Zonoids and sparsification of quantum measurements

Guillaume AUBRUN
(joint with Cécilia Lancien)

Université Lyon 1, France
Lyapounov convexity theorem

Let $\mu : (\Omega, \mathcal{F})$ be a vector measure, non-atomic. Then

$$\{\mu(A) : A \in \mathcal{F}\} \subset \mathbb{R}^n$$

is a convex set.
Lyapounov convexity theorem

Let $\mu : (\Omega, \mathcal{F})$ be a vector measure, non-atomic. Then

$$\{\mu(A) : A \in \mathcal{F}\} \subset \mathbb{R}^n$$

is a convex set.

Such convex sets are called zonoids.

Equivalently, a zonoid is a limit of zonotopes. A zonotope is a finite Minkowski sum of segments. The Minkowski sum is

$$A + B = \{a + b : a \in A, b \in B\}.$$
Lyapounov convexity theorem

Let $\mu : (\Omega, \mathcal{F})$ be a vector measure, non-atomic. Then

$$\{\mu(A) : A \in \mathcal{F}\} \subset \mathbb{R}^n$$

is a convex set.

Such convex sets are called zonoids.

Equivalently, a zonoid is a limit of zonotopes. A zonotope is a finite Minkowski sum of segments. The Minkowski sum is

$$A + B = \{a + b : a \in A, b \in B\}.$$ 

Also: for a vector measure, the convex hull of the range is a zonoid.
1. The cube is a zonoid.
2. The octahedron is not a zonoid.
3. Any planar compact convex set with a center of symmetry is a zonoid.
4. The Euclidean ball $B_2^n$ is a zonoid

$$B_2^n = \alpha_n \int_{S^{n-1}} [-u, -u] \, d\sigma(u).$$
A Positive Operator-Valued Measure (POVM) is a vector measure

\[ M : (\Omega, \mathcal{F}) \rightarrow \mathcal{M}_+ (\mathbb{C}^d) \]

such that \( M(\Omega) = \text{Id} \). Here \( \mathcal{M}_+ (\mathbb{C}^d) \) is the set of positive self-adjoint \( d \times d \) matrices.
A Positive Operator-Valued Measure (POVM) is a vector measure

\[ M : (\Omega, \mathcal{F}) \rightarrow \mathcal{M}_+ (\mathbb{C}^d) \]

such that \( M(\Omega) = \text{Id} \). Here \( \mathcal{M}_+ (\mathbb{C}^d) \) is the set of positive self-adjoint \( d \times d \) matrices.

POVMs corresponds to quantum measurements.
A Positive Operator-Valued Measure (POVM) is a vector measure

\[ M : (\Omega, \mathcal{F}) \rightarrow \mathcal{M}_+(\mathbb{C}^d) \]

such that \( M(\Omega) = \text{Id} \). Here \( \mathcal{M}_+(\mathbb{C}^d) \) is the set of positive self-adjoint \( d \times d \) matrices.

POVMs corresponds to quantum measurements.

We often consider the special case of discrete POVMs (the purely atomic case). They are given by operators \( (M_1, \ldots, M_N) \), where \( M_i \geq 0 \) and \( M_1 + \cdots + M_N = \text{Id} \). The range is

\[
\{ M(A) ; A \in \mathcal{F} \} = \left\{ \sum_{i \in I} M_i : I \subset \{1, \ldots, N\} \right\}.
\]
The convex hull of the range is a zonoid

\[ \text{conv}\{M(A); A \in \mathcal{F}\} = \sum_{i=1}^{N} [0, M_i]. \]

It is more natural to consider the 0-symmetric version

\[ K_M = 2 \text{conv}\{M(A); A \in \mathcal{F}\} - \text{Id} = \sum_{i=1}^{N} [-M_i, M_i] \]

This is a zonotope inside \( K = \{ A \in \mathcal{M}_+(\mathbb{C}^d) : \|A\|_\infty \leq 1 \}. \)
Zonoid associated to a POVM

The convex hull of the range is a zonoid

\[ \text{conv}\{ M(A) ; A \in \mathcal{F} \} = \sum_{i=1}^{N} [0, M_i]. \]

It is more natural to consider the 0-symmetric version

\[ K_M = 2 \text{conv}\{ M(A) ; A \in \mathcal{F} \} - \text{Id} = \sum_{i=1}^{N} [-M_i, M_i] \]

This is a zonotope inside \( K = \{ A \in \mathcal{M}_+(\mathbb{C}^d) : \|A\|_\infty \leq 1 \}. \)

Conversely, any zonoid inside \( K \) and containing \( \pm \text{Id} \) comes from a POVM.
Support function

Given a POVM $M$, the support function of the zonoid $K_M$ is a norm

$$\|\Delta\|_M = \sup_{A \in K_M} \mathrm{Tr}(\Delta A) = \sum_{i=1}^{N} |\mathrm{Tr} \Delta M_i|.$$
Given a POVM $M$, the support function of the zonoid $K_M$ is a norm

$$
\| \Delta \|_M = \sup_{A \in K_M} \text{Tr}(\Delta A) = \sum_{i=1}^{N} |\text{Tr} \Delta M_i|.
$$

Note that the normed space $(\mathcal{M}_+(\mathbb{C}^d), \| \cdot \|_M)$ embeds into $\ell^N_1 = (\mathbb{R}^N, \| \cdot \|_1)$ (another characterization of zonotopes/zonoids).
Given a POVM $M$, the support function of the zonoid $K_M$ is a norm

$$\|\Delta\|_M = \sup_{A \in K_M} \text{Tr}(\Delta A) = \sum_{i=1}^{N} |\text{Tr} \Delta M_i|.$$  

Note that the normed space $(\mathcal{M}_+(\mathbb{C}^d), \| \cdot \|_M)$ embeds into $\ell_1^N = (\mathbb{R}^N, \| \cdot \|_1)$ (another characterization of zonotopes/zonoids).

As we shall see this norm has an interpretation as distinguishability norms (Matthews–Wehner–Winter).
State discrimination

Let $\rho, \sigma$ two quantum states on $\mathbb{C}^d$. A referee chooses $\rho$ or $\sigma$ with equal probability. You have to guess which was chosen using the POVM $M$ with a single sample.
State discrimination

Let $\rho, \sigma$ two quantum states on $C^d$. A referee chooses $\rho$ or $\sigma$ with equal probability. You have to guess which was chosen using the POVM $M$ with a single sample.

Born’s rule: if $\rho$ was chosen, outcome $i$ is output with probability $\text{Tr} \rho M_i$; if $\sigma$ was chosen, outcome $i$ is output with probability $\text{Tr} \sigma M_i$. 

$$p = \frac{1}{2} - \frac{1}{4} \sum_{i=1}^{N} \min(\text{Tr} \rho M_i, \text{Tr} \sigma M_i)$$

$$= \frac{1}{2} - \frac{1}{4} \| \rho - \sigma \|_M$$
State discrimination

Let \( \rho, \sigma \) two quantum states on \( \mathbb{C}^d \). A referee chooses \( \rho \) or \( \sigma \) with equal probability. You have to guess which was chosen using the POVM \( M \) with a single sample.

Born’s rule: if \( \rho \) was chosen, outcome \( i \) is output with probability \( \text{Tr} \rho M_i \); if \( \sigma \) was chosen, outcome \( i \) is output with probability \( \text{Tr} \sigma M_i \).

The best strategy is of course, given the outcome, to guess the most likely state. The probability of error is

\[
p = \frac{1}{2} \sum_{i=1}^{N} \min(\text{Tr} \rho M_i, \text{Tr} \sigma M_i)
\]

\[
= \frac{1}{2} - \frac{1}{4} \sum_{i=1}^{N} |\text{Tr} \rho M_i - \text{Tr} \sigma M_i|
\]

\[
= \frac{1}{2} - \frac{1}{4} \|\rho - \sigma\|_M
\]
Let $U_d$ be the uniform POVM, defined on $(S_{\mathbb{C}^d}, Borel)$ by

$$U_d(A) = d \int_{A} |\psi\rangle \langle \psi| \, d\sigma(\psi).$$
The uniform POVM

Let $U_d$ be the uniform POVM, defined on $(S_{\mathbb{C}^d}, \text{Borel})$ by

$$U_d(A) = d \int_A \langle \psi | \psi \rangle \, d\sigma(\psi).$$

We would like sparsifications of $U_d$, i.e. POVMs $M$ with as few outcomes as possible and such that

$$(1 - \varepsilon) \| \cdot \|_M \leq \| \cdot \| U_d \leq (1 + \varepsilon) \| \cdot \|_M.$$
Start from the identity \((t \in \mathbb{N})\)

\[
\pi := \int_{S_{\mathbb{C}^d}} |\psi\rangle\langle\psi|^\otimes t \, d\sigma = \frac{1}{\dim \text{Sym}^t(\mathbb{C}^d)} P_{\text{Sym}^t(\mathbb{C}^d)}.
\]
Start from the identity \((t \in \mathbb{N})\)

\[
\pi := \int_{S_{cd}} |\psi\rangle\langle\psi|^{\otimes t} \, d\sigma = \frac{1}{\dim \operatorname{Sym}^t(\mathbb{C}^d)} P_{\operatorname{Sym}^t(\mathbb{C}^d)}.
\]

An \(\varepsilon\)-approximate \(t\)-design is a finitely supported measure \(\mu\) on \(S_{\mathbb{C}^d}\) such that

\[
(1 - \varepsilon)\pi \leq \int_{S_{\mathbb{C}^d}} |\psi\rangle\langle\psi|^{\otimes t} \, d\mu \leq (1 + \varepsilon)\pi.
\]
$t$-designs

Start from the identity ($t \in \mathbb{N}$)

$$
\pi := \int_{S_{\mathbb{C}^d}} |\psi\rangle \langle \psi| \otimes^t d\sigma = \frac{1}{\dim \text{Sym}^t(\mathbb{C}^d)} P_{\text{Sym}^t(\mathbb{C}^d)}.
$$

An $\varepsilon$-approximate $t$-design is a finitely supported measure $\mu$ on $S_{\mathbb{C}^d}$ such that

$$
(1 - \varepsilon)\pi \leq \int_{S_{\mathbb{C}^d}} |\psi\rangle \langle \psi| \otimes^t d\mu \leq (1 + \varepsilon)\pi.
$$

Example: $\varepsilon = 0$ gives an exact integration formula (cubature formula) for homogeneous polynomial of degree $t$. 
Ambainis–Emerson (2007) showed that if $\mu$ is a (exact or approximate) 4-design, then the corresponding POVM $M$ satisfies

$$c \| \cdot \|_M \leq \| \cdot \|_{U_d} \leq \| \cdot \|_M.$$
Ambainis–Emerson (2007) showed that if $\mu$ is a (exact or approximate) 4-design, then the corresponding POVM $M$ satisfies

$$c \| \cdot \|_M \leq \| \cdot \|_{U_d} \leq \| \cdot \|_M.$$  

Idea: the 1-norm can be controlled from 2- and 4-norms

$$\frac{\|X\|_2^3}{\|X\|_4^2} \leq \|X\|_1 \leq \|X\|_2.$$  

This approach requires $\text{card supp}(\mu) \geq \dim \text{Sym}^t(\mathbb{C}^d) = \Omega(d^4)$.
Ambainis–Emerson (2007) showed that if $\mu$ is a (exact or approximate) 4-design, then the corresponding POVM $M$ satisfies

$$c \| \cdot \|_M \leq \| \cdot \|_{U_d} \leq \| \cdot \|_M.$$  

Idea: the 1-norm can be controlled from 2- and 4-norms

$$\frac{\|X\|_{L^2}^3}{\|X\|_{L^4}^2} \leq \|X\|_{L^1} \leq \|X\|_{L^2}.$$  

This approach requires $\text{card supp}(\mu) \geq \dim \text{Sym}^t(\mathbb{C}^d) = \Omega(d^4)$.

Similar to Rudin (1960): $\ell_2^n \subset \ell_4^n$ isometrically and therefore $\ell_2^n \subset \ell_1^n$ with distortion $\sqrt{3}$. Equivalently, gives a zonotope $Z$ with $n^2$ summands such that $Z \subset B_2^n \subset \sqrt{3}Z$. 

Guillaume Aubrun (Lyon)
Rudin’s result can be improved via random constructions based on the concentration of measure phenomenon.
Rudin’s result can be improved via random constructions based on the concentration of measure phenomenon.

Figiel–Lindenstrauss–Milman (1977): given $\varepsilon > 0$, $\ell_2^n$ embeds with distortion $1 + \varepsilon$ in $\ell_1^N$ with $N = C\varepsilon^{-2}n$. Equivalently, there is a zonoid $Z$ with $C\varepsilon^{-2}n$ summands such that $Z \subset B_2^n \subset (1 + \varepsilon)Z$. 

Proof: choose the directions of the $N$ segments independently and uniformly at random.
Rudin’s result can be improved via random constructions based on the concentration of measure phenomenon.

Figiel–Lindenstrauss–Milman (1977): given $\varepsilon > 0$, $\ell^2_n$ embeds with distortion $1 + \varepsilon$ in $\ell^N_1$ with $N = C\varepsilon^{-2}n$.

Equivalently, there is a zonoid $Z$ with $C\varepsilon^{-2}n$ summands such that $Z \subset B^n_2 \subset (1 + \varepsilon)Z$.

Proof: choose the directions of the $N$ segments independently and uniformly at random.
Theorem 1: optimal sparsifications of the uniform POVM

Theorem (A.-Lancien)

Given $d \in \mathbb{N}$ and $\varepsilon \in (0, 1)$, there is a POVM $M$ on $\mathbb{C}^d$ with $N$ outcomes such that $N \leq C\varepsilon^{-2}|\log \varepsilon|d^2$ and

$$(1 - \varepsilon)\| \cdot \|_M \leq \| \cdot \|_{u_d} \leq (1 + \varepsilon)\| \cdot \|_M.$$ 

The size $d^2 = \dim M_{sa}(\mathbb{C}^d)$ is obviously optimal.
Theorem (A.-Lancien)

Given \( d \in \mathbb{N} \) and \( \varepsilon \in (0, 1) \), there is a POVM \( M \) on \( \mathbb{C}^d \) with \( N \) outcomes such that \( N \leq C \varepsilon^{-2} |\log \varepsilon| d^2 \) and

\[
(1 - \varepsilon) \| \cdot \|_M \leq \| \cdot \|_{u_d} \leq (1 + \varepsilon) \| \cdot \|_M.
\]

The size \( d^2 = \dim \mathcal{M}_{sa}(\mathbb{C}^d) \) is obviously optimal.

The construction is random: take \((\psi_i)\) independent, uniform on the sphere \( S_{\mathbb{C}^d} \). Let

\[
S = \sum_{i=1}^{N} |\psi_i\rangle\langle\psi_i|.
\]

The POVM is the family

\[
\left( |S^{-1/2} \psi_i\rangle \langle S^{-1/2} \psi_i| \right)_{1 \leq i \leq N}.
\]
Theorem 1, ideas of the proof

The proof uses standard tools

1. Net arguments (discrete approximation of the unit sphere)
2. Deviation inequalities for sum of sub-exponential random variables.
3. Random matrix estimates to show that the matrix $S$ is well-conditioned.
Theorem 1, ideas of the proof

The proof uses standard tools

1. Net arguments (discrete approximation of the unit sphere)
2. Deviation inequalities for sum of sub-exponential random variables.
3. Random matrix estimates to show that the matrix $S$ is well-conditioned.

What about derandomization?
Tensor product of POVM

There is natural notion of tensor product for POVMs: given (discrete) POVMs \((M_i)_{i \in I}\) and \((N_j)_{j \in J}\) on \(\mathbb{C}^d\), consider \((M_i \otimes N_j)_{i \in I, j \in J}\). Accordingly there is a notion of tensor products for zonoids.
There is a natural notion of tensor product for POVMs: given (discrete) POVMs \((M_i)_{i \in I}\) and \((N_j)_{j \in J}\) on \(\mathbb{C}^d\), consider \((M_i \otimes N_j)_{i \in I, j \in J}\). Accordingly there is a notion of tensor products for zonoids.

Simple fact: if \((1 - \varepsilon)\parallel \cdot \parallel_M \leq \parallel \cdot \parallel_M' \leq (1 + \varepsilon)\parallel \cdot \parallel_M\) and \((1 - \varepsilon)\parallel \cdot \parallel_N \leq \parallel \cdot \parallel_N' \leq (1 + \varepsilon)\parallel \cdot \parallel_N\), then

\[(1 - \varepsilon)^2 \parallel \cdot \parallel_{M \otimes N} \leq \parallel \cdot \parallel_{M' \otimes N'} \leq (1 + \varepsilon)^2 \parallel \cdot \parallel_{M \otimes N}.
\]

It follows from Theorem 1 that there are optimal local sparsifications of the “local uniform POVM” \(LU = U_d \otimes U_d\).
There is natural notion of tensor product for POVMs: given (discrete) POVMs \((M_i)_{i \in I}\) and \((N_j)_{j \in J}\) on \(\mathbb{C}^d\), consider \((M_i \otimes N_j)_{i \in I, j \in J}\). Accordingly there is a notion of tensor products for zonoids.

Simple fact: if \((1 - \varepsilon) \| \cdot \|_M \leq \| \cdot \|_{M'} \leq (1 + \varepsilon) \| \cdot \|_M\) and \((1 - \varepsilon) \| \cdot \|_N \leq \| \cdot \|_{N'} \leq (1 + \varepsilon) \| \cdot \|_N\), then

\[
(1 - \varepsilon)^2 \| \cdot \|_{M \otimes N} \leq \| \cdot \|_{M' \otimes N'} \leq (1 + \varepsilon)^2 \| \cdot \|_{M \otimes N}.
\]

It follows from Theorem 1 that there are optimal local sparsifications of the “local uniform POVM” \(LU = U_d \otimes U_d\).

Note that \(\| \cdot \|_{LU}\) is equivalent to the following norm (Lancien–Winter)

\[
\|\Delta\|_{2(2)}^2 = (\text{Tr} \, \Delta)^2 + \text{Tr}_2(\text{Tr}_1 \, \Delta)^2 + \text{Tr}_1(\text{Tr}_2 \, \Delta)^2 + \text{Tr}(\Delta^2).
\]
A series of results from the late ’80s (Schechtman, Bourgain–Lindenstrauss–Milman, Talagrand) culminating in the following: Any zonoid $K \subset \mathbb{R}^n$ can be $\varepsilon$-approximated by a zonotope $Z$ with $N \leq C\varepsilon^{-2}n \log n$ summands, in the sense

$$K \subset Z \subset (1 + \varepsilon)K.$$
A series of results from the late ’80s (Schechtman, Bourgain–Lindenstrauss–Milman, Talagrand) culminating in the following: Any zonoid $K \subset \mathbb{R}^n$ can be $\varepsilon$-approximated by a zonotope $Z$ with $N \leq C \varepsilon^{-2} n \log n$ summands, in the sense

$$K \subset Z \subset (1 + \varepsilon)K.$$ 

Important open question: can we remove the log $n$ factor?
The POVM version of the previous theorem is the following.

**Theorem (A.-Lancien)**

Any POVM $M$ on $\mathbb{C}^d$ can be $\varepsilon$-approximated by a sub-POVM $M'$ with $N \leq C\varepsilon^{-2}d^2 \log d$ outcomes, in the sense

$$
(1 - \varepsilon)\|\cdot\|_{M'} \leq \|\cdot\|_M \leq (1 + \varepsilon)\|\cdot\|_{M'}.
$$

A sub-POVM is a finite family $(M_i)$ of positive operators with $\sum M_i \leq \text{Id}$. 