rho^{\otimes n}|\sigma^{\otimes n}}}(\Omega_{\rho^{\otimes n}}) \rangle = \prod_{k=1}^{n} \langle \Omega_\rho, e^{iu \log \Delta_{\rho|\sigma}}(\Omega_\rho) \rangle = \mathbb{E}\left[e^{iu \sum_{k=1}^n \tilde{X}_k}\right],$$

  where $(\tilde{X}_k)$ i.i.d. random variables so that $\mathbb{E}[\tilde{X}_k] = D(\rho \| \sigma)$.

From error exponents to concentration

Lemma (Li14, Datta&Pautrat&R16)

For all $L > 0$ there exists a test T such that

$$\mathrm{Tr}(\rho^{\otimes n}(\mathbf{1} - T)) \leq \langle (\rho^{\otimes n})^{1/2}, \Pi_{(-\infty, \log L]}(\log \Delta_{\rho^{\otimes n}|\sigma^{\otimes n}})((\rho^{\otimes n})^{1/2}) \rangle \quad \textit{and} \quad \mathrm{Tr}(\sigma^{\otimes n} T) \leq L^{-1}. \tag{1}$$

- Via spectral theorem, define the random variable X_n of law $(\Omega_\tau = \tau^{1/2})$

$$\mathbb{P}(X_n = x) := \langle \Omega_{\rho^{\otimes n}}, \Pi_x(\log \Delta_{\rho^{\otimes n}|\sigma^{\otimes n}})(\Omega_{\rho^{\otimes n}}) \rangle$$

i.e. $\forall f : \mathrm{sp}(\log(\Delta_{\rho^{\otimes n}|\sigma^{\otimes n}})) \to \mathbb{R}$

$$\langle \Omega_{\rho^{\otimes n}}, f(\log \Delta_{\rho^{\otimes n}|\sigma^{\otimes n}})(\Omega_{\rho^{\otimes n}}) \rangle = \sum_{x \in \mathcal{X}} f(x) \langle \Omega_{\rho^{\otimes n}}, \Pi_x(\log \Delta_{\rho^{\otimes n}|\sigma^{\otimes n}})(\Omega_{\rho^{\otimes n}}) \rangle \equiv \mathbb{E}[f(X_n)].$$

- X_n has characteristic function

$$\mathbb{E}\left[\mathrm{e}^{iuX_n}\right] = \langle \Omega_{\rho^{\otimes n}}, \mathrm{e}^{iu \log \Delta_{\rho^{\otimes n}|\sigma^{\otimes n}}}(\Omega_{\rho^{\otimes n}}) \rangle = \prod_{k=1}^{n} \langle \Omega_\rho, \mathrm{e}^{iu \log \Delta_{\rho|\sigma}}(\Omega_\rho) \rangle = \mathbb{E}\left[\mathrm{e}^{iu \sum_{k=1}^n \tilde{X}_k}\right],$$

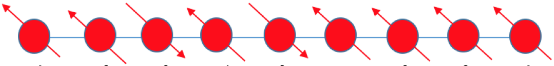
where $(\tilde{X}_k)$ i.i.d. random variables so that $\mathbb{E}[\tilde{X}_k] = D(\rho\|\sigma)$.

- $\{X_k - kD(\rho\|\sigma)\}_{1 \leq k \leq n}$ is a centered martingale, hence by Azuma-Hoeffding inequality:

$$\langle \Omega_{\rho^{\otimes n}}, \Pi_{(-\infty, -r]}(\log \Delta_{\rho^{\otimes n}|\sigma^{\otimes n}} - nD(\rho\|\sigma)\,\mathrm{id})(\Omega_{\rho^{\otimes n}}) \rangle = \mathbb{P}(X_n - nD(\rho\|\sigma) \leq -r) \leq \mathrm{e}^{-\frac{r^2}{2nd^2}},$$

$$d := \| \log \Delta_{\rho|\sigma} - D(\rho\|\sigma)\,\mathrm{id} \|_\infty .$$

Example 1: quantum Gibbs states on spin chains



- $\mathcal{H}_n := \mathcal{H}^{\otimes n}$, $\mathcal{H}_X := \bigotimes_{i \in X \subseteq \mathbb{Z}} \mathcal{H}$
- Translation invariant interaction:

$$\Phi : \mathbb{Z} \supseteq X \mapsto \Phi_X \in \mathcal{B}_{sa}(\mathcal{H}_X).$$

- Finite range: $\exists R > 0 : |X| > R \Rightarrow \Phi_X = 0$.
- Local Gibbs state: $H^{\Phi}_{[n]} := \sum_{Y \subset [n]} \Phi_Y$, $\quad \rho_n := \frac{e^{-\beta H^{\Phi}_{[n]}}}{\mathrm{Tr}(e^{-\beta H^{\Phi}_{[n]}})}$.
- Such states satisfy the *factorization property*: $\exists\ R > 0$ such that (Hiai&Mosonyi&Ogawa08):

$$\rho_n \leq R\ \rho_{n-1} \otimes \rho_1 \qquad \textit{upper factorization},$$

$$\rho_n \geq R^{-1} \rho_{n-1} \otimes \rho_1 \qquad \textit{lower factorization}.$$

Example 2: finitely correlated states

- Defined by Fannes&Nachtergaele&Werner92
- Generalizes the notion of a Markov chain to the quantum regime.
- $\mathcal{E} : \mathcal{B}(\mathcal{K}) \to \mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathcal{K})$ CPTP map
- ρ full-rank state on $\mathcal{B}(\mathcal{K})$ such that $\mathrm{Tr}_{\mathcal{H}}\, \mathcal{E}(\rho) = \rho$.

$$\rho_n := \mathrm{Tr}_{\mathcal{K}}(\mathrm{id}_{\mathcal{B}(\mathcal{H})}^{\otimes n-1} \otimes \mathcal{E}) \circ \ \dots \ \circ (\mathrm{id}_{\mathcal{B}(\mathcal{H})} \otimes \mathcal{E}) \circ \mathcal{E}(\rho) \in \mathcal{D}_+(\mathcal{H}^{\otimes n}).$$

- ω state whose restriction to each $\mathcal{H}^{\otimes n}$ is ρ_n is a finitely correlated state.
- Ex: ground state of AKLT Hamiltonian on spin 1 chain (Affleck&Kennedy&Lieb&Tasaki87).
- Finitely correlated states satisfy the upper factorization property (Hiai&Mosonyi&Ogawa08)

$$\exists R > 1 : \quad \rho_n \leq R\ \rho_{n-1} \otimes \rho_1 .$$

Finite sample size bounds beyond i.i.d.: correlated states

Definition (upper factorization)

$\{\rho_n\}_{n\in\mathbb{N}}$ satisfies the *upper factorization property* if $\exists R > 0$ such that $\forall n \geq 1$,

$$\rho_n \leq R\ \rho_{n-1} \otimes \rho_1 \qquad \textit{upper factorization},$$

Theorem (R&Datta17)

If $\{\rho_n\}_{n\in\mathbb{N}}$ and $\{\sigma_n\}_{n\in\mathbb{N}}$ satisfy the upper factorization property, with $R \geq 1$, for any $0 \leq \varepsilon \leq 1$,

$$\max_{\alpha(T_n)\leq\varepsilon} R_n \geq \begin{cases} D(\rho_1\|\sigma_1) - c\sqrt{\dfrac{2\log(R^n\varepsilon^{-1})}{n}} & \textit{if } \varepsilon \geq R^n \mathrm{e}^{-nc^2/2}, \\ D(\rho_1\|\sigma_1) - c^2/2 - \dfrac{1}{n}\log(R^n\varepsilon^{-1}) & \textit{else}, \end{cases}$$

where $c := \|\log \Delta_{\rho_1\|\sigma_1} - D(\rho_1\|\sigma_1)\,\mathrm{id}\,\|_\infty$.

- The proof consists of an adaptation of the proof of Azuma-Hoeffding's inequality.

Moderate deviations beyond i.i.d.: correlated states

Definition (factorization property)

$\exists R > 0$ such that

$$\rho_n \leq R\ \rho_{n-1} \otimes \rho_1 \qquad \text{(upper factorization)},$$
$$\rho_n \geq R^{-1}\ \rho_{n-1} \otimes \rho_1 \qquad \text{(lower factorization)}.$$

Theorem (R&Datta17)

Let $\rho_n, \sigma_n \in \mathcal{D}_+(\mathcal{H}^{\otimes n})$ *satisfying the upper factorization property with*

$$0 \leq \log R < (4 - \mathrm{e}^c)(\mathrm{e}^c - 1)^2 V(\rho_1 \| \sigma_1)/(6c^2),$$

where $c := \| \log \Delta_{\rho_1 \| \sigma_1} - D(\rho_1 \| \sigma_1)\, \mathrm{id}\, \|_\infty < \log 4$. *Let* $\{a_n\}_{n\in\mathbb{N}}$ *moderate,* $\varepsilon_n := \mathrm{e}^{-na_n^2}$. *Then, for n large enough:*

$$\max_{\alpha(T_n)\leq\varepsilon_n} R_n \geq D(\rho_1\|\sigma_1) - \sqrt{2V(\rho_1\|\sigma_1)} \left(\frac{3}{4 - e^c} (\log R + a_n^2) \right)^{1/2}$$

If $\{\rho_n\}_{n\in\mathbb{N}}$ *and* $\{\sigma_n\}_{n\in\mathbb{N}}$ *satisfy the lower factorization property with* $R > 1$, *then*

$$\max_{\alpha(T_n)\leq\varepsilon_n} R_n \leq D(\rho_1\|\sigma_1) - \sqrt{2V(\rho_1\|\sigma_1)}\, a_n + o(a_n).$$

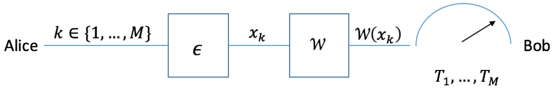
- The proof can be rearranged to get back the results of Chubb et al and Chen et al in the i.i.d. scenario.

Information transmission via memoryless c-q channels

- Assume Alice wants to communicate with Bob using a c-q channel:

$$\mathcal{W} : \mathcal{X} \to \mathcal{D}(\mathcal{H})$$

- Alice encodes message $k \in \mathcal{M} := \{1, \ldots, M\}$ into codeword $x_k \in \mathcal{X}$, into state $\mathcal{W}(x_k)$.
- Bob performs a measurement modeled by a POVM $T = \{T_1, \ldots, T_M\}$.



- The average probability of success is given by

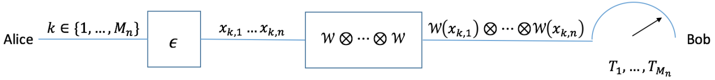
$$\mathbb{P}(\text{success}|\mathcal{W}, T) = \frac{1}{M} \sum_{k=1}^{M} \mathrm{Tr}(\mathcal{W}(x_k) T_k).$$

- The *one-shot ε-error capacity* of $\mathcal{W}$ is given by

$$C(\mathcal{W}, \varepsilon) := \log M^*(\mathcal{W}, \varepsilon),$$
$$M^*(\mathcal{W}, \varepsilon) := \max\{M \in \mathbb{N} | \; \exists \text{ POVM } T : \; \mathbb{P}(\text{success}|\mathcal{W}, T) \geq 1 - \varepsilon\}.$$

Finite sample size analysis of memoryless c-q channels

- n independent uses of $\mathcal{W}$: Alice encodes message $k \in \{1, \dots, M_n\} \mapsto (x_{k,1}, \dots, x_{k,n}) \in \mathcal{X}^n$.



- Holevo-Schumacher-Westmoreland theorem: for all $0 < \varepsilon < 1$,

$$\frac{C(\mathcal{W}^{\otimes n}, \varepsilon)}{n} \to \chi^*(\mathcal{W}) := \sup_{p_\mathcal{X}} \sum_{x \in \mathcal{X}} p_\mathcal{X}(x) D\left(\mathcal{W}(x) \| \sum_{y \in \mathcal{X}} p_\mathcal{X}(y)\mathcal{W}(y)\right)$$

- Second order & moderate deviation: Tan&Tomamichel15, Chubb&Tan&Tomamichel17, Cheng&Hsieh17, using Wang&Renner12 to relate c-q capacity to type II error exponent.

Proposition (Finite sample size analysis, R&Datta17)

For any $\varepsilon \in (0,1)$, any $p_\mathcal{X}$ achieving the Holevo capacity, and any $\varepsilon' \in (0, \varepsilon)$:

$$\frac{C(\mathcal{W}^{\otimes n}, \varepsilon)}{n} \geq \chi^*(\mathcal{W}) - \sqrt{\frac{2 \log \varepsilon'^{-1}}{n}} c_{p_\mathcal{X}} - \frac{1}{n} \log \frac{4\varepsilon}{\varepsilon - \varepsilon'},$$

where $c_{p_\mathcal{X}} := \| \log \Delta_{\rho_{p_\mathcal{X}} | \sigma_{p_\mathcal{X}}} - D(\rho_{p_\mathcal{X}} \| \sigma_{p_\mathcal{X}}) \,\mathrm{id} \|_\infty$, *with*

$$\rho_{p_\mathcal{X}} = \sum_{x \in \mathcal{X}} p_\mathcal{X}(x) |x\rangle\langle x| \otimes \mathcal{W}(x), \quad \sigma_{p_\mathcal{X}} = \sum_{x \in \mathcal{X}} p_\mathcal{X}(x) |x\rangle\langle x| \otimes \sum_y p_\mathcal{X}(y) \mathcal{W}(y)$$

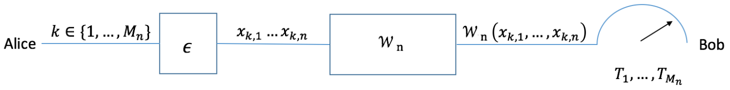
.

c-q channels with memory

- A sequence of c-q channels $\mathcal{W}_n : \mathcal{X}^n \to \mathcal{D}(\mathcal{H}^{\otimes n})$ satisfies the *channel upper factorization property* if $\exists R > 0$:

$$\forall (x_1, \dots, x_n) \in \mathcal{X}^n : \quad \mathcal{W}_n(x_1, \dots, x_n) \leq R\, \mathcal{W}_{n-1}(x_1, \dots, x_{n-1}) \otimes \mathcal{W}_1(x_n).$$

- Assume the following communication scheme:



Proposition (Finite sample size analysis)

If $R \geq 1$, then for any $\varepsilon \in (0,1)$, any $p_{\mathcal{X}}$ achieving $\chi^*(\mathcal{W}_1)$, and any $\varepsilon' \in (0, \varepsilon)$:

$$\frac{C(\mathcal{W}_n, \varepsilon)}{n} \geq \begin{cases} \chi^*(\mathcal{W}_1) - c_{p_{\mathcal{X}}} \sqrt{\frac{2 \log(R^n \varepsilon'^{-1})}{n}} - \frac{1}{n} \log \frac{4\varepsilon}{\varepsilon - \varepsilon'} & \text{for } R^n e^{-n c_{p_{\mathcal{X}}}^2 /2} \leq \varepsilon' < \varepsilon, \\ \chi^*(\mathcal{W}_1) - \frac{c_{p_{\mathcal{X}}}^2}{2} - \frac{1}{n} \log R^n \varepsilon'^{-1} - \log \frac{4\varepsilon}{\varepsilon - \varepsilon'} & \text{else.} \end{cases}$$

- Similarly, a moderate deviation analysis can be carried out.
- Open question: find a c-q channel with memory satisfying the factorization property.

Summary

- Martingale concentration inequalities can be used to find a finite sample size lower bound for the optimal type II error exponent in the i.i.d. case and beyond. In the i.i.d. case, this bound is tighter than the previous bound of Audenaert&Mosonyi&Verstraete12.
- Similar techniques can be used to derive moderate deviation bounds for sequences of states satisfying a factorization property.
- The one-shot bound of Wang&Renner12 can be applied to find finite sample size and moderate deviation results for c-q channels.
- An example of a quantum channel with memory satisfying the channel upper factorization property is lacking at the moment.

- Thank you for your attention!

References

- Audenaert&Mosonyi&Verstraete12: Quantum state discrimination bounds for finite sample size. Journal of Mathematical Physics, Vol. 53, 122205 (2012).
- Bjelakovic&Siegmund-Schultze04: An ergodic theorem for quantum relative entropy. Comm. Math. Phys. 247, 697 (2004)
- Bjelakovic&Deuschel,&Krueger&Seiler&Siegmund-Schultze&Szkola08: Typical support and Sanov large deviations of correlated states. Comm. Math. Phys. 279, 559 (2008).
- Brandão&Plenio: A Generalization of Quantum Stein's Lemma. Commun. Math. Phys. 295: 791 (2010).
- Datta&Pautrat&Rouze16: Second-order asymptotics for quantum hypothesis testing in settings beyond i.i.d.—quantum lattice systems and more. Journal of Mathematical Physics, 57, 6, 062207 (2016)
- Fannes&Nachtergaele&Werner92: Finitely correlated states on quantum spin chains. Commun. Math. Physics, 144, 443 490 (1992)
- Hayashi07: Error exponent in asymmetric quantum hypothesis testing and its application to classical-quantum channel Phys. Rev. A 76, 062301
- Hiai&Mosonyi&Ogawa08: Large deviations and Chernoff bound for certain correlated states on the spin chain. J. Math. Phys.
- Hiai&Mosonyi&Ogawa08: Error exponents in hypothesis testing for correlated states on a spin chain. J. Math. Phys. 49, 032112 (2008).
- Hiai&Petz91: The proper formula for the relative entropy an its asymptotics in quantum probability. Comm. Math. Phys. 143, 99 (1991).

- Jaksic&Ogata&Piller&Seiringer12: Quantum hypothesis testing and non-equilibrium statistical mechanics. Rev. Math. Phys. 24, 1230002 (2012).
- Li14: Second-order asymptotics for quantum hypothesis testing. Ann. Statist. Volume 42, Number 1, 171-189 (2014).
- Mosonyi&Hiai&Ogawa&Fannes08: Asymptotic distinguishability measures for shiftinvariant quasi-free states of fermionic lattice systems. J. Math. Phys. 49, 032112 (2008).
- Nagaoka06: The converse part of the theorem for quantum Hoeffding bound. arXiv:quant-ph/0611289
- Ogawa&Nagaoka00: Strong Converse and Stein's Lemma in the Quantum Hypothesis Testing. IEEE Trans. Inf. Theo. 46, 2428 (2000).
- Sadghi&Moslehian14: Noncommutative martingale concentration inequalities. Illinois Journal of Mathematics, Volume 58, Number 2, 561-575 (2014).
- Sason11: Moderate deviations analysis of binary hypothesis testing. 2012 IEEE International Symposium on Information Theory Proceedings, Cambridge, MA, pp. 821-825 (2012).
- Tomamichel&Hayashi13: A Hierarchy of Information Quantities for Finite Block Length Analysis of Quantum Tasks. IEEE Transactions on Information Theory, vol. 59, no. 11, pp. 7693-7710 (2013).

Proof (1)

- Define $V_{n-1} : \mathcal{B}(\mathcal{H}^{\otimes n})({\rho_{n-1}}^{1/2} \otimes {\rho_1}^{1/2}) \to \mathcal{B}(\mathcal{H}^{\otimes n}){\rho_n}^{1/2}$ by
$$V_{n-1}(X_n({\rho_{n-1}}^{1/2} \otimes {\rho_1}^{1/2})) = R^{-1/2} X_n {\rho_n}^{1/2}, \qquad \forall X_n \in \mathcal{B}(\mathcal{H}^{\otimes n}).$$
- V_{n-1} is a contraction (upper factorization property). Moreover,
$$V_{n-1}^* \Delta_{\sigma_n|\rho_n} V_{n-1} \leq \Delta_{\sigma_{n-1}\otimes\sigma_1|\rho_{n-1}\otimes\rho_1} = \Delta_{\sigma_{n-1}|\rho_{n-1}} \otimes \Delta_{\sigma_1|\rho_1}.$$

$$\begin{aligned}
\langle {\rho_n}^{1/2}, \Delta^t_{\sigma_n|\rho_n}({\rho_n}^{1/2})\rangle &= R\ \langle(\rho_{n-1}^{1/2} \otimes {\rho_1}^{1/2}), V_{n-1}^* \Delta^t_{\sigma_n|\rho_n} V_{n-1}({\rho_{n-1}}^{1/2} \otimes {\rho_1}^{1/2})\rangle \\
&\leq R\ \langle(\rho_{n-1}^{1/2} \otimes {\rho_1}^{1/2}), (V_{n-1}^* \Delta_{\sigma_n|\rho_n} V_{n-1})^t(\rho_{n-1}^{1/2} \otimes {\rho_1}^{1/2})\rangle && \text{operator Jensen} \\
&\leq R\ \langle(\rho_{n-1}^{1/2} \otimes {\rho_1}^{1/2}), (\Delta_{\sigma_{n-1}|\rho_{n-1}} \otimes \Delta_{\sigma_1|\rho_1})^t(\rho_{n-1}^{1/2} \otimes {\rho_1}^{1/2})\rangle && t \in [0,1],\ x \mapsto x^t \\
&= R\ \langle \rho_{n-1}^{1/2}, \Delta^t_{\sigma_{n-1}|\rho_{n-1}}({\sigma_{n-1}}^{1/2})\rangle \quad \langle {\rho_1}^{1/2}, \Delta^t_{\sigma_1|\rho_1}({\rho_1}^{1/2})\rangle && \text{operator monotone} \\
&\leq R^n \mathrm{e}^{-ntD(\rho_1\|\sigma_1)}\ \langle \Omega_{\rho_1}, \mathrm{e}^{t(\log(\Delta_{\sigma_1|\rho_1})+D(\rho_1\|\sigma_1))}(\Omega_{\rho_1})\rangle^n && \text{iteration.}
\end{aligned}$$

- By convexity of the exponential function, with $c \equiv \|\log \Delta_{\sigma_1|\rho_1} + D(\rho_1\|\sigma_1)\ \mathrm{id}\,\|_\infty$:
$$\mathrm{e}^{ts} \leq \frac{1}{2c}\left(\mathrm{e}^{tc} - \mathrm{e}^{-tc}\right)s + \frac{1}{2}\left(\mathrm{e}^{tc} + \mathrm{e}^{-tc}\right) \qquad -c \leq s \leq c.$$

$$\begin{aligned}
&\langle \Omega_{\rho_1},\ \Delta^t_{\sigma_1|\rho_1}(\Omega_{\rho_1})\rangle \\
&\leq \mathrm{e}^{-tD(\rho_1\|\sigma_1)}\left[\frac{1}{2c}\left(\mathrm{e}^{tc} - \mathrm{e}^{-tc}\right)(\langle \Omega_{\rho_1}, \log \Delta_{\sigma_1|\rho_1}(\Omega_{\rho_1})\rangle + D(\rho_1\|\sigma_1)) + \frac{1}{2}\left(\mathrm{e}^{tc} + \mathrm{e}^{-tc}\right)\right] \\
&= \mathrm{e}^{-tD(\rho_1\|\sigma_1)}\frac{1}{2}\left(\mathrm{e}^{tc} + \mathrm{e}^{-tc}\right) \leq \mathrm{e}^{-tD(\rho_1\|\sigma_1)+t^2c^2/2}, \qquad \text{using } \frac{1}{2}(\mathrm{e}^u + \mathrm{e}^{-u}) \leq \mathrm{e}^{u^2/2}.
\end{aligned}$$

Proof (2)

- By functional calculus, the following Markov-type inequality holds:

$$\Pi_{[\lambda,\infty)}(\log(\Delta_{\sigma_n|\rho_n})) = \Pi_{[e^{\lambda t},\infty)}(\Delta^t_{\sigma_n|\rho_n}) \le e^{-\lambda t}\Delta^t_{\sigma_n|\rho_n}.$$

- Therefore

$$\langle \Omega_{\rho_n}, \Pi_{[\lambda,\infty)}(\log \Delta_{\sigma_n|\rho_n})(\Omega_{\rho_n})\rangle \le R^n e^{-(\lambda+nD(\rho_1\|\sigma_1))t} e^{nt^2c^2/2},$$

- Optimizing over t,

$$\langle \Omega_{\rho_n}, \Pi_{[\lambda,\infty)}(\log \Delta_{\sigma_n|\rho_n})(\Omega_{\rho_n})\rangle \le \begin{cases} R^n e^{-(\lambda+nD(\rho_1\|\sigma_1))^2/(2nc^2)} & \text{if } \lambda \le n(c^2 - D(\rho_1\|\sigma_1)) \\ R^n e^{-(\lambda+nD(\rho_1\|\sigma_1))+nc^2/2} & \text{else.} \end{cases}$$