Available online at www.sciencedirect.com
ScienceDirect
Advances in Mathematics 209 (2007) 212–240
ADVANCES IN Mathematics
www.elsevier.com/locate/aim

Second order freeness and fluctuations of random matrices: II. Unitary random matrices

James A. Mingo [a,*,1], Piotr Śniady [b,2], Roland Speicher [a,1,3]

[a] *Queen's University, Department of Mathematics and Statistics, Jeffery Hall, Kingston, ON, K7L 3N6, Canada*
[b] *Instytut Matematyczny, Uniwersytet Wroclawski, plac Grunwaldzki 2/4, 50-384 Wroclaw, Poland*

Received 6 July 2005; accepted 4 May 2006

Available online 14 June 2006

Communicated by Dan Voiculescu

Abstract

We extend the relation between random matrices and free probability theory from the level of expectations to the level of fluctuations. We show how the concept of "second order freeness", which was introduced in Part I, allows one to understand global fluctuations of Haar distributed unitary random matrices. In particular, independence between the unitary ensemble and another ensemble goes in the large N limit over into asymptotic second order freeness. Two important consequences of our general theory are: (i) we obtain a natural generalization of a theorem of Diaconis and Shahshahani to the case of several independent unitary matrices; (ii) we can show that global fluctuations in unitarily invariant multi-matrix models are not universal.

Keywords: Free probability; Haar distributed unitary random matrices; Second order freeness

* Corresponding author.
E-mail addresses: mingo@mast.queensu.ca (J.A. Mingo), piotr.sniady@math.uni.wroc.pl (P. Śniady), speicher@mast.queensu.ca (R. Speicher).
[1] Research supported by Discovery Grants and a Leadership Support Initiative Award from the Natural Sciences and Engineering Research Council of Canada.
[2] Research supported by State Committee for Scientific Research (KBN) grant 2 P03A 007 23, RTN network: QP-Applications contract No. HPRN-CT-2002-00279, and KBN-DAAD project 36/2003/2004. The author is a holder of a scholarship from the European Post-Doctoral Institute for Mathematical Sciences.
[3] Research supported by a Premier's Research Excellence Award from the Province of Ontario.

0001-8708/$ – see front matter
doi:10.1016/j.aim.2006.05.003

1. Introduction

In Part I of this series [11] we introduced the concept of second order freeness as the mathematical concept for dealing with the large N limit of fluctuations of $N \times N$-random matrices. Whereas Voiculescu's freeness (of first order) (see [13,18] and [19]) provides the crucial notion behind the leading order of expectations of traces, our second order freeness is intended to describe in a similar way the structure of leading orders of global fluctuations, i.e., of variances of traces. In Part I we showed how fluctuations of Gaussian and Wishart random matrices can be understood from this perspective. Here we give the corresponding treatment for fluctuations of unitary random matrices. Global fluctuations of unitary random matrices have received much attention in the last decade, see, e.g, the survey article of Diaconis [5].

Our main concern will be to understand the relation between unitary random matrices and some other ensemble of random matrices which is independent from the unitary ensemble. This includes in particular the case that the second ensemble consists of constant (i.e., non-random) matrices. A basic result of Voiculescu tells us that on the level of expectations, independence between the ensembles goes over into asymptotic freeness. We will show that this result remains true on the level of fluctuations: *independence between the ensembles implies that we have asymptotic second order freeness between their fluctuations*.

Theorem 1. *Let $\{A_1, \dots, A_s\}_N$ be a sequence of $N \times N$-random matrices which has a second order limit distribution and let $\{B_1, \dots, B_t\}_N$ be a sequence of $\mathcal{U}(N)$-invariant $N \times N$-random matrices which has a second order limit distribution. Furthermore assume that the matrices $\{A_1, \dots, A_s\}_N$ and the matrices $\{B_1, \dots, B_t\}_N$ are independent. Then the sequences $\{A_1, \dots, A_s\}_N$ and $\{B_1, \dots, B_t\}_N$ are asymptotically free of second order.*

Two important consequences of our investigations are the following.

We get a generalization to the case of several independent unitary random matrices of a classical result of Diaconis and Shahshahani [6]. Their one-dimensional case states that, for a unitary random matrix U, the family of traces $\mathrm{Tr}(U^n)$ converge towards a Gaussian family where the covariance between $\mathrm{Tr}(U^m)$ and $\mathrm{Tr}(U^{*n})$ is given by $n \cdot \delta_{mn}$. In the case of several independent unitary random matrices, one has to consider traces in reduced words of these random matrices, and again these converge to a Gaussian family, where the covariance between two such reduced words is now given by the number of cyclic rotations which match one word with the other. This result was also independently derived by Rădulescu [15] in the course of his investigations around Connes's embedding problem.

Theorem 2. *Let $\{U_{(1)}\}_N, \dots, \{U_{(r)}\}_N$ be independent sequences of Haar distributed unitary $N \times N$-random matrices. Then, the collection $\{\mathrm{Tr}(U_{i(1)}^{k(1)} \cdots U_{i(n)}^{k(n)})\}$ of non-normalized traces in cyclically reduced words in these random matrices converges to a Gaussian family of centered random variables whose covariance is given by the number of matchings between the two reduced words*:

$$\lim_{N\to\infty} \mathrm{k}_2\big(\mathrm{Tr}\big(U_{i(1)}^{k(1)} \cdots U_{i(m)}^{k(m)}\big), \mathrm{Tr}\big(U_{j(n)}^{l(n)} \cdots U_{j(1)}^{l(1)}\big)\big)$$
$$= \delta_{mn} \cdot \#\big\{r \in \{1, \dots, n\} \mid i(s) = j(s+r),\ k(s) = -l(s+r)\ \forall s = 1, \dots, n\big\}.$$

We can show that we do *not* have universality of fluctuations in multi-matrix models. For unitarily invariant one-matrix models it was shown by Johansson [7] (compare also [1]) that many random matrix ensembles have the same fluctuations as the ensemble of Gaussian random matrices. A main motivation for our investigations was the expectation that many unitarily invariant models of multi-matrix random ensembles should have the same fluctuations as the ensemble of independent Gaussian random matrices. However, our theory of second order freeness shows that this is not the case.

The paper is organized as follows. In Section 2, we recall all the necessary definitions and results around permutations, unitary random matrices, and second order freeness. We will recall all the relevant notions from Part I, so that our presentation will be self-contained. However, for getting more background information on the concept of second order freeness one should consult [11]. In addition, a diagrammatic investigation of the fluctuations of Wishart matrices is given in [8].

In Section 3, we derive our main result about the asymptotic second order freeness between unitary random matrices and another independent random matrix ensemble. This yields as corollary that independent unitary random matrices are asymptotically free of second order, implying the above mentioned generalization of the result of Diaconis and Shahshahani [6]. Section 4 shows how our results imply the failure of universality of global fluctuations in multi-matrix models.

2. Preliminaries

2.1. The lattice of partitions

For natural numbers $m, n \in \mathbb{N}$ with $m < n$, we denote by $[m, n]$ the interval of natural numbers between m and n, i.e.,

$$[m, n] := \{m, m+1, m+2, \dots, n-1, n\}$$

and $[m] = [1, m]$. For a matrix $A = (a_{ij})_{i,j=1}^{N}$, we denote by Tr the unnormalized and by tr the normalized trace,

$$\operatorname{Tr}(A) := \sum_{i=1}^{N} a_{ii}, \qquad \operatorname{tr}(A) := \frac{1}{N} \operatorname{Tr}(A).$$

We say that $A = \{A_1, \dots, A_k\}$ is a partition of the set $[1, n]$ if the subsets A_i are disjoint, nonempty, and their union is equal to $[1, n]$. We call $A_1, \dots, A_k$ the blocks of the partition A. For a permutation $\pi \in S_n$ we say that a partition A is π-invariant if π leaves invariant each block A_i. Let 1_n denote that partition of $[n]$ with one block and $\mathcal{P}(n)$ denote the partitions of $[n]$. Given positive integers m and n let $1_{m,n}$ be the partition of $[m+n]$ with the two blocks: $[m]$ and $[m+1, m+n]$.

If $A = \{A_1, \dots, A_k\}$ and $B = \{B_1, \dots, B_l\}$ are partitions of the same set, we say that $A \leqslant B$ if for every block A_i there exists some block B_j such that $A_i \subseteq B_j$. For a pair of partitions A, B we denote by $A \vee B$ the smallest partition C such that $A \leqslant C$ and $B \leqslant C$.

If we are considering classical random variables on a probability space, then we denote by E the expectation with respect to the corresponding probability measure and by k_r the corresponding classical cumulants (as multi-linear functionals in r arguments); in particular,

$$\mathrm{k}_1(a) = \mathrm{E}(a) \quad \text{and} \quad \mathrm{k}_2(a_1, a_2) = \mathrm{E}(a_1 a_2) - \mathrm{E}(a_1)\mathrm{E}(a_2).$$

If $a_1, \dots, a_n$ are random variables and $C = \{C_1, \dots, C_k\}$ is in $\mathcal{P}(n)$ we let

$$\mathrm{E}_C(a_1, \dots, a_n) = \prod_{i=1}^{k} \mathrm{E}\biggl(\prod_{j \in C_i} a_j\biggr).$$

On the lattice $\mathcal{P}(n)$ moments to cumulants are related by the Möbius function: Möb. In particular the nth cumulant k_n is given by

$$\mathrm{k}_n(a_1, \dots, a_r) = \sum_{C \in \mathcal{P}(n)} \mathrm{M\ddot{o}b}(C, 1_r)\mathrm{E}_C(a_1, \dots, a_n),$$

where $\mathrm{M\ddot{o}b}(C, 1_n) = (-1)^{k-1}(k-1)!$ and where k is the number of blocks of C. We shall need the following formula for the second cumulant of the product of random variables, see for example [9, 3.2],

$$\mathrm{k}_2(a_1 \cdots a_m, b_1 \cdots b_n) = \sum_{\substack{\tau \in \mathcal{P}(m+n) \\ \tau \vee 1_{m,n} = 1_{m+n}}} \mathrm{k}_\tau(a_1, \dots, a_m, b_1, \dots, b_n). \tag{1}$$

The sum is over all partitions of $[m+n]$ which have at least one block which connects the two blocks of $1_{m,n}$.

2.2. *Permutations*

We will denote the set of permutations on n elements by S_n. We will quite often use the cycle notation for such permutations, i.e., $\pi = (i_1, i_2, \dots, i_r)$ is a cycle which sends i_k to i_{k+1} ($k = 1, \dots, r$), where $i_{r+1} = i_1$.

2.2.1. *Length function*

For a permutation $\pi \in S_n$ we denote by $\#(\pi)$ the number of cycles of π and by $|\pi|$ the minimal number of transpositions needed to write π as a product of transpositions. Note that one has

$$|\pi| + \#(\pi) = n \quad \text{for all } \pi \in S_n.$$

2.2.2. *Non-crossing permutations*

Let us denote by $\gamma_n \in S_n$ the cycle

$$\gamma_n = (1, 2, \dots, n).$$

For all $\pi \in S_n$ one has that

$$n - 1 \leqslant |\pi| + \bigl|\gamma_n \pi^{-1}\bigr|.$$

If we have equality then we call π *non-crossing*, see [2] for the basic properties of non-crossing permutations. Note that this is equivalent to

$$\#(\pi) + \#\big(\gamma_n \pi^{-1}\big) = n + 1.$$

If π is non-crossing, then so are $\gamma_n \pi^{-1}$ and $\pi^{-1}\gamma_n$; the latter is called the (*Kreweras*) *complement* of π.

We will denote the set of non-crossing permutations in S_n by $NC(n)$. Note that such a non-crossing permutation can be identified with a non-crossing partition, by forgetting the order on the cycles. There is exactly one cyclic order on the blocks of a non-crossing partition which makes it into a non-crossing permutation.

2.2.3. Annular non-crossing permutations

Fix $m, n \in \mathbb{N}$ and denote by $\gamma_{m,n}$ the product of the two cycles

$$\gamma_{m,n} = (1, 2, \ldots, m)(m+1, m+2, \ldots, m+n).$$

More generally, we shall denote by $\gamma_{m_1,\ldots,m_k}$ the product of the corresponding k cycles.

We call a $\pi \in S_{m+n}$ *connected* if the pair π and $\gamma_{m,n}$ generates a transitive subgroup in S_{m+n}. A connected permutation $\pi \in S_{m+n}$ always satisfies

$$m + n \leqslant |\pi| + \big|\gamma_{m,n}\pi^{-1}\big|. \tag{2}$$

If π is connected and if we have equality in that equation then we call π *annular non-crossing*. Note that with π also $\gamma_{m,n}\pi^{-1}$ is annular non-crossing. Again, we call the latter the *complement* of π. Of course, all the above notations depend on the pair (m, n); if we want to emphasize this dependency we will also speak about (m, n)-connected permutations and (m, n)-annular non-crossing permutations.

We will denote the set of (m, n)-annular non-crossing permutations by $S_{NC}(m, n)$. Again one can go over to annular non-crossing partitions by forgetting the cyclic orders on cycles; however, in the annular case, the relation between non-crossing permutation and non-crossing partition is not one-to-one. Since we will not use the language of annular partitions in the present paper, this is of no relevance here.

Annular non-crossing permutations and partitions were introduced in [10]; there, many different characterizations—in particular, the one (2) above in terms of the length function—were given.

2.3. A triangle inequality

Let $\{A_1, \ldots, A_k\}$ be a partition of $[n]$. If, for $1 \leqslant i \leqslant k$, π_i is a permutation of the set A_i we denote by $\pi_1 \times \cdots \times \pi_k \in S_n$ the concatenation of these permutations. We say that $\pi = \pi_1 \times \cdots \times \pi_k$ is a cycle decomposition if additionally every factor π_i is a cycle.

Notation 2.1. (1) For $A \in \mathcal{P}(n)$ we put $|A| := n - \#(A)$.

(2) For any $\pi \in S_n$ and any π-invariant $A \in \mathcal{P}(n)$ we put

$$\big|(A, \pi)\big| := 2|A| - |\pi|.$$

Lemma 2.2.

(1) *For all* $A, B \in \mathcal{P}(n)$ *we have*

$$|A \vee B| \leqslant |A| + |B|.$$

(2) *If* π *and* σ *are in* S_n *and* $A, B \in \mathcal{P}(n)$ *are* π *and* σ *invariant respectively, then*

$$\left|(A \vee B, \pi\sigma)\right| \leqslant \left|(A, \pi)\right| + \left|(B, \sigma)\right|.$$

Proof. (1) Each block of B with k points can glue together at most k blocks of A, thereby reducing the number of blocks of A by at most $k-1$. Thus B can reduce by at most $n - \#(B)$ the number of blocks of A. Hence the difference between $\#(A)$ and $\#(A \vee B)$ cannot exceed $n - \#(B)$ and hence

$$\#(A) - \#(A \vee B) \leqslant n - \#(B).$$

This is equivalent to our assertion.

(2) We prove this, for fixed π and σ by induction on $|A| + |B|$. The smallest possible value of $|A| + |B|$ occurs when $|A| = |\pi|$ and $|B| = |\sigma|$. But then we have (since $A \vee B \geqslant \pi\sigma$)

$$2|A \vee B| - |\pi\sigma| \leqslant |A \vee B| \leqslant |A| + |B| \quad \text{(by (1))}$$

which is exactly our assertion for this case. For the induction step note that we have just shown that

$$2|A \vee B| - |\pi\sigma| \leqslant 2|A| - |\pi| + 2|B| - |\sigma|$$

when $|A| = |\pi|$ and $|B| = |\sigma|$. Now one only has to observe that if one increases $|A|$ (or $|B|$) by 1 then $|A \vee B|$ can also increase by at most 1. □

2.4. Haar distributed unitary random matrices and the Weingarten function

In the following we will be interested in the asymptotics of special matrix integrals over the group $\mathcal{U}(N)$ of unitary $N \times N$-matrices. We always equip the compact group $\mathcal{U}(N)$ with its Haar probability measure and accordingly distributed random matrices we shall call *Haar distributed unitary random matrices*. Thus the expectation E over this ensemble is given by integrating with respect to the Haar measure.

The expectation of products of entries of Haar distributed unitary random matrices can be described in terms of a special function on the permutation group. Since such considerations go back to Weingarten [20], Collins [3] calls this function the Weingarten function and denotes it by Wg. We will follow his notation. In the following we just recall the relevant information about this Weingarten function, for more details we refer to [3,4,21].

We use the following definition of the Weingarten function. For $\pi \in S_n$ and $N \geqslant n$ we put

$$\mathrm{Wg}(N, \pi) = \mathrm{E}(U_{11} \cdots U_{nn} \overline{U}_{1\pi(1)} \cdots \overline{U}_{n\pi(n)}),$$

where $U = (U_{ij})_{i,j=1}^{N}$ is an $N \times N$ Haar distributed unitary random matrix. Sometimes we will suppress the dependence on N and just write $\mathrm{Wg}(\pi)$. This $\mathrm{Wg}(N,\pi)$ depends on π only through its conjugacy class. General matrix integrals over the unitary groups can be calculated as follows:

$$\begin{aligned} &\mathrm{E}(U_{i'_1 j'_1} \cdots U_{i'_n j'_n} \overline{U}_{i_1 j_1} \cdots \overline{U}_{i_n j_n}) \\ &\quad = \sum_{\alpha,\beta \in S_n} \delta_{i_1 i'_{\alpha(1)}} \cdots \delta_{i_n i'_{\alpha(n)}} \delta_{j_1 j'_{\beta(1)}} \cdots \delta_{j_n j'_{\beta(n)}} \mathrm{Wg}\big(\beta\alpha^{-1}\big). \end{aligned} \tag{3}$$

The Weingarten function is a quite complicated object, and its full understanding is at the basis of questions around Itzykson–Zuber integrals. For our purposes, only the behaviour of leading orders in N of $\mathrm{Wg}(N,\pi)$ is important. One knows (see, e.g., [3,4]) that the leading order in $1/N$ is given by $|\pi| + n$ and increases in steps of 2.

Let us use the following notation for the first two orders ($\pi \in S(n)$):

$$\mathrm{Wg}(N,\pi) = \mu(\pi) N^{-(|\pi|+n)} + \phi(\pi) N^{-(|\pi|+n+2)} + O\big(N^{-(|\pi|+n+4)}\big).$$

One knows that μ is multiplicative with respect to the cycle decomposition, i.e.,

$$\mu(\pi_1 \times \pi_2) = \mu(\pi_1) \cdot \mu(\pi_2).$$

The important part of the second order information is contained in the leading order of $\mathrm{Wg}(\pi_1 \times \pi_2) - \mathrm{Wg}(\pi_1)\mathrm{Wg}(\pi_2)$, which is given by $\mu_2(\pi_1,\pi_2) N^{-(|\pi_1|+|\pi_2|+m+n+2)}$ for $\pi_1 \in S_m$ and $\pi_2 \in S_n$ and where

$$\mu_2(\pi_1,\pi_2) := \phi(\pi_1 \times \pi_2) - \mu(\pi_1)\phi(\pi_2) - \phi(\pi_1)\mu(\pi_2).$$

Note that we have

$$\mu_2(\pi_1,\pi_2) = \mu_2(\pi_2,\pi_1).$$

Collins [3] has general counting formulas for the calculation of μ and μ_2 (and also higher order analogues); however, a conceptual explanation of μ_2 seems still to be missing. μ is the Möbius function of the lattice of non-crossing partitions (thus determined by Catalan numbers), and this fact is quite well understood by the relation between μ and asymptotic freeness of unitary random matrices. In a similar way, one should get a conceptual understanding of μ_2 by the relation with second order freeness. In the present paper we will not pursue further this direction, but we will come back to it in forthcoming investigations. Here we will not rely on the concrete values of μ or μ_2, but will only use the basic properties mentioned above.

2.5. *Second order freeness*

In [11], we introduced the concept of second order freeness which is intended to capture the structure of the fluctuation functionals for random matrices arising in the limit $N \to \infty$, in the same way as the usual freeness captures the structure of the expectation of the trace in the limit. We recall the relevant notations and definitions.

Definition 2.3. A *second order non-commutative probability space* $(\mathcal{A}, \varphi_1, \varphi_2)$ consists of a unital algebra $\mathcal{A}$, a tracial linear functional

$$\varphi_1 : \mathcal{A} \to \mathbb{C} \quad \text{with} \quad \varphi_1(1) = 1$$

and a bilinear functional

$$\varphi_2 : \mathcal{A} \times \mathcal{A} \to \mathbb{C},$$

which is tracial in both arguments and which satisfies

$$\varphi_2(a, 1) = 0 = \varphi_2(1, b) \quad \text{for all } a, b \in \mathcal{A}.$$

Notation 2.4. Let unital subalgebras $\mathcal{A}_1, \ldots, \mathcal{A}_r \subset \mathcal{A}$ be given.

(1) We say that a tuple $(a_1, \ldots, a_n)$ $(n \geqslant 1)$ of elements from $\mathcal{A}$ is *cyclically alternating* if, for each k, we have an $i(k) \in \{1, \ldots, r\}$ such that $a_k \in \mathcal{A}_{i(k)}$ and, if $n \geqslant 2$, we have $i(k) \neq i(k+1)$ for all $k = 1, \ldots, n$. We count indices in a cyclic way modulo n, i.e., for $k = n$ the above means $i(n) \neq i(1)$. Note that for $n = 1$, we do not impose any condition on neighbours.

(2) We say that a tuple $(a_1, \ldots, a_n)$ of elements from $\mathcal{A}$ is *centered* if we have

$$\varphi_1(a_k) = 0 \quad \text{for all } k = 1, \ldots, n.$$

Definition 2.5. Let $(\mathcal{A}, \varphi_1, \varphi_2)$ be a second order non-commutative probability space. We say that unital subalgebras $\mathcal{A}_1, \ldots, \mathcal{A}_r \subset \mathcal{A}$ are *free with respect to* (φ_1, φ_2) or *free of second order*, if they are free (in the usual sense [19]) with respect to φ_1 and if the following condition for φ_2 is satisfied. Whenever we have, for $n, m \geqslant 1$, tuples $(a_1, \ldots, a_n)$ and $(b_m, \ldots, b_1)$ from $\mathcal{A}$ such that both are centered and cyclically alternating then we have:

(1) If $n \neq m$, then

$$\varphi_2(a_1 \cdots a_n, b_m \cdots b_1) = 0.$$

(2) If $n = m = 1$ and $a \in \mathcal{A}_i$, $b \in \mathcal{A}_j$, with $i \neq j$, then

$$\varphi_2(a, b) = 0.$$

(3) If $n = m \geqslant 2$, then

$$\varphi_2(a_1 \cdots a_n, b_n \cdots b_1) = \sum_{k=0}^{n-1} \varphi_1(a_1 b_{1+k}) \cdot \varphi_1(a_2 b_{2+k}) \cdots \varphi_1(a_n b_{n+k}).$$

For a visualization of this formula, one should think of two concentric circles with the a's on one of them and the b's on the other. However, whereas on one circle we have a clockwise orientation of the points, on the other circle the orientation is counter-clockwise. Thus, in order to match up these points modulo a rotation of the circles, we have to pair the indices as in the sum above.

Recall that in the combinatorial description of freeness [12], the extension of φ_1 to a multiplicative function on non-crossing partitions plays a fundamental role. In the same way, second order freeness will rely on a suitable extension of φ_2.

Notation 2.6. Let $(\mathcal{A}, \varphi_1, \varphi_2)$ be a second order non-commutative probability space. Then we extend the definition of φ_1 and φ_2 as follows:

$$\varphi_1 : \bigcup_{n=1}^{\infty} (S_n \times \mathcal{A}^n) \to \mathbb{C}$$
$$(\pi, a_1, \dots, a_n) \mapsto \varphi_1(\pi)[a_1, \dots, a_n]$$

is, for a cycle $\pi = (i_1, i_2, \dots, i_r)$, given by

$$\varphi_1(\pi)[a_1, \dots, a_n] := \varphi_1(a_{i_1} a_{i_2} a_{i_3} \cdots a_{i_r})$$

and extended to general $\pi \in S_n$ by multiplicativity

$$\varphi_1(\pi_1 \times \pi_2)[a_1, \dots, a_n] = \varphi_1(\pi_1)[a_1, \dots, a_n] \cdot \varphi_1(\pi_2)[a_1, \dots, a_n].$$

In a similar way,

$$\varphi_2 : \bigcup_{m,n=1}^{\infty} \left(S_m \times S_n \times \mathcal{A}^m \times \mathcal{A}^n\right) \to \mathbb{C}$$
$$(\pi_1, \pi_2, a_1, \dots, a_m, b_1, \dots, b_m) \mapsto \varphi_2(\pi_1, \pi_2)[a_1, \dots, a_m; b_1, \dots, b_n]$$

is defined, for two cycles $\pi_1 = (i_1, i_2, \dots, i_p)$ and $\pi_2 = (j_1, j_2, \dots, j_r)$, by

$$\varphi_2(\pi_1, \pi_2)[a_1, \dots, a_m; b_1, \dots, b_n] := \varphi_2(a_{i_1} a_{i_2} \cdots a_{i_p}, b_{j_1} b_{j_2} \cdots b_{j_r})$$

and extended to the general situation by the derivation property

$$\begin{aligned} &\varphi_2(\pi_1 \times \pi_2, \pi_3)[a_1, \dots, a_m; b_1, \dots, b_n] \\ &\quad = \varphi_2(\pi_1, \pi_3)[a_1, \dots, a_m; b_1, \dots, b_n] \cdot \varphi_1(\pi_2)[a_1, \dots, a_m, b_1, \dots, b_n] \\ &\qquad + \varphi_2(\pi_2, \pi_3)[a_1, \dots, a_m; b_1, \dots, b_n] \cdot \varphi_1(\pi_1)[a_1, \dots, a_m, b_1, \dots, b_n] \end{aligned} \tag{4}$$

and

$$\begin{aligned} &\varphi_2(\pi_1, \pi_2 \times \pi_3)[a_1, \dots, a_m; b_1, \dots, b_n] \\ &\quad = \varphi_2(\pi_1, \pi_2)[a_1, \dots, a_m; b_1, \dots, b_n] \cdot \varphi_1(\pi_3)[a_1, \dots, a_m, b_1, \dots, b_n] \\ &\qquad + \varphi_2(\pi_1, \pi_3)[a_1, \dots, a_m; b_1, \dots, b_n] \cdot \varphi_1(\pi_2)[a_1, \dots, a_m, b_1, \dots, b_n]. \end{aligned} \tag{5}$$

Remark 2.7. Let $(A_i)_{i\in I}$ be a family of unital subalgebras of the second order probability space (A,φ_1,φ_2) which are free of second order. Suppose that for each i we have $(B_{i,j})_{j\in K_i}$ a family of unital subalgebras of A_i which are free of second order. By [19, Proposition 2.5.5(iii)] $(B_{i,j})_{j\in\bigcup_i K_i}$ are free of first order. We leave as an exercise for the reader to show that the proof of [19] can be adapted to show that $(B_{i,j})_{j\in\bigcup_i K_i}$ are free of second order.

3. Asymptotic second order freeness for unitary random matrices

Notation 3.1. Suppose $\epsilon:[2l]\to\{-1,1\}$ is such that $\sum_{i=1}^{2l}\epsilon_i=0$. We write $\epsilon^{-1}(1)=\{p_1,p_2,\dots,p_l\}$ and $\epsilon^{-1}(-1)=\{q_1,q_2,\dots,q_l\}$, with $p_1<p_2<\dots<p_l$ and $q_1<q_2<\dots<q_l$. Let $S_{2l}^{(\epsilon)}$ be the permutations π in S_{2l} such that π takes $\{p_1,\dots,p_l\}$ onto $\{q_1,\dots,q_l\}$ and vice versa. Given a π in $S_{2l}^{(\epsilon)}$ we may extract a pair of permutations α_π and β_π in S_l from the equations

$$\pi(p_{\alpha_\pi(k)})=q_k\quad\text{and}\quad\pi(q_k)=p_{\beta_\pi(k)}\tag{6}$$

and conversely: $(\alpha,\beta)\mapsto\pi_{\alpha,\beta}$. Thus we have a bijection of sets between $S_{2l}^{(\epsilon)}$ and $S_l\times S_l$.

Given $\pi\in S_{2l}^{(\epsilon)}$ we let $\tilde\pi\in S_l$ be defined by

$$\pi^2(p_k)=p_{\tilde\pi(k)}.$$

Note that $\tilde\pi_{\alpha,\beta}=\beta\alpha^{-1}$.

Also we have

$$\#(\pi)=\#(\tilde\pi),$$

and thus

$$|\pi|=|\tilde\pi|+l.$$

Lemma 3.2. *Fix $l\in\mathbb{N}$ and $\gamma\in S_{2l}$. Let, for $N\in\mathbb{N}$, U be a Haar distributed unitary $N\times N$-random matrix. Let $\epsilon:[2l]\to\{-1,1\}$ such that $\sum_{i=1}^{2l}\epsilon_i=0$. Then we have for all $1\leqslant r_1,\dots,r_{2l},s_1,\dots,s_{2l}\leqslant N$ that*

$$\mathrm{E}\big(U_{r_1,s_{\gamma(1)}}^{\epsilon_1}\cdots U_{r_{2l},s_{\gamma(2l)}}^{\epsilon_{2l}}\big)=\sum_{\pi\in S_{2l}^{(\epsilon)}}\prod_{k=1}^{2l}\delta_{r_k,s_{\gamma(\pi(k))}}\mathrm{Wg}(N,\tilde\pi).\tag{7}$$

Proof. Let i_k,i'_k,j_k,j'_k $(1\leqslant k\leqslant l)$ be such that

$$\mathrm{E}\big(U_{r_1,s_{\gamma(1)}}^{\epsilon_1}\cdots U_{r_{2l},s_{\gamma(2l)}}^{\epsilon_{2l}}\big)=\mathrm{E}\big(U_{i'_1,j'_1}\cdots U_{i'_l,j'_l}U_{j_1,i_1}^{-1}\cdots U_{j_l,i_l}^{-1}\big),$$

i.e., let $\epsilon^{-1}(1)=\{p_1,\dots,p_l\}$ with $p_1<\dots<p_l$ and $\epsilon^{-1}(-1)=\{q_1,\dots,q_l\}$ with $q_1<\dots<q_l$ and $i'_k=r_{p_k}$, $j'_k=s_{\gamma(p_k)}$, $i_k=s_{\gamma(q_k)}$, and $j_k=r_{q_k}$.

Now suppose that α and β in S_l and $\pi\in S_{2l}^{(\epsilon)}$ is as in Eq. (6) above.

Thus we have

$$i_k=s_{\gamma(q_k)}=s_{\gamma(\pi(p_{\alpha(k)}))},\quad\text{and}\quad i'_{\alpha(k)}=r_{p_{\alpha(k)}},$$

and

$$j'_{\beta(k)} = s_{\gamma(p_{\beta(k)})} = s_{\gamma(\pi(q_k))}, \quad \text{and} \quad j_k = r_{q_k}$$

which shows that

$$i_k = i'_{\alpha(k)} \quad \Longleftrightarrow \quad r_{p_{\alpha(k)}} = s_{\gamma(\pi(p_{\alpha(k)}))}$$

and

$$j_k = j'_{\beta(k)} \quad \Longleftrightarrow \quad r_{q_k} = s_{\gamma(\pi(q_k))}.$$

Thus

$$\prod_{k=1}^{l} \delta_{i_k, i'_{\alpha(k)}} \delta_{j_k, j'_{\beta(k)}} = \prod_{k=1}^{2l} \delta_{r_k, s_{\gamma(\pi(k))}}.$$

Hence

$$\begin{aligned} \mathrm{E}\big(U^{\epsilon_1}_{r_1, s_{\gamma(1)}} \cdots U^{\epsilon_{2l}}_{r_{2l}, s_{\gamma(2l)}}\big) &= \mathrm{E}\big(U_{i'_1, j'_1} \cdots U_{i'_l, j'_l} U^{-1}_{j_1, i_1} \cdots U^{-1}_{j_l, i_l}\big) \\ &= \sum_{\alpha, \beta \in S_n} \delta_{i_1 i'_{\alpha(1)}} \cdots \delta_{i_n i'_{\alpha(n)}} \delta_{j_1 j'_{\beta(1)}} \cdots \delta_{j_n j'_{\beta(n)}} \mathrm{Wg}\big(\beta\alpha^{-1}\big) \\ &= \sum_{\pi \in S^{(\epsilon)}_{2l}} \prod_{k=1}^{2l} \delta_{r_k, s_{\gamma(\pi(k))}} \mathrm{Wg}(\tilde{\pi}). \qquad \square \end{aligned}$$

We can now address the question how to calculate expectations of products of traces of our matrices. The following result is exact for each N; later on we will look on its asymptotic version.

Note that the notation $\mathrm{Tr}_\pi(D_1, \dots, D_n)$ for $\pi \in S_n$ is defined in the usual multiplicative way, as was done in Notation 2.6 for φ_1.

We shall need the following standard lemma. For $D_1, \dots, D_p \in M_N(\mathcal{A})$ let the entries of D_i be $(D^{(i)}_{r,s})$.

Lemma 3.3. *Let $\pi \in S_n$ and $D_1, \dots, D_n \in M_N(\mathcal{A})$. Then*

$$\mathrm{Tr}_\pi(D_1, D_2, \dots, D_n) = \sum_{j_1, j_2, \dots, j_n} D^{(1)}_{j_1, j_{\pi(1)}} D^{(2)}_{j_2, j_{\pi(2)}} \cdots D^{(n)}_{j_n, j_{\pi(n)}}.$$

Given $m_1, \dots, m_k$, let $\gamma_{m_1, \dots, m_k}$ be the permutation of $[m_1 + \cdots + m_k]$ with k cycles where the ith cycle is $(m_1 + \cdots + m_{i-1} + 1, \dots, m_1 + \cdots + m_i)$.

Proposition 3.4. *Fix $m_1, \dots, m_k \in \mathbb{N}$ such that $m_1 + \cdots + m_k = 2l$ is even. Let, for fixed $N \in \mathbb{N}$, U be a Haar distributed unitary $N \times N$-random matrix and $D_1, \dots, D_{2l}$ be $N \times N$-random matrices which are independent from U. Let $\epsilon : [2l] \to \{-1, 1\}$ with $\sum_{i=1}^{2l} \epsilon_i = 0$. Put $\gamma = \gamma_{m_1, \dots, m_k}$. Then*

$$\begin{aligned}
&\mathrm{E}\big(\mathrm{Tr}\big(D_1 U^{\epsilon_1} \cdots D_{m_1} U^{\epsilon_{m_1}}\big) \\
&\quad \times \mathrm{Tr}\big(D_{m_1+1} U^{\epsilon_{m_1+1}} \cdots D_{m_1+m_2} U^{\epsilon_{m_1+m_2}}\big) \times \cdots \\
&\quad \times \mathrm{Tr}\big(D_{m_1+\cdots+m_{k-1}+1} U^{\epsilon_{m_1+\cdots+m_{k-1}+1}} \cdots D_{m_1+\cdots+m_k} U^{\epsilon_{m_1+\cdots+m_k}}\big)\big) \\
&\qquad = \sum_{\pi \in S_{2l}^{(\epsilon)}} \mathrm{Wg}(N, \tilde{\pi}) \cdot \mathrm{E}\big(\mathrm{Tr}_{\gamma\pi^{-1}}(D_1, \ldots, D_{2l})\big). \qquad (8)
\end{aligned}$$

Proof. Summations over r's and s's in the following formulas are from 1 to N.

$$\begin{aligned}
&\mathrm{E}\big(\mathrm{Tr}\big(D_1 U^{\epsilon_1} \cdots D_{m_1} U^{\epsilon_{m_1}}\big) \mathrm{Tr}\big(D_{m_1+1} U^{\epsilon_{m_1+1}} \cdots D_{m_1+m_2} U^{\epsilon_{m_1+m_2}}\big) \times \cdots \\
&\quad \times \mathrm{Tr}\big(D_{m_1+\cdots+m_{k-1}+1} U^{\epsilon_{m_1+\cdots+m_{k-1}+1}} \cdots D_{m_1+\cdots+m_k} U^{\epsilon_{m_1+\cdots+m_k}}\big)\big) \\
&\quad = \sum_{\substack{r_1,\ldots,r_{2l} \\ s_1,\ldots,s_{2l}}} \mathrm{E}\big(U^{\epsilon_1}_{s_1, r_{\gamma(1)}} \cdots U^{\epsilon_{2l}}_{s_{2l}, r_{\gamma(2l)}}\big) \cdot \mathrm{E}\big(D^{(1)}_{r_1, s_1} \cdots D^{(2l)}_{r_{2l}, s_{2l}}\big) \\
&\quad = \sum_{\substack{r_1,\ldots,r_{2l} \\ s_1,\ldots,s_{2l}}} \sum_{\pi \in S_{2l}^{(\epsilon)}} \prod_{k=1}^{2l} \delta_{s_k, r_{\gamma(\pi(k))}} \mathrm{Wg}(\tilde{\pi}) \cdot \mathrm{E}\big(D^{(1)}_{r_1, s_1} \cdots D^{(2l)}_{r_{2l}, s_{2l}}\big) \\
&\quad = \sum_{\pi \in S_{2l}^{(\epsilon)}} \mathrm{Wg}(\tilde{\pi}) \sum_{\substack{r_1,\ldots,r_{2l} \\ s_1,\ldots,s_{2l}}} \prod_{k=1}^{2l} \delta_{s_k, r_{\gamma(\pi(k))}} \cdot \mathrm{E}\big(D^{(1)}_{r_1, s_1} \cdots D^{(2l)}_{r_{2l}, s_{2l}}\big) \\
&\quad = \sum_{\pi \in S_{2l}^{(\epsilon)}} \mathrm{Wg}(\tilde{\pi}) \sum_{r_1,\ldots,r_{2l}} \mathrm{E}\big(D^{(1)}_{r_1, r_{\gamma(\pi(1))}} \cdots D^{(2l)}_{r_{2l}, r_{\gamma(\pi(2l))}}\big) \\
&\quad = \sum_{\pi \in S_{2l}^{(\epsilon)}} \mathrm{Wg}(\tilde{\pi}) \mathrm{E}\big(\mathrm{Tr}_{\gamma\pi}(D_1, \ldots, D_{2l})\big) \\
&\quad = \sum_{\pi \in S_{2l}^{(\epsilon)}} \mathrm{Wg}(\tilde{\pi}) \mathrm{E}\big(\mathrm{Tr}_{\gamma\pi^{-1}}(D_1, \ldots, D_{2l})\big).
\end{aligned}$$

In the last equality we used that $\mathrm{Wg}(\tilde{\pi})$ depends only on the conjugacy class of π. $\square$

Motivated by the result of Voiculescu [16,17] that Haar distributed unitary random matrices and constant matrices are asymptotically free, we want to investigate now the corresponding question for second order freeness. It will turn out that one can replace the constant matrices by another ensemble of random matrices, as long as those are independent from the unitary random matrices. Of course, we have to assume that the second ensemble has some asymptotic limit distribution. This is formalized in the following definition. Note that we make a quite strong requirement on the vanishing of the higher order cumulants. This is however in accordance with the observation that in many cases the unnormalized traces converge to Gaussian random variables. Of course, if we have a non-probabilistic ensemble of constant matrices, then the only requirement is the convergence of k_1; all other cumulants are automatically zero.

Definition 3.5. (1) Let $\{A_1, \dots, A_s\}_N$ be a sequence of $N \times N$-random matrices. We say that they have a *second order limit distribution* if there exists a second order non-commutative probability space $(\mathcal{A}, \varphi_1, \varphi_2)$ and $a_1, \dots, a_s \in \mathcal{A}$ such that for all polynomials $p_1, p_2, \dots$ in s non-commuting indeterminates we have

$$\lim_{N\to\infty} \mathrm{k}_1\big(\mathrm{tr}\big(p_1(A_1, \dots, A_s)\big)\big) = \varphi_1\big(p_1(a_1, \dots, a_s)\big), \tag{9}$$

$$\lim_{N\to\infty} \mathrm{k}_2\big(\mathrm{Tr}\big(p_1(A_1, \dots, A_s)\big), \mathrm{Tr}\big(p_2(A_1, \dots, A_s)\big)\big)$$
$$= \varphi_2\big(p_1(a_1, \dots, a_s), p_2(a_1, \dots, a_s)\big), \tag{10}$$

and, for $r \geqslant 3$,

$$\lim_{N\to\infty} \mathrm{k}_r\big(\mathrm{Tr}\big(p_1(A_1, \dots, A_s)\big), \dots, \mathrm{Tr}\big(p_r(A_1, \dots, A_s)\big)\big) = 0. \tag{11}$$

(2) We say that two sequences of $N \times N$-random matrices, $\{A_1, \dots, A_s\}_N$ and $\{B_1, \dots, B_t\}_N$, are *asymptotically free of second order* if the sequence $\{A_1, \dots, A_s, B_1, \dots, B_t\}_N$ has a second order limit distribution, given by $(\mathcal{A}, \varphi_1, \varphi_2)$ and $a_1, \dots, a_s, b_1, \dots, b_t \in \mathcal{A}$, and if the unital algebras

$$\mathcal{A}_1 := \mathrm{alg}(1, a_1, \dots, a_s) \quad \text{and} \quad \mathcal{A}_2 := \mathrm{alg}(1, b_1, \dots, b_t)$$

are free with respect to (φ_1, φ_2).

Notation 3.6. Fix $m, n \in \mathbb{N}$ and let $\epsilon : [1, m+n] \to \{-1, +1\}$. We defined S_{m+n}^{ϵ} in Notation 3.1, for the case where $\sum_{k=1}^{m+n} \epsilon(k) = 0$, as those permutations in S_{m+n} for which ϵ alternates cyclically between -1 and $+1$ on all cycles. Note that this definition also makes sense in the case where the sum of the ϵ's is not equal to zero, then we just have $S_{m+n}^{\epsilon} = \emptyset$. Let ϵ_1 and ϵ_2 be the restrictions of ϵ to $[1, m]$ and to $[m+1, m+n]$, respectively. Then we put

$$S_{NC}^{(\epsilon)}(m, n) := S_{m+n}^{\epsilon} \cap S_{NC}(m, n)$$

and

$$NC^{(\epsilon_1)}(m) := S_m^{(\epsilon_1)} \cap NC(m), \qquad NC^{(\epsilon_2)}(n) := S_n^{(\epsilon_2)} \cap NC(n).$$

Theorem 3.7. *Let $\{U\}_N$ be a sequence of Haar distributed unitary $N \times N$-random matrices and $\{A_1, \dots, A_s\}_N$ a sequence of $N \times N$-random matrices which has a second order limit distribution, given by $(\mathcal{A}, \varphi_1, \varphi_2)$ and $a_1, \dots, a_s \in \mathcal{A}$. Furthermore, assume that $\{U\}_N$ and $\{A_1, \dots, A_s\}_N$ are independent. Fix now $m, n \in \mathbb{N}$ and consider polynomials $p_1, \dots, p_{m+n}$ in s non-commuting indeterminates. If we put (for $i = 1, \dots, m+n$)*

$$D_i := p_i(A_1, \dots, A_s) \quad \text{and} \quad d_i := p_i(a_1, \dots, a_s),$$

then we have for all $\epsilon(1), \dots, \epsilon(m+n) \in \{-1, +1\}$ that

$$\lim_{N\to\infty} \mathrm{k}_2\big(\mathrm{Tr}\big(D_1U^{\epsilon_1}\cdots D_mU^{\epsilon_m}\big), \mathrm{Tr}\big(D_{m+1}U^{\epsilon_{m+1}}\cdots D_{m+n}U^{\epsilon_{m+n}}\big)\big)$$
$$= \sum_{\pi\in S_{NC}^{(\epsilon)}(m,n)} \mu(\tilde{\pi})\cdot\varphi_1\big(\gamma_{m,n}\pi^{-1}\big)[d_1,\ldots,d_{m+n}]$$
$$+ \sum_{\substack{\pi_1\in NC^{(\epsilon_1)}(m)\\ \pi_2\in NC^{(\epsilon_2)}(n)}} \big(\mu_2(\tilde{\pi}_1,\tilde{\pi}_2)\cdot\varphi_1\big(\gamma_m\pi_1^{-1}\times\gamma_n\pi_2^{-1}\big)[d_1,\ldots,d_{m+n}]$$
$$+ \mu(\tilde{\pi}_1\times\tilde{\pi}_2)\cdot\varphi_2\big(\gamma_m\pi_1^{-1},\gamma_n\pi_2^{-1}\big)[d_1,\ldots,d_{m+n}]\big). \tag{12}$$

Note that in the case where the sum of the ϵ's is different from zero this just states that the limit of k_2 vanishes.

Proof. For notational convenience, we will sometimes write $m+n=2l$ in the following, and also use $\gamma := \gamma_{m,n}$.

We have

$$\mathrm{k}_2\big(\mathrm{Tr}\big(D_1U^{\epsilon_1}\cdots D_mU^{\epsilon_m}\big), \mathrm{Tr}\big(D_{m+1}U^{\epsilon_{m+1}}\cdots D_{2l}U^{\epsilon_{2l}}\big)\big)$$
$$= \mathrm{E}\big(\mathrm{Tr}\big(D_1U^{\epsilon_1}\cdots D_mU^{\epsilon_m}\big)\mathrm{Tr}\big(D_{m+1}U^{\epsilon_{m+1}}\cdots D_{2l}U^{\epsilon_{2l}}\big)\big)$$
$$- \mathrm{E}\big(\mathrm{Tr}\big(D_1U^{\epsilon_1}\cdots D_mU^{\epsilon_m}\big)\big)\cdot\mathrm{E}\big(\mathrm{Tr}\big(D_{m+1}U^{\epsilon_{m+1}}\cdots D_{2l}U^{\epsilon_{2l}}\big)\big)$$
$$= \sum_{\pi\in S_{2l}^{(\epsilon)}} \mathrm{Wg}(\tilde{\pi})\cdot\mathrm{E}\big(\mathrm{Tr}_{\gamma\pi^{-1}}(D_1,\ldots,D_{2l})\big)$$
$$- \sum_{\substack{\pi_1\in S_m^{(\epsilon_1)}\\ \pi_2\in S_n^{(\epsilon_2)}}} \mathrm{Wg}(\tilde{\pi}_1)\mathrm{Wg}(\tilde{\pi}_2)\cdot\mathrm{E}\big(\mathrm{Tr}_{\gamma_m\pi_1^{-1}}(D_1,\ldots,D_m)\big)\cdot\mathrm{E}\big(\mathrm{Tr}_{\gamma_n\pi_2^{-1}}(D_{m+1},\ldots,D_{2l})\big)$$
$$= \sum_{\substack{\pi\in S_{2l}^{(\epsilon)}\\ \pi \text{ connected}}} \mathrm{Wg}(\tilde{\pi})\cdot\mathrm{E}\big(\mathrm{Tr}_{\gamma\pi^{-1}}(D_1,\ldots,D_{2l})\big)$$
$$+ \sum_{\substack{\pi_1\in S_m^{(\epsilon_1)}\\ \pi_2\in S_n^{(\epsilon_2)}}} \big(\mathrm{Wg}(\tilde{\pi}_1\times\tilde{\pi}_2)\cdot\mathrm{E}\big(\mathrm{Tr}_{\gamma_m\pi_1^{-1}\times\gamma_n\pi_2^{-1}}(D_1,\ldots,D_{2l})\big)$$
$$- \mathrm{Wg}(\tilde{\pi}_1)\mathrm{Wg}(\tilde{\pi}_2)\mathrm{E}\big(\mathrm{Tr}_{\gamma_m\pi_1^{-1}}(D_1,\ldots,D_m)\big)\cdot\mathrm{E}\big(\mathrm{Tr}_{\gamma_n\pi_2^{-1}}(D_{m+1},\ldots,D_{2l})\big)\big).$$

Note that if either m or n is odd then the last two terms are zero, which is consistent with Eq. (12), as in this case $NC^{(\epsilon_1)}(m)$ and $NC^{(\epsilon_2)}(n)$ are empty. So for the remainder of the proof we shall assume that m and n are even.

The leading order in the first summand for a connected π is given by

$$\mu(\tilde{\pi})N^{-(|\tilde{\pi}|+(m+n)/2)}\cdot N^{\#(\gamma\pi^{-1})}\cdot\mathrm{E}\big(\mathrm{tr}_{\gamma\pi^{-1}}(D_1,\ldots,D_{m+n})\big)$$
$$= N^{m+n-|\pi|-|\gamma\pi^{-1}|}\cdot\mu(\tilde{\pi})\cdot\mathrm{E}\big(\mathrm{tr}_{\gamma\pi^{-1}}(D_1,\ldots,D_{m+n})\big).$$

Recall that, for a connected π, we always have

$$m+n-|\pi|-\left|\gamma\pi^{-1}\right|\leqslant 0,$$

and equality is exactly achieved in the case where π is annular non-crossing. Thus, in the limit $N\to\infty$ the first sum gives the contribution

$$\sum_{\pi\in S_{NC}^{(\epsilon)}(m,n)}\mu(\tilde{\pi})\cdot\varphi_1\big(\gamma\pi^{-1}\big)[d_1,\ldots,d_{m+n}].$$

We can rewrite the second sum as

$$\begin{aligned}&\sum_{\substack{\pi_1\in S_m^{(\epsilon_1)}\\ \pi_2\in S_n^{(\epsilon_2)}}}\big\{\mathrm{Wg}(\tilde{\pi}_1\times\tilde{\pi}_2)-\mathrm{Wg}(\tilde{\pi}_1)\mathrm{Wg}(\tilde{\pi}_2)\big\}\mathrm{E}\big(\mathrm{Tr}_{\gamma_m\pi_1^{-1}\times\gamma_n\pi_2^{-1}}(D_1,\ldots,D_{m+n})\big)\\ &\quad+\mathrm{Wg}(\tilde{\pi}_1)\mathrm{Wg}(\tilde{\pi}_2)\big\{\mathrm{E}\big(\mathrm{Tr}_{\gamma_m\pi_1^{-1}\times\gamma_n\pi_2^{-1}}(D_1,\ldots,D_{m+n})\big)\\ &\quad-\mathrm{E}\big(\mathrm{Tr}_{\gamma_m\pi_1^{-1}}(D_1,\ldots,D_m)\big)\cdot\mathrm{E}\big(\mathrm{Tr}_{\gamma_n\pi_2^{-1}}(D_{m+1},\ldots,D_{m+n})\big)\big\}.\end{aligned}\tag{13}$$

For a disconnected $\pi_1\times\pi_2$ the leading orders in N of all relevant terms are given as follows: $\mathrm{Wg}(\tilde{\pi}_1\times\tilde{\pi}_2)$ and $\mathrm{Wg}(\tilde{\pi}_1)\mathrm{Wg}(\tilde{\pi}_2)$ both have leading order (note that μ is multiplicative)

$$\mu(\tilde{\pi}_1)\mu(\tilde{\pi}_2)N^{-(m+n)+\#(\pi_1)+\#(\pi_2)};$$

$\mathrm{E}(\mathrm{Tr}_{\gamma_m\pi_1^{-1}}(D_1,\ldots,D_m))\cdot\mathrm{E}(\mathrm{Tr}_{\gamma_n\pi_2^{-1}}(D_{m+1},\ldots,D_{m+n}))$ and $\mathrm{E}(\mathrm{Tr}_{\gamma_m\pi_1^{-1}\times\gamma_n\pi_2^{-1}}(D_1,\ldots,D_{m+n}))$ are both asymptotic to

$$\varphi_1\big(\gamma_m\pi_1^{-1}\times\gamma_n\pi_2^{-1}\big)[d_1,\ldots,d_{m+n}]N^{\#(\gamma_m\pi_1^{-1})+\#(\gamma_n\pi_2^{-1})};$$

$\mathrm{Wg}(\tilde{\pi}_1\times\tilde{\pi}_2)-\mathrm{Wg}(\tilde{\pi}_1)\mathrm{Wg}(\tilde{\pi}_2)$ has leading order

$$\mu_2(\tilde{\pi}_1,\tilde{\pi}_2)\cdot N^{-(m+n)+\#(\pi_1)+\#(\pi_2)-2}.$$

Now

$$\begin{aligned}&-(m+n)+\#(\pi_1)+\#(\pi_2)-2+\#\big(\gamma_m\pi_1^{-1}\big)+\#\big(\gamma_n\pi_2^{-1}\big)\\ &\quad=-\big(m+1-\#(\pi_1)-\#\big(\gamma_m\pi_1^{-1}\big)\big)-\big(n+1-\#(\pi_2)-\#\big(\gamma_n\pi_2^{-1}\big)\big)\leqslant 0\end{aligned}$$

with equality only if both $\pi_1\in NC^{(\epsilon_1)}(m)$ and $\pi_2\in NC^{(\epsilon_2)}(n)$.

Thus

$$\begin{aligned}&\lim_N\big\{\mathrm{Wg}(\tilde{\pi}_1\times\tilde{\pi}_2)-\mathrm{Wg}(\tilde{\pi}_1)\mathrm{Wg}(\tilde{\pi}_2)\big\}\mathrm{E}\big(\mathrm{Tr}_{\gamma_{m,n}(\pi_1\times\pi_2)^{-1}}(D_1,\ldots,D_{m+n})\big)\\ &\quad=\begin{cases}\mu_2(\tilde{\pi}_1,\tilde{\pi}_2)\varphi_1(\gamma_m\pi_1^{-1}\times\gamma_n\pi_2^{-1})[d_1,\ldots,d_{m+n}] & \left\{\begin{array}{l}\pi_1\in NC^{(\epsilon_1)}(m)\\ \text{and}\\ \pi_2\in NC^{(\epsilon_2)}(n),\end{array}\right.\\ 0 & \text{otherwise.}\end{cases}\end{aligned}$$

To deal with the second term of the second sum (13) we will use the following notation. Let the cycles of $\gamma_m\pi_1^{-1}$ be $c_1\cdots c_r$ and the cycles of $\gamma_n\pi_2^{-1}$ be $c_{r+1}\cdots c_{r+s}$. Let $a_i=\mathrm{Tr}_{c_i}(D_1,\ldots,D_m)$ for $1\leqslant i\leqslant r$ and $b_j=\mathrm{Tr}_{c_{r+j}}(D_{m+1},\ldots,D_{m+n})$ for $1\leqslant j\leqslant s$. Then

$$\begin{aligned}&\mathrm{k}_2\big(\mathrm{Tr}_{\gamma_m\pi_1^{-1}}(D_1,\ldots,D_m),\mathrm{Tr}_{\gamma_n\pi_2^{-1}}(D_{m+1},\ldots,D_{m+n})\big)\\&\quad=\mathrm{k}_2(a_1\cdots a_r,b_1\cdots b_s).\end{aligned}$$

So let us find for which $\pi_1\in S_m^{(\epsilon_1)}$, $\pi_2\in S_n^{(\epsilon_2)}$, and $\tau\in\mathcal{P}(r+s)$ we have a non-zero limit of

$$\mathrm{Wg}(\tilde{\pi}_1)\mathrm{Wg}(\tilde{\pi}_2)\mathrm{k}_\tau(a_1,\ldots,a_r,b_1,\ldots,b_s).\tag{14}$$

As noted above the order of $\mathrm{Wg}(\tilde{\pi}_1)\mathrm{Wg}(\tilde{\pi}_2)$ is $N^{-(m+n)+\#(\pi_1)+\#(\pi_2)}$. By Eq. (1) and the definition of a second order limit distribution, $\mathrm{k}_\tau(a_1,\ldots,a_r,b_1,\ldots,b_s)$ is $O(N^c)$ where c is the number of singletons of τ. Thus the order of (14) is

$$\begin{aligned}&-(m+n)+\#(\pi_1)+\#(\pi_2)+c\\&\quad=-\big(m+1-\#(\pi_1)-\#\big(\gamma_m\pi_1^{-1}\big)\big)-\big(n+1-\#(\pi_2)-\#\big(\gamma_n\pi_2^{-1}\big)\big)+c+2-(r+s).\end{aligned}$$

Hence (14) will vanish unless three conditions are satisfied: we must have that π_1 and π_2 are non-crossing and $c=r+s-2$, i.e., τ has one pair and the rest of its blocks are singletons.

Thus

$$\begin{aligned}&\lim_N\sum_{\pi_1\in S_m^{(\epsilon_1)},\pi_2\in S_n^{(\epsilon_2)}}\mathrm{Wg}(\tilde{\pi}_1)\mathrm{Wg}(\tilde{\pi}_2)\mathrm{k}_2\big(\mathrm{Tr}_{\gamma_m\pi_1^{-1}}(D_1,\ldots,D_m),\mathrm{Tr}_{\gamma_n\pi_2^{-1}}(D_{m+1},\ldots,D_{2l})\big)\\&\quad=\sum_{\substack{\pi_1\in NC^{(\epsilon_1)}(m)\\ \pi_2\in NC^{(\epsilon_2)}(n)}}\mu(\tilde{\pi}_1)\mu(\tilde{\pi}_2)\lim_N N^{r+s-2}\sum_{\tau\in\mathcal{P}(r+s)}\mathrm{k}_\tau(a_1,\ldots,a_r,b_1,\ldots,b_s),\end{aligned}$$

where the τ's in the sum have one pair and the remainder are singletons and $\tau\vee 1_{r,s}=1_{r+s}$.

So the remainder of the proof is to show that

$$\begin{aligned}&\lim_N N^{r+s-2}\sum_{\tau\in\mathcal{P}(r+s)}\mathrm{k}_\tau(a_1,\ldots,a_r,b_1,\ldots,b_s)\\&\quad=\varphi_2\big(\gamma_m\pi_1^{-1},\gamma_n\pi_2^{-1}\big)[d_1,\ldots,d_{m+n}],\end{aligned}$$

where the sum runs over τ's as above.

Let $\tau\in\mathcal{P}(r+s)$ be as above with pair (i,j) where $1\leqslant i\leqslant r$ and $1\leqslant j\leqslant s$ and all other blocks singletons. Then

$$\begin{aligned}&\mathrm{k}_\tau(a_1,\ldots,a_r,b_1,\ldots,b_s)\\&\quad=\mathrm{k}_1(a_1)\cdots\widehat{\mathrm{k}_1(a_i)}\cdots\mathrm{k}_1(a_r)\mathrm{k}_1(b_1)\cdots\widehat{\mathrm{k}_1(b_j)}\cdots\mathrm{k}_1(b_s)\mathrm{k}_2(a_i,b_j),\end{aligned}$$

where the hatted elements are deleted. So after multiplying by N^{r+s-2} and taking a limit we get (after omitting the arguments $d_1, \dots, d_{m+n}$ which are the same for each factor)

$$\varphi_1(c_1)\cdots\widehat{\varphi_1(c_i)}\cdots\varphi_1(c_r)\varphi_1(c_{r+1})\cdots\widehat{\varphi_1(c_{r+j})}\cdots\varphi_1(c_{r+s})\varphi_2(c_i, c_{r+j}).$$

Now summing over all τ, which is equivalent to summing over all i and j, we get via the derivation property of φ_2 (see Eqs. (4) and (5))

$$\varphi_2\big(\gamma_m\pi_1^{-1}, \gamma_n\pi_2^{-1}\big)[d_1, \dots, d_{m+n}]$$

as required. $\square$

Remark 3.8. When all the D's are equal to 1, Eq. (12) implies the following well-known result of Diaconis and Shahshahani [6]: for integers r and s

$$\lim_N \mathrm{k}_2\big(\mathrm{Tr}\big(U^r\big), \mathrm{Tr}\big(U^s\big)\big) = \begin{cases} 0 & r \neq -s, \\ |r| & r = -s. \end{cases} \tag{15}$$

Indeed let $m = |r|$ and for $1 \leqslant i \leqslant m$, let $\epsilon_i = \mathrm{sgn}(r)$, where $\mathrm{sgn}(r)$ denotes the sign of r; let $n = |s|$ and for $m+1 \leqslant i \leqslant m+n$, let $\epsilon_i = \mathrm{sgn}(s)$. Then $\epsilon_1 + \dots + \epsilon_{m+n} = r + s$. So if $r + s \neq 0$ then Eq. (12) says that

$$\lim_n \mathrm{k}_2\big(\mathrm{Tr}\big(U^r\big), \mathrm{Tr}\big(U^s\big)\big) = 0.$$

Suppose that $r + s = 0$. The second term on the right-hand side of (12) is zero since both $NC^{(\epsilon_1)}(m)$ and $NC^{(\epsilon_2)}(n)$ are empty. For the first term in (12), note that the only elements of $S_{NC}^{(\epsilon)}(m, n)$ which connect in this alternating way are pairings, where each block must contain one U and one U^*. This forces m and n to be equal. In that case, we have the freedom of pairing the first U with any of the n U^*'s. After this choice is made, the rest is determined. Thus there are n possibilities for such pairings. Since $\mu(\tilde{\pi})$ is always 1 for a pairing we get the claimed formula.

Let $\epsilon : [2l] \to \{-1, 1\}$ be such that $\sum_{i=1}^{2l} \epsilon_i = 0$. For $\pi \in S_{2l}^{(\epsilon)}$ let $\tilde{\pi} \in S_l$ be as in 3.1. A π-invariant partition A of $[2l]$ gives $\tilde{A}$, a $\tilde{\pi}$-invariant partition of $[l]$ as follows. For each block V of A let $\tilde{V} = \{k \mid p_k \in V\}$, where we have used the notation of 3.1. Also each $\tilde{\pi}$-invariant partition of $[l]$ comes from a unique π-invariant partition of $[2l]$.

Let Möb be the Möbius function on the partially ordered set of partitions of $[l]$ ordered by inclusion. Let $\pi \in S_l$ and A be a π-invariant partition of $[l]$. In [3, §2.3] Collins denotes the relative cumulant by $C_{\Pi_\pi, A}(\pi, N)$, which we will denote by $C_{\pi,A}$. In our notation

$$C_{\pi,A} = \sum_{\substack{C \in [\pi, A] \\ C = \{V_1, \dots, V_k\}}} \mathrm{Möb}(C, A)\mathrm{Wg}(\pi|_{V_1})\cdots\mathrm{Wg}(\pi|_{V_k}),$$

where $\pi|_{V_i}$ denotes the restriction of π to the invariant subset V_i and where necessary we have identified π with the partition given by its cycles. Conversely given $A = \{V_1, \dots, V_k\}$

a π-invariant partition of $[l]$ we write $\mathrm{Wg}_A(\pi)$ for $\mathrm{Wg}(\pi|_{V_1})\cdots\mathrm{Wg}(\pi|_{V_k})$. Then by Möbius inversion we have

$$\mathrm{Wg}_A(\pi) = \sum_{C\in[\pi,A]} C_{\pi,C}.$$

Remark 3.9. When $\pi \in S_{2l}^{(\epsilon)}$ and $A \in \mathcal{P}(2l)$ is π-invariant the equation above can also be written

$$\mathrm{Wg}_A(\tilde{\pi}) = \sum_{C\in[\pi,A]} C_{\tilde{\pi},\tilde{C}}. \tag{16}$$

In [3, Corollary 2.9] Collins showed that the order of $C_{\tilde{\pi},\tilde{C}}$ is at most $N^{-2l-\#(\pi)+2\#(C)}$.

In the following we address the estimates for higher order cumulants, k_r for $r \geqslant 3$.

If $D_1,\dots,D_{2l}$ are random matrices and $\pi \in S_{2l}$ is a permutation with cycle structure $\pi = \pi_1 \times \cdots \times \pi_r$ with $\pi_i = (\pi_{i,1},\dots,\pi_{i,l(i)})$ we denote

$$\mathrm{k}_\pi(D_1,\dots,D_{2l}) = \mathrm{k}_r\big(\mathrm{Tr}(D_{\pi_{1,1}}\cdots D_{\pi_{1,l(1)}}), \mathrm{Tr}(D_{\pi_{2,1}}\cdots D_{\pi_{2,l(2)}}),\dots\big).$$

When $A = \{A_1,\dots,A_k\}$ is a π-invariant partition of $[2l]$ we can write $\pi = \pi_1 \times \cdots \times \pi_k$ where $\pi_i = \pi|_{A_i}$ is a permutation of the set A_i. We denote the multiplicative extension of k_π by

$$\mathrm{k}_{\pi,A}(D_1,\dots,D_{2l}) = \mathrm{k}_{\pi_1}(D_1,\dots,D_{2l})\cdots\mathrm{k}_{\pi_k}(D_1,\dots,D_{2l}).$$

Möbius inversion gives us that

$$\mathrm{E}\big(\mathrm{Tr}_\pi(D_1,\dots,D_{2l})\big) = \sum_{\substack{A\in\mathcal{P}(2l)\\ A\ \pi\text{-inv.}}} \mathrm{k}_{\pi,A}(D_1,\dots,D_{2l}),$$

where the sums run over all π-invariant partitions A in $\mathcal{P}(2l)$.

Theorem 3.10. *Let $\{U\}_N$ be a sequence of Haar distributed unitary $N \times N$-random matrices and $\{A_1,\dots,A_s\}_N$ a sequence of $N \times N$-random matrices which has a second order limit distribution, given by $(\mathcal{A},\varphi_1,\varphi_2)$ and $a_1,\dots,a_s \in \mathcal{A}$. Furthermore, assume that $\{U\}_N$ and $\{A_1,\dots,A_s\}_N$ are independent.*

Suppose $r > 1$ and $m_1,\dots,m_r$ are positive integers such that $m_1 + \cdots + m_r = 2l$ and $\epsilon_1,\dots,\epsilon_{2l} \in \{-1,+1\}$ are such that $\sum_{i=1}^{2l} \epsilon_i = 0$. Consider polynomials $p_1,\dots,p_{2l}$ in s non-commuting indeterminates. For $i = 1,\dots,2l$ we set

$$D_i := p_i(A_1,\dots,A_s)$$

and for $1 \leqslant i \leqslant r$ let

$$X_i = \mathrm{Tr}\big(D_{m_1+\cdots+m_{i-1}+1}U^{\epsilon(m_1+\cdots+m_{i-1}+1)}\cdots D_{m_1+\cdots+m_i}U^{\epsilon(m_1+\cdots+m_i)}\big).$$

Then

$$\mathrm{k}_r(X_1,\dots,X_r)=\sum_{\pi\in S_n^{(\epsilon)}}\sum_{\substack{A,B\\ A\vee B=1_{[1,2l]}}} C_{\tilde\pi,\tilde A}\cdot \mathrm{k}_{\gamma\pi^{-1},B}(D_1,\dots,D_{2l}),\tag{17}$$

where the second sum runs over pairs (A,B) of partitions of $[1,2l]$ such that A is π-invariant and B is $\gamma\pi^{-1}$-invariant and furthermore $A\vee B=1_{[1,2l]}$.

Secondly, we have for $r\geqslant 3$ that

$$\lim_{N\to\infty}\mathrm{k}_r(X_1,\dots,X_r)=0.\tag{18}$$

If we have $m_1,\dots,m_r$ for which $m_1+\cdots+m_r$ is odd or $\epsilon_1,\dots,\epsilon_{2l}$ for which $\sum_{i=1}^{2l}\epsilon_i\neq 0$ then $\mathrm{k}_r(X_1,\dots,X_r)=0$.

Proof. In order to simplify the writing we shall write $\vec D$ for $(D_1,\dots,D_{2l})$. Let I_i be the interval $[m_1+\cdots+m_{i-1}+1, m_1+\cdots+m_i]$ and for any subset $V\subset[r]$, $I_V=\bigcup_{j\in V}I_j$.

If $C=\{V_1,\dots,V_k\}$ is a partition of $[r]$ we let $S_{V_i}^{(\epsilon)}$ the set of permutations of I_{V_i} that take $\epsilon^{-1}(1)\cap I_{V_i}$ onto $\epsilon^{-1}(-1)\cap I_{V_i}$. If these two sets have different cardinalities then $S_{V_i}^{(\epsilon)}$ is empty. Let 1_C be the partition $\{I_{V_1},\dots,I_{V_k}\}$ of $[2l]$. Let γ_i be the cyclic permutation of I_i given by $(m_1+\cdots+m_{i-1}+1,\dots,m_1+\cdots+m_i)$.

With this notation

$$\begin{aligned}
\mathrm{E}_C(X_1,\dots,X_r)&=\mathrm{E}_{V_1}(X_1,\dots,X_r)\cdots\mathrm{E}_{V_k}(X_1,\dots,X_r)\\
&=\sum_{\pi_1\in S_{V_1}^{(\epsilon)}}\sum_{\pi_k\in S_{V_k}^{(\epsilon)}}\mathrm{Wg}(\tilde\pi_1)\cdots\mathrm{Wg}(\tilde\pi_k)\mathrm{E}\big(\mathrm{Tr}_{\gamma_1\pi_1^{-1}}(\vec D)\big)\cdots\mathrm{E}\big(\mathrm{Tr}_{\gamma_k\pi_k^{-1}}(\vec D)\big)\\
&=\sum_{\substack{\pi\in S_{2l}^{(\epsilon)}\\ 1_C\ \pi\text{-inv.}}}\mathrm{Wg}_{1_C}(\tilde\pi)\mathrm{E}_C\big(\mathrm{Tr}_{\gamma\pi^{-1}}(\vec D)\big)\\
&=\sum_{\substack{\pi\in S_{2l}^{(\epsilon)}\\ 1_C\ \pi\text{-inv.}}}\sum_{A\in[\pi,1_C]}C_{\tilde\pi,\tilde A}\sum_{\substack{B\in\mathcal{P}(2l)\\ B\ \gamma\pi^{-1}\text{-inv.}}}\mathrm{k}_{\gamma\pi^{-1},B}(\vec D).
\end{aligned}$$

Thus

$$\begin{aligned}
&\mathrm{k}_r(X_1,\dots,X_r)\\
&\quad=\sum_{C\in\mathcal{P}(r)}\mathrm{M\ddot{o}b}(C,1_r)\mathrm{E}_C(X_1,\dots X_r)\\
&\quad=\sum_{C\in\mathcal{P}(r)}\mathrm{M\ddot{o}b}(C,1_r)\sum_{\substack{\pi\in S_{2l}^{(\epsilon)}\\ 1_C\ \pi\text{-inv.}}}\sum_{A\in[\pi,1_C]}C_{\tilde\pi,\tilde A}\sum_{\substack{B\in\mathcal{P}(2l)\\ B\ \gamma\pi^{-1}\text{-inv.}}}\mathrm{k}_{\gamma\pi^{-1},B}(\vec D)
\end{aligned}$$

$$= \sum_{\pi \in S_{2l}^{(\epsilon)}} \sum_{\substack{C \in \mathcal{P}(r) \\ 1_C\ \pi\text{-inv.}}} \sum_{A \in [\pi, 1_C]} \sum_{\substack{B \in \mathcal{P}(2l) \\ B\ \gamma\pi^{-1}\text{-inv.}}} \text{Möb}(C, 1_r) C_{\tilde{\pi}, \tilde{A}} \mathrm{k}_{\gamma\pi^{-1}, B}(\vec{D})$$

$$= \sum_{\pi \in S_{2l}^{(\epsilon)}} \sum_{\substack{A \in \mathcal{P}(2l) \\ A\ \pi\text{-inv.}}} \sum_{\substack{B \in \mathcal{P}(2l) \\ B\ \gamma\pi^{-1}\text{-inv.}}} \sum_{\substack{C \in \mathcal{P}(r) \\ A, B \leqslant 1_C}} \text{Möb}(C, 1_r) C_{\tilde{\pi}, \tilde{A}} \mathrm{k}_{\gamma\pi^{-1}, B}(\vec{D})$$

$$= \sum_{\pi \in S_{2l}^{(\epsilon)}} \sum_{\substack{A, B \in \mathcal{P}(2l) \\ A \vee B = 1_{2l}}} C_{\tilde{\pi}, \tilde{A}}, \mathrm{k}_{\gamma\pi^{-1}, B}(\vec{D}),$$

where the sum is over all A and B which are π and $\gamma\pi^{-1}$-invariant, respectively. The last equality followed from the identity

$$\sum_{\substack{C \in \mathcal{P}(r) \\ A, B \leqslant 1_C}} \text{Möb}(C, 1_r) = \begin{cases} 1 & A \vee B = 1_{2l}, \\ 0 & \text{otherwise.} \end{cases}$$

This proves (17).

We know that the order of $C_{\tilde{\pi}, \tilde{A}}$ is $N^{-2l - \#(\pi) + 2\#(A)}$. Let c_i be the number of blocks of B that contain i cycles of $\gamma\pi^{-1}$. By our assumption on the second order limiting distribution of $\{A_1, \ldots, A_s\}_N$

$$\mathrm{k}_{\gamma\pi^{-1}, B}(\vec{D}) = \begin{cases} O(N^{c_1}) & c_3 + c_4 + \cdots = 0, \\ o(N^{c_1}) & c_3 + c_4 + \cdots > 0. \end{cases}$$

Suppose first that $c_3 + c_4 + \cdots > 0$. Then

$$\sum_{i \geqslant 2} i c_i = (c_2 + c_3 + \cdots) + \sum_{i \geqslant 1} (i - 1) c_i \geqslant 1 + \#\big(\gamma\pi^{-1}\big) - \#(B).$$

So

$$c_1 = \#\big(\gamma\pi^{-1}\big) - \sum_{i \geqslant 2} i c_i \leqslant \#(B) - 1.$$

By Lemma 2.2(1), $\#(A) + \#(B) \leqslant 2l + 1$. Thus

$$-2l - \#(\pi) + 2\#(A) + c_1 \leqslant -2l - \#(\pi) + 2\#(A) + \#(B) - 1 \leqslant 0.$$

Hence $C_{\tilde{\pi}, \tilde{A}} \cdot \mathrm{k}_{\gamma\pi^{-1}}(\vec{D}) = o(N^0)$ as required.

So now suppose that $c_3 + c_4 + \cdots = 0$. Then $\#(\gamma\pi^{-1}) = c_2 + \#(B)$. In this case

$$C_{\tilde{\pi}, \tilde{A}} \cdot \mathrm{k}_{\gamma\pi^{-1}}(\vec{D}) = O\big(N^{-2l - \#(\pi) + 2\#(A) + c_1}\big)$$

and thus it remains to show that

$$-2l - \#(\pi) + 2\#(A) + c_1 \leqslant 2 - r. \tag{19}$$

Note that

$$\left|(1_{2l},\gamma)\right| = 2|1_{2l}| - |\gamma| = 2l - 2 + r,$$
$$\left|(A,\pi)\right| = 2|A| - |\pi| = 2l - 2\#(A) + \#(\pi)$$

and

$$\left|\left(B,\gamma\pi^{-1}\right)\right| = 2|B| - \left|\gamma\pi^{-1}\right| = 2l - 2\#(B) + \#\left(\gamma\pi^{-1}\right).$$

So by Lemma 2.2(2)

$$2\left(\#(A)+\#(B)\right) - 2l - \#(\pi) - \#\left(\gamma\pi^{-1}\right) \leqslant 2 - r. \tag{20}$$

However

$$2\#(B) - \#\left(\gamma\pi^{-1}\right) = \#(B) - c_2 = c_1$$

together with (20) proves (19). □

Remark 3.11. As a corollary of Theorem 3.10 we obtain that if $\{U\}_N$ is a sequence of Haar distributed unitary random matrices, then $\{U\}_N$ has a second order limit distribution given by Eq. (15). Indeed, relative to $\mathrm{E}(\mathrm{Tr}(\cdot))$, U is already a Haar unitary so condition (9) of Definition 3.5 is satisfied. We have observed in Remark 3.8 that condition (10) is satisfied and by Theorem 3.10 above we have that condition (11) is satisfied.

Then $\{U\}_N$ has a second order limit distribution which is given by

$$\lim_{N\to\infty} \mathrm{k}_2\left(\mathrm{Tr}\left(U^{\epsilon_1}\cdots U^{\epsilon_m}\right), \mathrm{Tr}\left(U^{\epsilon_{m+1}}\cdots U^{\epsilon_{m+n}}\right)\right)$$
$$= \sum_{\pi\in S_{NC}^{(\epsilon)}(m,n)} \mu(\tilde{\pi}) + \sum_{\substack{\pi_1\in NC^{(\epsilon_1)}(m)\\ \pi_2\in NC^{(\epsilon_2)}(n)}} \mu_2(\tilde{\pi}_1,\tilde{\pi}_2). \tag{21}$$

Combining this formula with Eq. (15) allows one to derive the values of μ_2. These kind of questions will be considered elsewhere.

Theorem 3.12. *Let $\{U\}_N$ be a sequence of Haar distributed unitary $N\times N$-random matrices and $\{A_1,\dots,A_s\}_N$ a sequence of $N\times N$-random matrices which has a second order limit distribution. If $\{U\}_N$ and $\{A_1,\dots,A_s\}_N$ are independent, then they are asymptotically free of second order.*

Proof. The asymptotic freeness with respect to $\mathrm{k}_1(\mathrm{tr}(\cdot))$ is essentially the same argument as Voiculescu's proof [16,17] for the case of constant matrices, see also the proof of Collins [3].

Theorem 3.10 provides the bound on higher order cumulants so we need to prove now only the second order statement.

We have to consider cyclically alternating and centered words in the U's and the A's. For the U's, every centered word is a linear combination of non-trivial powers of U, thus it suffices to consider such powers. Thus we have to look at expressions of the form

$$\mathrm{k}_2\big(\mathrm{Tr}\big(B_1U^{i(1)}\cdots B_pU^{i(p)}\big), \mathrm{Tr}\big(U^{j(r)}C_r\cdots U^{j(1)}C_1\big)\big), \tag{22}$$

where the B's and the C's are centered polynomials in the A's and $i(1),\dots,i(p), j(1),\dots,j(r)$ are integers different from zero. We have to show that in the limit $N\to\infty$ the expression (22) converges to

$$\lim_N \delta_{pr}\sum_{k=0}^{p-1}\varphi_1(B_1C_{1+k})\varphi_1\big(U^{i(1)}U^{j(1+k)}\big)\cdots\varphi_1(B_pC_{p+k})\varphi_1\big(U^{i(p)}U^{j(p+k)}\big). \tag{23}$$

We can bring the expression (22) into the form considered in Theorem 3.7 by inserting 1's between neighbouring factors U or neighbouring factors U^*. If we relabel the B's, C's, and 1's as D's then we have to look at the following situation: For polynomials p_i in s non-commuting indeterminates we consider

$$D_i := p_i(A_1,\dots,A_s),$$

which are either asymptotically centered or equal to 1. The latter case can only appear if we have cyclically the pattern $\dots UD_iU\dots$ or $\dots U^*D_iU^*\dots$. Formally, this means:

- if $\epsilon_{\gamma^{-1}(i)}=\epsilon_i$ then either $D_i=1$ (for all N, i.e., $p_i=1$) or

$$\lim_{N\to\infty}\mathrm{k}_1\big(\mathrm{tr}[D_i]\big)=0;$$

- if $\epsilon_{\gamma^{-1}(i)}\neq\epsilon_i$ then

$$\lim_{N\to\infty}\mathrm{k}_1\big(\mathrm{tr}[D_i]\big)=0.$$

We can now use Theorem 3.7 for calculating the limit

$$\lim_{N\to\infty}\mathrm{k}_2\big(\mathrm{Tr}\big(D_1U^{\epsilon_1}\cdots D_mU^{\epsilon_m}\big), \mathrm{Tr}\big(D_{m+1}U^{\epsilon_{m+1}}\cdots D_{m+n}U^{\epsilon_{m+n}}\big)\big),$$

and we will argue that most terms appearing there will vanish. First consider the last two sums in Eq. (12), corresponding to $\pi_1\in NC(m)$ and $\pi_2\in NC(n)$. Since π_1 is non-crossing we have that $\#(\pi_1)+\#(\gamma_m\pi_1^{-1})=m+1$. Since each cycle of π_1 must contain at least one U and one U^*, we have

$$\#(\pi_1)\leqslant\frac{m}{2},$$

which implies $\#(\gamma_m\pi_1^{-1}) \geqslant m/2+1$. However, this can only be true if $\gamma_m\pi_1^{-1}$ contains at least two singletons. Note that if (i) is a singleton of $\gamma_m\pi_1^{-1}$ and if we have $D_i = 1$ for that i, then we have

$$\gamma_m\pi_1^{-1}(i) = i, \quad \text{thus} \quad \pi_1^{-1}(i) = \gamma_m^{-1}(i) = \gamma^{-1}(i),$$

and hence

$$\epsilon_{\pi_1^{-1}(i)} = \epsilon_{\gamma^{-1}(i)} = \epsilon_i,$$

which is not allowed because π_1 is from $NC^{(\epsilon_1)}(m)$, i.e., it must connect alternatingly U with U^*. Hence $D_i \neq 1$ and so $\varphi_1(d_i) = \lim_N \mathrm{k}_1(\mathrm{tr}(D_i)) = 0$. Thus, both

$$\varphi_1\big(\gamma_m\pi_1^{-1} \times \gamma_n\pi_2^{-1}\big)[d_1, \ldots, d_{m+n}]$$

and

$$\varphi_2\big(\gamma_m\pi_1^{-1}, \gamma_n\pi_2^{-1}\big)[d_1, \ldots, d_{m+n}]$$

are zero, because at least one singleton (i) gives the contribution $\varphi_1(d_i) = 0$.

Consider now the first summand of Eq. (12). Suppose $\pi \in S_{NC}^{(\epsilon)}(m,n)$. Let us again put $\gamma := \gamma_{m,n}$. Since π is annular non-crossing we have

$$|\pi| + \big|\gamma\pi^{-1}\big| = m+n,$$

or

$$\#(\pi) + \#\big(\gamma\pi^{-1}\big) = m+n.$$

Again, each cycle of π must contain at least two elements, i.e.,

$$\#(\pi) \leqslant \frac{m+n}{2},$$

thus

$$\#\big(\gamma\pi^{-1}\big) \geqslant \frac{m+n}{2}.$$

If $\gamma\pi^{-1}$ has a singleton (i), then this will contribute $\varphi_1(d_i)$ and since, as above the case $d_i = 1$ is excluded for a singleton, we get a vanishing contribution in this case. This implies that, in order to get a non-vanishing contribution, $\gamma\pi^{-1}$ must contain no singletons, which, however, means that we must have

$$\#\big(\gamma\pi^{-1}\big) = \frac{m+n}{2}, \quad \text{and thus also} \quad \#(\pi) = \frac{m+n}{2},$$

i.e., all cycles of $\gamma\pi^{-1}$ and of π contain exactly two elements. This, however, can only be the case if each cycle connects one point on the outer circle to one point on the inner circle. Being non-crossing fixes the permutation up to a rotation of the inner circle. Thus, in order to get a non-vanishing contribution, we need $m = n$ and

$$\pi = \big(1, \gamma^k(2n)\big)\big(2, \gamma^k(2n-1)\big), \ldots, \big(n, \gamma^k(n+1)\big)$$

for some $k = 0, 1, \ldots, n-1$. Note that π must always couple a U with a U^* and the factor $\mu(\tilde{\pi})$ is always 1 for such pairings. This gives exactly the contribution as needed for second order freeness. □

Of course, a natural question in this context is how the result of Diaconis and Shahshahani (Remark 3.8) generalizes to the case of several independent unitary random matrices. Note that as we have established the existence of a second order limit distribution for Haar distributed unitary random matrices we can use an independent copy of them as the ensemble $\{A_1, \ldots, A_s\}$ in our Theorem 3.12. By Remark 2.7 this can be iterated to give the following.

Theorem 3.13. *Let $\{U^{(1)}\}_N, \ldots, \{U^{(r)}\}_N$ be r sequences of Haar distributed unitary $N \times N$-random matrices. If $\{U^{(1)}\}_N, \ldots, \{U^{(r)}\}_N$ are independent, then they are asymptotically free of second order.*

This contains the information about the fluctuation of several independent Haar distributed unitary random matrices. Again, it suffices to consider traces of reduced words in our random matrices, i.e., expressions of the form

$$\mathrm{Tr}\big(U_{i(1)}^{k(1)} \cdots U_{i(n)}^{k(n)}\big) \tag{24}$$

for $n \in \mathbb{N}$, and $k(r) \in \mathbb{Z} \setminus \{0\}$ and $i(r) \neq i(r+1)$ for all $r = 1, \ldots, n$ (where $i(n+1) = i(1)$). But these are now products in cyclically alternating and centered variables, so that by the very definition of second order freeness we get

$$\begin{aligned}
&\lim_{N\to\infty} \mathrm{k}_2\big(\mathrm{Tr}\big(U_{i(1)}^{k(1)} \cdots U_{i(m)}^{k(m)}\big), \mathrm{Tr}\big(U_{j(n)}^{l(n)} \cdots U_{j(1)}^{l(1)}\big)\big) \\
&\quad = \delta_{mn} \sum_{r=0}^{n-1} \varphi_1\big(U_{i(1)}^{k(1)} U_{j(1+r)}^{l(1+r)}\big) \cdots \varphi_1\big(U_{i(n)}^{k(n)} U_{j(n+r)}^{l(n+r)}\big).
\end{aligned} \tag{25}$$

The contribution of φ_1 in these terms vanishes unless the matrices and their powers match. Note also that the vanishing of higher cumulants can be rephrased in a more probabilistic language by saying that the random variables (24) converge to a Gaussian family.

Corollary 3.14. *Let $\{U_{(1)}\}_N, \ldots, \{U_{(r)}\}_N$ be independent sequences of Haar distributed unitary $N \times N$-random matrices. Then, the collection* (24) *of unnormalized traces in cyclically reduced words in these random matrices converges to a Gaussian family of centered random variables whose covariance is given by the number of matchings between the two reduced words,*

$$\begin{aligned}
&\lim_{N\to\infty} \mathrm{k}_2\big(\mathrm{Tr}\big(U_{i(1)}^{k(1)} \cdots U_{i(m)}^{k(m)}\big), \mathrm{Tr}\big(U_{j(n)}^{l(n)} \cdots U_{j(1)}^{l(1)}\big)\big) \\
&\quad = \delta_{mn} \cdot \#\big\{r \in \{1, \ldots, n\} \mid i(s) = j(s+r),\ k(s) = -l(s+r)\ \forall s = 1, \ldots, n\big\}.
\end{aligned} \tag{26}$$

This result was also obtained independently in the recent work of Rădulescu [15] around Connes's embedding problem.

The following theorem gives an easy way to construct families of random matrices which are asymptotically free of second order.

Theorem 3.15. *Let $\{U\}_N$ be a sequence of Haar distributed unitary $N \times N$-random matrices. Suppose that $\{A_1, \dots, A_s\}_N$ and $\{B_1, \dots, B_t\}_N$ are sequences of $N \times N$-random matrices each of which has a second order limit distribution. Furthermore, assume that $\{A_1, \dots, A_s\}_N$, $\{B_1, \dots, B_t\}_N$, and $\{U\}_N$ are independent. Then the sequences $\{A_1, \dots, A_s\}_N$ and $\{UB_1U^*, \dots, UB_tU^*\}_N$ are asymptotically free of second order.*

Proof. The proof is a repetition of the proof of Theorem 3.12 except that we cannot assume that $\{A_1, \dots, A_s, B_1, \dots, B_t\}_N$ has a second order limit distribution. Instead we have the independence of the A_i from the B_i's and a special ϵ. So we shall only indicate how the proof has to be modified.

The first order freeness follows as in the proof of Theorem 3.12. In the proofs of Theorems 3.7, 3.10, and 3.12 the cumulants we need are all sums over $S_{2n}^{(\epsilon)}$ for various n's. Now we have a special form of ϵ, namely

$$\epsilon_i = (-1)^{i+1}.$$

Thus if $\pi \in S_{2n}^{(\epsilon)}$ then π takes even numbers to odd numbers and vice versa. Since the same applies to any of the γ's, we have that $\gamma\pi^{-1}$ takes even numbers to even numbers and odd numbers to odd numbers. Thus the orbits of $\gamma\pi^{-1}$ consist either of all odd numbers or of all even numbers. Hence if $P_1, \dots, P_n$ are words in $A_1, \dots, A_s$ and $Q_1, \dots, Q_n$ are words in $B_1, \dots, B_t$, then by the independence of the A's and the B's

$$\begin{aligned}
&\mathrm{E}\big(\mathrm{Tr}_{\gamma\pi^{-1}}(P_1, Q_1, \dots, P_n, Q_n)\big) \\
&\quad = \mathrm{E}\big(\mathrm{Tr}(W_1) \cdots \mathrm{Tr}(W_j)\big) \cdot \mathrm{E}\big(\mathrm{Tr}(W_{j+1}) \cdots \mathrm{Tr}(W_k)\big),
\end{aligned}$$

where $k = \#(\gamma\pi^{-1})$ and W_i for $1 \leqslant i \leqslant j$ is a word only in A's and for $j+1 \leqslant i \leqslant k$ is a word only in B's. This means that as far as the asymptotic behaviour of $\mathrm{E}(\mathrm{Tr}_{\gamma\pi^{-1}}(P_1, Q_1, \dots, P_n, Q_n))$ is concerned we may assume that $\{A_1, \dots, A_s, B_1, \dots, B_t\}_N$ has a second order limit distribution. Hence our claim follows from Theorem 3.12. $\square$

We say that a tuple $\{B_1, \dots, B_s\}$ of $N \times N$-random matrices is $\mathcal{U}(N)$-invariant if for every $U \in \mathcal{U}(N)$ the joint probability distribution of the random matrices $\{B_1, \dots, B_s\}$ coincides with the joint probability distribution of the random matrices $\{UB_1U^*, \dots, UB_sU^*\}$.

Corollary 3.16. *Let $\{A_1, \dots, A_s\}_N$ be a sequence of $N \times N$-random matrices which has a second order limit distribution and let $\{B_1, \dots, B_t\}_N$ be a sequence of $\mathcal{U}(N)$-invariant $N \times N$-random matrices which has a second order limit distribution. Furthermore, assume that the matrices $\{A_1, \dots, A_s\}_N$ and the matrices $\{B_1, \dots, B_t\}_N$ are independent. Then the sequences $\{A_1, \dots, A_s\}_N$ and $\{B_1, \dots, B_t\}_N$ are asymptotically free of second order.*

4. Failure of universality for multi-matrix models

In this section we want to make the meaning of second order freeness for fluctuations of random matrices more explicit and relate this with the question of universality of such fluctuations. There has been much interest in global fluctuations of random matrices, in particular, since it was observed that for large classes of one-matrix models these fluctuations are universal. In the physical literature this observation goes at least back to Politzer [14], culminating in the paper of Ambjørn et al. [1], whereas a proof on the mathematical level of rigour is due to Johansson [7]. Universality for one-matrix models lets one expect that one would also have such universality for multi-matrix models. Indeed, this expectation was one of the starting points of our investigations. However, our machinery around second order freeness shows that such universality is *not* present in multi-matrix models.

Before we address multi-matrix models let us first recall the relevant result of Johansson [7]. We consider Hermitian $N \times N$-random matrices $A = (a_{ij})_{i,j=1}^N$ equipped with the probability measure

$$d\mu_N(A) = \frac{1}{Z_N} \exp\bigl\{-N \operatorname{Tr}\bigl[P(A)\bigr]\bigr\}\, dA, \tag{27}$$

where

$$dA = \prod_{1 \leqslant i < j \leqslant N} d\operatorname{Re} a_{ij}\, d\operatorname{Im} a_{ij} \prod_{i=1}^{N} da_{ii}.$$

Here, P is a polynomial in one variable, which we will address as "potential" in the following, and Z_N is a normalization constant. If one restricts to a special class $\mathcal{V}$ of potentials P (the most important condition being that the limit eigenvalue distribution has a single interval as support—which we normalize in the following to $[-2,2]$) then Johansson proved the following universality of fluctuations for this class: Consider the sequence of $N \times N$-random matrices $\{A_N\}_N$ given by (27). Then this sequence has a second order limit distribution which can be described as follows:

(1) We have

$$\lim_{N\to\infty} \mathrm{k}_1\bigl\{\operatorname{tr}\bigl[A_N^n\bigr]\bigr\} = \int t^n \, d\nu_P(t) \quad (n \in \mathbb{N}),$$

where ν_P is a probability measure on $\mathbb{R}$ ("limiting eigenvalue distribution") which is given as the solution of the singular integral equation

$$\int \frac{d\nu_P(t)}{t-s} = -\frac{1}{2} P'(s) \quad \text{for all } s \in \operatorname{supp} \nu_P. \tag{28}$$

(2) Let T_n ($n \in \mathbb{N}$) be the Chebyshev polynomials of first kind (which are the orthogonal polynomials with respect to the arcsine law on $[-2,2]$). Then

$$\lim_{N\to\infty} \mathrm{k}_2\bigl\{\operatorname{Tr}\bigl[T_n(A_N)\bigr], \operatorname{Tr}\bigl[T_m(A_N)\bigr]\bigr\} = \delta_{mn} \cdot n.$$

Whereas the limiting eigenvalue distribution ν_P depends on the form of the potential P, the fluctuations are the same for all potentials in the class $\mathcal{V}$—they are always diagonalized by the same polynomials $\{T_n \mid n \in \mathbb{N}\}$. Note that the most prominent example for the considered class of random matrices is given by $P(A) = A^2$, which corresponds to the Gaussian random matrix ensemble. So one can phrase this universality also in the way that all considered random matrices have the same fluctuations as Gaussian random matrices (but their eigenvalue distributions are of course different from Wigner's semi-circle law).

Let us now consider multi-matrix models. For notational convenience we restrict to the case of two-matrix models. Take now a polynomial $P(A, B)$ in two non-commuting variables and consider pairs of Hermitian $N \times N$-matrices $A = (a_{ij})_{i,j=1}^N$ and $B = (b_{ij})_{i,j=1}^N$ equipped with the probability measure

$$d\mu_N(A, B) = \frac{1}{Z_N} \exp\bigl\{-N \operatorname{Tr}\bigl[P(A, B)\bigr]\bigr\}\, dA\, dB, \tag{29}$$

where Z_N is a normalization constant. The simplest ensemble of this kind is the case of two independent Gaussian random matrices which corresponds to the choice $P(A, B) = A^2 + B^2$. The above mentioned universality result for the one-matrix case lets one expect that one might also have universality for multi-matrix ensembles which are close to the ensemble of independent Gaussian random matrices. However, we will now show that even restricted to potentials of the form $P(A, B) = P_1(A) + P_2(B)$ we do not have universal fluctuations.

Let us first observe that our concept of second order freeness tells us how to diagonalize fluctuations. We spell this out in the following theorem which is an easy consequence of Definition 2.5 of second order freeness.

Theorem 4.1. *Let $\{A_N\}_N$ and $\{B_N\}_N$ be two sequences of $N \times N$-random matrices which are asymptotically free of second order. Let $\{Q_n^A \mid n \in \mathbb{N}\}$ and $\{Q_n^B \mid n \in \mathbb{N}\}$ be the orthogonal polynomials for the limiting eigenvalue distribution of A_N and B_N, respectively, determined by the requirements that Q_n^A and Q_n^B are polynomials of degree n and that*

$$\lim_{N\to\infty} \mathrm{k}_1\bigl\{\operatorname{tr}\bigl[Q_n^A(A_N) Q_m^A(A_N)\bigr]\bigr\} = \delta_{nm},$$
$$\lim_{N\to\infty} \mathrm{k}_1\bigl\{\operatorname{tr}\bigl[Q_n^B(B_N) Q_m^B(B_N)\bigr]\bigr\} = \delta_{nm}.$$

Then the fluctuations of mixed traces in A_N and B_N are diagonalized by cyclically alternating products of Q_n^A and Q_n^B and the covariances are given by the number of cyclic matchings of these products:

$$\begin{aligned}
&\lim_{N\to\infty} \mathrm{k}_2\bigl\{\operatorname{Tr}\bigl[Q_{i(1)}^A(A_N) Q_{j(1)}^B(B_N) \cdots Q_{i(m)}^A(A_N) Q_{j(m)}^B(B_N)\bigr],\\
&\qquad \operatorname{Tr}\bigl[Q_{l(n)}^B(B_N) Q_{k(n)}^A(A_N) \cdots Q_{l(1)}^B(B_N) Q_{k(1)}^A(A_N)\bigr]\bigr\}\\
&\quad = \delta_{mn} \cdot \#\bigl\{r \in \{1, \dots, n\} \mid i(s) = k(s + r),\ j(s) = l(s + r)\ \forall s = 1, \dots, n\bigr\}.
\end{aligned}$$

In the case of Wishart matrices the polynomials are shown in [8] to have an interpretation in terms of planar diagrams whose linear span forms a basis for the irreducible representations of the of the annular Temperly–Lieb algebra.

To come back to our problem of universality for multi-matrix models we only have to observe that Corollary 3.16 tells us that we have asymptotic freeness of second order if we choose a potential of the form $P(A, B) = P_1(A) + P_2(B)$.

Theorem 4.2. *Let $P(A, B) = P_1(A) + P_2(B)$ where P_1 and P_2 are polynomials from the class $\mathcal{V}$. Consider the two-matrix model $\{A_N, B_N\}_N$ given by the probability measure* (29). *Let ν_1 and ν_2 be the limiting eigenvalue distribution for P_1 and P_2, respectively* (*as described in Eq.* (28)) *and denote by $\{Q_n^1 \mid n \in \mathbb{N}\}$ and $\{Q_n^2 \mid n \in \mathbb{N}\}$ the respective orthogonal polynomials. Then the sequence $\{A_N, B_N\}_N$ has a second order limit distribution given by $(\mathcal{A}, \varphi_1, \varphi_2)$ and $a, b \in \mathcal{A}$ which can be described as follows*:

(1) *φ_1 is the free product of ν_1 and ν_2.*
(2) *φ_2 is diagonalized by the following collection of polynomials*:

$$\big\{T_n(a) \mid n \in \mathbb{N}\big\}, \qquad \big\{T_n(b) \mid n \in \mathbb{N}\big\},$$
$$\big\{Q^1_{i(1)}(a) Q^2_{j(1)}(b) \cdots Q^1_{i(n)}(a) Q^2_{j(n)}(b)\big\},$$

where in the last set we choose one representative from each cyclic equivalence class, i.e., for all $n \in \mathbb{N}$, n-tuples $((i(1), j(1)), \ldots, (i(n), j(n)))$ which are different modulo cyclic rotation.

Proof. Note that the additive form of the potential $P(A, B) = P_1(A) + P_2(B)$ means that A_N and B_N are independent, A_N is a one-matrix ensemble corresponding, via (27), to a potential P_1, and B_N is a one-matrix ensemble corresponding to a potential P_2. Thus the statement about the diagonalizing polynomials in either only A_N or in only B_N is just Johansson's result. For getting the statement about the diagonalizing polynomials in both A_N and B_N we have to note that the random matrices A_N (and also B_N) are $\mathcal{U}(N)$-invariant, thus Corollary 3.16 implies that $\{A_N\}_N$ and $\{B_N\}_N$ are asymptotically free of second order. Hence we can apply Theorem 4.1 above. □

Note that whereas the polynomials in only one of the matrices are universal (namely equal to the Chebyshev polynomials $\{T_n\}_n$), the polynomials which involve both matrices are not universal but depend on the eigenvalue distributions ν_1 and ν_2. Since the latter vary with the potentials P_1 and P_2, the polynomials Q_n^1 and Q_n^2, and thus also the alternating products in them, depend on the choice of P. Thus we can conclude that even within the very restricted class of P's of the form $P(A, B) = P_1(A) + P_2(B)$ we have no universality of global fluctuations in multi-matrix models.

References

[1] J. Ambjørn, J. Jurkiewicz, Y. Makeenko, Multiloop correlators for two-dimensional quantum gravity, Phys. Lett. B 251 (1990) 517–524.
[2] P. Biane, Some properties of crossings and partitions, Discrete Math. 175 (1–3) (1997) 41–53.
[3] B. Collins, Moments and cumulants of polynomial random variables of unitary groups, the Itzykson–Zuber integral, and free probability, IMRN 17 (2003) 953–982.
[4] B. Collins, P. Śniady, Integration with respect to the Haar measure on unitary, orthogonal and symplectic group, Comm. Math. Phys. 264 (2006) 773–795.
[5] P. Diaconis, Patterns in eigenvalues: The 70th Josiah Willlard Gibbs lecture, Bull. Amer. Math. Soc. 40 (2003) 155–178.

[6] P. Diaconis, M. Shahshahani, On the eigenvalues of random matrices, J. Appl. Probab. A 31 (1994) 49–62.
[7] K. Johansson, On fluctuations of eigenvalues of random hermitian matrices, Duke Math. J. 91 (1998) 151–204.
[8] T. Kusalik, J. Mingo, R. Speicher, Orthogonal polynomials and fluctuations of random matrices, J. Reine Angew. Math., in press.
[9] F. Lehner, Cumulants in noncommutative probability I, Noncommutative exchangeability systems, Math. Z. 248 (2004) 67–100.
[10] J. Mingo, A. Nica, Annular non-crossing permutations and partitions, and second-order asymptotics for random matrices, IMRN 28 (2004) 1413–1460.
[11] J. Mingo, R. Speicher, Second order freeness and fluctuations of random matrices: I. Gaussian and Wishart matrices and cyclic Fock spaces, J. Funct. Anal. 235 (2006) 226–270.
[12] A. Nica, R. Speicher, Lectures on the Combinatorics of Free Probability Theory, Paris, 1999.
[13] F. Hiai, D. Petz, The Semicircle Law, Free Random Variables and Entropy, Amer. Math. Soc., Providence, RI, 2000.
[14] D. Politzer, Random-matrix description of the distribution of mesoscopic conductance, Phys. Rev. B 40 (1989) 11917–11919.
[15] F. Rădulescu, Combinatorial aspects of Connes's embedding conjecture, preprint math.OA/0404308, 2004.
[16] D. Voiculescu, Limit laws for random matrices and free products, Invent. Math. 104 (1991) 201–220.
[17] D. Voiculescu, A strengthened asymptotic freeness result for random matrices with applications to free entropy, IMRN 1 (1998) 41–63.
[18] D. Voiculescu, Lectures on free probability theory, in: Lecture Notes in Math., vol. 1738, Springer, 2000, pp. 279–349.
[19] D. Voiculescu, K. Dykema, A. Nica, Free Random Variables, Amer. Math. Soc., Providence, RI, 1992.
[20] D. Weingarten, Asymptotic behavior of group integrals in the limit of infinite rank, J. Math. Phys. 19 (1978) 999–1001.
[21] Feng Xu, A random matrix model from two-dimensional Yang–Mills theory, Comm. Math. Phys. 190 (1997) 287–307.