

Distinguishability of quantum states on high-dimensional multi-partite quantum systems

Master degree research project achieved at the Université Claude Bernard Lyon 1 (France)
under the supervision of Guillaume Aubrun

Cécilia LANCIEN
Université Pierre et Marie Curie Paris 6 and Ecole Polytechnique (France)

June 18th 2013

My Master degree research project was carried out in the Mathematics Department of the Université Claude Bernard Lyon 1 (France) under the supervision of Guillaume Aubrun, professor in the Probability Group there.

The task of distinguishing quantum states from accessible experimental data (i.e. from the outcomes of measurements which are performed on the studied quantum system) lies at the heart of quantum physics and quantum information theories. It is nonetheless by now a well-known fact that, when dealing with large composite quantum systems, the discrimination ability of observers might drop. One could thus roughly say that the guiding lead of the work achieved during my research placement was to use techniques from high-dimensional convex geometry and probabilities to investigate some of the numerous questions in this area.

The remainder of the current report is hence organized as follows.

Section 1 might be seen as a quick panorama of the mathematical framework in which quantum physics and quantum information theories develop. In section 2, the cornerstone our work relies on, namely the way one may systematically associate a norm to a measurement performed on a quantum system, is precisely stated. The link is made with the general issue of distinguishing quantum states under some allowed measurements. Section 3 is devoted to studying certain restricted classes of measurements on high-dimensional multi-partite quantum systems. Quantitative results are obtained regarding the capacity observers may have of discriminating between two quantum states when the only measurements they are able to perform are limited by locality constraints. The question studied in section 4 may be described in straightforward terms in the following manner: what would the minimal requirements on a randomly chosen measurement be so that it approximates the uniform measurement? All these plain words are of course given a rigorous mathematical meaning, and an accurate answer is provided too. As for section 5, the problem it deals with is not anymore the one of emulating one single measurement but instead the one of emulating the set of all possible measurements on a given quantum system.

To finish with, section 6 establishes a summary of the various results obtained and enumerates a few open questions, amongst many non-cited others.

Appendices A, B and C present required mathematical tools from geometric and probabilistic functional analysis. They contain more than the strictly necessary ideas to our aim, but nevertheless remain far from being exhaustive, proofs being either straightly omitted or only sketched. Appendix D exposes a few of the fundamental already known results within the wide and still extensively studied field of quantum states' geometry. As for appendix E, it is of much more technical nature: it provides a rather detailed proof of a norm inequality on the space of Hermitians on a tensor product Hilbert space (and which is needed at some point to generalize a result from the one-partite to the multi-partite setting).

Acknowledgements

First and foremost, I have to say how immensely grateful I am towards Guillaume Aubrun for having been so marvellous in supervising me. It is indeed tricky for a master degree student who still needs some guidance to achieve a research thesis at a distance. And without Guillaume's amazing reactivity and availability, things would for sure have turned out differently. What a luxury indeed that, whenever I was stuck into any long-lasting wandering, I could just send him out an SOS email and no doubt I would be "rescued" within an hour by just the appropriate remark of his. And what about my regular working days with him in Lyon, from which I was coming out simply stunned, since each single minute I had spent there had been dedicated to the progressing of my project.

I would also like to thank Andreas Winter, under the supervision of whom I had carried out a research thesis last year, for having been since then on so present. If ever I had a question flashing through my mind, on any random subject, I knew that I could submit it to him... and that he would even pretend it had some worth.

I am definitely entering into my PhD, co-advised by Guillaume and Andreas, with full confidence in those I am entrusting to and a lot of enthusiasm!

Contents

1	Introduction: The postulates of quantum mechanics and its mathematical formalism	5
2	Correspondence between measurements and symmetric convex bodies	6
2.1	General setting	6
2.2	Link with the task of distinguishing two quantum states under restricted families of measurements	7
2.3	What can one say about the value of a given measurement norm knowing the mean-width of its associated symmetric convex body?	8
3	Locally restricted measurements on a multi-partite quantum system	9
3.1	Different classes of locally restricted POVMs	9
3.2	Bi-partite case	10
3.2.1	PPT-measurements	11
3.2.2	SEP-measurements	12
3.2.3	“Typical” value of the PPT-norm and the SEP-norm	16
3.3	Multi-partite case	16
3.3.1	PPT-measurements	16
3.3.2	SEP-measurements	17
3.3.3	“Typical” value of the PPT-norm and the SEP-norm	17
4	POVMs with “few” outcomes whose measurement norm is equivalent to the one of the uniform POVM	18
4.1	One-partite case	18
4.1.1	First “rough” bound	19
4.1.2	Improved bound	21
4.2	Multi-partite case	23
5	Sets of POVMs with minimal cardinality whose measurement norm approximates the one of the set of all POVMs	25
5.1	Set of 2-outcome projective POVMs	25
5.2	Set of general POVMs	28
6	Conclusion and open questions	30
	Appendices	31
A	Convex geometry and functional analysis	31
A.1	Duality between norms and convex bodies	31
A.2	“Classic” geometric inequalities involving volumes	32
A.3	Volume-radius and mean-width of a convex body	33
A.4	Estimates on entropy numbers by a volumic approach	36
B	Gaussian variables	36
B.1	Generalities	36
B.2	Gaussian variables and mean-width	37
B.3	Brief incursion into random matrix theory: the GUE	38
C	Large deviations	39
C.1	Concentration rate and deviation inequalities on a probability metric space	39
C.2	Orlicz spaces and ψ_α -random variables	41
C.3	Tail bounds for sums of random matrices	42
D	Geometry of quantum states	43

D.1	Separability	43
D.2	Random states	45
E	Some properties of a family of norms	46
E.1	Special case $p = 2q$ even	47
E.1.1	Special case $q = 2$	47
E.1.2	General case $q \geq 2$	50
E.2	General case $p \geq 2$	51
	References	52

1 Introduction: The postulates of quantum mechanics and its mathematical formalism

Quantum mechanics does not tell what laws a physical system must obey but only provides a conceptual framework for the development of such laws. It relies on a few basic postulates which connect the physical world to the mathematical formalism that enables its description. The reader is referred to [1] or [2] for general and detailed references on this topic, the account made here being clearly minimalist.

Postulate 1: Associated to any isolated physical system is a Hilbert space \mathbb{H} known as its *state space*. The system is then completely described by its *state*, or *density operator*, which is a positive (hence Hermitian) operator with trace one acting on \mathbb{H} .

A state ρ is said to be *pure* if there exists a unit vector $|\psi\rangle \in \mathbb{H}$ such that $\rho = |\psi\rangle\langle\psi|$. It is otherwise referred to as being *mixed*.

If a system is known to be in state ρ_i with probability p_i for $i \in I$, then it may be described by the density operator $\rho = \sum_{i \in I} p_i \rho_i$ which is called a *mixture* of the density operators ρ_i .

Postulate 2: The state space of a composite physical system is the tensor product of the state spaces of the component physical systems. Moreover, if we have K sub-systems, numbered 1 through K , with sub-system i in state ρ_i for all $1 \leq i \leq K$, then the joint state of the total system is $\rho_1 \otimes \cdots \otimes \rho_n$.

Postulate 3: The evolution of a closed quantum system (i.e. a system that is not interacting in any way with other systems) is described by a *unitary transformation*: if the system is in state ρ at time t and in state ρ' at time t' , then there exists a unitary operator U acting on the system's state space (that depends only on t and t' , not on ρ and ρ') such that $\rho' = U\rho U^\dagger$.

Postulate 4: A quantum measurement performed on a physical system is described by a set $\{M_i, i \in I\}$ of *Positive Operator-Valued Measure (POVM) elements*, which are positive operators acting on the system's state space satisfying the *completeness equation* $\sum_{i \in I} M_i = 1$. The index $i \in I$ refers to

the measurement outcomes that may occur in the experiment. If the state of the system immediately before the measurement is ρ , then, for all $i \in I$, the probability that result i occurs is given by $\mathbb{P}_\rho(i) = \text{Tr}(M_i \rho)$ (so that the completeness equation simply expresses the fact that probabilities sum to one). The fact that each state ρ generates a probability distribution \mathbb{P}_ρ on the outputs $i \in I$ of a given measurement $\{M_i, i \in I\}$ is known as the *Born rule for measurements*.

We can actually be more precise: M_i being positive, $\sqrt{M_i}$ is well defined, and the state of the system just after the measurement that yielded outcome i is $\frac{\sqrt{M_i} \rho \sqrt{M_i}^\dagger}{\text{Tr}(M_i \rho)}$.

It may be pointed out that the free evolution $\rho \mapsto U\rho U^\dagger$ and the measurement $\rho \mapsto \frac{\sqrt{M} \rho \sqrt{M}^\dagger}{\text{Tr}(M \rho)}$ are two particular examples of *quantum operations*, i.e. operations that transform a quantum state into another. The most general way of describing such transformations is by a *Completely Positive and Trace Preserving (CPTP)* map.

- $\Lambda : \mathcal{H}(\mathbb{C}^m) \rightarrow \mathcal{H}(\mathbb{C}^n)$ is *Completely Positive (CP)* if:

$$\forall p \in \mathbb{N}, \forall \rho \in \mathcal{H}(\mathbb{C}^{m \times p}), \rho \geq \mathbf{O} \Rightarrow (\Lambda \otimes \text{Id})(\rho) \geq \mathbf{O}$$

\mathbb{C}^m here describes the state space of the input principal system and \mathbb{C}^n the state space of the output principal system, whereas \mathbb{C}^p should be thought of as the state space of any environment the system of interest might be coupled with. Thus, positivity of operators on the space of the global composite system is preserved when applying Λ to the part that acts on the principal system's space and leaving the part that acts on the environment's space invariant.

- $\Lambda : \mathcal{H}(\mathbb{C}^m) \rightarrow \mathcal{H}(\mathbb{C}^n)$ is *Trace Preserving (TP)* if:

$$\forall \rho \in \mathcal{H}(\mathbb{C}^m), \rho \geq \mathbf{O} \Rightarrow \text{Tr} \Lambda(\rho) = \text{Tr} \rho$$

Λ being a CPTP map is actually equivalent to the existence of so-called *Kraus operators* $(V_i)_{i \in I}$ that satisfy the completeness relation $\sum_{i \in I} V_i V_i^\dagger = 1$ and that are such that Λ can be written in the operator-sum representation as $\Lambda(\rho) = \sum_{i \in I} V_i \rho V_i^\dagger$.

2 Correspondence between measurements and symmetric convex bodies

2.1 General setting

Let $\mathbf{M} = (M_i)_{i \in I}$ be a POVM on \mathbb{C}^d . Denoting by $\{|i\rangle, i \in I\}$ an orthonormal basis of $\mathbb{C}^{|I|}$, we may associate to \mathbf{M} the following CPTP map (as just defined in section 1 above):

$$\mathcal{M} : \Delta \in \mathcal{H}(\mathbb{C}^d) \mapsto \sum_{i \in I} \text{Tr}(M_i \Delta) |i\rangle\langle i| \in \mathcal{H}(\mathbb{C}^{|I|})$$

The measurement norm associated to \mathbf{M} is then defined as:

$$\forall \Delta \in \mathcal{H}(\mathbb{C}^d), \|\Delta\|_{\mathbf{M}} := \|\mathcal{M}(\Delta)\|_1 = \sum_{i \in I} |\text{Tr}(M_i \Delta)|$$

More generally, one can define the measurement norm associated to a whole set \mathbf{M} of POVMs on \mathbb{C}^d as: $\|\cdot\|_{\mathbf{M}} := \sup_{M \in \mathbf{M}} \|\cdot\|_M$.

Remark 2.1 *Such designation seems to presume that the quantity we defined above is a norm. It is actually, whatever the set of POVMs \mathbf{M} , a semi-norm: it is non-negative, homogeneous and obeys the triangle inequality. It may however vanish on non-zero Hermitians in the general case. This is excluded when the set of POVMs \mathbf{M} is informationally complete, i.e. when:*

$$\forall \Delta \in \mathcal{H}(\mathbb{C}^d), \Delta \neq 0 \Rightarrow \exists (M_i)_{i \in I} \in \mathbf{M}, \exists i_0 \in I : \text{Tr}(M_{i_0} \Delta) \neq 0$$

This is equivalent to demanding that: $\text{Span}(\{M_i, i \in I, (M_i)_{i \in I} \in \mathbf{M}\}) = \mathcal{H}(\mathbb{C}^d)$, so that any density operator ρ on \mathbb{C}^d can be reconstructed from its outcome statistics $\{\text{Tr}(M_i \rho), i \in I, (M_i)_{i \in I} \in \mathbf{M}\}$ when measures from the set \mathbf{M} are carried on (which justifies the naming informationally complete). This especially implies that the total number of (distinct) POVM operators in \mathbf{M} is greater than $d^2 = \dim \mathcal{H}(\mathbb{C}^d)$. All the sets of POVMs we will later be lead to consider will have such property.

Something that is worth pointing at is that, for any set \mathbf{M} of POVMs on \mathbb{C}^d , there exists a set $\widetilde{\mathbf{M}}$ of 2-outcome POVMs on \mathbb{C}^d which is such that $\|\cdot\|_{\mathbf{M}} = \|\cdot\|_{\widetilde{\mathbf{M}}}$. It may be explicitly defined as:

$$\widetilde{\mathbf{M}} := \left\{ (M, 1 - M), \exists (M_i)_{i \in I} \in \mathbf{M}, \exists \tilde{I} \subset I : M = \sum_{i \in \tilde{I}} M_i \right\}$$

Indeed, for each $\Delta \in \mathcal{H}(\mathbb{C}^d)$, defining $I(\Delta) := \{i \in I, \text{Tr}(M_i \Delta) \geq 0\}$, and just recalling that $\sum_{i \in I(\Delta)} M_i + \sum_{i \in I \setminus I(\Delta)} M_i = \sum_{i \in I} M_i = 1$, we see that:

$$\sum_{i \in I} |\text{Tr}(M_i \Delta)| = \left| \sum_{i \in I(\Delta)} \text{Tr}(M_i \Delta) \right| + \left| \sum_{i \in I \setminus I(\Delta)} \text{Tr}(M_i \Delta) \right| = \left| \text{Tr} \left(\left(\sum_{i \in I(\Delta)} M_i \right) \Delta \right) \right| + \left| \text{Tr} \left(\left(1 - \sum_{i \in I(\Delta)} M_i \right) \Delta \right) \right|$$

That is precisely: $\|\Delta\|_{(M_i)_{i \in I}} = \|\Delta\|_{(M(\Delta), 1 - M(\Delta))}$ for $M(\Delta) := \sum_{i \in I(\Delta)} M_i$.

Hence, one may associate to any set \mathbf{M} of POVMs on \mathbb{C}^d a symmetric convex body $K_{\mathbf{M}}$, which is contained in the ∞ -norm unit ball of $\mathcal{H}(\mathbb{C}^d)$ (i.e. the operator interval $[-1; 1]$) and which contains ± 1 :

$$K_{\mathbf{M}} := \text{Conv}\{2M - 1, (M, 1 - M) \in \widetilde{\mathbf{M}}\}$$

This actually defines a one-to-one correspondence between informationally complete sets of POVMs on \mathbb{C}^d and symmetric convex bodies with non-empty interior contained in the ∞ -norm unit ball of $\mathcal{H}(\mathbb{C}^d)$ and containing ± 1 .

Furthermore: $\forall \Delta \in \mathcal{H}(\mathbb{C}^d)$, $\|\Delta\|_{\mathbf{M}} = \|\Delta\|_{\widetilde{\mathbf{M}}} = \sup_{A \in K_{\mathbf{M}}} |\text{Tr}(A\Delta)| = g_{(K_{\mathbf{M}})^\circ}(\Delta)$.

This means that: $B_{\|\cdot\|_{\mathbf{M}}} = B_{g_{(K_{\mathbf{M}})^\circ}} = (K_{\mathbf{M}})^\circ$, or equivalently that: $K_{\mathbf{M}} = (B_{\|\cdot\|_{\mathbf{M}}})^\circ$.

The reader is referred to appendix A.1 for all the used notations regarding norms associated to symmetric convex bodies.

Example 2.2 *The symmetric convex body associated with the set **ALL** of all POVMs on \mathbb{C}^d is nothing else than the ∞ -norm unit ball: $K_{\mathbf{ALL}} = B_{\|\cdot\|_{\infty}}^d$.*

Indeed: $(M, 1 - M)$ POVM $\Leftrightarrow 0 \leq M \leq 1 \Leftrightarrow -1 \leq 2M - 1 \leq 1$.

*As a consequence, the measurement norm associated with **ALL** is nothing else than the 1-norm:*

$$\|\cdot\|_{\mathbf{ALL}} = g_{(K_{\mathbf{ALL}})^\circ} = g_{(B_{\|\cdot\|_{\infty}}^d)^\circ} = g_{B_{\|\cdot\|_1}^d} = \|\cdot\|_1 \text{ (cf appendix A.1).}$$

2.2 Link with the task of distinguishing two quantum states under restricted families of measurements

Let us consider the situation where a system (with associated Hilbert space \mathbb{C}^d) can be either in state ρ or in state σ , with equal prior probabilities $\frac{1}{2}$. We would like to guess with the smallest probability of error in which of those two states it is by only performing one given POVM $M = (M_i)_{i \in I}$ on it. We therefore base our decision on the so-called *maximum likelihood rule*. Namely, knowing that $\text{Tr}(M_i \rho) > \text{Tr}(M_i \sigma)$ for $i \in \tilde{I}$ and $\text{Tr}(M_i \rho) < \text{Tr}(M_i \sigma)$ for $i \in I \setminus \tilde{I}$, we decide on ρ if outcome $i \in \tilde{I}$ is observed and on σ otherwise. The probability of error is thus, denoting by s the random variable “effective state of the system” and by d the random variable “state of the system we decide to be more likely after carrying out the measurement”:

$\mathbb{P}_e = \mathbb{P}(s = \sigma, d = \rho) + \mathbb{P}(s = \rho, d = \sigma) = \mathbb{P}(s = \sigma)\mathbb{P}(d = \rho | s = \sigma) + \mathbb{P}(s = \rho)\mathbb{P}(d = \sigma | s = \rho)$, that is:

$$\mathbb{P}_e = \frac{1}{2} \sum_{i \in \tilde{I}} \text{Tr}(M_i \sigma) + \frac{1}{2} \sum_{i \in I \setminus \tilde{I}} \text{Tr}(M_i \rho) = \frac{1}{2} \left(1 - \sum_{i \in \tilde{I}} \left| \text{Tr} \left[M_i \left(\frac{1}{2} \rho - \frac{1}{2} \sigma \right) \right] \right| \right) = \frac{1}{2} \left(1 - \left\| \frac{1}{2} \rho - \frac{1}{2} \sigma \right\|_{\mathbf{M}} \right)$$

In this context, the quantity $\left\| \frac{1}{2} \rho - \frac{1}{2} \sigma \right\|_{\mathbf{M}}$ is therefore called the *bias* of the POVM M on the state pair (ρ, σ) .

Remark 2.3 *We can easily generalize the discrimination task described above to states ρ and σ with non necessarily equal prior probabilities, q and $1 - q$ respectively. Indeed, the only change in that case is that we are now dealing with the general Hermitian $q\rho - (1 - q)\sigma$ instead of the traceless one $\frac{1}{2}\rho - \frac{1}{2}\sigma$.*

So for instance, the probability of error is then equal to: $\mathbb{P}_e = \frac{1}{2} (1 - \|q\rho - (1 - q)\sigma\|_{\mathbf{M}})$.

This result is actually nothing more than the generalization of a classical statistics’ result in hypothesis testing (see for instance [3] for a general reference). There, the optimal discrimination between two hypotheses modelled as probability distributions $\{P(i), i \in I\}$ and $\{Q(i), i \in I\}$, with prior probabilities q and $1 - q$ respectively, is in fact given by the maximum likelihood rule, so that the minimum probability of error takes value: $\mathbb{P}_e = \frac{1}{2} (1 - \|qP - (1 - q)Q\|_1)$, where $\|f\|_1 := \sum_{i \in I} |f(i)|$.

Suppose we are now interested in looking for the maximum bias achievable on a state pair (ρ, σ) (which corresponds to the minimum probability of error when trying to discriminate between states ρ and σ) when we are allowed POVMs in a given set \mathbf{M} . The quantity we will be lead to consider is then precisely: $\left\| \frac{1}{2} \rho - \frac{1}{2} \sigma \right\|_{\mathbf{M}}$.

One general issue in the field of quantum state discrimination is to compare, for various informationally complete sets of POVMs \mathbf{M} , the maximum bias achievable in discriminating two states when only measurements in \mathbf{M} are allowed to the one achievable when all measurements are allowed. As just stated, this boils down to comparing the distinguishability norm associated with \mathbf{M} to the one associated with \mathbf{ALL} , i.e. as pointed out in example 2.2 to the 1-norm. And actually, the result $\|\cdot\|_{\mathbf{ALL}} = \|\cdot\|_1$ was one of the seminal observations by Holevo [4] and Helstrom [5] on optimal quantum state distinction.

2.3 What can one say about the value of a given measurement norm knowing the mean-width of its associated symmetric convex body?

Let \mathbf{M} and \mathbf{M}' be two informationally complete sets of POVMs on \mathbb{C}^d .

As explained in section 2.1, showing that $\forall \Delta \in \mathcal{H}(\mathbb{C}^d)$, $\|\Delta\|_{\mathbf{M}} \leq \|\Delta\|_{\mathbf{M}'}$ amounts to showing that $K_{\mathbf{M}} \subset K_{\mathbf{M}'}$.

Now, it may happen that such inclusion does not hold although $K_{\mathbf{M}}$ is “much smaller” than $K_{\mathbf{M}'}$. In such case, one would expect that for “most” $\Delta \in \mathcal{H}(\mathbb{C}^d)$, $\|\Delta\|_{\mathbf{M}} \leq \|\Delta\|_{\mathbf{M}'}$.

It is precisely this intuitive idea that we will try to formalize rigorously in this section.

Let \mathbf{M} be an informationally complete set of POVMs on \mathbb{C}^d .

First of all, we know that: $\|\cdot\|_{\mathbf{M}} \leq \|\cdot\|_{\mathbf{ALL}} = \|\cdot\|_1 \leq \sqrt{d}\|\cdot\|_2$. So, denoting by $S_{\|\cdot\|_2}^d$ the unit sphere for the 2-norm on $\mathcal{H}(\mathbb{C}^d)$, it holds that $\|\cdot\|_{\mathbf{M}} : S_{\|\cdot\|_2}^d \mapsto \mathbb{R}^+$ is a \sqrt{d} -lipschitz function.

What is more, by definition of the mean-width w (cf appendix A.3) we have:

$\mathbb{E}_{\mathcal{U}(S_{\|\cdot\|_2}^d)} \|\cdot\|_{\mathbf{M}} = w\left(\left(B_{\|\cdot\|_{\mathbf{M}}}^d\right)^\circ\right) = w(K_{\mathbf{M}})$, where $\mathcal{U}(S_{\|\cdot\|_2}^d)$ stands for the probability distribution over $S_{\|\cdot\|_2}^d$ induced by the Hilbert-Schmidt distance on $\mathcal{H}(\mathbb{C}^d)$.

Hence, by the concentration inequality for lipschitz functions on the d^2 -dimensional real euclidean unit sphere $S_{\|\cdot\|_2}^d \equiv S_2^{d^2}(\mathbb{R})$ (cf example C.2) we get:

$$\forall 0 < \epsilon < 1, \mathbb{P}_{\Delta \sim \mathcal{U}(S_{\|\cdot\|_2}^d)} (\|\Delta\|_{\mathbf{M}} \notin [(1 - \epsilon)w(K_{\mathbf{M}}); (1 + \epsilon)w(K_{\mathbf{M}})]) \leq 2e^{-d^2(w(K_{\mathbf{M}})/\sqrt{d})^2 \epsilon^2/2}$$

Taking for instance $\epsilon = \frac{1}{2}$, this yields equivalently:

$$\mathbb{P}_{\Delta \sim \mathcal{U}(S_{\|\cdot\|_2}^d)} \left(\frac{1}{2}w(K_{\mathbf{M}}) \leq \|\Delta\|_{\mathbf{M}} \leq \frac{3}{2}w(K_{\mathbf{M}}) \right) \geq 1 - 2e^{-dw(K_{\mathbf{M}})^2/8} \quad (1)$$

By homogeneity, this implies that for any probability distribution $\nu_{HS(d)}$ on $\mathcal{H}(\mathbb{C}^d)$ which is induced by the Hilbert-Schmidt distance on $\mathcal{H}(\mathbb{C}^d)$:

$$\mathbb{P}_{\Delta \sim \nu_{HS(d)}} \left(\frac{1}{2}w(K_{\mathbf{M}})\|\Delta\|_2 \leq \|\Delta\|_{\mathbf{M}} \leq \frac{3}{2}w(K_{\mathbf{M}})\|\Delta\|_2 \right) \geq 1 - 2e^{-dw(K_{\mathbf{M}})^2/8} \quad (2)$$

Remark 2.4 *In the case when the set of POVMs under consideration is made of one single POVM, the result provided by equation 2 turns out to be, in some sense, quite disappointing.*

Indeed, one given rank-1 POVM M on \mathbb{C}^d (informationally complete or not) may be generically written in the form $M := \{d|\psi\rangle\langle\psi|dp_M(\psi), |\psi\rangle \in S_2^d(\mathbb{C})\}$ with dp_M an isotropic probability distribution over $S_2^d(\mathbb{C})$ (which means that $\int_{|\psi\rangle \in S_2^d(\mathbb{C})} dp_M(\psi) = 1$ and $\int_{|\psi\rangle \in S_2^d(\mathbb{C})} |\psi\rangle\langle\psi|dp_M(\psi) = \frac{1}{d}$).

With these notations, we have: $\forall \Delta \in \mathcal{H}(\mathbb{C}^d)$, $\|\Delta\|_M = d \int_{|\psi\rangle \in S_2^d(\mathbb{C})} |\langle\psi|\Delta|\psi\rangle|dp_M(\psi)$.

And subsequently: $w(K_M) = \mathbb{E}_{\Delta \sim \mathcal{U}(S_{\|\cdot\|_2}^d)} (\|\Delta\|_M) = d \int_{|\psi\rangle \in S_2^d(\mathbb{C})} \mathbb{E}_{\Delta \sim \mathcal{U}(S_{\|\cdot\|_2}^d)} (|\langle\psi|\Delta|\psi\rangle|)dp_M(\psi)$.

Now, referring to appendix B.3 for all useful definitions and statements, we see that for all $|\psi\rangle \in S_2^d(\mathbb{C})$:

$$\mathbb{E}_{\Delta \sim \mathcal{U}(S_{\|\cdot\|_2}^d)}(|\langle \psi | \Delta | \psi \rangle|) = \frac{1}{\gamma_{d^2}} \mathbb{E}_{G \sim GUE(d)}(|\langle \psi | G | \psi \rangle|) = \frac{1}{\gamma_{d^2}} \mathbb{E}_{g \sim \mathcal{N}(0,1)}(|g|) = \frac{1}{\gamma_{d^2}} \sqrt{\frac{2}{\pi}}$$

Hence, for any rank-1 POVM M on \mathbb{C}^d , we have: $w(K_M) = \frac{d}{\gamma_{d^2}} \sqrt{\frac{2}{\pi}} \underset{d \rightarrow +\infty}{\sim} \sqrt{\frac{2}{\pi}}$.

Inserting this into equation 2 shows that, for any rank-1 POVM M on \mathbb{C}^d :

$$\mathbb{P}_{\Delta \sim \nu_{HS(d)}} \left(\frac{1}{\sqrt{2\pi}} \|\Delta\|_2 \leq \|\Delta\|_M \leq \frac{3}{\sqrt{2\pi}} \|\Delta\|_2 \right) \geq 1 - 2e^{-d/4\pi}$$

What is more, this result remains true under restriction to traceless Hermitians (cf remark B.7):

$$\mathbb{P}_{\substack{\Delta \sim \nu_{HS(d)} \\ \text{Tr } \Delta = 0}} \left(\frac{1}{\sqrt{2\pi}} \|\Delta\|_2 \leq \|\Delta\|_M \leq \frac{3}{\sqrt{2\pi}} \|\Delta\|_2 \right) \geq 1 - 2e^{-d/4\pi}$$

Besides, we know that in the particular case of the uniform POVM U on \mathbb{C}^d , its measurement-norm is dimension-independently equivalent to the 2-norm on traceless Hermitians (which was really stressed upon for the first time in [42]):

$$\forall \Delta \in \mathcal{H}(\mathbb{C}^d), \text{Tr } \Delta = 0, \frac{1}{3} \|\Delta\|_2 \leq \|\Delta\|_U \leq \|\Delta\|_2$$

What we thus see is that, for a general rank-1 POVM M on \mathbb{C}^d , this dimension-independent equivalence with the 2-norm of course does not hold for all traceless Hermitians, but nonetheless for “most” of them (and for a growing proportion of them when the dimension d increases).

3 Locally restricted measurements on a multi-partite quantum system

Let H_1, \dots, H_K be K finite dimensional Hilbert spaces (with $d_i := \dim H_i$ for all $1 \leq i \leq K$) and $H = H_1 \otimes \dots \otimes H_K$ be their tensor product Hilbert space (of dimension $D := d_1 \times \dots \times d_K$).

3.1 Different classes of locally restricted POVMs

Several classes of POVMs can be defined on the K -partite Hilbert space H due to various levels of locality restrictions (consult [34] or [35] for further information).

The most restricted class of POVMs on H is the one of *local measurements* whose elements are tensor products of measurements on each of the sub-systems:

$$\mathbf{LO} := \left\{ \left(M_{j_1}^{(1)} \otimes \dots \otimes M_{j_K}^{(K)} \right)_{j_1 \in J_1, \dots, j_K \in J_K}, M_{j_i}^{(i)} \geq 0, j_i \in J_i, \sum_{j_i \in J_i} M_{j_i}^{(i)} = 1, 1 \leq i \leq K \right\}$$

More generally, **LOCC** is the class of measurements that can be implemented by a finite sequence of local operations on the sub-systems followed by classical communication between the parties.

Then, there is the class of *separable measurements* whose elements are the measurements on H made of operators that can be factorized as a tensor product of operators acting only on one sub-system:

$$\mathbf{SEP} := \left\{ \left(M_j^{(1)} \otimes \dots \otimes M_j^{(K)} \right)_{j \in J}, M_j^{(i)} \geq 0, j \in J, 1 \leq i \leq K, \sum_{j \in J} M_j^{(1)} \otimes \dots \otimes M_j^{(K)} = 1 \right\}$$

And finally, there is the class of the *positive under partial transpose measurements* whose elements are the measurements on H made of operators that remain positive when partially transposed on any combination of the sub-systems:

$$\mathbf{PPT} := \left\{ (M_j)_{j \in J}, M_j^{\Gamma_I} \geq 0, j \in J, I \subset \{1, \dots, K\}, \sum_{j \in J} M_j = 1 \right\}$$

where, for all $I \subset \{1, \dots, K\}$ the partial transposition on $H_I := \bigotimes_{i \in I} H_i$ is defined by its action on factorized operators on H : $(M_1 \otimes \dots \otimes M_K)^{\Gamma_I} := \left(\bigotimes_{i \in I} M_i^T \right) \otimes \left(\bigotimes_{i \notin I} M_i \right)$, M_i^T denoting the usual transpose of M_i .

Let us point out that, even though the expression of a matrix's transpose depends on the chosen basis, its eigenvalues on the contrary are intrinsic. So the PPT notion is indeed well defined.

Remark 3.1 *It is clear from the definitions that we have the chain of inclusions:*

$$\mathbf{LO} \subset \mathbf{LOCC} \subset \mathbf{SEP} \subset \mathbf{PPT} \subset \mathbf{ALL}$$

The most widely used inclusions in many questions dealing with operations on multi-partite quantum systems are certainly $\mathbf{LOCC} \subset \mathbf{SEP}$ and $\mathbf{LOCC} \subset \mathbf{PPT}$. Indeed, however natural it might seem in this context, the class of \mathbf{LOCC} operations is mathematically hard to characterize, contrary to the ones of \mathbf{SEP} operations and even more so of \mathbf{PPT} operations.

The reader may look at [27] for a general overview of some issues related to this topic, and at [28], [29] or [30] for a specific description of some striking phenomena one has to deal with when trying to grasp the class of \mathbf{LOCC} operations.

3.2 Bi-partite case

We consider here the case when $H = H_1 \otimes H_2 \equiv \mathbb{C}^d \otimes \mathbb{C}^d$ is a bi-partite Hilbert space whose parties have equal finite dimension d . The sets of 2-outcome POVMs associated to the sets of POVMs \mathbf{LO} , \mathbf{SEP} and \mathbf{PPT} on H are then (consult section 2.1 for the required definitions)

$$\begin{aligned} \widetilde{\mathbf{LO}} &= \left\{ (M, 1 - M), M = \sum_{(j_1, j_2) \in I} M_{j_1}^{(1)} \otimes M_{j_2}^{(2)}, I \subset J_1 \times J_2, M_{j_k}^{(k)} \geq 0, j_k \in J_k, \sum_{j_k \in J_k} M_{j_k}^{(k)} = 1, k \in \{1, 2\} \right\} \\ \widetilde{\mathbf{SEP}} &= \left\{ (M, 1 - M), M = \sum_{j \in I} M_j^{(1)} \otimes M_j^{(2)}, I \subset J, M_j^{(k)} \geq 0, j \in J, k \in \{1, 2\}, \sum_{j \in J} M_j^{(1)} \otimes M_j^{(2)} = 1 \right\} \\ \widetilde{\mathbf{PPT}} &= \left\{ (M, 1 - M), M = \sum_{j \in I} M_j, I \subset J, M_j \geq 0, M_j^\Gamma \geq 0, j \in J, \sum_{j \in J} M_j = 1 \right\} \end{aligned}$$

The main results the two coming sections 3.2.1 and 3.2.2 will lead us to are summarized below (consult appendix A.3 for the definition of vrad and w , and section 2.1 for the definition of $K_{\mathbf{SEP}}$ and $K_{\mathbf{PPT}}$):

Theorem 3.2 *(Volume-radii and mean-widths of the symmetric convex bodies associated with the sets of POVMs \mathbf{SEP} and \mathbf{PPT} on $\mathbb{C}^d \otimes \mathbb{C}^d$)*

$$d \simeq \text{vrad}(K_{\mathbf{PPT}}) \leq w(K_{\mathbf{PPT}}) \simeq d \quad \text{and} \quad \sqrt{d} \simeq \text{vrad}(K_{\mathbf{SEP}}) \leq w(K_{\mathbf{SEP}}) \simeq \sqrt{d}$$

A very simple idea we will use in an essential way to get both estimates in theorem 3.2 is the following: If one wants to evaluate the mean-width and the volume-radius of a given convex body K , it is enough to find an upper-bound on its mean-width and a lower-bound on its volume-radius, and to show that those two bounds are of the same order of magnitude, since we know thanks to Urysohn's inequality (theorem A.12) that $\text{vrad}(K) \leq w(K)$ always holds.

One theorem we shall also make repeated use of in the sequel, in order to come to the statements in theorem 3.2, is the one below (*cf* [19] for the original statement and proof):

Theorem 3.3 (*Milman-Pajor inequality*)

Let K, L be convex bodies with the same center of gravity.

Then: $\text{vrad}(K \cap L)\text{vrad}(K - L) \geq \text{vrad}(K)\text{vrad}(L)$.

As important special instances of the general statement from theorem 3.3, we have that, for any convex body K with center of gravity at the origin:

First of all: $\text{vrad}(K \cap -K)\text{vrad}(K + K) \geq \text{vrad}(K)\text{vrad}(-K)$. But since $\text{vrad}(-K) = \text{vrad}(K)$ and $\text{vrad}(K + K) = \text{vrad}(2K) = 2\text{vrad}(K)$, we get in the end: $\text{vrad}(K \cap -K) \geq \frac{1}{2}\text{vrad}(K)$.

More generally, for any orthogonal transformation θ : $\text{vrad}(K \cap \theta(K))\text{vrad}(K - \theta(K)) \geq \text{vrad}(K)\text{vrad}(\theta(K))$. Now: $\text{vrad}(\theta(K)) = \text{vrad}(K)$ and $\text{vrad}(K - \theta(K)) \leq w(K - \theta(K)) = w(K) + w(-\theta(K)) = 2w(K)$.

So eventually: $\text{vrad}(K \cap \theta(K)) \geq \frac{1}{2} \frac{\text{vrad}(K)^2}{w(K)}$.

Remark 3.4 *Theorem 3.3 is actually itself a corollary of a result that applies in a much wider context, namely the one of rotation invariant and log-concave measures:*

Let $\alpha, \beta \geq 0$ and consider the measure μ on \mathbb{R}^n with density $d\mu(x) = \alpha e^{-\beta\|x\|_2^2} dx$.

Let also $0 < \theta < \frac{\pi}{2}$ and $K, L \subset \mathbb{R}^n$ two convex bodies.

Set $z := \frac{\sin\theta}{\mu(K)} \int_K x d\mu(x) - \frac{\cos\theta}{\mu(L)} \int_L y d\mu(y)$ and $C(z) := \left(\frac{1}{\cos\theta}K - \frac{\sin\theta}{\cos\theta}z\right) \cap \left(\frac{1}{\sin\theta}L + \frac{\cos\theta}{\sin\theta}z\right)$.

Then: $\mu(K)\mu(L) \leq \mu(\sin\theta K - \cos\theta L)\mu(C(z))$.

3.2.1 PPT-measurements

Let us focus first on the set **PPT**. We see that:

$(M, 1 - M) \in \mathbf{PPT} \Leftrightarrow 0 \leq M, M^\Gamma, 1 - M, (1 - M)^\Gamma \leq 1 \Leftrightarrow 0 \leq M, M^\Gamma \leq 1$, so that:

$$K_{\mathbf{PPT}} = \text{Conv}\{2M - 1, 0 \leq M, M^\Gamma \leq 1\} = \left(B_{\|\cdot\|_\infty}^{d^2}\right) \cap \left(B_{\|\cdot\|_\infty}^{d^2}\right)^\Gamma$$

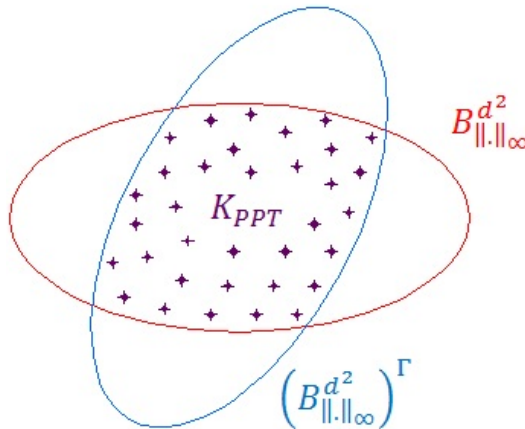


Figure 1: Symmetric convex body associated with the set of POVMs **PPT** on $\mathbb{C}^d \otimes \mathbb{C}^d$

We thus have the immediate upper-bound on the mean-width of $K_{\mathbf{PPT}}$: $w(K_{\mathbf{PPT}}) \leq w(B_{\|\cdot\|_\infty}^{d^2}) \simeq d$.

To get a lower-bound on the volume-radius of $K_{\mathbf{PPT}}$, we may apply Milman-Pajor inequality (theorem 3.3) to the convex body $B_{\|\cdot\|_\infty}^{d^2}$ (which indeed has the origin as center of gravity) and to the orthogonal

transformation Γ :
$$\text{vrad}(K_{\mathbf{PPT}}) \geq \frac{1}{2} \frac{\text{vrad}(B_{\|\cdot\|_\infty}^{d^2})}{w(B_{\|\cdot\|_\infty}^{d^2})} \text{vrad}(B_{\|\cdot\|_\infty}^{d^2}).$$

Hence, recalling that $\text{vrad}(B_{\|\cdot\|_\infty}^{d^2}) \simeq w(B_{\|\cdot\|_\infty}^{d^2})$ (cf example B.8): $\text{vrad}(K_{\mathbf{PPT}}) \gtrsim \text{vrad}(B_{\|\cdot\|_\infty}^{d^2}) \simeq d$.

Remark 3.5 *In this precise case, it is actually quite easy to be much more definite.*

In fact, we know that: $\text{vrad}(B_{\|\cdot\|_\infty}^{d^2}) \underset{d \rightarrow +\infty}{\sim} \frac{e^{1/4}}{\pi} d$ (cf theorem A.10), whereas: $w(B_{\|\cdot\|_\infty}^{d^2}) \underset{d \rightarrow +\infty}{\sim} \frac{8}{3\pi} d$ (cf example B.8).

We consequently have, when $d \rightarrow +\infty$, the quantitative estimate:

$$\frac{3e^{1/2}}{16\pi} (1 + o(1))d \leq \text{vrad}(K_{\mathbf{PPT}}) \leq w(K_{\mathbf{PPT}}) \leq \frac{8}{3\pi} (1 + o(1))d$$

3.2.2 SEP-measurements

Let us now look at the set **SEP**. Denoting by \mathcal{CS} the cone of separable positive operators on $\mathbb{C}^d \otimes \mathbb{C}^d$ (definition in appendix D.1) we see that: $(M, 1 - M) \in \widetilde{\mathbf{SEP}} \Leftrightarrow M, 1 - M \in \mathcal{CS} \cap B_{\|\cdot\|_\infty}^{d^2}$, so that:

$$K_{\mathbf{SEP}} = \text{Conv} \left\{ 2M - 1, M, 1 - M \in \mathcal{CS} \cap B_{\|\cdot\|_\infty}^{d^2} \right\} = \left\{ 2\mathcal{CS} \cap B_{\|\cdot\|_\infty}^{d^2} - 1 \right\} \cap \left\{ 1 - 2\mathcal{CS} \cap B_{\|\cdot\|_\infty}^{d^2} \right\}$$

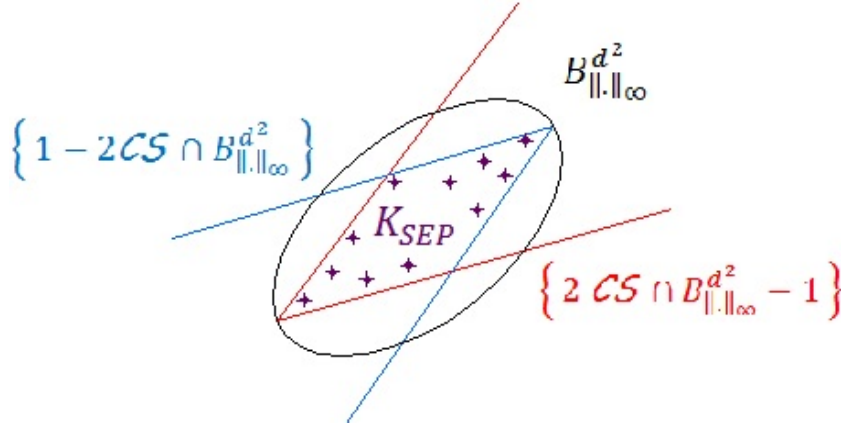


Figure 2: Symmetric convex body associated with the set of POVMs **SEP** on $\mathbb{C}^d \otimes \mathbb{C}^d$

In the sequel, we will denote by \mathcal{S} the set of separable states on $\mathbb{C}^d \otimes \mathbb{C}^d$ (definition in section 1 or appendix D.1), which is a convex body that is included in the hyperplane of $\mathcal{H}(\mathbb{C}^d \otimes \mathbb{C}^d)$ of trace 1 Hermitians on $\mathbb{C}^d \otimes \mathbb{C}^d$.

More generally, for any convex body S which is included in a hyperplane of $\mathcal{H}(\mathbb{C}^d \otimes \mathbb{C}^d)$, we will define the two following associated convex sets of full dimension (see again appendix D.1 for further information):

- \mathcal{C}_S the cone of basis S in $\mathcal{H}(\mathbb{C}^d \otimes \mathbb{C}^d)$.
- Σ_S the symmetrization of S in $\mathcal{H}(\mathbb{C}^d \otimes \mathbb{C}^d)$.

Firstly, we may notice that: $(\mathcal{CS} \cap B_{\|\cdot\|_\infty}^{d^2}) \subset (\mathcal{CS} \cap d^2 B_{\|\cdot\|_1}^{d^2}) = \mathcal{C}_{d^2 S}$.

Consequently: $K_{\text{SEP}} \subset \{2\mathcal{C}_{d^2 S} - 1\} \cap \{1 - 2\mathcal{C}_{d^2 S}\} \subset 2\Sigma_{d^2 S}$.

We thus have the immediate upper-bound: $w(K_{\text{SEP}}) \leq w(2\Sigma_{d^2 S}) \simeq d^2 w(\Sigma_S)$.

Now, we know that $w(\Sigma_S) \simeq \frac{1}{d^{3/2}}$ (cf theorem D.3), so we get in the end the upper-bound on the mean-width of K_{SEP} : $w(K_{\text{SEP}}) \lesssim \sqrt{d}$.

Secondly, for any fixed $1 < \alpha < 2$, we have: $(\mathcal{CS} \cap B_{\|\cdot\|_\infty}^{d^2}) \supset (\mathcal{CS} \cap \frac{d^2}{\alpha} B_{\|\cdot\|_1}^{d^2} \cap B_{\|\cdot\|_\infty}^{d^2}) = \frac{1}{\alpha} \mathcal{C}_{\alpha \tilde{B}_{d^2/\alpha} \cap d^2 S}$,

where for all $s \in \mathbb{R}$, we have defined \tilde{B}_s as: $\tilde{B}_s := \{M \in \mathcal{H}(\mathbb{C}^d \otimes \mathbb{C}^d), -1 \leq M \leq 1, \text{Tr}M = s\}$.

Hence: $\left\{ \frac{2}{\alpha} \mathcal{C}_{\alpha \tilde{B}_{d^2/\alpha} \cap d^2 S} - 1 \right\} \cap \left\{ 1 - \frac{2}{\alpha} \mathcal{C}_{\alpha \tilde{B}_{d^2/\alpha} \cap d^2 S} \right\} \subset K_{\text{SEP}}$.

Yet, for any convex body K in $\mathcal{H}(\mathbb{C}^d \otimes \mathbb{C}^d)$, we have by Fubini:

$\text{Vol}(K) = \int_{-\infty}^{+\infty} \text{Vol}(K \cap H_s) \frac{ds}{d} = \int_{-\infty}^{+\infty} d \text{Vol}(K \cap H_{td^2}) dt$, where for all $s \in \mathbb{R}$, we have defined H_s as the hyperplane of trace s Hermitians on $\mathbb{C}^d \otimes \mathbb{C}^d$: $H_s := \{M \in \mathcal{H}(\mathbb{C}^d \otimes \mathbb{C}^d), \text{Tr}M = s\}$ (whose Hilbert-Schmidt distance to the origin is $\frac{|s|}{d}$).

And what is more, if S is a convex body in $\mathcal{H}(\mathbb{C}^d \otimes \mathbb{C}^d)$ which is included in the hyperplane H_{d^2} and which is symmetric with respect to 1 in H_{d^2} , then for any $\lambda > 0$ and $t \in \mathbb{R}$, we have:

On the one hand, $\{\lambda \mathcal{C}_S - 1\} \cap H_{td^2} = \begin{cases} \{(1+t)S - 1\} & \text{if } -1 \leq t \leq \lambda - 1 \\ \emptyset & \text{otherwise} \end{cases}$, as shown on figure 3.

And analogously on the other hand, $\{1 - \lambda \mathcal{C}_S\} \cap H_{td^2} = \begin{cases} \{1 - (1-t)S\} & \text{if } -(\lambda - 1) \leq t \leq 1 \\ \emptyset & \text{otherwise} \end{cases}$.

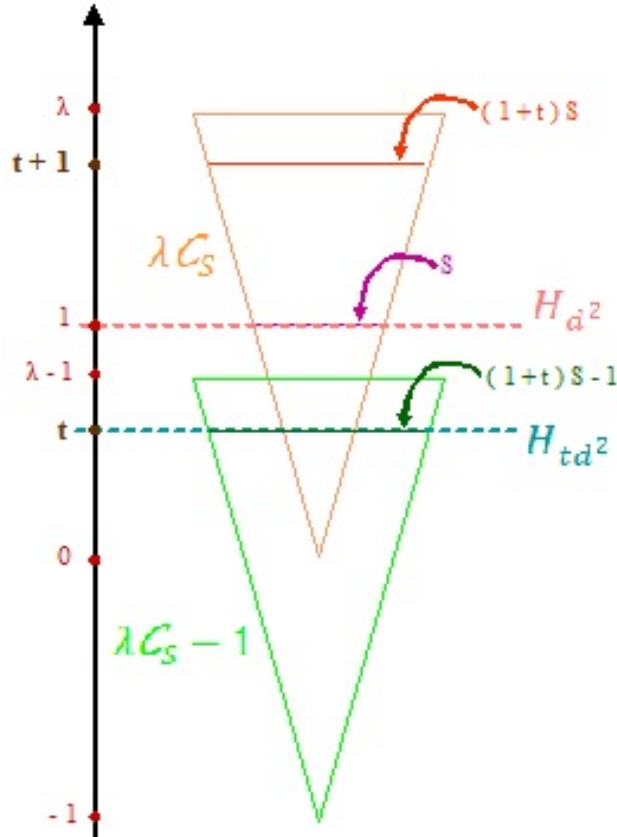


Figure 3: Section of the cone $\{\lambda \mathcal{C}_S - 1\}$ by the hyperplane H_{td^2} in $\mathcal{H}(\mathbb{C}^d \otimes \mathbb{C}^d)$

Subsequently, regarding the convex body K_{SEP} , we have for any $t \in \mathbb{R}$:

$$\begin{aligned} K_{\text{SEP}} \cap H_{td^2} &\supset \left\{ \frac{2}{\alpha} \mathcal{C}_{\alpha \tilde{B}_{d^2/\alpha} \cap d^2 \mathcal{S}} - 1 \right\} \cap \left\{ 1 - \frac{2}{\alpha} \mathcal{C}_{\alpha \tilde{B}_{d^2/\alpha} \cap d^2 \mathcal{S}} \right\} \cap H_{td^2} \\ &= \begin{cases} \tilde{B}_{td^2} \cap \{(1+t)d^2 \mathcal{S} - 1\} \cap \{1 - (1-t)d^2 \mathcal{S}\} & \text{if } -\left(\frac{2}{\alpha} - 1\right) \leq t \leq \frac{2}{\alpha} - 1 \\ \emptyset & \text{otherwise} \end{cases} \end{aligned}$$

$$\text{And therefore: } \text{Vol}(K_{\text{SEP}}) \geq \int_{-\left(\frac{2}{\alpha}-1\right)}^{\frac{2}{\alpha}-1} d \text{Vol} \left(\tilde{B}_{td^2} \cap \{(1+t)d^2 \mathcal{S} - 1\} \cap \{1 - (1-t)d^2 \mathcal{S}\} \right) dt$$

Furthermore, for any fixed $\delta > 0$ and any convex bodies $K_t \subset H_{td^2}$, $-\delta \leq t \leq \delta$, first applying Jensen's inequality to the concave function $(\cdot)^{1/d^4}$, and then setting $v := \sup_{-\delta \leq t \leq \delta} [\text{Vol}(K_t)]$, one gets:

$$\left[\int_{-\delta}^{\delta} d \text{Vol}(K_t) dt \right]^{1/d^4} \geq 2\delta \left(\frac{d}{2\delta} \right)^{1/d^4} \int_{-\delta}^{\delta} [\text{Vol}(K_t)]^{1/d^4} dt \geq \left(\frac{(2\delta)^{d^4-1} d}{v^{1/(d^4-1)}} \right)^{1/d^4} \int_{-\delta}^{\delta} [\text{Vol}(K_t)]^{1/(d^4-1)} dt$$

Since it holds additionally that $[\text{Vol}(B_2^{d^4})]^{1/d^4} \simeq [\text{Vol}(B_2^{d^4-1})]^{1/(d^4-1)} \simeq \frac{1}{d^4}$, one has in the end:

$$\left[\frac{\int_{-\delta}^{\delta} d \text{Vol}(K_t) dt}{\text{Vol}(B_2^{d^4})} \right]^{1/d^4} \gtrsim \int_{-\delta}^{\delta} \left[\frac{\text{Vol}(K_t)}{\text{Vol}(B_2^{d^4-1})} \right]^{1/(d^4-1)} dt.$$

As far as K_{SEP} is concerned, this implies in terms of volume-radius:

$$\text{vrad}(K_{\text{SEP}}) \gtrsim \int_{-\left(\frac{2}{\alpha}-1\right)}^{\frac{2}{\alpha}-1} d^2 \text{vrad} \left(\frac{1}{d^2} \tilde{B}_{td^2} \cap \left\{ (1+t)\mathcal{S} - \frac{1}{d^2} \right\} \cap \left\{ \frac{1}{d^2} - (1-t)\mathcal{S} \right\} \right) dt$$

Now, for all $-\left(\frac{2}{\alpha} - 1\right) \leq t \leq \frac{2}{\alpha} - 1$, we may apply Milman-Pajor inequality (theorem 3.3) to the convex bodies $\frac{1}{d^2} \tilde{B}_{td^2}$ and $\Omega_t := \{(1+t)\mathcal{S} - \frac{1}{d^2}\} \cap \{\frac{1}{d^2} - (1-t)\mathcal{S}\}$ (which indeed have same center of gravity

$$\frac{t}{d^2} \mathbf{1} \text{ as justified by theorem D.1): } \text{vrad} \left(\frac{1}{d^2} \tilde{B}_{td^2} \cap \Omega_t \right) \geq \frac{\text{vrad} \left(\frac{1}{d^2} \tilde{B}_{td^2} \right) \text{vrad}(\Omega_t)}{\text{vrad} \left(\frac{1}{d^2} \tilde{B}_{td^2} - \Omega_t \right)}$$

And applying this same inequality (theorem 3.3) once more, this time to the convex bodies $\{(1+t)\mathcal{S} - \frac{1}{d^2}\}$ and $\{\frac{1}{d^2} - (1-t)\mathcal{S}\}$ (which indeed have same center of gravity $\frac{t}{d^2} \mathbf{1}$ as justified by theorem D.1) gives:

$$\text{vrad}(\Omega_t) \geq \frac{\text{vrad}(\{(1+t)\mathcal{S} - \frac{1}{d^2}\}) \text{vrad}(\{\frac{1}{d^2} - (1-t)\mathcal{S}\})}{\text{vrad}(\{(1+t)\mathcal{S} - \frac{1}{d^2}\} - \{\frac{1}{d^2} - (1-t)\mathcal{S}\})} = \frac{(1+t)(1-t)}{2} \text{vrad}(\mathcal{S})$$

In order to go any further, we shall need the following result:

Lemma 3.6 (*Mean-width of hyperplane sections of the ∞ -norm unit ball*)

$$\forall 0 \leq t \leq 1, w(\tilde{B}_{td^2}) \gtrsim (1-t)w(B_{\|\cdot\|_\infty}^{d^2}) \quad \text{and} \quad \forall -1 \leq t \leq 0, w(\tilde{B}_{td^2}) \gtrsim (1+t)w(B_{\|\cdot\|_\infty}^{d^2})$$

Proof: Two preliminary statements will be necessary to come to the content of lemma 3.6:

- $\forall 0 \leq t \leq 1$, $\{1 + (1-t)\tilde{B}_0\} \subset \tilde{B}_{td^2}$ and $\forall -1 \leq t \leq 0$, $\{-1 + (1+t)\tilde{B}_0\} \subset \tilde{B}_{td^2}$
Indeed, for $0 \leq t \leq 1$: $M \in \{1 + (1-t)\tilde{B}_0\} \Rightarrow \begin{cases} \text{Tr } M = td^2 \\ -1 \leq -(1-2t)1 \leq M \leq 1 \end{cases} \Rightarrow M \in \tilde{B}_{td^2}$.
And for $-1 \leq t \leq 0$: $M \in \{-1 + (1+t)\tilde{B}_0\} \Rightarrow \begin{cases} \text{Tr } M = td^2 \\ -1 \leq -(1+2t)1 \leq M \leq 1 \end{cases} \Rightarrow M \in \tilde{B}_{td^2}$.
- $\text{vrad}(\tilde{B}_0) \gtrsim \text{vrad}(B_{\|\cdot\|_\infty}^{d^2})$
Indeed, by Brunn's principle (corollary A.6), $B_{\|\cdot\|_\infty}^{d^2}$ being a symmetric convex body, the function

$s \in \mathbb{R} \mapsto \text{Vol}(\tilde{B}_s) = \text{Vol}(B_{\|\cdot\|_\infty}^{d^2} \cap H_s)$ is maximal in 0.

Hence: $\text{Vol}(B_{\|\cdot\|_\infty}^{d^2}) = \int_{-1}^1 d \text{Vol}(\tilde{B}_{td^2}) dt \leq 2d \text{Vol}(\tilde{B}_0)$, which implies the advertized fact after

noting that: $[\text{Vol}(B_2^{d^4})]^{1/d^4} \simeq [\text{Vol}(B_2^{d^4-1})]^{1/(d^4-1)}$ and $[2d [\text{Vol}(B_{\|\cdot\|_\infty}^{d^2})]^{1/(d^4-1)}]^{-1/d^4} \simeq 1$.

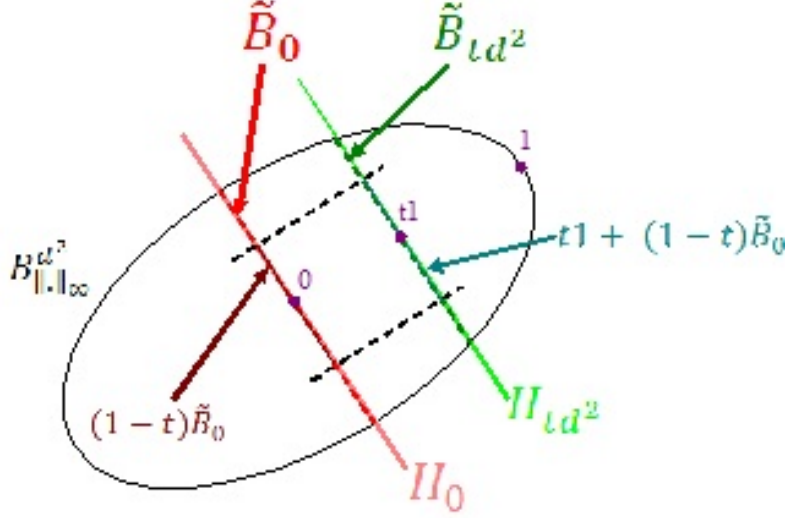


Figure 4: Section of the ball $B_{\|\cdot\|_\infty}^{d^2}$ by the hyperplane H_{td^2} in $\mathcal{H}(\mathbb{C}^d \otimes \mathbb{C}^d)$

These two assertions are illustrated by figure 4, and putting them together yields:

If $0 \leq t \leq 1$: $w(\tilde{B}_{td^2}) \geq \text{vrad}(\tilde{B}_{td^2}) \geq (1-t)\text{vrad}(\tilde{B}_0) \gtrsim (1-t)\text{vrad}(B_{\|\cdot\|_\infty}^{d^2}) \simeq (1-t)w(B_{\|\cdot\|_\infty}^{d^2})$, the first inequality being by Urysohn's inequality (theorem A.12) and the last equality by example B.8.

And similarly, if $-1 \leq t \leq 0$: $w(\tilde{B}_{td^2}) \gtrsim (1+t)w(B_{\|\cdot\|_\infty}^{d^2})$.

Which are precisely the results stated in lemma 3.6.

Keeping this in mind and coming back to our initial issue, we see that:

For $0 \leq t \leq \frac{2}{\alpha} - 1$: $\begin{cases} w\left(\frac{1}{d^2}\tilde{B}_{td^2}\right) \gtrsim \frac{(1-t)}{d^2}w\left(B_{\|\cdot\|_\infty}^{d^2}\right) \simeq \frac{(1-t)}{d}$ (because $w(B_{\|\cdot\|_\infty}^{d^2}) \simeq d$ by example B.8)
 $w(\Omega_t) \leq w((1-t)\mathcal{S}) \simeq \frac{1-t}{d^{3/2}}$ (because $w(\mathcal{S}) \simeq \frac{1}{d^{3/2}}$ by theorem D.3)

And for $-\left(\frac{2}{\alpha} - 1\right) \leq t \leq 0$: $\begin{cases} w\left(\frac{1}{d^2}\tilde{B}_{td^2}\right) \gtrsim \frac{(1+t)}{d^2}w\left(B_{\|\cdot\|_\infty}^{d^2}\right) \simeq \frac{(1+t)}{d}$
 $w(\Omega_t) \leq w((1+t)\mathcal{S}) \simeq \frac{1+t}{d^{3/2}}$

Subsequently, for all $-\left(\frac{2}{\alpha} - 1\right) \leq t \leq \frac{2}{\alpha} - 1$: $w(\Omega_t) \leq w\left(\frac{1}{d^2}\tilde{B}_{td^2}\right)$, and therefore:

$\text{vrad}\left(\frac{1}{d^2}\tilde{B}_{td^2} - \Omega_t\right) \leq w\left(\frac{1}{d^2}\tilde{B}_{td^2} - \Omega_t\right) = w\left(\frac{1}{d^2}\tilde{B}_{td^2}\right) + w(\Omega_t) \lesssim w\left(\frac{1}{d^2}\tilde{B}_{td^2}\right)$, where we used Urysohn's inequality (theorem A.12) to get the first inequality.

Hence, since $\text{vrad}\left(\frac{1}{d^2}\tilde{B}_{td^2}\right) \simeq w\left(\frac{1}{d^2}\tilde{B}_{td^2}\right)$ (cf example B.8), we finally get for all $-\left(\frac{2}{\alpha} - 1\right) \leq t \leq \frac{2}{\alpha} - 1$:

$$\text{vrad}\left(\frac{1}{d^2}\tilde{B}_{td^2} \cap \Omega_t\right) \gtrsim \frac{\text{vrad}\left(\frac{1}{d^2}\tilde{B}_{td^2}\right)}{w\left(\frac{1}{d^2}\tilde{B}_{td^2}\right)} \frac{(1+t)(1-t)}{2} \text{vrad}(\mathcal{S}) \gtrsim \frac{(1+t)(1-t)}{2} \text{vrad}(\mathcal{S})$$

Consequently, what we come to in the end is: $\text{vrad}(K_{\text{SEP}}) \gtrsim d^2 \text{vrad}(\mathcal{S})$.

And since we know that $\text{vrad}(\mathcal{S}) \simeq \frac{1}{d^{3/2}}$ (cf theorem D.3), we eventually get the lower-bound on the volume-radius of K_{SEP} : $\text{vrad}(K_{\text{SEP}}) \gtrsim \sqrt{d}$.

3.2.3 “Typical” value of the PPT-norm and the SEP-norm

It was mentioned earlier on (*cf* section 2.3) that, if one is interested in the most probable value of a given measurement norm, it is the mean-width rather than the volume-radius of its associated symmetric convex body which is the relevant quantity.

What theorem 3.2 especially tells us is that the mean-widths of the symmetric convex bodies associated with the sets of POVMs **PPT** and **SEP** on $\mathbb{C}^d \otimes \mathbb{C}^d$ are of order of magnitude:

$$w(K_{\mathbf{PPT}}) \simeq d \quad \text{and} \quad w(K_{\mathbf{SEP}}) \simeq \sqrt{d}$$

In terms of the measurement norms $\|\cdot\|_{\mathbf{PPT}}$ and $\|\cdot\|_{\mathbf{SEP}}$ on $\mathcal{H}(\mathbb{C}^d \otimes \mathbb{C}^d)$, this implies by equation 1:

$$\exists C, C' > 0 : \begin{cases} \mathbb{P}_{\Delta \sim \mathcal{U}(S_{\|\cdot\|_2}^{d^2})} \left(\frac{C}{2}d \leq \|\Delta\|_{\mathbf{PPT}} \leq \frac{3C}{2}d \right) \geq 1 - 2e^{-C^2 d^4/8} \\ \mathbb{P}_{\Delta \sim \mathcal{U}(S_{\|\cdot\|_2}^{d^2})} \left(\frac{C'}{2}\sqrt{d} \leq \|\Delta\|_{\mathbf{SEP}} \leq \frac{3C'}{2}\sqrt{d} \right) \geq 1 - 2e^{-C'^2 d^3/8} \end{cases}$$

3.3 Multi-partite case

The results obtained in the bi-partite setting $\mathbb{H} = \mathbb{H}_1 \otimes \mathbb{H}_2 \equiv \mathbb{C}^d \otimes \mathbb{C}^d$ may in fact be quite directly generalized to the multi-partite one $\mathbb{H} = \mathbb{H}_1 \otimes \cdots \otimes \mathbb{H}_K \equiv (\mathbb{C}^d)^{\otimes K}$.

Denoting by **PPT**(d, K) and **SEP**(d, K) the sets of, respectively, positive under partial transpose and separable POVMs on $(\mathbb{C}^d)^{\otimes K}$ we have:

$$K_{\mathbf{PPT}(d,K)} = \bigcap_{I \subset \{1, \dots, K\}} \left(B_{\|\cdot\|_\infty}^{d^K} \right)^{\Gamma_I}$$

$$K_{\mathbf{SEP}(d,K)} = \left\{ 2\mathcal{CS}^{(d,K)} \cap B_{\|\cdot\|_\infty}^{d^K} - 1 \right\} \cap \left\{ 1 - 2\mathcal{CS}^{(d,K)} \cap B_{\|\cdot\|_\infty}^{d^K} \right\}$$

where $\mathcal{CS}^{(d,K)}$ stands for the cone of separable positive operators on $(\mathbb{C}^d)^{\otimes K}$ (definition in appendix D.1).

3.3.1 PPT-measurements

To begin with, let us look at $K_{\mathbf{PPT}(d,K)}$.

First of all, we still have the obvious upper-bound on the mean-width of $K_{\mathbf{PPT}(d,K)}$:

$$w(K_{\mathbf{PPT}(d,K)}) \leq w\left(B_{\|\cdot\|_\infty}^{d^K}\right) \simeq \sqrt{d^K}$$

Moreover, iterating Milman-Pajor inequality (theorem 3.3), we see that if K_1, \dots, K_m are m convex

bodies with the same center of gravity, then:
$$\text{vrad}\left(\bigcap_{i=1}^m K_i\right) \geq \frac{\prod_{i=1}^m \text{vrad}(K_i)}{\prod_{i=1}^{m-1} \text{vrad}\left(K_i - \bigcap_{j=i+1}^m K_j\right)}.$$

Hence, in the particular case when $K_i = \theta_i(K)$, $1 \leq i \leq m$, for a given convex body K with center of gravity at the origin and given orthogonal transformations $\theta_1, \dots, \theta_m$, we get:

$$\text{vrad}\left(\bigcap_{i=1}^m \theta_i(K)\right) \geq \frac{1}{2^{m-1}} \frac{[\text{vrad}(K)]^m}{[w(K)]^{m-1}},$$
 due to the fact that:

$$\begin{cases} \text{vrad}(\theta_i(K)) = \text{vrad}(K), \quad 1 \leq i \leq m \\ \text{vrad}\left(\theta_i(K) - \bigcap_{j=i+1}^m \theta_j(K)\right) \leq w\left(\theta_i(K) - \bigcap_{j=i+1}^m \theta_j(K)\right) \leq w(\theta_i(K)) + w(\theta_{i+1}(K)) = 2w(K), \quad 1 \leq i \leq m-1 \end{cases}$$

If additionally, $w(K) \simeq \text{vrad}(K)$, we eventually come to: $\text{vrad}\left(\bigcap_{i=1}^m \theta_i(K)\right) \gtrsim \frac{1}{2^{m-1}} \text{vrad}(K)$.

Applying this general result to the convex body $B_{\|\cdot\|_\infty}^{d^K}$ and to the orthogonal transformations Γ_I , $I \subset \{1, \dots, K\}$, we obtain the lower-bound on the volume-radius of $K_{\text{PPT}(d,K)}$:

$$\text{vrad}(K_{\text{PPT}(d,K)}) \gtrsim \frac{1}{2^{2^K}} \text{vrad}(B_{\|\cdot\|_\infty}^{d^K}) \simeq \frac{1}{2^{2^K}} \sqrt{d^K}$$

Theorem 3.7 (*Volume-radius and mean-width of the symmetric convex body associated with the set of POVMs PPT on $(\mathbb{C}^d)^{\otimes K}$*)

$$\frac{1}{2^{2^K}} d^{K/2} \lesssim \text{vrad}(K_{\text{PPT}(d,K)}) \leq w(K_{\text{PPT}(d,K)}) \lesssim d^{K/2}$$

Remark 3.8 *As in the bi-partite case, we can be more precise (thanks to the explicit results of theorem A.10 and example B.8) and provide the quantitative estimate, for K fixed and $d \rightarrow +\infty$:*

$$\left(\frac{3e^{1/4}}{16}\right)^{2^{K-1}-1} \frac{e^{1/4}}{\pi} d^{K/2} (1 + o(1)) \leq \text{vrad}(K_{\text{PPT}(d,K)}) \leq w(K_{\text{PPT}(d,K)}) \leq \frac{8}{3\pi} d^{K/2} (1 + o(1))$$

3.3.2 SEP-measurements

Let us look now at $K_{\text{SEP}(d,K)}$.

In complete analogy to the bi-partite case, denoting by $\mathcal{S}^{(d,K)}$ the convex set of separable states on $(\mathbb{C}^d)^{\otimes K}$ (definition in section 1 or appendix D.1) we have that for any fixed $1 < \alpha < 2$:

$$\left\{ \frac{2}{\alpha} \mathcal{C}_{\alpha \tilde{B}_{d^K/\alpha}^+ \cap d^K \mathcal{S}^{(d,K)}} - 1 \right\} \cap \left\{ 1 - \frac{2}{\alpha} \mathcal{C}_{\alpha \tilde{B}_{d^K/\alpha}^+ \cap d^K \mathcal{S}^{(d,K)}} \right\} \subset K_{\text{SEP}(d,K)} \subset 2\Sigma_{d^K \mathcal{S}^{(d,K)}}$$

Then:
$$\left\{ \begin{array}{l} w(2\Sigma_{d^K \mathcal{S}^{(d,K)}}) \simeq d^K w(\mathcal{S}^{(d,K)}) \\ \text{vrad}\left(\left\{ \frac{2}{\alpha} \mathcal{C}_{\alpha \tilde{B}_{d^K/\alpha}^+ \cap d^K \mathcal{S}^{(d,K)}} - 1 \right\} \cap \left\{ 1 - \frac{2}{\alpha} \mathcal{C}_{\alpha \tilde{B}_{d^K/\alpha}^+ \cap d^K \mathcal{S}^{(d,K)}} \right\}\right) \simeq d^K \text{vrad}(\mathcal{S}^{(d,K)}) \end{array} \right. .$$

Besides, we know (*cf* theorem D.3 for a brief overview and [22] for a detailed account) that:

$$\frac{2^{-K}}{d^{K-1/2}} \lesssim \text{vrad}(\mathcal{S}^{(d,K)}) \leq w(\mathcal{S}^{(d,K)}) \lesssim \frac{\sqrt{K \log K}}{d^{K-1/2}}$$

Theorem 3.9 (*Volume-radius and mean-width of the symmetric convex body associated with the set of POVMs SEP on $(\mathbb{C}^d)^{\otimes K}$*)

$$2^{-K} \sqrt{d} \lesssim \text{vrad}(K_{\text{SEP}(d,K)}) \leq w(K_{\text{SEP}(d,K)}) \lesssim \sqrt{K \log K} \sqrt{d}$$

3.3.3 ‘‘Typical’’ value of the PPT-norm and the SEP-norm

Regarding the mean-widths of $K_{\text{PPT}(d,K)}$ and $K_{\text{SEP}(d,K)}$, what we have shown is that, if we only focus on the dimensional dependence and not on the number of party dependence, they scale as:

$$w(K_{\text{PPT}(d,K)}) \simeq d^{K/2} \quad \text{and} \quad w(K_{\text{SEP}(d,K)}) \simeq \sqrt{d}$$

Such information may then be plugged into equation 1 and hence translated into statements on the “typical” values of the measurement norms $\|\cdot\|_{\mathbf{PPT}(d,K)}$ and $\|\cdot\|_{\mathbf{SEP}(d,K)}$ on $\mathcal{H}((\mathbb{C}^d)^{\otimes K})$:

$$\exists C_K, C'_K > 0 : \begin{cases} \mathbb{P}_{\Delta \sim \mathcal{U}(S_{\|\cdot\|_2}^{d^K})} \left(\frac{C_K}{2} d^{K/2} \leq \|\Delta\|_{\mathbf{PPT}(d,K)} \leq \frac{3C_K}{2} d^{K/2} \right) \geq 1 - 2e^{-C_K^2 d^{2K}/8} \\ \mathbb{P}_{\Delta \sim \mathcal{U}(S_{\|\cdot\|_2}^{d^K})} \left(\frac{C'_K}{2} \sqrt{d} \leq \|\Delta\|_{\mathbf{SEP}(d,K)} \leq \frac{3C'_K}{2} \sqrt{d} \right) \geq 1 - 2e^{-C_K'^2 d^{K+1}/8} \end{cases}$$

What this result means is that, for “most” $\Delta \in \mathcal{H}((\mathbb{C}^d)^{\otimes K})$, $\left\{ \begin{array}{l} \|\Delta\|_{\mathbf{PPT}(d,K)} \simeq d^{K/2} \|\Delta\|_2 \simeq \|\Delta\|_1 \\ \|\Delta\|_{\mathbf{SEP}(d,K)} \simeq \sqrt{d} \|\Delta\|_2 \simeq \|\Delta\|_{2K/(K+1)} \end{array} \right.$

Indeed, we know from example B.8 that, for all $1 \leq p \leq +\infty$, $w(B_{\|\cdot\|_p}^{d^K}) \simeq (d^K)^{1/2-1/p}$, so that $\|\cdot\|_p \simeq w(B_{\|\cdot\|_{p/(p-1)}}^{d^K}) \|\cdot\|_2 \simeq (d^K)^{1/p-1/2} \|\cdot\|_2$.

It may also be rephrased in the following way: For a fixed number K of parties and a growing dimension d of each of these parties, the set of PPT-measurements behaves (with probability tending to 1) roughly like the set of all measurements. On the contrary, the set of SEP-measurements is far from reaching this same discriminating power since $\|\cdot\|_{\mathbf{SEP}(d,K)} \simeq \frac{1}{d^{(K-1)/2}} \|\cdot\|_{\mathbf{ALL}(d,K)}$.

Remark 3.10 *It was established in [35] that: $\forall \Delta \in \mathcal{H}((\mathbb{C}^d)^{\otimes K})$, $\left\{ \begin{array}{l} \|\Delta\|_{\mathbf{PPT}(d,K)} \geq \|\Delta\|_2 \geq \frac{1}{d^{K/2}} \|\Delta\|_1 \\ \|\Delta\|_{\mathbf{SEP}(d,K)} \geq \frac{2}{2^{K/2}} \|\Delta\|_2 \geq \frac{2}{(2d)^{K/2}} \|\Delta\|_1 \end{array} \right.$*

These lower-bounds were furthermore shown to be first order optimal, at least in their dimensional dependence, since: $\exists \Delta \in \mathcal{H}((\mathbb{C}^d)^{\otimes K})$, $\Delta \neq 0$: $\|\Delta\|_{\mathbf{SEP}(d,K)} \leq \|\Delta\|_{\mathbf{PPT}(d,K)} \leq \frac{2}{d^{K/2+1}} \|\Delta\|_1$.

Nevertheless, what we have just demonstrated here is that these so-called data-hiding Hermitians on $\mathcal{H}((\mathbb{C}^d)^{\otimes K})$ (in the sense introduced, among others, by [39], [40] or [41]), even though they exist, remain “exceptionnal”.

4 POVMs with “few” outcomes whose measurement norm is equivalent to the one of the uniform POVM

4.1 One-partite case

Let $d \in \mathbb{N}^*$ and denote by $U := \{d|\psi\rangle\langle\psi|d\psi, |\psi\rangle \in S_2^d(\mathbb{C})\}$ the uniform POVM on \mathbb{C}^d . The measurement norm associated with U is by definition (cf section 2.1):

$$\forall \Delta \in \mathcal{H}(\mathbb{C}^d), \|\Delta\|_U := \int_{|\psi\rangle \in S_2^d(\mathbb{C})} |\mathrm{Tr}(d|\psi\rangle\langle\psi|\Delta)| d\psi := d\mathbb{E}|\mathrm{Tr}(P_U \Delta)|$$

with P_U the random rank-1 projector taking value $|\psi\rangle\langle\psi|$ with probability $d\psi$.

Remark 4.1 *Let $1 \leq p \leq +\infty$.*

For any informationally complete set of POVMs \mathbf{M} on \mathbb{C}^d , define: $\left\{ \begin{array}{l} \lambda_p(\mathbf{M}) := \inf_{\|\Delta\|_p=1} \|\Delta\|_{\mathbf{M}} \\ \mu_p(\mathbf{M}) := \sup_{\|\Delta\|_p=1} \|\Delta\|_{\mathbf{M}} \end{array} \right.$

Those are respectively the largest constant λ and smallest constant μ such that: $\lambda\|\cdot\|_p \leq \|\cdot\|_{\mathbf{M}} \leq \mu\|\cdot\|_p$, or equivalently such that: $\lambda B_{\|\cdot\|_p}^d \subset K_{\mathbf{M}} \subset \mu B_{\|\cdot\|_p}^d$ (cf section 2.1).

Due to the convex structure, if $\{\mathbf{M}_i, i \in I\}$ are informationally complete sets of POVMs on \mathbb{C}^d , then

$$\text{for any probability distribution } \{p_i, i \in I\}, \text{ we have: } \left\{ \begin{array}{l} \lambda_p \left(\sum_{i \in I} p_i \mathbf{M}_i \right) \geq \sum_{i \in I} p_i \lambda_p(\mathbf{M}_i) \\ \mu_p \left(\sum_{i \in I} p_i \mathbf{M}_i \right) \leq \sum_{i \in I} p_i \mu_p(\mathbf{M}_i) \end{array} \right.$$

And due to the unitary invariance, if \mathbf{M} is an informationally complete set of POVMs on \mathbb{C}^d , then

for any unitary V on \mathbb{C}^d , we have: $\begin{cases} \lambda_p(\mathbf{VMV}^\dagger) = \lambda_p(\mathbf{M}) \\ \mu_p(\mathbf{VMV}^\dagger) = \mu_p(\mathbf{M}) \end{cases}$.

Consequently we get in the end that, if \mathbf{M} is an informationally complete set of POVMs on \mathbb{C}^d , then for any probability distribution $\{dp(V), V \in \mathfrak{U}(d)\}$ on the unitaries of \mathbb{C}^d , we have:

$$\lambda_p \left(\int_{\mathfrak{U}(d)} \mathbf{VMV}^\dagger dp(V) \right) \geq \lambda_p(\mathbf{M}) \quad \text{and} \quad \mu_p \left(\int_{\mathfrak{U}(d)} \mathbf{VMV}^\dagger dp(V) \right) \leq \mu_p(\mathbf{M})$$

What we thus see is that, for any single informationally complete POVM M on \mathbb{C}^d , $\lambda_p(M) \leq \lambda_p(U)$ and $\mu_p(M) \geq \mu_p(U)$. To put it in more trivial terms, the uniform POVM on \mathbb{C}^d is the “best” single POVM on \mathbb{C}^d , and that is why we shall take a special interest in it from now on.

The question we address in this section is: what is the minimal number of outcomes a POVM on \mathbb{C}^d must have in order for its measurement norm to “behave like” the one of the uniform POVM on \mathbb{C}^d ?

Let us be more precise. Defining the “modified 2-norm” on $\mathcal{H}(\mathbb{C}^d)$ as: $\|\Delta\|_{2(1)} := \sqrt{\text{Tr}(\Delta^2) + (\text{Tr}\Delta)^2}$, we know from [35] that the following inequalities hold:

$$\frac{1}{\sqrt{18}} \|\cdot\|_{2(1)} \leq \|\cdot\|_U \leq \|\cdot\|_{2(1)} \quad (3)$$

In light of equation 3, our question thus becomes more specifically: what is the minimal number of outcomes a POVM on \mathbb{C}^d should have in order for its measurement norm to be in this way dimension-independently equivalent to the “modified 2-norm”?

4.1.1 First “rough” bound

Let $n \in \mathbb{N}^*$ and $\{P_k, 1 \leq k \leq n\}$ independent random rank-1 projectors with the same probability distribution as P_U . Set $S := \sum_{k=1}^n P_k$ (which is a random positive operator that is almost surely invertible for $n \geq d$) and consider $P := \{\tilde{P}_k := S^{-1/2} P_k S^{-1/2}, 1 \leq k \leq n\}$, random POVM on \mathbb{C}^d made of n random rank-1 operators, and whose associated measurement norm on $\mathcal{H}(\mathbb{C}^d)$ is: $\|\Delta\|_P = \sum_{k=1}^n |\text{Tr}(\tilde{P}_k \Delta)|$.

• **Step 1:** Large deviation probability for $\frac{d}{n} \sum_{k=1}^n |\text{Tr}(P_k \Delta)|$, $\Delta \in S_{\|\cdot\|_{2(1)}}^d$

Let $0 < \delta < 1$ and consider \mathcal{M}_δ a δ -net for $\|\cdot\|_{2(1)}$ within $S_{\|\cdot\|_{2(1)}}^d$, the unit sphere for $\|\cdot\|_{2(1)}$ in $\mathcal{H}(\mathbb{C}^d)$. We may choose \mathcal{M}_δ such that $|\mathcal{M}_\delta| \leq (1 + \frac{2}{\delta})^{d^2}$ (cf example A.13).

Let $0 < \epsilon < 1$.

For any fixed $\Delta \in \mathcal{M}_\delta$, $\{|\text{Tr}(P_k \Delta)|, 1 \leq k \leq n\}$ are i.i.d. random variables taking values in $[0; 1]$ (because $\forall 1 \leq k \leq n, 0 \leq |\text{Tr}(P_k \Delta)| \leq \|P_k\|_2 \|\Delta\|_2 \leq \|P_k\|_2 \|\Delta\|_{2(1)} = 1$).

Hence, denoting by μ there common expectancy, we have by Chernoff’s inequality (corollary C.11):

$$\mathbb{P} \left(\frac{1}{n} \sum_{k=1}^n |\text{Tr}(P_k \Delta)| \notin [(1 - \epsilon)\mu; (1 + \epsilon)\mu] \right) \leq 2e^{-n\mu\epsilon^2/4}.$$

Yet: $\mu = \mathbb{E}|\text{Tr}(P_U \Delta)| = \frac{1}{d} \|\Delta\|_U$. So by equation 3, we get the estimate: $\frac{1}{\sqrt{18}d} \leq \mu \leq \frac{1}{d}$.

$$\text{Therefore: } \mathbb{P} \left(\frac{1}{n} \sum_{k=1}^n |\text{Tr}(P_k \Delta)| \notin \left[\frac{1 - \epsilon}{\sqrt{18}d}; \frac{1 + \epsilon}{d} \right] \right) \leq 2e^{-n\epsilon^2/4\sqrt{18}d}.$$

Which implies by the union bound that:

$$\mathbb{P} \left(\exists \Delta \in \mathcal{M}_\delta : \frac{1}{n} \sum_{k=1}^n |\text{Tr}(P_k \Delta)| \notin \left[\frac{1-\epsilon}{\sqrt{18d}}; \frac{1+\epsilon}{d} \right] \right) \leq |\mathcal{M}_\delta| 2e^{-n\epsilon^2/4\sqrt{18}d} \leq 2 \left(1 + \frac{2}{\delta}\right)^{d^2} e^{-n\epsilon^2/4\sqrt{18}d}$$

$$\text{Or equivalently: } \mathbb{P} \left(\forall \Delta \in \mathcal{M}_\delta, \frac{d}{n} \sum_{k=1}^n |\text{Tr}(P_k \Delta)| \in \left[\frac{1-\epsilon}{\sqrt{18}}; 1+\epsilon \right] \right) \geq 1 - 2 \left(1 + \frac{2}{\delta}\right)^{d^2} e^{-n\epsilon^2/4\sqrt{18}d}.$$

Yet, if we have: $\forall \Delta \in \mathcal{M}_\delta, \frac{1-\epsilon}{\sqrt{18}} \leq \frac{d}{n} \sum_{k=1}^n |\text{Tr}(P_k \Delta)| \leq 1+\epsilon$, then we necessarily have:

$$\forall \Delta \in S_{\|\cdot\|_{2(1)}}^d, \frac{1-\epsilon}{\sqrt{18}} - \frac{\delta(1+\epsilon)}{1-\delta} \leq \frac{d}{n} \sum_{k=1}^n |\text{Tr}(P_k \Delta)| \leq \frac{1+\epsilon}{1-\delta}.$$

Indeed, let us denote by $\{|k\rangle, 1 \leq k \leq n\}$ an orthonormal basis of \mathbb{C}^n and suppose that the operator $\text{T} : \Delta \in \mathcal{H}(\mathbb{C}^d) \mapsto \frac{d}{n} \sum_{k=1}^n \text{Tr}(P_k \Delta) |k\rangle\langle k| \in \mathcal{H}(\mathbb{C}^n)$ satisfies: $\forall \tilde{\Delta} \in \mathcal{M}_\delta, \frac{1-\epsilon}{\sqrt{18}} \leq \|\text{T}(\tilde{\Delta})\|_1 \leq 1+\epsilon$.

Consider first $\Delta_0 \in S_{\|\cdot\|_{2(1)}}^d$ such that $\|\text{T}(\Delta_0)\|_1 = \sup_{\Delta \in S_{\|\cdot\|_{2(1)}}^d} \|\text{T}(\Delta)\|_1 = \|\text{T}\|_{(\mathcal{H}(\mathbb{C}^d), \|\cdot\|_{2(1)}) \rightarrow (\mathcal{H}(\mathbb{C}^n), \|\cdot\|_1)}$.

By assumption: $\exists \tilde{\Delta}_0 \in \mathcal{M}_\delta : \|\Delta_0 - \tilde{\Delta}_0\|_{2(1)} \leq \delta$.

So: $\|\text{T}(\Delta_0)\|_1 \leq \|\text{T}(\Delta_0 - \tilde{\Delta}_0)\|_1 + \|\text{T}(\tilde{\Delta}_0)\|_1 \leq \|\text{T}(\Delta_0)\|_1 \delta + (1+\epsilon)$, that is: $\|\text{T}(\Delta_0)\|_1 \leq \frac{1+\epsilon}{1-\delta}$.

And thus: $\forall \Delta \in S_{\|\cdot\|_{2(1)}}^d, \|\text{T}(\Delta)\|_1 \leq \|\text{T}(\Delta_0)\|_1 \leq \frac{1+\epsilon}{1-\delta}$.

Now, consider any $\Delta \in S_{\|\cdot\|_{2(1)}}^d$. By assumption: $\exists \tilde{\Delta} \in \mathcal{M}_\delta : \|\Delta - \tilde{\Delta}\|_{2(1)} \leq \delta$.

So: $\|\text{T}(\Delta)\|_1 \geq \|\text{T}(\tilde{\Delta})\|_1 - \|\text{T}(\Delta - \tilde{\Delta})\|_1 \geq \frac{1-\epsilon}{\sqrt{18}} - \frac{1+\epsilon}{1-\delta} \delta$.

Hence, we actually get as advertized: $\forall \Delta \in S_{\|\cdot\|_{2(1)}}^d, \frac{1-\epsilon}{\sqrt{18}} - \frac{\delta(1+\epsilon)}{1-\delta} \leq \|\text{T}(\Delta)\|_1 \leq \frac{1+\epsilon}{1-\delta}$.

Subsequently, what we eventually come to is:

$$\mathbb{P} \left(\forall \Delta \in S_{\|\cdot\|_{2(1)}}^d, \frac{d}{n} \sum_{k=1}^n |\text{Tr}(P_k \Delta)| \in \left[\frac{1-\epsilon}{\sqrt{18}} - \frac{\delta(1+\epsilon)}{1-\delta}; \frac{1+\epsilon}{1-\delta} \right] \right) \geq 1 - 2 \left(1 + \frac{2}{\delta}\right)^{d^2} e^{-n\epsilon^2/4\sqrt{18}d}$$

Choosing for instance $\delta = \frac{1}{1+2\sqrt{18}}$ (so that $\frac{\delta}{1-\delta} = \frac{1}{2\sqrt{18}}$ and $\frac{1}{1-\delta} = 1 + \frac{1}{2\sqrt{18}} \leq \frac{3}{2}$) and $\epsilon = \frac{1}{6}$, we get:

$$\mathbb{P} \left(\forall \Delta \in S_{\|\cdot\|_{2(1)}}^d, \frac{d}{n} \sum_{k=1}^n |\text{Tr}(P_k \Delta)| \in \left[\frac{1}{4\sqrt{18}}; \frac{7}{4} \right] \right) \geq 1 - 2(3 + 4\sqrt{18})^{d^2} e^{-n/144\sqrt{18}d} \quad (4)$$

• **Step 2:** Large deviation probability for $\sum_{k=1}^n |\text{Tr}(\tilde{P}_k \Delta)|, \Delta \in S_{\|\cdot\|_{2(1)}}^d$

Let $0 < \eta < 1$.

Since $\{P_k, 1 \leq k \leq n\}$ are independent random variables taking values in the operator interval $[0; 1]$, the matrix Chernoff's inequality (corollary C.12) yields, denoting by $M \geq \mu 1$ there common

expectancy: $\mathbb{P} \left(\frac{1}{n} \sum_{k=1}^n P_k \notin [(1-\eta)M; (1+\eta)M] \right) \leq 2de^{-n\mu\eta^2/4}$.

Yet: $M = \mathbb{E}P_U = \frac{1}{d}1 = \frac{1}{d} \sum_{k=1}^n \tilde{P}_k$, so that: $\mathbb{P} \left(\frac{d}{n} \sum_{k=1}^n P_k \notin \left[(1-\eta) \sum_{k=1}^n \tilde{P}_k; (1+\eta) \sum_{k=1}^n \tilde{P}_k \right] \right) \leq 2de^{-n\eta^2/4d}$.

Now, if $\left| \frac{d}{n} \sum_{k=1}^n P_k - \sum_{k=1}^n \tilde{P}_k \right| \leq \eta 1$, then for any $\Delta \in S_{\|\cdot\|_{2(1)}}^d, \|\Delta\|_1 \leq \sqrt{d}\|\Delta\|_2 \leq \sqrt{d}\|\Delta\|_{2(1)} = \sqrt{d}$, so

$$\left| \frac{d}{n} \sum_{k=1}^n |\text{Tr}(P_k \Delta)| - \sum_{k=1}^n |\text{Tr}(\tilde{P}_k \Delta)| \right| \leq \text{Tr} \left| \left(\frac{d}{n} \sum_{k=1}^n P_k - \sum_{k=1}^n \tilde{P}_k \right) \Delta \right| \leq \left\| \frac{d}{n} \sum_{k=1}^n P_k - \sum_{k=1}^n \tilde{P}_k \right\|_{\infty} \|\Delta\|_1 \leq \eta \sqrt{d}$$

Hence in the end:

$$\mathbb{P} \left(\forall \Delta \in S_{\|\cdot\|_{2(1)}}^d, \left| \frac{d}{n} \sum_{k=1}^n |\text{Tr}(P_k \Delta)| - \sum_{k=1}^n |\text{Tr}(\tilde{P}_k \Delta)| \right| \geq \sqrt{d\eta} \right) \leq 2de^{-n\eta^2/4d} \quad (5)$$

Choosing for instance $\eta = \frac{1}{8\sqrt{18}\sqrt{d}}$ in equation 5 (so that $\left[\frac{1}{4\sqrt{18}} - \sqrt{d\eta}; \frac{7}{4} + \sqrt{d\eta}\right] \subset \left[\frac{1}{8\sqrt{18}}; \frac{15}{8}\right]$), and combining it with equation 4, we finally get:

$$\mathbb{P} \left(\forall \Delta \in S_{\|\cdot\|_{2(1)}}^d, \sum_{k=1}^n |\text{Tr}(\tilde{P}_k \Delta)| \in \left[\frac{1}{8\sqrt{18}}; \frac{15}{8}\right] \right) \geq 1 - 2 \left((3 + 4\sqrt{18})^{d^2} e^{-n/144\sqrt{18}d} + de^{-n/4608d^2} \right)$$

• Step 3: Conclusion

By homogeneity, what we obtain in the end is:

$$\mathbb{P} \left(\forall \Delta \in \mathcal{H}(\mathbb{C}^d), \frac{1}{8\sqrt{18}} \|\Delta\|_{2(1)} \leq \|\Delta\|_P \leq \frac{15}{8} \|\Delta\|_{2(1)} \right) \geq 1 - 2 \left((3 + 4\sqrt{18})^{d^2} e^{-n/144\sqrt{18}d} + de^{-n/4608d^2} \right)$$

This implies that:

$$\forall 0 < \alpha < 1, \exists C_\alpha > 0 : n \geq C_\alpha d^3 \Rightarrow \mathbb{P} \left(\forall \Delta \in \mathcal{H}(\mathbb{C}^d), \frac{1}{8\sqrt{18}} \|\Delta\|_{2(1)} \leq \|\Delta\|_P \leq \frac{15}{8} \|\Delta\|_{2(1)} \right) \geq 1 - \alpha$$

The result we have come to might be formulated in the following way: a POVM on \mathbb{C}^d obtained from $\Omega(d^3)$ appropriately renormalized randomly chosen rank-1 projectors will behave (with high probability) as well as the uniform POVM on \mathbb{C}^d . However, nothing in our reasoning guarantees that this lower-bound is actually tight, the only *a priori* lower-bound being that a POVM on \mathbb{C}^d has to have at least d^2 outcomes to be informationally complete...

4.1.2 Improved bound

Actually, it may be shown that, indeed, $\Omega(d^2)$ randomly chosen rank-1 operators are enough to get a POVM on \mathbb{C}^d that emulates the uniform one. To come to this optimal result, one has to make use of refined deviation inequalities.

• Step 1: Preliminary technical result

Let $\Delta \in S_{\|\cdot\|_{2(1)}}^d$ and consider the centered random variable $X := d|\text{Tr}(P_U \Delta)| - \|\Delta\|_U$.

For each $p \in \mathbb{N}^*$, we may upper-bound its p -order moment by:

$$\mathbb{E}|X|^p = \mathbb{E} \left| \sum_{q=0}^p \binom{p}{q} d^q |\text{Tr}(P_U \Delta)|^q (-1)^{p-q} \|\Delta\|_U^{p-q} \right| \leq \sum_{q=0}^p \binom{p}{q} d^q \mathbb{E} |\text{Tr}(P_U \Delta)|^q \|\Delta\|_U^{p-q}$$

Furthermore, by Jensen's inequality: $\forall 0 \leq q \leq p$, $\mathbb{E} |\text{Tr}(P_U \Delta)|^q \leq \sqrt{\mathbb{E} [\text{Tr}(P_U \Delta)]^{2q}}$. And:

$$\mathbb{E} [\text{Tr}(P_U \Delta)]^{2q} = \mathbb{E} [\text{Tr}(P_U^{\otimes 2q} \Delta^{\otimes 2q})] = \int_{|\psi\rangle \in S_2^q(\mathbb{C})} \text{Tr}(|\psi\rangle\langle\psi|^{\otimes 2q} \Delta^{\otimes 2q}) d\psi = \text{Tr} \left(\left(\int_{|\psi\rangle \in S_2^q(\mathbb{C})} |\psi\rangle\langle\psi|^{\otimes 2q} d\psi \right) \Delta^{\otimes 2q} \right)$$

Yet: $\forall r \in \mathbb{N}$, $\int_{|\psi\rangle \in S_2^q(\mathbb{C})} |\psi\rangle\langle\psi|^{\otimes r} d\psi = \frac{r!}{(d+r-1) \times \dots \times d} P_{\text{Sym}(d,r)} = \frac{1}{(d+r-1) \times \dots \times d} \sum_{\pi \in \mathfrak{S}_r} U(\pi)$, where $P_{\text{Sym}(d,r)}$

denotes the orthogonal projector onto the completely symmetric subspace of $(\mathbb{C}^d)^{\otimes r}$, and for each permutation $\pi \in \mathfrak{S}_r$, $U(\pi)$ denotes the associated permutation unitary on $(\mathbb{C}^d)^{\otimes r}$.

Thus: $\mathbb{E} [\text{Tr}(P_U \Delta)]^{2q} = \frac{1}{(d+2q-1) \times \dots \times d} \sum_{\pi \in \mathfrak{S}_{2q}} \text{Tr}(U(\pi) \Delta^{\otimes 2q})$.

Now: $\forall \pi \in \mathfrak{S}_{2q}$, $\text{Tr}(U(\pi) \Delta^{\otimes 2q}) = \left(\prod_{i=1}^{c(\pi)} \text{Tr}(\Delta^{l(i)}) \right) (\text{Tr} \Delta)^{2q-s(\pi)}$, where $c(\pi)$ is the number of non-trivial cycles of π , $l(i)$, $1 \leq i \leq c(\pi)$, are the respective lengths of those non-trivial cycles, and

$s(\pi) = \sum_{i=1}^{c(\pi)} l(i)$ is the number of non-fixed points of π .

Yet if $l = 2t$ is even: $|\text{Tr}(\Delta^{2t})| \leq [\text{Tr}(\Delta^2)]^t$.

And if $l = 2t+1$ is odd: $|\text{Tr}(\Delta^{2t+1})| \leq [\text{Tr}(\Delta^{4t})\text{Tr}(\Delta^2)]^{1/2} \leq [\text{Tr}(\Delta^2)]^{t+1/2} \leq \frac{[\text{Tr}(\Delta^2)]^t + [\text{Tr}(\Delta^2)]^{t+1}}{2}$.

Whereas if $s = 2t$ is even: $|\text{Tr}(\Delta)^{2q-2t}| = [\text{Tr}(\Delta)^{2(q-t)}]$.

And if $s = 2t + 1$ is odd: $|\text{Tr}(\Delta)^{2q-(2t+1)}| \leq \frac{[\text{Tr}(\Delta)^{2(q-t-1)} + [\text{Tr}(\Delta)^{2(q-t)}]}{2}$.

Hence, recalling that $\|\Delta\|_{2(1)}^{2q} = (\text{Tr}(\Delta^2) + [\text{Tr}(\Delta)^2]^q = \sum_{j=0}^q \binom{q}{j} [\text{Tr}(\Delta^2)]^j [\text{Tr}(\Delta)^{2(q-j)}]$, we finally get

that: $\forall \pi \in \mathfrak{S}_{2q}$, $|\text{Tr}(U(\pi)\Delta^{\otimes 2q})| \leq \|\Delta\|_{2(1)}^{2q}$.

So in the end: $\sum_{\pi \in \mathfrak{S}_{2q}} \text{Tr}(U(\pi)\Delta^{\otimes 2q}) \leq \sum_{\pi \in \mathfrak{S}_{2q}} |\text{Tr}(U(\pi)\Delta^{\otimes 2q})| \leq (2q)! \|\Delta\|_{2(1)}^{2q}$.

Thus, just noticing that $\frac{(2q)!}{(d+2q-1) \times \dots \times d} \leq \frac{(2q)^{2q}}{d^{2q}}$, we eventually come to:

$$\mathbb{E}|X|^p \leq \sum_{q=0}^p \binom{p}{q} (2q)^q \|\Delta\|_{2(1)}^q \|\Delta\|_U^{p-q} \leq p^p (2\|\Delta\|_{2(1)} + \|\Delta\|_U)^p$$

And since by equation 3: $\|\Delta\|_U \leq \|\Delta\|_{2(1)} = 1$, we actually have: $\mathbb{E}|X|^p \leq (3p)^p$.

This implies that X is a centered ψ_1 random variable (see appendix C.2 for all definitions and statements concerning this matter) with ψ_1 -norm satisfying: $\|X\|_{\psi_1} \leq 2e^2 \sup_{p \geq 1} \frac{(\mathbb{E}|X|^p)^{1/p}}{p} \leq 2e^2 \times 3 = 6e^2$.

With this result in mind, let us now turn back to our initial strategy. Just as before, we first draw $\{P_k, 1 \leq k \leq n\}$ independently with the same probability distribution as P_U , and then consider the random POVM $P := \{\tilde{P}_k := S^{-1/2}P_k S^{-1/2}, 1 \leq k \leq n\}$, where $S := \sum_{k=1}^n P_k$.

• **Step 2:** Large deviation probability for $\frac{d}{n} \sum_{k=1}^n |\text{Tr}(P_k \Delta)|$, $\Delta \in S_{\|\cdot\|_{2(1)}}^d$

For a given $0 < \delta < 1$, we consider here again \mathcal{M}_δ a δ -net for $\|\cdot\|_{2(1)}$ within $S_{\|\cdot\|_{2(1)}}^d$, that we choose such that $|\mathcal{M}_\delta| \leq (1 + \frac{2}{\delta})^{d^2}$ (cf example A.13).

By what precedes, for any $\Delta \in \mathcal{M}_\delta$, $\{X_k := d|\text{Tr}(P_k \Delta)| - \|\Delta\|_U, 1 \leq k \leq n\}$ are i.i.d. centered ψ_1 random variables with ψ_1 -norm bounded by $6e^2$. So by Bernstein's inequality (theorem C.8):

$$\forall \epsilon > 0, \mathbb{P}\left(\left|\frac{1}{n} \sum_{k=1}^n X_k\right| \geq \epsilon\right) \leq 2 \exp\left(-\frac{n}{4} \min\left(\frac{\epsilon^2}{(6e^2)^2}, \frac{\epsilon}{6e^2}\right)\right)$$

And whenever $\epsilon \leq 6e^2$, we have: $\min\left(\frac{\epsilon^2}{(6e^2)^2}, \frac{\epsilon}{6e^2}\right) = \frac{\epsilon^2}{(6e^2)^2}$.

Hence, reasoning exactly as in section 4.1.1, what we come to is that, for all $0 < \epsilon < 1$:

$$\text{First: } \mathbb{P}\left(\forall \Delta \in \mathcal{M}_\delta, \frac{d}{n} \sum_{k=1}^n |\text{Tr}(P_k \Delta)| \in \left[\frac{1-\epsilon}{\sqrt{18}}; 1+\epsilon\right]\right) \geq 1 - 2\left(1 + \frac{2}{\delta}\right)^{d^2} e^{-n\epsilon^2/144e^4}.$$

$$\text{Then: } \mathbb{P}\left(\forall \Delta \in S_{\|\cdot\|_{2(1)}}^d, \frac{d}{n} \sum_{k=1}^n |\text{Tr}(P_k \Delta)| \in \left[\frac{1-\epsilon}{\sqrt{18}} - \frac{\delta(1+\epsilon)}{1-\delta}; \frac{1+\epsilon}{1-\delta}\right]\right) \geq 1 - 2\left(1 + \frac{2}{\delta}\right)^{d^2} e^{-n\epsilon^2/144e^4}$$

Choosing, again just as in section 4.1.1, $\delta = \frac{1}{1+2\sqrt{18}}$ and $\epsilon = \frac{1}{6}$, we thus get:

$$\mathbb{P}\left(\forall \Delta \in S_{\|\cdot\|_{2(1)}}^d, \frac{d}{n} \sum_{k=1}^n |\text{Tr}(P_k \Delta)| \in \left[\frac{1}{4\sqrt{18}}; \frac{7}{4}\right]\right) \geq 1 - 2(5\sqrt{18})^{d^2} e^{-n/36 \times 144e^4} \quad (6)$$

• **Step 3:** Large deviation probability for $\sum_{k=1}^n |\text{Tr}(\tilde{P}_k \Delta)|$, $\Delta \in S_{\|\cdot\|_2(1)}^d$

In order to conclude, we must at last bound $\mathbb{P}\left(\left\|\frac{d}{n}\sum_{k=1}^n P_k - \sum_{k=1}^n \tilde{P}_k\right\|_{\infty} \geq \eta\right)$ for any fixed $\eta > 0$.

Yet: $\|\cdot\|_{\infty} = \|\cdot\|_{(\mathbb{C}^d, \|\cdot\|_2) \rightarrow (\mathbb{C}^d, \|\cdot\|_2)}$, so: $\forall M \in \mathcal{H}(\mathbb{C}^d)$, $\|M\|_{\infty} = \sup_{x \in S_2^d(\mathbb{C})} \|Mx\|_2 = \sup_{x \in S_2^d(\mathbb{C})} |\langle x|M|x\rangle|$.

Besides, considering \mathcal{N} a $\frac{1}{3}$ -net for $\|\cdot\|_2$ within $S_2^d(\mathbb{C})$, we have: $\forall x \in S_2^d(\mathbb{C})$, $\exists \tilde{x} \in \mathcal{N} : \|x - \tilde{x}\|_2 \leq \frac{1}{3}$, so that: $|\langle x|M|x\rangle| \leq |\langle x|M|x - \tilde{x}\rangle| + |\langle x - \tilde{x}|M|\tilde{x}\rangle| + |\langle \tilde{x}|M|\tilde{x}\rangle|$
 $\leq \|x - \tilde{x}\|_2 \|Mx\|_2 + \|x - \tilde{x}\|_2 \|M\tilde{x}\|_2 + |\langle \tilde{x}|M|\tilde{x}\rangle|$
 $\leq \frac{1}{3} \|M\|_{\infty} + \frac{4}{3} \sup_{y \in \mathcal{N}} |\langle y|M|y\rangle|$

Thus in the end: $\|M\|_{\infty} \leq 2 \sup_{y \in \mathcal{N}} |\langle y|M|y\rangle|$, which implies by the union bound that:

$$\forall \eta > 0, \mathbb{P}(\|M\|_{\infty} \geq \eta) \leq |\mathcal{N}| \sup_{y \in \mathcal{N}} \mathbb{P}(|\langle y|M|y\rangle| \geq \frac{\eta}{2})$$

Now, let $y \in \mathcal{N}$ and consider the random variable $Y := \langle y|dP_U - 1|y\rangle = d\langle y|P_U|y\rangle - 1$.

It is centered (because $\mathbb{E}(y|P_U|y) = \frac{1}{d}$), and following the exact same lines as above, we might upper-bound its p -order moment, $p \in \mathbb{N}^*$, by: $\mathbb{E}|Y|^p \leq p^p (2\|y\|_2 \langle y\|_2(1) + 1)^p = p^p (2\sqrt{2} + 1)^p \leq (4p)^p$.

This implies that Y is a centered ψ_1 random variable with ψ_1 -norm satisfying: $\|Y\|_{\psi_1} \leq 8e^2$.

Hence, for any $y \in \mathcal{N}$, $\{Y_k := \langle y|dP_k - 1|y\rangle, 1 \leq k \leq n\}$ are i.i.d. centered ψ_1 random variables with ψ_1 -norm bounded by $8e^2$. So by Bernstein's inequality (theorem C.8):

$$\forall 0 < \eta < 8e^2, \mathbb{P}\left(\left|\langle y\left|\frac{1}{n}\sum_{k=1}^n (dP_k - 1)\right|y\rangle\right| \geq \eta\right) = \mathbb{P}\left(\left|\frac{1}{n}\sum_{k=1}^n Y_k\right| \geq \eta\right) \leq 2 \exp\left(-\frac{n}{4} \frac{\eta^2}{(8e^2)^2}\right)$$

Consequently, since we may choose \mathcal{N} such that $|\mathcal{N}| \leq 7^{2d}$ (cf example A.13), we get:

$$\forall 0 < \eta < 1, \mathbb{P}\left(\left\|\frac{d}{n}\sum_{k=1}^n P_k - \sum_{k=1}^n \tilde{P}_k\right\|_{\infty} \geq \eta\right) = \mathbb{P}\left(\left\|\frac{1}{n}\sum_{k=1}^n (dP_k - 1)\right\|_{\infty} \geq \eta\right) \leq 2 \times 49^d e^{-n\eta^2/1024e^4} \quad (7)$$

Choosing, once more as in section 4.1.1, $\eta = \frac{1}{8\sqrt{18}\sqrt{d}}$ in equation 7 and combining it with equation 6, what we eventually come to is:

$$\mathbb{P}\left(\forall \Delta \in S_{\|\cdot\|_2(1)}^d, \sum_{k=1}^n |\text{Tr}(\tilde{P}_k \Delta)| \in \left[\frac{1}{8\sqrt{18}}; \frac{15}{8}\right]\right) \geq 1 - 2 \left((5\sqrt{18})^{d^2} e^{-n/36 \times 144e^4} + 49^d e^{-n/1152 \times 1024e^4 d} \right)$$

• **Step 4:** Conclusion

By homogeneity, what we obtain in the end is:

$$\mathbb{P}\left(\forall \Delta \in \mathcal{H}(\mathbb{C}^d), \frac{1}{8\sqrt{18}} \|\Delta\|_{2(1)} \leq \|\Delta\|_P \leq \frac{15}{8} \|\Delta\|_{2(1)}\right) \geq 1 - 2 \left((5\sqrt{18})^{d^2} e^{-n/36 \times 144e^4} + 49^d e^{-n/1152 \times 1024e^4 d} \right)$$

This straightforwardly implies the advertized improvement:

$$\forall 0 < \alpha < 1, \exists C_{\alpha} > 0 : n \geq C_{\alpha} d^2 \Rightarrow \mathbb{P}\left(\forall \Delta \in \mathcal{H}(\mathbb{C}^d), \frac{1}{8\sqrt{18}} \|\Delta\|_{2(1)} \leq \|\Delta\|_P \leq \frac{15}{8} \|\Delta\|_{2(1)}\right) \geq 1 - \alpha$$

4.2 Multi-partite case

Actually, the preceding result can be quite directly generalized to the multi-partite setting.

Let $\mathbb{H} \equiv \mathbb{C}^{d_1} \otimes \cdots \otimes \mathbb{C}^{d_K}$ and consider $U = U_1 \otimes \cdots \otimes U_K$ the so-called local uniform POVM on \mathbb{H} , which is the tensor product of the uniform POVMs U_i on \mathbb{C}^{d_i} , $1 \leq i \leq K$.

We define the ‘‘modified K -partite 2-norm’’ on $\mathcal{H}(\mathbb{H})$ as: $\|\Delta\|_{2(K)} := \sqrt{\sum_{I \subset \{1, \dots, K\}} \text{Tr}_{\mathbb{H} \setminus H_I} \left[(\text{Tr}_{H_I} \Delta)^2 \right]}$,

where $\mathbb{H}_{\{i_1, \dots, i_p\}} := \mathbb{C}^{d_{i_1}} \otimes \cdots \otimes \mathbb{C}^{d_{i_p}}$.

We then know from [35] that the dimension-independent norm equivalence below holds:

$$\frac{1}{\sqrt{18^K}} \|\cdot\|_{2(K)} \leq \|\cdot\|_U \leq \|\cdot\|_{2(K)} \quad (8)$$

What is more, we might show as in the one-partite case that:

$$\forall q \in \mathbb{N}, \forall \pi_1, \dots, \pi_K \in \mathfrak{S}_{2q}, \left| \text{Tr} \left(U(\pi_1) \otimes \cdots \otimes U(\pi_K) \Delta^{\otimes 2q} \right) \right| \leq \|\Delta\|_{2(K)}^{2q}$$

Hence, denoting by $D := d_1 \times \cdots \times d_K$ the dimension of \mathbb{H} , we have for all $q \in \mathbb{N}$:

$$\mathbb{E} \left[\text{Tr}(P_U \Delta) \right]^{2q} = \frac{\sum_{\pi_1, \dots, \pi_K \in \mathfrak{S}_{2q}} \text{Tr} \left(\bigotimes_{i=1}^K U(\pi_i) \Delta^{\otimes 2q} \right)}{\prod_{i=1}^K (d_i + 2q - 1) \times \cdots \times d_i} \leq \frac{[(2q)!]^K}{\prod_{i=1}^K d_i^{2q}} \|\Delta\|_{2(K)}^{2q} \leq \left(\frac{(2q)^K}{D} \|\Delta\|_{2(K)} \right)^{2q}$$

(see appendix E for a complete proof and additional comments on that matter)

Analogously to the one-partite case, this implies that:

- For any $\Delta \in S_{\|\cdot\|_{2(K)}}^D$, $X := D|\text{Tr}(P_U \Delta)| - \|\Delta\|_U$ is a centered ψ_1 random variable with ψ_1 -norm satisfying $\|X\|_{\psi_1} \leq 2e^2 \times (2^K + 1)$.
- For any $y \in S_2^D(\mathbb{C})$, $Y := D\langle y|P_U|y\rangle - 1$ is a centered ψ_1 random variable with ψ_1 -norm satisfying $\|Y\|_{\psi_1} \leq 2e^2 \times (2^K \sqrt{2} + 1)$.

Thus, following step by step the exact same lines as in section 4.1.2, we may, for each $1 \leq i \leq K$, draw $\{P_k^{(i)}, 1 \leq k \leq n\}$, independent random rank-1 projectors with the same probability distribution as P_{U_i} , then set $S_i := \sum_{k=1}^n P_k^{(i)}$ and at last $\tilde{P}_k^{(i)} := S_i^{-1/2} P_k^{(i)} S_i^{-1/2}$, $1 \leq k \leq n$. The POVM we next look at is $P := \{\tilde{P}_k^{(1)} \otimes \cdots \otimes \tilde{P}_k^{(K)}, 1 \leq k \leq n\}$, random POVM on \mathbb{H} made of n random rank-1 operators. And the final result we eventually come to is:

$$\mathbb{P} \left(\forall \Delta \in \mathcal{H}(\mathbb{C}^D), \frac{1}{8\sqrt{18^K}} \|\Delta\|_{2(K)} \leq \|\Delta\|_P \leq \frac{15}{8} \|\Delta\|_{2(K)} \right) \geq 1 - 2 \left((5\sqrt{18^K})^{D^2} e^{-\frac{n}{9K576e^4}} + 49^D e^{-\frac{n}{288K4096e^4 D}} \right)$$

Which leads to a similar statement as in the one-partite case:

$$\forall 0 < \alpha < 1, \exists C_\alpha > 0 : n \geq C_\alpha D^2 \Rightarrow \mathbb{P} \left(\forall \Delta \in \mathcal{H}(\mathbb{C}^D), \frac{1}{8\sqrt{18^K}} \|\Delta\|_{2(K)} \leq \|\Delta\|_P \leq \frac{15}{8} \|\Delta\|_{2(K)} \right) \geq 1 - \alpha$$

Remark 4.2 As drawn to attention by remark 4.1, the tensor product of the uniform POVMs on $\mathbb{C}^{d_1}, \dots, \mathbb{C}^{d_K}$ is the local POVM with the ‘‘best’’ discriminating power on $\mathbb{C}^{d_1} \otimes \cdots \otimes \mathbb{C}^{d_K}$. That is why it is of practical interest to have at hand an ‘‘almost as efficient’’ POVM that is ‘‘easily implementable’’ (which is the case when it may be constructed by picking up ‘‘few’’ projectors ‘‘at random’’).

A t -design POVM M on \mathbb{C}^d is by definition a POVM on \mathbb{C}^d which behaves ‘‘up to a certain extent quantified by t ’’ as the uniform one: $M := (dp_x P_x)_{x \in \mathcal{X}}$ with $(p_x)_{x \in \mathcal{X}}$ a probability distribution and $(P_x)_{x \in \mathcal{X}}$ rank-1 projectors on \mathbb{C}^d such that $\sum_{x \in \mathcal{X}} p_x P_x^{\otimes t} = \int_{|\psi\rangle \in S_2^d(\mathbb{C})} |\psi\rangle\langle\psi|^{\otimes t} d\psi = \frac{1}{d \times \cdots \times (d+t-1)} \sum_{\sigma \in \mathfrak{S}_t} U(\sigma)$.

The uniform POVM is thus an ∞ -design POVM. The reader is referred, for instance, to [36] or [37] for much more on that extensively studied theory (both ‘‘classically’’ and ‘‘quantumly’’).

It was shown in [35] that it is actually sufficient for a local POVM M on $\mathbb{C}^{d_1} \otimes \dots \otimes \mathbb{C}^{d_K}$ to be a tensor product of 4-design POVMs on $\mathbb{C}^{d_1}, \dots, \mathbb{C}^{d_K}$ so that the norm equivalence from equation 8 hold: $\frac{1}{\sqrt{18}^K} \|\cdot\|_{2(K)} \leq \|\cdot\|_M \leq \|\cdot\|_{2(K)}$. This result found diverse applications. It was notably used in [45] to describe an algorithm that would decide in a quasipolynomial time whether a bi-partite state is separable or whether it is “far away” from the set of separable states. It was also used in [42] to put bounds on the possibility of “compressing” quantum states into smaller dimension ones.

Now, there are two main problems that one might encounter following this exact and deterministic approach: firstly, a 4-design POVM on \mathbb{C}^d must have at least $\Omega(d^4)$ outcomes, and secondly, no explicit constructions of such POVMs are known. Hence enlightened perhaps’ the benefit of the approximate and probabilistic point of view adopted here.

5 Sets of POVMs with minimal cardinality whose measurement norm approximates the one of the set of all POVMs

The issue we focus on in this section is the one of determining how many distinct POVMs a set \mathbf{M} of POVMs on \mathbb{C}^d must contain in order for its measurement norm to approximate the one of the set \mathbf{ALL} of all POVMs on \mathbb{C}^d (in the sense that: $\exists 0 < \lambda < 1 : \lambda \|\cdot\|_{\mathbf{ALL}} \leq \|\cdot\|_{\mathbf{M}} \leq \|\cdot\|_{\mathbf{ALL}}$).

In return, we will wonder if it is possible to exhibit sets of POVMs on \mathbb{C}^d containing this minimal number of distinct POVMs and whose measurement norm approximates the one of the set of all POVMs on \mathbb{C}^d .

Before getting at the heart of our problem, let us introduce a few general notions we will later need.

Let $d \in \mathbb{N}^*$. For all $0 \leq k \leq d$, we denote by $G_{d,k}$ the so-called *Grassmannian* of dimension- k subspaces of \mathbb{C}^d (over \mathbb{C}). It is a manifold of dimension $k(d-k)$ over \mathbb{C} , and $2k(d-k)$ over \mathbb{R} .

For all $1 \leq p \leq +\infty$, $\bigsqcup_{0 \leq k \leq d} G_{d,k}$ may be equipped with the metric d_p defined by:

$\forall E, F \in \bigsqcup_{0 \leq k \leq d} G_{d,k}$, $d_p(E, F) := \|P_E - P_F\|_p$, where P_E and P_F are the orthogonal projectors onto E and F respectively.

Each $(G_{d,k}, d_p)$, $0 \leq k \leq d$, $1 \leq p \leq +\infty$, is then a compact manifold of diameter:

$$D_p(G_{d,k}) := \sup_{E, F \in G_{d,k}} d_p(E, F) = \begin{cases} (2k)^{1/p} & \text{if } k \leq \frac{d}{2} \\ (2(d-k))^{1/p} & \text{if } k \geq \frac{d}{2} \end{cases}$$

Let $0 \leq k \leq \frac{d}{2}$, $1 \leq p \leq +\infty$ and $0 < \epsilon < (2k)^{1/p}$. Denote by $N(G_{d,k}, d_p, \epsilon)$ the minimal cardinality of an ϵ -net and by $K(G_{d,k}, d_p, \epsilon)$ the maximal cardinality of an ϵ -separated set for d_p within $G_{d,k}$ (see appendix A.4 for precise definitions). We have the important following fact (proved in [21]):

There exist universal constants $0 < c < c'$ (independent of d , k , p and ϵ) such that:

$$\left(c \frac{(2k)^{1/p}}{\epsilon} \right)^{2k(d-k)} \leq N(G_{d,k}, d_p, \epsilon) \leq K(G_{d,k}, d_p, \epsilon) \leq \left(c' \frac{(2k)^{1/p}}{\epsilon} \right)^{2k(d-k)} \quad (9)$$

5.1 Set of 2-outcome projective POVMs

Considering $d+1$ instead of d if need be, we might assume without loss of generality that d is even, and look first at $G_{d,d/2}$ the Grassmannian of dimension- $d/2$ subspaces of \mathbb{C}^d (over \mathbb{C}). Applying the general result provided by equation 9 to this special case yields:

$$\forall 1 \leq p \leq +\infty, \forall 0 < \epsilon < d^{1/p}, \left(c \frac{d^{1/p}}{\epsilon} \right)^{d^2/2} \leq N(G_{d,d/2}, d_p, \epsilon) \leq K(G_{d,d/2}, d_p, \epsilon) \leq \left(c' \frac{d^{1/p}}{\epsilon} \right)^{d^2/2}$$

Let $\{E_\alpha, \alpha \in A\} \subset G_{d,d/2}$ be a set of $|A|$ dimension- $d/2$ subspaces of \mathbb{C}^d .

We consider $\mathbf{M} := \{P_\alpha := (P_{E_\alpha}, P_{E_\alpha^\perp}), \alpha \in A\}$, set of $|A|$ 2-outcome POVMs on \mathbb{C}^d whose operators

are rank- $d/2$ projectors, and we assume that it is such that:

$$\exists 0 < \lambda < 1 : \forall \Delta \in \mathcal{H}(\mathbb{C}^d), \|\Delta\|_{\mathbf{M}} \geq \lambda \|\Delta\|_{\mathbf{ALL}} = \lambda \|\Delta\|_1$$

Let $0 < \epsilon < \sqrt{d}$ and denote by $\mathcal{M}_\epsilon := \{F_\beta, \beta \in B\}$ a maximal ϵ -separated set for d_2 within $G_{d,d/2}$.

By assumption on \mathbf{M} , we have in particular: $\forall \beta \in B, \|\Delta_{F_\beta}\|_{\mathbf{M}} \geq \lambda \|\Delta_{F_\beta}\|_1$, where $\Delta_{F_\beta} := P_{F_\beta} - P_{F_\beta^\perp}$, that is: $\forall \beta \in B, \exists \alpha \in A : \|\Delta_{F_\beta}\|_{P_\alpha} \geq \lambda \|\Delta_{F_\beta}\|_1$.

Yet, on the one hand: $\|\Delta_{F_\beta}\|_1 = \text{Tr}|P_{F_\beta} - P_{F_\beta^\perp}| = d$.

And on the other: $\|\Delta_{F_\beta}\|_{P_\alpha} = |\text{Tr}[(P_{F_\beta} - P_{F_\beta^\perp})P_{E_\alpha}]| + |\text{Tr}[(P_{F_\beta} - P_{F_\beta^\perp})P_{E_\alpha^\perp}]| = 2|\text{Tr}[P_{E_\alpha}(1 - 2P_{F_\beta})]|$.

Now: $[d_2(E_\alpha, F_\beta)]^2 = \|P_{E_\alpha} - P_{F_\beta}\|_2^2 = \text{Tr}(P_{E_\alpha}^2) + \text{Tr}(P_{F_\beta}^2) - 2\text{Tr}(P_{E_\alpha}P_{F_\beta}) = \text{Tr}[P_{E_\alpha}(1 - 2P_{F_\beta})] + \frac{d}{2}$.

So: $\|\Delta_{F_\beta}\|_{P_\alpha} = 2 \left| \frac{d}{2} - [d_2(E_\alpha, F_\beta)]^2 \right|$.

And since $[d_2(E_\alpha^\perp, F_\beta)]^2 = d - [d_2(E_\alpha, F_\beta)]^2$, we actually have: $\|\Delta_{F_\beta}\|_{P_\alpha} = \begin{cases} 2 \left(\frac{d}{2} - [d_2(E_\alpha, F_\beta)]^2 \right) \\ 2 \left(\frac{d}{2} - [d_2(E_\alpha^\perp, F_\beta)]^2 \right) \end{cases}$,

depending on which one of those two quantities is positive and which one is negative.

We must therefore have, either $d_2(E_\alpha, F_\beta) \leq \sqrt{\frac{1-\lambda}{2}d}$, or $d_2(E_\alpha^\perp, F_\beta) \leq \sqrt{\frac{1-\lambda}{2}d}$.

But by ϵ -separation of \mathcal{M}_ϵ , a ball of radius $\frac{\epsilon}{2}$ for d_2 centered at some given point of $G_{d,d/2}$ contains at most one point of \mathcal{M}_ϵ . Hence, choosing $\epsilon = \sqrt{2(1-\lambda)d}$, we get by what precedes that, for each $\alpha \in A$, there exist at most two $\beta \in B$ such that $\|\Delta_{F_\beta}\|_{P_\alpha} \geq \lambda \|\Delta_{F_\beta}\|_1$ (because the ball of radius $\sqrt{\frac{1-\lambda}{2}d}$ for d_2 centered at E_α contains at most one point of $\mathcal{M}_{\sqrt{2(1-\lambda)d}}$, and similarly for E_α^\perp).

This is enlightened perhaps' by figure 5 below.

Consequently, we must have: $|A| \times 2 \geq |\mathcal{M}_{\sqrt{2(1-\lambda)d}}|$.

Now, by maximality of $\mathcal{M}_{\sqrt{2(1-\lambda)d}}$, equation 9 implies: $|\mathcal{M}_{\sqrt{2(1-\lambda)d}}| \geq \left(\frac{c \frac{\sqrt{d}}{\sqrt{2(1-\lambda)d}}}{\frac{c}{\sqrt{2(1-\lambda)d}}} \right)^{d^2/2} = \left(\frac{c}{\sqrt{2(1-\lambda)}} \right)^{d^2/2}$.

So in the end, we get the following lower-bound on the cardinality of the considered set of POVMs:

$$|A| \geq \frac{1}{2} \left(\frac{c}{\sqrt{2(1-\lambda)}} \right)^{d^2/2}.$$

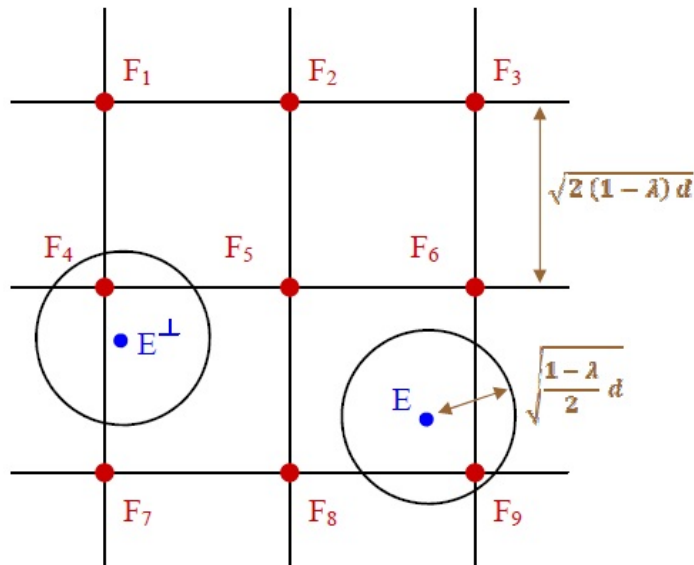


Figure 5: $(F_\beta)_{\beta \in B}$ a $\sqrt{2(1-\lambda)d}$ -separated set for d_2 within $G_{d,d/2}$: for any $\tilde{E} \in G_{d,d/2}$, there exists at most one $\beta \in B$ such that $d_2(\tilde{E}, F_\beta) \leq \sqrt{\frac{1-\lambda}{2}d}$

Conversely, let us consider as set of 2-outcome POVMs on \mathbb{C}^d $\mathbf{M} := \{P_\alpha := (P_{E_\alpha}, P_{E_\alpha^\perp}), \alpha \in A\}$ where $\mathcal{M}_{\sqrt{2(1-\lambda)d}} := \{E_\alpha, \alpha \in A\}$ is a minimal $\sqrt{2(1-\lambda)d}$ -net for d_2 within $G_{d,d/2}$.

By minimality of $\mathcal{M}_{\sqrt{2(1-\lambda)d}}$, we have by equation 9: $|A| \leq \left(c' \frac{\sqrt{d}}{\sqrt{2(1-\lambda)d}} \right)^{d^2/2} = \left(\frac{c'}{\sqrt{2(1-\lambda)}} \right)^{d^2/2}$.

And by $\sqrt{2(1-\lambda)d}$ -covering of $\mathcal{M}_{\sqrt{2(1-\lambda)d}}$: $\forall F \in G_{d,d/2}, \exists \alpha \in A : d_2(F, E_\alpha) \leq \sqrt{2(1-\lambda)d}$, which implies, following the exact same lines as above backwards, that: $\|\Delta_F\|_{P_\alpha} \geq \lambda \|\Delta_{F_\beta}\|_1$, and thus that: $\|\Delta_F\|_{\mathbf{M}} \geq \lambda \|\Delta_F\|_1$.

Hence, to sum things up, there are basically two things that we have shown at that point:

- If \mathbf{M} is a set of POVMs on \mathbb{C}^d which is composed of 2-outcome POVMs of the form (P_E, P_{E^\perp}) with $E \in G_{d,d/2}$, it must have at least $\left(\frac{C}{\sqrt{1-\lambda}} \right)^{d^2/2}$ elements in order to be such that $\forall F \in G_{d,d/2}, \|P_F - P_{F^\perp}\|_{\mathbf{M}} \geq \lambda \|P_F - P_{F^\perp}\|_1$.
- There exists a set $\overline{\mathbf{M}}$ of POVMs on \mathbb{C}^d which is composed of 2-outcome POVMs of the form (P_E, P_{E^\perp}) with $E \in G_{d,d/2}$, which has less than $\left(\frac{C'}{\sqrt{1-\lambda}} \right)^{d^2/2}$ elements, and which is such that $\forall F \in G_{d,d/2}, \|P_F - P_{F^\perp}\|_{\overline{\mathbf{M}}} \geq \lambda \|P_F - P_{F^\perp}\|_1$.

And this result may in fact be generalized to subspaces of \mathbb{C}^d of any dimension.

Let indeed $1 \leq k \leq \frac{d}{2}$. For all $E, F \in G_{d,k}$, we have:

$$\begin{aligned} \|P_F - P_{F^\perp}\|_{(P_E, P_{E^\perp})} &= |\text{Tr}[P_E(1 - 2P_F)]| + |\text{Tr}[P_E(1 - 2P_{F^\perp})]| + d - 2k \\ &= |[d_2(E, F)]^2 - k| + |[d_2(E, F)]^2 + d - 3k| \end{aligned}$$

Now, there exists \tilde{E}^\perp a k -dimensional subspace of E^\perp such that $[d_2(\tilde{E}^\perp, F)]^2 = 2k - [d_2(E, F)]^2$.

So what we eventually come to is: $\|P_F - P_{F^\perp}\|_{(P_E, P_{E^\perp})} = \begin{cases} 2 \left((d-k) - [d_2(E, F)]^2 \right) \\ 2 \left((d-k) - [d_2(\tilde{E}^\perp, F)]^2 \right) \end{cases}$, depending on which one of those two quantities is positive and which one is negative.

Hence: $\|P_F - P_{F^\perp}\|_{(P_E, P_{E^\perp})} \geq \lambda \|P_F - P_{F^\perp}\|_1 \Leftrightarrow \begin{cases} d_2(E, F) \leq \sqrt{(d-k) - \lambda \frac{d}{2}} \\ \text{or} \\ d_2(\tilde{E}^\perp, F) \leq \sqrt{(d-k) - \lambda \frac{d}{2}} \end{cases}$

Consequently, considering, similarly to what was done in the particular case $k = \frac{d}{2}$ above, either a maximal $\sqrt{2(2(d-k) - \lambda d)}$ -separated set or a minimal $\sqrt{2(2(d-k) - \lambda d)}$ -net for d_2 within $G_{d,k}$, we get the following result:

- If \mathbf{M}_k is a set of POVMs on \mathbb{C}^d which is composed of 2-outcome POVMs of the form (P_E, P_{E^\perp}) with $E \in G_{d,k}$, it must have at least $\left(\frac{C}{\sqrt{\frac{d-k}{k} - \lambda \frac{d/2}{k}}} \right)^{2k(d-k)}$ elements in order to be such that $\forall F \in G_{d,k}, \|P_F - P_{F^\perp}\|_{\mathbf{M}} \geq \lambda \|P_F - P_{F^\perp}\|_1$.
- There exists a set $\overline{\mathbf{M}}_k$ of POVMs on \mathbb{C}^d which is composed of 2-outcome POVMs of the form (P_E, P_{E^\perp}) with $E \in G_{d,k}$, which has less than $\left(\frac{C'}{\sqrt{\frac{d-k}{k} - \lambda \frac{d/2}{k}}} \right)^{2k(d-k)}$ elements, and which is such that $\forall F \in G_{d,k}, \|P_F - P_{F^\perp}\|_{\overline{\mathbf{M}}_k} \geq \lambda \|P_F - P_{F^\perp}\|_1$.

As a consequence, $\overline{\mathbf{M}} := \bigsqcup_{1 \leq k \leq \frac{d}{2}} \overline{\mathbf{M}}_k$ is a set of POVMs on \mathbb{C}^d which is composed of 2-outcome POVMs of the form (P_E, P_{E^\perp}) with E subspace of \mathbb{C}^d , and which is such that:

On the one hand, $|\overline{\mathbf{M}}| \leq \sum_{1 \leq k \leq \frac{d}{2}} \left(\frac{C'}{\sqrt{\frac{d-k}{k} - \lambda \frac{d/2}{k}}} \right)^{2k(d-k)} \leq \frac{d}{2} \left(\frac{C'}{\sqrt{1-\lambda}} \right)^{d^2/2} \leq \left(\frac{C}{\sqrt{1-\lambda}} \right)^{d^2/2}$.

And on the other, for any subspace F of \mathbb{C}^d , $\|P_F - P_{F^\perp}\|_{\overline{\mathbf{M}}} \geq \lambda \|P_F - P_{F^\perp}\|_1$.

5.2 Set of general POVMs

We now consider $\mathbf{M} := \left\{ M_\alpha := \left(M_{i_\alpha}^{(\alpha)} \right)_{i_\alpha \in I_\alpha}, \alpha \in A \right\}$ a set of $|A|$ general POVMs on \mathbb{C}^d , and we assume, just as was done in section 5.1, that it is such that:

$$\exists 0 < \lambda < 1 : \forall \Delta \in \mathcal{H}(\mathbb{C}^d), \|\Delta\|_{\mathbf{M}} \geq \lambda \|\Delta\|_{\mathbf{ALL}} = \lambda \|\Delta\|_1$$

First of all, let us notice that for any $0 \leq M \leq 1$ and any $E \in G_{d,d/2}$, defining the traceless Hermitian $\Delta_E := P_E - P_{E^\perp} = 1 - 2P_E$, we have:

$$\|\Delta_E\|_{(M,1-M)} = |\text{Tr}[\Delta_E M]| + |\text{Tr}[\Delta_E(1-M)]| = \begin{cases} 2\text{Tr}[(1-2P_E)M] \\ 2\text{Tr}[(1-2P_E)(1-M)] \end{cases}, \text{ depending on which}$$

one of these two quantities is positive and which one is negative.

Now: $\|P_E - M\|_2^2 = \text{Tr}(P_E^2) + \text{Tr}(M^2) - 2\text{Tr}(P_E M) \leq \text{Tr}P_E + \text{Tr}M - 2\text{Tr}(P_E M) = \frac{d}{2} + \text{Tr}[(1-2P_E)M]$.

And similarly: $\|P_E - (1-M)\|_2^2 \leq \frac{d}{2} + \text{Tr}[(1-2P_E)(1-M)]$.

Consequently, either $\|P_E - M\|_2 \leq \sqrt{\frac{d}{2} - \frac{1}{2}\|\Delta_E\|_{(M,1-M)}}$ or $\|P_E - (1-M)\|_2 \leq \sqrt{\frac{d}{2} - \frac{1}{2}\|\Delta_E\|_{(M,1-M)}}$.

With this preliminary result in mind, let us turn back to our initial concern.

Consider $\mathcal{M}_{\sqrt{8(1-\lambda)d}} := \{E_\beta, \beta \in B\}$ a maximal $\sqrt{8(1-\lambda)d}$ -separated set for d_2 within $G_{d,d/2}$.

By assumption on \mathbf{M} , we have in particular:

$$\forall \beta \in B, \exists \alpha \in A, \exists J_\alpha \subset I_\alpha : \|\Delta_{E_\beta}\|_{(M_{J_\alpha}^{(\alpha)}, 1-M_{J_\alpha}^{(\alpha)})} \geq \lambda \|\Delta_{E_\beta}\|_1, \text{ where } M_{J_\alpha}^{(\alpha)} := \sum_{\alpha \in J_\alpha} M_{i_\alpha}^{(\alpha)}.$$

As pointed out above, since $\|\Delta_{E_\beta}\|_1 = d$, this implies that either $\|P_{E_\beta} - M_{J_\alpha}^{(\alpha)}\|_2 \leq \sqrt{\frac{1-\lambda}{2}d}$ or $\|P_{E_\beta} - (1 - M_{J_\alpha}^{(\alpha)})\|_2 \leq \sqrt{\frac{1-\lambda}{2}d}$.

But by $\sqrt{8(1-\lambda)d}$ -separation of $\mathcal{M}_{\sqrt{8(1-\lambda)d}}$, the ball of radius $\sqrt{\frac{1-\lambda}{2}d}$ for $\|\cdot\|_2$ centered at $M_{J_\alpha}^{(\alpha)}$ or at $1 - M_{J_\alpha}^{(\alpha)}$ contains at most one point of $\mathcal{M}_{\sqrt{8(1-\lambda)d}}$. Hence, for each $\alpha \in A$ and each $J_\alpha \subset I_\alpha$, there exist at most 2 $\beta \in B$ such that $\|\Delta_{E_\beta}\|_{(M_{J_\alpha}^{(\alpha)}, 1-M_{J_\alpha}^{(\alpha)})} \geq \lambda \|\Delta_{E_\beta}\|_1$.

Now, consider one given POVM $M = (M_i)_{i \in I}$ on \mathbb{C}^d and assume that it is such that:

$$\exists I' \subset I, |I'| = n : \forall J \subset I', \exists \beta \in B : \|\Delta_{E_\beta}\|_{(M_J, 1-M_J)} \geq \lambda \|\Delta_{E_\beta}\|_1$$

By what precedes, this implies that the 2^n distinct $M_J, J \subset I'$, are each $\sqrt{\frac{1-\lambda}{2}d}$ -close of a given P_{E_β} .

And since those are $\sqrt{8(1-\lambda)d}$ -separated, this entails in turn that the 2^n distinct $M_J, J \subset I'$, are $\sqrt{2(1-\lambda)d}$ -separated.

Subsequently, the symmetric convex body $K_{M'} := \text{Conv}\{2M_J - 1, J \subset I'\} \subset \mathcal{H}(\mathbb{C}^d)$ is such that $K_{M'} = \text{Conv}(T)$ with $|T| = 2^n$ and $\forall X, Y \in T, \|X - Y\|_2 \geq \sqrt{2(1-\lambda)d}$.

Therefore, by theorem B.4: $w(K_{M'}) \gtrsim \sqrt{2(1-\lambda)d} \sqrt{\frac{n}{d^2}} = \sqrt{2(1-\lambda)} \sqrt{\frac{n}{d}}$.

Yet, it also holds that: $w(K_{M'}) \lesssim 1$ (cf remark 2.4). Hence necessarily: $n \lesssim d$.

What we have thus shown is that, for each $\alpha \in A$, there exist at most C^d $\beta \in B$ such that $\|\Delta_{E_\beta}\|_{M_\alpha} \geq \lambda \|\Delta_{E_\beta}\|_1$.

As a consequence, we must have: $|A| \times C^d \geq \left| \mathcal{M}_{\sqrt{8(1-\lambda)d}} \right|$.

Now, by maximality of $\mathcal{M}_{\sqrt{8(1-\lambda)d}}$, equation 9 implies: $|\mathcal{M}_{\sqrt{2(1-\lambda)d}}| \geq \left(c \frac{\sqrt{d}}{\sqrt{8(1-\lambda)d}} \right)^{d^2/2} = \left(\frac{c}{\sqrt{8(1-\lambda)}} \right)^{d^2/2}$.
So in the end, we get the following lower-bound on the cardinality of the considered set of POVMs:

$$|A| \geq \frac{1}{C^d} \left(\frac{c}{\sqrt{8(1-\lambda)}} \right)^{d^2/2} \geq \left(\frac{\tilde{c}}{\sqrt{1-\lambda}} \right)^{d^2/2}.$$

Let us summarize: a set of POVMs on \mathbb{C}^d whose associated measurement norm would be, on any Hermitian, at least, say, half the one associated with the set of all POVMs on \mathbb{C}^d has to be composed of at least C^{d^2} distinct POVMs (and this whatever the number and the type of the operators composing each POVM).

Remark 5.1 *It is in fact possible to come to a quite similar result from a rather different (and when all is said and done, probably more straightforward) approach.*

Let us denote by $\mathcal{HU}(\mathbb{C}^d)$ the set of Hermitian unitaries on \mathbb{C}^d . It should be noted that:

$$U \in \mathcal{HU}(\mathbb{C}^d) \Leftrightarrow U = \frac{1}{2}(1 + P) \text{ with } P \text{ an orthogonal projector on } \mathbb{C}^d$$

$\mathcal{HU}(\mathbb{C}^d)$ may thus be identified with $\bigsqcup_{0 \leq k \leq d} G_{d,k}$. This implies that for all $1 \leq p \leq +\infty$ and $0 < \epsilon < 2^{1/p}$, one can take as ϵ -net (or as ϵ -separated set) for $\|\cdot\|_p$ within $\mathcal{HU}(\mathbb{C}^d)$: $\bigsqcup_{1 \leq k \leq d-1} \left\{ \frac{1}{2}(1 + P_E), E \in \mathcal{M}_\epsilon(k) \right\}$, where for each $1 \leq k \leq d-1$, $\mathcal{M}_\epsilon(k)$ is an ϵ -net (or an ϵ -separated set) for d_p within $G_{d,k}$.

Therefore, by equation 9, it holds regarding the entropy numbers of $\mathcal{HU}(\mathbb{C}^d)$ that:

There exist universal constants $0 < c < c'$ (independent of d, k, p and ϵ) such that:

$$2 \sum_{k=1}^{d/2} \left(c \frac{(2k)^{1/p}}{\epsilon} \right)^{2k(d-k)} \leq N(\mathcal{HU}(\mathbb{C}^d), \|\cdot\|_p, \epsilon) \leq K(\mathcal{HU}(\mathbb{C}^d), \|\cdot\|_p, \epsilon) \leq 2 \sum_{k=1}^{d/2} \left(c' \frac{(2k)^{1/p}}{\epsilon} \right)^{2k(d-k)}$$

Hence, there actually exist universal constants $0 < \bar{c} < \bar{c}'$ (independent of d, p and ϵ) such that:

$$\left(\frac{\bar{c} d^{1/p}}{\epsilon} \right)^{d^2/2} \leq N(\mathcal{HU}(\mathbb{C}^d), \|\cdot\|_p, \epsilon) \leq K(\mathcal{HU}(\mathbb{C}^d), \|\cdot\|_p, \epsilon) \leq \left(\frac{\bar{c}' d^{1/p}}{\epsilon} \right)^{d^2/2} \quad (10)$$

With this point in mind, we assume as before that we have a set $\mathbf{M} := \{M_\alpha, \alpha \in A\}$ of $|A|$ POVMs on \mathbb{C}^d which is such that: $\lambda K_{\text{ALL}} \subset K_{\mathbf{M}}$, i.e. $\lambda B_{\|\cdot\|_\infty}^d \subset \text{Conv} \left(\bigcup_{\alpha \in A} K_{M_\alpha} \right)$, for some $0 < \lambda < 1$.

We next fix $\epsilon > 0$ and consider $\mathcal{M}_\epsilon := \{U_\beta, \beta \in B\}$ a maximal ϵ -separated set for $\|\cdot\|_2$ within $\mathcal{HU}(\mathbb{C}^d)$. By maximality of \mathcal{M}_ϵ equation 10 entails that: $|B| \geq \left(\frac{\bar{c}\sqrt{d}}{\epsilon} \right)^{d^2/2}$.

What is more, by assumption on \mathbf{M} : $\lambda \mathcal{M}_\epsilon \subset \lambda B_{\|\cdot\|_\infty}^d \subset K_{\mathbf{M}}$.

And by extremality of the unitaries in $B_{\|\cdot\|_\infty}^d$, this in fact implies that: $\lambda \mathcal{M}_\epsilon \subset \bigcup_{\alpha \in A} K_{M_\alpha}$.

Now, let $\alpha \in A$ and suppose that there exist n_α distinct $\beta \in B$ such that $\lambda U_\beta \in K_{M_\alpha}$.

By ϵ -separation of \mathcal{M}_ϵ : $\forall \beta \neq \beta' \in B, \|\lambda U_\beta - \lambda U_{\beta'}\|_2 \geq \lambda \epsilon$. So by theorem B.4: $w(K_{M_\alpha}) \gtrsim \lambda \epsilon \frac{\sqrt{\log n_\alpha}}{d}$. Since it also stands by remark 2.4 that: $w(K_{M_\alpha}) \lesssim 1$, it follows that: $n_\alpha \leq C^{d^2/\lambda^2 \epsilon^2}$.

Choosing $\epsilon = \sqrt{\frac{d}{\lambda}}$, what we come to in the end is that, on the one hand: $|B| \geq \left(\bar{c}\sqrt{\lambda} \right)^{d^2/2}$, and on the other: for each $\alpha \in A$, there are at most C^d distinct $\beta \in B$ such that $\lambda U_\beta \in K_{M_\alpha}$.

Hence necessarily: $|A| \geq \frac{\left(\bar{c}\sqrt{\lambda} \right)^{d^2/2}}{C^d} \geq \left(\tilde{c}\sqrt{\lambda} \right)^{d^2/2}$.

6 Conclusion and open questions

What spurred us into the investigation carried on in section 3 was the will to get quantitative estimates on the discriminating power of some classes of locally restricted measurements on “large” composite systems. The results we obtained, namely that, on $(\mathbb{C}^d)^{\otimes K}$, $w(K_{\mathbf{PPT}}) \simeq d^{K/2}$ and $w(K_{\mathbf{SEP}}) \simeq d^{1/2}$, are valid for a fixed K and $d \rightarrow +\infty$, i.e. for a “small” number of “large” subsystems. Now, the opposite setting in which the local dimension d is fixed and the number of local parties $K \rightarrow +\infty$ might be equally naturally considered: it corresponds to the situation of a “large” number of “small” subsystems. It is a weakness of our approach that it only allows us to deal with one of the two “regular” high-dimensional multi-partite systems one could think of.

Besides, the reformulation we get in terms of the “typical” value of $\|\cdot\|_{\mathbf{PPT}}$ and $\|\cdot\|_{\mathbf{SEP}}$ states that for $\Delta \sim \mathcal{U}(S_{\|\cdot\|_2}^{d^K})$, $\|\Delta\|_{\mathbf{PPT}} \simeq d^{K/2}$ and $\|\Delta\|_{\mathbf{SEP}} \simeq \sqrt{d}$ with high probability. However, the initial motivation for taking a closer look at those measurement norms was the discrimination task described in section 2.2. Our results should therefore be translated now into statements on the “typical” value of the biases $\|\frac{1}{2}\rho - \frac{1}{2}\sigma\|_{\mathbf{PPT}}$ and $\|\frac{1}{2}\rho - \frac{1}{2}\sigma\|_{\mathbf{SEP}}$ for “random” states ρ and σ (see appendix D.2 for more details on how to define rigorously what a “random” state could be).

In section 4.1, it was shown that the uniform POVM on \mathbb{C}^d could be emulated by drawing $\Omega(d^2)$ uniformly distributed rank-1 projectors. From there, one could then legitimately ask the following question: given a rank-1 POVM which already approximates the uniform POVM, how many rank-1 projectors should be sampled from the corresponding probability distribution in order, once again, to emulate the uniform POVM? One would expect that $\Omega(d^2)$ rank-1 projectors would not be enough anymore in that case, but perhaps’ $\Omega(d^2(\log d)^\alpha)$ or $\Omega(d^{2+\epsilon})$... The reason why this wonder could be relevant is that POVMs with a finite number of outcomes are known to “behave almost as well as” the uniform POVM, for instance 4-design POVMs (the reader is referred to [38] for definitions and one-partite results, to [34] and [35] for multi-partite generalizations).

Regarding now section 4.2 and the approximation of the local uniform POVM in the multi-partite case, there is something else that might be worth pointing at: the random approximating POVM which is constructed there is a separable POVM but not a local POVM (in the sense defined in section 3.1). Indeed, on $\mathbb{C}^{d_1} \otimes \dots \otimes \mathbb{C}^{d_K} \equiv \mathbb{C}^D$, it is of the form $P_{\mathbf{SEP}} = \{\tilde{P}_k^{(1)} \otimes \dots \otimes \tilde{P}_k^{(K)}, 1 \leq k \leq CD^2\}$ and not of the form $P_{\mathbf{LO}} = \{\tilde{P}_{k_1}^{(1)} \otimes \dots \otimes \tilde{P}_{k_K}^{(K)}, 1 \leq k_i \leq Cd_i^2, 1 \leq i \leq K\}$. Both types of POVMs are made of $\Omega(D^2)$ random operators, but the problem in the latter case is that those random operators are not independent anymore, which forbids an application of the “usual” large deviation estimates without change...

As for the approximation of the set of all POVMs on \mathbb{C}^d , it was proved in section 5 that it requires C^{d^2} POVMs. One could then naturally wonder how many POVMs would be needed to approximate many other sets of POVMs (for instance the sets of local, separable or positive under partial transpose POVMs on $(\mathbb{C}^d)^{\otimes K}$).

Appendices

A Convex geometry and functional analysis

A.1 Duality between norms and convex bodies

The reader is referred, for instance, to [6] or [10] for a complete exposition of all the basic convex geometry notions presented succinctly in this section.

Let $(\mathbb{H}, \|\cdot\|)$ be a real Hilbert space (with the norm $\|\cdot\|$ deriving from an inner product $\langle \cdot | \cdot \rangle$ on \mathbb{H}). For any norm η on \mathbb{H} and any $r > 0$ we will denote by $B_\eta(r) := \{x \in \mathbb{H}, \eta(x) \leq r\}$ the closed ball of radius r (centered at the origin) for η . When $r = 1$, we will generally omit it and write $B_\eta := B_\eta(1)$ to denote the closed unit ball (centered at the origin) for η .

Proposition A.1 *Let K be a symmetric convex body of \mathbb{H} with non-empty interior.*

Define its gauge or Minkowski functional $g_K : x \in \mathbb{H} \mapsto \inf \{t > 0, x \in tK\} = \inf \left\{ t > 0, \frac{1}{t}x \in K \right\}$.

Then g_K is a norm on \mathbb{H} that is such that $B_{g_K} = K$.

Proof: The subadditivity of g_K is guaranteed by the convexity of K and its homogeneity is guaranteed by the symmetry of K . Due to the fact that K is compact, g_K is additionally positive definite.

Proposition A.2 *Conversely, for any norm η on \mathbb{H} and any $r > 0$, $B_\eta(r)$ is a symmetric convex body of \mathbb{H} with non-empty interior, and $g_{B_\eta} = \eta$.*

Definition/Proposition A.3 *For all $K \subset \mathbb{H}$ we define its polar as $K^\circ := \{x \in \mathbb{H}, \forall y \in K, |\langle y|x \rangle| \leq 1\}$. If K is a symmetric convex body of \mathbb{H} , then so is K° , and $(K^\circ)^\circ = K$.*

And in such case, the following duality formulas stand:

$$\forall x \in \mathbb{H}, g_K(x) = \sup_{y \in K^\circ} |\langle y|x \rangle| \text{ and } g_{K^\circ}(x) = \sup_{y \in K} |\langle y|x \rangle|$$

Let $n \in \mathbb{N}^*$. We define the inner-product $\langle \cdot | \cdot \rangle$ on \mathbb{R}^n by:

$$\forall x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in \mathbb{R}^n, \langle x|y \rangle := \sum_{1 \leq i \leq n} x_i y_i$$

We more generally define, for all $1 \leq p \leq +\infty$, the p -norm $\|\cdot\|_p$ on \mathbb{R}^n (with associated unit ball later denoted by B_p^n) as:

$$\forall x = (x_1, \dots, x_n) \in \mathbb{R}^n, \|x\|_p := \left(\sum_{1 \leq i \leq n} |x_i|^p \right)^{1/p}, \quad 1 \leq p < +\infty, \text{ and } \|x\|_\infty := \max_{1 \leq i \leq n} |x_i|$$

In the so-called *classical* or *commutative* space $(\mathbb{R}^n, \langle \cdot | \cdot \rangle)$, the following duality holds:

$$\forall 1 \leq p, p' \leq +\infty, \frac{1}{p} + \frac{1}{p'} = 1, (B_p^n)^\circ = B_{p'}^n \quad (11)$$

Let $d \in \mathbb{N}^*$. We define the Hilbert-Schmidt inner-product $\langle \cdot | \cdot \rangle$ on the space of complex Hermitian $d \times d$ matrices $\mathcal{H}(\mathbb{C}^d) \equiv \mathbb{R}^{d^2}$ by:

$$\forall M, N \in \mathcal{H}(\mathbb{C}^d), \langle M|N \rangle := \text{Tr}(MN)$$

We more generally define, for all $1 \leq p \leq +\infty$, the Schatten p -norm $\|\cdot\|_p$ on $\mathcal{H}(\mathbb{C}^d)$ (with associated unit ball later denoted by $B_{\|\cdot\|_p}^d$) as:

$$\forall M \in \mathcal{H}(\mathbb{C}^d), \|M\|_p := (\text{Tr}|M|^p)^{1/p}, \quad 1 \leq p < +\infty, \quad \text{and} \quad \|M\|_\infty := \|M\|$$

Denoting, for each $M \in \mathcal{H}(\mathbb{C}^d)$, by $\lambda(M) = (\lambda_1(M), \dots, \lambda_d(M)) \in \mathbb{R}^d$ the real-valued d -tuple of eigenvalues of M , we have: $\forall 1 \leq p \leq +\infty, \|M\|_p = \|\lambda(M)\|_p$.

In the so-called *quantum* or *non-commutative* space $(\mathcal{H}(\mathbb{C}^d), \langle \cdot | \cdot \rangle)$, the following duality holds:

$$\forall 1 \leq p, p' \leq +\infty, \frac{1}{p} + \frac{1}{p'} = 1, \quad \left(B_{\|\cdot\|_p}^d\right)^\circ = B_{\|\cdot\|_{p'}}^d \quad (12)$$

A.2 “Classic” geometric inequalities involving volumes

In this section are exposed the most basic inequalities involving volumes. Since those really are fundamental, rather detailed proofs are for once included. The reader is referred to any standard geometric functional analysis textbook for an enlarged outline of the fruitful interplay between convex geometry and functional analysis (e.g. [7], [8] or [10]).

In the sequel, the Lebesgue measure on \mathbb{R}^n will be denoted indiscriminately by either $\text{Vol}(\cdot)$ or $|\cdot|$.

Theorem A.4 (Prekopa-Leindler functional inequality)

Let $f, g, h : \mathbb{R}^n \rightarrow [0; +\infty]$ and $0 \leq \lambda \leq 1$.

Assume that those are such that: $\forall x, y \in \mathbb{R}^n, h(\lambda x + (1 - \lambda)y) \geq f(x)^\lambda g(y)^{1-\lambda}$

Then: $\int_{\mathbb{R}^n} h \geq \left(\int_{\mathbb{R}^n} f\right)^\lambda \left(\int_{\mathbb{R}^n} g\right)^{1-\lambda}$

Proof: The preliminary observation in showing Prekopa-Leindler inequality is that for compact sets $A, B \subset \mathbb{R}, |A| + |B| \leq |A + B|$. Indeed, by translation invariance of the Lebesgue measure on \mathbb{R} , one might assume without loss of generality that $A \subset \mathbb{R}^-$ and $B \subset \mathbb{R}^+$, so that $A, B \subset A + B$ and $|A \cap B| = 0$, which implies that $|A| + |B| \leq |A + B|$.

Then comes the initialization step which consists in proving Prekopa-Leindler inequality on \mathbb{R} .

By the definition of Lebesgue’s integral itself: $\forall F : \mathbb{R} \rightarrow [0; +\infty], \int_{\mathbb{R}} F = \int_0^{+\infty} |\{F \geq a\}| da$.

Yet, if $f, g, h : \mathbb{R} \rightarrow [0; +\infty]$ and $0 \leq \lambda \leq 1$ satisfy the assumptions of Prekopa-Leindler inequality on \mathbb{R} , then: $\forall a > 0, \lambda\{f \geq a\} + (1 - \lambda)\{g \geq a\} \subset \{h \geq a\}$, so that by the preliminary observation: $\lambda|\{f \geq a\}| + (1 - \lambda)|\{g \geq a\}| \leq |\{h \geq a\}|$.

Integrating on both sides over $a > 0$ yields: $\lambda \int_{\mathbb{R}} f + (1 - \lambda) \int_{\mathbb{R}} g \leq \int_{\mathbb{R}} h$.

And by arithmetic-geometric mean inequality: $\left(\int_{\mathbb{R}} f\right)^\lambda \left(\int_{\mathbb{R}} g\right)^{1-\lambda} \leq \int_{\mathbb{R}} h$.

To finish with is the induction step on the dimension.

Assume Prekopa-Leindler inequality holds on \mathbb{R}^{n-1} and consider $f, g, h : \mathbb{R}^n \rightarrow [0; +\infty]$ and $0 \leq \lambda \leq 1$ satisfying the assumptions of Prekopa-Leindler inequality on \mathbb{R}^n .

Defining for any $F : \mathbb{R}^n \rightarrow [0; +\infty]$ and any $t \in \mathbb{R}, F_t : x \in \mathbb{R}^{n-1} \mapsto F(t, x)$, we then have:

$t = \lambda r + (1 - \lambda)s \Rightarrow \forall x, y \in \mathbb{R}^{n-1}, h_t(\lambda x + (1 - \lambda)y) \geq f_r(x)^\lambda g_s(y)^{1-\lambda}$.

So by induction hypothesis: $\int_{\mathbb{R}^{n-1}} h_t \geq \left(\int_{\mathbb{R}^{n-1}} f_r\right)^\lambda \left(\int_{\mathbb{R}^{n-1}} g_s\right)^{1-\lambda}$.

Hence, by Prekopa-Leindler inequality on $\mathbb{R}: \int_{\mathbb{R}} \left(\int_{\mathbb{R}^{n-1}} h_t\right) \geq \left(\int_{\mathbb{R}} \left(\int_{\mathbb{R}^{n-1}} f_r\right)\right)^\lambda \left(\int_{\mathbb{R}} \left(\int_{\mathbb{R}^{n-1}} g_s\right)\right)^{1-\lambda}$, which is precisely Prekopa-Leindler inequality on \mathbb{R}^n .

Theorem A.5 (Brunn-Minkowski geometric inequality)

- *Additive dimensional form:* For any non-empty measurable sets $A, B \subset \mathbb{R}^n$:

$$\text{Vol}(A + B)^{1/n} \geq \text{Vol}(A)^{1/n} + \text{Vol}(B)^{1/n}$$

- *Multiplicative adimensional form:* For any measurable sets $A, B \subset \mathbb{R}^n$ and any $0 \leq \lambda \leq 1$:

$$\text{Vol}(\lambda A + (1 - \lambda)B) \geq \text{Vol}(A)^\lambda \text{Vol}(B)^{1-\lambda}$$

Proof: Brunn-Minkowski geometric inequality is actually nothing more than Prekopa-Leindler functional inequality (theorem A.4) applied to $f = \mathbb{1}_A$, $g = \mathbb{1}_B$ and $h = \mathbb{1}_{\lambda A + (1-\lambda)B}$.

Corollary A.6 (*Brunn's principle: Hyperplane sections of a convex body*)

Let $K \subset \mathbb{R}^n$ be a convex body and $u \in \mathbb{R}^n$.

Define $f_{K,u} : t \in \mathbb{R} \mapsto \text{Vol}(K \cap \{tu + u^\perp\})$. Then, $[f_{K,u}]^{1/(n-1)}$ is concave on its support.

If K is additionally symmetric, then $f_{K,u}$ is even, and hence maximal in 0.

Proof: Consider the hyperplane sections of K $K_t := K \cap \{tu + u^\perp\}$, $t \in \mathbb{R}$.

By convexity of K , for all $r, s \in \mathbb{R}$ and all $0 \leq \lambda \leq 1$: $\lambda K_r + (1 - \lambda)K_s \subset K_{\lambda r + (1-\lambda)s}$, which implies by Brunn-Minkowski inequality (theorem A.5):

$$\lambda \text{Vol}(K_r)^{1/(n-1)} + (1 - \lambda) \text{Vol}(K_s)^{1/(n-1)} \leq \text{Vol}(\lambda K_r + (1 - \lambda)K_s)^{1/(n-1)} \leq \text{Vol}(K_{\lambda r + (1-\lambda)s})^{1/(n-1)}$$

And this means precisely, as wanted, that for all $r, s \in \mathbb{R}$ and all $0 \leq \lambda \leq 1$:

$$f_{K,u}(\lambda r + (1 - \lambda)s)^{1/(n-1)} \geq \lambda f_{K,u}(r)^{1/(n-1)} + (1 - \lambda)f_{K,u}(s)^{1/(n-1)}$$

A.3 Volume-radius and mean-width of a convex body

For any real m -dimensional Hilbert space $(H, \langle \cdot | \cdot \rangle)$, we will later denote by $\text{Vol}(\cdot)$ the m -dimensional Lebesgue measure on H .

Let $n \in \mathbb{N}^*$. For a given convex body $K \subset \mathbb{R}^n$, we define its *volume-radius* as the radius of the euclidean ball which has the same volume: $\text{vrad}(K) := \left(\frac{\text{Vol}(K)}{\text{Vol}(B_2^n)} \right)^{1/n}$.

Theorem A.7 (*Volume-radii of the commutative p -norm unit balls*)

Let $n \in \mathbb{N}^*$ and $1 \leq p \leq +\infty$.

The volume of the unit ball of \mathbb{R}^n for $\|\cdot\|_p$ is: $\text{Vol}(B_p^n) = \frac{[2\Gamma(1 + 1/p)]^n}{\Gamma(1 + n/p)}$.

As a consequence: $\text{vrad}(B_p^n) \underset{n \rightarrow +\infty}{\sim} \frac{\Gamma(1 + 1/p) p^{1/p}}{\Gamma(1 + 1/2) 2^{1/2}} \left(\frac{e}{n} \right)^{1/p-1/2}$.

In particular: $\text{vrad}(B_1^n) \underset{n \rightarrow +\infty}{\sim} \sqrt{\frac{2e}{\pi}} \frac{1}{\sqrt{n}}$ and $\text{vrad}(B_\infty^n) \underset{n \rightarrow +\infty}{\sim} \sqrt{\frac{2e}{\pi}} \sqrt{n}$.

Let us recall that the Gamma function is defined by: $\forall x > 0$, $\Gamma(x) := \int_0^{+\infty} t^{x-1} e^{-t} dt$.

Remark A.8 *Theorem A.7 is actually a special instance of the more general result:*

For all symmetric convex body $K \subset \mathbb{R}^n$ and all $p > 0$, $|K| = \frac{\int_{\mathbb{R}^n} e^{-g_K(x)^p} dx}{\Gamma(1 + n/p)}$.

The latter is a consequence of the computation:

$$\int_{\mathbb{R}^n} e^{-g_K(x)^p} dx = \int_{\mathbb{R}^n} \int_{t > g_K(x)^p} e^{-t} dt dx = \int_0^{+\infty} \int_{g_K(x) < t^{1/p}} e^{-t} dx dt = \int_0^{+\infty} e^{-t} |t^{1/p} K| dt = |K| \int_0^{+\infty} t^{n/p} e^{-t} dt$$

One thing we retrieve from theorem A.7 is:

$$\forall 1 \leq p, p' \leq +\infty, \frac{1}{p} + \frac{1}{p'} = 1, \text{vrad}(B_{p'}^n) \simeq n^{1/2-1/p'} = \frac{1}{n^{1/2-1/p}} \simeq \frac{1}{\text{vrad}(B_p^n)}$$

Hence, in light of the duality formula 11: $\forall 1 \leq p \leq +\infty$, $\text{vrad}((B_p^n)^\circ) \simeq \frac{1}{\text{vrad}(B_p^n)}$.

We shall see in theorem A.9 below that such inverse scaling between the volume-radius of a symmetric convex body and the one of its polar is in fact not specific to the p -norm balls.

Theorem A.9 (*Santaló's inequalities*)

There exists $0 < c < 1$ such that for all $n \in \mathbb{N}^*$ and all $K \subset \mathbb{R}^n$ symmetric convex body:

$$\left[c^n \text{Vol}(B_2^n) \right]^2 \leq \text{Vol}(K) \text{Vol}(K^\circ) \leq \left[\text{Vol}(B_2^n) \right]^2, \text{ i.e. } c \leq \text{vrad}(K) \text{vrad}(K^\circ) \leq 1.$$

Proof: The so-called *direct Santaló inequality*, that is the upper-bound on $\text{Vol}(K) \text{Vol}(K^\circ)$, is amongst the “standard” geometric inequalities. Let us sketch the proof due to Meyer and Pajor [17], which makes essential use of the *Steiner symmetrization* (see e.g. [10] for a detailed presentation).

The *Steiner symmetrization* of a symmetric convex body $K \subset \mathbb{R}^n$ with respect to a hyperplane H of \mathbb{R}^n is defined as: $K_H := \left\{ \frac{1}{2}(x - y), x, y \in K, x - y \in H^\perp \right\}$. It satisfies the two main properties:

- *Volume-preservingness:* $\text{Vol}(K_H) = \text{Vol}(K)$
- *Convergence to an euclidean ball* (in geometric distance $d_g(K, L) := \inf\{\beta/\alpha, \alpha L \subset K \subset \beta L\}$): $\exists (H_k)_{k \in \mathbb{N}} : K_{H_k} \xrightarrow{k \rightarrow +\infty} K^*$ where $K^* = \text{vrad}(K) B_2^n$.

To complete the proof, it thus suffices to show that:

$$\text{Vol}((K_H)^\circ) \geq \text{Vol}(K^\circ) \tag{13}$$

Indeed, taking then $(H_k)_{k \in \mathbb{N}}$ such that $K_{H_k} \xrightarrow{k \rightarrow +\infty} K^*$, we also have $(K_{H_k})^\circ \xrightarrow{k \rightarrow +\infty} (K^*)^\circ$.

So, since already $\forall k \in \mathbb{N}$, $\text{Vol}(K_{H_k}) = \text{Vol}(K)$, if additionally $\forall k \in \mathbb{N}$, $\text{Vol}((K_{H_k})^\circ) \geq \text{Vol}(K^\circ)$, then $\forall k \in \mathbb{N}$, $\text{Vol}(K_{H_k}) \text{Vol}((K_{H_k})^\circ) \geq \text{Vol}(K) \text{Vol}(K^\circ)$.

And consequently $\text{Vol}(K^*) \text{Vol}((K^*)^\circ) \geq \text{Vol}(K) \text{Vol}(K^\circ)$ i.e. $[\text{Vol}(B_2^n)]^2 \geq \text{Vol}(K) \text{Vol}(K^\circ)$.

Now, considering for any convex body \tilde{K} its slices $\tilde{K}[s] := \{x \in H, x + su_H \in \tilde{K}\}$, $s \in \mathbb{R}$, where $H^\perp = \mathbb{R}u_H$ and $\|u_H\|_2 = 1$, we have by Fubini: $\text{Vol}(\tilde{K}) = \int_{\mathbb{R}} \text{Vol}(\tilde{K}[s]) ds$. So to get equation 13, it is actually enough to show that:

$$\forall s \in \mathbb{R}, \text{Vol}((K_H)^\circ[s]) \geq \text{Vol}(K^\circ[s]) \tag{14}$$

Yet, one may check that, by symmetry of K° : $\forall s \in \mathbb{R}$, $\frac{1}{2}(K^\circ[s] - K^\circ[-s]) \subset (K_H)^\circ[s]$, so that:

$\forall s \in \mathbb{R}$, $\text{Vol}((K_H)^\circ[s]) \geq \text{Vol}\left(\frac{1}{2}(K^\circ[s] - K^\circ[-s])\right) \geq (\text{Vol}(K^\circ[s]) \text{Vol}(-K^\circ[-s]))^{1/2} = \text{Vol}(K^\circ[s])$, where the next to last inequality is by Brunn-Minkowski inequality (theorem A.5).

This is precisely equation 14.

Proving the so-called *reverse Santaló inequality*, that is the lower-bound on $\text{Vol}(K) \text{Vol}(K^\circ)$, is much more involved. It relies on Milman's *isomorphic symmetrization* introduced in [18]. The reader is referred again to [10] for a complete proof.

Let $d \in \mathbb{N}^*$. $\mathcal{H}(\mathbb{C}^d)$ is a d^2 -dimensional real vector space. Thus, if we still define the *volume-radius* of a convex body $K \subset \mathcal{H}(\mathbb{C}^d)$ as the radius of the euclidean ball which has the same volume as K , we

$$\text{get: } \text{vrad}(K) := \left(\frac{\text{Vol}(K)}{\text{Vol}(B_2^{d^2})} \right)^{1/d^2}.$$

The results of theorem A.10 below are taken from [20].

Theorem A.10 (*Volume-radii of the non-commutative p -norm unit balls*)

Let $d \in \mathbb{N}^*$.

The unit ball of $\mathcal{H}(\mathbb{C}^d)$ for $\|\cdot\|_2$ has the same volume as the unit ball of \mathbb{R}^{d^2} for $\|\cdot\|_2$:

$\text{Vol}\left(B_{\|\cdot\|_2}^d\right) = \text{Vol}(B_2^{d^2}) = \frac{\sqrt{\pi}^{d^2}}{\Gamma(1+d^2/2)}$, so that: $\left(\text{Vol}\left(B_{\|\cdot\|_2}^d\right)\right)^{1/d^2} \underset{d \rightarrow +\infty}{\sim} \frac{\sqrt{2\pi e}}{d}$.

And more generally, for all $1 \leq p \leq +\infty$, the volume of the unit ball of $\mathcal{H}(\mathbb{C}^d)$ for $\|\cdot\|_p$ satisfies:

$$\left(\text{Vol}\left(B_{\|\cdot\|_p}^d\right)\right)^{1/d^2} \underset{d \rightarrow +\infty}{\sim} \frac{\sqrt{2\pi e^{3/2}\Delta(p/2)}}{d^{1/2+1/p}}, \text{ with } \frac{1}{4} \leq \Delta(q) \leq 4 \text{ for } \frac{1}{2} \leq q \leq +\infty, \text{ and } \begin{cases} \Delta(1) = e^{-1/2} \\ \Delta(+\infty) = \frac{1}{4} \end{cases}.$$

As a consequence: $\forall 1 \leq p \leq +\infty$, $\text{vrad}\left(B_{\|\cdot\|_p}^d\right) \underset{d \rightarrow +\infty}{\sim} \sqrt{e^{1/2}\Delta(p/2)}d^{1/2-1/p}$.

Hence, regarding the volume-radii of both the commutative and non-commutative p -norm unit balls, one fact that may be worth pointing out is the following:

$$\forall 1 \leq p \leq +\infty, \text{vrad}\left(B_{\|\cdot\|_p}^d\right) \simeq \text{vrad}\left(B_p^d\right) \simeq d^{1/2-1/p}$$

This is a striking phenomenon. Indeed, the non-commutative Hilbert space $\mathcal{H}(\mathbb{C}^d)$ has real dimension d^2 . Nonetheless, the relation between the volumes of its p -norm balls and the one its euclidean ball is not the same as the relation which holds in the d^2 -dimensional commutative Hilbert space \mathbb{R}^{d^2} , but instead the same as the relation which holds in the d -dimensional commutative Hilbert space \mathbb{R}^d .

Remark A.11 Other quantities that might be of interest regarding the volume considerations in the geometry of a given convex body $K \subset \mathbb{R}^n$ are its in-radius $\text{inrad}(K)$ and its out-radius $\text{outrad}(K)$: $\text{inrad}(K)$ is the radius of the largest euclidean ball (centered at the center of gravity of K) which is contained into K , whereas $\text{outrad}(K)$ is the radius of the smallest euclidean ball (centered at the center of gravity of K) which contains K .

Let $K \subset \mathbb{R}^n$ be a convex body containing 0 in its interior.

For all unit vector $u \in S_2^n$, we define the width of K in the direction u as: $w(K, u) := \sup_{x \in K} \langle x | u \rangle$.

We then define the mean-width of K as: $w(K) := \int_{S_2^n} w(K, u) du$, where du denotes the uniform probability distribution over S_2^n .

When K is symmetric, we may write more simply (consult appendix A.1 for needed notations):

$$w(K, u) = g_{K^\circ}(u) \quad \text{and} \quad w(K) = \int_{S_2^n} g_{K^\circ}(u) du$$

Theorem A.12 (Urysohn's inequality)

Let $K \subset \mathbb{R}^n$ be a symmetric convex body. Then: $\text{vrad}(K) \leq w(K)$.

Proof: By integrating $\mathbb{1}_{B_2^n}$ and $\mathbb{1}_{K^\circ}$ in polar coordinates, we see that:

$$\text{Vol}(B_2^n) = \int_{\mathbb{R}^n} \mathbb{1}_{B_2^n}(x) dx = \int_{S_2^n} \int_0^1 r^{n-1} dr du$$

$$\text{Vol}(K^\circ) = \int_{\mathbb{R}^n} \mathbb{1}_{K^\circ}(x) dx = \int_{S_2^n} \int_0^{1/g_{K^\circ}(u)} r^{n-1} dr du = \int_{S_2^n} \int_0^1 g_{K^\circ}(u)^{-n} s^{n-1} ds du \quad (s = g_{K^\circ}(u)r)$$

Hence, by Jensen's inequality applied to the convex function $(\cdot)^{-1/n}$, we get:

$$\left(\frac{\text{Vol}(K^\circ)}{\text{Vol}(B_2^n)}\right)^{-1/n} = \left(\int_{S_2^n} g_{K^\circ}(u)^{-n} du\right)^{-1/n} \leq \int_{S_2^n} g_{K^\circ}(u) du = w(K)$$

Now, by Santaló's inequality (theorem A.9), we know that: $\left(\frac{\text{Vol}(K)}{\text{Vol}(B_2^n)}\right)^{1/n} \leq \left(\frac{\text{Vol}(K^\circ)}{\text{Vol}(B_2^n)}\right)^{-1/n}$, which yields as advertized: $\text{vrad}(K) \leq w(K)$.

A.4 Estimates on entropy numbers by a volumic approach

Let (X, d) be a metric space and $T \subset X$ be a subset of X . Fix $\epsilon > 0$.

A collection of points $\{x_i, 1 \leq i \leq n\} \subset T$ is said to be:

- an ϵ -net of T for d if $T \subset \bigcup_{1 \leq i \leq n} B_d(x_i, \epsilon)$.
- an ϵ -separated set in T for d if $\forall 1 \leq i \neq j \leq n, d(x_i, x_j) \geq \epsilon$.

The cardinality of a minimal ϵ -net of T for d is called the *covering number* $N(T, d, \epsilon)$.

The cardinality of a maximal ϵ -separated set in T for d is called the *packing number* $K(T, d, \epsilon)$.

The logarithms of the covering and packing numbers are often referred to as the *entropy numbers*.

Noticing that a maximal ϵ -separated set is an ϵ -net, and that an $\frac{\epsilon}{2}$ -net may be mapped to an ϵ -separated set, one gets the relations:

$$N(T, d, \epsilon) \leq K(T, d, \epsilon) \leq N(T, d, \epsilon/2)$$

Example A.13 For all A, B symmetric convex bodies of \mathbb{R}^n and $\epsilon > 0$, set $N(A, \epsilon B) := N(A, g_B, \epsilon)$ (consult appendix A.1 for needed notations).

By simply estimating the volumes of either a minimal ϵ -net of A for g_B or a maximal ϵ -separated set in A for g_B , one gets the bounds:

$$\frac{\text{Vol}(\frac{1}{\epsilon}A)}{\text{Vol}(B)} \leq N(A, \epsilon B) \leq \frac{\text{Vol}(B + \frac{2}{\epsilon}A)}{\text{Vol}(B)}$$

In particular, if $B \subset A$: $\left(\frac{1}{\epsilon}\right)^n \frac{\text{Vol}(A)}{\text{Vol}(B)} \leq N(A, \epsilon B) \leq \left(1 + \frac{2}{\epsilon}\right)^n \frac{\text{Vol}(A)}{\text{Vol}(B)}$.

B Gaussian variables

We will assume throughout this section that we have at hand an abstract probability space (Ω, \mathbb{P}) on which all random variables are defined.

B.1 Generalities

- A real-valued random variable $g : \Omega \rightarrow \mathbb{R}$ is said to be (*standard*) *gaussian* if:

$$\forall t \in \mathbb{R}, \mathbb{P}(g > t) = \frac{1}{\sqrt{2\pi}} \int_t^{+\infty} e^{-x^2/2} dx$$

We will write in such case: $g \sim \mathcal{N}(0, 1)$.

- Let $n \in \mathbb{N}^*$. A vector-valued random variable $g = (g_1, \dots, g_n) : \Omega \rightarrow \mathbb{R}^n$ is said to be (*standard*) *gaussian* if the $g_i : \Omega \rightarrow \mathbb{R}, 1 \leq i \leq n$, are independent (*standard*) gaussian real-valued random variables, i.e. if:

$$\forall (t_1, \dots, t_n) \in \mathbb{R}^n, \mathbb{P}(g_1 > t_1, \dots, g_n > t_n) = \frac{1}{\sqrt{2\pi}^n} \int_{t_1}^{+\infty} \dots \int_{t_n}^{+\infty} e^{-(x_1^2 + \dots + x_n^2)/2} dx_1 \dots dx_n$$

We will write in such case: $g \sim \mathcal{N}^n(0, 1)$.

The (*standard*) *gaussian probability measure* on \mathbb{R}^n is thus defined as:

$$d\nu_n(x) := \frac{1}{\sqrt{2\pi}^n} e^{-(x_1^2 + \dots + x_n^2)/2} dx_1 \dots dx_n$$

Its two essential features are being a product measure and invariant under orthogonal transformations.

Remark B.1 More generally, given $\mu \in \mathbb{R}$ and $\sigma > 0$, one will say that a real-valued random variable $g : \Omega \rightarrow \mathbb{R}$ has law $\mathcal{N}(\mu, \sigma^2)$ if: $\forall t \in \mathbb{R}, \mathbb{P}(g > t) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_t^{+\infty} e^{-(x-\mu)^2/2\sigma^2} dx$.

For any real-valued random variable Z , we define its p -norm, $1 \leq p \leq +\infty$, as:

$$\|Z\|_p := (\mathbb{E}|Z|^p)^{1/p} \text{ if } 1 \leq p < +\infty, \text{ and } \|Z\|_\infty := \sup |Z|.$$

Then, for any collection $\mathbf{Z} := \{Z_i, i \in I\}$ of real-valued random variables and any $\epsilon > 0$, we may define its covering number: $N_{\mathbf{Z}}(\epsilon) := N(\mathbf{Z}, \|\cdot\|_2, \epsilon)$ (cf appendix A.4).

Definition B.2 A collection $\{Z_i, i \in I\}$ of real-valued random variables is called a Gaussian process if all the linear combinations of the variables $Z_i, i \in I$, are Gaussian.

The following theorem, which provides bounds on the supremum of a Gaussian process indexed by a countable set, is quite fundamental. A proof and additional comments can be found in [9].

Theorem B.3 There exist constants $C, C' > 0$ such that for any Gaussian process $\mathbf{Z} := \{Z_i, i \in I\}$ indexed by a countable set I , we have:

$$C \sup_{\epsilon > 0} \left(\epsilon \sqrt{\log N_{\mathbf{Z}}(\epsilon)} \right) \leq \mathbb{E} \left(\sup_{i \in I} Z_i \right) \leq C' \int_0^{+\infty} \sqrt{\log N_{\mathbf{Z}}(\epsilon)} d\epsilon$$

This general result is actually obtained by iterating the finite-case theorem below (which will often be sufficient for the applications we have in mind):

Theorem B.4 Let $\{Z_i, i \in I\}$ be a Gaussian process indexed by a finite set I and such that: $\exists \alpha, \beta > 0 : \forall i \in I, \|Z_i\|_2 \leq \alpha$ and $\forall i \neq j \in I, \|Z_i - Z_j\|_2 \geq \beta$.

Then: $C\beta\sqrt{\log |I|} \leq \mathbb{E} \left(\sup_{i \in I} Z_i \right) \leq C'\alpha\sqrt{\log |I|}$.

Furthermore, if the Gaussian process is additionally centered, then one may choose $C' = \sqrt{2}$.

B.2 Gaussian variables and mean-width

Within the framework of Gaussian variables, one may give an alternative definition of the mean-width as the one provided in appendix A.3.

Let $K \subset \mathbb{R}^n$ be a convex body containing 0 in its interior.

Its mean-width may actually be equivalently defined as: $w(K) := \mathbb{E}_{G \sim \mathcal{N}^n(0,1)} \left[\sup_{x \in K} \left\langle x \mid \frac{G}{\|G\|_2} \right\rangle \right]$.

Yet, for $G \sim \mathcal{N}^n(0,1)$, $\|G\|_2$ and $\frac{G}{\|G\|_2}$ are independent random variables, so that:

$$\mathbb{E} \left[\sup_{x \in K} \langle x | G \rangle \right] = \mathbb{E} \left[\|G\|_2 \sup_{x \in K} \left\langle x \mid \frac{G}{\|G\|_2} \right\rangle \right] = \mathbb{E} \|G\|_2 \mathbb{E} \left[\sup_{x \in K} \left\langle x \mid \frac{G}{\|G\|_2} \right\rangle \right] = \gamma_n \mathbb{E} \left[\sup_{x \in K} \left\langle x \mid \frac{G}{\|G\|_2} \right\rangle \right], \text{ where}$$

$$\gamma_n := \mathbb{E}_{G \sim \mathcal{N}^n(0,1)} \|G\|_2 = \frac{\sqrt{2}\Gamma(n/2+1/2)}{\Gamma(n/2)}, \text{ so that } \sqrt{n-1} \leq \gamma_n \leq \sqrt{n}.$$

We thus get in the end: $w(K) = \frac{1}{\gamma_n} \mathbb{E}_{G \sim \mathcal{N}^n(0,1)} \left[\sup_{x \in K} \langle x | G \rangle \right]$.

When K is symmetric, we may write more simply: $w(K) = \frac{1}{\gamma_n} \mathbb{E}_{G \sim \mathcal{N}^n(0,1)} [g_{K^\circ}(G)]$.

Theorem B.4 then provides a quite often efficient way of bounding the mean-width of convex bodies that are the convex envelope of a finite number of points.

Let indeed $T \subset \tau B_2^n$ be a finite set. Then: $\text{vrad}(\text{Conv}(T)) \leq w(\text{Conv}(T)) \leq \frac{\tau\sqrt{2\log |T|}}{\sqrt{n}}$, where the first inequality is by Urysohn's inequality (theorem A.12) and the second inequality is by theorem B.4.

If additionally: $\forall x \neq y \in T, \|x - y\|_2 \geq \delta$, then by theorem B.4 again: $w(\text{Conv}(T)) \gtrsim \frac{\delta\sqrt{\log |T|}}{\sqrt{n}}$.

Example B.5 *Mean-width of the commutative p -norm unit balls.*

- For all $1 < p \leq +\infty$ and $1 \leq p' < +\infty$ such that $\frac{1}{p} + \frac{1}{p'} = 1$, we have:
 $w(B_p^n) = \frac{1}{\gamma_n} \mathbb{E}_{G \sim \mathcal{N}^n(0,1)} \|G\|_{p'} \simeq \frac{n^{1/p'}}{n^{1/2}} = n^{1/2-1/p} \simeq \text{vrad}(B_p^n)$
- $w(B_1^n) = \frac{1}{\gamma_n} \mathbb{E}_{G \sim \mathcal{N}^n(0,1)} \|G\|_\infty \simeq \frac{\sqrt{\log n}}{\sqrt{n}} \simeq \sqrt{\log n} \text{vrad}(B_1^n)$

B.3 Brief incursion into random matrix theory: the GUE

The results from random matrix theory presented in this section are by far not the most general ones: we have focussed on the limited statements that were of use for our purpose. The reader is referred to [11] for a complete exposition and proofs.

- A complex matrix-valued random variable $G = (G_{j,k})_{1 \leq j,k \leq d} : \Omega \rightarrow \mathbb{C}^{d \times d}$ is said to be (*standard*) *gaussian* if the $G_{j,k} : \Omega \rightarrow \mathbb{C}$, $1 \leq j, k \leq d$, are independent (standard) gaussian complex-valued random variables.
 We shall write in such case: $\forall 1 \leq j, k \leq d$, $G_{j,k} \sim \mathcal{N}_{\mathbb{C}}(0, 1)$ and $G \sim \mathcal{N}_{\mathbb{C}}^{d \times d}(0, 1)$.
- A Hermitian-valued random variable $H : \Omega \rightarrow \mathcal{H}(\mathbb{C}^d)$ is said to belong to the *Gaussian Unitary Ensemble (GUE)* if $H = \frac{1}{2}(G + G^\dagger)$ with $G \sim \mathcal{N}_{\mathbb{C}}^{d \times d}(0, 1)$.
 We shall write in such case: $H \sim \text{GUE}(d)$.

For any $H \in \mathcal{H}(\mathbb{C}^d)$, we denote by $\lambda_1(H), \dots, \lambda_d(H) \in \mathbb{R}$ its eigenvalues, and by $\lambda_{\max}(H)$ its largest eigenvalue.

We then define its *empirical eigenvalue distribution* as: $N_H := \frac{1}{d} \sum_{j=1}^d \delta_{\lambda_j(H)} : \mathbb{R} \rightarrow [0; 1]$.

In other words, N_H is the probability measure on \mathbb{R} which is characterized by the property that for any $I \subset \mathbb{R}$, $N_H(I)$ is the proportion of eigenvalues of H belonging to I .

We also introduce the so-called *semi-circular probability measure* on \mathbb{R} :

$$d\nu_{SC}(x) := \frac{1}{2\pi} \sqrt{4 - x^2} \mathbb{1}_{[-2;2]}(x) dx$$

Theorem B.6 (*Wigner's semi-circular law*)

For every $d \in \mathbb{N}^*$, let $H_d \sim \text{GUE}(d)$. Then $N_{\frac{1}{\sqrt{d}}H_d} \xrightarrow{d \rightarrow +\infty} \nu_{SC}$ in the following sense:

$$\forall \epsilon > 0, \forall I \subset \mathbb{R}, \lim_{d \rightarrow +\infty} \mathbb{P} \left(\left| N_{\frac{1}{\sqrt{d}}H_d}(I) - \nu_{SC}(I) \right| > \epsilon \right) = 0$$

Furthermore, $\lambda_{\max} \left(\frac{1}{\sqrt{d}}H_d \right) \xrightarrow{d \rightarrow +\infty} 2$ in the following sense:

$$\forall \epsilon > 0, \lim_{d \rightarrow +\infty} \mathbb{P} \left(\left| \lambda_{\max} \left(\frac{1}{\sqrt{d}}H_d \right) - 2 \right| > \epsilon \right) = 0$$

Remark B.7 *Wigner's semi-circular law remains actually valid under much weaker hypotheses on the considered sequence of random matrices (namely that for all $d \in \mathbb{N}^*$, $H_d \in \mathcal{H}(\mathbb{C}^d)$ has independent centered and suitably bounded entries).*

For instance, if not looking at a sequence of GUE matrices but instead at a sequence of GUE matrices conditioned to have trace 0, the limit eigenvalue distribution will not be affected.

One consequence of Wigner's semi-circular law (theorem B.6) is that, for all $p \in \mathbb{N}^*$:

$$\mathbb{E}_{\Delta \sim GUE(d)} \left[\text{Tr} \left| \frac{1}{\sqrt{d}} \Delta \right|^p \right] = \mathbb{E}_{\Delta \sim GUE(d)} \left[\sum_{j=1}^d \left| \lambda_j \left(\frac{1}{\sqrt{d}} \Delta \right) \right|^p \right] \underset{d \rightarrow +\infty}{\sim} d \int_{-\infty}^{+\infty} |x|^p d\nu_{SC}(x)$$

Hence, for all $p \in \mathbb{N}^*$, the expectancy of the p -norm of a GUE matrix follows the asymptotics:

$$\mathbb{E}_{\Delta \sim GUE(d)} \|\Delta\|_p \underset{d \rightarrow +\infty}{\sim} \left(\int_{-\infty}^{+\infty} |x|^p d\nu_{SC}(x) \right)^{1/p} d^{1/p+1/2}$$

Besides: $\mathbb{E}_{\Delta \sim GUE(d)} \|\Delta\|_\infty \underset{d \rightarrow +\infty}{\sim} 2d^{1/2}$

Now, the moments of the semi-circular probability distribution may be computed:

$$\begin{aligned} \int_{-\infty}^{+\infty} |x|^p d\nu_{SC}(x) &= \frac{1}{\pi} \int_0^2 x^p \sqrt{4-x^2} dx \\ &= \frac{4 \times 2^p}{\pi} \int_0^{\pi/2} \sin^p \theta (1 - \sin^2 \theta) d\theta \\ &= \frac{4 \times 2^p}{\pi} \left(4(p-1) \int_0^{\pi/2} \sin^{p-2} \theta (1 - \sin^2 \theta) d\theta - (p+1) \int_0^{\pi/2} \sin^{p+2} \theta (1 - \sin^2 \theta) d\theta \right) \\ &= 4(p-1) \int_{-\infty}^{+\infty} |x|^{p-2} d\nu_{SC}(x) - (p+1) \int_{-\infty}^{+\infty} |x|^{p+2} d\nu_{SC}(x) \end{aligned}$$

where we operated the change of variables $x = \sin \theta$ from the first to the second line, and an integration by parts from the second to the third.

We thus get the recursivity relation: $\forall p \in \mathbb{N}$, $\int_{-\infty}^{+\infty} |x|^{p+2} d\nu_{SC}(x) = \frac{2(p+1)}{|p/2|+2} \int_{-\infty}^{+\infty} |x|^p d\nu_{SC}(x)$.

Which implies that for all $q \in \mathbb{N}$:

$$\begin{aligned} \int_{-\infty}^{+\infty} |x|^{2q} d\nu_{SC}(x) &= \frac{1}{q+1} \binom{2q}{q} \int_{-\infty}^{+\infty} d\nu_{SC}(x) = \frac{1}{q+1} \binom{2q}{q} \\ \int_{-\infty}^{+\infty} |x|^{2q+1} d\nu_{SC}(x) &= \frac{1}{q+1} \binom{2q}{q} \int_{-\infty}^{+\infty} |x| d\nu_{SC}(x) = \frac{1}{q+1} \binom{2q}{q} \frac{8}{3\pi} \end{aligned}$$

In fact: $\int_{-\infty}^{+\infty} d\nu_{SC}(x) = 1$ and $\int_{-\infty}^{+\infty} |x| d\nu_{SC}(x) = \frac{8}{\pi} \int_0^{\pi/2} \sin \theta \cos^2 \theta d\theta = \frac{8}{\pi} \left[-\frac{\cos^3 \theta}{3} \right]_0^{\pi/2} = \frac{8}{3\pi}$.

Example B.8 *Mean-width of the non-commutative p -norm unit balls.*

For all $1 \leq p \leq +\infty$ and $1 \leq p' \leq +\infty$ such that $\frac{1}{p} + \frac{1}{p'} = 1$, we have:

$$w \left(B_{\|\cdot\|_p}^d \right) = \frac{1}{\gamma_{d^2}} \mathbb{E}_{G \sim GUE(d)} \|G\|_{p'} \simeq \frac{d^{1/2+1/p'}}{d} = d^{1/2-1/p} \simeq \text{vrad} \left(B_{\|\cdot\|_p}^d \right)$$

C Large deviations

C.1 Concentration rate and deviation inequalities on a probability metric space

The so-called *concentration of measure phenomenon* and the counter-intuitive results it yields in high-dimensional convex geometry have been extensively studied already. Those subjects are however still continuously shaped by new developments. A both general and detailed presentation of this field may be found in the very accessible [7] or the more technical [8]. [12] presents a more probability-orientated account of these ideas.

Let (X, d, μ) be a probability metric space. We define:

- The *diameter* of (X, d) : $D_{(X,d)} := \sup_{x,y \in X} d(x, y) \in [0; +\infty]$.
- The *concentration rate* of (X, d, μ) : $\alpha_{(X,d,\mu)} : r \in]0; +\infty[\mapsto \sup_{A \subset X, \mu(A) \geq \frac{1}{2}} \mu(X \setminus A_r) \in [0; 1]$,
where $A_r := \{x \in X, d(x, A) \geq r\}$ is the r -extension (or r -neighbourhood) of A .

A function $f : (X, d) \rightarrow \mathbb{R}$ is said to be L -lipschitz, $L \in \mathbb{R}^+$, if: $\forall x, y \in X, |f(x) - f(y)| \leq L d(x, y)$.

A *median* for μ of a function $f : (X, \mu) \rightarrow \mathbb{R}$ is any $m_f \in \mathbb{R}$ such that: $\begin{cases} \mu(\{f \leq m_f\}) \geq \frac{1}{2} \\ \mu(\{f \geq m_f\}) \geq \frac{1}{2} \end{cases}$.

The *average* for μ of a function $f : (X, \mu) \rightarrow \mathbb{R}$ is: $M_f := \int_X f d\mu$.

Theorem C.1 *Let $f : (X, d, \mu) \rightarrow \mathbb{R}$ be a L -lipschitz function. Then:*

$$\forall r > 0, \mu(\{|f - m_f| \geq r\}) \leq 2 \alpha_{(X,d,\mu)}\left(\frac{r}{L}\right) \text{ and } \mu(\{|f - M_f| \geq r\}) \leq 2 \alpha_{(X,d,\mu)}\left(\frac{r}{L}\right)$$

$$\text{Conversely: } \forall r > 0, \alpha_{(X,d,\mu)}(r) = \sup_{\substack{f:(X,d,\mu) \rightarrow \mathbb{R} \\ f \text{ 1-lipschitz}}} \frac{1}{2} \mu(\{|f - m_f| \geq r\}) = \sup_{\substack{f:(X,d,\mu) \rightarrow \mathbb{R} \\ f \text{ 1-lipschitz}}} \frac{1}{2} \mu(\{|f - M_f| \geq r\})$$

What theorem C.1 brings to light is that there is a strong connection between the so-called *concentration* of lipschitz functions $f : (X, d, \mu) \rightarrow \mathbb{R}$ around their median or their average and the *isoperimetric problem* on (X, d, μ) , namely: given $0 < \epsilon < D_{(X,d)}$ and $0 < m < 1$, find $\tilde{A} \subset X$ such that $\mu(\tilde{A}) = m$ and $\mu(\tilde{A}_\epsilon) = \inf_{A \subset X, \mu(A)=m} \mu(A_\epsilon)$.

Example C.2 *Concentration of measure phenomenon on the real euclidean unit sphere.*

Denote by μ_n the n -dimensional (volumic) Lebesgue measure on \mathbb{R}^n and by σ_n the induced normalized $(n-1)$ -dimensional (surfacic) measure on S_2^n . σ_n is also the unique rotationally invariant probability measure on S_2^n (unique Haar probability measure for the action of the orthogonal group $\mathfrak{D}(n)$ on S_2^n). The isoperimetric problem has been solved by Lévy on $(S_2^n, \|\cdot\|_2, \sigma_n)$:

For all $0 < \epsilon < 1$ and $0 < m < 1$, $\inf_{A \subset S_2^n, \sigma_n(A)=m} \sigma_n(A_\epsilon)$ exists and is attained by the spherical cap

$C(x, r) := \{y \in S_2^n, \|x - y\|_2 \leq r\}$ for any $x \in S_2^n$ and r such that $\sigma_n(C(x, r)) = m$.

Now, it holds that: $\forall x \in S_2^n, \forall 0 < r < \sqrt{2}, \sigma_n(C(x, r)) \leq e^{-nr^2/2}$ where $\epsilon_r := 1 - \frac{r^2}{2}$.

We thus get as a corollary of Lévy's isoperimetric theorem the following concentration inequality for lipschitz functions on $(S_2^n, \|\cdot\|_2, \sigma_n)$:

If $f : (S_2^n, \|\cdot\|_2, \sigma_n) \rightarrow \mathbb{R}$ is a L -lipschitz function, then: $\forall r > 0, \sigma_n(\{|f - m_f| \geq r\}) \leq 2e^{-nr^2/2L^2}$.

The so-called *Laplace transform method* enables one to get more user-friendly concentration results whenever (X, d) has finite diameter. It may be summarized as follows:

First of all, we have by Markov's inequality that for any function $g : (X, d, \mu) \rightarrow \mathbb{R}$ and any $a \in \mathbb{R}$:

$$\mu(\{g \geq a\}) \leq \inf_{\lambda \geq 0} \left(e^{-\lambda a} \int_X e^{\lambda g} d\mu \right)$$

Then, we have by Jensen's inequality that for any function $f : (X, d, \mu) \rightarrow \mathbb{R}$ and any $\lambda \geq 0$:

$$\int_X e^{\lambda[f(x) - M_f]} d\mu(x) \leq \int_{X \times X} e^{\lambda[f(x) - f(y)]} d\mu(x) d\mu(y)$$

Moreover, if $\forall x, y \in X, |f(x) - f(y)| \leq k$, then $\int_{X \times X} e^{\lambda[f(x) - f(y)]} d\mu(x) d\mu(y) \leq e^{k^2 \lambda^2 / 2}$.

And the last thing we should notice is that whenever $D_{(X,d)} < +\infty$, then any L -lipschitz function $f : (X, d, \mu) \rightarrow \mathbb{R}$ satisfies: $\forall x, y \in X, |f(x) - f(y)| \leq LD_{(X,d)}$.

Putting everything together, the result we eventually come to is:

Theorem C.3 If $D_{(X,d)} < +\infty$, then for any L -lipschitz function $f : (X, d, \mu) \rightarrow \mathbb{R}$, we have:

$$\forall r > 0, \mu(\{|f - M_f| \geq r\}) \leq 2e^{-r^2/2[D_{(X,d)}]^2 L^2}$$

Remark C.4 As a consequence, we also get by theorem C.1 the corresponding inequality for the concentration rate of (X, d, μ) : if $D_{(X,d)} < +\infty$, then $\forall r > 0$, $\alpha_{(X,d,\mu)}(r) \leq e^{-r^2/8[D_{(X,d)}]^2}$.

The concentration inequality from theorem C.3 itself generally does not provide a good estimate (the main problem being that it does not depend on the measure μ but only on the distance d). Nonetheless, a nice property of this theorem is that it may actually be recursively generalized to spaces with a product structure, on which it might yield much more accurate concentration inequalities.

Theorem C.5 Let $(X_1, d_1), \dots, (X_n, d_n)$ be n metric spaces with finite diameter.

We consider the product space $X := X_1 \times \dots \times X_n$ equipped with the product distance $d := d_1 \oplus \dots \oplus d_n$ and any product probability measure $\mu = \mu_1 \otimes \dots \otimes \mu_n$. Then, setting $D^2 := [D_{(X_1,d_1)}]^2 + \dots + [D_{(X_n,d_n)}]^2$, we have that for any L -lipschitz function $f : (X, d, \mu) \rightarrow \mathbb{R}$:

$$\forall r > 0, \mu(\{|f - M_f| \geq r\}) \leq 2e^{-r^2/2D^2 L^2}$$

Deviation inequalities for sums of bounded random variables can be directly deduced from theorem C.5. The first, and perhaps' most important, one being:

Theorem C.6 (*Hoeffding's inequality*)

Let X_1, \dots, X_n be n independent random variables such that $\forall 1 \leq k \leq n$, $a_k \leq X_k \leq b_k$.

Setting $\Delta^2 := \frac{1}{n} \sum_{1 \leq k \leq n} (b_k - a_k)^2$, we have:

$$\forall t > 0, \mathbb{P} \left(\left| \frac{1}{n} \sum_{k=1}^n (X_k - \mathbb{E}X_k) \right| \geq t \right) \leq 2 \exp \left(-\frac{nt^2}{2\Delta^2} \right)$$

C.2 Orlicz spaces and ψ_α -random variables

Amongst the numerous properties related to Orlicz spaces, only a few ones which are necessary to our purpose are exposed here. This short description should be fleshed by referring, for instance, to the very complete course [14].

Definition C.7 A function $\varphi : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is said to be an Orlicz function if it is convex, increasing, with closed support, and such that $\varphi(0) = 0$, $\varphi(x) \xrightarrow{x \rightarrow +\infty} +\infty$.

If φ is an Orlicz function, a random variable X is said to be a φ -random variable if its φ -norm, defined by $\|X\|_\varphi := \inf \left\{ c > 0, \mathbb{E} \left[\varphi \left(\frac{|X|}{c} \right) \right] \leq \varphi(1) \right\}$ is finite.

The only examples we shall be interested in are the following:

- For all $p \geq 1$, $\phi_p : x \in \mathbb{R}^+ \mapsto \frac{x^p}{p} \in \mathbb{R}^+$ is an Orlicz function. The ϕ_p -norm $\|\cdot\|_{\phi_p}$ is actually the p -order moment $\mathbb{E}|\cdot|^p$, so that the space of ϕ_p -random variable is nothing else than the L_p -space.
- For all $\alpha \geq 1$, $\psi_\alpha : x \in \mathbb{R}^+ \mapsto e^{x^\alpha} - 1$ is an Orlicz function. The ψ_1 random variables are sometimes referred to as *sub-exponential*, and the ψ_2 random variables as *sub-gaussian*.

There are actually very precise connections between the L_p and ψ_α norms of a given random variable. Explicitly, the following estimate holds:

$$\forall \alpha \geq 1, \frac{1}{2e^2} \|\cdot\|_{\psi_\alpha} \leq \sup_{p \geq \alpha} \frac{(\mathbb{E}|\cdot|^p)^{1/p}}{p^{1/\alpha}} \leq 2e \|\cdot\|_{\psi_\alpha}$$

Furthermore: $\forall \alpha \geq 1$, $L_\infty \subset L_{\psi_\alpha}$ and $\|\cdot\|_{\psi_\alpha} \leq \|\cdot\|_\infty$.

Hence, the large deviation inequalities that one gets for sums of independent ψ_α random variables may be seen as generalizations of the ‘‘classical’’ Hoeffding-type deviation inequalities that one has for sums of independent bounded random variables (theorem C.6).

For instance in the sub-exponential case, the following Bernstein-type deviation inequality holds:

Theorem C.8 (*Bernstein’s inequality*)

Let X_1, \dots, X_n be n independent centered ψ_1 random variables.

Setting $M := \max_{1 \leq k \leq n} \|X_k\|_{\psi_1}$ and $\sigma^2 := \frac{1}{n} \sum_{1 \leq k \leq n} \|X_k\|_{\psi_1}^2$, we have:

$$\forall t > 0, \mathbb{P} \left(\left| \frac{1}{n} \sum_{k=1}^n X_k \right| \geq t \right) \leq 2 \exp \left(-\frac{n}{2(2e-1)} \min \left(\frac{t^2}{\sigma^2}, \frac{t}{M} \right) \right)$$

C.3 Tail bounds for sums of random matrices

Regarding the theory of large deviations for real-valued random variables, a standard reference is [13]. For its extension to matrix-valued random variables, one might refer to [15] or [16].

Let us define the *Kullback-Leibler divergence* between $0 \leq x \leq 1$ and $0 \leq y \leq 1$ as:

$$D(x||y) := x \log \left(\frac{x}{y} \right) + (1-x) \log \left(\frac{1-x}{1-y} \right)$$

Theorem C.9 (*Chernoff’s inequality*)

Let X_1, \dots, X_n be n independent real-valued random variables.

Assume that: $\forall 1 \leq k \leq n$, $\begin{cases} 0 \leq X_k \leq 1 \\ \mathbb{E}X_k = \mu_k \text{ with } \mu \leq \mu_k \leq \mu' \end{cases}$. Then:

$$\forall \epsilon > 0, \mathbb{P} \left(\frac{1}{n} \sum_{k=1}^n X_k \geq \mu' + \epsilon \right) \leq e^{-nD(\mu'+\epsilon||\mu')} \text{ and } \mathbb{P} \left(\frac{1}{n} \sum_{k=1}^n X_k \leq \mu - \epsilon \right) \leq e^{-nD(\mu-\epsilon||\mu)}$$

Theorem C.10 (*Matrix Chernoff’s inequality*)

Let X_1, \dots, X_n be n independent $\mathcal{H}(\mathbb{C}^d)$ -valued random variables.

Assume that: $\forall 1 \leq k \leq n$, $\begin{cases} 0 \leq X_k \leq 1 \\ \mathbb{E}X_k = M_k \text{ with } \mu 1 \leq M_k \leq \mu' 1 \end{cases}$. Then:

$$\forall \epsilon > 0, \mathbb{P} \left(\frac{1}{n} \sum_{k=1}^n X_k \geq (\mu' + \epsilon) 1 \right) \leq de^{-nD(\mu'+\epsilon||\mu')} \text{ and } \mathbb{P} \left(\frac{1}{n} \sum_{k=1}^n X_k \leq (\mu - \epsilon) 1 \right) \leq de^{-nD(\mu-\epsilon||\mu)}$$

Noticing that $D(\eta(1+\delta)||\eta) \geq \frac{\eta\delta^2}{4}$ for $-\frac{1}{2} \leq \delta \leq \frac{1}{2}$, we get as corollaries of Chernoff’s inequalities:

Corollary C.11 Let X_1, \dots, X_n be n independent real-valued random variables.

Assume that: $\forall 1 \leq k \leq n$, $\begin{cases} 0 \leq X_k \leq R \\ \mathbb{E}X_k = \mu_k \geq \mu 1 \end{cases}$. Setting $M := \frac{1}{n} \sum_{k=1}^n M_k$, we then have:

$$\forall 0 < \delta < \frac{1}{2}, \mathbb{P} \left(\frac{1}{n} \sum_{k=1}^n X_k \notin [(1-\delta)M; (1+\delta)M] \right) \leq 2e^{-n\mu\delta^2/4R^2}$$

Corollary C.12 Let X_1, \dots, X_n be n independent $\mathcal{H}(\mathbb{C}^d)$ -valued random variables.

Assume that: $\forall 1 \leq k \leq n$, $\begin{cases} 0 \leq X_k \leq R1 \\ \mathbb{E}X_k = M_k \geq \mu 1 \end{cases}$. Setting $M := \frac{1}{n} \sum_{k=1}^n M_k$, we then have:

$$\forall 0 < \delta < \frac{1}{2}, \mathbb{P} \left(\frac{1}{n} \sum_{k=1}^n X_k \notin [(1 - \delta)M; (1 + \delta)M] \right) \leq 2de^{-n\mu\delta^2/4R^2}$$

D Geometry of quantum states

The idea of having a geometric approach to questions concerning, for instance, the purity and the separability of high dimensional quantum systems (consult section 1 for further information) is a quite recent one. [26] provides a concise introduction to this trend.

D.1 Separability

Let $\mathbb{H}_i \equiv \mathbb{C}^{d_i}$, $1 \leq i \leq K$, be K finite-dimensional complex Hilbert spaces, and denote by $\mathbb{H} = \mathbb{H}_1 \otimes \dots \otimes \mathbb{H}_K$ their tensor product complex Hilbert space ($\mathbb{H} \equiv \mathbb{C}^D$ where $D = d_1 \times \dots \times d_K$). Equipped with the Hilbert-Schmidt inner product, the Hilbert space $\mathcal{H}(\mathbb{H})$ of Hermitians on \mathbb{H} inherits a real D^2 -dimensional euclidean structure. When later speaking about volumes of subsets of $\mathcal{H}(\mathbb{H})$, we will always refer to the corresponding D^2 -dimensional Lebesgue volume on $\mathcal{H}(\mathbb{H})$ (or to the induced D' -dimensional Lebesgue volume on subspaces of $\mathcal{H}(\mathbb{H})$ of real dimension $D' < D^2$).

The set of *states* on \mathbb{H} is defined as:

$$\mathcal{D}(\mathbb{H}) := \{ \rho \in \mathcal{H}(\mathbb{H}), \rho \geq 0, \text{Tr}\rho = 1 \} = \text{Conv}\{ |\psi\rangle\langle\psi|, |\psi\rangle \in \mathbb{H}, \langle\psi|\psi\rangle = 1 \}$$

The set of *separable states* on \mathbb{H} is defined as:

$$\begin{aligned} \mathcal{S}(\mathbb{H}_1 : \dots : \mathbb{H}_K) &:= \text{Conv}\{ \rho_1 \otimes \dots \otimes \rho_K, \rho_i \in \mathcal{D}(\mathbb{H}_i), 1 \leq i \leq K \} \\ &= \text{Conv}\{ |\psi_1\rangle\langle\psi_1| \otimes \dots \otimes |\psi_K\rangle\langle\psi_K|, |\psi_i\rangle \in \mathbb{H}_i, \langle\psi_i|\psi_i\rangle = 1, 1 \leq i \leq K \} \end{aligned}$$

Note that the definition of $\mathcal{S}(\mathbb{H}_1 : \dots : \mathbb{H}_K)$, contrary to the one of $\mathcal{D}(\mathbb{H})$, brings in the *local* structure of the *global* Hilbert space \mathbb{H} (i.e. its particular tensor product decomposition).

Both $\mathcal{D}(\mathbb{H})$ and $\mathcal{S}(\mathbb{H}_1 : \dots : \mathbb{H}_K)$ are convex subsets of $\mathcal{H}(\mathbb{H})$ of real dimension $D^2 - 1$ (they belong to the hyperplane of $\mathcal{H}(\mathbb{H})$ of trace 1 Hermitians).

$\mathcal{D}(\mathbb{H})$ is invariant under conjugacy by unitaries: $\forall U \in \mathfrak{U}(\mathbb{H}), \forall \rho \in \mathcal{D}(\mathbb{H}), U\rho U^\dagger = \rho$

$\mathcal{S}(\mathbb{H}_1 : \dots : \mathbb{H}_K)$ is invariant under conjugacy by local unitaries:

$$\forall U_1 \in \mathfrak{U}(\mathbb{H}_1), \dots, U_K \in \mathfrak{U}(\mathbb{H}_K), \forall \rho \in \mathcal{S}(\mathbb{H}_1 : \dots : \mathbb{H}_K), (U_1 \otimes \dots \otimes U_K)\rho(U_1^\dagger \otimes \dots \otimes U_K^\dagger) = \rho$$

Theorem D.1 $\frac{1}{D}$, the so-called *maximally mixed state* on \mathbb{H} , is the only element of $\mathcal{D}(\mathbb{H})$ that is fixed by all the isometries of $\mathcal{D}(\mathbb{H})$, and also the only element of $\mathcal{S}(\mathbb{H}_1 : \dots : \mathbb{H}_K)$ that is fixed by all the isometries of $\mathcal{S}(\mathbb{H}_1 : \dots : \mathbb{H}_K)$.

It is thus the center of gravity of both $\mathcal{D}(\mathbb{H})$ and $\mathcal{S}(\mathbb{H}_1 : \dots : \mathbb{H}_K)$.

Proof: Multiples of the identity are the only operators which are fixed by all unitaries.

Since all unitaries are isometries of $\mathcal{D}(\mathbb{H})$, it is thus clear that $\frac{1}{D}$ is the only operator in $\mathcal{D}(\mathbb{H})$ which is fixed by all the isometries of $\mathcal{D}(\mathbb{H})$ (the multiplicative factor being imposed by the trace constraint).

Regarding $\mathcal{S}(\mathbb{H}_1 : \dots : \mathbb{H}_K)$, things are not as obvious.

Let $\rho := \sum_{j \in J} \lambda_j \rho_j^{(1)} \otimes \dots \otimes \rho_j^{(K)}$ be an operator in $\mathcal{S}(\mathbb{H}_1 : \dots : \mathbb{H}_K)$ which is fixed by all the isometries

of $\mathcal{S}(\mathbb{H}_1 : \dots : \mathbb{H}_K)$.

In particular, ρ is fixed by all local unitaries, and consequently:

$$\begin{aligned} \rho &= \mathbb{E}_{U_i \sim \mathcal{U}(\mathfrak{u}(\mathbb{H}_i)), 1 \leq i \leq K} \left[(U_1 \otimes \dots \otimes U_K) \rho (U_1^\dagger \otimes \dots \otimes U_K^\dagger) \right] \\ &= \sum_{j \in J} \lambda_j \mathbb{E}_{U_1 \sim \mathcal{U}(\mathfrak{u}(\mathbb{H}_1))} \left[U_1 \rho_j^{(1)} U_1^\dagger \right] \otimes \dots \otimes \mathbb{E}_{U_K \sim \mathcal{U}(\mathfrak{u}(\mathbb{H}_K))} \left[U_K \rho_j^{(K)} U_K^\dagger \right] \\ &= \sum_{j \in J} \lambda_j \frac{1}{d_1} \otimes \dots \otimes \frac{1}{d_K} \\ &= \frac{1}{D} \end{aligned}$$

the next before last equality being due to the fact that: $\forall A \in \mathcal{H}(\mathbb{C}^d)$, $\mathbb{E}_{U \sim \mathcal{U}(\mathfrak{u}(\mathbb{C}^d))} [U A U^\dagger] = \frac{\text{Tr } A}{d} \mathbb{1}$.

The volume of $\mathcal{D}(\mathbb{H})$ may be computed exactly:

Theorem D.2 (*Volume of the states on \mathbb{C}^D*)

$$\text{Vol}(\mathcal{D}(\mathbb{C}^D)) = \sqrt{D} (2\pi)^{D(D-1)/2} \frac{\Gamma(1) \times \dots \times \Gamma(D)}{\Gamma(D^2)}$$

As a consequence, the volume-radius of $\mathcal{D}(\mathbb{C}^D)$ satisfies: $\text{vrad}(\mathcal{D}(\mathbb{C}^D)) \underset{D \rightarrow +\infty}{=} \frac{1}{e^{1/4} \sqrt{D}} (1 + O(\frac{1}{D}))$.

The reader is referred to [23] for a complete and rigorous proof.

Nevertheless, it is perhaps' worth mentioning that the root idea underlying such result is that any state ρ on \mathbb{C}^D may be unitarily diagonalized: $\rho = U \Lambda U^\dagger$, where $\Lambda = \text{Diag}(\lambda_1(\rho), \dots, \lambda_D(\rho))$ is ρ 's diagonal *matrix of eigenvalues*, and U is ρ 's unitary *matrix of eigenvectors*.

Due to the hermiticity condition $\forall 1 \leq k \leq D$, $\lambda_k(\rho) \in \mathbb{R}$ and to the trace condition $\sum_{k=1}^D \lambda_k(\rho) = 1$, Λ is determined by $D - 1$ real parameters.

Due to the phase invariance $U \equiv VU$ for any diagonal unitary V (in the generic case of a non-degenerate spectrum), U is determined by $D^2 - D$ real parameters.

Denoting by \mathcal{L}^D the space of eigenvalue D -tuples and by \mathcal{U}^D the space of eigenvector D -tuples, we thus get: $\text{Vol}(\mathcal{D}(\mathbb{H})) = \sqrt{D} \text{Vol}(\mathcal{L}^D) \text{Vol}(\mathcal{U}^D)$.

Now: $\text{Vol}(\mathcal{L}^D) = \text{Vol}((S_1^D)^+) = \frac{1}{D!} \frac{\prod_{k=0}^{D-1} \Gamma(D-k) \Gamma(D-k+1)}{\Gamma(N^2)}$ and $\text{Vol}(\mathcal{U}^D) = \frac{\text{Vol}(\mathfrak{u}(D))}{[\text{Vol}(\mathfrak{u}(1))]^D} = \frac{(2\pi)^{D(D-1)/2}}{1! \times \dots \times (D-1)!}$.

Which leads in the end to the advertized result.

Regarding the volume-radius of $\mathcal{S}(\mathbb{H}_1 : \dots : \mathbb{H}_K)$, in the special case when all the \mathbb{H}_i , $1 \leq i \leq K$, have same dimension, we have the following estimate (established in [22]):

Theorem D.3 (*Volume of the set of separable states on $(\mathbb{C}^d)^{\otimes K} \equiv \mathbb{C}^D$*)

$$\exists c, c' > 1 : \frac{c^{-K}}{D^{1-1/2K}} \leq \text{vrad} \left(\mathcal{S} \left((\mathbb{C}^d)^{\otimes K} \right) \right) \leq \frac{c' \sqrt{K \log K}}{D^{1-1/2K}}$$

where the constants c, c' depend neither on d nor on K .

Let us give an outline of the groundwork this result relies on.

One general technique to handle with a n -dimensional convex body S into an underlying $n + 1$ -dimensional vector space is to look first at its so-called *symmetrization* $\Sigma := \text{Conv}(S \cup -S)$, which is a $n + 1$ -dimensional convex body, instead of working directly with S . Indeed, knowing for instance the $n + 1$ -dimensional volume of Σ will provide us with a quite accurate estimation of the n -dimensional volume of S , as the following theorem specifies it:

Theorem D.4 (*Rogers-Shephard inequality*)

Let $H \subset \mathbb{R}^{n+1}$ be an hyperplane such that $h := \inf_{x \in H} \|x\|_2 > 0$ and $S \subset H$ be a convex body. Denoting by $\Sigma := \text{Conv}(S \cup -S)$ the symmetrization of S , we have:

$$2h \text{Vol}_n(S) \leq \text{Vol}_{n+1}(\Sigma) \leq 2h \frac{2^n}{n+1} \text{Vol}_n(S)$$

In the specific case we are interested in, namely $S = \mathcal{S}(H)$, we have $n = D^2 - 1$ and $h = \frac{1}{\sqrt{D}}$, so that we get the estimates: $\frac{D^{3/2}}{2D^2} \text{Vol}(\Sigma(H)) \leq \text{Vol}(\mathcal{S}(H)) \leq \frac{D^{1/2}}{2} \text{Vol}(\Sigma(H))$.

Hence, passing to the volume-radii: $\frac{1}{2} \text{vrad}(\Sigma(H)) \lesssim \text{vrad}(\mathcal{S}(H)) \lesssim \text{vrad}(\Sigma(H))$.

To obtain an upper-bound on the volume of $\Sigma((\mathbb{C}^d)^{\otimes K})$, one possible strategy might be the following: Let $0 < \delta < \frac{1}{4}$ and consider \mathcal{M}_δ a δ -net for $\|\cdot\|_2$ within the complex 2-norm unit sphere $S_2^d(\mathbb{C})$.

Then, defining $\mathcal{P}(\mathcal{M}_\delta, K) := \{\pm |\psi_1\rangle\langle\psi_1| \otimes \cdots \otimes |\psi_K\rangle\langle\psi_K|, |\psi_i\rangle \in \mathcal{M}_\delta, 1 \leq i \leq K\}$, one may check that: $\Sigma((\mathbb{C}^d)^{\otimes K}) \subset \frac{1}{(1-2\delta^2)^K} \text{Conv}\{\mathcal{P}(\mathcal{M}_\delta, K)\}$.

Now, since $S_2^d(\mathbb{C}) \equiv S_2^{2d}(\mathbb{R})$, \mathcal{M}_δ may be chosen such that $|\mathcal{M}_\delta| \leq (1 + \frac{2}{\delta})^{2d}$ (cf appendix A.4), so that: $|\mathcal{P}(\mathcal{M}_\delta, K)| \leq 2|\mathcal{M}_\delta|^K \leq 2(1 + \frac{2}{\delta})^{2Kd}$.

Consequently, by theorem B.4: $\text{vrad}(\text{Conv}\{\mathcal{P}(\mathcal{M}_\delta, K)\}) \leq \frac{\sqrt{2 \times \log |\mathcal{P}(\mathcal{M}_\delta, K)|}}{\sqrt{(d^K)^2}} \leq \frac{2\sqrt{2K \log(1 + \frac{2}{\delta})}}{d^{K-1/2}}$.

And hence in the end: $\text{vrad}(\Sigma((\mathbb{C}^d)^{\otimes K})) \leq \inf_{0 < \delta < 1/4} \frac{1}{(1-2\delta^2)^K} \text{vrad}(\text{Conv}\{\mathcal{P}(\mathcal{M}_\delta, K)\}) \leq \frac{c' \sqrt{K \log K}}{d^{K-1/2}}$ (choosing for instance $\delta = \frac{1}{2K}$).

To obtain a lower-bound on the volume of $\Sigma((\mathbb{C}^d)^{\otimes K})$, one may reason as follows:

Defining $\Omega((\mathbb{C}^d)^{\otimes K}) := \text{Conv}\{|\psi_1\rangle\langle\phi_1| \otimes \cdots \otimes |\psi_K\rangle\langle\phi_K|, |\psi_i\rangle, |\phi_i\rangle \in B_2^d(\mathbb{C}), 1 \leq i \leq K\}$, and denoting by Π the projection onto Hermitian part, it may be shown that: $\frac{1}{h_K} \Pi(\Omega((\mathbb{C}^d)^{\otimes K})) \subset \Sigma((\mathbb{C}^d)^{\otimes K})$, with $h_K \leq 6^{K/2}$.

Now, $\text{Conv}\{|\psi\rangle\langle\phi|, |\psi\rangle, |\phi\rangle \in B_2^d(\mathbb{C})\}$ may be identified with $(B_2^d(\mathbb{C}))^{\otimes 2}$, providing the identifications: $\Omega((\mathbb{C}^d)^{\otimes K}) \equiv (B_2^d(\mathbb{C}))^{\otimes 2K}$ and $\Pi(\Omega((\mathbb{C}^d)^{\otimes K})) \equiv (B_2^d(\mathbb{R}))^{\otimes 2K}$.

Moreover, for any $m, k \in \mathbb{N}^*$, one has the inclusion: $\frac{1}{m^{(k-1)/2}} B_2^{mk}(\mathbb{R}) \subset (B_2^m(\mathbb{R}))^{\otimes k}$.

So eventually: $\frac{1}{6^{K/2}} \frac{1}{d^{(2K-1)/2}} B_2^{d^{2K}}(\mathbb{R}) \subset \Sigma((\mathbb{C}^d)^{\otimes K})$, which implies: $\frac{\sqrt{6}^{-K}}{d^{K-1/2}} \leq \text{vrad}(\Sigma((\mathbb{C}^d)^{\otimes K}))$.

Concerning the in-radii of $\mathcal{D}(H)$ and $\mathcal{S}(H_1 : \cdots : H_K)$ we have (see [31] and [32]):

$$\text{inrad}(\mathcal{D}(H)) = \frac{1}{\sqrt{D(D-1)}} \quad \text{and} \quad \text{inrad}(\mathcal{S}(H_1 : \cdots : H_K)) \geq \frac{1}{2^{K/2-1} \sqrt{D(D-2^{2-K})}}$$

In the bi-partite case, it thus stands quite remarkably that: $\text{inrad}(\mathcal{S}(H_1 : H_2)) = \text{inrad}(\mathcal{D}(H_1 \otimes H_2))$.

D.2 Random states

Most of the material exposed in this section may be found wrapped and proved in [24] and [25].

Let $H \equiv \mathbb{C}^d$ be a finite-dimensional Hilbert space. One may be interested in looking at various properties of a ‘‘typical’’ state on H . Nevertheless, it is known that there is no distinguished probability distribution over the set $\mathcal{D}(H)$ of density operators on H , so that defining what a ‘‘random’’ state on H would be might be tricky.

The question however does not arise when only looking at pure states: there is indeed a distinguished probability distribution over the unit vectors of H , namely the uniform one (unique normalized Haar measure over $S_2^d(\mathbb{C})$).

Therefore, one option if one is to define random states on H would be to consider a bigger composite Hilbert space $H \otimes H' \equiv \mathbb{C}^d \otimes \mathbb{C}^{d'}$ and to look at random mixed states on H obtained by partial-tracing

on \mathbb{H}' random pure states on $\mathbb{H} \otimes \mathbb{H}'$: $\rho = \text{Tr}_{\mathbb{H}'}(|\psi\rangle\langle\psi|)$, where $|\psi\rangle$ is drawn according to the uniform probability distribution over the unit vectors of $\mathbb{H} \otimes \mathbb{H}'$.

Let us introduce first a few notations:

- For all $d \in \mathbb{N}^*$, we denote by $\sigma_{FS}(d)$ the probability measure over the unit vectors of \mathbb{C}^d induced by the Fubini-Study distance $d_{FS}(|\psi\rangle, |\tilde{\psi}\rangle) := \|\tilde{|\psi\rangle} - |\psi\rangle\|_2$.
- For all $d \in \mathbb{N}^*$, we denote by $\mu_{HS}(d)$ the probability measure over the density operators on \mathbb{C}^d induced by the Hilbert-Schmidt distance $d_{HS}(\rho, \tilde{\rho}) := \|\rho - \tilde{\rho}\|_2$.
- For all $d, d' \in \mathbb{N}^*$, we denote by $\mu_{d,d'}$ the probability law of the normalized (d, d') -Wishart matrices, that is $G \sim \mathcal{N}_{\mathbb{C}}^{d \times d'}(0, 1) \Rightarrow \frac{GG^\dagger}{\text{Tr}(GG^\dagger)} \sim \mu_{d,d'}$ (consult appendix B.3 for definitions and notations regarding complex matrix-valued Gaussian variables).

Theorem D.5 *Let $d, d' \in \mathbb{N}^*$, and suppose $|\psi\rangle \sim \sigma_{FS}(d \times d')$. Then, identifying $\mathbb{C}^{d \times d'}$ with $\mathbb{C}^d \otimes \mathbb{C}^{d'}$: $\text{Tr}_{\mathbb{C}^{d'}}(|\psi\rangle\langle\psi|) \sim \mu_{d,d'}$.*

The following important theorem illustrates how the Hilbert-Schmidt measure may be viewed as arising from the Fubini-Study measure.

Theorem D.6 *For all $d \in \mathbb{N}^*$, $\mu_{d,d} = \mu_{HS}(d)$.*

Given $M \in \mathcal{H}(\mathbb{C}^d)$ we denote by $\lambda(M) = (\lambda_1(M), \dots, \lambda_d(M)) \in \mathbb{R}^d$ its eigenvalue-vector. Let $d \in \mathbb{N}^*$ and $s \geq d$. Denoting by $P_{\mu_{d,s}} : (\mathbb{R}^+)^d \rightarrow [0; 1]$ the joint probability distribution of $\lambda(\rho)$ for $\rho \sim \mu_{d,s}$, we have:

$$P_{\mu_{d,s}}(\lambda) = \frac{\Gamma(ds)}{\prod_{0 \leq j \leq d-1} \Gamma(s-j)\Gamma(d-j+1)} \delta\left(1 - \sum_{1 \leq k \leq d} \lambda_k\right) \prod_{1 \leq k \leq d} \lambda_k^{s-d} \prod_{1 \leq k < k' \leq d} (\lambda_k - \lambda_{k'})^2$$

From this formula, one may derive the expectation values of various quantities which are computed from a state's spectrum.

For instance, the mean value of a random state's purity is: $\mathbb{E}_{\rho \sim \mu_{d,s}}[\text{Tr}(\rho^2)] = \frac{d+s}{ds+1}$.

And more generally, for any $q > 0$, the mean value of a random state's q -order moment has the following asymptotics: $\mathbb{E}_{\rho \sim \mu_{d,d}}[\text{Tr}(\rho^q)] \underset{d \rightarrow +\infty}{=} d^{1-q} \frac{\Gamma(1+2q)}{\Gamma(1+q)\Gamma(2+q)} (1 + O(\frac{1}{d}))$.

E Some properties of a family of norms

We consider here the case when $\mathbb{H} = \mathbb{H}_1 \otimes \dots \otimes \mathbb{H}_K \equiv (\mathbb{C}^d)^{\otimes K}$.

For all $p \in \mathbb{N}^*$, we define the norm $\|\cdot\|_{p[K]}$ on $\mathcal{H}(\mathbb{H})$ by:

$$\forall \Delta \in \mathcal{H}(\mathbb{H}), \|\Delta\|_{p[K]} := \left(\int_{\substack{|\psi_i\rangle \in \mathbb{H}_i, \\ 1 \leq i \leq K}} |\text{Tr}(|\psi_1\rangle\langle\psi_1| \otimes \dots \otimes |\psi_K\rangle\langle\psi_K| \Delta)|^p d\psi_1 \dots d\psi_K \right)^{1/p}$$

We may point out in particular that the norm $\|\cdot\|_{2[K]}$ is related to the ‘‘modified K -partite 2-norm’’

$$\|\cdot\|_{2(K)} := \sqrt{\sum_{I \subset \{1, \dots, K\}} \text{Tr}_{\mathbb{H} \setminus \mathbb{H}_I} \left[(\text{Tr}_{\mathbb{H}_I}[\cdot])^2 \right]} \text{ by: } \|\cdot\|_{2[K]} = \frac{1}{[d(d+1)]^{K/2}} \|\cdot\|_{2(K)}.$$

E.1 Special case $p = 2q$ even

Theorem E.1 For all $q \in \mathbb{N}^*$ and all $\Delta \in \mathcal{H}(\mathbb{H})$:

$$\|\Delta\|_{2q[K]} \leq \left(\frac{d^q (d+1)^q}{d \times \dots \times (d+2q-1)} (2q)! \right)^{K/2q} \|\Delta\|_{2[K]} \underset{q \rightarrow \infty}{\sim} \left(\frac{2q}{e} \right)^K \|\Delta\|_{2[K]}$$

Let us notice that, setting $U_\sigma := \bigotimes_{i=1}^K U_{\sigma_i}$ for every K -tuple $\sigma = (\sigma_1, \dots, \sigma_K) \in \mathfrak{S}_{2q}^K$, we have:

$$\left(\|\Delta\|_{2q[K]} \right)^{2q} = \frac{1}{[d \times \dots \times (d+2q-1)]^K} \text{Tr}_{\mathbb{H}^{\otimes 2q}} \left(\left(\sum_{\sigma \in \mathfrak{S}_{2q}^K} U_\sigma \right) \Delta^{\otimes 2q} \right)$$

Hence, the only thing we actually have to show in order to prove theorem E.1 is:

$$\text{Tr}_{\mathbb{H}^{\otimes 2q}} \left(\left(\sum_{\sigma \in \mathfrak{S}_{2q}^K} U_\sigma \right) \Delta^{\otimes 2q} \right) \leq ((2q)!)^K \|\Delta\|_{2(K)}^{2q} \quad (15)$$

And since \mathfrak{S}_{2q}^K contains $((2q)!)^K$ elements, equation 15 may be obtained by simply showing that, for all $\sigma \in \mathfrak{S}_{2q}^K$, defining $t(\sigma) := |\text{Tr}_{\mathbb{H}^{\otimes 2q}} (U_\sigma \Delta^{\otimes 2q})|$, we have:

$$t(\sigma) \leq \max_{I \subset \{1, \dots, K\}} \left[\text{Tr}_{\mathbb{H} \setminus \mathbb{H}_I} (\text{Tr}_{\mathbb{H}_I} \Delta)^2 \right]^q \leq \sum_{I \subset \{1, \dots, K\}} \left[\text{Tr}_{\mathbb{H} \setminus \mathbb{H}_I} (\text{Tr}_{\mathbb{H}_I} \Delta)^2 \right]^q \leq \left[\sum_{I \subset \{1, \dots, K\}} \text{Tr}_{\mathbb{H} \setminus \mathbb{H}_I} (\text{Tr}_{\mathbb{H}_I} \Delta)^2 \right]^q \quad (16)$$

and then summing over \mathfrak{S}_{2q}^K .

E.1.1 Special case $q = 2$

Since we cannot proceed by inspection of the 24^K K -tuples of \mathfrak{S}_4^K , our first task will be to find a way of restricting our attention to only a few elements of \mathfrak{S}_4 without any loss of generality. In that end, our strategy can be described as follows.

Let $M_1, M_2, M_3, M_4 \in \mathcal{H}(\mathbb{H})$ and $\sigma \in \mathfrak{S}_4$. We write: $\text{Tr}_{\mathbb{H}^{\otimes 4}} (U_\sigma (M_1 \otimes M_2 \otimes M_3 \otimes M_4)) = \text{Tr}_{\mathbb{H}^{\otimes 4}} (XY^\dagger)$,

with $X, Y \in \mathcal{H}(\mathbb{H}^{\otimes 4})$ such that $\exists \sigma', \sigma'' \in \mathfrak{S}_4$:
$$\begin{cases} \text{Tr}_{\mathbb{H}^{\otimes 4}} (XX^\dagger) = \text{Tr}_{\mathbb{H}^{\otimes 4}} (U_{\sigma'} (M_1 \otimes M_2 \otimes M_2 \otimes M_1)) \\ \text{Tr}_{\mathbb{H}^{\otimes 4}} (YY^\dagger) = \text{Tr}_{\mathbb{H}^{\otimes 4}} (U_{\sigma''} (M_4 \otimes M_3 \otimes M_3 \otimes M_4)) \end{cases} .$$

In order to easily visualize into which pair $(\sigma', \sigma'') \in \mathfrak{S}_4 \times \mathfrak{S}_4$ each $\sigma \in \mathfrak{S}_4$ splits, we can make use of Penrose's ingenious tensor diagrams (introduced first in [33]), which we briefly explain here:


Let $M \in \mathcal{H}(\mathbb{H})$ and $|i\rangle, |j\rangle \in \mathbb{H}$. We represent the matrix element $\langle i|M|j\rangle$ by a diagram with terminals:

$$i \text{ --- } \boxed{M} \text{ --- } j$$

Then, summing matrix elements for a unit vector running through an orthonormal basis of \mathbb{H} is represented by joining the corresponding terminals.

Hence for instance, $\text{Tr}_{\mathbb{H}}(M) = \sum_j \langle j|M|j\rangle$ is represented by:

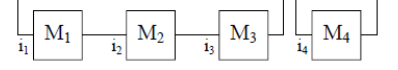
$$j \text{ --- } \boxed{\boxed{M}} \text{ --- } j$$

And in the same way, $\langle i|MN|k\rangle = \sum_j \langle i|M|j\rangle \langle j|N|k\rangle$ is represented by: 

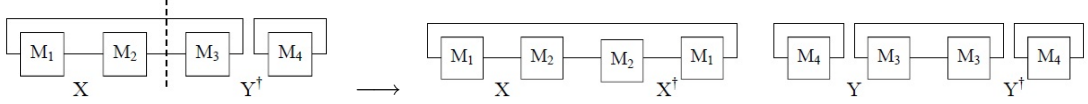
Yet, for any $M_1, M_2, M_3, M_4 \in \mathcal{H}(\mathbb{H})$ and $\sigma \in \mathfrak{S}_4$, we have:

$$\text{Tr}_{\mathbb{H}^{\otimes 4}}(U_\sigma(M_1 \otimes M_2 \otimes M_3 \otimes M_4)) = \sum_{i_1, i_2, i_3, i_4} \langle i_1|M_1|i_{\sigma(1)}\rangle \langle i_2|M_2|i_{\sigma(2)}\rangle \langle i_3|M_3|i_{\sigma(3)}\rangle \langle i_4|M_4|i_{\sigma(4)}\rangle$$

So for example, $\text{Tr}_{\mathbb{H}^{\otimes 4}}(U_{(123)}(M_1 \otimes M_2 \otimes M_3 \otimes M_4))$ is represented by:



And in this case, the splitting procedure described above can be schematically represented by:



which means that $\sigma = (123)$ splits into $\sigma' = (1234)$ and $\sigma'' = (23)$

What we have gained by doing so is that σ' and σ'' cannot be any permutation: they necessarily belong to the subset $\overline{\mathfrak{S}}_4 := \{\text{id}, (14), (23), (1234), (1432), (12)(34), (14)(23)\}$ of \mathfrak{S}_4 containing the permutations that are equal to their opposite under the exchange $1 \leftrightarrow 4$ and $2 \leftrightarrow 3$ (i.e. under the conjugation by $(14)(23)$).

We now have to see more precisely in which pair $(\sigma', \sigma'') \in \overline{\mathfrak{S}}_4 \times \overline{\mathfrak{S}}_4$ each of the elements $\sigma \in \mathfrak{S}_4$ breaks down. First of all, it is clear that the seven elements of $\overline{\mathfrak{S}}_4$ split into twice themselves. Similarly, if σ splits into (σ', σ'') , then its conjugate $(14)(23)\sigma(14)(23)$ splits into (σ'', σ') . We are thus left with actually looking at 9 permutations, one of which being invariant under the conjugation by $(14)(23)$ and the 8 others providing the result for their 8 respective conjugates by switching σ' and σ'' . The resulting splitting map $\text{Split} : \sigma \in \mathfrak{S}_4 \mapsto (\sigma', \sigma'') \in \overline{\mathfrak{S}}_4 \times \overline{\mathfrak{S}}_4$ for each $\sigma \in \mathfrak{S}_4$ can then easily be constructed and looked up in the table of Figure 6.

Let us now turn back to the problem we are dealing with. Let $\sigma = (\sigma_1, \dots, \sigma_K) \in \mathfrak{S}_4^K$. Applying the splitting map $\mathfrak{S}_4 \rightarrow \overline{\mathfrak{S}}_4 \times \overline{\mathfrak{S}}_4$ to all the σ_i , $1 \leq i \leq K$, and then using the Cauchy-Schwarz inequality and the arithmetic-geometric mean inequality, we get:

$$t(\sigma) = |\text{Tr}_{\mathbb{H}^{\otimes 4}}(U_\sigma \Delta^{\otimes 4})| \leq \sqrt{|\text{Tr}_{\mathbb{H}^{\otimes 4}}(U_{\sigma'} \Delta^{\otimes 4})| |\text{Tr}_{\mathbb{H}^{\otimes 4}}(U_{\sigma''} \Delta^{\otimes 4})|} = \sqrt{t(\sigma')t(\sigma'')} \leq \frac{1}{2}t(\sigma') + \frac{1}{2}t(\sigma'') \quad (17)$$

Now, since $\Delta^{\otimes 4}$ is invariant under conjugation by elements of the form $(U_\sigma)^{\otimes K}$, $\sigma \in \mathfrak{S}_4$, it holds that t is invariant under conjugation by elements from the subgroup $G := \{(\sigma, \dots, \sigma), \sigma \in \mathfrak{S}_4\}$ of \mathfrak{S}_4^K .

Yet, we can notice that the subset $\tilde{\mathfrak{S}}_4$ of $\overline{\mathfrak{S}}_4$ defined by $\tilde{\mathfrak{S}}_4 := \{\text{id}, (12)(34), (14)(23)\}$ is such that:

- $\forall \sigma \in \tilde{\mathfrak{S}}_4^K, \forall \varsigma \in G, \text{Split}(\varsigma \sigma \varsigma^{-1}) \in \tilde{\mathfrak{S}}_4^K \times \tilde{\mathfrak{S}}_4^K$
- $\forall \sigma \in \overline{\mathfrak{S}}_4^K, \exists k \in \mathbb{N}^*, \exists \varsigma_1, \dots, \varsigma_k \in G : \text{Split}(\varsigma_k \cdots \text{Split}(\varsigma_1 \sigma \varsigma_1^{-1}) \cdots \varsigma_k) \in (\tilde{\mathfrak{S}}_4^K)^{2k}$

Thus, using equation 17, we see that for any $\sigma \in \mathfrak{S}_4^K$, by repeatedly conjugating by elements of G and splitting, we get in the end the upper bound $t(\sigma) \leq \sum_\alpha p_\alpha t(\sigma^{(\alpha)})$, with certain $p_\alpha = \frac{1}{2^{k_\alpha}}$ that sum to 1, and the $\sigma^{(\alpha)}$ that belong to $\tilde{\mathfrak{S}}_4^K$. Hence eventually:

$$\forall \sigma \in \mathfrak{S}_4^K, t(\sigma) \leq \max_{\pi \in \tilde{\mathfrak{S}}_4^K} t(\pi) \quad (18)$$

In order to upper bound the traces on the right hand side of equation 18, let us deal with the following auxiliary problem.

Conjugacy class	σ	σ'	σ''
(1 ⁴)	id	id	id
(2 ¹ , 1 ²)	(12)	(12)(34)	id
	(13)	(14)	(23)
	(14)	(14)	(14)
	(23)	(23)	(23)
	(24)	(23)	(14)
	(34)	id	(12)(34)
	(12)(34)	(12)(34)	(12)(34)
(2 ²)	(12)(34)	(12)(34)	(12)(34)
	(13)(24)	(14)(23)	(14)(23)
	(14)(23)	(14)(23)	(14)(23)
(3 ¹ , 1 ¹)	(123)	(1234)	(23)
	(132)	(1432)	(23)
	(124)	(1234)	(14)
	(142)	(1432)	(14)
	(134)	(14)	(1234)
	(143)	(14)	(1432)
	(234)	(23)	(1234)
	(243)	(23)	(1432)
	(1234)	(1234)	(1234)
(4 ¹)	(1234)	(1234)	(1234)
	(1243)	(1234)	(1432)
	(1324)	(14)(23)	(14)(23)
	(1342)	(1432)	(1234)
	(1432)	(1432)	(1432)
	(1423)	(14)(23)	(14)(23)

Figure 6: Table of the splitting map $\text{Split} : \mathfrak{S}_4 \longrightarrow \overline{\mathfrak{S}}_4 \times \overline{\mathfrak{S}}_4$, $\text{Split}(\sigma) = (\sigma', \sigma'')$, grouped according to conjugacy classes of σ .

Let $H = A \otimes B \otimes C$ be a finite dimensional 3-partite Hilbert space. For all $P \in \mathcal{H}(H)$ and all $|a\rangle, |a'\rangle \in A$, $|b\rangle, |b'\rangle \in B$ and $|c\rangle, |c'\rangle \in C$ we set $P_{a,b,c}^{a',b',c'} := \langle c| \otimes \langle b| \otimes \langle a| P |a'\rangle \otimes |b'\rangle \otimes |c'\rangle$.

Let $\sigma = (\sigma_A, \sigma_B, \sigma_C) \in \mathfrak{S}_4^3$ be a 3-tuple of permutations. For all $\Delta \in \mathcal{H}(H)$, we have, with the $|a_q\rangle$, $|b_q\rangle$ and $|c_q\rangle$, $1 \leq q \leq 4$, respectively running through an orthonormal basis of A , B and C :

$$\text{Tr}_{H^{\otimes 4}}((U_{\sigma_A} \otimes U_{\sigma_B} \otimes U_{\sigma_C})\Delta^{\otimes 4}) = \sum_{\substack{a_1, b_1, c_1 \\ a_2, b_2, c_2 \\ a_3, b_3, c_3 \\ a_4, b_4, c_4}} \prod_{q=1}^4 \Delta_{a_q, b_q, c_q}^{a_{\sigma_A(q)}, a_{\sigma_B(q)}, c_{\sigma_C(q)}}$$

We now consider the particular case $\sigma_A = \text{id}$, $\sigma_B = (12)(34)$ and $\sigma_C = (14)(23)$, in which we have:

$$\begin{aligned} \text{Tr}_{H^{\otimes 4}}((U_{\sigma_A} \otimes U_{\sigma_B} \otimes U_{\sigma_C})\Delta^{\otimes 4}) &= \sum_{\substack{a_1, b_1, c_1 \\ a_2, b_2, c_2 \\ a_3, b_3, c_3 \\ a_4, b_4, c_4}} \Delta_{a_1 b_1 c_1}^{a_1 b_2 c_4} \Delta_{a_2 b_2 c_2}^{a_2 b_1 c_3} \Delta_{a_3 b_3 c_3}^{a_3 b_4 c_2} \Delta_{a_4 b_4 c_4}^{a_4 b_3 c_1} \\ &= \sum_{\substack{b_1, c_1 \\ b_2, c_2 \\ b_3, c_3 \\ b_4, c_4}} [\text{Tr}_A \Delta]_{b_1, c_1}^{b_2, c_4} [\text{Tr}_A \Delta]_{b_2, c_2}^{b_1, c_3} [\text{Tr}_A \Delta]_{b_3, c_3}^{b_4, c_2} [\text{Tr}_A \Delta]_{b_4, c_4}^{b_3, c_1} \end{aligned}$$

Let us introduce the maximally entangled matrix on $C \otimes C$: $M_{C \otimes C} := \sum_{c, \tilde{c}} |c\rangle \otimes |c\rangle \langle \tilde{c}| \otimes \langle \tilde{c}|$.

Now, let $R := (\text{Tr}_A \Delta \otimes 1_C)(1_B \otimes M_{C \otimes C})(\text{Tr}_A \Delta \otimes 1_C)$.

We notice that, for all $b, b', c, c', \tilde{c}, \tilde{c}'$: $R_{b, c, \tilde{c}}^{b', c', \tilde{c}'} = \sum_{b''} [\text{Tr}_A \Delta]_{b, c}^{b'', \tilde{c}} [\text{Tr}_A \Delta]_{b'', c'}^{b', \tilde{c}'}$.

So that: $\text{Tr}_{H^{\otimes 4}}((U_{\sigma_A} \otimes U_{\sigma_B} \otimes U_{\sigma_C})\Delta^{\otimes 4}) = \sum_{\substack{b_1, b_3 \\ c_1, c_2, c_3, c_4}} R_{b_1, c_1, c_4}^{b_1, c_2, c_3} R_{b_3, c_2, c_3}^{b_3, c_1, c_4} = \text{Tr}_{C \otimes C}(\text{Tr}_B R)^2$.

Yet, defining $P := (\text{Tr}_A \Delta \otimes 1_C) \left(1_B \otimes \sum_c |c\rangle \otimes |c\rangle \right)$, we see that $R = PP^\dagger$. Hence R is a positive matrix, and so is $\text{Tr}_B R$. Thus, using the fact that, for a positive matrix V , $\text{Tr}(V^2) \leq (\text{Tr}V)^2$, we get: $\text{Tr}_{C \otimes C}(\text{Tr}_B R)^2 \leq [\text{Tr}_{B \otimes C \otimes C} R]^2 = [\text{Tr}_{B \otimes C}(\text{Tr}_A \Delta)^2]^2$.

So eventually: $\text{Tr}_{H^{\otimes 4}}((U_{\sigma_A} \otimes U_{\sigma_B} \otimes U_{\sigma_C})\Delta^{\otimes 4}) \leq [\text{Tr}_{H \setminus A}(\text{Tr}_A \Delta)^2]^2$.

We can now turn back to our initial problem.

For all $\pi \in \tilde{\mathfrak{S}}_4^K$, we can define the following factors of the global Hilbert space H :

- $A(\pi) := H_{i_1} \otimes \cdots \otimes H_{i_a}$ with $\pi_{i_1}, \dots, \pi_{i_a} = \text{id}$
- $B(\pi) := H_{i_{a+1}} \otimes \cdots \otimes H_{i_b}$ with $\pi_{i_{a+1}}, \dots, \pi_{i_b} = (12)(34)$
- $C(\pi) := H_{i_{b+1}} \otimes \cdots \otimes H_{i_K}$ with $\pi_{i_{b+1}}, \dots, \pi_{i_K} = (14)(23)$

H can then be written as: $H = A(\pi) \otimes B(\pi) \otimes C(\pi)$.

And hence: $t(\pi) = |\text{Tr}_{H^{\otimes 4}}(U_\pi \Delta^{\otimes 4})| \leq [\text{Tr}_{H \setminus A(\pi)}(\text{Tr}_{A(\pi)} \Delta)^2]^2 \leq \max_{I \subset \{1, \dots, K\}} [\text{Tr}_{H \setminus H_I}(\text{Tr}_{H_I} \Delta)^2]^2$.

Plugging this result in equation 18, we get equation 16 as wanted.

E.1.2 General case $q \geq 2$

We may now extend the method used above. Let us define the following subset of \mathfrak{S}_{2q} , containing the identity and the permutations made of q disjoint transpositions that are invariant under the exchange $j \leftrightarrow 2q+1-j$, $1 \leq j \leq q$, i.e. under the conjugation by the product of transpositions $\prod_{j=1}^q (j, 2q+1-j)$:

$$\tilde{\mathfrak{S}}_{2q} := \left\{ \text{id}, \prod_{k=1}^p (i_k, i'_k)(2q+1-i_k, 2q+1-i'_k) \prod_{l=1}^m (j_l, 2q+1-j_l), 2p+m=q, 1 \leq i_k, i'_k, j_l \leq q \right\}$$

Just as in the special case $q = 2$, letting $G := \{(\sigma, \dots, \sigma), \sigma \in \mathfrak{S}_{2q}\}$, we have that conjugating an element of $\tilde{\mathfrak{S}}_{2q}^K$ by an element of G and then splitting gives two elements of $\tilde{\mathfrak{S}}_{2q}^K$, and that any element of \mathfrak{S}_{2q}^K can be transformed into a tuple of elements of $\tilde{\mathfrak{S}}_{2q}^K$ by repeatedly conjugating by elements of G and splitting.

Thus, by repeated use of the Cauchy-Schwarz inequality and arithmetic-geometric mean inequality, we get that for all $\sigma \in \mathfrak{S}_{2q}^K$: $|\text{Tr}_{H^{\otimes 2q}}(U_\sigma \Delta^{\otimes 2q})| \leq \sum_\alpha p_\alpha |\text{Tr}_{H^{\otimes 2q}}(U_{\sigma^{(\alpha)}} \Delta^{\otimes 2q})|$, with certain $p_\alpha = \frac{1}{2^{k_\alpha}}$

that sum to 1, and the $\sigma^{(\alpha)}$ that belong to $\tilde{\mathfrak{S}}_{2q}^K$. Hence in the end:

$$\forall \sigma \in \mathfrak{S}_{2q}^K, |\text{Tr}_{H^{\otimes 2q}}(U_\sigma \Delta^{\otimes 2q})| \leq \max_{\pi \in \tilde{\mathfrak{S}}_{2q}^K} |\text{Tr}_{H^{\otimes 2q}}(U_\pi \Delta^{\otimes 2q})| \quad (19)$$

Yet, once again similarly to the special case $q = 2$, for all $\pi \in \tilde{\mathfrak{S}}_{2q}^K$, we have the upper bound: $|\text{Tr}_{H^{\otimes 2q}}(U_\pi \Delta^{\otimes 2q})| \leq \max_{I \subset \{1, \dots, K\}} [\text{Tr}_{H \setminus H_I}(\text{Tr}_{H_I} \Delta)^2]^q$, which, plugged into equation 19, gives equation 16 as desired.

E.2 General case $p \geq 2$

Theorem E.1 relates the norm $\|\cdot\|_{p[K]}$ to the norm $\|\cdot\|_{2[K]}$ whenever p is even. One might now wonder what can be said for p odd.

Yet, by Hölder's inequality, we have:

$$\forall p, q, r \in \mathbb{N}^*, \frac{1}{p} + \frac{1}{q} = \frac{1}{r}, \forall \Delta \in \mathcal{H}(\mathbb{H}), \|\Delta\|_{r[K]} \leq \|\Delta\|_{p[K]} \|\Delta\|_{q[K]} \text{ i.e. } \|\Delta\|_{r[K]} \leq \|\Delta\|_{q[K]}$$

Thus: $\forall p \leq p' \in \mathbb{N}^*, \forall \Delta \in \mathcal{H}(\mathbb{H}), \|\Delta\|_{p[K]} \leq \|\Delta\|_{p'[K]}$.

Combining this monotonicity result for $p \mapsto \|\cdot\|_{p[K]}$ to theorem E.1, we finally get:

Theorem E.2 *For all $q \in \mathbb{N}^*$ and all $\Delta \in \mathcal{H}(\mathbb{H})$:*

$$\|\Delta\|_{2q-1[K]} \leq \|\Delta\|_{2q[K]} \leq \left(\frac{d^q (d+1)^q}{d \times \dots \times (d+2q-1)} (2q)! \right)^{K/2q} \|\Delta\|_{2[K]} \underset{q \rightarrow \infty}{\sim} \left(\frac{2q}{e} \right)^K \|\Delta\|_{2[K]}$$

Remark E.3 *It is actually possible to relate the norm $\|\cdot\|_{p[K]}$, $p \geq 2$, to the norm $\|\cdot\|_{2[K]}$ by a completely different approach described in [43]. Indeed, using a hypercontractive inequality of Beckner, one gets that for all $p \geq 2$ and all $\Delta \in \mathcal{H}(\mathbb{H})$:*

$$\|\Delta\|_{p[K]} \leq (p-1)^K \|\Delta\|_{2[K]}$$

This upper bound is however asymptotically worse than the one obtained by the method described above.

These norms occur in many other issues related to quantum information theory than the one of distinguishing quantum states. One example amongst others appears in [44], with the description of a test which tells whether or not a multi-partite quantum state is a product state. The probability of acceptance of the generalized $2q$ -copy product test on the K -partite state ρ described there is:

$$P_{(2q,K)}(\rho) := \left(\frac{d \times \dots \times (d+2q-1)}{(2q)!} \right)^K (\|\rho\|_{2q[K]})^{2q}$$

Using theorem E.1, the latter can be directly related to the probability of acceptance of the generalized 2-copy product test on the K -partite state ρ :

$$P_{(2q,K)}(\rho) \leq \left(2^q \frac{d \times \dots \times (d+2q-1)}{d^q (d+1)^q} \right)^K P_{(2,K)}(\rho) \leq [(2q)!]^K P_{(2,K)}(\rho)$$

References

- [1] **M.A. Nielsen, I.L. Chuang**, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, 2000.
- [2] **J. Watrous**, *Theory of Quantum Information*.
- [3] **W. Feller**, *An Introduction to Probability Theory and its Applications*, Wiley Series in Probability and Mathematical Statistics, New York, 1966.
- [4] **A.S. Holevo**, “Statistical decision theory for quantum systems”, *J. Multivariate Analysis* 3:337-394 (1973).
- [5] **C.W. Helstrom**, *Quantum detection and estimation theory*, Academic Press, New York, 1976.
- [6] **M. Reed, B. Simon**, *Methods of Modern Mathematical Physics - Functional Analysis*, Academic Press, New York, 1972.
- [7] **K. Ball**, *An Elementary Introduction to Modern Convex Geometry*, Flavors of Geometry Volume 31, MSRI Publications (S. Levy editor), Cambridge, 1997.
- [8] **K. Ball**, *Convex Geometry and Functional Analysis*, Handbook of the Geometry of Banach Spaces Volume 1 161-194, Elsevier (W.B Johnson and J. Lindenstrauss editors), Amsterdam, 2001.
- [9] **G. Pisier**, *The Volume of Convex Bodies and Banach Spaces Geometry*, Cambridge Tracts in Mathematics Volume 94, Cambridge University Press, Cambridge, 1989.
- [10] **R. Vershynin**, *Lectures in Geometric Functional Analysis*.
- [11] **G.W. Anderson, A. Guionnet, O. Zeitouni**, *An Introduction to Random Matrices*, Cambridge Studies in Advanced Mathematics Volume 118, Cambridge University Press, Cambridge, 2010.
- [12] **M. Ledoux**, *The Concentration of Measure Phenomenon*, Mathematical Surveys and Monographs Volume 89, American Mathematical Society, Providence, 2001.
- [13] **A. Dembo, O. Zeitouni**, *Large Deviations: Techniques and Applications*, Applications of Mathematics Volume 38, Springer-Verlag, Berlin Heidelberg, 2010.
- [14] **D. Chafaï, O. Guédon, G. Lecué, A. Pajor**, *Interactions between compressed sensing, random matrices and high dimensional geometry*.
- [15] **R. Ahlswede, A. Winter**, “Strong Converse for Identification Via Quantum Channels”, *IEEE Trans. Inf. Theory* (2002); arXiv:quant-ph/001212.
- [16] **J.A. Tropp**, “User-friendly Tail Bounds for Sums of Random Matrices”, *Found. Comput. Math.* 10.1007 (2011); arXiv1004.4389[math.PR].
- [17] **M. Meyer, A. Pajor**, “On the Blaschke-Santaló inequality”, *Archiv der Mathematik* 55(1):82-93 (1990).
- [18] **J. Bourgain, V.D. Milman**, “New volume ratio properties for convex symmetric bodies in \mathbb{R}^n ”, *Invent. Math.* 88:319-340 (1987).
- [19] **V.D. Milman, A. Pajor**, “Entropy and asymptotic geometry of non-symmetric convex bodies”, *Advances in Math.* 152:314-335 (2000).
- [20] **J. Saint Raymond**, “Le volume des idaux d’opérateurs classiques”, *ICM Studia Mathematica T LXXX* (1984).
- [21] **S.J. Szarek**, “Nets of Grassmann manifold and the orthogonal group”, *Proceedings of Banach Space Workshop* 169-185, University of Iowa Press, 1982.

- [22] **G. Aubrun, S.J. Szarek**, “Tensor product of convex sets and the volume of separable states on N qudits”, *Phys. Rev. A.* 73 (2006); arXiv:quant-ph/0503221.
- [23] **H-J. Sommers, K. Życzkowski**, “Hilbert-Schmidt volume of the set of mixed quantum states”, *J. Phys. A.* 36:10115-10130 (2003); arXiv:quant-ph/0302197.
- [24] **H-J. Sommers, K. Życzkowski**, “Induced measures in the space of mixed quantum states”, *J. Phys. A.* 34:7111-7124 (2001); arXiv:quant-ph/0012101.
- [25] **H-J. Sommers, K. Życzkowski**, “Statistical properties of random density matrices”, *J. Phys. A.* 37:8457-8466 (2004); arXiv:quant-ph/0405031.
- [26] **I. Bengtsson, K. Życzkowski**, “An Introduction to Quantum Entanglement: A Geometric Approach”; arXiv:quant-ph/0606228.
- [27] **R. Horodecki, P. Horodecki, M. Horodecki, K. Horodecki**, “Quantum entanglement”, *Rev. Mod. Phys.* 81:865-942 (2009); arXiv:quant-ph/0702225.
- [28] **C.H. Bennett, D.P. DiVincenzo, C.A. Fuchs, T. Mor, E. Rains, P.W. Shor, J.A. Smolin, W.K. Wootters**, “Quantum Nonlocality without entanglement”, *Phys. Rev. A.* 59:1070 (1999); arXiv:quant-ph/9804053.
- [29] **A.M. Childs, D. Leung, L. Mančinska, M. Ozols**, “A framework for bounding nonlocality of state discrimination”; arXiv:1206.5822[quant-ph] (2012).
- [30] **E. Chitambar, D. Leung, L. Mančinska, M. Ozols, A. Winter**, “Everything you always wanted to know about LOCC (but were afraid to ask)”; arXiv:1210.4583[quant-ph] (2012).
- [31] **H.N. Barnum, L. Gurvits**, “Largest separable balls around the maximally mixed bipartite quantum states”, *Phys. Rev. A.* 66:062311 (2002); arXiv:quant-ph/0204159.
- [32] **H.N. Barnum, L. Gurvits**, “Separable balls around the maximally mixed multipartite quantum states”, *Phys. Rev. A.* 68:042312 (2003); arXiv:quant-ph/0302102.
- [33] **R. Penrose, W. Rindler**, *Spinors and Space-Time, Volume 1: Two-spinor calculus and relativistic fields*, Cambridge University Press, Cambridge, 1986.
- [34] **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].
- [35] **C. Lancien, A. Winter**, “Distinguishing multi-partite states by local measurements”; arXiv[quant-ph]:1206.2884.
- [36] **G. Zauner**, “Quantum designs: Foundations of a noncommutative design theory”, *International Journal of Quantum Information* 9(1):445-507 (2011).
- [37] **A.J. Scott**, “Tight informationally complete quantum measurements”, *J. Phys. A.* 39:13507-13526 (2006); arXiv:quant-ph/0604049.
- [38] **A. Ambainis, J. Emerson**, “Quantum t-designs: t-wise independence in the quantum world”, *Proc. 22nd IEEE Conf. Computational Complexity (CCC007)* 129-140, Piscataway, NJ, 2007; arXiv:quant-ph/0701126.
- [39] **T. Eggeling, R.F. Werner**, “Hiding classical data in multi-partite quantum states”, *Phys. Rev. Lett.* 89:097905 (2002); arXiv:quant-ph/0203004.
- [40] **D.P. DiVincenzo, D. Leung, B.M. Terhal**, “Hiding Bits in Bell States”, *Phys. Rev. Lett.* 86(25):5807-5810 (2001); arXiv:quant-ph/0011042.
- [41] **D.P. DiVincenzo, D. Leung, B.M. Terhal**, “Quantum Data Hiding”, *IEEE Trans. Inf Theory* 48(3):580-599 (2002); arXiv:quant-ph/0103098.

- [42] **A.W. Harrow, A. Montanaro, A.J. Short**, “Limitations on quantum dimensionality reduction”, Proc. ICALP’11 LNCS 6755 86-97, Springer-Verlag, Berlin Heidelberg, 2011; arXiv[quant-ph]:1012.2262.
- [43] **A. Montanaro**, “Some applications of hypercontractive inequalities in quantum information theory”; arXiv[quant-ph]:1208.0161.
- [44] **A.W. Harrow, A. Montanaro**, “An efficient test for product states, with applications to quantum Merlin-Arthur games”, Proc. 51st ASFCS 633-642, 2010; arXiv[quant-ph]:1001.0017.
- [45] **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.