Uncertainty relations for high dimensional random unitary matrices

Radosław Adamczak

University of Warsaw

High Dimensional Probability VIII
Oaxaca 2017

Some information theory

- X - a random variable with values in a finite set A_X, with law p_X.
- **Shannon's entropy**

$$H(X) = H(p_X) = -\sum_{x \in A_X} p_X(x) \log p_X(x).$$

- **Conditional entropy**

$$\begin{aligned} H(Y|X) &= \sum_{x \in A_X} p_X(x)\Big(-\sum_{y \in A_Y} p_{Y|X}(y|x) \log(p_{Y|X}(y|x))\Big) \\ &= \mathbb{E} H(\mathbb{P}(Y \in \cdot|X)). \end{aligned}$$

- **Mutual information**

$$\begin{aligned} I(Y : X) &= H(Y) - H(Y|X) = H(X) - H(X|Y) \\ &= H(X) + H(Y) - H(X, Y). \end{aligned}$$

- **Mutual information**

$$
\begin{aligned}
I(X:Y) &= H(Y) - H(Y|X) = H(X) - H(X|Y) \\
&= H(X) + H(Y) - H(X,Y).
\end{aligned}
$$

- **A simple observation**

$$
I(X,Y:Z,Y) \leq I(X,Y:Z) + H(Y),
$$

  i.e. sending *k*-bits cannot increase the mutual information by more than *k*-bits.

Quantum setting

- **Pure states** – unit elements of a complex Hilbert space H (in our case of dimension d, $\simeq \mathbb{C}^d$)
- We identify a state $x \in H$ with the projection on $span(x)$, i.e. xx^*
- **Mixed states** - convex combinations of pure states, i.e. positive self-adjoint operators of trace one

$$\psi = \sum_{i=1}^{n} p_i x_i x_i^*. \quad |x_i| = 1, p_i \geq 0, \sum_{i=1}^{n} p_i = 1.$$

- **A measurement, POVM** – $\{P_i\}_{i \in I}$ – a collection of positive operators on $\mathbb{C}^d$, such that

$$\sum_{i \in I} P_i = Id.$$

A measurement on a system in state ψ gives output i with probability $p_\psi(i) = \mathrm{tr} P_i \psi$.

- **A measurement, POVM** – $\{P_i\}_{i \in I}$ – a collection of positive operators on $\mathbb{C}^d$, such that
$$\sum_{i \in I} P_i = Id.$$
A measurement on a system in state ψ gives output i with probability $p_\psi(i) = \mathrm{tr} P_i \psi$.
- **Nondegenerate von Neumann measurement**, $P_i = e_i e_i^*$, where $e_1, \ldots, e_d$ - an orthonormal basis. For a pure state x,
$$p_x(i) = |\langle x, e_i \rangle|^2, \quad i = 1, \ldots, d.$$

Bipartite systems

A system composed of two subsystems is described by a tensor product of corresponding Hilbert spaces. Typically:

- Alice has access to a part of the system (some particles) modeled on a Hilbert space H_A, $\dim H_A = d_A$
- Bob has access to the remaining part of the system – H_B, $\dim H_B = d_B$.
- The whole system is $H = H_A \otimes H_B$, with $\dim H = d_A d_B$.

Local measurements

- $\{P_i \otimes Q_j\}_{i \in I, j \in J}$
- Alice and Bob measure only their parts of the system, gives rise to a pair of random variables X, Y with values in I, J resp.

Local measurements

- $\{P_i \otimes Q_j\}_{i \in I, j \in J}$
- Alice and Bob measure only their parts of the system, gives rise to a pair of random variables X, Y with values in I, J resp.

Classical mutual information of a bipartite state ψ.

$$I_c(\psi) = \max_{(X,Y)} I(X : Y),$$

i.e. Alice and Bob measure their parts of the systems and one looks at the measurements which maximize the mutual information between their results.

Information locking (very informally)

DiVincenzo et. al. (2003) found a state $\psi \in \mathbb{C}^{2d} \otimes \mathbb{C}^{d}$ shared between Alice and Bob, s.t.

- if Alice sends to Bob a single bit (which changes the state $\psi \to \psi'$) the classical mutual information increases by $\frac{1}{2} \log d \overset{d\to\infty}{\to} \infty$, i.e.

$$I_c(\psi') - I_c(\psi) \geq \frac{1}{2} \log d.$$

- Physicists say that a single bit 'unlocks' $\frac{1}{2} \log d$ bits of correlation locked in ψ.
- This cannot happen in classical information theory.

A rough description of the protocol

- $\{U_1, U_2, \ldots, U_t\}$ – unitaries (specially chosen), $\{e_1, \ldots, e_d\}$ - orth. basis in $\mathbb{C}^d$.
- Alice chooses uniformly at random $k \in \{1, \ldots, t\}$ and $m \in \{1, \ldots, d\}$, prepares two systems, one in state e_m, the other in state $U_k e_m$ and sends the latter to Bob.
- If Bob doesn't know k, he can only say very little about (m, k). For $t = 2$:

$$I_c(\psi) \leq \frac{1}{2} \log d.$$

- If Bob knows k, he can invert U_k and measure m

$$I_c(\psi') = 1 + \log d.$$

Entropic uncertainty

For the construction to work, one needs a lower bound on

$$\min_{x\in\mathbb{C}^n,|x|=1} \frac{1}{t}\sum_{k=1}^{t} H(p_{U_k x}),$$

where $p_y = (p_y(1), \ldots, p_y(d))$ with

$$p_y(m) = |\langle y, e_m\rangle|^2.$$

Theorem (Maassen-Uffink)

U_1, U_2 – unitary matrices. Then

$$\min_{|x|=1} \frac{1}{2}\Big(H(p_{U_1 x}) + H(p_{U_2 x})\Big) \geq -\log c,$$

where $c = \max_{i,j\leq n} |\langle U_0 U_1^* e_i, e_j\rangle|$.

Theorem (Maassen-Uffink)

U_1, U_2 – unitary matrices. Then

$$\min_{|x|=1} \frac{1}{2}\Big(H(p_{U_1 x}) + H(p_{U_2 x})\Big) \geq -\log c,$$

where $c = \max_{i,j \leq n} |\langle U_0 U_1^* e_i, e_j\rangle|$.

Example

If U_1, U_2 are mutually unbiased (e.g. $c = 1/\sqrt{d}$), e.g. $U_1 = Id$, U_2 - Fourier, then

$$\min_{|x|=1} \frac{1}{2}\Big(H(p_{U_1 x}) + H(p_{U_2 x})\Big) \geq \frac{1}{2}\log d$$

This is best possible since for $x = U_1^{-1} e_1$, $H(p_{U_1 x}) = 0$ and $H(p_{U_2 x}) \leq \log d$.

What happens for general t?

Can you find $U_1, \dots, U_t$ such that

$$\Theta(d,t) := \min_{x \in \mathbb{C}^n, |x|=1} \frac{1}{t} \sum_{k=1}^{t} H(p_{U_k x}) \geq (1 - \frac{1}{t}) \log d?$$

- For $3 \leq t \leq \sqrt{d}$ mutually unbiased bases do not work (Balister-Wehner, Ambainis). You get again $\frac{1}{2} \log d$.
- For $t = d + 1$ you get (Ivanovic, Sanchez, 1992) $\Theta(d,t) \geq \log(d+1) - 1$.
- In general random constructions only
  - Hayden et al. (2004)

$$\Theta(d, \log^4 d) \geq \log d - O(1)$$

  - Fawzi-Hayden-Sen (2013)

$$\Theta(d,t) \geq \left(1 - \sqrt{\frac{O(1) \log t}{t}}\right) \log d - \log\left(\frac{t}{\log t}\right).$$

Theorem (Latała, Puchała, Życzkowski, A. 2014)

If $U_1, \ldots, U_t$ are i.i.d. (Haar) random unitary matrices, then with probability 1 - o(1), as d $\to \infty$,

$$\min_{x \in \mathbb{C}^n, |x|=1} \frac{1}{t} \sum_{k=1}^{t} H(p_{U_k x}) \geq \left(1 - \frac{1}{t}\right) \log d - C,$$

where C is a universal constant.

In particular this answers the question of Leung-Wehner-Winter (2009) about identifying for fixed t the limit

$$\liminf_{d \to \infty} \frac{1}{\log d} \max_{U_1, \ldots, U_t} \min_{x \in \mathbb{C}^n, |x|=1} \frac{1}{t} \sum_{k=1}^{t} H(p_{U_k x}),$$

which turns out to be $1 - 1/t$.

Sketch of proof

- **Majorization** $p = (p(1), \dots, p(N))$, $q = (q(1), \dots, q(N))$. We say that q majorizes p ($p \prec q$) if for all $k \leq N$,

$$\sum_{i=1}^{k} p^{\downarrow}(i) \leq \sum_{i=1}^{k} q^{\downarrow}(i),$$

  with equality for $k = N$.
- The function $p \mapsto F(p) = -\sum_{i=1}^{N} p(i) \log p(i)$ is **Schur concave**, i.e.

$$p \prec q \implies F(p) \geq F(q).$$

- **The main idea**: Find a sequence $q = (q(1), \dots, q(td))$ such that for all x,

$$p := p_{U_1 x} \oplus \cdots \oplus p_{U_t x} \prec q.$$

- An observation due to Rudnicki-Puchała-Życzkowski (2014)

$$\sum_{i=1}^{k} p^{\downarrow}(i) \leq s_k^2,$$

where s_k is the maximum operator norm of a matrix formed by choosing k columns from $[U_1^*|U_2^*|\dots|U_t^*]$.
- Standard concentration of measure + ϵ-net + union bound approach gives

$$s_k \leq 1 + C\sqrt{\frac{k}{d} \ln\left(\frac{edt}{k}\right)}$$

- This allows you to define $q(k) \simeq s_k^2 - s_{k-1}^2$. Estimating the 'entropy' of q ends the proof.

A different perspective - towards metric uncertainty relations

The inequality

$$\min_{x\in\mathbb{C}^n,|x|=1} \frac{1}{t}\sum_{k=0}^{t-1} H(p_{U_k x}) \geq \Big(1-\varepsilon\Big)\log d$$

can be rewritten as

$$\max_{x\in\mathbb{C}^n,|x|=1} \frac{1}{t}\sum_{k=0}^{t-1} d_{KL}\Big(p_{U_k x}, unif([d])\Big) \leq \varepsilon \log d,$$

where $d_{KL}(\nu,\mu) = \int \log(\frac{d\nu}{d\mu})d\nu$.

Question:

Can you replace d_{KL} with something else, e.g. the total variation or Hellinger distance?

Total variation uncertainty relations (Fawzi-Hayden-Sen)

A change of setting,

- a bipartite system $H = H_A \otimes H_B$, with $H_A = \mathbb{C}^{d_A}$, $H_B = \mathbb{C}^{d_B}$.
- $\{e_i\}_{i\in[d_A]}$, $\{f_j\}_{j\in[d_B]}$, $\{e_i \otimes f_j\}_{i\in[d_A],j\in[d_B]}$ - orth. bases in H_A, H_B, H.
- For $x \in H$, define $p_\psi^A = (p_\psi^A(1), \dots, p_\psi^A(d_A))$ by

$$p_x^A(i) = \sum_{j=1}^{d_B} |\langle x, e_i \otimes f_j \rangle|^2.$$

- $p_x^A(i)$ is the probability of getting outcome i, when measuring the A part of the system in the basis $e_1, \dots, e_{d_A}$.

Question

Can we find unitaries $U_1, \dots, U_t$ acting on H so that

$$\max_{x\in H, |x|=1} \frac{1}{t} \sum_{k=1}^{t} d_{TV}(p_{U_k x}^A, unif([d_A])) \le \varepsilon?$$

Question

Can we find unitaries $U_1, \ldots, U_t$ acting on H so that

$$\max_{x \in H, |x|=1} \frac{1}{t} \sum_{k=1}^{t} d_{TV}(p^A_{U_k x}, unif([d_A])) \leq \varepsilon?$$

Geometrically:

Can we find t decompositions of $\mathbb{C}^{d_A d_B}$ into d_A orthogonal subspaces of dimension d_B, such that for any x in most decompositions $|x|^2$ is evenly distributed among the subspaces?

Theorem (Fawzi-Hayden-Sen, 2013)

If $d_B \geq \frac{C}{\varepsilon^2}$ *and* $t \geq C \log(1/\varepsilon)/\varepsilon^2$ *and* $U_1, \ldots, U_t$ *are i.i.d. random unitary matrices, then with high probability*

$$\max_{x \in H, |x|=1} \frac{1}{t} \sum_{k=1}^{t} d_{TV}(p^A_{U_k x}, unif([d_A])) \leq \varepsilon \tag{1}$$

It is not difficult to eliminate $\log(1/\varepsilon)$ in the assumption on t.

Proposition (A. 2016)

If (1) *holds for some (deterministic) matrices* $U_1, \ldots, U_t$ *then* $d_B, t \geq c/\varepsilon^2$

- for d_{KL}, $t = O(1/\varepsilon)$, no need for H_B
- for d_{TV}, $t = O(1/\varepsilon^2)$, one needs an auxiliary system H_B.

Hellinger distance

- p, q - distributions on $\{1, \ldots, N\}$

$$d_H(p,q) = \sqrt{\sum_{i=1}^{N} (\sqrt{p(i)} - \sqrt{q(i)})^2}$$

- $$d_{TV}(p,q) \leq \sqrt{2} d_H(p,q)$$

Theorem (A. 2016)

If $t, d_B \geq C/\varepsilon^2$ and $U_1, \ldots, U_d$ are i.i.d. random unitaries, then with high probability

$$\max_{|x|=1} \sqrt{\frac{1}{t} \sum_{k=1}^{t} d_H\left(p^A_{U_k x}, unif([d_A])\right)^2} \leq \varepsilon.$$

Theorem (A. 2016)

If $t, d_B \geq C/\varepsilon^2$ and $U_1, \ldots, U_d$ are i.i.d. random unitaries, then with high probability

$$\max_{|x|=1} \sqrt{\frac{1}{t}\sum_{k=1}^{t} d_H\Big(p^A_{U_k x}, unif([d_A])\Big)^2} \leq \varepsilon.$$

- Adapting G. Schechtman's proof of (Gaussian) Dvoretzky theorem to random unitary matrices.
- $x \mapsto \sqrt{\frac{1}{t}\sum_{k=1}^{t} d_H\Big(p^A_{U_k x}, unif([d_A])\Big)^2}$ is subgaussian
- Comparison with a Gaussian process via the Majorizing measure theorem
- A byproduct: improved dependence on ε in Dvoretzky thm. for $\ell_1^n(\ell_2^m)$.
- Weaker conditions on t if restricting x to a subset,
- $t = 1$ is enough for separable states $x = x_A \otimes x_B$ (Applications to Quantum Data Hiding).

Thank you