Concentration estimates for random subspaces of a tensor product, and application to Quantum Information Theory

Benoît Collins, Félix Parraud

Concentration estimates for random subspaces of a tensor product, and application to Quantum Information Theory

Benoît Collins, Félix Parraud

Abstract

Given a random subspace H_n chosen uniformly in a tensor product of Hilbert spaces $V_n \otimes W$, we consider the collection K_n of all singular values of all norm one elements of H_n with respect to the tensor structure. A law of large numbers has been obtained for this random set in the context of W fixed and the dimension of H_n, V_n tending to infinity at the same speed in [3].

In this paper, we provide measure concentration estimates in this context. The probabilistic study of K_n was motivated by important questions in Quantum Information Theory, and allowed to provide the smallest known dimension (184) for the dimension an an ancilla space allowing Minimum Output Entropy (MOE) violation. With our estimates, we are able, as an application, to provide actual bounds for the dimension of spaces where violation of MOE occurs.

1 Introduction

One of the most important questions in Quantum Information Theory (QIT) was to figure out whether one can find two quantum channels Φ_1 and Φ_2 such that

$$H_{min}(\Phi_1 \otimes \Phi_2) < H_{min}(\Phi_1) + H_{min}(\Phi_2),$$

where H_{min} is the Minimum Output Entropy (MOE), defined in section 4. This problem was solved by [14], with important preliminary work by [13] (see also references therein). This was especially important in QIT, since there was hope it would be true and in this case, it would give a systematic way to compute the classical capacity of a quantum channel. For more explanation we refer to [7].

All proofs available so far are not constructive in the sense that constructions rely on the probabilistic method. After the initial construction of [14], the probabilistic tools involved in the proof have been found to have deep relation with random matrix theory in many respects, including large deviation principle [2], Free probability [4], convex geometry [1] and Operator Algebra [9]. The last two probably give the most conceptual proofs, and in particular convex geometry gives explicit numbers. Free probability gives the best numbers for the output dimension [4] but was unable to give estimates for the input dimension so far. More generally, the optimal violation obtained in [4] relates to a LLN obtained in [3] whose speed of convergence was not explicit, and in turn, did not give any estimate on the smallest dimension of the input space. In order to obtain explicit parameters, measure concentration estimates, ideally large deviation estimates, are required. And from a theoretical point of view, this is the goal of this paper. The main result is as follows:

Theorem 1.1 (For the precise statement, see Theorem 2.2)**.** *Given the output space $K_{n,k,t}$ of our quantum model and its limit $K_{k,t}$ (see Equation (2) for a definition), for $n \geq 3^4 \times 2^{29} \times \ln^2(kn) \times k^3\varepsilon^{-4}$,*

$$\mathbb{P}\left(K_{n,k,t} \not\subset K_{k,t} + \varepsilon\right) \leq e^{k^2(\ln(3k^2\varepsilon^{-1})) - \frac{n}{k} \times \frac{\varepsilon^2}{576}}.$$

They are based on the far reaching approach of [15] – see as well [6]. As a corollary, we obtain the following important application in Quantum Information Theory:

Theorem 1.2 (For the precise statement, see Theorem 4.3)**.** *There exists a quantum channel from $\mathbb{M}_{184\times 10^{51}}(\mathbb{C})$ to $\mathbb{M}_{184}(\mathbb{C})$ that yields violation of MOE.*

The paper is organized as follows. After this introduction, section 2 is devoted to introducing necessary notations and state the main theorem. Section 3 contains the proof of the main theorem, and section 4 contains application to Quantum Information Theory.

Acknowledgements: B.C. was supported by JSPS KAKENHI 17K18734, 17H04823 and 20K20882. F.P. was supported by a JASSO fellowship and Labex Milyon (ANR-10-LABX-0070) of Université de Lyon. This work was initiated while the second author was doing his MSc under the supervision of Alice Guionnet and he would like to thank her for insightful comments and suggestions on this work.

2 Notation and main theorem

We denote by H a Hilbert space, which we assume to be finite dimensional. $B(H)$ is the set of bounded linear operators on H, and $D(H) \subset B(H)$ is the collection of trace 1, positive operators – known as *density matrices*. In the case of matrices, we denote it by $\mathcal{D}_k \subset \mathbb{M}_k(\mathbb{C})$.

Let $d, k, n \in \mathbb{N}$, let U be distributed according to the Haar measure on the unitary group of $\mathbb{M}_{kn}(\mathbb{C})$, let P_n be the canonical projection from $\mathbb{C}^d$ to $\mathbb{C}^{kn}$, that is the matrix with kn lines and d columns with 1 on the diagonal and 0 elsewhere. With Tr_n the trace on $\mathbb{M}_n(\mathbb{C})$, we define the following random linear map,

$$\Phi_n : X \in \mathbb{M}_d(\mathbb{C}) \mapsto id_k \otimes \mathrm{Tr}_n(UP_nXP_n^*U^*) \in \mathbb{M}_k(\mathbb{C}). \tag{1}$$

This map is trace preserving, linear and completely positive and as such, it is known as a quantum channel. Let $t \in [0,1]$. We fix d be an integer sequence (depending on n) such that $d \sim tkn$, and define

$$K_{n,k,t} = \Phi_n(\mathcal{D}_d). \tag{2}$$

There is a much more geometric definition of $K_{n,k,t}$ thanks to the following proposition.

Proposition 2.1. *We have,*

$$K_{n,k,t} = \{X \in \mathcal{D}_k \mid \forall A \in \mathcal{D}_k, \mathrm{Tr}_k(XA) \leq \|P_n^*U^*A \otimes I_nUP_n\|\}. \tag{3}$$

Besides for any $A \in \mathcal{D}_k$, $\{X \in K_{n,k,t} \mid \mathrm{Tr}_k(XA) = \|P_n^*U^*A \otimes I_nUP_n\|\}$ is non-empty.

Proof. Let $Y \in \mathcal{D}_d$, $A \in \mathcal{D}_k$, then

$$\begin{aligned}
\mathrm{Tr}_k(\Phi_n(Y)A) &= \mathrm{Tr}_{kn}(UP_nYP_n^*U^* \cdot A \otimes I_n) \\
&= \mathrm{Tr}_d(\sqrt{Y}P_n^*U^* \cdot A \otimes I_n \cdot UP_n\sqrt{Y}) \\
&\leq \mathrm{Tr}_d(Y)\,\|P_n^*U^* \cdot A \otimes I_n \cdot UP_n\| \\
&= \|P_n^*U^* \cdot A \otimes I_n \cdot UP_n\|\,.
\end{aligned}$$

Let us write E for the right member of the equation (3), we just showed that $K_{n,k,t} = \Phi_n(\mathcal{D}_d) \subset E$. Besides if P_x is the orthogonal projection on the vector x, we have that

$$\|P_n^*U^* \cdot A \otimes I_n \cdot UP_n\| = \max_{x\in\mathbb{C}^d} \mathrm{Tr}_d(P_n^*U^* \cdot A \otimes I_n \cdot UP_nP_x) = \max_{x\in\mathbb{C}^d} \mathrm{Tr}_k(A\ \Phi_n(P_x)).$$

Thus, for every $\varepsilon > 0$ and $A \in \mathcal{D}_k$, we can find an element of $K_{n,k,t}$ in $\{X \in \mathcal{D}_k \mid \mathrm{Tr}_k(XA) \geq \|P_n^*U^*A \otimes I_nUP_n\| - \varepsilon\}$. By compacity of $K_{k,n,t}$, we can even find an element of $K_{n,k,t}$ in $\{X \in \mathcal{D}_k \mid \mathrm{Tr}_k(XA) = \|P_n^*U^*A \otimes I_nUP_n\|\}$.

If we see E as a convex set of $\mathbb{M}_k(\mathbb{C})_{sa}$, let $X \in E$ be an exposed point of E, that is there exists $A \in \mathbb{M}_k(\mathbb{C})_{sa}$ and C such that the intersection of E and $\{Y \in \mathbb{M}_k(\mathbb{C})_{sa} \mid \mathrm{Tr}_k(AY) = C\}$ is reduced to $\{X\}$ and that E is included in $\{Y \in \mathbb{M}_k(\mathbb{C})_{sa} \mid \mathrm{Tr}_k(AY) \leq C\}$. We have the following equality for λ large enough since if $Y \in \mathcal{D}_k$, $\mathrm{Tr}_k(Y) = 1$,

$$\{Y \in \mathcal{D}_k \mid \mathrm{Tr}_k(AY) = C\} = \left\{Y \in \mathcal{D}_k \mid \mathrm{Tr}_k\left(\frac{A + \lambda I_k}{\mathrm{Tr}_k(A + \lambda I_k)}Y\right) = \frac{C + \lambda}{\mathrm{Tr}_k(A + \lambda I_k)}\right\}.$$

Thus, we can find $B \in \mathcal{D}_k$ and c such that such that the intersection of E and $\{Y \in \mathcal{D}_k \mid \mathrm{Tr}_k(BY) = c\}$ is reduced to $\{X\}$ and that E is included in $\{Y \in \mathcal{D}_k \mid \mathrm{Tr}_k(BY) \leq c\}$. To summarize:

- The intersection of $K_{n,k,t}$ and $\{Y \in \mathcal{D}_k \mid \operatorname{Tr}_k(BY) = \|P_n^* U^* \cdot B \otimes I_n \cdot U P_n\|\}$ is non-empty.
- $K_{n,k,t} \subset E$, so the intersection of E and $\{Y \in \mathcal{D}_k \mid \operatorname{Tr}_k(BY) = \|P_n^* U^* \cdot B \otimes I_n \cdot U P_n\|\}$ is non-empty.
- The intersection of E and $\{Y \in \mathcal{D}_k \mid \operatorname{Tr}_k(BY) = c\}$ is a exactly $\{X\}$.
- E is included in both $\{Y \in \mathcal{D}_k \mid \operatorname{Tr}_k(BY) \leq c\}$ and $\{Y \in \mathcal{D}_k \mid \operatorname{Tr}_k(BY) \leq \|P_n^* U^* \cdot B \otimes I_n \cdot U P_n\|\}$.

Hence it implies that $c = \|P_n^* U^* B \otimes I_n U P_n\|$ and that $X \in K_{n,k,t}$. Thus we showed that the exposed point of E belongs to $K_{n,k,t}$. By a result of Straszewicz ([8],theorem 18.6) the set of exposed points is dense in the set of extremal points, so the set of extremal points of E is included in $K_{n,k,t}$. Since $K_{n,k,t}$ is convex, E is included in $K_{n,k,t}$. $\square$

Thanks to Theorem 1.4 of [10], we know that $\|P_n^* U^* \cdot A \otimes I_n \cdot U P_n\|$ converges almost surely towards a limit $\|A\|_{(t)}$, which we now describe in terms of free probability (for the interested reader we refer to [17], but a non expert reader can take $\lim_n \|P_n^* U^* \cdot A \otimes I_n \cdot U P_n\|$ as the definition of $\|A\|_{(t)}$ without loss of generality). We view $B(L^2([0,1]))$ as the $\mathcal{C}^*$-algebra endowed with the state $\tau(u) = \langle u(\mathbf{1}_{[0,1]}), \mathbf{1}_{[0,1]} \rangle_{L^2([0,1]}$. The endomorphism $p_t : f \mapsto \mathbf{1}_{[0,t]} f$ is a self-adjoint projection of rank t, that is $\tau(p_t) = t$. We consider $\mathcal{A}$ the $\mathcal{C}^*$-algebra generated by p_t, restricted to $\mathcal{A}$, τ is a faithful tracial state. Hence thanks to Teorem 7.9 from [18], we can consider the free product of $\mathcal{A}$ and $\mathbb{M}_k(\mathbb{C})$. For $A \in \mathbb{M}_k(\mathbb{C})$, we introduce the following quantity, called the *(t)-norm*:

$$\|A\|_{(t)} := \|p_t A p_t\|\,, \tag{4}$$

where on the right side we took the operator norm in the free product of $\mathcal{A}$ and $\mathbb{M}_k(\mathbb{C})$. We naturally define the asymptotic limit of $K_{n,k,t}$ by

$$K_{k,t} = \{X \in \mathcal{D}_k \mid \forall A \in \mathcal{D}_k, \operatorname{Tr}_k(XA) \leq \|A\|_{(t)}\}. \tag{5}$$

The main result of this paper is a measure concentration estimate and can be stated as follows:

Theorem 2.2. *If we endow $\mathbb{M}_k(\mathbb{C})$ with the norm $M \mapsto \sqrt{\operatorname{Tr}_k(M^*M)}$, and that we assume $d \leq tkn$, then for $n \geq 3^4 \times 2^{29} \times \ln^2(kn) \times k^3 \varepsilon^{-4}$,*

$$\mathbb{P}\left(K_{n,k,t} \not\subset K_{k,t} + \varepsilon\right) \leq e^{k^2(\ln(3k^2\varepsilon^{-1})) - \frac{n}{k} \times \frac{\varepsilon^2}{576}}.$$

The convergence of $K_{n,k,t}$ towards $K_{k,t}$ for the Hausdorff distance has already been proved in [3], Theorem 5.2. More precisely the authors proved that given a random subspace of size d, $F_{n,k,t}$ the collection of singular values of unit vectors in this subspace converges for the Hausdorff distance towards a deterministic set $F_{k,t}$. It turns out that $K_{n,k,t}$ (respectively $K_{k,t}$) is the convex hull of the self-adjoint matrices whose eigenvalues are in $F_{n,k,t}$ (respectively $F_{k,t}$). However our paper is self-contained and we do not use this theorem.

3 Proof of main theorem

We will combine this geometrical description with the following lemma to get an estimate.

Proposition 3.1. *If we endow $\mathbb{M}_k(\mathbb{C})_{sa}$ with the norm $\|.\|_2$ which comes from the scalar product $(U,V) \mapsto \operatorname{Tr}_k(UV)$, then the following implication is true,*

$$\forall A \in \mathcal{D}_k,\ \|P_n^* U^* A \otimes I_n U P_n\| \leq \|A\|_{(t)} + \frac{\varepsilon}{k} \quad \Longrightarrow \quad K_{n,k,t} \subset K_{k,t} + \varepsilon.$$

Before proving it, we need a small lemma on the structure of $K_{k,t}$.

Lemma 3.2. *Let $A \in \mathcal{D}_k$, then $\{X \in K_{k,t} \mid \operatorname{Tr}_k(XA) = \|A\|_{(t)}\}$ is non-empty.*

Proof. Thanks to Proposition 2.1 we know that for any n, $\{X \in K_{n,k,t} \mid \operatorname{Tr}_k(XA) = \|P_n^* U^* A \otimes I_n U P_n\|\}$ is non-empty. Hence there exists X_n such that:

- $\mathrm{Tr}_k(X_n A) = \|P_n^* U^* A \otimes I_n U P_n\|\}$,
- $\forall B \in \mathcal{D}_k$, $\mathrm{Tr}_k(X_n B) \leq \|P_n^* U^* B \otimes I_n U P_n\|\}$.

By compacity of $\mathcal{D}_k$, we can assume that X_n converges towards a limit X. But then as we said in the previous section, thanks to Theorem 1.4 from [10], $\|P_n^* U^* B \otimes I_n U P_n\|$ converges towards $\|B\|_{(t)}$. Thus X is such that:

- $\mathrm{Tr}_k(XA) = \|A\|_{(t)}\}$,
- $\forall B \in \mathcal{D}_k$, $\mathrm{Tr}_k(XB) \leq \|B\|_{(t)}\}$.

That is, X belongs to $\{X \in K_{k,t} \mid \mathrm{Tr}_k(XA) = \|A\|_{(t)}\}$. □

We can now prove Proposition 3.1.

Proof of Proposition 3.1. We assume that $K_{n,k,t} \not\subset K_{k,t} + \varepsilon$, then thanks to the compacity of $K_{n,k,t}$ and $K_{k,t}$, we can find $X \in K_{k,t}$ and $Y \in K_{n,k,t}$ such that $\|X - Y\|_2 \geq \varepsilon$, and $K_{k,t} \cap B(Y, \|X - Y\|_2)$ is empty. We set $U = \frac{Y-X}{\|Y-X\|_2}$, $A = \frac{1}{k}(U + I_k)$, then $A \in \mathcal{D}_k$. We are going to show that $\|P_n^* U^* A \otimes I_n U P_n\| > \|A\|_{(t)} + \frac{\varepsilon}{k}$. To do so we define

$$P_C = \left\{ B \in K_{k,t} \;\middle|\; \mathrm{Tr}_k(AB) = \frac{C+1}{k} \right\} = \left\{ B \in K_{k,t} \mid \mathrm{Tr}_k(UB) = C \right\}.$$

Let us assume that for $C > \mathrm{Tr}_k(UX)$, P_C is not empty, then let $S \in P_C$. We can write $C = \mathrm{Tr}_k(U(X + tU))$ for some $t > 0$, thus $\mathrm{Tr}_k(US) = \mathrm{Tr}_k(U(X + tU))$, that is $\mathrm{Tr}_k((Y - X)(X - S)) = -t\|Y - X\|_2$. Hence the following estimate:

$$\begin{aligned} \|Y - (\alpha X + (1-\alpha)S)\|_2^2 &= \mathrm{Tr}_k\left(\left(Y - X + (1-\alpha)(X - S)\right)^2 \right) \\ &= \|Y - X\|_2^2 - 2t(1-\alpha)\|Y - X\|_2 + \bigcirc((1-\alpha)^2). \end{aligned}$$

Consequently since $K_{k,t}$ is convex, for any α, $\alpha X + (1-\alpha)S \in K_{k,t}$, thus for $1-\alpha$ small enough we could find an element of $K_{k,t}$ in $B(Y, \|X - Y\|_2)$. Hence the contradiction. Thus for $C > \mathrm{Tr}_k(UX)$, P_C is empty. By Lemma 3.2, we get that $\frac{\mathrm{Tr}_k(UX)+1}{k} \geq \|A\|_{(t)}$. Next we define

$$Q_C = \left\{ B \in K_{n,k,t} \;\middle|\; \mathrm{Tr}_k(AB) = \frac{C+1}{k} \right\} = \left\{ B \in K_{n,k,t} \mid \mathrm{Tr}_k(UB) = C \right\}.$$

Then clearly for $C = \mathrm{Tr}_k(UY)$, Q_C is non-empty since $Y \in Q_{\mathrm{Tr}_k(UY)}$. Hence thanks to the geometric definition (3) of $K_{n,k,t}$, we have that $\frac{\mathrm{Tr}_k(UY)+1}{k} \leq \|P_n^* U^* \cdot A \otimes I_n \cdot U P_n\|$. Thus we have,

$$\|P_n^* U^* A \otimes I_n U P_n\| \geq \frac{\mathrm{Tr}_k(U(Y - X))}{k} + \|A\|_{(t)} = \frac{\|Y - X\|_2}{k} + \|a\|_{(t)} \geq \frac{\varepsilon}{k} + \|A\|_{(t)}.$$

□

Actually with a very similar proof, we could even show that almost surely there exist $A \in \mathcal{D}_k$ such that

$$d_H(K_{n,k,t}, K_{k,t}) = k \times \left| \|P_n^* U^* \cdot A \otimes I_n \cdot U P_n\| - \|A\|_{(t)} \right|,$$

where d_H is the Hausdorff distance associated to the norm $\|.\|_2$ which comes from the scalar product $(U, V) \mapsto \mathrm{Tr}_k(UV)$. However this result will not be useful in this paper since the absolute value would be detrimental for the computation of our estimate. The following lemma is a rather direct consequence of the previous proposition.

Lemma 3.3. *Let $u > 0$, let $\mathcal{S}_u = \{uM \mid M \in \mathbb{M}_k(\mathbb{C})_{sa},\ \forall i \geq j,\ \Re(m_{i,j}) \in \left\{\mathbb{N} + \frac{1}{2}\right\} \cap [0, \lceil u^{-1} \rceil],\ \forall i > j,\ \Im(m_{i,j}) \in \left\{\mathbb{N} + \frac{1}{2}\right\} \cap [0, \lceil u^{-1} \rceil]\ \}$, let $P_{\mathcal{D}_k}$ be the convex projection on $\mathcal{D}_k$. Then with $u = \frac{\sqrt{2}\varepsilon}{3k^2}$,*

$$\mathbb{P}\left(K_{n,k,t} \not\subset K_{k,t} + \varepsilon\right) \leq \sum_{M \in \mathcal{S}_u} \mathbb{P}\left(\|P_n^* U^* (P_{\mathcal{D}_k} M \otimes I_n) U P_n\| > \|P_{\mathcal{D}_k} M\|_{(t)} + \frac{\varepsilon}{3k} \right).$$

Proof. We immediately get from proposition 3.1 that

$$\mathbb{P}\left(K_{n,k,t} \not\subset K_{k,t}+\varepsilon\right) \leq \mathbb{P}\left(\exists A \in \mathcal{D}_k,\ \|P_n^* U^* \cdot A \otimes I_n \cdot U P_n\| > \|A\|_{(t)} + \frac{\varepsilon}{k}\right).$$

Now, let $A \in \mathcal{D}_k$, by construction of $\mathcal{S}_u$, there exists $M \in \mathcal{S}_u$ such that the real part and the imaginary part of the coefficients of M are $u/2$-close from those of A. Thus we have $\|A-M\|_2 \leq \frac{ku}{\sqrt{2}}$. Hence if we fix $u = \frac{\sqrt{2}\varepsilon}{3k^2}$, then we can always find $M \in \mathcal{S}_u$ such that $\|A-M\|_2 \leq \frac{\varepsilon}{3k}$. Besides we have,

$$\left| \|A\|_{(t)} - \|P_{\mathcal{D}_k} M\|_{(t)} \right| \leq \|A - P_{\mathcal{D}_k} M\| \leq \|A - P_{\mathcal{D}_k} M\|_2\,,$$

$$\Big| \|P_n^* U^* A \otimes I_n U P_n\| - \|P_n^* U^* (P_{\mathcal{D}_k} M \otimes I_n) U P_n\| \Big| \leq \|A - P_{\mathcal{D}_k} M\| \leq \|A - P_{\mathcal{D}_k} M\|_2\,.$$

Hence since $P_{\mathcal{D}_k} A = A$ and that $P_{\mathcal{D}_k}$ is 1-lipschitz, we have $\|A - P_{\mathcal{D}_k} M\|_2 \leq \|A-M\|_2 \leq \frac{\varepsilon}{3k}$. Consequently,

$$\left\{ \|P_n^* U^* \cdot A \otimes I_n \cdot U P_n\| > \|A\|_{(t)} + \frac{\varepsilon}{k} \right\} \subset \left\{ \|P_n^* U^* (P_{\mathcal{D}_k} M \otimes I_n) U P_n\| > \|P_{\mathcal{D}_k} M\|_{(t)} + \frac{\varepsilon}{3k} \right\}.$$

Hence,

$$\left\{ \exists A \in \mathcal{D}_k,\ \|P_n^* U^* \cdot A \otimes I_n \cdot U P_n\| > \|A\|_{(t)} + \frac{\varepsilon}{k} \right\} \subset \bigcup_{M \in \mathcal{S}_u} \left\{ \|P_n^* U^* (P_{\mathcal{D}_k} M \otimes I_n) U P_n\| > \|P_{\mathcal{D}_k} M\|_{(t)} + \frac{\varepsilon}{3k} \right\}.$$

The conclusion follows.

□

The next lemma shows that there exist a smooth function which verifies some assumptions on the infinite norm of its derivatives.

Lemma 3.4. *There exists g a $\mathcal{C}^6$ function which takes value 0 on $(-\infty,0]$ and value 1 on $[1,\infty)$, and in $[0,1]$ otherwise. Besides for any $j \leq 6$, $\left\|g^{(j)}\right\|_\infty = 2^{\frac{j(j+1)}{2}}$.*

Proof. Firstly we define,

$$f : t \in [0,1] \mapsto \begin{cases} 2t & \text{if } t \leq 1/2 \\ 2(1-t) & \text{if } t \geq 1/2 \end{cases},$$

$$\begin{array}{rccc} H & : & \mathbb{C}^0([0,1]) & \to & \mathbb{C}^0([0,1]) \\ & & f & \mapsto & t \mapsto \begin{cases} f(2t) & \text{if } t \leq 1/2 \\ -f(2t-1) & \text{if } t \geq 1/2 \end{cases} \end{array}.$$

Inspired by Taylor's Theorem, we define

$$h : x \in [0,1] \mapsto \int_0^x \frac{(x-t)^5}{5!}\ H^5 f(t)\ dt.$$

It is easy to see that $h \in \mathcal{C}^6([0,1])$ with

$$\forall j \leq 5, \quad h^{(j)} : x \in [0,1] \mapsto \int_0^x \frac{(x-t)^{5-j}}{(5-j)!} H^5 f(t)\ dt, \qquad h^{(6)} = H^5 f.$$

Thus one can easily extend h by 0 on $\mathbb{R}^-$ and h remains $\mathcal{C}^6$ in 0, as for what happens in 1 it is way less obvious. In order to build g we want to show that

$$\forall 1 \leq j \leq 6, \quad h^{(j)}(1) = 0, \qquad h(1) > 0.$$

To do so let $w \in \mathcal{C}^0([0,1])$, then for any $k \geq 0$,

$$\begin{aligned} \int_0^1 (1-t)^k H w(t) dt &= \int_0^{1/2} (1-t)^k w(2t) dt - \int_{1/2}^1 (1-t)^k w(2t-1) dt \\ &= \frac{1}{2^{k+1}} \int_0^1 \left((2-t)^k - (1-t)^k \right) w(t) dt \\ &= \frac{1}{2^{k+1}} \int_0^1 \sum_{0 \leq i < k} \binom{k}{i} (1-t)^i w(t) dt. \end{aligned}$$

Thus recursively one can show that $\forall 1 \leq j \leq 6$, $h^{(j)}(1) = 0$. We also get that

$$h(1) = \int_0^x \frac{(1-t)^5}{5!}\ H^5 f(t)\ dt = 2^{-\sum_{2\leq i\leq 6} i} \int_0^x f(t)\ dt = 2^{-21}.$$

Hence we fix $g = 2^{21}h$, further studies show that $\left\|g^{(j)}\right\|_\infty = 2^{\frac{j(j+1)}{2}}$. ◻

In the next lemma, we prove a first rough estimate on the deviation of the norm with respect to its limit. It is the only one where we use that $d \leq tkn$.

Lemma 3.5. *For any $A \in \mathcal{D}_k$, $\varepsilon > 0$,*

$$\mathbb{P}\left(\|P_n^* U^* A \otimes I_n U P_n\| \geq \|A\|_{(t)} + \varepsilon\right) \leq 3 \times 2^{21} \times \frac{\ln^2(kn)}{kn}\varepsilon^{-4}. \tag{6}$$

Proof. For a better understanding of the notations and tools used in this proof, such as free stochastic calculus, we refer to [15]. In particular τ_{kn} is the trace on the free product of $\mathbb{M}_{kn}(\mathbb{C})$ with a $\mathcal{C}^*$-algebra which contains two free unitary Brownian motions, see Definition 2.8 of [15]. As for $\delta, \mathcal{D}$ and $\boxtimes$, see Definition 2.8 and 2.10, as well as Proposition 2.2 of the same paper. If you are not familiar with free probability, it is possible to simply admit equation (8) to avoid having to understand the previous notations.

Since $\|P_n^* U^* \cdot A \otimes I_n \cdot U P_n\| = \|P_n P_n^* U^* \cdot A \otimes I_n \cdot U P_n P_n^*\|$, we will rather work with $P = P_n P_n^*$ since it is a square matrix. To simplify notation, instead of $A \otimes I_n$ we simply write A. Thus thanks to Lemma 4.1 and 4.4 from [15], given a function f such that

$$\forall x \in \mathbb{R}, \quad f(x) = \int_{\mathbb{R}} e^{\mathbf{i}xy} d\mu(y), \tag{7}$$

for some measure μ. Then with $u_t = U a_t$ and $v_t = U b_t$ where a_t and b_t are free unitary Brownian motions started in 1 and free between each other, we have the following expression,

$$\begin{aligned}
&\mathbb{E}\left[\frac{1}{kn}\mathrm{Tr}_{kn}\Big(f(P\ U^* A\ U\ P)\Big)\right] - \tau\Big(f(P\ u_T^*\ A\ u_T\ P)\Big) \\
&= \frac{1}{2(kn)^2}\int\int_0^T t\ \tau_{kn}\left(\mathcal{D}\circ\delta^1\circ\mathcal{D}\Big(e^{\mathbf{i}yPU^*AUP}\Big)(u_t,u_t^*,A,P)\ \boxtimes\ \mathcal{D}\circ\delta^2\circ\mathcal{D}\Big(e^{\mathbf{i}yPU^*AUP}\Big)(v_t,v_t^*,A,P)\right) \\
&\hspace{30em} dt\ d\mu(y).
\end{aligned}$$

Then if we set $R_1^t = Pu_t^* A u_t P$ and $R_2^t = Pv_t^* A v_t P$, after a tedious computation,

$$\begin{aligned}
&\tau_{kn}\left(\mathcal{D}\circ\delta^1\circ\mathcal{D}\Big(e^{\mathbf{i}yPU^*AUP}\Big)(u_t,u_t^*,A,P)\ \boxtimes\ \mathcal{D}\circ\delta^2\circ\mathcal{D}\Big(e^{\mathbf{i}yPU^*AUP}\Big)(v_t,v_t^*,A,P)\right) \qquad (8)\\
&= 2\mathbf{i}y^3\int_0^1 s\ \tau_{kn}\Bigg(\left(e^{\mathbf{i}yR_1^ts}Pu_t^*Au_t - u_t^*Au_tPe^{\mathbf{i}yR_1^ts}\right)\times\Bigg(e^{\mathbf{i}yR_2^t(1-s)}Pv_t^*A^2v_t - v_t^*A^2v_tPe^{\mathbf{i}yR_2^t(1-s)} \\
&\quad + \mathbf{i}y(1-s)\int_0^1 e^{\mathbf{i}yR_2^t(1-s)(1-r)}Pv_t^*A^2v_tPe^{\mathbf{i}yR_2^t(1-s)r}v_t^*Av_t \\
&\qquad\qquad - v_t^*Av_te^{\mathbf{i}yR_2^t(1-s)(1-r)}Pv_t^*A^2v_tPe^{\mathbf{i}yR_2^t(1-s)r}\ dr\Bigg)\Bigg)ds \\
&\quad - 2y^2\int_0^1 \tau_{kn}\Bigg(\left(\mathbf{i}ys(R_1^te^{\mathbf{i}yR_1^ts}u_t^*Au_t - u_t^*Au_te^{\mathbf{i}yR_1^ts}R_1^t) - u_t^*Au_tPe^{\mathbf{i}yR_1^ts} + e^{\mathbf{i}yR_1^ts}Pu_t^*Au_t\right) \\
&\qquad \times\Big(\mathbf{i}y(1-s)(R_2^te^{\mathbf{i}yR_2^t(1-s)}v_t^*Av_t - v_t^*Av_te^{\mathbf{i}yR_2^t(1-s)}R_2^t) \\
&\qquad\qquad - v_t^*Av_tPe^{\mathbf{i}yR_2^t(1-s)} + e^{\mathbf{i}yR_2^t(1-s)}Pv_t^*Av_t\Big)\Bigg)ds
\end{aligned}$$

One have that given f a continuous function, X and Y such that $XP = X$ and $PY = Y$, then

$$\tau_{kn}\left((f(R_1^t)Pu_t^*Au_t - u_t^*Au_tPf(R_1^t))X\right) = \tau_{kn}\left(Xf(R_1^t)Pu_t^*Au_t(1-P)\right),$$

$$\tau_{kn}\left((f(R_1^t)Pu_t^*Au_t - u_t^*Au_tPf(R_1^t))Y\right) = \tau_{kn}\left((P-1)u_t^*Au_tPf(R_1^t)Y\right).$$

Since the norm of $A, P, 1-P$ and u_t are smaller than 1, we finally get that for any t,

$$\begin{aligned}
&\left|\tau_{kn}\left(\mathcal{D}\circ\delta^1\circ\mathcal{D}\Big(e^{\mathbf{i}yPU^*AUP}\Big)(u_t,u_t^*,A,P)\ \boxtimes\ \mathcal{D}\circ\delta^2\circ\mathcal{D}\Big(e^{\mathbf{i}yPU^*AUP}\Big)(v_t,v_t^*,A,P)\right)\right| \\
&\leq 4y^2\int_0^1 s|y|(1+|y|(1-s)) + (1+|y|s)(|y|(1-s)+1)ds \\
&= 4y^2 + 6|y|^3 + \frac{4}{3}y^4.
\end{aligned}$$

Thanks to Proposition 3.3 from [15], we get that

$$\begin{aligned}
&\left|\tau\Big(e^{\mathbf{i}yPu_T^*Au_TP}\Big) - \tau\Big(e^{\mathbf{i}yPu^*AuP}\Big)\right| \\
&= \left|y\int_0^1 \tau\left(e^{\mathbf{i}syPu_T^*Au_TP}P\ (u_T^*Au_T - \widetilde{u}_T^*A\widetilde{u}_T)P\ e^{\mathbf{i}(1-s)yP\ \widetilde{u}_T^*A\widetilde{u}_TP}\right)ds\right| \\
&\leq 8e^2\pi e^{-T/2}|y|.
\end{aligned}$$

Consequently, if the support of f and the spectrum of Pu^*AuP are disjoint, then $\tau\Big(f(P\ u^*\ A\otimes I_n\ u\ P)\Big) = 0$, and

$$\begin{aligned}
&\left|\mathbb{E}\left[\frac{1}{kn}\operatorname{Tr}_{kn}\Big(f(P\ U^*\ A\otimes I_n\ U\ P)\Big)\right]\right| \\
&\leq 8e^2\pi e^{-T/2}\int |y|\ d|\mu|(y) + \left(\frac{T}{kn}\right)^2\int y^2 + \frac{3}{2}|y|^3 + \frac{1}{3}y^4\ d|\mu|(y).
\end{aligned}\tag{9}$$

Let g be a $\mathcal{C}^6$ function which takes value 0 on $(-\infty,0]$ and value 1 on $[1,\infty)$, and in $[0,1]$ otherwise. We set $f_\varepsilon : t\mapsto g(2\varepsilon^{-1}(t-\alpha)-1)g(2\varepsilon^{-1}(1-t)+1)$ with $\alpha = \|A\|_{(t)}$. Then the support of f is included in $[\|A\|_{(t)},\infty)$, whereas the spectrum of Pu^*AuP is bounded by $\|Pu^*AuP\| = \|A\|_{d(kn)^{-1}} \leq \|A\|_{(t)}$ since $d\leq tkn$. Hence f_ε satisfies (9). Setting $h: t\mapsto g(t-2\varepsilon^{-1}\alpha-1)g(2\varepsilon^{-1}+1-t)$, we have with convention $\widehat{f}(x) = (2\pi)^{-1}\int_{\mathbb{R}} f(y)e^{-\mathbf{i}xy}dy$, for $0\leq k\leq 4$ and any $\beta>0$,

$$\begin{aligned}
\int |y|^k|\hat{f}_\varepsilon(y)|\ dy &= \frac{1}{2\pi}\int |y|^k\left|\int g(2\varepsilon^{-1}(t-\alpha)-1)g(2\varepsilon^{-1}(1-t)+1)e^{-\mathbf{i}yt}\ dt\right|\ dy \\
&= \frac{1}{2\pi}\int |y|^k\left|\int h(\beta t)e^{-\mathbf{i}y\varepsilon\beta t/2}\ \frac{\varepsilon\beta}{2}dt\right|\ dy \\
&= \frac{1}{2\pi}2^k\varepsilon^{-k}\beta^{-k}\int |y|^k\left|\int h(\beta t)e^{-\mathbf{i}yt}\ dt\right|\ dy \\
&\leq \frac{1}{2\pi}2^k\varepsilon^{-k}\int\frac{1}{1+y^2}\ dy\int(|h^{(k)}(\beta t)| + \beta^2|h^{(k+2)}(\beta t)|)\ dt \\
&\leq 2^{k-1}\varepsilon^{-k}\left(\beta^{-1}\left\|g^{(k)}\right\|_\infty + \beta\left\|g^{(k+2)}\right\|_\infty\right).
\end{aligned}$$

In the last line we used the fact that we can always assume that $\alpha+\varepsilon\leq 1$ (otherwise $\mathbb{P}(\|P_n^*U^*\cdot A\otimes I_n\cdot UP_n\| \geq \|A\|_{(t)}+\varepsilon) = 0$ and there is no need to do any computation) and thus that the support of $t\mapsto g(t-2\varepsilon^{-1}\alpha-1)$ and the derivative of $t\mapsto g(2\varepsilon^{-1}+1-t)$ are disjoint. Thus by fixing $\beta = \sqrt{\left\|g^{(k)}\right\|_\infty\left\|g^{(k+2)}\right\|_\infty^{-1}}$ we get

$$\int |y|^k|\hat{f}_\varepsilon(y)|\ dy \leq 2^k\varepsilon^{-k}\sqrt{\left\|g^{(k)}\right\|_\infty\left\|g^{(k+2)}\right\|_\infty}.$$

Consequently, since f_ε satisfies (7) with $d\mu(y) = \widehat{f_\varepsilon}(y)dy$, by using (9) we get

$$\begin{aligned}
&\left|\mathbb{E}\left[\frac{1}{kn}\operatorname{Tr}_{kn}\left(f_\varepsilon(P\ U^*\ A\otimes I_n\ U\ P)\right)\right]\right| \\
&\leq 16e^2\pi e^{-T/2}\sqrt{\left\|g^{(1)}\right\|_\infty\left\|g^{(3)}\right\|_\infty}\varepsilon^{-1} + \left(\frac{T}{kn}\right)^2 4\varepsilon^{-2}\sqrt{\left\|g^{(2)}\right\|_\infty\left\|g^{(4)}\right\|_\infty} \\
&\quad + \left(\frac{T}{kn}\right)^2 12\varepsilon^{-3}\sqrt{\left\|g^{(3)}\right\|_\infty\left\|g^{(5)}\right\|_\infty} + \left(\frac{T}{kn}\right)^2 \frac{16}{3}\varepsilon^{-4}\sqrt{\left\|g^{(4)}\right\|_\infty\left\|g^{(6)}\right\|_\infty}.
\end{aligned}$$

Combined with Lemma 3.4 and fixing $T = 4\ln(kn)$, we get

$$\begin{aligned}
&\left|\mathbb{E}\left[\frac{1}{kn}\operatorname{Tr}_{kn}\left(f_\varepsilon(P\ U^*\ A\otimes I_n\ U\ P)\right)\right]\right| \\
&\leq 2^{\frac{15}{2}}e^2\pi\ \frac{\varepsilon^{-1}}{(kn)^2} + 2^{25/2}\left(\frac{\ln(kn)}{kn}\right)^2\varepsilon^{-2} + 3\times 2^{33/2}\left(\frac{\ln(kn)}{kn}\right)^2\varepsilon^{-3} + \frac{2^{47/2}}{3}\left(\frac{\ln(kn)}{kn}\right)^2\varepsilon^{-4}.
\end{aligned}$$

Since for any n, almost surely $\|P_n^* U^*\cdot A\otimes I_n\cdot UP_n\| \leq 1$, we have

$$\begin{aligned}
\mathbb{P}\left(\|P_n^*U^*A\otimes I_nUP_n\| \geq \|A\|_{(t)} + \varepsilon\right) &= \mathbb{P}\Big(\exists\lambda\in\sigma(PU^*A\otimes I_nUP),\ f_\varepsilon(\lambda) = 1\Big) \\
&\leq \mathbb{P}\Big(\operatorname{Tr}_{kn}\left(f_\varepsilon(P^*U^*A\otimes I_nUP)\right) \geq 1\Big) \\
&\leq \mathbb{E}\Big[\operatorname{Tr}_{kn}\left(f_\varepsilon(P\ U^*\ A\otimes I_n\ U\ P)\right)\Big] \\
&\leq 2^{\frac{15}{2}}e^2\pi\ \frac{\varepsilon^{-1}}{kn} + \frac{\ln^2(kn)}{kn}\left(2^{25/2}\varepsilon^{-2} + 3\times 2^{33/2}\varepsilon^{-3} + \frac{2^{47/2}}{3}\varepsilon^{-4}\right).
\end{aligned}$$

One can always assume that $\ln^2(kn) \geq 1$ since for small value of k and n, (6) is easily verified since the right member of the inequality is larger than 1. One can also assume that $\varepsilon < 1$ since almost surely $\|P_n^*U^*A\otimes I_nUP_n\| \leq 1$. We get the conclusion by a numerical computation.

□

We can now refine this inequality by relying on corollary 4.4.28 of [16], we state the part that we will be using in the next proposition.

Proposition 3.6. *We set $S\mathbb{U}_N = \{X\in\mathbb{U}_N \mid \det(X) = 1\}$, let f be a continuous, real-valued function on $\mathbb{U}_N$. We assume that there exists a constant C such that for every $X,Y\in\mathbb{U}_N$,*

$$|f(X) - f(Y)| \leq C\,\|X - Y\|_2 \tag{10}$$

Then if we set ν_G the law of the Haar measure on G, for all $\delta > 0$:

$$\nu_{\mathbb{U}_N}\left(\left|f(.) - \int f(Y.)d\nu_{S\mathbb{U}_N}(Y)\right| \geq \delta\right) \leq 2e^{-\frac{N\delta^2}{4C^2}} \tag{11}$$

Lemma 3.7. *For any $A\in\mathcal{D}_k$, $\varepsilon > 0$, if $kn \geq 2^{29}\times\ln^2(kn)\times\varepsilon^{-4}$, we have*

$$\mathbb{P}\left(\|P_n^*U^*\cdot A\otimes I_n\cdot UP_n\| \geq \|A\|_{(t)} + \varepsilon\right) \leq 2e^{-kn\times\frac{\varepsilon^2}{64}}.$$

Proof. We set,

$$f : U \mapsto \|P_n^*U^*A\otimes I_nUP_n\|,$$

$$h : X\in\mathbb{U}_n \mapsto \int f(YX)d\nu_{S\mathbb{U}_{kn}}(Y).$$

If U^1 is a random matrix of law $\nu_{S\mathbb{U}_{kn}}$, and α a scalar of law $\nu_{\mathbb{U}_1}$ independent of U^1. Then the law of αU^1 is $\nu_{\mathbb{U}_{kn}}$ since its law is invariant by multiplication by a unitary matrix. Consequently for any $X\in\mathbb{U}_{kn}$,

$$h(X) = \mathbb{E}[f(U^1X)] = \mathbb{E}[f(\alpha U^1 X)] = \mathbb{E}[f(\alpha U^1)] = \int f(Y) d\nu_{U_{kn}}(Y).$$

The third inequality is true since for any scalar α and $X \in \mathbb{U}_{kn}$, $f(X) = f(\alpha X)$. Besides we also have that for any $U, V \in \mathbb{U}_{kn}$,

$$|f(U) - f(V)| \leq 2 \|U - V\| \leq 2 \|U - V\|_2 .$$

Thus by using Proposition 3.6, we get

$$\mathbb{P}\left(\Big| \|P_n^* U^* \cdot A \otimes I_n \cdot U P_n\| - \mathbb{E}\Big[\|P_n^* U^* A \otimes I_n U P_n\| \Big] \Big| \geq \delta \right) \leq 2e^{-\frac{kn\delta^2}{16}}.$$

Besides if for $x \in \mathbb{R}$, we denote $x_+ = \max(0, x)$, then

$$\begin{aligned}
&\mathbb{P}\left(\|P_n^* U^* A \otimes I_n U P_n\| \geq \|A\|_{(t)} + \varepsilon \right) \\
&\leq \mathbb{P}\left(\|P_n^* U^* A \otimes I_n U P_n\| - \mathbb{E}\left[\|P_n^* U^* A \otimes I_n U P_n\|\right] \geq \varepsilon - \mathbb{E}\left[\left(\|P_n^* U^* A \otimes I_n U P_n\| - \|A\|_{(t)} \right)_+ \right] \right) \\
&\leq \mathbb{P}\left(\Big| \|P_n^* U^* A \otimes I_n U P_n\| - \mathbb{E}\Big[\|P_n^* U^* A \otimes I_n U P_n\| \Big] \Big| \geq \varepsilon - \mathbb{E}\left[\left(\|P_n^* U^* A \otimes I_n U P_n\| - \|A\|_{(t)} \right)_+ \right] \right) \\
&\leq 2e^{-kn\left(\varepsilon - \mathbb{E}\left[\left(\|P_n^* U^* A \otimes I_n U P_n\| - \|A\|_{(t)}\right)_+\right]\right)^2/16}.
\end{aligned}$$

Besides thanks to our first estimate, i.e. Lemma 3.5, we get that for any $r > 0$,

$$\begin{aligned}
\mathbb{E}\left[\left(\|P_n^* U^* \cdot A \otimes I_n \cdot U P_n\| - \|A\|_{(t)} \right)_+ \right] &\leq r + 2^{21} \times \frac{\ln^2(kn)}{kn} \int_r^1 3 \times \alpha^{-4} d\alpha \\
&\leq r + 2^{21} \times \frac{\ln^2(kn)}{kn} r^{-3}
\end{aligned}$$

And after fixing $r = \left(2^{21} \times \frac{\ln^2(kn)}{kn}\right)^{1/4}$, we get that

$$\mathbb{E}\left[\left(\|P_n^* U^* \cdot A \otimes I_n \cdot U P_n\| - \|A\|_{(t)} \right)_+ \right] \leq \left(2^{25} \times \frac{\ln^2(kn)}{kn} \right)^{1/4}.$$

Hence if $kn \geq 2^{29} \times \ln^2(kn) \times \varepsilon^{-4}$, we have

$$\mathbb{P}\left(\|P_n^* U^* \cdot A \otimes I_n \cdot U P_n\| \geq \|A\|_{(t)} + \varepsilon \right) \leq 2e^{-kn \times \frac{\varepsilon^2}{64}}.$$

□

We can finally prove Theorem 2.2 by using the former lemma in combination with Lemma 3.3.

Proof of Theorem 2.2. If we set $u = \frac{\sqrt{2}\varepsilon}{3k^2}$, then with $\mathcal{S}_u = \{uM \mid M \in \mathbb{M}_k(\mathbb{C})_{sa},\ \forall i \geq j,\ \Re(m_{i,j}) \in \{\mathbb{N} + \frac{1}{2}\} \cap [0, \lceil u^{-1} \rceil],\ \forall i > j,\ \Im(m_{i,j}) \in \{\mathbb{N} + \frac{1}{2}\} \cap [0, \lceil u^{-1} \rceil]\ \}$, Lemma 3.3 tells us that

$$\mathbb{P}\left(K_{n,k,t} \not\subset K_{k,t} + \varepsilon\right) \leq \sum_{M \in \mathcal{S}_u} \mathbb{P}\left(\|P_n^* U^* (P_{\mathcal{D}_k} M \otimes I_n) U P_n\| > \|P_{\mathcal{D}_k} M\|_{(t)} + \frac{\varepsilon}{3k} \right).$$

But thanks to Lemma 3.7, we know that for any $A \in \mathcal{D}_k$, if $n \geq 3^4 \times 2^{29} \times \ln^2(kn) \times k^3 \varepsilon^{-4}$, then

$$\mathbb{P}\left(\|P_n^* U^* A \otimes I_n U P_n\| \geq \|A\|_{(t)} + \frac{\varepsilon}{3k} \right) \leq 2e^{-\frac{n}{k} \times \frac{\varepsilon^2}{576}}.$$

Thus since the cardinal of $\mathcal{S}_u$ can be bounded by $(u^{-1}+1)^{k^2}$, we get that for $n \geq 3^4 \times 2^{29} \times \ln^2(kn) \times k^3 \varepsilon^{-4}$,

$$\mathbb{P}\left(K_{n,k,t} \not\subset K_{k,t} + \varepsilon\right) \leq 2(u^{-1}+1)^{k^2} e^{-\frac{n}{k} \times \frac{\varepsilon^2}{576}} \leq e^{k^2(\ln(3k^2\varepsilon^{-1})) - \frac{n}{k} \times \frac{\varepsilon^2}{576}}.$$

□

4 Application to Quantum Information Theory

4.1 Preliminaries on entropy

For $X \in D(H)$, its *von Neumann entropy* is defined by functional calculus by $H(X) = -\operatorname{Tr}(X \ln X)$, where $0 \ln 0$ is assumed by continuity to be zero. In other words, $H(X) = \sum_{\lambda \in spec(X)} -\lambda \ln \lambda$ where the sum is counted with multiplicity. A *quantum channel* $\Phi : B(H_1) \to B(H_2)$ is a completely positive trace preserving linear map. The *Minimum Output Entropy* (MOE) of Φ is

$$H_{min}(\Phi) = \min_{X \in D(H_1)} H(\Phi(X)). \tag{12}$$

During the last decade, a crucial problem in Quantum Information Theory was to determine whether one can find two quantum channels

$$\Phi_i : B(H_{j_i}) \to B(H_{k_i}), i = \{1, 2\},$$

such that

$$H_{min}(\Phi_1 \otimes \Phi_2) < H_{min}(\Phi_1) + H_{min}(\Phi_2).$$

Let $e_1 = (1, 0, \ldots, 0) \in \mathbb{R}^k$ and let

$$x_t^* = \left(||e_1||_t, \underbrace{\frac{1 - ||e_1||_t}{k-1}, \ldots, \frac{1 - ||e_1||_t}{k-1}}_{k-1 \text{ times}} \right). \tag{13}$$

If we view x_t^* as a diagonal matrix, then it can be easily checked that $x_t^* \in K_{k,t}$, and the following is the main result of [4]:

Theorem 4.1. *For any $p > 1$, the maximum of the l^p norm on $K_{k,t}$ is reached at the point x_t^*.*

By letting $p \to 1$ it implies that the minimum of the entropy on $K_{k,t}$ is reached at the point x_t^* and this is what we will be using. For the sake of making actual computation, it will be useful to recall the value of $||e_1||_t$. For this, we use the following notation:

$$(1^j 0^{k-j}) = (\underbrace{1, 1, \ldots, 1}_{j \text{ times}}, \underbrace{0, 0, \ldots, 0}_{k-j \text{ times}}) \in \mathbb{R}^k, \tag{14}$$

and $1^k = (1^k 0^0)$. It was proved in the early days of free probability theory (see [17]) that for $j = 1, 2, \ldots, k$, one has

$$\left\| (1^j 0^{k-j}) \right\|_{(t)} = \phi(u, t) = \begin{cases} t + u - 2tu + 2\sqrt{tu(1-t)(1-u)} & \text{if } t + u < 1, \\ 1 & \text{if } t + u \geq 1, \end{cases}$$

where $u = j/k$.

4.2 Corollary of the main result

The following is a direct consequence of the main theorem in terms of possible entropies of the output set.

Theorem 4.2. *With $S_{n,k,t} = \min_{A \in K_{n,k,t}} H(A) = H_{\min}(\Phi_n)$ and $S_{k,t} = \min_{A \in K_{k,t}} H(A)$, if we assume $d \leq tkn$, then for $n \geq 3^4 \times 2^{29} \times \ln^2(kn) \times k^3 \varepsilon^{-4}$ where $0 < \varepsilon \leq e^{-1}$,*

$$\mathbb{P}\Big(S_{n,k,t} \leq S_{k,t} - 3k\varepsilon |\ln(\varepsilon)| \Big) \leq e^{k^2 (\ln(3k^2 \varepsilon^{-1})) - \frac{n}{k} \times \frac{\varepsilon^2}{576}}.$$

Proof. Let $A, B \in \mathcal{D}_k$ such that $\|A - B\|_2 \le \varepsilon$ with $\|M\|_2 = \sqrt{\mathrm{Tr}_k(M^*M)}$, with eigenvalues $(\lambda_i)_i$ and $(\mu_i)_i$. Then with $\widetilde{x} = \max\{\varepsilon, x\}$,

$$\begin{aligned}
\Big| |\,\mathrm{Tr}_k(A\ln(A))| - |\,\mathrm{Tr}_k(B\ln(B))| \Big| &= \Big| \sum_i \lambda_i \ln(\lambda_i) - \sum_i \mu_i \ln(\mu_i) \Big| \\
&\le 2k \sup_{x\in[0,\varepsilon]} |x\ln(x)| + \Big| \sum_i \widetilde{\lambda_i} \ln(\widetilde{\lambda_i}) - \sum_i \widetilde{\mu_i} \ln(\widetilde{\mu_i}) \Big| \\
&\le 2k\varepsilon |\ln(\varepsilon)| + \sum_i |\lambda_i - \mu_i| |\ln(\varepsilon)| \\
&\le 2k\varepsilon |\ln(\varepsilon)| + k\,\|A - B\|\,|\ln(\varepsilon)| \\
&\le 3k\varepsilon |\ln(\varepsilon)|
\end{aligned}$$

Thus if we endow $\mathbb{M}_k(\mathbb{C})$ with the norm $M \mapsto \sqrt{\mathrm{Tr}_k(M^*M)}$, then if $K_{n,k,t} \subset K_{k,t} + \varepsilon$, then $S_{n,k,t} \ge \min_{A\in K_{k,t}+\varepsilon} |\,\mathrm{Tr}_k(A\ln(A))| \ge S_{k,t} - 3k\varepsilon|\ln(\varepsilon)|$. Hence

$$\mathbb{P}\Big(S_{n,k,t} \le S_{k,t} - 3k\varepsilon|\ln(\varepsilon)|\Big) \le \mathbb{P}\left(K_{n,k,t} \not\subset K_{k,t} + \varepsilon\right).$$

Theorem 2.2 then allows us to conclude.

□

4.3 Application to violation of the Minimum Output Entropy of Quantum Channels

In order to obtain violations for the additivity relation of the minimum output entropy, one needs to obtain upper bounds for the quantity $H_{\min}(\Phi \otimes \Psi)$ for some quantum channels Φ and Ψ. The idea of using conjugate channels ($\Psi = \bar{\Phi}$) and bounding the minimum output entropy by the value of the entropy at the Bell state dates back to Werner, Winter and others (we refer to [13] for references). To date, it has been proven to be the most successful method of tackling the additivity problem. The following inequality is elementary and lies at the heart of the method

$$H_{\min}(\Phi \otimes \bar{\Phi}) \le H([\Phi \otimes \bar{\Phi}](E_d)), \tag{15}$$

where E_d is the maximally entangled state over the input space $(\mathbb{C}^d)^{\otimes 2}$. More precisely, E_d is the projection on the Bell vector

$$Bell_d = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} e_i \otimes e_i, \tag{16}$$

where $\{e_i\}_{i=1}^d$ is a fixed basis of $\mathbb{C}^d$.

For random quantum channels $\Phi = \Phi_n$, the random output matrix $[\Phi_n \otimes \bar{\Phi}_n](E_d)$ was thoroughly studied in [5] in the regime $d \sim tkn$; we recall here one of the main results of that paper. There, it was proved that almost surely, as n tends to infinity, the random matrix $[\Phi_n \otimes \bar{\Phi}_n](E_{tkn}) \in M_{k^2}(\mathbb{C})$ has eigenvalues

$$\gamma_t^* = \left(t + \frac{1-t}{k^2}, \underbrace{\frac{1-t}{k^2}, \ldots, \frac{1-t}{k^2}}_{k^2-1 \text{ times}} \right). \tag{17}$$

This result improves on a bound [13] via linear algebra techniques, which states that the largest eigenvalue of the random matrix $[\Phi_n \otimes \bar{\Phi}_n](E_d)$ is at least $d/(kn) \sim t$. Although it might be possible to work directly with the bound provided by (17) with additional probabilistic consideration, for the sake of concreteness we will work with the bound of [13]. Thus if the largest eigenvalue of $[\Phi_n \otimes \bar{\Phi}_n](E_d)$ is $d/(kn)$, since $\mathrm{Tr}_k \otimes \mathrm{Tr}_k([\Phi_n \otimes \bar{\Phi}_n](E_d)) = 1$, the entropy is maximized if we take the remaining $k^2 - 1$ eigenvalues equal to $\frac{1-d/(kn)}{k^2-1}$, thus it follows that

$$H_{\min}(\Phi \otimes \bar{\Phi}) \le H([\Phi \otimes \bar{\Phi}](E_d)) \le H\left(\frac{d}{kn}, \underbrace{\frac{1-\frac{d}{kn}}{k^2-1}, \ldots, \frac{1-\frac{d}{kn}}{k^2-1}}_{k^2-1 \text{ times}} \right)$$

Therefore, it is enough to find n, k, d, t such that

$$-\frac{d}{kn}\log\left(\frac{d}{kn}\right)-\left(1-\frac{d}{kn}\right)\log\left[\left(1-\frac{d}{kn}\right)/(k^2-1)\right]<2H(x_t^*). \tag{18}$$

In [4] it was proved with the assistance of a computer that this can be done for any $k \geq 184$, as long as we take t around 1/10, see figure 1 from [4]. However for k large enough, the difference between the right and left term of (18) is maximal for $t = 1/2$. As soon as we take ε such that $3k\varepsilon|\ln(\varepsilon)|$ is less than the difference, we are done. For example, we obtain the following theorem

Theorem 4.3. *For the following values* $(k,t,n) = (184, 1/10, 10^{52}), (185, 1/10, 2\times10^{51}), (200, 1/10, 10^{47}), (500, 1/10, 4\times 10^{45}), (500, 1/2, 6\times 10^{44})$ *violation of additivity is achieved with probability at least* $1-\exp(-10^{20})$.

Proof. We make sure to work with n a multiple of 10 so that we can set $d = tkn$, then since $H_{\min}(\Phi_n) = H_{\min}(\bar{\Phi}_n)$,

$$\begin{aligned}
&\mathbb{P}\Big(H_{\min}(\Phi_n\otimes\bar{\Phi}_n) < H_{\min}(\Phi_n)+H_{\min}(\bar{\Phi}_n)\Big)\\
&=\mathbb{P}\Big(H_{\min}(\Phi_n\otimes\bar{\Phi}_n) < 2H_{\min}(\Phi_n)\Big)\\
&\leq \mathbb{P}\Big(-t\log(t)-(1-t)\log\left[(1-t)/(k^2-1)\right] < 2H_{\min}(\Phi_n)\Big)\\
&=1-\mathbb{P}\left(H_{\min}(\Phi_n)\leq -\frac{t}{2}\log(t)-\frac{1-t}{2}\log\left(\frac{1-t}{k^2-1}\right)\right)\\
&=1-\mathbb{P}\left(S_{n,k,t}\leq S_{k,t}-\delta_{k,t}\right),
\end{aligned}$$

with

$$\delta_{k,t}=\frac{t}{2}\log(t)+\frac{1-t}{2}\log\left(\frac{1-t}{k^2-1}\right)-\|e_1\|_{(t)}\log(\|e_1\|_{(t)})-(1-\|e_1\|_{(t)})\log\left(\frac{1-\|e_1\|_{(t)}}{k-1}\right).$$

Then we conclude with Theorem 4.2 to compute explicit parameters.

□

To conclude, since our bound is explicit, we solve the problem of supplying actual input dimensions for any valid output dimension, for which the violation of MOE will occur. From a point of view of theoretical probability, this is a step towards a large deviation principle. And although our bound is far from optimal, our results presumably give the right speed of deviation. However conjecturing a complete large deviation principle and a rate function seems to be beyond the scope of our techniques.

References

[1] Aubrun, G, Szarek, S; Werner, E; Hastings's additivity counterexample via Dvoretzky's theorem. *Comm. Math. Phys.* 305 (2011), no. 1, 85–97.

[2] Brandao, F., Horodecki, M. S. L. On Hastings's counterexamples to the minimum output entropy additivity conjecture. *Open Systems & Information Dynamics*, 2010, 17:01, 31–52.

[3] Belinschi, S., Collins, B. and Nechita, I. Eigenvectors and eigenvalues in a random subspace of a tensor product. *Invent. math.* 190, 647-697 (2012).

[4] Belinschi, S., Collins, B. and Nechita, I. CMP Almost one bit violation for the additivity of the minimum output entropy *Comm. Math. Phys.* 341 (2016), no. 3, 885–909.

[5] Collins, B. and Nechita Random quantum channels I: graphical calculus and the Bell state phenomenon *Comm. Math. Phys.* 297, 2 (2010) 345-370.

[6] Collins, B., Guionnet, A and Parraud, F. On the operator norm of non-commutative polynomials in deterministic matrices and iid GUE matrices *arXiv:1912.04588*, 2019.

[7] Collins, B. and Youn, S-G. Additivity violation of the regularized Minimum Output Entropy *arXiv:1907.07856* .

[8] Rockafellar, R. T. Convex analysis. *Princeton Mathematical Series*, No. 28 Princeton University Press, Princeton, N.J. 1970 xviii+451 pp.

[9] Collins, B. Haagerup's inequality and additivity violation of the Minimum Output Entropy *Houston J. Math. 44* (2018), no. 1, 253-261.

[10] B. Collins, and C. Male, The strong asymptotic freeness of Haar and deterministic matrices. *Ann. Sci. Éc. Norm. Supér.* (4) 47, 1, 147–163, 2014.

[11] Fukuda, M. and King, C. Entanglement of random subspaces via the Hastings bound. *J. Math. Phys.* 51, 042201 (2010).

[12] Fukuda, M., King, C. and Moser, D. Comments on Hastings' Additivity Counterexamples. *Commun. Math. Phys.*, vol. 296, no. 1, 111 (2010).

[13] Hayden, P. and Winter, A. Counterexamples to the maximal p-norm multiplicativity conjecture for all $p > 1$. *Comm. Math. Phys.* 284 (2008), no. 1, 263–280.

[14] Hastings, M. B. Superadditivity of communication capacity using entangled inputs. *Nature Physics* 5, 255 (2009).

[15] F. Parraud, On the operator norm of non-commutative polynomials in deterministic matrices and iid Haar unitary matrices, *arXiv:2005.13834*, 2020.

[16] G. W. Anderson, A. Guionnet, and O. Zeitouni, An introduction to random matrices, *Cambridge University Press*, volume 118 of Cambridge Studies in Advanced Mathematics Cambridge, 2010.

[17] Voiculescu, D.V., Dykema. K.J. and Nica, A. Free random variables, *AMS* (1992).

[18] A. Nica, and R. Speicher, *Lectures on the combinatorics of free probability*, volume 335 of *London Mathematical Society Lecture Note Series.* Cambridge University Press, Cambridge, 2006.