J. OPERATOR THEORY
73:2(2015), 333–360
doi: 10.7900/jot.2013dec04.2008

ON THE CONVERGENCE OF OUTPUT SETS OF QUANTUM CHANNELS

BENOÎT COLLINS, MOTOHISA FUKUDA and ION NECHITA

Communicated by Kenneth R. Davidson

ABSTRACT. We study the asymptotic behavior of the output states of sequences of quantum channels. Under a natural assumption, we show that the output set converges to a compact convex set, clarifying and substantially generalizing results in S.T. Belinschi, B. Collins, I. Nechita, *Invent. Math.* **190**(2012), 647–697. Random mixed unitary channels satisfy the assumption; we give a formula for the asymptotic maximum output infinity norm and we show that the minimum output entropy and the Holevo capacity have a simple relation for the complementary channels. We also give non-trivial examples of sequences Φ_n such that along with any other quantum channel Ξ, we have convergence of the output set of Φ_n and $\Phi_n \otimes \Xi$ simultaneously; the case when Ξ is entanglement breaking is investigated in details.

KEYWORDS: *Random matrices, quantum information theory, random quantum channel.*

MSC (2010): Primary 15A52; Secondary 94A17, 94A40.

INTRODUCTION

Quantum channels are of central importance in quantum information theory, and in the meantime, many mathematical quantities that are associated to quantum channels are still not very well understood. This is the case, for example, for the maximum output infinity norm, the minimum output entropy and Holevo capacity. These quantities as well as other important quantities turn out to actually depend only on the image of the collection of all possible output states. Incidentally, the output set — a compact convex subset in the set of all states — turns out to be an interesting geometric object that has nice interpretations in the theory of entanglement, statistics, free probability and others. However, identifying the output set for a given channel turns out to be a difficult task. So, instead, we analyze sequences of quantum channels which have nice asymptotic properties. Recently, research on random quantum channels in terms of eigenvalues has

led to important advances in the understanding of quantum channels, in particular in relation to the problem of additivity of the minimum output entropy, see for example [3], [4] and [13].

In this paper we elaborate an axiomatic and systematic study of properties of sequences of quantum channels that ensure the convergence of the output set towards a limit. It turns out that the sufficient conditions that we unveil are not of random nature, although most of all examples available so far rely on random constructions. The main results of this paper are Theorems 2.8 and 2.9. The idea underlying these theorems was already available in [3], but we considerably simplify and conceptualize the argument, and we remove all probabilistic considerations from our main argument. We start with examples of deterministic quantum channels (or projections) which fit our axiomatic framework. Then, we treat random quantum channels and random mixed unitary channels as examples of this axiomatic approach. To do so, we rely on recent results of [9] and [19] where the strong asymptotic freeness of Haar unitary matrices and constant matrices or extension of strong convergence to polynomials with matrix coefficients is proved.

Our paper is organised as follows. Section 1 contains definitions and reminders about quantum channels and quantities associated to them. Then, our main result is stated and proved in Section 2. Then, in Section 3 our main result is applied to the convergence of entropies. Section 4 introduces some results from random matrix theory and free probability. A subclass of entanglement-breaking channels is investigated in Section 5, then we discuss in Section 6 examples of our main result: random Stinespring channels (Subsection 6.1) and random mixed unitary channels (Subsection 6.2). Finally, in Section 7, we discuss tensor products of channels, and especially entanglement-breaking channels are discussed in details.

1. QUANTUM CHANNELS AND THEIR IMAGE

1.1. NOTATION. Following the quantum information theoretic notation, we call *quantum states* semidefinite positive matrices of unit trace

$$D_k = \{A \in M_k : A \geqslant 0 \text{ and } \mathrm{Tr}A = 1\},$$

where we write $M_k = M_k(\mathbb{C})$. A *quantum channel* is a completely positive and trace preserving linear map $\Phi : M_N \to M_k$. Following Stinespring's picture [33], we view any quantum channel Φ as an isometric embedding of $\mathbb{C}^N$ into $\mathbb{C}^k \otimes \mathbb{C}^n$, to which we apply the partial trace:

$$V : \mathbb{C}^N \to \mathbb{C}^k \otimes \mathbb{C}^n.$$

That is,

$$\Phi = (\mathrm{id}_k \otimes \mathrm{Tr}_n) \circ E \tag{1.1}$$

where $E(\cdot) = V \cdot V^*$ is a non-unital embedding of M_N in $M_k \otimes M_n$. We define the complementary channel of Φ by

$$\widetilde{\Phi} = (\mathrm{Tr}_k \otimes \mathrm{id}_n) \circ E.$$

Note that for a pure input, its outputs via Φ and $\widetilde{\Phi}$ share the same non-zero eigenvalues, but it is not the case in general for a mixed input. We also define the adjoint channel Φ^*, which is the adjoint of Φ with respect to the Hilbert–Schmidt scalar product in M_k:

$$\mathrm{Tr}[\Phi(X)^* Y] = \mathrm{Tr}[X^* \Phi^*(Y)]. \tag{1.2}$$

If Φ is defined via a Stinespring dilation using an isometry V as in (1.1), then

$$\Phi^*(Y) = V^*(Y \otimes I_n)V.$$

In quantum information language, $\mathbb{C}^n$ is called the environment. In this paper, we are interested in a sequence of quantum channels, that we will index by the environment n, $\Phi_n : M_N \to M_k$. From now on, our setting is as follows: $k, n, N \in \mathbb{N}$ are such that k is fixed and $N \in \mathbb{N}$ is any function of $n \in \mathbb{N}$. Importantly, a quantum channel is defined, up to a unitary conjugation on the input, by $P_n = V_n V_n^*$, which is the unit of M_N embedded in $M_k \otimes M_n$.

Let D_N be the collection of states in M_N, and $D_N^{\mathrm{p}} \subset D_N$ be the collection of extremal (pure) states, i.e. rank-one (self-adjoint) projections. We are interested in $L_n = \Phi_n(D_N^{\mathrm{p}})$, which is the image of all the pure states under the quantum channel Φ_n. One can see that L_n is a compact subset of D_k, but not always convex, although so is $K_n = \Phi_n(D_N)$.

The task to classify all the possible sets K_n and L_n arising from this construction seems to be out of reach. Instead, we focus our attention on possible asymptotic behaviours of K_n and L_n as $n \to \infty$.

In Section 2, we identify some assumption with which K_n and L_n converge to some well-described compact convex set as $n \to \infty$. Then, we present examples of sequences of random projections $\{P_n\}_{n\in\mathbb{N}}$ which satisfy this assumption with probability one.

1.2. ENTROPIES AND CAPACITIES. We introduce three quantities associated with quantum channels.

Firstly, the maximum output infinity norm of channel Φ is defined as

$$\|\Phi\|_{1,\infty} = \max_{\rho \in D_N} \|\Phi(\rho)\|_\infty$$

where 1 and ∞ represent norms used for the input and output spaces respectively.

Secondly, the minimal output entropy (MOE) of channel Φ is defined as

$$S^{\min}(\Phi) = \min_{\rho \in D_N} S(\Phi(\rho)).$$

Here, $S(\cdot)$ is the von Neumann entropy.

Thirdly, the Holevo capacity (HC) of channel Φ is defined as

$$\chi(\Phi) = \max_{\{p_i,\rho_i\}} S(\Phi(\widehat{\rho})) - \sum_i p_i S(\Phi(\rho_i)).$$

Here, $\{p_i\}_i$ is a probability distribution, $\{\rho_i\}_i \subset D_N$ and $\widehat{\rho} = \sum_i p_i\rho_i$. Note that

$$S^{\min}(\widetilde{\Phi}) = S^{\min}(\Phi)$$

but

$$\chi(\widetilde{\Phi}) \neq \chi(\Phi)$$

in general.

It is a rather direct observation that $S_{\min}$ and χ depend only on the output of the channel. Therefore, for any convex set $M \subset D_k$, it is natural to define

$$S^{\min}(M) = \min_{X\in M} S(X), \tag{1.3}$$

$$\chi(M) = \max_{\{p_i,X_i\}} S\Big(\sum_i p_i X_i\Big) - \sum_i p_i S(X_i), \tag{1.4}$$

where $X_i \in M$.

For those quantities one can think of additivity questions:

$$\chi(\Phi\otimes\Omega) \stackrel{?}{=} \chi(\Phi) + \chi(\Omega),$$
$$S^{\min}(\Phi\otimes\Omega) \stackrel{?}{=} S^{\min}(\Phi) + S^{\min}(\Omega),$$

for two quantum channels. These equalities are not true in general; additivity of MOE was disproved by Hastings [20] and this non-additivity can be translated to be the one for HC [30]. In terms of information theory, the additivity of HC is important. Suppose in particular that

$$\chi(\Phi^{\otimes r}) = r\chi(\Phi)$$

for some channel Φ, then the classical capacity over this channel Φ has a one-shot formula (in the sense that the regularization is not necessary):

$$\lim_{r\to\infty} \frac{1}{r}\chi(\Phi^{\otimes r}) = \chi(\Phi).$$

Additivity of MOE itself is also interesting because it measures purity of channels, but caught more attention when Shor proved the equivalence between the two additivity questions [30]. Moreover, a breakthrough was made by disproving additivity of MOE [20] with a use of random matrix theory as MOE concerns eigenvalues of matrices whereas HC depends on the geometry of output states. By contrast, our paper sheds light on not only MOE but also HC because we consider geometry of output states (at least, of single channels).

2. MAIN RESULT — LINEAR ALGEBRA AND CONVEX ANALYSIS

2.1. PRELIMINARY. For a sequence of sets $\{S_n\}_{n\in\mathbb{N}}$ we use the following standard notations of lim-inf and lim-sup:

$$\varliminf_{n\to\infty} S_n = \bigcup_{N\in\mathbb{N}} \bigcap_{n\geqslant N} S_n \quad \text{and} \quad \varlimsup_{n\to\infty} S_n = \bigcap_{N\in\mathbb{N}} \bigcup_{n\geqslant N} S_n.$$

If S is a subset of a topological space, we denote the *interior* of S by S° and the *closure* S^{cl}. Suppose we have a *convex* set K in a real vector space; any line segment joining two points of K is included in K. Then, we have the following two definitions:

(i) A point $x \in K$ is an *extreme point* if x does not lie in any open line segment joining two points of K.

(ii) A point $x \in K$ is an *exposed point* if there exists a supporting hyperplane which intersects with K only at one point.

We also use the following notation:

$$\mathrm{hull}(\{x_i\}_{i=1}^m) = \Big\{ \sum_{i=1}^m \lambda_i x_i : \lambda_i \geqslant 0 \text{ and } \sum_{i=1}^m \lambda_i = 1 \Big\}$$

which is called the *convex hull* of $\{x_i\}_{i=1}^m$.

THEOREM 2.1 ([32]). *Suppose we have a convex and compact set $K \subset \mathbb{R}^d$. Then, for any interior point of K we can choose $2d$ (or less) extreme points of K whose convex hull includes the point within the interior.*

THEOREM 2.2 ([34]). *For any closed convex set K, the set of exposed points is dense in the set of extreme points.*

We define conditions which ensure the limiting convex set of output states.

DEFINITION 2.3. A sequence of projections $P_n \subset M_{kn}$ is said to satisfy the condition $\mathcal{C}_m$ if for all $A \in D_k$, the following m infinite sequences in $n \in \mathbb{N}$:

$$\lambda_1(P_n(A \otimes I_n)P_n), \ldots, \lambda_m(P_n(A \otimes I_n)P_n) \tag{2.1}$$

converge to a common limit, which we denote $f(A)$. Here, $\lambda_i(\cdot)$ is the i-th largest eigenvalue. Note that $\mathcal{C}_m \implies \mathcal{C}_l$ for $l < m$.

Let (Φ_n) be a sequence of quantum channels associated with the sequence of projections (P_n), that is Φ_n is defined by (1.1) and $P_n = V_n V_n^*$. One can then define the function f in terms of the adjoint channels Φ_n^* defined in (1.2):

PROPOSITION 2.4. *Let $\Phi : M_N \to M_k$ be a quantum channel defined by an isometry $V : \mathbb{C}^N \to \mathbb{C}^k \otimes \mathbb{C}^n$, and put $P = VV^*$. Then, for any $A \in D_k$, the non-zero eigenvalues of the matrices $P(A \otimes I_n)P$ and $\Phi^*(A)$ are identical. In particular, a sequence of projections (P_n) satisfies condition $\mathcal{C}_m$ if and only if, for all $A \in D_k$, the m largest eigenvalues of $\Phi_n^*(A)$ converge to $f(A)$.*

Proof. We have

$$P(A \otimes I_n)P = VV^*(A \otimes I_n)VV^* = V\Phi^*(A)V^*.$$

The conclusion follows from the fact that V is an isometry, hence the matrices $V\Phi^*(A)V^*$ and $\Phi^*(A)$ have the same non-zero eigenvalues. ∎

Les us start by recording an obvious upper bound:

LEMMA 2.5. *If a sequence of projections P_n satisfies the condition $\mathcal{C}_1$, then for any sequence $x_n \in \mathbb{C}^k \otimes \mathbb{C}^n$ such that $x_n x_n^* \leqslant P_n$ we have*

$$\overline{\lim_{n\to\infty}} \operatorname{Tr}[X_n A] \leqslant f(A) \quad \forall A \in D_k$$

where $X_n = \operatorname{Tr}_{\mathbb{C}^n}[x_n x_n^*]$.

Proof. For all $n \in \mathbb{N}$,

$$\operatorname{Tr}[X_n A] = \operatorname{Tr}[x_n x_n^*(A\otimes I_n)] \leqslant \max_{vv^* \leqslant P_n} \langle v, (A\otimes I_n)v\rangle = \|P_n(A\otimes I_n)P_n\|_\infty \to f(A). \quad ∎$$

Next, we prove existence of sequence of optimal vectors:

LEMMA 2.6. *If a sequence of projections $P_n \subset M_{kn}$ satisfies the condition $\mathcal{C}_m$, then for any $A \in D_k$ there exists a sequence of m-dimensional subspaces $W_n \subset \operatorname{range} P_n$ for large enough $n \in \mathbb{N}$ such that any sequence of unit vectors $x_n \in W_n$ satisfies*

$$\operatorname{Tr}[X_n A] \to f(A) \quad \text{as } n \to \infty. \tag{2.2}$$

Here, $X_n = \operatorname{Tr}_{\mathbb{C}^n}[x_n x_n^*]$.

Proof. Let v_i be eigenvectors of λ_i in (2.1), and define $W_n = \operatorname{span}\{v_i : 1 \leqslant i \leqslant m\}$. Then, for any unit vector $x_n \in W_n$ we have

$$\lambda_m \leqslant x_n^*(A \otimes I_n)x^* = \operatorname{Tr}[X_n A] \leqslant \lambda_1$$

where the both bounds converges to $f(A)$. ∎

If $\mathcal{C}_m$ holds with large enough m with respect to k, we can have a sequence x_i in Lemma 2.6 with orthogonal property in $\mathbb{C}^n$, which is useful in proving Theorem 2.9:

LEMMA 2.7. *Given two subspaces $W \subset \mathbb{C}^k \otimes \mathbb{C}^n$ and $T \subset \mathbb{C}^n$ such that $\dim W > k \dim T$, there exists $x \in W$ having the following Schmidt decomposition:*

$$x = \sum_{i=1}^{r} \sqrt{\lambda_i} e_i \otimes f_i.$$

Here, $\{e_i\}$ and $\{f_i\}$ are orthonormal in $\mathbb{C}^k$ and $\mathbb{C}^n$ respectively and moreover $f_i \perp T$, for all $i = 1, \ldots, r$.

Proof. Define $\widetilde{T} = \mathbb{C}^k \otimes T$. Since $\dim \widetilde{T} = k \dim T$, there exists a unit vector $x \in W$ such that $x \perp \widetilde{T}$. Consider the Schmidt decomposition of x, as in the statement, with $\lambda_i > 0$, for all i. For any $f \in T$, we have

$$\langle f, f_i \rangle = \lambda_i^{-1/2} \langle e_i \otimes f, x \rangle = 0,$$

since $x \perp e_i \otimes f \in \widetilde{T}$. ∎

If the condition $\mathcal{C}_1$ is satisfied we define a compact convex set:

$$K = \{B \in D_k : \mathrm{Tr}[BA] \leqslant f(A) \ \forall A \in D_k\}. \tag{2.3}$$

In the following sections, we prove that both the images of mixed input states and pure input states converge to this convex set K. Especially, the latter statement is interesting because the set of pure input states itself is not a convex set.

2.2. LIMITING IMAGE FOR MIXED INPUT STATES. Our first result is as follows:

THEOREM 2.8. *If a sequence of projections $P_n \subset M_{kn}$ satisfies the condition $\mathcal{C}_1$, then*

$$K^\circ \subseteq \varinjlim_{n\to\infty} K_n \subseteq \overline{\varinjlim_{n\to\infty}} K_n \subseteq K.$$

Here, as before, K_n is the image of all the mixed states by the quantum channels defined by P_n.

Proof. Firstly, we show that $\overline{\varinjlim_{n\to\infty}} K_n \subseteq K$ by showing $\overline{\varinjlim_{n\to\infty}} L_n \subseteq K$ because $K_n = \mathrm{hull}(L_n)$. For any $X \in \overline{\varinjlim_{n\to\infty}} L_n$, there exists a subsequence $\{n_j\}_j$ such that $X \in L_{n_j}$. Since X is an output of the channel P_{n_j}, there exists a unit vector x_{n_j} which lives in the support of P_{n_j} such that $\mathrm{Tr}_{\mathbb{C}^{n_j}}[x_{n_j} x_{n_j}^*] = X$. By Lemma 2.5, we have $\mathrm{Tr}[XA] \leqslant f(A)$, and the first inclusion follows.

Secondly, we prove that $K^\circ \subseteq \varinjlim_{n\to\infty} K_n$. Take $X \in K^\circ$. Since K is a compact and convex set embedded into $\mathbb{R}^{k^2-1}$, writing $r = 2k^2 - 2$, by Theorem 2.1, there exist r extreme points of K, say, $(E_1, \ldots, E_r)$ such that

$$X \in (\mathrm{hull}\{E_1, \ldots, E_r\})^\circ.$$

Also, by Theorem 2.2, there exists r-tuple of exposed points of K, say, $(F_1, \ldots, F_r)$ such that

$$X \in (\mathrm{hull}\{F_1, \ldots, F_r\})^\circ. \tag{2.4}$$

Note that since each F_i is an exposed point of K, there exists A_i such that

$$\begin{aligned} &\mathrm{Tr}[F_i A_i] = f(A_i), \\ &\mathrm{Tr}[Y A_i] < f(A_i) \quad \forall Y \in K \setminus \{F_i\}. \end{aligned} \tag{2.5}$$

On the other hand, by Lemma 2.6, for each A_i, there exists a sequence $X_i^{(n)} \in K_n$ such that

$$\mathrm{Tr}[X_i^{(n)} A_i] \to f(A_i) \quad \text{as } n \to \infty. \tag{2.6}$$

We claim that $X_i^{(n)} \to F_i$ as $n \to \infty$. Take a converging subsequence $X_i^{(n_j)} \to G$. Then, the first statement: $\overline{\lim_{n\to\infty}} K_n \subseteq K$ implies $G \in K$ because K is closed. Moreover, (2.6) implies that $\mathrm{Tr}[GA_i] = f(A_i)$. Hence, the equation (2.5) implies the above claim.

Therefore, for large enough n, we have

$$X \in \mathrm{hull}\{X_i^{(n)} : 1 \leqslant i \leqslant r\} \subset K_n.$$

Here, the first inclusion follows from (2.4) and the second holds because K_n is convex. Therefore, $K^\circ \subseteq \underline{\lim_{n\to\infty}} K_n$. ∎

2.3. LIMITING IMAGE FOR PURE INPUT STATES. The second theorem we prove is about the image L_n of the set of pure states.

THEOREM 2.9. *If a sequence of projections $P_n \subset M_{kn}$ satisfies the condition $\mathcal{C}_m$, with $m = (2k^2-3)k^2+1$, then*

$$K^\circ \subseteq \underline{\lim_{n\to\infty}} L_n \subseteq \overline{\lim_{n\to\infty}} L_n \subseteq K.$$

Here, as before, L_n is the image of all the pure states by the quantum channels defined by P_n.

Proof. Since condition $\mathcal{C}_m$ is stronger than $\mathcal{C}_1$, the last inclusion follows from the proof of Theorem 2.8 and from the fact that $L_n \subseteq K_n$. We shall now show that the first inclusion holds.

Comparing this statement to the one in the proof of Theorem 2.8, we see that the difficulty comes from the fact that L_n is not always a convex set. As before, for a fixed $X \in K^\circ$, choose a set of r exposed points $F_1, \ldots, F_r$ of K, with $r \leqslant 2k^2-2$ such that $X \in (\mathrm{hull}\{F_1, \ldots, F_r\})^\circ$. The main idea here is to build approximating sequences $L_n \ni X_i^{(n)} \to F_i$ with an additional orthogonality property with respect to $\mathbb{C}^n$. More precisely, we want sequences $x_i^{(n)} \in \mathbb{C}^k \otimes \mathbb{C}^n$, such that

$$X_i^{(n)} = \mathrm{Tr}_{\mathbb{C}^n}[x_i^{(n)} x_i^{(n)*}] \to F_i,$$

and their Schmidt decompositions:

$$x_i^{(n)} = \sum_a \sqrt{\lambda_a^{(i,n)}}\, e_a^{(i,n)} \otimes f_a^{(i,n)} \tag{2.7}$$

have the additional property that the families $\{f_a^{(i,n)}\}_{a,i}$ are all orthogonal to each other for large enough n.

Taking advantage of this orthogonality condition, we claim that, for n large enough,

$$\mathrm{hull}\{X_i^{(n)} : 1 \leqslant i \leqslant r\} \subset L_n.$$

Indeed, for $X = \sum_{i=1}^{r} t_i X_i^{(n)}$ in the hull, the orthogonality condition implies that the unit vector

$$x = \sum_{i=1}^{r} \sqrt{t_i} x_i^{(n)} \in \text{range} P_n$$

turns out to give $\text{Tr}_{\mathbb{C}^n}[xx^*] = X$. Then, as before, for large enough n,

$$X \in \text{hull}\{X_i^{(n)} : 1 \leqslant i \leqslant r\} \subset L_n$$

proving $L^\circ \subseteq \varliminf_{m\to\infty} L_m$.

To finish the proof, we shall construct the approximating sequences (2.7), inductively, for $i = 1, 2, \ldots, r$. The first step is identical to the one in the proof of Theorem 2.8: since P_n satisfies $\mathcal{C}_1$; choose a sequence $x_1^{(n)}$ such that $X_1^{(n)} = \text{Tr}_{\mathbb{C}^n}[x_1^{(n)} x_1^{(n)*}]$ satisfies $X_1^{(n)} \to F_1$. Suppose now we have constructed the first s approximating sequences $x_i^{(n)}$, $1 \leqslant i \leqslant s$. For each n, Lemma 2.6, provides us with an m-dimensional subspace $W_n \subset \mathbb{C}^k \otimes \mathbb{C}^n$ of vectors verifying equation (2.2) for $A = A_{s+1}$. As before, one can show that for all sequences $x_n \in W_n$, the reduced states $X_n = \text{Tr}_{\mathbb{C}^n}[x_n x_n^*]$ converge to F_{s+1}. Define now $T_s^{(n)} = \operatorname{span}_{a,i}\{f_a^{(i,n)}\}$ with $1 \leqslant i \leqslant s$, the span of all vectors $f \in \mathbb{C}^n$ appearing in the Schmidt decompositions of the vectors $x_1^{(n)}, \ldots, x_s^{(n)}$ (see equation (2.7)). Since $\dim T_s^{(n)} \leqslant sk$ and $m = (2k^2 - 2 - 1)k^2 + 1 > sk^2$, by Lemma 2.7, one can find a sequence of vectors $x_{s+1}^{(n)} \in W_n$ such that the vectors f appearing in the Schmidt decompositions are orthogonal to $T_s^{(n)}$.

To summarize, we have constructed a sequence $x_{s+1}^{(n)}$ with the following two properties:

- the reduced states $X_{s+1}^{(n)}$ converge to F_{s+1};
- the vectors f_i appearing in the SVD of $x_{s+1}^{(n)}$ are all orthogonal to $T_s^{(n)}$.

In such a way, one constructs recursively a family of approximating vectors with the required orthogonality condition. ∎

3. ASYMPTOTIC BEHAVIOUR OF SOME ENTROPIC QUANTITIES

3.1. THE $S_1 \to S_\infty$ NORM.

Our first result is as follows

PROPOSITION 3.1. *Let $(\Phi_n)_n$ be a sequence of channels satisfying condition $\mathcal{C}_1$. Then,*

$$\|\Phi_n\|_{1,\infty} \to \max_{a \in \mathbb{C}^k,\, \|a\|_2 = 1} f(aa^*) \quad \text{as } n \to \infty.$$

Proof. It is easy to see from the definition in (1.2) that

$$\|\Phi_n\|_{1,\infty} = \|\Phi_n^*\|_{1,\infty}.$$

The fact that $f(\cdot)$ is convex and Proposition 2.4 complete the proof. ∎

3.2. THE MINIMUM OUTPUT ENTROPY AND THE HOLEVO QUANTITY. The following proposition is a rather direct observation.

PROPOSITION 3.2. *Let $(P_n)_n$ be a sequence of orthogonal projections satisfying condition $\mathcal{C}_1$. Then, one has*

$$\lim_{n\to\infty} S_p^{\min}(\Phi_n) = S_p^{\min}(K), \quad \lim_{n\to\infty} \chi(\Phi_n) = \chi(K).$$

Proof. By theorem 2.8 and continuity of the von Neumann entropy, the first statement is proved. For the second, note that Carathéodory's theorem implies that optimal ensemble can always consist of $2k^2+1$. Therefore, the first statement also implies the second. ∎

The following proposition gives a necessarily and sufficient condition for the Holevo capacity to be written nicely:

PROPOSITION 3.3. *We have*

$$\chi(K) = \log k - S^{\min}(K), \tag{3.1}$$

if and only if the limiting convex set K has the property that

$$\frac{I}{k} \in \text{hull}\,(\text{argmin}\, S),$$

where $X \in \text{argmin}\, S$ if and only if $S(X) = S^{\min}(K)$,

Proof. The first and second terms in (1.4) have the upper bounds respectively $\log k$ and $-S^{\min}(K)$ for the convex set K. So, our assumption lets Φ achieve both the bounds. The converse is obvious from this argument. ∎

Suppose f is G-invariant for some group G and its unitary representation $\{U_g\}_{g\in G}$, i.e., $f(U_g A U_g^*) = f(A)$ for all $g \in G$ and $A \in D_k$. Then K is invariant with respect to those rotations: $U_g K U_g^* = K$ for all $g \in G$. This, in particular, implies that the set of optimal points argmin S is also invariant. In addition, if the unitary representation $\{U_g\}_{g\in G}$ is irreducible so that $\int U_g A U_g^* = I/k$ for all $A \in D_k$ [22], then we get the formula (3.1). For example, consider the additive group $\mathbb{Z}_k \times \mathbb{Z}_k$ and define unitary operators, which are called *discrete Weyl operators*, by

$$W_{a,b} = X^a Y^b. \tag{3.2}$$

Here, $(a,b) \in \mathbb{Z}_k \times \mathbb{Z}_k$, and X and Y act on the canonical basis vectors $\{e_l\}_{l=1}^k$ of $\mathbb{C}^k$ as follows:

$$Xe_l = e_{l+1} \quad \text{and} \quad Ye_l = \exp\{\tfrac{2\pi \mathrm{i}}{k} \cdot l\} e_l.$$

This is an irreducible unitary adjoint representation of the group $\mathbb{Z}_k \times \mathbb{Z}_k$ on $\mathbb{C}^k$. Although this argument only gives a sufficient condition, it turns out to be useful. We have proven the following corollary.

COROLLARY 3.4. *Suppose, as is in* (2.3)*, a convex set K is defined by a function f which is invariant with respect to the discrete Weyl operators:*

$$f(W_{a,b}\, A\, W_{a,b}^*) = f(A) \quad \forall A \in D_k,\ \forall (a,b) \in \mathbb{Z}_k \times \mathbb{Z}_k. \tag{3.3}$$

Then, the formula (3.1) *holds.*

4. FREE PROBABILITY

A $*$-*non-commutative probability space* is a unital $*$-algebra $\mathcal{A}$ endowed with a linear map $\varphi\colon \mathcal{A} \to \mathbb{C}$ satisfying $\varphi(ab) = \varphi(ba)$, $\varphi(aa^*) \geqslant 0$, $\varphi(1) = 1$. The map φ is called a trace, and an element of $\mathcal{A}$ is called a *non-commutative random variable*.

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 self-adjoint 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 self-adjoint indeterminates $X_1, \ldots, X_k$. The *joint distribution* of the k-tuple $(a_i)_{i=1}^k$ is the linear form

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

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.

Considering a sequence of k-tuples $(a_i^{(n)})_{i=1}^k$ in $*$-non-commutative probability spaces $(\mathcal{A}_n, \varphi_n)$, we say that it converges *in distribution* to the distribution of $(a_1, \ldots, a_k) \in (\mathcal{A}, \varphi)$ if and only if $\mu_{(a_1^{(n)},\ldots,a_k^{(n)})}$ converges point wise to $\mu_{(a_1,\ldots,a_k)}$. Likewise, a sequence is said to converge *strongly in distribution* if and only if it converges in distribution, and in addition, for any non-commutative polynomial P, its operator norm converges

$$\|P(a_1^{(n)}, \ldots, a_k^{(n)})\| \to \|P(a_1, \ldots, a_k)\|.$$

In this definition, we assume that the operator norm is given by the distribution, i.e.

$$\|P(a_1^{(n)}, \ldots, a_k^{(n)})\| = \lim_{p\to\infty} \|P(a_1^{(n)}, \ldots, a_k^{(n)})\|_p,$$

and

$$\|P(a_1,\ldots,a_k)\| = \lim_{p\to\infty} \|P(a_1,\ldots,a_k)\|_p. \tag{4.1}$$

For the purpose of this paper, let us record two important theorems which extend strong convergence. I.e., let $(a_i^{(n)})_{i=1}^k$ be a sequence of $n\times n$ matrices, viewed as elements of the non-commutative probability space $(M_n, n^{-1}\mathrm{Tr})$ and assume that it converges strongly in distribution towards a k-tuple of random variables $(a_1,\ldots,a_k)\in(\mathcal{A},\varphi)$, then we have the following extension theorems.

THEOREM 4.1. *Let U_n be an $n\times n$ Haar distributed unitary random matrix. Then the family*

$$(a_1^{(n)},\ldots,a_k^{(n)},U_n,U_n^*)$$

almost surely converges strongly too, towards the $k+2$-tuple of random variables $(a_1,\ldots,a_k,u,u^*)$, where u,u^* are unitary elements free from $(a_1,\ldots,a_k)$.

Historically, the convergence of distribution is due to Voiculescu, [35]. A simpler proof was given by [8]. The strong convergence relies on [9] — it relies heavily on preliminary works by [19] and [26].

Actually, although this is counterintuitive, Theorem 4.1 is equivalent to a stronger statement where, in the conclusion, the non-commutative polynomial is not taken with complex coefficients, but with any matrix coefficient of fixed size. This follows from the "linearization lemma" as proved by Haagerup and Thorbjørnsen [19]. We state this result below, as it will be useful to widen our range of examples.

THEOREM 4.2. *Let P be a non-commutative polynomial in k variables with coefficients in $M_l(\mathbb{C})$ instead of $\mathbb{C}$. Then the operator norm of $P((a_i^{(n)}))_{i=1}^k\in M_l\otimes M_n$ still converges as $n\to\infty$. The limit is obtained by taking the limit as $p\to\infty$ of the limit as $n\to\infty$ of the p-norms.*

5. EXAMPLE OF NON-RANDOM PROJECTIONS

In this section we consider some elementary examples of deterministic sequence of projections which satisfy the condition C_m.

Let's start with the completely depolarizing channel $\Phi_n : M_N\to M_k$:

$$\Phi(\rho) = \mathrm{Tr}[\rho]\cdot\frac{I_k}{k}. \tag{5.1}$$

Its adjoint channel is written as

$$\Phi^*(\sigma) = \mathrm{Tr}[\sigma]\cdot\frac{I_N}{k}.$$

This immediately implies via Proposition 2.4 the following result:

PROPOSITION 5.1. *The sequence of depolarizing channels defined in* (5.1) *satisfies the condition $\mathcal{C}_m$ for all $m \geqslant 1$.*

Note that the above example is trivial, since the image of the channels Φ_n consists of a single point, $\{\frac{I_k}{k}\}$, so the convergence result is obvious.

We now generalize the above example by considering a subclass of entanglement-breaking channels. In general, any entanglement-breaking channel has the Holevo form [23]:

$$\Xi(X) = \sum_{i=1}^{l} \mathrm{Tr}[XM_i]\sigma_i \tag{5.2}$$

where $\{M_i\}_i$ are positive operators which sum up to the identity and σ_i are fixed states. Note that the set of the operators $\{M_i\}_i$ is called *positive operator valued measure* in quantum information theory and $p_i = \mathrm{Tr}[XM_i]$ constitute a probability distribution. As the name suggests, those channels break entanglement through measurements.

PROPOSITION 5.2. *Let Φ_n be a sequence of entanglement-breaking channels:*

$$\Phi_n(\rho) = \sum_{i=1}^{l} \mathrm{Tr}[M_i^{(n)}\rho]\sigma_i$$

where $l > 0$ and $(\sigma_i)_{i=1}^{l} \in D_k^l$ do not depend on n, and, for all $n \in \mathbb{N}$, $(M_i^{(n)})_{i=1}^{l}$ is a POVM *such that $\|M_i^{(n)}\| = 1$ for all $i = 1, \ldots, l$. Then, the sequence of projections P_n associated to Φ_n satisfies the condition $\mathcal{C}_m$ where*

$$m = \liminf_{n\to\infty} \min_{1\leqslant i\leqslant l} \dim_1 M_i^{(n)} \geqslant 1.$$

Here, for a given operator X, $\dim_1 X$ denotes the dimension of the eigenspace corresponding to the eigenvalue $\lambda = 1$ of X.

Proof. A direct computation shows that the adjoint channel of Φ_n is

$$\Phi_n^*(A) = \sum_{i=1}^{l} \mathrm{Tr}[A\sigma_i]M_i^{(n)}.$$

First, note that the operator $\Phi_n^*(A)$ has eigenvalue $\mathrm{Tr}[A\sigma_i]$ with multiplicity $\dim_1 M_i^{(n)}$. Define now

$$f(A) = \max_{1\leqslant i\leqslant l} \mathrm{Tr}[A\sigma_i].$$

It follows that $\Phi_n^*(A)$ has eigenvalue $f(A)$ with multiplicity at least

$$m_n = \min_{1\leqslant i\leqslant l} \dim_1 M_i^{(n)} \geqslant 1.$$

Also, we claim that $\|\Phi_n^*(A)\| = f(A)$:

$$\Phi_n^*(A) = \sum_{i=1}^{l} \mathrm{Tr}[A\sigma_i]M_i^{(n)} \leqslant \sum_{i=1}^{l} f(A)M_i^{(n)} = f(A)I_N.$$

We conclude that $f(A)$ is the largest eigenvalue of $\Phi^*(A)$ and that it has multiplicity at least m_n; the conclusion follows by Proposition 2.4. ∎

REMARK 5.3. The condition $\|M_i^{(n)}\| = 1$ ensures that the image of the channel Φ_n is precisely $\text{hull}(\sigma_i)_{i=1}^l$, and thus the convergence to the limiting set K is again obvious.

6. EXAMPLES OF RANDOM PROJECTIONS

In this section we look at random projection operators and we show how Theorem 2.8 together with Theorems 4.1 and 4.2 give interesting examples.

6.1. RANDOM STINESPRING CHANNELS. Let us first study channels coming from random isometries. Such random channels were used by Hayden and Winter [21] to show violations of additivity for minimum p-Rényi entropy, for $1 < p$. Following Hastings' counterexample [20] (see the next subsection), it was shown that they also violate additivity for the von Neumann entropy ($p = 1$) [2], [6], [15], [16]. More recently, the output of these channels has been fully characterized using free probability theory [3] and macroscopic violations (or order of 1 bit) for the additivity of the MOE have been observed [4].

We construct the channel from the Stinespring dilation

$$\Phi_n(X) = [\text{id} \otimes \text{Tr}](VXV^*), \tag{6.1}$$

where $V : \mathbb{C}^N \to \mathbb{C}^k \otimes \mathbb{C}^n$ is a random Haar isometry. In particular, the operator $P_n = V_n V_n^*$ projects onto a random Haar N-dimensional subspace of $\mathbb{C}^k \otimes \mathbb{C}^n$. The asymptotic regime is as follows: we fix a parameter $t \in (0,1)$, and N is any function of n that satisfies $N \sim tnk$.

Under these circumstances, the convex set K defined in (2.3) is renamed $K_{k,t}$ and it was studied at length in [3] and [4].

PROPOSITION 6.1. *Consider the free product $\mathcal{M}$ of the von Neumann non-commutative probability spaces $(M_k(\mathbb{C}), \text{tr})$ and $(\mathbb{C}^2, t\delta_1 + (1-t)\delta_2)$. The element $p = (1,0)$ of $\mathbb{C}^2$ in $\mathcal{M}$ is a selfadjoint projection of trace t, free from elements in $M_k(\mathbb{C})$. For any $A \in M_k(\mathbb{C})$, we define $f_t(A) = \|pAp\|$. Then, the sequence of projections P_n defining the quantum channels* (6.1) *satisfies condition C_m for any m, with limiting function f_t.*

A proof can be deduced from the next section on mixed unitary channels. We also refer the reader to [3], [4] for the proof of the following theorem, gathering some of the most important properties of the set $K_{k,t}$. As an original motivation, let us state the following theorem, in which the element with least entropy inside $K_{k,t}$ is identified.

THEOREM 6.2. *The convex set $K_{k,t}$ has the following properties:*

(i) *It is conjugation invariant:* $A \in K_{k,t} \iff UAU^* \in K_{k,t}$, *for all* $U \in \mathcal{U}(k)$. *In particular, one only needs the eigenvalues of a selfadjoint element in order to decide if it belongs to* $K_{k,t}$ *or not.*

(ii) *Its boundary is smooth if and only if* $t < k^{-1}$.

(iii) *Any self-adjoint element with eigenvalues*

$$\lambda = (a, \underbrace{b, b, \ldots, b}_{k-1 \text{ times}}),$$

where

$$a = \begin{cases} t + \frac{1}{k} - \frac{2t}{k} + 2\frac{\sqrt{t(1-t)(k-1)}}{k} & \text{if } t + \frac{1}{k} < 1, \\ 1 & \text{if } t + \frac{1}{k} \geqslant 1, \end{cases}$$

and $b = (1-a)/(k-1)$ *is a joint minimizer for all the p-Rényi entropies on* $K_{k,t}$, *for all* $p \geqslant 1$.

6.2. RANDOM MIXED UNITARY CHANNELS. In this section, we are interested in *random mixed unitary channels*, namely, convex combinations of random automorphisms of $M_n(\mathbb{C})$ (note that deterministic incarnations of these channels are also known in the literature as "random unitary channels"; in this work, we prefer the term "mixed", since the unitary operators appearing in the channel are themselves random). After the setup, we argue that this class of channel has the property $\mathcal{C}_m$ for all m. Based on this result, we identify the limiting maximum output infinity norm of this class. This section ends with the assertion that this class satisfies the property (3.1), which gives a simple relation between MOE and HC.

To set up our model, we recall that these channels can be written as follows

$$\widetilde{\Phi}_{n,k}^{(w)} : M_n(\mathbb{C}) \to M_n(\mathbb{C}) \quad \widetilde{\Phi}_{n,k}^{(w)}(X) = \sum_{i=1}^{k} w_i U_i X U_i^*,$$

and we are interested in the complementary channels; $\widetilde{\widetilde{\Phi}} = \Phi$. Here, $\{U_i\}_{i=1}^k$ are i.i.d. $n \times n$ Haar distributed random unitary matrices and w_i are positive weights which sum up to one (we shall consider the probability vector w a parameter of the model). Here, $N = n$ and the corresponding isometric embedding is the block column matrix whose i-th block is $\sqrt{w_i}U_i$. Then, the corresponding projection $P_n^{(w)} \in M_n(\mathbb{C}) \otimes M_k(\mathbb{C})$ is given by

$$P_n^{(w)} = \sum_{i,j=1}^{k} \sqrt{w_i w_j}\, e_i e_j^* \otimes U_i U_j^*,$$

where $\{e_i\}$ is the canonical basis of $\mathbb{C}^k$. Our model of this paper corresponds to the complementary channel of this channel $\Phi_{n,k}^{(w)} : M_n(\mathbb{C}) \to M_k(\mathbb{C})$, such that the matrix entries of its output are as follows

$$(\Phi_{n,k}^{(w)}(X))_{ij} = \sqrt{w_i w_j}\, \mathrm{Tr}[U_i X U_j^*]. \tag{6.2}$$

Firstly, we claim that these sequences of channels almost surely have property $\mathcal{C}_m$ with $m \geqslant 1$, and moreover, the limiting function can be written explicitly as follows. Let $L(F_k)$ be the free group von Neumann algebra with k free generators $u_1, \ldots, u_k$. Consider the algebra $M_k(L(F_k))$. This algebra contains $M_k(\mathbb{C})$ in a natural way, and for $A \in D_k$ with respect to this inclusion, define

$$f_w(A) = \|P^{(w)} A P^{(w)}\|, \tag{6.3}$$

where, for all i, j, $P_{ij}^{(w)} = \sqrt{w_i \overline{w_j}} u_i u_j^*$. Then, our first claim is:

PROPOSITION 6.3. *The sequence of orthogonal projections $P_n^{(w)}$ almost surely satisfies condition $\mathcal{C}_m$ for all m with limiting function f_w defined in* (6.3).

Proof. First, notice that $P_n^{(w)}(A \otimes I_n)P_n^{(w)}$ can be understood as polynomials of P_n's with coefficients in M_k. Indeed,

$$P_n^{(w)}(A \otimes I_n)P_n^{(w)} = \sum_{i,j,s,t=1}^{k} \sqrt{w_i \overline{w_j} w_s \overline{w_t}} (e_i e_j^* \, A \, e_s e_t^*) \otimes U_i U_j^* U_s U_t^*.$$

Hence, Theorem 4.2 implies that, for any fixed matrix $A \in M_k$, the operator norm of $P_n^{(w)}(A \otimes I_n)P_n^{(w)}$ converges to the operator norm of $P^{(w)} A P^{(w)}$ because it follows from Theorem 4.1 that k independent random unitary matrices $(U_i^{(n)})_{i=1}^k$ strongly converges to a k-tuple of free unitary elements $(u_i)_{i=1}^k$ almost surely.

We have thus shown that, for every matrix $A \in D_k$, almost surely, $\|P_n^{(w)}(A \otimes I_n)P_n^{(w)}\| \to f_w(A)$. To conclude that the property $\mathcal{C}_1$ holds, we have to show the above convergence simultaneously, for all A. To do this, consider a countable set $(A_i) \subset D_k$ with the property that for all $A \in D_k$ and for all $\varepsilon > 0$, there is some i such that $\|A - A_i\| < \varepsilon$. By taking a countable intersection of probability one events, the convergence $\|P_n^{(w)}(A_i \otimes I_n)P_n^{(w)}\| \to f_w(A_i)$ holds almost surely, for all $i \geqslant 1$. Since the function $f(\cdot)$ is continuous from the definition, the above chosen sequences of projections show the convergence $\|P_n^{(w)}(A \otimes I_n)P_n^{(w)}\| \to f_w(A)$ for all matrices $A \in D_k$.

Next, we show $\mathcal{C}_m$ property with $m > 1$. Remember that the infinity norm is defined by the limit of p-norms as in (4.1), the limiting density function yields non-vanishing measure around the limiting infinity norm. More precisely, for any $\varepsilon > 0$ there exists a ratio $0 < \eta_\varepsilon < 1$ such that the measure of the ε-neighborhood of the infinity norm is η_ε. Hence, fix $A \in D_k$ and for large enough n, we have $\eta_\varepsilon \cdot n$ eigenvalues of $P_n^{(w)}(A \otimes I_n)P_n^{(w)}$ which are 2ε-close to the limiting infinity norm. As $\varepsilon > 0$ is arbitrary, for any $m \geqslant 1$, the largest m eigenvalues converge to the operator norm almost surely. Again, we can prove that almost surely all the sequences show this convergence for all $A \in M_k$. This proves $\mathcal{C}_m$ property with $m > 1$. ∎

Secondly, we characterize the limiting value of the maximal output infinity norm via Proposition 6.3. To do so, we recall the following result from [1], generalizing questions that can be traced back to [24] (for a matricial coefficient version, see [25]):

PROPOSITION 6.4 ([1], Theorems IV G and IV K). *Consider an integer $k \geqslant 2$ and let $\{u_1, \ldots, u_k\}$ be a family of free unitary random variables and $a = (a_1, \ldots, a_k)$ a scalar vector. Then,*

$$\psi(a) := \Big\| \sum_{i=1}^{k} a_i u_i \Big\| = \min_{x \geqslant 0} \Big[2x + \sum_{i=1}^{k} \Big(\sqrt{x^2 + |a_i|^2} - x \Big) \Big].$$

Moreover,

$$\min_{\|a\|_1 = 1} \psi(a) = \frac{2\sqrt{k-1}}{k}, \quad \max_{\|a\|_2 = 1} \psi(a) = \frac{2\sqrt{k-1}}{\sqrt{k}},$$

with both extrema being achieved on "flat" vectors, i.e. vectors with $|a_i| = const.$

REMARK 6.5. In [1], the minimum in the formula for ψ is taken over all values $x \geqslant 0$, but one can show, by considering the derivative of the above function at $x = 0$, that the minimum is achieved at a strictly positive value $x > 0$ for $k \geqslant 3$.

Let us introduce the following notation: for a given vector $b \in \mathbb{C}^k$, let

$$\psi_*(b) = \sup_{\|a\|_2 = 1} \psi(a \cdot b), \tag{6.4}$$

where $(a \cdot b)_i = a_i b_i$. From the result above, we have that

$$\psi_*(\underbrace{(1, \ldots, 1)}_{k \text{ times}}) = \frac{2\sqrt{k-1}}{\sqrt{k}}.$$

Next, from this proposition, we can show the following results:

THEOREM 6.6. *For $k \geqslant 3$,*

(i) *The function f_w defined in* (6.3) *satisfies*

$$\max_{A \in D_k, \, \mathrm{rk}\, A = 1} f_w(A) = \psi_*(\sqrt{w}),$$

where $\sqrt{w} \in \ell^2$ is the vector with coordinates $(\sqrt{w})_j = \sqrt{w_j}$. The general formula for $\psi_*(\sqrt{w})$ is described in the Appendix, Proposition A.1.

(ii) *This implies that, with probability one, $\|\Phi_{n,k}^{(w)}\|_{1,\infty}$ converges to $\psi_*(\sqrt{w})^2$ as $n \to \infty$.*

(iii) *In the particular case of the flat distribution $w = w_{\mathrm{flat}} = (1/k, \ldots, 1/k)$, we have*

$$\psi_*(\sqrt{w_{\mathrm{flat}}}) = \psi(w_{\mathrm{flat}}) = \frac{2\sqrt{k-1}}{k}, \quad \lim_{n \to \infty} \|\Phi_{n,k}^{(w_{\mathrm{flat}})}\|_{1,\infty} = \frac{4(k-1)}{k^2}.$$

Proof. Since $A \in D_k$ is a pure state, it can be written as $A = aa^*$ for some unit vector $a = (a_1, \ldots, a_k) \in \mathbb{C}^k$. Then, since we work in C^*-algebra,

$$\begin{aligned} f_w(A) &= \|P^{(w)} a a^* P^{(w)}\|_{M_k(L(F_k))} = \|P^{(w)} a\|^2_{\ell^2(L(F_k))} \\ &= \sum_{i=1}^k \|[P^{(w)} a]_i\|^2_{L(F_k)} = \sum_{i=1}^k \Big\| \sqrt{w_i} u_i \sum_{j=1}^k \sqrt{w_j} \bar{a}_j u_j^* \Big\|^2 \\ &= \sum_{i=1}^k w_i \Big\| \sum_{j=1}^k a_j \sqrt{w_j} u_j \Big\|^2 = \Big\| \sum_{j=1}^k a_j \sqrt{w_j} u_j \Big\|^2 = \psi(a \cdot \sqrt{w})^2. \end{aligned}$$

Taking the supremum over all $a \in \mathbb{C}^k$ with $\|a\|_2 = 1$ proves the first claim. The second one is a consequence of Proposition 3.1 and the third is shown by Proposition 6.4. ∎

REMARK 6.7. Note that the function f_w is not "spectral", i.e. it does not depend only on the spectrum of its input, as it is the case for the function f_t from Proposition 6.1. Indeed, notice that, with the choice of the unit vectors

$$a^{(1)} = (1, 0, \ldots, 0), \quad a^{(2)} = (\tfrac{1}{\sqrt{k}}, \tfrac{1}{\sqrt{k}}, \ldots, \tfrac{1}{\sqrt{k}}),$$

one has

$$1 = f_w(a^{(1)}(a^{(1)})^*) \neq f_w(a^{(2)}(a^{(2)})^*) = \frac{2\sqrt{k-1}}{\sqrt{k}},$$

although the matrices $A_i = a_i a_i^*$ are isospectral.

Thirdly, we claim that, in the limit, the minimum output entropy and the Holevo capacity of the channel (4.2) identify each other:

THEOREM 6.8. *The convex set K for $\Phi_n^{(w)}$ has the property* (3.1).

Proof. Take the Weyl operators $W_{a,b}$ as defined in (3.2) and calculate

$$\|P_n(W_{a,b} A W_{a,b}^* \otimes I_n) P_n\|_\infty = \|(\sqrt{A} \otimes I_n) \underbrace{(W_{a,b}^* \otimes I_n) P_n (W_{a,b} \otimes I_n)}_{(\star)} (\sqrt{A} \otimes I_n)\|_\infty \tag{6.5}$$

while we have

$$\begin{aligned} (\star) &= (W_{a,b}^* \otimes I_n) \Big(\sum_{s,t=1}^k \sqrt{w_s w_t}\, e_s e_t^* \otimes U_s U_t^* \Big) (W_{a,b} \otimes I_n) \\ &= \sum_{s,t=1}^k \sqrt{w_s w_t} \exp\{\tfrac{2\pi \mathrm{i}}{n} b(t-s)\} e_{s-a} e_{t-a}^* \otimes U_s U_t^* \\ &= \sum_{s,t=1}^k \sqrt{w_s w_t} e_{s-a} e_{t-a}^* \otimes (\exp\{-bs\tfrac{2\pi \mathrm{i}}{k}\} U_s)(\exp\{-bt\tfrac{2\pi \mathrm{i}}{k}\} U_t)^*. \end{aligned}$$

This implies that $(W_{a,b}^* \otimes I_n)P_n(W_{a,b} \otimes I_n)$ have the same law for all $(a,b) \in \mathbb{Z}_k \times \mathbb{Z}_k$ because $\{U\}_{i=1}^k$ are i.i.d. with respect to the Haar measure. Therefore, (3.3) is true and then Corollary 3.4 completes the proof. ∎

7. IMAGE OF TENSOR PRODUCT OF CHANNELS

In this section, we investigate the image of tensor products of two channels. Section 7.1 describes general theory when one channel has a nice asymptotic behavior and the other is fixed. In Section 7.2, we consider cases where the fixed channel is entanglement-breaking.

7.1. TENSOR WITH ANY FINITE DIMENSIONAL QUANTUM CHANNEL.

Our setting is as follows. Let Ψ_n be a quantum channel obtained from P_n a sequence of projections in $M_k \otimes M_n$ of rank $N = N(n)$, namely,

$$\Psi_n : M_N \to M_k.$$

Then,

THEOREM 7.1. *If the family $(P_n, E_{ij} \otimes I_n : i,j \in \{1,\dots,k\})$ converges strongly as in the definition of Section 4 then, for any quantum channel $\Xi : M_p \to M_q$, there exists a convex body K in D_{kq} such that*

$$\Xi \otimes \Psi_n(D_{pN}) \to K$$

as in the sense of Theorem 2.9.

REMARK 7.2. In this setting, existence of the limiting convex set of output states of the tensor products depends only on the asymptotic behavior of Ψ_n.

Proof. First, we choose $m \in \mathbb{N}$ such that there exists a projection P of rank p on $M_q \otimes M_m$ which is associated to Ξ. This construction can be made uniquely up to an isometry.

Next, it follows from Theorem 4.2 that the fact that $(P_n, E_{ij} \otimes I_n : i,j \in \{1,\dots,k\})$ converges strongly as $n \to \infty$ implies also that

$$(P_n \otimes P, E_{i_1 j_1} \otimes 1_n \otimes E_{i_2 j_2} : i_1, j_1 \in \{1,\dots,k\}, i_2, j_2 \in \{1,\dots,qm\})$$

converges strongly. This strong convergence implies that for any $A \in M_k \otimes M_q$, the sequence $P_n \otimes PA \otimes 1_{nm} P_n \otimes P$ satisfies the condition $\mathcal{C}_l$ (see Definition 2.3) for any l. Note that in the above equation, we viewed $A \otimes 1_{nm}$ as an element of $M_k \otimes M_n \otimes M_q \otimes M_n$.

Finally, the proof then follows from Theorem 2.9. ∎

We want to point out that it remains difficult to analyze the limiting outputs sets K of Theorem 7.1 in general. For example, even in the simple case where Ξ is the identity map, we are unable to describe the collection of limiting output sets.

7.2. TENSOR WITH ENTANGLEMENT BREAKING CHANNEL. It seems difficult in general to compute K explicitly in the tensor product case. However, when Ξ is an entanglement-breaking channel of certain type, we can write down the image explicitly. In this section, channels tensored with entanglement-breaking channels are fixed and we do not use the asymptotic behavior to get results, in the first place.

Let us start with an interesting example among entanglement-breaking channels, which is called *pinching map*:

$$\begin{aligned} \Xi : M_l &\to M_l \\ [m_{i,j}]_{i,j=1}^l &\mapsto [\delta_{i,j} m_{i,j}]_{i,j=1}^l \end{aligned}$$

where $m_{i,j}$ is the (i,j)-element of square matrices.

PROPOSITION 7.3. *Let $\Xi : M_l \to M_l$ be the pinching map. Then the image $K_{\Xi\otimes\Psi}$ can be described as follows:*

$$\widetilde{K} = \{a_1 K_\Psi \oplus \cdots \oplus a_l K_\Psi ; (a_i) \in \Delta_l\}.$$

Proof. This follows directly from the fact that the image of S_{lN} under $\Xi \otimes 1_N$ is exactly $\{a_1 S_N \oplus \cdots \oplus a_l S_N, (a_i) \in \Delta_l\}$. This can be readily seen by double inclusion. ∎

This has an immediate corollary:

COROLLARY 7.4. *Let Ψ_n be a sequence of quantum channels obeying the hypotheses of Theorem* 2.8 *(or Theorem* 2.9*). Then, for any integer l, the conclusion of Theorem* 2.8 *(or Theorem* 2.9*) still holds true for $\Psi_n^{\oplus l}$, where K is replaced by $K^{\oplus l}$.*

The images of entanglement-breaking channels are described as follows:

LEMMA 7.5. *For an entanglement-breaking channel Ξ defined in* (5.2)*. Then, $K_\Xi = \Xi(S_p)$ is written as*

$$K_\Xi = \Big\{ \sum_{i=1}^l p_i \sigma_i : (p_i) \in \Delta_\Xi \Big\}.$$

Here, we denote possible probability distributions by channel Ξ by Δ_Ξ.

A straightforward application of Lemma 7.5 implies:

LEMMA 7.6. *Suppose we have two quantum channels Ξ and Ψ. Let Ξ be an entanglement-breaking channel defined in* (5.2)*. Then, the set of images of all the states via $\Xi \otimes \Psi$ is given by*

$$K_{\Xi\otimes\Psi} = \operatorname{hull}\Big\{ \sum_{i=1}^l \sigma_i \otimes \Psi(B M_i^{\mathrm{T}} B^*) : B \in M_{N,p} \text{ with } \operatorname{Tr}[BB^*] = 1 \Big\}. \tag{7.1}$$

Proof. Let the input spaces of Ξ and Ψ be $\mathbb{C}^p$ and $\mathbb{C}^N$, respectively. Take a bipartite vector b in $\mathbb{C}^N \otimes \mathbb{C}^p$ and calculate as follows:

$$(\Xi \otimes \Psi)(bb^*) = \sum_{i=1}^{l} \sigma_i \otimes \Psi(BM_i^{\mathrm{T}}B^*)$$

where we used the canonical isomorphism $\mathbb{C}^N \otimes \mathbb{C}^p \ni b \leftrightarrow B \in M_{N,p}(\mathbb{C})$. Indeed,

$$\mathrm{Tr}_{\mathbb{C}^l}[bb^*(M_i \otimes I_N)] = BM_i^{\mathrm{T}}B^*. \quad \blacksquare$$

Then, we define

$$K_\Xi \otimes K_\Psi \tag{7.2}$$

such that $\otimes$ in the formula yields the smallest convex set which contains all the simple tensors. It is easy to see that (7.2) $\subseteq$ (7.1):

$$K_\Xi \otimes K_\Psi \subseteq K_{\Xi\otimes\Psi}.$$

These two sets turn out to be identical under some assumption:

THEOREM 7.7. *Suppose we have two quantum channels Ξ and Ψ. Let Ξ be an entanglement-breaking channel defined in* (5.2) *such that $\|M_i\|_\infty = 1$ for $1 \leqslant i \leqslant l$. Then:*

$$K_{\Xi\otimes\Psi} = K_\Xi \otimes K_\Psi.$$

Proof. We show $K_\Xi \otimes K_\Psi \supseteq K_{\Xi\otimes\Psi}$. This is true if for all B there exist $\{r_k\} \in \Delta_d$, $\{p_i^{(k)}\}_i \in \Delta_\Xi$, $\rho^{(k)} \in S$ such that

$$\sum_{k=1}^{d} r_k \Big(\sum_{i=1}^{l} p_i^{(k)} \sigma_i \otimes \Psi(\rho^{(k)}) \Big) = \sum_{i=1}^{l} \sigma_i \otimes \Psi(BM_i^{\mathrm{T}}B^*) \tag{7.3}$$

and this is true if

$$\sum_k r_k p_i^{(k)} \rho^{(k)} = BM_i^{\mathrm{T}}B^* \quad \forall i \in \{1, \ldots, l\}.$$

This can be written, by abusing notations, as

$$P \cdot \Gamma = \begin{pmatrix} p_1^{(1)} & \ldots & p_1^{(d)} \\ \vdots & \ddots & \vdots \\ p_l^{(1)} & \ldots & p_l^{(d)} \end{pmatrix} \begin{pmatrix} \gamma^{(1)} \\ \vdots \\ \gamma^{(d)} \end{pmatrix} = \begin{pmatrix} BM_1^{\mathrm{T}}B^* \\ \vdots \\ BM_l^{\mathrm{T}}B^* \end{pmatrix} \tag{7.4}$$

with $\gamma^{(k)} = r_k \rho^{(k)}$. Since each M_i has an eigenvalue 1, $\Delta_\Xi = \Delta_l$. Hence we set $d = l$ and

$$P = I_l; \quad \gamma^{(k)} = BM_k^{\mathrm{T}}B^*. \quad \blacksquare$$

Entanglement breaking channels having POVM operators of unit operator norm have been investigated in [18] in relation to the convex geometry of the output of quantum channels.

We think that the above condition $\Delta_\Xi = \Delta_l$ should be close to a necessary condition too. We set $d = l$ and think whether each block of

$$P^{-1} \times \begin{pmatrix} M_1^{\mathrm{T}} \\ \vdots \\ M_l^{\mathrm{T}} \end{pmatrix}$$

is positive or not. Suppose we have chosen P as

$$\lambda I + (1-\lambda)\psi\psi^*.$$

Here, $0 < \lambda \leqslant 1$ and $\psi = (1/\sqrt{l})(1, \ldots, 1)^{\mathrm{T}}$. Set

$$Q = I - P = (1-\lambda)(I - \psi\psi^*).$$

Then,

$$P^{-1} = (I-Q)^{-1} = \sum_{i=0}^{\infty} Q^i = \frac{1}{\lambda}(I - \psi\psi^*).$$

However then this always gives a non-positive block. Indeed, the i-th block, rescaled, will be

$$M_i - \frac{1}{l}\sum_{j=1}^{l} M_j = M_i - \frac{1}{l} I$$

and one of them should be non-positive.

Appendix A. THE OPTIMIZATION PROBLEM FOR RANDOM MIXED UNITARY CHANNELS

In this technical appendix, we provide the details of the proof for the optimization problem appearing in Theorem 6.6. Let us recall it here, for the convenience of the reader. Let $\mathcal{S}_{\mathbb{C}}^{k-1}$ be the unit sphere of $\mathbb{C}^k$ and define

$$g : \mathcal{S}_{\mathbb{C}}^{k-1} \times (0, \infty) \to \mathbb{R} \tag{A.1}$$

$$(a, x) \mapsto (2-k)x + \sum_{i=1}^{k} \sqrt{x^2 + |a_i|^2 w_i} \tag{A.2}$$

where $k > 2$ is an integer parameter and $(w_1, \ldots, w_k)$ is a strictly positive probability vector: $w_i > 0$ and $\sum_i w_i = 1$. Since only the absolute values $|a_i|^2$ appear in the above formula, we shall assume, without loss of generality, that the numbers a_i are real and satisfy $\sum_i a_i^2 = 1$.

In what follows, we prove the following result:

PROPOSITION A.1. *Let g be the function defined in* (A.1)*, but on $\mathcal{S}_{\mathbb{R}}^{k-1} \times (0, \infty)$ as is described above. Then*

(i) *We have the formula:*

$$\psi_*(\sqrt{w}) = \max_{J \in \mathcal{J}} h(J). \tag{A.3}$$

Here, remember that $\psi_*(\sqrt{w}) = \max_{a\in\mathcal{S}^{k-1}} \min_{x>0} g(a,x)$ *defined in* (6.4). *In the above formula,* $\mathcal{J}$ *is a collections of subsets of* $[k] = \{1,\dots,k\}$, *defined as*

$$\mathcal{J} = \Big\{ J \subset [k] : \min_{j\in J} w_j \geqslant \gamma |\#J - 2| \Big\} \tag{A.4}$$

elements of which we call valid subsets. Also, the function $h(\cdot)$ *is defined on* $\mathcal{J}$ *as*

$$h(J) = \sqrt{\beta - \gamma(\#J-2)^2}$$

where β *and* γ *are*

$$\beta = \sum_{j\in J} w_j \frac{1}{\gamma} = \sum_{j\in J} \frac{1}{w_j}. \tag{A.5}$$

Note that $\mathcal{J}$ *contains all subsets with cardinality less than or equal to* 3.

(ii) *The function* $h(\cdot)$ *is well-defined on on* $2^{[k]}$ *and non-decreasing with respect to the canonical partial order. As a result, if the full set* $J = [k]$ *is valid, i.e.* $\min_{j\in[k]} w_j \geqslant \gamma_0(k-2)$ *with* $\gamma_0^{-1} = \sum_{j=1}^{k} w_j^{-1}$, *then the optimum is* $\sqrt{1-\gamma_0(k-2)^2}$. *In particular, when* w_i *is the flat distribution,* $w_i = 1/k$, *we get* $a_i = 1/k$ *and the optimum is* $2\sqrt{k-1}/k$.

Proof. Let us start by giving an outline of the proof. First, we notice that the minimization problem in x is convex, hence a unique minimum exists. Moreover, this minimum X_a depends smoothly on a and thus we are left with a smooth maximization problem in $a \in \mathcal{S}^{k-1}_{\mathbb{R}}$. Next, we use Lagrange multipliers to solve this problem, and we find a set of critical points indexed by subsets $J \subset [k] = \{1,\dots,k\}$, where the coordinates a_i are non-zero. Not all subsets J yield critical points and one has to take a maximum over the set of valid subsets J to conclude. Finally, we show monotonic property of the function $h(\cdot)$ with respect to the partial order in $2^{[k]}$.

Step 1. Let us start by noticing that, at fixed a, the function $x \mapsto g(a,x)$ is convex, so it admits a unique minimum $X_a \in [0,\infty)$. Since $\partial g/\partial x$ is negative at $x = 0$, we have $X_a > 0$ (see also Remark 6.5). The value X_a is defined by the following implicit equation

$$\frac{\partial g}{\partial x}|_{x=X_a} = 0,$$

which is equivalent to $F(a, X_a) = 0$, for

$$F(a,x) = \frac{\partial g}{\partial x} = 2 - k + \sum_{i=1}^{k} \frac{x}{\sqrt{x^2 + a_i^2 w_i}}. \tag{A.6}$$

It follows from the implicit function theorem that the map $a \mapsto X_a$ is C^1 because

$$\frac{\partial F}{\partial x} = \sum_{i=1}^{k} \Big[\frac{1}{(x^2 + a_i^2 w_i)^{1/2}} - \frac{x^2}{(x^2 + a_i^2 w_i)^{3/2}} \Big] = \sum_{i=1}^{k} \frac{a_i^2 w_i}{(x^2 + a_i^2 w_i^2)^{1/2}} \neq 0.$$

Step 2. Now we want to solve

$$\max_{a\in\mathcal{S}_{\mathbb{R}}^{k-1}} g(a, X_a)$$

by introducing the Lagrange multiplier functional

$$G(a,\lambda) = g(a,X_a) - \frac{\lambda}{2}\sum_{i=1}^{k} a_i^2 = (2-k)X_a + \sum_{i=1}^{k}\sqrt{X_a^2 + a_i^2 w_i} - \frac{\lambda}{2}\sum_{i=1}^{k} a_i^2.$$

The criticality condition, the normalization for a and the restriction of X_a translate to

$$\forall j\in[k],\quad \frac{\partial X_a}{\partial a_j}\Big(2-k+\sum_{i=1}^{k}\frac{X_a}{\sqrt{X_a^2+a_i^2w_i}}\Big) + \frac{a_j w_j}{\sqrt{X_a^2+a_j^2w_j}} - \lambda a_j = 0, \tag{A.7}$$

$$\sum_{i=1}^{k} a_i^2 = 1, \tag{A.8}$$

$$F(a, X_a) = 0. \tag{A.9}$$

Below, we get candidates for the solutions for this system of equations.

Firstly, (A.7) and (A.9) imply that

$$\forall j\in[k],\quad \frac{a_j w_j}{\sqrt{X_a^2+a_j^2w_j}} = \lambda a_j.$$

Let us now introduce the index sets $I = \{i : a_i = 0\}$ and $J = [k]\setminus I$. Then, for $j\in J$ we have

$$w_j = \lambda\sqrt{X_a^2 + a_j^2 w_j}\,.$$

This implies two equations: (A.9) gives

$$0 = 2 - k + \#I + \lambda X_a\sum_{j\in J}\frac{1}{w_j} \quad\text{or}\quad \#J - 2 = \frac{\lambda X_a}{\gamma}$$

and, squaring the both sides yields

$$a_j^2 = \frac{w_j}{\lambda^2} - \frac{X_a^2}{w_j} = \frac{1}{\lambda^2}\Big(w_j - \frac{(\lambda X_a)^2}{w_j}\Big) = \frac{1}{\lambda^2}\Big(w_j - \frac{\gamma^2(\#J-2)^2}{w_j}\Big).$$

Secondly, with (A.8), we have

$$1 = \sum_{j\in J} a_i^2 = \frac{1}{\lambda^2}\Big(\beta - \frac{(\lambda X_a)^2}{\gamma}\Big) = \frac{1}{\lambda^2}(\beta - \gamma(\#J-2)^2).$$

This leads to

$$a_j^2 = \frac{1}{\beta-\gamma(\#J-2)^2}\cdot\Big(w_j - \frac{\gamma^2(\#J-2)^2}{w_j}\Big). \tag{A.10}$$

Also,

$$X_a = \frac{\gamma(\#J-2)}{|\lambda|} = \frac{\gamma(\#J-2)}{\sqrt{\beta-\gamma(\#J-2)^2}}.$$

Thirdly, for those candidates the function $g(\cdot,\cdot)$ can be simplified:

$$g(a,X_a) = (2-\#J)X_a + \sum_{j\in J}\sqrt{X_a^2+a_j^2 w_j} = (2-\#J)X_a + \frac{\beta}{\lambda}$$
$$= -\frac{\gamma(\#J-2)^2}{\sqrt{\beta-\gamma(\#J-2)^2}} + \frac{\beta}{\sqrt{\beta-\gamma(\#J-2)^2}} = \sqrt{\beta-\gamma(\#J-2)^2}.$$

Since this function only depends on set J, we redefine this function to be $h(J)$ as in the statement of theorem.

Step 3. So far, we get a set of candidates for solutions, but we get the actual solutions, and hence the precise set of critical points, by thinking positivity issues for a_j^2 with $j \in J$. The inequality between the harmonic and the arithmetic means, applied for $\{w_j\}_{j\in J}$ reads

$$\frac{\#J}{\sum_{j\in J} 1/w_j} \leqslant \frac{\sum_{j\in J} w_j}{\#J} \tag{A.11}$$

hence we have that $(\#J)^2\gamma \leqslant \beta$ for all choices of J. This implies that the first factor in (A.10) is always strictly positive, except for $\#J = 1$, when it is zero. Hence, looking into the second factor in (A.10), the condition that $a_j^2 \geqslant 0$ for all $j \in J$ is equivalent to the condition that J is a valid subset, as in (A.4) with respect to those candidates in (A.10). Therefore, maximizing $h(J)$ over $\mathcal{J}$ gives the maximum of $g(a,X_a)$ under the normalization condition on a.

Note that when $\#J = 1,2$ the condition for J to be valid, $\min\limits_{j\in J} w_j \geqslant \gamma|\#J-2|$ is trivially satisfied. When $\#J = 3$, the condition reads

$$\frac{3-2}{\min_{j\in J} w_j} \leqslant \frac{1}{\gamma} = \frac{1}{w_1}+\frac{1}{w_2}+\frac{1}{w_3}$$

which is also always fulfilled. Thus, every subset J with $\#J \leqslant 3$ is valid.

Step 4. The mean inequality (A.11) implies also that the quantity $h(J)$ is well defined for all subsets $J \subset [k]$, even if J is not valid. Let us show next h is an increasing function of J with the canonical partial order. To this end, consider a subset J, an element $s \notin J$ and put $J' = J \cup \{s\}$. With $p = \#J$, we have the following sequence of equivalent inequalities

$$h(J)^2 \leqslant h(J')^2,$$
$$\beta-\gamma(p-2)^2 \leqslant \beta'-\gamma'(p-1)^2,$$
$$-\frac{(p-2)^2}{1/\gamma} \leqslant w_s - \frac{(p-1)^2}{1/\gamma+1/w_s},$$

$$-(p-2)^2\Big(\frac{1}{\gamma}+\frac{1}{w_s}\Big) \leqslant \frac{w_s}{\gamma}\Big(\frac{1}{\gamma}+\frac{1}{w_s}\Big) - \frac{(p-1)^2}{\gamma},$$
$$\frac{1}{\gamma}\Big(\frac{2(p-2)}{w_s} - \frac{1}{\gamma}\Big) \leqslant \frac{(p-2)^2}{w_s^2},$$

where the last one is true by the inequality: $\sqrt{ab} \leqslant (a+b)/2$ for $a, b > 0$. In particular, we conclude that if $J = [k]$ is valid, then

$$\max_{J\in\mathcal{J}} h(J) = h([k]) = \sqrt{1-\gamma_0(k-2)^2}. \quad \blacksquare$$

As an illustration of the above result, let us consider the case $k = 4$ and

$$w_r = [r, \tfrac{1-r}{3}, \tfrac{1-r}{3}, \tfrac{1-r}{3}]$$

with $r \in (0, \frac{1}{4})$. For $J = \{1,2,3,4\}$ to be valid, one must have $r \geqslant 2c$. By direct computation, one finds $c = r(1-r)/(8r+1)$, thus $J = [4]$ is valid if and only if $r \in (0, 1/10)$. We conclude that, for $r \geqslant 1/10$, the optimum is $h([4]) = (2r+1)/\sqrt{8r+1}$.

Let us now study the other regime, where $r < 1/10$. There are only two distinct choices for J with $\#J = 3$: $J_1 = \{1,2,3\}$ and $J_2 = \{2,3,4\}$, both valid since they have cardinality 3. One computes directly

$$h(J_1) = \frac{\sqrt{2}}{\sqrt{3}}\frac{2r+1}{\sqrt{5r+1}} < \frac{2\sqrt{2}}{3}\sqrt{1-r} = h(J_2).$$

We conclude that

$$\psi_*(\sqrt{w_r}) = \begin{cases} \frac{2r+1}{\sqrt{8r+1}} & \text{if } r \geqslant \frac{1}{10}, \\ \frac{2\sqrt{2}}{3}\sqrt{1-r} & \text{if } r < \frac{1}{10}. \end{cases}$$

Acknowledgements. The authors had opportunities to meet at the LPT in Toulouse, the ICJ in Lyon, the Department of Mathematics of University of Ottawa, the TU München and the Isaac Newton Institute in Cambridge to complete their research, and thank these institutions for a fruitful working environment.

Benoît Collins's research was supported by NSERC discovery grants, Ontario's ERA and AIMR, Tohoku university. Motohisa Fukuda's research was financially supported by the CHIST-ERA/BMBF project CQC. Ion Nechita's research has been supported by the ANR grants "OSQPI" 2011 BS01 008 01 and "RMTQIT" ANR-12-IS01-0001-01, and by the PEPS-ICQ CNRS project "Cogit".

REFERENCES

[1] C. AKEMANN, P. OSTRAND, Computing norms in group C^*-algebras, *Amer. J. Math.* **98**(1976), 1015–1047.

[2] G. AUBRUN, S. SZAREK, E. WERNER, Hastings's additivity counterexample via Dvoretzky's theorem, *Comm. Math. Phys.* **305**(2011), 85–97.

[3] S.T. BELINSCHI, B. COLLINS, I. NECHITA, Laws of large numbers for eigenvectors and eigenvalues associated to random subspaces in a tensor product, *Invent. Math.* **190**(2012), 647–697.

[4] S.T. BELINSCHI, B. COLLINS, I. NECHITA, Almost one bit violation for the additivity of the minimum output entropy, arXiv:1305.1567 [math-ph].

[5] I. BENGTSSON, K. ŻYCZKOWSKI, *Geometry of Quantum States. An Introduction to Quantum Entanglement*, Cambridge Univ. Press, Cambridge 2006.

[6] F. BRANDAO, M.S.L. HORODECKI, On Hastings's counterexamples to the minimum output entropy additivity conjecture, *Open Sys. Inf. Dyn.* **17**(2010), 31–52.

[7] S.L. BRAUNSTEIN, Geometry of quantum inference, *Phys. Lett. A* **219**(1996), 169–174.

[8] B. COLLINS, Moments and cumulants of polynomial random variables on unitary groups, the Itzykson–Zuber integral and free probability, *Int. Math. Res. Not.* **17**(2003), 953–982.

[9] B. COLLINS, C. MALE, The strong asymptotic freeness of Haar and deterministic matrices, *Ann. Sci. École Norm. Sup.*, to appear, arXiv:1105.4345 [math.OA].

[10] B. COLLINS, I. NECHITA, Eigenvalue and entropy statistics for products of conjugate random quantum channels, *Entropy* **12**(2010), 1612–1631.

[11] B. COLLINS, I. NECHITA, Random quantum channels I: Graphical calculus and the Bell state phenomenon, *Comm. Math. Phys.* **297**(2010), 345–370.

[12] B. COLLINS, I. NECHITA, Gaussianization and eigenvalue statistics for Random quantum channels (III), *Ann. Appl. Probab.* **21**(2011), 1136–1179.

[13] B. COLLINS, I. NECHITA, Random quantum channels II: Entanglement of random subspaces, Rényi entropy estimates and additivity problems, *Adv. Math.* **226**(2011), 1181–1201.

[14] B. COLLINS, I. NECHITA, K. ŻYCZKOWSKI, Random graph states, maximal flow and Fuss–Catalan distributions, *J. Phys. A Math. Theor.* **43**(2010), no. 27, 275303.

[15] M. FUKUDA, Revisiting additivity violation of quantum channels, *Comm. Math. Phys.* **332**(2014), 713–728.

[16] M. FUKUDA, C. KING, Entanglement of random subspaces via the Hastings bound, *J. Math. Phys.* **51**(2010), no. 4, 042201.

[17] M. FUKUDA, C. KING, D. MOSER, Comments on Hastings' additivity counterexamples, *Comm. Math. Phys.* **296**(2010), 111–143.

[18] M. FUKUDA, I. NECHITA, M. WOLF, Quantum channels with polytopic images and image additivity, *IEEE Trans. Inform. Theory* **61**(2015), 1851–1859.

[19] 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**(2005), 711–775.

[20] M.B. HASTINGS, Superadditivity of communication capacity using entangled inputs, *Nature Phys.* **5**(2009), 255.

[21] P. HAYDEN, A. WINTER, Counterexamples to the maximal p-norm multiplicativity conjecture for all $p > 1$, *Comm. Math. Phys.* **284**(2008), 263–280.

[22] A. HOLEVO, Remarks on the classical capacity of quantum channel, arXiv:quant-ph/0212025.

[23] M. HORODECKI, P.W. SHOR, M.B. RUSKAI, General entanglement breaking channels, *Rev. Math. Phys.* **15**(2003), 629–641.

[24] H. KESTEN, Symmetric random walks on groups, *Trans. Amer. Math. Soc.* **92**(1959), 336–354.

[25] F. LEHNER, Computing norms of free operators with matrix coefficients, *Amer. J. Math.* **121**(1999), 453–486.

[26] C. MALE, The norm of polynomials in large random and deterministic matrices, *Probab. Theory Related Fields* **154**(2012), 477–532.

[27] I. NECHITA, Asymptotics of random density matrices, *Ann. Henri Poincaré* **8**(2007), 1521–1538.

[28] A. NICA, R. SPEICHER, *Lectures on the Combinatorics of Free Probability*, London Math. Soc. Lecture Note Ser., vol. 335, Cambridge Univer. Press, Cambridge 2006.

[29] D. PAGE, Average entropy of a subsystem, *Phys. Rev. Lett.* **71**(1993), 1291–1294.

[30] P.W. SHOR, Equivalence of additivity questions in quantum information theory, *Comm. Math. Phys.* **246**(2004), 453–472.

[31] H.-J. SOMMERS, K. ŻYCZKOWSKI, Statistical properties of random density matrices, *J. Phys. A* **37**(2004), 8457–8466.

[32] E. STEINITZ, Bedingt konvergente Reihen und Konvexe Systeme, *J. Reine Angrew. Math.* **143**(1913), 128–175; **146**(1916), 1–52; **144**(1924), 1–40.

[33] W.F. STINESPRING, Positive functions on C^*-algebras, *Proc. Amer. Math. Soc.* **6**(1955), 211–216.

[34] S. STRASZEWICS, Über exponierte Punkte abgeschlossener Punktmengen, *Fund. Math.* **24**(1935), 139–143.

[35] D. VOICULESCU, A strengthened asymptotic freeness result for random matrices with applications to free entropy, *Internat. Math. Res. Notices* **1**(1998), 41–63.

[36] K. ŻYCZKOWSKI, H.-J. SOMMERS, Induced measures in the space of mixed quantum states, *J. Phys. A* **34**(2001), 7111–7125.

BENOîT COLLINS, DÉPARTEMENT DE MATHÉMATIQUE ET STATISTIQUE, UNIVERSITÉ D'OTTAWA, 585 KING EDWARD, OTTAWA, ON, K1N6N5 CANADA, *and* KYOTO UNIVERSITY, DEPARTMENT OF MATHEMATICS, JAPAN *and* CNRS, INSTITUT CAMILLE JORDAN UNIVERSITÉ LYON 1, FRANCE

E-mail address: bcollins@uottawa.ca

MOTOHISA FUKUDA, ZENTRUM MATHEMATIK, M5, TECHNISCHE UNIVERSITÄT MÜNCHEN, BOLTZMANNSTRASSE 3, 85748 GARCHING, GERMANY

E-mail address: m.fukuda@tum.de

ION NECHITA, CNRS, LABORATOIRE DE PHYSIQUE THÉORIQUE, IRSAMC, UNIVERSITÉ DE TOULOUSE, UPS, 31062 TOULOUSE, FRANCE

E-mail address: nechita@irsamc.ups-tlse.fr

Received December 4, 2013.