Available online at www.sciencedirect.com

ScienceDirect

Advances in Mathematics 226 (2011) 1181–1201

ADVANCES IN Mathematics

www.elsevier.com/locate/aim

Random quantum channels II: Entanglement of random subspaces, Rényi entropy estimates and additivity problems

Benoît Collins [a,b,c,*], Ion Nechita [c]

[a] *Département de Mathématique et Statistique, Université d'Ottawa, 585 King Edward, Ottawa, ON, K1N6N5 Canada*
[b] *CNRS, France*
[c] *Institut Camille Jordan Université Lyon 1, 43 Bd du 11 Novembre 1918, 69622 Villeurbanne, France*

Received 15 December 2009; accepted 3 August 2010

Available online 21 August 2010

Communicated by Dan Voiculescu

Abstract

In this paper we obtain new bounds for the minimum output entropies of random quantum channels. These bounds rely on random matrix techniques arising from free probability theory. We then revisit the counterexamples developed by Hayden and Winter to get violations of the additivity equalities for minimum output Rényi entropies. We show that random channels obtained by randomly coupling the input to a qubit violate the additivity of the p-Rényi entropy, for all $p > 1$. For some sequences of random quantum channels, we compute almost surely the limit of their Schatten $S_1 \to S_p$ norms.

MSC: 15A52; 94A17; 94A40

Keywords: Random matrices; Quantum information theory; Random channel; Additivity conjecture; Schatten p-norm

1. Introduction

The relationship between random matrix theory and free probability theory lies in the asymptotic freeness of random matrices. Asymptotic freeness, as it was discovered by Voiculescu (see

* Corresponding author at: Département de Mathématique et Statistique, Université d'Ottawa, 585 King Edward, Ottawa, ON, K1N6N5 Canada.
E-mail addresses: bcollins@uottawa.ca (B. Collins), nechita@math.univ-lyon1.fr (I. Nechita).

0001-8708/$ – see front matter
doi:10.1016/j.aim.2010.08.002

e.g. [19]), usually predicts the asymptotic pointwise behavior of joint non-commutative moments. However in some cases it can also predict more. For example, in the case of i.i.d. GUE random matrices, it was showed by Haagerup and Thorbjørnsen [10] that even the norms have an almost sure behavior predicted by free probability theory.

Quantum information theory is the analogue of classical information theory, where classical communication protocols are replaced by quantum information protocols, known as quantum channels. Despite the apparent simplicity of some mathematical question related to information theory, their resistance to various attempts to (dis)prove them have led to the study of their statistical properties.

The Holevo conjecture is arguably the most important conjecture in quantum information theory, and the theory of random matrices has been used here with success by Hayden and Winter [12,14] to produce counterexamples to the additivity conjecture of Rényi entropy for $p > 1$. These counterexamples are of great theoretical importance, as they depict the likely behavior of a random channel. In [11], Hastings gave a counterexample for the case $p = 1$. It is also of probabilistic nature, but uses a very different and less canonical measure.

In our previous paper [6], we introduced a graphical model that allowed us to understand more systematically the computation of expectation and covariance of random channels and their powers. In particular we studied at length the output of the Bell state under a random conjugate bi-channel and obtained the explicit asymptotic behavior of this random matrix.

In the present paper, we focus on the mono-channel case. Our main result is Theorem 4.1. It relies on a result obtained by one author in [5] and can be stated as follows:

Theorem 1.1. *Let k be an integer and t be a real number in $(0, 1)$. Let Φ_n be a sequence of random channels defined according to Eq.* (3)*. Then there exists a probability vector $\beta^{(t)}$ (defined in Eq.* (4)*) such that, for all $\varepsilon > 0$, almost surely when $n \to \infty$, for all input density matrix ρ, the inequality*

$$\operatorname{spec}\big(\Phi(\rho)\big) \prec \beta^{(t)} \tag{1}$$

is ε-close to being satisfied. Moreover, $\beta^{(t)}$ is optimal in the sense that any other probability vector $\beta \in \Delta_k$ satisfying the same property must satisfy $\beta^{(t)} \prec \beta$.

We combine this result with bi-channel bounds to obtain new counterexamples to the additivity conjectures for $p > 1$.

An illustration of our result is Corollary 5.6, which one can reformulate as follows:

Theorem 1.2. *For each $p > 1$ and each Hilbert space A of dimension $k' \geqslant 2$, there exists an integer such that for each Hilbert space B of dimension larger than this integer, the quantum channel arising from a random isometry into $A \otimes B$ (whose image is of appropriate relative dimension, depending on p and k') has a high probability to be Rényi strict subadditive when coupled with its conjugate.*

From a quantum information theory point of view, the true novelty of this result is that any dimension $k' \geqslant 2$ is acceptable. This result does not seem to be attainable with the alternative proofs available in [12,11,4,8].

Our techniques rely on *free probability theory*. They allow us to understand entanglement of random subspaces, and do not rely on a specific choice of a measure of entanglement. Even

though the von Neumann entropy is the most natural measure of entanglement in general, this subtlety is important as the papers [1,2] imply that all the $p \geqslant 1$ Rényi entropies don't enclose enough data to fully understand entanglement.

Our paper is organized as follows. We first recall a few basics and useful results of free probability theory of random matrix theoretical flavor in Section 2. In Section 3, we describe the random quantum channels we study. Section 4 describes the behavior of the eigenvalues of the outputs of random channels. In Section 5, we use results of the previous sections and of [6] to obtain new counterexamples to the additivity conjectures.

2. A reminder of free probability

The following is a summary of results contained in [5,18,19,7].

2.1. Asymptotic freeness

A *non-commutative probability space* is an algebra $\mathcal{A}$ with unit endowed with a tracial state φ. An element of $\mathcal{A}$ is called a (non-commutative) random variable. In this paper we shall be mostly concerned with the non-commutative probability space of *random matrices* $(\mathcal{M}_n(L^{\infty-}(\Omega, \mathbb{P})), \mathbb{E}[n^{-1}\operatorname{Tr}(\cdot)])$ (we use the standard notation $L^{\infty-}(\Omega, \mathbb{P}) = \bigcap_{p \geqslant 1} L^p(\Omega, \mathbb{P})$).

Let $\mathcal{A}_1, \ldots, \mathcal{A}_k$ be subalgebras of $\mathcal{A}$ having the same unit as $\mathcal{A}$. They are said to be *free* if for all $a_i \in \mathcal{A}_{j_i}$ $(i = 1, \ldots, k)$ such that $\varphi(a_i) = 0$, one has

$$\varphi(a_1 \cdots a_k) = 0$$

as soon as $j_1 \neq j_2$, $j_2 \neq j_3, \ldots, j_{k-1} \neq j_k$. Collections $S_1, S_2, \ldots$ of random variables are said to be free if the unital subalgebras they generate are free.

Let $(a_1, \ldots, a_k)$ be a k-tuple of selfadjoint random variables and let $\mathbb{C}\langle X_1, \ldots, X_k\rangle$ be the free $*$-algebra of non-commutative polynomials on $\mathbb{C}$ generated by the k indeterminates $X_1, \ldots, X_k$. The *joint distribution* of the family $\{a_i\}_{i=1}^k$ is the linear form

$$\mu_{(a_1,\ldots,a_k)} : \mathbb{C}\langle X_1, \ldots, X_k\rangle \to \mathbb{C},$$
$$P \mapsto \varphi\big(P(a_1, \ldots, a_k)\big).$$

Given a k-tuple $(a_1, \ldots, a_k)$ of free random variables such that the distribution of a_i is μ_{a_i}, the joint distribution $\mu_{(a_1,\ldots,a_k)}$ is uniquely determined by the μ_{a_i}'s. A family $(a_1^n, \ldots, a_k^n)_n$ of k-tuples of random variables is said to *converge in distribution* towards $(a_1, \ldots, a_k)$ iff for all $P \in \mathbb{C}\langle X_1, \ldots, X_k\rangle$, $\mu_{(a_1^n,\ldots,a_k^n)}(P)$ converges towards $\mu_{(a_1,\ldots,a_k)}(P)$ as $n \to \infty$. Sequences of random variables $(a_1^n)_n, \ldots, (a_k^n)_n$ are called *asymptotically free* as $n \to \infty$ iff the k-tuple $(a_1^n, \ldots, a_k^n)_n$ converges in distribution towards a family of free random variables.

The following result was contained in [18] (see also [7]).

Theorem 2.1. *Let $\{U_k^{(n)}\}_{k\in\mathbb{N}}$ be a collection of independent Haar distributed random matrices of $\mathcal{M}_n(\mathbb{C})$ and $\{W_k^{(n)}\}_{k\in\mathbb{N}}$ be a set of constant matrices of $\mathcal{M}_n(\mathbb{C})$ admitting a joint limit distribution as $n \to \infty$ with respect to the state $n^{-1}\operatorname{Tr}$. Then, almost surely, the family $\{U_k^{(n)}, W_k^{(n)}\}_{k\in\mathbb{N}}$ admits a limit $*$-distribution $\{u_k, w_k\}_{k\in\mathbb{N}}$ with respect to $n^{-1}\operatorname{Tr}$, such that $u_1, u_2, \ldots, \{w_1, w_2, \ldots\}$ are free.*

2.2. Free projectors

Let us fix real numbers $0 \leqslant \alpha, \beta \leqslant 1$, and consider, for all n, a selfadjoint projector $\pi_n \in \mathcal{M}_n(\mathbb{C})$ of rank q_n such that asymptotically $q_n \sim \alpha n$ as $n \to \infty$. Let π_n' be a projector of rank q_n' such that $q_n' \sim \beta n$, and assume that it can be written under the form $U\pi_n^0 U^*$, where U is a Haar distributed unitary random matrix and π_n^0 is a deterministic selfadjoint projector.

It is a consequence of Theorem 2.1, that π_n and π_n' are asymptotically free. Therefore $\pi_n\pi_n'\pi_n$ has an empirical eigenvalues distribution converging towards a probability measure. This measure is usually denoted by $\mu_1 \boxtimes \mu_2$, where μ_1, μ_2 are the limit empirical eigenvalue distributions of the projectors π_n and π_n' respectively:

$$\mu_1 = (1-\alpha)\delta_0 + \alpha\delta_1, \qquad \mu_2 = (1-\beta)\delta_0 + \beta\delta_1.$$

In this specific case, we can compute explicitly $\mu_1 \boxtimes \mu_2$. For this purpose, we introduce two 2-variable functions which will be of great importance in what follows.

$$\begin{aligned}
\varphi^+ : \big\{(x,y) \in [0,1]^2\big\} &\to [0,1], \\
(x,y) &\mapsto 1 - \big[\sqrt{(1-x)(1-y)} - \sqrt{xy}\big]^2, \\
\varphi^- : \big\{(x,y) \in [0,1]^2\big\} &\to [0,1], \\
(x,y) &\mapsto 1 - \big[\sqrt{(1-x)(1-y)} + \sqrt{xy}\big]^2.
\end{aligned}$$

Let us omit the variables of $\varphi^{+/-}$ and rewrite

$$\varphi^{+/-} = \varphi^{+/-}(\alpha,\beta) = \alpha + \beta - 2\alpha\beta \pm \sqrt{4\alpha\beta(1-\alpha)(1-\beta)}.$$

It follows then from [19, Example 3.6.7], that

$$\mu_1 \boxtimes \mu_2 = \big[1 - \min(\alpha,\beta)\big]\delta_0 + \big[\max(\alpha+\beta-1,0)\big]\delta_1 + \frac{\sqrt{(\varphi^+ - x)(x - \varphi^-)}}{2\pi x(1-x)} 1_{[\varphi^-,\varphi^+]}\,dx.$$

The proof relies on a technique introduced by Voiculescu to compute $\mu_1 \boxtimes \mu_2$ in general, called the S-transform. For more details, we refer the interested reader to [19]. Since we are only interested in φ^+, we consider the two-variable function $\varphi : [0,1]^2 \to [0,1]$

$$\varphi(x,y) = \begin{cases} 0 & \text{if } x = 0 \text{ or } y = 0; \\ \varphi^+(x,y) & \text{if } x, y > 0 \text{ and } x + y \leqslant 1; \\ 1 & \text{if } x + y > 1. \end{cases}$$

In the case where $\alpha + \beta < 1$, the ranges of π_n and π_n' do not (generically) overlap and $\varphi(\alpha,\beta) < 1$. The previous asymptotic freeness results imply that almost surely,

$$\liminf_n \big\|\pi_n\pi_n'\pi_n\big\|_\infty \geqslant \varphi(\alpha,\beta).$$

We are interested in whether we actually have

$$\lim_n \big\|\pi_n\pi_n'\pi_n\big\|_\infty = \varphi(\alpha,\beta) < 1.$$

This turns out to be true, and can be found in Proposition 4.9 of [5]. For the purposes of this paper, we state this result as follows:

Theorem 2.2. *In $\mathbb{C}^n$, choose at random according to the Haar measure two independent subspaces V_n and V_n' of respective dimensions $q_n \sim \alpha n$ and $q_n' \sim \beta n$ where $\alpha, \beta \in (0,1)$. Let π_n (resp. π_n') be the orthogonal projection onto V_n (resp. V_n'). Then,*

$$\lim_n \left\| \pi_n \pi_n' \pi_n \right\|_\infty = \varphi(\alpha, \beta).$$

The proof of this result is technical and is contained in [5], so we do not discuss it here in detail. Let us just mention that it follows from the fact that the eigenvalues of $\pi_n \pi_n' \pi_n$ are a random set whose points follow the law of a determinantal point process (see [17] for a definition). The kernel of this determinantal point process can be computed explicitly in terms of Jacobi polynomials, and the study of their asymptotics together with explicit formulas for the probability that a determinantal point process never intersects an ensemble gives the result. Actually, Proposition 4.8 of [5] shows that the probability that an eigenvalue of $\pi_n \pi_n' \pi_n$ belongs to $[\varphi(\alpha, \beta) + \varepsilon, 1]$ for any $\varepsilon > 0$ decays quicker than $e^{-C_\varepsilon n}$ for some $C_\varepsilon > 0$, and the result follows by the Borel–Cantelli lemma.

Note that a moment approach towards this result would also be possible, cf. [15].

3. Quantum channels and additivity conjectures

3.1. Rényi entropies and minimum output entropies

Let $\Delta_k = \{x \in \mathbb{R}_+^k \mid \sum_{i=1}^k x_i = 1\}$ be the $(k-1)$-dimensional probability simplex. For a positive real number $p > 0$, define the *Rényi entropy of order p* of a probability vector $x \in \Delta_k$ to be

$$H^p(x) = \frac{1}{1-p} \log \sum_{i=1}^k x_i^p.$$

Since $\lim_{p \to 1} H^p(x)$ exists, we define the *Shannon entropy* of x to be this limit, namely:

$$H(x) = H^1(x) = -\sum_{i=1}^k x_i \log x_i.$$

We extend these definitions to density matrices by functional calculus:

$$H^p(\rho) = \frac{1}{1-p} \log \operatorname{Tr} \rho^p;$$

$$H(\rho) = H^1(\rho) = -\operatorname{Tr} \rho \log \rho.$$

Given a vector $x \in \mathbb{C}^n$, $\|x\| = 1$, we call P_x the rank one orthogonal projection onto the span of x. Using Dirac's bra-ket notation, $P_x = |x\rangle\langle x|$. More generally, for a subspace $V \subset \mathbb{C}^n$, we denote by P_V the orthogonal projection onto V in $\mathcal{M}_n(\mathbb{C})$.

A *quantum channel* is a linear completely positive trace preserving map $\Phi : \mathcal{M}_n(\mathbb{C}) \to \mathcal{M}_k(\mathbb{C})$. The trace preservation condition means that density matrices are mapped to density matrices, and the complete positivity reads:

$$\forall d \geqslant 1, \quad \Phi \otimes \mathrm{I}_d : \mathcal{M}_{nd}(\mathbb{C}) \to \mathcal{M}_{kd}(\mathbb{C}) \text{ is a positive map.}$$

We recall that according to Stinespring theorem, a linear map $\Phi : \mathcal{M}_n(\mathbb{C}) \to \mathcal{M}_k(\mathbb{C})$ is a quantum channel if and only if there exists a finite dimensional Hilbert space $\mathcal{K} = \mathbb{C}^d$, and a partial isometry $V \in \mathrm{End}(\mathbb{C}^n, \mathbb{C}^{kd})$ (satisfying $V^*V = \mathrm{I}_n$) such that

$$\Phi(X) = \mathrm{Tr}_{\mathcal{K}}\big[VXV^*\big], \quad \forall X \in \mathcal{M}_n(\mathbb{C}). \tag{2}$$

For a quantum channel $\Phi : \mathcal{M}_n(\mathbb{C}) \to \mathcal{M}_k(\mathbb{C})$, we define its *minimum output Rényi entropy (of order p)* by

$$H^p_{\min}(\Phi) = \min_{\substack{\rho \in \mathcal{M}_n(\mathbb{C}) \\ \rho \geqslant 0,\, \mathrm{Tr}\,\rho = 1}} H^p\big(\Phi(\rho)\big).$$

Since the Rényi entropies are Schur-concave functions, their minima are attained on the extremal points of the set of density matrices and hence

$$H^p_{\min}(\Phi) = \min_{\substack{x \in \mathbb{C}^n \\ \|x\| = 1}} H^p\big(\Phi(P_x)\big).$$

3.2. The random quantum channel model

We fix an integer k and a real number $t \in (0, 1)$. For each n, let $U_n \in \mathcal{M}_{nk}(\mathbb{C})$ be a random unitary matrix distributed according to the Haar measure, and q_n be a projection of $\mathcal{M}_{nk}(\mathbb{C})$ of trace p_n such that $p_n/(kn) \sim t$ as $n \to \infty$. To q_n we associate a non-unital matrix algebra map $\chi_n : \mathcal{M}_{p_n}(\mathbb{C}) \to \mathcal{M}_{nk}(\mathbb{C})$ satisfying $\chi_n(1) = q_n$. The choice of χ_n is unique up to unitary conjugation, and the actual choice of χ_n is irrelevant for the computations we want to perform – in the sense that any choice will yield the same results.

We study the sequence of random channels

$$\Phi_n : \mathcal{M}_{p_n}(\mathbb{C}) \to \mathcal{M}_k(\mathbb{C})$$

given by

$$\Phi_n(X) = \mathrm{Tr}_n\big(U_n\big(\chi_n(X)\big)U_n^*\big), \tag{3}$$

where $\mathrm{Tr}_n(\cdot)$ is the partial trace operation over $\mathbb{C}^n$.

Remark 3.1. In our previous paper [6], we considered exactly the same model of random quantum channels, with one small difference: the partial trace was taken with respect to $\mathbb{C}^k$. However, it is well known that, when partial tracing a rank one projector, the non-zero eigenvalue of the resulting matrix do not depend on which space is traced out. Hence, from the point of view of eigenvalue statistics, the model we consider here is identical with the one in [6, Section 6.2].

We are interested in the random process given by the *set of all possible* eigenvalues of $\Phi(X)$ as $n \to \infty$. In our setup, we deal with k eigenvalues.

Let V_n be the image of $U\chi_n(1)U^*$. This is a random vector space in $\mathbb{C}^n \otimes \mathbb{C}^k$ of dimension p_n distributed according to the uniform measure on the Grassmannian space $\mathrm{Gr}_{p_n}(\mathbb{C}^{nk})$.

If we can ensure that the entanglement of *every* norm one vector $x \in V_n$ in $\mathbb{C}^n \otimes \mathbb{C}^k$ is large with high probability for the uniform measure on $V_n \in \mathrm{Gr}_{p_n}(\mathbb{C}^{nk})$, this will yield new entropy bounds. The entropy of entanglement of a vector $x \in \mathbb{C}^n \otimes \mathbb{C}^k$ is a concave function of the principal values of x. We recall that for an element $x \in \mathbb{C}^n \otimes \mathbb{C}^k$ we denote $\lambda(x)$ and $\mathrm{rk}(x)$ the singular values and the rank of x, when viewed as a matrix $x \in \mathcal{M}_{n\times k}(\mathbb{C})$. In quantum information theory, these quantities are also called the *Schmidt coefficients* and the *Schmidt rank* of x respectively:

$$x = \sum_{i=1}^{\mathrm{rk}(x)} \sqrt{\lambda_i(x)} e_i \otimes f_i,$$

where $\{e_i\}$ and $\{f_i\}$ are orthonormal families from $\mathbb{C}^n$ and $\mathbb{C}^k$ respectively. Both quantities can also be expressed as the rank and respectively the spectrum of the reduced density matrix $\mathrm{Tr}_n P_x$. The strategy adopted in this paper is to describe a convex polyhedron such that with high probability, for a vector subspace V chosen at random, for all input $x \in V$, the eigenvalue vector $\lambda(x)$ belongs to a neighborhood of this convex set.

3.3. Known bounds

Some results are already available in order to quantify the entanglement of generic spaces in $\mathrm{Gr}_{p_n}(\mathbb{C}^n \otimes \mathbb{C}^k)$ [16]. The best result known so far is arguably the following theorem of Hayden, Leung and Winter in [13]:

Theorem 3.2. *(See Hayden, Leung, Winter [13, Theorem IV.1].) Let A and B be quantum systems of dimension d_A and d_B with $d_B \geqslant d_A \geqslant 3$. Let $0 < \alpha < \log d_A$. Then there exists a subspace $S \subset A \otimes B$ of dimension*

$$d \sim d_A d_B \frac{\Gamma \alpha^{2.5}}{(\log d_A)^{2.5}}$$

such that all states $x \in S$ have entanglement satisfying

$$H\big(\lambda(x)\big) \geqslant \log d_A - \alpha - \beta,$$

where $\beta = d_A/(d_B \log 2)$ and $\Gamma = 1/1753$.

To prove this result, the authors require sophisticated methods from asymptotic geometry theory. In particular, they need estimates on the covering numbers of unitary groups by balls of radius ε and results of concentration of measure. The results of concentration of measure are applied to a specific measure of entanglement (e.g. one entropy H^p), therefore the measure of entanglement does not deal directly with the behavior of the Schmidt coefficients, but rather with the behavior of a function of them.

4. Confining the eigenvalues almost surely

4.1. Main result

Our strategy is to describe a convex polyhedron K inside the probability simplex Δ_k with the property that, for all $\varepsilon > 0$, almost surely when n goes to infinity, all input density matrices $\rho \geqslant 0$, $\mathrm{Tr}(\rho) = 1$, are mapped to output states $\Phi(\rho)$ whose spectra are contained $K + \varepsilon$, the ε-neighborhood of K.

For $t \in (0, 1)$, let us first define the vector $\beta^{(t)} \in \mathbb{R}^k$, where

$$\beta_j^{(t)} = \varphi\left(\frac{j}{k}, t\right) - \varphi\left(\frac{j-1}{k}, t\right), \quad \forall 1 \leqslant j \leqslant k. \tag{4}$$

One can check directly that $\beta^{(t)}$ is a probability vector and that it is a non-increasing sequence. Moreover, $\beta_1^{(t)} = \varphi(1/k, t)$ and $\beta_j^{(t)} = 0$ for $j \geqslant \lfloor k(1-t) \rfloor + 2$.

Since the *majorization* partial order plays an important role in this situation, let us remind here the definition and some basic properties of this relation. For two probability vectors $x, y \in \Delta_k$, we say that x is *majorized* by y (and we write $x \prec y$) iff for all $j \in \{1, \ldots, k\}$

$$s_j(x) = \sum_{i=1}^{j} x_i^{\downarrow} \leqslant \sum_{i=1}^{j} y_i^{\downarrow} = s_j(y), \tag{5}$$

where $x^{\downarrow}$ and $y^{\downarrow}$ are the decreasing rearrangements of x and y; note that for $j = k$ we actually have an equality, since x and y are probability vectors. We extend the functions s_j, by functional calculus, to selfadjoint matrices $X \in \mathcal{M}_k(\mathbb{C})$; these quantities are called the *Ky Fan norms* of the matrix X, see [3]. The majorization relation can also be characterized in the following way: for a probability vector y and a permutation $\sigma \in \mathfrak{S}_k$, denote by $\sigma.y$ the vector obtained by permuting the coordinates of y along σ: $(\sigma.y)_i = y_{\sigma(i)}$. Then

$$x \prec y \quad \text{iff} \quad x \in S(y),$$

where $S(y)$ is the convex hull of the set $\{\sigma.y \mid \sigma \in \mathfrak{S}_k\}$. Moreover, the extremal points of $S(y)$ are exactly y and its permutations $\sigma.y$. In Fig. 1, we plot Δ_3, the 2-dimensional simplex together with the sets $S(\beta^{(t)})$, for $t = 1/k'$ and $k' = 2, 3, 4, 5, 10, 20, 50, 100$. Notice that for $k' = 2, 3$, the set $S(\beta^{(1/k')})$ touches the triangle Δ_3, because of the fact that $\beta^{(t)}$ has in this case a zero coordinate.

We can now state the main result of this section:

Theorem 4.1. *Let t be a parameter in $(0, 1)$ and $\varepsilon > 0$. Let $S(\beta^{(t)}) + \varepsilon$ be the ε-ball around $S(\beta^{(t)})$ in Δ_k. Then, almost surely when $n \to \infty$, for all input density matrix ρ,*

$$\mathrm{spec}\big(\Phi_n(\rho)\big) \in S\big(\beta^{(t)}\big) + \varepsilon. \tag{6}$$

Moreover, $\beta^{(t)}$ is optimal: a probability vector $\beta \in \Delta_k$ such that, with positive probability,

$$\mathrm{spec}\big(\Phi_n(\rho)\big) \in S(\beta) + \varepsilon, \quad \forall \rho \tag{7}$$

must satisfy $\beta^{(t)} \prec \beta$.

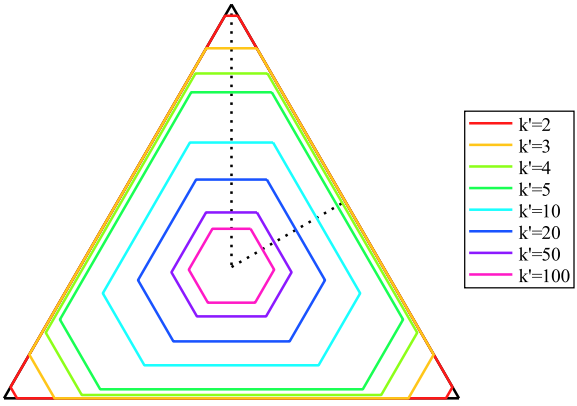

Fig. 1. The 2-dimensional probability simplex with the sets $S(\beta^{(t)})$, for $t = 1/k'$ and $k' = 2, 3, 4, 5, 10, 20, 50, 100$.

We split the proof of Theorem 4.1 into several lemmas. The first one is an easy consequence of the definition of the operator norm.

Lemma 4.2. *Let Q, R be two selfadjoint projections in $\mathcal{M}_n(\mathbb{C})$. Then*

$$\|QRQ\|_\infty = \max_{x \in \operatorname{Im} Q} \operatorname{Tr}(P_x R).$$

Proof. Since QRQ is a selfadjoint operator, we have:

$$\begin{aligned} \|QRQ\|_\infty &= \sup_{\|y\| \leqslant 1} \langle QRQy, y\rangle = \sup_{\|y\| \leqslant 1} \langle RQy, Qy\rangle \\ &= \sup_{\substack{x \in \operatorname{Im} Q \\ \|x\| \leqslant 1}} \langle Rx, x\rangle = \max_{x \in \operatorname{Im} Q} \operatorname{Tr}(P_x R). \qquad \square \end{aligned}$$

The following lemma is a reformulation of the min–max theorem:

Lemma 4.3. *Let $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_k$ be the Schmidt coefficients of a vector $x \in \mathbb{C}^{nk}$. Then, for all $1 \leqslant j \leqslant k$,*

$$s_j(x) = \lambda_1 + \lambda_2 + \cdots + \lambda_j = \max_{F \in \operatorname{Gr}_j(\mathbb{C}^k)} \operatorname{Tr}(P_x P_{\mathbb{C}^n \otimes F}).$$

Proof. Since λ_i are the eigenvalues of $\operatorname{Tr}_n P_x \in \mathcal{M}_k(\mathbb{C})$, the min–max theorem for $\operatorname{Tr}_n P_x$ can be stated as:

$$s_j(x) = \max_{F \in \operatorname{Gr}_j(\mathbb{C}^k)} \operatorname{Tr}(P_F \operatorname{Tr}_n P_x).$$

The conditional expectation property of the partial trace implies that

$$s_j(x) = \max_{F \in \operatorname{Gr}_j(\mathbb{C}^k)} \operatorname{Tr}(P_x \cdot \mathrm{I}_n \otimes P_F) = \max_{F \in \operatorname{Gr}_j(\mathbb{C}^k)} \operatorname{Tr}(P_x \cdot P_{\mathbb{C}^n \otimes F}). \qquad \square$$

We are interested in majorization inequalities which hold uniformly for all norm one elements of a subspace V. In other words, we are interested in the quantity

$$\max_{\substack{x\in V\\ \|x\|=1}} s_j(x) = \max_{\substack{x\in V\\ \|x\|=1}} \max_{F\in \mathrm{Gr}_j(\mathbb{C}^k)} \mathrm{Tr}(P_x P_{\mathbb{C}^n\otimes F}).$$

Since k is a fixed parameter of our model, in order to compute the maximum over the Grassmannian, it suffices to consider only a finite number of subspaces F:

Lemma 4.4. *For all $\varepsilon > 0$, for all j, there exists a finite number of j-dimensional subspaces $F_1, \ldots, F_N \in \mathrm{Gr}_j(\mathbb{C}^k)$ such that, for all $x \in \mathbb{C}^{nk}$,*

$$\max_{i=1}^{N} \mathrm{Tr}(P_x P_{\mathbb{C}^n\otimes F_i}) \leqslant s_j(x) \leqslant \max_{i=1}^{N} \mathrm{Tr}(P_x P_{\mathbb{C}^n\otimes F_i}) + \varepsilon.$$

Note that in Lemma 4.4, N does depend on ε but can be chosen to be finite for any $\varepsilon > 0$.

Proof. We only need to prove the second inequality. Since the Grassmannian $\mathrm{Gr}_j(\mathbb{C}^k)$ is compact and metric for $d(E, F) = \|P_E - P_F\|_\infty$, for all $\varepsilon > 0$ there exists a covering of $\mathrm{Gr}_j(\mathbb{C}^k)$ by a finite number of balls of radius ε centered in $F_1, \ldots, F_N$. Fix some $x \in \mathbb{C}^{nk}$ and consider the element $F \in \mathrm{Gr}_j(\mathbb{C}^k)$ for which the maximum in the definition of $s_j(x)$ is attained. F is inside some ball centered at F_i and we have

$$\begin{aligned}\mathrm{Tr}(P_x P_{\mathbb{C}^n\otimes F}) &\leqslant \mathrm{Tr}(P_x P_{\mathbb{C}^n\otimes F_i}) + \left|\mathrm{Tr}\left(P_x (P_{\mathbb{C}^n\otimes F} - P_{\mathbb{C}^n\otimes F_i})\right)\right| \\ &= \mathrm{Tr}(P_x P_{\mathbb{C}^n\otimes F_i}) + \|P_F - P_{F_i}\|_\infty \leqslant \mathrm{Tr}(P_x P_{\mathbb{C}^n\otimes F_i}) + \varepsilon,\end{aligned}$$

and the conclusion follows. □

Now we are ready to prove Theorem 4.1.

Proof of Theorem 4.1. First, notice that it suffices to show (6) holds for rank one projectors $\rho = P_x$. The general case follows from the convexity of the functions $s_1, \ldots, s_k$.

Let $\varepsilon > 0$ and $j \in \{1, \ldots, k\}$. For a random subspace $V \subset \mathbb{C}^{nk}$ of dimension $p_n \sim tnk$,

$$\max_{\substack{x\in V\\ \|x\|=1}} s_j(x) = \max_{\substack{x\in V\\ \|x\|=1}} \max_{F\in \mathrm{Gr}_j(\mathbb{C}^k)} \mathrm{Tr}(P_x P_{\mathbb{C}^n\otimes F}).$$

Using the compactness argument in Lemma 4.4, one can consider (at a cost of ε) only a finite number of subspaces F_i:

$$\max_{\substack{x\in V\\ \|x\|=1}} s_j(x) \leqslant \max_{i=1}^{N} \max_{\substack{x\in V\\ \|x\|=1}} \mathrm{Tr}(P_x P_{\mathbb{C}^n\otimes F_i}) + \varepsilon.$$

According to Theorem 2.2, for all $i \in \{1, \ldots, N\}$, almost surely when $n \to \infty$,

$$\lim_n \|P_V P_{\mathbb{C}^n\otimes F_i} P_V\|_\infty = \varphi(j/k, t).$$

Since N is finite, with probability one, the above equality is true *for all* i. Next, using Lemmas 4.2 and 4.3, one has that, almost surely,

$$\limsup_n \max_{\substack{x\in V\\ \|x\|=1}} s_j(x) \leqslant \varphi(j/k,t)+\varepsilon,$$

which concludes the proof of the direct implication.

Conversely, let $\beta \in \Delta_k$ be a probability vector which satisfies Eq. (7). For $j \in \{1,2,\ldots,k\}$ fixed, let F_0 be a subspace of $\mathbb{C}^k$ of dimension j. We have

$$\begin{aligned}\max_{\substack{x\in V\\ \|x\|=1}} s_j(x) &= \max_{\substack{x\in V\\ \|x\|=1}} \max_{F\in \mathrm{Gr}_j(\mathbb{C}^k)} \mathrm{Tr}(P_x P_{\mathbb{C}^n\otimes F})\\ &\geqslant \max_{\substack{x\in V\\ \|x\|=1}} \mathrm{Tr}(P_x P_{\mathbb{C}^n\otimes F_0}) = \|P_V P_{\mathbb{C}^n\otimes F_0} P_V\|_\infty \xrightarrow[n\to\infty]{\text{a.s.}} \varphi(j/k,t).\end{aligned}$$

Since, with positive probability, $\max_{x\in V,\,\|x\|=1} s_j(x) \leqslant s_j(\beta)+\varepsilon$, we conclude that $s_j(\beta) \geqslant \varphi(j/k,t) = s_j(\beta^{(t)})$ and the proof is complete. □

The interest of Theorem 4.1 in comparison to Theorem 3.2 is that it does not rely specifically on one measurement of entanglement, as we are able to confine almost surely the eigenvalues in a convex set. Also, our argument relies neither on concentration inequalities nor on net estimates, as we fix k. However, unlike Theorem 3.2, Theorem 4.1 does not give explicit control on n. It is theoretically possible to give an explicit control on n, but this would require to work out explicitly many constants in the paper [5] and apart from being potentially quite cumbersome, this amounts to computational saddle point analysis type refinement of the paper [5], rather than to conceptual free probability theory.

4.2. Application to entropies

Once the eigenvalues of the output of a channel have been confined inside a fixed convex polyhedron, entropy inequalities follow easily. Indeed, the confining polyhedron is defined in terms of the majorization partial order, and thus the notion of Schur-convexity (see [3]) is crucial in what follows.

A function $f:\mathbb{R}^k \to \mathbb{R}$ is said to be Schur-convex if $x \prec y$ implies $f(x) \leqslant f(y)$. The Rényi entropies H^p are Schur-concave, and thus majorization relations $x \prec y$ imply $H^p(x) \geqslant H^p(y)$ for all $p \geqslant 1$. The reciprocal implication has been studied in [1,2]: entropy inequalities $H^p(x) \geqslant H^p(y)$ (for all $p \geqslant 1$) characterize a weaker form of majorization called *catalytic majorization*, which has applications in LOCC protocols for the transformation of bipartite states.

For the purposes of this paper, the main corollary of Theorem 4.1 is the following:

Theorem 4.5. *For a fixed parameter t, almost surely when $n \to \infty$,*

$$\liminf_n H^p_{\min}(\Phi_n) \geqslant H^p\big(\beta^{(t)}\big).$$

Proof. This follows directly from Theorem 4.1 and from the Schur-concavity of the Rényi entropies. □

5. New examples and counterexamples of strictly subadditive channels

Since our main result, Theorem 4.1, is valid almost surely in the limit $n \to \infty$, the limiting objects depend only on the (*a priori* fixed) parameters k and t. In what follows, we consider large values of the parameter k, and introduce the "little-o" notation $o(\cdot)$ with respect to the limit $k \to \infty$.

5.1. Strict subadditivity

We start with a crucial recent series of result, which we summarize into the following theorem:

Theorem 5.1. *For every $p \geqslant 1$, there exist quantum channels Φ_1 and Φ_2 such that*

$$H_{\min}^{p}(\Phi_1 \otimes \Phi_2) < H_{\min}^{p}(\Phi_1) + H_{\min}(\Phi_2). \tag{8}$$

This theorem results mainly from the papers [11,12,14]. Even nowadays, no explicit non-random counterexamples are known for $1 \leqslant p \leqslant 2$.

5.2. The Bell phenomenon

In order to provide counterexamples for the additivity conjectures, one has to produce lower bounds for the minimum output entropy of single copies of the channels (and this is where Theorem 4.1 is useful) and upper bounds for the minimum output entropy of the tensor product of the quantum channels. The latter task is somewhat easier, since one has to exhibit a particular input state such that the output has low entropy.

The choice of the input state for the product channel is guided by the following observation. It is clear that if one chooses a product input state $\rho = \rho_1 \otimes \rho_2$, then the output state is still in product form, and the entropies add up:

$$H^{p}\big([\Phi_1 \otimes \Phi_2](\rho_1 \otimes \rho_2)\big) = H^{p}\big(\Phi_1(\rho_1) \otimes \Phi_2(\rho_2)\big) = H^{p}\big(\Phi_1(\rho_1)\big) + H^{p}\big(\Phi_2(\rho_2)\big).$$

Hence, such choices cannot violate the additivity of Rényi entropies. Instead, one has to look at *entangled* states, and the maximally entangled states are obvious candidates.

All our examples rely on the study of the product of conjugate channels

$$\Phi_n \otimes \overline{\Phi}_n$$

where

$$\Phi_n(X) = \mathrm{Tr}_n\big(U_n \chi_n(X) U_n^*\big), \qquad \overline{\Phi}_n(X) = \mathrm{Tr}_n\big(\overline{U}_n \chi_n(X) U_n^t\big)$$

have been introduced in Section 3.2. Our task is to obtain a good upper bound for

$$\limsup_n H_{\min}^{p}(\Phi_n \otimes \overline{\Phi}_n).$$

Our strategy is systematically to write

$$\limsup_n H_{\min}^{p}(\Phi_n \otimes \overline{\Phi}_n) \leqslant H_{\min}^{p}\big(\Phi_n \otimes \overline{\Phi}_n(E_{tnk})\big)$$

where E_{tnk} is the maximally entangled state over the input space $(\mathbb{C}^{tnk})^{\otimes 2}$, where we tacitly assume that tnk is an integer. More precisely, E_{tnk} is the projection on the Bell vector

$$\mathrm{Bell}_{tnk} = \frac{1}{\sqrt{tnk}} \sum_{i=1}^{tnk} e_i \otimes e_i,$$

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

The random matrix $\Phi_n \otimes \overline{\Phi}_n(E_{tnk})$ was thoroughly studied in our previous paper [6] and we recall here one of the main results of this paper:

Theorem 5.2. *Almost surely, as $n \to \infty$, the random matrix $\Phi_n \otimes \overline{\Phi}_n(\mathrm{Bell}_{tnk}) \in \mathcal{M}_{k^2}(\mathbb{C})$ has eigenvalues*

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

From this we deduce the following corollary, which gives an upper bound for the minimum output entropy for the product channel $\Phi \otimes \overline{\Phi}$:

Corollary 5.3. *Almost surely, as $n \to \infty$,*

$$\limsup_n H_{\min}^p(\Phi_n \otimes \overline{\Phi}_n) \leqslant \frac{1}{1-p} \log \left[\left(t + \frac{1-t}{k^2} \right)^p + \left(k^2 - 1\right) \left(\frac{1-t}{k^2} \right)^p \right].$$

In the case $p = 1$ the upper bound is modified to

$$\limsup_n H_{\min}(\Phi_n \otimes \overline{\Phi}_n) \leqslant -\left(t + \frac{1-t}{k^2} \right) \log \left(t + \frac{1-t}{k^2} \right) - \left(k^2 - 1\right) \frac{1-t}{k^2} \log \left(\frac{1-t}{k^2} \right).$$

5.3. Macroscopic counterexamples for the Rényi entropy

In this section, we start by fixing $t = 1/2$. We assume that k is even, in order to avoid non-integer dimensions. A value of $1/2$ for t means that the environment to which the input of the channel is coupled is 2-dimensional, i.e. a single qubit. The main result of this section is that we obtain a violation of the Rényi entropy in this simplest purely quantum case, $k' = 2$.

Using Theorem 5.2, the asymptotic eigenvalue vector for the output of the product channel is

$$\gamma = \gamma^{(1/2)} = \left(\frac{1}{2} + \frac{1}{2k^2}, \frac{1}{2k^2}, \ldots, \frac{1}{2k^2} \right).$$

The series expansion for $H^p(\gamma)$ when $k \to \infty$ and Corollary 5.3 imply that, almost surely,

$$\limsup_n H_{\min}^p(\Phi \otimes \overline{\Phi}) \leqslant \frac{p}{p-1} \log 2 + o(1). \tag{9}$$

In the case of a single channel, since $\varphi(x, 1/2) = 1/2 + \sqrt{x(1-x)}$, the vector $\beta = \beta^{(1/2)}$ has a particularly simple form:

$$\beta_1 = \frac{1}{2} + \frac{\sqrt{k-1}}{k},$$
$$\beta_j = \psi\left(\frac{j}{k}\right) - \psi\left(\frac{j-1}{k}\right), \quad \forall 2 \leqslant j \leqslant k/2,$$
$$\beta_j = 0, \quad \forall k/2 < j \leqslant k,$$

where $\psi(x) = \sqrt{x(1-x)}$. Note that the first eigenvalue is large (of order $1/2$) and that the others are small:

$$\beta_j \leqslant \frac{1}{k}\psi'\left(\frac{1}{k}\right) \leqslant \frac{1}{\sqrt{k}}, \quad \forall 2 \leqslant j \leqslant k/2.$$

Theorem 5.4. *Almost surely as $n \to \infty$,*

$$\liminf_n H_{\min}^p(\Phi) = \liminf_n H_{\min}^p(\overline{\Phi}) \geqslant \frac{p}{p-1}\log 2 + o(1).$$

Since

$$\limsup_n H_{\min}^p(\Phi \otimes \overline{\Phi}) \leqslant \frac{p}{p-1}\log 2 + o(1),$$

the additivity of the Rényi p-norms is violated for all $p > 1$.

Proof. We shall provide a lower bound for $H^p(\beta)$. Notice that the main contribution is given by the largest eigenvalue: $\beta_1^p = 2^{-p} + o(1)$. Next, we show that the contribution of the smaller eigenvalues is asymptotically zero. We consider three cases: $p > 2$, $p = 2$ and $1 < p < 2$. If $p > 2$, then

$$\sum_{j \geqslant 2} \beta_j^p \leqslant \sum_{j \geqslant 2} k^{-p/2} \leqslant k^{1-p/2} = o(1).$$

For $p = 2$, one has

$$\sum_{j \geqslant 2} \beta_j^2 = \sum_{j=2}^{k/2}\left[\psi\left(\frac{j}{k}\right) - \psi\left(\frac{j-1}{k}\right)\right]^2 \leqslant \sum_{j=2}^{k/2}\left[\frac{1}{k} \cdot \sup_{(j-1)/k \leqslant x \leqslant j/k} \psi'(x)\right]^2$$
$$= \sum_{j=2}^{k/2}\left[\frac{1}{k}\psi'\left(\frac{j-1}{k}\right)\right]^2 = \frac{1}{k}\sum_{j=1}^{k/2-1} \frac{(1-2j/k)^2}{4j(1-j/k)} = o(1).$$

The case $1 < p < 2$ is more involved:

$$\sum_{j \geqslant 2} \beta_j^p \leqslant \sum_{j=2}^{k/2}\left[\frac{1}{k}\psi'\left(\frac{j-1}{k}\right)\right]^p \leqslant k^{1-p}\left[\int_0^{1/2} \psi'(t)^p\, dt\right] = o(1).$$

Hence, in all three cases, $H^p(\beta) \geqslant \frac{p}{p-1}\log 2 + o(1)$. This inequality and Eq. (9) provide the announced violation of the additivity conjecture for Rényi entropies.

$$H^p_{\min}(\Phi_n \otimes \overline{\Phi}_n) \leqslant \frac{p}{p-1}\log 2 + o(1) < 2 \cdot \left[\frac{p}{p-1}\log 2 + o(1)\right] \leqslant 2H^p_{\min}(\Phi_n). \qquad \square$$

Let us now come back to the more general case of arbitrary $t \in (0, 1)$ fixed. It is natural to ask whether the bound $H^p(\beta^{(t)})$ is optimal. Even though this is an open question for fixed k, the corollary below implies that it is asymptotically optimal for large k. More precisely, let $\Phi_{k,n}$ be the random quantum channel Φ_n introduced in Section 3.2 (since k will vary in the statement below, we need to keep track of it). We can then state the following

Corollary 5.5. *For all $p > 1$, there exists a sequence n_k tending to infinity as k tends to infinity, such that, almost surely*

$$\lim_k H^p_{\min}(\Phi_{k,n_k} \otimes \overline{\Phi}_{k,n_k}) = \lim_k H^p_{\min}(\Phi_{k,n_k}) = \frac{p}{1-p}\log t.$$

In particular this means that we can almost surely estimate the Schatten $S_1 \to S_p$ norm of that quantum channel:

$$\lim_k \|\Phi_{k,n_k} \otimes \overline{\Phi}_{k,n_k}\|_{S_1 \to S_p} = \lim_k \|\Phi_{k,n_k}\|_{S_1 \to S_p} = t.$$

Proof. For $t = 1/2$, this follows directly by a diagonal argument from Theorem 5.4 and Eq. (9) together with the simple fact that the entropy increases when one takes tensor products:

$$H^p_{\min}(\Phi_{k,n_k}) \leqslant H^p_{\min}(\Phi_{k,n_k} \otimes \overline{\Phi}_{k,n_k}).$$

The asymptotic estimates of Theorem 5.4 are readily adapted to arbitrary $t \in (0, 1)$. As for the norm estimate, it follows from the definition of the Schatten norm and the Rényi entropy, as well as the fact that the $S_1 \to S_p$ norm is attained on density matrices. $\square$

It is remarkable that the norm estimate for $\|\Phi \otimes \overline{\Phi}\|_{S_1 \to S_p}$ given by $\|\Phi \otimes \overline{\Phi}(E_{tnk})\|_{S_p}$ is actually optimal. The above corollary stands as a mathematical evidence that the Bell states asymptotically maximize the $S_1 \to S_p$ norm of $\Phi \otimes \overline{\Phi}$.

The first example of 'Rényi subadditive' quantum channel was obtained by Holevo and Werner in [20] using a deterministic channel. However, their example violated the additivity conjecture only for $p > 4.79$. Hayden and Winter found a class of random counter examples for the whole range of parameters $p > 1$ in [12,14] and for $p = 1$ in [11]. Our being able to prescribe t in the counterexample of Theorem 5.4 is an improvement to the counterexamples provided in the paper [14] (even though there is evidence that the very recent techniques of [8,4] could be applied for $p > 1$ and *finite* t – yet perhaps not as big as $1/2$ or $1/3$).

Physically, this means that to obtain a counterexample, it is enough to couple randomly the input to a qubit ($k' = 1/t = 2$) to obtain a counterexample. The above reasoning applies actually for any t. In the following corollary we focus on the case $t = 1/k'$ for integer k', as it is more relevant physically.

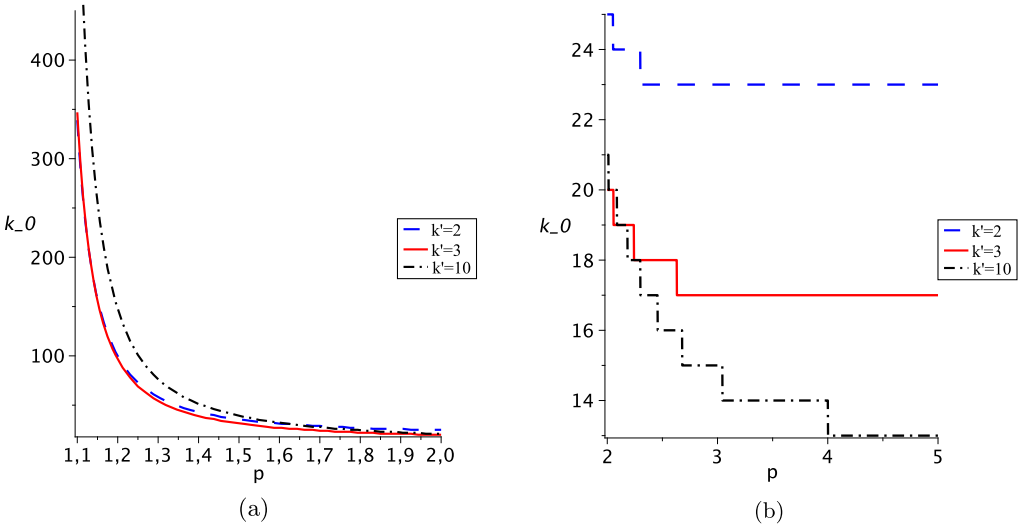

Fig. 2. Plots of k_0, i.e. acceptable values of k for which we get an asymptotic additivity violation, in function of p, for different values of $k' = 1/t$. Two ranges for p are plotted separately: $p \in [1.1, 2]$ in (a) and $p \in [2, 5]$ in (b).

Corollary 5.6. *For each $p > 1$ and each integer $k' \geqslant 2$, let $t = 1/k'$. There exists an integer k_0 such that for all $k \geqslant k_0$, one has almost surely*

$$\limsup_n H_{\min}^p(\Phi \otimes \overline{\Phi}) < 2 \liminf_n H_{\min}^p(\Phi).$$

Since the proof is very similar to the case $k' = 2$, instead of providing the details, we plot in Fig. 2 *acceptable* values for k_0 as functions of p, for several values of $k' = 1/t$:

$$k_0(t, p) = \min\left\{k \in \mathbb{N} \,\middle|\, \limsup_n H^p(\Phi \otimes \overline{\Phi})(E_{tnk}) < 2H^p\left(\beta^{(t)}\right)\right\}.$$

Note that k_0 as defined above may not be the *smallest* dimension yielding a violation of p-Rényi additivity. It may be that a better choice for the input state of the product channel could yield a smaller value for $H_{\min}^p(\Phi \otimes \overline{\Phi})$. As the plots suggest, the values of k_0 are not bounded when $p \to 1$. This fact is independent on the choice of the parameter $t = 1/k'$. The results of [4,8] suggest that there should be a k' large enough for which it is possible to keep k_0 bounded as $p \to 1$. This improvement is due to their better bounds on $H_{\min}^p(\Phi)$, obtained using the techniques developed by Hastings in [11].

We finish this section by a computation showing that the above bounds are not good enough to obtain the violation of the additivity conjecture in the case $p = 1$. We start with the entropy of the product channel:

$$H(\gamma) = \log k + \log 2 + o(1).$$

For the case of the single channel, we need an upper bound for $H(\beta^{(1/2)})$ (recall that $h(x) = -x \log x$):

$$H(\beta) = h(\beta_1) + \sum_{j=2}^{k/2} h(\beta_j) = \frac{\log 2}{2} + \sum_{j=2}^{k/2} h\left[\psi\left(\frac{j}{k}\right) - \psi\left(\frac{j-1}{k}\right)\right] + o(1)$$
$$\leqslant \frac{\log 2}{2} + \sum_{j=2}^{k/2} h\left[\frac{1}{k} \cdot \sup_{(j-1)/k \leqslant x \leqslant j/k} \psi'(x)\right] + o(1)$$
$$= \frac{\log 2}{2} + \sum_{j=2}^{k/2} h\left[\frac{1}{k}\psi'\left(\frac{j-1}{k}\right)\right] + o(1).$$

Using

$$h\left[\frac{1}{k}\psi'\left(\frac{j}{k}\right)\right] = \frac{\log k}{k}\psi'\left(\frac{j}{k}\right) + \frac{1}{k}\left[h \circ \psi'\right]\left(\frac{j}{k}\right)$$

and

$$\int_0^{1/2} \left[h \circ \psi'\right](t)\, dt = -\frac{\log 2}{2},$$

we obtain

$$H\left(\beta^{(1/2)}\right) \leqslant \frac{\log 2}{2} + \sum_{j=1}^{k/2-1} \frac{\log k}{k}\psi'\left(\frac{j}{k}\right) + \sum_{j=1}^{k/2-1} \frac{1}{k}\left[h \circ \psi'\right]\left(\frac{j}{k}\right) + o(1)$$
$$= \frac{\log 2}{2} + \frac{1}{2}\log k - \frac{\log 2}{2} + o(1) = \frac{\log k}{2} + o(1).$$

At the end, the entropy deficit is

$$H(\gamma) - 2H(\beta) \geqslant \log 2 + o(1) > 0,$$

which does not yield a violation of the minimum output von Neumann entropy.

5.4. The case $t = k^{-\alpha}$

We conclude this paper with the study of the case $t = k^{-\alpha}$, where $\alpha > 0$ is a fixed parameter. This corresponds to an exploration of a larger environment size $\mathbb{C}^{k^\alpha}$. To simplify the computations, we consider only the case of the minimum output von Neumann entropy. As before, we provide estimates, when k is fixed but large, for the minimum output entropies of $\Phi \otimes \overline{\Phi}$ and Φ.

We start with the simpler case of the product channel $\Phi \otimes \overline{\Phi}$. Theorem 5.2 provides the almost sure eigenvalues of $[\Phi \otimes \overline{\Phi}](E_{tmk})$:

$$\gamma = \gamma^{(k^{-\alpha})} = \left(\frac{1}{k^\alpha} + \frac{1}{k^2} - \frac{1}{k^{\alpha+2}}, \underbrace{\frac{1}{k^2} - \frac{1}{k^{\alpha+2}}, \ldots, \frac{1}{k^2} - \frac{1}{k^{\alpha+2}}}_{k^2-1 \text{ times}}\right).$$

Using the series expansion $h(1-x) = x - x^2/2 + o(x^2)$, one can compute the asymptotics for the minimum output entropy:

Proposition 5.7. *For the product channel $\Phi \otimes \overline{\Phi}$, the following upper bounds hold almost surely:*

$$H_{\min}(\Phi \otimes \overline{\Phi}) \leqslant H(\gamma) = \begin{cases} 2\log k - \frac{(2-\alpha)\log k}{k^\alpha} + o(\frac{\log k}{k^\alpha}) & \text{if } 0 < \alpha < 2; \\ 2\log k - \frac{2\log 2 - 1}{k^2} + o(\frac{1}{k^2}) & \text{if } \alpha = 2; \\ 2\log k - \frac{1}{2k^{2\alpha-2}} + o(\frac{1}{k^{2\alpha-2}}) & \text{if } \alpha > 2. \end{cases} \tag{10}$$

Our estimate for the single channel case is as follows:

Proposition 5.8. *For all $\alpha > 0$, the following lower bound holds true almost surely:*

$$H_{\min}(\Phi) \geqslant H_{\min}(\beta) = \log k - \frac{\log k}{k^\alpha} + o\left(\frac{\log k}{k^\alpha}\right).$$

For the purposes of this proof, we define $\varphi_k : [0, 1 - k^{-\alpha}] \to [0, 1]$, $\varphi_k(x) = \varphi(x, k^{-\alpha})$ and $h(x) = -x \log x$. We have

$$\frac{\partial}{\partial x}\varphi(x, y) = 1 - 2y + \sqrt{y(1-y)}\left[\frac{\sqrt{1-x}}{\sqrt{x}} - \frac{\sqrt{x}}{\sqrt{1-x}}\right] = (1-2y)\left(1 + \frac{g(x)}{g(y)}\right),$$

where the function $g : (0, 1) \to \mathbb{R}$ is defined by

$$g(x) = \frac{\sqrt{1-x}}{\sqrt{x}} - \frac{\sqrt{x}}{\sqrt{1-x}}.$$

Proof. According to Theorem 4.5, for all $\varepsilon > 0$,

$$H_{\min}(\Phi) \geqslant H(\beta) - \varepsilon,$$

where $\beta = \beta^{(k^{-\alpha})}$ is the k-dimensional vector defined by

$$\begin{aligned} &\beta_1 = \varphi_k\left(\frac{1}{k}\right); \\ &\beta_j = \varphi_k\left(\frac{j}{k}\right) - \varphi_k\left(\frac{j-1}{k}\right) = \frac{1}{k}\varphi_k'\left(\frac{\xi_j}{k}\right), \quad \forall 2 \leqslant j \leqslant J; \\ &\beta_j = 0, \quad \forall J < j \leqslant k. \end{aligned}$$

The index J is the number of non-trivial inequalities we get by using Theorem 4.1, and it is equal to $k-1$ if $\alpha \geqslant 1$ and to $\lfloor k - k^{1-\alpha} \rfloor$ if $\alpha < 1$.

Our purpose in what follows is to provide a "good" estimate for $H(\beta)$. We start by rescaling the eigenvalues: $H(\beta) = \log k + \frac{1}{k} H(k\beta)$. In this way, we can focus on the "entropy defect" $\log k - H(\beta)$ and reduce our problem to showing that

$$\frac{k^{\alpha-1}}{\log k} H(k\beta) = \frac{k^{\alpha-1}}{\log k} \sum_{j=1}^{J} h(k\beta_j) \xrightarrow[k\to\infty]{} -1. \tag{11}$$

The next step in our asymptotic computation is to replace the unknown points ξ_j by simpler estimates of the type j/k. Notice that the largest eigenvalue β_1 is of order k^{-1}. By the continuity of the function h, there exists a constant $C > 0$ such that $|h(k\beta_j)| \leqslant C$ and thus, individual terms in the sum (11) have no asymptotic contribution. Moreover, we can assume $J = k - 1$, ignoring at most $k^{1-\alpha}$ terms which have again no asymptotic contribution. It is clear that the function $x \mapsto \varphi_k'(x)$ is decreasing at fixed k and since the entropy function h is increasing for $x \in (0, e^{-1})$ and decreasing for $x \geqslant e^{-1}$, we can bound $h(\varphi_k'(\xi_j/k))$ by $h(\varphi_k'(j/k))$, and we reduce our problem to showing that

$$\frac{k^{\alpha-1}}{\log k} \sum_{j=1}^{k-1} h\left(\varphi_k'\left(\frac{j}{k}\right)\right) \xrightarrow[k\to\infty]{} -1, \tag{12}$$

or, equivalently,

$$\frac{k^{\alpha-1}}{\log k} \sum_{j=1}^{\lfloor k/2 \rfloor} h\left(\varphi_k'\left(\frac{j}{k}\right)\right) + h\left(\varphi_k'\left(1 - \frac{j}{k}\right)\right) \xrightarrow[k\to\infty]{} -1.$$

Now,

$$\begin{aligned} &h\left[(1-2y)\left(1 + \frac{g(x)}{g(y)}\right)\right] + h\left[(1-2y)\left(1 - \frac{g(x)}{g(y)}\right)\right] \\ &\quad = 2h(1-2y) + (1-2y)\left[h\left(1 + \frac{g(x)}{g(y)}\right) + h\left(1 - \frac{g(x)}{g(y)}\right)\right]. \end{aligned}$$

The term $2h(1-2y) = 2h(1-2k^{-\alpha}) \sim 4k^{-\alpha}$ has no asymptotic contribution and, using $h(1+t) + h(1-t) = -t^2 + O(t^4)$, we are left with computing the limit of the main contribution

$$\frac{k^{\alpha-1}}{\log k} \sum_{j=1}^{\lfloor k/2 \rfloor} -\frac{g(j/k)^2}{g(y)^2} \sim \frac{k^{\alpha-1}}{\log k} \sum_{j=1}^{\lfloor k/2 \rfloor} -g(j/k)^2 k^{-\alpha}.$$

Finally,

$$\frac{1}{k \log k} \sum_{j=1}^{\lfloor k/2 \rfloor} -g(j/k)^2 = \frac{-1}{\log k} \sum_{j=1}^{\lfloor k/2 \rfloor} \frac{1}{j} \frac{(1-2j/k)^2}{1-j/k} \sim -\frac{\log(k/2)}{\log k} \xrightarrow[k\to\infty]{} -1.$$

The error term

$$\frac{k^{\alpha-1}}{\log k}\sum_{j=1}^{\lfloor k/2\rfloor} -\frac{g(j/k)^4}{g(y)^4} \sim \frac{k^{\alpha-1}}{\log k}\sum_{j=1}^{\lfloor k/2\rfloor} -g(j/k)^4 k^{-2\alpha} \sim \frac{1}{k^{\alpha+1}\log k}\sum_{j=1}^{\lfloor k/2\rfloor}\frac{1}{j^2}\frac{(1-2j/k)^4}{(1-j/k)^2}$$

converges to zero. In conclusion, we have shown that Eq. (11) holds and we deduce that

$$H(\beta) = \log k - \frac{\log k}{k^\alpha} + o\left(\frac{\log k}{k^\alpha}\right). \qquad \square$$

The bounds obtained in this section do not yield a violation of the Holevo additivity conjecture. However, after the first version of this paper was released, Brandao and Horodecki [4] and Fukuda and King [8,9] used the same model as ours and adapted original ideas from Hastings [11] to prove that this model can also lead to a violation of the minimum output entropy additivity.

The techniques in [4,8] yield more information on the possibility of large values of the minimum output entropy for the model under discussion. However, our proofs are of free probabilistic nature and yield results of almost sure nature. In addition, [4,8] rely very much on the actual properties of Shannon's entropy function, whereas our techniques attack directly the question of the behavior of the eigenvalues.

We conjecture that the set $S(\beta^{(t)})$ (having the property that for any $\varepsilon > 0$, $S(\beta^{(t)}) + \varepsilon$ contains almost surely the eigenvalues of outputs of random quantum channels) can be made smaller and actually optimal, thus yielding as a byproduct that all the values $H_{\min}(\Phi)$ converge almost surely. However, the results of this paper show that the notion of majorization is not sufficient to achieve this goal.

Acknowledgments

Both authors would like to thank an anonymous referee for many useful comments on a preliminary version of our manuscript.

This paper was completed while one author (B.C.) was visiting the university of Tokyo and then the university of Wrocław and he thanks these two institutions for providing him with a very fruitful working environment. I.N. thanks Guillaume Aubrun for useful discussions.

B.C. was partly funded by ANR GranMa and ANR Galoisint. The research of both authors was supported in part by NSERC grants including grant RGPIN/341303-2007.

References

[1] G. Aubrun, I. Nechita, Catalytic majorization and l_p norms, Comm. Math. Phys. 278 (1) (2008) 133–144.

[2] G. Aubrun, I. Nechita, Stochastic domination for iterated convolutions and catalytic majorization, Ann. Inst. H. Poincaré Probab. Statist. 45 (3) (2009) 611–625.

[3] R. Bhatia, Matrix Analysis, Grad. Texts in Math., vol. 169, Springer-Verlag, New York, 1997.

[4] F. Brandao, M.S.L. Horodecki, On Hastings' counterexamples to the minimum output entropy additivity conjecture, Open Syst. Inf. Dyn. 17 (2010) 31–52.

[5] B. Collins, Product of random projections, Jacobi ensembles and universality problems arising from free probability, Probab. Theory Related Fields 133 (3) (2005) 315–344.

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

[7] B. Collins, P. Śniady, Integration with respect to the Haar measure on unitary, orthogonal and symplectic group, Comm. Math. Phys. 264 (3) (2006) 773–795.
[8] M. Fukuda, C. King, Entanglement of random subspaces via the Hastings bound, J. Math. Phys. 51 (2010) 042201.
[9] M. Fukuda, C. King, A. Moser, Comments on Hastings' additivity counterexamples, Comm. Math. Phys. 296 (1) (2010) 111–143.
[10] U. Haagerup, S. Thorbjørnsen, A new application of random matrices: $\mathrm{Ext}(C^*_{\mathrm{red}}(F_2))$ is not a group, Ann. of Math. (2) 162 (2) (2005) 711–775.
[11] M.B. Hastings, Superadditivity of communication capacity using entangled inputs, Nat. Phys. 5 (2009) 255–257.
[12] P. Hayden, The maximal p-norm multiplicativity conjecture is false, arXiv:0707.3291v1.
[13] P. Hayden, D. Leung, A. Winter, Aspects of generic entanglement, Comm. Math. Phys. 265 (2006) 95–117.
[14] P. Hayden, A. Winter, Counterexamples to the maximal p-norm multiplicativity conjecture for all $p > 1$, Comm. Math. Phys. 284 (1) (2008) 263–280.
[15] M. Ledoux, Differential operators and spectral distributions of invariant ensembles from the classical orthogonal polynomials part I: the continuous case, Electron. J. Probab. 9 (2004) 177–208.
[16] I. Nechita, Asymptotics of random density matrices, Ann. H. Poincaré 8 (8) (2007) 1521–1538.
[17] A. Soshnikov, Determinantal random point fields, Uspekhi Mat. Nauk 55 (5(335)) (2000) 107–160 (in Russian); translation in: Russian Math. Surveys 55 (5) (2000) 923–975.
[18] D.V. Voiculescu, A strengthened asymptotic freeness result for random matrices with applications to free entropy, Int. Math. Res. Not. IMRN 1 (1998) 41–63.
[19] D.V. Voiculescu, K.J. Dykema, A. Nica, Free Random Variables, Amer. Math. Soc., 1992.
[20] R. Werner, A. Holevo, Counterexample to an additivity conjecture for output purity of quantum channels, J. Math. Phys. 43 (2002) 4353–4357.