Noncommutative Λ_p-sets and the additivity problems

Stanislaw Szarek

Case Western Reserve U./Sorbonne U.

Institut Henri Poincaré, Paris, October 18, 2024

Talk summary

- additivity problems and instances of Dvoretzky's theorem (DT)
- Milman's tangible version of DT
- derandomizing/partial derandomizing of DT for classical L_p-spaces
- two feeble attempts for non-commutative L_p's (Schatten classes)
  - finite geometries
  - random Pauli matrices

Talk summary

- additivity problems, minimum output entropy
- connection to Dvoretzky's theorem (DT)
- Milman's tangible version of DT
- derandomizing/partial derandomizing of DT for classical L_p-spaces
- two feeble attempts for non-commutative L_p's (Schatten classes)
  - finite geometries (pseudo-random tensors ??)
  - random Pauli matrices (random tensors ??)

Additivity problems, minimal output entropy

Let $\mathcal{H}, \mathcal{K}$ be finite-dimensional complex Hilbert spaces and let $\Phi : B(\mathcal{H}) \to B(\mathcal{K})$ be a CPTP map (a quantum channel).

In 1970's through 1990's, various forms of capacity of CPTPs for transmitting information were defined and studied, one of them being the capacity $\chi(\Phi)$ for transmitting classical information (Holevo, Shumacher, Westmoreland etc.) Inevitably, a question was raised whether such capacity is additive, i.e., whether

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

This appeared a hopeless problem until P. Shor showed in early 2000's that an affirmative answer would follow from the additivity of a much more tractable quantity, the minimal output entropy, defined by

$$S^{\min}(\Phi) = \min_{\rho \in \mathcal{D}(\mathcal{H})} S(\Phi(\rho)).$$

where $\mathcal{D}(\mathcal{H})$ are the states on $\mathcal{H}$ and $S(\cdot)$ is the von Neumann entropy.

Connection to functional analysis, part 1

The connection to classical functional analysis (and ultimately to Dvoretzky's theorem) is based on several observations, all of which were well-known.

First, by the Stinespring-Kraus-Choi theorem, for any CPTP map $\Phi : B(\mathcal{H}) \to B(\mathcal{K})$ there exists another space $\mathcal{E}$ and an isometry $V : \mathcal{H} \to \mathcal{K} \otimes \mathcal{E}$ such that Φ can be represented as follows

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

where $\mathrm{tr}_{\mathcal{E}}$ is the partial trace. Next, by the concavity of entropy, the infimum in the definition of $S^{\min}(\Phi)$ is attained on pure states $\rho = |\psi\rangle\langle\psi|$, which means that

$$S^{\min}(\Phi) = \min_{|\psi\rangle \in \mathcal{H}} S\big(\mathrm{tr}_{\mathcal{E}}(V|\psi\rangle\langle\psi|V^{\dagger})\big) = \min_{|\varphi\rangle \in \mathcal{W}} S\big(\mathrm{tr}_{\mathcal{E}}(|\varphi\rangle\langle\varphi|)\big),$$

where $\mathcal{W} = V\mathcal{H}$ and ψ, φ are unit vectors. In other words, the calculation reduces to an analysis of a linear subspace $\mathcal{W} \subset \mathcal{K} \otimes \mathcal{E}$.

Connection to functional analysis, part 2

The next observation is that the entropy of $\tau := \mathrm{tr}_{\mathcal{E}}(|\varphi\rangle\langle\varphi|)$ depends only on its eigenvalues, which are squares of the Schmidt coefficients of $|\varphi\rangle$. We now cheat a little and pretend that $|\varphi\rangle \in \mathcal{K} \otimes \mathcal{E}$ is an operator from $\mathcal{K}$ to $\mathcal{E}$, or just a $d \times k$ matrix A for appropriate d, k, and we will restrict out attention to $k = d$. In other words, we just need to understand the patterns of singular values of operators in an m-dimensional subspace of the space of $d \times d$ matrices, where $m = \dim \mathcal{H}$.

Finally, we note that the von Neumann entropy $S(\sigma)$ is the limit, as $p \to 1$, of p-Rényi entropies

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

where $\|\tau\|_p = \left(\mathrm{tr}\left(\tau^*\tau\right)^{p/2}\right)^{1/p}$ is the Schatten p-norm. That is, we want to understand the geometry of the m-dimensional subspaces $\mathcal{W}$ of the q-Schatten space of $d \times d$ matrices, where $q = 2p$. More precisely, we want to find $R \geq 1$ such that, for $A \in \mathcal{W}$,

$$d^{1/q-1/2}\|A\|_2 \leq \|A\|_q \leq R\, d^{1/q-1/2}\|A\|_2.$$

Dvoretzky's theorem (1961)

There exist sequences $k_n \to +\infty$ and $\varepsilon_n \to 0^+$ such that, for every normed space X with $\dim X \geq n$ there exists subspace $E \subset X$ with $\dim E \geq k_n$, which is Euclidean, up to ε_n.

Equivalently

For every 0-symmetric convex body. $K \subset \mathbb{R}^n$ there is a central section $E \cap K$ of dimension at least k_n, which is – up to ε_n – a Euclidean ball.

Or, back to the language of norms (say, on $\mathbb{R}^n$ or $\mathbb{C}^n$)

$$r^{-1} \leq \frac{\|x\|}{|x|} \leq (1+\varepsilon_n) r^{-1} \quad \text{for all } \ x \in E$$

for some scaling constant r, where $|\cdot|$ is the Euclidean norm.

The symmetry hypothesis is not essential and specific (optimal or near optimal) formulae for (k_n) and (ε_n) are known.

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

Dvoretzky theorem as stated doesn't tell us the value of the scaling constant $r^{-1} = R$, but this issue is taken care of by the following "tangible" version due to Milman (1971):

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

In the specific case of q-Schatten spaces, calculating the parameter b is more or less trivial, and finding the average M is routine using *Gaussian random matrices* and *Chevet's inequality*. In particular, if $q = 2p > 2$, $k = d$, and $\varepsilon = \frac{1}{2}$ we get $m = \Omega(d^{1+2/q}) = \Omega(d^{1+1/p})$, so that for $A \in \mathcal{W}$, a generic m-dimensional subspace of the space of $d \times d$ matrices, we have

$$d^{1/q-1/2}\|A\|_2 \leq \|A\|_q \leq Cd^{1/q-1/2}\|A\|_2$$

Since the lower and the upper bounds are of the same order, this is "as good as it gets," and leads to an unexpectedly sharp upper bound for the minimal output Rényi entropy

$$S_p^{\min}(\Phi) := \min_{\rho \in \mathcal{D}(\mathcal{H})} S_p(\Phi(\rho)).$$

This sharp bound for the random channel Φ (and some other tricks) allowed P. Hayden and A. Winter (2008) to produce a counterexample to additivity of $S_p^{\min}(\cdot)$ for $p > 1$ by considering the product $\Phi \otimes \overline{\Phi}$.

The counterexample to additivity

The counterexample to additivity of $S^{\min}(\cdot)$ is more subtle, but is based on the same general ideas. It was found by M. Hastings (2009).

Subsequent to their work, G. Aubrun, E. Werner and I realized that their arguments involved in fact proving instances of Dvoretzky's theorem, allowing to simplify the proofs by using "off-the-shelf" technology, and making them conceptual rather than *ad hoc*.

Is the failure of additivity 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.

Explicit examples of additivity violations?

The channels Φ obtained via Dvoretzky's theorem are random, high-dimensional, and even certifying their validity may be difficult. To the best of my knowledge, explicit examples are known only for violations of $S_p^{\min}$ for $p > 2$.

Here is a (narrowly failed) attempt for an example for a violation $p = 2$.

The starting point are the so-called Λ_q-sets of Rudin for $q = 4$. Those are explicit subsets $S \subset \{0, 1, 2, \ldots, N-1\}$, $\#S = \Theta(N^{1/2})$ such that the space $E = \operatorname{span}\{e^{int} : n \in S\}$ is a Dvoretzky subspace of $L_4([0, 2\pi])$. That is, for every $f \in E$,

$$\|f\|_2 \leq \|f\|_4 \leq C\|f\|_2,$$

where $C > 1$ is a constant independent of f and of N. The dimension $m = \Theta(N^{1/2})$ is likewise given by the (probabilistic) proof of the Milman-Dvoretzky theorem and is optimal.

Λ_4-sets via finite geometries, an outline

We first note that in order to control

$$\|f\|_4^4 = \frac{1}{2\pi}\int_0^{2\pi}\Big|\sum_{n\in S} a_n e^{int}\Big|^4 dt = \frac{1}{2\pi}\int_0^{2\pi}\sum_{n_j\in S} a_{n_1}a_{n_2}\overline{a}_{n_3}\overline{a}_{n_4}e^{i(n_1+n_2-n_3-n_4)t}dt$$

we need to control, for a given α, the number of solutions of $n_1 + n_2 = \alpha$, ideally have only one solution (modulo order).

Assume $N = p^2$ (p prime, or a prime power) and identify $\{0, 1, 2, \ldots, N-1\}$ with $\mathbb{F}_p^2$, where $\mathbb{F}_p$ is the field with p elements. Let $S = \{(k, k^2) : k \in \mathbb{F}_p\}$, then $(k_1, k_1^2) + (k_2, k_2^2) = (\alpha, \alpha')$ can have at most one solution in $\mathbb{Z}$ and just a few in $\mathbb{F}_p$. So such S works.

Similar ideas work for other $q \in 2\mathbb{N}$, but not for $q \notin 2\mathbb{N}$ that. More about the latter case later.

Λ_4-sets, 2nd attempt

Here is a version of the construction for the group $\mathbb{Z}_2^k$ (rather than $\mathbb{Z}$ or $\mathbb{Z}_N$). Let $k = 2r$, so $\mathbb{Z}_2^k = \mathbb{Z}_2^r \times \mathbb{Z}_2^r$. Consider now the set of characters on $\mathbb{Z}_2^k$, indexed by $\alpha \subset \{1, 2, \ldots, k\}$, which can be identified with Walsh functions $w_\alpha = \prod_{j \in \alpha} \epsilon_j$, where ϵ_j is a ± 1-valued Bernoulli variable depending on the j-th coordinate in $\mathbb{Z}_2^k$. There are $N = 2^k = 2^{2r}$ such functions, and so their span is of dimension N.

Let us further identify $\mathbb{Z}_2^r$ with a field $\mathbb{F}_s$ of cardinality $s = 2^r$. Then each $\alpha \subset \{1, 2, \ldots, r\}$ can be thought of as an element of $\mathbb{F}_s$. We now set

$$S := \{(\alpha, \alpha^3) : \alpha \in \mathbb{F}_s\} \subset \mathbb{F}_s \times \mathbb{F}_s = \mathbb{Z}_2^k.$$

Here the power α^3 is meant in the sense of the multiplicative structure of $\mathbb{F}_s$, which is different from that of the ring $\mathbb{Z}_2^r$. Then $\#S = 2^r = N^{1/2}$ and a simple calculation shows that S is a Λ_4-set, whose size is of largest possible order of magnitude.

Non-commutative Λ_4-sets

We now replace Bernoulli variables with Pauli matrices and consider, for $\alpha \in \{0,1\}^k$,

$$X_\alpha := \bigotimes_{j=1}^{k} \sigma_x^{\alpha_j},$$

i.e., a tensor product where the jth factor equals σ_x if $\alpha_j = 1$ and equals $\mathbf{I}$ if $\alpha_j = 0$. These behave exactly like Walsh functions, so again $\{X_\alpha : \alpha \in S\}$ is a Λ_4-set. Analogously, one could introduce the $\Lambda(4)$-sets $\{Y_\beta : \beta \in S\}$ and $\{Z_\gamma : \gamma \in S\}$. One is now tempted to try $\{X_\alpha Y_\beta Z_\gamma : \alpha, \beta, \gamma \in S\}$, which – had it worked – would give a set of cardinality $m = (2^r)^3$. Since the dimension of the Hilbert space is here $d = 2^k = 2^{2r}$, we would have had $m = d^{3/2} = d^{1+2/q}$, exactly as in the Dvoretzky theorem. Unfortunately, this doesn't quite work.

However, if rather than using the previous construction as a black box, we start from scratch and consider $\{X_\alpha Y_\beta : \beta = \alpha^3, \alpha \in \mathbb{F}_d\}$, then we get a $\Lambda(4)$-set, whose size is just on the border of yielding a violation of additivity for $p = 2$.

Random commutative and non-commutative Λ_4-sets

For $q > 2$, $q \notin 2\mathbb{N}$, there are two Acta Math. papers (J. Bourgain (1989), and Talagrand (1995), simplifying the original proof) that basically show that – in the case of characters – random choice of a subset S of the proper size works. Both papers are a *tour the force*, using the state of the art in the analysis of sub-Gaussian processes.

I went through a simpler variant of the Talagrand argument, and it appears to give a Λ_q-subset of Pauli matrices of cardinality $m = \Omega(d^{(1-\varepsilon)(1+2/q)})$. This would be enough to produce a violation for additivity of p-Renyi entropy for any $p > 1$.

This is not an explicit procedure, but requires much less randomness than the original Dvoretzky theorem : $\dim G_{d^2,m} = \Theta(md^2)$ bits vs. $\Theta(m \log d)$ bits. For general q, full derandomization seems beyond reach at this stage. However, for $q = 4$, perhaps some Clifford group magic could help to boost the construction from the previous page...

THANK YOU!