Stein's Lemma for Classical-Quantum Channels

Mario Berta

arXiv:1808.01498
with Hirche, Kaur, Wilde

IEEE ISIT Paris, 2019

Imperial College London

Hypothesis Testing

- Discriminate between two sequences of quantum states ρ_n, σ_n on $\mathcal{H}^{\otimes n}$ — **null and alternative hypothesis** — with errors

  $\alpha_n(M_n) := \mathrm{Tr}\big[\rho_n(1 - M_n)\big]$ Type 1 error $\quad \beta_n(M_n) := \mathrm{Tr}\big[\sigma_n M_n\big]$ Type 2 error

  for two outcome POVM $\{M_n, (1 - M_n)\}$.

- **Symmetric setting** for $\rho_n = \rho^{\otimes n}, \sigma_n = \sigma^{\otimes n}$ with

$$\xi_n(\rho, \sigma) := -\frac{1}{n} \log \inf_{0 \leq M_n \leq 1} \left(\frac{\alpha_n(M_n)}{2} + \frac{\beta_n(M_n)}{2} \right)$$ leads to

Quantum Chernoff Bound [Audenaert *et al.* 07]

$$\xi(\rho, \sigma) := \lim_{n \to \infty} \xi_n(\rho, \sigma) = -\log \min_{0 \leq s \leq 1} \mathrm{Tr}\left[\rho^s \sigma^{1-s}\right]$$

Asymmetric Hypothesis Testing

- Same two type of errors $\alpha_n(M_n), \beta_n(M_n)$ and $\rho_n = \rho^{\otimes n}, \sigma_n = \sigma^{\otimes n}$ but **asymmetric setting** with

$$D_h^{\varepsilon,n}(\rho\|\sigma) := -\frac{1}{n}\log \inf_{0\leq M_n\leq 1}\left\{\beta_n(M_n)\middle|\alpha_n(M_n)\leq\varepsilon\right\}$$

leads to asymptotic error exponent

Quantum Stein's Lemma [Hiai & Petz 91, Ogawa & Nagaoka 10]

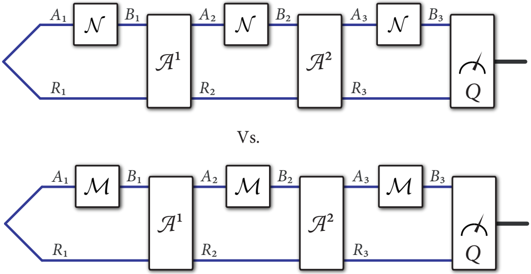
$$D(\rho\|\sigma) := \lim_{n\to\infty} D_h^{\varepsilon,n}(\rho\|\sigma) = \mathrm{tr}\left[\rho\,(\log\rho - \log\sigma)\,\right]$$

$\Rightarrow$ quantum relative entropy (independent of ε)

- Main question: What happens for **channel discrimination**?

Channel Discrimination

- **Adaptive protocols**: After each invocation of the channel $\mathcal{N}_{A\to B}$ or $\mathcal{M}_{A\to B}$, an (adaptive) channel $\mathcal{A}^i_{B_iR_i\to A_{i+1}R_{i+1}}$ is applied to the registers B_i and R_i



- Adaptive protocols help for finite rounds [Harrow *et al.* 10]

Channel Discrimination (continued)

- General strategy $\left\{\mathcal{A}^i_{B_iR_i\to A_{i+1}R_{i+1}}\right\}$ with intermediate states

  $$\rho_{B_iR_i} := \mathcal{N}_{A_i\to B_i}(\rho_{A_iR_i}) \text{ and } \tau_{B_iR_i} := \mathcal{M}_{A_i\to B_i}(\tau_{A_iR_i}) \text{ (set } \tau_{A_1R_1} = \rho_{A_1R_1})$$

  and final two outcome POVM $\{Q_{B_nR_n}, 1_{B_nR_n} - Q_{B_nR_n}\}$ gives rise to **type I and type II errors**

  $$\alpha_n(\{Q,\mathcal{A}\}) := \mathrm{Tr}\left[(1_{B_nR_n} - Q_{B_nR_n})\rho_{B_nR_n}\right] \quad \beta_n(\{Q,\mathcal{A}\}) := \mathrm{Tr}\left[Q_{B_nR_n}\tau_{B_nR_n}\right].$$

- Asymmetric setting in the sense of **Stein**

  $$D_h^{\varepsilon,n}(\mathcal{N}\|\mathcal{M}) := -\frac{1}{n}\log\inf_{\{Q,\mathcal{A}\}}\{\beta_n(\{Q,\mathcal{A}\})|\alpha_n(\{Q,\mathcal{A}\}) \le \varepsilon\}$$

  with the **asymptotic question**

  $$\lim_{n\to\infty} D_h^{\varepsilon,n}(\mathcal{N}\|\mathcal{M}) = ? \quad \text{and with that} \quad D(\mathcal{N}\|\mathcal{M}) := ?$$

Channel Discrimination (continued)

- For **classical channels** $\mathcal{N}, \mathcal{M}$ with trivial quantum memory R we have for $\varepsilon \in (0,1)$ that [Hayashi 09, Polyanski 09]

$$\lim_{n\to\infty} D_h^{\varepsilon,n}(\mathcal{N}\|\mathcal{M}) = \max_x D(N(|x\rangle\langle x|)\|M(|x\rangle\langle x|)) =: D(\mathcal{N}\|\mathcal{M}).$$

$\Rightarrow$ adaptive protocols are of no asymptotic help

Stein's Lemma for Classical-Quantum Channels

For channels $\mathcal{N}_{X\to B}(\cdot) = \sum_x \langle x| \cdot |x\rangle \nu_B^x$, $\mathcal{M}_{X\to B}(\cdot) = \sum_x \langle x| \cdot |x\rangle \mu_B^x$ we have for $\varepsilon \in (0,1)$ that

$$\begin{aligned}\lim_{n\to\infty} D_h^{\varepsilon,n}(\mathcal{N}\|\mathcal{M}) &= \max_\rho D((\mathcal{N}_{A\to X}\otimes \mathcal{I}_R)(\rho_{AR})\|(\mathcal{M}_{A\to X}\otimes \mathcal{I}_R)(\rho_{AR}))\\ &= \max_x D(\nu_B^x\|\mu_B^x) =: D(\mathcal{N}\|\mathcal{M}).\end{aligned}$$

Proof Stein's Lemma

- Achievability directly from Stein's lemma for quantum states
- In the following converse proof for $\varepsilon \to 0$ (weak converse)

Measures of distinguishability:

- Channel relative entropy [Leditzky *et al.* 18]

$$D(\mathcal{N}\|\mathcal{M}) := \sup_{\rho_{AR}} D((\mathcal{N}_{A\to B} \otimes \mathcal{I}_R)(\rho_{AR})\|(\mathcal{M}_{A\to B} \otimes \mathcal{I}_R)(\rho_{AR}))$$

- Amortized channel relative entropy

$$D^{\mathcal{A}}(\mathcal{N}\|\mathcal{M})$$
$$:= \sup_{\rho_{AR},\sigma_{AR}} D((\mathcal{N}_{A\to B} \otimes \mathcal{I}_R)(\rho_{AR})\|(\mathcal{M}_{A\to B} \otimes \mathcal{I}_R)(\sigma_{AR})) - D(\rho_{AR}\|\sigma_{AR})$$

Proof Stein's Lemma (continued)

- Weak converse quantum channel discrimination

$$D_h^{\varepsilon,n}(\mathcal{N}\|\mathcal{M}) \leq \frac{1}{1-\varepsilon}\left(D^{\mathcal{A}}(\mathcal{N}\|\mathcal{M}) + \frac{h_2(\varepsilon)}{n}\right)$$

from monotonicity of quantum relative entropy under channels.

Classical-Quantum Amortization Collapse

For channels $\mathcal{N}_{X\to B}(\cdot) = \sum_x \langle x| \cdot |x\rangle \nu_B^x$, $\mathcal{M}_{X\to B}(\cdot) = \sum_x \langle x| \cdot |x\rangle \mu_B^x$ we have

$$D^{\mathcal{A}}(\mathcal{N}\|\mathcal{M}) = \max_x D\left(\nu_B^x \| \mu_B^x\right) = D(\mathcal{N}\|\mathcal{M}).$$

- Proof is as follows.

Amortization Collapse $D^{\mathcal{A}}(\mathcal{N}\|\mathcal{M}) = \max_x D\left(\nu_B^x\|\mu_B^x\right)$

$$D^{\mathcal{A}}(\mathcal{N}\|\mathcal{M}) := \sup_{\rho_{AR},\sigma_{AR}} D((\mathcal{N}_{A\to B}\otimes\mathcal{I}_R)(\rho_{AR})\|(\mathcal{M}_{A\to B}\otimes\mathcal{I}_R)(\sigma_{AR})) - D(\rho_{AR}\|\sigma_{AR})$$

We have $D^{\mathcal{A}}(\mathcal{N}\|\mathcal{M}) \geq D(\nu_B^x\|\mu_B^x)$ for $\rho_{AR} = \sigma_{AR} = |x\rangle\langle x|_A \otimes |x\rangle\langle x|_R$ and it remains to show

$$D^{\mathcal{A}}(\mathcal{N}\|\mathcal{M}) \leq \max_x D(\nu_B^x\|\tau_B^x).$$

Proof: By monotonicity of quantum relative entropy under channels

$$\begin{aligned}
&D(\mathcal{N}_{A\to B}(\rho_{AR})\|\mathcal{M}_{A\to B}(\sigma_{AR})) - D(\rho_{AR}\|\sigma_{AR})\\
&\leq D\left(\sum_x p_x\nu_B^x\otimes\rho_R^x\middle\|\sum_x q_x\mu_B^x\otimes\sigma_R^x\right) - D\left(\sum_x p_x|x\rangle\langle x|_X\otimes\nu_B^x\otimes\rho_R^x\middle\|\sum_x q_x|x\rangle\langle x|_X\otimes\nu_B^x\otimes\sigma_R^x\right)\\
&\leq D\left(\sum_x p_x|x\rangle\langle x|_X\otimes\nu_B^x\otimes\rho_R^x\middle\|\sum_x q_x|x\rangle\langle x|_X\otimes\mu_B^x\otimes\sigma_R^x\right)\\
&\quad - D\left(\sum_x p_x|x\rangle\langle x|_X\otimes\nu_B^x\otimes\rho_R^x\middle\|\sum_x q_x|x\rangle\langle x|_X\otimes\nu_B^x\otimes\sigma_R^x\right)
\end{aligned}$$

Amortization Collapse $D^{\mathcal{A}}(\mathcal{N}\|\mathcal{M}) = \max_x D\left(\nu_B^x\|\mu_B^x\right)$

$$= \sum_x p_x \left(\mathrm{Tr}\left[(\nu_B^x \otimes \rho_R^x) \log\left(q_x \nu_B^x \otimes \sigma_R^x\right) \right] - \mathrm{Tr}\left[\left(\nu_B^x \otimes \rho_R^x\right) \log\left(q_x \mu_B^x \otimes \sigma_R^x\right) \right] \right)$$

$$= \sum_x p_x \,\mathrm{Tr}\left[\left(\nu_B^x \otimes \rho_R^x\right) \left(\left(\log \nu_B^x - \log \mu_B^x\right) \otimes 1_R\right) \right]$$

$$= \sum_x p_x D(\nu_B^x \| \mu_B^x)$$

$$\leq \max_x D(\nu_B^x \| \mu_B^x) \quad \square$$

- Converse proof for $\varepsilon \in (0,1)$ (strong converse) via standard techniques together with classical-quantum amortization collapse for **channel Rényi divergences** based on

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

Conclusion

Take Home Message

Stein's lemma for classical-quantum channels by showing that adaptive protocols are of no help (strong converse exponent formula)

Unresolved **quantum additivity questions**:

- Stein's lemma for general quantum channels remains open, do adaptive protocols help?
- Chernoff bound and error exponent for classical channels [Hayashi 09]
- Symmetric setting remains open even for classical-quantum channels — but already for entanglement breaking channels adaptive protocols help [Harrow *et al.* 10]