Quantum information theory (MAT4430) **Spring 2021**

Lecture 14: The Haar measure and Hasting's counterexample

Lecturer: Alexander Müller-Hermes

In the previous lecture, we have seen how to associate quantum channels to subspaces of tensor product spaces. By choosing subspaces in a clever way, we can use such constructions to show that certain operator norms are not additive under tensor products. Unfortunately, such concrete constructions have so far been too weak to give a counterexample to the additivity of the minimum output entropy, and hence the Holevo information. In this lecture, we will show how random subspaces can lead to counterexamples for these conjectures. For this, we will first introduce the Haar measure on the unitary group $\mathcal{U}\left(\mathbb{C}^d\right)$, which is used to construct random subspaces. We will then sketch the main lines of argument behind Hasting's proof.

1 The Haar measure on the unitary group $\mathcal{U}\left(\mathbb{C}^d\right)$

1.1 Preliminaries from measure theory

We will start with some preliminaries from measure theory, which we will specialize to the finite-dimensional setting of complex Euclidean spaces for simplicity. Note that these concepts are also valid in more general settings.

Definition 1.1 (Borel σ-algebras and Borel measures)**.** *Let $\mathcal{A} \subset \mathcal{H}$ denote a subset of a complex Euclidean space $\mathcal{H}$.*

1. *The* Borel σ-algebra $\text{Borel}(\mathcal{A})$ *is the σ-algebra generated by the open subsets of $\mathcal{A}$ in the subset topology inherented from the topology of $\mathcal{H}$. Specifically,* $\text{Borel}(\mathcal{A})$ *is the smallest σ-agebra containing the open subsets of $\mathcal{A}$ and which is closed under taking complements and countable unions. The elements of* $\text{Borel}(\mathcal{A})$ *are called* Borel sets*.*

2. *A* Borel measure *on $\mathcal{A}$ is a function $\mu : \text{Borel}(\mathcal{A}) \rightarrow [0, \infty]$ such that*

$$\mu(\emptyset) = 0,$$

and

$$\mu\left(\bigcup_{k=1}^{\infty} \mathcal{S}_k\right) = \sum_{k=1}^{\infty} \mu(\mathcal{S}_k),$$

for any family $\{\mathcal{S}_k\}_{k=1}^{\infty} \subset \text{Borel}(\mathcal{A})$ of pairwise disjoint Borel sets. A Borel measure on $\mathcal{A}$ is called a probability measure *if $\mu(\mathcal{A}) = 1$.*

Having defined the Borel σ-algebra, we may consider *Borel functions* (or Borel measurable functions) $f : \mathcal{A} \rightarrow \mathcal{B}$ between subsets $\mathcal{A} \subseteq \mathcal{H}$ and $\mathcal{B} \subseteq \mathcal{H}'$ of complex Euclidean spaces $\mathcal{H}$ and $\mathcal{H}'$, i.e., such that $f^{-1}(\mathcal{S}') \in \text{Borel}(\mathcal{A})$ for any $\mathcal{S}' \in \text{Borel}(\mathcal{B})$. In particular, we may consider Borel functions $f : \mathcal{A} \rightarrow \mathbb{C}$. From the definition of the Borel σ-algebra it is easy to see that any continuous function $f : \mathcal{A} \rightarrow \mathcal{B}$ is a Borel function. Given any Borel measure μ on $\mathcal{A}$ and a Borel function $f : \mathcal{A} \rightarrow \mathcal{B}$, we may define a Borel measure ν on $\mathcal{B}$ by setting $\nu(\mathcal{S}) = \mu(f^{-1}(\mathcal{S}))$ for any Borel set $\mathcal{S}$. This measure is called the *pushforward measure*

For any subset $\mathcal{A} \subseteq \mathcal{H}$ and any Borel measure $\mu : \text{Borel}(\mathcal{A}) \rightarrow [0, \infty]$ we will define integrals

$$\int_{\mathcal{A}} f(x) d\mu(x),$$

for certain measurable functions $f : \mathcal{A} \rightarrow \mathbb{R}$ in the usual way:

1. Consider a characteristic function $\chi_{\mathcal{S}} : \mathcal{A} \rightarrow \mathbb{R}$ of some Borel set $\mathcal{S} \in \text{Borel}(\mathcal{A})$ given by
$$\chi_{\mathcal{S}}(x) = \begin{cases} 1, & \text{if } x \in \mathcal{S} \\ 0, & \text{otherwise.} \end{cases}$$
For any such function, we define
$$\int_{\mathcal{A}} \chi_{\mathcal{S}}(x) d\mu(x) = \mu(\mathcal{S}).$$

2. A function $f : \mathcal{A} \rightarrow \mathbb{R}$ is called a *simple function* if it can be written as
$$f = \sum_{k=1}^{K} \alpha_k \chi_{\mathcal{S}_k},$$
with $\alpha_k \geq 0$ and $\mathcal{S}_k \in \text{Borel}(\mathcal{A})$ for each $k \in \{1, \ldots, K\}$. For such a function, we define
$$\int_{\mathcal{A}} f(x) d\mu(x) = \sum_{k=1}^{K} \alpha_k \mu(\mathcal{S}_k).$$

3. For every non-negative Borel function $f : \mathcal{A} \rightarrow [0, \infty)$ we define
$$\int_{\mathcal{A}} f(x) d\mu(x) = \sup_{g \leq f} \int_{\mathcal{A}} g(x) d\mu(x),$$
where the supremum goes over all simple functions $g : \mathcal{A} \rightarrow [0, \infty)$ satisfying $f(x) \leq g(x)$ for every $x \in \mathcal{A}$. We say that f is *integrable* if the supremum in the last equation is finite.

4. A Borel function $f : \mathcal{A} \rightarrow \mathbb{R}$ is called *integrable* if it can be written as $f = f_+ - f_-$ for non-negative integrable Borel functions $f_+, f_- : \mathcal{A} \rightarrow [0, \infty)$ and in this case we define
$$\int_{\mathcal{A}} f(x) d\mu(x) = \int_{\mathcal{A}} f_+(x) d\mu(x) - \int_{\mathcal{A}} f_-(x) d\mu(x).$$
Similarly, a Borel function $f : \mathcal{A} \rightarrow \mathbb{C}$ is called *integrable* if it can be written as $f = g + ih$ for integrable Borel functions $g, h : \mathcal{A} \rightarrow \mathbb{R}$ and in this case we define
$$\int_{\mathcal{A}} f(x) d\mu(x) = \int_{\mathcal{A}} g(x) d\mu(x) + i \int_{\mathcal{A}} h(x) d\mu(x).$$
It can be shown that the values of the integrals defined above are independent of the decompositions chosen for the function f.

The following lemma is an easy consequence of the third point from above:

Lemma 1.2. *If $f : \mathcal{A} \rightarrow [0, \infty)$ is an integrable non-negative Borel function, then we have*
$$\int_{\mathcal{A}} f(x) d\mu(x) \geq 0,$$
for any Borel measure μ.

We will also need the following theorem showing that continuous functions are integrable in the cases we are interested in.

Theorem 1.3. *If $\mathcal{A} \subset \mathcal{H}$ is a compact subset and $\mu : \mathrm{Borel}(\mathcal{A}) \to [0, \infty)$ a finite Borel measure, then any continuous function $f : \mathcal{A} \to \mathbb{C}$ is μ-integrable.*

Proof. For a continuous function $f : \mathcal{A} \to \mathbb{C}$ we define $f_{\max} = \max_{z\in\mathcal{A}} |f(z)|$, which exists by compactness of $\mathcal{A}$ and continuity of f. We may decompose

$$f = g + ih = g_+ - g_- + i(h_+ - h_-),$$

where $g = (f + \overline{f})/2$ and $h = (f - \overline{f})/(2i)$, and $g_+ = g \cdot \chi_{g^{-1}([0,\infty))}$ and $h_+ = h \cdot \chi_{h^{-1}([0,\infty))}$, and $g_- = -(g - g_+)$ and $h_- = -(h - h_+)$. Clearly, all involved functions are Borel functions. Now, note that

$$g_+(x) \leq |\mathrm{Re}\,(f(x))\,| \leq f_{\max},$$

for all $x \in \mathcal{A}$, which shows that

$$\int_{\mathcal{A}} g_+(x) d\mu(x) = \sup_{q\leq g_+} \int_{\mathcal{A}} q(x) d\mu(x) \leq \mu(\mathcal{A}) f_{\max} < \infty,$$

where the supremum goes over all non-negative simple functions $q : \mathcal{A} \to [0, \infty)$ satisfying $q(x) \leq g_+(x) \leq f_{\max}$ for every $x \in \mathcal{A}$. We conclude that g_+ is integrable. In the same way we can show that g_-, h_+ and h_- are also integrable. This shows that f is integrable. □

Finally, we will need operator-valued integrals, which in our (finite-dimensional) setting may be defined component-wise. Specifically, given a subset $\mathcal{A} \subseteq \mathcal{H}$ of a complex Euclidean space $\mathcal{H}$, and another complex Euclidean space $\mathcal{H}'$ we will call a function $f : \mathcal{A} \to B(\mathcal{H}')$ integrable if its entries $f_{kl} : \mathcal{A} \to \mathbb{C}$ in the computational basis given by $f_{kl} = \langle k|f(\cdot)|l\rangle$ are integrable. In this case, we denote by

$$X = \int_{\mathcal{A}} f(x) d\mu(x) \in B(\mathcal{H}'),$$

the operator with entries

$$X_{kl} = \int_{\mathcal{A}} f_{kl}(x) d\mu(x),$$

in the computational basis. We finish this section with another useful observation concerning the integrals of operator-valued functions mapping into a closed cone:

Theorem 1.4. *Consider complex Euclidean spaces $\mathcal{H}$ and $\mathcal{H}'$, a compact subset $\mathcal{A} \subset \mathcal{H}$, and a closed convex cone $\mathcal{C} \subset B(\mathcal{H}')_{sa}$. If $f : \mathcal{A} \to B(\mathcal{H}')$ is continuous and if $f(x) \in \mathcal{C}$ for every $x \in \mathcal{A}$, then*

$$\int_{\mathcal{A}} f(x) d\mu(x) \in \mathcal{C}.$$

Proof. Consider the dual cone $\mathcal{C}^* \subset B(\mathcal{H}')_{sa}$ and recall that

$$\mathcal{C} = \{y \in B(\mathcal{H}')_{sa} \ : \ \langle z, y\rangle_{HS} \geq 0 \ : \ \text{for all } z \in \mathcal{C}^*\}.$$

Since $x \mapsto \langle z, f(x)\rangle_{HS}$ is continuous and non-negative, we conclude by Theorem 1.3 and Lemma 1.2 that

$$\int_{\mathcal{A}} \langle z, f(x)\rangle_{HS} d\mu(x) \geq 0,$$

Using the linearity of the integral we have

$$\langle z, \int_{\mathcal{A}} f(x) d\mu(x)\rangle_{HS} = \int_{\mathcal{A}} \langle z, f(x)\rangle_{HS} d\mu(x) \geq 0,$$

for any $z \in \mathcal{C}^*$ which, by duality, finishes the proof. □

1.2 The Haar measure on $\mathcal{U}\left(\mathbb{C}^d\right)$

We will first state a definition/theorem, which can be seen as the main result of this section.

Definition 1.5 (Haar measure on $\mathcal{U}\left(\mathbb{C}^d\right)$)**.** *For $d \in \mathbb{N}$, the* unitary Haar measure η *is the unique Borel measure on $\mathcal{U}\left(\mathbb{C}^d\right)$ satisfying the following conditions:*

1. ***Normalization:***
$$\eta\left(\mathcal{U}\left(\mathbb{C}^d\right)\right) = 1.$$

2. ***Unitary invariance:***
$$\eta(\mathcal{S}) = \eta\left(\mathcal{S}U\right) = \eta\left(U\mathcal{S}\right),$$
for any Borel subset $\mathcal{S} \subseteq \mathcal{U}\left(\mathbb{C}^d\right)$ and any unitary $U \in \mathcal{U}\left(\mathbb{C}^d\right)$.

We will now go through the main steps to see how the Haar measure can be constructed and why it is unique. Identifying $\mathbb{C}$ with $\mathbb{R}^2$, we may consider the usual Lebesgue measure λ on $\mathbb{C}$, which is a Borel measure. Next, we consider the standard Gaussian measure γ defined by the density function
$$z \mapsto \frac{1}{\pi}e^{-|z|^2},$$
i.e., such that
$$\gamma(\mathcal{S}) = \frac{1}{\pi}\int_{\mathcal{S}} e^{-|z|^2} d\lambda(z),$$
for any Borel set $\mathcal{S}$. Note that γ is a probability measure, which is quantifies the probability of events involving standard Gaussian random variables. Using the product of d^2 of such standard Gaussian measures, we may define a Borel probability measure on $B(\mathbb{C}^d) \simeq \mathbb{C}^{d^2}$ by
$$\Gamma(\mathcal{S}) = \frac{1}{\pi^{d^2}}\int_{\mathcal{S}} \prod_{k,l=1}^{d} e^{-|z_{k,l}|^2} d\Lambda(Z) = \frac{1}{\pi^{d^2}}\int_{\mathcal{S}} e^{\mathrm{Tr}\left[Z^\dagger Z\right]} d\Lambda(Z),$$
where we used the notation $Z = [z_{k,l}]_{k,l=1}^d$ and a simple trace formula. Here, Λ arises as the product measure of d^2 Lebesgue measures λ on $\mathbb{C}$ and hence is the usual Lebesgue measure on $\mathbb{C}^{d^2} \simeq \mathbb{R}^{2d^2}$. Again, we could have chosen a more probabilistic point of view by defining random matrices $M \in B(\mathbb{C}^d)$ with entries M_{kl} chosen i.i.d. according to the standard normal distribution on $\mathbb{C}$. The measure Γ is the probability measure corresponding to these random matrices. In the following, we will need two properties of the measure Γ:

Lemma 1.6. *For any $d \in \mathbb{N}$ let Γ denote the Borel probability measure defined above. Then, we have:*

1. *We have*
$$\Gamma\left(\{M \in B(\mathbb{C}^d) \ : \ \det(M) = 0\}\right) = 0.$$

2. *For any Borel set $\mathcal{S} \in \mathrm{Borel}\left(B(\mathbb{C}^d)\right)$ and any unitary $U \in \mathcal{U}\left(\mathbb{C}^d\right)$ we have*
$$\Gamma\left(U\mathcal{S}\right) = \Gamma\left(\mathcal{S}\right).$$

Before proving the lemma, we will need a result about Lebesgue null sets:

Lemma 1.7. *For a multivariate polynomial $p \in \mathbb{C}\left[x_1, \ldots, x_n\right]$ we consider the set $\mathcal{S} = \{x \in \mathbb{C}^n \ : \ p(x) = 0\}$. Then, the set $\mathcal{S}$ is Borel and we either have $\lambda(\mathcal{S}) = 0$ or $\mathcal{S} = \mathbb{C}^n$.*

Proof. Since the function $x \mapsto p(x)$ is continuous the set $\mathcal{S}$ is closed and hence Borel. In the following, we assume that $p \neq 0$ and we will show that $\lambda(\mathcal{S}) = 0$ in this case. The proof proceeds by induction in n. The statement is certainly true if $n = 1$ since p has an at most finite number of zeros forming a Lebesgue null set. Assume that the statement is true for some $n \in \mathbb{N}$ and consider a non-zero polynomial $p \in \mathbb{C}\left[x_1, \ldots, x_{n+1}\right]$. For $(x, x_{n+1}) \in \mathbb{C}^{n+1}$ we can write

$$p(x, x_{n+1}) = \sum_{i=0}^{k} p_i(x) x_{n+1}^i,$$

with polynomials $p_i \in \mathbb{C}\left[x_1, \ldots, x_n\right]$ for $i \in \{1, \ldots, k\}$. Next, we define two sets:

$$\mathcal{A} = \{(x, x_{n+1}) \in \mathbb{C}^{n+1} \ : \ p_0(x) = p_1(x) = \cdots = p_k(x) = 0\},$$

and

$$\mathcal{B} = \{(x, x_{n+1}) \in \mathbb{C}^{n+1} \ : \ \sum_{i=0}^{k} p_i(x) x_{n+1}^i = 0 \text{ and } p_i(x) \neq 0 \text{ for at least one } i\}.$$

Clearly, we have $\mathcal{S} = \mathcal{A} \cup \mathcal{B}$. By assumption at least one of the polynomials $p_0, \ldots, p_k$ say p_l is non-zero and by the induction hypothesis we can conclude that

$$\begin{aligned}\lambda\left(\mathcal{A}\right) &\leq \lambda\left(\{(x, x_{n+1}) \in \mathbb{C}^{n+1} \ : \ p_l(x) = 0\}\right) \\ &= \int \lambda\left(\{x \in \mathbb{C}^n \ : \ p_l(x) = 0\}\right) d\lambda(x_{n+1}) = 0.\end{aligned}$$

Therefore, we have $\lambda(\mathcal{A}) = 0$.

Consider now the set $\mathcal{B}$. For any $x \in \mathbb{C}^n$ such that $p_i(x) \neq 0$ for some i there are at most finitely many values x_{n+1} such that $\sum_{i=0}^{k} p_i(x) x_{n+1}^i = 0$. We conclude that

$$\begin{aligned}\lambda(\mathcal{B}) = &\int_{\{x \in \mathbb{C}^n \ : \ p_l(x) \neq 0 \text{ for some } l\}} \lambda\left(\{x_{n+1} \in \mathbb{C} \ : \ \sum_{i=0}^{k} p_i(x) x_{n+1}^i = 0\}\right) d\lambda(x) \\ &= 0.\end{aligned}$$

We conclude that

$$\lambda(\mathcal{S}) \leq \lambda(\mathcal{A}) + \lambda(\mathcal{B}) = 0,$$

and hence $\lambda(\mathcal{S}) = 0$. □

Now, we proceed with the proof of the properties of the measure Γ:

Proof of Lemma 1.6. For the first statement, note that $M \mapsto \det(M)$ is a polynomial and therefore the set $\mathcal{S} = \{M \in B(\mathbb{C}^d) \ : \ \det(M) = 0\}$ is Borel (polynomials are continuous). Now, we note that if $\mathcal{S}$ is any set of zeros of a multivariate polynomial, then we have either $\Lambda(\mathcal{S}) = 0$ or $\mathcal{S} = B(\mathbb{C}^d)$, where dZ denotes the usual Lebesgue measure on $B(\mathbb{C}^d)$ viewed as $\mathbb{C}^{d^2}$. Since $\mathcal{S} \neq B(\mathbb{C}^d)$ we have $\Lambda(\mathcal{S}) = 0$ and we conclude that $\Gamma(\mathcal{S}) = 0$ as well since Γ is absolutely continuous with respect to Λ since it is defined from Λ using a density function.

For the second statement consider a Borel set $\mathcal{S} \in \text{Borel}\left(B(\mathbb{C}^d)\right)$ and a unitary $U \in \mathcal{U}\left(\mathbb{C}^d\right)$. Note that the Jacobian of the transformation $X \mapsto UX$ (viewed as a transformation of $\mathbb{C}^{d^2}$ to itself) equals $J_U = U^{\oplus d}$ such that $|\det\left(J_U\right)| = 1$. Now, we can apply the transformation formula for the Lebesgue integral to compute

$$\begin{aligned}\Gamma\left(U\mathcal{S}\right) &= \frac{1}{\pi^{d^2}} \int_{U\mathcal{S}} e^{\text{Tr}\left[Z^\dagger Z\right]} d\Lambda(Z) \\ &= \frac{1}{\pi^{d^2}} \int_{\mathcal{S}} |\det\left(J_U\right)| e^{\text{Tr}\left[(U^\dagger Z)^\dagger (U^\dagger Z)\right]} d\Lambda(Z) \\ &= \frac{1}{\pi^{d^2}} \int_{\mathcal{S}} e^{\text{Tr}\left[Z^\dagger Z\right]} d\Lambda(Z) = \Gamma\left(\mathcal{S}\right).\end{aligned}$$

□

We can now introduce the Haar measure on the unitary group $\mathcal{U}\left(\mathbb{C}^d\right)$.

Definition 1.8 (The Haar measure – concrete construction)**.** *For $d \in \mathbb{N}$, we consider the continous function $F : B(\mathbb{C}^d) \setminus \{M \ : \ \det(M) = 0\} \to \mathcal{U}\left(\mathbb{C}^d\right)$ given by*

$$F(M) = M\left(M^\dagger M\right)^{-1/2}.$$

Then, the Haar measure $\eta : \text{Borel}\left(\mathcal{U}\left(\mathbb{C}^d\right)\right) \to [0,\infty)$ is defined as the pushforward measure of Γ under F, i.e.,

$$\eta\left(\mathcal{S}\right) = \Gamma\left(F^{-1}\left(\mathcal{S}\right)\right),$$

for any $\mathcal{S} \in \text{Borel}\left(\mathcal{U}\left(\mathbb{C}^d\right)\right)$.

We will first show that the previous definition gives rise to a Borel measure with the properties of the Haar measure.

Theorem 1.9. *The function $\eta : \text{Borel}\left(\mathcal{U}\left(\mathbb{C}^d\right)\right) \to [0,\infty)$ from the previous definition is a normalized and unitarily invariant Borel measure.*

Proof. Note that the function F is continuous and hence a Borel function. Since Γ is a Borel measure, we conclude that $\eta = \Gamma \circ F^{-1}$ is a Borel measure as well. It is easy to see that η is normalized since $M\left(M^\dagger M\right)^{-1/2} \in \mathcal{U}\left(\mathbb{C}^d\right)$ if and only if M is invertible. Therefore, we have

$$\eta\left(\mathcal{U}\left(\mathbb{C}^d\right)\right) = \Gamma\left(\{M \in B(\mathbb{C}^d) \ : \ \det(M) \neq 0\}\right) = 1,$$

by Lemma 1.6. Finally, we need to show that η is unitarily invariant. To see this, note that

$$F(UM) = UM\left((UM)^\dagger UM\right)^{-1/2} = UM\left(M^\dagger M\right)^{-1/2} = UF(M),$$

for any unitary $U \in \mathcal{U}\left(\mathbb{C}^d\right)$ and any invertible M. Therefore, we have

$$\eta\left(U\mathcal{S}\right) = \Gamma\left(F^{-1}\left(U\mathcal{S}\right)\right) = \Gamma\left(UF^{-1}\left(\mathcal{S}\right)\right) = \Gamma\left(F^{-1}\left(\mathcal{S}\right)\right) = \eta\left(\mathcal{S}\right),$$

for any Borel set $\mathcal{S}$, where we used Lemma 1.6, and the proof is finished. □

It is not difficult to show that the Haar measure is unique if it exists:

Lemma 1.10 (Uniqueness of the Haar measure)**.** *For $d \in \mathbb{N}$ let $\nu : \text{Borel}\left(\mathcal{U}\left(\mathbb{C}^d\right)\right) \to [0,\infty)$ denote a Borel probability measure such that at least one of the following is true:*

- *For every $\mathcal{S} \subseteq \mathcal{U}\left(\mathbb{C}^d\right)$ and every $U \in \mathcal{U}\left(\mathbb{C}^d\right)$ we have $\nu\left(U\mathcal{S}\right) = \nu\left(\mathcal{S}\right)$.*
- *For every $\mathcal{S} \subseteq \mathcal{U}\left(\mathbb{C}^d\right)$ and every $U \in \mathcal{U}\left(\mathbb{C}^d\right)$ we have $\nu\left(\mathcal{S}U\right) = \nu\left(\mathcal{S}\right)$.*

Then, ν equals the Haar measure η on $\mathcal{U}\left(\mathbb{C}^d\right)$.

Proof. We will state the proof in the case where ν satisfies the first property. The case of the second property works in the same way. Consider a Borel set $\mathcal{S} \in \text{Borel}\left(\mathcal{U}\left(\mathbb{C}^d\right)\right)$ and denote by $\chi_{\mathcal{S}} : \mathcal{U}\left(\mathbb{C}^d\right) \to \mathbb{C}$ the characteristic function of $\mathcal{S}$. Note that

$$\nu\left(\mathcal{S}\right) = \int \chi_{\mathcal{S}}\left(U\right) d\nu(U) = \int \chi_{\mathcal{S}}\left(VU\right) d\nu(U),$$

for every unitary $V \in \mathcal{U}\left(\mathbb{C}^d\right)$. Since the Haar measure η is normalized and unitarily invariant, we have

$$\nu\left(\mathcal{S}\right) = \int\int \chi_{\mathcal{S}}\left(VU\right) d\nu(U) d\eta(V) = \int\int \chi_{\mathcal{S}}\left(VU\right) d\eta(V) d\nu(U) = \int \eta(\mathcal{S}U) d\nu(U) = \eta\left(\mathcal{S}\right),$$

where we used the Fubini-Tonelli theorem for the second equality. □

2 Random constructions and extremely entangled subspaces

We have seen in the previous lecture that some of the norms $\| \cdot \|_{1\to p}$ are not multiplicative. The proof of this result relied on choosing nice subspaces of $\mathbb{C}^{d_E} \otimes \mathbb{C}^{d_B}$. Unfortunately, nobody has so far come up with a explicit subspace showing multiplicativity of $\| \cdot \|_{1\to p}$ for values p close to 1, or to show additivity of the minimum output entropy. The constructions known to show these results are all probabilistic, i.e., the subspace is chosen at random. In the following, we will comment on some aspects of these constructions, but many details exceed the scope of this course.

2.1 How can we choose subspaces at random?

To define random subspaces we need the notion of a Haar-random unitary, which is simply a random unitary matrix distributed according to the Haar measure. The construction of the previous section suggests a way to generate such unitaries in practice:

Theorem 2.1 (Concrete Haar-random unitaries)**.** *Consider a random unitary $U \in \mathcal{U}\left(\mathbb{C}^d\right)$ constructed as follows:*

1. *Consider a random matrix $M \in B(\mathbb{C}^d)$ with entries $M_{kl} = X_{kl} + iY_{kl}$, where X_{kl} and Y_{kl} are chosen i.i.d. normally distributed.*
2. *Compute the singular value decomposition $M = WSV$ and define a unitary $U = WV \in \mathcal{U}\left(\mathbb{C}^d\right)$.*

Then, U is a Haar-random unitary.

We will use the following definition of a random subspace:

Definition 2.2 (Random subspaces)**.** *A random k-dimensional subspace of $\mathbb{C}^d$ is given by $U(\mathbb{C}^k \oplus 0_{d-k})$, where $U \in \mathcal{U}\left(\mathbb{C}^d\right)$ is a Haar-random unitary.*

2.2 Many subspaces are extremely entangled

We start by defining a measure of entanglement for pure states:

Definition 2.3. *For a bipartite pure state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ we define the* entropy of entanglement *as*

$$E\left(|\psi\rangle\right) = H\left(\mathrm{Tr}_B\left[|\psi\rangle\langle\psi|\right]\right).$$

The entropy of entanglement measures the entanglement in pure states, and it can be verified easily that

$$E\left(|\psi_A\rangle \otimes |\psi_B\rangle\right) = 0,$$

for pure states $|\psi_A\rangle \in \mathcal{H}_A$ and $|\psi_B\rangle \in \mathcal{H}_B$. We also have

$$E\left(|\Omega\rangle\right) = \log(\dim(\mathcal{H})),$$

if $|\Omega\rangle \in \mathcal{H} \otimes \mathcal{H}$ is the maximally entangled state. The following theorem is the key part of the counterexample to the additivity of the minimum output entropy. We will not prove it here. Its proof relies on the so-called Dvoretzky-Milman theorem in asymptotic geometric analysis, which goes beyond the scope of this course.

Theorem 2.4 (Extremely entangled subspaces)**.** *There are absolute constants $c, C > 0$ such that we have the following: With $d_B = d_E^2$ and $d_A = cd_E^2$ a random d_A-dimensional subspace $\mathcal{S} \subset \mathbb{C}^{d_E} \otimes \mathbb{C}^{d_B}$ satisfies*

$$\min_{|\psi\rangle\in\mathcal{S},\langle\psi|\psi\rangle=1} E\left(|\psi\rangle\right) \geq \log(d_E) - \frac{C}{d_E},$$

with high probability.

Let us comment briefly on the main idea behind the proof. The main tool used is the following theorem by Dvoretzky and Milman:

Theorem 2.5. *There are absolute constants $c, c' > 0$ such that we have the following: Fix an $\epsilon \in (0,1]$. For any circled*[1] *convex body $K \subset \mathbb{C}^n$, there exists an $M \in \mathbb{R}$ and a number $d(K) \in \mathbb{R}$ such that a random k-dimensional subspace E with $k = c\epsilon^2 d(K)$ satisfies*

$$(1-\epsilon)M\|x\|_2 \leq \|x\|_K \leq (1+\epsilon)M\|x\|_2,$$

for all $x \in E$, with probability larger than $1 - \exp(-c'\epsilon^2 d(K))$. Here, we used the norm

$$\|x\|_K = \inf\{t \geq 0 \;:\; x \in tK\}.$$

In the previous theorem, the number $d(K)$ is also called the Dvoretzky-dimension of K and it can be computed using geometric properties of K. Geometrically, the Dvoretzky-Milman theorem (which also has a version over the reals) states that taking intersections of a sufficiently low-dimensional hyperplane with a sufficiently high-dimensional circled (or symmetric in the real case) convex body results in almost Euclidean balls with high probability. Let us now bring the statement of Theorem 2.4 into a form where it becomes more clear that the Dvoretzky-Milman theorem plays a role. We need the following lemma:

Lemma 2.6. *For any quantum state $\rho \in D(\mathcal{H})$ with $d = \dim(\mathcal{H})$ we have*

$$H(\rho) \geq \log(d) - d\|\rho - \frac{\mathbb{1}_{\mathcal{H}}}{d}\|_2^2.$$

The following theorem implies Theorem 2.7 when combined with the previous lemma:

Theorem 2.7. *There are absolute constants $c, C > 0$ such that we have the following: Set $d_B = d_E^2$ and $d_A = cd_E^2$ and let $B_{\|\cdot\|_2} \subset B(\mathbb{C}^{d_E}, \mathbb{C}^{d_B})$ denote the Hilbert-Schmidt unit ball. We define a function $g : B_{\|\cdot\|_2} \to \mathbb{R}$ by*

$$g(X) = \|XX^\dagger - \frac{\mathbb{1}_{d_E}}{d_E}\|_2.$$

With large probability, a random d_A-dimensional subspace $E \subset B(\mathbb{C}^{d_E}, \mathbb{C}^{d_B})$ satifies

$$\sup_{X \in B_{\|\cdot\|_2} \cap E} g(M) \leq \frac{C}{d_E}.$$

How is this related to the Dvoretzky-Milman theorem? The statemement of the theorem can be reformulated as follows: For every $X \in B_{\|\cdot\|_2} \cap E$ we have

$$\frac{C^2}{d_E^2} \geq \|XX^\dagger - \frac{\mathbb{1}_{d_E}}{d_E}\|_2^2 = \mathrm{Tr}\left[|X|^4\right] - \frac{2\,\mathrm{Tr}\left[X^\dagger X\right]}{d_E} + \frac{\mathrm{Tr}\left[\mathbb{1}_{d_E}\right]}{d_E} \geq 0.$$

Rearranging these inequalities yields

$$d_E^{-1/4}\|X\|_2 \leq \|X\|_4 \leq d_E^{-1/4}\left(1 + \frac{C^2}{d_E}\right)^{1/4}\|X\|_2 \leq d_E^{-1/4}\left(1 + \frac{C^2}{4d_E}\right)^{1/4}\|X\|_2,$$

which is a statement as in Theorem 2.5. Unfortunately, going through the constants (and using that $k(B_{\|\cdot\|_4}) = d_E^{1/2} d_B$) only gives an $\epsilon \frac{1}{d_E^{1/4}}$, which is not good enough to prove the previous inequalities and Theorem 2.7. Getting the correct scaling is a bit more involved and the interested reader might want to check out the book "Alice and Bob meet Banach" by Guillaume Aubrun and Stanislaw Szarek (from which the material in this lecture was mostly adapted from).

[1] such that $e^{i\alpha}x \in K$ for every $\alpha \in \mathbb{R}$ whenever $x \in K$

2.3 Counterexample to the additivity of the Holevo quantity

Using this theorem, we can find a counterexample to the additivity problem:

Corollary 2.8. *There exist quantum channels $S, T : B(\mathbb{C}^{d_A}) \to B(\mathbb{C}^{d_B})$ such that*

$$H_{\min}(S \otimes T) < H_{\min}(S) + H_{\min}(T).$$

Proof. Choose a quantum channel $T : B(\mathbb{C}^{d_A}) \to B(\mathbb{C}^{d_B})$ with Stinespring dilation

$$T(X) = \mathrm{Tr}_E \left[V X V^\dagger \right],$$

with an isometry $V : \mathbb{C}^{d_A} \to \mathbb{C}^{d_E} \otimes \mathbb{C}^{d_B}$ such that $\mathrm{Im}(V) \subset \mathbb{C}^{d_E} \otimes \mathbb{C}^{d_B}$ is extremely entangled, as in Theorem 2.4. Moreover, we choose $S = \overline{T}$ as the conjugate channel. By the properties of the subspace $\mathrm{Im}(V)$, we have

$$H_{\min}(T) = H_{\min}(S) = \min_{|\psi\rangle \in \mathcal{S}, \langle\psi|\psi\rangle = 1} E(|\psi\rangle) \geq \log(d_E) - \frac{C}{d_E}.$$

Using Lemma **??**, we find that the quantum state $(S \otimes T)(\omega_{A'A})$ has an eigenvalue larger than

$$\frac{d_A}{d_B d_E} = \frac{c}{d_E}.$$

By the exercises, we find that

$$H_{\min}(S \otimes T) \leq H((S \otimes T)(\omega_{A'A})) \leq 2\log(d_E) - \frac{c\log(d_E)}{d_E} + \frac{1}{d_E}.$$

For large enough values of d_E we conclude that

$$H_{\min}(S \otimes T) < H_{\min}(S) + H_{\min}(T).$$

□

The previous theorem shows that the additivity is violated for two different channels. Using two corollaries from the previous lecture shows the following:

Theorem 2.9. *There exists a quantum channel $T : B(\mathcal{H}_A) \to B(\mathcal{H}_B)$ such that*

$$\chi(T \otimes T) > \chi(T) + \chi(T).$$

A consequence of this theorem is that the regularization in the HSW theorem is necessary, and it cannot be replaced by the single-letter formula $\chi(T)$.