The distinguishability of random quantum states

Ashley Montanaro[1]

[1]Department of Computer Science
University of Bristol
Bristol, UK

25th October 2006

Distinguishing quantum states

We will consider a basic question in quantum measurement theory.

Question

Consider a known ensemble $\mathcal{E}$ of n quantum states $\{|\psi_i\rangle\}$ with a priori probabilities p_i. Given an unknown state $|\psi_?\rangle$, picked at random from $\mathcal{E}$, what is the optimal probability $P^{opt}(\mathcal{E})$ of identifying $|\psi_?\rangle$? That is,

$$P^{opt}(\mathcal{E}) = \max_M \sum_{i=1}^{n} p_i \langle \psi_i | M_i | \psi_i \rangle$$

where we maximise over all POVMs $M = \{M_i\}$.

Previous work

- This problem has been considered by many authors since the 1970s, under titles like “quantum hypothesis testing”, “quantum detection”, “quantum state discrimination” etc.
- Many other optimality criteria have also been considered (e.g.: maximise information gain).

[1] C. Helstrom, *Quantum detection and estimation theory* (1978)
[2] Y. Eldar, A. Megretski, G. Verghese, quant-ph/0205178 (2002)

Previous work

- This problem has been considered by many authors since the 1970s, under titles like "quantum hypothesis testing", "quantum detection", "quantum state discrimination" etc.
- Many other optimality criteria have also been considered (e.g.: maximise information gain).
- Helstrom derived an analytic expression for $P^{opt}(\mathcal{E})$ in the case where $\mathcal{E}$ contains 2 states [1].
- In general, producing an analytic expression for $P^{opt}(\mathcal{E})$ appears to be intractable (although good numerical solutions can be found[2])

[1] C. Helstrom, *Quantum detection and estimation theory* (1978)

[2] Y. Eldar, A. Megretski, G. Verghese, quant-ph/0205178 (2002)

Previous work

- This problem has been considered by many authors since the 1970s, under titles like "quantum hypothesis testing", "quantum detection", "quantum state discrimination" etc.
- Many other optimality criteria have also been considered (e.g.: maximise information gain).
- Helstrom derived an analytic expression for $P^{opt}(\mathcal{E})$ in the case where $\mathcal{E}$ contains 2 states [1].
- In general, producing an analytic expression for $P^{opt}(\mathcal{E})$ appears to be intractable (although good numerical solutions can be found[2])
- We are therefore led to producing lower bounds on $P^{opt}(\mathcal{E})$.

[1] C. Helstrom, *Quantum detection and estimation theory* (1978)

[2] Y. Eldar, A. Megretski, G. Verghese, quant-ph/0205178 (2002)

This talk

I will discuss:

1. Part I: the distinguishability of quantum states
   1. Using a specific measurement to lower bound $P^{opt}(\mathcal{E})$
   2. Two lower bounds on $P^{opt}(\mathcal{E})$: a "local" bound and a "global" bound
   3. Extending the lower bounds to mixed states
2. Part II: random quantum states
   1. Random quantum states and random matrix theory
   2. Lower bounds on the distinguishability of random quantum states
   3. Application: how mixed is my subsystem?
   4. Application: the "oracle identification problem" in quantum computation

Notation

I will use the following notation throughout the talk:

- $\mathcal{E} = \{|\psi_i\rangle\}$: the ensemble of states to distinguish
- p_i: the *a priori* probability of the i'th state
- $n = |\mathcal{E}|$: the number of states in $\mathcal{E}$
- d: the dimension of the states in $\mathcal{E}$

Notation

I will use the following notation throughout the talk:

- $\mathcal{E} = \{|\psi_i\rangle\}$: the ensemble of states to distinguish
- p_i: the *a priori* probability of the i'th state
- $n = |\mathcal{E}|$: the number of states in $\mathcal{E}$
- d: the dimension of the states in $\mathcal{E}$

- S: the $d \times n$ state matrix $S = \left(\sqrt{p_1}|\psi_1\rangle \sqrt{p_2}|\psi_2\rangle \cdots \sqrt{p_n}|\psi_n\rangle\right)$
- ρ: the density matrix $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$
- G: the Gram matrix $G_{ij} = \sqrt{p_i}\sqrt{p_j}\langle\psi_i|\psi_j\rangle$

- $P^M(\mathcal{E})$: the probability of success of measurement M applied to $\mathcal{E}$

Part I: the distinguishability of quantum states

1. Using a specific measurement to lower bound $P^{opt}(\mathcal{E})$
2. Two lower bounds on $P^{opt}(\mathcal{E})$: a "local" bound and a "global" bound
3. Extending the lower bounds to mixed states

Methods

- The lower bounds are obtained by putting a lower bound on the probability of success of a specific measurement that can be defined for any ensemble of states, the *Pretty Good Measurement* (PGM)[3].
- For pure states, the PGM is defined by the set of measurement operators $\{|\mu_i\rangle\langle\mu_i|\}$, where $|\mu_i\rangle = \sqrt{p_i}\rho^{-1/2}|\psi_i\rangle$.
- It's easy to show that this always gives a valid measurement ($\sum_i |\mu_i\rangle\langle\mu_i| = I$)

[3] P. Hausladen, W. Wootters (1994)

The canonical nature of the PGM

The PGM has a number of desirable properties, including that:

- It can be defined analytically for any ensemble of states
- It's almost optimal for any ensemble $\mathcal{E}$ [4]:
  - $P^{pgm}(\mathcal{E}) \geq P^{opt}(\mathcal{E})^2$

For us, the important fact is that it's easy to analyse.

[4] H. Barnum and E. Knill, quant-ph/0004088 (2000)

The canonical nature of the PGM

The PGM has a number of desirable properties, including that:

- It can be defined analytically for any ensemble of states
- It's almost optimal for any ensemble $\mathcal{E}$ [4]:
  - $P^{pgm}(\mathcal{E}) \geq P^{opt}(\mathcal{E})^2$

For us, the important fact is that it's easy to analyse.

Key fact

Let G be the rescaled Gram matrix of the ensemble $\mathcal{E}$, $G_{ij} = \sqrt{p_i p_j}\langle\psi_i|\psi_j\rangle$. Then the probability of success of the PGM is

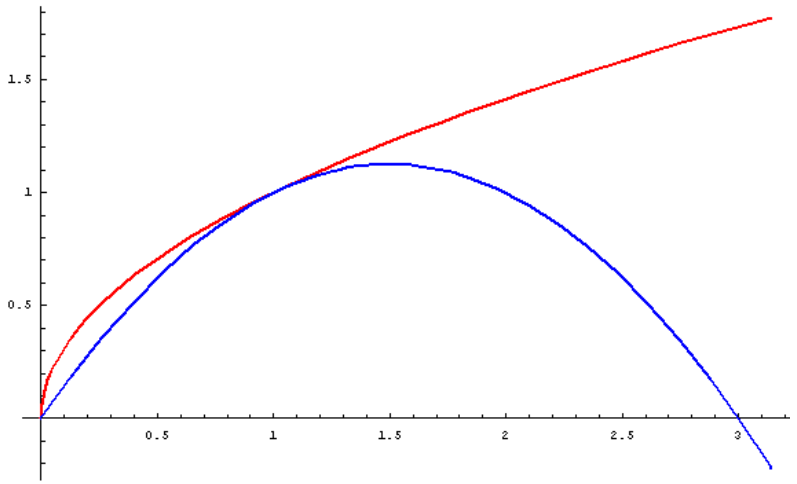
$$P^{pgm}(\mathcal{E}) = \sum_{i=1}^{n} p_i |\langle\psi_i|\mu_i\rangle|^2 = \sum_{i=1}^{n} (\sqrt{G})_{ii}^2$$

Our two lower bounds will be based on lower bounding this sum.

[4] H. Barnum and E. Knill, quant-ph/0004088 (2000)

The pairwise inner product bound

- The first lower bound is based on a strategy used by Hausladen et al.[5]
- The strategy is to put a lower bound on the square root function by an "easier" function (a parabola)
- Works because $\sqrt{x} \geq ax + bx^2 \Rightarrow (\sqrt{G})_{ii} \geq aG_{ii} + b\sum_j |G_{ij}|^2$.



Red: $\sqrt{x}$. Blue: $\frac{3}{2}x - \frac{1}{2}x^2$

[5] P. Hausladen, R. Jozsa, B. Schumacher, W. Wootters (1996)

The pairwise inner product bound

- We can improve their bound by producing (for a given set of states) a set of optimal parabolae.
- For each i, we look for a and b such that $\sqrt{x} \geq ax + bx^2$ for $x \geq 0$, and $aG_{ii} + b\sum_j |G_{ij}|^2$ is maximised.

The pairwise inner product bound

- We can improve their bound by producing (for a given set of states) a set of optimal parabolae.
- For each i, we look for a and b such that $\sqrt{x} \geq ax + bx^2$ for $x \geq 0$, and $aG_{ii} + b\sum_j |G_{ij}|^2$ is maximised.
- Only basic calculus is required to find these values of a and b, and substituting in gives the result:

Pairwise inner product bound

Let $\mathcal{E}$ be an ensemble of n states $\{|\psi_i\rangle\}$ with a priori probabilities p_i.

$$\text{Then} \quad P^{pgm}(\mathcal{E}) \geq \sum_{i=1}^{n} \frac{p_i^2}{\sum_{j=1}^{n} p_j |\langle\psi_i|\psi_j\rangle|^2}$$

The eigenvalue bound

The second lower bound is based on a global measure of distinguishability of the states in $\mathcal{E}$: the eigenvalues $\{\lambda_i\}$ of the Gram matrix G.

The eigenvalue bound

The second lower bound is based on a global measure of distinguishability of the states in $\mathcal{E}$: the eigenvalues $\{\lambda_i\}$ of the Gram matrix G.

$$\sum_{i=1}^{n} (\sqrt{G})_{ii} \quad = \quad \sum_{i=1}^{n} \sqrt{\lambda_i} \qquad (1)$$

The eigenvalue bound

The second lower bound is based on a global measure of distinguishability of the states in $\mathcal{E}$: the eigenvalues $\{\lambda_i\}$ of the Gram matrix G.

$$\sum_{i=1}^{n} (\sqrt{G})_{ii} = \sum_{i=1}^{n} \sqrt{\lambda_i} \quad (1)$$

$$\Rightarrow \left(\sum_{i=1}^{n} (\sqrt{G})_{ii} \right)^2 = \left(\sum_{i=1}^{n} \sqrt{\lambda_i} \right)^2 \quad (2)$$

The eigenvalue bound

The second lower bound is based on a global measure of distinguishability of the states in $\mathcal{E}$: the eigenvalues $\{\lambda_i\}$ of the Gram matrix G.

$$\sum_{i=1}^{n} (\sqrt{G})_{ii} = \sum_{i=1}^{n} \sqrt{\lambda_i} \quad (1)$$

$$\Rightarrow \left(\sum_{i=1}^{n} (\sqrt{G})_{ii} \right)^2 = \left(\sum_{i=1}^{n} \sqrt{\lambda_i} \right)^2 \quad (2)$$

$$\Rightarrow n \sum_{i=1}^{n} (\sqrt{G})_{ii}^2 \geq \left(\sum_{i=1}^{n} \sqrt{\lambda_i} \right)^2 \quad (3)$$

The eigenvalue bound

The second lower bound is based on a global measure of distinguishability of the states in $\mathcal{E}$: the eigenvalues $\{\lambda_i\}$ of the Gram matrix G.

$$\sum_{i=1}^{n} (\sqrt{G})_{ii} = \sum_{i=1}^{n} \sqrt{\lambda_i} \quad (1)$$

$$\Rightarrow \left(\sum_{i=1}^{n} (\sqrt{G})_{ii} \right)^2 = \left(\sum_{i=1}^{n} \sqrt{\lambda_i} \right)^2 \quad (2)$$

$$\Rightarrow n \sum_{i=1}^{n} (\sqrt{G})_{ii}^2 \geq \left(\sum_{i=1}^{n} \sqrt{\lambda_i} \right)^2 \quad (3)$$

$$\Rightarrow P^{pgm}(\mathcal{E}) \geq \frac{1}{n} \left(\sum_{i=1}^{n} \sqrt{\lambda_i} \right)^2 \quad (4)$$

The eigenvalue bound

The second lower bound is based on a global measure of distinguishability of the states in $\mathcal{E}$: the eigenvalues $\{\lambda_i\}$ of the Gram matrix G.

$$\sum_{i=1}^{n} (\sqrt{G})_{ii} = \sum_{i=1}^{n} \sqrt{\lambda_i} \quad (1)$$

$$\Rightarrow \left(\sum_{i=1}^{n} (\sqrt{G})_{ii} \right)^2 = \left(\sum_{i=1}^{n} \sqrt{\lambda_i} \right)^2 \quad (2)$$

$$\Rightarrow n \sum_{i=1}^{n} (\sqrt{G})_{ii}^2 \geq \left(\sum_{i=1}^{n} \sqrt{\lambda_i} \right)^2 \quad (3)$$

$$\Rightarrow P^{pgm}(\mathcal{E}) \geq \frac{1}{n} \left(\sum_{i=1}^{n} \sqrt{\lambda_i} \right)^2 \quad (4)$$

In terms of the trace norm, $P^{pgm}(\mathcal{E}) \geq \frac{1}{n} \|S\|_1^2 = \frac{1}{n} \left(\sum_i \sigma_i(S) \right)^2$.

Comparison with previous bounds

- Previous authors (e.g. Burnashev and Holevo [6]) have used bounds based on similar principles.
- But the bounds here are stronger, especially for low values of $P^{pgm}(\mathcal{E})$, and always give a non-trivial value.

[6] M. V. Burnashev and A. S. Holevo, On reliability function of quantum communication channel, quant-ph/9703013

Comparison with previous bounds

- Previous authors (e.g. Burnashev and Holevo [6]) have used bounds based on similar principles.
- But the bounds here are stronger, especially for low values of $P^{pgm}(\mathcal{E})$, and always give a non-trivial value.
- Assuming the states in $\mathcal{E}$ have equal probabilities:

Comparison of bounds

Previously known lower bound	**New lower bound**
$P^{pgm}(\mathcal{E}) \geq 1 - \frac{1}{n}\sum_{i \neq j} \lvert\langle\psi_i\rvert\psi_j\rangle\rvert^2$	$P^{pgm}(\mathcal{E}) \geq \frac{1}{n}\sum_{i=1}^{n} \frac{1}{\sum_{j=1}^{n} \lvert\langle\psi_i\rvert\psi_j\rangle\rvert^2}$
$P^{pgm}(\mathcal{E}) \geq \frac{2}{\sqrt{n}}\mathrm{tr}(\sqrt{G}) - 1$	$P^{pgm}(\mathcal{E}) \geq \frac{1}{n}\mathrm{tr}(\sqrt{G})^2$

[6] M. V. Burnashev and A. S. Holevo, On reliability function of quantum communication channel, quant-ph/9703013

A local bound and a global bound

- It is interesting to note that the inner product bound only considers the pairwise distinguishability between states, while the eigenvalue bound is based on global features of the ensemble.
- We might therefore expect the latter to be stronger...

A local bound and a global bound

- It is interesting to note that the inner product bound only considers the pairwise distinguishability between states, while the eigenvalue bound is based on global features of the ensemble.
- We might therefore expect the latter to be stronger...

- Consider an ensemble of n states, each pair of which have the same inner product, $k \in \mathbb{R}^+$. Then it is possible to show that:
  - The inner product bound gives an almost trivial bound: $P^{pgm}(\mathcal{E}) \geq O(1/n)$
  - The eigenvalue bound gives a strong bound: $P^{pgm}(\mathcal{E}) \geq (1-k) - o(1)$

A local bound and a global bound

- It is interesting to note that the inner product bound only considers the pairwise distinguishability between states, while the eigenvalue bound is based on global features of the ensemble.
- We might therefore expect the latter to be stronger...

- Consider an ensemble of n states, each pair of which have the same inner product, $k \in \mathbb{R}^+$. Then it is possible to show that:
  - The inner product bound gives an almost trivial bound: $P^{pgm}(\mathcal{E}) \geq O(1/n)$
  - The eigenvalue bound gives a strong bound: $P^{pgm}(\mathcal{E}) \geq (1-k) - o(1)$
- (NB: in this trivial case we can actually diagonalise the Gram matrix and calculate the probability of success of the PGM exactly)

Mixed states

- We have only considered the distinguishability of pure states.
- It turns out that we can give a lower bound on the distinguishability of mixed states too...
  - If we have an ensemble of states $\{\rho_i\}$, and we know their eigendecompositions $\rho_i = \sum_k \lambda_{ik} |e_{ik}\rangle\langle e_{ik}|$, then we can relate the problem of distinguishing the $\{\rho_i\}$ to distinguishing the $|e_{ik}\rangle$.

Mixed states

- We have only considered the distinguishability of pure states.
- It turns out that we can give a lower bound on the distinguishability of mixed states too...
  - If we have an ensemble of states $\{\rho_i\}$, and we know their eigendecompositions $\rho_i = \sum_k \lambda_{ik} |e_{ik}\rangle\langle e_{ik}|$, then we can relate the problem of distinguishing the $\{\rho_i\}$ to distinguishing the $|e_{ik}\rangle$.
- Even if we don't know the eigendecompositions, we can convert the inner product bound to a bound based on the pairwise *fidelities* of the states in $\mathcal{E}$.

Pairwise fidelity bound

Let $\mathcal{E}$ be an ensemble of n states $\{\rho_i\}$ with a priori probabilities p_i.

$$\text{Then} \quad P^{pgm}(\mathcal{E}) \geq \sum_{i=1}^{n} \frac{p_i^2 \operatorname{tr}(\rho_i^2)}{\sum_{j=1}^{n} p_j F(\rho_i, \rho_j)}$$

Part II: random quantum states

1. Random quantum states and random matrix theory
2. Lower bounds on the distinguishability of random quantum states
3. Application: how mixed is my subsystem?
4. Application: the "oracle identification problem" in quantum computation

Random quantum states

- We will now apply the eigenvalue bound to the case where the states in $\mathcal{E}$ are random.
- To be precise, for all i:
  - $|\psi_i\rangle$ is distributed uniformly at random on the d-dimensional complex unit sphere (according to Haar measure)
  - $p_i = 1/n$ (the states are equiprobable)
- We will calculate the expected probability of success of identifying $|\psi_?\rangle$ in this case.

Random quantum states

- We will now apply the eigenvalue bound to the case where the states in $\mathcal{E}$ are random.
- To be precise, for all i:
  - $|\psi_i\rangle$ is distributed uniformly at random on the d-dimensional complex unit sphere (according to Haar measure)
  - $p_i = 1/n$ (the states are equiprobable)
- We will calculate the expected probability of success of identifying $|\psi_?\rangle$ in this case.

So we are given a state picked at random from a known set of states which are themselves randomly picked, and asked to determine which random state our randomly picked state actually is ☺

Random quantum states

How do we produce a state $|\psi\rangle$ distributed uniformly at random?

- Generate a vector v whose components v_i are complex Gaussians, then set $|\psi\rangle = v/\|v\|$.
  - i.e. v_i's real and complex parts are independently normally distributed with variance $1/2$; both parts have probability density function $\frac{1}{2\sqrt{\pi}} e^{-x^2/2}$ and $\mathbb{E}(|v_i|^2) = 1$.
- This works because of the spherical symmetry of the multivariate normal distribution.

Random quantum states

How do we produce a state $|\psi\rangle$ distributed uniformly at random?

- Generate a vector v whose components v_i are complex Gaussians, then set $|\psi\rangle = v/\|v\|$.
  - i.e. v_i's real and complex parts are independently normally distributed with variance $1/2$; both parts have probability density function $\frac{1}{2\sqrt{\pi}} e^{-x^2/2}$ and $\mathbb{E}(|v_i|^2) = 1$.
- This works because of the spherical symmetry of the multivariate normal distribution.

- It turns out that the normalisation step becomes "almost" unnecessary in high dimension (qv): rescaling v by $1/\sqrt{d}$ will give a complex vector whose norm is approximately 1.
- So the state matrix S is (almost!) a rescaled matrix of Gaussians: $S_{ij} \sim \widetilde{N}(0, 1/nd)$, and we need to calculate $\mathbb{E}(\frac{1}{n}\|S\|_1^2)$.

Random matrix theory

- Random matrix theory deals with the properties of matrices whose entries are random variables.

[7] V. A. Marčenko and L. A. Pastur (1967)

Random matrix theory

- Random matrix theory deals with the properties of matrices whose entries are random variables.
- In particular, infinite-dimensional random matrix theory allows us to answer questions like "what is the limiting density of the eigenvalues of a family of $n \times n$ random matrices, as $n \to \infty$?".
  - By density, we mean the function $f(x)$ which integrates to $F(x) = \frac{1}{n}(\#eigenvalues < x)$
  - It's not a priori obvious that such a limit should exist!

[7] V. A. Marčenko and L. A. Pastur (1967)

Random matrix theory

- Random matrix theory deals with the properties of matrices whose entries are random variables.
- In particular, infinite-dimensional random matrix theory allows us to answer questions like "what is the limiting density of the eigenvalues of a family of $n \times n$ random matrices, as $n \to \infty$?".
  - By density, we mean the function $f(x)$ which integrates to $F(x) = \frac{1}{n}(\#eigenvalues < x)$
  - It's not a priori obvious that such a limit should exist!

- Statisticians have long studied the density of eigenvalues of the matrix $G = SS^\dagger$, where S is a random matrix: under certain conditions, it's given by the Marčenko-Pastur law [7].
  - This is the equivalent of the famous Wigner semicircle law for random Hermitian matrices...

[7] V. A. Marčenko and L. A. Pastur (1967)

The Marčenko-Pastur law

The Marčenko-Pastur law gives the limiting density of the eigenvalues of a sample covariance matrix $G = SS^\dagger$ under very weak conditions.

The Marčenko-Pastur law

The Marčenko-Pastur law gives the limiting density of the eigenvalues of a sample covariance matrix $G = SS^\dagger$ under very weak conditions.

Marčenko-Pastur law

Let R_r be a family of $d \times n$ matrices with $n \geq d$ and $d/n \to r \in (0, 1]$ as $n, d \to \infty$, where the entries of R_r are i.i.d. complex random variables with mean 0 and variance 1. Then, as $n, d \to \infty$, the eigenvalues of the rescaled matrix $\frac{1}{n} R_r R_r^\dagger$ tend almost surely to a limiting distribution with density

$$p_r(x) = \frac{\sqrt{(x - A^2)(B^2 - x)}}{2\pi r x}$$

for $A^2 \leq x \leq B^2$ (where $A = 1 - \sqrt{r}$, $B = 1 + \sqrt{r}$), and density 0 elsewhere.

The Marčenko-Pastur law

The Marčenko-Pastur law gives the limiting density of the eigenvalues of a sample covariance matrix $G = SS^\dagger$ under very weak conditions.

Marčenko-Pastur law

Let R_r be a family of $d \times n$ matrices with $n \geq d$ and $d/n \to r \in (0, 1]$ as $n, d \to \infty$, where the entries of R_r are i.i.d. complex random variables with mean 0 and variance 1. Then, as $n, d \to \infty$, the eigenvalues of the rescaled matrix $\frac{1}{n} R_r R_r^\dagger$ tend almost surely to a limiting distribution with density

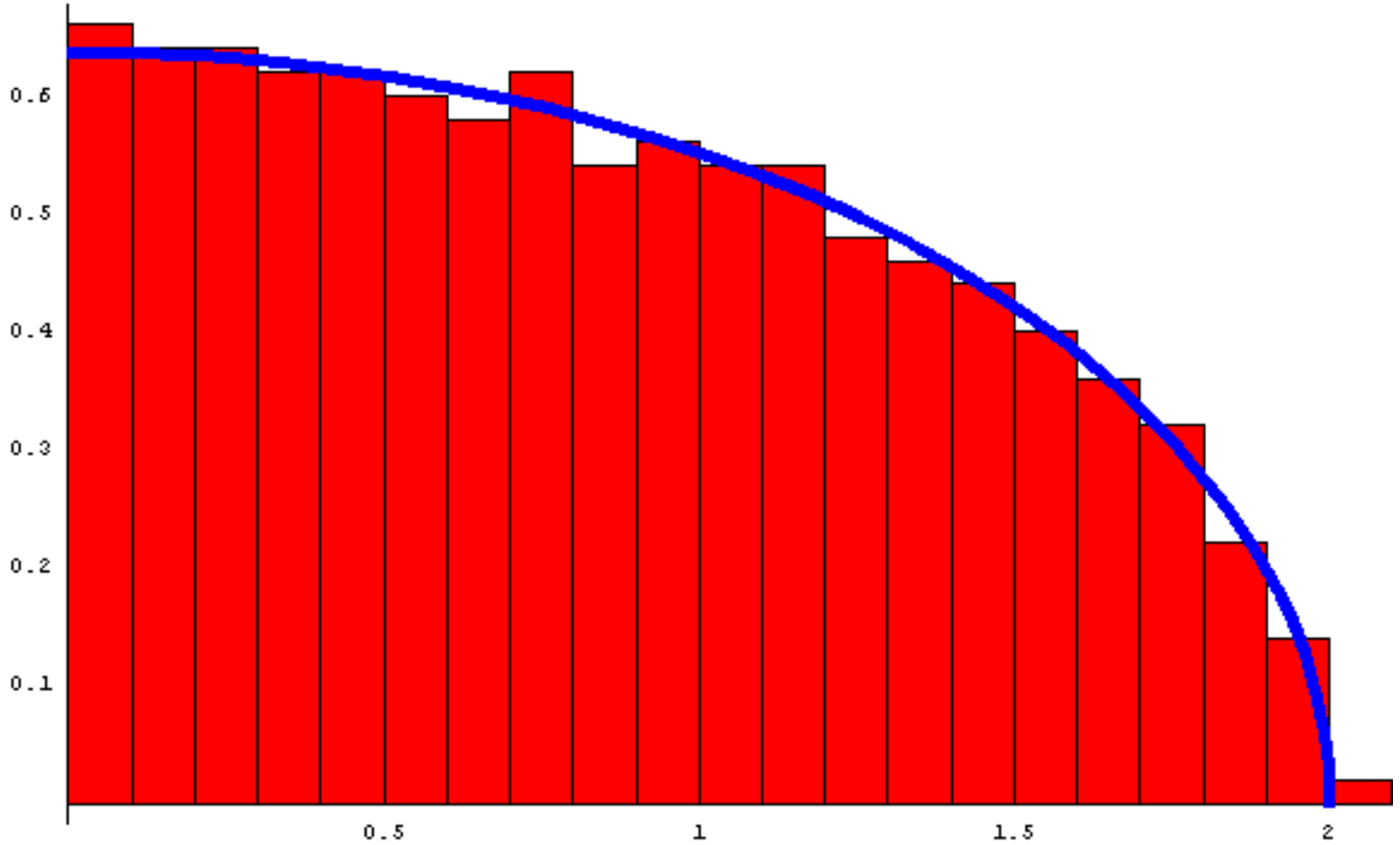
$$p_r(x) = \frac{\sqrt{(x - A^2)(B^2 - x)}}{2\pi r x}$$

for $A^2 \leq x \leq B^2$ (where $A = 1 - \sqrt{r}$, $B = 1 + \sqrt{r}$), and density 0 elsewhere.

We can easily tweak this result to tell us the density of the singular values of R_r instead!

Experimental results

Who says theorists don't know how to do experiments? ☺



Blue: singular value density predicted by Marčenko-Pastur law

Red: empirical singular value distribution of a 500x500 matrix

Applying the Marčenko-Pastur law

We can use the M-P law to give us the expected trace norm of a random matrix, again under very weak conditions.

Expected trace norm

Let R_r be a family of $d \times n$ matrices with $k/m \to r \in (0, 1]$ as $n, d \to \infty$, where $k = \min(n, d)$ and $m = \max(n, d)$, and the entries of R_r are i.i.d. complex random variables with mean 0 and variance 1. Then, as $n, d \to \infty$, the expected trace norm of R_r tends almost surely to

$$\mathbb{E}(\|R_r\|_1) = \frac{m^{3/2}}{\pi} \int_A^B \sqrt{(y^2 - A^2)(B^2 - y^2)}\, dy$$

where $A = 1 - \sqrt{r}$, $B = 1 + \sqrt{r}$.

Applying the Marčenko-Pastur law (2)

We want to evaluate the following integral:

$$\int_A^B \sqrt{(y^2 - A^2)(B^2 - y^2)}\, dy$$

Unfortunately, this is an elliptic integral with no analytic solution. But we can find a good lower bound on the integral...

Applying the Marčenko-Pastur law (2)

We want to evaluate the following integral:

$$\int_A^B \sqrt{(y^2 - A^2)(B^2 - y^2)}\, dy$$

Unfortunately, this is an elliptic integral with no analytic solution. But we can find a good lower bound on the integral...

Elliptic integral lower bound

Let $0 \leq r \leq 1$ and $A = 1 - \sqrt{r}$, $B = 1 + \sqrt{r}$. Then

$$\int_A^B \sqrt{(y^2 - A^2)(B^2 - y^2)}\, dy \geq r\pi \sqrt{1 - r\left(1 - \frac{64}{9\pi^2}\right)}$$

with equality at $r = 0$, $r = 1$.

(The proof is quite long and involves representing the integral as the difference of two hypergeometric series and performing several transformations on these hypergeometric series...)

The asymptotic lower bound

Main theorem

Let $\mathcal{E}$ be an ensemble of n equiprobable d-dimensional quantum states $\{|\psi_i\rangle\}$ with $n/d \to r \in (0, \infty)$ as $n, d \to \infty$, and let the components of $|\psi_i\rangle$ in some basis be i.i.d. complex random variables with mean 0 and variance $1/d$. Then, as $n, d \to \infty$,

$$\mathbb{E}(P^{pgm}(\mathcal{E})) \geq \begin{cases} \frac{1}{r}\left(1 - \frac{1}{r}\left(1 - \frac{64}{9\pi^2}\right)\right) & \text{if } n \geq d \\ 1 - r\left(1 - \frac{64}{9\pi^2}\right) & \text{otherwise} \end{cases}$$

and in particular $\mathbb{E}(P^{pgm}(\mathcal{E})) > 0.720$ when $n \leq d$.

Comparison with numerical results (1)
$(0 \leq n \leq 2d)$

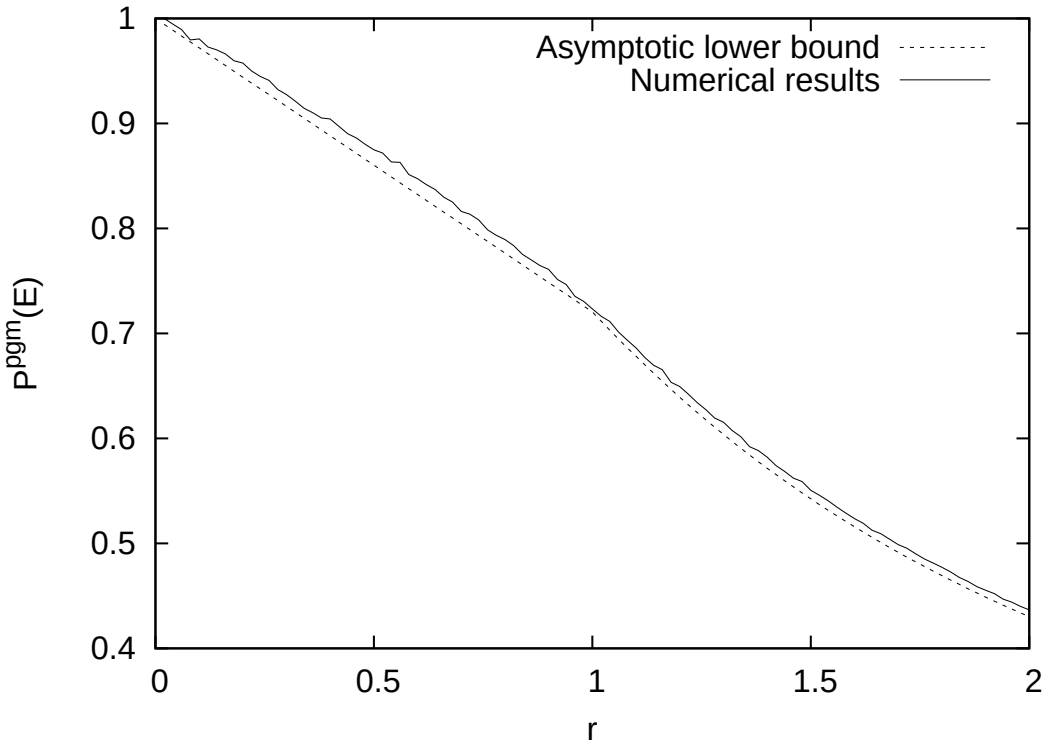


Figure: Asymptotic bound on $P^{pgm}(\mathcal{E})$ vs. numerical results (averaged over 10 runs) for ensembles of $n = 50r$ 50-dimensional uniformly random states.

Comparison with numerical results (2)
$(0 \leq n \leq 10d)$

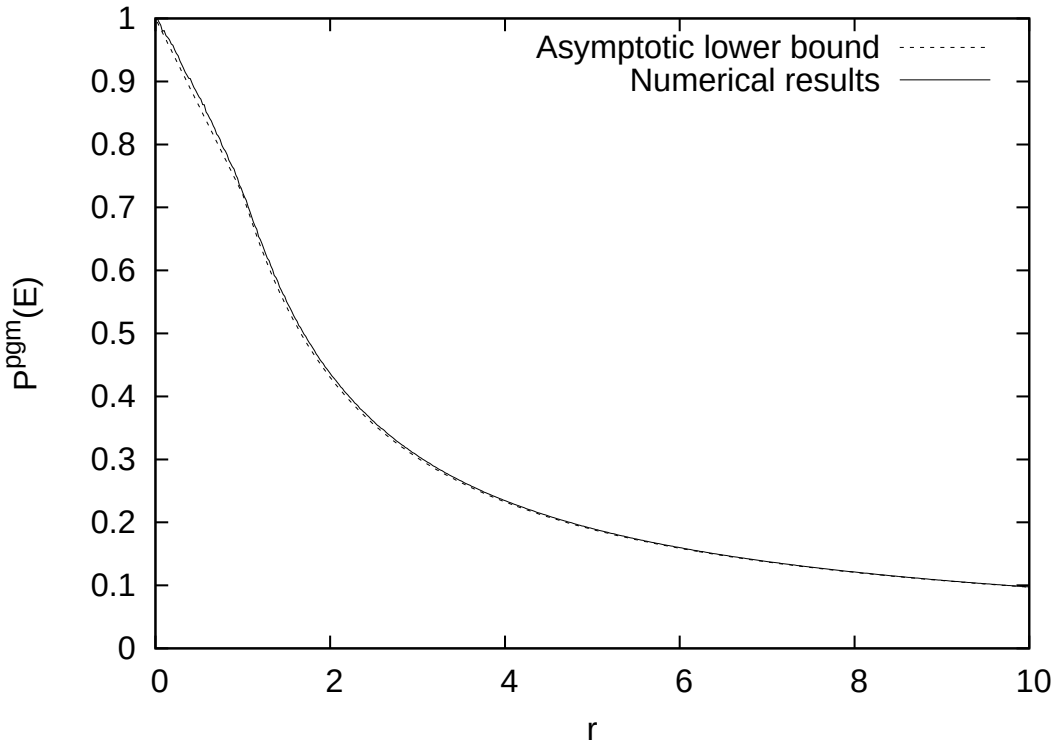


Figure: Asymptotic bound on $P^{pgm}(\mathcal{E})$ vs. numerical results (averaged over 10 runs) for ensembles of $n = 50r$ 50-dimensional uniformly random states.

A finite-dimensional lower bound

- The M-P law holds in the asymptotic limit. Can we find a lower bound on the expected distinguishability of an ensemble of finite-dimensional random states?
- Also, I glossed over the issue of normalising the states we produce...

A finite-dimensional lower bound

- The M-P law holds in the asymptotic limit. Can we find a lower bound on the expected distinguishability of an ensemble of finite-dimensional random states?
- Also, I glossed over the issue of normalising the states we produce...
- There are two "bad events" that we have to take into account:
  1. The eigenvalue distribution in finite dimension d will not be given by the M-P law, but some approximation
  2. The normalisation of the states might perturb the state matrix excessively

A finite-dimensional lower bound

- The M-P law holds in the asymptotic limit. Can we find a lower bound on the expected distinguishability of an ensemble of finite-dimensional random states?
- Also, I glossed over the issue of normalising the states we produce...
- There are two "bad events" that we have to take into account:
  1. The eigenvalue distribution in finite dimension d will not be given by the M-P law, but some approximation
  2. The normalisation of the states might perturb the state matrix excessively
- Actually, both of these problems can be overcome:
  1. There is a convergence result bounding the rate at which the eigenvalues converge to the M-P law
  2. We can produce a tail bound that says that the normalisation step makes little difference

A finite-dimensional lower bound (2)

Convergence rate to the M-P law:

- We can use a convergence result [8] that gives that the expected probability of success we calculated needs adjusting by a term of order $O(d^{-5/48})$.

[8] Z. D. Bai, 1993

[9] C. H. Bennett, P. Hayden, D. Leung, P. Shor, A. Winter (2003)

A finite-dimensional lower bound (2)

Convergence rate to the M-P law:

- We can use a convergence result [8] that gives that the expected probability of success we calculated needs adjusting by a term of order $O(d^{-5/48})$.

Effect of replacing the normalisation step with rescaling:

- The norm of the original d-dimensional vector v is sharply concentrated around its expected value[9]:
  $\Pr[|\|v\|^2 - 1| \geq \epsilon] < 2e^{-d\epsilon^2/12}$

[8] Z. D. Bai, 1993

[9] C. H. Bennett, P. Hayden, D. Leung, P. Shor, A. Winter (2003)

A finite-dimensional lower bound (2)

Convergence rate to the M-P law:

- We can use a convergence result [8] that gives that the expected probability of success we calculated needs adjusting by a term of order $O(d^{-5/48})$.

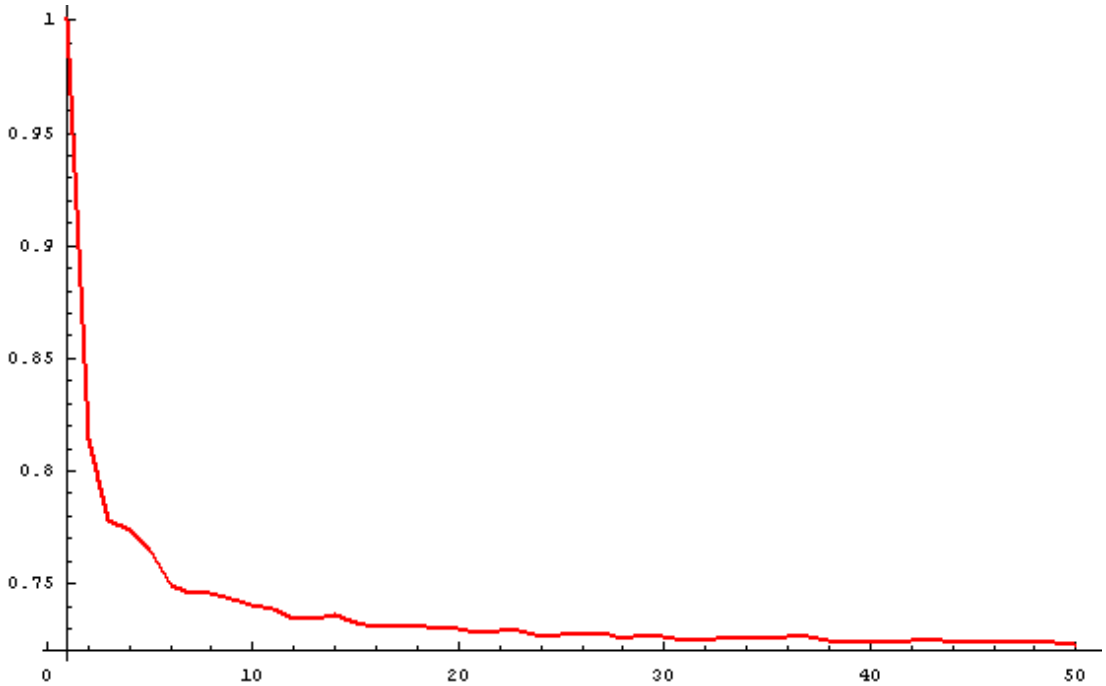
Effect of replacing the normalisation step with rescaling:

- The norm of the original d-dimensional vector v is sharply concentrated around its expected value[9]:
  $\Pr[|\|v\|^2 - 1| \geq \epsilon] < 2e^{-d\epsilon^2/12}$
- $\Rightarrow$ Each column of the state matrix S will be close to the vector we would get by blindly rescaling v
- $\Rightarrow$ The probability that the trace norm of S is far from the "blindly rescaled" matrix S' is exponentially small in d

[8] Z. D. Bai, 1993

[9] C. H. Bennett, P. Hayden, D. Leung, P. Shor, A. Winter (2003)

Convergence in low dimension



Empirical probability of success of the PGM applied to n states in n dimensions (averaged over 100 runs).

Concentration of measure

- So far we've only calculated the *expected* distinguishability of a set of random states.
- How close are most states to this expected value?
- Concentration of measure results allow us to answer this question!

Concentration of measure

- So far we've only calculated the *expected* distinguishability of a set of random states.
- How close are most states to this expected value?
- Concentration of measure results allow us to answer this question!

Concentration of measure

Let p be a point in $\mathbb{R}^d$ picked in accordance with standard Gaussian measure (i.e. $\mathbb{E}(p_i^2) = 1$). Then

$$\Pr[|f(p) - \mathbb{E}(f)| \geq \epsilon] \leq 2e^{-\epsilon^2/2\eta^2} \tag{5}$$

where η is the Lipschitz constant of f,
$\eta = \sup_{x,y} |f(x) - f(y)|/\|x - y\|_2$.

Concentration of measure (2)

- So we identify the state matrix S with a (rescaled) $2nd$-dimensional vector of real Gaussians and consider the "distinguishability" function $f(S) = \frac{1}{n}\|S\|_1^2$: then $\eta \leq 2\|S\|_2$.

Concentration of measure (2)

- So we identify the state matrix S with a (rescaled) $2nd$-dimensional vector of real Gaussians and consider the "distinguishability" function $f(S) = \frac{1}{n}\|S\|_1^2$: then $\eta \leq 2\|S\|_2$.
- Plugging this in, and considering the finite-dimensional "correction factors" discussed previously, after some fiddly algebra we end up with:

Concentration of measure (2)

- So we identify the state matrix S with a (rescaled) $2nd$-dimensional vector of real Gaussians and consider the "distinguishability" function $f(S) = \frac{1}{n}\|S\|_1^2$: then $\eta \leq 2\|S\|_2$.
- Plugging this in, and considering the finite-dimensional "correction factors" discussed previously, after some fiddly algebra we end up with:

Probability of success in finite dimension

Let $\mathcal{E}$ be an ensemble of n equiprobable d-dimensional quantum states picked uniformly at random. Set
$p = \frac{1}{r}\left(1 - \frac{1}{r}\left(1 - \frac{64}{9\pi^2}\right)\right) - O(n^{-5/48})$ if $n \geq d$, and
$p = 1 - r\left(1 - \frac{64}{9\pi^2}\right) - O(d^{-5/48})$ otherwise. Then, for any $\epsilon \leq p/2$,

$$\Pr[P^{pgm}(\mathcal{E}) \leq p - 2\epsilon] \leq 2\left((n+1)e^{-d\epsilon^4/K} + e^{-nd\epsilon^2/5}\right)$$

where K is a constant ≤ 300.

Why study random states anyway?

Why study random states anyway?

Almost all states are random!

Why study random states anyway?

Almost all states are random!

Other reasons:

- Random states provide an interesting case where we can determine the distinguishability of an ensemble based only on two parameters: n and d.

Why study random states anyway?

Almost all states are random!

Other reasons:

- Random states provide an interesting case where we can determine the distinguishability of an ensemble based only on two parameters: n and d.
- We can get very tight analytic results in this case: even for low dimensions, the bound seems to be within 1% of the observed probability of success of the PGM.

Why study random states anyway?

Almost all states are random!

Other reasons:

- Random states provide an interesting case where we can determine the distinguishability of an ensemble based only on two parameters: n and d.
- We can get very tight analytic results in this case: even for low dimensions, the bound seems to be within 1% of the observed probability of success of the PGM.
- These results allow one to say: my states are like random states $\Rightarrow$ they're (quite) distinguishable.

Why study random states anyway?

Almost all states are random!

Other reasons:

- Random states provide an interesting case where we can determine the distinguishability of an ensemble based only on two parameters: n and d.
- We can get very tight analytic results in this case: even for low dimensions, the bound seems to be within 1% of the observed probability of success of the PGM.
- These results allow one to say: my states are like random states $\Rightarrow$ they're (quite) distinguishable.

But what if we don't care about quantum measurement theory?

Application: how mixed is my subsystem?

- There is another interpretation of these results which doesn't come from quantum measurement.
- It turns out that $\frac{1}{n}\|S\|_1^2$ gives the fidelity of the Gram matrix G with the n-dimensional maximally mixed state I/n.
  - where the fidelity $F(\rho,\sigma) = (\mathrm{tr}\sqrt{\rho^{1/2}\,\sigma\,\rho^{1/2}})^2$
- We may thus interpret the lower bound on the distinguishability of a set of states as how close its Gram matrix is to the maximally mixed state.

Application: how mixed is my subsystem?

- Let $\rho_{n,d}$ be the density matrix obtained by picking a pure state uniformly at random from a $n \times d$-dimensional Hilbert space, and tracing out the n-dimensional portion of it.
  - It's easy to show that $\rho_{n,d} \approx \frac{1}{n} \sum_{i=1}^{n} |\psi_i\rangle\langle\psi_i|$, where $|\psi_i\rangle$ is picked uniformly at random in the d-dimensional space

[10] R. Jozsa, J. Schlienz, quant-ph/9911009

Application: how mixed is my subsystem?

- Let $\rho_{n,d}$ be the density matrix obtained by picking a pure state uniformly at random from a $n \times d$-dimensional Hilbert space, and tracing out the n-dimensional portion of it.
  - It's easy to show that $\rho_{n,d} \approx \frac{1}{n} \sum_{i=1}^{n} |\psi_i\rangle\langle\psi_i|$, where $|\psi_i\rangle$ is picked uniformly at random in the d-dimensional space
- It's possible to show that the non-zero eigenvalues of $\rho_{n,d}$ are the same as those of the Gram matrix of a set of n equiprobable d-dimensional random states[10]
- Using this, one can show that $\frac{1}{d}\|S\|_1^2$ gives the approximate fidelity of $\rho_{n,d}$ with I/d!

[10] R. Jozsa, J. Schlienz, quant-ph/9911009

Application: how mixed is my subsystem?

- Let $\rho_{n,d}$ be the density matrix obtained by picking a pure state uniformly at random from a $n \times d$-dimensional Hilbert space, and tracing out the n-dimensional portion of it.
  - It's easy to show that $\rho_{n,d} \approx \frac{1}{n} \sum_{i=1}^{n} |\psi_i\rangle\langle\psi_i|$, where $|\psi_i\rangle$ is picked uniformly at random in the d-dimensional space
- It's possible to show that the non-zero eigenvalues of $\rho_{n,d}$ are the same as those of the Gram matrix of a set of n equiprobable d-dimensional random states[10]
- Using this, one can show that $\frac{1}{d}\|S\|_1^2$ gives the approximate fidelity of $\rho_{n,d}$ with I/d!
- The previous results thus predict the distance of $\rho_{n,d}$ from the maximally mixed state very closely.
- (Popescu, Short, and Winter previously obtained a similar result by different methods)

[10] R. Jozsa, J. Schlienz, quant-ph/9911009

Application: oracle identification

Problem

Given an unknown Boolean function f, picked uniformly at random from a set S of N Boolean functions on n bits, identify f with the minimum number of uses of f.

[11] A. Ambainis et al, Quantum identification of Boolean oracles, quant-ph/0403056

Application: oracle identification

Problem

Given an unknown Boolean function f, picked uniformly at random from a set S of N Boolean functions on n bits, identify f with the minimum number of uses of f.

- This is a particular case of the oracle identification problem studied by Ambainis et al[11].
- We consider the case where we are allowed a bounded probability of error in our quest to identify f.
- Many important problems fit into this framework (eg. unstructured search).

[11] A. Ambainis et al, Quantum identification of Boolean oracles, quant-ph/0403056

Oracle identification: classical

- A classical algorithm must make at least $\log N$ queries
  - (each query can only reduce the size of the search space by half)
- Note that being allowed some probability of error $< 1/2$ is useless for classical algorithms.
- We can actually show a classical upper bound of $O(\log N)$ queries in the random oracle case.
  - (because in this case every query will reduce the search space by almost half whp)

Oracle identification: quantum

We will show that, when 2^n is large relative to N, for almost all sets of functions f can be identified with a constant number of quantum queries.

Oracle identification: quantum

We will show that, when 2^n is large relative to N, for almost all sets of functions f can be identified with a constant number of quantum queries.

- Consider the following single-query quantum "algorithm":
  1. Create the state $|\psi_f\rangle = \frac{1}{\sqrt{2^n}} \sum_x (-1)^{f(x)} |x\rangle$ using one query to f.
  2. Use the PGM to distinguish the states in the ensemble $\mathcal{E} = \{|\psi_f\rangle\}$.

Oracle identification: quantum

We will show that, when 2^n is large relative to N, for almost all sets of functions f can be identified with a constant number of quantum queries.

- Consider the following single-query quantum "algorithm":
  1. Create the state $|\psi_f\rangle = \frac{1}{\sqrt{2^n}} \sum_x (-1)^{f(x)} |x\rangle$ using one query to f.
  2. Use the PGM to distinguish the states in the ensemble $\mathcal{E} = \{|\psi_f\rangle\}$.
- When the functions are random, the state matrix $S = (\{|\psi_f\rangle / \sqrt{N}\})$ is random, in the sense that the M-P law can be applied to it.
- Why? Because each entry of $\sqrt{N 2^n} S$ is i.i.d. with mean 0 and expected value 1.

Oracle identification: quantum (2)

- So the results here can be used to put the same lower bound on the probability of success of distinguishing these states.
- And in particular, the input size and the number of functions determine this probability (unlike the classical case where we can't use all the input)...

Oracle identification: quantum (2)

- So the results here can be used to put the same lower bound on the probability of success of distinguishing these states.
- And in particular, the input size and the number of functions determine this probability (unlike the classical case where we can't use all the input)...
- Concentration of measure can be used again (but on the hypercube this time) to show that this bound holds for almost all sets of functions.
    - In fact, the proof is easier as there is no difficulty with normalisation.

Oracle identification: quantum (2)

- So the results here can be used to put the same lower bound on the probability of success of distinguishing these states.
- And in particular, the input size and the number of functions determine this probability (unlike the classical case where we can't use all the input)...
- Concentration of measure can be used again (but on the hypercube this time) to show that this bound holds for almost all sets of functions.
  - In fact, the proof is easier as there is no difficulty with normalisation.
- When the probability of success is a constant $> 1/2$, we can repeat the algorithm a constant number of times for an arbitrarily good probability of success.

Summary and further work

- Good lower bounds have been obtained on the probability of distinguishing pure quantum states.
- These bounds can be applied to distinguishing random quantum states. For example:
  - For large n, n random states in n dimensions can be distinguished with probability > 0.72.
  - Almost all sets of 2^n Boolean functions on n bits can be distinguished with a constant number of quantum queries.

Summary and further work

- Good lower bounds have been obtained on the probability of distinguishing pure quantum states.
- These bounds can be applied to distinguishing random quantum states. For example:
  - For large n, n random states in n dimensions can be distinguished with probability > 0.72.
  - Almost all sets of 2^n Boolean functions on n bits can be distinguished with a constant number of quantum queries.

Possible future directions:

- Upper bounds on $P^{pgm}(\mathcal{E})$?
- Multiple copies?
- Further applications to quantum computation?

The End

- Further reading:
  "On the distinguishability of random quantum states"
  Communications in Mathematical Physics, to appear
  quant-ph/0607011

- Thanks for your time!