Notes on asymptotics of quantum hypothesis testing

Anna Jenčová*

1 Preliminaries

Let $\mathcal{H}$ be a finite dimensional Hilbert space.

1.1 Pinching

Let $A \in B(\mathcal{H})$ be self-djoint, with spectral decomposition $A = \sum_i \lambda_i P_i$. We will need the pinching map $B(\mathcal{H}) \to B(\mathcal{H})$, defined as

$$\mathcal{E}_A(X) = \sum_i P_i X P_i.$$

Then A is a cp unital map. Moreover, $\mathcal{E}_A(X)$ commutes with X and we have the pinching inequality [?]

$$\mathcal{E}_A(X) \le |\mathrm{spec}(A)| X, \qquad X \ge 0. \tag{1}$$

1.2 Relative entropies

Let ρ and σ be density operators. The (Umegaki) relative entropy is defined as

$$D(\rho\|\sigma) := \begin{cases} \mathrm{Tr}\,[\rho(\log\rho - \log\sigma)], & \mathrm{supp}(\rho) \le \mathrm{supp}(\sigma) \\ \infty, & \text{otherwise.} \end{cases}$$

The standard Rényi relative entropy for $\alpha \in [0,1] \setminus \{1\}$ is defined as

$$D_\alpha(\rho\|\sigma) := \begin{cases} \frac{1}{\alpha-1} \log \mathrm{Tr}\,[\rho^\alpha \sigma^{1-\alpha}], & \mathrm{supp}(\rho) \le \mathrm{supp}(\sigma) \text{ or } \alpha \in (0,1) \\ \infty, & \text{otherwise.} \end{cases}$$

The sandwiched Rényi relative entropy for $\alpha \in [1/2, \infty] \setminus \{1\}$ is defined as

$$\hat{D}_\alpha(\rho\|\sigma) := \begin{cases} \frac{1}{\alpha-1} \log \mathrm{Tr}\,[\sigma^{\frac{1-\alpha}{2\alpha}} \rho \sigma^{\frac{1-\alpha}{2\alpha}}], & \mathrm{supp}(\rho) \le \mathrm{supp}(\sigma) \text{ or } \alpha \in [1/2,1) \\ \infty, & \text{otherwise.} \end{cases}$$

*Mathematical Institute, Slovak Academy of Sciences, Štefánikova 49, 814 73 Bratislava, Slovakia, jenca@mat.savba.sk

1.3 The functions ϕ and ψ

We assume $\mathrm{supp}(\rho) \leq \mathrm{supp}(\sigma)$**, so that** $D(\rho\|\sigma) < \infty$**.**

Let us define

$$\phi(s) = \log \mathrm{Tr}\,[\rho^{1-s}\sigma^s], \qquad s \in \mathbb{R}.$$

Then ϕ is a strictly convex and smooth function, with derivative

$$\phi'(s) = (\mathrm{Tr}\,[\rho^{1-s}\sigma^s])^{-1}\mathrm{Tr}\,[\rho^{1-s}\sigma^s(\log\sigma - \log\rho)],$$

[**?** , Exercise 3.5] In particular, $\phi'(0) = -D(\rho\|\sigma)$ and $\phi'(1) = D(\sigma\|\rho)$ Let us define

$$\psi(\lambda) = \min_{s\in\mathbb{R}} \lambda s + \phi(s).$$

Lemma 1. *Let* $-D(\sigma\|\rho) = -\phi'(1) \leq \lambda \leq -\phi'(0) = D(\rho\|\sigma)$*. Then*

1. $\psi(\lambda) = \inf_{s\in[0,1]} \lambda s + \phi(s)$
2. ψ *is monotone increasing*
3. $\psi(-D(\sigma\|\rho)) = -D(\sigma\|\rho)$, $\psi(D(\rho\|\sigma)) = 0$.

Proof. By strict convexity, the derivative $\phi'(s)$ is increasing, so that there is some $s_0 \in [0,1]$ such that $\lambda = -\phi'(s_0)$ and the function $s \mapsto \lambda s + \phi(s)$ is decreasing for $s \leq s_0$ and increasing for $s_0 \leq s$. It follows that the infimum is attained at $s_0 \in [0,1]$. This also implies (3).

Assume that $\lambda \in (-\phi'(1), \phi'(0))$, then $\lambda = -\phi'(s_0)$ for some $s_0 \in (0,1)$ and we have

$$\psi(\lambda) = \lambda s_0 + \phi(s_0) = \phi(s_0) - \phi'(s_0)s_0 < \phi(0) = 0 = \psi(-\phi(0)),$$

the inequality follows by strict convexity of ϕ. If $-\phi'(1) \leq \lambda_1 < \lambda$, then clearly

$$\psi(\lambda_1) \leq \lambda_1 s_0 + \phi(s_0) < \lambda s_0 + \phi(s_0) = \psi(\lambda).$$

This proves (2). □

Lemma 2. *Let* $\lambda \in [-D(\sigma\|\rho), D(\rho\|\sigma)]$ *and* $0 \leq r \leq D(\rho\|\sigma)$*. Then*

$$\lambda - \psi(\lambda) = r \iff \psi(\lambda) = b(r) := \inf_{s\in[0,1]} \frac{s}{1-s}r + \frac{1}{1-s}\phi(s).$$

Proof. Let $s_\lambda \in [0,1]$ be such that $\lambda = -\phi'(s_\lambda)$, then the assuptions imply that

$$\psi(\lambda) = -\phi'(s_\lambda)s_\lambda + \phi(s_\lambda) = -\phi'(s_\lambda) - r.$$

Solving for $\phi'(s_\lambda)$, we get $-\phi'(s_\lambda) = \frac{1}{1-s_\lambda}(r + \phi(s_\lambda))$, so that

$$\psi(\lambda) = \frac{s_\lambda}{1-s_\lambda}r + \frac{1}{1-s_\lambda}\phi(s_\lambda).$$

Let

$$g(s) = \frac{s}{1-s}r + \frac{1}{1-s}\phi(s).$$

Then $g'(s) = \frac{1}{(1-s)^2}(r + \phi(s) + \phi'(s)(1-s))$. Now note that $h(s) := \phi(s) + \phi'(s)(1-s)$ satisfies $h'(s) = \phi''(s)(1-s)$ so that h is increasing on $(0,1)$ (strict) convexity of ϕ, so that $-D(\rho\|\sigma) = h(0) \leq h(s) \leq h(1) = 0$. It follows that the derivative $g'(s)$ changes sign at a unique point $s_r \in [0,1]$, such that $r + \phi(s_r) + \phi'(s_r)(1-s_r) = 0$. Comparing this to the above computation, we see that $s_r = s_\lambda$ and

$$\psi(\lambda) = \min_{s\in[0,1]} g(s) = b(r).$$

□

We define

$$\psi^*(\lambda) = \inf_{t\in[-1,0]} t\lambda + \phi(t).$$

Again, if $\lambda > D(\rho\|\sigma) = -\phi'(0)$, then $t \mapsto t\lambda + \phi(t)$ is strictly increasing at $t = 0$, which implies that $\phi^*(\lambda) < 0$.

1.4 Inequalities

We have two basic inequalities. For $A, B \geq 0$, let $\{A \geq B\}$ be the sum of eigenprojections of $A - B$ corresponding to nonnegative eigenvalues, similarly $\{A \leq B\}$, $\{A > B\}$ etc. Then

Lemma 3 (Quantum Neyman-Pearson). *We have*

$$\min_{0\leq T\leq I} \mathrm{Tr}\,[A(I-T)] + \mathrm{Tr}\,[BT] = \mathrm{Tr}\,[A\{A \leq B\}] + \mathrm{Tr}\,[B\{A > B\}].$$

Lemma 4 (Audenaert et al). *We have for any $s \in [0,1]$,*

$$\mathrm{Tr}\,[A\{A \leq B\}] + \mathrm{Tr}\,[B\{A > B\}] \leq \mathrm{Tr}\,[A^{1-s}B^s].$$

These statements hold in the von Neumann algebra case as well.

1.5 Nussbaum-Szkola probability distributions

Let $\rho = \sum_i \lambda_i |x_i\rangle\langle x_i|$ and $\sigma = \sum_j \mu_j |y_j\rangle\langle y_j|$ be the spectral decompositions. The pair (P,Q) of Nussbaum-Szkola probability distributions related to (ρ,σ) is defined on $[n]\times[n]$, here $n = \dim(\mathcal{H})$ and $[n] = \{1,\ldots,n\}$. We put

$$P_{ij} = \lambda_i |\langle x_i|y_j\rangle|^2, \qquad Q_{ij} = \mu_j |\langle x_i|y_j\rangle|^2.$$

We then have $D(\rho\|\sigma) = D(P\|Q)$ and $D_\alpha(\rho\|\sigma) = D_\alpha(P\|Q)$ for all α.

For $(\rho^{\otimes n}, \sigma^{\otimes n})$ we get the iid distributions (P^n, Q^n). We also have the following result:

Lemma 5 (Nussbaum-Szkola). *For any test T and $c > 0$ we have*

$$\alpha(T) + c\beta(T) \geq \frac{1}{2}(P(\{P \leq cQ\}) + cQ(\{P > cQ\})) = \frac{1}{2}\sum_{ij} \min\{P_{ij}, cQ_{ij}\}$$

Proof. Let T be a projection, then

$$\mathrm{Tr}\,[\rho T] = \sum_i \lambda_i \langle\, x_i | TT | x_i\,\rangle = \sum_{ij} \lambda_i \langle\, x_i | T | y_j\,\rangle \langle\, y_j | T | y_j\,\rangle = \sum_{ij} \lambda_i |\langle\, x_i | T | y_j\,\rangle|^2$$

and similarly for σ. It follows that

$$\alpha(T) + c\beta(T) = \sum_{ij} \lambda_i |\langle\, x_i | I - T | y_j\,\rangle|^2 + c \sum_{ij} \mu_j |\langle\, x_i | T | y_j\,\rangle|^2.$$

Now we use the inequality

$$a|u-v|^2 + b|v|^2 \geq \frac{1}{2}|u|^2 \min\{a, b\}$$

to lower bound

$$\alpha(T) + c\beta(T) \geq \frac{1}{2} \sum_{ij} |\langle\, x_i | y_j\,\rangle|^2 \min\{\lambda_i, c\mu_j\} = \frac{1}{2} \sum_{i,j} \min\{P_{ij}, cQ_{ij}\}$$

By Lemma 3, this inequality holds for all tests. The equality $\min\{P_{ij}, cQ_{ij}\} = P(\{P \leq cQ\}) + cQ(\{P > cQ\})$ can be easily seen.

□

2 QHT

Let ρ, σ be a pair of density matrices. We test the hypothesis $H_0 = \rho$ against the alternative $H_1 = \sigma$. A test is given by an operator $0 \leq T \leq I$, corresponding to accepring H_0. The two error probabilities are

$$\alpha(T) = \mathrm{Tr}\,[(I - T)\rho], \qquad \beta(T) = \mathrm{Tr}\,[T\sigma].$$

We will consider the asymptotic behaviour of the error probabilities

$$\alpha_n(T_n) = \mathrm{Tr}\,[(I - T_n)\rho_n], \qquad \beta_n(T_n) = \mathrm{Tr}\,[T_n\sigma_n]$$

in testing $H_0 = \rho_n := \rho^{\otimes n}$ against $H_1 = \sigma_n := \sigma^{\otimes n}$.

2.1 Quantum Stein's lemma

We assume $\mathrm{supp}(\rho) \leq \mathrm{supp}(\sigma)$**, so that** $D(\rho\|\sigma) < \infty$**.**

- Quantum Stein's lemma states that if the I. kind error probabilities are constrained as $\alpha_n(T_n) \leq \epsilon$, then the II. kind error probabilities go to zero exponentially, with optimal decay rate equal to the relative entropy $D(\rho\|\sigma)$.
- The strong converse says that if the decay rate of $\beta_n(T_n)$ is greater than $D(\rho\|\sigma)$, then $\alpha_n(T_n) \to 0$ exponentially.

We will need the following inequalities (3) and (4).

Let $\lambda \in \mathbb{R}$ and let $S_n := \{\rho^{\otimes n} > e^{n\lambda}\sigma^{\otimes n}\}$. Then using Lemma 4 (Audenaert) with $A = \rho^{\otimes n}$ and $B = e^{\lambda n}\sigma^{\otimes n}$, we get for any $s \in [0, 1]$

$$\alpha_n(S_n) + e^{n\lambda}\beta_n(S_n) \leq \mathrm{Tr}\, e^{n\lambda s}[(\rho^{\otimes n})^{1-s}(\sigma^{\otimes n})^s] = e^{n\lambda s}(\mathrm{Tr}\,[\rho^{1-s}\sigma^s])^n = e^{n(\lambda s + \phi(s))}. \tag{2}$$

Hence by taking the infimum over $s \in [0,1]$,

$$\alpha_n(S_n) \le e^{n\psi(\lambda)}, \qquad \beta_n(S_n) \le e^{n(-\lambda+\psi(\lambda))} \tag{3}$$

On the other hand, put $p_n = \mathrm{Tr}\,[\rho^{\otimes n} S_n]$ and $q_n = \mathrm{Tr}\,[\sigma^{\otimes n} S_n]$. Then $p_n \ge e^{n\lambda} q_n$ and therefore $p_n^t \le e^{n\lambda t} q_n^t$ for for any $t \in [-1,0]$. We get

$$\begin{aligned} 1-\alpha_n(S_n) = p_n &\le e^{n\lambda t} p_n^{1-t} q_n^t \le e^{n\lambda t}(p_n^{1-t}q_n^t + (1-p_n)^{1-t}(1-q_n)^t) \le e^{n\lambda t}\mathrm{Tr}\,[(\rho^{\otimes n})^{1-t}(\sigma^{\otimes n})^t] \\ &= e^{n(\lambda t+\phi(t))} \end{aligned}$$

for all $t \in [-1,0]$. It follows that for any test T_n, we have

$$\begin{aligned} 1-\alpha_n(T_n) = \mathrm{Tr}\,[\rho^{\otimes n} T_n] &= \mathrm{Tr}\,[(\rho^{\otimes n} - e^{\lambda n}\sigma^{\otimes n})T_n] + e^{\lambda n}\beta_n(T_n) \le \mathrm{Tr}\,[(\rho^{\otimes n} - e^{\lambda n}\sigma^{\otimes n})S_n] + e^{\lambda n}\beta_n(T_n) \\ &\le 1-\alpha_n(S_n) + e^{\lambda n}\beta_n(T_n) \le e^{n(\lambda t+\phi(t))} + e^{\lambda n}\beta_n(T_n) \end{aligned}$$

and hence

$$\beta_n(T_n) \ge e^{-n\lambda}(1-\alpha_n(T_n) - e^{n\psi^*(\lambda)}) \tag{4}$$

Let

$$\beta_n(\epsilon) := \min_{0\le T_n \le I}\{\beta_n(T_n) \mid \alpha_n(T_n) \le \epsilon\}, \qquad \epsilon > 0.$$

Lemma 6 (Quantum Stein's lemma)**.** *[? ?] For all $\epsilon \in (0,1)$, we have*

$$\lim_{n\to\infty} -\frac{1}{n}\log\beta_n(\epsilon) = D(\rho\|\sigma).$$

In particular, there exists a sequence T_n of tests such that $\alpha_n(T_n) \to 0$ and $\lim_n \frac{1}{n}\log\beta_n(T_n) = -D(\rho\|\sigma)$.

Proof. Let $\lambda < D(\rho\|\sigma)$, then by Lemma 1, $\psi(\lambda) < 0$, so that in this case (3), $\alpha_n(S_n) \to 0$ and

$$-\frac{1}{n}\log\beta_n(S_n) \ge \lambda - \psi(\lambda) > \lambda.$$

For $\epsilon \in (0,1)$ we have $\alpha_n(S_n) \le \epsilon$ for large enough n, so that $\beta_n(\epsilon) \le \beta_n(S_n)$. It follows that

$$\liminf_n -\frac{1}{n}\log\beta_n(\epsilon) \ge \liminf_n -\frac{1}{n}\log\beta_n(S_n) \ge \lambda.$$

Conversely, by (4) we have for any sequence of tests such that $\alpha_n(T_n) \le \epsilon$ that

$$\beta_n(T_n) \ge e^{-n\lambda}(1-\epsilon - e^{n\psi^*(\lambda)}).$$

Since $\psi^*(\lambda) < 0$ for $\lambda > D(\rho\|\sigma)$, this implies that, for such λ,

$$\limsup_n -\frac{1}{n}\beta_n(\epsilon) \le \lambda.$$

Choosing any $\delta > 0$, we obtain

$$D(\rho\|\sigma) - \delta \le \liminf_n -\frac{1}{n}\beta_n(\epsilon) \le \limsup_n -\frac{1}{n}\beta_n(\epsilon) \le D(\rho\|\sigma) + \delta.$$

Since δ was arbitrary, this implies the first statement. For the second statement, we can choose sequences $\delta_n, \epsilon_n > 0$, $\delta_n, \epsilon_n \to 0$, then we can find a sequence of tests T_n such that $\alpha_n(T_n) \le \epsilon_n$ and $|\frac{1}{n}\log\beta_n(T_n) + D(\rho\|\sigma)| < \delta_n$.

□

Lemma 7 (Strong converse)**.** *[?] Let T_n be a sequence of tests and let $r > D(\rho\|\sigma)$. If*

$$\limsup_n \frac{1}{n}\log\beta_n(T_n) \le -r$$

then $\alpha_n(T_n) \to 1$ exponentially fast.

Proof. By (4), we have for any $\lambda \in \mathbb{R}$, $\delta > 0$ and large enough n,

$$1-\alpha_n(T_n) \le e^{n\psi^*(\lambda)} + e^{n\lambda}\beta_n(T_n) \le e^{n\psi^*(\lambda)} + e^{n(\lambda-r+\delta)}.$$

We then may choose δ and λ such that $D(\rho\|\sigma) < \lambda < r-\delta$, in which case both of the above exponents are negative.

□

2.2 Hoeffding bound

The Hoeffding bound studies the exponential decay of α_n under an exponential constraint on β_n. Specifically, we look at the value of

$$B_e(r) := \sup\{-\limsup_n \frac{1}{n}\log\alpha_n(T_n) \mid \limsup_n \frac{1}{n}\log\beta_n(T_n) \le -r\}, \qquad r > 0$$

The following result was proved in [?] (direct part) and [?].

Lemma 8 (Quantum Hoeffding bound)**.** *For $0 < r \le D(\rho\|\sigma)$, we have*

$$B_e(r) = \sup_{0\le s\le 1} \frac{-sr-\phi(s)}{1-s} = -b(r).$$

Proof. By (2), we see that for any λ,

$$\alpha_n(S_n) \le e^{n(s\lambda+\phi(s))}, \qquad \beta_n(S_n) \le e^{n(-(1-s)\lambda+\phi(s))}$$

Let us choose λ such that $-(1-s)\lambda+\phi(s) = -r$, that is, $\lambda = \frac{r+\phi(s)}{1-s}$, then we obtain

$$\alpha_n(S_n) \le e^{-n\frac{-sr-\phi(s)}{1-s}}, \qquad \beta_n(S_n) \le e^{-nr}.$$

. This implies that $B_e(r) \ge \sup_{0\le s\le 1} \frac{-sr-\phi(s)}{1-s} = -b(r)$. To show the lower bound, we will use the pair of Nussbaum-Szkola probability distributions (P, Q) related to the pair (ρ, σ) and Lemma 5.

What we need to prove is that, for a sequence of tests T_n,

$$\limsup_n \frac{1}{n}\log\beta_n(T_n) \le -r \implies \limsup_n \frac{1}{n}\log\alpha_n(T_n) \ge b(r).$$

By Lemma 5, we have for any $b \in \mathbb{R}$,

$$\alpha_n(T_n) + e^{nb}\beta_n(T_n) \ge \frac{1}{2}[P^n(\{P^n \le e^{nb}Q^n\}) + e^{-nb}Q^n(\{P^n > e^{nb}Q^n\})] \tag{5}$$

Now note that $\{P^n \le e^{nb}Q^n\} = \{\frac{1}{n}\log\frac{Q^n}{P^n} \ge -b\}$ and

$$\log\frac{Q^n(\omega^n)}{P^n(\omega^n)} = \sum_k \log\frac{Q(\omega_k)}{P(\omega_k)}, \qquad \omega^n = (\omega_1,\dots,\omega_n) \in \Omega^n,$$

where $\Omega = [n]\times[n]$. Put $X(\omega) = \log \frac{Q(\omega)}{P(\omega)}$, then $E_P[X] = -D(P\|Q) = -D(\rho\|\sigma)$ and the cumulant generating function of X at P is

$$\log E_P[e^{sX}] = \log E_P[Q^s P^{-s}] = \phi_{P\|Q}(s) = \phi_{\rho\|\sigma}(s) = \phi(s).$$

Now the Cramér theorem of tha large deviation theory implies that

$$\lim_n \frac{1}{n}\log(P(\frac{1}{n}\sum_i X_i \geq -b)) = \psi(b)$$

for all $-b > E_P(X) = -D(\rho\|\sigma)$. Similarly, $\{P^n > e^{nb}Q^n\} = \{\frac{1}{n}\sum_i X_i(\omega_i) < -b\}$, and we have $E_Q[X] = D(Q\|P) = D(\sigma\|\rho)$ and

$$\log E_Q[e^{sX}] = \log \mathrm{Tr}\,[\sigma^{1+s}\rho^{-s}] = \phi(1+s).$$

Now note that

$$\inf_s bs + \phi(1+s) = \inf_s b(1+s) + \phi(1+s) - b = \psi(b) - b$$

so that

$$\lim_n \frac{1}{n}\log(e^{nb}Q(\frac{1}{n}\sum_i X_i < -b)) = \psi(b)$$

for $-b < E_Q[X] = D(\sigma\|\rho)$. It follows that for any $b \in (-D(\sigma\|\rho), D(\rho\|\sigma))$, we have

$$\lim_n \frac{1}{n}\log[P^n(\{P^n \leq e^{nb}Q^n\}) + e^{-nb}Q^n(\{P^n > e^{nb}Q^n\})] = \psi(b),$$

this follows from

$$\log 2 + \min\{\log x, \log y\} = \log(2\min\{x,y\}) \leq \log(x+y) \leq \log 2 + \max\{\log x, \log y\}$$

and the fact that the two limits above are the same. From (5) and the assumption on $\beta_n(T_n)$, we now get

$$\psi(b) \leq \liminf_n \frac{1}{n}\log(\alpha_n(T_n) + e^{nb}\beta_n(T_n)) \leq \max\{\limsup_n \frac{1}{n}\log\alpha_n(T_n), b - r\}.$$

Let us now assume that $0 < r \leq D(\rho\|\sigma)$ and let $\lambda \in (-D(\sigma\|\rho), D(\rho\|\sigma)]$ be such that $r = \lambda - \psi(\lambda)$. Choose a small $\epsilon > 0$ such that $b = \lambda - \epsilon > -D(\sigma\|\rho)$ (we clearly must have $\lambda > -D(\sigma\|\rho)$, since otherwise $r = 0$). Then we get

$$\psi(\lambda - \epsilon) \leq \max\{\limsup_n \frac{1}{n}\log\alpha_n(T_n), \psi(\lambda) - \epsilon\}.$$

Since clearly

$$(\lambda - \epsilon)s + \phi(s) > \lambda s + \phi(s) - \epsilon, \qquad s \in (0,1)$$

and by the assumptions both infima are attained in $(0,1)$, we see that $\phi(\lambda-\epsilon) > \phi(\lambda) - \epsilon$, so that we must have

$$\phi(\lambda - \epsilon) \leq \limsup_n \frac{1}{n}\log\alpha_n(T_n)$$

Taking the limit $\epsilon \to 0$ and noting that $\phi(\lambda) = b(r)$ finishes the proof.

$\square$

2.3 Strong converse exponents