Discriminating quantum states: the *multiple Chernoff distance*

Ke Li
California Institute of Technology

QMath 13, Georgia Tech

K. Li, Annals of Statistics 44: 1661-1679 (2016); arXiv:1508.06624

Outline

1. The problem
2. The answer
3. History review
4. Proof sketch
5. One-shot case
6. Open questions

Accessing quantum systems: quantum measurement

- Quantum measurement: formulated as positive operator-valued measure (POVM)

  $$\mathcal{M} = \{M_i\}_i, \text{ with } 0 \leq M_i \leq \mathbb{1} \text{ and } \sum_i M_i = \mathbb{1};$$

  when performing the POVM on a system in the state ω, we obtain outcome "i" with probability $\mathrm{Tr}(\omega M_i)$.

- von Neumann measurement: special case of POVM, with the POVM elements being orthogonal projectors: $M_i M_j = \delta_{ij} M_i$, where δ_{ij} is the Kronecker delta.

Quantum state discrimination (quantum hypothesis testing)

- Suppose a quantum system is in one of a set of states $\{\omega_1, \ldots, \omega_r\}$, with a given prior $\{p_1, \ldots, p_r\}$. The task is to detect the true state with a minimal error probabality.
- Method: making quantum measurement $\{M_i\}_{i=1}^r$.
- Error probability (let $A_i := p_i \omega_i$)

$$P_e(\{A_1, \ldots, A_r\}; \{M_1, \ldots, M_r\}) := \sum_{i=1}^r \operatorname{Tr} A_i(\mathbb{1} - M_i).$$

- Optimal error probability

$$P_e^*(\{A_1, \ldots, A_r\}) :=$$

$$\min \left\{ P_e(\{A_1, \ldots, A_r\}; \{M_1, \ldots, M_r\}) : \mathrm{POVM}\ \{M_1, \ldots, M_r\} \right\}.$$

Asymptotics in quantum hypothesis testing

- What's the asymptotic behavior of
  $P_e^*\left(\{p_1\rho_1^{\otimes n},\ldots,p_r\rho_r^{\otimes n}\}\right)$, as $n \to \infty$?

- Exponentially decay! (Parthasarathy '2001)

$$P_e^* \sim \exp(-\xi n)$$

- But, what's the error exponent

$$\xi = \liminf_{n\to\infty} \frac{-1}{n} \log P_e^*\left(\{p_1\rho_1^{\otimes n},\ldots,p_r\rho_r^{\otimes n}\}\right) \text{ ?}$$

It has been an open problem (except for r=2)!

Outline

1. The problem
2. The answer
3. History review
4. Proof sketch
5. One-shot case
6. Open questions

Our result: error exponent = multiple Chernoff distance

- We prove that

Theorem *Let $\{\rho_1, \ldots, \rho_r\}$ be a finite set of quantum states on a finite-dimensional Hilbert space $\mathcal{H}$. Then the asymptotic error exponent for testing $\{\rho_1^{\otimes n}, \ldots, \rho_r^{\otimes n}\}$, for an arbitrary prior $\{p_1, \ldots, p_r\}$, is given by the multiple quantum Chernoff distance:*

$$\lim_{n\to\infty} \frac{-1}{n} \log P_e^* \left(\{p_1\rho_1^{\otimes n}, \ldots, p_r\rho_r^{\otimes n}\}\right) = \min_{(i,j):i\neq j} \max_{0\leq s\leq 1} \left\{-\log \operatorname{Tr} \rho_i^s \rho_j^{1-s}\right\}. \quad (1)$$

Remarks

- Remark 1: Our result is a multiple-hypothesis generalization of the r=2 case. Denote the multiple quantum Chernoff distance (r.h.s. of eq. (1)) as $C(\rho_1, \ldots, \rho_r)$, then

$$C(\rho_1, \ldots, \rho_r) = \min_{(i,j):i\neq j} C(\rho_i, \rho_j),$$

with the binary quantum Chernoff distance is defined as

$$C(\rho_1, \rho_2) := \max_{0\leq s\leq 1} \{-\log \mathrm{Tr}\, \rho_1^s \rho_2^{1-s}\}.$$

- Remark 2: when $\rho_1, \ldots, \rho_r$ commute, the problem reduces to classical statistical hypothesis testing. Compared to the classical case, the difficulty of quantum statistics comes from noncommutativity & entanglement.

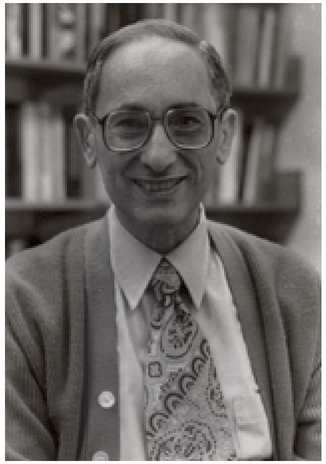
Outline

1. The problem
2. The answer
3. History review
4. Proof sketch
5. One-shot case
6. Open questions

Some history review

- The classical Chernoff distance as the opimal error exponent for testing two probability distributions was given in

  H. Chernoff, Ann. Math. Statist. 23, 493 (1952).



- The multipe generalizations were subsequently made in

  N. P. Salihov, Dokl. Akad. Nauk SSSR 209, 54 (1973);
  E. N. Torgersen, Ann. Statist. 9, 638 (1981);
  C. C. Leang and D. H. Johnson, IEEE Trans. Inf. Theory 43, 280 (1997);
  N. P. Salihov, Teor. Veroyatn. Primen. 43, 294 (1998).

Some history review

- Quantum hypothesis testing (state discrimination) was the main topic in the early days of quantum information theory in 1970s.
- Maximum likelihood estimation
  - for two states: Holevo-Helstrom tests

$$(\{\rho_1 - \rho_2 > 0\}, \mathbb{1} - \{\rho_1 - \rho_2 > 0\})$$

C. W. Helstrom, Quantum Detection and Estimation Theory, Academic Press (1976); A. S. Holevo, Theor. Prob. Appl. 23, 411 (1978).

  - for more than two states: only formulated in a complex and implicit way. Competitions between pairs make the problem complicated!

A. S. Holevo, J. Multivariate Anal. 3, 337 (1973); H. P. Yuen, R. S. Kennedy and M. Lax, IEEE Trans. Inf. Theory 21, 125 (1975).

Some history review

- In 2001, Parthasarathy showed exponential decay.

K. R. Parthasarathy, in Stochastics in Finite and Infinite Dimensions 361 (2001).

- In 2006, two groups [Audenaert et al] and [Nussbaum & Szkola] together solved the r=2 case.

K. Audenaert et al, arXiv: quant-ph/0610027; Phys. Rev. Lett. 98, 160501 (2007);
M. Nussbaum and A. Szkola, arXiv: quant-ph/0607216 ; Ann. Statist. 37, 1040 (2009).

- In 2010/2011, Nussbaum & Szkola conjectured the solution (our theorem), and proved that $C/3 \leq \xi \leq C$.

M. Nussbaum and A. Szkola, J. Math. Phys. 51, 072203 (2010); Ann. Statist. 39, 3211 (2011).

- In 2014, Audenaert & Mosonyi proved that $C/2 \leq \xi \leq C$.

K. Audenaert and M. Mosonyi, J. Math. Phys. 55, 102201 (2014).

Outline

1. The problem
2. The answer
3. History review
4. Proof sketch
5. One-shot case
6. Open questions

Sketch of proof

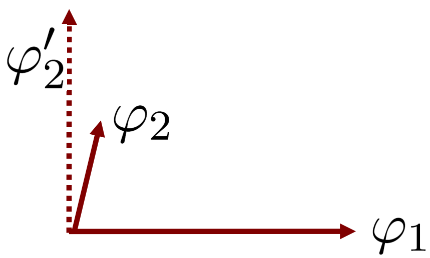
- We only need to prove the achievability part " $\xi \geq C$ ".

  For this purpose, we construct an asymptotically optimal quantum measurement, and show that it achieves the quantum multiple Chernoff distance as the error exponent.

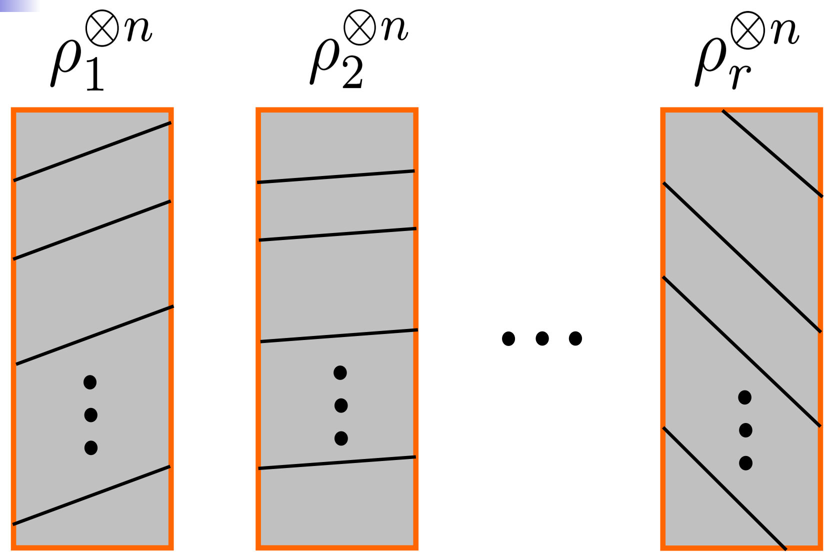
- Motivation: consider detecting two weighted pure states.

  Big overlap: give up the light one;

  Small overlap: make a projective measurement, using orthonormalized version of the two states.



Sketch of proof



Spectral decomposition:

$$\rho_i^{\otimes n} = \bigoplus_{k=1}^{T_i} \lambda_{ik}^{(n)} Q_{ik}^{(n)},$$

$$T := \max\{T_i\}_i \leq (n+1)^d$$

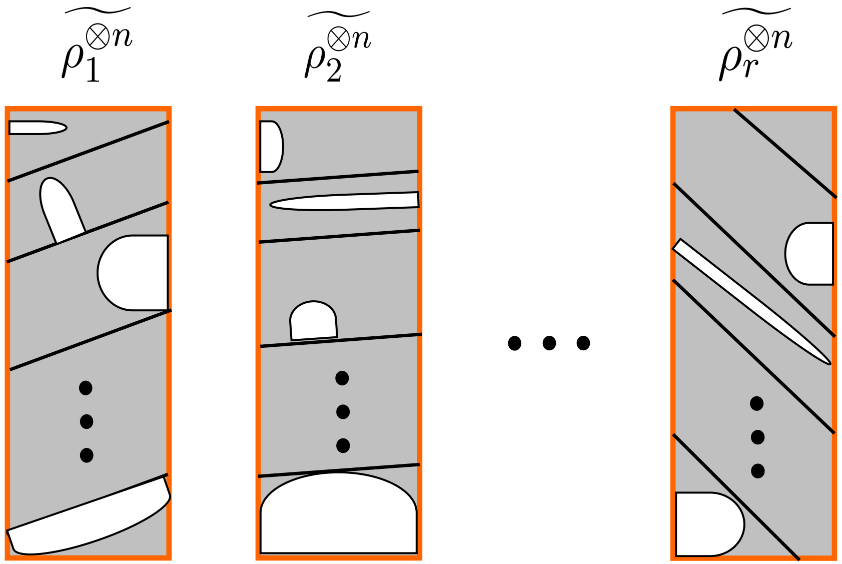
Overlap between eigenspaces:

$$\mathrm{Olap}\left(\mathrm{supp}\left(Q_{ik}^{(n)}\right), \mathrm{supp}\left(Q_{j\ell}^{(n)}\right)\right)$$

$$:= \max\left\{ |\langle\varphi|\phi\rangle| : |\varphi\rangle \in \mathrm{supp}\left(Q_{ik}^{(n)}\right), |\phi\rangle \in \mathrm{supp}\left(Q_{j\ell}^{(n)}\right)\right\}$$

Sketch of proof

"Dig holes" in every eigenspaces to reduce overlaps



ϵ-subtraction:

Let $P_1 P_2 P_1 = \bigoplus_x \lambda_x Q_x$

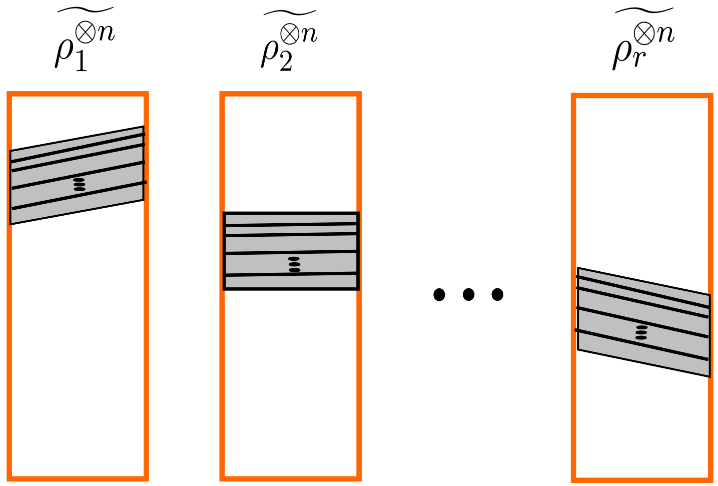
Define $P_1 \ominus_\epsilon P_2 := P_1 - \sum_{x:\lambda_x \geq \epsilon^2} Q_x$

$$\widetilde{\rho_i^{\otimes n}} = \bigoplus_{k=1}^{T_i} \lambda_{ik}^{(n)} \widetilde{Q_{ik}^{(n)}} \;, \quad \mathrm{Olap}\left(\mathrm{supp}\left(\widetilde{Q_{ik}^{(n)}}\right), \mathrm{supp}\left(\widetilde{Q_{j\ell}^{(n)}}\right)\right) \leq \epsilon$$

Sketch of proof

- Now the supporting space of the hypothetic states have small overlaps. For $i \neq j$,

$$\mathrm{Olap}\left(\mathrm{supp}\left(\widetilde{\rho_i^{\otimes n}}\right), \mathrm{supp}\left(\widetilde{\rho_j^{\otimes n}}\right)\right) \leq T\epsilon$$



- The next step is to orthogonalize these eigenspaces
  1. Order the eigenspaces according to the their eigenvalues, in the decreasing order.
  2. Orthogonalization using the Gram-Schmidt process.

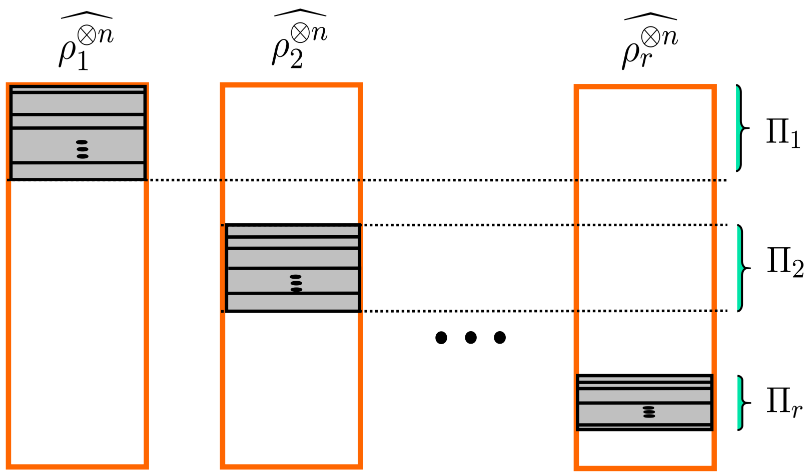
Sketch of proof

- Now the eigenspaces are all orthogonal.

$$\widehat{\rho_i^{\otimes n}} = \bigoplus_{k=1}^{T_i} \lambda_{ik}^{(n)} \widehat{Q_{ik}^{(n)}}$$

- We construct a projective measurement

$$\left\{ \Pi_i = \bigoplus_k \widehat{Q_{ik}^{(n)}} \right\}_{i=1}^{r}$$



- Use this to discriminate the original states:

$$P_{succ} = \sum_{i=1}^{r} p_i \operatorname{Tr} \rho_i^{\otimes n} \Pi_i$$

Sketch of proof

$$Q_{ik}^{(n)} \xrightarrow{\text{“digging holes”}} \widetilde{Q_{ik}^{(n)}} \xrightarrow{\text{orthogonalization}} \widehat{Q_{ik}^{(n)}}$$

- Loss in "digging holes":

$$\mathrm{Tr}\left(Q_{ik}^{(n)} - \widetilde{Q_{ik}^{(n)}}\right) \le \frac{1}{\epsilon^2} \sum_{(j,\ell):\lambda_{j\ell}^{(n)} > \lambda_{ik}^{(n)}} \mathrm{Tr}\, Q_{ik}^{(n)} Q_{j\ell}^{(n)}$$

- Mismatch due to orthogonalization:

$$\mathrm{Tr}\left[\widetilde{Q_{ik}^{(n)}}\left(\mathbb{1} - \widehat{Q_{ik}^{(n)}}\right)\right] \le \frac{1-(r-1)T\epsilon}{1-2(r-1)T\epsilon} \sum_{(j,\ell):\lambda_{j\ell}^{(n)} > \lambda_{ik}^{(n)}} \mathrm{Tr}\, Q_{ik}^{(n)} Q_{j\ell}^{(n)}$$

- Estimation of the total error:

$$P_e \le \sum_{(i,k)} \lambda_{ik}^{(n)} \mathrm{Tr}\left[Q_{ik}^{(n)}\left(\mathbb{1} - \widehat{Q_{ik}^{(n)}}\right)\right] \le \sum_{(i,k)} \lambda_{ik}^{(n)} \left\{\mathrm{Tr}\left(Q_{ik}^{(n)} - \widetilde{Q_{ik}^{(n)}}\right) + \mathrm{Tr}\left[\widetilde{Q_{ik}^{(n)}}\left(\mathbb{1} - \widehat{Q_{ik}^{(n)}}\right)\right]\right\}$$

Sketch of proof

$$P_e \le \underbrace{\left(\frac{1}{\epsilon^2} + \frac{1-(r-1)T\epsilon}{1-2(r-1)T\epsilon}\right)}_{\le p(n)} \sum_{(i,j):i\neq j} \sum_{k,\ell} \underbrace{\min\{\lambda_{ik}^{(n)}, \lambda_{j\ell}^{(n)}\}}_{\le \left(\lambda_{ik}^{(n)}\right)^s \left(\lambda_{j\ell}^{(n)}\right)^{(1-s)}} \operatorname{Tr} Q_{ik}^{(n)} Q_{j\ell}^{(n)}$$

$$\le p(n) \sum_{(i,j):i\neq j} \min_{0\le s\le 1} \left(\operatorname{Tr} \rho_i^s \rho_j^{(1-s)}\right)^n$$

$$\sim \exp\left\{-n\left(\min_{(i,j):i\neq j} \max_{0\le s\le 1} \left\{-\log \operatorname{Tr} \rho_i^s \rho_j^{1-s}\right\}\right)\right\}$$

■

Outline

1. The problem
2. The answer
3. History review
4. Proof sketch
5. One-shot case
6. Open questions

Result for the one-shot case

Theorem *Let $A_1, \ldots, A_r \in \mathcal{P}(\mathcal{H})$ be nonnegative matrices on a finite-dimensional Hilbert space $\mathcal{H}$. For all $1 \le i \le r$, let $A_i = \bigoplus_{k=1}^{T_i} \lambda_{ik} Q_{ik}$ be the spectral decomposition of A_i, and write $T := \max\{T_1, \ldots, T_r\}$. Then*

$$P_e^*(\{A_1, \ldots, A_r\}) \le 10(r-1)^2 T^2 \sum_{(i,j):i<j} \sum_{k,\ell} \min\{\lambda_{ik}, \lambda_{j\ell}\} \operatorname{Tr} Q_{ik} Q_{j\ell}.$$

- Remark 1: It matches a lower bound up to some states-dependent factors:

$$P_e^*(\{A_1, \ldots, A_r\}) \ge \frac{1}{2(r-1)} \sum_{(i,j):i<j} \sum_{k,\ell} \min\{\lambda_{ik}, \lambda_{j\ell}\} \operatorname{Tr} Q_{ik} Q_{j\ell}.$$

Obtained by combining [M. Nussbaum and A. Szkola, Ann. Statist. 37, 1040 (2009)] and [D.-W. Qiu, PRA 77. 012328 (2008)].

Result for the one-shot case

- Remark 2: for the case r=2, we have

$$P_e^*(\{A_1, A_2\}) \leq 10T^2 \sum_{k,\ell} \min\{\lambda_{1k}, \lambda_{2\ell}\} \operatorname{Tr} Q_{1k} Q_{2\ell}.$$

On the other hand, it is proved in [K. Audenaert et al, PRL, 2007] that

$$P_e^*(\{A_1, A_2\}) \leq \min_{0 \leq s \leq 1} \operatorname{Tr} A_1^s A_2^{1-s}.$$

(note that it is always true that

$$\sum_{k,\ell} \min\{\lambda_{1k}, \lambda_{2\ell}\} \operatorname{Tr} Q_{1k} Q_{2\ell} \leq \min_{0 \leq s \leq 1} \operatorname{Tr} A_1^s A_2^{1-s}.$$)

Outline

1. The problem
2. The answer
3. History review
4. Proof sketch
5. One-shot case
6. Open questions

Open questions

1. Applications of the bounds:

$$P_e^*(\{A_1,\ldots,A_r\}) \begin{cases} \leq 10(r-1)^2T^2 \sum_{(i,j):i<j} \sum_{k,\ell} \min\{\lambda_{ik},\lambda_{j\ell}\} \operatorname{Tr} Q_{ik}Q_{j\ell} \\ \geq \frac{1}{2(r-1)} \sum_{(i,j):i<j} \sum_{k,\ell} \min\{\lambda_{ik},\lambda_{j\ell}\} \operatorname{Tr} Q_{ik}Q_{j\ell} \end{cases}$$

2. Strenthening the states-dependent factors

3. Testing composite hypotheses:

$$\rho^{\otimes n} \qquad \text{Vs} \qquad \sum_i q_i \sigma_i^{\otimes n} \;\; (\text{or}, \; \int \sigma^{\otimes n}\, \mathrm{d}\mu(\sigma))$$

K. Audenaert and M. Mosonyi, J. Math. Phys. 55, 102201 (2014).

Brandao, Harrow, Oppenheim and Strelchuk, PRL 115, 050501 (2015).

Thank you !