On the Chernoff distance for asymptotic LOCC discrimination of bipartite quantum states

William Matthews and Andreas Winter

University of Bristol

CQIT 2008 Santa Fe
March 28th

Outline

- Background
  - Classical Chernoff Distance
  - Quantum Chernoff Distance (Global Measurements)
- State Discrimination and Chernoff Distances on Bipartite Systems
- Data Hiding States
  - Single Copy
  - Dimension Dependence of Error
  - Shared Entanglement
- LOCC Chernoff Distances

Classical Chernoff Distance

- Given n i.i.d. samples drawn from one of two probability distributions over an alphabet A: $p(x)$ and $q(x)$ ($x \in A$). Equally likely the distribution p or q is used.
- Guess which distribution has been used based on the n samples.
- Probability of error is $P_{err}(p, q; n) = \frac{1}{2}P(\text{guess q}|n \text{ samples from } p) + \frac{1}{2}P(\text{guess p}|n \text{ samples from } q)$
- Guessing according to maximum likelihood rule minimizes this error probability.

Classical Chernoff Distance

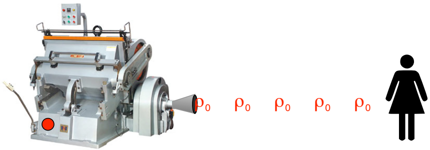
- Large n asymptotic behaviour derived by Chernoff (1952)[1].
- $P_{err}\left(p,q;n\right) \sim 2^{-\xi(p,q)n}$.
- Where $\xi(p,q) = \lim_{n\to\infty}\left(-\frac{1}{n}\log P_{err}(p,q;n)\right)$, is the (classical) Chernoff distance and has the following simple form in terms of the probability distributions:

$$\xi\left(p,q\right) = -\log\left(\min_{0\leq s\leq 1}\sum_{x\in A} p(x)^s q(x)^{1-s}\right)$$

[1] The Annals of Mathematical Statistics, Vol. 23, No. 4, pp. 493-507

Quantum Chernoff Distance

- Source produces copies of state ρ_0 or state ρ_1. What is best asymptotic behaviour of error?

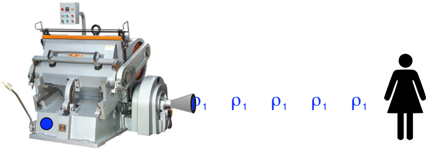


- A decision procedure (for given n) can be written as a two element POVM, $\{M, \mathbb{1} - M\}$; If the outcome corresponding to M occurs guess ρ_1, otherwise guess ρ_0.
- In terms of M:
$P_{err}(\rho_0^{\otimes n}, \rho_1^{\otimes n}; M) = \frac{1}{2}\left(\mathrm{Tr}\left(M\rho_0^{\otimes n}\right) + \mathrm{Tr}\left(\left(\mathbb{1} - M\right)\rho_1^{\otimes n}\right)\right)$.
- Well known that the optimal POVM is the Holevo-Helstrom measurement[2].

[2]C.W. Helstrom, Quantum Detection and Estimation Theory, Academic Press, New York (1976)

Quantum Chernoff Distance

- Source produces copies of state ρ_0 or state ρ_1. What is best asymptotic behaviour of error?



- A decision procedure (for given n) can be written as a two element POVM, $\{M, \mathbb{1} - M\}$; If the outcome corresponding to M occurs guess ρ_1, otherwise guess ρ_0.
- In terms of M:
  $P_{err}(\rho_0^{\otimes n}, \rho_1^{\otimes n}; M) = \frac{1}{2}\left(\mathrm{Tr}\left(M\rho_0^{\otimes n}\right) + \mathrm{Tr}\left(\left(\mathbb{1} - M\right)\rho_1^{\otimes n}\right)\right).$
- Well known that the optimal POVM is the Holevo-Helstrom measurement[3].

[3]C.W. Helstrom, Quantum Detection and Estimation Theory, Academic Press, New York (1976)

Quantum Chernoff Distance

- $P_{err}(\rho_0^{\otimes n}, \rho_1^{\otimes n}) = \frac{1}{2}\left(1 - \frac{1}{2}||\rho_1^{\otimes n} - \rho_0^{\otimes n}||_1\right)$.
- What is the asymptotic dependence on n?
- Answer was only recently discovered. [4]
- $P_{err}\left(\rho_0^{\otimes n}, \rho_1^{\otimes n}\right) \sim 2^{-\xi(\rho_0, \rho_1) n}$.
- Where the Quantum Chernoff distance is:
  $\xi\left(\rho_0, \rho_1\right) = -\log\left(\min_{0 \leq s \leq 1} \mathrm{Tr}\left(\rho_0^s \rho_1^{1-s}\right)\right)$.
- Remarkably straightforward generalization of the classical expression.
- Motivates the question: What happens when the states to be distinguished are distributed between multiple parties?
- I will only talk about the bipartite case here.

---

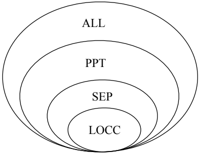
[4] "Asymptotic Error Rates in Quantum Hypothesis Testing", Audenaert et al., arXiv:0708.4282

Classes of Operations on Bipartite Systems

- LOCC: Local Operations and Classical Communication.
- Separable Operations (SEP):

$$L \in \text{SEP} \iff L(\rho) = \sum_i A_i \otimes B_i \rho A_i^\dagger \otimes B_i^\dagger .$$

- PPT Operations (PPT) [5]: $L \in \text{PPT} \iff \Gamma \circ L \circ \Gamma$ is completely positive. Where $\Gamma = \mathbb{1} \otimes T$ is the (linear) partial transpose map.



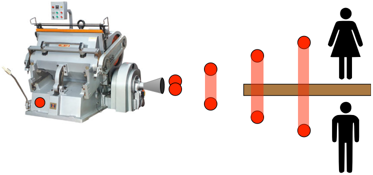
$$\text{LOCC} \subset \text{SEP} \subset \text{PPT} \subset \text{ALL(CPTP)}.$$

[5] E. M. Rains, "A semidefinite program for distillable entanglement", IEEE Trans. Inf. Theory, **47**(7):2921-2933 (2001).

Measurements on Bipartite Systems

- Which measurements can be performed with operations in one of these classes?
- LOCC - hard to characterise.
- A POVM (M_i) can be implemented in SEP iff $M_i = \sum_i X_i \otimes Y_i$.
- A POVM (M_i) can be implemented in PPT iff $M_i^\Gamma \geq 0$.

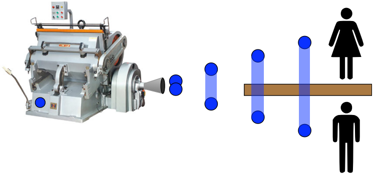
State Discrimination and Chernoff Distances for Bipartite Systems



- $P_{err}^{X}(\rho_0,\rho_1) := \min_{(M,\mathbb{1}-M)\in X} \frac{1}{2}\left(\mathrm{Tr}\left(M\rho_0\right) + \mathrm{Tr}\left(\left(\mathbb{1}-M\right)\rho_1\right)\right).$
- $\xi^{X}(\rho_0,\rho_1) := \lim_{n\to\infty}\left(-\frac{1}{n}\log P_{err}^{X}(\rho_0^{\otimes n},\rho_1^{\otimes n})\right).$
- Containment of classes implies ordering

$$P_{err}^{\mathrm{LOCC}}(\rho_0,\rho_1) \geq P_{err}^{\mathrm{SEP}}(\rho_0,\rho_1) \geq P_{err}^{\mathrm{PPT}}(\rho_0,\rho_1) \geq P_{err}^{\mathrm{ALL}}(\rho_0,\rho_1).$$

State Discrimination and Chernoff Distances for Bipartite Systems



- $P_{err}^{X}(\rho_0, \rho_1) := \min_{(M, \mathbb{1}-M) \in X} \frac{1}{2} \left(\mathrm{Tr}(M \rho_0) + \mathrm{Tr}((\mathbb{1} - M) \rho_1) \right).$
- $\xi^{X}(\rho_0, \rho_1) := \lim_{n \to \infty} \left(-\frac{1}{n} \log P_{err}^{X}(\rho_0^{\otimes n}, \rho_1^{\otimes n}) \right).$
- Containment of classes implies ordering

$$P_{err}^{\mathrm{LOCC}}(\rho_0, \rho_1) \geq P_{err}^{\mathrm{SEP}}(\rho_0, \rho_1) \geq P_{err}^{\mathrm{PPT}}(\rho_0, \rho_1) \geq P_{err}^{\mathrm{ALL}}(\rho_0, \rho_1).$$

State Discrimination and Chernoff Distances for Bipartite Systems

$$\xi^{\mathrm{LOCC}}(\rho_0,\rho_1) \leq \xi^{\mathrm{SEP}}(\rho_0,\rho_1) \leq \xi^{\mathrm{PPT}}(\rho_0,\rho_1) \leq \xi^{\mathrm{ALL}}(\rho_0,\rho_1)$$

Define $\xi^{\mathrm{SC}}(\rho_0,\rho_1)$ to be the *classical* Chernoff distance between the statistics generated by the optimal single copy LOCC measurement:

$$\begin{aligned}\xi^{\mathrm{SC}}(\rho_0,\rho_1) = -\log \min_{0\leq s\leq 1} &\left(\mathrm{Tr}\left(M^*\rho_0\right)^{1-s}\mathrm{Tr}\left(M^*\rho_1\right)^{s}\right. \\ &\left.+\,\mathrm{Tr}\left((\mathbb{1}-M^*)\rho_0\right)^{1-s}\mathrm{Tr}\left((\mathbb{1}-M^*)\rho_1\right)^{s}\right)\end{aligned}$$

Clearly we have the lower bound:

$$\xi^{\mathrm{SC}}(\rho_0,\rho_1) \leq \xi^{\mathrm{LOCC}}(\rho_0,\rho_1).$$

State Discrimination and Chernoff Distances for Bipartite Systems

- ξ^{LOCC} not necessarily $+\infty$ for orthogonal states.
- $\xi^{\mathrm{ALL}}(\rho_0, \rho_1) = \xi^{\mathrm{ALL}}(\rho_0 \otimes \tau, \rho_1 \otimes \tau)$.
  Not always true in the bipartite LOCC case:
  E.g. Let Φ_K denote a maximally entangled state of Schmidt rank K.
  It can be the case that $\xi^{\mathrm{LOCC}}(\rho \otimes \Phi_K, \sigma \otimes \Phi_K) < \xi^{\mathrm{LOCC}}(\rho, \sigma)$, because for some n Alice and Bob will share enough copies of Φ to teleport and apply global measurements.
  If $K \geq$ dimension of states then $\xi^{\mathrm{LOCC}}(\rho \otimes \Phi_K, \sigma \otimes \Phi_K) = \xi^{\mathrm{ALL}}(\rho \otimes \Phi_K, \sigma \otimes \Phi_K) = \xi^{\mathrm{ALL}}(\rho, \sigma)$.

State Discrimination and Chernoff Distances for Bipartite Systems

- For pure states, LOCC can do just as well as global measurements [6]. So, $\xi^{\mathrm{LOCC}}(|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|) = \xi^{\mathrm{ALL}}(|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|)$.
- For mixed states, LOCC can be much worse than global measurements, e.g. 'Data hiding states'[7].
- Calculate LOCC Chernoff distance for states which give different behaviour from global measurements.

---

[6] Walgate et al. Phys. Rev. Lett. 8(23):4972-4975 (2000); (quant-ph/0007098); Virmani et al. Phys. Lett. A. 288, p.62 (quant-ph/0102073)

[7] DiVincenzo et al. Information Theory, IEEE Transactions on, Vol.48, Iss.3, Mar 2002 Pages 580-598 (quant-ph/0103098)

Strategy

- Finding $P_{err}^{\mathrm{PPT}}(\rho_0, \rho_1)$ is a semidefinite programming problem:

$$P_{err}^{\mathrm{PPT}}(\rho_0, \rho_1) = \min \mathrm{Tr}\, \frac{1}{2} \left((M \rho_0) + \mathrm{Tr} \left((\mathbb{1} - M) \, \rho_1 \right) \right)$$

$$M \geq 0, \mathbb{1} - M \geq 0, M^{\Gamma} \geq 0, (\mathbb{1} - M)^{\Gamma} \geq 0$$

- Feasible points of the *dual* SDP provide lower bounds on $P_{err}^{\mathrm{PPT}}(\rho_1, \rho_2)$.
- Guess dual optimal solution.
- Guess LOCC protocol which matches the lower bound.
- If we can do this then we have shown that this protocol is optimal.

Strategy

- Generally quite hard to do this.
- Use symmetries which are: Shared by ρ_1 and ρ_2 and generated by LOCC.
- In the cases we shall look at this simplifies the problem to a linear program.

Data Hiding States

$$\sigma_d = \frac{2}{d(d+1)}\mathcal{S}_d \in B(\mathbb{C}^d \otimes \mathbb{C}^d)$$

$$\alpha_d = \frac{2}{d(d-1)}\mathcal{A}_d \in B(\mathbb{C}^d \otimes \mathbb{C}^d)$$

- These are the extremal $d \times d$ Werner states.
- Invariant under bi-unitary transformations: $U \otimes U$.
- A generalization of the data hiding states of DiVincenzo *et al.*
- Orthogonal, and therefore perfectly distinguishable globally, but...
- hard to distinguish using LOCC.

Data Hiding States

Let F_d denote the flip operator on $\mathbb{C}^d \otimes \mathbb{C}^d$:
$F_d|\psi\rangle_A \otimes |\phi\rangle_B = |\phi\rangle_A \otimes |\psi\rangle_B$.

$$\Phi_K^\Gamma = \frac{1}{d}(\sum_{i,j} |i\rangle\langle j| \otimes |i\rangle\langle j|)^\Gamma = \frac{1}{d}\sum_{i,j} |i\rangle\langle j| \otimes |j\rangle\langle i| = \frac{1}{d}F.$$

$$\mathcal{S}_d = (\mathbb{1} + F_d)/2, \mathcal{A}_d = (\mathbb{1} - F_d)/2$$

$$\mathcal{S}_d^\Gamma = (\mathbb{1} + d\Phi_d)/2, \mathcal{A}_d^\Gamma = (\mathbb{1} - d\Phi_d)/2$$

Data Hiding: Single Copy Linear Program

- POVM elements can be written as linear combinations of $\mathcal{S}_d$ and $\mathcal{A}_d$: $M = x_0 \mathcal{S}_d + x_1 \mathcal{A}_d$.
- Noting that $\left(x_0 \mathcal{S}_d + x_1 \mathcal{A}_d\right)^\Gamma = \frac{1}{2}\left(\left(\mathbb{1} - \Phi_d\right), \Phi_d\right) \begin{pmatrix} 1 & 1 \\ d+1 & 1-d \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}$, we have

$$P_{err}^{\mathrm{PPT}}\left(\sigma_d, \alpha_d\right) = \min\left(1 + x_0 - x_1\right)$$

subject to

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} \leq \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \leq \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} \leq \frac{1}{2} \begin{pmatrix} 1 & 1 \\ d+1 & 1-d \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \leq \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

- There is a dual feasible point where the dual objective is $\frac{1}{2}\left(\frac{d-1}{d+1}\right)$, so

$$P_{err}^{\mathrm{PPT}}\left(\sigma_d, \alpha_d\right) \geq \frac{1}{2}\left(\frac{d-1}{d+1}\right).$$

Data Hiding: Single Copy LOCC Protocol

- Alice and Bob both measure in the computational basis, obtaining outcomes a and b from $\{1, ..., d\}$, respectively.
- If $a \neq b$, then they guess α_d.
- If $a = b$, then they guess σ_d.
- The wrong guess is made with probability

$$P^*_{err}(\sigma_d, \alpha_d) = \frac{1}{2}P(a \neq b|\sigma_d) + \frac{1}{2}P(a = b|\alpha_d) = \frac{1}{2}\left(\frac{d-1}{d+1}\right) + 0$$

- This achieves the lower bound for PPT, so:

$$P^{\mathrm{LOCC}}_{err}(\sigma_d, \alpha_d) = P^{\mathrm{PPT}}_{err}(\sigma_d, \alpha_d) = \frac{1}{2}\left(\frac{d-1}{d+1}\right).$$

Data Hiding: Single Copy LOCC Bias Dimension Dependence - Worst Case?

- By increasing d we can make the achievable bias arbitrarily small:

$$B^{\mathrm{LOCC}}(\sigma_d, \alpha_d) = 1 - 2P_{err}^{\mathrm{LOCC}}(\sigma_d, \alpha_d) = \frac{2}{d+1} \sim \frac{1}{d}.$$

- Is this the worst (best) possible dimension dependence?
- For separable measurements it is...

Generic Separable Measurement

- Barnum and Gurvits [8]: Every operator in the ball of radius one in H-S norm centred on the identity is separable.
- Take Holevo-Helstrom POVM $(M, \mathbb{1} - M)$ and mix in just enough identity with the elements to ensure that they are in this ball.

$$\left(\frac{1}{2}\left(\mathbb{1} + \frac{M}{\|M\|_2}\right), \frac{1}{2}\left(\mathbb{1} - \frac{M}{\|M\|_2}\right)\right).$$

- Using $\|M\|_2 \leq \sqrt{D}$, (where D is total dimension of the system), we find

$$B^{\mathrm{SEP}} \geq \frac{1}{2\sqrt{D}} B^{\mathrm{ALL}}.$$

[8] H. Barnum, L. Gurvits, "Largest separable balls around the maximally mixed bipartite quantum state", Phys. Rev. A **66**, 062311 (2002)

Data Hiding: Single Copy + Shared Entanglement

- What if Alice and Bob share a maximally entangled state of Schmidt rank $K \leq d$?
- Φ_K has $U \otimes \bar{U}$ invariance.
- $M = x.(\mathcal{S}_d \otimes \Phi_K, \mathcal{A}_d \otimes (\mathbb{1} - \Phi_K), \mathcal{A}_d \otimes \Phi_K, \mathcal{A}_d \otimes (\mathbb{1} - \Phi_K))$.
- Again, we can simplify to a linear program.
- A dual feasible point can be found yielding the bound

$$P_{err}^{\mathrm{PPT}}(\sigma_d \otimes \Phi_K, \alpha_d \otimes \Phi_K) \geq \frac{1}{2}\left(\frac{d-K}{d+1}\right).$$

Data Hiding: Single Copy + Shared Entanglement

- Again, this bound can be achieved by an LOCC protocol:
- Alice performs the POVM $\left(\Pi^A_{K,j}/K\right)_{j=1,\ldots,d}$ on her half of the data hiding state, where $\Pi^A_{K,j} = \sum_{m=0}^{K-1} |j \oplus m\rangle\langle j \oplus m|$, and tells Bob the outcome j.
- Bob does the projective measurement $\left(\Pi^B_{K,j}, \mathbb{1} - \Pi^B_{K,j}\right)$ on his half of the data-hiding state.
- If the first outcome occurs, the resulting state is the completely symmetric or anti-symmetric Werner state on a $K \times K$ subspace. Bob teleports his half to Alice with the entangled state and Alice identifies it without error.
- If the second outcome occurs, they guess that they have the anti-symmetric Werner state.

Data Hiding: Linear Program for Many Copies

- $U \otimes U$ invariance on each copy - SDP to LP.
- Invariance under permutations of copies - 2^n variables to $n+1$ variables.
- The dual linear program has a feasible point which gives the bound

$$P_{err}^{\mathrm{PPT}}\left(\sigma^{\otimes n}, \alpha^{\otimes n}\right) \geq \frac{1}{2}\left(\frac{d-1}{d+1}\right)^{n}.$$

Data Hiding: Protocol for Many Copies

- The following protocol achieves the PPT bound:
  - Alice and Bob take each copy separately and measure in the computational basis, obtaining on the i^{th} copy the outcomes a_i and b_i from $\{1, ..., d\}$.
  - If $a_i \neq b_i$ for all i, then they guess α_d.
  - If $a_i = b_i$ for some i, then they guess σ_d.

$$P^*_{err}(\sigma_d, \alpha_d; n) = \frac{1}{2} P(\forall i : a_i \neq b_i | \sigma_d^{\otimes n}) + \frac{1}{2} P(\exists i : a_i = b_i | \alpha_d^{\otimes n}).$$

$$P^*_{err}(\sigma_d, \alpha_d; n) = \frac{1}{2} \left(\frac{d-1}{d+1} \right)^n .$$

Data Hiding: LOCC Chernoff Distance

- Whereas $P^{\mathrm{ALL}}\left(\sigma_d^{\otimes n}, \alpha_d^{\otimes n}\right) = 0$

$$\begin{aligned} P^{\mathrm{PPT}}\left(\sigma_d^{\otimes n}, \alpha_d^{\otimes n}\right) =& P^{\mathrm{SEP}}\left(\sigma_d^{\otimes n}, \alpha_d^{\otimes n}\right) = P^{\mathrm{LOCC}}\left(\sigma_d^{\otimes n}, \alpha_d^{\otimes n}\right) \\ =& \frac{1}{2}\left(\frac{d-1}{d+1}\right)^n. \end{aligned}$$

- So, we have,

$$\begin{aligned} \xi^{\mathrm{PPT}}(\sigma_d, \alpha_d) =& \xi^{\mathrm{SEP}}(\sigma_d, \alpha_d) = \xi^{\mathrm{LOCC}}(\sigma_d, \alpha_d) \\ =& \xi^{\mathrm{SC}}(\sigma_d, \alpha_d) = \log \frac{d+1}{d-1}. \end{aligned}$$

- It is notable that we do not need joint measurements to achieve the optimal result.

LOCC Chernoff Distance for Extremal Isotropic States

- $\Phi_d^{\perp} := \frac{\mathbb{1} - \Phi_d}{d^2 - 1}$.
- $U \otimes \bar{U}$ invariance.
- Copy permutation invariance.
- Again, we can use the dual SDP for $P_{err}^{\mathrm{PPT}}\left(\Phi_d, \Phi_d^{\perp}\right)$ show that the following protocol is optimal:
  - Alice and Bob measure each copy in the computational basis.
  - If for every copy they get the same result then they guess that they have n copies of Φ_d, otherwise they know that they have $\Phi_d^{\perp}$.

LOCC Chernoff Distance for Extremal Isotropic States

- Similar to the extremal Werner state case, all of the non-global min. errors are equal

$$\begin{aligned} P_{err}^{\mathrm{LOCC}}\left(\Phi_d, \Phi_d^{\perp}\right) =& P_{err}^{\mathrm{SEP}}\left(\Phi_d, \Phi_d^{\perp}\right) = P_{err}^{\mathrm{PPT}}\left(\Phi_d, \Phi_d^{\perp}\right) \\ =& \frac{1}{2}\frac{1}{(d+1)^n}. \end{aligned}$$

- Again there is an optimal many-copy measurement which can be performed one copy at a time.

$$\begin{aligned} \xi^{\mathrm{LOCC}}\left(\Phi_d, \Phi_d^{\perp}\right) =& \xi^{\mathrm{SEP}}\left(\Phi_d, \Phi_d^{\perp}\right) = \xi^{\mathrm{PPT}}\left(\Phi_d, \Phi_d^{\perp}\right) \\ =& \xi^{\mathrm{SC}}\left(\Phi_d, \Phi_d^{\perp}\right) = \log\left(d+1\right). \end{aligned}$$

In Summary

- Dimensional dependence of bias is optimal for separable measurements - is it for LOCC?
- Data hiding property fails gradually in the presence of shared entanglement.
- Optimal LOCC protocols determined for discriminating between the extremal Werner states and between the extremal isotropic states, when n copies are available.
- $\xi^{\mathrm{LOCC}}(\sigma_d, \alpha_d) = \xi^{\mathrm{SC}}(\sigma_d, \alpha_d) = \log \frac{d+1}{d-1}$.
- $\xi^{\mathrm{LOCC}}\left(\Phi_d, \Phi_d^{\perp}\right) = \xi^{\mathrm{SC}}\left(\Phi_d, \Phi_d^{\perp}\right) = \log\left(d+1\right)$.
- Thank you.

WM acknowledges support of EPSRC. Authors would like to acknowledge useful discussions with Keiji Matsumoto, Chris King and Mike Nathanson.
arXiv:0710.4113