Norms on quantum states...
... and quantum data hiding

Andreas Winter : University of Bristol
+
NUS (CQT)

[with Will Matthews & Stephanie Wehner]

arXiv [quant-ph]: 0810.2327

Hypothesis testing is at the heart of information processing:

- distinguishing signals in communication
- detect cheating players in cryptographic games
- hide information from eavesdroppers
- statistical applications

⋮

Basic case of 2 hypotheses ≡ statistical scenarios

Probability distribution

P_0 or P_1

on a sample space Ω

Basic case of 2 hypotheses ≡ statistical scenarios

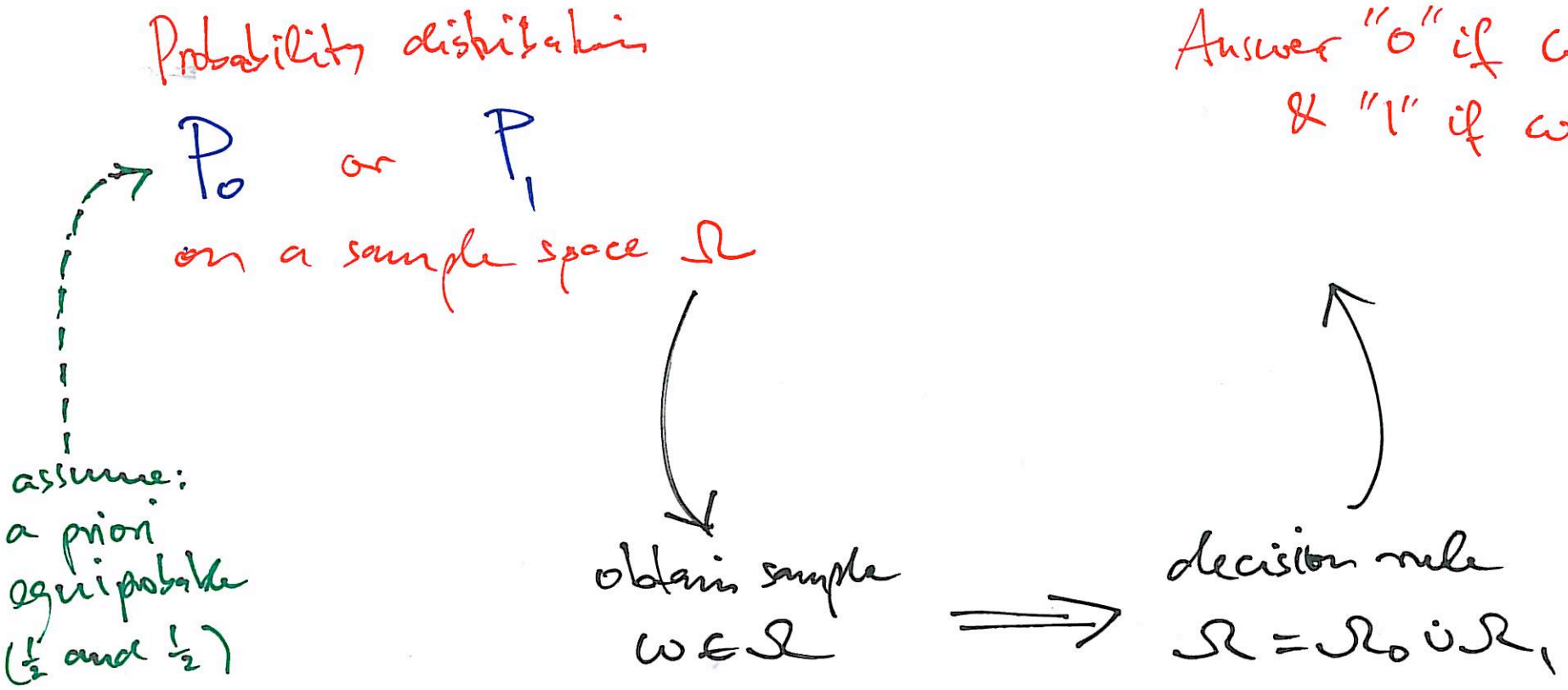
Probability distribution

P_0 or P_1

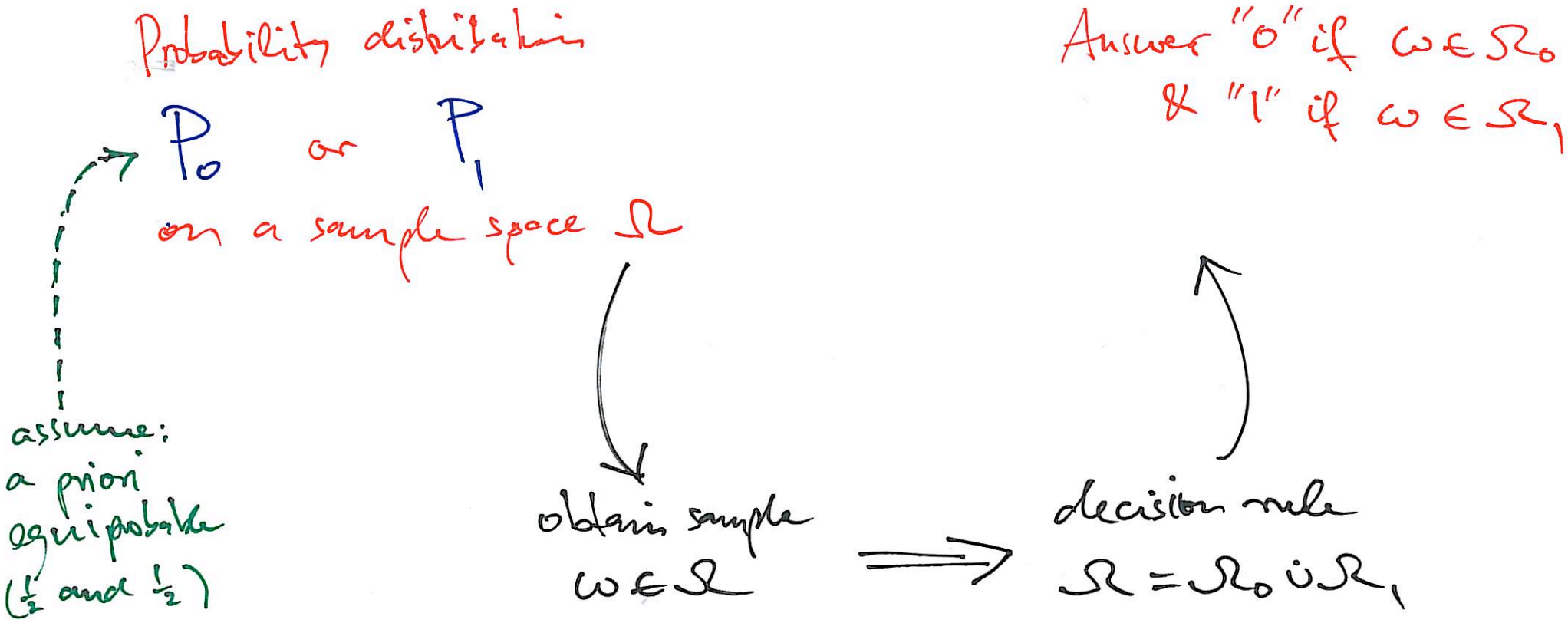
on a sample space Ω

assume: a priori equiprobable ($\frac{1}{2}$ and $\frac{1}{2}$)

Basic case of 2 hypotheses ≡ statistical scenarios



Basic case of 2 hypotheses ≡ statistical scenarios



Error probability $P_{err} = \frac{1}{2} Pr\{"0" | P_1\} + \frac{1}{2} Pr\{"1" | P_0\}$

$= \frac{1}{2} P_1(\Omega_0) + \frac{1}{2} P_0(\Omega_1)$

$$P_{err} = \frac{1}{2} P_1(\Omega_0) + \frac{1}{2} P_0(\Omega_1)$$

$$= \frac{1}{2} + \frac{1}{2}(P_1 - P_0)(\Omega_0)$$

Minimized when $\Omega_0 = \{\omega : P_0(\omega) \geq P_1(\omega)\}$

["maximum likelihood decision" (MLD)]

$$P_{err} = \frac{1}{2} P_1(\Omega_0) + \frac{1}{2} P_0(\Omega_1)$$

$$= \frac{1}{2} + \frac{1}{2}(P_1 - P_0)(\Omega_0)$$

Minimized when $\Omega_0 = \{\omega : P_0(\omega) \geq P_1(\omega)\}$

["maximum likelihood decision" (MLD)]

$$\Rightarrow \min P_{err} = \frac{1}{2} - \frac{1}{4}\|P_0 - P_1\|_1$$

$$\|P_0 - P_1\|_1 = \sum_\omega |P_0(\omega) - P_1(\omega)|$$

ℓ^1-distance (norm!)

[$\frac{1}{2}\|P_0 - P_1\|_1$ aka "Kolmogorov distance"]

Quantum case: two hypotheses associated with quantum states, ρ_0 or ρ_1, on a system with Hilbert space $\mathcal{H}$

state (in general mixed):
density operator ρ on $\mathcal{H}$, i.e.
$\rho = \rho^\dagger \geq 0$, $\mathrm{tr}\,\rho = 1$

Quantum case: two hypotheses associated with quantum states, ρ_0 or ρ_1, on a system with Hilbert space H

state (in general mixed):
density operator ρ on H, i.e.
$\rho = \rho^\dagger \geq 0$, $\mathrm{tr}\,\rho = 1$

New problem: need to make a measurement to obtain data for the decision ⟹ some measurements not as informative as others...

Recall: quantum measurement ≡ positive operator valued measure (POVM) ≡ partition of $\mathbb{1}$

$$\mathbb{1} = \sum_{\omega \in \Omega} M_\omega \text{ s.t. } M_\omega = M_\omega^\dagger \geq 0$$

Measurement is probabilistic: $\Pr\{\omega \mid e\} = \mathrm{Tr}\, e M_\omega$ [Born!]

Recall: quantum measurement ≡ positive operator valued measure (POVM) ≡ partition of $\mathbb{1}$

$$\mathbb{1} = \sum_{\omega \in \Omega} M_\omega \text{ s.t. } M_\omega = M_\omega^\dagger \geq 0$$

Measurement is probabilistic: $\Pr\{\omega | \rho\} = \mathrm{Tr}\, \rho M_\omega$ [Born!]

Here: want two-outcome measurement → decision rule $\Omega = \Omega_0 \dot\cup \Omega_1$ leads to quantum decision rule

$$M_0 = \sum_{\omega \in \Omega_0} M_\omega$$

$$M_1 = \sum_{\omega \in \Omega_1} M_\omega = \mathbb{1} - M_0$$

C. Helstrom 1969 (see book 1976): quantum MLD

$$p_{err} = \frac{1}{2} \mathrm{Tr}\, \rho_0 M_1 + \frac{1}{2} \mathrm{Tr}\, \rho_1 M_0$$

$$= \frac{1}{2} + \frac{1}{2} \mathrm{Tr} \left[(\rho_1 - \rho_0) M_0 \right]$$

Hermitian, so has spectral decomposition

$$\sum_i \lambda_i P_i$$

P_i ← eigenprojections

λ_i ← eigenvalues

C. Helstrom 1969 (see book 1976): quantum MLD

$$p_{err} = \frac{1}{2}\mathrm{Tr}\,\rho_0 M_1 + \frac{1}{2}\mathrm{Tr}\,\rho_1 M_0$$

$$= \frac{1}{2} + \frac{1}{2}\mathrm{Tr}\left[(\rho_1 - \rho_0) M_0\right]$$

Hermitian, so has spectral decomposition $\sum_i \lambda_i P_i$ ← eigenprojections; ← eigenvalues

Minimized when M_0 = negative part of $\rho_1 - \rho_0$

$$= \sum_{i \text{ s.t. } \lambda_i < 0} P_i := \{\rho_1 - \rho_0 < 0\}$$

$$\Rightarrow \min_{\text{POVMs}} p_{err} = \frac{1}{2} - \frac{1}{4}\|\rho_0 - \rho_1\|_1$$

ℓ^1/trace norm: $\|\xi\|_1 = \mathrm{Tr}\,|\xi| = \mathrm{Tr}\sqrt{\xi^\dagger \xi}$

This paper/talk: What if we don't have access to all measurements?

→ characterize loss of distinguishability

Motivation:

- Reality of the lab
  (... sometimes there's only one experiment, a "full tomography" available)
- System distributed spatially & only "local" operations allowed/available
  (→ later)

Outline for rest of the talk:

- Every set of POVMs defines a norm on states
- Examples & questions to ask
- Comparing the norms: "Constants of domination"
  - → tomographic measurements
  - → data hiding

Consider some fixed set of POVMs : $\mathbf{M} = \{(M_\omega)_{\omega \in \Omega}\}$

Then: $\|\alpha\|_{\mathbf{M}} := \sup_{(M_\omega)_\omega \in \mathbf{M}} \sum_\omega |\mathrm{Tr}\, \alpha M_\omega|$

defines a norm on density operators

Consider some fixed set of POVMs : $\mathbf{M} = \{(M_\omega)_{\omega \in \Omega}\}$

Then: $\|\alpha\|_{\mathbf{M}} := \sup_{(M_\omega)_\omega \in \mathbf{M}} \sum_\omega |\mathrm{Tr}\, \alpha M_\omega|$

defines a norm on density operators

[Assuming the set of all M_ω is separable,
i.e. $\forall \alpha \neq 0 \;\; \exists (M_\omega)_\omega \in \mathbf{M} \;\; \exists \omega \;\; \mathrm{Tr}\, \alpha M_\omega \neq 0$]

Notes: * Can do the maximization

$$\|\alpha\|_{\mathbf{M}} = \sup_{(M_\omega)\in\mathbf{M}} \sum_\omega |\mathrm{Tr}\,\alpha M_\omega|$$

as well as all 2-outcome coarse-grainings of $\mathbf{M}$:

$$\mathbf{M}_2 := \Big\{ (M_0, M_1) : \exists (M_\omega)_{\omega\in\Omega} \in \mathbf{M}\ \exists \Omega = \Omega_0 \cup \Omega_1 \text{ s.t. } M_i = \sum_{\omega\in\Omega_i} M_\omega \Big\}$$

* $\mathbb{M} :=$ closure of $\mathrm{conv}\{ 2M_0 - \mathbb{1} : (M_0, M_1) \in \mathbf{M}_2 \}$ is a symmetric convex body contained in $[-\mathbb{1}, \mathbb{1}]$ (w. operator order $A \le B :\Leftrightarrow B - A \ge 0$)

It is exactly the polar of the unit ball of $\|\cdot\|_{\mathbf{M}}$ under the inner product $\mathrm{Tr}\,\alpha X$ ...

* Clearly, for discriminating ρ from σ using only $(M_w)_{w \in \mathbb{R}} \in \mathcal{M}$:

$$\max p_{\text{err}} = \frac{1}{2} - \frac{1}{4} \|\rho - \sigma\|_{\mathcal{M}}$$

* Clearly, for discriminating ℓ from σ using only $(M\omega)_{\omega \in \Omega} \in \mathbb{M}$:

$$\min p_{err} = \frac{1}{2} - \frac{1}{4}\|\ell - \sigma\|_{\mathbb{M}}$$

* Equally clearly, in general $\|\ell - \sigma\|_{\mathbb{M}} \leq \|\ell - \sigma\|_1$

(and typically strict $<$)

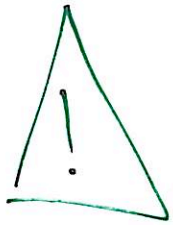
How much smaller can it be, for given $\mathbb{M}$?

Example 1: A single, "tomographically complete" POVM

$(M_\omega)_{\omega \in \Omega}$ [Requires that the M_ω span the set of all Hermitian operators on H!]

{ Several are known:
* complete sets of mutually unbiased bases
* "symmetric informationally complete POVMs"
* $U(H)$-invariant continuous POVM
⋮ }

Example 2: $H = A \otimes B$, $LOCC = \{ (M_\omega)_{\omega \in \Omega}$ can be implemented by multi-round protocol of local operations & classical comm. $\}$



Unlike classical statistics, in quantum mechanics LOCC does not contain all observables

Of course, all norms on finite-dim. normed spaces are equivalent. If $\dim H < \infty$, then exists $\lambda, \mu > 0$ s.t.

$$\lambda \|\alpha\|_1 \leq \|\alpha\|_M \leq \mu \|\alpha\|_1$$

Of course, all norms on finite-dim. normed spaces are equivalent. If $\dim H < \infty$, then exists $\lambda_M > 0$ s.t.

$$\lambda \|\alpha\|_1 \leq \|\alpha\|_M \leq \mu \|\alpha\|_1$$

largest such λ: $\lambda(M)$ — "constants of domination"
smallest such μ: $\mu(M)$ — of $\|\cdot\|_M$ w.r.t. $\|\cdot\|_1$

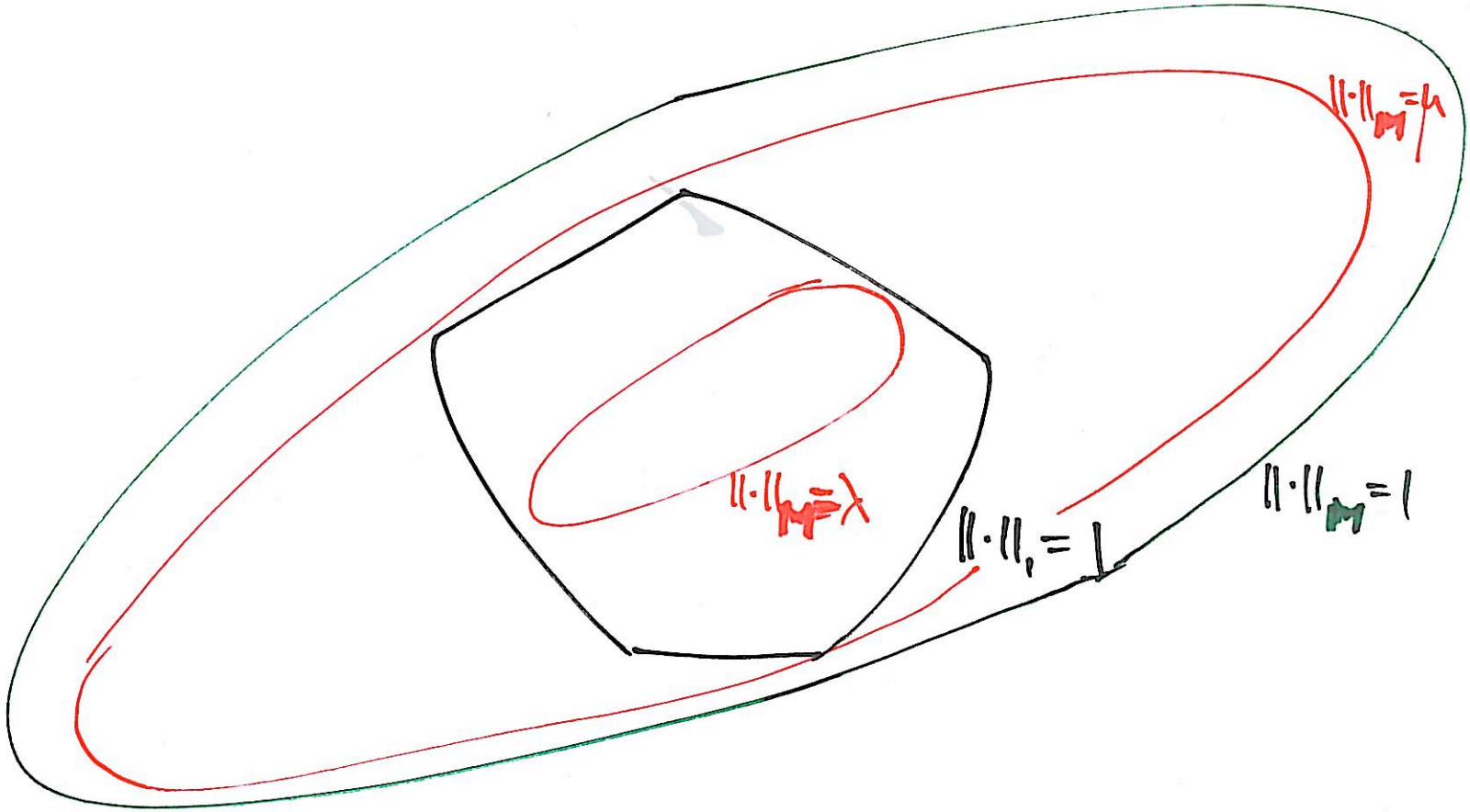
Of course, all norms on finite-dim. normed spaces are equivalent. If $\dim \mathcal{H} < \infty$, then exists $\lambda, \mu > 0$ s.t.

$$\lambda \|\alpha\|_1 \leq \|\alpha\|_M \leq \mu \|\alpha\|_1$$

largest such λ: $\lambda(M)$ — "constants of domination"
smallest such μ: $\mu(M)$ — of $\|\cdot\|_M$ w.r.t. $\|\cdot\|_1$

Note: We restrict to $\mathrm{Tr}\,\alpha = 0$ here! ($\leftrightarrow \alpha = \rho - \sigma$ for state discrimination)

Geometric illustration of constants of domination:



Significance: For orthogonal states $\rho \perp \sigma$ (i.e. error-free distinguishable under all POVMs),

$$\frac{1}{2} - \frac{1}{2}\mu(\mathbf{M}) \leq p_{\text{err}} \leq \frac{1}{2} - \frac{1}{2}\lambda(\mathbf{M})$$

for measurements in $\mathbf{M}$.

Significance: for orthogonal states $\rho \perp \sigma$ (i.e. error-free distinguishable under all POVMs),

$$\frac{1}{2} - \frac{1}{2}\mu(\mathbb{M}) \leq p_{err} \leq \frac{1}{2} - \frac{1}{2}\lambda(\mathbb{M})$$

for measurements in $\mathbb{M}$.

I.e., $\frac{1}{2}\|\rho - \sigma\|_{\mathbb{M}}$ is the best distinguishability bias in guessing the correct state — and $\lambda(\mathbb{M})$, $\mu(\mathbb{M})$ are the worst, best case over all $\rho \perp \sigma$.

Example 1 — single POVM (in $d = \dim \mathcal{H}$):

First result: the largest/smallest λ/p for all single measurements are $\lambda_{max} = \sqrt{\frac{2}{\pi d}}\,(H \circ (1))$

$p_{min} = \frac{1}{2}$

... Attained at the unitary invariant POVM:

$$\left(d \cdot d\psi\, |\psi\rangle\langle\psi| \right)_{|\psi\rangle \in \mathcal{H}}$$

unit vector

But what about more manageable (= fewer outcomes) measurements?

(a) $\left(\frac{1}{d+1} |\varphi_i^{(k)}\rangle\langle\varphi_i^{(k)}| \right)_{\substack{i=1,\dots,d \\ k=0,\dots,d}}$ POVM of $d+1$

"mutually unbiased bases" [see e.g. Wootters/Fields, Ann. Phys. 1989]

$$\left\{ \text{I.e.} \quad |\langle \varphi_i^{(k)} | \varphi_j^{(l)} \rangle|^2 = \begin{cases} \frac{1}{d} & k \neq l \\ \delta_{ij} & k = l \end{cases} \right\}$$

(a) $\left(\frac{1}{d+1}|\varphi_i^{(k)}\rangle\langle\varphi_i^{(k)}|\right)_{\substack{i=1,\dots,d \\ k=0,\dots,d}}$ POVM of $d+1$

"mutually unbiased bases" [see e.g. Wootters/Fields, Ann. Phys. 1989]

$$\left\{ \text{I.e.} \quad |\langle\varphi_i^{(k)}|\varphi_j^{(l)}\rangle|^2 = \begin{cases} \frac{1}{d} & k \neq l \\ \delta_{ij} & k = l \end{cases} \right\}$$

&

$\left(\frac{1}{d}|\chi_i\rangle\langle\chi_i|\right)_{i=1..d^2}$ SIC-POVM [see e.g. Flammia, quant-ph/0605050 or Appleby et al. 0707.2074]

$$\left\{ \text{I.e.} \quad |\langle\chi_i|\chi_j\rangle|^2 = \begin{cases} \frac{1}{d+1} & i \neq j \\ 1 & i = j \end{cases} \right\}$$

... are both examples of (complex projective) 2-designs

These are POVMs $\left(d p_i |\varphi_i\rangle\langle\varphi_i|\right)_{i=1,\dots,n}$

s.t. $\sum_{i=1}^{n} p_i |\varphi_i\rangle\langle\varphi_i| = \frac{1}{d}\mathbb{1}$

and $\sum_{i=1}^{n} p_i |\varphi_i\rangle\langle\varphi_i|^{\otimes 2} = \frac{2}{d(d+1)} \Pi_{\text{sym}}^{d\times d}$

[see e.g. Scott, J. Phys. A 2006]

... are both examples of (complex projective) 2-designs

These are POVMs $\left(d\,p_i\,|\varphi_i\rangle\langle\varphi_i|\right)_{i=1,\dots,n}$

s.t. $\sum_{i=1}^{n} p_i\,|\varphi_i\rangle\langle\varphi_i| = \frac{1}{d}\mathbb{1}$

and $\sum_{i=1}^{n} p_i\,|\varphi_i\rangle\langle\varphi_i|^{\otimes 2} = \frac{2}{d(d+1)}\Pi_{\mathrm{sym}}^{d\times d}$

[see e.g. Scott, J. Phys. A 2006]

For 2-design POVMs, can show:

$$\Omega\left(\frac{1}{d}\right) \leq \lambda \leq O\left(\frac{1}{d}\right)$$

for some 2-designs, e.g. MUBs + SIC-POVMs ...

(b) By a result of Ambainis/Emerson, Proc. CCC '07:

If $(d\,p_i\,|\varphi_i\rangle\langle\varphi_i|)_{i=1\dots n}$ is a 4-design,

[I.e. $\sum_{i=1}^{n} p_i\,|\varphi_i\rangle\langle\varphi_i|^{\otimes 4} = \frac{24}{(d+3)(d+2)(d+1)d}\,\Pi^{d\times d\times d\times d}_{\mathrm{sym}}$]

then $\lambda \geq \frac{1}{3\sqrt{d}}$ — up to a small factor the same as for the uniform POVM $(d\cdot d\,|\psi\rangle\langle\psi|)_{|\psi\rangle}$!

(b) By a result of Ambainis/Emerson, Proc. CCC '07:

If $(d p_i |\varphi_i\rangle\langle\varphi_i|)_{i=1\dots n}$ is a 4-design,

[i.e. $\sum_{i=1}^{n} p_i |\varphi_i\rangle\langle\varphi_i|^{\otimes 4} = \frac{24}{(d+3)(d+2)(d+1)d} \Pi_{sym}^{d\times d\times d\times d}$]

then $\lambda \geq \frac{1}{3\sqrt{d}}$ — up to a small factor the same as for the uniform POVM $(d \cdot d\psi\, |\psi\rangle\langle\psi|)_{(\psi)}$!

(Open question:
performance of 3-designs?)

How it is done: to lower bound

$\sum_i d_{p_i} \left| \mathrm{Tr}\left(\xi |\varphi_i\rangle\langle\varphi_i| \right) \right|$, they use this nice inequality: $\mathbb{E}|S| \geq \sqrt{\frac{(\mathbb{E}S^2)^3}{\mathbb{E}S^4}}$ (proved by Hölder inequality)

How it is done: to lower bound

$\sum_i d_{p_i} \left| \mathrm{Tr}\left(\xi |\varphi_i\rangle\langle\varphi_i| \right) \right|$, they use this nice inequality: $\mathbb{E}|S| \geq \sqrt{\frac{(\mathbb{E}S^2)^3}{\mathbb{E}S^4}}$ (proved by Hölder inequality)

↑ Modulus: difficult

↑ 2nd & 4th moments: can be easier

How it is done: to lower bound

$$\sum_i d_{p_i} \left| \mathrm{Tr}\left(\xi |\varphi_i\rangle\langle\varphi_i| \right) \right|$$, they use this nice inequality : $\mathbb{E}|S| \geq \sqrt{\frac{(\mathbb{E}S^2)^3}{\mathbb{E}S^4}}$ (proved by Hölder inequality)

↑ Modulus: difficult

↑ 2nd & 4th moments: can be easier

... then use 4-design property (≡ up to 4th moments same as under $\mathcal{U}(H)$-inv. averaging)

and representation theory of $\mathcal{U}^{\otimes 4}$...

Example 2 — LOCC (in $d^2 = \dim H = \dim A \cdot \dim B$):

Recall: all measurements that can be built iteratively by local operations and classical communication

For $\varrho = \frac{2}{d(d-1)} \Pi^{d\times d}_{\text{anti}}$ & $\sigma = \frac{2}{d(d+1)} \Pi^{d\times d}_{\text{sym}}$ it is shown by Terhal/DiVincenzo/Leung, PRL 2001 (see also Matthews/AW, arXiv:0710.4113)

that $\| \varrho - \sigma \|_{\text{LOCC}} = \frac{4}{d+1}$

Example 2 — LOCC (in $d^2 = \dim H = \dim A \cdot \dim B$):

Recall: all measurements that can be built iteratively by local operations and classical communication

For $\varrho = \frac{2}{d(d-1)} \Pi^{d\times d}_{\text{anti}}$ & $\sigma = \frac{2}{d(d+1)} \Pi^{d\times d}_{\text{sym}}$ it is shown by Terhal/DiVincenzo/Leung, PRL 2001 (see also Matthews/AW, arXiv:0710.4113)

that $\| \varrho - \sigma \|_{\text{LOCC}} = \frac{4}{d+1}$

"Data hiding": despite being orthogonal, e.g. look almost the same for LOCC-meas. (cf. secret sharing...)

Example 2 — LOCC (in $d^2 = \dim H = \dim A \cdot \dim B$):

Recall: all measurements that can be built iteratively by local operations and classical communication

For $\varrho = \frac{2}{d(d-1)} \Pi^{d\times d}_{\text{anti}}$ & $\sigma = \frac{2}{d(d+1)} \Pi^{d\times d}_{\text{sym}}$ it is shown by Terhal/DiVincenzo/Leung, PRL 2001 (see also Matthews/AW, arXiv:0710.4113)

that $\| \varrho - \sigma \|_{\text{LOCC}} = \frac{4}{d+1}$

"Data hiding": despite being orthogonal, pair of states almost the same for LOCC-meas. (cf. secret sharing...)

$\Rightarrow \lambda(\text{LOCC}) \leq \frac{2}{d+1}$

Notes: * Thanks to Wolgate et al., PRL 2000, we know that if ρ & σ are pure states, then

$$\|\rho-\sigma\|_{LOCC} = \|\rho-\sigma\|_1$$

Notes: * Thanks to Wolgate et al., PRL 2000, we know that if ρ & σ are pure states, then

$$\|\rho-\sigma\|_{LOCC} = \|\rho-\sigma\|_1$$

→ explains why for security in data hiding we need mixedness

→ compare to secret sharing: need randomness to hide information from local access

↑

[à la Shamir, Comm. ACM 1979]

Notes: * Thanks to Wolgate et al., PRL 2000, we know that if ρ & σ are pure states, then

$$\|\rho-\sigma\|_{LOCC} = \|\rho-\sigma\|_1$$

→ explains why for security in data hiding we need mixedness

→ compare to secret sharing: need randomness to hide information from local access

[à la Shamir, Comm. ACM 1979]

* From DiVincenzo et al., IEEE-IT 2002: $\chi(LOCC) \geq \Omega\left(\frac{1}{d^2}\right)$

We show: $\lambda(\mathrm{LOCC}) \geq \lambda\left(\underbrace{(\text{unif. POVM})}_{(d \cdot d_A |\psi\rangle\langle\psi|)_{|\psi\rangle \in A}} \otimes \underbrace{(\text{unif. POVM})}_{\text{same for } B}\right) \geq \frac{1}{13d}$

... by an adaptation of the technique of Aubrun/Szarek, Proc. CCC 2007, some extra tricks & a gruelling calculation...

We show: $\lambda(\mathrm{LOCC}) \geq \lambda\left(\underbrace{(\text{unif. POVM})}_{(d \cdot d_A |\psi\rangle\langle\psi|)_{|\psi\rangle \in A}} \otimes \underbrace{(\text{unif. POVM})}_{\text{same for } B}\right) \geq \frac{1}{13d}$

... by an adaptation of the technique of Aubrun/Szarek, Proc. CCC 2007, some extra tricks & a gruelling calculation...

[In fact: $\|\rho - \sigma\|_{\mathrm{LOCC}} \geq \frac{1}{13}\|\rho - \sigma\|_2 = \frac{1}{13}\sqrt{\mathrm{Tr}(\rho-\sigma)^2}$]

→ Implies: both states need to be highly mixed for secure data hiding!

Conclusion & Outlook:

- Formalized the power of restricted measurements as norms on states

Conclusion & Outlook:

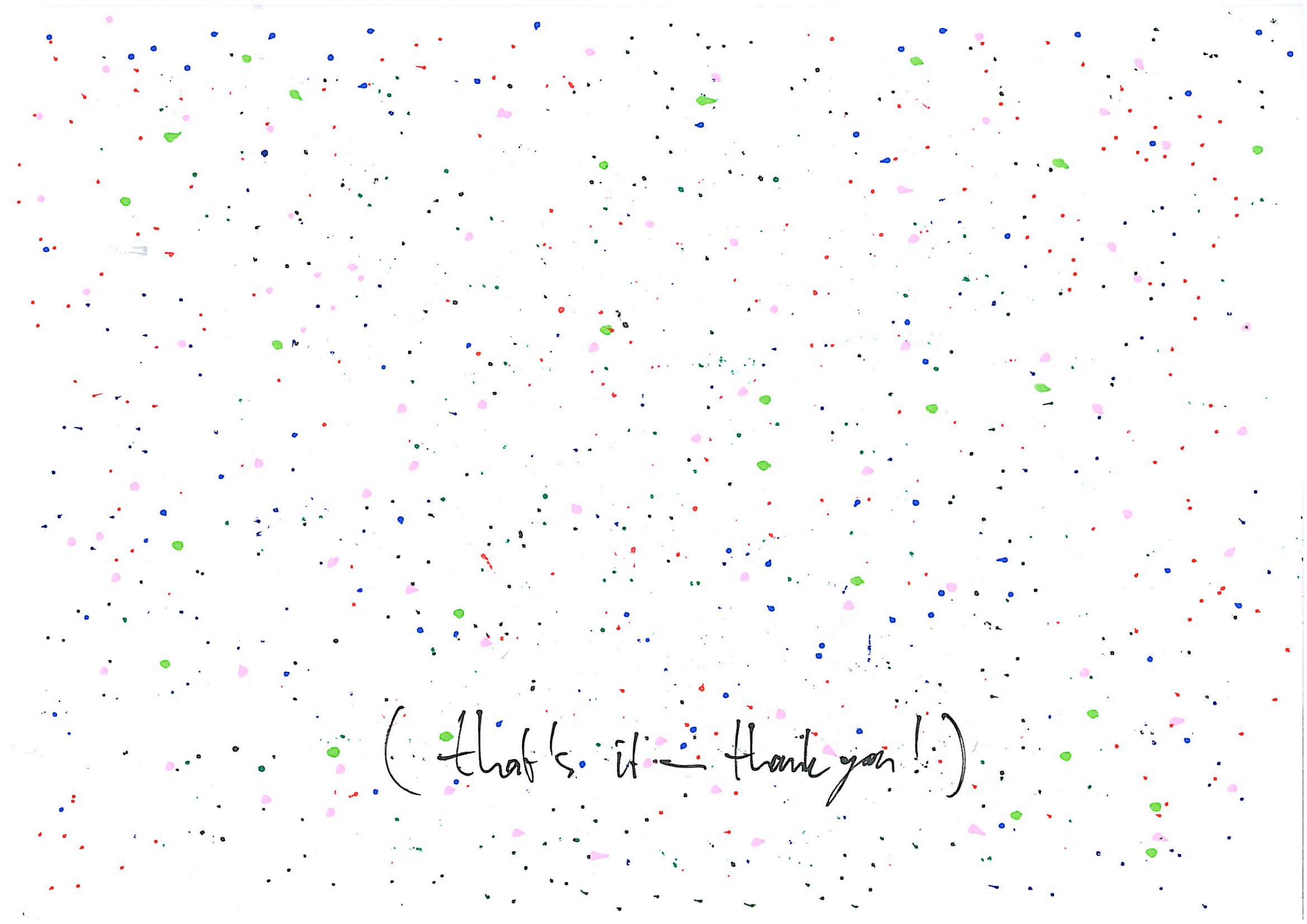
- Formalized the power of restricted measurements as norms on states
- Found (asymptotic) best & worst cases for some interesting models

Conclusion & Outlook:

- Formalized the power of restricted measurements as norms on states
- Found (asymptotic) best & worst cases for some interesting models
- ...in particular: smallest achievable bias in data hiding

Conclusion & Outlook:

- Formalized the power of restricted measurements as norms on states
- Found (asymptotic) best & worst cases for some interesting models
- ...in particular: smallest achievable bias in data hiding
- Many open questions: constant(s) of domination only one parameter of the geometric shape of the unit ball of $\|\cdot\|_{\mathcal{M}}$...


(that's it — thank you!)