



**HAL**  
open science

## Zonoids and sparsification of quantum measurements

Guillaume Aubrun, Cécilia Lancien

► **To cite this version:**

Guillaume Aubrun, Cécilia Lancien. Zonoids and sparsification of quantum measurements. *Positivity*, 2016, 20 (1), pp.1 - 23. 10.1007/s11117-015-0337-5 . hal-01890603

**HAL Id: hal-01890603**

**<https://hal.science/hal-01890603v1>**

Submitted on 16 Jan 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# ZONOIDS AND SPARSIFICATION OF QUANTUM MEASUREMENTS

GUILLAUME AUBRUN AND CÉCILIA LANCINIEN

**ABSTRACT.** In this paper, we establish a connection between zonoids (a concept from classical convex geometry) and the distinguishability norms associated to quantum measurements or POVMs (Positive Operator-Valued Measures), recently introduced in quantum information theory.

This correspondence allows us to state and prove the POVM version of classical results from the local theory of Banach spaces about the approximation of zonoids by zonotopes. We show that on  $\mathbf{C}^d$ , the uniform POVM (the most symmetric POVM) can be sparsified, i.e. approximated by a discrete POVM having only  $O(d^2)$  outcomes. We also show that similar (but weaker) approximation results actually hold for any POVM on  $\mathbf{C}^d$ .

By considering an appropriate notion of tensor product for zonoids, we extend our results to the multipartite setting: we show, roughly speaking, that local POVMs may be sparsified locally. In particular, the local uniform POVM on  $\mathbf{C}^{d_1} \otimes \cdots \otimes \mathbf{C}^{d_k}$  can be approximated by a discrete POVM which is local and has  $O(d_1^2 \times \cdots \times d_k^2)$  outcomes.

## INTRODUCTION

A classical result by Lyapounov ([26], Theorem 5.5) asserts that the range of a non-atomic  $\mathbf{R}^n$ -valued vector measure is closed and convex. Convex sets in  $\mathbf{R}^n$  obtained in this way are called zonoids. Zonoids are equivalently characterized as convex sets which can be approximated by finite sums of segments.

In this paper we consider a special class of vector measures: Positive Operator-Valued Measures (POVMs). In the formalism of quantum mechanics, POVMs represent the most general form of a quantum measurement. Recently, Matthews, Wehner and Winter [22] introduced the distinguishability norm associated to a POVM. This norm has an operational interpretation as the bias of the POVM for the state discrimination problem (a basic task in quantum information theory) and is closely related to the zonoid arising from Lyapounov's theorem.

A well-studied question in high-dimensional convexity is the approximation of zonoids by zonotopes. The series of papers [11, 28, 7, 31] culminates in the following result: any zonoid in  $\mathbf{R}^n$  can be approximated by the sum of  $O(n \log n)$  segments. The aforementioned connection between POVMs and zonoids allows us to state and prove approximation results for POVMs, which improve on previously known bounds. Precise statements appear as Theorem 4.3 and 4.4.

Our article is organized as follows. Section 1 introduces POVMs and their associated distinguishability norms. Section 2 connects POVMs with zonoids. Section 3 introduces a notion of tensor product for POVMs, and the corresponding notion for zonoids. Section 4 pushes forward this connection to state the POVM version of approximation results for zonoids, which are proved in Sections 6, 7 and 8. Section 5 provides sparsification results for local POVMs on multipartite systems.

The reader may have a look at Table 1, which summarizes analogies between zonoids and POVMs.

**Notation.** We denote by  $\mathcal{H}(\mathbf{C}^d)$  the space of Hermitian operators on  $\mathbf{C}^d$ , and by  $\mathcal{H}_+(\mathbf{C}^d)$  the subset of positive operators. We denote by  $\|\cdot\|_1$  the trace class norm, by  $\|\cdot\|_\infty$  the operator norm and by  $\|\cdot\|_2$  the Hilbert–Schmidt norm. Notation  $[-\text{Id}, \text{Id}]$  stands for the set of self-adjoint operators  $A$  such that  $-\text{Id} \leq A \leq \text{Id}$ . In other words  $[-\text{Id}, \text{Id}]$  is the self-adjoint part of the unit ball for  $\|\cdot\|_\infty$ . We denote by  $S(\mathbf{C}^d)$  the set of states on  $\mathbf{C}^d$  (a state is a positive operator with trace 1).

Let us recall a few standard concepts from classical convex geometry that we will need throughout our proofs. The support function  $h_K$  of a convex compact set  $K \subset \mathbf{R}^n$  is the function defined for  $x \in \mathbf{R}^n$  by  $h_K(x) = \sup\{\langle x, y \rangle : y \in K\}$ . Moreover, for a pair  $K, L$  of convex compact sets, the inclusion  $K \subset L$  is equivalent to the inequality  $h_K \leq h_L$ . The polar of a convex set  $K \subset \mathbf{R}^n$  is  $K^\circ = \{x \in \mathbf{R}^n : \langle x, y \rangle \leq 1 \text{ whenever } y \in K\}$ . The bipolar theorem (see e.g. [3]) states that  $(K^\circ)^\circ$  is the closed convex hull of  $K$  and  $\{0\}$ . A convex body is a convex compact set with non-empty interior. Whenever we apply tools from convex geometry in the (real) space  $\mathcal{H}(\mathbf{C}^d)$  (e.g. polar or support function), we use the Hilbert–Schmidt inner product  $(A, B) \mapsto \text{Tr } AB$  to define the Euclidean structure.

1991 *Mathematics Subject Classification.* 52A21, 81P15, 81P45.

*Key words and phrases.* positive operator-valued measure, zonoid, sparsification.

This research was supported by the ANR project OSQPI ANR-11-BS01-0008.

The letters  $C, c, c_0, \dots$  denote numerical constants, independent from any other parameters such as the dimension. The value of these constants may change from occurrence to occurrence. Similarly  $c(\varepsilon)$  denotes a constant depending only on the parameter  $\varepsilon$ . We also use the following convention: whenever a formula is given for the dimension of a (sub)space, it is tacitly understood that one should take the integer part.

## 1. POVMs AND DISTINGUISHABILITY NORMS

In quantum mechanics, the state of a  $d$ -dimensional system is described by a positive operator on  $\mathbf{C}^d$  with trace 1. The most general form of a measurement that may be performed on such a quantum system is encompassed by the formalism of Positive Operator-Valued Measures (POVMs). Given a set  $\Omega$  equipped with a  $\sigma$ -algebra  $\mathcal{F}$ , a POVM on  $\mathbf{C}^d$  is a map  $M : \mathcal{F} \rightarrow \mathcal{H}_+(\mathbf{C}^d)$  which is  $\sigma$ -additive and such that  $M(\Omega) = \text{Id}$ . In this definition the space  $(\Omega, \mathcal{F})$  could potentially be infinite, so that the POVMs defined on it would be continuous. However, we often restrict ourselves to the subclass of discrete POVMs, and a main point of this article is to substantiate this “continuous to discrete” transition.

A discrete POVM is a POVM in which the underlying  $\sigma$ -algebra  $\mathcal{F}$  is required to be finite. In that case there is a finite partition  $\Omega = A_1 \cup \dots \cup A_n$  generating  $\mathcal{F}$ . The positive operators  $M_i = M(A_i)$  are often referred to as the elements of the POVM, and they satisfy the condition  $M_1 + \dots + M_n = \text{Id}$ . We usually identify a discrete POVM with the set of its elements by writing  $M = (M_i)_{1 \leq i \leq n}$ . The index set  $\{1, \dots, n\}$  labels the outcomes of the measurement. The integer  $n$  is thus the number of outcomes of  $M$  and can be seen as a crude way to measure the complexity of  $M$ .

What happens when measuring with a POVM  $M$  a quantum system in a state  $\rho$ ? In the case of a discrete POVM  $M = (M_i)_{1 \leq i \leq n}$ , we know from Born’s rule that the outcome  $i$  is output with probability  $\text{Tr}(\rho M_i)$ . This simple formula can be used to quantify the efficiency of a POVM to perform the task of state discrimination. State discrimination can be described as follows: a quantum system is prepared in an unknown state which is either  $\rho$  or  $\sigma$  (both hypotheses being a priori equally likely), and we have to guess the unknown state. After measuring it with the discrete POVM  $M = (M_i)_{1 \leq i \leq n}$ , the optimal strategy, based on the maximum likelihood probability, leads to a probability of wrong guess equal to [17, 16]

$$\mathbf{P}_{error} = \frac{1}{2} \left( 1 - \frac{1}{2} \sum_{i=1}^n |\text{Tr}(\rho M_i) - \text{Tr}(\sigma M_i)| \right).$$

In this context, the quantity  $\frac{1}{2} \sum_{i=1}^n |\text{Tr}(\rho M_i) - \text{Tr}(\sigma M_i)|$  is therefore called the bias of the POVM  $M$  on the state pair  $(\rho, \sigma)$ .

Following [22], we introduce a norm on  $\mathcal{H}(\mathbf{C}^d)$ , called the distinguishability norm associated to  $M$ , and defined for  $\Delta \in \mathcal{H}(\mathbf{C}^d)$  by

$$(1) \quad \|\Delta\|_M = \sum_{i=1}^n |\text{Tr}(\Delta M_i)|.$$

It is such that  $\mathbf{P}_{error} = \frac{1}{2} (1 - \frac{1}{2} \|\rho - \sigma\|_M)$ , and thus quantifies how powerful the POVM  $M$  is in discriminating one state from another with the smallest probability of error.

The terminology “norm” is slightly abusive since one may have  $\|\Delta\|_M = 0$  for a nonzero  $\Delta \in \mathcal{H}(\mathbf{C}^d)$ . The functional  $\|\cdot\|_M$  is however always a semi-norm, and it is easy to check that  $\|\cdot\|_M$  is a norm if and only if the POVM elements  $(M_i)_{1 \leq i \leq n}$  span  $\mathcal{H}(\mathbf{C}^d)$  as a vector space. Such POVMs are called informationally complete in the quantum information literature.

Similarly, the distinguishability norm associated to a general POVM  $M$ , defined on a set  $\Omega$  equipped with a  $\sigma$ -algebra  $\mathcal{F}$ , is described for  $\Delta \in \mathcal{H}(\mathbf{C}^d)$  by

$$(2) \quad \|\Delta\|_M = \|\text{Tr}(\Delta M(\cdot))\|_{\text{TV}} = \sup_{A \in \mathcal{F}} [\text{Tr}(\Delta M(A)) - \text{Tr}(\Delta M(\Omega \setminus A))] = \sup_{M \in \mathcal{M}(\mathcal{F})} \text{Tr}(\Delta(2M - \text{Id})).$$

Here  $\|\mu\|_{\text{TV}}$  denotes the total variation of a measure  $\mu$ . When  $M$  is discrete, formulae (1) and (2) coincide. Note also that the inequality  $\|\cdot\|_M \leq \|\cdot\|_1$  holds for any POVM  $M$ , with equality on  $\mathcal{H}_+(\mathbf{C}^d)$ .

Given a POVM  $M$ , we denote by  $B_M = \{\|\cdot\|_M \leq 1\}$  the unit ball for the distinguishability norm, and  $K_M = (B_M)^\circ$  its polar, i.e.

$$K_M = \{A \in \mathcal{H}(\mathbf{C}^d) : \text{Tr}(AB) \leq 1 \text{ whenever } \|B\|_M \leq 1\}.$$

The set  $K_M$  is a compact convex set. Moreover  $K_M$  has nonempty interior if and only if the POVM  $M$  is informationally complete. It follows from the inequality  $\|\cdot\|_M \leq \|\cdot\|_1$  that  $K_M$  is always included in the operator interval  $[-\text{Id}, \text{Id}]$ .

Zonotope which is the Minkowski sum of $N$ segments	Discrete POVM with $N$ outcomes
Zonoid = limit of zonotopes	General POVM = limit of discrete POVMs
Tensor product of zonoids	Local POVM on a multipartite system
Euclidean ball $B_2^n$ = most symmetric zonoid in $\mathbf{R}^n$	Uniform POVM $U_d$ = most symmetric POVM on $\mathbf{C}^d$
“4th moment method” ([25], explicit): $cB_2^n \subset Z \subset CB_2^n$ , with $Z$ a zonotope which is the sum of $O(n^2)$ segments.	“Approximate 4-design POVM” [1]: explicit sparsification of $U_d$ with $O(d^4)$ outcomes.
Measure concentration ([11], non-explicit): $(1 - \varepsilon)B_2^n \subset Z \subset (1 + \varepsilon)B_2^n$ , with $Z$ a zonotope which is the sum of $O_\varepsilon(n)$ segments.	Theorem 4.3: a randomly chosen POVM with $O(d^2)$ outcomes is a sparsification of $U_d$ .
Derandomization [12, 21, 19]	?
Any zonoid in $\mathbf{R}^n$ can be approximated by a zonotope which is the sum of $O(n \log n)$ segments [31].	Theorem 4.4: any POVM on $\mathbf{C}^d$ can be sparsified into a sub-POVM with $O(d^2 \log d)$ outcomes.

TABLE 1. A “dictionary” between zonoids and POVMs

On the other hand, it follows from (2) that  $B_M = (2M(\mathcal{F}) - \text{Id})^\circ$ , and the bipolar theorem implies that

$$(3) \quad K_M = 2 \text{conv}(M(\mathcal{F})) - \text{Id}.$$

By Lyapounov’s theorem, the convex hull operation is not needed when  $M$  is non-atomic. For a discrete POVM  $M = (M_i)_{1 \leq i \leq n}$ , equation (3) may be rewritten in the form

$$(4) \quad K_M = \text{conv}\{\pm M_1\} + \cdots + \text{conv}\{\pm M_n\},$$

where the addition of convex sets should be understood as the Minkowski sum:  $A + B = \{a + b : a \in A, b \in B\}$ .

We are going to show that POVMs can be sparsified, i.e approximated by discrete POVMs with few outcomes. The terminology “approximation” here refers to the associated distinguishability norms: a POVM  $M$  is considered to be “close” to a POVM  $M'$  when their distinguishability norms satisfy inequalities of the form

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

This notion of approximation has an operational significance: two POVMs are comparable when both lead to comparable biases when used for any state discrimination task. Let us perhaps stress that point: if one has additional information on the states to be discriminated, it may of course be used to design a POVM specifically efficient for those (one could for instance be interested in the problem of distinguishing pairs of low-rank states [30, 1]).

In this paper, we study the distinguishability norms from a functional-analytic point of view. We are mostly interested in the asymptotic regime, when the dimension  $d$  of the underlying Hilbert space is large.

## 2. POVMs AND ZONONDS

**2.1. POVMs as probability measures on states.** The original definition of a POVM involves an abstract measure space, and the specification of this measure space is irrelevant when considering the distinguishability norms. The following proposition, which is probably well-known, gives a more concrete look at POVMs as probability measures on the set  $S(\mathbf{C}^d)$  of states on  $\mathbf{C}^d$ .

**Proposition 2.1.** *Let  $M$  be a POVM on  $\mathbf{C}^d$ . There is a unique Borel probability measure  $\mu$  on  $S(\mathbf{C}^d)$  with barycenter equal to  $\text{Id}/d$  and such that, for any  $\Delta \in \mathcal{H}(\mathbf{C}^d)$ ,*

$$(5) \quad \|\Delta\|_M = d \int_{S(\mathbf{C}^d)} |\text{Tr}(\Delta\rho)| \, d\mu(\rho).$$

*Conversely, given a Borel probability measure  $\mu$  with barycenter equal to  $\text{Id}/d$ , there is a POVM  $M$  such that (5) is satisfied.*

*Proof.* We use the polar decomposition for vector measures, which follows from applying the Radon–Nikodym theorem to vector measures (see [27], Theorem 6.12): a vector measure  $\mu$  defined on a  $\sigma$ -algebra  $\mathcal{F}$  on  $\Omega$  and taking

values in a normed space  $(\mathbf{R}^n, \|\cdot\|)$  satisfies  $d\mu = hd|\mu|$  for some measurable function  $h : \Omega \rightarrow \mathbf{R}^n$ . Moreover, one has  $\|h\| = 1$   $|\mu|$ -a.e. Here  $|\mu|$  denotes the total variation measure of  $\mu$ .

Let  $M$  be a POVM on  $\mathbf{C}^d$ , defined on a  $\sigma$ -algebra  $\mathcal{F}$  on  $\Omega$ . We equip  $\mathcal{H}(\mathbf{C}^d)$  with the trace norm, so that we simply have  $|M| = \text{Tr } M$  and  $|M|(\Omega) = d$ . The polar decomposition yields a measurable function  $h : \Omega \rightarrow \mathcal{H}(\mathbf{C}^d)$  such that  $\|h\|_1 = 1$   $|M|$ -a.e. Moreover, the fact that  $M(\mathcal{F}) \subset \mathcal{H}_+(\mathbf{C}^d)$  implies that  $h \in \mathcal{H}_+(\mathbf{C}^d)$   $|M|$ -a.e. Let  $\mu$  be the push forward of  $\frac{1}{d}|M|$  under the map  $h$ . We have

$$\text{Id} = M(\Omega) = \int_{\Omega} h \, d|M| = d \int_{\mathcal{H}(\mathbf{C}^d)} \rho \, d\mu(\rho).$$

And since  $h \in S(\mathbf{C}^d)$  a.e.,  $\mu$  is indeed a Borel probability measure on  $S(\mathbf{C}^d)$ , with barycenter equal to  $\text{Id}/d$ . Finally, for any  $\Delta \in \mathcal{H}(\mathbf{C}^d)$ ,

$$\|\Delta\|_M = \int_{\Omega} |\text{Tr}(\Delta h)| \, d|M| = d \int_{S(\mathbf{C}^d)} |\text{Tr}(\Delta \rho)| \, d\mu(\rho).$$

We postpone the proof of uniqueness to the next subsection (see after Proposition 2.5).

Conversely, given a Borel probability measure  $\mu$  on  $S(\mathbf{C}^d)$  with barycenter at  $\text{Id}/d$ , consider the vector measure  $M : \mathcal{B} \rightarrow \mathcal{H}(\mathbf{C}^d)$ , where  $\mathcal{B}$  is the Borel  $\sigma$ -algebra on  $S(\mathbf{C}^d)$ , defined by

$$M(A) = d \int_A \rho \, d\mu(\rho).$$

It is easily checked that  $M$  is a POVM and that formula (5) is satisfied.  $\square$

Note that in the case of a discrete POVM  $M = (M_i)_{1 \leq i \leq n}$ , the corresponding probability measure is

$$\mu = \frac{1}{d} \sum_{i=1}^n (\text{Tr } M_i) \delta_{\frac{M_i}{\text{Tr } M_i}}.$$

**Corollary 2.2.** *Given a POVM  $M$  on  $\mathbf{C}^d$ , there is a sequence  $(M_n)$  of discrete POVMs such that  $K_{M_n}$  converges to  $K_M$  in Hausdorff distance. Moreover, if  $\mu$  (resp.  $\mu_n$ ) denotes the probability measure on  $S(\mathbf{C}^d)$  associated to  $M$  (resp. to  $M_n$ ) as in (5), we can guarantee that the support of  $\mu_n$  is contained into the support of  $\mu$ .*

*Proof.* Let  $\mu$  be the probability measure associated to  $M$ . Given  $n$ , let  $(Q_k)$  be a finite partition of  $S(\mathbf{C}^d)$  into sets of diameter at most  $1/n$  with respect to the trace norm. Let  $\rho_k \in S(\mathbf{C}^d)$  be the barycenter of the restriction of  $\mu$  to  $Q_k$  (only defined when  $\mu(Q_k) > 0$ ). The probability measure

$$\mu_n = \sum_k \mu(Q_k) \delta_{\rho_k}$$

has the same barycenter as  $\mu$ , and the associated POVM  $M_n$  satisfies

$$|h_{K_M}(\Delta) - h_{K_{M_n}}(\Delta)| \leq d \frac{\|\Delta\|_{\infty}}{n},$$

and therefore  $K_{M_n}$  converges to  $K_M$ .

The condition on the supports can be enforced by changing slightly the definition of  $\mu_n$ . For each  $k$  we can write  $\rho_k = \sum \lambda_{k,j} \rho_{k,j}$ , where  $(\lambda_{k,j})$  is a convex combination and  $(\rho_{k,j})$  belong to the support of  $\mu$  restricted to  $Q_k$ . The measure

$$\mu'_n = \sum_k \mu(Q_k) \sum_j \lambda_{k,j} \delta_{\rho_{k,j}}$$

satisfies the same properties as  $\mu_n$ , and its support is contained into the support of  $\mu$ .  $\square$

**2.2. POVMs and zonoids.** We connect here POVMs with zonoids, which form an important family of convex bodies (see [6, 29, 12] for surveys on zonoids to which we refer for all the material presented here). A zonotope  $Z \subset \mathbf{R}^n$  is a closed convex set which can be written as the Minkowski sum of finitely many segments, i.e. such that there exist finite sets of vectors  $(u_i)_{1 \leq i \leq N}$  and  $(v_i)_{1 \leq i \leq N}$  in  $\mathbf{R}^n$  such that

$$(6) \quad Z = \text{conv}\{u_1, v_1\} + \cdots + \text{conv}\{u_N, v_N\}.$$

A zonoid is a closed convex set which can be approximated by zonotopes (with respect to the Hausdorff distance). Every zonoid has a center of symmetry, and therefore can be translated into a (centrally) symmetric zonoid. Note that for a centrally symmetric zonotope, we can choose  $v_i = -u_i$  in (6).

Here are equivalent characterizations of zonoids.

**Proposition 2.3.** *Let  $K \subset \mathbf{R}^n$  be a symmetric closed convex set. The following are equivalent.*

- (i)  $K$  is a zonoid.

- (ii) There is a Borel positive measure  $\nu$  on the sphere  $S^{n-1}$  which is even (i.e. such that  $\nu(A) = \nu(-A)$  for any Borel set  $A \subset S^{n-1}$ ) and such that, for every  $x \in \mathbf{R}^n$ ,

$$(7) \quad h_K(x) = \int_{S^{n-1}} |\langle x, \theta \rangle| d\nu(\theta).$$

- (iii) There is a vector measure  $\mu : (\Omega, \mathcal{F}) \rightarrow \mathbf{R}^n$  such that  $K = \mu(\mathcal{F})$ .

Moreover, when these conditions are satisfied, the measure  $\nu$  is unique.

**Remark 2.4.** Having the measure  $\nu$  supported on the sphere and be even is only a matter of normalization and a way to enforce uniqueness: if  $\nu$  is a Borel measure on  $\mathbf{R}^n$  for which linear forms are integrable, there is a symmetric zonoid  $K \subset \mathbf{R}^n$  such that

$$h_K(x) = \int_{\mathbf{R}^n} |\langle x, y \rangle| d\nu(y).$$

As an immediate consequence, we characterize which subsets of  $[-\text{Id}, \text{Id}]$  arise as  $K_M$  for some POVM  $M$ .

**Proposition 2.5.** Let  $K \subset \mathcal{H}(\mathbf{C}^d)$  be a symmetric closed convex set. Then the following are equivalent.

- (i)  $K$  is a zonoid such that  $K \subset [-\text{Id}, \text{Id}]$  and  $\pm \text{Id} \in K$ .  
 (ii) There exists a POVM  $M$  on  $\mathbf{C}^d$  such that  $K = K_M$ .

Moreover,  $K$  is a zonotope only if the POVM  $M$  can be chosen to be discrete.

*Proof.* Let  $K$  be a zonoid such that  $\pm \text{Id} \in K \subset [-\text{Id}, \text{Id}]$ . From Proposition 2.3, there is a vector measure  $\mu$  defined on a  $\sigma$ -algebra  $\mathcal{F}$  on a set  $\Omega$ , whose range is  $K$ . Let  $A \in \mathcal{F}$  such that  $\mu(A) = -\text{Id}$ . The vector measure  $M$  defined for  $B \in \mathcal{F}$  by

$$M(B) = \frac{1}{2}(\mu(B \setminus A) - \mu(B \cap A)) = \frac{1}{2}(\mu(B \Delta A) + \text{Id})$$

is a POVM. Indeed, its range, which equals  $\frac{1}{2}(K + \text{Id})$ , lies inside the positive semidefinite cone, and contains  $\text{Id}$ . We get from (3) that  $K_M = K$ .

Conversely, for any POVM  $M$ , formula (3) implies that  $\pm \text{Id} \in K \subset [-\text{Id}, \text{Id}]$ . The fact that  $K$  is a zonoid follows, using the general fact that the convex hull of the range of a vector measure is a zonoid (see [6], Theorem 1.6).

In the case of zonotopes and discrete POVMs, these arguments have more elementary analogues which we do not repeat.  $\square$

We can now argue about the uniqueness part in Proposition 2.1. This is indeed a consequence of the uniqueness of the measure associated to a zonoid in Proposition 2.3: after rescaling and symmetrization, a measure  $\mu$  on  $S(\mathbf{C}^d)$  satisfying (5) naturally induces a measure  $\nu$  on the Hilbert–Schmidt sphere satisfying (7) for  $K = K_M$ .

Another characterization of zonoids involves the Banach space  $L^1 = L^1([0, 1])$ . A symmetric convex body  $K$  is a zonoid if and only if the normed space  $(\mathbf{R}^n, h_K)$  embeds isometrically into  $L^1$ . Therefore, Proposition 2.5 can be restated as a characterization of distinguishability norms on  $\mathcal{H}(\mathbf{C}^d)$ .

**Corollary 2.6.** Let  $\|\cdot\|$  be a norm on  $\mathcal{H}(\mathbf{C}^d)$ . The following are equivalent

- (1) There is POVM  $M$  on  $\mathbf{C}^d$  such that  $\|\cdot\| = \|\cdot\|_M$ .  
 (2) The normed space  $(\mathcal{H}(\mathbf{C}^d), \|\cdot\|)$  is isometric to a subspace of  $L^1$ , and the following inequality is satisfied for any  $\Delta \in \mathcal{H}(\mathbf{C}^d)$

$$|\text{Tr } \Delta| \leq \|\Delta\| \leq \text{Tr } |\Delta|.$$

### 3. LOCAL POVMs AND TENSOR PRODUCTS OF ZONONDS

**3.1. Tensor products for zonoids.** There is a natural notion of tensor product for subspaces of  $L^1$  which appeared in the Banach space literature (see e.g. [10]).

**Definition 3.1.** Let  $X, Y$  be two Banach spaces which can be embedded isometrically into  $L^1$ , i.e. such that there exist linear norm-preserving maps  $i : X \rightarrow L^1(\mu)$  and  $j : Y \rightarrow L^1(\nu)$ . Then, the 1-tensor product of  $X$  and  $Y$  is defined as the completion of the algebraic tensor product  $X \otimes Y$  for the norm

$$\left\| \sum_k x_k \otimes y_k \right\|_{X \otimes Y} = \int \int \left| \sum_k i(x_k)(s)j(y_k)(t) \right| d\mu(s)d\nu(t).$$

It can be checked that the norm above is well-defined and does not depend on the particular choice of the embeddings  $i, j$  (see e.g. [10] or Lemma 2 in [24]).

In the finite-dimensional case, subspaces of  $L^1$  are connected to zonoids. Therefore, Definition 3.1 leads naturally to a notion of tensor product for (symmetric) zonoids.

**Definition 3.2.** Let  $K \subset \mathbf{R}^m$  and  $L \subset \mathbf{R}^n$  be two symmetric zonoids, and suppose that  $\nu_K$  and  $\nu_L$  are Borel measures on  $S^{m-1}$  and  $S^{n-1}$  respectively, such that for any  $x \in \mathbf{R}^m$  and  $y \in \mathbf{R}^n$ ,

$$h_K(x) = \int_{S^{m-1}} |\langle x, \theta \rangle| d\nu_K(\theta) \quad \text{and} \quad h_L(y) = \int_{S^{n-1}} |\langle y, \phi \rangle| d\nu_L(\phi).$$

The zonoid tensor product of  $K$  and  $L$  is defined as the zonoid  $K \otimes^Z L \subset \mathbf{R}^m \otimes \mathbf{R}^n$  whose support function satisfies

$$(8) \quad h_{K \otimes^Z L}(z) = \int_{S^{m-1}} \int_{S^{n-1}} |\langle z, \theta \otimes \phi \rangle| d\nu_K(\theta) d\nu_L(\phi)$$

for any  $z \in \mathbf{R}^m \otimes \mathbf{R}^n$ .

As in Definition 3.1, this construction does not depend on the choice of the measures  $\nu_K$  and  $\nu_L$ . This can be seen directly: given  $z \in \mathbf{R}^m \otimes \mathbf{R}^n$  and  $\phi \in S^{n-1}$ , set  $\tilde{z}(\phi) = (\text{Id} \otimes \langle \phi |)(z)$ . We have

$$(9) \quad h_{K \otimes^Z L}(z) = \int_{S^{n-1}} h_K(\tilde{z}(\phi)) d\nu_L(\phi),$$

and therefore  $K \otimes^Z L$  does not depend on  $\nu_K$ . The same argument applies for  $\nu_L$ .

In the case of zonotopes, the zonoid tensor product takes a simpler form :

$$\left( \sum_i \text{conv}\{\pm v_i\} \right) \otimes^Z \left( \sum_j \text{conv}\{\pm w_j\} \right) = \sum_i \sum_j \text{conv}\{\pm v_i \otimes w_j\}.$$

Here is a first simple property of the zonoid tensor product.

**Lemma 3.3.** Given symmetric zonoids  $K, L$  and linear maps  $S, T$ , we have

$$S(K) \otimes^Z T(L) = (S \otimes T)(K \otimes^Z L)$$

Additionally, and crucially for the applications we have in mind, the zonoid tensor product is compatible with inclusions.

**Lemma 3.4.** Let  $K, K'$  be two symmetric zonoids in  $\mathbf{R}^m$  with  $K \subset K'$ , and let  $L, L'$  be two symmetric zonoids in  $\mathbf{R}^n$  with  $L \subset L'$ . Then

$$K \otimes^Z L \subset K' \otimes^Z L'.$$

*Proof.* This is a special case of Lemma 2 in [24]. Here is a proof in the language of zonoids. We may assume that  $L = L'$ , the general case following then by arguing that  $K \otimes^Z L \subset K' \otimes^Z L \subset K' \otimes^Z L'$ .

In terms of support functions, we are thus reduced to showing that the inequality  $h_K \leq h_{K'}$  implies the inequality  $h_{K \otimes^Z L} \leq h_{K' \otimes^Z L}$ , which is an easy consequence of (9).  $\square$

Suppose that  $(X, \|\cdot\|_X)$  and  $(Y, \|\cdot\|_Y)$  are Banach spaces with Euclidean norms, i.e. induced by some inner products  $\langle \cdot, \cdot \rangle_X$  and  $\langle \cdot, \cdot \rangle_Y$ . Their Euclidean tensor product  $X \otimes^2 Y$  is defined (after completion) by the norm induced by the inner product on the algebraic tensor product which satisfies

$$\langle x \otimes y, x' \otimes y' \rangle = \langle x, x' \rangle_X \langle y, y' \rangle_Y.$$

It turns out that, for Euclidean norms, the tensor norms  $\otimes^1$  and  $\otimes^2$  are equivalent.

**Proposition 3.5** (see [24, 5]). If  $X$  and  $Y$  are two Banach spaces equipped with Euclidean norms, then

$$\sqrt{\frac{2}{\pi}} \|\cdot\|_{X \otimes^2 Y} \leq \|\cdot\|_{X \otimes^1 Y} \leq \|\cdot\|_{X \otimes^2 Y}.$$

**3.2. Local POVMs.** In quantum mechanics, when a system is shared by several parties, the underlying global Hilbert space is the tensor product of the local Hilbert spaces corresponding to each of the subsystems. A physically relevant class of POVMs on such a multipartite system is the one of local POVMs, describing the situation where each party is only able to perform measurements on his own subsystem.

**Definition 3.6.** For  $i = 1, 2$ , let  $M_i$  denote a POVM on  $\mathbf{C}^{d_i}$ , defined on a  $\sigma$ -algebra  $\mathcal{F}_i$  on a set  $\Omega_i$ . The tensor POVM  $M_1 \otimes M_2$  is the unique map defined on the product  $\sigma$ -algebra  $\mathcal{F}_1 \otimes \mathcal{F}_2$  on  $\Omega_1 \times \Omega_2$ , and such that

$$(M_1 \otimes M_2)(A_1 \times A_2) = M_1(A_1) \otimes M_2(A_2)$$

for every  $A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2$ . By construction,  $M_1 \otimes M_2$  is a POVM on  $\mathbf{C}^{d_1} \otimes \mathbf{C}^{d_2}$ .

In the discrete case, this definition becomes more transparent: if  $M = (M_i)_{1 \leq i \leq m}$  and  $N = (N_j)_{1 \leq j \leq n}$  are discrete POVMs, then  $M \otimes N$  is also discrete, and

$$M \otimes N = (M_i \otimes N_j)_{1 \leq i \leq m, 1 \leq j \leq n}.$$

POVMs on  $\mathbf{C}^{d_1} \otimes \mathbf{C}^{d_2}$  which can be decomposed as tensor product of two POVMs are called local POVMs. If we identify the POVMs  $M_1$  and  $M_2$  with measures  $\mu_1$  and  $\mu_2$  as in Proposition 2.1, then the measure corresponding to  $M_1 \otimes M_2$  is the image of the product measure  $\mu_1 \times \mu_2$  under the map  $(\rho, \sigma) \mapsto \rho \otimes \sigma$ . It thus follows that

**Proposition 3.7.** *If  $M$  and  $N$  are two POVMs, then  $\|\cdot\|_{M \otimes N} = \|\cdot\|_M \otimes^1 \|\cdot\|_N$  and  $K_{M \otimes N} = K_M \otimes^Z K_N$ .*

These definitions and statements are given here only in the bipartite case for the sake of clarity, but can be extended to the situation where a system is shared between any number  $k$  of parties.

#### 4. SPARSIFYING POVMs

**4.1. The uniform POVM.** It has been proved in [22] that, in several senses, the “most efficient” POVM on  $\mathbf{C}^d$  is the “most symmetric” one, i.e. the uniform POVM  $U_d$ , which corresponds to the uniform measure on the set of pure states in the representation (5) from Proposition 2.5.

The corresponding norm is

$$(10) \quad \|\Delta\|_{U_d} = d \mathbf{E} |\langle \psi | \Delta | \psi \rangle|,$$

where  $\psi$  is a random Haar-distributed unit vector.

An important property is that the norm  $\|\cdot\|_{U_d}$  is equivalent to a “modified” Hilbert–Schmidt norm.

**Proposition 4.1** ([15, 20]). *For every  $\Delta \in \mathcal{H}(\mathbf{C}^d)$ , we have*

$$(11) \quad \frac{1}{\sqrt{18}} \|\Delta\|_{2(1)} \leq \|\Delta\|_{U_d} \leq \|\Delta\|_{2(1)},$$

where the norm  $\|\cdot\|_{2(1)}$  is defined as

$$(12) \quad \|\Delta\|_{2(1)} = \sqrt{\text{Tr}(\Delta^2) + (\text{Tr} \Delta)^2}.$$

One can check that  $\|\Delta\|_{2(1)}$  equals the  $L^2$  norm of the random variable  $\langle g | \Delta | g \rangle$ , where  $g$  is a standard Gaussian vector in  $\mathbf{C}^d$ , while the  $L^1$  norm of this random variable is nothing else than  $\|\Delta\|_{U_d}$ . Therefore Proposition 4.1 can be seen as a reverse Hölder inequality, and an interesting problem would be to find the optimal constant in that inequality (the factor  $\sqrt{18}$  is presumably far from optimal).

This dimension-free lower bound on the distinguishing power of the uniform POVM is of interest in quantum information theory. One could cite as one of its applications the possibility to establish lower-bounds on the dimensionality reduction of quantum states [15]. However, from a computational or algorithmic point of view, this statement involving a continuous POVM is of no practical use. There has been interest therefore in the question of sparsifying  $U_d$ , i.e. of finding a discrete POVM, with as few outcomes as possible, which would be equivalent to  $U_d$  in terms of discriminating efficiency. Examples of such constructions arise from the theory of projective 4-designs.

Given an integer  $t \geq 1$ , an (exact)  $t$ -design is a finitely supported probability measure  $\mu$  on  $S_{\mathbf{C}^d}$  such that

$$\int_{S_{\mathbf{C}^d}} |\psi\rangle\langle\psi|^{\otimes t} d\mu(\psi) = \int_{S_{\mathbf{C}^d}} |\psi\rangle\langle\psi|^{\otimes t} d\sigma(\psi) = \binom{d+t-1}{t}^{-1} P_{\text{Sym}^t(\mathbf{C}^d)}.$$

Here,  $\sigma$  denotes the Haar probability measure on  $S_{\mathbf{C}^d}$ , and  $P_{\text{Sym}^t(\mathbf{C}^d)}$  denotes the orthogonal projection onto the symmetric subspace  $\text{Sym}^t(\mathbf{C}^d) \subset (\mathbf{C}^d)^{\otimes t}$ .

Note that a  $t$ -design is also a  $t'$ -design for any  $t' \leq t$ . Let  $\mu$  be a 1-design. The map  $\psi \mapsto |\psi\rangle\langle\psi|$  pushes forward  $\mu$  into a measure  $\tilde{\mu}$  on the set of (pure) states, with barycenter equal to  $\text{Id}/d$ . By Proposition 2.5, this measure corresponds to a POVM, and in the following we identify  $t$ -designs with the associated POVMs. For example the uniform POVM  $U_d$  is a  $t$ -design for any  $t$ .

This notion can be relaxed: define an  $\varepsilon$ -approximate  $t$ -design to be a finitely supported measure  $\mu$  on  $S_{\mathbf{C}^d}$  such that

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

It has been proved in [1] that a 4-design (exact or approximate) supported on  $N$  points yields a POVM  $M$  with  $N$  outcomes such that

$$(13) \quad C^{-1} \|\cdot\|_{U_d} \leq \|\cdot\|_M \leq C \|\cdot\|_{U_d}$$



for some constant  $C$ . The proof is based on the fourth moment method, which is used to control the first absolute moment of a random variable by its second and fourth moments.

Now, what is the minimal cardinality of a 4-design? The support of any exact or  $\varepsilon$ -approximate (provided  $\varepsilon < 1$ ) 4-design must contain at least  $\dim(\text{Sym}^4(\mathbf{C}^d)) = \binom{d+3}{4} = \Omega(d^4)$  points. Conversely, an argument based on Carathéodory's theorem shows that there exist exact 4-designs with  $O(d^8)$  points. Starting from such an exact 4-design, the sparsification procedure from [4] gives a deterministic and efficient algorithm which outputs an  $\varepsilon$ -approximate 4-design supported by  $O(d^4/\varepsilon^2)$  points.

However, this approach has two drawbacks: the constant  $C$  from (13) cannot be taken close to 1, and the number of outcomes has to be  $\Omega(d^4)$ . We are going to remove both inconveniences in our Theorem 4.3.

**4.2. Euclidean subspaces.** How do these ideas translate into the framework of zonoids? The analogue of  $U_d$  is the most symmetric zonoid, namely the Euclidean ball  $B_2^n \subset \mathbf{R}^n$ . To connect with literature from functional analysis, it is worth emphasizing that approximating  $B_2^n$  by a zonotope which is the sum of  $N$  segments is equivalent to embedding the space  $\ell_2^n = (\mathbf{R}^n, \|\cdot\|_2)$  into the space  $\ell_1^N = (\mathbf{R}^N, \|\cdot\|_1)$ . Indeed, assume that  $x_1, \dots, x_N$  are points in  $\mathbf{R}^n$  such that, for some constants  $c, C$ ,

$$cZ \subset B_2^n \subset CZ,$$

where  $Z = \text{conv}\{\pm x_1\} + \dots + \text{conv}\{\pm x_N\}$ . Then the map  $u : \mathbf{R}^n \rightarrow \mathbf{R}^N$  defined by

$$u(x) = \left( \langle x, x_1 \rangle, \dots, \langle x, x_N \rangle \right)$$

satisfies  $c\|u(x)\|_1 \leq \|x\|_2 \leq C\|u(x)\|_1$  for any  $x \in \mathbf{R}^n$ . In this context, the ratio  $C/c$  is often called the distortion of the embedding.

An early result by Rudin [25] shows an explicit embedding of  $\ell_2^n$  into  $\ell_1^{O(n^2)}$  with distortion  $\sqrt{3}$ . This is proved by the fourth moment method and can be seen as the analogue of the constructions based on 4-designs. The following theorem (a variation on Dvoretzky's theorem) has been a major improvement on Rudin's result, showing that  $\ell_1^N$  has almost Euclidean sections of proportional dimension.

**Theorem 4.2** ([11]). *For every  $0 < \varepsilon < 1$ , there exists a subspace  $E \subset \mathbf{R}^N$  of dimension  $n = c(\varepsilon)N$  such that for any  $x \in E$ ,*

$$(14) \quad (1 - \varepsilon)M\|x\|_2 \leq \|x\|_1 \leq (1 + \varepsilon)M\|x\|_2,$$

where  $M$  denotes the average of the 1-norm over the Euclidean unit sphere  $S^{N-1}$ .

Theorem 4.2 was first proved in [11], making a seminal use of measure concentration in the form of Lévy's lemma. The argument shows that a generic subspace  $E$  (i.e. picked uniformly at random amongst all  $c(\varepsilon)N$ -dimensional subspaces of  $\mathbf{R}^N$ ) satisfies the conclusion of the theorem with high probability for  $c(\varepsilon) = O(\varepsilon^2 |\log \varepsilon|^{-1})$ . This was later improved in [13] to  $c(\varepsilon) = O(\varepsilon^2)$ .

**4.3. Sparsification of the uniform POVM.** Translated in the language of zonotopes, Theorem 4.2 states that the sum of  $O(n)$  randomly chosen segments in  $\mathbf{R}^n$  is close to the Euclidean ball  $B_2^n$ . More precisely, for any  $0 < \varepsilon < 1$ , the zonotope  $Z = \text{conv}\{\pm x_1\} + \dots + \text{conv}\{\pm x_N\}$ , with  $N = c(\varepsilon)^{-1}n$  and  $x_1, \dots, x_N$  randomly chosen points in  $\mathbf{R}^n$ , is  $\varepsilon$ -close to the Euclidean ball  $B_2^n$ , in the sense that  $(1 - \varepsilon)Z \subset B_2^n \subset (1 + \varepsilon)Z$ .

By analogy, we expect a POVM constructed from  $O(d^2)$  randomly chosen elements to be close to the uniform POVM. This random construction can be achieved as follows: let  $(|\psi_i\rangle)_{1 \leq i \leq n}$  be independent random vectors, uniformly chosen on the unit sphere of  $\mathbf{C}^d$ . Set  $P_i = |\psi_i\rangle\langle\psi_i|$ ,  $1 \leq i \leq n$ , and  $S = P_1 + \dots + P_n$ . When  $n \geq d$ ,  $S$  is almost surely invertible, and we may consider the random POVM

$$(15) \quad M = (S^{-1/2} P_i S^{-1/2})_{1 \leq i \leq n}.$$

**Theorem 4.3.** *Let  $M$  be a random POVM on  $\mathbf{C}^d$  with  $n$  outcomes, defined as in (15), and let  $0 < \varepsilon < 1$ . If  $n \geq C\varepsilon^{-2} |\log \varepsilon| d^2$ , then with high probability the POVM  $M$  satisfies the inequalities*

$$(1 - \varepsilon)\|\Delta\|_{U_d} \leq \|\Delta\|_M \leq (1 + \varepsilon)\|\Delta\|_{U_d}$$

for every  $\Delta \in \mathcal{H}(\mathbf{C}^d)$ .

By "with high probability" we mean that the probability that the conclusion fails is less than  $\exp(-c(\varepsilon)d)$  for some constant  $c(\varepsilon)$ . Theorem 4.3 is proved in Section 6, the proof being based on a careful use of  $\varepsilon$ -nets and deviation inequalities. It does not seem possible to deduce formally Theorem 4.3 from the existing Banach space literature.

Theorem 4.3 shows that the uniform POVM on  $\mathbf{C}^d$  can be  $\varepsilon$ -approximated (in the sense of closeness of distinguishability norms) by a POVM with  $n = O(\varepsilon^{-2} |\log \varepsilon| d^2)$  outcomes. Note that the dependence of  $n$  with respect

to  $d$  is optimal: since a POVM on  $\mathbf{C}^d$  must have at least  $d^2$  outcomes to be informationally complete, one cannot hope for a tighter dimensional dependence. The dependence with respect to  $\varepsilon$  is less clear: the factor  $|\log \varepsilon|$  can probably be removed but we do not pursue this direction.

Our construction is random and a natural question is whether deterministic constructions yielding comparable properties exist. A lot of effort has been put in derandomizing Theorem 4.2. We refer to [19] for bibliography and mention two of the latest results. Given any  $0 < \gamma < 1$ , it is shown in [19] how to construct, from  $cn^\gamma$  random bits (i.e. an amount of randomness sub-linear in  $n$ ) a subspace of  $\ell_1^N$  satisfying (14) with  $N \leq (\gamma\varepsilon)^{-C\gamma}n$ . A completely explicit construction appears in [18], with  $N \leq n2^{C(\varepsilon)(\log \log n)^2} = n^{1+C(\varepsilon)o(n)}$ . It is not obvious how to adapt these constructions to obtain sparsifications of the uniform POVM using few or no randomness.

**4.4. Sparsification of any POVM.** Theorem 4.2 initiated intensive research in the late 80's [28, 7, 31] on the theme of ‘‘approximation of zonoids by zonotopes’’, trying to extend the result for the Euclidean ball (the most symmetric zonoid) to an arbitrary zonoid. This culminated in Talagrand’s proof [31] that for any zonoid  $Y \subset \mathbf{R}^n$  and any  $0 < \varepsilon < 1$ , there exists a zonotope  $Z \subset \mathbf{R}^n$  which is the sum of  $O(\varepsilon^{-2}n \log n)$  segments and such that  $(1 - \varepsilon)Y \subset Z \subset (1 + \varepsilon)Y$ . A more precise version is stated in Section 8. Whether the  $\log n$  factor can be removed is still an open problem.

This result easily implies a similar result for POVMs, provided we consider the larger class of sub-POVMs. A discrete sub-POVM with  $n$  outcomes is a finite family  $M = (M_i)_{1 \leq i \leq n}$  of  $n$  positive operators such that  $S = \sum_{i=1}^n M_i \leq \text{Id}$ . As for POVMs, the norm associated to a sub-POVM  $M$  is defined for  $\Delta \in \mathcal{H}(\mathbf{C}^d)$  by

$$\|\Delta\|_M = \sum_{i=1}^n |\text{Tr}(\Delta M_i)|.$$

We prove the following result in Section 8.

**Theorem 4.4.** *Given any POVM  $M$  on  $\mathbf{C}^d$  and any  $0 < \varepsilon < 1$ , there is a sub-POVM  $M' = (M'_i)_{1 \leq i \leq n}$ , with  $n \leq C\varepsilon^{-2}d^2 \log(d)$  such that, for any  $\Delta \in \mathcal{H}(\mathbf{C}^d)$ ,*

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

Moreover, we can guarantee that the states  $M'_i / \text{Tr}(M'_i)$  belong to the support of the measure  $\mu$  associated to  $M$ .

We do not know whether Theorem 4.4 still holds if we want  $M'$  to be a POVM. Given a sub-POVM  $(M_i)_{1 \leq i \leq n}$ , there are at least two natural ways to modify it into a POVM. A solution is to add an extra outcome corresponding to the operator  $\text{Id} - S$ , and another one is to substitute  $S^{-1/2}M_i S^{-1/2}$  in place of  $M_i$ , as we proceeded in (15). However for a general POVM, the error terms arising from this renormalization step may exceed the quantity to be approximated.

## 5. SPARSIFYING LOCAL POVMs

Proposition 5.1 below is an immediate corollary of Lemma 3.4 and Proposition 3.7. In words, it shows that, on a multipartite system, a local POVM can be sparsified by tensorizing sparsifications of each of its factors.

**Proposition 5.1.** *Let  $0 < \varepsilon < 1$ . Let  $M_1, \dots, M_k$  be POVMs and  $M'_1, \dots, M'_k$  be (sub-)POVMs, on  $\mathbf{C}^{d_1}, \dots, \mathbf{C}^{d_k}$  respectively, satisfying, for all  $1 \leq i \leq k$ , and for all  $\Delta \in \mathcal{H}(\mathbf{C}^{d_i})$ ,*

$$(1 - \varepsilon)\|\Delta\|_{M_i} \leq \|\Delta\|_{M'_i} \leq (1 + \varepsilon)\|\Delta\|_{M_i}.$$

Then, for any  $\Delta \in \mathcal{H}(\mathbf{C}^{d_1} \otimes \dots \otimes \mathbf{C}^{d_k})$ ,

$$(1 - \varepsilon)^k \|\Delta\|_{M_1 \otimes \dots \otimes M_k} \leq \|\Delta\|_{M'_1 \otimes \dots \otimes M'_k} \leq (1 + \varepsilon)^k \|\Delta\|_{M_1 \otimes \dots \otimes M_k}.$$

Let us give a concrete application of Proposition 5.1. We consider  $k$  finite-dimensional Hilbert spaces  $\mathbf{C}^{d_1}, \dots, \mathbf{C}^{d_k}$  and define the local uniform POVM on the  $k$ -partite Hilbert space  $\mathbf{C}^{d_1} \otimes \dots \otimes \mathbf{C}^{d_k}$  as the tensor product of the  $k$  uniform POVMs  $U_{d_1}, \dots, U_{d_k}$ . We will denote it by LU. The corresponding distinguishability norm can be described, for any  $\Delta \in \mathcal{H}(\mathbf{C}^{d_1} \otimes \dots \otimes \mathbf{C}^{d_k})$ , as

$$\|\Delta\|_{\text{LU}} = d \mathbf{E} |\langle \psi_1 \otimes \dots \otimes \psi_k | \Delta | \psi_1 \otimes \dots \otimes \psi_k \rangle|,$$

where  $d = d_1 \times \dots \times d_k$  is the dimension of the global Hilbert space, and where the random unit vectors  $\psi_1, \dots, \psi_k$  are independent and Haar-distributed in  $\mathbf{C}^{d_1}, \dots, \mathbf{C}^{d_k}$  respectively.

The following multipartite generalization of Proposition 4.1 shows that the norm  $\|\cdot\|_{\text{LU}}$ , in analogy to the norm  $\|\cdot\|_U$ , is equivalent to a ‘‘modified’’ Hilbert–Schmidt norm.

**Proposition 5.2** ([20]). *For every  $\Delta \in \mathcal{H}(\mathbf{C}^{d_1} \otimes \cdots \otimes \mathbf{C}^{d_k})$ , we have*

$$(16) \quad \frac{1}{18^{k/2}} \|\Delta\|_{2(k)} \leq \|\Delta\|_{\text{LU}} \leq \|\Delta\|_{2(k)},$$

where the norm  $\|\cdot\|_{2(k)}$  is defined as

$$(17) \quad \|\Delta\|_{2(k)} = \sqrt{\sum_{I \subset \{1, \dots, k\}} \text{Tr} \left[ (\text{Tr}_I \Delta)^2 \right]}.$$

Here  $\text{Tr}_I$  denotes the partial trace over all parties  $I \subset \{1, \dots, k\}$ .

*Proof of Proposition 5.2.* A direct proof appears in [20], but we find interesting to show that it can be deduced (with a worst constant) from Proposition 4.1. If we denote by  $\langle \cdot, \cdot \rangle_H$  the inner product inducing a Euclidean norm  $\|\cdot\|_H$ , we have

$$\langle A_1 \otimes \cdots \otimes A_k, B_1 \otimes \cdots \otimes B_k \rangle_{2(k)} = \langle A_1, B_1 \rangle_{2(1)} \times \cdots \times \langle A_k, B_k \rangle_{2(1)}$$

which is equivalent to saying that

$$\|\cdot\|_{2(k)} = \|\cdot\|_{2(1)} \otimes^2 \cdots \otimes^2 \|\cdot\|_{2(1)}.$$

We thus get by Proposition 3.5,

$$c_0^{k-1} \|\cdot\|_{2(k)} \leq \|\cdot\|_{2(1)} \otimes^1 \cdots \otimes^1 \|\cdot\|_{2(1)} \leq \|\cdot\|_{2(k)}$$

with  $c_0 = \sqrt{2/\pi}$ . Now, we also know by Proposition 3.7 that on  $\mathcal{H}(\mathbf{C}^{d_1} \otimes \cdots \otimes \mathbf{C}^{d_k})$ ,  $\|\cdot\|_{\text{LU}} = \|\cdot\|_{U_{d_1}} \otimes^1 \cdots \otimes^1 \|\cdot\|_{U_{d_k}}$ , and by Proposition 4.1 that  $c\|\cdot\|_{2(1)} \leq \|\cdot\|_{U_d} \leq \|\cdot\|_{2(1)}$  for some constant  $c$  ( $c = 1/\sqrt{18}$  works). So by Lemma 3.4,

$$c^k \|\cdot\|_{2(1)} \otimes^1 \cdots \otimes^1 \|\cdot\|_{2(1)} \leq \|\cdot\|_{\text{LU}} \leq \|\cdot\|_{2(1)} \otimes^1 \cdots \otimes^1 \|\cdot\|_{2(1)},$$

and therefore

$$c_0^{k-1} c^k \|\cdot\|_{2(k)} \leq \|\cdot\|_{\text{LU}} \leq \|\cdot\|_{2(k)}. \quad \square$$

Remarkably, local dimensions do not appear in equation (16). This striking fact that local POVMs can have asymptotically non-vanishing distinguishing power can be used to construct an algorithm that solves the Weak Membership Problem for separability in quasi-polynomial time (see [8] for a description in the bipartite case). Hence the importance of being able to sparsify the local uniform POVM by a POVM for which the locality property is preserved and which has a number of outcomes that optimally scales as the square of the global dimension. We state the corresponding multipartite version of Theorem 4.3, which is straightforwardly obtained by combining the unipartite version with Proposition 5.1.

**Theorem 5.3.** *Let  $0 < \varepsilon < 1$ . For all  $1 \leq i \leq k$ , let  $M_i$  be a random POVM on  $\mathbf{C}^{d_i}$  with  $n_i \geq C\varepsilon^{-2} |\log \varepsilon| d_i^2$  outcomes, defined as in (15). Then, with high probability, the local POVM  $M_1 \otimes \cdots \otimes M_k$  on  $\mathbf{C}^{d_1} \otimes \cdots \otimes \mathbf{C}^{d_k}$  is such that, for any  $\Delta \in \mathcal{H}(\mathbf{C}^{d_1} \otimes \cdots \otimes \mathbf{C}^{d_k})$ ,*

$$(1 - \varepsilon)^k \|\Delta\|_{\text{LU}} \leq \|\Delta\|_{M_1 \otimes \cdots \otimes M_k} \leq (1 + \varepsilon)^k \|\Delta\|_{\text{LU}}.$$

Let us rephrase the content of Theorem 5.3: the local uniform POVM on  $\mathbf{C}^{d_1} \otimes \cdots \otimes \mathbf{C}^{d_k}$  can be  $k\varepsilon$ -approximated (in terms of distinguishability norms) by a POVM which is also local and has a total number of outcomes  $n = O(C^k \varepsilon^{-2k} |\log \varepsilon|^k d^2)$ , where  $d = d_1 \times \cdots \times d_k$ . Note that the dimensional dependence of  $n$  is optimal. On the contrary, the dependence of  $n$  on  $\varepsilon$  deteriorates as  $k$  grows. The high-dimensional situation our result applies to is thus really the one of a “small” number of “large” subsystems (i.e.  $k$  fixed and  $d_1, \dots, d_k \rightarrow +\infty$ ), and not of a “large” number of “small” subsystems.

## 6. PROOF OF THEOREM 4.3

In this section we prove Theorem 4.3. Let  $n \in \mathbf{N}$  and  $(|\psi_i\rangle)_{1 \leq i \leq n}$  be independent random unit vectors, uniformly distributed on the unit sphere of  $\mathbf{C}^d$ . Our main technical estimates are a couple of probabilistic inequalities. Proposition 6.1 is an immediate consequence of Theorem 1 in [2]. Proposition 6.2 is a consequence of Bernstein inequalities. However, its proof requires some careful estimates which we postpone to Section 7.

**Proposition 6.1.** *If  $(|\psi_i\rangle)_{1 \leq i \leq n}$  are independent random vectors, uniformly distributed on the unit sphere of  $\mathbf{C}^d$ , then for every  $0 < \eta < 1$*

$$\mathbf{P} \left( (1 - \eta) \frac{\text{Id}}{d} \leq \frac{1}{n} \sum_{i=1}^n |\psi_i\rangle\langle\psi_i| \leq (1 + \eta) \frac{\text{Id}}{d} \right) \geq 1 - C^d \exp(-cn\eta^2).$$

**Proposition 6.2.** *Let  $\Delta \in \mathcal{H}(\mathbf{C}^d)$ , and  $(|\psi_i\rangle)_{1 \leq i \leq n}$  be independent random vectors, uniformly distributed on the unit sphere of  $\mathbf{C}^d$ . For  $1 \leq i \leq n$ , consider the random variables  $X_i = d|\langle \psi_i | \Delta | \psi_i \rangle|$  and  $Y_i = X_i - \mathbf{E} X_i = X_i - \|\Delta\|_{\mathcal{U}_d}$ . Then, for any  $t > 0$ ,*

$$\mathbf{P} \left( \left| \frac{1}{n} \sum_{i=1}^n Y_i \right| \geq t \|\Delta\|_{\mathcal{U}_d} \right) \leq 2 \exp(-c'_0 n \min(t, t^2)).$$

We now show how to derive Theorem 4.3 from the estimates in Propositions 6.1 and 6.2. For each  $1 \leq i \leq n$ , set  $P_i = |\psi_i\rangle\langle \psi_i|$ , and introduce the (random) norm defined for any  $\Delta \in \mathcal{H}(\mathbf{C}^d)$  as

$$\|\|\Delta\|\| = \frac{d}{n} \sum_{i=1}^n |\mathrm{Tr}(\Delta P_i)|.$$

We will now prove that  $\|\|\cdot\|\|$  is, with probability close to 1, a good approximation to  $\|\cdot\|_{\mathcal{U}_d}$ . First, using Proposition 6.2, we obtain that for any  $0 < \varepsilon < 1$  and any  $\Delta \in \mathcal{H}(\mathbf{C}^d)$

$$(18) \quad \mathbf{P}((1 - \varepsilon)\|\Delta\|_{\mathcal{U}_d} \leq \|\|\Delta\|\| \leq (1 + \varepsilon)\|\Delta\|_{\mathcal{U}_d}) \geq 1 - 2 \exp(-c'_0 n \varepsilon^2).$$

We next use a net argument. Fix  $0 < \varepsilon < 1/3$  and a  $\varepsilon$ -net  $\mathcal{N}$  inside the unit ball for the norm  $\|\cdot\|_{\mathcal{U}_d}$ , with respect to the distance induced by  $\|\cdot\|_{\mathcal{U}_d}$ . A standard volumetric argument (see [23], Lemma 4.10) shows that we may assume  $\mathrm{card}(\mathcal{N}) \leq (1 + 2/\varepsilon)^{d^2} \leq (3/\varepsilon)^{d^2}$ . Introduce the quantities

$$A := \sup\{\|\|\Delta\|\| : \|\Delta\|_{\mathcal{U}_d} \leq 1\},$$

$$A' := \sup\{\|\|\Delta\|\| : \Delta \in \mathcal{N}\}.$$

Given  $\Delta$  such that  $\|\Delta\|_{\mathcal{U}_d} \leq 1$ , there is  $\Delta_0 \in \mathcal{N}$  with  $\|\Delta - \Delta_0\|_{\mathcal{U}_d} \leq \varepsilon$ . By the triangle inequality, we have  $\|\|\Delta\|\| \leq A' + \|\|\Delta - \Delta_0\|\| \leq A' + \varepsilon A$ . Taking supremum over  $\Delta$  yields  $A \leq A' + \varepsilon A$  i.e.  $A \leq \frac{A'}{1 - \varepsilon}$ .

If we introduce  $B := \inf\{\|\|\Delta\|\| : \|\Delta\|_{\mathcal{U}_d} = 1\}$  and  $B' := \inf\{\|\|\Delta\|\| : \Delta \in \mathcal{N}\}$ , a similar argument shows that  $B \geq B' - \varepsilon A$ , so that in fact  $B \geq B' - \frac{\varepsilon A}{1 - \varepsilon}$ . We therefore have the implications

$$(19) \quad 1 - \varepsilon \leq B' \leq A' \leq 1 + \varepsilon \implies 1 - \varepsilon - \frac{\varepsilon(1 + \varepsilon)}{1 - \varepsilon} \leq B \leq A \leq \frac{1 + \varepsilon}{1 - \varepsilon} \implies 1 - 3\varepsilon \leq B \leq A \leq 1 + 3\varepsilon.$$

By the union bound, we get from (18) that  $\mathbf{P}(1 - \varepsilon \leq B' \leq A' \leq 1 + \varepsilon) \geq 1 - 2 \mathrm{card}(\mathcal{N}) \exp(-c'_0 n \varepsilon^2)$ . Combined with (19), and using homogeneity of norms, this yields

$$(20) \quad \mathbf{P}\left((1 - 3\varepsilon)\|\cdot\|_{\mathcal{U}_d} \leq \|\|\cdot\|\| \leq (1 + 3\varepsilon)\|\cdot\|_{\mathcal{U}_d}\right) \geq 1 - 2 \left(\frac{3}{\varepsilon}\right)^{d^2} \exp(-c'_0 n \varepsilon^2).$$

This probability estimate is non-trivial, and can be made close to 1, provided  $n \gtrsim d^2 \varepsilon^{-2} |\log \varepsilon|$ .

Whenever  $n \geq d$ , the vectors  $(|\psi_i\rangle)_{1 \leq i \leq n}$  generically span  $\mathbf{C}^d$ , and therefore the operator  $S = P_1 + \dots + P_n$  is invertible. We may then define  $\tilde{P}_i = S^{-1/2} P_i S^{-1/2}$  so that  $\mathbf{M} = (\tilde{P}_i)_{1 \leq i \leq n}$  is a POVM. The norm associated to  $\mathbf{M}$  is, for any  $\Delta \in \mathcal{H}(\mathbf{C}^d)$ ,

$$\|\Delta\|_{\mathbf{M}} = \sum_{i=1}^n |\mathrm{Tr}(\Delta \tilde{P}_i)|.$$

We now argue that the norms  $\|\|\cdot\|\|$  and  $\|\cdot\|_{\mathbf{M}}$  are similar enough (modulo normalization), because the modified operators  $\tilde{P}_i$  are close enough to the initial ones  $P_i$ . This is achieved by showing that  $T := \left(\frac{d}{n} S\right)^{-1/2}$  is close to  $\mathrm{Id}$  (in operator-norm distance). We use Proposition 6.1 for  $\eta = \varepsilon \|\Delta\|_{\mathcal{U}_d} / \|\Delta\|_1$ . By Proposition 4.1, we have  $\eta \geq \varepsilon / \sqrt{18d}$ . Proposition 6.1 implies that

$$(21) \quad \mathbf{P}(\|T - \mathrm{Id}\|_{\infty} \geq \eta) \leq \mathbf{P}(\|T^{-2} - \mathrm{Id}\|_{\infty} \geq \eta) \leq C^d \exp(-c' n \varepsilon^2 / d).$$

This upper bound is much smaller than 1 provided  $n \geq C_1 \varepsilon^{-2} d^2$ . Also, note that the event  $\|T - \mathrm{Id}\|_{\infty} \leq \eta$  implies that

$$\|\Delta - T \Delta T\|_{\mathbf{M}} \leq \|\Delta - T \Delta T\|_1 \leq \|\Delta\|_1 \|\mathrm{Id} - T\|_{\infty} (1 + \|T\|_{\infty}) \leq 2\eta \|\Delta\|_1 = 2\varepsilon \|\Delta\|_{\mathcal{U}_d}.$$

Using the cyclic property of the trace, we check that  $\|T \Delta T\|_{\mathbf{M}} = \|\|\Delta\|\|$ . Now, choose  $n$  larger than both  $C_0 \varepsilon^{-2} |\log \varepsilon| d^2$  and  $C_1 \varepsilon^{-2} d^2$ . With high probability, the events from equations (20) and (21) both hold. We then obtain for every  $\Delta \in \mathcal{H}(\mathbf{C}^d)$ ,

$$\|\Delta\|_{\mathbf{M}} \leq \|T \Delta T\|_{\mathbf{M}} + \|\Delta - T \Delta T\|_{\mathbf{M}} \leq \|\|\Delta\|\| + 2\varepsilon \|\Delta\|_{\mathcal{U}_d} \leq (1 + 5\varepsilon) \|\Delta\|_{\mathcal{U}_d}$$

and similarly  $\|\Delta\|_{\mathbf{M}} \geq (1 - 5\varepsilon) \|\Delta\|_{\mathcal{U}_d}$ . This is precisely the result from Theorem 4.3 with  $5\varepsilon$  instead of  $\varepsilon$ , which of course can be absorbed by renaming the constants appropriately.

## 7. PROOF OF PROPOSITION 6.2

The proof is a direct application of a large deviation inequality for sums of independent sub-exponential (or  $\psi_1$ ) random variables. Recall that the  $\psi_1$ -norm of a random variable  $X$  (which quantifies the exponential decay of the tail) may be defined via the growth of even moments

$$\|X\|_{\psi_1} := \sup_{q \in \mathbf{N}} \frac{1}{2q} (\mathbf{E} |X|^{2q})^{1/2q}.$$

This definition is more practical than the standard definition through the Orlicz function  $x \mapsto \exp(x) - 1$ , and leads to an equivalent norm (see [9], Corollary 1.1.6). The large deviation inequality for a sum of independent  $\psi_1$  random variables is known as Bernstein's inequality.

**Theorem 7.1** (Bernstein's inequality, see [9], Theorem 1.2.5.). *Let  $X_1, \dots, X_n$  be  $n$  independent  $\psi_1$  random variables with mean zero. Setting  $M = \max_{1 \leq i \leq n} \|X_i\|_{\psi_1}$  and  $\sigma^2 = \frac{1}{n} \sum_{1 \leq i \leq n} \|X_i\|_{\psi_1}^2$ , we have*

$$\forall t > 0, \mathbf{P} \left( \left| \frac{1}{n} \sum_{i=1}^n X_i \right| \geq t \right) \leq 2 \exp \left( -c_0 n \min \left( \frac{t^2}{\sigma^2}, \frac{t}{M} \right) \right),$$

$c_0 > 0$  being a universal constant.

For  $\Delta \in \mathcal{H}(\mathbf{C}^d)$ , consider the random variables  $X_i = d |\operatorname{Tr}(\Delta P_i)|$  with  $P_i = |\psi_i\rangle\langle\psi_i|$ , and  $Y_i = X_i - \mathbf{E} X_i = d |\operatorname{Tr}(\Delta P_i)| - \|\Delta\|_{\mathcal{U}_d}$ . The random variables  $Y_i$  are independent and have mean zero. The key lemma is a bound on their  $\psi_1$  norm.

**Lemma 7.2.** *Let  $\Delta \in \mathcal{H}(\mathbf{C}^d)$  and consider the random variable  $X := d |\operatorname{Tr}(\Delta P)|$ , where  $P = |\psi\rangle\langle\psi|$ , with  $\psi$  uniformly distributed on the unit sphere of  $\mathbf{C}^d$ . Then  $\|X\|_{\psi_1} \leq \|\Delta\|_{2(1)}$  and  $\|X - \mathbf{E} X\|_{\psi_1} \leq 3\|\Delta\|_{2(1)} \leq 3\sqrt{18}\|\Delta\|_{\mathcal{U}_d}$ .*

Therefore, we may apply Bernstein's inequality with  $M = \sigma \leq 3\sqrt{18}\|\Delta\|_{\mathcal{U}_d}$ , yielding Proposition 6.2.

*Proof of Lemma 7.2.* For each integer  $q$ , we compute

$$\mathbf{E} [\operatorname{Tr}(\Delta P)]^{2q} = \mathbf{E} \operatorname{Tr} (\Delta^{\otimes 2q} P^{\otimes 2q}) = \operatorname{Tr} (\Delta^{\otimes 2q} [\mathbf{E} P^{\otimes 2q}]).$$

We use the fact (see e.g. [14]) that

$$\mathbf{E} P^{\otimes 2q} = \frac{(2q)!}{(d+2q-1) \times \dots \times d} P_{\operatorname{Sym}^{2q}(\mathbf{C}^d)} = \frac{1}{(d+2q-1) \times \dots \times d} \sum_{\pi \in \mathfrak{S}_{2q}} U(\pi),$$

where  $P_{\operatorname{Sym}^{2q}(\mathbf{C}^d)}$  denotes the orthogonal projection onto the symmetric subspace  $\operatorname{Sym}^{2q}(\mathbf{C}^d) \subset (\mathbf{C}^d)^{\otimes 2q}$ , and for each permutation  $\pi \in \mathfrak{S}_{2q}$ ,  $U(\pi)$  denotes the associated permutation unitary on  $(\mathbf{C}^d)^{\otimes 2q}$ . This yields

$$\mathbf{E} [\operatorname{Tr}(\Delta P)]^{2q} = \frac{1}{(d+2q-1) \times \dots \times d} \sum_{\pi \in \mathfrak{S}_{2q}} \operatorname{Tr} (\Delta^{\otimes 2q} U(\pi)).$$

If  $\ell_1, \dots, \ell_k$  denote the lengths of the cycles appearing in the cycle decomposition of a permutation  $\pi \in \mathfrak{S}_{2q}$ , we have  $\ell_1 + \dots + \ell_k = 2q$  and

$$\operatorname{Tr} (\Delta^{\otimes 2q} U(\pi)) = \prod_{i=1}^k \operatorname{Tr} (\Delta^{\ell_i}).$$

Now, for any integer  $\ell \geq 2$ , we have  $|\operatorname{Tr}(\Delta^\ell)| \leq [\operatorname{Tr}(\Delta^2)]^{\ell/2} \leq \|\Delta\|_{2(1)}^\ell$ . The inequality  $|\operatorname{Tr}(\Delta^\ell)| \leq \|\Delta\|_{2(1)}^\ell$  is also (trivially) true for  $\ell = 1$ . Therefore  $|\operatorname{Tr}(\Delta^{\otimes 2q} U(\pi))| \leq \|\Delta\|_{2(1)}^{2q}$ . It follows that

$$\mathbf{E} [\operatorname{Tr}(\Delta P)]^{2q} \leq \frac{(2q)!}{d^{2q}} \|\Delta\|_{2(1)}^{2q} \leq \left( \frac{2q \|\Delta\|_{2(1)}}{d} \right)^{2q},$$

so that  $(\mathbf{E} X^{2q})^{1/2q} \leq 2q \|\Delta\|_{2(1)}$ , and thus  $\|X\|_{\psi_1} \leq \|\Delta\|_{2(1)}$ . The last part of the Lemma follows from the triangle inequality, since  $\|\mathbf{E} X\|_{\psi_1} = |\mathbf{E} X| \leq 2\|X\|_{\psi_1}$ , and from the equivalence (11) between the norms  $\|\cdot\|_{\mathcal{U}_d}$  and  $\|\cdot\|_{2(1)}$ .  $\square$

## 8. PROOF OF THEOREM 4.4

Here is a version of Talagrand’s theorem which is suitable for our purposes.

**Theorem 8.1** ([31]). *Let  $Z \subset \mathbf{R}^n$  be a symmetric zonotope, with*

$$Z = \sum_{i \in I} \text{conv}\{\pm u_i\}$$

*for a finite family of vectors  $(u_i)_{i \in I}$ . Then for every  $\varepsilon > 0$  there exists a subset  $J \subset I$  with  $\text{card } J \leq Cn \log n / \varepsilon^2$ , and positive numbers  $(\lambda_i)_{i \in J}$  such that the zonotope*

$$Z' = \sum_{i \in J} \text{conv}\{\pm \lambda_i u_i\}$$

*satisfies  $Z' \subset Z \subset (1 + \varepsilon)Z'$ .*

Theorem 4.4 is a very simple consequence of Theorem 8.1. Let  $M$  be a POVM to be sparsified. Using Corollary 2.2, we may assume that  $M = (M_i)_{i \in I}$  is discrete. Applying Theorem 8.1 to the zonotope  $K_M = \sum_{i \in I} \text{conv}\{\pm M_i\}$  (which lives in a  $d^2$ -dimensional space), we obtain a zonotope  $Z' = \sum_{i \in J} \text{conv}\{\pm \lambda_i M_i\}$  with  $\text{card } J \leq Cd^2 \log d / \varepsilon^2$  such that  $Z' \subset K_M \subset (1 + \varepsilon)Z'$ . It remains to show that  $M' = (\lambda_i M_i)_{i \in J}$  is a sub-POVM. We know that  $h_{Z'} \leq h_{K_M}$ . Therefore, given a unit vector  $x \in \mathbf{C}^d$ , the inequality  $h_{Z'}(\Delta) \leq h_{K_M}(\Delta)$  applied with  $\Delta = |x\rangle\langle x|$  shows that

$$\sum_{i \in J} \lambda_i |\langle x | M_i | x \rangle| \leq \| |x\rangle\langle x| \|_M \leq \| |x\rangle\langle x| \|_1 = 1,$$

and therefore  $\sum_{i \in J} \lambda_i M_i \leq \text{Id}$ , as required. Since the inclusions  $Z' \subset K_M \subset (1 + \varepsilon)Z'$  are equivalent to the inequalities  $\| \cdot \|_{M'} \leq \| \cdot \|_M \leq (1 + \varepsilon) \| \cdot \|_{M'}$ , Theorem 4.4 follows.

## ACKNOWLEDGEMENTS

We thank Andreas Winter for having first raised the general question of finding POVMs with few outcomes but good discriminating power. We also thank Marius Junge for suggesting the possible connection between POVMs and zonoids, and for pointing out to us relevant literature.

## REFERENCES

- [1] **A. Ambainis, J. Emerson**, “Quantum t-designs: t-wise independence in the quantum world”, Proc. 22nd IEEE Conference on Computational Complexity, 129–140, Piscataway, NJ (2007); arXiv:quant-ph/0701126.
- [2] **G. Aubrun**, “On almost randomizing channels with a short Kraus decomposition”, Commun. Math. Phys. 288(3), 1103–1116 (2009); arXiv:0805.2900.
- [3] **A. Barvinok**, *A course in convexity*, Vol. 54. American Mathematical Soc., 2002.
- [4] **J. Batson, D.A. Spielman, N. Srivatsava**, “Twice-Ramanujan sparsifiers”; arXiv:0808.0163.
- [5] **G. Bennett**, “Schur multipliers”, Duke Math. J. 44.3, 603–639 (1977).
- [6] **E.D. Bolker**, “A class of convex bodies”, Trans. AMS 145, 323–345 (1969).
- [7] **J. Bourgain, J. Lindenstrauss, V. Milman**, “Approximation of zonoids by zonotopes”, Acta Mathematica 162.1, 73–141 (1989).
- [8] **F.G.S.L. Brandão, M. Christandl, J.T. Yard**, “Faithful Squashed Entanglement”, Commun. Math. Phys. 306, 805–830 (2011); arXiv[quant-ph]:1010.1750.
- [9] **D. Chafaï, O. Guédon, G. Lecué, A. Pajor**, *Interactions between compressed sensing, random matrices and high dimensional geometry*.
- [10] **T. Figiel, W. B. Johnson**, “Large subspaces of  $\ell_\infty^n$  and estimates of the Gordon–Lewis constant”, Israel J. Math. 37.1-2, 92–112 (1980).
- [11] **T. Figiel, J. Lindenstrauss, V.D. Milman**, “The dimension of almost spherical sections of convex bodies”, Acta Mathematica 139.1-2, 53–94 (1977).
- [12] **P. Goodey, W. Weil**, “Zonoids and Generalizations”, Handbook of Convex Geometry, Vol. B, 1296–1326, North-Holland, Amsterdam (1993).
- [13] **Y. Gordon**, “Some inequalities for Gaussian processes and applications”, Israel J. Math. 50.4, 265–289 (1985).
- [14] **A.W. Harrow**, “The Church of the Symmetric Subspace”; arXiv[quant-ph]:1308.6595.
- [15] **A.W. Harrow, A. Montanaro, A.J. Short**, “Limitations on quantum dimensionality reduction”, Proceedings of ICALP’11 LNCS 6755, 86–97, Springer-Verlag, Berlin Heidelberg (2011); arXiv[quant-ph]:1012.2262.
- [16] **C.W. Helstrom**, *Quantum detection and estimation theory*, Academic Press, New York (1976).
- [17] **A.S. Holevo**, “Statistical decision theory for quantum systems”, J. Mult. Anal. 3, 337–394 (1973).
- [18] **P. Indyk**, “Uncertainty Principles, Extractors, and Explicit Embeddings of  $L_2$  into  $L_1$ ”, 39th ACM Symposium on Theory of Computing (2007).
- [19] **P. Indyk, S. Szarek**, “Almost-Euclidean subspaces of  $l_1^N$  via tensor products: a simple approach to randomness reduction”, RANDOM 2010, LNCS 6302, 632–641, Springer-Verlag, Berlin Heidelberg (2010); arXiv[math.MG]1001.0041.
- [20] **C. Lancien, A. Winter**, “Distinguishing multi-partite states by local measurements”, Commun. Math. Phys. 323, 555–573 (2013); arXiv[quant-ph]:1206.2884.

- [21] **S. Lovett, S. Sodin**, “Almost Euclidean sections of the  $N$ -dimensional cross-polytope using  $O(N)$  random bits”, *Commun. Contemp. Math.* 10.4, 477–489 (2008); arXiv:math/0701102.
- [22] **W. Matthews, S. Wehner, A. Winter**, “Distinguishability of quantum states under restricted families of measurements with an application to data hiding”, *Comm. Math. Phys.* 291(3) (2009); arXiv:0810.2327[quant-ph].
- [23] **G. Pisier**, *The Volume of Convex Bodies and Banach Spaces Geometry*, Cambridge Tracts in Mathematics Volume 94, Cambridge University Press, Cambridge (1989).
- [24] **H. Rosenthal, S. Szarek**, “On tensor products of operators from  $L^p$  to  $L^q$ ”, *Functional Analysis*, 108–132, Springer-Verlag, Berlin Heidelberg (1991).
- [25] **W. Rudin**, “Trigonometric series with gaps”. *J. Math. Mech.* 9(2), 203–227.
- [26] **W. Rudin**, *Functional analysis*, McGraw-Hill International Series in pure and applied Mathematics, Singapore (1973).
- [27] **W. Rudin**, *Real and complex analysis*, McGraw-Hill International Editions, Mathematics Series, Singapore (1987).
- [28] **G. Schechtman**, “More on embedding subspaces of  $L_p$  in  $l_r^n$ ”, *Compositio Math.* 61.2, 159–169 (1987).
- [29] **R. Schneider, W. Weil**, “Zonoids and related topics”, *Convexity and its Applications*, 296–317 (1983).
- [30] **P. Sen**, “Random measurement bases, quantum state distinction and applications to the hidden subgroup problem”, *Proc. 21st IEEE Conference on Computational Complexity*, Piscataway, NJ (2006); arXiv:quant-ph/0512085.
- [31] **M. Talagrand**, “Embedding subspaces of  $L_1$  into  $\ell_1^N$ ”, *Proceedings of the American Mathematical Society* 108.2, 363–369 (1990).

INSTITUT CAMILLE JORDAN, UNIVERSITÉ CLAUDE BERNARD LYON 1, 43 BOULEVARD DU 11 NOVEMBRE 1918, 69622 VILLEURBANNE CEDEX, FRANCE

*E-mail address:* aubrun@math.univ-lyon1.fr

INSTITUT CAMILLE JORDAN, UNIVERSITÉ CLAUDE BERNARD LYON 1, 43 BOULEVARD DU 11 NOVEMBRE 1918, 69622 VILLEURBANNE CEDEX, FRANCE AND FÍSICA TEÒRICA: INFORMACIÓ I FENOMENS QUÀNTICS, UNIVERSITAT AUTÒNOMA DE BARCELONA, ES-08193 BELLATERRA (BARCELONA), SPAIN

*E-mail address:* lancien@math.univ-lyon1.fr