The Strong Asymptotic Freeness of Haar and Deterministic Matrices

B. Collins  
C. Male

Vienna, Preprint ESI 2320 (2011)  
May 31, 2011

Supported by the Austrian Federal Ministry of Education, Science and Culture
Available online at http://www.esi.ac.at
THE STRONG ASYMPTOTIC FREENESS OF HAAR AND DETERMINISTIC MATRICES

B. COLLINS AND C. MALE

Abstract. In this paper, we are interested in sequences of q-tuple of $N \times N$ random matrices having a strong limiting distribution (i.e. given any noncommutative polynomial in the matrices and their conjugate transpose, its normalized trace and its norm converge). We start with such a sequence having this property, and we show that this property pertains if the q-tuple is enlarged with independent unitary Haar distributed random matrices. Besides, the limit of norms and traces in non-commutative polynomials in the enlarged family can be computed with reduced free product construction. This extends results of one author (C. M.) and of Haagerup and Thorbjørnsen. We also show that a p-tuple of independent orthogonal and symplectic Haar matrices have a strong limiting distribution, extending a recent result of Schultz.

1. Introduction and statement of the main results

Following random matrix notation, we call GUE the Gaussian Unitary Ensemble, i.e. any sequence $(X_N)_{N \geq 1}$ of random variables where $X_N$ is an $N \times N$ selfadjoint random matrix whose distribution is proportional to the measure exp $\left(-\frac{N}{2} \text{Tr}(A^2)\right) dA$, where $dA$ denotes the Lebesgue measure on the set of $N \times N$ Hermitian matrices.

We recall for readers convenience the following definitions from free probability theory (see [3, 17]).

Definition 1.1. (1) A $C^*$-probability space $(\mathcal{A}, \tau, ||\cdot||)$ consists of a unital $C^*$-algebra $(\mathcal{A}, \tau, ||\cdot||)$ endowed with a state $\tau$, i.e. a linear map $\tau : \mathcal{A} \rightarrow C$ satisfying $\tau(1_A) = 1$ and $\tau(aa^*) \geq 0$ for all $a$ in $\mathcal{A}$. In this paper, we always assume that $\tau$ is a trace, i.e. that it satisfies $\tau(ab) = \tau(ba)$ for every $a, b$ in $\mathcal{A}$. A trace is said to be faithful if $\tau(aa^*) > 0$ whenever $a \neq 0$. An element of $\mathcal{A}$ is called a (non commutative) random variable.

(2) Let $A_1, \ldots, A_k$ be $*$-subalgebras of $\mathcal{A}$ having the same unit as $\mathcal{A}$. They are said to be free if for all $a_i \in A_{j_i}$ $(i = 1, \ldots, k, j_i \in \{1, \ldots, k\})$ such that $\tau[a_i] = 0$, one has

$$\tau(a_1 \cdots a_k) = 0$$

as soon as $j_1 \neq j_2$, $j_2 \neq j_3$, $\ldots$, $j_{k-1} \neq j_k$. Collections of random variables are said to be free if the unital subalgebras they generate are free.

(3) Let $a = (a_1, \ldots, a_k)$ be a k-tuple of random variables. The joint distribution of the family $a$ is the linear form $P \mapsto \tau[P(a, a^*)]$ on the set of polynomials in $2p$ non commutative indeterminates. By convergence in distribution, for a sequence of families of variables $(a_N)_{N \geq 1} = (a_1^{(N)}, \ldots, a_p^{(N)})_{N \geq 1}$, we mean the pointwise convergence of the map

$$P \mapsto \tau[P(a_N, a_N^*)],$$

and by strong convergence in distribution, we mean convergence in distribution, and pointwise convergence of the map

$$P \mapsto \|P(a_N, a_N^*)\|.$$
Theorem 1.2 In their seminal paper [12], Haagerup and Thorbjørnsen proved the following result. For any integer $N \geq 1$, let $X_1^{(N)}, \ldots, X_p^{(N)}$ be $N \times N$ independent GUE matrices and let $(x_1, \ldots, x_p)$ be a free semicircular system in a $C^*$-probability space with faithful state. Then, almost surely, for all polynomials $P$ in $p$ non commutative indeterminates, one has

$$\tau[p^{(N)}] = \begin{cases} 1 & \text{if } n = 0, \\ 0 & \text{otherwise.} \end{cases}$$

In their seminal paper [12], Haagerup and Thorbjørnsen proved the following result.

Theorem 1.2 ([12] The strong asymptotic freeness of independent GUE matrices). For any integer $N \geq 1$, let $X_1^{(N)}, \ldots, X_p^{(N)}$ be $N \times N$ independent GUE matrices and let $(x_1, \ldots, x_p)$ be a free semicircular system in a $C^*$-probability space with faithful state. Then, almost surely, for all polynomials $P$ in $p$ non commutative indeterminates, one has

$$\|P(X_1^{(N)}, \ldots, X_p^{(N)})\| \xrightarrow{N \to \infty} \|P(x_1, \ldots, x_p)\|,$$

where $\| \cdot \|$ denotes the operator norm in the left hand side and the $C^*$-algebra in the right hand side.

This theorem is a very deep result in random matrix theory, and had an important impact. Firstly, it had significant applications to $C^*$-algebra theory [12, 18], and more recently to quantum information theory [4, 7]. Secondly, it was generalized in many directions. Schultz [19] has shown that Theorem 1.2 is true when the GUE matrices are replaced by matrices of the Gaussian Orthogonal Ensemble (GOE) or by matrices of the Gaussian Symplectic Ensemble (GSE). Capitaine and Donati-Martin [5] and, very recently, Anderson [2] has shown the analogue for certain Wigner matrices.

An other significant extension of Haagerup and Thorbjørnsen’s result was obtained by one author (C. M.) in [15], where he managed to show that if in addition to independent GUE matrices, one also has an extra family of independent matrices with strong limiting distribution, the result still holds.

Theorem 1.3 ([15] The strong asymptotic freeness of $X_1^{(N)}, \ldots, X_p^{(N)}, Y_N$). For any integer $N \geq 1$, we consider

- a family $X_N = (X_1^{(N)}, \ldots, X_p^{(N)})$ of $N \times N$ independent GUE matrices,
- a family $Y_N = (Y_1^{(N)}, \ldots, Y_q^{(N)})$ of $N \times N$ matrices, possibly random but independent of $X_N$.

In a $C^*$-probability space $(\mathcal{A}, \cdot^*, \tau, \| \cdot \|)$ with faithful trace, we consider

- a free semicircular system $x = (x_1, \ldots, x_p),$
- a family $y = (y_1, \ldots, y_q)$ of non commutative random variables, free from $x$.

Then, if $y$ is the strong limit in distribution of $Y_N$, we have that $(x, y)$ is the strong limit in distribution of $(X_N, Y_N)$. In other words, if we assume that almost surely, for all polynomials $P$ in $2q$ non commutative indeterminates, one has

(1.1) $\tau_N[P(Y_N, Y_N^*)] \xrightarrow{N \to \infty} \tau[P(y, y^*)],$

(1.2) $\|P(Y_N, Y_N^*)\| \xrightarrow{N \to \infty} \|P(y, y^*)\|,$

then, almost surely, for all polynomials $P$ in $p + 2q$ non commutative indeterminates, one has

(1.3) $\tau_N[P(X_N, Y_N, Y_N^*)] \xrightarrow{N \to \infty} \tau[P(x, y, y^*)],$

(1.4) $\|P(X_N, Y_N, Y_N^*)\| \xrightarrow{N \to \infty} \|P(x, y, y^*)\|.$

It is natural to wonder whether the same property holds for unitary Haar matrices, instead of GUE matrices. The main result of this paper is the following theorem.

Theorem 1.4 (The strong asymptotic freeness of $U_1^{(N)}, \ldots, U_p^{(N)}, Y_N$). For any integer $N \geq 1$, we consider

- a family $U_N = (U_1^{(N)}, \ldots, U_p^{(N)})$ of $N \times N$ independent unitary Haar matrices,
- a family $Y_N = (Y_1^{(N)}, \ldots, Y_q^{(N)})$ of $N \times N$ matrices, possibly random but independent of $U_N$.

In a $C^*$-probability space $(\mathcal{A}, \cdot^*, \tau, \| \cdot \|)$ with faithful trace, we consider

- a family $u = (u_1, \ldots, u_p)$ of free Haar unitaries,
- a family $y = (y_1, \ldots, y_q)$ of non commutative random variables, free from $u$. 

...
The strong asymptotic freeness of Haar and deterministic matrices

2. Proof of Theorems 1.4 and 1.5

2.1. Idea of the proof. The cornerstone of the proof is the existence of an explicit coupling \((U_N, X_N)\) of an \(N \times N\) Haar matrix \(U_N\) and an \(N \times N\) GUE matrix \(X_N\), consisting of
- a trivial coupling of the eigenvalues of \(U_N\) and \(X_N\) (they are independent),
- a deterministic coupling of their eigenvectors \((U_N, X_N)\) are diagonalizable in a same basis), such that the relative orders of the eigenvalues of \(X_N\) and of the arguments of the eigenvalues of \(U_N\) with respect to a numeration of their eigenvectors are consistent. Such a coupling is possible thanks to the unitary invariance of the GUE law and of the Haar measure. Moreover, we can construct a function \(h_N : \mathbb{R} \to S^1\), referred as the folding map, such that almost surely one has

\[
U_N = h_N(X_N).
\]

Formally, the function \(h_N\) depends measurably on the pair \((U_N, X_N)\), but we will make a slight abuse of notation and denote it \(h_N\) (note that actually the dependence of \(h_N\) on \((U_N, X_N)\) becomes negligible as \(N \to \infty\) with probability one - this observation will be made rigorous in the proof). Recall that for a map \(f : \mathbb{C} \to \mathbb{C}\) and a normal matrix \(M = V \text{diag}(x_1, \ldots, x_N)V^*\), with \(V\) unitary, the symbol \(f(M)\) denotes the normal matrix \(V \text{diag}(f(x_1), \ldots, f(x_N))V^*\). The map \(h_N\) is not continuous and is random. It is obtained by combination of the empirical cumulative functions of the eigenvalues of \(X_N\) and of the arguments of the eigenvalues of \(U_N\) and \(X_N\) (see definition (2.5) below). The construction of \(h_N\) is quite a classical trick in probability on the real line, sometimes referred as the folding/unfolding of random variables, hence the name.

At the level of non commutative random variables, we have an analogue coupling

\[
u = h(x),
\]

between a Haar unitary \(u\) and a semicircular variable \(x\) in a \(C^*\)-probability space. The map \(h : \mathbb{R} \to S^1\) is continuous. In particular, the symbol \(h(x)\) is computed by functional calculus. If we consider \(\tilde{U}_N = h(X_N)\), we can deduce from Theorem 1.3 that \((\tilde{U}_N, Y_N)\) converges strongly to \((u, y)\) (i.e. we have the convergence of normalized trace and norm for any polynomial). This idea is used in [12, Part
8] to deduce results of \(C^\ast\)-algebra theory from the convergence of random matrices.

Now, knowing the coupling \((U_N, X_N)\) described above, it is actually possible to get directly the strong convergence for \((U_N, Y_N)\). We only have to estimate \(\|U_N - \bar{U}_N\|\). This amounts to show the uniform convergence of the empirical cumulative function of the eigenvalues of \(X_N\) and of the general inverse of the empirical cumulative function of the arguments of the eigenvalues of \(U_N\), which is obtained as a byproduct of Wigner’s theorem and Dini’s type theorems.

2.2. An almost sure coupling for random matrices. We first recall, in Proposition 2.1 below, the spectral theorem for unitary invariant random matrices, a well known result of random matrices theory.

**Proposition 2.1** (Spectral theorem for unitary invariant random matrices). Let \(M_N\) be an \(N \times N\) Hermitian or unitary random matrix whose distribution is invariant under conjugacy by unitary matrices. Then, \(M_N\) can be written \(M_N = V_N \Delta_N V_N^\ast\) almost surely, where

- \(V_N\) is distributed according to the Haar measure on the unitary group,
- \(\Delta_N\) is the diagonal matrix of the eigenvalues of \(M_N\), arranged in increasing order if \(M_N\) Hermitian, and in increasing order with respect to the set of arguments in \([-\pi, \pi]\) if \(M_N\) is unitary,
- \(V_N\) and \(\Delta_N\) are independent.

We recall a proof for the convenience of the readers. We actually use the proposition only for unitary Haar and GUE matrices, which are two cases where almost surely the eigenvalues are distinct. This fact brings slight conceptual simplifications, which nevertheless do not change the proof. Hence, we prefer to state the proposition without any restriction on the multiplicity of the matrices.

**Proof.** By reasoning conditionally, one can always assume that the multiplicities of the eigenvalues of \(M_N\) is almost surely constant. We denote by \((N_1, \ldots, N_K)\) the sequence of multiplicities when the eigenvalues are considered in the natural order in \(\mathbb{R}\) or in increasing order with respect to their argument in \([-\pi, \pi]\).

Since almost surely \(M_N\) is normal, it can be written \(M_N = \bar{V}_N \Delta_N \bar{V}_N\), where \(\bar{V}_N\) is a random unitary matrix and \(\Delta_N\) is as announced. The choice of \(\bar{V}_N\) can be made in a measurable way, for instance by requiring that the first nonzero element of each column of \(V_N\) is a positive real number.

Let \((u_1, \ldots, u_K)\) be a family of independent random matrices, independent of \((\Delta_N, \bar{V}_N)\) and such that for any \(k = 1, \ldots, K\), the matrix \(u_k\) is distributed according to the Haar measure on \(U(N_k)\), the group of \(N_k \times N_k\) unitary matrices. We set

\[
V_N = \bar{V}_N \text{ diag } (u_1, \ldots, u_K),
\]

and claim that the law of \(V_N\) depends only on the law of \(M_N\), not in the choice of the random matrix \(\bar{V}_N\). Indeed, let \(M_N = \bar{V}_N \Delta_N \bar{V}_N\) be an other decomposition, where \(\bar{V}_N\) is a unitary random matrix, independent of \((u_1, \ldots, u_K)\). The multiplicities of the eigenvalues being \(N_1, \ldots, N_K\), there exists \((v_1, \ldots, v_K)\) in \(U(N_1) \times \cdots \times U(N_K)\), independent of \((u_1, \ldots, u_K)\), such that \(\bar{V}_N = \bar{V}_N \text{ diag } (v_1, \ldots, v_K)\). Hence, we get \(V_N = \bar{V}_N \text{ diag } (u_1, \ldots, u_K) = V_N \text{ diag } (v_1 u_1, \ldots, v_K u_K)\), which is equal in law to \(V_N\). This proves the claim.

Let \(W_N\) be an \(N \times N\) unitary matrix. Then \(W_N M_N W_N^\ast = (W_N \bar{V}_N) \Delta_N (W_N \bar{V}_N)^\ast\). By the above, since \(M_N\) and \(W_N M_N W_N^\ast\) are equal in law, then \(V_N\) and \(W_N V_N\) are also equal in law. Hence \(V_N\) is Haar distributed in \(U(N)\).

It remains to show the independence between \(V_N\) and \(\Delta_N\). Let \(f : U(N) \to \mathbb{C}\) and \(g : M_N(\mathbb{C}) \to \mathbb{C}\) two bounded measurable functions such that \(g\) depends only on the eigenvalues of its entries. Then one as \(E[f(V_N)g(\Delta_N)] = E[f(V_N)g(M_N)]\). Let \(W_N\) be Haar distributed in \(U(N)\), independent of \((V_N, \Delta_N)\). Then by the invariance under unitary conjugacy of the law of \(M_N\), one has

\[
E[f(V_N)g(\Delta_N)] = E[f(W_N V_N)g(W_N M_N W_N^\ast)]
\]

\[
= E[f(W_N V_N)g(\Delta_N)]
\]

\[
= E[E[f(W_N V_N)|V_N, \Delta_N]g(\Delta_N)]
\]

\[
= E[f(W_N)]E[g(\Delta_N)] = E[f(V_N)]E[g(\Delta_N)].
\]

□
We are ready to construct the desired coupling. For the purposes of this paper, we start with a Haar unitary matrix, and then construct a GUE matrix.

Let $U_N$ be an $N \times N$ unitary Haar matrix. By Proposition 2.1, we can write $U_N = V_N \Delta_N V_N^*$, where $V_N$ is a Haar unitary matrix, independent of $\Delta_N = \text{diag} (e^{i\theta_1(N)}, \ldots, e^{i\theta_N(N)})$, and

$$-\pi \leq \theta_1(N) \leq \ldots \leq \theta_N(N) < \pi.$$  

We consider a random diagonal matrix $\tilde{\Delta}_N = \text{diag} (\lambda_1(N), \ldots, \lambda_N(N))$, independent of $(V_N, \Delta_N)$ and such that the random vector $(\lambda_1(N), \ldots, \lambda_N(N))$ has the law of the eigenvalues of a GUE matrix, sorted in increasing order. We set

$$X_N := V_N \tilde{\Delta}_N V_N^*,$$

which is a GUE matrix by Proposition 2.1. Hence the announced coupling $(U_N, X_N)$.

We now define the map $h_N$ which gives $U_N = h_N(X_N)$. In the sequel, we will omit the superscript $(N)$ and replace the notations $\lambda_1(N), \ldots, \lambda_N(N)$ by $\lambda_1, \ldots, \lambda_N$ and $\theta_1(N), \ldots, \theta_N(N)$ by $\theta_1, \ldots, \theta_N$. Let $F_{X_N} : \mathbb{R} \to [0, 1]$ be the empirical cumulative distribution function of $\{\lambda_1, \ldots, \lambda_N\}$, i.e. for all $t$ in $\mathbb{R}$,

$$F_{X_N}(t) = N^{-1} \sum_{j=1}^{N} 1_{[-\infty, \lambda_j]}(t).$$  

The eigenvalues of a GUE matrix are distinct with probability one, and $\lambda_1, \ldots, \lambda_N$ are arranged in increasing order. Then, almost surely and for any $j = 1, \ldots, N$, one has $F_{X_N}(\lambda_j) = j/N$. Remark that the push forward of the uniform measure on the spectrum of $X_N$ is the uniform measure on $\{1/N, 2/N, \ldots, 1\}$, a phenomenon sometimes referred as the unfolding trick.

Let $F_{U_N} : [-\pi, \pi] \to [0, 1]$ be the empirical cumulative distribution function of $\{\theta_1, \ldots, \theta_N\}$ (defined as in (2.3) with the $\lambda_j$’s replaced by the $\theta_j$’s). Let $F_{U_N}^{-1} : [0, 1] \to [-\pi, \pi]$ be its generalized inverse i.e. for all $s$ in $[0, 1]$,

$$F_{U_N}^{-1}(s) = \inf \{ t \in [-\pi, \pi] \mid F_{U_N}(t) \geq s \}.$$  

By the arrangement of the eigenvalues of $U_N$, for any $j = 1, \ldots, N$, one has $F_{U_N}^{-1}(j/N) = \theta_j$. Remark that the push forward of the uniform measure on $\{1/N, 2/N, \ldots, 1\}$ is the uniform measure on the spectrum of $U_N$. This step is sometimes called the folding trick.

We set the random function

$$h_N : \mathbb{R} \to S^1$$

$$t \mapsto \exp (i F_{U_N}^{-1} \circ F_{X_N}(t)).$$

By construction, almost surely for any $j = 1, \ldots, N$, one has $h_N(\lambda_j) = e^{i \theta_j}$, and hence, we get the expected relation between $U_N$ and $X_N$: almost surely one has

$$h_N(X_N) = V_N \text{diag} (h_N(\lambda_1), \ldots, h_N(\lambda_N)) V_N^* = U_N.$$  

In the following, we call $h_N$ the folding map associated to the coupling $(U_N, X_N)$.

### 2.3. A coupling for noncommutative random variables.

Let $F_x : \mathbb{R} \to [0, 1]$ be the cumulative distribution function of the semicircular law with radius two, i.e. for all $t$ in $\mathbb{R}$,

$$F_x(t) = \int_{-\infty}^{t} \frac{1}{2\pi} \sqrt{4-y^2} dy.$$  

Let $F_u^{-1} : [0, 1] \to [-\pi, \pi]$ be the inverse of the cumulative distribution function of the Lebesgue measure on $[-\pi, \pi]$, i.e. for all $s$ in $[0, 1]$,

$$F_u^{-1}(s) = 2\pi \left( s - \frac{1}{2} \right).$$

We define the continuous function

$$h : \mathbb{R} \to S^1$$

$$t \mapsto \exp (i F_u^{-1} \circ F_x(t)).$$
By construction, the push forward of the semicircular law with radius two is the uniform measure on the unit circle. Let $u$ be a Haar unitary and $x$ be a semicircular variable in a $\mathcal{C}^*$-probability space $(\mathcal{A}, \tau, \|\cdot\|)$ (we do not care about the possible relation between $u$ and $x$). Let $y$ be a family of non-commutative random variables in $\mathcal{A}$, free from $u$ and $x$. Then, one has the equality in non-commutative law

\begin{equation}
(h(x), y) \overset{\text{ex.}}{=} (u, y),
\end{equation}

In other words, for any polynomial $P$ in $2+q$ non-commutative indeterminates, one has $\tau[P(h(x), h(x)^*, y)] = \tau[P(u, u^*, y)]$ and then $\|P(h(x), h(x)^*, y)\| = \|P(u, u^*, y)\|$ if $\tau$ is faithful. The symbol $h(x)$ is computed by functional calculus [see [17, Lecture 3]].

2.4. Proof of Theorem 1.4. Let $\mathbf{U}_N, \mathbf{Y}_N, \mathbf{u}, \mathbf{y}$ be as in Theorem 1.4. Without loss of generality, one can assume that the matrices $\mathbf{Y}_N$ are Hermitian, at the possible cost of replacing the collection of matrices by the collection their real and imaginary parts.

Let $X_N = (X_1^{(N)}, \ldots, X_p^{(N)})$ be a family of independent $N \times N$ GUE matrices such that

- $(U_1^{(N)}, X_1^{(N)}), \ldots, (U_p^{(N)}, X_p^{(N)}), Y_N$ are independent,

- for any $j = 1, \ldots, p$, $(U_j^{(N)}, X_j^{(N)})$ is a coupling constructed by the method of Section 2.2, whose folding map is denoted $h_j^{(N)}$.

Let $h$ the function defined in Section 2.3 by formula (2.9). For any $j = 1, \ldots, p$, we set the $N \times N$ unitary random matrix $U_j^{(N)} = h(X_j^{(N)})$. We denote $\tilde{U}_N = (U_1^{(N)}, \ldots, U_p^{(N)})$. Theses matrices are not Haar distributed: for instance, as it is noticed in [12, Remark 8.3], the matrix $U_1^{(N)}$ is the identity matrix with (small but) nonzero probability. Nevertheless, it is a known consequence of Theorem 1.3 that the family of matrices $\tilde{U}_N$ converges strongly to the family $\mathbf{u}$ of free Haar unitaries (see [12, Section 8]). We only need here the norm convergence, and we recall a proof for the convenience of the readers.

Lemma 2.2. Almost surely, for every polynomial $P$ in $2+q$ non-commutative indeterminates, one has

$$\|P(\tilde{U}_N, \tilde{U}_N, Y_N)\| \xrightarrow{N \to \infty} \|P(\mathbf{u}, \mathbf{u}^*, Y_N)\|,$$

where $\tilde{U}_N = (h(X_1^{(N)}), \ldots, h(X_p^{(N)}))$.

We shall need the following lemma.

Lemma 2.3. Let $a = (a_1, \ldots, a_n)$ and $b = (b_1, \ldots, b_n)$ be families of elements in a $\mathcal{C}^*$-algebra $(\mathcal{A}, \|\cdot\|)$. Denote $D = \sup(\|a_1\|, \ldots, \|a_n\|, \|b_1\|, \ldots, \|b_n\|, 1)$. Then for every polynomial $P$ in $n$ non-commutative indeterminates one has

$$\|P(a) - P(b)\| \leq \beta D^{\alpha-1}\sum_{i=1}^{n} |a_i - b_i|,$$

where the constant $\beta$ depends only on $P$ and $\alpha$ is the total degree of $P$.

Proof of Lemma 2.3. It is sufficient to show that there exist $\beta$ such that, for any $a, b, c = (c_1, \ldots, c_{n-1})$ in $\mathcal{A}$, with $D = \sup(\|a\|, \|b\|, \|c_1\|, \ldots, \|c_{n-1}\|)$, one has

$$\|P(a, c) - P(b, c)\| \leq \beta D^{\alpha-1}\|a - b\|,$$

and then apply $n$ times this fact. Moreover, it is sufficient to show this inequality when $P$ is a monic monomial, of positive degree in the first indeterminate. For such a polynomial $P$, there exist two monic monomial $L$ and $R$ such that $P(a, c) = L(c)aR(a, c)$, $P(b, c) = L(c)bR(b, c)$. Then, one has

$$\|P(a, c) - P(b, c)\| \leq \|L(c)\|\|aR(a, c) - bR(b, c)\| \leq D^{\alpha-1}\|a - b\| + \|L(c)\|\|aR(a, c) - bR(b, c)\| \leq \|L(c)\||\|aR(a, c) - bR(b, c)\|.$$

By induction on the degree of the monomials, we get the result. 

Proof of Lemma 2.2. In the following we use the notation $f(a) = (f(a_1), \ldots, f(a_k))$ whenever $a = (a_1, \ldots, a_k)$ is a family of normal elements of a $\mathcal{C}^*$-algebra and $f : \mathbb{C} \to \mathbb{C}$ a continuous map. For
any \( \varepsilon > 0 \), let \( h_\varepsilon \) be a polynomial such that \( |h(x) - h_\varepsilon(x)| \leq \varepsilon \) for all \( x \) in \([-3, 3]\). For any polynomial \( P \) in \( 2p + q \) non commutative indeterminates, one has

\[
\left\| P(\tilde{U}_N, \tilde{U}_N^*, Y_N) - P(u, u^*, y) \right\| = \left\| P(h(X_N), \tilde{h}(X_N), Y_N) - P(h(x), \tilde{h}(x), y) \right\| \\
\leq \left\| P(h(X_N), \tilde{h}(X_N), Y_N) - P(h_\varepsilon(X_N), \tilde{h}_\varepsilon(X_N), Y_N) \right\| \\
+ \left\| P(h_\varepsilon(X_N), \tilde{h}_\varepsilon(X_N), Y_N) - P(h_\varepsilon(x), \tilde{h}_\varepsilon(x), y) \right\| \\
+ \left\| P(h_\varepsilon(x), \tilde{h}_\varepsilon(x), y) - P(h_\varepsilon(x), \tilde{h}_\varepsilon(x), y) \right\|
\]

By Theorem 1.3, one has almost surely

\[
\left\| P(h_\varepsilon(X_N), \tilde{h}_\varepsilon(X_N), Y_N) - P(h_\varepsilon(x), \tilde{h}_\varepsilon(x), y) \right\| \xrightarrow{N \to \infty} 0.
\]

On the other hand, by Lemma 2.3, we have almost surely

\[
(2.11) \quad \left\| P(h(X_N), \tilde{h}(X_N), Y_N) - P(h_\varepsilon(X_N), \tilde{h}_\varepsilon(X_N), Y_N) \right\| \leq C \sum_{j=1}^p \left\| h(X_j^{(N)}) - h_\varepsilon(X_j^{(N)}) \right\|
\]

\[
(2.12) \quad \left\| P(h(x), \tilde{h}(x), y) - P(h_\varepsilon(x), \tilde{h}_\varepsilon(x), y) \right\| \leq C \sum_{j=1}^p \left\| h(x_j) - h_\varepsilon(x_j) \right\|
\]

where \( C \) is a constant that only depends on \( P \) and on a (random) bound \( D \) such that for any \( j = 1, \ldots, q \), one has \( \|Y_j^{(N)}\| \leq D \). By Theorem 1.2, almost surely there exists \( N_0 \) such that for any \( N \geq N_0 \) and \( j = 1, \ldots, p \), one has \( \|X_j^{(N)}\| \leq 3 \). Moreover, the support of the semicircular distribution is \([-2, 2]\). Then, almost surely for \( N \) large enough, the two quantities (2.11) and (2.12) are bounded by \( C\varepsilon \). Hence, we have shown that almost surely,

\[
(2.13) \quad \| P(\tilde{U}_N, \tilde{U}_N^*, Y_N) \| \xrightarrow{N \to \infty} \| P(u, u^*, y) \|.
\]

Since a countable intersection of probability one sets is again of probability one, we get that almost surely, (2.13) holds for all polynomials \( P \) with coefficients in \( \mathbb{Q} \). Both sides in (2.13) are continuous in \( P \), hence we obtain the expected result by density of polynomials with rational coefficients.

Let \( P \) be a polynomial in \( 2p + q \) non commutative indeterminates. We want to show that: almost surely one has

\[
\| P(U_N, U_N^*, Y_N) \| \xrightarrow{N \to \infty} \| P(u, u^*, y) \|,
\]

which will be enough to show Theorem 1.4 by the same reasoning as in the end of the proof of Lemma 2.2. We set the random variable

\[
\varepsilon_N = \| P(\tilde{U}_N, \tilde{U}_N^*, Y_N) \| - \| P(u, u^*, y) \|,
\]

which tends to zero almost surely by Lemma 2.2. Now, one has by Lemma 2.3

\[
(2.14) \quad \left\| P(U_N, U_N^*, Y_N) - P(u, u^*, y) \right\| \leq \left\| P(U_N, U_N^*, Y_N) - P(\tilde{U}_N, \tilde{U}_N^*, Y_N) \right\| + \varepsilon_N
\]

\[
(2.15) \quad \leq C \sum_{j=1}^p \| U_j^{(N)} - \tilde{U}_j^{(N)} \| + \varepsilon_N,
\]

where \( C \) is a constant that only depends on \( P \) and on a bound \( D \) such that for any \( j = 1, \ldots, q \), one has \( \|Y_j^{(N)}\| \leq D \). It remains to show that, for any \( j = 1, \ldots, p \), almost surely \( \| U_j^{(N)} - \tilde{U}_j^{(N)} \| \) tends to zero as \( N \) goes to infinity. For any \( j = 1, \ldots, p \), recall that almost surely

\[
U_j^{(N)} = h_j^{(N)}(X_j^{(N)}), \quad \tilde{U}_j^{(N)} = h(X_j^{(N)}),
\]
where $h_j^{(N)}$ is the folding map associated to the coupling $(U_j^{(N)}, X_j^{(N)})$ and $h$ is given by formula (2.9). For any $j = 1, \ldots, p$, we denote by $\lambda_1(j), \ldots, \lambda_N(j)$ the eigenvalues of $X_j^{(N)}$. Hence, one has

$$\|U_j^{(N)} - \tilde{U}_j^{(N)}\| = \|h_j^{(N)}(X_j^{(N)}) - h(X_j^{(N)})\| = \left\| \exp(i F_{U_j^{(N)}}^{-1}(X_j^{(N)})) - \exp(i F_{\tilde{U}_j^{(N)}}^{-1}(X_j^{(N)})) \right\|$$

$$\leq \sup_{n=1, \ldots, N} \left| \exp(i F_{U_j^{(N)}}^{-1}(X_j^{(N)})) - \exp(i F_{\tilde{U}_j^{(N)}}^{-1}(X_j^{(N)})) \right|$$

$$\leq \sup_{n=1, \ldots, N} \left| F_{U_j^{(N)}}^{-1}(X_j^{(N)})(\lambda_n(j)) - F_{\tilde{U}_j^{(N)}}^{-1}(X_j^{(N)})(\lambda_n(j)) \right|$$

$$\leq \sup_{n=1, \ldots, N} \left| F_{U_j^{(N)}}^{-1}(X_j^{(N)})(\lambda_n(j)) - F_{\tilde{U}_j^{(N)}}^{-1}(X_j^{(N)})(\lambda_n(j)) \right|$$

$$+ \sup_{n=1, \ldots, N} \left| F_{U_j^{(N)}}^{-1}(X_j^{(N)})(\lambda_n(j)) - F_{\tilde{U}_j^{(N)}}^{-1}(X_j^{(N)})(\lambda_n(j)) \right|$$

(2.16)

$$< \left\| F_{U_j^{(N)}}^{-1} - F_{\tilde{U}_j^{(N)}}^{-1} \right\|_{L^\infty([0,1])} + 2\pi \left\| F_{X_j^{(N)}}^{-1} - F_{\tilde{X}_j^{(N)}}^{-1} \right\|_{L^\infty([0,1])}.$$ 

We shall need two lemmas in order to conclude the proof. The first one is famous in real analysis and is known as Dini's lemma.

**Lemma 2.4.** For any $n$ in $\mathbb{N} \cup \{\infty\}$, let $f_n : \mathbb{R} \to [0, 1]$ be a non decreasing function such that $\lim_{x \to -\infty} f_n(x) = 0$ and $\lim_{x \to +\infty} f_n(x) = 1$. Assume that $f_\infty$ is continuous and that $f_n$ converges pointwise to $f_\infty$ on $\mathbb{R}$. Then $f_n$ converges uniformly to $f_\infty$ on $\mathbb{R}$.

**Proof.** Let $\varepsilon > 0$. We set $K$ the ceiling of $2/\varepsilon$. For any $j = 1, \ldots, K - 1$, we set $x_j = f_\infty^{-1}(\frac{j}{K})$, where $f_\infty^{-1}$ denotes the generalized inverse of $f_\infty$ defined as in (2.4). We also set $x_0 = -\infty$ and $x_K = +\infty$. In the following we use the convention $f_\infty(-\infty) = f_\infty(-\infty) = 0$ and $f_\infty(+\infty) = f_\infty(+\infty) = 1$. By the pointwise convergence of $f_n$ to $f_\infty$ at the points $x_1, \ldots, x_{K-1}$: there exists $n_0$ such that for any $n \geq n_0$ and $j = 1, \ldots, K - 1$, one has

$$|f_n(x_j) - f_\infty(x_j)| \leq \frac{\varepsilon}{2}.$$  

(2.17)

Let $n \geq n_0$. For any $x$ in $\mathbb{R}$, let $j$ in $\{0, \ldots, K\}$ such that $x_j \leq x \leq x_{j+1}$. Since the functions are non decreasing, one has $f_n(x_i) - f_\infty(x_{i+1}) \leq f_n(x) - f_\infty(x) \leq f_n(x_{i+1}) - f_\infty(x_i)$, and so, by (2.17), we get

$$-\frac{\varepsilon}{2} - f_\infty(x_i) + f_n(x_{i+1}) \leq f_n(x) - f_\infty(x) \leq \frac{\varepsilon}{2} + f_\infty(x_{i+1}) - f_\infty(x_i).$$

The continuity of $f_\infty$ implies that $f_\infty(x_i) = i/K$. Hence we get $|f_n(x) - f_\infty(x)| \leq 1/K + \varepsilon/2 \leq \varepsilon$. \hfill $\Box$

**Lemma 2.5.** For any $n$ in $\mathbb{N} \cup \{\infty\}$, let $f_n : [a, b] \to [0, 1]$ be a non decreasing function. Assume that $f_\infty$ is differentiable in $[a, b]$, its derivative is positive and $f_n$ converges uniformly to $f_\infty$ as $n$ goes to infinity. Then $f_n^{-1}$ converges uniformly to $f_\infty^{-1}$ as $n$ goes to infinity, where $f^{-1}$ stands for the generalized inverse of $f_n$, defined as in (2.4).

**Proof.** It is sufficient to prove the pointwise convergence of $f_n^{-1}$ to $f_\infty^{-1}$. Indeed, $f_n^{-1}$ is continuous on $[0, 1]$. So, the pointwise convergence granted, we can extend for any $n$ in $\mathbb{N} \cup \{\infty\}$ the map $f_n^{-1}$ on $\mathbb{R}$ by $f_n(x) = a$ if $x < 0$ and $f_n(x) = b$ if $x > 1$, and then apply Lemma 2.4 to $(f_n^{-1} - a)/(b - a)$.

Let $\alpha > 0$ such that $f'_\infty(x) \geq \alpha$ for any $x$ in $[a, b]$. By the mean value theorem, we get that for any $\varepsilon > 0$

$$U_\varepsilon := \left\{(x, y) \in [a, b] \times [0, 1] \left| y - f_\infty(x) \leq \varepsilon \right. \right\} \subseteq V_\varepsilon := \left\{(x, y) \in [a, b] \times [0, 1] \left| x - f_\infty^{-1}(y) \leq \frac{\varepsilon}{\alpha} \right. \right\}.$$  

(2.18)

Let $\varepsilon > 0$. By the uniform convergence, there exists $n_0$ such that for any $n \geq n_0$, the graph of $f_n$ is contained in $U_{\alpha\varepsilon}$. Let $n \geq n_0$ and $t$ in $[0, 1]$. If $f_n^{-1}(t)$ is a point of continuity for $f_n$, then $f_n \circ f_n^{-1}(t) = t$. So $(f_n^{-1}(t), t)$ is in the graph of $f_n$ and it belongs to $U_{\alpha\varepsilon}$.

Otherwise, denote by $t_1$, respectively $t_2$, the left limit, respectively the right limit, of $f_n$ in $f_n^{-1}(t)$. These limits exist since $f_n$ is non decreasing. By definition of the generalized inverse, $t$ belongs to the interval $[t_1, t_2]$. Moreover, the vertical sections of $U_{\alpha\varepsilon}$ are convex. Hence, if we show that $(f_n^{-1}(t), t_1)$ and $(f_n^{-1}(t), t_2)$ are in $U_{\alpha\varepsilon}$, we get that $(f_n^{-1}(t), t)$ also belongs to this set. Since $f_\infty$ is continuous then $U_{\alpha\varepsilon}$ is closed in $\mathbb{R}^2$. On the other hand, we can find $\eta > 0$ arbitrary small such that $f_n^{-1}(t) - \eta$ is a
point of continuity for \( f_n \), and hence \( (f_n^{-1}(t) - \eta, f_n(f_n^{-1}(t) - \eta) \) belongs to \( U_{0\varepsilon} \). As \( \eta \) goes to zero, \( (f_n^{-1}(t) - \eta, f_n(f_n^{-1}(t) - \eta) \) converges to \((f_n^{-1}(t), t_1)\) and hence \((f_n^{-1}(t), t_1)\) belongs to \( U_{0\varepsilon} \). With the same reasoning with \( t_2 \), we get as expected that \((f_n^{-1}(t), t)\) in \( U_{0\varepsilon} \). Hence by (2.18) we obtain that \((f_n^{-1}(t), t)\) belongs to \( \mathcal{V}_\varepsilon \), i.e. \(|f_n^{-1}(t) - f_n^{-1}(t)| \leq \varepsilon\). \hfill \Box

By Wigner’s theorem [10, Theorem 1.13], almost surely the empirical eigenvalue distribution of \( X_j^{(N)} \) converges to the semicircular law with radius two, and hence \( F_{X_j^{(N)}} \) converges pointwise to \( F_2 \). By Lemma 2.4, we get that almost surely \( \|F_{X_j^{(N)}} - F_2\|_{L^\infty([0,1])} \) goes to zero as \( N \) goes to infinity.

Similarly, almost surely the empirical eigenvalue distribution of \( U_j^{(N)} \) converges to the uniform measure on the unit circle [3, Theorem 5.4.10]. Hence we get that almost surely \( \|F_{U_j^{(N)}} - F_2\|_{L^\infty([0,1])} \) tends to zero and by Lemma 2.5 we have that almost surely \( \|F_{U_j^{(N)}}^{-1} - F_{U_j}^{-1}\|_{L^\infty([0,1])} \) goes to zero as \( N \) goes to infinity.

Hence, by (2.15) and (2.16) we obtain that: for any polynomial \( P \), almost surely one has

\[
\lim_{N \to \infty} \|P(U_N, U_N, Y_N)\| = \|P(u, u^*, y)\|,
\]

which completes the proof.

2.5. **Proof of Theorem 1.5.** The proof of Theorem 1.5 is obtained by changing the words unitary, Hermitian and GUE into orthogonal, symmetric and GOE, respectively symplectic, self dual and GSE, by taking \( Y_N = 0 \) and citing the main results of [19] instead of Theorem 1.3. In the symplectic case, we also have to consider matrices of even size.

### 3. Applications

Our main result has the potential for many applications in random matrix theory.

#### 3.1. The spectrum of the sum and the product of Hermitian random matrices.

**Corollary 3.1.** Let \( A_N, B_N \) be two \( N \times N \) independent Hermitian random matrices. Assume that:

1. the law of one of the matrices is invariant under unitary conjugacy,
2. almost surely, the empirical eigenvalue distribution of \( A_N \) (respectively \( B_N \)) converges to a compactly supported probability measure \( \mu \) (respectively \( \nu \)),
3. almost surely, for any neighborhood of the support of \( \mu \) (respectively \( \nu \)), for \( N \) large enough, the eigenvalues of \( A_N \) (respectively \( B_N \)) belong to the respective neighborhood.

Then, one has

- almost surely, for \( N \) large enough, the eigenvalues of \( A_N + B_N \) belong to a small neighborhood of the support of \( \mu \boxplus \nu \), where \( \boxplus \) denotes the free additive convolution (see [17, Lecture 12]).
- if moreover \( B_N \) is nonnegative, then the eigenvalues of \((B_N)^{1/2}A_N(B_N)^{1/2}\) belong to a small neighborhood of the support of \( \mu \boxtimes \nu \), where \( \boxtimes \) denotes the free multiplicative convolution (see [17, Lecture 14]).

Corollary 3.1 can be applied in the following situation. Let \( A_N \) be an \( N \times N \) Hermitian random matrix whose law is invariant under unitary conjugacy. Assume that, almost surely, the empirical eigenvalue distribution of \( A_N \) converges to a compactly supported probability measure \( \mu \) and its eigenvalues belong to the support of \( \mu \) for \( N \) large enough. Let \( \Pi_N \) be the matrix of the projection on first \( p_N \) coordinates, \( \Pi_N = \text{diag}(1_{p_N}, 0_{N-p_N}) \), where \( p_N \sim tN, t \in (0,1) \). We consider the empirical eigenvalue distribution \( \mu_N \) of the Hermitian random matrix

\[
\Pi_N A_N \Pi_N.
\]

Then, it follows from a Theorem of Voiculescu [21] (see also [6]) that almost surely \( \mu_N \) converges weakly to the probability measure \( \mu^{(t)} = \mu \boxtimes [(1-t)\delta_0 + t\delta_1] \). This distribution is important in free probability theory because of its close relationship to the free additive convolution semigroup (see [17, Exercise 14.21]). Besides, the eigenvalue counting measure \( \mu_N \) was proved to be a determinantal point process obtained as the push forward of a uniform measure in a Gelfand-Cetlin cone [9]. Very recently, it was proved by Metcalfe [16] that the eigenvalues satisfy universality property inside the bulk of the spectrum.
Our result complement his, by showing that almost surely, for $N$ large enough there is no eigenvalue outside of any neighborhood of the spectrum of $\mu^{(t)}$.

**Proof of Corollary 3.1.** Without loss of generality, assume that the law of $A_N$ is invariant under unitary conjugacy. Let $D_1^{(N)} = \text{diag} (\lambda_1^{(N)}, \ldots, \lambda_N^{(N)})$ be the diagonal matrix whose entries are the eigenvalues of $B_N$, sorted in non decreasing order. For any $\rho$ in $[0,1]$, we set

$$D_1^{(N)}(\rho) = \text{diag} (\lambda_1^{(N)}_{1+\lfloor \rho N \rfloor}, \ldots, \lambda_N^{(N)}_{N+\lfloor \rho N \rfloor}),$$

with indices modulo $N$.

By the spectral theorem, we can write

$$B_N = V_N(\rho) D_1^{(N)}(\rho) V_N(\rho)^*,$$

where $V_N(\rho)$ is unitary, $(V_N(\rho), D_1^{(N)}(\rho))$ being independent of $A_N$. The law of the Hermitian matrix $V_N(\rho)^* A_N V_N(\rho)$ is still invariant under unitary conjugacy. Then, by Proposition 2.1, we can write $V_N(\rho)^* A_N V_N(\rho) = U_N D_2^{(N)} U_N^*$, where $U_N$ is a Haar unitary matrix, $D_2^{(N)}$ is a real diagonal matrix whose entries are non decreasing along the diagonal, $U_N, D_1^{(N)}, D_2^{(N)}$ are independent.

By [15, Corollary 2.1], there exists $\rho$ in $[0,1]$ such that, almost surely, the non commutative law of $(D_1^{(N)}(\rho), D_2^{(N)})$ converges strongly to the law of a couple of non commutative random variables $(d_1, d_2)$ in a $C^*$-probability space $(\mathcal{A}, \cdot, \tau, \|\cdot\|)$ with faithful trace. Let $u$ be a Haar unitary in $\mathcal{A}$, free from $(d_1, d_2)$.

By Theorem 1.4, we get that almost surely $U_N D_1^{(N)}(\rho) U_N^* + D_2^{(N)}$ converges strongly to $u d_1 u^* + d_2$. The spectrum of $A_N + B_N$ being the spectra of $U_N D_1^{(N)}(\rho) U_N^* + D_2^{(N)}$, we get the first point of Corollary 3.1 since strong convergence of random matrices implies convergence of the support.

We get the second point of Corollary 3.1 with the same reasoning on $((D_1^{(N)}(\rho))^{1/2}, D_2^{(N)})$. The application stated after Corollary 3.1 follows by taking $\Pi_N = B_N$, which satisfies the assumptions since $t \in (0,1)$, and remarking that $\Pi_N^{1/2} = \Pi_N$.

3.2. Questions from operator space theory. The following question was raised by Gilles Pisier to one author (B.C.) ten years ago: Let $U_1^{(N)}, \ldots, U_p^{(N)}$ be $N \times N$ independent unitary Haar random matrices. Is it true that

$$(3.1) \quad \left\| \sum_{i=1}^p U_i^{(N)} \right\| \xrightarrow{N \to \infty} 2\sqrt{p-1}$$

almost surely? This question is very natural from the operator space theory point of view, and although at least ten years old, it was still open before this paper. Our main theorem implies immediately that the answer is positive since $2\sqrt{p-1}$ is the norm of the sum of $p$ free Haar unitaries, a computation that goes back to a paper of Akemann and Ostrand [1]. We can give some generalizations of (3.1).

From [1], we can deduce more generally that for any complex numbers $a_1, \ldots, a_p$, almost surely one has

$$\left\| \sum_{i=1}^p a_i U_i^{(N)} \right\| \xrightarrow{N \to \infty} \min_{t \geq 0} \left\{ 2t + \sum_{i=1}^p \left( \sqrt{t^2 + |a_i|^2} - t \right) \right\}.$$ 

By a result of Kesten [14], the norm of the sum of $p$ free Haar unitaries and of their conjugate equals $2\sqrt{2k-1}$. Hence, we get from our result that almost surely one has

$$\left\| \sum_{i=1}^p (U_i^{(N)} + U_i^{(N)*}) \right\| \xrightarrow{N \to \infty} 2\sqrt{2p-1}.$$ 

Furthermore, recall that from Theorem 1.4 we can deduce the following corollary (see [15, Proposition 7.3] for a proof). We use the notations of Theorem 1.4.

**Corollary 3.2.** Let $k \geq 1$ be an integer. For any polynomial $P$ with coefficients in $\mathbb{M}_k(\mathbb{C})$, almost surely one has

$$\|P(U_N, U_N^*, Y_N, Y_N^*)\| \xrightarrow{N \to \infty} \|P(u, u^*, y, y^*)\|,$$

where $\| \cdot \|$ stands in the left hand side for the operator norm in $\mathbb{M}_k(N)(\mathbb{C})$ and in the right hand side for the $C^*$-algebra norm in $\mathbb{M}_k(\mathcal{A})$. 
By Corollary 3.2 and Fell’s absorption principle [18, Proposition 8.1], we can answer the question asked by Pisier in [18, Chapter 20]: for any \( k \times k \) unitary matrices \( a_1, \ldots, a_p \), almost surely one has
\[
\left\| \sum_{i=1}^p a_i \otimes U_i^{(N)} \right\| \xrightarrow{N \to \infty} 2\sqrt{p-1}.
\]

### 3.3 Haagerup's inequalities.

Let \( \mathbf{u} = (u_1, \ldots, u_p) \) be free Haar unitaries in a \( \mathcal{C}^* \)-probability space \((\mathcal{A}, *, \tau, \| \cdot \|)\). For any integer \( d \geq 1 \), we denote by \( W_d \) the set of elements of \( \mathcal{A} \) of length \( d \) in \((\mathbf{u}, \mathbf{u}^*)\), i.e.
\[
W_d = \left\{ u_{j_1}^{(1)} \cdots u_{j_d}^{(d)} \mid j_1 \neq \cdots \neq j_d, \ v_e \in \{1, *) \ \forall j = 1, \ldots, d \right\}.
\]

In 1979, Haagerup [11] has shown that one has
\[
\left\| \sum_{n \geq 1} \alpha_n x_n \right\| \leq (d + 1) \| \alpha \|_2,
\]
for any sequence \((x_n)_{n \geq 1}\) of elements in \( W_d \) and sequence \( \alpha = (\alpha_n)_{n \geq 1} \) of complex numbers whose \( \ell^2 \)-norm is denoted by
\[
\| \alpha \|_2 = \sqrt{\sum_{n \geq 1} |\alpha|^2}.
\]

This result, known as Haagerup’s inequality, has many applications and has been generalized in many ways. For instance, Buchholz has generalized (3.2) in an estimate of \( \sum_{n \geq 1} a_n \otimes x_n \), where the \( a_n \) are new \( k \times k \) matrices. Let \( U_N \) be a family of \( p \) independent \( N \times N \) unitary Haar matrices. As a byproduct of our main result, we then get from (3.2) an estimate of the norm of matrices of the form
\[
\sum_{n \geq 1} \alpha_n X_n^{(N)}
\]
where for any \( n \geq 1 \), the matrix \( X_n^{(N)} \) is a word of fixed length in \((U_N, U_N^*)\).

Kemp and Speicher [13] have generalized Haagerup’s inequality for \( \mathcal{R} \)-diagonal elements in the so-called holomorphic case. Theorem 1.4 established, the consequence for random matrices sounds relevant since it allows to consider combinations of Haar and deterministic matrices. The result of [13] that states below has been generalized by de la Salle [8] in the case where the non commutative random variables have matrix coefficients. This situation could be interesting for practical applications, where block random matrices are sometimes considered (see [20] for applications of random matrices in telecommunication).

Nevertheless, we only consider the scalar version for simplicity.

Recall that a non commutative random variable \( a \) is called an \( \mathcal{R} \)-diagonal element if it can be written \( a = a y \), for \( y \) a Haar unitary free from \( y \) (see [17]). Let \( a = (a_1, \ldots, a_p) \) be a family of free, identically distributed \( \mathcal{R} \)-diagonal elements in a \( \mathcal{C}^* \)-probability space \((\mathcal{A}, *, \tau, \| \cdot \|)\). We denote by \( W_d^+ \) the set of elements of \( \mathcal{A} \) of length \( d \) in \( a \) (and not its conjugate), i.e.
\[
W_d^+ = \left\{ a_{j_1} \cdots a_{j_d} \mid j_1 \neq \cdots \neq j_d \right\}.
\]

Kemp and Speicher have shown the following, where the interesting fact is that the constant \((d + 1)\) is replace by a constant of order \( \sqrt{d + 1} \): for any sequence \((x_n)_{n \geq 1}\) of elements of \( W_d^+ \) and any sequence \( \alpha = (\alpha_n)_{n \geq 1} \), one has
\[
\left\| \sum_{n \geq 1} \alpha_n x_n \right\| \leq e \sqrt{d + 1} \left\| \sum_{n \geq 1} \alpha_n x_n \right\|_2,
\]
where \( \| \cdot \|_2 \) denotes the \( L^2 \)-norm in \( \mathcal{A} \), given by \( \|x\|_2 = \tau[x^* x]^{1/2} \) for any \( a \) in \( \mathcal{A} \). In particular, if \( a \) is a family of free unitaries (i.e. \( y = 1 \)) then we get \( \| \sum_{n \geq 1} \alpha_n x_n \|_2 = \| \alpha \|_2 \), so that (3.3) is already an improvement of (3.2) without the generalization on \( \mathcal{R} \)-diagonal elements.

Now let \( U_N = (U_1^{(N)}, \ldots, U_p^{(N)}) \), \( V_N = (V_1^{(N)}, \ldots, V_p^{(N)}) \) be families of \( N \times N \) independent unitary Haar matrices and \( Y_N = (Y_1^{(N)}, \ldots, Y_p^{(N)}) \) be a family of \( N \times N \) deterministic Hermitian matrices. Assume that for any \( j = 1, \ldots, p \), the empirical spectral distribution of \( Y_j^{(N)} \) converges weakly to a measure
\(\mu\) (that does not depend on \(j\)) and that for \(N\) large enough, the eigenvalues of \(Y^{(N)}_j\) belong to a small neighborhood of the support of \(\mu\). We set for any \(j = 1, \ldots, p\) the random matrix
\[A_j^{(N)} = U_j^{(N)} Y_j^{(N)} V_j^{(N)*}.\]
From Theorem 1.4 and [15, Corollary 2.1], we can deduce that almost surely the family \((A_1, \ldots, A_p)\) converges strongly in law to a family of free \(R\)-diagonal elements \((a_1, \ldots, a_p)\), identically distributed. Hence, inequality (3.3) gives an asymptotic bound for the norm of a random matrix of the form
\[\sum_{n \geq 1} \alpha_n X_n^{(N)},\]
where for any \(n \geq 1\), the matrix \(X_n^{(N)}\) is a word of fixed length in \(A_j^{(N)}, \ldots, A_j^{(N)*}\).

4. Acknowledgments

B.C. would like to thank Gilles Pisier for asking him Question (3.1) many years ago, as it was a source of inspiration for this paper. C.M. would like thank his Ph.D. supervisor Alice Guionnet for submitting him as a part of his Ph.D. project a question that lead to the main theorem of this paper and for her insightful guidance. He would also like to thank Ofer Zeitouni and Mikael de la Salle for useful discussions during the preparation of this article.

This paper was initiated and mostly completed during the Erwin Schrödinger International Institute for Mathematical Physics workshop “Bialgebras in Free Probability” in April 2011. Both authors gratefully acknowledge the ESI and the organizers of the meeting for providing an inspiring working environment.

Both authors acknowledge the financial support of the ESI. The research of B.C. was supported in part by a Discovery grant from the Natural Science and Engineering Research Council of Canada, an Early Research Award of the Government of Ontario, and the ANR Grâma.

References

[1] C. A. Akemann and P. A. Ostrand. Computing norms in group \(C^*\)-algebras. Amer. J. Math., 98(4):1015–1047, 1976.
[2] G. W. Anderson. Convergence of the largest singular value of a polynomial in independent wigner matrices. arXiv:1103.4825v1, preprint, http://arxiv.org/abs/1008.3099.
[3] G. W. Anderson, A. Guionnet, and O. Zeitouni. An Introduction to Random Matrices, volume 118 of Cambridge studies in advanced mathematics. Cambridge University Press, 2010.
[4] S. Belinschi, B. Collins, and I. Nechita. Laws of large numbers for eigenvectors and eigenvalues associated to random subspaces in a tensor product. arXiv:1008.3099v2, preprint, http://arxiv.org/abs/1008.3099v2.
[5] M. Capitaine and C. Donati-Martin. Strong asymptotic freeness for Wigner and Wishart matrices. Indiana Univ. Math. J., 56(2):767–803, 2007.
[6] B. Collins. Moments and cumulants of polynomial random variables on unitary groups, the Itzykson-Zuber integral, and free probability. Int. Math. Res. Not., (17):953–982, 2003.
[7] B. Collins and I. Nechita. Random quantum channels II: entanglement of random subspaces, Rényi entropy estimates and additivity problems. Adv. Math., 226(2):1181–1201, 2011.
[8] M. de la Salle. Strong Haagerup inequalities with operator coefficients. J. Funct. Anal., 257(12):3968–4002, 2009.
[9] M. Defosseux. Orbit measures, random matrix theory and interlaced determinantal processes. Ann. Inst. Henri Poincaré Probab. Stat., 46(1):209–249, 2010.
[10] A. Guionnet. Large random matrices: lectures on macroscopic asymptotics, volume 1957 of Lecture Notes in Mathematics. Springer-Verlag, Berlin, 2009. Lectures from the 36th Probability Summer School held in Saint-Flour, 2006.
[11] U. Haagerup. An example of a nonnuclear \(C^*\)-algebra, which has the metric approximation property. Invent. Math., 50(3):279–293, 1978/79.
[12] U. Haagerup and S. Thorbjørnsen. A new application of random matrices: Ext(\(C^*_{\text{red}}(F_2)\)) is not a group. Ann. of Math. (2), 162(2):711–775, 2005.
[13] T. Kemp and R. Speicher. Strong Haagerup inequalities for free \(R\)-diagonal elements. J. Funct. Anal., 251(1):141–173, 2007.
[14] H. Kesten. Symmetric random walks on groups. Trans. Amer. Math. Soc., 92:336–354, 1959.
[15] C. Male. The norm of polynomials in large random and deterministic matrices. arXiv:1004.4155v5, preprint, http://arxiv.org/abs/1004.4155, accepted in Probab. Theory Related Fields.
[16] A. Metcalfe. Universality properties of gelfand-tsetlin patterns. arXiv:1105.1272v1, preprint, http://arxiv.org/abs/1105.1272.
[17] A. Nica and R. Speicher. Lectures on the combinatorics of free probability, volume 335 of London Mathematical Society Lecture Note Series. Cambridge University Press, Cambridge, 2006.
[18] G. Pisier. Introduction to operator space theory, volume 294 of London Mathematical Society Lecture Note Series. Cambridge University Press, Cambridge, 2003.
[19] H. Schultz. Non-commutative polynomials of independent Gaussian random matrices. The real and symplectic cases. *Probab. Theory Related Fields*, 131(2):261–309, 2005.

[20] A. M. Tulino and S. Verdú. *Random matrices and wireless communications*, volume vol. 1 of *Fundations and Trends in Communications and Information Theory*. Now Publishers Inc., 2004.

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

Département de Mathématique et Statistique, Université d’Ottawa, 585 King Edward, Ottawa, ON, K1N6N5 Canada. AND CNRS, Institut Camille Jordan Université Lyon 1, 43 Bd du 11 Novembre 1918, 69622 Villeurbanne, France

E-mail address: bcollins@uottawa.ca

École Normale Supérieure de Lyon, Unité de Mathématiques Pures et Appliquées, UMR CNRS 5669, 46 allée d’Italie, 69364 Lyon Cedex 07, FRANCE

E-mail address: camille.male@umpa.ens-lyon.fr