The interface between quantum information theory and functional analysis. Additivity conjectures and Dvoretzky's theorem.

Stanislaw Szarek

Paris 6/Case Western Reserve

BIRS, June 28, 2010

Collaborators:
G. Aubrun, E. Werner

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

http://www.math.jussieu.fr/∼szarek/

Talk summary

• overview of certain aspects of quantum information theory: paradigms, concepts, notation

• additivity/multiplicativity problems

• an approach to those problems via tools of geometric functional analysis, notably Dvoretzky's theorem

Conference “Perspectives in High Dimensions”

Cleveland, August 2 until August 6, 2010

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

Quantum information theory

(from the geometric functional analysis angle)

• A complex Hilbert space $\mathcal{H}$, usually $\mathcal{H} = \mathbb{C}^d$, and 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 positive semi-definite cone $\mathcal{PSD} \subset \mathcal{M}_d^{sa}$

• The base of $\mathcal{PSD}$ consisting of density matrices: $\mathcal{D}(\mathcal{H}) := \mathcal{PSD} \cap \{\mathrm{tr}(\cdot) = 1\}$ $\sim$ the states of $\mathcal{B}(\mathcal{H})$ = the positive face of the unit ball in the trace class (1-Schatten) norm

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

More context and more notation

Unit vector $\psi \in \mathcal{H} = \mathbb{C}^d$ (or $|\psi\rangle$) : "state" of a quantum system with d levels

$d = 2$ $\rightarrow$ qubits

$\rho = \psi\psi^\dagger = |\psi\rangle\langle\psi|$: the corresponding rank one projection, or

• a pure state of $\mathcal{B}(\mathcal{H})$, an element of $\mathcal{B}(\mathcal{H})^*$ via duality $(A, \rho) := \mathrm{tr}(A\rho^\dagger)$ or

• an element of the projective space $\mathbb{CP}^{d-1}$

Mixed states: $\rho = \sum_\alpha p_\alpha |\psi_\alpha\rangle\langle\psi_\alpha|$ with $\sum_\alpha p_\alpha = 1$

The set of mixed states coincides with $\mathcal{D}(\mathcal{H}) = \mathcal{PSD} \cap \{\mathrm{tr}(\cdot) = 1\}$

Measurements

$|\langle\psi|e_j\rangle|^2 = \langle e_j|\psi\rangle\langle\psi|e_j\rangle = \langle e_j|\rho|e_j\rangle = \mathrm{tr}\Big(\rho|e_j\rangle\langle e_j|\Big)$:

the probability of jth outcome under measurement "in the basis (e_j)" for $\rho = |\psi\rangle\langle\psi|$,or general ρ

More general measurements schemes (POVM):

Given $P_i \in \mathcal{PSD}$ with $\sum_i P_i = \mathrm{Id}$, the probability of the ith outcome is $\mathrm{tr}(\rho P_i)$

In general, P_i's do not need to be projections

Bi- or multipartite systems, entanglement

m systems (or particles) : $\mathcal{K} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \ldots \otimes \mathcal{H}_m$

Example: our apparatus and environment $\mathcal{K} = \mathcal{H} \otimes \mathcal{E}$

Pure separable state (product vector): $\psi = \xi \otimes \eta$

General separable states:

$$\mathcal{S} = \{\textstyle\sum_\alpha p_\alpha \, |\psi_\alpha\rangle\langle\psi_\alpha| \;\; : \;\; \psi_\alpha \text{ product vectors}\}$$

Entangled states: $\mathcal{D} \setminus \mathcal{S}$

$\mathrm{conv}\big(-\mathcal{D} \cup \mathcal{D}\big)$ = the unit ball of trace class

$\mathrm{conv}\Big(-\mathcal{S} \cup \mathcal{S}\Big)$ = the unit ball of the projective tensor product of trace class spaces on respective subsystems

Partial transpose, Peres-Horodecki criterion

Bipartite system: $\mathcal{K} = \mathcal{H}_1 \otimes \mathcal{H}_2$

Partial transpose $\mathcal{B}(\mathcal{K}) \xrightarrow{T_2} \mathcal{B}(\mathcal{K})$: $T_2(\rho_1 \otimes \rho_2) = \rho_1 \otimes \rho_2^t$ etc.

Easy: ρ separable $\Rightarrow$ $T_2(\rho)$ separable $\Rightarrow$ $T_2(\rho) \in \mathcal{PSD}$

Criterion: $T_2(\rho) \notin \mathcal{PSD}$ $\Rightarrow$ ρ entangled

"$\Leftrightarrow$" only for 2×2 and 2×3 systems

(Størmer-Woronowicz)

PPT states: $\mathcal{PPT} := \mathcal{D} \cap T_2^{-1}(\mathcal{D})$

Entangled PPT states: example of undistillable entanglement (not defined)

Quantum vs. classical correlations, Tsirelson bound

$X_1, X_2, \ldots,\ \ Y_1, Y_2, \ldots$ random variables; $\|X_j\|_\infty, \|Y_k\|_\infty \le 1$

Covariance matrix: $\big(\mathbb{E} X_j Y_k\big)_{j,k}$

Possible covariance matrices: $\mathcal{C} := \mathrm{conv}\big\{\big(\delta_j \eta_k\big)_{j,k}\ :\ \delta_j, \eta_k = \pm 1\big\}$

$\mathcal{C}$ - a polytope; faces $\sim$ Bell inequalities

Quantum covariance matrices:

$\mathcal{Q} := \Big\{\Big(\mathrm{tr}\big(\rho(U_j \otimes V_k)\big)\Big)_{j,k} : \rho \in \mathcal{D}(\mathcal{H}_1 \otimes \mathcal{H}_2), \|U_j\|_\infty, \|V_k\|_\infty \le 1\Big\}$

Tsirelson: $\mathcal{Q} = \mathrm{conv}\big\{\big(\langle u_j | v_k\rangle\big)_{j,k}\ :\ u_j, v_k \in \mathcal{H}, |u_j|, |v_k| \le 1\big\}$

In particular, $\mathcal{C} \subsetneq \mathcal{Q} \subset K_G^{\mathbb{R}} \mathcal{C}$

Quantum operations, channels

Evolution of a (closed) system in discrete time :
$\psi = |\psi\rangle$ input, $U\psi = U|\psi\rangle$ output, U unitary (or an isometry)

In the language of states : $|\psi\rangle\langle\psi| \to U|\psi\rangle\langle\psi|U^\dagger$

Quantum operation (channel) $\rho \to \Phi(\rho) = U\rho U^\dagger$
(valid also for mixed states)

These are examples of "elementary" completely positive maps. For open systems, quantum formalism allows also other CP maps as quantum operations. However, by Stinespring-Kraus-Choi theorem all such maps can be "reduced" to elementary ones

$$\rho \to \Phi(\rho) = \sum_j B_j \rho B_j^\dagger$$

Quantum operations via partial trace

$\mathcal{K} = \mathcal{H} \otimes \mathcal{E}$ (e.g., our apparatus and environment)

Accessible part of a product state $\xi \otimes \eta$ is just ξ

Accessible part of φ is $\mathrm{tr}_{\mathcal{E}}(|\varphi\rangle\langle\varphi|)$, where $\mathrm{tr}_{\mathcal{E}}$ is the partial trace induced by $\mathrm{tr}_{\mathcal{E}}(\sigma \otimes \tau) = \mathrm{tr}(\tau)\sigma$,and similarly for general states

Let $V : \mathcal{H} \to \mathcal{K} = \mathcal{H} \otimes \mathcal{E}$ an isometry, $|\psi\rangle \to V|\psi\rangle$

Consider the following quantum operation :

$\Phi(|\psi\rangle\langle\psi|) = \mathrm{tr}_{\mathcal{E}}(V|\psi\rangle\langle\psi|V^{\dagger}) = \mathrm{tr}_2(V|\psi\rangle\langle\psi|V^{\dagger})$ and, generally,

$\Phi(\rho) = \mathrm{tr}_{\mathcal{E}}(V\rho V^{\dagger}) = \mathrm{tr}_2(V\rho V^{\dagger})$

Equivalent to Stinespring-Kraus-Choi representation
$\Phi(\rho) = \sum_i B_i \rho B_i^{\dagger}$:$V = \sum_i B_i \otimes e_i$, so this is the general case

Channels as subspaces

Quantum operations on $\mathcal{H} = \mathbb{C}^d$ are really d-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 $\mathcal{W}$

Examples:

- $k = 1$ or, more generally, $V(\xi) = \xi \otimes \eta$ (fixed η) $\Rightarrow$ $\Phi(|\xi\rangle\langle\xi|) = \mathrm{tr}_2(|\xi \otimes \eta\rangle\langle\xi \otimes \eta|) = |\xi\rangle\langle\xi|\, \mathrm{tr}(|\eta\rangle\langle\eta|) = |\xi\rangle\langle\xi|$, or $\Phi = I_{\mathcal{M}_d}$
- $V(\xi) = \eta \otimes \xi \;\Rightarrow \forall\rho\; \Phi(\rho) = |\eta\rangle\langle\eta|$
- $V = k^{-1/2} \sum_{i=1}^k U_i \otimes e_i$, U_i's i.i.d. random unitaries

If instead of U_i's we had i.i.d. Gaussian matrices, the range of V would be a Haar-random subspace of $\mathbb{C}^d \otimes \mathbb{C}^k$
$\Phi(\rho) = k^{-1} \sum_i U_i \rho U_i^\dagger$

Range of a channel and the Schmidt decomposition

$\mathcal{W}$ associated to Φ

For a pure state $\varphi = V\psi \in \mathcal{W}$, the accessible part $\mathrm{tr}_2(|\varphi\rangle\langle\varphi|)$ of φ, or $\Phi(|\psi\rangle\langle\psi|)$, is simply encoded in its "Schmidt decomposition"

$$\varphi = \sum_j s_j\, u_j \otimes v_j$$

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

This is more or less SVD of the matrix

$$A = \sum_j s_j\, |u_j\rangle\langle v_j|$$

that can be identified with φ

The image of a pure state $|\psi\rangle\langle\psi|$ under Φ

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

Verification:

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

Morale: important 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

For future reference

If $A = \sum_j s_j \, |u_j\rangle\langle v_j|$ is the matrix identified with φ, then

$$\mathrm{tr}_2(|\varphi\rangle\langle\varphi|) = \sum_j s_j^2 \, |u_j\rangle\langle u_j| = AA^\dagger$$

Quantum channels, capacities and such

"One-shot" capacity of Φ (for transmitting classical information)

$$\chi(\Phi) := \max_{p_\alpha, \rho_\alpha} S\Big(\Phi\Big(\sum_\alpha p_\alpha \rho_\alpha\Big)\Big) - \sum_\alpha p_\alpha S(\Phi(\rho_\alpha))$$

where $S(\rho) = -\mathrm{tr}(\rho \log \rho)$ is the von Neumann entropy
$(= \sum_j q_j \log(1/q_j)$, if q_j's are eigenvalues of $\rho)$

The "true" capacity is

$$\chi^\infty(\Phi) := \lim_{n\to\infty} \frac{1}{n} \chi\big(\Phi \otimes \Phi \otimes \ldots \otimes \Phi\big) \qquad (n \text{ fold product})$$

Additivity problems

Is $\chi^{\infty}(\cdot)$ additive? I.e., is $\chi^{\infty}(\Phi \otimes \Psi) = \chi^{\infty}(\Phi) + \chi^{\infty}(\Psi)$?

This would follow if $\chi(\cdot)$ was additive or even (Shor 2004 and others) if the following much simpler quantity was additive

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

$S_{\min}$ is called the "minimum output entropy"

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

where $\|\tau\|_p = \left(\mathrm{tr}\left(\tau^\dagger \tau\right)^{p/2}\right)^{1/p}$ is the Schatten p-norm.

(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 $\|\Phi\|_{1\to p}$.

Additivity of $S_p^{\min}(\Phi)$ is equivalent to multiplicativity of $\|\Phi\|_{1\to p}$.

Additivity/multiplicativity problems - recapitulation

For completely positive (trace preserving) maps

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

$$\|\Phi \otimes \Psi\|_{1\to p} \stackrel{?}{=} \|\Phi\|_{1\to p}\|\Psi\|_{1\to p} \quad (p > 1)$$

The mins and the norms are attained on pure states, so all these quantities depend on the patterns of eigenvalues of $\Phi(|\psi\rangle\langle\psi|)$.

In view of prior remarks, this is equivalent to understanding the patterns of singular numbers of matrices varying over m-dimensional subspaces $\mathcal{W}$ of the space of $d \times k$ matrices.

" No" and " No" (Hayden-Winter 2008, Hastings 2009)

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

Let $\mathcal{W}$ be the m-dimensional subspace of $\mathbb{C}^d \otimes \mathbb{C}^k$ (or $\mathcal{M}_{d\times k}$) associated with Φ

$$\|\Phi\|_{1\to p} = \max_{\varphi\in\mathcal{W},\ |\varphi|=1} \|\mathrm{tr}_2(|\varphi\rangle\langle\varphi|)\|_p$$

If $\varphi = \sum_j s_j\, u_j \otimes v_j$,this becomes

$$\| \sum_j s_j^2\, |u_j\rangle\langle u_j| \|_p = \left(\sum_j s_j^{2p}\right)^{1/p} = \|A\|_{2p}^2 = \|AA^\dagger\|_p,$$

where $A = \sum_j s_j\, |u_j\rangle\langle v_j|$ is the $d \times k$ matrix identified with φ.

In other words

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

Milman's version of Dvoretzky's theorem

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 $m \leq c\varepsilon^2(M/b)^2 n$, 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 $q = 2p > 2$, $k = d$ and $\varepsilon = \frac{1}{2}$ we get $b = 1$ and $M \sim d^{1/q - 1/2}$, hence if

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

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.

So it is clear that we are up to something.

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)

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

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

The counterexample to multiplicativity

Need $\|\Phi \otimes \Psi\|_{1\to p} > \|\Phi\|_{1\to p}\|\Psi\|_{1\to p}$

$\Psi = \Phi$? $\Psi = \Phi'$ (independent copy)?

What works is $\Psi = \overline{\Phi}$!

Fact 1 : If $\Phi : \mathcal{B}(\mathbb{C}^m) \to \mathcal{B}(\mathbb{C}^d)$ is associated to an m-dimensional subspace of $\mathbb{C}^d \otimes \mathbb{C}^k$, then there is an input state $\sigma \in \mathcal{D}(\mathbb{C}^m \otimes \mathbb{C}^m)$ such that $(\Phi \otimes \overline{\Phi})(\sigma)$ has an eigenvalue $\geq \frac{m}{kd}$, hence $\|\Phi \otimes \overline{\Phi}\|_{1\to p} \geq \frac{m}{kd}$

In our setting $\frac{m}{kd} \sim \frac{d^{1+1/p}}{d^2} = d^{1/p-1}$, so

$$\|\Phi \otimes \overline{\Phi}\|_{1\to p} \geq cd^{1/p-1}$$

while

$$\|\Phi\|_{1\to p} \cdot \|\overline{\Phi}\|_{1\to p} = \left(\|\Phi\|_{1\to p}\right)^2 \leq \left(C^2 d^{1/p-1}\right)^2 \ll cd^{1/p-1}$$

The counterexample to additivity

of $S_{\min}(\cdot)$ is more subtle. The analysis of a single random channel is based on two facts

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

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

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

This reduces the study of the somewhat involved quantity $S_{\min}(\cdot)$ to upper-bounding $\left\|\sigma - \frac{\mathrm{Id}}{d}\right\|_{HS}$ 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 $\mathcal{W} \subset \mathcal{M}_{d \times k}$,

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

Recall: $AA^\dagger = \Phi(|\psi\rangle\langle\psi|)$, where ψ 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 eigenvalue" argument gives for the composite channel

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

Payback to geometric functional analysis

Fact 3 essentially says that $\mathcal{W}$, when endowed with the Schatten 4-norm, is $1 + O(\frac{1}{d^2})$-Euclidean.

On the other hand, applying *directly* Dvoretzky's theorem for that choice of parameters gives only $1 + O\left(\frac{1}{\sqrt{d}}\right)$

Is this 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