Title: Hastings' additivity counterexample and a sharp version of Dvoretzky theorem

Date: Jul 06, 2010 01:30 PM

URL: http://pirsa.org/10070019

Abstract: In this talk we will explain how the main step technical steps in the proofs by Hastings and Hayden-Winter of the non-additivity of the minimal output von Neumann and p-Renyi entropy (for any $p>1$) can be reduced to a sharp version of Dvoretzky's theorem on almost spherical sections of convex bodies. This substantially simplifies their analysis, at least on the conceptual level, and provides an alternative point of view on these and related questions.

Joint work with G. Aubrun and E. Werner

Additivity conjectures for quantum channels and a sharp version of Dvoretzky's theorem

Stanislaw Szarek

Paris 6/Case Western Reserve

Perimeter Institute, July 6, 2010

Collaborators:
G. Aubrun, E. Werner

J. Math. Phys. 51, 022102 (2010); arxiv:1003.4925

http://www.math.jussieu.fr/~szarek/

Talk summary

Talk summary

- the setup : quantum channels as subspaces
- p-Rényi entropy and a link to Dvoretzky's theorem

Talk summary

- the setup : quantum channels as subspaces
- *p*-Rényi entropy and a link to Dvoretzky's theorem
- Dvoretzky's theorem and its various forms
- the Hayden-Winter counterexample

Talk summary

- the setup : quantum channels as subspaces
- p-Rényi entropy and a link to Dvoretzky's theorem
- Dvoretzky's theorem and its various forms
- the Hayden-Winter counterexample
- the Hastings's counterexample
- measure concentration: the union bound vs. chaining

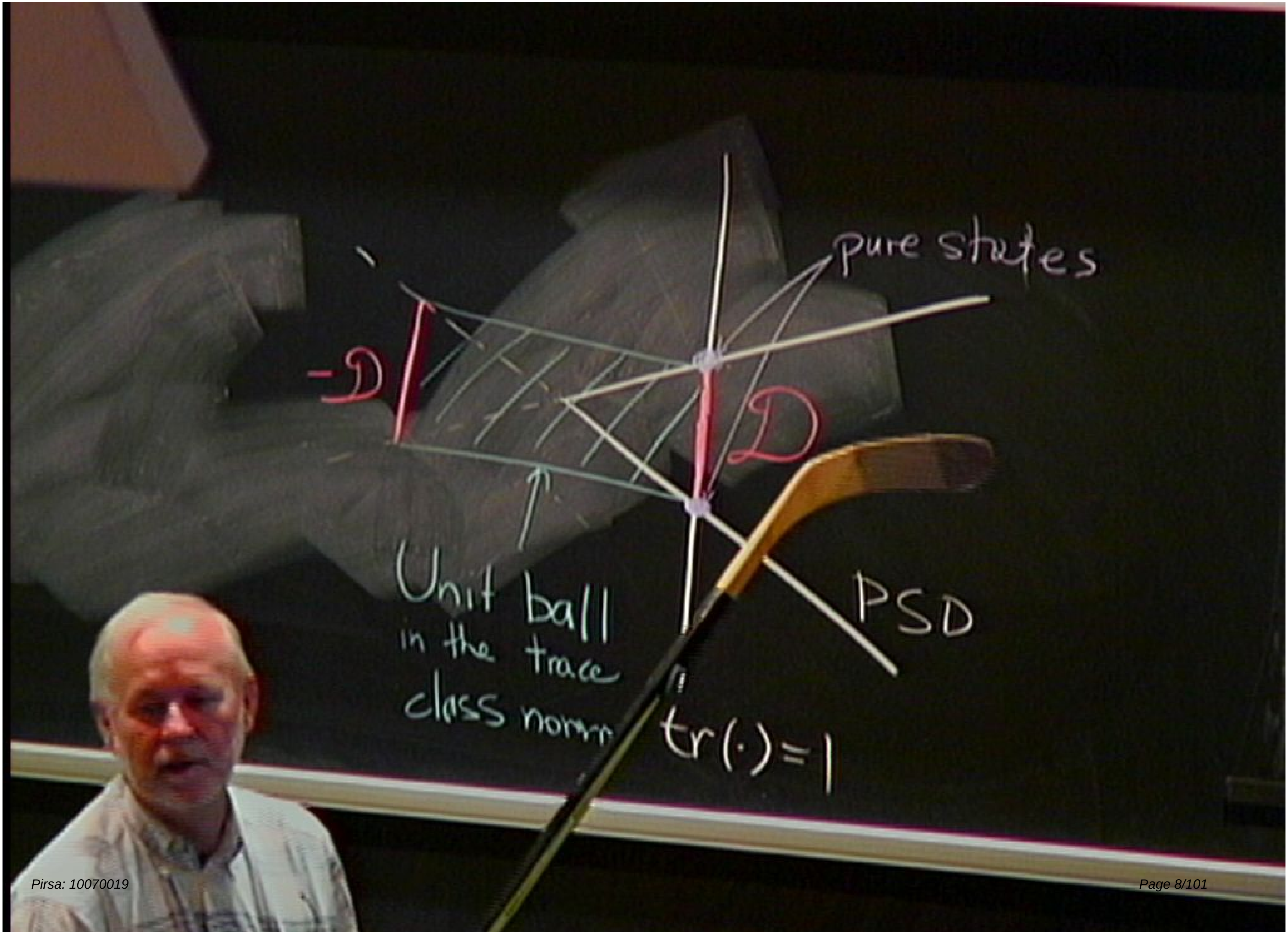
QI objects (geometric functional analysis angle)

- a complex Hilbert space $\mathcal{H}$, usually $\mathcal{H} = \mathbb{C}^d$
- the C^*-algebra $\mathcal{B}(\mathcal{H})$, $\mathcal{B}(\mathbb{C}^d) = \mathcal{M}_d$
- the real space $\mathcal{M}_d^{sa}$ of $d \times d$ Hermitian matrices
- the Schatten p-norm on $\mathcal{M}_d$ or $\mathcal{M}_d^{sa}$ or $\mathcal{M}_{d\times k}$

$$\|\sigma\|_p = \left(\mathrm{tr}(\sigma^\dagger \sigma)^{p/2}\right)^{1/p}$$

- the positive semi-definite cone $\mathcal{PSD} \subset \mathcal{M}_d^{sa}$
- $\mathcal{D} = \mathcal{D}(\mathcal{H})$, the set of states of $\mathcal{B}(\mathcal{H})$, or of density matrices
  - ◇ $\mathcal{D}(\mathcal{H}) = \mathcal{PSD} \cap \{\mathrm{tr}(\cdot) = 1\}$, or the base of $\mathcal{PSD}$
  - ◇ $\mathcal{D}(\mathcal{H}) =$ the positive face of the unit ball in the trace class (Schatten 1-norm)
  - ◇ $\mathcal{D}(\mathcal{H}) = \mathrm{conv}\{|\psi\rangle\langle\psi| : \psi \in \mathcal{H}, |\psi| = 1\}$

- $\mathrm{conv}(-\mathcal{D} \cup \mathcal{D})$ = the unit ball of the trace class in $\mathcal{M}^{sa}$


pure states
-D
D
Unit ball
in the trace
class norm
PSD
tr(·)=1

QI morphisms : quantum operations, or channels

Completely positive (CP) maps $\Phi : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$, or $\Phi : \mathcal{M}_m \to \mathcal{M}_d$, usually also required to be trace preserving (TP)

Fact 0 : For $\Phi : \mathcal{M}_m \to \mathcal{M}_d$ T.F.A.E. (Stinespring-Kraus-Choi)

- Φ is CPTP
- for some $B_1, \ldots, B_k \in \mathcal{M}_{d\times m}$ with $\sum_i B_i^\dagger B_i = \mathrm{Id}_{\mathbb{C}^d}$

$$\Phi(\rho) = \sum_i B_i \rho B_i^\dagger$$

- for some isometry $V : \mathbb{C}^m \to \mathbb{C}^d \otimes \mathbb{C}^k$

$$\Phi(\rho) = \mathrm{tr}_{\mathbb{C}^k}(V\rho V^\dagger) = \mathrm{tr}_2(V\rho V^\dagger)$$

The relation between V and B_j's: $V = \sum_i B_i \otimes e_i$

QI objects (geometric functional analysis angle)

- a complex Hilbert space $\mathcal{H}$, usually $\mathcal{H} = \mathbb{C}^d$
- the C^*-algebra $\mathcal{B}(\mathcal{H})$, $\mathcal{B}(\mathbb{C}^d) = \mathcal{M}_d$
- the real space $\mathcal{M}_d^{sa}$ of $d \times d$ Hermitian matrices
- the Schatten p-norm on $\mathcal{M}_d$ or $\mathcal{M}_d^{sa}$ or $\mathcal{M}_{d\times k}$

$$\|\sigma\|_p = \left(\mathrm{tr}(\sigma^\dagger \sigma)^{p/2}\right)^{1/p}$$

- the positive semi-definite cone $\mathcal{PSD} \subset \mathcal{M}_d^{sa}$
- $\mathcal{D} = \mathcal{D}(\mathcal{H})$, the set of states of $\mathcal{B}(\mathcal{H})$, or of density matrices
  - ◇ $\mathcal{D}(\mathcal{H}) = \mathcal{PSD} \cap \{\mathrm{tr}(\cdot) = 1\}$, or the base of $\mathcal{PSD}$
  - ◇ $\mathcal{D}(\mathcal{H}) =$ the positive face of the unit ball in the trace class (Schatten 1-norm)
  - ◇ $\mathcal{D}(\mathcal{H}) = \mathrm{conv}\{|\psi\rangle\langle\psi| : \psi \in \mathcal{H}, |\psi| = 1\}$

- $\mathrm{conv}(-\mathcal{D} \cup \mathcal{D}) =$ the unit ball of the trace class in $\mathcal{M}_d^{sa}$

QI morphisms : quantum operations, or channels

Completely positive (CP) maps $\Phi : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$, or $\Phi : \mathcal{M}_m \to \mathcal{M}_d$, usually also required to be trace preserving (TP)

Fact 0 : For $\Phi : \mathcal{M}_m \to \mathcal{M}_d$ T.F.A.E. (Stinespring-Kraus-Choi)

- Φ is CPTP
- for some $B_1, \ldots, B_k \in \mathcal{M}_{d\times m}$ with $\sum_i B_i^\dagger B_i = \mathrm{Id}_{\mathbb{C}^d}$

$$\Phi(\rho) = \sum_i B_i \rho B_i^\dagger$$

- for some isometry $V : \mathbb{C}^m \to \mathbb{C}^d \otimes \mathbb{C}^k$

$$\Phi(\rho) = \mathrm{tr}_{\mathbb{C}^k}(V\rho V^\dagger) = \mathrm{tr}_2(V\rho V^\dagger)$$

The relation between V and B_j's: $V = \sum_i B_i \otimes e_i$

Channels as subspaces, range of a channel

Quantum operations $\Phi : \mathbb{C}^m \to \mathbb{C}^d$ are really m-dimensional subspaces $\mathcal{W} = V(\mathbb{C}^d) \subset \mathbb{C}^d \otimes \mathbb{C}^k$

Channels as subspaces, range of a channel

Quantum operations $\Phi : \mathbb{C}^m \to \mathbb{C}^d$ are really m-dimensional subspaces $\mathcal{W} = V(\mathbb{C}^d) \subset \mathbb{C}^d \otimes \mathbb{C}^k$

The isometry V is not important: corresponds to fixing a basis of $\mathcal{W}$

Channels as subspaces, range of a channel

Quantum operations $\Phi : \mathbb{C}^m \to \mathbb{C}^d$ are really m-dimensional subspaces $\mathcal{W} = V(\mathbb{C}^d) \subset \mathbb{C}^d \otimes \mathbb{C}^k$

The isometry V is not important: corresponds to fixing a basis of $\mathcal{W}$
Alternatively, one may consider $\Phi : \mathcal{B}(\mathcal{W}) \to \mathcal{B}(\mathbb{C}^d)$, with $\Phi = \mathrm{tr}_2$

Channels as subspaces, range of a channel

Quantum operations $\Phi : \mathbb{C}^m \to \mathbb{C}^d$ are really m-dimensional subspaces $\mathcal{W} = V(\mathbb{C}^d) \subset \mathbb{C}^d \otimes \mathbb{C}^k$

The isometry V is not important: corresponds to fixing a basis of $\mathcal{W}$
Alternatively, one may consider $\Phi : \mathcal{B}(\mathcal{W}) \to \mathcal{B}(\mathbb{C}^d)$, with $\Phi = \mathrm{tr}_2$

For a pure state $\varphi = V\psi \in \mathcal{W}$, its image by Φ is simply encoded in its "Schmidt decomposition" :

If $\varphi = \sum_j s_j \, u_j \otimes v_j$, then $\Phi(|\psi\rangle\langle\psi|) = \mathrm{tr}_2(|\varphi\rangle\langle\varphi|) = \sum_j s_j^2 \, |u_j\rangle\langle u_j|$

Verification:

$$\mathrm{tr}_2(|\varphi\rangle\langle\varphi|) \;=\; \mathrm{tr}_2\Big(|\sum_i s_i\, u_i \otimes v_i\rangle\langle\sum_j s_j\, u_j \otimes v_j|\Big)$$

Channels as subspaces, range of a channel

Quantum operations $\Phi : \mathbb{C}^m \to \mathbb{C}^d$ are really m-dimensional subspaces $\mathcal{W} = V(\mathbb{C}^d) \subset \mathbb{C}^d \otimes \mathbb{C}^k$

The isometry V is not important: corresponds to fixing a basis of $\mathcal{W}$
Alternatively, one may consider $\Phi : \mathcal{B}(\mathcal{W}) \to \mathcal{B}(\mathbb{C}^d)$, with $\Phi = \mathrm{tr}_2$

For a pure state $\varphi = V\psi \in \mathcal{W}$, its image by Φ is simply encoded in its "Schmidt decomposition" :

If $\varphi = \sum_j s_j\, u_j \otimes v_j$, then $\Phi(|\psi\rangle\langle\psi|) = \mathrm{tr}_2(|\varphi\rangle\langle\varphi|) = \sum_j s_j^2\, |u_j\rangle\langle u_j|$

Verification:

$$\mathrm{tr}_2(|\varphi\rangle\langle\varphi|) \;=\; \mathrm{tr}_2\Big(|\sum_i s_i\, u_i \otimes v_i\rangle\langle\sum_j s_j\, u_j \otimes v_j|\Big)$$

Verification:

$$
\begin{aligned}
\mathrm{tr}_2(|\varphi\rangle\langle\varphi|) &= \mathrm{tr}_2\Big(|\sum_i s_i\, u_i \otimes v_i\rangle\langle\sum_j s_j\, u_j \otimes v_j|\Big) \\
&= \sum_{i,j} s_i s_j\, |u_i\rangle\langle u_j|\, \mathrm{tr}(|v_i\rangle\langle v_j|) \\
&= \sum_j s_j^2\, |u_j\rangle\langle u_j|
\end{aligned}
$$

Recall: $(u_j), (v_j)$ are orthonormal sequences in $\mathbb{C}^d$ and $\mathbb{C}^k$

Recapitulation in linear algebra terms

Φ associated to $\mathcal{W} \subset \mathbb{C}^d \otimes \mathbb{C}^k$

Recapitulation in linear algebra terms

Φ associated to $\mathcal{W} \subset \mathbb{C}^d \otimes \mathbb{C}^k \sim \mathcal{M}_{d\times k}$

The identification $\sim$ is induced by $|u\rangle \otimes |v\rangle \sim |u\rangle\langle v|$

Unit vector $\varphi \sim$ matrix A with $\|A\|_2 = 1$

Recapitulation in linear algebra terms

Φ associated to $\mathcal{W} \subset \mathbb{C}^d \otimes \mathbb{C}^k \sim \mathcal{M}_{d\times k}$

The identification $\sim$ is induced by $|u\rangle \otimes |v\rangle \sim |u\rangle\langle v|$

Unit vector $\varphi \sim$ matrix A with $\|A\|_2 = 1$

Schmidt decomposition $\varphi = \sum_j s_j\, u_j \otimes v_j \sim$ SVD $A = \sum_j s_j\, |u_j\rangle\langle v_j|$

Recapitulation in linear algebra terms

Φ associated to $\mathcal{W} \subset \mathbb{C}^d \otimes \mathbb{C}^k \sim \mathcal{M}_{d\times k}$

The identification $\sim$ is induced by $|u\rangle \otimes |v\rangle \sim |u\rangle\langle v|$

Unit vector $\varphi \sim$ matrix A with $\|A\|_2 = 1$

Schmidt decomposition $\varphi = \sum_j s_j \, u_j \otimes v_j \sim$ SVD $A = \sum_j s_j \, |u_j\rangle\langle v_j|$

The output $\quad \Phi(|\varphi\rangle\langle\varphi|) = \sum_j s_j^2 \, |u_j\rangle\langle u_j| = AA^\dagger$

The bottom line: To understand quantum channels, we need to understand the patterns of singular numbers of A as A varies over an m-dimensional subspace $\mathcal{W}$ of the space of $d \times k$ matrices

The additivity conjecture

Von Neumann entropy of a state ρ:

$$S(\rho) = -\mathrm{tr}(\rho \log \rho)$$

The additivity conjecture

Von Neumann entropy of a state ρ:

$$S(\rho) = -\mathrm{tr}(\rho \log \rho) = \sum_j q_j \log(1/q_j)$$

if q_j's are eigenvalues of ρ

Minimum output entropy of a channel Φ:

$$S^{\min}(\Phi) := \min_{\rho \in \mathcal{D}(\mathbb{C}^m)} S(\Phi(\rho))$$

By concavity of $S(\cdot)$, $S^{\min}$ is always attained on a pure state.

The additivity conjecture

Von Neumann entropy of a state ρ:

$$S(\rho) = -\mathrm{tr}(\rho \log \rho) = \sum_j q_j \log(1/q_j)$$

if q_j's are eigenvalues of ρ

Minimum output entropy of a channel Φ:

$$S^{\min}(\Phi) := \min_{\rho \in \mathcal{D}(\mathbb{C}^m)} S(\Phi(\rho))$$

By concavity of $S(\cdot)$, $S^{\min}$ is always attained on a pure state.

Additivity conjecture: For CPTP maps Φ, Ψ, do we have

$$S_{\min}(\Phi \otimes \Psi) \stackrel{?}{=} S_{\min}(\Phi) + S_{\min}(\Psi)$$

"No" (Hastings 2009)

Rényi entropy and multiplicativity problems

Rényi entropy and multiplicativity problems

Additivity of the minimum output entropy would follow from additivity of the minimum output p-Rényi entropy

$$S_p^{\min}(\Phi) := \min_{\rho \in \mathcal{D}(\mathbb{C}^m)} S_p(\Phi(\rho))$$

for $p > 1$, where $S_p(\sigma) := \frac{1}{1-p} \log(\mathrm{tr}\sigma^p)$

Rényi entropy and multiplicativity problems

Additivity of the minimum output entropy would follow from additivity of the minimum output p-Rényi entropy

$$S_p^{\min}(\Phi) := \min_{\rho \in \mathcal{D}(\mathbb{C}^m)} S_p(\Phi(\rho))$$

for $p > 1$, where $S_p(\sigma) := \frac{1}{1-p} \log(\mathrm{tr}\sigma^p) = \frac{p}{1-p} \log \|\sigma\|_p$
(let $p \to 1$)

Modulo normalizing factors and logarithmic change of variables, $S_p^{\min}(\Phi)$ is equivalent to $\max_{\rho \in \mathcal{D}(\mathbb{C}^m)} \|\Phi(\rho)\|_p$, or $\|\Phi\|_{1 \to p}$

Rényi entropy and multiplicativity problems

Additivity of the minimum output entropy would follow from additivity of the minimum output p-Rényi entropy

$$S_p^{\min}(\Phi) := \min_{\rho \in \mathcal{D}(\mathbb{C}^m)} S_p(\Phi(\rho))$$

for $p > 1$, where $S_p(\sigma) := \frac{1}{1-p} \log(\mathrm{tr}\sigma^p) = \frac{p}{1-p} \log \|\sigma\|_p$
(let $p \to 1$)

Modulo normalizing factors and logarithmic change of variables, $S_p^{\min}(\Phi)$ is equivalent to $\max_{\rho \in \mathcal{D}(\mathbb{C}^m)} \|\Phi(\rho)\|_p$, or $\|\Phi\|_{1\to p}$
(again, attained on a pure state).

Additivity of $S_p^{\min}(\Phi)$ is equivalent to multiplicativity of $\|\Phi\|_{1\to p}$
"No" (Hayden-Winter 2008)

Focus on $\|\Phi\|_{1\to p}$

Let $\mathcal{W}$ be the m-dimensional subspace of $\mathcal{M}_{d\times k}$ associated with Φ

Focus on $\|\Phi\|_{1\to p}$

Let $\mathcal{W}$ be the m-dimensional subspace of $\mathcal{M}_{d\times k}$ associated with Φ

$$\|\Phi\|_{1\to p} = \max_{A\in\mathcal{W},\ \|A\|_2=1} \|AA^\dagger\|_p = \|A\|_{2p}^2$$

In other words

$$\|\Phi\|_{1\to p}^{1/2} = \max_{A\in\mathcal{W}\setminus\{0\}} \frac{\|A\|_{2p}}{\|A\|_2}$$

Conference "Perspectives in High Dimensions"

Cleveland, August 2 until August 6, 2010

http://www.case.edu/artsci/math/perspectivesInHighDimensions/

Dvoretzky's theorem

High-dimensional convex bodies have almost spherical sections of large dimension.

Dvoretzky's theorem

High-dimensional convex bodies have almost spherical sections of large dimension.

Given $m \in \mathbb{N}$ *and* $\varepsilon > 0$ *there is* $N = N(m, \varepsilon)$ *such that, for any norm on* $\mathbb{R}^N$ *(or* $\mathbb{C}^N$*) there is an m-dimensional subspace on which the ratio between that norm and the Euclidean norm is (approximately) constant, up to a multiplicative factor* $1 + \varepsilon$.

Milman's "tangible" version of Dvoretzky's theorem

Milman's "tangible" version of Dvoretzky's theorem

Fact 1 : Consider the N-dimensional Euclidean space (over $\mathbb{R}$ or $\mathbb{C}$) endowed with the Euclidean norm $|\cdot|$ and some other norm $\|\cdot\|$ such that, for some $b > 0$, $\|\cdot\| \leq b|\cdot|$. Denote $M = \mathbb{E}\|X\|$, where X is a random variable uniformly distributed on the unit Euclidean sphere. Let $\varepsilon > 0$ and let $m \leq c\varepsilon^2(M/b)^2N$, where $c > 0$ is an appropriate (computable) universal constant. Then, for most m-dimensional subspaces E we have

$$\forall x \in E, \quad (1-\varepsilon)M|x| \leq \|x\| \leq (1+\varepsilon)M|x|.$$

Milman's "tangible" version of Dvoretzky's theorem

Fact 1 : Consider the N-dimensional Euclidean space (over $\mathbb{R}$ or $\mathbb{C}$) endowed with the Euclidean norm $|\cdot|$ and some other norm $\|\cdot\|$ such that, for some $b > 0$, $\|\cdot\| \leq b|\cdot|$. Denote $M = \mathbb{E}\|X\|$, where X is a random variable uniformly distributed on the unit Euclidean sphere. Let $\varepsilon > 0$ and let $m \leq c\varepsilon^2(M/b)^2N$, where $c > 0$ is an appropriate (computable) universal constant. Then, for most m-dimensional subspaces E we have

$$\forall x \in E, \quad (1-\varepsilon)M|x| \leq \|x\| \leq (1+\varepsilon)M|x|.$$

A similar statement holds for Lipschitz functions in place of norms.

Dvoretzky's theorem for Schatten classes (FLM '77)

For the Schatten norm $\|\cdot\|_q$ with $q = 2p > 2$

Dvoretzky's theorem for Schatten classes (FLM '77)

For the Schatten norm $\|\cdot\|_q$ with $q = 2p > 2$, $k = d$ and $\varepsilon = \frac{1}{2}$ we get $b = 1$ and $M \sim d^{1/q-1/2}$ (next slide), hence when

$$m \sim (M/b)^2 N = M^2 d^2 \sim \left(d^{1/q-1/2}\right)^2 d^2 = d^{1+2/q} = d^{1+1/p}$$

Dvoretzky's theorem for Schatten classes (FLM '77)

For the Schatten norm $\|\cdot\|_q$ with $q = 2p > 2$, $k = d$ and $\varepsilon = \frac{1}{2}$ we get $b = 1$ and $M \sim d^{1/q-1/2}$ (next slide), hence when

$$m \sim (M/b)^2 N = M^2 d^2 \sim \left(d^{1/q-1/2}\right)^2 d^2 = d^{1+2/q} = d^{1+1/p},$$

then for a generic m-dimensional subspace $\mathcal{W}$ of $\mathcal{M}_d$

$$\forall A \in \mathcal{W} \quad d^{1/q-1/2}\|A\|_2 \leq \|A\|_q \leq C d^{1/q-1/2}\|A\|_2$$

Dvoretzky's theorem for Schatten classes (FLM '77)

For the Schatten norm $\|\cdot\|_q$ with $q = 2p > 2$, $k = d$ and $\varepsilon = \frac{1}{2}$ we get $b = 1$ and $M \sim d^{1/q-1/2}$ (next slide), hence when

$$m \sim (M/b)^2 N = M^2 d^2 \sim \left(d^{1/q-1/2}\right)^2 d^2 = d^{1+2/q} = d^{1+1/p},$$

then for a generic m-dimensional subspace $\mathcal{W}$ of $\mathcal{M}_d$

$$\forall A \in \mathcal{W} \quad d^{1/q-1/2}\|A\|_2 \leq \|A\|_q \leq C d^{1/q-1/2}\|A\|_2$$

Accordingly, for the associated (random) channel Φ

$$\|\Phi\|_{1\to p} = \left(\max_{A\in\mathcal{W}} \frac{\|A\|_{2p}}{\|A\|_2}\right)^2$$

Dvoretzky's theorem for Schatten classes (FLM '77)

For the Schatten norm $\|\cdot\|_q$ with $q = 2p > 2$, $k = d$ and $\varepsilon = \frac{1}{2}$ we get $b = 1$ and $M \sim d^{1/q-1/2}$ (next slide), hence when

$$m \sim (M/b)^2 N = M^2 d^2 \sim \left(d^{1/q-1/2}\right)^2 d^2 = d^{1+2/q} = d^{1+1/p},$$

then for a generic m-dimensional subspace $\mathcal{W}$ of $\mathcal{M}_d$

$$\forall A \in \mathcal{W} \quad d^{1/q-1/2}\|A\|_2 \leq \|A\|_q \leq C d^{1/q-1/2}\|A\|_2$$

Accordingly, for the associated (random) channel Φ

$$\|\Phi\|_{1\to p} = \left(\max_{A\in\mathcal{W}} \frac{\|A\|_{2p}}{\|A\|_2}\right)^2 \leq \left(C d^{1/q-1/2}\right)^2 = C^2 d^{1/p-1}$$

which is $\ll 1$ for large d and nearly as small as it can be: $\|\Phi\|_{1\to p} \geq d^{1/p-1}$ always.

Dvoretzky's theorem for Schatten classes (FLM '77)

For the Schatten norm $\|\cdot\|_q$ with $q = 2p > 2$, $k = d$ and $\varepsilon = \frac{1}{2}$ we get $b = 1$ and $M \sim d^{1/q-1/2}$ (next slide), hence when

$$m \sim (M/b)^2 N = M^2 d^2 \sim \left(d^{1/q-1/2}\right)^2 d^2 = d^{1+2/q} = d^{1+1/p},$$

then for a generic m-dimensional subspace $\mathcal{W}$ of $\mathcal{M}_d$

$$\forall A \in \mathcal{W} \quad d^{1/q-1/2}\|A\|_2 \leq \|A\|_q \leq Cd^{1/q-1/2}\|A\|_2$$

Accordingly, for the associated (random) channel Φ

$$\|\Phi\|_{1\to p} = \left(\max_{A\in\mathcal{W}} \frac{\|A\|_{2p}}{\|A\|_2}\right)^2 \leq \left(Cd^{1/q-1/2}\right)^2 = C^2 d^{1/p-1}$$

which is $\ll 1$ for large d and nearly as small as it can be: $\|\Phi\|_{1\to p} \geq d^{1/p-1}$ always.

The counterexample follows by showing that for the composite channel $\Phi \otimes \overline{\Phi}$ we have a nontrivial lower bound $\frac{m}{dk} = \frac{m}{d^2} \sim d^{1/p-1}$.

Why $M \sim d^{1/q-1/2}$?

If $q = \infty$, $\|\cdot\|_\infty = \|\cdot\|_{op}$, so $\mathbb{E}\|X\|_{op} \sim 2d^{-1/2}$

Why $M \sim d^{1/q-1/2}$?

If $q = \infty$, $\|\cdot\|_\infty = \|\cdot\|_{op}$, so $\mathbb{E}\|X\|_{op} \sim 2d^{-1/2}$

- 2 is the same as in the Wigner semi-circle law
- $d^{-1/2}$ is due to the normalization in the Frobenius norm $\|\cdot\|_2$

Why $M \sim d^{1/q-1/2}$?

If $q = \infty$, $\|\cdot\|_\infty = \|\cdot\|_{op}$, so $\mathbb{E}\|X\|_{op} \sim 2d^{-1/2}$

- 2 is the same as in the Wigner semi-circle law
- $d^{-1/2}$ is due to the normalization in the Frobenius norm $\|\cdot\|_2$

Obviously $\mathbb{E}\|X\|_2 = 1$

For $q \in (2, \infty)$ we interpolate (Hölder inequality)

The counterexample to additivity

of $S^{\min}(\cdot)$ is more subtle (follows Brandao-Horodecki).
The analysis of a single random channel is based on two facts

The counterexample to additivity

of $S^{\min}(\cdot)$ is more subtle (follows Brandao-Horodecki).
The analysis of a single random channel is based on two facts

Fact 2 : $\forall \sigma \in \mathcal{D}(\mathbb{C}^d)$ $S(\sigma) \geq S\left(\frac{\mathrm{Id}}{d}\right) - d\left\|\sigma - \frac{\mathrm{Id}}{d}\right\|_2^2$

Consequently $\forall \Phi \ : \mathcal{M}_m \to \mathcal{M}_d$

$$S^{\min}(\Phi) \geq \log(d) - d \cdot \max_{\rho \in \mathcal{D}(\mathbb{C}^d)} \left\|\Phi(\rho) - \frac{\mathrm{Id}}{d}\right\|_2^2$$

The counterexample to additivity

of $S^{\min}(\cdot)$ is more subtle (follows Brandao-Horodecki).
The analysis of a single random channel is based on two facts

Fact 2 : $\forall \sigma \in \mathcal{D}(\mathbb{C}^d)$ $S(\sigma) \geq S\left(\frac{\mathrm{Id}}{d}\right) - d\left\|\sigma - \frac{\mathrm{Id}}{d}\right\|_2^2$

Consequently $\forall \Phi \; : \mathcal{M}_m \to \mathcal{M}_d$

$$S^{\min}(\Phi) \geq \log(d) - d \cdot \max_{\rho \in \mathcal{D}(\mathbb{C}^d)} \left\|\Phi(\rho) - \frac{\mathrm{Id}}{d}\right\|_2^2$$

This reduces the study of the not-so-regular and somewhat involved quantity $S^{\min}(\cdot)$ to upper-bounding $\left\|\sigma - \frac{\mathrm{Id}}{d}\right\|_2$ for σ in the range of Φ

Fact 3 : If $k \sim d^2$, $m \sim d^2$, then, for a typical m-dimensional subspace $\mathcal{W} \subset \mathcal{M}_{d\times k}$,

$$\max_{A\in\mathcal{W}, \|A\|_2=1} \left\| AA^\dagger - \frac{\mathrm{Id}}{d} \right\|_2 \leq \frac{C}{d}$$

The counterexample to additivity

of $S^{\min}(\cdot)$ is more subtle (follows Brandao-Horodecki).
The analysis of a single random channel is based on two facts

Fact 2 : $\forall \sigma \in \mathcal{D}(\mathbb{C}^d)$ $S(\sigma) \geq S\left(\frac{\mathrm{Id}}{d}\right) - d \left\|\sigma - \frac{\mathrm{Id}}{d}\right\|_2^2$

Consequently $\forall \Phi \ : \mathcal{M}_m \to \mathcal{M}_d$

$$S^{\min}(\Phi) \geq \log(d) - d \cdot \max_{\rho \in \mathcal{D}(\mathbb{C}^d)} \left\|\Phi(\rho) - \frac{\mathrm{Id}}{d}\right\|_2^2$$

This reduces the study of the not-so-regular and somewhat involved quantity $S^{\min}(\cdot)$ to upper-bounding $\left\|\sigma - \frac{\mathrm{Id}}{d}\right\|_2$ for σ in the range of Φ

Fact 3 : If $k \sim d^2$, $m \sim d^2$, then, for a typical m-dimensional subspace $\mathcal{W} \subset \mathcal{M}_{d\times k}$,

$$\max_{A\in\mathcal{W}, \|A\|_2=1} \left\| AA^\dagger - \frac{\mathrm{Id}}{d} \right\|_2 \le \frac{C}{d}$$

Fact 3 : If $k \sim d^2, m \sim d^2$, then, for a typical m-dimensional subspace $\mathcal{W} \subset \mathcal{M}_{d\times k}$,

$$\max_{A\in\mathcal{W}, \|A\|_2=1} \left\| AA^\dagger - \frac{\mathrm{Id}}{d} \right\|_2 \leq \frac{C}{d}$$

Recall: $AA^\dagger = |A|^2 = \Phi(|\varphi\rangle\langle\varphi|)$, where φ is the unit vector corresponding to A and Φ is the channel associated to $\mathcal{W}$.

Fact 3 : If $k \sim d^2, m \sim d^2$, then, for a typical m-dimensional subspace $\mathcal{W} \subset \mathcal{M}_{d\times k}$,

$$\max_{A\in\mathcal{W}, \|A\|_2=1} \left\| AA^\dagger - \frac{\mathrm{Id}}{d} \right\|_2 \leq \frac{C}{d}$$

Recall: $AA^\dagger = |A|^2 = \Phi(|\varphi\rangle\langle\varphi|)$, where φ is the unit vector corresponding to A and Φ is the channel associated to $\mathcal{W}$.

Combining the estimates

$$S^{\min}(\Phi) \geq \log(d) - d\left(\frac{C}{d}\right)^2 = \log(d) - O\left(\frac{1}{d}\right)$$

Fact 3 : If $k \sim d^2, m \sim d^2$, then, for a typical m-dimensional subspace $\mathcal{W} \subset \mathcal{M}_{d\times k}$,

$$\max_{A\in\mathcal{W},\|A\|_2=1} \left\| AA^\dagger - \frac{\mathrm{Id}}{d} \right\|_2 \leq \frac{C}{d}$$

Recall: $AA^\dagger = |A|^2 = \Phi(|\varphi\rangle\langle\varphi|)$, where φ is the unit vector corresponding to A and Φ is the channel associated to $\mathcal{W}$.

Combining the estimates

$$S^{\min}(\Phi) \geq \log(d) - d\left(\frac{C}{d}\right)^2 = \log(d) - O\left(\frac{1}{d}\right)$$

On the other hand, the "large subspace/large eigenvalue" argument gives for the composite channel

$$S^{\min}(\Phi\otimes\bar{\Phi}) \leq \log(d^2) - \Omega\left(\frac{\log d}{d}\right)$$

Focus on Fact 3

For $A \in \mathcal{W}, \|A\|_2 = 1$, we have

$$\frac{C^2}{d^2} \geq \left\| |A|^2 - \frac{\mathrm{Id}}{d} \right\|_2^2$$

Focus on Fact 3

For $A \in \mathcal{W}$, $\|A\|_2 = 1$, we have

$$\frac{C^2}{d^2} \geq \left\| |A|^2 - \frac{\mathrm{Id}}{d} \right\|_2^2 = \mathrm{tr}|A|^4 - \frac{2\,\mathrm{tr}|A|^2}{d} + \frac{\mathrm{tr}\mathrm{Id}}{d^2} = \mathrm{tr}|A|^4 - \frac{1}{d} \geq 0$$

Consequently,

$$d^{-1/4}\|A\|_2 \leq \|A\|_4 \leq d^{-1/4}\left(1 + \frac{C^2}{d}\right)^{1/4}\|A\|_2 \leq d^{-1/4}\left(1 + \frac{C^2}{4d}\right)\|A\|_2$$

Focus on Fact 3

For $A \in \mathcal{W}$, $\|A\|_2 = 1$, we have

$$\frac{C^2}{d^2} \geq \left\| |A|^2 - \frac{\mathrm{Id}}{d} \right\|_2^2 = \mathrm{tr}|A|^4 - \frac{2\,\mathrm{tr}|A|^2}{d} + \frac{\mathrm{tr}\mathrm{Id}}{d^2} = \mathrm{tr}|A|^4 - \frac{1}{d} \geq 0$$

Consequently,

$$d^{-1/4}\|A\|_2 \leq \|A\|_4 \leq d^{-1/4}\left(1+\frac{C^2}{d}\right)^{1/4}\|A\|_2 \leq d^{-1/4}\left(1+\frac{C^2}{4d}\right)\|A\|_2$$

Focus on Fact 3

For $A \in \mathcal{W}$, $\|A\|_2 = 1$, we have

$$\frac{C^2}{d^2} \geq \left\| |A|^2 - \frac{\mathrm{Id}}{d} \right\|_2^2 = \mathrm{tr}|A|^4 - \frac{2\,\mathrm{tr}|A|^2}{d} + \frac{\mathrm{tr}\mathrm{Id}}{d^2} = \mathrm{tr}|A|^4 - \frac{1}{d} \geq 0$$

Consequently,

$$d^{-1/4}\|A\|_2 \leq \|A\|_4 \leq d^{-1/4}\left(1 + \frac{C^2}{d}\right)^{1/4}\|A\|_2 \leq d^{-1/4}\left(1 + \frac{C^2}{4d}\right)\|A\|_2$$

Focus on Fact 3

For $A \in \mathcal{W}$, $\|A\|_2 = 1$, we have

$$\frac{C^2}{d^2} \geq \left\| |A|^2 - \frac{\mathrm{Id}}{d} \right\|_2^2 = \mathrm{tr}|A|^4 - \frac{2\,\mathrm{tr}|A|^2}{d} + \frac{\mathrm{tr}\mathrm{Id}}{d^2} = \mathrm{tr}|A|^4 - \frac{1}{d} \geq 0$$

Consequently,

$$d^{-1/4}\|A\|_2 \leq \|A\|_4 \leq d^{-1/4}\left(1+\frac{C^2}{d}\right)^{1/4}\|A\|_2 \leq d^{-1/4}\left(1+\frac{C^2}{4d}\right)\|A\|_2$$

In other words, $\mathcal{W}$ is $1 + O(\frac{1}{d})$-Euclidean as a subspace of the Schatten 4-class.

Focus on Fact 3

For $A \in \mathcal{W}$, $\|A\|_2 = 1$, we have

$$\frac{C^2}{d^2} \geq \left\| |A|^2 - \frac{\mathrm{Id}}{d} \right\|_2^2 = \mathrm{tr}|A|^4 - \frac{2\,\mathrm{tr}|A|^2}{d} + \frac{\mathrm{trId}}{d^2} = \mathrm{tr}|A|^4 - \frac{1}{d} \geq 0$$

Consequently,

$$d^{-1/4}\|A\|_2 \leq \|A\|_4 \leq d^{-1/4}\left(1+\frac{C^2}{d}\right)^{1/4}\|A\|_2 \leq d^{-1/4}\left(1+\frac{C^2}{4d}\right)\|A\|_2$$

In other words, $\mathcal{W}$ is $1 + O(\frac{1}{d})$-Euclidean as a subspace of the Schatten 4-class.

Note : Applying *directly* Dvoretzky's theorem for the parameters in question ($k \sim d^2, m \sim d^2$) gives only $1 + O\left(\frac{1}{\sqrt[4]{d}}\right)$

A bootstrap : Dvoretzky $\times$ 2

The trick : $g(A) := \left\| AA^\dagger - \frac{\mathrm{Id}}{d} \right\|_2$ is only 2-Lipschitz on the Frobenius sphere S_F, but much more regular on a large subset

A bootstrap : Dvoretzky $\times$ 2

The trick : $g(A) := \|AA^\dagger - \frac{\mathrm{Id}}{d}\|_2$ is only 2-Lipschitz on the Frobenius sphere S_F, but much more regular on a large subset

Fact 4 : g is $6/\sqrt{d}$-Lipschitz when restricted to the set

$$\Omega = \{A \in S_F \ : \ \|A\|_\infty \leq 3/\sqrt{d}\}$$

verifying $\mathbb{P}(\Omega^c) \leq \exp(-ck)$. Moreover, for a typical m-dimensional subspace $E \subset \mathcal{M}_{d\times k}$, $E \cap S_F \subset \Omega$.

A bootstrap : Dvoretzky $\times$ 2

The trick : $g(A) := \left\|AA^\dagger - \frac{\mathrm{Id}}{d}\right\|_2$ is only 2-Lipschitz on the Frobenius sphere S_F, but much more regular on a large subset

Fact 4 : (1) g is $6/\sqrt{d}$-Lipschitz when restricted to the set

$$\Omega = \{A \in S_F \; : \; \|A\|_\infty \leq 3/\sqrt{d}\}$$

verifying (2)$\mathbb{P}(\Omega^c) \leq \exp(-ck)$. (3) Moreover, for a typical m-dimensional subspace $E \subset \mathcal{M}_{d\times k}$, $E \cap S_F \subset \Omega$.

(1)

A bootstrap : Dvoretzky $\times$ 2

The trick : $g(A) := \|AA^\dagger - \frac{\mathrm{Id}}{d}\|_2$ is only 2-Lipschitz on the Frobenius sphere S_F, but much more regular on a large subset

Fact 4 : (1) g is $6/\sqrt{d}$-Lipschitz when restricted to the set

$$\Omega = \{A \in S_F \,:\, \|A\|_\infty \leq 3/\sqrt{d}\}$$

verifying (2)$\mathbb{P}(\Omega^c) \leq \exp(-ck)$. (3) Moreover, for a typical m-dimensional subspace $E \subset \mathcal{M}_{d\times k}$, $E \cap S_F \subset \Omega$.

$$\begin{aligned}(1)\ g(A) - g(B) &= \|AA^\dagger - \tfrac{\mathrm{Id}}{k}\|_2 - \|BB^\dagger - \tfrac{\mathrm{Id}}{k}\|_2 \\ &\leq \|AA^\dagger - BB^\dagger\|_2 \\ &\leq \|A(A^\dagger - B^\dagger) + (A-B)B^\dagger\|_2\end{aligned}$$

A bootstrap : Dvoretzky $\times$ 2

The trick : $g(A) := \left\|AA^\dagger - \frac{\mathrm{Id}}{d}\right\|_2$ is only 2-Lipschitz on the Frobenius sphere S_F, but much more regular on a large subset

Fact 4 : (1) g is $6/\sqrt{d}$-Lipschitz when restricted to the set

$$\Omega = \{A \in S_F \ : \ \|A\|_\infty \leq 3/\sqrt{d}\}$$

verifying (2) $\mathbb{P}(\Omega^c) \leq \exp(-ck)$. (3) Moreover, for a typical m-dimensional subspace $E \subset \mathcal{M}_{d\times k}$, $E \cap S_F \subset \Omega$.

(3)

A bootstrap : Dvoretzky $\times$ 2

The trick : $g(A) := \|AA^\dagger - \frac{\mathrm{Id}}{d}\|_2$ is only 2-Lipschitz on the Frobenius sphere S_F, but much more regular on a large subset

Fact 4 : (1) g is $6/\sqrt{d}$-Lipschitz when restricted to the set

$$\Omega = \{A \in S_F \; : \; \|A\|_\infty \leq 3/\sqrt{d}\}$$

verifying (2) $\mathbb{P}(\Omega^c) \leq \exp(-ck)$. (3) Moreover, for a typical m-dimensional subspace $E \subset \mathcal{M}_{d\times k}$, $E \cap S_F \subset \Omega$.

(3) follows from Dvoretzky's theorem with $\|\cdot\| = \|\cdot\|_\infty$ and $\varepsilon = 1/2$ once we establish that the average M of $\|\cdot\|_\infty$ verifies $M \sim 1/\sqrt{d} + 1/\sqrt{k} \leq 2/\sqrt{d}$

A bootstrap : Dvoretzky $\times$ 2

The trick : $g(A) := \|AA^\dagger - \frac{\mathrm{Id}}{d}\|_2$ is only 2-Lipschitz on the Frobenius sphere S_F, but much more regular on a large subset

Fact 4 : (1) g is $6/\sqrt{d}$-Lipschitz when restricted to the set

$$\Omega = \{A \in S_F \,:\, \|A\|_\infty \leq 3/\sqrt{d}\}$$

verifying (2) $\mathbb{P}(\Omega^c) \leq \exp(-ck)$. (3) Moreover, for a typical m-dimensional subspace $E \subset \mathcal{M}_{d\times k}$, $E \cap S_F \subset \Omega$.

(3) follows from Dvoretzky's theorem with $\|\cdot\| = \|\cdot\|_\infty$ and $\varepsilon = 1/2$ once we establish that the average M of $\|\cdot\|_\infty$ verifies $M \sim 1/\sqrt{d} + 1/\sqrt{k} \leq 2/\sqrt{d}$ (Marchenko-Pastur)

Note: $m \sim \varepsilon^2 (M/b)^2 N = (1/2)^2 M^2 kd \sim k \sim d^2$

A bootstrap : Dvoretzky $\times$ 2

The trick : $g(A) := \left\| AA^\dagger - \frac{\mathrm{Id}}{d} \right\|_2$ is only 2-Lipschitz on the Frobenius sphere S_F, but much more regular on a large subset

Fact 4 : (1) g is $6/\sqrt{d}$-Lipschitz when restricted to the set

$$\Omega = \{A \in S_F \, : \, \|A\|_\infty \leq 3/\sqrt{d}\}$$

verifying (2) $\mathbb{P}(\Omega^c) \leq \exp(-ck)$. (3) Moreover, for a typical m-dimensional subspace $E \subset \mathcal{M}_{d \times k}$, $E \cap S_F \subset \Omega$.

(3) follows from Dvoretzky's theorem with $\| \cdot \| = \| \cdot \|_\infty$ and $\varepsilon = 1/2$ once we establish that the average M of $\| \cdot \|_\infty$ verifies $M \sim 1/\sqrt{d} + 1/\sqrt{k} \leq 2/\sqrt{d}$ (Marchenko-Pastur)

Note: $m \sim \varepsilon^2 (M/b)^2 N = (1/2)^2 M^2 kd \sim k \sim d^2$

(2) follows from Levy's lemma

Levy's lemma

Fact 5 : If $f : S^{N-1} \to \mathbb{R}$ is a 1-Lipschitz function, then for every $\varepsilon > 0$,

$$\mathbb{P}(|f(x) - \mu| > \varepsilon) \leq C_1 \exp(-c_1 N \varepsilon^2),$$

where x is uniformly distributed on S^{N-1}, μ is any central value of f, and $C_1, c_1 > 0$ are absolute constants.

Levy's lemma

Fact 5 : If $f : S^{N-1} \to \mathbb{R}$ is a 1-Lipschitz function, then for every $\varepsilon > 0$,

$$\mathbb{P}(|f(x) - \mu| > \varepsilon) \leq C_1 \exp(-c_1 N \varepsilon^2),$$

where x is uniformly distributed on S^{N-1}, μ is any central value of f, and $C_1, c_1 > 0$ are absolute constants.

Central value : the median, the mean, or any number between (say) the 1st and the 3rd quartile.

Levy's lemma

Fact 5 : If $f : S^{N-1} \to \mathbb{R}$ is a 1-Lipschitz function, then for every $\varepsilon > 0$

$$\mathbb{P}(|f(x) - \mu| > \varepsilon) \leq C_1 \exp(-c_1 N \varepsilon^2),$$

where x is uniformly distributed on S^{N-1}, μ is any central value of f, and $C_1, c_1 > 0$ are absolute constants.

Central value : the median, the mean, or any number between (say) the 1st and the 3rd quartile.

Here : $\mu = M \leq 2/\sqrt{d}$, $\varepsilon = 1/\sqrt{d}$, $N = 2kd$

$\Rightarrow \mu + \varepsilon \leq 3/\sqrt{d}$ and $c_1 N \varepsilon^2 = 4 c_1 k$

The final step

Restrict $g(A) = \|AA^\dagger - \frac{\mathrm{Id}}{d}\|_2$ to Ω, and then extend the restriction $g_{|\Omega}$ to $\tilde{g} : S_F \to \mathbb{R}$ without increasing the Lipschitz constant, which is $\leq 6/\sqrt{d}$

The final step

Restrict $g(A) = \|AA^\dagger - \frac{\mathrm{Id}}{d}\|_2$ to Ω, and then extend the restriction $g_{|\Omega}$ to $\tilde{g} : S_F \to \mathbb{R}$ without increasing the Lipschitz constant, which is $\leq 6/\sqrt{d}$

Next, apply the "non-linear" Dvoretzky's theorem to $\tilde{g}$.

The final step

Restrict $g(A) = \|AA^\dagger - \frac{\mathrm{Id}}{d}\|_2$ to Ω, and then extend the restriction $g_{|\Omega}$ to $\tilde{g} : S_F \to \mathbb{R}$ without increasing the Lipschitz constant, which is $\leq 6/\sqrt{d}$

Next, apply the "non-linear" Dvoretzky's theorem to $\tilde{g}$.

Since for a typical m-dimensional subspace E we have $E \cap S_F \subset \Omega$, it follows that $\tilde{g}$ and g coincide on $E \cap S_F$, and the conclusion obtained from the theorem is valid also for g, as required.

The final step

Restrict $g(A) = \left\| AA^\dagger - \frac{\mathrm{Id}}{d} \right\|_2$ to Ω, and then extend the restriction $g_{|\Omega}$ to $\tilde{g} : S_F \to \mathbb{R}$ without increasing the Lipschitz constant, which is $\leq 6/\sqrt{d}$

Next, apply the "non-linear" Dvoretzky's theorem to $\tilde{g}$.

Since for a typical m-dimensional subspace E we have $E \cap S_F \subset \Omega$, it follows that $\tilde{g}$ and g coincide on $E \cap S_F$, and the conclusion obtained from the theorem is valid also for g, as required.

Needed to verify : Since we want to upper-bound $\tilde{g}$ on $E \cap S_F$ by $O(\frac{1}{d})$, we need to check that

The final step

Restrict $g(A) = \left\| AA^\dagger - \frac{\mathrm{Id}}{d} \right\|_2$ to Ω, and then extend the restriction $g_{|\Omega}$ to $\tilde{g} : S_F \to \mathbb{R}$ without increasing the Lipschitz constant, which is $\leq 6/\sqrt{d}$

Next, apply the "non-linear" Dvoretzky's theorem to $\tilde{g}$.

Since for a typical m-dimensional subspace E we have $E \cap S_F \subset \Omega$, it follows that $\tilde{g}$ and g coincide on $E \cap S_F$, and the conclusion obtained from the theorem is valid also for g, as required.

Needed to verify : Since we want to upper-bound $\tilde{g}$ on $E \cap S_F$ by $O(\frac{1}{d})$, we need to check that

- the median of g on S_F is $O(\frac{1}{d})$
- allowing $O(\frac{1}{d})$ deviation from a central value of $\tilde{g}$ leads to a correct value of m, i.e., $m \sim k \sim d^2$

The "non-linear" Dvoretzky's theorem

Fact 1' : Let $f : S_{\mathbb{C}^N} \to \mathbb{R}$ is a 1-Lipschitz circled function and let $\varepsilon > 0$. Let $E \subset \mathbb{C}^n$ be a random subspace (Haar-distributed) of dimension $m = c_0 N \varepsilon^2$. Then, with large probability,

$$|f - \mu| \leq \varepsilon \;\; \text{on} \;\; S_{\mathbb{C}^N} \cap E,$$

where μ is any central value of f (with respect to the normalized Lebesgue measure on $S_{\mathbb{C}^N}$) and c_0 is an absolute constant.

The "non-linear" Dvoretzky's theorem

Fact 1' : Let $f : S_{\mathbb{C}^N} \to \mathbb{R}$ is a 1-Lipschitz circled function and let $\varepsilon > 0$. Let $E \subset \mathbb{C}^n$ be a random subspace (Haar-distributed) of dimension $m = c_0 N \varepsilon^2$. Then, with large probability,

$$|f - \mu| \leq \varepsilon \ \text{ on } \ S_{\mathbb{C}^N} \cap E,$$

where μ is any central value of f (with respect to the normalized Lebesgue measure on $S_{\mathbb{C}^N}$) and c_0 is an absolute constant. If the function is L-Lipschitz, the dimension changes to $m = c_0 N (\varepsilon / L)^2$.

Final verifications

- $L = 6/\sqrt{d}$, $\varepsilon = O(1/d)$ $\Rightarrow m \sim N\left(\frac{\varepsilon}{L}\right)^2 \sim kd\left(\frac{1/d}{6/\sqrt{d}}\right)^2 \sim k$

Final verifications

- $L = 6/\sqrt{d}$, $\varepsilon = O(1/d)$ $\Rightarrow m \sim N\left(\frac{\varepsilon}{L}\right)^2 \sim kd\left(\frac{1/d}{6/\sqrt{d}}\right)^2 \sim k$ OK

- the median of $\tilde{g} \stackrel{?}{=} O(1/d)$

Marchenko-Pastur: with large probability, all singular values of A are in the interval $\left[\frac{1}{\sqrt{d}} - \frac{1}{\sqrt{k}}, \frac{1}{\sqrt{d}} + \frac{1}{\sqrt{k}}\right]$
$\Rightarrow$ all eigenvalues of $AA^\dagger$ are within $O\left(\frac{1}{\sqrt{kd}}\right)$ of $\frac{1}{d}$

The union bound vs. chaining

The standard proofs of Dvoretzky-like statements go as follows:

The union bound vs. chaining

The standard proofs of Dvoretzky-like statements go as follows:

- for a fixed subspace E_0, choose an $\varepsilon/2$-net $\mathcal{N}$ of $S^{N-1} \cap E_0$

The union bound vs. chaining

The standard proofs of Dvoretzky-like statements go as follows:

- for a fixed subspace E_0, choose an $\varepsilon/2$-net $\mathcal{N}$ of $S^{N-1} \cap E_0$
- if U is a random unitary, Levy's lemma implies that for fixed $z \in \mathcal{N}$ with large probability $|f(Uz) - \mu| \leq \varepsilon/2$

The union bound vs. chaining

The standard proofs of Dvoretzky-like statements go as follows:

- for a fixed subspace E_0, choose an $\varepsilon/2$-net $\mathcal{N}$ of $S^{N-1} \cap E_0$
- if U is a random unitary, Levy's lemma implies that for fixed $z \in \mathcal{N}$ with large probability $|f(Uz) - \mu| \leq \varepsilon/2$
- the union bound $\Rightarrow$ the same property for all $z \in \mathcal{N}$
- f is 1-Lipshitz $\Rightarrow |f(Uz) - \mu| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$ on $S^{N-1} \cap UE_0$ for "most of" choices of U

The union bound vs. chaining

The standard proofs of Dvoretzky-like statements go as follows:

• for a fixed subspace E_0, choose an $\varepsilon/2$-net $\mathcal{N}$ of $S^{N-1} \cap E_0$

• if U is a random unitary, Levy's lemma implies that for fixed $z \in \mathcal{N}$ with large probability $|f(Uz) - \mu| \leq \varepsilon/2$

• the union bound $\Rightarrow$ the same property for all $z \in \mathcal{N}$

• f is 1-Lipshitz $\Rightarrow |f(Uz) - \mu| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$ on $S^{N-1} \cap UE_0$ for "most of" choices of U

Difficulty : necessarily $\mathrm{card}\mathcal{N} \geq \left(\frac{1}{\varepsilon}\right)^{\dim E_0}$, so for the procedure to work one needs $m := \dim E_0 = O\left(\frac{\varepsilon^2}{\log(2/\varepsilon)} N\right)$

The union bound vs. chaining

The standard proofs of Dvoretzky-like statements go as follows:

- for a fixed subspace E_0, choose an $\varepsilon/2$-net $\mathcal{N}$ of $S^{N-1} \cap E_0$
- if U is a random unitary, Levy's lemma implies that for fixed $z \in \mathcal{N}$ with large probability $|f(Uz) - \mu| \leq \varepsilon/2$
- the union bound $\Rightarrow$ the same property for all $z \in \mathcal{N}$
- f is 1-Lipshitz $\Rightarrow |f(Uz) - \mu| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$ on $S^{N-1} \cap UE_0$ for "most of" choices of U

Difficulty : necessarily $\mathrm{card}\mathcal{N} \geq \left(\frac{1}{\varepsilon}\right)^{\dim E_0}$, so for the procedure to work one needs $m := \dim E_0 = O\left(\frac{\varepsilon^2}{\log(2/\varepsilon)} N\right)$

Remedy : a chaining argument, or a more economical usage of nets

The chaining argument

- S compact metric space of radius R, $h : S \to \mathbb{R}$ continuous

The chaining argument

- S compact metric space of radius R, $h : S \to \mathbb{R}$ continuous
- for $k = 0, 1, 2, \ldots$, $\mathcal{N}_k$ is an $2^{-k}R$-net of S; $\mathcal{N}_0 = \{s_0\}$
- given $s \in S$ and $k = 0, 1, 2, \ldots,$, let $z_k(s) \in \mathcal{N}_k$ verify
$\mathrm{dist}\big(s, z_k(s)\big) \leq 2^{-k}R$

The chaining argument

- S compact metric space of radius R, $h : S \to \mathbb{R}$ continuous

- for $k = 0, 1, 2, \ldots$, $\mathcal{N}_k$ is an $2^{-k}R$-net of S; $\mathcal{N}_0 = \{s_0\}$

- given $s \in S$ and $k = 0, 1, 2, \ldots,$, let $z_k(s) \in \mathcal{N}_k$ verify $\mathrm{dist}\big(s, z_k(s)\big) \leq 2^{-k}R$

- then $h(s) = h(s_0) + \sum_{k=0}^{\infty} h\big(z_{k+1}(s)\big) - h\big(z_k(s)\big)$, hence $|h(s) - h(s_0)| \leq \sum_{k=0}^{\infty} \max_{t \in \mathcal{N}_k, t' \in \mathcal{N}_{k+1}} |h(t) - h(t')|$

The chaining argument

- S compact metric space of radius R, $h : S \to \mathbb{R}$ continuous

- for $k = 0, 1, 2, \ldots$, $\mathcal{N}_k$ is an $2^{-k}R$-net of S; $\mathcal{N}_0 = \{s_0\}$

- given $s \in S$ and $k = 0, 1, 2, \ldots,$, let $z_k(s) \in \mathcal{N}_k$ verify
$\mathrm{dist}\big(s, z_k(s)\big) \leq 2^{-k}R$

- then $h(s) = h(s_0) + \sum_{k=0}^{\infty} h\big(z_{k+1}(s)\big) - h\big(z_k(s)\big)$, hence
$|h(s) - h(s_0)| \leq \sum_{k=0}^{\infty} \max_{t \in \mathcal{N}_k, t' \in \mathcal{N}_{k+1}} |h(t) - h(t')|$

- if $\big(X_s\big)_{s \in S}$ are random variables, then
$\mathbb{E}|X_s - X_{s_0})| \leq \sum_{k=0}^{\infty} \mathbb{E} \max_{t \in \mathcal{N}_k, t' \in \mathcal{N}_{k+1}} |X_t - X_{t'}|$

The chaining argument

- S compact metric space of radius R, $h : S \to \mathbb{R}$ continuous
- for $k = 0, 1, 2, \ldots$, $\mathcal{N}_k$ is an $2^{-k}R$-net of S; $\mathcal{N}_0 = \{s_0\}$
- given $s \in S$ and $k = 0, 1, 2, \ldots,$, let $z_k(s) \in \mathcal{N}_k$ verify $\mathrm{dist}(s, z_k(s)) \leq 2^{-k}R$
- then $h(s) = h(s_0) + \sum_{k=0}^{\infty} h(z_{k+1}(s)) - h(z_k(s))$, hence $|h(s) - h(s_0)| \leq \sum_{k=0}^{\infty} \max_{t \in \mathcal{N}_k, t' \in \mathcal{N}_{k+1}} |h(t) - h(t')|$
- if $(X_s)_{s \in S}$ are random variables, then $\mathbb{E}|X_s - X_{s_0})| \leq \sum_{k=0}^{\infty} \mathbb{E} \max_{t \in \mathcal{N}_k, t' \in \mathcal{N}_{k+1}} |X_t - X_{t'}|$

Under appropriate assumptions on the continuity of the process $(X_s)_{s \in S}$, this argument leads to surprisingly sharp bounds

Dudley's inequality

Fact 6 : S compact metric space, $(X_s)_{s\in S}$ a subgaussian process, i.e., there are $\alpha, \beta > 0$ such that, for all $t, t' \in S$ and for all $\lambda \geq 0$,

$$\mathbb{P}(|X_t - X_{t'}| \geq \lambda) \leq \beta \exp\left(-\alpha \frac{\lambda^2}{\mathrm{dist}(t,t')^2}\right)$$

Then

$$\mathbb{E} \sup_{t,t'\in S} |X_t - X_{t'}| \leq C\beta\alpha^{-1/2} \int_0^R \sqrt{\log N(S,\eta)}\, d\eta,$$

where $N(S,\eta)$ is the minimal cardinality of a η-net of S and R is the radius of S.

In our context : $f : S_{\mathbb{C}^N} \to \mathbb{R}$, 1-Lipschitz

- $S := \left(S_{\mathbb{C}^N} \cap E_0\right) \cup \{0\}$, $\dim E_0 =: m$

In our context : $f : S_{\mathbb{C}^N} \to \mathbb{R}$, 1-Lipschitz

- $S := (S_{\mathbb{C}^N} \cap E_0) \cup \{0\}$, $\dim E_0 =: m$
- randomness : $U \in U(N)$, the Haar measure
- $X_s := f(Us)$
- subgaussian property, with $\alpha = cN$ and $\beta = O(1)$, is given by Levy's lemma + some standard tricks

In our context : $f : S_{\mathbb{C}^N} \to \mathbb{R}$, 1-Lipschitz

- $S := (S_{\mathbb{C}^N} \cap E_0) \cup \{0\}$, $\dim E_0 =: m$
- randomness : $U \in U(N)$, the Haar measure
- $X_s := f(Us)$
- subgaussian property, with $\alpha = cN$ and $\beta = O(1)$, is given by Levy's lemma + some standard tricks
- elementary bound $N(S, \eta) \leq \left(\frac{3}{\eta}\right)^m$
- singularity at 0 integrates out; in other words, no $\log(2/\varepsilon)$ effect

Is negative solution to the additivity conjecture good or bad?

Is negative solution to the additivity conjecture good or bad?

An affirmative answer would greatly simplify the theory: BAD

On the other hand, a negative answer means that entanglement allows using quantum channels more efficiently than previously thought: GOOD

But to exploit this opportunity one would need *explicit* maps for *reasonable* values of the parameters m, d

Is negative solution to the additivity conjecture good or bad?

An affirmative answer would greatly simplify the theory: BAD

On the other hand, a negative answer means that entanglement allows using quantum channels more efficiently than previously thought: GOOD

But to exploit this opportunity one would need *explicit* maps for *reasonable* values of the parameters m, d

Is negative solution to the additivity conjecture good or bad?

An affirmative answer would greatly simplify the theory: BAD

On the other hand, a negative answer means that entanglement allows using quantum channels more efficiently than previously thought: GOOD

But to exploit this opportunity one would need *explicit* maps for *reasonable* values of the parameters m, d