The Annals of Applied Probability
2011, Vol. 21, No. 3, 1136–1179
DOI: 10.1214/10-AAP722

GAUSSIANIZATION AND EIGENVALUE STATISTICS FOR RANDOM QUANTUM CHANNELS (III)

BY BENOÎT COLLINS[1,2] AND ION NECHITA[2]

University of Ottawa

In this paper, we present applications of the calculus developed in Collins and Nechita [*Comm. Math. Phys.* **297** (2010) 345–370] and obtain an exact formula for the moments of random quantum channels whose input is a pure state thanks to Gaussianization methods. Our main application is an in-depth study of the random matrix model introduced by Hayden and Winter [*Comm. Math. Phys.* **284** (2008) 263–280] and used recently by Brandao and Horodecki [*Open Syst. Inf. Dyn.* **17** (2010) 31–52] and Fukuda and King [*J. Math. Phys.* **51** (2010) 042201] to refine the Hastings counterexample to the additivity conjecture in quantum information theory. This model is exotic from the point of view of random matrix theory as its eigenvalues obey two different scalings simultaneously. We study its asymptotic behavior and obtain an asymptotic expansion for its von Neumann entropy.

1. Introduction. In the paper [9] we developed a calculus permitting the computation of any moments of random quantum channels. It has already proven useful in understanding the random matrix models involved in the additivity violation theorems and in improving lower bounds of dimensions needed to obtain violation of the additivity of entropy estimates (developed in [9–11]), as well as in the study of random quantum states associated with graphs [12]. In the present work we study two more applications of our calculus, to new random matrix models introduced for quantum information theoretic purposes.

The first application is of theoretical interest and of a nonasymptotic nature: we extend our calculus to Gaussian matrices and show that it yields explicit formulas for the moments of Wishart matrices and of outputs of random quantum channels. The formulas are of a purely combinatorial nature and make it possible to bypass Weingarten calculus, whose asymptotic estimates can be involved. For this we use a "Gaussianization" method.

The second application is an extended study of the random matrix model that was introduced by Hayden and Winter in [20] and used recently in [4, 14, 15] to refine the results of Hastings [19]. As a motivation, let us recall the quantum information theoretic context of this random matrix. A *quantum channel* is a linear

Received April 2010.
[1]Supported in part by ANR GranMa and ANR Galoisint.
[2]Supported in part by NSERC Grant RGPIN/341303-2007.
MSC2010 subject classifications. Primary 15A52; secondary 94A17, 94A40.
Key words and phrases. Random matrices, Weingarten calculus, quantum information theory, random quantum channel.

completely positive trace-preserving map $\Phi : \mathcal{M}_n(\mathbb{C}) \to \mathcal{M}_k(\mathbb{C})$. A *density matrix* is a self-adjoint positive semidefinite matrix with trace 1. Let $\Delta_k = \{x \in \mathbb{R}_+^k \mid \sum_{i=1}^k x_i = 1\}$ be the $(k-1)$-dimensional probability simplex. The *Shannon entropy* of x is defined to be

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

These definitions are extended to density matrices by functional calculus:

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

For a quantum channel $\Phi : \mathcal{M}_n(\mathbb{C}) \to \mathcal{M}_k(\mathbb{C})$, its *minimum output entropy* is defined by

$$H_{\min}(\Phi) = \min_{\substack{\rho \in \mathcal{M}_n(\mathbb{C}) \\ \rho \geq 0, \operatorname{Tr} \rho = 1}} H(\Phi(\rho)).$$

The additivity conjecture for minimum output entropies is arguably one of the most important in quantum information theory, and it can be stated as follows.

CONJECTURE 1.1. *For all quantum channels Φ_1 and Φ_2, we have*

$$H_{\min}(\Phi_1 \otimes \Phi_2) = H_{\min}(\Phi_1) + H_{\min}(\Phi_2). \tag{1}$$

This conjecture was disproven by Hastings in [19] as follows.

THEOREM 1.2. *There exists a counterexample to the conjecture for the choice $\Phi_1 = \overline{\Phi_2}$.*

In the proof of [19], one reason why $\Phi_1 = \overline{\Phi_2}$ yields a counterexample is that it ensures that the largest eigenvalue of outputs of well-chosen inputs—Bell states—is much bigger than the other eigenvalues. The counterexamples to the additivity conjecture obtained thus far use a random matrix model which we redefine in Section 6.3 and call Z_n. The main result of this paper is as follows (the dimension ratio c is a fixed positive constant).

THEOREM 1.3. *As $n \to \infty$, $k \sim cn$, the eigenvalues $\lambda_1 \geq \cdots \geq \lambda_{n^2}$ of Z_n are such that*:

- *in probability*, $cn\lambda_1 \to 1$;
- *almost surely*, $\frac{1}{n^2-1} \sum_{i=2}^{n^2} \delta_{c^2 n^2 \lambda_i}$ *converges to a Marchenko–Pastur distribution of parameter* c^2;

- *almost surely*,

$$H(Z_n) = \begin{cases} 2\log n - \dfrac{1}{2c^2} + o(1), & \text{if } c \geq 1, \\ 2\log(cn) - \dfrac{c^2}{2} + o(1), & \text{if } 0 < c < 1, \end{cases}$$

as $n \to \infty$, *where* H *is the von Neumann entropy.*

The interest of this result is that it yields improvements to the results of [4, 14, 15, 19], as the only data that these papers were using was a lower bound on the largest eigenvalue of Z_n, whereas the above theorem gives a full understanding of the eigenvalue behavior of Z_n.

In addition, the matrix model Z_n has the novel property that it has two different regimes for its eigenvalues (one in n^{-1} and one in n^{-2}). As far as we know, it is the first model in random matrix theory whose eigenvalues have two regimes simultaneously.

The proof of the main theorem uses a mix of moment methods and functional calculus methods. It is very instructive, as the moment method is used to prove the convergence in distribution of the eigenvalues of smaller decay, and this goes beyond the standard intuition that moment methods instead give results about the larger eigenvalues. Actually, our Theorem 6.10 shows new kinds of cancellation properties, going beyond those which are usually expected with standard "moments–cumulants" and "connectedness" arguments.

This paper is organized as follows. We first recall some known facts about Wick calculus, Weingarten calculus and noncommutative and free probability theory. We also recall our graphical calculus introduced in [9] and extend it to Gaussian graphical calculus. We use this to obtain new nonasymptotic results for the moments of some single random channels. We obtain further asymptotic results in the single random channel setting, and we then return to the random matrix model introduced in the bi-channel setting by Hayden and Winter, computing the asymptotics of the subleading eigenvalues.

2. Wick calculus and Weingarten calculus.

In this section we recall known results which allow the computation of expectations against Gaussian measures and Haar measures on unitary groups, as well as some standard facts in free probability theory.

2.1. *Wick calculus*.

A *Gaussian space* V is a real vector space of random variables with moments of all orders such that each of these random variables are centered Gaussian distributions. Such a Gaussian space comes with a positive symmetric bilinear form $(x, y) \to \mathbb{E}[xy]$. Gaussian spaces are in one-to-one correspondence with Euclidean spaces, and isomorphisms of Gaussian spaces correspond to the notion of isomorphisms of Euclidean spaces. In particular, the Euclidean norm of a random variable determines it fully (via its variance) and if two random variables are given, then their joint distribution is determined by their angle. The following is usually called the Wick lemma.

LEMMA 2.1. *Let V be a Gaussian space and $x_1, \dots, x_k$ be elements in V. If $k = 2l + 1$, then $\mathbb{E}[x_1 \cdots x_k] = 0$, and if $k = 2l$, then*

$$\mathbb{E}[x_1 \cdots x_k] = \sum_{\substack{p=\{\{i_1,j_1\},\dots,\{i_l,j_l\}\} \\ \text{pairing of } \{1,\dots,k\}}} \prod_{m=1}^{l} \mathbb{E}[x_{i_m} x_{j_m}]. \tag{2}$$

In particular, it follows that if $x_1, \dots, x_p$ are independent standard Gaussian random variables, then

$$\mathbb{E}[x_1^{k_1} \cdots x_p^{k_p}] = \prod_{i=1}^{p} (2k_i)!!.$$

For a proof see, for instance, [29]. It is possible to extend the notion of a Gaussian space to a complex Gaussian space. A complex-valued vector space V is called a Gaussian space if and only if for any real structure on V, the pair $(\mathrm{Re}(V), \mathrm{Im}(V))$ is a real-valued Gaussian space. It can be readily checked that in the case of a complex Gaussian space, the Wick Lemma 2.1 holds with exactly the same statement.

We will usually denote by $G_{n,m}$ (or G when there is no ambiguity) the standard complex Gaussian random matrix $n \times m$. It has the distribution $\exp(-N \times \mathrm{Tr}(GG^*))\, dG$, where dG is the Lebesgue measure on the space of the $n \times m$ complex matrices properly rescaled, and $G^* = \overline{G}^t$ is the standard algebraic adjoint operator.

Since we shall mostly be concerned with traces of products of random matrices in this paper, we need to introduce one last item of notation for generalized traces, which we borrow from [6]. For some matrices $A_1, A_2, \dots, A_s \in \mathcal{M}_n(\mathbb{C})$, some permutation $\sigma \in \mathcal{S}_p$ and some function $t : \{1, \dots, p\} \to \{1, \dots, s\}$, we define

$$\mathrm{Tr}_{\sigma,t}(A_1, \dots, A_s) = \prod_{c \in \mathcal{C}(\sigma)} \mathrm{Tr}\Biggl(\overrightarrow{\prod_{j \in c}} A_{t(j)}\Biggr),$$

where $\mathcal{C}(\sigma)$ is the set of cycles of σ. When $s = p$, we use the simplified notation $\mathrm{Tr}_{\sigma,t}(A_1, \dots, A_p) = \mathrm{Tr}_{\sigma,\mathrm{id}}(A_1, \dots, A_p)$. We also put $\mathrm{Tr}_\sigma(A) = \mathrm{Tr}_\sigma(A, A, \dots, A)$.

2.2. *Weingarten calculus.* In this section, we recall a few facts about Weingarten calculus.

DEFINITION 2.2. The unitary Weingarten function $\mathrm{Wg}(n, \sigma) : \mathbb{N} \times \bigcup_{p \in \mathbb{N}^*} \mathcal{S}_p \to \mathbb{R}$ is a function of a dimension parameter n and of a permutation σ in the symmetric group $\mathcal{S}_p$. It is the pseudo-inverse of the function $\sigma \mapsto n^{\#\sigma}$

under the convolution for the symmetric group ($\#\sigma$ denotes the number of cycles of the permutation σ).

Notice that the function $\sigma \mapsto n^{\#\sigma}$ is invertible as $n \geq p$. We refer to [13] for historical references and further details. We shall use the shorthand notation $\mathrm{Wg}(\sigma) = \mathrm{Wg}(n, \sigma)$ when the dimension parameter n is obvious.

The function Wg is used to compute integrals with respect to the Haar measure on the unitary group.

THEOREM 2.3. *Let n be a positive integer and $(i_1, \dots, i_p)$, $(i'_1, \dots, i'_p)$, $(j_1, \dots, j_p)$, $(j'_1, \dots, j'_p)$ be p-tuples of positive integers from $\{1, 2, \dots, n\}$. Then,*

$$\begin{aligned}
&\int_{\mathcal{U}(n)} U_{i_1 j_1} \cdots U_{i_p j_p} \overline{U_{i'_1 j'_1}} \cdots \overline{U_{i'_p j'_p}} \, dU \\
&\quad = \sum_{\sigma, \tau \in \mathcal{S}_p} \delta_{i_1 i'_{\sigma(1)}} \cdots \delta_{i_p i'_{\sigma(p)}} \delta_{j_1 j'_{\tau(1)}} \cdots \delta_{j_p j'_{\tau(p)}} \mathrm{Wg}(n, \tau\sigma^{-1}).
\end{aligned} \tag{3}$$

If $p \neq p'$, then

$$\int_{\mathcal{U}(n)} U_{i_1 j_1} \cdots U_{i_p j_p} \overline{U_{i'_1 j'_1}} \cdots \overline{U_{i'_{p'} j'_{p'}}} \, dU = 0. \tag{4}$$

We are interested in the values of the Weingarten function in the limit $n \to \infty$. The following result encloses all the data we need for our computations relating to the asymptotics of the Wg function; see [8] for a proof.

THEOREM 2.4. *For a permutation $\sigma \in \mathcal{S}_p$, let $\mathrm{Cycles}(\sigma)$ denote the set of cycles of σ. Then,*

$$\mathrm{Wg}(n, \sigma) = \prod_{c \in \mathrm{Cycles}(\sigma)} \mathrm{Wg}(n, c)\bigl(1 + O(n^{-2})\bigr) \tag{5}$$

and

$$\mathrm{Wg}(n, (1, \dots, d)) = (-1)^{d-1} c_{d-1} \prod_{-d+1 \leq j \leq d-1} (n - j)^{-1}, \tag{6}$$

where $c_i = \frac{(2i)!}{(i+1)!i!}$ is the ith Catalan number.

A shorthand for this theorem is the introduction of a function Mob on the symmetric group, invariant under conjugation and multiplicative over the cycles, satisfying, for any permutation $\sigma \in \mathcal{S}_p$,

$$\mathrm{Wg}(n, \sigma) = n^{-(p+|\sigma|)}\bigl(\mathrm{Mob}(\sigma) + O(n^{-2})\bigr), \tag{7}$$

where $|\sigma| = p - \#\sigma$ is the *length* of σ, that is, the minimal number of transpositions that multiply to σ. We refer to [13] for details about the function Mob.

2.3. *Elementary review of noncommutative and free probability theory.* A *noncommutative probability space* is an algebra $\mathcal{A}$ with unit endowed with a tracial state φ. An element of $\mathcal{A}$ is called a (noncommutative) random variable. In this paper we shall be mostly concerned with the noncommutative 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 $L^{\infty-}(\Omega,\mathbb{P}) = \bigcap_{p\geq 1} L^p(\Omega,\mathbb{P})$].

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

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

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

The following result is from [24] and will be crucial for us. In what follows, $NC(p)$ denotes the set of noncrossing partitions on p elements, endowed with the reversed refinement partial order (see [24], Lecture 9), which makes it into a lattice.

LEMMA 2.5. *The function $d(\sigma,\tau) = |\sigma^{-1}\tau|$ is an integer-valued distance on $\mathcal{S}_p$. Further, it has the following properties*:

- *the diameter of $\mathcal{S}_p$ is $p-1$*;
- *$d(\cdot,\cdot)$ is left and right translation invariant*;
- *for three permutations $\sigma_1,\sigma_2,\tau\in\mathcal{S}_p$, the quantity $d(\tau,\sigma_1)+d(\tau,\sigma_2)$ has the same parity as $d(\sigma_1,\sigma_2)$*;
- *the set of geodesic points* (*elements which saturate the triangle inequality*) *between the identity permutation* id *and some permutation $\sigma\in\mathcal{S}_p$ is in bijection with the set of noncrossing partitions smaller than π, where the partition π encodes the cycle structure of σ. Moreover, the preceding bijection preserves the lattice structure.*

We finish by collecting the bare minimum of free probability theory results needed for the development of the main results of this paper. We skip the definition of freeness, as we will not need it. Free cumulants are multilinear maps indexed by noncrossing partitions $\sigma\in NC(p)$ on p elements

$$\kappa_\sigma : \underbrace{A\times\cdots\times A}_{p\text{ times}} \to \mathbb{C}$$

such that

$$\sum_{\pi \le \sigma \in NC(p)} \kappa_\pi(x_1, \dots, x_p) = \mathbb{E}_\sigma[x_1, \dots, x_p] \tag{8}$$

for all noncrossing partitions $\sigma \in NC(p)$, where $\mathbb{E}_\sigma[x_1, \dots, x_p]$ is the product over the blocks $\{x_{i_1}, \dots, x_{i_j}\}$ of σ, of $\mathbb{E}[x_{i_1}, \dots, x_{i_j}]$. Cumulants are known to be multiplicative over blocks and therefore a special role is played by the cumulant corresponding to the maximal partition $\mathbf{1}_p$, which we denote by $\kappa(a_1, \dots, a_p) := \kappa_{\mathbf{1}_p}(a_1, \dots, a_p)$.

We will need free cumulants for computational purposes, in order to identify free Poisson distributions. Let us mention, for the benefit of the interested reader, that the main property of the free cumulants is that mixed cumulants of free variables vanish.

We recall that the *free Poisson distribution* of parameter c is given by

$$\pi_c = \max(1-c, 0)\delta_0 + \frac{\sqrt{4c - (x-1-c)^2}}{2\pi x} \mathbf{1}_{[1+c-2\sqrt{c}, 1+c+2\sqrt{c}]}(x)\, dx.$$

It is characterized by the fact that all its free cumulants are equal to c. Although we will not need this fact, it is worth mentioning that it has a semigroup structure with respect to the additive free convolution of Voiculescu (see, e.g., [24]). It is also sometimes called the Marchenko–Pastur distribution. One can compute (minus) the entropy of this probability distribution:

$$K_c = \int x \log x \, d\pi_c(x) = \begin{cases} \dfrac{1}{2} + c \log c, & \text{if } c \ge 1, \\ \dfrac{c^2}{2}, & \text{if } 0 < c < 1. \end{cases} \tag{9}$$

3. Unitary and Gaussian graphical calculi. In this section we briefly recall the results of [9] for the convenience of the reader and in order to make the paper self-contained. We then introduce the Gaussian graphical calculus and present a first application of it to Wishart matrices.

3.1. *Axioms of unitary graphical calculus.* The purpose of the graphical calculus introduced in [9] is to yield an effective method to evaluate the expectations of random tensors with respect to the Haar measure on a unitary group. The tensors under consideration can be constructed from a few elementary tensors such as the Bell state, fixed kets and bras, and random unitary matrices. In graphical language, a tensor corresponds to a *box*, and an appropriate Hilbertian structure yields a correspondence between boxes and tensors. However, the calculus yielding expectations only relies on diagrammatic operations.

Each box B is represented as a rectangle with decorations on its boundary. The decorations are either white [elements of the set of white decorations $S(B)$] or

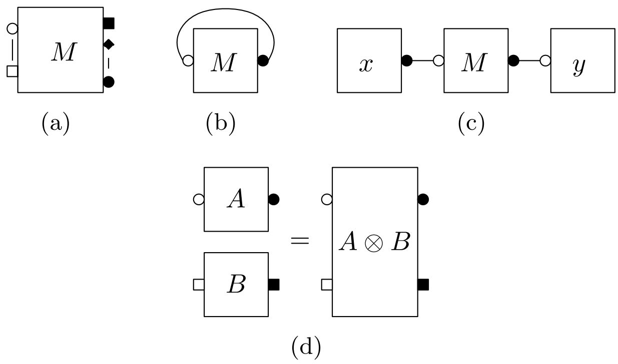

FIG. 1. *Basic diagrams and axioms.*

black [elements of the dual set of black decorations, $S^*(B)$]. In the Hilbertian picture, decorations correspond to complex vector spaces, dual decorations being associated to dual spaces. Figure 1 depicts an example of a box.

It is possible to construct new boxes from old ones by formal algebraic operations such as sums or products. We call a *diagram* a picture consisting of boxes and wires according to the following rule: a wire may link a white decoration in $S(B)$ to its black counterpart in $S^*(B)$. A diagram can be turned into a box by choosing an orientation and a starting point.

Regarding the Hilbertian structure, wires correspond to tensor contractions. There exists an involution for boxes and diagrams. It is antilinear and turns a decoration in $S(B)$ into its counterpart in $S^*(B)$. Our conventions are close to those of [7, 21], and we hope that they are familiar to the reader acquainted with existing graphical calculi of various types (planar algebra theory, Feynman diagram, traced category theory). Our notation is designed to conform well to the problem of computing expectations, as shown in the next section. In Figure 1(b)–(d) we depict the trace of a matrix, multiplication of tensors and the tensor product operation, respectively. For details, we refer to [9].

3.2. *Planar expansion.* In this subsection we describe the main application of our calculus. For this, we need a concept of *removal* of boxes U and $\overline{U}$. A removal r is a way to pair decorations of the U and $\overline{U}$ boxes appearing in a diagram. It therefore consists of a pairing α of the white decorations of U boxes with the white decorations of $\overline{U}$ boxes, together with a pairing β between the black decorations of U boxes and the black decorations of $\overline{U}$ boxes. Assuming that $\mathcal{D}$ contains p boxes of type U and that the boxes U (resp., $\overline{U}$) are labeled from 1 to p, then $r = (\alpha, \beta)$, where α, β are permutations of $\mathcal{S}_p$. The set of all removals of U and $\overline{U}$ boxes is denoted by $\mathrm{Rem}_U(\mathcal{D})$.

Given a removal $r \in \mathrm{Rem}_U(\mathcal{D})$ we construct a new diagram $\mathcal{D}_r$ associated with r, one which has the important property that it no longer contains boxes of type U or $\overline{U}$. We start by erasing the boxes U and $\overline{U}$, but keep the decorations attached to them. Assuming that we have labeled the erased boxes U and $\overline{U}$ with integers from $\{1, \dots, p\}$, we connect *all* the (inner parts of the) *white* decorations of the ith erased U box with the corresponding (inner parts of the) *white* decorations of the $\alpha(i)$th erased $\overline{U}$ box. In a similar manner, we use the permutation β to connect black decorations.

In [9], we proved the following result.

THEOREM 3.1.

$$\mathbb{E}_U(\mathcal{D}) = \sum_{r=(\alpha,\beta)\in \mathrm{Rem}_U(\mathcal{D})} \mathcal{D}_r \, \mathrm{Wg}(n, \alpha\beta^{-1}).$$

3.3. *Gaussian planar expansion.* We now consider the case where we allow a new special box G in our diagrams, corresponding to a *Gaussian random matrix*. We shall address the same issue as in the unitary case: computing the expected value of a random diagram with respect to the Gaussian probability measure.

To begin, consider $\mathcal{D}$, a diagram which contains, among other constant tensors, boxes corresponding to independent Gaussian random matrices of *covariance one* (identity). We can deal with more general Gaussian matrices by multiplying the standard ones by constant matrices. Note that a box can appear several times, adjoints of boxes are allowed and the diagram may be disconnected. Also, Gaussian matrices need not be square.

The expectation value of such a random diagram $\mathcal{D}$ can be computed by a *removal* procedure, as in the unitary case. Without loss of generality, we assume that we do not have adjoints of Gaussian matrices in our diagram, but instead their complex conjugate box. This assumption allows for a more straightforward use of the Wick Lemma 2.1. As in the unitary case, we can assume that $\mathcal{D}$ contains only one type of random Gaussian box G; the other independent random Gaussian matrices are assumed to be constant at this stage as they shall afterward be removed in the same manner.

A removal of the diagram $\mathcal{D}$ is a pairing between *Gaussian boxes* G and their conjugates $\overline{G}$. The set of removals is denoted by $\mathrm{Rem}_G(\mathcal{D})$, and it may be empty: if the number of G boxes is different from the number of $\overline{G}$ boxes, then $\mathrm{Rem}_G(\mathcal{D}) = \varnothing$ [this is consistent with the first case of the Wick formula (2)]. Otherwise, a removal r can identified with a permutation $\alpha \in \mathcal{S}_p$, where p is the number of G and $\overline{G}$ boxes. Let us stress here the main difference between the notion of a removal in the Gaussian and the Haar unitary cases. In the Haar unitary (or the Weingarten) case, a removal was associated with a *pair of permutations*: one had to pair white decorations of U and $\overline{U}$ boxes and, independently, black decorations of conjugate boxes. On the other hand, in the Gaussian/Wick case, one pairs conjugate boxes: white and black decorations are paired in an identical manner, hence only one permutation is needed to encode the removal.

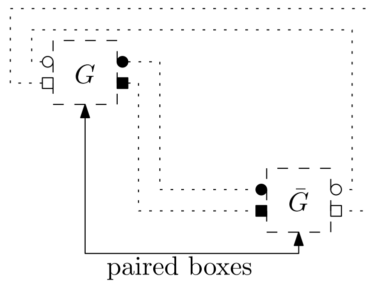

FIG. 2. *Pairing of boxes in the Gaussian case.*

To each removal r associated with a permutation $\alpha \in \mathcal{S}_p$ there corresponds a removed diagram $\mathcal{D}_r$, constructed as follows. We starts by erasing the boxes G and $\overline{G}$, but keep the decorations attached to these boxes. Then, the decorations (white *and* black) of the ith G box are paired with the decorations of the $\alpha(i)$th $\overline{G}$ box in a coherent manner; see Figure 2.

The graphical reformulation of the Wick Lemma 2.1 becomes the following theorem, which we state without proof.

THEOREM 3.2.

$$\mathbb{E}_G[\mathcal{D}] = \sum_{r \in \mathrm{Rem}_G(\mathcal{D})} \mathcal{D}_r.$$

3.4. *Moments of Wishart matrices.* As a first application of our Gaussian graphical calculus, we compute the moments of traces of products of Wishart matrices. By definition, a *Wishart matrix* of parameters (n, k) is a positive random matrix $W \in \mathcal{M}_n(\mathbb{C})$ such that

$$W = G \cdot G^*,$$

where $G \in \mathcal{M}_{n \times k}(\mathbb{C})$ is a standard Gaussian random matrix. In our graphical formalism, since we only consider Gaussian random matrices, the previous equation corresponds to the graphical substitution in Figure 3; round decorations correspond to n-dimensional complex Hilbert spaces $\mathbb{C}^n$ and square-shaped labels correspond to $\mathbb{C}^k$.

The same problem of computing expected values of traces of Wishart matrices was considered in [6, 16, 18, 22], and we shall rederive Corollary 3 of Theorem 2 from [6]. The general covariance case (Theorem 2 in [6]) can be easily derived from the result below.

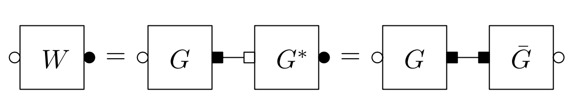

FIG. 3. *Diagram of a Wishart matrix.*

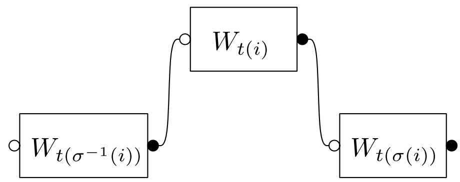

FIG. 4. *Monomials of traces of Wishart matrices.*

PROPOSITION 3.3. *Let $W_1, W_2, \dots, W_s$ be independent Wishart matrices with unit covariance and parameters $(n, k_1), (n, k_2), \dots, (n, k_s)$, respectively. For a permutation $\sigma \in \mathcal{S}_p$ and a function $t : \{1, \dots, p\} \to \{1, \dots, s\}$, we have*

$$\mathbb{E}[\mathrm{Tr}_{\sigma,t}(W_1, \dots, W_s)] = \sum_{\alpha \in \mathcal{S}_p(t)} \prod_{j=1}^{s} k_j^{\#\alpha_j} n^{\#(\sigma^{-1}\alpha)}, \tag{10}$$

where $\mathcal{S}_p(t) = \{\alpha \in \mathcal{S}_n \mid t = t \circ \alpha\}$. Every permutation $\alpha \in \mathcal{S}_p(t)$ leaves the level sets of t invariant and induces on each set $t^{-1}(j)$ a permutation α_j $(j = 1, \dots, s)$.

PROOF. We consider the diagram $\mathcal{D}$ corresponding to the left-hand side of equation (10). It contains n Wishart boxes from the set $\{W_1, \dots, W_s\}$ which are wired according to the permutation σ (see Figure 4). Computing the expectation of the diagram $\mathcal{D}$ is rather straightforward using our graphical calculus. Since we are dealing with s independent Gaussian matrices $G_1, \dots, G_s$ (recall that $W_j = G_j G_j^*$), we need to apply Theorem 3.2 s times, once for each Gaussian matrix G_j. Each box G_j appears $|t^{-1}(j)|$ times and, using Theorem 3.2, we get

$$\mathbb{E}[\mathcal{D}] = \sum \mathcal{D}_{\alpha_1,\dots,\alpha_s},$$

where each permutation $\alpha_j \in \mathcal{S}_{|t^{-1}(j)|}$ encodes the removal procedure for the G_j boxes.

Diagrams obtained after the successive removal procedures $\mathcal{D}_{\alpha_1,\dots,\alpha_s}$ are made of loops of two types: loops associated with the n-dimensional space $\mathbb{C}^n$ and loops associated with "internal spaces" $\mathbb{C}^{k_j}$. In order to count the number of loops of each dimensionality, let us first observe that the set of s-tuples of permutations $(\alpha_1, \dots, \alpha_s)$ is in bijection with the set of permutations $\alpha \in \mathcal{S}_p(t)$ defined in the statement of the theorem.

For such a permutation $\alpha \in \mathcal{S}_p(t)$, let us count the number of loops corresponding to traces over $\mathbb{C}^{k_j}$. Initially, the p_j decorations of the G_j boxes are connected in the simplest manner: the k_j decoration of the ith G_j box is connected to the corresponding decoration of the $\overline{G_j}$ box with the same index i. The jth removal procedure, encoded by the permutation α_j, then produces a number of $\#(\mathrm{id}^{-1} \alpha_j) = \#\alpha_j$ loops. Hence, the contribution of the $\mathbb{C}^{k_j}$-type loops is $k_j^{\#\alpha_j}$.

The computation of the loops associated with $\mathbb{C}^n$ is more involved since the decorations are already nontrivially linked by the permutation σ. Since σ may not respect the level sets of the function t, we need to consider the global action of α, the restrictions α_j not being sufficient in this case. Since the boxes are initially connected by σ and the removal procedures add wires according to the permutation α, the total number of loops is $\#(\sigma^{-1}\alpha)$. Adding all loop contributions, we obtain the announced formula (10). □

REMARK 3.4. We can consider more general covariances in the graphical model and obtain Theorem 2 of [6] in its full generality. All there is to be done is to add constant tensors associated with covariance matrices in our diagrams. After the successive removal procedures, we are left with loops *and* traces of monomials in these constant matrices. Since our purpose in this section was to illustrate the Gaussian graphical calculus, we leave the details of this more technical generalization to the interested reader.

4. Application of Gaussianization: Pure states through random quantum channels.

4.1. *Single random channel model.* In this section we present an important application of the Gaussian diagrammatic calculus: we compute eigenvalue statistics for the action of a random quantum channel on a pure quantum state. By definition, a *quantum channel* $\Phi : \mathcal{M}_n(\mathbb{C}) \to \mathcal{M}_n(\mathbb{C})$ is a trace-preserving, completely positive map. According to the Stinespring theorem, such a linear application can be written as

$$\Phi(X) = \Phi^{U,Y}(X) = \mathrm{Tr}_k[U(X \otimes Y)U^*],$$

where U is a unitary matrix in $\mathcal{U}(nk)$ and Y is a k-dimensional rank-one projector. A diagrammatic representation of the above formula is presented in Figure 5. The set of quantum channels can be endowed with a natural probability measure by fixing the projection Y and picking U uniformly with respect to the Haar measure on the unitary group $\mathcal{U}(nk)$. This is the model of randomness we refer to when we speak or random quantum channels, and it has received a lot of attention from the quantum information community [9, 20]. From the definition of Φ we can see that the Weingarten calculus developed in [9] may be applied to this situation since random unitary matrices are a key element in the problem. However, when random

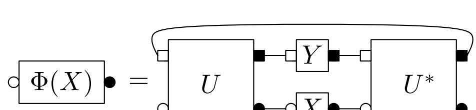

FIG. 5. *Diagram for a quantum channel.*

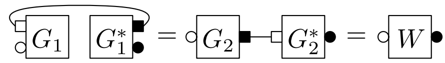

FIG. 6. *An equivalent diagram for quantum channels with rank one X and Y.*

quantum channels are presented with rank-one inputs (or pure states), we show that the simpler Gaussian calculus can be used, see Figure 6. Using this approach, we shall recover some exact formulas for the moments of the output from [9], as well as some asymptotic results from [23].

We are interested in the output random matrix

$$Z = \Phi^{U,Y}(X), \tag{11}$$

where X is a rank-one projector. The main result, obtained in [23], is as follows.

PROPOSITION 4.1. *Let $W = G \cdot G^* \in \mathcal{M}_n(\mathbb{C})$ be a Wishart matrix with parameters (n, k). Then,*

$$Z = \Phi(X) = W/\operatorname{Tr}(W).$$

Observe that this result does not depend on the choice of X, Y due to the invariance of the Haar measure.

The main point is that we can show (see [23]) that the eigenvalues of Z, that is, the normalized eigenvalues of W, are independent of the trace of W. This implies that we we can apply the results on Wishart matrices developed in Section 3.4 to this particular case.

4.2. *Exact moments.* In this section we provide exact formulas for the moments $\mathbb{E}[\operatorname{Tr}(Z^p)]$ of the output of a random quantum channel. Other formulas for the same quantities (as well as some recursion relations) have been obtained in [23, 26, 28].

Using the Gaussianization trick, we have

$$\mathbb{E}[\operatorname{Tr}(Z^p)] = \frac{\mathbb{E}[\operatorname{Tr}(W^p)]}{\mathbb{E}[\operatorname{Tr}(W)^p]},$$

where W is a Wishart matrix with parameters (n, k). One uses Proposition 3.3 to compute $\mathbb{E}[\operatorname{Tr}(W^p)]$ and $\mathbb{E}[\operatorname{Tr}(W)^p]$:

$$\mathbb{E}[\operatorname{Tr}(W^p)] = \sum_{\alpha \in \mathcal{S}_p} k^{\#\alpha} n^{\#(\gamma^{-1}\alpha)},$$

$$\mathbb{E}[\operatorname{Tr}(W)^p] = \sum_{\alpha \in \mathcal{S}_p} (nk)^{\#\alpha},$$

where $\gamma = (p\ p-1\ \cdots\ 2\ 1) \in \mathcal{S}_p$ is the full cycle. In the second formula above, we recognize the generating polynomial for the number of cycles of a permutation of

p objects evaluated at nk. This is known to be equal to $nk(nk+1)(nk+2)\cdots(nk+p-1)$ (see [27], Proposition 1.3.4), and we obtain the following theorem.

THEOREM 4.2.

$$\mathbb{E}[\operatorname{Tr}(Z^p)] = \left(\prod_{j=0}^{p-1}(nk+j)\right)^{-1} \sum_{\alpha\in\mathcal{S}_p} k^{\#\alpha} n^{\#(\gamma^{-1}\alpha)}. \tag{12}$$

This is exactly like formula (10) from [9], which was obtained via the Weingarten formula. The approach followed here is more straightforward and does not use unitary integration: It is based on the purely combinatorial Wick formula and the Gaussianization trick.

4.3. *Asymptotics*. We now look at the probability distribution of the output random matrix Z when one (or both) of the parameters n and k grow to infinity. The asymptotic behavior of random matrices has been one of the main objects of study in random matrix theory. For instance, it is in this large-dimension regime that the freeness phenomenon appears. In the particular case of random quantum channels under study here, this question has an interesting physical motivation: large-dimensional Hilbert spaces model physical systems with large numbers of degrees of freedom. This point of view has been discussed in the quantum information theory literature (see [2, 5, 23, 28]). Although some of what follows has already been treated in [23], the approach of this paper has the merit of being self-contained and illustrates perfectly the power and range of the Gaussian graphical calculus.

We split the results according to three possible asymptotic regimes, depending on which of the parameters n and/or k is large. Of special interest is the third regime, when *both* parameters grow to infinity, but at a constant positive ratio $c > 0$. We use the equivalence symbol $x(n) \sim y(n)$ for nonzero sequences $x(n)$ and $y(n)$ which are such that $x(n)/y(n) \to 1$ when $n \to \infty$.

THEOREM 4.3. *Let $Z = \Phi^{U,Y}(X)$ denote the output of a random quantum channel Φ, where X and Y are rank-one projectors.*

(I) *In the regime where n is fixed and $k \to \infty$, the limiting spectral distribution of Z is almost surely $\delta_{1/n}$.*
(II) *In the regime where k is fixed and $n \to \infty$, Z tends almost surely to a variable that has eigenvalues $1/k$ with multiplicity k and 0 with multiplicity $n-k$.*
(III) *In the regime where $n, k \to \infty$, $k/n \to c > 0$, cnZ converges almost surely to a free Poisson distribution with parameter c.*

PROOF. In the first regime,

$$\mathbb{E}[\operatorname{Tr}(Z^p)] \overset{k\to\infty}{\sim} \frac{1}{n}(nk)^{-p} \sum_{\alpha\in\mathcal{S}_p} k^{\#\alpha} n^{\#(\gamma^{-1}\alpha)}.$$

Permutations α which give nonvanishing contributions are those such that $\#\alpha = p$, hence $\alpha = \mathrm{id}$. In the end, we obtain

$$\lim_{k\to\infty} \mathbb{E}[\mathrm{Tr}(Z^p)] = n^{1-p},$$

hence the limiting spectral distribution of Z is $\delta_{1/n}$.

In order to prove the almost sure convergence, we show that the empirical measures

$$\mu_{n,k}(Z) = \frac{1}{n}\sum_{i=1}^{n} \lambda_i(Z)$$

converge almost surely to the limit $\delta_{1/n}$ (which is equivalent to the fact that, almost surely, every eigenvalue of Z converges to $1/n$—recall that n is fixed). As usual, almost sure convergence of moments suffices and we aim to prove that for all p,

$$\text{a.s.} \qquad \lim_{k\to\infty} \mathrm{Tr}(Z^p) = n^{1-p}.$$

A standard application of Chebyshev's inequality and the Borel–Cantelli lemma shows that it is enough to verify that for all integers p, the series of variances is summable:

$$\sum_{k=1}^{\infty} \mathbb{E}\big[\big(\mathrm{Tr}(Z^p) - \mathbb{E}\,\mathrm{Tr}(Z^p)\big)^2\big] < \infty.$$

Let us separately compute $\mathbb{E}[\mathrm{Tr}(Z^p)^2]$ and $\mathbb{E}[\mathrm{Tr}(Z^p)]^2$ using formula (12). For the first expectation, we need to introduce the permutation

$$\gamma_2 = \big(p\ (p-1)\ \cdots\ 2\ 1\big)\big(2p\ (2p-1)\ \cdots\ (p+2)\ (p+1)\big) \in \mathcal{S}_{2p}. \tag{13}$$

We then have

$$\begin{aligned}\mathbb{E}[\mathrm{Tr}(Z^p)^2] &= \left(\prod_{j=0}^{2p-1}(nk+j)\right)^{-1} \sum_{\alpha\in\mathcal{S}_{2p}} k^{\#\alpha} n^{\#(\gamma_2^{-1}\alpha)} \\ &= \left(\prod_{j=0}^{2p-1}\left(1+\frac{j}{nk}\right)\right)^{-1} \sum_{\alpha\in\mathcal{S}_{2p}} k^{-|\alpha|} n^{-|\gamma_2^{-1}\alpha|}.\end{aligned}$$

The first contribution (of order k^0) in the last sum is given by $\alpha = \mathrm{id}$ and is equal to n^{2-2p} (recall that γ_2 has two cycles). The second-order in k is given by transpositions $\alpha = (ij)$. In this case, $|\gamma_2^{-1}\alpha| = 2p-3$ if i and j belong to the same cycle of γ_2 and $|\gamma_2^{-1}\alpha| = 2p-1$ otherwise. Hence, we obtain

$$\begin{aligned}\mathbb{E}[\mathrm{Tr}(Z^p)^2] &= \left[1 - \frac{2p(2p-1)}{2nk} + O\left(\frac{1}{k^2}\right)\right] \\ &\quad \times \left[n^{2-2p} + \frac{1}{k}\big(p^2 n^{1-2p} + p(p-1)n^{3-2p}\big) + O\left(\frac{1}{k^2}\right)\right] \\ &= n^{2-2p} + \frac{1}{k}p(p-1)n^{1-2p}(n^2-1) + O\left(\frac{1}{k^2}\right).\end{aligned}$$

Using the same ideas, $\mathbb{E}[\mathrm{Tr}(Z^p)]^2$ is easily computed:

$$\begin{aligned}
&\mathbb{E}[\mathrm{Tr}(Z^p)]^2 \\
&\quad = \left(\prod_{j=0}^{2p-1}(nk+j)\right)^{-2}\left(\sum_{\alpha\in\mathcal{S}_p} k^{\#\alpha} n^{\#(\gamma^{-1}\alpha)}\right)^2 \\
&\quad = \left[1-\frac{p(p-1)}{2nk}+O\left(\frac{1}{k^2}\right)\right]^2\cdot\left[n^{1-p}+\frac{1}{k}\frac{p(p-1)}{2}n^{2-p}+O\left(\frac{1}{k^2}\right)\right]^2 \\
&\quad = n^{2-2p}+\frac{1}{k}p(p-1)n^{1-2p}(n^2-1)+O\left(\frac{1}{k^2}\right),
\end{aligned}$$

and we conclude that $\mathbb{E}[\mathrm{Tr}(Z^p)^2-\mathbb{E}[\mathrm{Tr}(Z^p)]^2 = O(k^{-2})$. Thus, the covariance series converges, completing the proof.

In the second regime,

$$\mathbb{E}[\mathrm{Tr}(Z^p)] \overset{n\to\infty}{\sim} \sum_{\alpha\in\mathcal{S}_p} k^{-|\alpha|} n^{-|\gamma^{-1}\alpha|}.$$

The nonvanishing contribution is given by $\alpha=\gamma$ and thus

$$\lim_{n\to\infty}\mathbb{E}[\mathrm{Tr}(Z^p)] = k^{1-p}.$$

In other words, for large n, Z has the following eigenvalues:

- $1/k$ with multiplicity k;
- 0 with multiplicity $n-k$.

The proof of the almost sure convergence follows the same lines as in the previous case and is left to the reader.

In the third regime, after making the substitution $k=cn$, the asymptotics are

$$\mathbb{E}[\mathrm{Tr}(Z^p)] \sim n^{-2p}c^{-p}\sum_{\alpha\in\mathcal{S}_p} c^{\#\alpha} n^{\#\alpha+\#(\gamma^{-1}\alpha)}. \tag{14}$$

Since

$$\#\alpha+\#(\gamma^{-1}\alpha) = 2p-(|\alpha|+|\gamma^{-1}\alpha|)\le p+1, \tag{15}$$

we should rescale the matrix Z by a factor of n. In fact, in order to avoid some unnecessary complications, we shall rescale Z by cn. We get

$$\mathbb{E}[\mathrm{tr}_n((cnZ)^p)] \sim n^{-p-1}\sum_{\alpha\in\mathcal{S}_p} c^{\#\alpha} n^{\#\alpha+\#(\gamma^{-1}\alpha)}.$$

Contributing permutations are those for which we have equality in equation (15), that is, $|\alpha|+|\gamma^{-1}\alpha| = |\gamma| = p-1$. These are permutations on the geodesic

$\mathrm{id} \to \gamma$ and are known to be in bijection with noncrossing partitions $\sigma \in NC(p)$. Thus,

$$\mathbb{E}[\mathrm{tr}_n((cnZ)^p)] \sim \sum_{\sigma \in NC(p)} c^{\#\sigma}.$$

One recognizes the moment–cumulant formula from free probability theory. Hence, the limiting distribution of cnZ has cumulants of all orders equal to c and we identify the free Poisson distribution of parameter c. Let us now show that almost sure convergence holds:

$$\lim_{n\to\infty} \mathbb{E}[\mathrm{tr}_n((cnZ)^p)] = \sum_{\sigma \in NC(p)} c^{\#\sigma} \qquad \text{almost surely.}$$

Using the same classical technique as in the first regime, we show that the series

$$\sum_n \bigl(\mathbb{E}[(\mathrm{tr}_n((cnZ)^p)^2)] - \mathbb{E}[\mathrm{tr}_n((cnZ)^p)]^2\bigr)$$

converges. We start by evaluating $\mathbb{E}[(\mathrm{tr}_n((cnZ)^p)^2)]$ up to the second-order in n. Using the permutation γ_2 defined in (13) and the Gaussian graphical calculus, we have

$$\mathbb{E}[(\mathrm{tr}_n((cnZ)^p)^2)] \sim \sum_{\alpha \in \mathcal{S}_{2p}} c^{\#\alpha} n^{2p-2-(|\alpha|+|\gamma_2^{-1}\alpha|)}.$$

Using similar ideas as before, $|\alpha| + |\gamma_2^{-1}\alpha| \geq |\gamma_2| = 2p-2$ with equality iff α is on the geodesic between id and γ_2. Given the 2-cycle structure of γ_2, geodesic permutations α admit a decomposition $\alpha = \alpha' + \alpha''$, where $\alpha' \in \mathcal{S}\{1, 2, \dots, p\} = \mathcal{S}_p$ and $\alpha'' \in \mathcal{S}\{p+1, p+2, \dots, 2p\} \simeq \mathcal{S}_p$ are themselves geodesic permutations $\mathrm{id} \to \alpha' \to \gamma$ and $\mathrm{id} \to \alpha'' \to \gamma$, respectively. Of course, in this case, $\#\alpha = \#\alpha' + \#\alpha''$ and thus

$$\mathbb{E}[(\mathrm{tr}_n((cnZ)^p)^2)] \sim \sum_{\substack{\mathrm{id}\to\alpha'\to\gamma \\ \mathrm{id}\to\alpha''\to\gamma}} c^{\#\alpha'+\#\alpha'} = \Bigl(\sum_{\mathrm{id}\to\tilde{\alpha}\to\gamma} c^{\#\tilde{\alpha}}\Bigr)^2.$$

By a standard parity argument, the function $\mathcal{S}_{2p} \ni \alpha \mapsto (|\alpha| + |\gamma_2^{-1}\alpha|) \bmod 2$ is constant and thus there is no n^{-1} term in the asymptotic development of $\mathbb{E}[(\mathrm{tr}_n((cnZ)^p)^2)]$:

$$\mathbb{E}[(\mathrm{tr}_n((cnZ)^p)^2)] = \Bigl(\sum_{\mathrm{id}\to\tilde{\alpha}\to\gamma} c^{\#\tilde{\alpha}}\Bigr)^2 + O(n^{-2}).$$

Similar ideas applied to formula (14) yield the same conclusion:

$$\mathbb{E}[\mathrm{tr}_n((cnZ)^p)] = \sum_{\mathrm{id}\to\alpha\to\gamma} c^{\#\alpha} + O(n^{-2}).$$

Taking the square of this last equation and comparing it with the previous one, we conclude that the general term of the covariance series behaves asymptotically as $O(n^{-2})$. This implies that the series is convergent and we conclude that the almost sure convergence holds. □

Even though Gaussianization results are exact and do not require a detour through Weingarten calculus, it is not clear how to apply them when the input is not one-dimensional. However, it is natural to wonder about the asymptotics in this case as well. The calculus that we introduced in [9] is crucial for this and that is the subject of Section 5.

4.4. *Almost sure convergence for entropies.* In this section, we improve the almost sure convergence of moments to the almost sure convergence of any continuous function with polynomial growth. Since the set of functions that it applies to is larger, this type of convergence is stronger than the weak convergence. We deduce corollaries for quantum information theory, and the techniques developed in this section will be useful toward the end of the paper. The technique of proof for this result is inspired by [17].

PROPOSITION 4.4. *Let f be a continuous function on $\mathbb{R}$ with polynomial growth and ν_n be a sequence of probability measures which converges in moments to a compactly supported measure ν. Then, $\int f\,d\nu_n \to \int f\,d\nu$.*

PROOF. Let $K > 1$ be a constant such that the interval $[-(K-1), K-1]$ contains the (compact) support of the limit measure ν. It follows that, for all integer powers $s \geq 0$,

$$\lim_{r\to\infty} K^{-2r} \int x^{2r+2s}\,d\nu(x) = 0. \tag{16}$$

Moreover, since the measures ν_n converge in moments to ν, for all $\varepsilon > 0$, there exists an r large enough such that for all n large enough,

$$K^{-2r} \int x^{2r+2s}\,d\nu_n(x) < \varepsilon. \tag{17}$$

For some fixed $\delta > 0$, the Weierstrass theorem produces a polynomial P such that $|f(x) - P(x)| < \delta$ for all $x \in [-K, K]$. We then have

$$\left|\int f\,d\nu_n - \int f\,d\nu\right| \leq \int |f-P|\,d\nu_n + \int |f-P|\,d\nu + \left|\int P\,d(\nu_n - \nu)\right|.$$

Since the polynomial approximation holds on the support of ν, the second term above is less than δ. Using the convergence in moments of the probability measures ν_n, the last term can be seen to be less than δ for n large enough. We focus now on the first term above, $\int |f-P|\,d\nu_n$. By the polynomial approximation, $\int |f -$

$P| \, d\nu_n \leq \delta + \int_{|x|\geq K} |f - P| \, d\nu_n$. Since f has polynomial growth, one can find a constant $q > 0$ such that $|f(x) - P(x)| \leq x^{2q}$ for all $|x| \geq K$. Using the Chebyshev inequality on the last integral, we have, for all $r \geq 1$,

$$\int_{|x|\geq K} |f - P| \leq \int_{\mathbb{R}} \frac{x^{2r}}{K^{2r}} x^{2q} \, d\nu_n = K^{-2r} \int x^{2q+2r} \, d\nu_n.$$

The convergence in moments, together with equations (16) and (17), implies that, for r and n large enough, the above expression can be made arbitrarily small, which completes the proof. □

REMARK 4.5. The conclusion of the above proposition still holds true under the weaker assumption that ν admits some finite exponential moment, thanks to the fact that weak convergence of measures implies convergence of integrals under uniform integrability. However, in this paper we only consider compactly supported measures.

COROLLARY 4.6. *Almost surely, in the regime $n \to \infty$, $k \sim cn$, the von Neumann entropy of the matrix Z from Theorem 4.3 satisfies*

$$H(Z) = \begin{cases} \log n - \dfrac{1}{2c} + o(1), & \text{if } c \geq 1, \\ \log(cn) - \dfrac{c}{2} + o(1), & \text{if } 0 < c < 1. \end{cases}$$

PROOF. Let us assume that $c \geq 1$, the other case being similar. We use Theorem 4.4 for the function $x \mapsto x \log x$, which is continuous and of polynomial growth on the domain $\mathbb{R}_+$, and for the empirical spectral measures of the matrices cnZ. It follows that, almost surely when $n \to \infty$,

$$\frac{1}{n} \sum_{i=1}^{n} cn\lambda_i \log(cn\lambda_i) = \int t \log t \, d\pi_c(t) + o(1),$$

where $\lambda_1 \geq \cdots \geq \lambda_n$ are the eigenvalues of Z.

Simplifying this expression and using the value of the right-hand side integral from equation (9), we have

$$H(Z) = -\sum_{i=1}^{n} \lambda_i \log \lambda_i = \log n - \frac{1}{2c} + o(1),$$

completing the proof. □

A formula of Page [25] states that the mean entropy of a random density matrix $Z^{(n,k)} \in \mathcal{M}_n(\mathbb{C})$ obtained by tracing out a k-dimensional environment is given by (here, $n \leq k$ are fixed)

$$\mathbb{E}H(Z^{(n,k)}) = \sum_{j=k+1}^{nk} \frac{1}{j} - \frac{n-1}{2k}.$$

We could obtain a weaker version of Corollary 4.6 from Page's formula by letting n tend to infinity and using the dominated convergence theorem.

5. Asymptotics of a single random quantum channel for general states.

5.1. *The model.* We are interested in single random quantum channels and study the asymptotic behavior of the output of such channels for more general input states than rank-one projectors. The Gaussian planar expansion cannot be used in the more general cases, so we need the Weingarten planar expansion. We may consider the general model

$$\mathrm{Tr}_\beta(X) \sim (n^s)^{\#\beta} u^{\#\beta} \varphi_\beta(x), \tag{18}$$

where $s, u \in \mathbb{R}$ are fixed parameters and x is a random variable in some noncommutative probability space with trace φ. In this section, we will deal only with two special cases of interest of the above formula. The first is motivated by quantum information theory: X is a rank-r projector. This choice corresponds to $s = 0$, $u = r$ and $x = r^{-1}$. The second special case we consider will seem natural to the reader with a free probabilistic background: X converges in moments to a noncommutative random variable x. To obtain this particular case from formula (18), we have to put $s = u = 1$ (this amounts to taking a normalized trace in the left-hand side). Note, however, that such an input matrix is not normalized, and we have to take into account the trace-one restriction for quantum states.

Let us recall here the formula for the moments of the output $Z = \Phi(X)$ of a random quantum channel (see [9]):

$$\mathbb{E}[\mathrm{Tr}(Z^p)] = \sum_{\alpha,\beta\in\mathcal{S}_p} k^{\#\alpha} n^{\#(\gamma^{-1}\alpha)} \mathrm{Tr}_\beta(X) \,\mathrm{Wg}(\alpha\beta^{-1}), \tag{19}$$

where γ is the full-cycle permutation $\gamma = (p\ p-1\ \cdots\ 2\ 1) \in \mathcal{S}_p$.

5.2. *Rank-r projectors.* Plugging, for all $\beta \in \mathcal{S}_p$, $\mathrm{Tr}_\beta(X) = r^{\#\beta} r^{-p} = r^{-|\beta|}$ into the previous equation, we obtain

$$\mathbb{E}[\mathrm{Tr}(Z^p)] = \sum_{\alpha,\beta\in\mathcal{S}_p} k^{\#\alpha} n^{\#(\gamma^{-1}\alpha)} r^{-|\beta|} \,\mathrm{Wg}(\alpha\beta^{-1}). \tag{20}$$

We study, as usual, the three following asymptotic regimes: n fixed, $k \to \infty$; k fixed, $n \to \infty$; $n, k \to \infty$, $k/n \to c$.

PROPOSITION 5.1. *Depending on the asymptotic regime, the almost sure behavior of Z is given as follows*:

(I) *when n is fixed and $k \to \infty$, the output density matrix Z converges almost surely to the maximally mixed state*

$$\rho_* = \frac{1}{n}\mathrm{I}_n;$$

(II) *when k is fixed and $n \to \infty$, the output density matrix Z, restricted to its support of dimension rk, converges to $1/(rk)\mathrm{I}_{rk}$;*

(III) *finally, in the third regime $k/n \to c$, the empirical spectral distribution of the matrix rkZ converges to a free Poisson distribution of parameter rc.*

PROOF. Using the Weingarten asymptotic $\mathrm{Wg}(\alpha\beta^{-1}) \sim (nk)^{-p-|\alpha\beta^{-1}|}$, the exponent of k in equation (20) is given by $\#\alpha - p - |\alpha\beta^{-1}|$. This reaches its maximum of zero when $\alpha = \beta = \mathrm{id}$. Hence, to the first-order in k, we have

$$\mathbb{E}[\mathrm{Tr}(Z^p)] = n^{1-p} + o(1),$$

and the conclusion follows.

The second regime is very similar, and we ultimately obtain (this time up to the first order in n)

$$\mathbb{E}[\mathrm{Tr}(Z^p)] = (rk)^{1-p} + o(1).$$

As for the third regime, making the substitution $k = cn$, we obtain the following asymptotic relation:

$$\mathbb{E}[\mathrm{Tr}(Z^p)] \sim \sum_{\alpha,\beta\in\mathcal{S}_p} r^{-|\beta|} c^{-(|\alpha|+|\alpha\beta^{-1}|)} n^{-(|\alpha|+|\gamma^{-1}\alpha|+2|\alpha\beta^{-1}|)} \mathrm{Mob}(\alpha\beta^{-1}).$$

The exponent of the large parameter n in the last formula is minimized when $\mathrm{id} \to \alpha = \beta \to \gamma$ is a geodesic in $\mathcal{S}_p$. Hence,

$$\mathbb{E}[\mathrm{Tr}(Z^p)] \sim n^{1-p} \sum_{\mathrm{id}\to\alpha\to\gamma} (rc)^{-|\beta|} \mathrm{Mob}(\alpha\beta^{-1}).$$

Thus, the normalized trace of the pth power of the matrix rkZ converges to

$$\sum_{\mathrm{id}\to\alpha\to\gamma} (rc)^{\#\beta} = \sum_{\sigma\in NC(p)} (rc)^{\#\sigma}$$

and we easily recognize the moment–cumulant formula for the Marchenko–Pastur distribution of parameter rc (see Section 2.3).

The above results have been proven to hold for the convergence in moments. Borel–Cantelli techniques (see [9] for a sample) can be easily used to show that the stronger almost sure convergence holds in all three cases. □

5.3. *Normalized macroscopic inputs*. We now consider matrices X which have a macroscopic scaling $\mathrm{Tr}(X^p) \sim n \cdot \varphi(x^p)$, where x is some noncommutative random variable. We have, of course, to normalize such input matrices and we shall consider

$$\tilde{X} = \frac{X}{\mathrm{Tr}\, X}.$$

With this normalization, the moments of the output matrix $Z = \Phi(\tilde{X})$ are given by

$$\mathbb{E}[\mathrm{Tr}(Z^p)] = \mathbb{E}[\mathrm{Tr}(\Phi(\tilde{X})^p)] = \mathbb{E}\left[\mathrm{Tr}\,\frac{\Phi(X)^p}{(\mathrm{Tr}\,X)^p}\right] = \frac{\mathbb{E}[\mathrm{Tr}(\Phi(X)^p)]}{(\mathrm{Tr}\,X)^p}.$$

As in the previous section, we consider different asymptotic regimes for the integer parameters n and k. However, it turns out that the k fixed, $n \to \infty$ regime is more involved, and its understanding requires some more advanced free probabilistic tools. To an integer k and a probability measure μ, we associate the measure $\mu_{(k)}$ defined by

$$\mu_{(k)} = \left(1 - \frac{1}{k}\right)\delta_0 + \frac{1}{k}\mu.$$

PROPOSITION 5.2. *The almost sure behavior of the output matrix $Z = \Phi(\tilde{X})$ is given as follows*:

(I) *when n is fixed and $k \to \infty$, Z converges almost surely to the maximally mixed state*

$$\rho_* = \frac{1}{n}\mathrm{I}_n;$$

(II) *when k is fixed and $n \to \infty$, the empirical spectral distribution of $\bar{\mu}knZ$ converges to the probability measure $\nu = [\mu_{(k)}]^{\boxplus k^2}$, where $\boxplus$ denotes the free additive convolution operation, μ is the probability distribution of x with respect to φ: $\varphi(x^p) = \int t^p \, d\mu(t)$ and $\bar{\mu}$ is the mean of μ, $\bar{\mu} = \varphi(x)$;*

(III) *when $n, k \to \infty$ and $k/n \to c$, the empirical spectral distribution of the matrix nZ converges to the Dirac mass δ_1.*

PROOF. We start with the simplest asymptotic regime, n fixed and $k \to \infty$. Plugging the scaling for $\mathrm{Tr}_\beta(X)$ into formula (19), we get

$$\mathbb{E}[\mathrm{Tr}(Z^p)] \sim n^{-p}\varphi(x)^{-p} \sum_{\alpha,\beta \in \mathcal{S}_p} k^{\#\alpha} n^{\#(\gamma^{-1}\alpha)} n^{\#\beta} \varphi_\beta(x)(nk)^{-p-|\alpha\beta^{-1}|}\,\mathrm{Mob}(\alpha\beta^{-1}).$$

In order to find the leading term in the preceding sum, we have to minimize the exponent of k, $|\alpha| + |\alpha\beta^{-1}|$. This expression attains its minimum 0 at $\alpha = \beta = \mathrm{id}$. In the end, we find $\mathbb{E}[\mathrm{Tr}(Z^p)] \sim n^{1-p}$ and conclude that the output matrix Z converges to the maximally mixed state $\rho_* = \mathrm{I}_n/n$.

Let us now look at the second regime, k fixed and $n \to \infty$. The asymptotic moments of Z are given by

$$\mathbb{E}[\mathrm{Tr}(Z^p)] \sim \varphi(x)^{-p} \sum_{\alpha,\beta \in \mathcal{S}_p} k^{-(|\alpha|+|\alpha\beta^{-1}|)} n^{-(|\beta|+|\alpha\beta^{-1}|+|\gamma^{-1}\alpha|)} \varphi_\beta(x)\,\mathrm{Mob}(\alpha\beta^{-1}).$$

The dominating terms in the preceding sum are given by permutations such that $|\beta| + |\alpha\beta^{-1}| + |\gamma^{-1}\alpha|$ is minimal. Permutations (α, β) which saturate the triangle inequality $|\beta| + |\alpha\beta^{-1}| + |\gamma^{-1}\alpha| \geq |\gamma| = p - 1$ are elements of the geodesic $\mathrm{id} \to \beta \to \alpha \to \gamma$ and can be put in bijection with noncrossing partitions $\sigma \leq \tau \in NC(p)$ using Lemma 2.5. We obtain

$$\frac{1}{n}\mathbb{E}[\mathrm{Tr}((\bar{\mu}knZ)^p)] \sim \sum_{\sigma \leq \tau \in NC(p)} k^{2\#\tau - \#\sigma} \varphi_\sigma(x) \operatorname{Mob}(\sigma, \tau).$$

Using the fact that $k^{-\#\sigma}\varphi_\sigma(x) = \varphi_\sigma(\mu_{(k)})$ and applying the moment–cumulant formula ([24], page 175), we get

$$\begin{aligned}\frac{1}{n}\mathbb{E}[\mathrm{Tr}((\bar{\mu}knZ)^p)] &\sim \sum_{\tau \in NC(p)} k^{2\#\tau} \sum_{\substack{\sigma \in NC(p)\\ \sigma \leq \tau}} \varphi_\sigma(\mu_{(k)}) \operatorname{Mob}(\sigma, \tau)\\ &= \sum_{\tau \in NC(p)} k^{2\#\tau} \kappa_\tau(\mu_{(k)}),\end{aligned}$$

where κ denotes the free cumulant. We conclude that the random matrix $\bar{\mu}knZ$ converges in distribution to a probability measure ν which has free cumulants $\kappa_p(\nu) = k^2\kappa_p(\mu_{(k)})$, and the conclusion follows.

We now turn to the third regime, where both n and k grow to infinity at a constant ratio $c > 0$. After making the substitution $k = cn$, we obtain the following equivalent:

$$\begin{aligned}&\mathbb{E}[\mathrm{Tr}(Z^p)]\\ &\quad\sim \varphi(x)^{-p} \sum_{\alpha, \beta \in \mathcal{S}_p} n^{-(|\alpha| + |\gamma^{-1}\alpha| + |\beta| + 2|\alpha\beta^{-1}|)} c^{-(|\alpha| + |\alpha\beta^{-1}|)} \varphi_\beta(x) \operatorname{Mob}(\alpha\beta^{-1}).\end{aligned}$$

The expression to minimize in this case is $|\alpha| + |\gamma^{-1}\alpha| + |\beta| + 2|\alpha\beta^{-1}|$. By the triangle inequality, (cf. Lemma 2.5), the sum of the first two terms is at least $|\gamma| = p - 1$ and the other terms are positive; hence, the (negative) exponent of n is at least $p - 1$, and the bound is reached for $\alpha = \beta = \mathrm{id}$. To the first-order in n, the asymptotic moments of Z are

$$\mathbb{E}[\mathrm{Tr}(Z^p)] \sim n^{1-p} \qquad \forall p \geq 1,$$

which is equivalent to the statement

$$\lim_{n \to \infty} \mathbb{E}[\mathrm{tr}_n((nZ)^p)] = 1 \qquad \forall p \geq 1.$$

In all the cases treated above, we leave the proof of the almost sure convergence to the reader. □

REMARK 5.3. Let us observe that for the regimes (I) and (III) studied above, the limit distribution of the output does not depend on the limit of the input distribution. The result obtained in the second regime could have been obtained in a more direct manner, using the powerful tools of free probability. For simplicity, let us forget about the normalization of the input matrix and observe that the limit distribution of $X \otimes Y$ is $\mu_{(k)}$, if μ is the limit distribution of X and Y is a $k \times k$ rank-one projector. The partial trace of the randomly rotated input matrix is equal to the sum of its k $n \times n$ diagonal blocks. Each block is a free compression of parameter $1/k$ (which accounts for a free additive convolution power of k), and the blocks are free. Taking the sum of the free blocks explains the other factor k appearing as an exponent for the free additive convolution.

6. Tensor products of quantum channels.

6.1. *Motivation and existing results.* When studying the question of the additivity of minimal output entropies, it is natural to consider products of random quantum channels.

Before looking in detail at some specific models, let us observe that if one chooses an input state which factorizes $X_{12} = X_1 \otimes X_2$, then

$$[\Phi_1 \otimes \Phi_2](X_{12}) = \Phi_1(X_1) \otimes \Phi_2(X_2),$$

and there is no correlation (classical or quantum) between the channels. In order to avoid such trivial situations, we must choose an input state which is entangled. An obvious choice (given that $n_1 = n_2 = n$) is to take $X_{12} = E_n$, the n-dimensional Bell state, and we shall use this state in what follows.

Winter and Hayden observed in [20] that it is relevant in this framework to introduce the further symmetry $U_2 = \overline{U}_1$, as it ensures that at least one eigenvalue is always large. In [9], using the channel model inspired by the ideas of Hayden and Winter, it was proven that the bounds on the eigenvalues could be improved as follows.

THEOREM 6.1. *In the k fixed, $n \to \infty$ regime, the eigenvalues of the matrix Z converge almost surely toward*:

- $\frac{1}{k} + \frac{1}{k^2} - \frac{1}{k^3}$, *with multiplicity one*;
- $\frac{1}{k^2} - \frac{1}{k^3}$, *with multiplicity $k^2 - 1$*;
- 0, *with multiplicity $n^2 - k^2$.*

In the asymptotic regime where n is fixed and $k \to \infty$, the random matrix Z converges to the chaotic state

$$\rho_* = \frac{\mathrm{I}_{n^2}}{n^2}.$$

If we look for optimal bounds for the minimum output entropy of $\Phi \otimes \overline{\Phi}$, then there is no mathematical proof that $U_2 = \overline{U}_1$ is the best choice. Actually, this choice of probability measure on $\mathcal{U}(n) \times \mathcal{U}(n)$ does not have full support and we cannot rule out that the maximum for the minimum output entropy is outside the support of the probability measure. This is what motivates the introduction of the example in which U_1 and U_2 are independent unitary matrices. As we will see, this does not yield improvements on the example of Winter with high probability. More strikingly, in the regimes that we consider, we will see that the constraint $U_2 = \overline{U}_1$ yields no significant improvement to the asymptotic behavior of the von Neumann entropies, and this suggests that the simpler random model where U_1 and U_2 are independent could be a candidate for additivity violation with high probability.

In the forthcoming subsections we analyze both models (independent and conjugate unitaries) in a different asymptotic regime, where both parameters n and k grow to infinity at a constant ratio $k/n \to c$. The model where the quantum channels are independent has received less attention from the quantum information community; here, we show that it is intimately connected to the (more interesting) case of conjugate channels, by comparing eigenvalue profiles for outputs of channels from the two families.

6.2. *Independent interaction unitaries.* Here, we consider two *independent* realizations $U_1 = U$ and $U_2 = V$ of Haar-distributed unitary random matrices on $\mathcal{U}(nk)$. For both channels the state of the environment is a rank-one projector and we are interested in the $n^2 \times n^2$ random matrix

$$Z = [\Phi^U \otimes \Phi^V](E_n),$$

where E_n is the maximal entangled Bell state

$$E_n = \frac{1}{n} \sum_{i,j=1}^{n} |e_i\rangle\langle e_j| \otimes |e_i\rangle\langle e_j|.$$

The diagram associated with the $(2,2)$ tensor Z appears in Figure 7.

We compute the moments $\mathbb{E}[\mathrm{Tr}(Z^p)]$ for all $p \geq 1$ using the graphical method. We start, as depicted in Figure 7, by replacing U^* (resp., V^*) boxes by $\bar{U}$ (resp., $\bar{V}$)

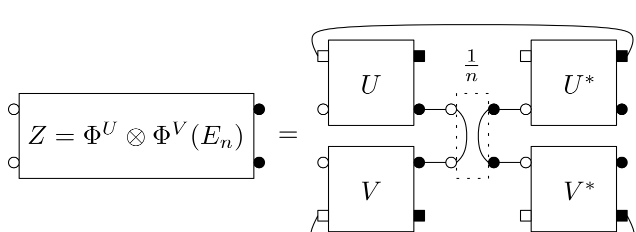

FIG. 7. $Z = \Phi^U \otimes \Phi^V(E_n)$.

boxes. Notice that there are two types of boxes corresponding to the independent random unitary matrices U and V (when computing the pth moment of Z, there are p boxes of each type). This has two important consequences: when expanding the diagram in order to compute the expectation of the trace, we can only pair U boxes with $\bar{U}$ boxes and V boxes with $\bar{V}$ boxes; "cross-pairings" between "U" boxes and "V" boxes are not allowed by the expansion algorithm. In addition, we have to index the Weingarten sum by two pairs of permutations, one for each type of box (we shall denote them by $\alpha_U, \beta_U, \alpha_V, \beta_V \in \mathcal{S}_p$). The four permutations are responsible for pairing boxes in the following ways ($1 \le i \le p$):

(1) the inputs of the ith U-box are paired with the inputs of the $\alpha_U(i)$th $\bar{U}$ box;
(2) the outputs of the ith U-box are paired with the outputs of the $\beta_U(i)$th $\bar{U}$ box;
(3) the inputs of the ith V-box are paired with the inputs of the $\alpha_V(i)$th $\bar{V}$ box;
(4) the outputs of the ith V-box are paired with the outputs of the $\beta_V(i)$th $\bar{V}$ box.

Since our diagram consists only of unitary matrices (there are no constant nontrivial tensors), the result of the graph expansion is a (sum over a) collection of loops, multiplied by some scalar factor. The different contributions of a general quadruple $(\alpha_U, \beta_U, \alpha_V, \beta_V) \in \mathcal{S}_p^4$ are given by (recall that circles correspond to n-dimensional spaces and squares correspond to k-dimensional spaces):

(1) loops from $\square U$ and $\bar{U}\square$: $k^{\#\alpha_U}$;
(2) loops from $\circ U$ and $\bar{U}\circ$: $n^{\#(\gamma^{-1}\alpha_U)}$;
(3) loops from $U\blacksquare$ and $\blacksquare\bar{U}$: none;
(4) loops from $U\bullet$, $\bullet\bar{U}$, $V\bullet$ and $\bullet\bar{V}$: $n^{\#(\beta_U^{-1}\beta_V)}$;
(5) loops from $\square V$ and $\bar{V}\square$: $k^{\#\alpha_V}$;
(6) loops from $\circ V$ and $\bar{V}\circ$: $n^{\#(\gamma^{-1}\alpha_V)}$;
(7) normalization factors $1/n$ from the Bell matrices E_n: n^{-p};
(8) Weingarten weights for the U-matrices: $\mathrm{Wg}(\alpha_U\beta_U^{-1})$;
(9) Weingarten weights for the V-matrices: $\mathrm{Wg}(\alpha_V\beta_V^{-1})$.

Adding all these contributions, we obtain an exact closed-form expression, as follows.

PROPOSITION 6.2. *The moments of the random variable Z can be computed as follows*:

$$\mathbb{E}[\mathrm{Tr}(Z^p)] = \sum_{\alpha_U,\beta_U,\alpha_V,\beta_V\in\mathcal{S}_p} k^{\#\alpha_U+\#\alpha_V} n^{\#(\gamma^{-1}\alpha_U)+\#(\gamma^{-1}\alpha_V)+\#(\beta_U^{-1}\beta_V)-p} \times \mathrm{Wg}(\alpha_U\beta_U^{-1})\,\mathrm{Wg}(\alpha_V\beta_V^{-1}). \tag{21}$$

Here, we study the asymptotic regime $n, k \to \infty$, $k/n \to c > 0$. Our main theorem is as follows.

THEOREM 6.3. *Almost surely, in the regime $n \to \infty$, $k \sim cn$, the distribution of the output matrix c^2n^2Z converges toward a free Poisson law with parameter c^2.*

PROOF. We start by replacing k by cn in equation (21) and obtain

$$\mathbb{E}[\mathrm{Tr}(Z^p)] \sim \sum_{\alpha_U,\beta_U,\alpha_V,\beta_V \in \mathcal{S}_p} n^{-\mathcal{P}_n} c^{-\mathcal{P}_c} \mathrm{Mob}(\alpha_U\beta_U^{-1}) \mathrm{Mob}(\alpha_V\beta_V^{-1}),$$

where

$$\mathcal{P}_n = |\alpha_U| + |\alpha_V| + |\gamma^{-1}\alpha_U| + |\gamma^{-1}\alpha_V| + |\beta_U^{-1}\beta_V| + 2|\alpha_U\beta_U^{-1}| + 2|\alpha_V\beta_V^{-1}|$$

and

$$\mathcal{P}_c = |\alpha_U| + |\alpha_V| + |\alpha_U\beta_U^{-1}| + |\alpha_V\beta_V^{-1}|.$$

Since we are interested in the asymptotic $n \to \infty$ (c is a constant), we want to minimize $\mathcal{P}_n$. The following inequalities are standard (cf. Lemma 2.5):

$$|\alpha_U| + |\gamma^{-1}\alpha_U| \geq p - 1; \tag{22}$$

$$|\alpha_V| + |\gamma^{-1}\alpha_V| \geq p - 1; \tag{23}$$

$$|\beta_U^{-1}\beta_V|, 2|\alpha_U\beta_U^{-1}|, 2|\alpha_V\beta_V^{-1}| \geq 0, \tag{24}$$

and thus $\mathcal{P}_n \geq 2p - 2$ with equality iff $\alpha_U = \beta_U = \alpha_V = \beta_V = \alpha$ and α is on a geodesic between id and γ. By choosing the obvious n^2 rescaling, we get

$$\lim_{n,k\to\infty} \mathbb{E}\left[\frac{1}{n^2}\mathrm{Tr}((c^2n^2Z)^p)\right] = \sum_{\alpha \text{ geodesic}} c^{2p-2|\alpha|} = \sum_{\alpha \text{ geodesic}} c^{2\#\alpha} = \sum_{\sigma \in NC(p)} c^{2\#\sigma},$$

and we recognize in the last sum the pth moment of the free Poisson distribution with parameter c^2. This shows that the the matrix Z converges *in moments* to the limiting Marchenko–Pastur distribution. The argument for the almost sure convergence relies on the Borel–Cantelli lemma and can be found in the Appendix. □

The von Neumann entropy of the output can be calculated in a fashion similar to Corollary 4.6.

PROPOSITION 6.4. *Almost surely, in the limit $n \to \infty$, the von Neumann entropy of the matrix Z satisfies*

$$H(Z) = \begin{cases} 2\log n - \dfrac{1}{2c^2} + o(1), & \text{if } c \geq 1, \\ 2\log(cn) - \dfrac{c^2}{2} + o(1), & \text{if } 0 < c < 1. \end{cases}$$

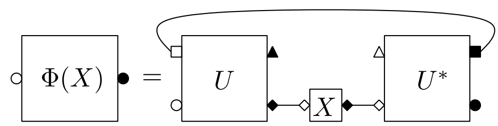

FIG. 8. *A quantum channel with asymmetric input and output tensor structure.*

Let us now consider a slightly generalized model of random quantum channels. We introduce channels $\Phi : \mathcal{M}_d(\mathbb{C}) \to \mathcal{M}_n(\mathbb{C})$ which have a different tensor product structure at their input and output. Here, d is an integer parameter, and we shall always suppose that $d|nk$. The diagram associated with such a channel is depicted in Figure 8, where diamond-shaped labels correspond to d-dimensional vector spaces and triangle-shaped decorations denote spaces of dimension $d' = nk/d$. The above analysis for a product of independent channels is easily adapted to this more general situation:

$$\mathbb{E}[\mathrm{Tr}(Z^p)] \sim \sum_{\alpha_U, \beta_U, \alpha_V, \beta_V \in \mathcal{S}_p} n^{-\tilde{\mathcal{P}}_n} d^{-\tilde{\mathcal{P}}_d} c^{-\tilde{\mathcal{P}}_c} \mathrm{Mob}(\alpha_U \beta_U^{-1}) \mathrm{Mob}(\alpha_V \beta_V^{-1}),$$

where

$$\tilde{\mathcal{P}}_n = |\alpha_U| + |\alpha_V| + |\gamma^{-1}\alpha_U| + |\gamma^{-1}\alpha_V| + 2|\alpha_U \beta_U^{-1}| + 2|\alpha_V \beta_V^{-1}|,$$

$$\tilde{\mathcal{P}}_d = |\beta_U^{-1}\beta_V| \quad \text{and} \quad \tilde{\mathcal{P}}_c = |\alpha_U| + |\alpha_V| + |\alpha_U \beta_U^{-1}| + |\alpha_V \beta_V^{-1}|.$$

REMARK 6.5. If $d = d(n)$ is a function of n such that $\lim_{n\to\infty} d(n) = \infty$, then the considerations in Theorem 6.3 carry over to this case and we obtain exactly the same limit, a free Poisson distribution with parameter c^2. The function $d = d(n)$ does not play any role in this situation.

On the other hand, if the parameter d is constant (inputs of fixed dimension), then the limiting behavior changes. Indeed, the minimizing constraint $|\beta_U^{-1}\beta_V| = 0$ disappears, and the contributing quadruples of permutations become uncoupled: $\mathrm{id} \to \alpha_U = \beta_U \to \gamma$ and $\mathrm{id} \to \alpha_V = \beta_V \to \gamma$. In conclusion, the asymptotic moments in this case are given by the formula, which we summarize in the following proposition.

PROPOSITION 6.6. *If d is constant, the limiting distribution of $c^2 n^2 Z$ also exists and its limit moments are given by*

$$\frac{1}{n^2}\mathbb{E}[\mathrm{Tr}((c^2 n^2 Z)^p)] \sim \sum_{\substack{\mathrm{id} \to \alpha_U = \beta_U \to \gamma \\ \mathrm{id} \to \alpha_V = \beta_V \to \gamma}} c^{\#\alpha_U + \#\alpha_V} d^{-|\alpha_U^{-1}\alpha_V|}.$$

QUESTION 6.7. We are not able to identify this distribution, even though its properties look new. We wonder whether this distribution could be related to generalized convolutions of Bożejko and coworkers (cf. [3]).

6.3. *Conjugate interaction unitaries.* To conclude, we consider the tensor product of two *conjugate* random quantum channels. As was emphasized in Section 6.1, product channels $\Phi_U \otimes \Phi_{\overline{U}}$ have very interesting eigenvalues statistics and have received a lot of attention in the last years because of their usefulness in providing counterexamples to different additivity conjectures. The purpose of this section is to obtain a description of the behavior of such channels in the regime where both n and k grow to infinity at a constant ratio $c \in (0, \infty)$.

Hayden and Winter remarked in [20] that such a conjugate product channel has a very important property: the output of the maximally entangled state over the input space has a "large" eigenvalue, of size at least $1/(cn)$. The results of [9] show that one expects for this model a large eigenvalue $\lambda_1 = 1/(cn) + o(1/n)$ and $(n^2 - 1)$ smaller eigenvalues. The purpose of this section is to show that this is indeed the case. Actually, we can prove that the random matrix under study has eigenvalues on two scalings: $1/n$ and $1/n^2$. In the next theorem, we compute the moments of the output matrix Z up to the first order in n.

THEOREM 6.8. *Fix some scaling constant $c > 0$ and consider a sequence of random quantum channels $\Phi_{n,k}$, where $n, k \to \infty$ and $k/n \to c$. The asymptotic moments of the output matrix $Z = \Phi \otimes \overline{\Phi}(E_n)$ are given by*

$$\begin{aligned} \operatorname{Tr}(Z) &= 1; \\ \mathbb{E} \operatorname{Tr}((cnZ)^2) &= 2 + c^2 + O(n^{-1}); \\ \mathbb{E} \operatorname{Tr}((cnZ)^p) &= 1 + O(n^{-1}) \qquad \forall p \geq 3. \end{aligned}$$

REMARK 6.9. Before we prove this result, we would like to point out to readers aware of random matrix theory and matrix integrals that the symbol $O(n^{-1})$ is actually optimal. One can check by inspection that there are terms of order n^{-1} in the expansion of the quantities of the theorem. This observation stresses the fact that the matrix model Z does not behave like a usual unitarily invariant matrix model, but rather like an orthogonal matrix model, even though the underlying invariance group is the unitary group. This technicality explains why we can only obtain convergence in probability of the rescaled largest eigenvalue and not the almost sure convergence.

PROOF OF THEOREM 6.8. We start from the exact expression for fixed n and k for the moments of Z (see [9]):

$$\mathbb{E}[\operatorname{Tr}(Z^p)] = \sum_{\alpha, \beta \in \mathcal{S}_{2p}} k^{\#\alpha} n^{\#(\alpha\gamma^{-1}) + \#(\beta\delta) - p} \operatorname{Wg}(\alpha\beta^{-1}). \tag{25}$$

Since in this "conjugate" case, the permutations α and β act on the whole set of $2p$ boxes, we introduce a special labeling on the boxes. The *top* row of boxes (corresponding to the channel Φ) shall be labeled by $1^T, 2^T, \dots, p^T$ and the *bottom* row by $1^B, 2^B, \dots, p^B$. With this notation, the permutations γ and δ have the following expressions:

$$\begin{aligned} \gamma &= \bigl(p^T\ (p-1)^T \cdots 1^T\bigr)\ (1^B\ 2^B \cdots p^B); \\ \delta &= (1^T\ 1^B)\ (2^T\ 2^B) \cdots (p^T\ p^B). \end{aligned} \tag{26}$$

Dropping the number-of-cycles statistics $\#(\cdot)$ in favor of permutation lengths $|\cdot|$, replacing $k \sim cn$ and using the standard asymptotic expansion for the Weingarten function, we have

$$\mathbb{E}[\mathrm{Tr}(Z^p)] \sim \sum_{\alpha,\beta \in \mathcal{S}_{2p}} c^{-(|\alpha|+|\alpha\beta^{-1}|)} n^{p-(|\alpha|+|\alpha\gamma^{-1}|+|\beta\delta|+2|\alpha\beta^{-1}|)} \mathrm{Mob}(\alpha\beta^{-1}).$$

In order to find the first order asymptotic (in n) of this expression, one has to minimize the quantity

$$|\alpha| + |\alpha\gamma^{-1}| + |\beta\delta| + 2|\alpha\beta^{-1}|$$

over all permutations $\alpha, \beta \in \mathcal{S}_{2p}$. We start by simplifying this optimization problem over two permutations by using the following two inequalities:

$$|\alpha| + |\alpha\beta^{-1}| \geq |\beta|; \tag{27}$$

$$|\alpha\gamma^{-1}| + |\alpha\beta^{-1}| \geq |\beta\gamma^{-1}|. \tag{28}$$

Note that these inequalities can be simultaneously saturated by choosing, for example, $\alpha = \beta$. So, one is left with the following minimization problem over $\beta \in \mathcal{S}_{2p}$:

$$\text{minimize} \qquad S_1(\beta) = |\beta| + |\beta\gamma^{-1}| + |\beta\delta^{-1}|. \tag{29}$$

The main ingredient in tackling this problem is the fact that both permutations δ and γ lie on the geodesic between the identity permutation id and the full-cycle permutation

$$\tilde{\gamma} = (p^T \cdots 2^T 1^T 1^B 2^B \cdots p^B).$$

This follows from the saturated triangle inequalities $|\delta| + |\delta^{-1}\tilde{\gamma}| = p + p - 1 = 2p - 1$ and $|\gamma| + |\gamma^{-1}\tilde{\gamma}| = 2(p-1) + 1 = 2p - 1$. If fact, one has $\tilde{\gamma} = (p^T 1^B) \cdot \gamma$. Under Biane's isomorphism, (cf. 2.5), the permutations δ and γ correspond to the noncrossing partitions in Figure 9.

We start by looking at the following simplified minimization problem:

$$\text{minimize} \qquad S_2(\beta) = |\beta| + |\beta\delta^{-1}| + |\beta\tilde{\gamma}^{-1}|.$$

Obviously, $|\beta| + |\beta\tilde{\gamma}^{-1}| \geq 2p - 1$, with equality iff β lies in the geodesic between id and $\tilde{\gamma}$. It follows from a parity argument that if β is not an element of the

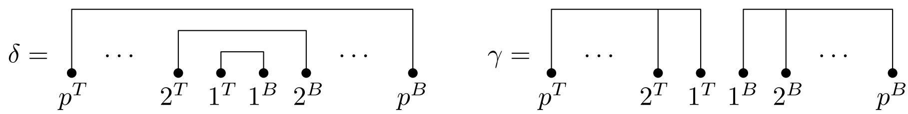

FIG. 9. *Noncrossing partitions associated with permutations δ and γ.*

geodesic set $\mathrm{id} \to \tilde{\gamma}$, then $|\beta| + |\beta\tilde{\gamma}^{-1}| \geq 2p + 1$ and hence, since in this case one has $\beta \neq \delta$, $S_2(\beta) \geq 2p + 2$. If β is a geodesic element, then $S_2(\beta) \geq 2p - 1$, with equality iff $\beta = \delta$.

Since the permutations γ and $\tilde{\gamma}$ are at distance one, the same holds for $\beta\gamma^{-1}$ and $\beta\tilde{\gamma}^{-1}$:

$$\beta\gamma^{-1} = \beta\tilde{\gamma}^{-1} \cdot (p^T 1^B).$$

We have $|\beta\gamma^{-1}| = |\beta\tilde{\gamma}^{-1}| \pm 1$ and, even more precisely,

$$|\beta\gamma^{-1}| = \begin{cases} |\beta\tilde{\gamma}^{-1}| - 1, & \text{if } p^T \text{ and } 1^B \text{ are in the same block of } \beta\tilde{\gamma}^{-1}, \\ |\beta\tilde{\gamma}^{-1}| + 1, & \text{otherwise.} \end{cases}$$

Note that the condition appearing in the first case can be restated in the following, simpler way. It is known that since β is a geodesic element, $\beta^{-1}\tilde{\gamma}$ is also on the geodesic, and the noncrossing partition associated with $\beta^{-1}\tilde{\gamma}$ is the Kreweras complement of the partition associated with β (see [24], Lecture 9, for a definition of the Kreweras complement of a noncrossing partition). The condition that p^T and 1^B belong to the same block of $K(\beta)$ is depicted in Figure 10 and it is easily seen to be equivalent to $\beta \leq \gamma$ (the permutations are compared here via their associated partitions).

It follows that $S_1(\beta) = S_2(\beta) \pm 1$ and, in order to conclude, one needs to look at the position of the permutation β with respect to the geodesic $\mathrm{id} \to \tilde{\gamma}$. If β is not an element of $\mathrm{id} \to \tilde{\gamma}$, then $S_1(\beta) \geq S_2(\beta) - 1 \geq 2p + 2 - 1 = 2p + 1$, which is enough to conclude. We assume from now on that β is a geodesic element. If $|\beta\gamma^{-1}| = |\beta\tilde{\gamma}^{-1}| + 1$, then $S_1(\beta) = S_2(\beta) + 1 \geq 2p$, with equality if and only if $\beta = \delta$, and the conclusion follows. The most difficult case is when $|\beta\gamma^{-1}| = |\beta\tilde{\gamma}^{-1}| - 1$, which is equivalent to the fact that p^T and 1^B are in different blocks of $\beta\tilde{\gamma}^{-1}$. We get $S_1(\beta) = 2p - 2 + |\beta\delta^{-1}|$. We claim that for any geodesic β such that $\beta \leq \gamma$, $|\beta\delta^{-1}| \geq p$. This follows from the fact that the permutation $\beta\delta^{-1} = \beta\delta$ has no fixed points: any index x^T is mapped by δ to x^B, which is then mapped by β to some $y^B \neq x^T$ (the same holds for bottom indices). Since it has no fixed

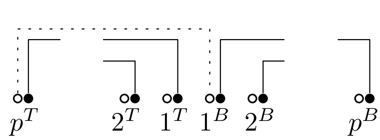

FIG. 10. *Kreweras complement of β.*

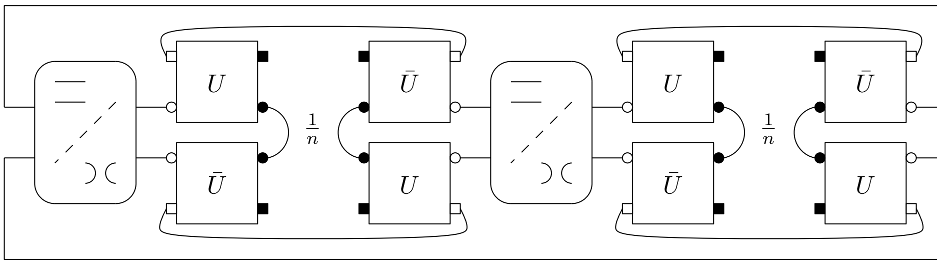

FIG. 11. *Developing* $\mathrm{Tr}(n^2 QZQ)^2$.

points, each cycle of $\beta\delta^{-1}$ has cardinality at least 2, and thus $\beta\delta^{-1}$ has at most p cycles, which implies $|\beta\delta^{-1}| = 2p - \#(\beta\delta^{-1}) \geq p$. We conclude that if a geodesic permutation β verifies $|\beta\gamma^{-1}| = |\beta\tilde{\gamma}^{-1}| - 1$, then $S_1(\beta) \geq 2p - 2 + p \geq 2p + 1$ for $p \geq 3$.

So far, we have shown that the inequality $S_1(\beta) \geq 2p$ holds for all β and p. Moreover, for $p \geq 3$, we have shown that equality holds if and only if $\beta = \delta$. For $p = 2$, using an exhaustive search in $\mathcal{S}_4$, we can identify the permutations which saturate the equality: $\beta \in \{\mathrm{id}, \delta, \gamma\}$.

Now that we have completely solved the minimization problem for β, let us go back to equations (27) and (28) and find, for each minimizing β, the values of α which saturate both inequalities. For $\beta = \delta$, the geodesic $\mathrm{id} \to \delta$ has a very simple expression since δ is a product of transpositions with disjoint support (see the proof of Theorem 6.3 in [9]):

$$\mathrm{id} \to \alpha \to \delta \quad \iff \quad \exists \varnothing \subseteq A \subseteq \{1, 2, \dots, p\} \quad \text{such that} \quad \alpha = \prod_{i \in A} (i^T i^B).$$

Obviously, we have $\alpha\delta^{-1} = \prod_{i \notin A} (i^T i^B)$, and thus formula (28) reads $|\alpha\gamma^{-1}| + p - |A| = p$. Writing explicitly $\alpha\gamma^{-1}$, we can show that

$$\#(\alpha\gamma^{-1}) = \begin{cases} 1, & \text{if } A = \varnothing, \\ |A|, & \text{otherwise.} \end{cases}$$

Obviously, $A = \varnothing$ does not verify the equality, so one is left with $2p - |A| = |A| \Rightarrow |A| = p$ and hence $\alpha = \delta = \beta$. The other two cases for $p = 2$ ($\beta = \mathrm{id}$ and $\beta = \gamma$) are trivial and yield the same result $\alpha = \beta$. In conclusion, for $p \geq 3$, we obtain

$$\mathbb{E}[\mathrm{Tr}(Z^p)] = c^{-p} n^{-p} + o(n^{-p})$$

and for $p = 2$,

$$\mathbb{E}[\mathrm{Tr}(Z^2)] = (1 + 2c^{-2}) n^{-2} + o(n^{-2}),$$

which completes the proof. □

At this point, the description of the random matrix Z is not complete: the moment information of the preceding theorem allows us to infer that there are at least some eigenvalues on the scale of $1/n$ and that the rest of the spectrum is distributed on lower scales, such as $1/n^2$. Hayden and Winter's proof of the existence of a *large* eigenvalue contains, as a byproduct, some information on the eigenvector for this particular eigenvalue. Indeed, they use the projection on the Bell state to obtain a lower bound for the largest eigenvalue of Z, so one can use this projector to obtain more precise information on the eigenvalue distribution of Z.

In order to obtain information on the rest of the spectrum, we introduce the orthogonal projection $Q = \mathrm{I} - E$, where E is the maximally entangled state. Using the (rank $n^2 - 1$) projector Q, we shall obtain some information on the smallest $n^2 - 1$ eigenvalues of the output matrix Z.

THEOREM 6.10. *Almost surely, the matrix c^2n^2QZQ converges in distribution, to a Marchenko–Pastur law with parameter c^2.*

PROOF. We compute the moments of the random matrix c^2n^2QZQ and show that they converge to the corresponding moments of the limit law:

$$\lim_{n\to\infty} \frac{1}{n^2}\mathbb{E}\,\mathrm{Tr}(c^2n^2QZQ)^p = \int x^p \, d\pi_{c^2}(x).$$

We start by replacing $Q = \mathrm{I} - E$ and expanding the product

$$\begin{aligned}\frac{1}{n^2}&\mathbb{E}\,\mathrm{Tr}(c^2n^2QZQ)^p \\ &= c^{2p}n^{2p-2}\mathbb{E}\,\mathrm{Tr}(\mathrm{I}-E)Z(\mathrm{I}-E)Z\cdots(\mathrm{I}-E)Z \\ &= c^{2p}n^{2p-2}\sum_{f\in\mathcal{F}_p}(-1)^{|f^{-1}(E)|}n^{-|f^{-1}(E)|}\mathbb{E}\,\mathrm{Tr}\,f(1)Zf(2)Z\cdots f(p)Z,\end{aligned}$$

where $\mathcal{F}$ is a set of the 2^p choice functions $f : \{1, 2, \ldots, p\} \to \{\mathrm{I}, E\}$. Notice that in the last formula, each Bell projector E is multiplied by a factor $-1/n$.

The moment $\mathbb{E}\,\mathrm{Tr}\,f(1)Zf(2)Z\cdots f(p)Z$ is computed with our graphical calculus, and the computation is similar to those in the proof of Theorem 6.3 (see Figure 11 for the case $p = 2$):

$$\mathbb{E}\,\mathrm{Tr}\,f(1)Zf(2)Z\cdots f(p)Z = \sum_{\alpha,\beta\in\mathcal{S}_{2p}} k^{\#\alpha}n^{\#(\alpha\hat{f}^{-1})+\#(\beta\delta)-p}\,\mathrm{Wg}(\alpha\beta^{-1}),$$

where $\hat{f} \in \mathcal{S}_{2p}$ is the permutation associated with the choice function $f \in \mathcal{F}_p$, describing the way f connects the different instances of the channel. The exact action of $\hat{f}$ can be easily computed:

$$i^T \overset{\hat{f}}{\mapsto} \begin{cases}(i-1)^T, & \text{if } f(i) = \mathrm{I},\\ i^B, & \text{if } f(i) = E,\end{cases}$$

$$i^B \overset{\hat{f}}{\mapsto} \begin{cases}(i+1)^B, & \text{if } f(i+1) = \mathrm{I},\\ i^T, & \text{if } f(i+1) = E,\end{cases}$$

where the arithmetic operations of indices i should be understood modulo p.

When trying to compute the leading order terms in the expression of $\mathbb{E}\,\mathrm{Tr}(n^2 QZQ)^p$, we have to understand the possible cancellations of high powers in n. When writing the exact formula for the pth moment and separating the (α, β) and f parts, we get

$$\frac{1}{n^2}\mathbb{E}\,\mathrm{Tr}(c^2 n^2 QZQ)^p = c^{2p} \sum_{\alpha,\beta\in\mathcal{S}_{2p}} n^{5p-2-|\beta\delta|} k^{2p-|\alpha|}\,\mathrm{Wg}(\alpha\beta^{-1}) \times \sum_{f\in\mathcal{F}_p} (-1)^{|f^{-1}(E)|} n^{-(|f^{-1}(E)|+|\alpha\hat{f}^{-1}|)}.$$

Note that the sum over $f \in \mathcal{F}_p$ depends only on the permutation α. Next, we show that for a large class of permutations α [the ones which are responsible for the *large eigenvalue* of size $1/(cn)$ of Theorem 6.8], this sum is zero. Let us introduce the set of "vertical line permutations,"

$$\begin{aligned}\mathcal{V} &= \{\sigma \in \mathcal{S}_{2p} \mid \exists i \in \{1, \dots, p\} \text{ such that } \sigma(i^T) = i^B \text{ or } \sigma(i^B) = i^T\}\\ &= \{\sigma \in \mathcal{S}_{2p} \mid \sigma\delta \text{ has at least one fixed point}\}.\end{aligned}$$

Fix a permutation $\alpha \in \mathcal{V}$ and some index i such that $\alpha(i^T) = i^B$ or $\alpha(i^B) = i^T$. Consider the "flip at position i" involution $T_i : \mathcal{F}_p \to \mathcal{F}_p$ which maps a choice function f to the function

$$T_i f : j \mapsto \begin{cases} \mathrm{I}, & \text{if } j = i \text{ and } f(j) = E,\\ E, & \text{if } j = i \text{ and } f(j) = \mathrm{I},\\ f(j), & \text{if } j \neq i.\end{cases}$$

We shall show that

$$\sum_{f\in\mathcal{F}_p} (-1)^{|f^{-1}(E)|} n^{-(|f^{-1}(E)|+|\alpha\hat{f}^{-1}|)} = \sum_{f\in\mathcal{F}_p} (-1)^{|(T_i f)^{-1}(E)|} n^{-(|(T_i f)^{-1}(E)|+|\alpha\widehat{(T_i f)}^{-1}|)},$$

which will imply that for all $\alpha \in \mathcal{V}$, both sums are zero. Since the cardinalities of the sets $f^{-1}(E)$ and $(T_i f)^{-1}(E)$ differ by exactly one, all we need to show is that for all $f \in \mathcal{F}_p$,

$$|f^{-1}(E)| + |\alpha\hat{f}^{-1}| = |(T_i f)^{-1}(E)| + |\alpha\widehat{(T_i f)}^{-1}|.$$

To this end, notice that (the order in which one multiplies the transpositions in not important)

$$\hat{f} = \prod_{j\,:\,f(j)=E} ((j-1)^T\ j^B)\tilde{\gamma}$$

and hence $\alpha\widehat{(T_i f)}^{-1} = \alpha \hat{f}^{-1} \cdot ((i-1)^T\, i^B)$. From this we find that

$$|\alpha\widehat{(T_i f)}^{-1}| = \begin{cases} |\alpha \hat{f}^{-1}| - 1, & \text{if } (i-1)^T \text{ and } i^B \text{ belong to the same orbit of } \alpha\hat{f}^{-1}, \\ |\alpha \hat{f}^{-1}| + 1, & \text{otherwise.} \end{cases}$$

Let us now suppose that $\alpha(i^T) = i^B$, the other case $\alpha(i^B) = i^T$ being similar. If $f(i) = \mathrm{I}$, then $\hat{f}(i^T) = (i-1)^T$, $\alpha\hat{f}^{-1}((i-1)^T) = i^B$ and thus $|\alpha\hat{f}^{-1}| - |\alpha\widehat{(T_i f)}^{-1}| = 1$. On the other hand, $f(i) = \mathrm{I} \Rightarrow (T_i f)(i) = E$ and then $|f^{-1}(E)| - |(T_i f)^{-1}(E)| = -1$, and we see that the differences compensate. The case $f(i) = E$ is treated in a similar manner.

We have proven that for all permutations $\alpha \in \mathcal{V}$, the sum over all choices $f \in \mathcal{F}_p$ is exactly zero; notice that the computations we have carried out up to this point are *nonasymptotic*; they are true at fixed matrix sizes n and k. We now interchange the sums over (α, β) and f, replace $k = cn$ and use the first-order asymptotic for the Weingarten function:

$$\begin{aligned} &\frac{1}{n^2}\mathbb{E}\,\mathrm{Tr}(c^2 n^2 QZQ)^p \\ &\quad\sim \sum_{\alpha,\beta \in \mathcal{S}_{2p}, \alpha \notin \mathcal{V}} n^{3p-2-(|\beta\delta|+|\alpha|+2|\alpha\beta^{-1}|)} c^{2p-(|\alpha|+|\alpha\beta^{-1}|)}\,\mathrm{Mob}(\alpha\beta^{-1}) \\ &\qquad\qquad \times \sum_{f \in \mathcal{F}_p} (-1)^{|f^{-1}(E)|} n^{-(|f^{-1}(E)|+|\alpha\hat{f}^{-1}|)}. \end{aligned}$$

To obtain the dominant power of n, we must minimize the following quantity over $(\alpha, \beta, f) \in (\mathcal{S}_{2p} \setminus \mathcal{V}) \times \mathcal{S}_{2p} \times \mathcal{F}_p$:

$$S(\alpha, \beta, f) = |\beta\delta| + |\alpha| + 2|\alpha\beta^{-1}| + |f^{-1}(E)| + |\alpha\hat{f}^{-1}|.$$

Since $\alpha \notin \mathcal{V}$, $\alpha\delta$ has no fixed point and hence $|\alpha\delta| \geq p$. Using the facts that $|\alpha\beta^{-1}| + |\beta\delta| \geq |\alpha\delta|$, $|\alpha\beta^{-1}| \geq 0$ and $|\alpha| + |\alpha\hat{f}^{-1}| \geq |\hat{f}|$, we obtain that

$$S(\alpha, \beta, f) \geq p + |f^{-1}(E)| + |\hat{f}|$$

with equality if and only if $\beta = \alpha$, $|\alpha\delta| = p$ and α is on the geodesic between id and $\hat{f}$. On the other hand, we can easily compute the number of cycles of $\hat{f}$:

$$\#\hat{f} = \begin{cases} 2, & \text{if } f \equiv \mathrm{I}, \\ |f^{-1}(E)|, & \text{otherwise}. \end{cases}$$

Hence, $S(\alpha, \beta, f) \geq 3p - 2$, with equality if and only if $f \equiv \mathrm{I}$, $\beta = \alpha$, $|\alpha\delta| = p$ and α is a permutation on the geodesic id $\to \hat{\mathrm{I}} = \gamma$. Since $\gamma = \gamma^T \oplus \gamma^B$ is a disjoint union of the two cycles $\gamma^T = (p^T \cdots 2^T 1^T)$ and $\gamma^B = (1^B 2^B \cdots p^B) = (\gamma^T)^{-1}$, the condition that α should be a geodesic permutation amounts to $\alpha = \alpha^T \oplus \alpha^B$, where $\alpha^{T,B} \in \mathcal{S}_p$ are geodesic permutations with respect to the cycles $\gamma^{T,B}$. We can easily show that $\#((\alpha^T \oplus \alpha^B)\delta) = \#(\alpha^T \alpha^B)$, where the first permutation is an element of $\mathcal{S}_{2p}$ and the second one is an element of $\mathcal{S}_p$. Using this equality, the condition $|\alpha\delta| = p$ implies that $\alpha^T \alpha^B = \mathrm{id}_p$. When putting all these considerations together, one obtains the final formula for the dominant term of the pth moment of QZQ:

$$\frac{1}{n^2}\mathbb{E}\,\mathrm{Tr}(c^2 n^2 QZQ)^p \sim \sum_{\mathrm{id}\to\alpha^T\to\gamma^T} c^{2p-2|\alpha^T|}\,\mathrm{Mob}(\mathrm{id}) = \sum_{\mathrm{id}\to\alpha^T\to\gamma^T} c^{2\#\alpha^T}.$$

Following the proof of Theorem 6.3, the moments of the Marchenko–Pastur distribution of parameter c^2 are easily recognized and the convergence in moments is settled. The proof of the almost sure convergence is more involved and can be found in the Appendix. □

From this we deduce the following theorem, which summarizes all the results obtained thus far in this section.

THEOREM 6.11. *The eigenvalues $\lambda_1 \geq \cdots \geq \lambda_{n^2}$ of Z are such that*:

- *in probability*, $cn\lambda_1 \to 1$;
- *almost surely*, $\frac{1}{n^2-1}\sum_{i=2}^{n^2} \delta_{c^2 n^2 \lambda_i}$ *converges to a free Poisson distribution with parameter* c^2.

PROOF. Let $\tilde{\lambda}_1 \geq \cdots \geq \tilde{\lambda}_{n^2-1}$ be the eigenvalues of QZQ, seen as a matrix in $\mathcal{M}_{n^2-1}(\mathbb{C})$. By Cauchy's interlacing theorem ([1], Corollary III.1.5), the eigenvalues of QZQ and those of Z are intertwined and satisfy

$$\lambda_1 \geq \tilde{\lambda}_1 \geq \lambda_2 \geq \cdots \geq \lambda_{n^2-1} \geq \tilde{\lambda}_{n^2-1} \geq \lambda_{n^2}.$$

Therefore, the second statement follows immediately from Theorem 6.10.

For the first statement, we have

$$1 \leq cn\lambda_1 \leq c^3 n^3 \lambda_1^3 \leq c^3 n^3 \lambda_1^3 + \cdots + c^3 n^3 \lambda_{n^2}^3,$$

so the inequality obtains if one takes expectations. In addition, we know from Theorem 6.8 that $\mathbb{E}[c^3 n^3 Z^3] = 1 + O(n^{-1})$, and therefore

$$\mathbb{E}[cn\lambda_1] = 1 + O(n^{-1}).$$

This proves the first statement. □

An important result for quantum information theoretic purposes is the following.

PROPOSITION 6.12. *Almost surely, in the limit $n \to \infty$, the von Neumann entropy of the matrix Z satisfies*

$$H(Z) = \begin{cases} 2\log n - \dfrac{1}{2c^2} + o(1), & \text{if } c \geq 1, \\ 2\log(cn) - \dfrac{c^2}{2} + o(1), & \text{if } 0 < c < 1. \end{cases}$$

PROOF. We use the fact that $cn\lambda_1 \geq 1$. Since $x \log x \leq x^3 - 1$ for any $x \geq 1$, we have

$$cn\lambda_1 \log(cn\lambda_1) \leq (cn\lambda_1)^3 - 1 \leq \sum_{i=1}^{n^2} (cn\lambda_i)^3 - 1.$$

Taking the expectation and using Theorem 6.8, we get

$$\mathbb{E}[cn\lambda_1 \log(cn\lambda_1)] = O(n^{-1}).$$

Similarly, we know by Theorem 6.8 that

$$\mathbb{E}[\lambda_1] = O(n^{-1}).$$

Putting this together, we obtain

$$\mathbb{E}[-\lambda_1 \log \lambda_1] = o(1).$$

We are now left with evaluating

$$\mathbb{E}[-\lambda_2 \log \lambda_2 - \cdots - \lambda_{n^2} \log \lambda_{n^2}].$$

This can be done exactly in the same way as in Corollary 4.6, and we then obtain the desired formula. □

REMARK 6.13. It is important to remark here that the estimate of Propositions 6.12 and 6.4 are asymptotically the same. This implies that in this scaling, the choice $U_1 = \overline{U_2}$ is irrelevant in the construction of counterexamples to the additivity problem. However, it remains to be checked whether this scaling indeed yields counterexamples with high probability, and this is not clear from our first-order asymptotics.

APPENDIX

In this appendix we present the complete proofs of the almost sure convergence statements in Theorems 6.3 and 6.10.

PROOF OF THEOREM 6.3, CONTINUED (ALMOST SURE CONVERGENCE). We have already proven the convergence *in moments*. To prove the almost sure convergence, it is sufficient to show that for each p, the series of covariance of the p moments is convergent. A classical application of the Borel–Cantelli lemma then suffices to complete the proof.

We start with the simplest term, $\mathbb{E}[\mathrm{tr}_{n^2}((c^2n^2Z)^p)]^2$. Since we need to compute its first two terms in the asymptotic expansion in n, we look at the subleading term (n^{-1}) of $\mathbb{E}[\mathrm{tr}_{n^2}((c^2n^2Z)^p)]$. Such terms come from permutations for which the exponent $\mathcal{P}_n$ has value $2(p-1)+1 = 2p-1$. Analyzing equations (22)–(24), we see that the permutations which "almost" saturate the bound are those which verify $\mathrm{id} \to \alpha_U = \beta_U \to \gamma$, $\mathrm{id} \to \alpha_V = \beta_V \to \gamma$ and $|\beta_U^{-1}\beta_V| = |\alpha_U^{-1}\alpha_V| = 1$. In conclusion, we have

$$\mathbb{E}[\mathrm{tr}_{n^2}((c^2n^2Z)^p)] \sim \sum_{\mathrm{id}\to\alpha_U=\beta_U=\alpha_V=\beta_V\to\gamma} c^{2\#\alpha_U} + n^{-1} \sum_{\substack{\mathrm{id}\to\alpha_U=\beta_U\to\gamma\\ \mathrm{id}\to\alpha_V=\beta_V\to\gamma\\ |\beta_U^{-1}\beta_V|=1}} c^{\#\alpha_U+\#\alpha_V} + O(n^{-2}).$$

Taking the square gives

$$\begin{aligned}\mathbb{E}[\mathrm{tr}_{n^2}((c^2n^2Z)^p)]^2 &= \Bigg[\sum_{\mathrm{id}\to\alpha_U=\beta_U=\alpha_V=\beta_V\to\gamma} c^{2\#\alpha_U}\Bigg]^2 \\ &\quad + n^{-1}2\Bigg[\sum_{\mathrm{id}\to\alpha_U=\beta_U=\alpha_V=\beta_V\to\gamma} c^{2\#\alpha_U}\Bigg]\cdot\Bigg[\sum_{\substack{\mathrm{id}\to\alpha_U=\beta_U\to\gamma\\ \mathrm{id}\to\alpha_V=\beta_V\to\gamma\\ |\beta_U^{-1}\beta_V|=1}} c^{\#\alpha_U+\#\alpha_V}\Bigg] \\ &\quad + O(n^{-2}).\end{aligned} \tag{30}$$

In order to compute the asymptotic expansion of $\mathbb{E}[(\mathrm{tr}_{n^2}((c^2n^2Z)^p)^2)]$, we must consider two copies of the diagram corresponding to the pth power of c^2n^2Z. The boxes are originally connected by the permutation

$$\gamma_2 = (p\ p-1\ \cdots\ 2\ 1)(2p\ 2p-1\ \cdots\ p+2\ p+1) \in \mathcal{S}_{2p}.$$

After counting the loops, we finds an analogous formula for the mean trace,

$$\mathbb{E}[(\mathrm{tr}_{n^2}((c^2n^2Z)^p)^2)] \sim \sum_{\alpha_U,\beta_U,\alpha_V,\beta_V\in\mathcal{S}_{2p}} n^{4p-4-\mathcal{P}_{n,2}} c^{4p-\mathcal{P}_{c,2}} \operatorname{Mob}(\alpha_U\beta_U^{-1}) \operatorname{Mob}(\alpha_V\beta_V^{-1}),$$

where

$$\mathcal{P}_{n,2} = |\alpha_U| + |\alpha_V| + |\gamma_2^{-1}\alpha_U| + |\gamma_2^{-1}\alpha_V| + |\beta_U^{-1}\beta_V| + 2|\alpha_U\beta_U^{-1}| + 2|\alpha_V\beta_V^{-1}|$$

and

$$\mathcal{P}_{c,2} = |\alpha_U| + |\alpha_V| + |\alpha_U\beta_U^{-1}| + |\alpha_V\beta_V^{-1}|.$$

Using the same inequalities and arguments as above, we find that $\mathcal{P}_{n,2} \geq 2|\gamma_2| = 4p - 4$ with equality iff $\mathrm{id}\, \alpha_U = \beta_U = \alpha_V = \beta_V \to \gamma_2$ is a geodesic. Since γ_2 contains two p-cycles, the preceding condition is equivalent to $\alpha_U = \beta_U = \alpha_V = \beta_V = \alpha \oplus \alpha'$, where $\alpha \in \mathcal{S}_p$ and $\alpha' \in \mathcal{S}\{p+1, p+2, \dots, 2p\} \simeq \mathcal{S}_p$ are such that $\mathrm{id} \to \alpha \to \gamma$ and $\mathrm{id} \to \alpha' \to \gamma$ are geodesics. Since Möbius functions vanish, this dominating term is equal to the first term in the asymptotic expansion (30) of $\mathbb{E}[\mathrm{tr}_{n^2}((c^2 n^2 Z)^p)]^2$. The term responsible for the n^{-1} contribution comes from permutations $\alpha_U, \beta_U, \alpha_V, \beta_V \in \mathcal{S}_{2p}$ such that $\mathrm{id} \to \alpha_U = \beta_U \to \gamma_2$, $\mathrm{id} \to \alpha_V = \beta_V \to \gamma_2$ and $|\beta_U^{-1}\beta_V| = |\alpha_U^{-1}\alpha_V| = 1$. Since, from the geodesic condition, $\alpha_U = \alpha'_U \oplus \alpha''_U$ and $\alpha_V = \alpha'_V \oplus \alpha''_V$, the condition $|\alpha_U^{-1}\alpha_V| = 1$ is equivalent to either

$$\alpha'_U = \alpha'_V \quad \text{and} \quad |(\alpha''_U)^{-1}\alpha''_V| = 1$$

or

$$|(\alpha'_U)^{-1}\alpha'_V| = 1 \quad \text{and} \quad \alpha''_U = \alpha''_V.$$

Summing these contributions, we find the term in n^{-1} from equation (30). Hence, the dominating (n^0) and the subdominating (n^{-1}) terms from $\mathbb{E}[\mathrm{tr}_{n^2}((c^2 n^2 Z)^p)]^2$ and $\mathbb{E}[(\mathrm{tr}_{n^2}((c^2 n^2 Z)^p)^2)]$ are equal, which implies that the general term of the series of covariances has order n^{-2}. The series is thus summable and a Borel–Cantelli-type argument completes the proof of the almost sure convergence from Theorem 6.3. □

PROOF OF THEOREM 6.10, CONTINUED (ALMOST SURE CONVERGENCE). We now prove the almost sure convergence statement of Theorem 6.10. We use the same technique as before, showing that the covariance series converges. The first step is to analyze the subleading terms (n^{-1}) in the expression of the pth moment for one copy of the channel. Recall that the exponent of n was given by the expression

$$S(\alpha, \beta, f) = |\beta\delta| + |\alpha| + 2|\alpha\beta^{-1}| + |f^{-1}(E)| + |\alpha \hat{f}^{-1}|.$$

Using the triangle inequality $|\alpha| + |\alpha \hat{f}^{-1}| \geq |\hat{f}|$, we split this minimization task into two independent problems:

$$\text{minimize} \qquad |\beta\delta| + 2|\alpha\beta^{-1}|$$

and

$$\text{minimize} \qquad |f^{-1}(E)| + |\hat{f}|.$$

The $2p - 2$ minimum in the second problem is reached for $f \equiv \mathrm{I}$; if f is different from I, it follows from the above analysis that $|f^{-1}(E)| + |\hat{f}| \geq 2p$ and thus only $f \equiv \mathrm{I}$ contributes to the subleading n^{-1} term. Moreover, a parity argument for the geodesic inequality $|\alpha| + |\alpha \hat{f}^{-1}| \geq |\hat{f}|$ implies that the permutation α must lie on the geodesic between id and $\hat{\mathrm{I}} = \gamma$. Let us now describe the couples $(\alpha, \beta) \in \mathcal{S}_{2p}^2$

such that $|\beta\delta| + 2|\alpha\beta^{-1}| = p + 1$. Since $|\beta\delta| + |\alpha\beta^{-1}| \geq |\alpha\delta| \geq p$, we need to consider two cases.

In the first case, we assume that $|\alpha\delta| = p + 1$ and $\alpha = \beta$. Since α is a geodesic permutation, $\alpha = \alpha^T \oplus \alpha^B$ and the condition $|\alpha\delta| = p + 1$ is equivalent to $|\alpha^T\alpha^B| = 1$. In conclusion, this case gives a contribution of

$$\frac{1}{n} \sum_{\substack{\mathrm{id}\to\alpha^T\to\gamma^T \\ \mathrm{id}\to\alpha^B\to\gamma^B \\ |\alpha^T\alpha^B|=1}} c^{\#\alpha^T+\#\alpha^B}.$$

In the second case, $|\alpha\delta| = p$ and $|\alpha\beta^{-1}| = 1$. This corresponds to $\alpha^T = (\alpha^B)^{-1}$, $|\alpha\beta^{-1}| = 1$ and $|\beta\delta| = p$. Since β is at distance 1 from α, $\beta = \alpha(i^s\ j^t)$ for some $i, j \in \{1, \ldots, p\}$ and $s, t \in T, B$. If $s = t$, then $\beta \notin \mathcal{V}$ and thus $|\beta\delta| \geq p$, which is impossible. We can now assume that $\beta = \alpha(i^T\ j^B)$ for some i, j. In order to have $|\beta\delta| < p$, the permutation $\beta\delta$ must have at least two fixed points. Using

$$[\alpha^T \oplus (\alpha^T)^{-1} \cdot (i^T\ j^B)](k^T) = \begin{cases} (\alpha^T(k))^T, & \text{if } k \neq i, \\ ((\alpha^T)^{-1}(j))^B, & \text{if } k = i \end{cases}$$

and

$$[\alpha^T \oplus (\alpha^T)^{-1} \cdot (i^T\ j^B)](k^B) = \begin{cases} ((\alpha^T)^{-1}(k))^B, & \text{if } k \neq j, \\ (\alpha^T(i))^T, & \text{if } k = j, \end{cases}$$

we conclude that in order to get an n^{-1} contribution, we must have $\alpha^T(i) = j$. Hence, for each geodesic permutation α, we can find p permutations β such that $|\alpha\beta^{-1}| = 1$ and $|\beta\delta| = p - 1$. We obtain a total contribution of [use Mob(transposition) $= -1$]

$$-\frac{p}{n} \sum_{\mathrm{id}\to\alpha^T\to\gamma^T} c^{2\#\alpha^T-1}.$$

Putting the first- and second-order contributions together, we obtain the asymptotic expansion for the square of the expected normalized trace:

$$\begin{aligned}\mathbb{E}[\mathrm{tr}_{n^2}(c^2n^2QZQ)^p]^2 &= \Big[\sum_{\mathrm{id}\to\alpha^T\to\gamma^T} c^{2\#\alpha^T}\Big]^2 \\ &\quad + \frac{2}{n}\Big[\sum_{\mathrm{id}\to\alpha^T\to\gamma^T} c^{2\#\alpha^T}\Big] \\ &\quad \times \Big[\sum_{\substack{\mathrm{id}\to\alpha^T\to\gamma^T \\ \mathrm{id}\to\alpha^B\to\gamma^B \\ |\alpha^T\alpha^B|=1}} c^{\#\alpha^T+\#\alpha^B} - p \sum_{\mathrm{id}\to\alpha^T\to\gamma^T} c^{2\#\alpha^T-1}\Big].\end{aligned} \tag{31}$$

Let us now analyze the second term in the expression of the covariance, $\mathbb{E}[(\mathrm{tr}_{n^2}(c^2n^2QZQ)^p)^2]$. The exponent we want to minimize in this situation is

$$S^{(2)}(\alpha,\beta,f)=|\beta\delta^{(2)}|+|\alpha|+2|\alpha\beta^{-1}|+|f^{-1}(E)|+|\alpha\hat{f}^{-1}|,$$

where α,β are permutations in $\mathcal{S}_{4p}$, and the choice function $f:\{1,\dots,p,p+1,\dots,2p\}\to\{\mathrm{I},E\}$ encodes the way the Z boxes are connected. Note, however, that in this case, the diagram under consideration has at least two connected components since we are dealing with a product of traces. Considerations similar to the ones in the proof of the convergence in moments lead to the conclusion that permutations $\alpha\in\mathcal{V}^{(2)}$ do not contribute, so we can restrain our minimization problem to the set $(\mathcal{S}_{4p}\setminus\mathcal{V}^{(2)})\times\mathcal{S}_{4p}\times\mathcal{F}_{2p}$. Using the triangle inequality $|\alpha|+|\alpha\hat{f}^{-1}|\geq|\hat{f}|$, we again split our problem into two independent parts: one minimization problem for the choice function f and another for the couple (α,β). The minimization problem for f is the same as in the single channel case, with the difference that f is now defined on a set of cardinality $2p$. The quantity $|f^{-1}(E)|+|\hat{f}|$ is minimized for $f\equiv\mathrm{I}$ and the minimum is equal to $4p-4$. Notice that in this case, the corresponding permutation $\hat{\mathrm{I}}$ has the following cycle structure: $\hat{\mathrm{I}}=\gamma^{T,1}\oplus\gamma^{T,2}\oplus\gamma^{B,1}\oplus\gamma^{B,2}$, where

$$\begin{aligned}
\gamma^{T,1}&=\bigl(p^T\ (p-1)^T\ \cdots\ 1^T\bigr);\\
\gamma^{T,2}&=\bigl((2p)^T\ (2p-1)^T\ \cdots\ (p+1)^T\bigr);\\
\gamma^{B,1}&=(1^B\ 2^B\ \cdots\ p^B);\\
\gamma^{B,2}&=\bigl((p+1)^B\ (p+2)^B\ \cdots\ (2p)^B\bigr).
\end{aligned}$$

Geodesic permutations $\mathrm{id}_{4p}\to\alpha\to\hat{\mathrm{I}}$ share the same cyclic decompositions and we can easily find the dominating term in this case:

$$\mathbb{E}[(\mathrm{tr}_{n^2}(c^2n^2QZQ)^p)^2]=\sum_{\substack{\mathrm{id}_p\to\alpha^{T,1}\to\gamma^{T,1}\simeq\gamma^T\\ \mathrm{id}\to\alpha^{T,2}\to\gamma^{T,2}\simeq\gamma^T}}c^{\#\alpha^{T,1}+\#\alpha^{T,2}}+o(1),$$

which is the same as the first term in equation (31). Let us now move on to the subleading term in the asymptotic expansion of $\mathbb{E}[(\mathrm{tr}_{n^2}(c^2n^2QZQ)^p)^2]$. As in the previous case, $f\equiv\mathrm{I}$ and contributing couples (α,β) are of two types: permutations such that $|\alpha\delta^{(2)}|=2p+1$ and $\alpha=\beta$, or couples such that $|\beta\delta^{(2)}|=2p-1$ and $|\alpha\beta^{-1}|=1$.

The analysis of the first situation is simpler: the cycle structure of the geodesic permutation α implies that $|\alpha\delta^{(2)}|=2p+|\alpha^{T,1}\alpha^{B,1}|+|\alpha^{T,2}\alpha^{B,2}|$. Hence, only one of $|\alpha^{T,1}\alpha^{B,1}|$ or $|\alpha^{T,2}\alpha^{B,2}|$ is equal to 1, the other being 0. This corresponds to a contribution of (we use the symmetry $1 \leftrightarrow 2$ of the problem)

$$\frac{2}{n}\Bigg[\sum_{\mathrm{id}\to\alpha^{T,1}\to\gamma^{T,1}} c^{2\#\alpha^{T,1}}\Bigg]\cdot\Bigg[\sum_{\substack{\mathrm{id}\to\alpha^{T,2}\to\gamma^{T,2}\\ \mathrm{id}\to\alpha^{B,2}\to\gamma^{B,2}\\ |\alpha^{T,2}\alpha^{B,2}|=1}} c^{\#\alpha^{T,2}+\#\alpha^{B,2}}\Bigg].$$

The second contribution is calculated in a similar manner to the case of a single trace. Permutations β at distance 1 from geodesic $\alpha = \alpha^{T,1}\oplus\alpha^{T,2}\oplus\alpha^{B,1}\oplus\alpha^{B,2}$ such that $\alpha^{B,1} = (\alpha^{T,1})^{-1}$ and $\alpha^{B,2} = (\alpha^{T,2})^{-1}$ are of the form $\beta = \alpha(i^s\ j^t)$. The condition $|\beta\delta^{(2)}| = 2p-1$ implies that we can choose the transposition $(i^T\ j^B)$ and that $[\alpha^{T,1}\oplus\alpha^{T,2}](i) = j$. This last condition implies that i and j have to be in the same half of the set $\{1,\dots,p,p+1,\dots,2p\}$ and, again using the symmetry between the first and the second trace, we can write the final contribution:

$$-\frac{p}{n}\Bigg[\sum_{\mathrm{id}\to\alpha^{T,1}\to\gamma^{T,1}} c^{2\#\alpha^{T,1}}\Bigg]\cdot\Bigg[\sum_{\mathrm{id}\to\alpha^{T,2}\to\gamma^{T,2}} c^{2\#\alpha^{T,2}-1}\Bigg].$$

Summing the leading (n^0) and the subleading (n^{-1}) contributions and comparing to equation (31), we find that

$$\mathbb{E}[(\mathrm{tr}_{n^2}(c^2n^2QZQ)^p)^2] - \mathbb{E}[\mathrm{tr}_{n^2}(c^2n^2QZQ)^p]^2 = O(n^{-2}).$$

The convergence of the covariance series follows, completing the proof. □

Acknowledgments. The authors would like to thank the organizers of the workshop *Thematic Program on Mathematics in Quantum Information* at the Fields Institute, where some of this work was done.

REFERENCES

[1] BHATIA, R. (1997). *Matrix Analysis*. *Graduate Texts in Mathematics* **169**. Springer, New York. MR1477662

[2] BENGTSSON, I. and ŻYCZKOWSKI, K. (2006). *Geometry of Quantum States*. Cambridge Univ. Press, Cambridge. MR2230995

[3] BOŻEJKO, M., KRYSTEK, A. D. and WOJAKOWSKI, Ł. J. (2006). Remarks on the r and Δ convolutions. *Math. Z.* **253** 177–196. MR2206642

[4] BRANDÃO, F. G. S. L. and HORODECKI, M. (2010). On Hastings' counterexamples to the minimum output entropy additivity conjecture. *Open Syst. Inf. Dyn.* **17** 31–52. MR2654951

[5] BRAUNSTEIN, S. L. (1996). Geometry of quantum inference. *Phys. Lett. A* **219** 169–174. MR1408400

[6] BRYC, W. (2008). Asymptotic normality for traces of polynomials in independent complex Wishart matrices. *Probab. Theory Related Fields* **140** 383–405. MR2365479

[7] COECKE, B. (2006). Kindergarten quantum mechanics—lecture notes. In *Quantum Theory: Reconsideration of Foundations—3*. *AIP Conf. Proc.* **810** 81–98. Amer. Inst. Phys., Melville, NY. MR2212821

[8] COLLINS, B. (2003). Moments and cumulants of polynomial random variables on unitary groups, the Itzykson–Zuber integral, and free probability. *Int. Math. Res. Not.* **17** 953–982. MR1959915

[9] COLLINS, B. and NECHITA, I. (2010). Random quantum channels I: Graphical calculus and the Bell state phenomenon. *Comm. Math. Phys.* **297** 345–370. MR2651902

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

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

[12] COLLINS, B., NECHITA, I. and ŻYCZKOWSKI, K. (2010). Random graph states, maximal flow and Fuss–Catalan distributions. *J. Phys. A Math. Theor.* **43** 275303.

[13] COLLINS, B. and ŚNIADY, P. (2006). Integration with respect to the Haar measure on unitary, orthogonal and symplectic group. *Comm. Math. Phys.* **264** 773–795. MR2217291

[14] FUKUDA, M. and KING, C. (2010). Entanglement of random subspaces via the Hastings bound. *J. Math. Phys.* **51** 042201. MR2662469

[15] FUKUDA, M., KING, C. and MOSER, D. K. (2010). Comments on Hastings' additivity counterexamples. *Comm. Math. Phys.* **296** 111–143. MR2606630

[16] GRACZYK, P., LETAC, G. and MASSAM, H. (2003). The complex Wishart distribution and the symmetric group. *Ann. Statist.* **31** 287–309. MR1962508

[17] GUIONNET, A. (2009). *Large Random Matrices: Lectures on Macroscopic Asymptotics. Lecture Notes in Math.* **1957**. Springer, Berlin. MR2498298

[18] HANLON, P. J., STANLEY, R. P. and STEMBRIDGE, J. R. (1992). Some combinatorial aspects of the spectra of normally distributed random matrices. In *Hypergeometric Functions on Domains of Positivity, Jack Polynomials, and Applications* (*Tampa, FL*, 1991). *Contemp. Math.* **138** 151–174. Amer. Math. Soc., Providence, RI. MR1199126

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

[20] HAYDEN, P. and WINTER, A. (2008). Counterexamples to the maximal p-norm multiplicity conjecture for all $p > 1$. *Comm. Math. Phys.* **284** 263–280. MR2443305

[21] JONES, V. F. R. (1999). *Planar Algebras*. Available at arXiv:math/9909027v1.

[22] MINGO, J. A. and NICA, A. (2004). Annular noncrossing permutations and partitions, and second-order asymptotics for random matrices. *Int. Math. Res. Not.* **28** 1413–1460. MR2052516

[23] NECHITA, I. (2007). Asymptotics of random density matrices. *Ann. Henri Poincaré* **8** 1521–1538. MR2374950

[24] NICA, A. and SPEICHER, R. (2006). *Lectures on the Combinatorics of Free Probability. London Mathematical Society Lecture Note Series* **335**. Cambridge Univ. Press, Cambridge. MR2266879

[25] PAGE, D. N. (1993). Average entropy of a subsystem. *Phys. Rev. Lett.* **71** 1291–1294. MR1232812

[26] SOMMERS, H.-J. and ŻYCZKOWSKI, K. (2004). Statistical properties of random density matrices. *J. Phys. A* **37** 8457–8466. MR2091254

[27] STANLEY, R. P. (1997). *Enumerative Combinatorics. Vol.* 1. *Cambridge Studies in Advanced Mathematics* **49**. Cambridge Univ. Press, Cambridge. MR1442260

[28] ŻYCZKOWSKI, K. and SOMMERS, H.-J. (2001). Induced measures in the space of mixed quantum states. *J. Phys. A* **34** 7111–7125. Quantum information and computation. MR1863143

[29] ZVONKIN, A. (1997). Matrix integrals and map enumeration: An accessible introduction. *Math*. *Comput*. *Modelling* **26** 281–304. MR1492512

DÉPARTEMENT DE MATHÉMATIQUE
ET STATISTIQUE
UNIVERSITÉ D'OTTAWA
585 KING EDWARD
OTTAWA, ONTARIO K1N6N5
CANADA
AND
CNRS
INSTITUT CAMILLE JORDAN
UNIVERSITÉ LYON 1
43 BD DU 11 NOVEMBRE 1918
69622 VILLEURBANNE
FRANCE
E-MAIL: bcollins@uottawa.ca

DÉPARTEMENT DE MATHÉMATIQUE
ET STATISTIQUE
UNIVERSITÉ D'OTTAWA
585 KING EDWARD
OTTAWA, ONTARIO K1N6N5
CANADA
E-MAIL: inechita@uottawa.ca