MPS-RR 1999-39
ISSN 1398-599X

Asymptotic Freeness Almost Everywhere for Random Matrices

Fumio Hiai[1] and Dénes Petz[2]

[1] Graduate School of Information Sciences
Tohoku University
Aoba-ku, Sendai 980-8577, Japan

[2] Department for Mathematical Analysis
Technical University of Budapest
H-1521 Budapest XI. Sztoczek u. 2, Hungary

Abstract

Voiculescu's asymptotic freeness result for random matrices is improved to the sense of almost everywhere convergence. The asymptotic freeness almost everywhere is first shown for standard unitary matrices based on the computation of multiple moments of their entries, and then it is shown for rather general unitarily invariant selfadjoint random matrices (in particular, standard selfadjoint Gaussian matrices) by applying the first result to the unitary parts of their diagonalization. Bi-unitarily invariant non-selfadjoint random matrices are also treated via polar decomposition.

1991 *Mathematics Subject Classification.* Primary 15A52; Secondary 62E20, 60F99.
Key words and phrases. random matrices, free probability, asymptotic freeness.

Introduction

A random matrix is a matrix whose entries are real or complex random variables (on a probability space). When tr_n denotes the normalized trace on the $n \times n$ matrices, the space of $n \times n$ random matrices admits a natural linear functional τ_n defined by

$$\tau_n(X) := E(\mathrm{tr}_n(X)) = \frac{1}{n}\sum_{i=1}^{n} E(X_{ii})$$

for random matrices $X = [X_{ij}]_{i,j=1}^{n}$ (whenever the expectations exist). Indeed, the $n \times n$ random matrices whose entries have all moments form a *-algebra (under usual matrix

[1]Supported by Grant-in-Aid for Scientific Research (C)09640152.
[2]Supported by OTKA F023447 and FKFP 502-121.

operations) which is a noncommutative probability space with the tracial functional τ_n (in the terminology in free probability theory [25]).

The classical Wigner theorem ([27], [28]) tells us that the mean spectral density of certain random symmetric matrices tends to the semicircle law if the matrix size goes to infinity. This convergence is concerned with the eigenvalue distribution with respect to the functionals τ_n. Later on, Arnold [1] proved that the empirical spectral density of real symmetric (also complex selfadjoint) random matrices with independent entries converges to the semicircle law almost everywhere, that is, the distribution with respect to tr_n converges almost surely. Currently, not only a random matrix model of the semicircle law, we know many random matrix models which produce particular distributions as the limiting eigenvalue density in the almost sure sense. For instance, certain non-selfadjoint random matrices admit the circular law as the limiting eigenvalue density ([10], [2]), and the Marchenko-Pastur distribution [11] appears as the limit distribution of Wishart matrices ([26], [14]). Furthermore, several results are known about the almost sure convergence of the largest/smallest eigenvalue or the norm (also the spectral radius) of symmetric or non-symmetric random matrices (see [7], [8], [3], [4] and also recent [9]). We also know that random unitary matrices sometimes play important roles in random matrix theory (see [13]), and the almost sure limit distribution of standard random unitary matrices is the uniform distribution on the unit circle.

The asymptotic free property of random matrices is central in recent breakthrough of free probability theory. It says that the purely algebraic concept of free relation of noncommutative random variables can be also modeled by random matrix ensembles if the matrix size goes to infinity. The asymptotic freeness result was first established by Voiculescu [22] in the case of Gaussian random matrices together with diagonal constant matrices. Further, Dykema [5] proved the same result in the case of general (non-Gaussian) random matrices together with block-diagonal constant matrices with bounded block-size, and recently Voiculescu [24] proved his asymptotic freeness result without restriction on the type of constant matrices. The inclusion of constant matrices in these results has played a crucial role in applications to von Neumann algebra theory (in particular, to problems on free group factors) ([21], [15]–[17], [6]) and to free entropy ([23], [24]). The paper [18] is concerned with the asymptotic freeness for matrices having bosonic and fermionic creations as entries.

Our motivation in the present paper is twofold. On one hand, we want to prove the asymptotic freeness for random matrices in the almost everywhere sense. This can be naturally expected from the above mentioned fact that many typical examples of random matrices (such as standard selfadjoint or non-selfadjoint Gaussian matrices, standard unitary matrices and so on) have the almost sure limit distribution when they are treated as a single sequence in the matrix size tending to infinity. In Sect. 2 we give the precise definition of the asymptotic freeness almost everywhere. On the other hand, we have believed that there should be a proof of the asymptotic freeness result starting from standard unitary random matrices. Indeed, in [22] Voiculescu obtained the asymptotic freeness of standard unitaries by taking the unitary parts in the polar decomposition of non-selfadjoint Gaussian matrices. Also, Speicher [19] used a similar method to show the almost sure limit spectral density of the sum of two selfadjoint matrices. Our approach is opposite. In this paper we first treat the asymptotic freeness of standard unitaries and then go to the case of certain selfadjoint matrices via the diagonalization process. An advantage of our approach is that standard unitaries themselves are Haar unitaries so that we consider only their monomials in proving the asymptotic freeness result. This simplifies the proof considerably. Another advantage is that we can treat unitarily invariant selfadjoint and bi-unitarily invariant non-selfadjoint random matrices more generally than Gaussian matrices. Note that a different approach was adopted by Xu [29] to obtain asymptotic freeness results for unitary random matrices. Moreover, almost sure convergence of mixed moments of random matrices was recently discussed in [20] too.

In Sect. 1 we start with computation of multiple moments of entries of a standard unitary. A convenient proof of the almost sure convergence of standard selfadjoint Gaussian matrices is also given. In Sect. 2, based on the computation in Sect. 1, the asymptotic freeness almost everywhere is established for independent standard unitaries together with constant matrices. In Sect. 3 we apply this result via diagonalization to obtain the same result for independent unitarily invariant selfadjoint random matrices. Furthermore, bi-unitarily invariant non-selfadjoint random matrices are treated via polar decomposition.

1 Preliminaries

Let $M_n(\mathbb{C})$ be the space of $n \times n$ complex matrices and $M_n(\mathbb{C})^{sa}$ the space of selfadjoint matrices in $M_n(\mathbb{C})$. The normalized trace on $M_n(\mathbb{C})$ is denoted by tr_n. Let $\mathcal{U}(n)$ be the compact group of $n \times n$ unitary matrices, and γ_n be the Haar probability measure on $\mathcal{U}(n)$. An $n \times n$ random unitary matrix is said to be *standard* if its distribution on $\mathcal{U}(n)$ is γ_n. The standard unitary random matrix U is a noncommutative random variable with respect to the functional τ_n given in Introduction. Due to the invariance of γ_n, $e^{\mathrm{i}\,\theta}U$ is standard for any $\theta \in \mathbb{R}$ so that the moments $\tau_n(U^k)$ vanish for all $k \in \mathbb{Z} \setminus \{0\}$, that is, U is a Haar unitary ([25], p. 58).

In this section we first compute higher order correlations (i.e. multiple moments) among the entries of the standard random unitary $U = [U_{ij}]_{i,j=1}^n$. We may of course consider on the probability space $(\mathcal{U}(n), \gamma_n)$ and have the expectation $E(f) := \int f(U)\, d\gamma_n(U)$ (whenever it exists) for a measurable function $f : \mathcal{U}(n) \to \mathbb{C}$. The two-sided invariance of γ_n guarantees that $E(f(U)) = E(f(VUW))$ is valid for any $V, W \in \mathcal{U}(n)$. When $V = \mathrm{diag}(e^{\mathrm{i}\,\theta_1}, \dots, e^{\mathrm{i}\,\theta_n})$ and $W = \mathrm{diag}(e^{\mathrm{i}\,\psi_1}, \dots, e^{\mathrm{i}\,\psi_n})$, we have

$$E(f) = E(f([e^{\mathrm{i}\,(\theta_i+\psi_j)} U_{ij}]_{i,j=1}^n)) \tag{1.1}$$

for all $\theta_i, \psi_j \in \mathbb{R}$. When V, W are permutation matrices, we have

$$E(f) = E(f([U_{\pi(i),\sigma(j)}]_{i,j=1}^n)) \tag{1.2}$$

for all permutations π, σ of $\{1, \dots, n\}$. These invariance (or symmetry) properties are enough for our purpose.

The next lemma says that most of multiple moments of the elements U_{ij} vanish.

Lemma 1.1 *Let $l \in \mathbb{N}$, $i_1, \ldots, i_l, j_1, \ldots, j_l \in \{1, \ldots, n\}$ and $k_1, \ldots, k_l, m_1, \ldots, m_l \in \mathbb{Z}^+$ $(:= \{0, 1, 2, \ldots\})$. If either $\sum_{i_r=i}(k_r - m_r) \neq 0$ for some $1 \leq i \leq n$ or $\sum_{j_r=j}(k_r - m_r) \neq 0$ for some $1 \leq j \leq n$, then*

$$E((U_{i_1j_1}^{k_1}\bar{U}_{i_1j_1}^{m_1})(U_{i_2j_2}^{k_2}\bar{U}_{i_2j_2}^{m_2})\cdots(U_{i_lj_l}^{k_l}\bar{U}_{i_lj_l}^{m_l})) = 0\,. \tag{1.3}$$

In particular, if $\sum_{r=1}^{l}(k_r - m_r) \neq 0$ (this is the case when $\sum_{r=1}^{l}(k_r + m_r)$ is odd), then (1.3) holds.

Proof. Suppose that $h := \sum_{i_r=i}(k_r - m_r) \neq 0$. One can apply (1.1) to get

$$E((U_{i_1j_1}^{k_1}\bar{U}_{i_1j_1}^{m_1})\cdots(U_{i_lj_l}^{k_l}\bar{U}_{i_lj_l}^{m_l})) = e^{\mathrm{i}\,h\theta}E((U_{i_1j_1}^{k_1}\bar{U}_{i_1j_1}^{m_1})\cdots(U_{i_lj_l}^{k_l}\bar{U}_{i_lj_l}^{m_l}))$$

for every $\theta \in \mathbb{R}$. This gives the conclusion.

The following is a list of multiple moments up to the fourth order. It is immediately seen from Lemma 1.1 that all other multiple moments up to the fourth order of elements $U_{ij}, \bar{U}_{ij}$ are zero.

Proposition 1.2 *If $1 \leq i, j, i', j' \leq n$, $i \neq i'$, $j \neq j'$, and $U = [U_{ij}]$ is a standard unitary matrix, then the following hold:*

(1) $E(|U_{ij}|^2) = \frac{1}{n}$,

(2) $E(|U_{ij}|^4) = \frac{2}{n(n+1)}$,

(3) $E(|U_{ij}|^2|U_{i'j}|^2) = E(|U_{ij}|^2|U_{ij'}|^2) = \frac{1}{n(n+1)}$,

(4) $E(|U_{ij}|^2|U_{i'j'}|^2) = \frac{1}{n^2-1}$,

(5) $E(U_{ij}U_{i'j'}\bar{U}_{ij'}\bar{U}_{i'j}) = -\frac{1}{n(n^2-1)}$.

Proof. The random variables U_{ij} are identically distributed thanks to (1.2) and hence (1) follows from $\sum_{j=1}^{n}|U_{ij}|^2 = 1$. Since the $(1,1)$ entries of U and

$$(\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \oplus I_{n-2})\,U$$

are identically distributed, we have

$$\begin{aligned} E(|U_{11}|^4) &= E(|U_{11}\cos\theta + U_{21}\sin\theta|^4) \\ &= E((|U_{11}|^2\cos^2\theta + |U_{21}|^2\sin^2\theta + (U_{11}\bar{U}_{21} + \bar{U}_{11}U_{21})\cos\theta\sin\theta)^2) \\ &= E(|U_{11}|^4)(\cos^4\theta + \sin^4\theta) + 4E(|U_{11}|^2|U_{21}|^2)\cos^2\theta\sin^2\theta\,, \end{aligned}$$

where we have used $E(|U_{11}|^2 U_{11}\bar{U}_{21}) = 0$, etc. due to Lemma 1.1. Hence $E(|U_{11}|^4) = 2E(|U_{11}|^2|U_{21}|^2) = 2E(|U_{i1}|^2|U_{i'1}|^2)$ for $i \neq i'$. Since $\sum_{i,i'=1}^n |U_{i1}|^2|U_{i'1}|^2 = 1$, this gives

$$\begin{aligned} 1 &= \sum_{i=1}^n E(|U_{i1}|^4) + \sum_{i\neq i'} E(|U_{i1}|^2|U_{i'1}|^2) \\ &= nE(|U_{11}|^4) + \frac{n(n-1)}{2}E(|U_{11}|^4) = \frac{n(n+1)}{2}E(|U_{11}|^4)\,. \end{aligned}$$

Hence $E(|U_{11}|^4) = \frac{2}{n(n+1)}$ and $E(|U_{11}|^2|U_{21}|^2) = \frac{1}{n(n+1)}$. In this way (2) and (3) are obtained. Apply (3) to $\sum_{i,i'=1}^n |U_{i1}|^2|U_{i'2}|^2 = 1$ to get (4). Proof of (5) is similar to the above and the details are left to the reader.

By the above proposition the correlation coefficients between two of the random variables $|U_{ij}|^2$ are computed as

$$\rho(|U_{ij}|^2, |U_{i'j}|^2) = \rho(|U_{ij}|^2, |U_{ij'}|^2) = -\frac{1}{n-1}\,, \quad \rho(|U_{ij}|^2, |U_{i'j'}|^2) = \frac{1}{(n-1)^2}\,,$$

where $i \neq i'$, $j \neq j'$.

Lemma 1.3 *For every $k_1, k_2, \ldots, k_n \in \mathbb{Z}^+$,*

$$E(|U_{ij}|^{2k}) = \frac{k!}{k_1!k_2!\cdots k_n!}E(|U_{11}|^{2k_1}|U_{21}|^{2k_2}\cdots|U_{n1}|^{2k_n})\,, \tag{1.4}$$

where $k := k_1 + k_2 + \cdots + k_n$.

Proof. For every $0 \leq t_1, \ldots, t_n \leq 1$ with $\sum_{i=1}^n t_i^2 = 1$, one can choose a unitary matrix V whose first row is $(t_1, \ldots, t_n)$. Since the $(1,1)$ entry of VU is $\sum_{i=1}^n t_i U_{i1}$, it follows as in the proof of Proposition 1.2 that

$$\begin{aligned} E(|U_{11}|^{2k}) &= E\Big(\Big(\Big(\sum_{i=1}^n t_i U_{i1}\Big)\Big(\sum_{i=1}^n t_i \bar{U}_{i1}\Big)\Big)^k\Big) \\ &= \sum_{k_1+\cdots+k_n=k} \Big(\frac{k!}{k_1!\cdots k_n!}\Big)^2 E(|U_{11}|^{2k_1}\cdots|U_{n1}|^{2k_n})\, t_1^{2k_1}\cdots t_n^{2k_n} \end{aligned}$$

for each $k \in \mathbb{Z}^+$. Since $\sum_{k_1+\cdots+k_n=k}(k!/k_1!\cdots k_n!)\, t_1^{2k_1}\cdots t_n^{2k_n} = 1$, one may compare the coefficients of $t_1^{2k_1}\cdots t_n^{2k_n}$ in the above to obtain (1.4).

Proposition 1.4 *For every $k \in \mathbb{Z}^+$ and $1 \leq i, j \leq n$,*

$$E(|U_{ij}|^{2k}) = \binom{n+k-1}{n-1}^{-1}. \tag{1.5}$$

Furthermore, the distribution of U_{ij} is $\frac{n-1}{\pi}(1-r^2)^{n-2} r\, dr\cdot d\theta$ ($0 \leq r \leq 1$, $0 \leq \theta \leq 2\pi$) in the polar coordinate $\zeta = re^{\mathrm{i}\theta}$.

Proof. For every $k \in \mathbb{Z}^+$ we have

$$\begin{aligned}
1 &= \sum_{i_1,\dots,i_k=1}^{n} E(|U_{i_1 1}|^2 |U_{i_2 1}|^2 \cdots |U_{i_k 1}|^2) \\
&= \sum_{\substack{k_1,\dots,k_n \ge 0 \\ k_1+\cdots+k_n=k}} \frac{k!}{k_1! k_2! \cdots k_n!} E(|U_{11}|^{2k_1} |U_{21}|^{2k_2} \cdots |U_{n1}|^{2k_n}) \\
&= E(|U_{11}|^{2k}) \times \#\{(k_1,\dots,k_n) : k_i \ge 0,\ k_1+\cdots+k_n = k\} \\
&= E(|U_{11}|^{2k}) \times \binom{n+k-1}{n-1}
\end{aligned}$$

using (1.4) for the third equality. Let μ denote the distribution supported on the unit disk given in the theorem. What remains to show is that $E(U_{11}^k \bar{U}_{11}^m) = \int \zeta^k \bar{\zeta}^m \, d\mu(\zeta)$ for all $k, m \in \mathbb{Z}^+$. Since both sides are 0 when $k \ne m$, it suffices to check that $\int |\zeta|^{2k} \, d\mu(\zeta) = \binom{n+k-1}{n-1}^{-1}$ for $k \in \mathbb{Z}^+$. But we compute

$$\begin{aligned}
\int |\zeta|^{2k} \, d\mu(\zeta) &= 2(n-1) \int_0^1 (1-r^2)^{n-2} r^{2k+1} \, dr \\
&= (n-1) \int_0^1 (1-t)^{n-2} t^k \, dt = \binom{n+k-1}{n-1}^{-1}
\end{aligned}$$

applying integration by parts repeatedly.

As a consequence of (1.5) we have

$$E(|U_{ij}|^{2k}) = O(n^{-k}) \quad (\text{as } n \to \infty), \tag{1.6}$$

which will play an important role in the next section. In fact, what we shall need in the sequel concerning the standard unitary matrix are only Lemma 1.1 and (1.6). It is worthwhile to note that the order in (1.6) is the same as the $2k$th moment of the normal distribution $N(0, 1/n)$ with variance $1/n$.

The rest of this section is a brief exposition on the almost sure convergence of standard selfadjoint Gaussian matrices. An $n \times n$ selfadjoint random matrix $H(n)$ is called a *standard selfadjoint Gaussian matrix* if $\{\operatorname{Re} H_{ij}(n) : 1 \le i \le j \le n\} \cup \{\operatorname{Im} H_{ij}(n) : 1 \le i < j \le n\}$ is an independent family of Gaussian random variables and if $E(H_{ij}(n)) = 0$ for $1 \le i \le j \le n$, $E(H_{ii}(n)^2) = 1/n$ for $1 \le i \le n$ and $E((\operatorname{Re} H_{ij}(n))^2) = E((\operatorname{Im} H_{ij}(n))^2) = 1/2n$ for $1 \le i < j \le n$. The word "standard" is used because of $\tau_n(H(n)) = 0$ and $\tau_n(H(n)^2) = 1$.

The *empirical spectral density* of a standard selfadjoint Gaussian matrix $H(n)$ is a random discrete probability measure given by

$$R_{H(n)} := \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i(n)}$$

(δ_λ is the Dirac measure at λ), where $\lambda_1(n), \lambda_2(n), \dots, \lambda_n(n)$ are the (random) eigenvalues of $H(n)$. According to a stronger form of the Wigner theorem due to [1], it is known that $R_{H(n)}$ converges in weak topology to the semicircle law

$$w_2 := \frac{1}{2\pi}\sqrt{4-x^2}\chi_{[-2,2]}\,dx$$

almost surely as $n \to \infty$. Moreover, the argument in [7] can be applied to complex selfadjoint random matrices as well, and in particular we know that

$$\lim_{n\to\infty} \|H(n)\| = 2 \quad \text{a.s.}$$

where $\|H(n)\|$ $(= \max_i |\lambda_i(n)|)$ is the operator norm of $H(n)$.

The above stated facts show that the distribution of $H(n)$ with respect to tr_n converges almost surely to w_2. In the next proposition we give a short proof using a graph in the case where we have already known that $H(n)$ has the limit distribution w_2 in expectation with respect to τ_n.

Proposition 1.5 *If $H(n)$ is an $n\times n$ standard selfadjoint Gaussian matrix, then $H(n)$ has the almost sure limit distribution w_2, that is,*

$$\lim_{n\to\infty} \mathrm{tr}_n(H(n)^k) = \frac{1}{2\pi}\int_{-2}^{2} x^k\sqrt{4-x^2}\,dx \quad a.s.$$

for every $k \in \mathbb{N}$.

Proof. As we have known the convergence of $H(n)$ with respect to τ_n, it suffices to show that

$$E\Big(\sum_{n=1}^{\infty}(\mathrm{tr}_n(H(n)^k) - \tau_n(H(n)^k))^2\Big) < +\infty \tag{1.7}$$

for any $k \in \mathbb{N}$. We write

$$\begin{aligned} E([\mathrm{tr}_n(H(n)^k) - \tau_n(H(n)^k)]^2) &= E([\mathrm{tr}_n(H(n)^k)]^2) - [\tau_n(H(n)^k)]^2 \\ &= \frac{1}{n^{k+2}}\sum Q_n(m_1,\dots,m_k;m_{k+1},\dots,m_{2k}), \end{aligned}$$

where the summation is over all $1 \le m_1,\dots,m_k,m_{k+1},\dots,m_{2k} \le n$ and

$$\begin{aligned} &Q_n(m_1,\dots,m_k;m_{k+1},\dots,m_{2k}) \\ &\quad := E(H_{m_1m_2}H_{m_2m_3}\cdots H_{m_km_1}H_{m_{k+1}m_{k+2}}H_{m_{k+2}m_{k+3}}\cdots H_{m_{2k}m_{k+1}}) \\ &\qquad -E(H_{m_1m_2}H_{m_2m_3}\cdots H_{m_km_1})E(H_{m_{k+1}m_{k+2}}H_{m_{k+2}m_{k+3}}\cdots H_{m_{2k}m_{k+1}}) \end{aligned}$$

(H_{ij} is for $H_{ij}(n)$). From the Hölder inequality one gets

$$|Q_n(m_1,\dots,m_k;m_{k+1},\dots,m_{2k})| \le C_k \tag{1.8}$$

for some $C_k < +\infty$ depending on k only. When $Q_n(m_1,\dots,m_k;m_{k+1},\dots,m_{2k})$ is nonzero, one can choose a pair partition $\mathcal{V}$ of $\{1,\dots,2k\}$, $p \in \{1,\dots,k\}$, $q \in \{k+1,\dots,2k\}$ and $\varepsilon \in \{\pm 1\}$ such that

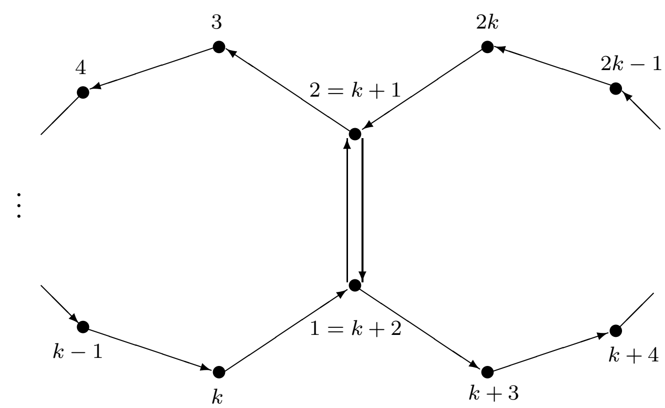
(a) $(m_i, m_{i+1}) = (m_{j+1}, m_j)$ if $\{i, j\} \in \mathcal{V}$,

(b) $(m_p, m_{p+1}) = (m_q, m_{q+1})$ if $\varepsilon = 1$ and $(m_p, m_{p+1}) = (m_{q+1}, m_q)$ if $\varepsilon = -1$.

In the above, $i+1$ for $1 \le i \le k$ (resp. $k+1 \le i \le 2k$) is understood in the cyclic order of $(1, \dots, k)$ (resp. $(k+1, \dots, 2k)$). For such $(\mathcal{V}, p, q, \varepsilon)$ we denote by $\Xi_n(\mathcal{V}, p, q, \varepsilon)$ the set of all $(m_1, \dots, m_{2k}) \in \{1, \dots, n\}^{2k}$ satisfying (a), (b). Let us show that

$$\#\Xi_n(\mathcal{V}, p, q, \varepsilon) \le n^k . \tag{1.9}$$

For this we may assume $p = 1$ and $q = k+1$ in view of the cyclicity of $(m_1, \dots, m_k)$ and $(m_{k+1}, \dots, m_{2k})$. When $\varepsilon = -1$ (hence $m_1 = m_{k+2}$ and $m_2 = m_{k+1}$), consider the graph G with $2k-2$ vertices $1\,(= k+2), 2\,(= k+1), 3, \dots, k, k+3, \dots, 2k$ and $2k$ edges $[1, 2], \dots, [k-1, k], [k, 1], [k+1, k+2], \cdots, [2k-1, 2k], [2k, k+1]$.



Picture of the graph G

Let $\tilde{G}$ be the quotient graph obtained by identifying edges $[i, i+1]$ and $[j, j+1]$ with orientation reversed if $\{i, j\} \in \mathcal{V}$. Then $\tilde{G}$ is a connected graph with k edges, and it is easy to see that $\tilde{G}$ has a loop passing through vertex 1 (even when edges $[1, 2]$ and $[k+1, k+2]$ are identified). This implies that $\tilde{G}$ has at most k vertices. We can argue in a similar way when $\varepsilon = 1$ (hence $m_1 = m_{k+1}$ and $m_2 = m_{k+2}$). In this way (1.9) is shown because the freedom in choosing $(m_1, \dots, m_{2k})$ from $\Xi_n(\mathcal{V}, p, q, \varepsilon)$ subject to (a), (b) is at most the number of vertices of $\tilde{G}$. From (1.8) and (1.9) we have

$$E([\mathrm{tr}_n(H(n)^k) - \tau_n(H(n)^k)]^2) = O(n^{-k-2}) \sum_{(\mathcal{V}, p, q, \varepsilon)} \#\Xi_n(\mathcal{V}, p, q, \varepsilon) = O(n^{-2})$$

as $n \to \infty$, which yields (1.7).

2 Standard unitary random matrices

When random matrices are concerned, besides the asymptotic freeness introduced in [22], one can consider the notion of asymptotic freeness almost everywhere as we define below. Let S be a set, and $\mathbb{C}\langle X_s | s \in S\rangle$ be the algebra of polynomials in noncommuting indeterminates X_s ($s \in S$) over $\mathbb{C}$. For $n \in \mathbb{N}$ let $(X(s,n))_{s\in S}$ be a family of $n \times n$ random matrices. It is said that $(X(s,n))_{s\in S}$ has the *limit distribution* μ (as $n \to \infty$) if μ is a distribution on $\mathbb{C}\langle X_s | s \in S\rangle$ (i.e. a linear functional $\mu : \mathbb{C}\langle X_s | s \in S\rangle \to \mathbb{C}$ with $\mu(1) = 1$) and

$$\mu(X_{s_1} X_{s_2} \cdots X_{s_m}) = \lim_{n\to\infty} \tau_n(X(s_1,n)X(s_2,n)\cdots X(s_m,n))$$

for all $s_1, \ldots, s_m \in S$. Let $\{S_j : j \in J\}$ be a partition of S. The meaning of the *asymptotic freeness* of $(\{X(s,n) : s \in S_j\})_{j\in J}$ introduced in [22] (also [25], p. 43) is that on one hand $(X(s,n))_{s\in S}$ has the limit distribution μ and on the other hand $(\mathbb{C}\langle X_s | s \in S_j\rangle)_{j\in J}$ is free in the noncommutative probability space $(\mathbb{C}\langle X_s | s \in S\rangle, \mu)$. This notion is concerned with the convergence under the tracial functionals τ_n. However, it is also natural to consider the convergence under the normalized traces tr_n almost everywhere.

Given $(X(s,n))_{s\in S}$ and $\{S_j : j \in J\}$ as above, we say that $(\{X(s,n) : s \in S_j\})_{j\in J}$ is *asymptotically free almost everywhere* if $(X(s,n))_{s\in S}$ has the (non-random) limit distribution almost surely, that is, there exists a distribution μ on $\mathbb{C}\langle X_s | s \in S\rangle$ such that

$$\lim_{n\to\infty} \mathrm{tr}_n(X(s_1,n)X(s_2,n)\cdots X(s_m,n)) = \mu(X_{s_1} X_{s_2} \cdots X_{s_m}) \quad \text{a.s.}$$

for all $s_1, \ldots, s_m \in S$, and if $(\mathbb{C}\langle X_s | s \in S_j\rangle)_{j\in J}$ is free in $(\mathbb{C}\langle X_s | s \in S\rangle, \mu)$. This is equivalent to saying that the following two conditions hold:

(i) for each $j \in J$, $(X(s,n))_{s\in S_j}$ has the (non-random) limit distribution almost surely,

(ii) for every $j_1, \ldots, j_l \in J$ with $j_1 \neq j_2 \neq \ldots \neq j_l$, if $P_r(X_{r1}, \ldots, X_{rm_r}) \in \mathbb{C}\langle X_s | s \in S_{j_r}\rangle$ satisfies

$$\lim_{n\to\infty} \mathrm{tr}_n(P_r(X(s_{r1},n), \ldots, X(s_{rm_r},n)) = 0 \quad \text{a.s.}$$

for $1 \le r \le l$, then

$$\lim_{n\to\infty} \mathrm{tr}_n\Big(\prod_{r=1}^{l} P_r(X(s_{r1},n), \ldots, X(s_{rm_r},n))\Big) = 0 \quad \text{a.s.}$$

Indeed, one can easily see by induction that (i) and (ii) together imply that the whole $(X(s,n))_{s\in S}$ has the limit distribution almost surely.

Random matrices $X(s,n)$ treated below mostly satisfy

$$\sup_n \tau_n((X(s,n)^* X(s,n))^k) < +\infty$$

for all $k \in \mathbb{N}$, and it implies that each sequence $\{\mathrm{tr}_n(\cdots)\}$ in (i) and (ii) above is uniformly integrable so that the almost everywhere convergence yields the convergence of the expectations. In this way we observe that the definition of asymptotic freeness almost everywhere is actually a stronger property than the plain asymptotic freeness (if restricted to uniformly integrable random matrices).

In the following let S, T be arbitrary sets. The main result of this section is

Theorem 2.1 *Let $(U(s,n))_{s\in S}$ be an independent family of $n \times n$ standard unitary random matrices. Let $(B(t,n))_{t\in T}$ be a family of $n \times n$ constant (i.e. non-random) matrices such that $\sup_n \|B(t,n)\| < +\infty$ ($\|\cdot\|$ being the operator norm) for each $t \in T$ and $(B(t,n), B(t,n)^*)_{t\in T}$ has the limit distribution. Then the family*

$$\Big((\{U(s,n), U(s,n)^*\})_{s\in S},\ \{B(t,n), B(t,n)^* : t \in T\}\Big) \tag{2.1}$$

is asymptotically free almost everywhere as $n \to \infty$.

Proof. As in the proof of [5], Theorem 2.1, we may assume without loss of generality that $\{(B(t,n))_{n\in\mathbb{N}} : t \in T\}$ forms a *-subalgebra of $\prod_{n\in\mathbb{N}} M_n(\mathbb{C})$. (In fact, the *-subalgebra of $\prod_{n\in\mathbb{N}} M_n(\mathbb{C})$ generated by $(B(t,n))_{n\in\mathbb{N}}$ ($t \in T$) and the identity $(I_n)_{n\in\mathbb{N}}$ may be considered as T itself.) Then it suffices to prove the following: If $s_1, \dots, s_l \in S$, $m_1, \dots, m_l \in \mathbb{Z} \setminus \{0\}$, and $t_1, \dots, t_l \in T$ are such that for each $1 \le r \le l$ either

(a) $\mathrm{tr}_n(B(t_r,n)) = 0$ ($n \in \mathbb{N}$), or

(b) $B(t_r,n) = I_n$ ($n \in \mathbb{N}$) and $s_r \ne s_{r+1}$ (with $s_{l+1} := s_1$),

then

$$E\biggl(\biggl|\mathrm{tr}_n\biggl(\prod_{r=1}^{l} U(s_r,n)^{m_r} B(t_r,n)\biggr)\biggr|^2\biggr) = O(n^{-2}) \quad (\text{as } n \to \infty). \tag{2.2}$$

(The assumption (a) can be put instead of $\lim_n \mathrm{tr}_n(B(t_r,n)) = 0$ because we may replace $B(t_r,n)$ by $B(t_r,n) - \mathrm{tr}_n(B(t_r,n))I_n$.)

When (2.2) has been proved, one obtains

$$E\biggl(\sum_{n=1}^{\infty} \biggl|\mathrm{tr}_n\biggl(\prod_{r=1}^{l} U(s_r,n)^{m_r} B(t_r,n)\biggr)\biggr|^2\biggr) < +\infty\,,$$

which implies that

$$\mathrm{tr}_n\biggl(\prod_{r=1}^{l} U(s_r,n)^{m_r} B(t_r,n)\biggr) \to 0 \ \text{ a.s.} \quad (\text{as } n \to \infty).$$

This says that the family (2.1) satisfies the condition (ii) stated above. When a single $U(s,n)$ is taken, it also follows that $(U(s,n), U(s,n)^*)$ has the almost sure limit distribution. Thus the result is concluded.

Now let us prove (2.2). Let

$$k := |m_1| + \cdots + |m_l|\,, \quad k(r) := |m_1| + \cdots + |m_r| \quad (1 \le r \le l),$$

and when $k(r-1)+1 \le h \le k(r)$ (with $k(0) := 0$) let

$$s(h) := s_r\,, \quad \varepsilon(h) := \begin{cases} 1 & \text{if } m_r > 0, \\ -1 & \text{if } m_r < 0. \end{cases}$$

Furthermore, set

$$k(l+r) := k + k(r)\,, \quad t_{l+r} := t_r \quad (1 \le r \le l),$$

$$s(k+h) := s(h)\,, \quad \varepsilon(k+h) := -\varepsilon(h) \quad (1 \le h \le k).$$

If we write

$$u_{ij}(s, \varepsilon, n) := \begin{cases} U_{ij}(s,n) & \text{if } \varepsilon = 1, \\ \bar{U}_{ji}(s,n) & \text{if } \varepsilon = -1, \end{cases}$$

then the left-hand side of (2.2) is expressed as

$$\begin{aligned} \left(\frac{1}{n}\right)^2 \sum_{i_1,\ldots,i_{2k}=1}^{n} \; \sum_{j_{k(1)},\ldots,j_{k(l)},j_{k(l+1)+1},\ldots,j_{k(2l)+1}=1}^{n} \Bigl(\prod_{r=1}^{l} B_{j_{k(r)} i_{k(r)+1}}(t_r, n)\Bigr) \\ \times \Bigl(\prod_{r=l+1}^{2l} \bar{B}_{i_{k(r)} j_{k(r)+1}}(t_r, n)\Bigr) E\Bigl(\prod_{h=1}^{2k} u_{i_h j_h}(s(h), \varepsilon(h), n)\Bigr) \end{aligned} \tag{2.3}$$

where $i_{k(l)+1} := i_1$, $j_{k(2l)+1} := j_{k+1}$, and

$$\begin{cases} j_h = i_{h+1} & \text{for } h \in \{1,\ldots,k\} \setminus \{k(1),\ldots,k(l)\}, \\ i_h = j_{h+1} & \text{for } h \in \{k+1,\ldots,2k\} \setminus \{k(l+1),\ldots,k(2l)\}. \end{cases} \tag{2.4}$$

Using the Hölder inequality and (1.6) one can estimate

$$\Bigl| E\Bigl(\prod_{h=1}^{2k} u_{i_h j_h}(s(h), \varepsilon(h), n)\Bigr)\Bigr| \le \prod_{h=1}^{2k} E(|u_{i_h j_h}(s(h), \varepsilon(h), n)|^{2k})^{1/2k} = O(n^{-k}) \tag{2.5}$$

as $n \to \infty$ uniformly for $i_1, j_1, \ldots, i_{2k}, j_{2k}$.

We want to analyze the structure of the nonzero terms of (2.3). Since $U(s,n)$ ($s \in S$) are independent, the term $E(\prod_{h=1}^{2k} u_{i_h j_h}(s(h), \varepsilon(h), n))$ becomes a product of the expectations factorized together with the same $s(h)$. When the term is nonzero, we can apply Lemma 1.1 to each factorized expectation of a product of $u_{i_h j_h}(s(h), \varepsilon(h), n))$ with the same $s(h)$, and hence for each $1 \le r \le l$ and $1 \le i, j \le n$ we have

$$\begin{aligned} \#\{h : i_h = i,\, s(h) = s_r,\, \varepsilon(h) = 1\} &= \#\{h : j_h = i,\, s(h) = s_r,\, \varepsilon(h) = -1\}\,, \\ \#\{h : j_h = j,\, s(h) = s_r,\, \varepsilon(h) = 1\} &= \#\{h : i_h = j,\, s(h) = s_r,\, \varepsilon(h) = -1\}\,. \end{aligned}$$

Thus, two pair partitions $\mathcal{U}$ and $\mathcal{V}$ can be chosen so that if $\{h, h'\} \in \mathcal{U}$ then $s(h) = s(h')$, $\varepsilon(h) = 1$, $\varepsilon(h') = -1$, $i_h = j_{h'}$, and if $\{h, h'\} \in \mathcal{V}$ then $s(h) = s(h')$, $\varepsilon(h) = -1$, $\varepsilon(h') = 1$, $i_h = j_{h'}$, where either $h < h'$ or $h > h'$ may be chosen. The pair partitions $\mathcal{U}$ and $\mathcal{V}$, together with the relations in (2.4), cause many equalities among $i_1, \dots, i_{2k}$, and they define the equivalence relation $\mathcal{R}(\mathcal{U}, \mathcal{V})$ on $\{1, \dots, 2k\}$ so that $i_h = i_{h'}$ whenever h, h' are in the same equivalence class of $\mathcal{R}(\mathcal{U}, \mathcal{V})$. Here, by assumptions (in particular, see (a), (b)) we can specify the case where a singleton $\{h\}$ is an equivalence class. In fact, if $\{h\}$ is an equivalence class of $\mathcal{R}(\mathcal{U}, \mathcal{V})$, then the following must hold: When $1 \le h \le k$, either $\{h, h-1\} \in \mathcal{U}$, $\varepsilon(h) = 1$, $\varepsilon(h-1) = -1$ or $\{h, h-1\} \in \mathcal{V}$, $\varepsilon(h) = -1$, $\varepsilon(h-1) = 1$, and for some $1 \le r \le l$

$$h = k(r) + 1\,, \quad s_r = s_{r+1}\,, \quad q_r = 0\,, \quad \mathrm{tr}_n(B(t_r, n)) = 0 \ (n \in \mathbb{N}), \tag{2.6}$$

where $h \in \{1, \dots, k\}$ and $r \in \{1, \dots, l\}$ are understood in the cyclic order. On the other hand, when $k + 1 \le h \le 2k$, either $\{h, h+1\} \in \mathcal{U}$, $\varepsilon(h) = 1$, $\varepsilon(h+1) = -1$ or $\{h, h+1\} \in \mathcal{V}$, $\varepsilon(h) = -1$, $\varepsilon(h+1) = -1$, and for some $l + 1 \le r \le 2l$

$$h = k(r)\,, \quad s_r = s_{r+1}\,, \quad q_r = 0\,, \quad \mathrm{tr}_n(B(t_r, n)) = 0 \ (n \in \mathbb{N}), \tag{2.7}$$

where $h \in \{k+1, \dots, 2k\}$ and $r \in \{l+1, \dots, 2l\}$ are understood in the cyclic order.

Now, fix pair partitions $\mathcal{U}, \mathcal{V}$ of $\{1, \dots, 2k\}$ and let $h(1), \dots, h(k_0)$ be the representatives from the equivalence classes of $\mathcal{R}(\mathcal{U}, \mathcal{V})$ where $h(1), \dots, h(l_0)$ are from the singleton equivalence classes (hence $0 \le l_0 \le k_0$). It is obvious that

$$k_0 \le l_0 + \frac{2k - l_0}{2} = k + \frac{l_0}{2}\,.$$

When $(i_1, \dots, i_{2k}) \in \{1, \dots, n\}^{2k}$ is subject to $\mathcal{R}(\mathcal{U}, \mathcal{V})$, the terms of (2.3) are determined by $(\iota_1, \dots, \iota_{k_0}) := (i_{h(1)}, \dots, i_{h(k_0)})$ so that one can set

$$\begin{aligned} \alpha_n(\iota_1, \dots, \iota_{k_0}) &:= E\Big(\prod_{h=1}^{2k} u_{i_h j_h}(s(h), \varepsilon(h), n)\Big)\,, \\ \beta_n(\iota_1, \dots, \iota_{k_0}) &:= \Big(\prod_{r=1}^{l} B_{j_{k(r)} i_{k(r)+1}}(t_r, n)\Big)\Big(\prod_{r=l+1}^{2l} \bar{B}_{i_{k(r)} j_{k(r)+1}}(t_r, n)\Big)\,, \end{aligned}$$

where $j_1, \dots, j_{2k}$ are determined subject to $\mathcal{U}, \mathcal{V}$, that is, $j_{h'} = i_h$ if $\varepsilon(h') = -1$ and $\{h, h'\} \in \mathcal{U}$, or if $\varepsilon(h') = 1$ and $\{h, h'\} \in \mathcal{V}$. Then it remains to prove that for any partition $\mathcal{W}$ of $\{1, \dots, k_0\}$ one has

$$\sum_{(\iota_1, \dots, \iota_{k_0}) : \mathcal{W}} \beta_n(\iota_1, \dots, \iota_{k_0}) \alpha_n(\iota_1, \dots, \iota_{k_0}) = O(1) \quad (\text{as } n \to \infty), \tag{2.8}$$

where the summation is over $(\iota_1, \dots, \iota_{k_0}) \in \{1, \dots, n\}^{k_0}$ such that $\iota_p = \iota_q$ if p, q are in the same block of $\mathcal{W}$ and otherwise $\iota_p \ne \iota_q$. Indeed, the sum in (2.3) can be divided into finite disjoint portions (independently of n) each of which is written as the sum in (2.8) subject to some possible triple $(\mathcal{U}, \mathcal{V}, \mathcal{W})$.

First assume $l_0 \le 1$ and so $k_0 \le k$. Since

$$\sum_{\iota_1,\dots,\iota_{k_0}=1}^{n} |\alpha_n(\iota_1,\dots,\iota_{k_0})| = n^{k_0} O(n^{-k}) = O(1)$$

thanks to (2.5), we have (2.8) for any $\mathcal{W}$.

Next assume $2 \le l_0 \le 3$ and so $k_0 \le k+1$. If $\#\mathcal{W}$ (the number of blocks of $\mathcal{W}$) $\le k_0 - 1$, then

$$\sum_{(\iota_1,\dots,\iota_{k_0}):\mathcal{W}} |\alpha_n(\iota_1,\dots,\iota_{k_0})| = n^{k_0-1} O(n^{-k}) = O(1)\,.$$

So assume $\#\mathcal{W} = k_0$ and choose $\{h\} \in \mathcal{R}(\mathcal{U},\mathcal{V})$. Then the summation is over all distinct $\iota_1,\dots,\iota_{k_0}$, and according to (2.6) and (2.7) either

$$\begin{aligned} &B_{j_{k(r)} i_{k(r)+1}}(t_r,n)\ (= B_{i_h i_h}(t_r,n)) \quad \text{(for some } 1 \le r \le l) \quad \text{or} \\ &\bar{B}_{i_{k(r)} j_{k(r)+1}}(t_r,n)\ (= \bar{B}_{i_h i_h}(t_r,n)) \quad \text{(for some } l+1 \le r \le 2l) \end{aligned}$$

appears only once in the product $\beta_n(\iota_1,\dots,\iota_{k_0})$. Hence we may write

$$\beta_n(\iota_1,\dots,\iota_{k_0}) = \tilde{B}_{\iota_1\iota_1}(t_{r_1},n)\tilde{\beta}_n(\iota_2,\dots,\iota_{k_0}) \tag{2.9}$$

for some $1 \le r_1 \le 2l$ where $\tilde{B}_{\iota\iota}(\cdot\,,\cdot)$ means $B_{\iota\iota}(\cdot\,,\cdot)$ or $\bar{B}_{\iota\iota}(\cdot\,,\cdot)$. Note that the permutation invariance (1.2) guarantees that $\alpha_n(\iota_1,\dots,\iota_{k_0})$ has a constant value for all distinct $\iota_1,\dots,\iota_{k_0}$. Therefore, we have

$$\begin{aligned} &\Big|\sum_{(\iota_1,\dots,\iota_{k_0}):\mathcal{W}} \beta_n(\iota_1,\dots,\iota_{k_0})\alpha_n(\iota_1,\dots,\iota_{k_0})\Big| \\ &\quad = \Big|\sum_{(\iota_2,\dots,\iota_{k_0}):\mathcal{W}} \Big(\sum_{\iota_1\ne\iota_2,\dots,\iota_{k_0}} \tilde{B}_{\iota_1\iota_1}(t_{r_1},n)\Big)\tilde{\beta}_n(\iota_2,\dots,\iota_{k_0})\alpha_n(\iota_1,\dots,\iota_{k_0})\Big| \\ &\quad = n^{k_0-1} O(n^{-k}) = O(1)\,, \end{aligned} \tag{2.10}$$

because

$$\sum_{\iota_1\ne\iota_2,\dots,\iota_{k_0}} B_{\iota_1\iota_1}(t_{r_1},n) = -\sum_{\iota_1=\iota_2,\dots,\iota_{k_0}} B_{\iota_1\iota_1}(t_{r_1},n) = O(1)$$

is valid thanks to $\mathrm{tr}_n(B(t_{r_1},n)) = 0$ $(n \in \mathbb{N})$.

Now assume $4 \le l_0 \le 5$ and so $k_0 \le k+2$. As above we get (2.8) for any $\mathcal{W}$ with $\#\mathcal{W} \le k_0 - 2$. When $\#\mathcal{W} = k_0 - 1$, we can choose $p \in \{1,\dots,l_0\}$ such that $\{p\}$ is a singleton block of $\mathcal{W}$. Then either $B_{\iota_p\iota_p}(t_r,n)$ or $\bar{B}_{\iota_p\iota_p}(t_r,n)$ (for some $1 \le r \le 2l$) appears only once in the product $\beta_n(\iota_1,\dots,\iota_{k_0})$. Letting $p = 1$ and writing $\beta_n(\iota_1,\dots,\iota_{k_0})$ as (2.9), we get (2.8) as in (2.10) because $\#\mathcal{W} = k_0 - 1$ yields $\#\{(\iota_2,\dots,\iota_{k_0}) : \mathcal{W}\} = O(n^{k_0-2})$. When $\#\mathcal{W} = k_0$, we can write

$$\beta_n(\iota_1,\dots,\iota_{k_0}) = \tilde{B}_{\iota_1\iota_1}(t_{r_1},n)\tilde{B}_{\iota_2\iota_2}(t_{r_2},n)\tilde{\beta}_n(\iota_3,\dots,\iota_{k_0})$$

for some $1 \le r_1, r_2 \le 2l$. Since $\alpha_n(\iota_1, \dots, \iota_{k_0})$ is constant for all distinct $\iota_1, \dots, \iota_{k_0}$, we have

$$\begin{aligned}
&\Big| \sum_{(\iota_1,\dots,\iota_{k_0}):\mathcal{W}} \beta_n(\iota_1, \dots, \iota_{k_0}) \alpha_n(\iota_1, \dots, \iota_{k_0}) \Big| \\
&\quad = \Big| \sum_{(\iota_3,\dots,\iota_{k_0}):\mathcal{W}} \Big(\sum_{\iota_2 \ne \iota_3,\dots,\iota_{k_0}} \tilde{B}_{\iota_2\iota_2}(t_{r_2}, n) \Big) \Big(\sum_{\iota_1 \ne \iota_2,\dots,\iota_{k_0}} \tilde{B}_{\iota_1\iota_1}(t_{r_1}, n) \Big) \\
&\qquad\qquad\qquad \times \tilde{\beta}_n(\iota_3, \dots, \iota_{k_0}) \alpha_n(\iota_1, \dots, \iota_{k_0}) \Big| \\
&\quad = n^{k_0-2} O(n^{-k}) = O(1)\,.
\end{aligned}$$

We can similarly proceed in the case $6 \le l_0 \le 7$ and so on, and the proof is completed.

Corollary 2.2 *Let $(U(s,n))_{s \in S}$ be an independent family of $n \times n$ standard unitary random matrices. Let $s_1, \dots, s_l \in S$, $m_1, \dots, m_l \in \mathbb{Z} \setminus \{0\}$, $R > 0$, and $B_r(n)$ $(1 \le r \le l)$ be $n \times n$ constant matrices such that for each $1 \le r \le l$ either*

(a) $\mathrm{tr}_n(B_r(n)) = 0$ *and* $\|B_r(n)\| \le R$ $(n \in \mathbb{N})$, *or*

(b) $B_r(n) = I_n$ $(n \in \mathbb{N})$ *and* $s_r \ne s_{r+1}$ (*with* $s_{r+1} := s_1$).

Then for every $1 \le p < \infty$,

$$E\Big(\Big| \mathrm{tr}_n \Big(\prod_{r=1}^{l} U(s_r, n)^{m_r} B_r(n) \Big) \Big|^p \Big)^{1/p} = O(n^{-1}) \quad (\text{as } n \to \infty) \tag{2.11}$$

uniformly for the choice of $B_r(n)$ $(1 \le r \le n)$ satisfying (a), (b) (*for given* $R > 0$). *Moreover,*

$$\mathrm{tr}_n \Big(\prod_{r=1}^{l} U(s_r, n)^{m_r} B_r(n) \Big) \to 0 \quad a.s. \quad (n \to \infty). \tag{2.12}$$

Proof. We notice that in the above proof of (2.2) for $p = 2$ the existence of the limit distribution of $(B(t,n), B(t,n)^*)_{t \in T}$ is no longer necessary and in fact the estimate (2.2) is uniform for the choice of $B(t,n)$ such that $\|B(t,n)\| \le R$ for given $R > 0$. Hence (2.11) for $p = 2$ has been proved. To prove (2.11) for general p, it is enough to assume that p is an even integer, say $2d$, because $E(|\cdot|^p)^{1/p} \le E(|\cdot|^{2d})^{1/2d}$ for $p \le 2d$. Now the above proof works for this case as well and we can get

$$E\Big(\Big| \mathrm{tr}_n \Big(\prod_{r=1}^{l} U(s_r, n)^{m_r} B_r(n) \Big) \Big|^{2d} \Big) = \Big(\frac{1}{n} \Big)^{2d} n^{dk} O(n^{-dk}) = O(n^{-2d})$$

as $n \to \infty$ where $k := |m_1| + \cdots + |m_l|$. The almost sure convergence (2.12) is a consequence of (2.11) for $p = 2$ as noted in the previous proof.

3 Unitarily invariant selfadjoint random matrices

We say that an $n \times n$ selfadjoint random matrix H is *unitarily invariant* if the distribution on $M_n(\mathbb{C})^{sa}$ of H is equal to that of the unitary transformation VHV^* for any $V \in \mathcal{V}(n)$. In particular, a standard selfadjoint Gaussian matrix is unitarily invariant. If H is unitarily invariant and f is a real continuous function on $\mathbb{R}$, then the random matrix $f(H)$ given via the functional calculus is also unitarily invariant.

In this section we prove the asymptotic freeness almost everywhere for an independent family of unitarily invariant selfadjoint random matrices together with constant matrices. To prove this we need the following technical lemma. In the case $p = 2$ this was shown in [23], Lemma 4.3 by making use of Lévy metric on the probability measures on $\mathbb{R}$. Proof for general $p \geq 1$ can be similarly done, so we omit the details.

Lemma 3.1 *For every $p \geq 1$, $R > 0$ and $\varepsilon > 0$, there exist $k_0 \in \mathbb{N}$ and $\delta > 0$ such that, for every $n \in \mathbb{N}$ and $(\xi_1, \ldots, \xi_n), (\eta_1, \ldots, \eta_n) \in \mathbb{R}^n$ with $\xi_1 \leq \ldots \leq \xi_n$ and $\eta_1 \leq \ldots \leq \eta_n$, if $|\eta_i| \leq R$ $(1 \leq i \leq n)$ and*

$$\Big| \sum_{i=1}^{n} \xi_i^k - \sum_{i=1}^{n} \eta_i^k \Big| \leq n\delta \quad (1 \leq k \leq k_0),$$

then $\sum_{i=1}^{n} |\xi_i - \eta_i|^p \leq n\varepsilon$.

In the proof of the next theorem we also use the Schatten p-norm (with respect to tr_n) $\|A\|_p := \mathrm{tr}_n(|A|^p)^{1/p}$ for $A \in M_n(\mathbb{C})$ where $1 \leq p < \infty$. Recall the inequalities $|\mathrm{tr}_n(A)| \leq \|A\|_1 \leq \|A\|_p \leq \|A\|$ and the Hölder inequality $\|AB\|_r \leq \|A\|_p \|B\|_q$ when $1/r = 1/p + 1/q$.

Theorem 3.2 *Let $(H(s,n))_{s \in S}$ be an independent family of $n \times n$ unitarily invariant selfadjoint random matrices and $(B(t,n))_{t \in T}$ be as in Theorem 2.1. If $H(s,n)$ converges in distribution (with respect to tr_n) almost surely to a compactly supported probability measure ρ_s on $\mathbb{R}$ for each $s \in S$, then the family*

$$\Big((\{H(s,n)\})_{s \in S},\ \{B(t,n), B(t,n)^* : t \in T\} \Big) \tag{3.1}$$

is asymptotically free almost everywhere as $n \to \infty$.

Proof. Take the diagonalization

$$H(s,n) = U(s,n)\Lambda(s,n)U(s,n)^*,$$

where $U(s,n)$ is a unitary random matrix and

$$\Lambda(s,n) = \mathrm{diag}(\lambda_1(s,n), \ldots, \lambda_n(s,n))$$

is a diagonal random matrix such that $\lambda_1(s,n) \leq \lambda_2(s,n) \leq \ldots \leq \lambda_n(s,n)$. We can make $(\{U(s,n), \Lambda(s,n)\})_{s \in S}$ an independent family. Choose an independent family

$(V(s,n))_{s\in S}$ of standard unitary matrices which are also independent of $U(s,n), \Lambda(s,n)$ $(s \in S)$. Then $(V(s,n)U(s,n))_{s\in S}$ becomes an independent family of standard unitary matrices, and $V(s,n)H(s,n)V(s,n)^*$ has the same distribution as $H(s,n)$ due to the unitary invariance. In this way, we may assume without loss of generality that $(U(s,n))_{s\in S}$ is an independent family of standard unitary matrices.

By assumption, for each $s \in S$ the empirical eigenvalue distribution $\frac{1}{n}\sum_{i=1}^n \delta_{\lambda_i(s,n)}$ converges in distribution to a compactly supported measure ρ_s almost surely as $n \to \infty$. We can choose (non-random) $\xi_1(s,n) \le \xi_2(s,n) \le \ldots \le \xi_n(s,n)$ in the support of ρ_s such that $\frac{1}{n}\sum_{i=1}^n \delta_{\xi_i(s,n)} \to \rho_s$ in distribution as $n \to \infty$. Now set $\Xi(s,n) := \mathrm{diag}(\xi_1(s,n),\ldots,\xi_n(s,n))$. Then for every $s \in S$ Lemma 3.1 implies that

$$\lim_{n\to\infty} \|\Lambda(s,n) - \Xi(s,n)\|_p = 0 \ \text{ a.s.} \qquad (p \ge 1). \tag{3.2}$$

For any $m \in \mathbb{N}$ and $p \ge 1$, using the Hölder inequality we get

$$\begin{aligned}&\|\Lambda(s,n)^m - \Xi(s,n)^m\|_p \\ &\quad \le \sum_{j=1}^m \|\Lambda(s,n)\|_{mp}^{m-j}\|\Lambda(s,n)-\Xi(s,n)\|_{mp}\|\Xi(s,n)\|^{j-1} \to 0 \ \text{ a.s.}\end{aligned}$$

due to (3.2). Hence for any polynomial P and $p \ge 1$ we have

$$\lim_{n\to\infty} \|P(\Lambda(s,n)) - P(\Xi(s,n))\|_p = 0 \ \text{ a.s.} \tag{3.3}$$

To prove the result, we may assume as in the proof of Theorem 2.1 that $\{(B(t,n))_{n\in\mathbb{N}} : t \in T\}$ forms a *-subalgebra of $\prod_{n\in\mathbb{N}} M_n(\mathbb{C})$. We have to prove that if $s_1,\ldots,s_l \in S$, $P_1,\ldots,P_l \in \mathbb{C}\langle X\rangle$ and $t_1,\ldots,t_l \in T$ are such that for $1 \le r \le l$

$$\lim_{n\to\infty} \mathrm{tr}_n(P_r(H(s_r,n))) = 0 \ \text{ a.s.} \tag{3.4}$$

and either

(a) $\lim_n \mathrm{tr}_n(B(t_r,n)) = 0$, or

(b) $B(t_r,n) = I_n$ $(n \in \mathbb{N})$ and $s_r \ne s_{r+1}$,

then

$$\lim_{n\to\infty} \mathrm{tr}_n\Bigl(\prod_{r=1}^l P_r(H(s_r,n))B(t_r,n)\Bigr) = 0 \ \text{ a.s.},$$

that is,

$$\lim_{n\to\infty} \mathrm{tr}_n\Bigl(\prod_{r=1}^l U(s_r,n)P_r(\Lambda(s_r,n))U(s_r,n)^*B(t_r,n)\Bigr) = 0 \ \text{ a.s.} \tag{3.5}$$

Using the Hölder inequality again we get

$$
\begin{aligned}
&\Big|\mathrm{tr}_n\Big(\prod_{r=1}^{l} U(s_r,n)P_r(\Lambda(s_r,n))U(s_r,n)^*B(t_r,n)\Big) \\
&\qquad\qquad - \mathrm{tr}_n\Big(\prod_{r=1}^{l} U(s_r,n)P_r(\Xi(s_r,n))U(s_r,n)^*B(t_r,n)\Big)\Big| \\
&\le \sum_{m=1}^{l}\Big(\prod_{r=1}^{m-1}\|P_r(\Xi(s_r,n))\|\Big)\|P_m(\Lambda(s_m,n))-P_m(\Xi(s_m,n))\|_l \\
&\qquad\qquad \times\Big(\prod_{r=m+1}^{l}\|P_r(\Lambda(s_r,n))\|_l\Big)\prod_{r=1}^{l}\|B(t_r,n)\| \\
&\to 0 \ \text{ a.s.}
\end{aligned}
$$

thanks to (3.3). On the other hand, since $\mathrm{tr}_n(P_r(\Xi(s_r,n)))\to 0$ by (3.3) and (3.4), Theorem 2.1 implies that

$$
\lim_{n\to\infty}\mathrm{tr}_n\Big(\prod_{r=1}^{l} U(s_r,n)P_r(\Xi(s_r,n))U(s_r,n)^*B(t_r,n)\Big)=0 \ \text{ a.s.}
$$

Here it should be remarked that each term $P_r(\Xi(s_r,n))$ is separated from $B(t_{r'},n)$ by $U(s_r,n)$ or $U(s_r,n)^*$, so we do not need to assume the existence of the limit distribution of $(B(t,n))_{t\in T}$ and $(\Xi(s,n))_{s\in S}$ combined. Thus (3.5) is concluded.

Proposition 1.5 and Theorem 3.2 yield

Corollary 3.3 *Let $(H(s,n))_{s\in S}$ be an independent family of $n\times n$ standard selfadjoint Gaussian matrices and $(B(t,n))_{t\in T}$ be as in Theorem 2.1. Then the family*

$$
\Big((\{H(s,n)\})_{s\in S},\ \{B(t,n),B(t,n)^*:t\in T\}\Big)
$$

is asymptotically free almost everywhere as $n\to\infty$.

Let $(a_s)_{s\in S}$ be a free family of semicircular elements (with distribution w_2) in a C^*-probability space. Then the corollary contains the almost sure convergence of the mixed moments:

$$
\mathrm{tr}_n(H(s_1,n)H(s_2,n)\cdots H(s_m,n))\to\varphi(a_{s_1}a_{s_2}\cdots a_{s_m}) \ \text{ a.s.}
$$

This result was proved independently by Thorbjørnsen [20] in a different method.

An $n\times n$ random matrix $X(n)$ is called a *standard non-selfadjoint Gaussian matrix* if $\{\mathrm{Re}\,X_{ij}(n):1\le i,j\le n\}\cup\{\mathrm{Im}\,X_{ij}(n):1\le i,j\le n\}$ is an independent family of Gaussian random variables with identical distribution $N(0,1/2n)$. It is clear that such $X(n)$ is written as $(H^{(1)}(n)+\mathrm{i}\,H^{(2)}(n))/\sqrt{2}$ where $H^{(1)}(n)$ and $H^{(2)}(n)$ are independent standard selfadjoint Gaussian matrices. Thus we know by Corollary 3.3 that the distribution of $(X(n),X(n)^*)$ converges almost surely to that of (c,c^*), where c is a (standard) circular element, that is, $c=(a+\mathrm{i}\,b)/\sqrt{2}$ with a free pair (a,b) of semicircular elements.

Corollary 3.4 *Let $(X(s,n))_{s\in S}$ be an independent family of $n\times n$ standard non-selfadjoint Gaussian matrices and $(B(t,n))_{t\in T}$ be as in Theorem* 2.1*. Then the family*

$$\Big((\{X(s,n),X(s,n)^*\})_{s\in S},\ \{B(t,n),B(t,n)^*:t\in T\}\Big)$$

is asymptotically free almost everywhere as $n\to\infty$. Moreover, the limit distribution of $(X(s,n),X(s,n)^*)_{s\in S}$ is the distribution of a free family $(c_s,c_s^*)_{s\in S}$ of circular elements.

Proof. The family $(X(s,n))_{s\in S}$ can be written as

$$X(s,n)=\frac{H^{(1)}(s,n)+\mathrm{i}\,H^{(2)}(s,n)}{\sqrt{2}}\,,$$

where $(H^{(1)}(s,n))_{s\in S}\cup(H^{(2)}(s,n))_{s\in S}$ is an independent family of $n\times n$ standard selfadjoint Gaussians. Corollary 3.3 says that

$$\Big((\{H^{(1)}(s,n)\})_{s\in S},\ (\{H^{(2)}(s,n)\})_{s\in S},\ \{B(t,n),B(t,n)^*:t\in T\}\Big)$$

is asymptotically free almost everywhere. This implies (see [25], Proposition 2.5.5 (ii)) that so is the family

$$\Big((\{H^{(1)}(s,n),H^{(2)}(s,n)\})_{s\in S},\ \{B(t,n),B(t,n)^*:t\in T\}\Big)\,,$$

which means the first assertion. The assertion on the limit distribution was already mentioned above the theorem.

The next result is essentially due to [19] where $A(n)$ and $B(n)$ are non-random. It says that the free convolution may be used to compute the limiting spectral distribution of random matrices and vice versa. (See [25], Chap. 3 for additive and multiplicative free convolutions.)

Proposition 3.5 *For $n\in\mathbb{N}$ let $A(n)$ and $B(n)$ be $n\times n$ selfadjoint random matrices, and $U(n)$ be an $n\times n$ standard unitary random matrix independent of $A(n),B(n)$. Assume that the distributions of $A(n),B(n)$ converge almost surely to compactly supported probability measures μ,ν, respectively. Then $(A(n),U(n)B(n)U(n)^*)$ is asymptotically free almost everywhere and the limit distribution of $A(n)+U(n)B(n)U(n)^*$ is the additive free convolution $\mu\boxplus\nu$. Moreover, when $A(n),B(n)\ge 0$, the limit distribution of $A(n)^{1/2}U(n)B(n)U(n)^*A(n)^{1/2}$ is the multiplicative free convolution $\mu\boxtimes\nu$.*

Proof. Write

$$\begin{aligned}A(n)&=V_1(n)\mathrm{diag}(a_1(n),\dots,a_n(n))V_1(n)^*\,,\\ B(n)&=V_2(n)\mathrm{diag}(b_1(n),\dots,b_n(n))V_2(n)^*\,,\end{aligned}$$

where $a_1(n)\le\dots\le a_n(n)$, $b_1(n)\le\dots\le b_n(n)$ and $V_1(n),V_2(n)$ are random unitaries. The assumption guarantees that $U(n)$ is independent of $V_1(n),V_2(n)$. Hence

$V_1(n)^* U(n) V_2(n)$ is still a standard unitary matrix. So we may assume that $A(n), B(n)$ are diagonal random matrices whose diagonals are ordered increasingly (though $U(n)$ is no longer independent of $A(n), B(n)$). As in the proof of Theorem 3.2 choose (non-random) $\varphi_1(n) \le \varphi_2(n) \le \ldots \le \varphi_n(n)$ in the support of μ such that $\frac{1}{n}\sum_{i=1}^n \delta_{\varphi_i(n)} \to \mu$ in distribution as $n \to \infty$, and set $\Phi(n) := \mathrm{diag}(\varphi_1(n), \ldots, \varphi_n(n))$. Similarly set $\Psi(n) := \mathrm{diag}(\psi_1(n), \ldots, \psi_n(n))$ for ν. Then Lemma 3.1 implies that $\lim_n \|A(n) - \Phi(n)\|_p = 0$ a.s. and $\lim_n \|B(n) - \Psi(n)\|_p = 0$ a.s. for all $p \ge 1$. Now the proof using Theorem 2.1 can be performed similarly to the proof of Theorem 3.2, so we omit the details. The assertions concerning the limit distributions are immediate from the asymptotic freeness.

We end this section with a few remarks.

Remark 3.6 Our discussion on Haar distributed unitary matrices in Sect. 1 can be repeated for Haar distributed orthogonal real matrices. Let $Q = [Q_{ij}]_{i,j=1}^n$ be an $n \times n$ random orthogonal matrix distributed according to the Haar probability measure on the orthogonal group $\mathcal{O}(n)$. The lemma taking the place of Lemma 1.1 is that if $i_1, \ldots, i_l, j_1, \ldots, j_l \in \{1, \ldots, n\}$ and $E(Q_{i_1 j_1} Q_{i_2 j_2} \cdots Q_{i_l j_l}) \ne 0$ then $\#\{r : i_r = i\}$ and $\#\{r : j_r = j\}$ are even for every $1 \le i, j \le n$. Also, it is not difficult to confirm $E(Q_{ij}^{2k}) = O(n^{-k})$ as $n \to \infty$ for $k \in \mathbb{N}$. Then the proof of Theorem 2.1 works well when $(U(s,n))_{s\in S}$ is replaced by an independent family $(Q(s,n))_{s\in S}$ of Haar distributed orthogonal matrices. In this way, one can show a version of Theorem 3.2 in the case where $(H(s,n))_{s\in S}$ is an independent family of real symmetric Gaussian matrices having an orthogonal invariant distribution and an almost sure limit distribution. This is the case in particular when $H(s,n)$'s are independent standard real symmetric Gaussian matrices.

Remark 3.7 In Theorem 3.2 one may ask if the plain asymptotic freeness of the family (3.1) holds under the weaker assumption that $H(s,n)$ converges in expectation (with respect to τ_n) to ρ_s for each $s \in S$. However, the following simple example shows that this is not true. Let $H(n)$ be a standard selfadjoint Gaussian matrix and ξ be a real random variable taking two values $\alpha \ne \beta$ with probability 1/2, respectively. Assume that $H(n)$ and ξ are independent. Then it is obvious that the unitarily invariant pair $(H(n), \xi I_n)$ has the limit distribution on $\mathbb{C}\langle X_1, X_2\rangle$ in expectation, for which X_1 and X_2 are not free (but independent). Here ξI_n does not have the almost sure limit distribution (note $\mathrm{tr}_n(\xi I_n) = \xi$). Thus we observe that the concept of asymptotic freeness almost everywhere should be more appropriate than the plain asymptotic freeness when unitarily invariant random matrices are concerned.

4 Bi-unitarily invariant non-selfadjoint random matrices

An $n \times n$ random matrix T is said to be *bi-unitarily invariant* if the distribution $M_n(\mathbb{C})$ of T is equal to that of $V_1 T V_2$ for any $V_1, V_2 \in \mathcal{U}(n)$. A standard non-selfadjoint

Gaussian matrix is bi-unitarily invariant. In this section we extend Corollary 3.4 to the case of bi-unitarily invariant random matrices.

The next lemma is a characterization of bi-unitarily invariant random matrices (up to distribution) in the form of polar decomposition. This may be rather known but we find no suitable reference, so the proof is given for convenience of the reader.

Lemma 4.1 *An $n \times n$ random matrix T is bi-unitarily invariant if and only if its distribution on $M_n(\mathbb{C})$ is equal to that of a random matrix of the form UH such that*

(1) *U is an $n \times n$ standard unitary random matrix,*

(2) *H is an $n \times n$ unitarily invariant positive semidefinite random matrix,*

(3) *U and H are independent.*

Proof. Let U and H be as stated in (1)–(3). For any $V_1, V_2 \in \mathcal{U}(n)$ it is clear that the distribution on $M_n(\mathbb{C})$ of $V_1(UH)V_2 = (V_1UV_2)(V_2^*HV_2)$ and UH are the same. Hence UH is bi-unitarily invariant. Conversely, assume that T is a bi-unitarily invariant random matrix defined on a probability space (Ω, P). Here we write an underlying probability space explicitly to make the proof precise. Let $T = U_0H$ be the polar decomposition with a unitary random matrix U_0 and $H = (T^*T)^{1/2}$. Remark that H is unique while U_0 is not. The bi-unitary invariance of T implies the unitary invariance of H. Now choose a standard unitary matrix V on another probability space (Ω', P') and define a unitary random matrix $U(\omega', \omega) := V(\omega')U_0(\omega)$ on $(\Omega' \times \Omega, P' \otimes P)$. It is immediate to see that the distribution of T and $VT = UH$ are the same. For any Borel sets $\Gamma \subset \mathcal{U}(n)$ and $\Xi \subset M_n(\mathbb{C})$ we have

$$\begin{aligned}(P' \otimes P)(U \in \Gamma, H \in \Xi) &= \int \Big(\int \chi_\Gamma(V(\omega')U_0(\omega))\, dP'(\omega') \Big) \chi_\Xi(H(\omega))\, dP(\omega) \\ &= \gamma_n(\Gamma) P(H \in \Xi)\,,\end{aligned}$$

where γ_n is the Haar measure on $\mathcal{U}(n)$. This shows that U is Haar distributed and U, H are independent. Hence the required properties of U, H are shown.

The notion of *R-diagonal elements* was introduced in [12]. In place of the definition we here state its characterization shown in [12], p. 155 as a lemma.

Lemma 4.2 *Let $(\mathcal{A}, \varphi)$ be a C^*-probability space such that φ is a tracial state. An element $x \in \mathcal{A}$ is R-diagonal if and only if there exist a Haar unitary u and a positive element h (in another $(\mathcal{A}', \varphi')$ such as $(\mathcal{A}, \varphi)$) such that h is free from $\{u, u^*\}$ and the distribution of (x, x^*) and (uh, hu^*) are the same.*

Theorem 4.3 *Let $(X(s,n))_{s \in S}$ be an independent family of $n \times n$ bi-unitarily invariant random matrices and $(B(t,n))_{t \in T}$ be as in Theorem 2.1. If $X(s,n)^*X(s,n)$ converges in distribution almost surely to a compactly supported measure ρ_s for each $s \in S$, then the family*

$$\Big((\{X(s,n), X(s,n)^*\})_{s \in S},\ \{B(t,n), B(t,n)^* : t \in T\} \Big)$$

is asymptotically free almost everywhere as $n \to \infty$. Moreover, the almost sure limit distribution of $(X(s,n), X(s,n)^*)_{s\in S}$ is the distribution of a free family $(x_s, x_s^*)_{s\in S}$ of R-diagonal elements where $x_s^* x_s$ has the distribution ρ_s.

Proof. According to Lemma 4.1 we may write $X(s,n) = U(s,n)H(s,n)$ where $U(s,n)$ and $H(s,n)$ are as stated in (1)–(3) of the lemma. Furthermore, as in the proof of Theorem 3.2, we may write $H(s,n) = V(s,n)\Lambda(s,n)V(s,n)^*$ where $V(s,n)$ is a standard unitary matrix and $\Lambda(s,n)$ is a diagonal random matrix with increasingly ordered diagonals. Here we can make $(U(s,n))_{s\in S} \cup (V(s,n))_{s\in S}$ an independent family. Since $H(s,n)^2 = X(s,n)^*X(s,n)$, it follows from Lemma 3.1 that $H(s,n)$ converges in distribution almost surely to the image measure of ρ_s by $t \mapsto t^{1/2}$ for each $s \in S$. Hence the method in proving Theorem 3.2 can be applied to show that

$$\Big((\{U(s,n)\})_{s\in S},\ (\{H(s,n)\})_{s\in S},\ \{B(t,n), B(t,n)^* : t \in T\}\Big)$$

is asymptotically free almost everywhere. Furthermore, this implies that $(X(s,n), X(s,n)^*)$ converges in distribution to $(u_s h_s, h_s u_s^*)$ almost surely, where u_s is a Haar unitary and h_s is a positive element (chosen in a C^*-probability space with a tracial state) such that h_s is free from $\{u_s, u_s^*\}$ and h_s^2 has the distribution ρ_s. Hence we have the conclusion by Lemma 4.2.

Lemma 4.2 is considered as the limiting form of Lemma 4.1. In fact, we see that for an element x in a C^*-probability space with a tracial state, x is R-diagonal if and only there exist bi-unitarily invariant random matrices $X(n)$ ($n \in \mathbb{N}$) such that the distribution of $(X(n), X(n)^*)$ with respect to tr_n converges almost surely to that of (x, x^*). The "if" part is included in Theorem 4.3 and the "only if" can be easily shown from Lemma 4.2 and Theorem 2.1. In this way, we observe that bi-unitarily invariant random matrices are almost everywhere random matrix models of R-diagonal elements.

References

[1] L. Arnold, On the asymptotic distribution of the eigenvalues of random matrices, *J. Math. Anal. Appl.* **20** (1967), 262–268.

[2] Z.D. Bai, Circular law, *Ann. Probab.* **25** (1997), 494–529.

[3] Z.D. Bai and Y.Q. Yin, Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix, *Ann. Probab.* **16** (1988), 1729–1741.

[4] Z.D. Bai and Y.Q. Yin, Limit of the smallest eigenvalue of a large dimensional sample covariance matrix, *Ann. Probab.* **21** (1993), 1275–1294.

[5] K. Dykema, On certain free product factors via an extended matrix model, *J. Funct. Anal.* **112** (1993), 31–60.

[6] K. Dykema, Interpolated free group factors, *Pacific J. Math.* **163** (1994), 123–135.

[7] S. Geman, A limit theorem for the norm of random matrices, *Ann. Probab.* **8** (1980), 252–261.

[8] S. Geman, The spectral radius of large random matrices, *Ann. Probab.* **14** (1986), 1318–1328.

[9] U. Haagerup and S. Thorbjørnsen, Random matrices with complex Gaussian entries, Preprint 1998/7, Institut for Matematik, Odense Universitet.

[10] C.R. Hwang, A brief survey on the spectral radius and the spectral distribution of large random matrices with i.i.d. entries, in *Random Matrices and Their Applications*, J.E. Cohen et al. (eds.), Contemp. Math. **50**, Amer. Math. Soc., 1986, pp. 145–152.

[11] V.A. Marchenko and L.A. Pastur, The distribution of eigenvalues in certain sets of random matrices, *Mat. Sb.* **72** (1967), 507–536.

[12] A. Nica and R. Speicher, R-diagonal pairs - a common approach to Haar unitaries and circular elements, in Free Probability Theory, D.V. Voiculescu (ed.), Fields Inst. Commun. **12**, Amer. Math. Soc, 1997, pp. 149–188.

[13] M.L. Mehta, *Random Matrices*, Academic Press, Boston, 1991.

[14] F. Oravecz and D. Petz, On the eigenvalue distribution of some symmetric random matrices, *Acta Sci. Math.* **63** (1997), 483–495.

[15] F. Rădulescu, The fundamental group of the von Neumann algebra of a free group with infinitely many generators is $\mathbb{R}\setminus\{0\}$, *J. Amer. Math. Soc.* **5** (1992), 517–532.

[16] F. Rădulescu, Stable equivalence of the weak closures of free groups convolution algebras, *Comm. Math. Phys.* **156** (1993), 17–36.

[17] F. Rădulescu, Random matrices, amalgamated free products and subfactors of the von Neumann algebra of a free group, of noninteger index, *Invent. Math.* **115** (1994), 347–389.

[18] D. Shlyakhtenko, Limit distributions of matrices with bosonic and fermionic entries, in *Free Probability Theory*, D.V. Voiculescu (ed.), Fields Institute Communications, Vol. 12, Amer. Math. Soc., 1997, pp. 241–252.

[19] R. Speicher, Free convolution and the random sum of matrices, *Publ. Res. Inst. Math. Sci.* **29** (1993), 731-744.

[20] S. Thorbjørnsen, Mixed moments of Voiculescu's Gaussian random matrices, Preprint 1999/6, Institut for Matematik, Odense Universitet.

[21] D. Voiculescu, Circular and semicircular systems and free product factors, in *Operator Algebras, Unitary Representations, Enveloping Algebras, and Invariant Theory*, A. Connes el al. (eds.), Birkhäuser, 1990, pp. 45–60.

[22] D. Voiculescu, Limit laws for random matrices and free products, *Invent. Math.* **104** (1991), 201–220.

[23] D. Voiculescu, The analogues of entropy and of Fisher's information measure in free probability theory, II, *Invent. Math.* **118** (1994), 411–440.

[24] D. Voiculescu, A strengthened asymptotic freeness result for random matrices with applications to free entropy, *Internat. Math. Res. Notices* **1998**, 41–63.

[25] D.V. Voiculescu, K.J. Dykema and A. Nica, *Free Random Variables*, CRM Monograph Ser., Vol. 1, Amer. Math. Soc., 1992.

[26] K.W. Wachter, The strong limits of random matrix spectra for sample matrices of independent elements, *Ann. Probab.* **6** (1978), 1–18.

[27] E.P. Wigner, Characteristic vectors of bordered matrices with infinite dimensions, *Ann. of Math.* **62** (1955), 548–564.

[28] E.P. Wigner, On the distribution of the roots of certain symmetric matrices, *Ann. of Math.* **67** (1958), 325–327.

[29] F. Xu, A random matrix model from two dimensional Yang-Mills theory, *Comm. Math. Phys.* **190** (1997), 287–307.