Commun. Math. Phys. 301, 627–659 (2011)
Digital Object Identifier (DOI) 10.1007/s00220-010-1164-y

Communications in **Mathematical Physics**

Asymptotic Infinitesimal Freeness with Amalgamation for Haar Quantum Unitary Random Matrices

Stephen Curran[1], **Roland Speicher**[2,3,★]

[1] Department of Mathematics, University of California, Los Angeles, CA 90095, USA.
E-mail: curransr@math.ucla.edu
[2] Saarland University, FR 6.1 - Mathematik, Campus E 2.4, Saarbrucken 66123, Germany.
E-mail: speicher@math.unisb.de
[3] Department of Mathematics and Statistics, Queen's University, Jeffery Hall, Kingston, Ontario K7L3N6, Canada. E-mail: speicher@mast.queensu.ca

Received: 4 January 2010 / Accepted: 27 June 2010
Published online: 21 November 2010 –

Abstract: We consider the limiting distribution of $U_N A_N U_N^*$ and B_N (and more general expressions), where A_N and B_N are $N \times N$ matrices with entries in a unital C*-algebra $\mathcal{B}$ which have limiting $\mathcal{B}$-valued distributions as $N \to \infty$, and U_N is a $N \times N$ Haar distributed quantum unitary random matrix with entries independent from $\mathcal{B}$. Under a boundedness assumption, we show that $U_N A_N U_N^*$ and B_N are asymptotically free with amalgamation over $\mathcal{B}$. Moreover, this also holds in the stronger infinitesimal sense of Belinschi-Shlyakhtenko.

We provide an example which demonstrates that this result may fail for classical Haar unitary random matrices when the algebra $\mathcal{B}$ is infinite-dimensional.

1. Introduction

One of the most important results in free probability theory is Voiculescu's asymptotic freeness for random matrices [20]. One simple form of this result is the following. Let A_N and B_N be (deterministic) $N \times N$ matrices with complex entries, and suppose that A_N and B_N have limiting distributions as $N \to \infty$ with respect to the normalized trace on $M_N(\mathbb{C})$. Let $(U_N)_{N \in \mathbb{N}}$ be a sequence of $N \times N$ unitary random matrices, distributed according to Haar measure. Then $U_N A_N U_N^*$ and B_N are asymptotically freely independent as $N \to \infty$. Moreover, when computing a fixed moment in $U_N A_N U_N^*$ and B_N, the error is $O(N^{-2})$ as $N \to \infty$ (see e.g. [10]), which can be interpreted as *asymptotic infinitesimal freeness* in the sense of Belinschi-Shlyakhtenko [6].

On the other hand, it is becoming increasingly apparent that in free probability, the roles of the classical groups are played by certain "free" quantum groups. This can most clearly be seen in the study of quantum distributional symmetries, originating with the free de Finetti theorem of Köstler and Speicher [16] and further developed in [4,12,13], in which the classical permutation, orthogonal and unitary groups are replaced by Wang's

★ Research supported by a Discovery grant from NSERC.

universal compact quantum groups [21,22]. For a general discussion of the passage from classical groups to free quantum groups, see [5].

In this paper, we will consider the limiting distribution of $U_N A_N U_N^*$ and B_N, where A_N and B_N are as above, but U_N is now a Haar distributed $N \times N$ *quantum* unitary random matrix, in the sense of Wang [21]. We will show that asymptotic (infinitesimal) freeness now holds even if the entries of A_N and B_N are allowed to take values in an arbitrary unital C*-algebra $\mathcal{B}$:

Theorem 1. *Let $\mathcal{B}$ be a unital C*-algebra and let $A_N, B_N \in M_N(\mathcal{B})$ for $N \in \mathbb{N}$. Assume that there is a finite constant C such that $\|A_N\| \leq C$, $\|B_N\| \leq C$ for all $N \in \mathbb{N}$. For each $N \in \mathbb{N}$, let U_N be a Haar distributed $N \times N$ quantum unitary random matrix, with entries independent from $\mathcal{B}$.*

(1) *Suppose that there are linear maps $\mu_A, \mu_B : \mathcal{B}\langle t\rangle \to \mathcal{B}$ such that for any $b_0, \dots, b_k \in \mathcal{B}$,*

$$\lim_{N\to\infty} \|(\mathrm{tr}_N \otimes \mathrm{id}_{\mathcal{B}})[b_0 A_N b_1 \cdots A_N b_k] - \mu_A[b_0 t b_1 \cdots t b_k]\| = 0,$$
$$\lim_{N\to\infty} \|(\mathrm{tr}_N \otimes \mathrm{id}_{\mathcal{B}})[b_0 B_N b_1 \cdots B_N b_k] - \mu_B[b_0 t b_1 \cdots t b_k]\| = 0,$$

where tr_N denotes the normalized trace on $M_N(\mathbb{C})$. Then $U_N A_N U_N^*$ and B_N are asymptotically free with amalgamation over $\mathcal{B}$.

(2) *Suppose that in addition, the limits*

$$\lim_{N\to\infty} N\,\{(\mathrm{tr}_N \otimes \mathrm{id}_{\mathcal{B}})[b_0 A_N b_1 \cdots A_N b_k] - \mu_A[b_0 t b_1 \cdots t b_k]\}$$
$$\lim_{N\to\infty} N\,\{(\mathrm{tr}_N \otimes \mathrm{id}_{\mathcal{B}})[b_0 B_N b_1 \cdots B_N b_k] - \mu_B[b_0 t b_1 \cdots t b_k]\}$$

converge in norm for any $b_0, \dots, b_k \in \mathcal{B}$. Then $U_N A_N U_N^*$ and B_N are asymptotically infinitesimally free with amalgamation over $\mathcal{B}$.

We will present more general asymptotic freeness results in Sect. 5, in particular Theorem 1 will be a special case of Corollary 5.9. We note that Theorem 5.1 holds equally well if U_N is a Haar distributed $N \times N$ quantum *orthogonal* random matrix [21], indeed it follows from the results of Banica in [1] that $U_N A_N U_N^*$ and B_N have the same joint distribution in both cases. However, the more general results given in Sect. 5 do require that we work in the unitary case.

For finite-dimensional $\mathcal{B}$, we show in Proposition 5.11 that classical Haar unitary random matrices are sufficient to obtain such a result. However, classical unitaries are in general insufficient for asymptotic freeness with amalgamation, even within the class of *approximately finite dimensional* C*-algebras, and so it is indeed necessary to allow quantum unitary transformations. We will discuss this further in the second part of Sect. 5, see in particular Example 5.12 and the remarks which follow.

We note that random matrix models for free products with amalgamation have also been considered by Brown, Dykema and Jung [8]. The difference between our frameworks is that we work with matrices whose entries take value in the algebra which we amalgamate over, while they consider random matrices with complex entries which approximate generating sets of certain amalgamated free products in distribution.

Our paper is organized as follows: Section 2 contains notations and preliminaries. Here we collect the basic notions from free and infinitesimally free probability and introduce the quantum unitary group $A_u(N)$. Section 3 contains some combinatorial results, related to the "fattening" operation on noncrossing partitions, which will be required in the sequel. In Sect. 4 we recall the Weingarten formula from [2] for computing integrals over $A_u(N)$, and prove a new estimate on the entries of the corresponding Weingarten matrix. Section 5 contains our main results, and a discussion of their failure for classical Haar unitaries.

2. Preliminaries and Notations

2.1. Free probability. We begin by recalling the basic notions of noncommutative probability spaces and distributions of random variables.

Definition 2.2. (1) *A noncommutative probability space is a pair* $(\mathcal{A}, \varphi)$*, where* $\mathcal{A}$ *is a unital algebra over* $\mathbb{C}$ *and* $\varphi : \mathcal{A} \to \mathbb{C}$ *is a linear functional such that* $\varphi(1) = 1$*. Elements in a noncommutative probability space will be called random variables.*
(2) *A* W^**-probability space* (M, τ) *is a von Neumann algebra* M *together with a faithful, normal, tracial state* τ*.*

The *joint distribution* of a family $(x_i)_{i\in I}$ of random variables in a noncommutative probability space $(\mathcal{A}, \varphi)$ is the collection of *joint moments*

$$\varphi(x_{i_1} \cdots x_{i_k})$$

for $k \in \mathbb{N}$ and $i_1, \ldots, i_k \in I$. This is nicely encoded in the linear functional $\varphi_x : \mathbb{C}\langle t_i | i \in I\rangle \to \mathbb{C}$ determined by

$$\varphi_x(p) = \varphi(p(x))$$

for $p \in \mathbb{C}\langle t_i | i \in I\rangle$, where $p(x)$ means of course to replace t_i by x_i for each $i \in I$.

These definitions have natural "operator-valued" extensions given by replacing $\mathbb{C}$ by a more general algebra of scalars, which we now recall.

Definition 2.3. *An operator-valued probability space* $(\mathcal{A}, E : \mathcal{A} \to \mathcal{B})$ *consists of a unital algebra* $\mathcal{A}$*, a subalgebra* $1 \in \mathcal{B} \subset \mathcal{A}$*, and a conditional expectation* $E : \mathcal{A} \to \mathcal{B}$*, i.e.,* E *is a linear map such that* $E[1] = 1$ *and*

$$E[b_1 a b_2] = b_1 E[a] b_2$$

for all $b_1, b_2 \in \mathcal{B}$ *and* $a \in \mathcal{A}$*.*

Example 2.4. Let $\mathcal{B}$ be a unital algebra over $\mathbb{C}$, and let $M_n(\mathcal{B}) = M_n(\mathbb{C}) \otimes \mathcal{B}$ be the algebra of $n \times n$ matrices over $\mathcal{B}$, with the natural inclusion of $\mathcal{B}$ as $I_n \otimes \mathcal{B}$. Let $\mathrm{tr}_n = n^{-1}\mathrm{Tr}_n$ denote the normalized trace on $M_n(\mathbb{C})$. Then $(M_n(\mathcal{B}), \mathrm{tr} \otimes \mathrm{id}_{\mathcal{B}})$ is a $\mathcal{B}$-valued probability space. Note that if $B = (b_{ij})_{i,j=1}^n \in M_n(\mathcal{B})$,

$$(\mathrm{tr}_n \otimes \mathrm{id}_{\mathcal{B}})(B) = \frac{1}{n}\sum_{i=1}^n b_{ii}.$$

The *$\mathcal{B}$-valued joint distribution* of a family $(x_i)_{i\in I}$ of random variables in an operator-valued probability space $(\mathcal{A}, E : \mathcal{A} \to \mathcal{B})$ is the collection of *$\mathcal{B}$-valued joint moments*

$$E[b_0 x_{i_1} \cdots x_{i_k} b_k]$$

for $k \in \mathbb{N}$, $i_1, \dots, i_k \in I$ and $b_0, \dots, b_k \in \mathcal{B}$. Again this is conveniently encoded in the $\mathcal{B}$-linear functional $E_x : \mathcal{B}\langle t_i | i \in I\rangle \to \mathcal{B}$ determined by

$$E_x[p] = E[p(x)]$$

for $p \in \mathcal{B}\langle t_i | i \in I\rangle$, the algebra of noncommutative polynomials with coefficients in $\mathcal{B}$.

Definition 2.5. *Let $(\mathcal{A}, E : \mathcal{A} \to \mathcal{B})$ be an operator-valued probability space, and let $(\mathcal{A}_i)_{i\in I}$ be a collection of subalgebras $\mathcal{B} \subset \mathcal{A}_i \subset A$. The algebras are said to be free with amalgamation over $\mathcal{B}$, or freely independent with respect to E, if*

$$E[a_1 \cdots a_k] = 0$$

whenever $E[a_j] = 0$ for $1 \le j \le k$ and $a_j \in \mathcal{A}_{i_j}$ with $i_j \ne i_{j+1}$ for $1 \le j < k$.

We say that subsets $\Omega_i \subset \mathcal{A}$ are free with amalgamation over $\mathcal{B}$ if the subalgebras $\mathcal{A}_i$ generated by $\mathcal{B}$ and Ω_i are freely independent with respect to E.

Remark 2.6. Voiculescu first defined freeness with amalgamation, and developed its basic theory in [19]. Freeness with amalgamation also has a rich combinatorial structure, developed in [18], which we now recall. For further information on the combinatorial theory of free probability, the reader is referred to the text [17].

Definition 2.7. (1) *A partition π of a set S is a collection of disjoint, non-empty sets $V_1, \dots, V_r$ such that $V_1 \cup \cdots \cup V_r = S$. $V_1, \dots, V_r$ are called the blocks of π, and we set $|\pi| = r$. If $s, t \in S$ are in the same block of π, we write $s \sim_\pi t$. The collection of partitions of S will be denoted $\mathcal{P}(S)$, or in the case that $S = \{1, \dots, k\}$ by $\mathcal{P}(k)$.*

(2) *Given $\pi, \sigma \in \mathcal{P}(S)$, we say that $\pi \le \sigma$ if each block of π is contained in a block of σ. There is a least element of $\mathcal{P}(S)$ which is larger than both π and σ, which we denote by $\pi \vee \sigma$.*

(3) *If S is ordered, we say that $\pi \in \mathcal{P}(S)$ is non-crossing if whenever V, W are blocks of π and $s_1 < t_1 < s_2 < t_2$ are such that $s_1, s_2 \in V$ and $t_1, t_2 \in W$, then $V = W$. The non-crossing partitions can also be defined recursively, a partition $\pi \in \mathcal{P}(S)$ is non-crossing if and only if it has a block V which is an interval, such that $\pi \backslash V$ is a non-crossing partition of $S \backslash V$. The set of non-crossing partitions of S is denoted by $NC(S)$, or by $NC(k)$ in the case that $S = \{1, \dots, k\}$.*

(4) *Given $\pi, \sigma \in NC(S)$, the join $\pi \vee \sigma$ taken in $\mathcal{P}(S)$ may not be non-crossing. However, there is a least element of $NC(S)$ which is larger than π and σ, which we will denote by $\pi \vee_{nc} \sigma$. Note that in this paper we will always use $\pi \vee \sigma$ to denote the join in $\mathcal{P}(S)$, even when π, σ are assumed noncrossing.*

(5) *Given $i_1, \dots, i_k$ in some index set I, we denote by* $\ker \mathbf{i}$ *the element of $\mathcal{P}(k)$ whose blocks are the equivalence classes of the relation*

$$s \sim t \Leftrightarrow i_s = i_t.$$

Note that if $\pi \in \mathcal{P}(k)$, then $\pi \le \ker \mathbf{i}$ is equivalent to the condition that whenever s and t are in the same block of π, i_s must equal i_t.

(6) *With 0_n and 1_n we will denote the smallest and largest element, respectively, in $\mathcal{P}(n)$; i.e., 0_n has n blocks, each consisting of one element, and 1_n has only one block. Of course, both 0_n and 1_n are in $NC(n)$.*

Definition 2.8. *Let $(\mathcal{A}, E : \mathcal{A} \to \mathcal{B})$ be an operator-valued probability space.*

(1) *A $\mathcal{B}$-functional is a n-linear map $\rho : \mathcal{A}^n \to \mathcal{B}$ such that*

$$\rho(b_0 a_1 b_1, a_2 b_2, \dots, a_n b_n) = b_0 \rho(a_1, b_1 a_2, \dots, b_{n-1} a_n) b_n$$

for all $b_0, \dots, b_n \in \mathcal{B}$ and $a_1, \dots, a_n \in \mathcal{A}$. Equivalently, ρ is a linear map from $\mathcal{A}^{\otimes_B n}$ to $\mathcal{B}$, where the tensor product is taken with respect to the obvious $\mathcal{B}$-$\mathcal{B}$-bi-module structure on $\mathcal{A}$.

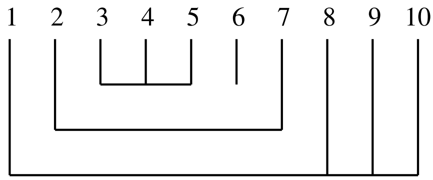
(2) *For each $k \in \mathbb{N}$, let $\rho^{(k)} : \mathcal{A}^k \to \mathcal{B}$ be a $\mathcal{B}$-functional. For $n \in \mathbb{N}$ and $\pi \in NC(n)$, we define a $\mathcal{B}$-functional $\rho^{(\pi)} : \mathcal{A}^n \to \mathcal{B}$ recursively as follows: If $\pi = 1_n$ is the partition containing only one block, we set $\rho^{(\pi)} = \rho^{(n)}$. Otherwise let $V = \{l+1, \dots, l+s\}$ be an interval of π and define*

$$\rho^{(\pi)}[a_1, \dots, a_n] = \rho^{(\pi \setminus V)}[a_1, \dots, a_l \rho^{(s)}(a_{l+1}, \dots, a_{l+s}), a_{l+s+1}, \dots, a_n]$$

for $a_1, \dots, a_n \in \mathcal{A}$.

Example 2.9. Let $(\mathcal{A}, E : \mathcal{A} \to \mathcal{B})$ be an operator-valued probability space, and for $k \in \mathbb{N}$ let $\rho^{(k)} : \mathcal{A}^k \to \mathcal{B}$ be a $\mathcal{B}$-functional as above. If

$$\pi = \{\{1, 8, 9, 10\}, \{2, 7\}, \{3, 4, 5\}, \{6\}\} \in NC(10),$$



then the corresponding $\rho^{(\pi)}$ is given by

$$\rho^{(\pi)}[a_1, \dots, a_{10}] = \rho^{(4)}(a_1 \cdot \rho^{(2)}(a_2 \cdot \rho^{(3)}(a_3, a_4, a_5), \rho^{(1)}(a_6) \cdot a_7), a_8, a_9, a_{10}).$$

Remark 2.10. Note that if $\mathcal{B}$ is commutative, then

$$\rho^{(\pi)}[a_1, \dots, a_n] = \prod_{V \in \pi} \rho(V)[a_1, \dots, a_n],$$

where if $V = (i_1 < \cdots < i_s)$ is a block of π, we set

$$\rho(V)[a_1, \dots, a_n] = \rho^{(s)}[a_{i_1}, \dots, a_{i_s}].$$

Definition 2.11. *Let $(\mathcal{A}, E : \mathcal{A} \to \mathcal{B})$ be an operator-valued probability space.*

(1) *For $k \in \mathbb{N}$, define the B-valued moment functions $E^{(k)} : \mathcal{A}^k \to \mathcal{B}$ by*

$$E^{(k)}[a_1, \dots, a_k] = E[a_1 \cdots a_k].$$

(2) *The operator-valued free cumulants $\kappa_E^{(k)} : \mathcal{A}^k \to \mathcal{B}$ are the $\mathcal{B}$-functionals defined by the moment-cumulant formula:*

$$E[a_1 \cdots a_n] = \sum_{\pi \in NC(n)} \kappa_E^{(\pi)}[a_1, \dots, a_n]$$

for $n \in \mathbb{N}$ and $a_1, \dots, a_n \in \mathcal{A}$.

Note that the right hand side of the moment-cumulant formula above is equal to $\kappa_E^{(n)}(a_1, \dots, a_n)$ plus products of lower order terms and hence can be solved recursively for $\kappa_E^{(n)}$. In fact the cumulant functions can be solved from the moment functions by the following formula from [18]: for each $n \in \mathbb{N}$, $\pi \in NC(n)$ and $a_1, \dots, a_n \in \mathcal{A}$,

$$\kappa_E^{(\pi)}[a_1, \dots, a_n] = \sum_{\substack{\sigma \in NC(n) \\ \sigma \leq \pi}} \mu_n(\sigma, \pi) E^{(\sigma)}[a_1, \dots, a_n],$$

where μ_n is the *Möbius function* on the partially ordered set $NC(n)$. The Möbius function $\mu_n(\sigma, \pi)$ is defined to be 0 unless $\sigma \leq \pi$, is 1 if $\sigma = \pi$, and for $\sigma < \pi$ is given by

$$-1 + \sum_{l \geq 1} (-1)^{l+1} \#\{(\nu_1, \dots, \nu_l) \in NC(n)^l : \sigma < \nu_1 < \cdots < \nu_l < \pi\}.$$

The key relation between operator-valued free cumulants and freeness with amalgamation is that freeness can be characterized in terms of the "vanishing of mixed cumulants".

Theorem 2.12 ([18]). Let $(\mathcal{A}, E : \mathcal{A} \to \mathcal{B})$ be an operator-valued probability space, and let $(\mathcal{A}_i)_{i \in I}$ be a collection of subalgebras $\mathcal{B} \subset \mathcal{A}_i \subset \mathcal{A}$. Then the family $(\mathcal{A}_i)_{i \in I}$ is free with amalgamation over $\mathcal{B}$ if and only if

$$\kappa_E^{(\pi)}[a_1, \dots, a_n] = 0$$

whenever $a_j \in \mathcal{A}_{i_j}$ for $1 \leq j \leq n$ and $\pi \in NC(n)$ is such that $\pi \not\leq \ker \mathbf{i}$.

2.13. Infinitesimal free probability. We will now introduce the notions of operator-valued infinitesimal probability spaces and infinitesimal freeness with amalgamation. This is a straightforward generalization of the framework of [6], and we refer the reader to that paper for further discussion of infinitesimal freeness and its relation to the type B free independence of Biane, Nica and Goodman [7]. See [14] for a more combinatorial treatment of infinitesimal freeness.

Definition 2.14. (1) *If $\mathcal{B}$ is a unital algebra, a $\mathcal{B}$-valued infinitesimal probability space is a triple $(\mathcal{A}, E, E')$ where $\mathcal{A}$ is a unital algebra which contains $\mathcal{B}$ as a unital subalgebra and E, E' are $\mathcal{B}$-linear maps from $\mathcal{A}$ to $\mathcal{B}$ such that $E[1] = 1$ and $E'[1] = 0$.*

(2) *Let $(\mathcal{A}, E, E')$ be a $\mathcal{B}$-valued infinitesimal probability space, and let $(\mathcal{A}_i)_{i \in I}$ be a collection of subalgebras $\mathcal{B} \subset \mathcal{A}_i \subset \mathcal{A}$. The algebras are said to be infinitesimally free with amalgamation over $\mathcal{B}$, or infinitesimally free with respect to (E, E'), if*
   (i) *$(\mathcal{A}_i)_{i \in I}$ are freely independent with respect to E.*
   (ii) *For any $a_1, \dots, a_k$ so that $a_j \in \mathcal{A}_{i_j}$ for $1 \leq j \leq k$ with $i_j \neq i_{j+1}$, we have*

$$E'[(a_1 - E[a_1]) \cdots (a_k - E[a_k])] = \sum_{j=1}^{k} E\left[(a_1 - E[a_1]) \cdots \left(E'[a_j]\right) \cdots (a_k - E[a_k])\right].$$

We say that subsets $(\Omega_i)_{i\in I}$ are infinitesimally free with amalgamation over $\mathcal{B}$ if the subalgebras $\mathcal{A}_i$ generated by $\mathcal{B}$ and Ω_i are infinitesimally free with respect to (E, E').

Remark 2.15. The motivating example is given by a family $(A_i(s))_{i\in I}$ of $\mathcal{B}$-valued random variables for $s > 0$ which are free "up to $o(s)$" as $s \to 0$. This is made precise in the next proposition. Note that there we make the notion "free up to $o(s)$" precise by comparing the family $(A_i(s))_{i\in I}$ with a family $(a_i(s))_{i\in I}$ which is free for all s. Infinitesimal freeness will then occur at $s = 0$ (both for the A_i and the a_i). Since 0 is not necessarily in K, we define the states E and E' on the free algebra $\mathcal{A} := \mathcal{B}\langle A_i | i \in I\rangle$ generated by non-commuting indeterminates $A_i \hat{=} A_i(0) \hat{=} a_i(0)$.

Proposition 2.16. *Let $\mathcal{B}$ be a unital C^*-algebra and K a subset of $\mathbb{R}$ for which 0 is an accumulation point. Suppose that for each $s \in K$ we have a $\mathcal{B}$-valued probability space $(\mathcal{A}(s), E_s : \mathcal{A}(s) \to \mathcal{B})$, where $\mathcal{A}(s)$ is a unital C^*-algebra which contains $\mathcal{B}$ as a unital subalgebra and E_s is contractive. Furthermore, suppose that, for each $s \in K$, there are variables $(A_i(s))_{i\in I}$ belonging to $\mathcal{A}(s)$ such that the following hold:*

(1) *There are $\mathcal{B}$-linear maps $E, E' : \mathcal{B}\langle A_i | i \in I\rangle \to \mathcal{B}$ such that*

$$E[p(A)] = \lim_{s\to 0} E_s\left[p(A(s))\right],$$
$$E'[p(A)] = \lim_{s\to 0} \frac{1}{s}\left\{E_s[p(A(s))] - E[p]\right\},$$

for $p \in \mathcal{B}\langle t_i | i \in I\rangle$, where the limits hold in norm.

(2) *For each $i \in I$,*

$$\limsup_{s\to 0} \|A_i(s)\| < \infty.$$

Let $I = \bigcup_{j\in J} I_j$ be a partition of I. For $s \in K$, let $(a_i(s))_{i\in I}$ be a family in some $\mathcal{B}$-valued probability space $(\mathcal{C}, F : \mathcal{C} \to \mathcal{B})$ and suppose that

(1) *For any $j \in J$, $p \in \mathcal{B}\langle t_i | i \in I_j\rangle$, and $s \in K$,*

$$E_s[p(A(s))] = F[p(a(s))].$$

(2) *The sets $(\{a_i(s) | s \in K, i \in I_j\})_{j\in J}$ are free with respect to F.*
(3) *For any $p \in \mathcal{B}\langle t_i | i \in I\rangle$ we have*

$$\|E_s[p(A(s))] - F[p(a(s))]\| = o(s) \quad (as\ s \to 0).$$

Then the sets $(\{A_i | i \in I_j\})_{j\in J} \subset \mathcal{B}\langle A_i | i \in I\rangle$ are infinitesimally free with respect to (E, E').

Proof. Since E, E' only depend on the distribution of the variables $A_i(s)$ up to first order, it clearly suffices to assume that the sets $(\{A_i(s) : i \in I_j\})_{j\in J}$ are freely independent with respect to E_s for all $s \in K$. It is then clear that the sets $(\{A_i : i \in I_j\})_{j\in J} \subset \mathcal{B}\langle A_i | i \in I\rangle$ are free with respect to E, so it suffices to show that E' satisfies condition (ii) of Definition 2.14. Let $j_1 \neq \cdots \neq j_k$ in J and $p_l \in \mathcal{B}\langle t_i | i \in I_{j_l}\rangle$ for $1 \leq l \leq k$, and consider

$$\begin{aligned} &E'\left[(p_1(A) - E[p_1(A)]) \cdots (p_k(A) - E[p_k(A)])\right] \\ &= \lim_{s\to 0} \frac{1}{s} \{E_s\left[(p_1(A(s)) - E[p_1(A)]) \cdots (p_k(A(s)) - E[p_k(A)])\right] \\ &\quad - E\left[(p_1(A) - E[p_1(A)]) \cdots (p_k(A) - E[p_k(A)])\right]\} \\ &= \lim_{s\to 0} \frac{1}{s} \{E_s\left[(p_1(A(s)) - E[p_1(A)]) \cdots (p_k(A(s)) - E[p_k(A)])\right]\}, \end{aligned}$$

where we have used freeness with respect to E. Rewrite this expression as

$$\begin{aligned} &\lim_{s\to 0} \frac{1}{s} \{E_s\left[((p_1(A(s)) - E_s[p_1(A(s))]) + (E_s[p_1(A(s))] - E[p_1(A)]))\right. \\ &\quad \left.\cdots ((p_k(A(s)) - E_s[p_k(A(s))]) + (E_s[p_k(A(s))] - E[p_k(A)]))\right]\}, \end{aligned}$$

and consider the terms which appear in the expansion. First observe that

$$\|E_s[p_l(A(s))] - E[p_l(A)]\|$$

is $O(s)$ for $1 \leq l \leq k$. By the boundedness assumption on the norms of $A_i(s)$, and the contractivity of E_s, it follows that those terms involving more than one expression $(E_s[p_l(A(s))] - E[p_l(A)])$ vanish in the limit.

The term involving none of these expressions is

$$E_s\left[(p_1(A(s)) - E_s[p_1(A(s))]) \cdots (p_k(A(s)) - E_s[p_k(A(s))])\right]$$

which is zero by freeness.

So we are left to consider only the terms involving one such expression, which gives

$$\begin{aligned} &\sum_{l=1}^{k} \lim_{s\to 0} \frac{1}{s} \{E_s\left[(p_1(A(s)) - E_s[p_1(A(s))])\right. \\ &\quad \left.\cdots (E_s[p_l(A(s))] - E[p_l(A)]) \cdots (p_k(A(s)) - E_s[p_k(A(s))])\right]\}, \end{aligned}$$

which again by invoking the boundedness assumptions on $A_i(s)$ and contractivity of E_s, converges to

$$\sum_{l=1}^{k} E\left[(p_1(A) - E[p_1(A)]) \cdots E'[p_l(A)] \cdots (p_k(A) - E[p_k])\right]$$

as desired.

2.17. Quantum unitary group. We now recall the definition of the quantum unitary group from [21], which is a compact quantum group in the sense of Woronowicz [23].

Definition 2.18. *$A_u(n)$ is the universal C^*-algebra generated by $\{U_{ij} : 1 \leq i, j \leq n\}$ such that the matrix $U = (U_{ij}) \in M_n(A_u(n))$ is unitary. $A_u(n)$ is a C^*-Hopf algebra with comultiplication, counit and antipode given by*

$$\begin{aligned} \Delta(U_{ij}) &= \sum_{k=1}^{n} U_{ik} \otimes U_{kj} \\ \epsilon(U_{ij}) &= \delta_{ij} \\ S(U_{ij}) &= U_{ji}^*. \end{aligned}$$

The existence of these maps is given by the universal property of $A_u(n)$.

Remark 2.19. A fundamental result of Woronowicz [23] guarantees the existence of a unique *Haar state* $\psi_n : A_u(n) \to \mathbb{C}$ which is left and right invariant in the sense that

$$(\psi_n \otimes \mathrm{id})\Delta(a) = \psi_n(a)1_{A_u(n)} = (\mathrm{id} \otimes \psi_n)\Delta(a)$$

for $a \in A_u(n)$. We will discuss this further in Sect. 4.

Wang also introduced the free product operation on compact quantum groups in [21]. We will use $A_u(n)^{*\infty}$ to denote the C*-algebraic free product (with amalgamation over $\mathbb{C}$) of countably many copies of $A_u(n)$. $A_u(n)^{*\infty}$ has a natural compact quantum group structure, given in Corollary 3.7 of [21]. The reader is referred to that paper for details, the only properties which we will use are the following:

(1) $A_u(n)^{*\infty}$ is generated (as a C*-algebra) by elements $\{U(l)_{ij} : l \in \mathbb{N}, 1 \leq i, j \leq n\}$, such that $U(l) \in M_n(A_u(n)^{*\infty})$ is unitary.
(2) The sets $(\{U(l)_{ij} : 1 \leq i, j \leq n\})_{l\in\mathbb{N}}$ are freely independent with respect to the Haar state $\psi_n^{*\infty}$ on $A_u(n)^{*\infty}$, and for each $l \in \mathbb{N}$, $(U(l)_{ij})$ has the same joint distribution in $(A_u(n)^{*\infty}, \psi_n^{*\infty})$ as (U_{ij}) in $(A_u(n), \psi_n)$. See Proposition 3.3 and Theorem 3.4 of [21].

3. Some Combinatorial Results

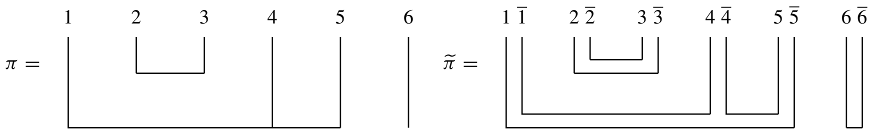
In this section we introduce several operations on partitions and prove some basic results which will be required throughout the remainder of the paper.

Notation 3.1. (1) Given $\pi \in NC(m)$, we define $\widetilde{\pi} \in NC_2(2m)$ as follows: For each block $V = (i_1, \ldots, i_s)$ of π, we add to $\widetilde{\pi}$ the pairings $(2i_1 - 1, 2i_s), (2i_1, 2i_2 - 1), \ldots, (2i_{s-1}, 2i_s - 1)$.
(2) Given $\pi \in NC(m)$, we define $\hat{\pi} \in NC(2m)$ by partitioning the m-pairs $(1, 2), (3, 4), \ldots, (2m - 1, 2m)$ according to π.
(3) Given $\pi, \sigma \in \mathcal{P}(m)$, we define $\pi \wr \sigma \in \mathcal{P}(2m)$ to be the partition obtained by partitioning the odd numbers $\{1, 3, \ldots, 2m - 1\}$ according to π and the even numbers $\{2, 4, \ldots, 2m\}$ according to σ.
(4) Given $\pi \in \mathcal{P}(m)$, let $\overleftarrow{\pi}$ denote the partition obtained by shifting k to $k - 1$ for $1 < k \leq m$ and sending 1 to m, i.e.,

$$s \sim_{\overleftarrow{\pi}} t \iff (s + 1) \sim_\pi (t + 1),$$

where we count modulo m on the right hand side. Likewise we let $\overrightarrow{\pi}$ denote the partition obtained by shifting k to $k + 1$ for $1 \leq k < m$ and sending m to 1.

Remark 3.2. The map $\pi \mapsto \widetilde{\pi}$ is easily seen to be a bijection, and corresponds to the well-known "fattening" operation. The following example shows this for $\pi = \{\{1, 4, 5\}, \{2, 3\}, \{6\}\}$.



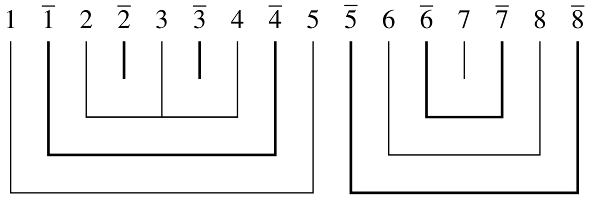
There is a simple description of the inverse, it sends $\sigma \in NC_2(2m)$ to the partition $\tau \in NC(m)$ such that $\sigma \vee \hat{0}_m = \hat{\tau}$, where $\hat{0}_m = \{\{1,2\},\ldots,\{2m-1,2m\}\}$. Thus we have for $\pi \in NC(m)$,

$$\hat{\pi} = \widetilde{\pi} \vee \hat{0}_m.$$

Note also that $\hat{0}_m = \widetilde{0}_m$ and that $\hat{1}_m = 1_{2m}$.

Definition 3.3. *Let $\pi \in NC(m)$. The Kreweras complement $K(\pi)$ is the largest partition in $NC(m)$ such that $\pi \wr K(\pi) \in NC(2m)$.*

Example 3.4. If $\pi = \{\{1,5\},\{2,3,4\},\{6,8\},\{7\}\}$ then $K(\pi) = \{\{1,4\},\{2\},\{3\},\{5,8\},\{6,7\}\}$, which can be seen follows:



The following lemma provides the relationship between the Kreweras complement on $NC(m)$ and the map $\pi \mapsto \widetilde{\pi}$.

Lemma 3.5. *If $\pi \in NC(m)$, then*

$$\widetilde{K(\pi)} = \overleftarrow{\widetilde{\pi}}.$$

Proof. We will prove this by induction on the number of blocks of π. If $\pi = 1_m$ has one block, the result is trivial from the definitions.

Suppose now that $V = \{l+1,\ldots,l+s\}$ is a block of π, $l \geq 1$. First note that $\widetilde{\pi}$ is obtained by taking $\widetilde{\pi\backslash V}$ then adding the pairs $(2l+1, 2(l+s))$, $(2l+2, 2l+3), \ldots, (2(l+s)-2, 2(l+s)-1)$.

Observe that $K(\pi)$ is obtained by taking $K(\pi\backslash V)$, adding singletons $\{l+1\},\ldots,\{l+s-1\}$, then placing $l+s$ in the block containing l. It follows that $\widetilde{K(\pi)}$ is the partition obtained by taking $\widetilde{K(\pi\backslash V)}$, which by induction is $\overleftarrow{\widetilde{\pi\backslash V}}$, then moving the leg connected to $2l$ to $2(l+s)$ and adding the pairs $(2l, 2(l+s)-1)$, $(2l+1, 2l+2), \ldots, (2(l+s)-3, 2(l+s)-2)$. The result now follows.

We will also need the following relationship between $\pi \mapsto \widetilde{\pi}$ and the Kreweras complement on $NC(2m)$. This is a generalization of the relation

$$K(\hat{\pi}) = K(\widetilde{0}_m \vee \widetilde{\pi}) = 0_m \wr K(\pi) \qquad (\pi \in NC(m)),$$

which is obvious from the definition of $\hat{\pi}$.

Lemma 3.6. *If $\pi, \sigma \in NC(m)$ and $\sigma \leq \pi$, then $\widetilde{\sigma} \vee \widetilde{\pi} \in NC(2m)$ and*

$$K(\widetilde{\sigma} \vee \widetilde{\pi}) = \sigma \wr K(\pi).$$

Proof. We will prove this by induction on the number of blocks of π. First suppose that $\pi = 1_m$, then we have

$$\widetilde{\sigma} \vee \widetilde{\pi} = \overrightarrow{\overleftarrow{\sigma} \vee \overleftarrow{\pi}} = \overrightarrow{\widehat{K(\sigma)} \vee \hat{0}_m} = \overrightarrow{\widehat{K(\sigma)}}$$

is noncrossing. Moreover,

$$K(\widetilde{\sigma} \vee \widetilde{\pi}) = K\left(\overrightarrow{\widehat{K(\sigma)}}\right) = \overrightarrow{0_m \wr K^2(\sigma)},$$

where for the last equality we used the equation for $K(\hat{\pi})$ mentioned before Lemma 3.6 and the fact that the Kreweras complement commutes with shifting. But, by [17, Exercise 9.23], we have that $K^2(\sigma) = \overleftarrow{\sigma}$ and thus we finally get

$$K(\widetilde{\sigma} \vee \widetilde{\pi}) = \overrightarrow{0_m \wr \overleftarrow{\sigma}} = \sigma \wr 0_m.$$

Now suppose that $V = \{l+1, \ldots, l+s\}$, $l \geq 1$ is an interval of π. Observe that $\widetilde{\sigma} \vee \widetilde{\pi}$ is the partition obtained by partitioning $\{1, \ldots, 2l\} \cup \{2(l+s)+1, \ldots, 2m\}$ according to $\widetilde{\sigma \backslash \sigma|_V \vee \pi \backslash V}$, and $\{2l+1, \ldots, 2(l+s)\}$ according to $\widetilde{\sigma|_V \vee 1_V}$. It follows that $\widetilde{\sigma} \vee \widetilde{\pi}$ is noncrossing and that $K(\widetilde{\sigma} \vee \widetilde{\pi})$ is the partition obtained by partitioning $\{1, \ldots, 2l\} \cup \{2(l+s)+1, \ldots, 2m\}$ according to $K(\widetilde{\sigma \backslash \sigma|_V \vee \pi \backslash V})$ and $\{2l+1, \ldots, 2(l+s)\}$ according to $K(\widetilde{\sigma|_V \vee 1_V})$, then joining the blocks containing $2l$ and $2(l+s)$. On the other hand, $K(\pi)$ is equal to the partition obtained by taking $K(\pi \backslash V)$ then adding $\{l+1\}, \ldots, \{l+s-1\}$ and joining $l+s$ to l, and the result now follows by induction.

We will need to compare the number of blocks in the join of two partitions before and after fattening. For this purpose we will use the following *linearization lemma* of Kodiyalam-Sunder [15] and, independently, Chen-Przytycki [9]. Note that the notation $S \mapsto \widetilde{S}$ used in [15] corresponds to the inverse of the fattening procedure $\pi \mapsto \widetilde{\pi}$ used here.

Theorem 3.7 ([15]). Let $\pi, \sigma \in NC(m)$. Then

$$|\widetilde{\pi} \vee \widetilde{\sigma}| = m + 2|\pi \vee \sigma| - |\pi| - |\sigma|.$$

In particular, if $\sigma \leq \pi$ then

$$|\widetilde{\pi} \vee \widetilde{\sigma}| = m + |\pi| - |\sigma|.$$

We now introduce some special classes of noncrossing partitions and prove some basic results. These are related to integration on the quantum unitary group via the Weingarten formula to be discussed in the next section.

Notation 3.8. Let $\epsilon_1, \ldots, \epsilon_{2m} \in \{1, *\}$.

(1) $NC_h^\epsilon(2m)$ denote the set of partitions $\pi \in NC(2m)$ such that each block V of π has an even number of elements, and $\epsilon|_V$ is alternating, i.e., $\epsilon|_V = 1 * 1 * \cdots 1*$ or $*1 * 1 \cdots * 1$.
(2) $NC_2^\epsilon(2m)$ will denote the collection of $\pi \in NC_2(2m)$ such that each pair in π connects a 1 with a $*$, i.e.,

$$s \sim_\pi t \Rightarrow \epsilon_s \neq \epsilon_t.$$

(3) $NC^\epsilon(m)$ will denote the collection of $\pi \in NC(m)$ such that $\widetilde{\pi} \in NC_2^\epsilon(m)$.

Lemma 3.9. *Let $\epsilon_1, \ldots, \epsilon_{2m} \in \{1, *\}$. If $\sigma, \pi \in NC^\epsilon(m)$ and $\sigma \leq \pi$, then $\widetilde{\sigma} \vee \widetilde{\pi}$ is in $NC_h^\epsilon(2m)$. Conversely, if $\tau \in NC_h^\epsilon(2m)$ then there are unique $\sigma, \pi \in NC^\epsilon(m)$ such that $\sigma \leq \pi$ and $\tau = \widetilde{\sigma} \vee \widetilde{\pi}$.*

Proof. First suppose that $\tau \in NC_h^\epsilon(2m)$. Since each block of τ has an even number of elements, we have $K(\tau) = \sigma \wr K(\pi)$ for some $\sigma, \pi \in NC(m)$ such that $\sigma \leq \pi$. By Lemma 3.6 we have $\tau = \widetilde{\sigma} \vee \widetilde{\pi}$, and this clearly determines σ and π uniquely. If V is a block of τ, then $\epsilon|_V$ is alternating and hence $\widetilde{\pi}|_V, \widetilde{\sigma}|_V \in NC_2^\epsilon(V)$. It follows that $\pi, \sigma \in NC^\epsilon(m)$.

Conversely, let $\sigma, \pi \in NC^\epsilon(m)$ such that $\sigma \leq \pi$. Let $\hat{\epsilon} = (\epsilon_1, \epsilon_1, \epsilon_2, \epsilon_2, \ldots, \epsilon_{2m}, \epsilon_{2m})$. Observe that if $\tau \in NC(2m)$, then $\tau \in NC_h^\epsilon(2m)$ if and only if $\widetilde{\tau} \in NC_2^{\hat{\epsilon}}(4m)$.

So let $\tau = \widetilde{\sigma} \vee \widetilde{\pi}$, we need to show $\widetilde{\tau} \in NC_2^{\hat{\epsilon}}(4m)$. Now

$$\overleftarrow{\widetilde{\tau}} = \widetilde{K(\tau)} = \sigma \wr \widetilde{K(\pi)},$$

where we have applied Lemmas 3.5 and 3.6. In other words, $\overleftarrow{\widetilde{\tau}}$ is the partition given by partitioning $\{1, 2, 5, 6, \ldots, 4m-3, 4m-2\}$ according to $\widetilde{\sigma}$ and $\{3, 4, 7, 8, \ldots, 4m-1, 4m\}$ according to $\widetilde{K(\pi)} = \overleftarrow{\widetilde{\pi}}$. Now since $\sigma, \pi \in NC^\epsilon(m)$, it follows that $\overleftarrow{\widetilde{\tau}} \in NC_2^{\overleftarrow{\hat{\epsilon}}}(4m)$, where $\overleftarrow{\hat{\epsilon}} = (\epsilon_1, \epsilon_2, \epsilon_2, \ldots, \epsilon_{2m}, \epsilon_{2m}, \epsilon_1)$, and hence $\widetilde{\tau} \in NC_2^{\hat{\epsilon}}(4m)$.

Lemma 3.10. *$NC^\epsilon(m)$ is closed under taking intervals in $NC(m)$, i.e., if $\sigma, \pi \in NC^\epsilon(m)$ and $\tau \in NC(m)$ is such that $\sigma < \tau < \pi$, then $\tau \in NC^\epsilon(m)$.*

Proof. Let $\sigma, \pi \in NC^\epsilon(m)$, and $\tau \in NC(m)$ such that $\sigma < \tau < \pi$. From the inductive definition of $\widetilde{\tau}$, to show that $\tau \in NC^\epsilon(m)$ it suffices to consider $\pi = 1_m$. Now by the previous lemma, we have $\widetilde{\sigma} \vee \widetilde{1_m} \in NC_h^\epsilon(2m)$. By Lemma 3.5,

$$\overleftarrow{\widetilde{\sigma} \vee \widetilde{1_m}} = \widetilde{K(\sigma)} \vee \hat{0}_m = \widehat{K(\sigma)}.$$

Since $\sigma \leq \tau$, we have $\widehat{0_m} \leq \widehat{K(\tau)} \leq \widehat{K(\sigma)}$. Let $\delta = (\epsilon_2, \ldots, \epsilon_{2m}, \epsilon_1)$, and suppose that $\widehat{K(\tau)} \notin NC_h^\delta(2m)$. Let V be a block of $\widehat{K(\tau)}$, and note that V is of the form $(2i_1 - 1, 2i_1, \ldots, 2i_s - 1, 2i_s)$ for some $i_1 < \cdots < i_s$. Since $\hat{0}_m \in NC_h^\delta(2m)$, it follows that there is a $1 \leq l < s$ with $\delta_{2i_l} = \delta_{2i_{l+1}-1}$. Now since $\widehat{0_m} \leq \widehat{K(\tau)} \leq \widehat{K(\sigma)}$, it follows that the block W of $\widehat{K(\sigma)}$ which contains V must have an even number of elements between $2i_l$ and $2i_{l+1} - 1$. But then $\delta|_W$ cannot be alternating, which contradicts $\widehat{K(\sigma)} \in NC_h^\delta(2m)$.

So we have shown that $\widehat{K(\tau)} \in NC_h^\delta(2m)$, and since

$$\overrightarrow{\widehat{K(\tau)}} = \overrightarrow{\widetilde{K(\tau)} \vee \hat{0}_m} = \widetilde{\tau} \vee \widetilde{1}_m,$$

we have $\widetilde{\tau} \vee \widetilde{1_m} \in NC_h^\epsilon(2m)$. But then by the previous lemma, there is a $\gamma \in NC^\epsilon(m)$ with $\widetilde{\gamma} \vee \widetilde{1_m} = \widetilde{\tau} \vee \widetilde{1_m}$, and by Lemma 3.6 this implies $\tau = \gamma$ is in $NC^\epsilon(m)$ as claimed.

4. Integration on the Quantum Unitary Group

We begin by recalling the *Weingarten formula* from [2] for computing integrals with respect to the Haar state on $A_u(n)$.

Let $\epsilon_1, \ldots, \epsilon_{2m} \in \{1, *\}$ and define, for $n \in \mathbb{N}$, the *Gram matrix*

$$G_{\epsilon n}(\pi, \sigma) = n^{|\pi \vee \sigma|} \qquad (\pi, \sigma \in NC_2^{\epsilon}(2m)).$$

It is shown in [2] that $G_{\epsilon n}$ is invertible for $n \geq 2$, let $W_{\epsilon n}$ denote its inverse.

Theorem 4.1 [2]. The Haar state on $A_u(n)$ is given by

$$\psi_n(U_{i_1 j_1}^{\epsilon_1} \cdots U_{i_{2m} j_{2m}}^{\epsilon_{2m}}) = \sum_{\substack{\pi,\sigma \in NC_2^{\epsilon}(2m) \\ \pi \leq \ker \mathbf{i} \\ \sigma \leq \ker \mathbf{j}}} W_{\epsilon n}(\pi, \sigma),$$

$$\psi_n(U_{i_1 j_1}^{\epsilon_1} \cdots U_{i_{2m+1} j_{2m+1}}^{\epsilon_{2m+1}}) = 0,$$

for $1 \leq i_1, j_1, \ldots, i_{2m+1}, j_{2m+1} \leq n$ and $\epsilon_1, \ldots, \epsilon_{2m+1} \in \{1, *\}$.

Remark 4.2. Note that the Weingarten formula above is effective for computing integrals of products of the entries in U and its conjugate $\overline{U}$, the matrix with (i, j)-entry U_{ij}^*. We will also need to compute integrals of products of entries from U and its adjoint U^*, whose (i, j)-entry we denote $(U^*)_{ij}$ to distinguish from the conjugate $\overline{U}$. To do this we will use the following proposition, which allows us to reduce to the former case. Note that such a formula clearly fails for the classical unitary group. Indeed we have

$$\int_{U_n} (U^*)_{21}(U^*)_{43}U_{12}U_{34} = \int_{U_n} \overline{U}_{12}\overline{U}_{34}U_{12}U_{34} = \frac{1}{n^2 - 1}$$

while

$$\int_{U_n} \overline{U}_{21}\overline{U}_{34}U_{12}U_{43} = 0,$$

as can be seen by using the Weingarten formula from [10].

Proposition 4.3. *Let* $1 \leq i_1, i_2, \ldots, i_{4m} \leq n$ *and* $\epsilon_1, \ldots, \epsilon_{2m} \in \{1, *\}$. *Then*

$$\psi_n\left((U^{\epsilon_1})_{i_1 i_2}(U^{\epsilon_2})_{i_3 i_4} \cdots (U^{\epsilon_{2m}})_{i_{4m-1} i_{4m}}\right) = \psi_n\left(U_{i_1 i_2}^{\epsilon_1} U_{i_4 i_3}^{\epsilon_2} \cdots U_{i_{4m} i_{4m-1}}^{\epsilon_{2m}}\right).$$

Proof. We will use the fact from [1] that the joint $*$-distribution of $(U_{ij})_{1 \leq i,j \leq n}$ with respect to ψ_n is the same as that of $(zO_{ij})_{1 \leq i,j \leq n}$, where z and (O_{ij}) are random variables in a $*$-probability space (M, τ) such that:

(1) z is $*$-freely independent from $\{O_{ij} : 1 \leq i, j \leq n\}$.
(2) z has a Haar unitary distribution.
(3) (O_{ij}) are self-adjoint, and have the same joint distribution as the generators of the quantum orthogonal group $A_o(n)$.

The joint distribution of (O_{ij}) can also be computed via a Weingarten formula, see [2] for details. The only fact that we will use is that the joint distribution is invariant under transposition, i.e., the families $(O_{ij})_{1\leq i,j\leq n}$ and $(O_{ji})_{1\leq i,j\leq n}$ have the same joint distribution.

Now let $\epsilon_1, \ldots, \epsilon_{2m} \in \{1, *\}$. Let $A = \{j : \text{j is even and } \epsilon_j = *\} \cup \{j : \text{j is odd and } \epsilon_j = 1\}$, and $B = \{1, \ldots, 2m\}\backslash A$. Let $1 \leq i_1, j_1, \ldots, i_{2m}, j_{2m} \leq n$. For $1 \leq k \leq 2m$, define

$$i'_k = \begin{cases} i_k, & k \in A \\ j_k, & k \in B \end{cases}, \quad j'_k = \begin{cases} j_k, & k \in A \\ i_k, & k \in B \end{cases}.$$

We claim that

$$\psi_n\left(U_{i_1 j_1}^{\epsilon_1} \cdots U_{i_{2m} j_{2m}}^{\epsilon_{2m}}\right) = \psi_n\left(U_{i'_1 j'_1}^{\epsilon_1} \cdots U_{i'_{2m} j'_{2m}}^{\epsilon_{2m}}\right),$$

from which the formula in the statement follows immediately.

As discussed above, we have

$$\psi_n\left(U_{i_1 j_1}^{\epsilon_1} \cdots U_{i_{2m} j_{2m}}^{\epsilon_{2m}}\right) = \tau\left((zO_{i_1 j_1})^{\epsilon_1} \cdots (zO_{i_{2m} j_{2m}})^{\epsilon_{2m}}\right).$$

Note that the expression $(zO_{i_1 j_1})^{\epsilon_1} \cdots (zO_{i_{2m} j_{2m}})^{\epsilon_{2m}}$ can be written as a product of terms of the form $zO_{i_k j_k}$ or $O_{i_k j_k} z^*$, depending if ϵ_k is 1 or $*$. After rewriting the expression in this form, let C be the subset of $\{1, \ldots, 4m\}$ consisting of those indices corresponding to z or z^*, and let D be its complement. Explicitly, if $\epsilon_k = 1$ then $2k-1$ is in C and $2k$ is in D, and if $\epsilon_k = *$ then $2k$ is in C and $2k-1$ is in D. Given partitions $\alpha, \beta \in NC(2m)$, let $\Theta(\alpha, \beta) \in P(4m)$ be given by partitioning C according to α and D according to β. By freeness, we have

$$\begin{aligned} &\tau\left((zO_{i_1 j_1})^{\epsilon_1} \cdots (zO_{i_{2m} j_{2m}})^{\epsilon_{2m}}\right) \\ &\quad = \sum_{\substack{\alpha,\beta \in NC(2m) \\ \Theta(\alpha,\beta) \in NC(4m)}} \kappa_\alpha[z^{\epsilon_1}, \ldots, z^{\epsilon_{2m}}]\kappa_\beta[O_{i_1 j_1}, \ldots, O_{i_{2m} j_{2m}}]. \end{aligned}$$

Now since Haar unitaries are R-diagonal, we have $\kappa_\alpha[z^{\epsilon_1}, \ldots, z^{\epsilon_{2m}}] = 0$ unless each block of α contains an even number of elements. So assume that α has this property, we claim that if β is such that $\Theta(\alpha, \beta)$ is noncrossing, then β does not join any element of A with an element of B. Indeed, suppose that β joins $k_1 < k_2$ and that one of k_1, k_2 is in A and the other is in B. If k_1, k_2 have the same parity, then it follows that one of $\epsilon_{k_1}, \epsilon_{k_2}$ is a 1 while the other is a $*$. Suppose that $\epsilon_{k_1} = 1, \epsilon_{k_2} = *$; the other case is similar. Then we have $2k_1$ connected to $2k_2 - 1$ in $\Theta(\alpha, \beta)$. Since $\Theta(\alpha, \beta)$ is noncrossing, α cannot join any element of $\{k_1 + 1, \ldots, k_2 - 1\}$ to an element outside of this set. But since this set contains an odd number of elements, we obtain a contradiction to the choice of α.

If k_1, k_2 have different parity, then it follows that $\epsilon_{k_1} = \epsilon_{k_2}$. Suppose that $\epsilon_{k_1} = \epsilon_{k_2} = 1$; the other case is similar. Then $2k_1$ is connected to $2k_2$ in $\Theta(\alpha, \beta)$. It follows that α cannot connect any element of $\{k_1 + 1, \ldots, k_2\}$ to an element outside of this set, and again this set has an odd number of elements which contradicts the choice of α.

So the only nonzero terms appearing in the expression above come from $\beta \in NC(2m)$ which split into noncrossing partitions π of A and σ of B. In this case, if $A = (a_1 < \cdots < a_s)$ and $B = (b_1 < \cdots < b_r)$, we have

$$\begin{aligned}\kappa_\beta[O_{i_1j_1},\ldots,O_{i_{2m}j_{2m}}] &= \kappa_\pi[O_{i_{a_1}j_{a_1}},\ldots,O_{i_{a_s}j_{a_s}}]\kappa_\sigma[O_{i_{b_1}j_{b_1}},\ldots,O_{i_{b_r}j_{b_r}}]\\ &= \kappa_\pi[O_{i_{a_1}j_{a_1}},\ldots,O_{i_{a_s}j_{a_s}}]\kappa_\sigma[O_{j_{b_1}i_{b_1}},\ldots,O_{j_{b_r}i_{b_r}}]\\ &= \kappa_\beta[O_{i'_1j'_1},\ldots,O_{i'_{2m}j'_{2m}}],\end{aligned}$$

where we have used the invariance of the distribution of (O_{ij}) under transposition.

Putting this all together, we have

$$\begin{aligned}\psi_n\left(U_{i_1j_1}^{\epsilon_1}\cdots U_{i_{2m}j_{2m}}^{\epsilon_{2m}}\right) &= \tau\left((zO_{i_1j_1})^{\epsilon_1}\cdots(zO_{i_{2m}j_{2m}})^{\epsilon_{2m}}\right)\\ &= \sum_{\substack{\alpha,\beta\in NC(2m)\\ \Theta(\alpha,\beta)\in NC(4m)}}\kappa_\alpha[z^{\epsilon_1},\ldots,z^{\epsilon_{2m}}]\kappa_\beta[O_{i_1j_1},\ldots,O_{i_{2m}j_{2m}}]\\ &= \sum_{\substack{\alpha,\beta\in NC(2m)\\ \Theta(\alpha,\beta)\in NC(4m)}}\kappa_\alpha[z^{\epsilon_1},\ldots,z^{\epsilon_{2m}}]\kappa_\beta[O_{i'_1j'_1},\ldots,O_{i'_{2m}j'_{2m}}]\\ &= \tau\left((zO_{i'_1j'_1})^{\epsilon_1}\cdots(zO_{i'_{2m}j'_{2m}})^{\epsilon_{2m}}\right)\\ &= \psi_n\left(U_{i'_1j'_1}^{\epsilon_1}\cdots U_{i'_{2m}j'_{2m}}^{\epsilon_{2m}}\right)\end{aligned}$$

as desired.

We can now extend this result to the free product $A_u(n)^{*\infty}$.

Corollary 4.4. *Let* $l_1,\ldots,l_{2m}\in\mathbb{N}$, $\epsilon_1,\ldots,\epsilon_{2m}\in\{1,*\}$ *and* $1\le i_1,j_1,\ldots,i_{2m},j_{2m}\le n$. *In* $A_u(n)^{*\infty}$, *we have*

$$\begin{aligned}&\psi_n^{*\infty}\left((U(l_1)^{\epsilon_1})_{i_1i_2}(U(l_2)^{\epsilon_2})_{i_3i_4}\cdots(U(l_{2m})^{\epsilon_{2m}})_{i_{4m-1}i_{4m}}\right)\\ &\quad=\psi_n^{*\infty}\left(U(l_1)_{i_1i_2}^{\epsilon_1}U(l_2)_{i_4i_3}^{\epsilon_2}\cdots U(l_{2m})_{i_{4m}i_{4m-1}}^{\epsilon_{2m}}\right).\end{aligned}$$

Proof. First we claim that in $A_u(n)$, we have

$$\begin{aligned}&\kappa^{(2m)}[(U^{\epsilon_1})_{i_1i_2},(U^{\epsilon_2})_{i_3i_4},\ldots,(U^{\epsilon_{2m}})_{i_{4m-1}i_{4m}}]\\ &\quad=\kappa^{(2m)}[U_{i_1i_2}^{\epsilon_1},U_{i_4i_3}^{\epsilon_2},\ldots,U_{i_{4m}i_{4m-1}}^{\epsilon_{2m}}].\end{aligned}$$

(Note that any cumulant of odd length is zero by Theorem 4.1).

Indeed, we have

$$\begin{aligned}&\kappa^{(2m)}[(U^{\epsilon_1})_{i_1i_2},(U^{\epsilon_2})_{i_3i_4},\ldots,(U^{\epsilon_{2m}})_{i_{4m-1}i_{4m}}]\\ &\quad=\sum_{\sigma\in NC(2m)}\mu_{2m}(\sigma,1_{2m})\prod_{V\in\sigma}\psi_n(V)[(U^{\epsilon_1})_{i_1i_2},(U^{\epsilon_2})_{i_3i_4},\ldots,(U^{\epsilon_{2m}})_{i_{4m-1}i_{4m}}].\end{aligned}$$

Now it is clear from Theorem 4.1 that

$$\psi_n(V)[(U^{\epsilon_1})_{i_1i_2},(U^{\epsilon_2})_{i_3i_4},\ldots,(U^{\epsilon_{2m}})_{i_{4m-1}i_{4m}}]=0$$

unless V has an even number of elements. So the nonzero terms in the expression above come from those $\sigma\in NC(2m)$ for which every block has an even number of elements.

For such a σ, the noncrossing condition implies that each block $V = (l_1 < \cdots < l_s)$ must be alternating in parity. By Proposition 4.3 we have

$$\begin{aligned}
&\psi_n(V)[(U^{\epsilon_1})_{i_1 i_2}, (U^{\epsilon_2})_{i_3 i_4}, \ldots, (U^{\epsilon_{2m}})_{i_{4m-1} i_{4m}}] \\
&= \psi_n\left((U^{\epsilon_{l_1}})_{i_{2l_1-1} i_{2l_1}} (U^{\epsilon_{l_2}})_{i_{2l_2-1} i_{2l_2}} \cdots (U^{\epsilon_{l_s}})_{i_{2l_s-1} i_{2l_s}}\right) \\
&= \psi_n\left(U^{\epsilon_{l_1}}_{i_{2l_1-1} i_{2l_1}} U^{\epsilon_{l_2}}_{i_{2l_2} i_{2l_2-1}} \cdots U^{\epsilon_{l_s}}_{i_{2l_s} i_{2l_s-1}}\right) \\
&= \psi_n\left(U^{\epsilon_{l_1}}_{i_{2l_1} i_{2l_1-1}} U^{\epsilon_{l_2}}_{i_{2l_2-1} i_{2l_2}} \cdots U^{\epsilon_{l_s}}_{i_{2l_s-1} i_{2l_s}}\right),
\end{aligned}$$

where the last equation follows from the invariance of the joint $*$-distribution of (U_{ij}) under transposition. It follows that

$$\begin{aligned}
&\kappa^{(2m)}[(U^{\epsilon_1})_{i_1 i_2}, (U^{\epsilon_2})_{i_3 i_4}, \ldots, (U^{\epsilon_{2m}})_{i_{4m-1} i_{4m}}] \\
&= \sum_{\sigma \in NC(2m)} \mu_{2m}(\sigma, 1_{2m}) \prod_{V \in \sigma} \psi_n(V)[(U^{\epsilon_1})_{i_1 i_2}, (U^{\epsilon_2})_{i_3 i_4}, \ldots, (U^{\epsilon_{2m}})_{i_{4m-1} i_{4m}}] \\
&= \sum_{\sigma \in NC(2m)} \mu_{2m}(\sigma, 1_{2m}) \prod_{V \in \sigma} \psi_n(V)[U^{\epsilon_1}_{i_1 i_2}, U^{\epsilon_2}_{i_4 i_3}, \ldots, U^{\epsilon_{2m}}_{i_{4m} i_{4m-1}}] \\
&= \kappa^{(2m)}[U^{\epsilon_1}_{i_1 i_2}, U^{\epsilon_2}_{i_4 i_3}, \ldots, U^{\epsilon_{2m}}_{i_{4m} i_{4m-1}}]
\end{aligned}$$

as claimed.

Now by free independence, in $A_u(n)^{*\infty}$ we have

$$\begin{aligned}
&\psi_n^{*\infty}\left((U(l_1)^{\epsilon_1})_{i_1 i_2} (U(l_2)^{\epsilon_2})_{i_3 i_4} \cdots (U(l_{2m})^{\epsilon_{2m}})_{i_{4m-1} i_{4m}}\right) \\
&= \sum_{\substack{\sigma \in NC(2m) \\ \sigma \leq \ker \mathbf{l}}} \prod_{V \in \sigma} \kappa(V)[(U(l_1)^{\epsilon_1})_{i_1 i_2}, (U(l_2)^{\epsilon_2})_{i_3 i_4}, \ldots, (U(l_{2m})^{\epsilon_{2m}})_{i_{4m-1} i_{4m}}].
\end{aligned}$$

Since $\kappa(V)$ is zero unless V has an even number of elements, the only terms which contribute to the sum above come again from $\sigma \in NC(2m)$ for which each block has an even number of elements. From the previous claim, we have

$$\begin{aligned}
&\kappa(V)[(U(l_1)^{\epsilon_1})_{i_1 i_2}, (U(l_2)^{\epsilon_2})_{i_3 i_4}, \ldots, (U(l_{2m})^{\epsilon_{2m}})_{i_{4m-1} i_{4m}}] \\
&= \kappa(V)[U(l_1)^{\epsilon_1}_{i_1 i_2}, U(l_2)^{\epsilon_2}_{i_4 i_3}, \ldots, U(l_{2m})^{\epsilon_{2m}}_{i_{4m} i_{4m-1}}]
\end{aligned}$$

for each block $V \in \sigma$, and the result follows immediately.

We will now give an estimate on the asymptotic behavior of the entries of $W_{\epsilon n}$ as $n \to \infty$. This improves the estimate given in [2]. Note that by taking $\epsilon = 1 * \cdots 1*$, this estimate also applies to the quantum orthogonal group, see [2].

Theorem 4.5. Let $\epsilon_1, \ldots, \epsilon_{2m} \in \{1, *\}$. Let $\pi, \sigma \in NC^{\epsilon}(m)$. Then

$$W_{\epsilon n}(\widetilde{\pi}, \widetilde{\sigma}) = O(n^{2|\pi \vee \sigma| - |\pi| - |\sigma| - m}).$$

Moreover,

$$n^{m + |\sigma| - |\pi|} W_{\epsilon n}(\widetilde{\pi}, \widetilde{\sigma}) = \mu_m(\sigma, \pi) + O(n^{-2}),$$

where μ_m is the Möbius function on $NC(m)$, and we use the convention that $\mu_m(\sigma, \pi) = 0$ if $\sigma \not\leq \pi$.

Proof. We use a standard method from [10,11], further developed in [2,3,12,4].
First observe that

$$G_{\epsilon n} = \Theta_{\epsilon n}^{1/2}(1 + B_{\epsilon n})\Theta_{\epsilon n}^{1/2},$$

where

$$\Theta_{\epsilon n}(\pi,\sigma) = \begin{cases} n^m, & \pi = \sigma \\ 0, & \pi \neq \sigma \end{cases},$$

$$B_{\epsilon n}(\pi,\sigma) = \begin{cases} 0, & \pi = \sigma \\ n^{|\pi \vee \sigma| - m}, & \pi \neq \sigma \end{cases}.$$

Note that the entries of $B_{\epsilon n}$ are $O(n^{-1})$, in particular for n large we have the geometric series expansion

$$(1 + B_{\epsilon n})^{-1} = 1 - B_{\epsilon n} + \sum_{l \geq 1} (-1)^{l+1} B_{\epsilon n}^{l+1}.$$

Hence

$$W_{\epsilon n}(\widetilde{\pi},\widetilde{\sigma}) = \sum_{l \geq 1} (-1)^{(l+1)} (\Theta_{\epsilon n}^{-1/2} B_{\epsilon n}^{l+1} \Theta_{\epsilon n}^{-1/2})(\widetilde{\pi},\widetilde{\sigma}) + \begin{cases} n^{-m} & \pi = \sigma, \\ -n^{|\widetilde{\pi} \vee \widetilde{\sigma}| - 2m} & \pi \neq \sigma. \end{cases}$$

Now for $l \geq 1$, we have

$$(\Theta_{\epsilon n}^{-1/2} B_{\epsilon n}^{l+1} \Theta_{\epsilon n}^{-1/2})(\widetilde{\pi},\widetilde{\sigma}) = \sum_{\substack{\nu_1,\ldots,\nu_l \in NC^{\epsilon}(m) \\ \pi \neq \nu_1 \neq \cdots \neq \nu_l \neq \sigma}} n^{|\widetilde{\pi} \vee \widetilde{\nu}_1| + |\widetilde{\nu}_1 \vee \widetilde{\nu}_2| + \cdots + |\widetilde{\nu}_l \vee \widetilde{\sigma}| - (l+2)m}.$$

Now we claim that

$$\begin{aligned} |\widetilde{\pi} \vee \widetilde{\nu}_1| + \cdots + |\widetilde{\nu}_l \vee \widetilde{\sigma}| &\leq |\widetilde{\pi} \vee \widetilde{\sigma}| + |\widetilde{\nu}_1| + \cdots + |\widetilde{\nu}_l| \\ &\leq |\widetilde{\pi} \vee \widetilde{\sigma}| + l \cdot m, \end{aligned}$$

from which the first equation follows from the above equation and Theorem 3.7.
Indeed, the case $l = 1$ follows from the semi-modular condition:

$$\begin{aligned} |\widetilde{\pi} \vee \widetilde{\nu}_1| + |\widetilde{\nu}_1 \vee \widetilde{\sigma}| &\leq |(\widetilde{\pi} \vee \widetilde{\nu}_1) \vee (\widetilde{\nu}_1 \vee \widetilde{\sigma})| + |(\widetilde{\pi} \vee \widetilde{\nu}_1) \wedge (\widetilde{\nu}_1 \vee \widetilde{\sigma})| \\ &\leq |\widetilde{\pi} \vee \widetilde{\sigma}| + |\widetilde{\nu}_1| \\ &= |\widetilde{\pi} \vee \widetilde{\sigma}| + m. \end{aligned}$$

The general case follows easily from induction on l.
For the second part, apply Theorem 3.7 to find that

$$\begin{aligned} |\widetilde{\pi} \vee \widetilde{\nu}_1| + \cdots + |\widetilde{\nu}_l \vee \widetilde{\sigma}| &= 2(|\nu_1 \vee \nu_2| + \cdots + |\nu_l \vee \sigma| - |\nu_1| - \cdots - |\nu_l|) \\ &\quad +2|\pi \vee \nu_1| - |\pi| - |\sigma| + (l+1)m \\ &\leq |\pi| - |\sigma| + (l+1)m, \end{aligned}$$

where equality holds if $\sigma < \nu_l < \cdots < \nu_1 < \pi$ and otherwise the difference is at least 2. It then follows from the equation above that, up to $O(n^{-2})$, $n^{m+|\sigma|-|\pi|}W_{\epsilon n}(\widetilde{\pi}, \widetilde{\sigma})$ is equal to 0 if $\sigma \not\leq \pi$, 1 if $\sigma = \pi$ and otherwise is given by

$$-1 + \sum_{l=1}^{\infty}(-1)^{l+1}|\{(\nu_1, \dots, \nu_l) \in (NC^{\epsilon}(m))^l : \sigma < \nu_l < \cdots < \nu_1 < \pi\}|.$$

Since $NC^{\epsilon}(m)$ is closed under taking intervals in $NC(m)$, this is equal to $\mu_m(\sigma, \pi)$.

As a corollary, we can give an estimate on the free cumulants of the generators U_{ij} of $A_u(n)$. (Note that the cumulants of odd length are all zero since the generators have an even joint distribution).

Corollary 4.6. *Let $\epsilon_1, \dots, \epsilon_{2m} \in \{1, *\}$ and $i_1, j_1, \dots, i_{2m}, j_{2m} \in \mathbb{N}$. For $\omega \in NC(2m)$, we have for the moment functions*

$$\psi_n^{(\omega)}[U_{i_1 j_1}^{\epsilon_1}, \dots, U_{i_{2m} j_{2m}}^{\epsilon_{2m}}] = \sum_{\substack{\sigma,\pi \in NC^{\epsilon}(m) \\ \widetilde{\pi} \leq \ker \mathbf{i} \wedge \omega \\ \widetilde{\sigma} \leq \ker \mathbf{j} \wedge \omega}} n^{|\pi|-|\sigma|-m}(\mu_m(\sigma, \pi) + O(n^{-2})),$$

and for the cumulant functions

$$\kappa^{(\omega)}[U_{i_1 j_1}^{\epsilon_1}, \dots, U_{i_{2m} j_{2m}}^{\epsilon_{2m}}] = \sum_{\substack{\pi,\sigma \in NC^{\epsilon}(m) \\ \widetilde{\pi} \leq \ker \mathbf{i} \\ \widetilde{\sigma} \leq \ker \mathbf{j} \\ \widetilde{\pi} \vee_{nc} \widetilde{\sigma} = \omega}} n^{|\pi|-|\sigma|-m}(\mu_m(\pi, \sigma) + O(n^{-2})).$$

Proof. First note that $\psi_n^{(\omega)}[U_{i_1 j_1}^{\epsilon_1}, \dots, U_{i_{2m} j_{2m}}^{\epsilon_{2m}}] = 0$ unless $\omega \in NC_h(2m)$, i.e., unless each block of ω has an even number of elements. So suppose this is the case, then by Lemma 3.9 we have $\omega = \widetilde{\alpha} \vee \widetilde{\beta}$ for some $\alpha, \beta \in NC(m)$ with $\alpha \leq \beta$. By the Weingarten formula, we have

$$\psi_n^{(\omega)}[U_{i_1 j_1}^{\epsilon_1}, \dots, U_{i_{2m} j_{2m}}^{\epsilon_{2m}}] = \sum_{\substack{\pi,\sigma \in NC^{\epsilon}(m) \\ \widetilde{\pi} \leq \ker \mathbf{i} \wedge \omega \\ \widetilde{\sigma} \leq \ker \mathbf{j} \wedge \omega}} \prod_{V \in \omega} W_{\epsilon|_V n}(\widetilde{\pi}|_V, \widetilde{\sigma}|_V).$$

Let $V = (l_1 < \cdots < l_s)$ be a block of ω. In order to apply Theorem 4.5 we have to write $\widetilde{\pi}|_V$ and $\widetilde{\sigma}|_V$ as $\widetilde{\pi_V}$ and $\widetilde{\sigma_V}$, respectively, for some $\pi_V, \sigma_V \in NC(|V|/2)$. Since $\mu_{|V|/2}(\sigma_V, \pi_V) = \mu_{|V|}(\widehat{\sigma_V}, \widehat{\pi_V})$, it suffices to recover the doubled versions $\widehat{\sigma_V}, \widehat{\pi_V}$ from $\widetilde{\pi}|_V$ and $\widetilde{\sigma}|_V$. But this can be achieved as follows:

$$\widehat{\pi_V} = \widetilde{\pi_V} \vee \hat{0}_{|V|/2} = \widetilde{\pi}|_V \vee \{(l_1, l_2), \dots, (l_{s-1}, l_s)\}.$$

So it remains to write $\{(l_1, l_2), \dots, (l_{s-1}, l_s)\}$ intrinsically in terms of ω.

Recall from Lemma 3.6 that we have $K(\omega) = \alpha \wr K(\beta)$. It follows that for $1 \leq r \leq s$ such that l_r is odd, α has a block whose least element is $\frac{l_r+1}{2}$ and greatest element is $\frac{l_{r+1}}{2}$. Therefore l_r is joined to l_{r+1} in $\widetilde{\alpha}$. So if l_1 is odd, then $\widetilde{\alpha}|_V$ is equal to $\{(l_1, l_2), (l_3, l_4), \dots, (l_{s-1}, l_s)\}$. In this case, from Theorem 4.5 we have

$$\begin{aligned} &W_{\epsilon|_V n}(\widetilde{\pi}|_V, \widetilde{\sigma}|_V) \\ &\quad = n^{|\widetilde{\pi}|_V \vee \widetilde{\alpha}|_V| - |\widetilde{\sigma}|_V \vee \widetilde{\alpha}|_V| - |V|/2}(\mu_{|V|}(\widetilde{\sigma}|_V \vee \widetilde{\alpha}|_V, \widetilde{\pi}|_V \vee \widetilde{\alpha}|_V) + O(n^{-2})). \end{aligned}$$

On the other hand, if l_1 is even then $\widetilde{\alpha}|_V = \{(l_1, l_s), (l_2, l_3), \ldots, (l_{s-2}, l_{s-1})\}$. In this case we have

$$\begin{aligned} &W_{\epsilon|_V n}(\widetilde{\pi}|_V, \widetilde{\sigma}|_V) \\ &= n^{\left|\widetilde{\sigma}|_V \vee \overrightarrow{\widetilde{\alpha}|_V}\right| - \left|\widetilde{\pi}|_V \vee \overrightarrow{\widetilde{\alpha}|_V}\right| - |V|/2} (\mu_{|V|}(\widetilde{\pi}|_V \vee \overrightarrow{\widetilde{\alpha}|_V}, \widetilde{\sigma}|_V \vee \overrightarrow{\widetilde{\alpha}|_V}) + O(n^{-2})) \\ &= n^{\left|\overleftarrow{\widetilde{\sigma}|_V} \vee \widetilde{\alpha}|_V\right| - \left|\overleftarrow{\widetilde{\pi}|_V} \vee \widetilde{\alpha}|_V\right| - |V|/2} (\mu_{|V|}(\overleftarrow{\widetilde{\pi}|_V} \vee \widetilde{\alpha}|_V, \overleftarrow{\widetilde{\sigma}|_V} \vee \widetilde{\alpha}|_V) + O(n^{-2})), \end{aligned}$$

where here the arrows act on the legs of V. Since this corresponds, by Lemma 3.5, to the Kreweras complement on $NC_{|V|/2}$, we have

$$\left|\overleftarrow{\widetilde{\sigma}|_V} \vee \widetilde{\alpha}|_V\right| = |V|/2 + 1 - |\widetilde{\sigma}|_V \vee \widetilde{\alpha}|_V|$$

and

$$\mu_{|V|}(\overleftarrow{\widetilde{\pi}|_V} \vee \widetilde{\alpha}|_V, \overleftarrow{\widetilde{\sigma}|_V} \vee \widetilde{\alpha}|_V) = \mu_{|V|}(\widetilde{\sigma}|_V \vee \widetilde{\alpha}|_V, \widetilde{\pi}|_V \vee \widetilde{\alpha}|_V).$$

So it follows that, as in previous case, we have

$$\begin{aligned} &W_{\epsilon|_V n}(\widetilde{\pi}|_V, \widetilde{\sigma}|_V) \\ &= n^{|\widetilde{\pi}|_V \vee \widetilde{\alpha}|_V| - |\widetilde{\sigma}|_V \vee \widetilde{\alpha}|_V| - |V|/2} (\mu_{|V|}(\widetilde{\sigma}|_V \vee \widetilde{\alpha}|_V, \widetilde{\pi}|_V \vee \widetilde{\alpha}|_V) + O(n^{-2})). \end{aligned}$$

Therefore,

$$\begin{aligned} &\psi_n^{(\omega)}[U_{i_1 j_1}^{\epsilon_1}, \ldots, U_{i_{2m} j_{2m}}^{\epsilon_{2m}}] \\ &= \sum_{\substack{\sigma,\pi \in NC^\epsilon(m) \\ \widetilde{\pi} \le \ker \mathbf{i} \wedge \omega \\ \widetilde{\sigma} \le \ker \mathbf{j} \wedge \omega}} \prod_{V \in \omega} n^{|\widetilde{\pi}|_V \vee \widetilde{\alpha}|_V| - |\widetilde{\sigma}|_V \vee \widetilde{\alpha}|_V| - |V|/2} \\ &\quad \cdot (\mu_{|V|}(\widetilde{\sigma}|_V \vee \widetilde{\alpha}|_V, \widetilde{\pi}|_V \vee \widetilde{\alpha}|_V) + O(n^{-2})) \\ &= \sum_{\substack{\sigma,\pi \in NC^\epsilon(m) \\ \widetilde{\pi} \le \ker \mathbf{i} \wedge \omega \\ \widetilde{\sigma} \le \ker \mathbf{j} \wedge \omega}} n^{|\widetilde{\pi} \vee \widetilde{\alpha}| - |\widetilde{\sigma} \vee \widetilde{\alpha}| - m} (\mu_{2m}(\widetilde{\sigma} \vee \widetilde{\alpha}, \widetilde{\pi} \vee \widetilde{\alpha}) + O(n^{-2})), \end{aligned}$$

where we have used the multiplicativity of the Möbius function on $NC(2m)$.

Now since $\widetilde{\sigma} = \widetilde{\sigma} \vee \widetilde{\sigma} \le \widetilde{\alpha} \vee \widetilde{\beta}$, taking the Kreweras complement and applying Lemma 3.6 gives $\alpha \wr K(\beta) \le \sigma \wr K(\sigma)$. So we have $\alpha \le \sigma \le \beta$. By Theorem 3.7, we then have $|\widetilde{\sigma} \vee \widetilde{\alpha}| = |\sigma| + m - |\alpha|$. Also, we have

$$\begin{aligned} \mu_{2m}(\widetilde{\sigma} \vee \widetilde{\alpha}, \widetilde{\pi} \vee \widetilde{\alpha}) &= \mu_{2m}(K(\widetilde{\pi} \vee \widetilde{\alpha}), K(\widetilde{\sigma} \vee \widetilde{\alpha})) \\ &= \mu_{2m}(\alpha \wr K(\pi), \alpha \wr K(\sigma)) \\ &= \mu_m(K(\pi), K(\sigma)) \\ &= \mu_m(\sigma, \pi). \end{aligned}$$

Plugging this into the equation above, we have

$$\psi_n^{(\omega)}[U_{i_1 j_1}^{\epsilon_1}, \ldots, U_{i_{2m} j_{2m}}^{\epsilon_{2m}}] = \sum_{\substack{\sigma,\pi \in NC^\epsilon(m) \\ \widetilde{\pi} \le \ker \mathbf{i} \wedge \omega \\ \widetilde{\sigma} \le \ker \mathbf{j} \wedge \omega}} n^{|\pi| - |\sigma| - m} (\mu_m(\sigma, \pi) + O(n^{-2})).$$

For the cumulant function this gives

$$\begin{aligned}
&\kappa^{(\tau)}[U_{i_1 j_1}^{\epsilon_1}, \ldots, U_{i_{2m} j_{2m}}^{\epsilon_{2m}}] \\
&= \sum_{\substack{\omega \in NC(2m)\\ \omega \le \tau}} \mu_{2m}(\omega, \tau) \psi_n^{(\omega)}[U_{i_1 j_1}^{\epsilon_1}, \ldots, U_{i_{2m} j_{2m}}^{\epsilon_{2m}}] \\
&= \sum_{\substack{\omega \in NC(2m)\\ \omega \le \tau}} \mu_{2m}(\omega, \tau) \sum_{\substack{\sigma, \pi \in NC^\epsilon(m)\\ \widetilde{\pi} \le \ker \mathbf{i} \wedge \omega \\ \widetilde{\sigma} \le \ker \mathbf{j} \wedge \omega}} n^{|\pi| - |\sigma| - m} (\mu_m(\sigma, \pi) + O(n^{-2})) \\
&= \sum_{\substack{\sigma, \pi \in NC^\epsilon(m)\\ \widetilde{\pi} \le \ker \mathbf{i} \\ \widetilde{\sigma} \le \ker \mathbf{j}}} n^{|\pi| - |\sigma| - m} (\mu_m(\sigma, \pi) + O(n^{-2})) \sum_{\substack{\omega \in NC(2m)\\ \widetilde{\pi} \vee_{nc} \widetilde{\sigma} \le \omega \le \tau}} \mu_{2m}(\omega, \tau).
\end{aligned}$$

Since

$$\sum_{\substack{\omega \in NC(2m)\\ \widetilde{\pi} \vee_{nc} \widetilde{\sigma} \le \omega \le \tau}} \mu_{2m}(\omega, \tau) = \begin{cases} 1, & \widetilde{\pi} \vee_{nc} \widetilde{\sigma} = \tau \\ 0, & \text{otherwise} \end{cases},$$

the result follows.

As a corollary, we can give an estimate on the Haar state on the free product $A_u(n)^{*\infty}$.

Corollary 4.7. *Let $l_1, \ldots, l_{2m} \in \mathbb{N}$, $\epsilon_1, \ldots, \epsilon_{2m} \in \{1, *\}$ and $i_1, j_1, \ldots, i_{2m}, j_{2m} \in \mathbb{N}$. In $A_u(n)^{*\infty}$, we have*

$$\psi_n^{*\infty}\left(U(l_1)_{i_1 j_1}^{\epsilon_1} \cdots U(l_{2m})_{i_{2m} j_{2m}}^{\epsilon_{2m}}\right) = \sum_{\substack{\pi, \sigma \in NC^\epsilon(m)\\ \widetilde{\pi} \le \ker \mathbf{i} \wedge \ker \mathbf{l} \\ \widetilde{\sigma} \le \ker \mathbf{j} \wedge \ker \mathbf{l}}} n^{|\pi| - |\sigma| - m} (\mu_m(\sigma, \pi) + O(n^{-2})).$$

Proof. Since the families $(\{U(l)_{ij}\})_{l \in \mathbb{N}}$ are freely independent, we have by the vanishing of mixed cumulants

$$\psi_n^{*\infty}\left(U(l_1)_{i_1 j_1}^{\epsilon_1} \cdots U(l_{2m})_{i_{2m} j_{2m}}^{\epsilon_{2m}}\right) = \sum_{\substack{\tau \in NC(2m)\\ \tau \le \ker \mathbf{l}}} \kappa^{(\tau)}[U(l_1)_{i_1 j_1}^{\epsilon_1}, \ldots, U(l_{2m})_{i_{2m} j_{2m}}^{\epsilon_{2m}}].$$

Since the families $(\{U(l)_{ij}\})_{l \in \mathbb{N}}$ are identically distributed, we have

$$\kappa^{(\tau)}[U(l_1)_{i_1 j_1}^{\epsilon_1}, \ldots, U(l_{2m})_{i_{2m} j_{2m}}^{\epsilon_{2m}}] = \kappa^{(\tau)}[U(1)_{i_1 j_1}^{\epsilon_1}, \ldots, U(1)_{i_{2m} j_{2m}}^{\epsilon_{2m}}]$$

for any $\tau \in NC(2m)$ such that $\tau \le \ker \mathbf{l}$. Applying the previous corollary, we have

$$\begin{aligned}
&\psi_n^{*\infty}\left(U(l_1)_{i_1 j_1}^{\epsilon_1} \cdots U(l_{2m})_{i_{2m} j_{2m}}^{\epsilon_{2m}}\right) \\
&= \sum_{\substack{\tau \in NC(2m)\\ \tau \le \ker \mathbf{l}}} \sum_{\substack{\pi, \sigma \in NC^\epsilon(m)\\ \widetilde{\pi} \le \ker \mathbf{i} \\ \widetilde{\sigma} \le \ker \mathbf{j} \\ \widetilde{\pi} \vee_{nc} \widetilde{\sigma} = \tau}} n^{|\pi| - |\sigma| - m} (\mu_m(\sigma, \pi) + O(n^{-2})) \\
&= \sum_{\substack{\pi, \sigma \in NC^\epsilon(m)\\ \widetilde{\pi} \le \ker \mathbf{i} \wedge \ker \mathbf{l} \\ \widetilde{\sigma} \le \ker \mathbf{j} \wedge \ker \mathbf{l}}} n^{|\pi| - |\sigma| - m} (\mu_m(\sigma, \pi) + O(n^{-2})).
\end{aligned}$$

5. Asymptotic Freeness Results

Throughout the first part of this section, the framework will be as follows: $\mathcal{B}$ will be a fixed unital C*-algebra, and $(D_N(i))_{i\in I}$ will be a family of matrices in $M_N(\mathcal{B})$ for $N \in \mathbb{N}$, which is a $\mathcal{B}$-valued probability space with conditional expectation $E_N = \mathrm{tr}_N \otimes \mathrm{id}_{\mathcal{B}}$. Consider the free product $A_u(N)^{*\infty}$, generated by the entries in the matrices $(U_N(l))_{l\in\mathbb{N}} \in M_N(A_u(N)^{*\infty})$. By a family of freely independent Haar quantum unitary random matrices, independent from $\mathcal{B}$, we will mean the family $(U_N(l) \otimes 1_{\mathcal{B}})_{l\in\mathbb{N}}$ in $M_N(A_u(N)^{*\infty} \otimes \mathcal{B}) = M_N(\mathbb{C}) \otimes A_u(N)^{*\infty} \otimes \mathcal{B}$, which we will still denote by $(U_N(l))_{l\in\mathbb{N}}$. We also identify $D_N(i) = D_N(i) \otimes 1_{A_u(N)^{*\infty}}$ for $i \in I$. We will consider the $\mathcal{B}$-valued joint distribution of the family of sets $(\{U_N(1), U_N(1)^*\}, \{U_N(2), U_N(2)^*\}, \ldots, \{D_N(i)|i \in I\})$ with respect to the conditional expectation

$$\psi_N^{*\infty} \otimes E_N = \mathrm{tr}_N \otimes \psi_N^{*\infty} \otimes \mathrm{id}_{\mathcal{B}}.$$

We can now state our main result.

Theorem 5.1. Let $\mathcal{B}$ be a unital C*-algebra, and let $(D_N(i))_{i\in I}$ be a family of matrices in $M_N(\mathcal{B})$ for $N \in \mathbb{N}$. Suppose that there is a finite constant C such that $\|D_N(i)\| \leq C$ for all $i \in I$ and $N \in \mathbb{N}$. Let $(U_N(l))_{l\in\mathbb{N}}$ be a family of freely independent $N \times N$ Haar quantum unitary random matrices, independent from $\mathcal{B}$. Let $(u(l), u(l)^*)_{l\in\mathbb{N}}$ and $(d_N(i))_{i\in I, N\in\mathbb{N}}$ be random variables in a $\mathcal{B}$-valued probability space $(\mathcal{A}, E : \mathcal{A} \to \mathcal{B})$ such that

(1) $(u(l), u(l)^*)_{l\in\mathbb{N}}$ is free from $(d_N(i))_{i\in I}$ with respect to E for each $N \in \mathbb{N}$.
(2) $(\{u(l), u(l)^*\})_{l\in\mathbb{N}}$ is a free family with respect to E, and $u(l)$ is a Haar unitary, independent from $\mathcal{B}$ for each $l \in \mathbb{N}$.
(3) $(d_N(i))_{i\in I}$ has the same $\mathcal{B}$-valued joint distribution with respect to E as $(D_N(i))_{i\in I}$ has with respect to E_N.

Then for any polynomials $p_1, \ldots, p_{2m} \in \mathcal{B}\langle t(i)|i \in I\rangle$, $l_1, \ldots, l_{2m} \in \mathbb{N}$ and $\epsilon_1, \ldots, \epsilon_{2m} \in \{1, *\}$,

$$\begin{aligned}\Big\|&(\psi_N^{*\infty} \otimes E_N)[U_N(l_1)^{\epsilon_1} p_1(D_N) \cdots U_N(l_{2m})^{\epsilon_{2m}} p_{2m}(D_N)]\\ &-E[u(l_1)^{\epsilon_1} p_1(d_N) \cdots u(l_{2m})^{\epsilon_{2m}} p_{2m}(d_N)]\Big\|\end{aligned}$$

is $O(N^{-2})$ as $N \to \infty$.

Observe that Theorem 5.1 makes no assumption on the existence of a limiting distribution for $(D_N(i))_{i\in I}$. If one assumes also the existence of a limiting (infinitesimal) $\mathcal{B}$-valued joint distribution, then asymptotic (infinitesimal) freeness follows easily. We will state this as Theorem 5.3 below, let us first recall the relevant notions.

Definition 5.2. *Let $\mathcal{B}$ be a unital C*-algebra, and for each $N \in \mathbb{N}$ let $(D_N(i))_{i\in I}$ be a family of noncommutative random variables in a $\mathcal{B}$-valued probability space $(\mathcal{A}(N), E_N : \mathcal{A}(N) \to \mathcal{B})$.*

(1) *We say that the joint distribution of $(D_N(i))_{i\in I}$ converges weakly in norm if there is a $\mathcal{B}$-linear map $E : \mathcal{B}\langle D(i)|i \in I\rangle \to \mathcal{B}$ such that*

$$\lim_{N\to\infty} \|E_N[b_0 D_N(i_1) \cdots D_N(i_k) b_k] - E[b_0 D(i_1) \cdots D(i_k) b_k]\| = 0$$

for any $i_1, \ldots, i_k \in I$ and $b_0, \ldots, b_k \in \mathcal{B}$. If $\mathcal{B}$ is a von Neumann algebra with faithful, normal trace state τ, we say the joint distribution of $(D_N(i))_{i\in I}$ converges weakly in L^2 if the equation above holds with respect to $|\ |_2$.

(2) *If* $I = \bigcup_{j \in J} I_j$ *is a partition of* I*, we say that the sequence of sets of random variables* $(\{D_N(i)|i \in I_j\})_{j\in J}$ *are asymptotically free with amalgamation over* $\mathcal{B}$ *if the sets* $(\{D(i)|i \in I_j\})_{j\in J}$ *are freely independent with respect to* E.

(3) *We say that the joint distribution of* $(D_N(i))_{i\in I}$ *converges infinitesimally in norm if there is a* $\mathcal{B}$*-linear map* $E' : \mathcal{B}\langle D(i)|i \in I\rangle \to \mathcal{B}$ *such that*

$$E'[b_0 D(i_1)\cdots D(i_k)b_k] = \lim_{N\to\infty} N\{E_N[b_0 D_N(i_1)\cdots D_N(i_k)b_k] - E[b_0 D(i_1)\cdots D(i_k)b_k]\}$$

with convergence in norm, for any $b_0, \ldots, b_k \in \mathcal{B}$ *and* $i_1, \ldots, i_k \in I$*. If* $\mathcal{B}$ *is a von Neumann algebra with faithful, normal trace state* τ*, we say the joint distribution of* $(D_N(i))_{i\in I}$ *converges infinitesimally in* L^2 *if the equation above holds with respect to* $|\ |_2$.

(4) *If* $I = \bigcup_{j \in J} I_j$ *is a partition of* I*, we say that the sequence of sets of random variables* $(\{D_N(i)|i \in I_j\})_{j\in J}$ *are asymptotically infinitesimally free with amalgamation over* $\mathcal{B}$ *if the sets* $(\{D(i)|i \in I_j\})_{j\in J}$ *are infinitesimally freely independent with respect to* (E, E').

Theorem 5.3. Let $\mathcal{B}$ be a unital C*-algebra, and let $(D_N(i))_{i\in I}$ be a family of matrices in $M_N(\mathcal{B})$ for $N \in \mathbb{N}$. Suppose that there is a finite constant C such that $\|D_N(i)\| \le C$ for all $i \in I$ and $N \in \mathbb{N}$. For each $N \in \mathbb{N}$, let $(U_N(l))_{l\in\mathbb{N}}$ be a family of freely independent $N \times N$ Haar quantum unitary random matrices, independent from $\mathcal{B}$.

(1) If the joint distribution of $(D_N(i))_{i\in I}$ converges weakly (in norm or in L^2 with respect to a faithful trace), then the sets

$$(\{U_N(1), U_N(1)^*\}, \{U_N(2), U_N(2)^*\}, \ldots, \{D_N(i)|i \in I\})$$

are asymptotically free with amalgamation over $\mathcal{B}$ as $N \to \infty$.

(2) If the joint distribution of $(D_N(i))_{i\in I}$ converges infinitesimally (in norm or in L^2 with respect to a faithful trace), then the sets

$$(\{U_N(1), U_N(1)^*\}, \{U_N(2), U_N(2)^*\}, \ldots, \{D_N(i)|i \in I\})$$

are asymptotically infinitesimally free with amalgamation over $\mathcal{B}$ as $N \to \infty$.

Theorem 5.3 follows immediately from Theorem 5.1 and Proposition 2.16. The proof of Theorem 5.1 will require some preparation, we begin by computing the limiting distribution appearing in the statement.

Proposition 5.4. *Let* $(u(l), u(l)^*)_{l\in\mathbb{N}}$ *and* $(d_N(i))_{i\in I, N\in\mathbb{N}}$ *be random variables in a* $\mathcal{B}$*-valued probability space* $(\mathcal{A}, E : \mathcal{A} \to \mathcal{B})$ *such that*

(1) $(u(l), u(l)^*)_{l\in\mathbb{N}}$ *is free from* $(d_N(i))_{i\in I}$ *with respect to* E *for each* $N \in \mathbb{N}$.

(2) $(\{u(l), u(l)^*\})_{l\in\mathbb{N}}$ *is a free family with respect to* E*, and* $u(l)$ *is a Haar unitary, independent from* $\mathcal{B}$ *for each* $l \in \mathbb{N}$.

Let $a(1), \ldots, a(2m)$ *be in the algebra generated by* $\mathcal{B}$ *and* $\{d(i)|i \in I\}$*, and let* $l_1, \ldots, l_{2m} \in \mathbb{N}$ *and* $\epsilon_1, \ldots, \epsilon_{2m} \in \{1, *\}$*. Then*

$$E[u(l_1)^{\epsilon_1}a(1)\cdots u(l_{2m})^{\epsilon_{2m}}a(2m)] = \sum_{\substack{\pi,\sigma \in NC^{\epsilon}(m)\\ \sigma\le\pi \\ \widetilde{\pi}\vee\widetilde{\sigma}\le \ker \mathbf{l}}} \mu_m(\sigma,\pi) E^{(\sigma \wr K(\pi))}[a(1), \ldots, a(2m)].$$

Note that elements of the form appearing in the statement of the proposition span the algebra generated by $(u(l), u(l)^*)_{l\in\mathbb{N}}$ and $(d(i))_{i\in I}$, and so this indeed determines the joint distribution.

Proof. We have

$$E[u(l_1)^{\epsilon_1}a(1)\cdots u(l_{2m})^{\epsilon_{2m}}a(2m)] = \sum_{\alpha\in NC(4m)} \kappa_E^{\alpha}[u(l_1)^{\epsilon_1}, a(1), \ldots, a(2m)].$$

By freeness, the only non-vanishing cumulants appearing above are those of the form $\tau \wr \gamma$, where $\tau, \gamma \in NC(2m)$, $\tau \le \ker \mathbf{l}$ and $\gamma \le K(\tau)$. So we have

$$\begin{aligned} &E[u(l_1)^{\epsilon_1}a(1)\cdots u(l_{2m})^{\epsilon_{2m}}a(2m)] \\ &= \sum_{\substack{\tau\in NC(2m)\\ \tau\le\ker\mathbf{l}}} \sum_{\substack{\gamma\in NC(2m)\\ \gamma\le K(\tau)}} \kappa_E^{(\tau\wr\gamma)}[u(l_1)^{\epsilon_1}, a(1), \ldots, a(2m)]. \end{aligned}$$

Since the expectation of any polynomial in $(u(l), u(l)^*)_{l\in\mathbb{N}}$ with complex coefficients is scalar-valued, it follows that

$$\begin{aligned} &E[u(l_1)^{\epsilon_1}a(1)\cdots u(l_{2m})^{\epsilon_{2m}}a(2m)] \\ &= \sum_{\substack{\tau\in NC(2m)\\ \tau\le\ker\mathbf{l}}} \kappa_E^{(\tau)}[u(l_1)^{\epsilon_1}, \ldots, u(l_{2m})^{\epsilon_{2m}}] \sum_{\substack{\gamma\in NC(2m)\\ \gamma\le K(\tau)}} \kappa_E^{(\gamma)}[a(1), \ldots, a(2m)] \\ &= \sum_{\substack{\tau\in NC(2m)\\ \tau\le\ker\mathbf{l}}} \kappa_E^{(\tau)}[u(l_1)^{\epsilon_1}, \ldots, u(l_{2m})^{\epsilon_{2m}}] E^{(K(\tau))}[a(1), \ldots, a(2m)]. \end{aligned}$$

Since Haar unitaries are R-diagonal ([17, Example 15.4]), we have

$$\kappa_E^{(\tau)}[u(l_1)^{\epsilon_1}, \ldots, u(l_{2m})^{\epsilon_{2m}}] = 0$$

unless $\tau \in NC_h^{\epsilon}(2m)$. By Lemmas 3.6 and 3.9, we have

$$\begin{aligned} &E[u(l_1)^{\epsilon_1}a(1)\cdots u(l_{2m})^{\epsilon_{2m}}a(2m)] \\ &= \sum_{\substack{\pi,\sigma\in NC^{\epsilon}(m)\\ \sigma\le\pi\\ \widetilde{\sigma}\vee\widetilde{\pi}\le\ker\mathbf{l}}} \kappa_E^{(\widetilde{\sigma}\vee\widetilde{\pi})}[u(l_1)^{\epsilon_1}, \ldots, u(l_{2m})^{\epsilon_{2m}}] E^{(\sigma\wr K(\pi))}[a(1), \ldots, a(2m)]. \end{aligned}$$

So it remains only to show that if $\sigma, \pi \in NC^{\epsilon}(m)$ and $\sigma \le \pi$ then

$$\mu_m(\sigma, \pi) = \kappa_E^{(\widetilde{\sigma}\vee\widetilde{\pi})}[u(l_1)^{\epsilon_1}, \ldots, u(l_{2m})^{\epsilon_{2m}}].$$

Since the Möbius function is multiplicative on $NC(m)$, we have

$$\mu_m(\sigma, \pi) = \prod_{W\in\pi} \mu_{|W|}(\sigma|_W, 1_W),$$

and so it suffices to consider the case $\pi = 1_m$.

By [17, Prop. 15.1],

$$\kappa_E^{(\widetilde{\sigma}\vee\widetilde{1_m})}[u(l_1)^{\epsilon_1}, \ldots, u(l_{2m})^{\epsilon_{2m}}] = \prod_{V\in\widetilde{\sigma}\vee\widetilde{1_m}} (-1)^{|V|/2-1} C_{|V|/2-1},$$

where C_n is the n^{th} Catalan number. Since

$$\widetilde{\sigma} \vee \widetilde{1_m} = \overrightarrow{\overleftarrow{\widetilde{\sigma}} \vee \overleftarrow{\widetilde{1}}_m} = \overrightarrow{\widetilde{K(\sigma)} \vee \widehat{0_m}} = \overrightarrow{\widehat{K(\sigma)}},$$

we have

$$\kappa_E^{(\widetilde{\sigma} \vee \widetilde{1_m})}[u(l_1)^{\epsilon_1}, \ldots, u(l_{2m})^{\epsilon_{2m}}] = \prod_{W \in K(\sigma)} (-1)^{|W|-1} C_{|W|-1}.$$

On the other hand, we have

$$\begin{aligned} \mu_m(\sigma, 1_m) &= \mu_m(0_m, K(\sigma)) \\ &= \prod_{W \in K(\sigma)} \mu_{|W|}(0_W, 1_W) \\ &= \prod_{W \in K(\sigma)} (-1)^{|W|-1} C_{|W|-1}, \end{aligned}$$

where we have used the formula for $\mu_m(0_m, 1_m)$ from [17, Prop. 10.15].

Proposition 5.5. *Let $\mathcal{B}$ be a unital algebra, $A(1), \ldots, A(2m) \in M_N(\mathcal{B})$ and $\pi, \sigma \in NC(m)$. Let $E_N = \mathrm{tr}_N \otimes \mathrm{id}_{\mathcal{B}}$. If $\sigma \leq \pi$, then*

$$\begin{aligned} &\sum_{\substack{1 \leq j_1, \ldots, j_{2m} \leq N \\ \widetilde{\sigma} \leq \ker \mathbf{j}}} \sum_{\substack{1 \leq i_1, \ldots, i_{2m} \leq N \\ \widetilde{K(\pi)} \leq \ker \mathbf{i}}} A(1)_{j_1 j_2} A(2)_{i_1 i_2} \cdots A(2m)_{i_{2m-1} i_{2m}} \\ &= N^{|\sigma| + |K(\pi)|} E_N^{(\sigma \wr K(\pi))}[A(1), \ldots, A(2m)]. \end{aligned}$$

Proof. First observe that the sum above can be rewritten as

$$\sum_{\substack{1 \leq i_1, \ldots, i_{4m} \leq N \\ \sigma \wr \widetilde{K(\pi)} \leq \ker \mathbf{i}}} A(1)_{i_1 i_2} \cdots A(2m)_{i_{4m-1} i_{4m}}.$$

So this will follow from the formula

$$\sum_{\substack{1 \leq i_1, \ldots, i_{2m} \leq N \\ \widetilde{\sigma} \leq \ker \mathbf{i}}} A(1)_{i_1 i_2} \cdots A(m)_{i_{2m-1} i_{2m}} = N^{|\sigma|} E_N^{(\sigma)}[A(1), \ldots, A(m)]$$

for any $\sigma \in NC(m)$.

We will prove this by induction on the number of blocks of m. If $\sigma = 1_m$ has only one block, then we have

$$\begin{aligned} \sum_{\substack{1 \leq i_1, \ldots, i_{2m} \leq N \\ \widetilde{\sigma} \leq \ker \mathbf{i}}} A(1)_{i_1 i_2} \cdots A(m)_{i_{2m-1} i_{2m}} &= \sum_{1 \leq i_1, \ldots, i_m \leq N} A(1)_{i_1 i_2} A(2)_{i_2 i_3} \cdots A(m)_{i_m i_1} \\ &= N \cdot E_N(A(1) \cdots A(m)). \end{aligned}$$

Suppose now that $V = \{l+1, \ldots, l+s\}$ is an interval of σ. Then

$$\sum_{\substack{1 \le i_1, \ldots, i_{2m} \le N \\ \widetilde{\sigma} \le \ker \mathbf{i}}} A(1)_{i_1 i_2} \cdots A(m)_{i_{2m-1} i_{2m}}$$

$$= \sum_{\substack{1 \le i_1, \ldots, i_{2l-2}, \\ i_{2(l+s)+1}, \ldots, i_{2m} \le N \\ \widetilde{\sigma \setminus V} \le \ker \mathbf{i}}} A(1)_{i_1 i_2} \cdots \left(\sum_{1 \le j_1, \ldots, j_s \le N} A(l+1)_{j_1 j_2} \cdots A(l+s)_{j_s j_1} \right)$$

$$\cdots A(m)_{i_{2m-1} i_{2m}}$$

$$= \sum_{\substack{1 \le i_1, \ldots, i_{2l-2}, \\ i_{2(l+s)+1}, \ldots, i_{2m} \le N \\ \widetilde{\sigma \setminus V} \le \ker \mathbf{i}}} A(1)_{i_1 i_2} \cdots (N \cdot E_N(A(l+1) \cdots A(l+s))) \cdots A(m)_{i_{2m-1} i_{2m}},$$

which by induction is equal to

$$N^{|\sigma|} E_N^{(\sigma \setminus V)}[A(1), \ldots, A(l) E_N(A(l+1) \cdots A(l+s)), \ldots, A(m)]$$
$$= N^{|\sigma|} E_N^{(\sigma)}[A(1), \ldots, A(m)].$$

Remark 5.6. We will also need to control the sum appearing in the proposition above for $\sigma, \pi \in NC(m)$ with $\sigma \not\le \pi$. If $\mathcal{B}$ is commutative this poses no difficulty, as then

$$\sum_{\substack{1 \le j_1, \ldots, j_{2m} \le N \\ \widetilde{\sigma} \le \ker \mathbf{j}}} \sum_{\substack{1 \le i_1, \ldots, i_{2m} \le N \\ \widetilde{K(\pi)} \le \ker \mathbf{i}}} A(1)_{j_1 j_2} A(2)_{i_1 i_2} \cdots A(2m)_{i_{2m-1} i_{2m}}$$

$$= \left(\sum_{\substack{1 \le j_1, \ldots, j_{2m} \le N \\ \widetilde{\sigma} \le \ker \mathbf{j}}} A(1)_{j_1 j_2} \cdots A(2m-1)_{j_{2m-1} j_{2m}} \right)$$

$$\cdot \left(\sum_{\substack{1 \le i_1, \ldots, i_{2m} \le N \\ \widetilde{K(\pi)} \le \ker \mathbf{i}}} A(2)_{i_1 i_2} \cdots A(2m)_{i_{2m-1} i_{2m}} \right)$$

$$= N^{|\sigma| + |K(\pi)|} E_N^{(\sigma)}[A(1), \ldots, A(2m-1)] E_N^{(K(\pi))}[A(2), \ldots, A(2m)].$$

However, when $\mathcal{B}$ is noncommutative it is not clear how to express this sum in terms of expectation functionals. Instead, we will use the following bound on the norm:

Proposition 5.7. *Let $\mathcal{B}$ be a unital C^*-algebra, and $A(1), \ldots, A(2m) \in M_N(\mathcal{B})$. If $\sigma, \pi \in NC(m)$ then*

$$\left\| \sum_{\substack{1 \le j_1, \ldots, j_{2m} \le N \\ \widetilde{\sigma} \le \ker \mathbf{j}}} \sum_{\substack{1 \le i_1, \ldots, i_{2m} \le N \\ \widetilde{K(\pi)} \le \ker \mathbf{i}}} A(1)_{j_1 j_2} A(2)_{i_1 i_2} \cdots A(2m)_{i_{2m-1} i_{2m}} \right\|$$
$$\le N^{|\sigma| + |K(\pi)|} \|A(1)\| \cdots \|A(2m)\|.$$

Proof. For this proof, we extend the definition of $\widetilde{\pi}$ to all partitions $\pi \in \mathcal{P}(m)$ in the obvious manner. We can rewrite the expression above as

$$\sum_{\substack{1\le i_1,\ldots,i_{4m}\le N\\ \sigma\wr\widetilde{K(\pi)}\le\ker\mathbf{i}}} A(1)_{i_1i_2}\cdots A(2m)_{i_{4m-1}i_{4m}},$$

and so the result will follow from

$$\left\|\sum_{\substack{1\le i_1,\ldots,i_{2m}\le N\\ \widetilde{\sigma}\le\ker\mathbf{i}}} A(1)_{i_1i_2}\cdots A(m)_{i_{2m-1}i_{2m}}\right\| \le N^{|\sigma|}\|A(1)\|\cdots\|A(m)\|$$

for any partition $\sigma \in \mathcal{P}(m)$.

The idea now is to realize this expression as the trace of a larger matrix. For each $V \in \sigma$, let M_N^V be a copy of $M_N(\mathbb{C})$. Consider the algebra

$$\bigotimes_{V\in\sigma} M_N^V \simeq M_{N^{|\sigma|}}(\mathbb{C}),$$

with the natural unital inclusions ι_V of M_N^V for $V \in \sigma$. For $1 \le l \le m$, let

$$X(l) = \left(\iota_{\sigma(l)} \otimes \mathrm{id}_{\mathcal{B}}\right) A(l) \in \left(\bigotimes_{V\in\sigma} M_N^V\right) \otimes \mathcal{B} \simeq M_{N^{|\sigma|}}(\mathcal{B}),$$

where we have used the notation $\sigma(l)$ for the block of σ which contains l.

In other words, $X(l)$ is the matrix indexed by maps $i : \sigma \to [N] = \{1,\ldots,N\}$ such that

$$X(l)_{ij} = A(l)_{i(\sigma(l))j(\sigma(l))} \prod_{\substack{V\in\sigma\\ l\notin V}} \delta_{i(V)j(V)}.$$

Consider now the trace

$$\begin{aligned}(\mathrm{Tr}_{N^{|\sigma|}} \otimes \mathrm{id}_{\mathcal{B}})(X(1)\cdots X(m)) &= \sum_{\substack{i_1,\ldots,i_m\\ i_l:\sigma\to[N]}} X(1)_{i_1i_2}\cdots X(m)_{i_mi_1}\\ &= \sum_{\substack{i_1,\ldots,i_m\\ i_l:\sigma\to[N]}} A(1)_{i_1(\sigma(1))i_2(\sigma(1))}\cdots A(m)_{i_m(\sigma(m))i_1(\sigma(m))} \prod_{1\le l\le m}\prod_{\substack{V\in\sigma\\ l\notin V}} \delta_{i_l(V)i_{\gamma(l)}(V)},\end{aligned}$$

where $\gamma \in S_m$ is the cyclic permutation $(123\cdots m)$. The nonzero terms in this sum are obtained as follows: For each block $V = (l_1 < \cdots < l_s)$ of σ, choose $1 \le i_{l_1}(V), i_{\gamma(l_1)}(V), \ldots, i_{l_s}(V), i_{\gamma(l_s)}(V) \le N$ with the restrictions $i_{\gamma(l_1)}(V) = i_{l_2}(V), \ldots, i_{\gamma(l_{s-1})}(V) = i_{l_s}(V)$ and $i_{\gamma(l_s)}(V) = i_{l_1}(V)$. Comparing with the definition of $\widetilde{\sigma}$, it follows that

$$(\mathrm{Tr}_{N^{|\sigma|}} \otimes \mathrm{id}_{\mathcal{B}})(X(1)\cdots X(m)) = \sum_{\substack{1\le i_1,\ldots,i_{2m}\le N\\ \widetilde{\sigma}\le\ker\mathbf{i}}} A(1)_{i_1i_2}\cdots A(m)_{i_{2m-1}i_{2m}}$$

is the expression to be bounded. However, $(\mathrm{tr}_{N^{|\sigma|}} \otimes \mathrm{id}_{\mathcal{B}}) = N^{-|\sigma|}(\mathrm{Tr}_{N^{|\sigma|}} \otimes \mathrm{id}_{\mathcal{B}})$ is a contractive conditional expectation onto $\mathcal{B}$ and so

$$\|(\mathrm{Tr}_{N^{|\sigma|}} \otimes \mathrm{id}_{\mathcal{B}})(X(1) \cdots X(m))\| \leq N^{|\sigma|}\|X(1)\| \cdots \|X(m)\|.$$

Since $(\iota_V \otimes \mathrm{id}_{\mathcal{B}})$ is a contractive $*$-homomorphism, we have $\|X(l)\| = \|(\iota_{\sigma(l)} \otimes \mathrm{id}_{\mathcal{B}})(A(l))\| \leq \|A(l)\|$ and the result follows.

We are now prepared to prove the main theorem.

Proof of Theorem 5.1. Fix $p_1, \ldots, p_{2m} \in \mathcal{B}\langle t(i)|i \in I\rangle$, and set $A_N(k) = p_k(D_N)$ for $1 \leq k \leq 2m$. For notational simplicity, we will suppress the subscript N in our computations.

Let $l_1, \ldots, l_{2m} \in \mathbb{N}$, $\epsilon_1, \ldots, \epsilon_{2m} \in \{1, *\}$ and consider

$$\begin{aligned}
&(\psi_N^{*\infty} \otimes E_N)[U(l_1)^{\epsilon_1} A(1) U(l_2)^{\epsilon_2} \cdots U(l_{2m})^{\epsilon_{2m}} A(2m)] \\
&\quad = (\psi_N^{*\infty} \otimes \mathrm{id}_{\mathcal{B}}) N^{-1} \sum_{1 \leq i_1, \ldots, i_{4m} \leq N} (U(l_1)^{\epsilon_1})_{i_1 i_2} A(1)_{i_2 i_3} (U(l_2)^{\epsilon_2})_{i_3 i_4} \cdots A(2m)_{i_{4m} i_1} \\
&\quad = \sum_{1 \leq i_1, \ldots, i_{4m} \leq N} N^{-1} \psi_N^{*\infty} \left[(U(l_1)^{\epsilon_1})_{i_1 i_2} \cdots (U(l_{2m})^{\epsilon_{2m}})_{i_{4m-1} i_{4m}}\right] \\
&\qquad \cdot A(1)_{i_2 i_3} \cdots A(2m)_{i_{4m} i_1}.
\end{aligned}$$

By Corollary 4.4, this is equal to

$$\begin{aligned}
&\sum_{1 \leq i_1, \ldots, i_{4m} \leq N} N^{-1} \psi_N^{*\infty} \left[U(l_1)^{\epsilon_1}_{i_1 i_2} U(l_2)^{\epsilon_2}_{i_4 i_3} \cdots U(l_{2m})^{\epsilon_{2m}}_{i_{4m} i_{4m-1}}\right] \\
&\qquad \cdot A(1)_{i_2 i_3} \cdots A(2m)_{i_{4m} i_1}.
\end{aligned}$$

After reindexing, this becomes

$$\begin{aligned}
&\sum_{1 \leq i_1, \ldots, i_{2m} \leq N} \sum_{1 \leq j_1, \ldots, j_{2m} \leq N} N^{-1} \psi_N^{*\infty} \left[U(l_1)^{\epsilon_1}_{i_{2m} j_1} U(l_2)^{\epsilon_2}_{i_1 j_2} \cdots U(l_{2m})^{\epsilon_{2m}}_{i_{2m-1} j_{2m}}\right] \\
&\qquad \cdot A(1)_{j_1 j_2} A(2)_{i_1 i_2} \cdots A(2m)_{i_{2m-1} i_{2m}}.
\end{aligned}$$

Applying Corollary 4.7, we have

$$\begin{aligned}
&\sum_{1 \leq i_1, \ldots, i_{2m} \leq N} \sum_{1 \leq j_1, \ldots, j_{2m} \leq N} \sum_{\substack{\pi, \sigma \in NC^{\epsilon}(m) \\ \widetilde{\pi} \leq \ker \overrightarrow{\mathbf{i}} \wedge \ker \mathbf{l} \\ \widetilde{\sigma} \leq \ker \mathbf{j} \wedge \ker \mathbf{l}}} N^{-|K(\pi)| - |\sigma|}(\mu_m(\sigma, \pi) + O(N^{-2})) \\
&\qquad \cdot A(1)_{j_1 j_2} A(2)_{i_1 i_2} \cdots A(2m)_{i_{2m-1} i_{2m}} \\
&\quad = \sum_{\substack{\pi, \sigma \in NC^{\epsilon}(m) \\ \widetilde{\pi} \leq \ker \mathbf{l} \\ \widetilde{\sigma} \leq \ker \mathbf{l}}} (\mu_m(\sigma, \pi) + O(N^{-2})) N^{-|K(\pi)| - |\sigma|} \\
&\qquad \times \sum_{\substack{1 \leq j_1, \ldots, j_{2m} \leq N \\ \widetilde{\sigma} \leq \ker \mathbf{j}}} \sum_{\substack{1 \leq i_1, \ldots, i_{2m} \leq N \\ \widetilde{K(\pi)} \leq \ker \mathbf{i}}} A(1)_{j_1 j_2} A(2)_{i_1 i_2} \cdots A(2m)_{i_{2m-1} i_{2m}}.
\end{aligned}$$

By Propositions 5.5 and 5.7, this is equal to

$$\sum_{\substack{\pi,\sigma\in NC^{\epsilon}(m)\\ \sigma\leq\pi\\ \widetilde{\pi}\vee\widetilde{\sigma}\leq\ker\mathbf{l}}} \mu_m(\sigma,\pi)E_N^{(\sigma\wr K(\pi))}[A(1),\dots,A(2m)],$$

up to $O(N^{-2})$ with respect to the norm on $\mathcal{B}$. Set $a(k) = p_k(d_N)$ for $1 \leq k \leq 2m$, then by Proposition 5.4 we have

$$\begin{aligned} &E[u(l_1)^{\epsilon_1}a(1)\cdots u(l_{2m})^{\epsilon_{2m}}a(2m)] \\ &= \sum_{\substack{\pi,\sigma\in NC^{\epsilon}(m)\\ \sigma\leq\pi\\ \widetilde{\pi}\vee\widetilde{\sigma}\leq\ker\mathbf{l}}} \mu_m(\sigma,\pi)E^{(\sigma\wr K(\pi))}[a(1),\dots,a(2m)] \\ &= \sum_{\substack{\pi,\sigma\in NC^{\epsilon}(m)\\ \sigma\leq\pi\\ \widetilde{\pi}\vee\widetilde{\sigma}\leq\ker\mathbf{l}}} \mu_m(\sigma,\pi)E_N^{(\sigma\wr K(\pi))}[A(1),\dots,A(2m)], \end{aligned}$$

and the result now follows immediately.

5.8. Randomly quantum rotated matrices. It follows easily from Theorem 5.3 and the definition of asymptotic freeness that under the hypotheses of the theorem, the sets

$$(\{D_N(i) : i \in I\}, (\{U_N(l)D_N(i)U_N(l)^* : i \in I\})_{l\in\mathbb{N}})$$

are asymptotically (infinitesimally) free with amalgamation over $\mathcal{B}$ as $N \to \infty$. The condition on existence of a limiting joint distribution can be weakened slightly as follows:

Corollary 5.9. *Let $\mathcal{B}$ be a unital C^*-algebra, and let $(D_N(i))_{i\in I}$ and $(D'_N(j))_{j\in J}$ be two families of matrices in $M_N(\mathcal{B})$ for $N \in \mathbb{N}$. Suppose that there is a finite constant C such that $\|D_N(i)\| \leq C$ and $\|D'_N(j)\| \leq C$ for $N \in \mathbb{N}$, $i \in I$ and $j \in J$. For each $N \in \mathbb{N}$, let U_N be a $N \times N$ Haar quantum unitary random matrix, independent from $\mathcal{B}$.*

(1) *If the joint distributions of $(D_N(i))_{i\in I}$ and $(D'_N(j))_{j\in J}$ both converge weakly (in norm or in L^2 with respect to a faithful trace), then $(U_N D_N(i)U_N^*)_{i\in I}$ and $(D'_N(j))_{j\in J}$ are asymptotically free with amalgamation over $\mathcal{B}$ as $N \to \infty$.*
(2) *If the joint distribution of $(D_N(i))_{i\in I}$ and $(D'_N(j))_{j\in J}$ both converge infinitesimally (in norm or in L^2 with respect to a faithful trace), then the families $(U_N D_N(i)U_N^*)_{i\in I}$ and $(D'_N(j))_{j\in J}$ are asymptotically infinitesimally free with amalgamation over $\mathcal{B}$.*

Proof. The only condition of Theorem 5.3 which is not satisfied is that $\{D_N(i) : i \in I\} \cup \{D'_N(j) : j \in J\}$ should have a limiting (infinitesimal) joint distribution as $N \to \infty$. We can see that this is not an issue as follows. Let $p_1,\dots,p_m \in B\langle t(i)|i \in I\rangle$ and $q_1,\dots,qm \in B\langle t(j)|j \in J\rangle$ and set $A_N(k) = p_k(D_N)$, $B_N(k) = q_k(D'_N)$ for $1 \leq k \leq m$. From the proof of Theorem 5.1, we have

$$\begin{aligned} &(\psi_N \otimes E_N)[UA(1)U^*B(1)\cdots UA(m)U^*B(m)] \\ &= \sum_{\substack{\pi,\sigma\in NC(m)\\ \sigma\leq\pi}} \mu_m(\sigma,\pi)E_N^{(\sigma\wr K(\pi))}[A(1),B(1),\dots,A(m),B(m)], \end{aligned}$$

up to $O(N^{-2})$. But the right-hand side depends only on the distributions of $(D(i))_{i\in I}$ and $(D'(j))_{j\in J}$, and so the result follows from Theorem 5.3.

5.10. Classical Haar unitary random matrices. In the remainder of this section, we will discuss the failure of these results for classical Haar unitaries. First we show that if $\mathcal{B}$ is finite dimensional, then classical Haar unitaries are sufficient.

Proposition 5.11. *Let $\mathcal{B}$ be a finite dimensional C^*-algebra, and let $(D_N(i))_{i\in I}$ be a family of matrices in $M_N(\mathcal{B})$ for each $N \in \mathbb{N}$. Assume that there is a finite constant C such that $\|D_N(i)\| \le C$ for all $N \in \mathbb{N}$ and $i \in I$. For each $N \in \mathbb{N}$, let $(U_N(l))_{l\in\mathbb{N}}$ be a family of independent $N \times N$ Haar unitary random matrices, independent from $\mathcal{B}$. Let $(u(l), u(l)^*)_{l\in\mathbb{N}}$ and $(d_N(i))_{i\in I, N\in\mathbb{N}}$ be random variables in a $\mathcal{B}$-valued probability space $(\mathcal{A}, E : \mathcal{A} \to \mathcal{B})$ such that*

1. *$(u(l), u(l)^*)_{l\in\mathbb{N}}$ is free from $(d_N(i))_{i\in I}$ with respect to E for each $N \in \mathbb{N}$.*
2. *$(\{u(l), u(l)^*\})_{l\in\mathbb{N}}$ is a free family with respect to E, and $u(l)$ is a Haar unitary, independent from $\mathcal{B}$ for each $l \in \mathbb{N}$.*
3. *$(d_N(i))_{i\in I}$ has the same $\mathcal{B}$-valued joint distribution with respect to E as $(D_N(i))_{i\in I}$ has with respect to E_N.*

Then for any polynomials $p_1, \ldots, p_{2m} \in \mathcal{B}\langle t(i) : i \in I\rangle$, $l_1, \ldots, l_{2m} \in \mathbb{N}$ and $\epsilon_1, \ldots, \epsilon_{2m} \in \{1, *\}$,

$$\Big\|(\psi_N^{*\infty} \otimes E_N)[U_N(l_1)^{\epsilon_1} p_1(D_N) \cdots U_N(l_{2m})^{\epsilon_{2m}} p_{2m}(D_N)] \\ - E[u(l_1)^{\epsilon_1} p_1(d_N) \cdots u(l_{2m})^{\epsilon_{2m}} p_{2m}(d_N)]\Big\|$$

is $O(N^{-2})$ as $N \to \infty$.

Proof. Let $e_1, \ldots, e_q$ be a basis for $\mathcal{B}$ with $\|e_r\| = 1$ for $1 \le r \le q$. Let $p_1, \ldots, p_{2m} \in \mathcal{B}\langle t(i)|i \in I\rangle$, let $A_N(k) = p_k(D_N)$ and let $A_N(k, r) \in M_N(\mathbb{C})$ be the matrix of coefficients of the entries of $A_N(k)$ on e_r for $1 \le k \le 2m$ and $1 \le r \le q$. Let $a_N(k, r)$ and $(u(l), u(l)^*)_{l\in\mathbb{N}}$ be random variables in a noncommutative probability space $(\mathcal{A}, \varphi)$ such that

1. $\{a_N(k, r) : 1 \le k \le 2m, 1 \le r \le q\}$ and $(u(l), u(l)^*)_{l\in\mathbb{N}}$ are free with respect to φ.
2. $(a_N(k, r))_{1\le k\le 2m, 1\le r\le q}$ has the same joint distribution with respect to φ as $(A_N(k, r))_{1\le k\le 2m, 1\le r\le q}$ with respect to tr_N.
3. $(u(l), u(l)^*)_{l\in\mathbb{N}}$ are freely independent with respect to φ and $u(l)$ has a Haar unitary distribution.

For $1 \le k \le 2m$ and $N \in \mathbb{N}$, let $a_N(k) = \sum a_N(k, r) \otimes e_r \in \mathcal{A} \otimes \mathcal{B}$, and note that the family $(a_n(k))_{1\le k\le 2m}$ has the same joint distribution with respect to $E = \varphi \otimes \mathrm{id}_{\mathcal{B}}$ as does $(A_N(k))_{1\le k\le 2m}$ with respect to E_N. Identifying $u(l) = u(l) \otimes 1$ in $\mathcal{A}\otimes\mathcal{B}$, it is also easy to see that $(u(l), u(l)^*)$ and $(a_N(k))_{1\le k\le 2m}$ are freely independent with respect to E.

Now let $\epsilon_1, \ldots, \epsilon_{2m} \in \{1, *\}$ and consider

$$(\mathrm{tr}_N \otimes \mathbb{E} \otimes \mathrm{id}_{\mathcal{B}})[U(l_1)^{\epsilon_1} A(1) \cdots A(2m) U(l_{2m})^{\epsilon_{2m}}] \\ = \sum_{1\le r_1,\ldots,r_{2m}\le q} (\mathrm{tr}_N \otimes \mathbb{E})[U(l_1)^{\epsilon_1} A(1, r_1) \cdots A(2m, r_{2m}) U(l_{2m})^{\epsilon_{2m}}] e_{r_1} \cdots e_{r_{2m}}.$$

Since $\|e_r\| = 1$, it follows that

$$\begin{aligned}&\Big\|(\mathrm{tr}_N \otimes \mathbb{E} \otimes \mathrm{id}_{\mathcal{B}})[U(l_1)^{\epsilon_1} A(1) \cdots U(l_{2m})^{\epsilon_{2m}} A(2m)] \\ &\quad - E[u(l_1)^{\epsilon_1} a(1) \cdots u(l_{2m})^{\epsilon_{2m}} a(2m)]\Big\| \\ &\le \sum_{1 \le r_1, \dots, r_{2m} \le q} \Big|(\mathrm{tr}_N \otimes \mathbb{E})[U(l_1)^{\epsilon_1} A(1, r_1) \cdots U(l_{2m})^{\epsilon_{2m}} A(2m, r_{2m})] \\ &\quad - \varphi[u(l_1)^{\epsilon_1} a(1, r_1) \cdots u(l_{2m})^{\epsilon_{2m}} a(2m, r_{2m})]\Big|.\end{aligned}$$

From standard asymptotic freeness results (see e.g. [10]), this expression is $O(N^{-2})$ as $N \to \infty$.

We will now give an example to show that Theorem 5.1 may fail for classical Haar unitaries if the algebra $\mathcal{B}$ is infinite dimensional. First we recall the Weingarten formula for computing the expectation of a word in the entries of a $N \times N$ Haar unitary random matrix and its conjugate:

$$\mathbb{E}[U_{i_1 j_1}^{\epsilon_1} \cdots U_{i_{2m} j_{2m}}^{\epsilon_{2m}}] = \sum_{\substack{\pi, \sigma \in \mathcal{P}_2^{\epsilon}(2m) \\ \pi \le \ker \mathbf{i} \\ \sigma \le \ker \mathbf{j}}} W_{\epsilon N}^{c}(\pi, \sigma),$$

where $\mathcal{P}_2^{\epsilon}(2m)$ is the set of pair partitions for which each pairing connects a 1 with a $*$ in the string $\epsilon_1, \dots, \epsilon_{2m}$, and $W_{\epsilon N}^{c}$ is the corresponding Weingarten matrix, see [5,10].

Example 5.12. Let $\mathcal{B}$ be a unital C*-algebra, and for each $N \in \mathbb{N}$ let $\{E_{ij}(N, l) : 1 \le i, j \le N, l = 1, 2\}$ be two commuting systems of matrix units in $\mathcal{B}$, i.e.,

(1) $E_{i_1 j_1}(N, 1) E_{i_2 j_2}(N, 2) = E_{i_2 j_2}(N, 2) E_{i_1 j_1}(N, 1)$ for $1 \le i_1, j_1, i_2, j_2 \le N$.
(2) $E_{ij}(N, l)^* = E_{ji}(N, l)$ for $1 \le i, j \le N$.
(3) $E_{ik_1}(N, l) E_{k_2 j}(N, l) = \delta_{k_1 k_2} E_{ij}(N, l)$ for $1 \le i, j, k_1, k_2 \le N$.
(4) $E_{ii}(N, l)$ is a projection for $1 \le i \le N$, and

$$\sum_{i=1}^{N} E_{ii}(N, l) = 1.$$

For $N \in \mathbb{N}$, define $A_N, B_N \in M_N(\mathcal{B})$ by

$$(A_N)_{ij} = E_{ji}(N, 1), \quad (B_N)_{ij} = E_{ji}(N, 2).$$

Note that A_N, B_N are self-adjoint and A_N^2, B_N^2 are the identity matrix, indeed

$$(A_N^2)_{ij} = \sum_{k=1}^{N} E_{ki}(N, 1) E_{jk}(N, 1) = \delta_{ij} \sum_{k=1}^{N} E_{kk}(N, 1) = \delta_{ij} \cdot 1,$$

and likewise for B_N. It follows that $\|A_N\| = \|B_N\| = 1$ for $N \in \mathbb{N}$.

For each $N \in \mathbb{N}$, let U_N be a $N \times N$ Haar unitary random matrix, independent from $\mathcal{B}$. Since

$$(\mathrm{tr}_N \otimes \mathrm{id}_{\mathcal{B}})[A_N] = \frac{1}{N} \sum_{i=1}^{N} E_{ii}(N, 1) = \frac{1}{N} \cdot 1$$

converges to zero as $N \to \infty$, and likewise for B_N, for asymptotic freeness we should have

$$\lim_{N\to\infty} (\mathrm{tr}_N \otimes \mathbb{E} \otimes \mathrm{id})[(U_N A_N U_N^* B_N)^3] = 0.$$

However, we will show that this limit is in fact equal to 1.

Indeed, suppressing the subindex N we have

$$\begin{aligned}
&(\mathrm{tr} \otimes \mathbb{E} \otimes \mathrm{id}_{\mathcal{B}})[(UAU^*B)^3] \\
&= \frac{1}{N} \sum_{1\le i_1,\dots,i_{12}\le N} \mathbb{E}[U_{i_1 i_2}\overline{U}_{i_4 i_3}\cdots \overline{U}_{i_{12} i_{11}}]A_{i_2 i_3}B_{i_4 i_5}\cdots B_{i_{12} i_1} \\
&= \sum_{1\le i_1, j_1,\dots,i_6, j_6\le N} \mathbb{E}[U_{i_6 j_1}\overline{U}_{i_1 j_2}\cdots \overline{U}_{i_5 j_6}]A_{j_1 j_2}A_{j_3 j_4}A_{j_5 j_6}B_{i_1 i_2}B_{i_3 i_4}B_{i_5 i_6}.
\end{aligned}$$

Applying the Weingarten formula, we obtain

$$\sum_{\pi,\sigma\in\mathcal{P}_2^{\epsilon}(6)} N^{-1} W^c_{\epsilon N}(\pi,\sigma) \left(\sum_{\substack{1\le j_1,\dots,j_6\le N \\ \sigma\le \ker \mathbf{j}}} A_{j_1 j_2}A_{j_3 j_4}A_{j_5 j_6} \right) \cdot \left(\sum_{\substack{1\le i_1,\dots,i_6\le N \\ \overleftarrow{\pi}\le \ker \mathbf{i}}} B_{i_1 i_2}B_{i_3 i_4}B_{i_5 i_6} \right).$$

Note that $\mathcal{P}_2^{\epsilon}(6)$ has 6 elements, namely the 5 noncrossing pair partitions and $\tau = \{(1,4),(2,5),(3,6)\}$. The noncrossing pair partitions can be expressed as $\widetilde{\sigma}$ for some $\sigma \in NC(3)$, in which case we have

$$\sum_{\substack{1\le j_1,\dots,j_6\le N \\ \widetilde{\sigma}\le \ker \mathbf{j}}} A_{j_1 j_2}A_{j_3 j_4}A_{j_5 j_6} = N^{|\sigma|} E_N^{(\sigma)}[A, A, A].$$

Using $E_N[A] = E_N[A^3] = N^{-1}$ and $E_N[A^2] = 1$, one easily sees that this expression is $O(N)$ for the 5 noncrossing pair partitions. For τ, we have

$$\begin{aligned}
\sum_{\substack{1\le j_1,\dots,j_6\le N \\ \tau\le \ker \mathbf{j}}} A_{j_1 j_2}A_{j_3 j_4}A_{j_5 j_6} &= \sum_{1\le j_1, j_2, j_3\le N} A_{j_1 j_2}A_{j_3 j_1}A_{j_2 j_3} \\
&= \sum_{1\le j_1, j_2, j_3\le N} E_{j_2 j_1}(N,1)E_{j_1 j_3}(N,1)E_{j_3 j_2}(N,1) \\
&= \sum_{1\le j_1, j_2, j_3\le N} E_{j_2 j_2}(N,1) \\
&= N^2 \cdot 1,
\end{aligned}$$

and likewise for B_N. Also we have $N^3 W^c_{\epsilon N}(\pi,\sigma) = \delta_{\pi\sigma} + O(N^{-1})$. Putting these statements together, we find that the only term which remains in the limit comes from $\pi = \sigma = \tau$, which gives 1.

Remarks. (1) We note that $M_{N^2}(\mathbb{C}) = M_N(\mathbb{C}) \otimes M_N(\mathbb{C})$ has a natural pair of commuting systems of matrix units, so this example demonstrates that Theorem 5.1 fails for any unital C*-algebra $\mathcal{B}$ which contains $M_{N_k^2}(\mathbb{C})$ as a unital subalgebra for some increasing sequence of natural numbers (N_k).

(2) It is a natural question whether the matrices A_N, B_N in the above example have limiting $\mathcal{B}$-valued distributions, which would demonstrate that Theorem 1 also fails for classical Haar unitaries. First observe that

$$\lim_{N\to\infty} (\mathrm{tr}_N \otimes \mathcal{B})[A_N^k] = \begin{cases} 1, & k \text{ is even} \\ 0, & k \text{ is odd} \end{cases},$$

which follows from the case $k = 1$ and the fact that A_N^2 is the identity matrix. However, it is not clear that moments of the form $b_0 A_N \cdots A_N b_k$ will converge for arbitrary $b_0, \dots, b_k \in \mathcal{B}$.

Let us point out a special case in which the limiting distribution does exist. Suppose that there is a dense $*$-subalgebra $\mathcal{F} \subset \mathcal{B}$ such that each element of $\mathcal{F}$ commutes with the matrix units $E_{ij}(N, l)$ for N sufficiently large. Then for any $b_0, \dots, b_k \in \mathcal{B}$ we have

$$\lim_{N\to\infty} (\mathrm{tr}_N \otimes \mathcal{B})[b_0 A_N \cdots A_N b_k] = \begin{cases} b_0 b_1 \cdots b_k, & k \text{ is even} \\ 0, & k \text{ is odd} \end{cases},$$

and likewise for B_N, indeed this holds for $b_0, \dots, b_k \in \mathcal{F}$ by hypothesis and for general $b_0, \dots, b_k$ by density.

In particular, we may take $\mathcal{B}$ to be the C*-algebraic infinite tensor product

$$\mathcal{B} = \bigotimes_{N\in\mathbb{N}} M_N(\mathbb{C})$$

with the obvious systems of matrix units $E(N, l)_{ij} \in M_{N^2} = M_N(\mathbb{C}) \otimes M_N(\mathbb{C}) \subset \mathcal{B}$, and $\mathcal{F} \subset \mathcal{B}$ to be the image of the purely algebraic tensor product. Note that $\mathcal{B}$ is *uniformly hyperfinite*, in particular *approximately finitely dimensional* in the C*-sense.

(3) Note that if $\mathcal{B}$ is a von Neumann algebra with a non-zero *continuous* projection p, then $p\mathcal{B}p$ contains $M_N(\mathbb{C})$ as a unital subalgebra for all $N \in \mathbb{N}$ and hence (1) applies to $p\mathcal{B}p$. It follows that Theorem 5.1 fails also for $\mathcal{B}$. To obtain a contradiction to Theorem 1 for classical Haar unitaries in the setting of a von Neumann algebra with faithful, normal trace, we may modify the example in (2) by taking $(\mathcal{B}, \tau)$ to be the infinite tensor product

$$(\mathcal{B}, \tau) = \bigotimes_{N\in\mathbb{N}} (M_N(\mathbb{C}), \mathrm{tr}_N)$$

taken with respect to the trace states tr_N on $M_N(\mathbb{C})$, which is the *hyperfinite* II_1 *factor*.

Acknowledgements. We would like to thank T. Banica, M. Neufang, and D. Shlyakhtenko for several useful discussions. S.C. would like to thank his thesis advisor, D.-V. Voiculescu, for his continued guidance and support while completing this project.

References

1. Banica, T.: Le groupe quantique compact libre U(n). Commun. Math. Phys. **190**, 143–172 (1997)
2. Banica, T., Collins, B.: Integration over compact quantum groups. Publ. Res. Inst. Math. Sci. **43**, 277–302 (2007)
3. Banica, T., Collins, B.: Integration over quantum permutation groups. J. Funct. Anal. **242**, 641–657 (2007)
4. Banica, T., Curran, S., Speicher, R.: *De Finetti theorems for easy quantum groups*. Ann. Probab. (in press)
5. Banica, T., Speicher, R.: Liberation of orthogonal Lie groups. Adv. Math. **222**, 1461–1501 (2009)
6. Belinschi, S.T., Shlyakhtenko, D.: *Free probability of type B: Analytic interpretation and applications*. http://arxiv.org/abs/0903.2721v1 [math.OA], 2009
7. Biane, P., Goodman, F., Nica, A.: Non-crossing cumulants of type B. Trans. Amer. Math. Soc. **355**, 2263–2303 (2003)
8. Brown, N.P., Dykema, K.J., Jung, K.: Free entropy dimension in amalgamated free products. Proc. Lond. Math. Soc. **97**(3), 339–367 (2008), With an appendix by Wolfgang Lück
9. Chen, Q., Przytycki, J.H.: The Gram matrix of a Temperley-Lieb algebra is similar to the matrix of chromatic joins. Commun. Contemp. Math. **10**, 849–855 (2008)
10. Collins, B.: Moments and cumulants of polynomial random variables on unitary groups, the Itzykson-Zuber integral, and free probability. Int. Math. Res. Not. 953–982 (2003)
11. Collins, B., Śniady, P.: Integration with respect to the Haar measure on unitary, orthogonal and symplectic group. Commun. Math. Phys. **264**, 773–795 (2006)
12. Curran, S.: Quantum exchangeable sequences of algebras. Indiana Univ. Math. J. **58**, 1097–1126 (2009)
13. Curran, S.: Quantum rotatability. Trans. Amer. Math. Soc. **362**, 4831–4851 (2010)
14. Fevrier, M., Nica, A.: *Infinitesimal non-crossing cumulants and free probability of type B*. http://arxiv.org/abs/0906.2017v1 [math.OA], 2009
15. Kodiyalam, V., Sunder, V.S.: *Temperley-Lieb and non-crossing partition planar algebras*. In: Noncommutative rings, group rings, diagram algebras and their applications, Vol. **456** of Contemp. Math., Providence, RI: Amer. Math. Soc., 2008, pp. 61–72
16. Köstler, C., Speicher, R.: A noncommutative de Finetti theorem: invariance under quantum permutations is equivalent to freeness with amalgamation. Commun. Math. Phys. **291**, 473–490 (2009)
17. Nica, A., Speicher, R.: *Lectures on the combinatorics of free probability*. Vol. **335** of London Mathematical Society Lecture Note Series. Cambridge University Press, Cambridge (2006)
18. Speicher, R.: Combinatorial theory of the free product with amalgamation and operator-valued free probability theory. Mem. Amer. Math. Soc. **132**, x+88 (1998)
19. Voiculescu, D.: *Symmetries of some reduced free product C^*-algebras*. In: Operator algebras and their connections with topology and ergodic theory (Buşteni, 1983), Vol. **1132** of Lecture Notes in Math., Berlin: Springer, 1985, pp. 556–588
20. Voiculescu, D.: Limit laws for random matrices and free products. Invent. Math. **104**, 201–220 (1991)
21. Wang, S.: Free products of compact quantum groups. Commun. Math. Phys. **167**, 671–692 (1995)
22. Wang, S.: Quantum symmetry groups of finite spaces. Commun. Math. Phys. **195**, 195–211 (1998)
23. Woronowicz, S.L.: Compact matrix pseudogroups. Commun. Math. Phys. **111**, 613–665 (1987)

Communicated by Y. Kawahigashi