Séminaire BOURBAKI Janvier 2023
75e année, 2022–2023, no 1202

STRONG CONVERGENCE OF THE SPECTRUM OF RANDOM PERMUTATIONS AND ALMOST-RAMANUJAN GRAPHS
[after Charles Bordenave and Benoît Collins]

by **Mylène Maïda**

1. INTRODUCTION

Consider the two following statements:

Independent random permutations, chosen uniformly among all permutations or all matchings of n points, are strongly asymptotically free (viewed as operators on the orthogonal of the constant vector $\mathbf{1}$).

versus

Random n-lifts of a fixed weighted base graph are close to being Ramanujan graphs.

They seem to belong to different mathematical landscapes, random matrix theory and free probability for the first one, theory of expander graphs for the second one. They are nevertheless two instances of the same result, due to Bordenave and Collins (2019) and that we will present hereafter. In particular, we will try to explain the meaning of the statement in each context and why it represents an important improvement with respect to the previous results, starting with the motivation from graph theory and then moving to free probability. This is not the only example of a result dealing with strong asymptotic freeness that can be applied to a completely different context and we will describe in detail, in the last part of these notes, some other applications of this notion.

Je remercie Charles Bordenave pour son aide lors de la préparation de ces notes. Elles ont été rédigées pendant mon séjour au Centre de recherches mathématiques de Montréal (IRL CNRS 3457) que je remercie pour son hospitalité.

2. FROM RAMANUJAN GRAPHS TO THE SYMMETRIC RANDOM PERMUTATION MODEL

The notion of *Ramanujan graph* was initially introduced by Lubotzky, Phillips, and Sarnak (1988) for d-regular graphs. The terminology *Ramanujan* comes from the fact that the construction of the regular graphs considered by Lubotzky, Phillips, and Sarnak (1988) was based on arithmetic properties of pairs of well-chosen prime numbers.

Let $G = (V, E)$ be an undirected graph, with countable vertex set V and edge set E. We allow loops but no multiple edges. The degree of a vertex $v \in V$ is defined as

$$\deg(v) := \sum_{u \in V} \mathbf{1}_{\{u,v\} \in E}.$$

If, for any $v \in V, \deg(v) < \infty$, the graph is said to be locally finite and its *adjacency operator* A is defined as follows: for any $\psi \in \ell_c(V)$, which is the subset of $\ell^2(V)$ of vectors with finite support,

$$A\psi(u) := \sum_{v \in V; \{u,v\} \in E} \psi(v).$$

In the case when V is a finite set, A can be seen as the usual *adjacency matrix* of G. As we are dealing with undirected graphs, the adjacency operator and matrix are self-adjoint. For any integer $d \geq 2$, we say that G is d-regular if all the vertices of G have degree d. For finite d-regular graphs with n vertices, if we denote by $\lambda_0 \geq \lambda_1 \geq \cdots \geq \lambda_{n-1}$ the eigenvalues of A in non-increasing order, one can check that $\lambda_0 = d$ and for all $j \leq n-1$, $|\lambda_j| \leq d$. The eigenvalue λ_0 is always simple and if G is bipartite, then λ_{n-1} is equal to $-d$ and is simple. They are often called the Perron–Frobenius eigenvalues and λ_0 is associated to the constant eigenvector $\mathbf{1}$. On the other hand, if we denote by

$$\lambda(G) := \max\{|\lambda_j| \text{ such that } |\lambda_j| < d\}$$

the largest eigenvalue in absolute value which is not equal to $\pm d$, we have the following result, known as the *Alon–Boppana bound*:

Theorem 2.1 (Alon, 1986). — *Let $(G_{n,d})_{n\geq 1}$ be any sequence of connected d-regular graphs such that, for any $n \in \mathbb{N}^*$, $G_{n,d}$ has n vertices. Then*

$$\liminf_{n \to \infty} \lambda(G_{n,d}) \geq 2\sqrt{d-1}.$$

This lead to the following definition:

Definition 2.2 (Ramanujan graph, d-regular case). — *A d-regular, connected, finite graph G is called* Ramanujan *if and only if any eigenvalue λ of its adjacency operator is such that $\lambda \in \{-d, d\}$ or $|\lambda| \leq 2\sqrt{d-1}$.*

Among connected d-regular graphs, Ramanujan graphs are the graphs with maximal spectral gap. For this reason, sequences of such graphs have very good properties as *expander graphs* (we refer to Kowalski (2019) for details on the link between spectral gap and expander properties of graphs). The question is then how to construct such sequences of graphs. In this direction, one has to mention the remarkable result of Friedman (2008):

Theorem 2.3. — *Let $d \geq 3$ be an integer. For each n such that nd is even, let G_n be a random graph chosen uniformly among d-regular graphs with n vertices. Then the sequence $(G_n)_{n\geq 1}$ is almost-Ramanujan*[(1)] *in the sense that, for any $\varepsilon > 0$,*

$$\lim_{n\to\infty} \mathbb{P}(\lambda(G_n) \geq 2\sqrt{d-1} + \varepsilon) = 0.$$

The notion of *random lift* is also very useful to construct almost-Ramanujan graphs. On the way of defining them, we also give a more general definition of Ramanujan graphs. Let G and H be undirected connected[(2)] graphs, with no self-loops and no multiple edges. A map $\pi\colon H \to G$ is called a *covering map* if for every vertex $h \in H$, π gives a bijection from the edges incident to h and those incident to $\pi(h)$. When $\pi\colon H \to G$ is a covering map, G is called the *base graph* and H is called a *covering graph* of G. If the base graph G is connected, then the cardinal of $\pi^{-1}(g)$ is the same for any vertex g of G. If this cardinal is equal to n, then H is called a *n-lift* of G. If H is a tree, it is the *universal cover* of G. In particular, the universal cover of any non-empty, connected, d-regular graph is the infinite d-regular tree $\mathbb{T}_d$. It is known that the spectrum of $\mathbb{T}_d$ is contained in the interval $[-2\sqrt{d-1}, 2\sqrt{d-1}]$[(3)]. Therefore, for a d-regular graph, being a Ramanujan graph means that all eigenvalues, expect $\pm d$, are contained inside the spectrum of its universal cover.

This led Greenberg (1995), in his thesis, to give a more general definition of a Ramanujan graph, not necessarily restricted to d-regular graphs.

Definition 2.4 (Ramanujan graph, general case). — *A finite connected graph X is Ramanujan if the spectrum of its adjacency operator is contained in $[-\rho, \rho] \cup \{-\lambda_0, \lambda_0\}$, where λ_0 is the largest eigenvalue of the graph and ρ is the spectral radius of its universal cover.*

Let us describe a standard model, due to Amit and Linial (2002) for constructing *random lifts*. Given a base graph $G = (V_G, E_G)$ and an integer $n \geq 2$, to each vertex v of G we associate a set of n vertices $(v, 1), \ldots, (v, n)$. For each edge $e = \{u, v\}$ of G, we choose an orientation, say (u, v), and a uniform permutation σ_e of $[n] := \{1, \ldots, n\}$, independent of all other edges. Then, if the vertex set of H is $\{(u, i), u \in V_G, i \in [n]\}$ and the edge set is $\{\{(u, i), (v, \sigma_{(u,v)}(i))\}, \{u, v\} \in E_G, i \in [n]\}$, then H is a *random n-lift* of G and its law is uniform over all possible n-lifts of G. Note that the choice of the orientations of the edges made for the construction does not change the distribution of the random lift H. Improving on Theorem 2.3, Bordenave (2020) showed that, under very general conditions on the base graph, the sequence of random n-lifts form a sequence of almost-Ramanujan graphs.

The model that we describe now is the main object studied by Bordenave and Collins (2019) and can be seen as a generalization of the notion of random lift; we will call it

(1) or weakly Ramanujan

(2) Covering maps can be defined in a more general framework but we only need the case of connected graphs.

(3) Its spectral measure is known as the Kesten–McKay distribution.

the *symmetric random permutation model*. Let X be a countable set. Let $\sigma_1, \ldots, \sigma_d$ be d permutations of the set X. We consider $\ell^2(X)$ the Hilbert space space spanned by the orthonormal basis $(\delta_x)_{x \in X}$. The identity operator on $\ell^2(X)$ is denoted by $\mathbf{1}$. A permutation σ_i acts naturally as a unitary operator S_i on $\ell^2(X)$ by $S_i(g)(x) = g(\sigma_i(x))$, for any $g \in \ell^2(X)$. Let $a_0, a_1, \ldots, a_d$ be matrices of size $r \times r$. The main object of interest in Bordenave and Collins (2019) is the operator $A := a_0 \otimes \mathbf{1} + \sum_{i=1}^{d} a_i \otimes S_i$ acting on $\ell^2(X)$. When $X = [n]$, we denote by $\sigma_{1,n}, \ldots, \sigma_{d,n}$ the permutations of X, by $S_{1,n}, \ldots, S_{d,n}$ the corresponding operators, $\mathbf{1}^{(n)}$ the identity operator and

$$A_n := a_0 \otimes \mathbf{1}^{(n)} + \sum_{i=1}^{d} a_i \otimes S_{i,n}. \tag{1}$$

Two symmetry conditions are added, one on the matrices $a_1, \ldots, a_d$ and one on the permutations $\sigma_{1,n}, \ldots, \sigma_{d,n}$.

Assumption 2.5 (Symmetric random permutation model)
We equip $[d]$ with the following involution: let $q \leq \frac{d}{2}$ be an integer; for any $i \in [q]$, set $i^* = i + q$, for $q + 1 \leq i \leq 2q$, set $i^* = i - q$, and for $2q + 1 \leq i \leq d$, set $i^* = i$. We assume that:

(H**a**) $a_0 = a_0^*$ *and* $\forall i \in \{1, \ldots, d\}, a_{i^*} = (a_i)^*$ *and* $\sigma_{i^*} = (\sigma_i)^{-1}$.

(H$\boldsymbol{\sigma}$) *The permutations $\{\sigma_{1,n}, \ldots, \sigma_{q,n}\} \cup \{\sigma_{2q+1,n}, \ldots, \sigma_{d,n}\}$ on $[n]$ are independent, $\{\sigma_{1,n}, \ldots, \sigma_{q,n}\}$ are uniformly distributed among the permutations of $[n]$ and $\{\sigma_{2q+1,n}, \ldots, \sigma_{d,n}\}$ are uniformly distributed among the* matchings[(4)] *of $[n]$ and for any $i \in [d]$, $\sigma_{i^*,n} = (\sigma_{i,n})^{-1}$.*

Let us explain how it can be seen as a generalization of the model of random lift. Assume that d is even, $q = d/2$, $a_0 = 0$ and the matrices $a_1, \ldots, a_d$ are of the form $a_i = E_{u_i v_i}$, with $u_i, v_i \in [r]$. The base graph will be the graph G with vertex set $[r]$ and adjacency matrix $A_1 = \sum_{i=1}^{d} a_i$. Under Assumption (H**a**), it will be undirected, with $d/2$ edges. The graph H with vertex set $[n] \times [r]$ and edges of the form $\{(x, u_i), (\sigma_i(x), v_i)\}$ is a n-lift of G. If the permutations $\sigma_{1,n}, \ldots, \sigma_{d,n}$ fulfill Assumption (H$\boldsymbol{\sigma}$), then the random lift we obtain has the same distribution as in the construction of Amit and Linial (2002).

Bordenave and Collins (2019) show that the $A_n, n \geq 1$, are the adjacency operators of an almost-Ramanujan sequence of weighted graphs, in a sense related to Definition 2.4 (we refer to Theorem 3.13 for a precise statement). But in parallel to this graph-theoretical motivation, the symmetric random permutation model is linked with asymptotic freeness properties of random permutations and we develop this point of view in the next section.

(4) A matching (or pair matching) is a permutation for which all the cycles are of length 2, that is an involution without fixed point.

3. FREENESS, ASYMPTOTIC FREENESS AND STRONG ASYMPTOTIC FREENESS

In the eighties, Dan Voiculescu introduced the concept of *freeness* (or *free independence*) in the context of operator algebras and created the field of *free probability theory*. In the early nineties, he discovered that many models of random matrices were *asymptotically free*, leading to model elements in operator algebras through random matrices. Since then, there has been a constant interplay between free probability theory and *random matrix theory* (RMT). We will try to give the main lines of these fruitful interactions. Among many nice references on free probability theory, we have chosen to follow the recent book of Mingo and Speicher (2017) and the lecture notes of Speicher (2019).

3.1. The notion of freeness

Let us start with the definition of freeness.

DEFINITION 3.1 (Freeness). — *Consider a unital algebra $\mathcal{A}$ over $\mathbb{C}$, equipped with a linear functional $\tau\colon \mathcal{A} \to \mathbb{C}$ such that $\tau(1) = 1$. $(\mathcal{A}, \tau)$ is called a* non-commutative probability space*. Unital subalgebras $(\mathcal{A}_i)_{i\in I}$ are called* free *(or* freely independent*) in $(\mathcal{A}, \tau)$ if, for any $a_1, \ldots, a_k$ such that $\forall j \in [k]$, $\tau(a_j) = 0, a_j \in \mathcal{A}_{i(j)}$ and $i(1) \neq i(2) \neq \cdots \neq i(k)$,*

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

An important example, which is particularly relevant in our context, is the following:

Example 3.2. — Let G be a group and $\mathbb{C}G$ its group algebra, that is

$$\mathbb{C}G = \Big\{\sum_{g\in G} \alpha_g g, \alpha_g \in \mathbb{C}, \forall g \in G, \alpha_g \neq 0 \text{ for finitely many } g\Big\}.$$

Then $\mathbb{C}G$ is a unital algebra and one can define $\tau_G\colon \mathbb{C}G \to \mathbb{C}$,

$$\tau_G\Big(\sum_{g\in G} \alpha_g g\Big) = \alpha_e,$$

where e is the neutral element in G, so that $(\mathbb{C}G, \tau_G)$ is a non-commutative probability space. It is related with the notion of freeness for subgroups in the algebraic sense : a family $(G_i)_{i\in I}$ of subgroups of G is free if, for any $g_1, \ldots, g_k$ such that $\forall j \in [k], g_j \in G_{i(j)}$ and $i(1) \neq i(2) \neq \cdots \neq i(k)$, $g_1 \ldots g_k \neq e$ whenever $g_1 \neq e, \ldots, g_k \neq e$. The link between free independence of subalgebras in the sense of Definition 3.1 and freeness for subgroups in the algebraic sense is made clear by the following proposition:

PROPOSITION 3.3. — *Let $(G_i)_{i\in I}$ be subgroups of a group G. Then the following statements are equivalent:*

- *The subgroups $(G_i)_{i\in I}$ are free in G.*
- *The subalgebras $(\mathbb{C}G_i)_{i\in I}$ are freely independent in the non-commutative probability space $(\mathbb{C}G, \tau_G)$.*

It is possible to enrich the structure of a non-commutative probability space as follows:

Definition 3.4. — *Let $(\mathcal{A}, \tau)$ be a non-commutative probability space.*

- *If τ is a trace, i.e. if $\tau(ab) = \tau(ba)$, for all $a, b \in \mathcal{A}$, then we call $(\mathcal{A}, \tau)$ a* tracial *non-commutative probability space.*
- *If $\mathcal{A}$ is a $*$-algebra (resp. a C^*-algebra) and τ is positive, i.e. if $\tau(a^*a) \geq 0$ for all $a \in \mathcal{A}$, then we call τ a* state *and $(\mathcal{A}, \tau)$ a $*$-probability space (resp. a C^*-probability space).*
- *A state τ is faithful if for all $a \in \mathcal{A}$, $\tau(a^*a) = 0$ implies $a = 0$.*

If $(\mathcal{A}, \tau)$ a $*$-probability space and $(a_1, \dots, a_m)$ is a family of m elements in $\mathcal{A}$, then the $*$-*distribution* (or simply the distribution) of $(a_1, \dots, a_m)$ is given by the collection $\{\tau(P(a_1, a_1^*, \dots, a_m, a_m^*)), P \in \mathbb{C}\langle X_1, \dots, X_{2m}\rangle\}$, where $\mathbb{C}\langle X_1, \dots, X_{2m}\rangle$ is the set of non-commutative polynomials[5] in $2m$ variables with complex coefficients. By analogy with classical probability theory, elements of a $*$-probability space are called *random variables* and the random variables $(a_i)_{i \in I}$ are freely independent if the subalgebras generated respectively by $(1, a_i, a_i^*)$ are freely independent.

3.2. The notion of asymptotic freeness

Roughly speaking, we say that a sequence of non-commutative random variables, say random matrices, are asymptotically free if they converge in distribution to a family of freely independent random variables. Let us give a precise meaning of this sentence.

Definition 3.5. — *Let $((\mathcal{A}_n, \tau_n))_{n \geq 1}$ be a sequence of $*$-probability spaces and $(\mathcal{A}, \tau)$ a $*$-probability space. If, for any $n \geq 1$, $a_{1,n}, \dots, a_{k,n}$ is a k-tuple of random variables in $(\mathcal{A}_n, \tau_n)$ and if there exists $a_1, \dots, a_k \in \mathcal{A}$ such that, for any non-commutative polynomial P in $2k$ variables, we have*

$$\tau_n(P(a_{1,n}, a_{1,n}^*, \dots, a_{k,n}, a_{k,n}^*)) \xrightarrow[n \to \infty]{} \tau(P(a_1, a_1^*, \dots, a_k, a_k^*)),$$

we say that $(a_{1,n}, \dots, a_{k,n})_{n \geq 1}$ converges in $*$-distribution *to $(a_1, \dots, a_k)$. If, in addition, $(a_1, \dots, a_k)$ is a family of random variables that are freely independent (with respect to τ), we say that $(a_{1,n}, \dots, a_{k,n})$ are* asymptotically free *(for $n \to \infty$).*

This notion appeared in Voiculescu's work (see e.g. Voiculescu, 1991), a few years after introducing free independence. Since random matrices are typical examples of asymptotically free random variables, it built an important bridge between operator space theory and random matrix theory. This was first identified on the most emblematic ensemble of random matrices: the Gaussian Unitary Ensemble (Gue). Before stating a precise result, let us explain how one can define an ensemble of random matrices in the framework of $*$-probability spaces.

[5] in the sense that e.g. $X_1X_2 \neq X_2X_1$

LEMMA 3.6. — *For every $n \geq 1$, if we set $(\mathcal{A}_n, \tau_n) := (M_n(L^{\infty-}(\Omega, \mathbb{P})), \mathrm{tr}_n \otimes \mathbb{E})$, where $(\Omega, \mathbb{P})$ is a classical probability space, where $L^{\infty-}(\Omega, \mathbb{P}) := \bigcap_{1 \leq p < \infty} L^p(\Omega, \mathbb{P})$ and for any complex algebra $\mathcal{A}$, $M_n(\mathcal{A}) \simeq M_n(\mathbb{C}) \otimes \mathcal{A}$ denotes the $n \times n$ matrices with entries from $\mathcal{A}$, where $\mathbb{E}$ denotes the expectation with respect to $\mathbb{P}$ and $\mathrm{tr}_n := \frac{1}{n}\mathrm{tr}$ the normalized trace on $M_n(\mathbb{C})$, then a $(\mathcal{A}_n, \tau_n)$ is a $*$-probability space.*

It means that we will consider in the sequel $n \times n$ random matrices of the form $A = (a_{ij})_{i,j \in [n]}$, with, for all $i, j \in [n]$, $a_{ij} \in L^{\infty-}(\Omega, \mathbb{P})$, and a state given by

$$\tau_n(A) = \mathrm{tr}_n \otimes \mathbb{E}(A) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}(a_{ii}).$$

An important example is provided by Gaussian random matrices:

Example 3.7. — For $n \geq 1$, we say that A belongs to the Gaussian unitary ensemble of dimension n, and write $A \in \mathrm{GUE}(n)$, if $A = (a_{ij})_{i,j \in [n]}$ is a random $n \times n$ matrix such that $A = A^*$ and such that the entries form a centred complex Gaussian family with covariance

$$\mathrm{Cov}(a_{ij}, a_{k\ell}) = \frac{1}{n} \delta_{i,\ell} \delta_{j,k}.$$

If we consider a sequence of matrices $(G_n)_{n \in \mathbb{N}}$ such that, for any $n \in \mathbb{N}^*$, $G_n \in \mathrm{GUE}(n)$, then one can show that, for any $k \in \mathbb{N}$,

$$\mathbb{E}\left(\mathrm{tr}_n(G_n^k)\right) \xrightarrow[n \to \infty]{} C_k := \begin{cases} 0 & \text{if } k \text{ is odd}, \\ \frac{1}{m+1}\binom{2m}{m} & \text{if } k = 2m, m \in \mathbb{N}. \end{cases}$$

In a $*$-probability space $(\mathcal{A}, \tau)$, one can define a self-adjoint element s such that, for any k in $\mathbb{N}$, $\tau(s^k) = C_k$. Then s is called a *semi-circular element* and $(G_n)_{n \in \mathbb{N}}$ converges in $*$-distribution to s. Now, asymptotic freeness for independent matrices from the $\mathrm{GUE}(n)$ can be stated as follows :

THEOREM 3.8 (Voiculescu, 1991). — *Let $\ell \in \mathbb{N}^*$ be fixed and, for every $n \in \mathbb{N}$, let $(A_{1,n}, \ldots, A_{\ell,n})$ be ℓ independent matrices from the $\mathrm{GUE}(n)$. Then the sequence $(A_{1,n}, \ldots, A_{\ell,n})_{n \geq 1}$ converges in $*$-distribution to a family $(s_1, \ldots, s_\ell)$ of freely independent semi-circular elements. The random matrices $(A_{1,n}, \ldots, A_{\ell,n})$ are therefore asymptotically free.*

An important property of the $\mathrm{GUE}(n)$ is its invariance by conjugation by a unitary matrix. This remark opened the way to asymptotic freeness for other ensembles of matrices. We denote by $\mathcal{U}(n) := \{U \in M_n(\mathbb{C}), UU^* = I_n\}$ the group of unitary matrices of size n. It is a compact group and therefore there exists a unique probability measure which is invariant by multiplication, called the Haar measure on $\mathcal{U}(n)$. If U_n is distributed according to the Haar measure on $\mathcal{U}(n)$, one can check that, for any $k \in \mathbb{Z}$,

$$\mathbb{E}\left(\mathrm{tr_n}(U_n^k)\right) \xrightarrow[n \to \infty]{} \delta_{k,0}.$$

By analogy, if u is a unitary element in a $*$-probability space $(\mathcal{A}, \tau)$ which satisfies $\tau(u^k) = \delta_{k,0}$, for any k in $\mathbb{Z}$, then u is called *Haar unitary.* A first asymptotic freeness property in this context is the following:

Proposition 3.9. — *Let $\ell \in \mathbb{N}^*$ be fixed and, for every $n \in \mathbb{N}$, let $(U_{1,n}, \ldots, U_{\ell,n})$ be ℓ independent random matrices distributed according to the Haar measure on $\mathcal{U}(n)$. Then the sequence $(U_{1,n}, \ldots, U_{\ell,n})_{n\geq 1}$ converges in $*$-distribution to a family $(u_1, \ldots, u_\ell)$ of freely independent Haar unitaries. The random matrices $(U_{1,n}, \ldots, U_{\ell,n})_{n\geq 1}$ are therefore asymptotically free.*

But Haar distributed random matrices can also convey asymptotic freeness to deterministic matrices.

Theorem 3.10. — *For every $n \in \mathbb{N}$, let U_n be a Haar unitary random $n \times n$-matrix, let $A_n, B_n \in M_n(\mathbb{C})$, and suppose that $(A_n)_{n\geq 1}$ converges in $*$-distribution to a and $(B_n)_{n\geq 1}$ converges in $*$-distribution to b for random variables a and b in some $*$-probability space, and that the sequences $(A_n)_{n\geq 1}$ and $(B_n)_{n\geq 1}$ are deterministic or independent of $(U_n)_{n\geq 1}$. Then, $(U_n A_n U_n^*, B_n)_{n\geq 1}$ converges in distribution to (a, b), where a and b are freely independent. In particular, the random matrices $(U_n A_n U_n^*)_{n\geq 1}$ and $(B_n)_{n\geq 1}$ are asymptotically free from each other.*

If we now go to the group $\mathfrak{S}_n$ of permutations of $[n]$, then the uniform law on $\mathfrak{S}_n$ is the Haar measure and it is natural to ask if a result similar to Theorem 3.10 holds true in this context. A positive answer has been brought by Nica (1993):

Theorem 3.11. — *Let $\ell \in \mathbb{N}^*$ be fixed and, for every $n \geq 1$, let $(S_{1,n}, \ldots, S_{\ell,n})$ be ℓ independent random matrices distributed uniformly over the set of permutation matrices of size n. Then the sequence $(S_{1,n}, \ldots, S_{\ell,n})_{n\geq 1}$ converges in $*$-distribution to a family $(u_1, \ldots, u_\ell)$ of freely independent Haar unitaries. The random matrices $(S_{1,n}, \ldots, S_{\ell,n})_{n\geq 1}$ are therefore asymptotically free.*

We will in fact use an extension of this result. Beyond the uniform distribution of the set of permutations of size n, we will also be interested in the uniform distribution of the set of matchings of n points, that is the subset of involutions without fixed point within the permutations of size n. Obviously, if $(R_n)_{n\geq 1}$ is a sequence of random variables with this distribution, it converges in $*$-distribution to a free random variable r such that $\tau(r^k) = 1$ if k is even and $\tau(r^k) = 0$ if k is odd. We say that r is a *free Rademacher* variable. We have the following asymptotic freeness result:

Proposition 3.12. — *With the notations of Theorem 3.11 and if $(R_{1,n}, \ldots, R_{m,n})$ are m independent random matrices distributed uniformly over the set of pair-matchings of n points, independent of $(S_{1,n}, \ldots, S_{\ell,n})$, the sequence $(S_{1,n}, \ldots, S_{\ell,n}, R_{1,n}, \ldots, R_{m,n})_{n\geq 1}$ converges in $*$-distribution to $(u_1, \ldots, u_\ell, r_1, \ldots, r_m)$, where $(r_1, \ldots, r_m)$ are freely independent free Rademacher variables, freely independent from $(u_1, \ldots, u_\ell)$.*

3.3. Limiting operator for the symmetric random permutation model

As explained above, we want to study the convergence of the sequence of operators $(A_n)_{n\geq 1}$ defined in (1), under Assumption 2.5. The goal of this section is to construct the limiting object it converges to. From Proposition 3.12, we know it will involve a family of freely independent Haar unitaries and free Rademacher variables. To describe them more concretely, we go back to our very first example 3.2 of non-commutative probability space but we will enrich the structure of the group algebra.

On a discrete group G, we define the inner product $\langle g, h\rangle = \delta_{g,h}$ and extend it to a bilinear form on $\mathbb{C}G$. Then

$$\ell^2(G) := \left\{\sum_{g\in G} \alpha_g g, \sum_{g\in G} |\alpha_g|^2 < \infty\right\}$$

is a Hilbert space. If $\mathcal{B}(\ell^2(G))$ is the space of bounded operators on $\ell^2(G)$, then we can define $\lambda\colon G \to \mathcal{B}(\ell^2(G))$ as follows: if $\sum_{g\in G} |\alpha_g|^2 < \infty, \forall g \in G$, we define

$$\lambda(g)\cdot\alpha = \lambda(g)\cdot \sum_{h\in G} \alpha_h h := \sum_{h\in G} \alpha_h\, g^{-1}h.$$

It is called the *left regular representation* and one can check that it is unitary. We then extend the application λ to $\mathbb{C}G$

$$\lambda(\alpha) := \sum_{g\in G} \alpha_g \lambda(g)$$

and the linear form τ_G to $\lambda(\mathbb{C}G)$ by

$$\tau_G(\lambda(g)) = \tau_G(g) = \delta_{g,e}.$$

We have $\lambda(\mathbb{C}G) \subset \mathcal{B}(\ell^2(G))$ and the closure of $\lambda(\mathbb{C}G)$ with respect to the operator norm topology is a C^*-algebra called the *reduced group C^*-algebra* of G, denoted by $C_{\mathrm{red}}(G)$.

In the context of the work of Bordenave and Collins (2019), the discrete group G (hereafter denoted by X_*) we start with takes the form of a free product:

$$X_* := \mathbb{Z}^{*q} * (\mathbb{Z}/2\mathbb{Z})^{*(d-2q)},$$

where $\mathbb{Z}^{*q}$ denotes the free product of q copies of $\mathbb{Z}$ and τ_{X_*} will be simply denoted by τ when there is no ambiguity.

More concretely, in our case, if $g_1, \ldots, g_d$ are generators of X_* such that, for $i \in [q]$, (g_i, g_{i+q}) generates the ith copy of $\mathbb{Z}$, then $((\lambda(g_i), \lambda(g_{i+q}))_{i\in[q]}, (\lambda(g_i))_{2q+1\leq i\leq d})$ is a family of freely independent random variables. Moreover, for $i \in [2q]$, for all $k \in \mathbb{Z}$, $\tau(\lambda(g_i)^k) = \tau(g_i^k) = \delta_{k,0}$, so that these random variables are Haar unitaries, whereas when $2q+1 \leq i \leq d, \tau(g_i^k) = 1$ if k is even and 0 if k is odd so that these random variables are free Rademacher variables. Otherwise stated, $(\lambda(g_i))_{i\in[d]}$ is a concrete realization of the limit in $*$-distribution of the operators $(S_{1,n}, \ldots, S_{d,n})_{n\geq 1}$ satisfying

Assumption (H$\boldsymbol{\sigma}$) and it is a straightforward consequence of Proposition 3.12 that the sequence of operators $(A_n)_{n\geq 1}$ we are interested in should converges to

$$A_* := a_0 \otimes \mathbf{1} + \sum_{i=1}^{d} a_i \otimes \lambda(g_i), \tag{2}$$

acting on $\mathbb{C}^r \otimes \ell^2(X_\star)$. This is an element of the unital $*$-algebra $\mathcal{A} := M_r(C_{\text{red}}(X_\star))$, equipped with the trace $\text{tr}_r \otimes \tau$. This implies also the convergence of the sequence of operators $(A_n)_{n\geq 1}$ to A_* in the topology of local convergence of Benjamini and Schramm but we don't want to insist here on this aspect. Bordenave and Collins (2019) rather focus on the convergence of the spectrum. We denote by $\sigma(T)$ the spectrum of an operator T, that is

$$\sigma(T) := \{\lambda \in \mathbb{C}, T - \lambda I \text{ is not invertible}\}.$$

The convergence in $*$-distribution of $(A_n)_{n\geq 1}$ to A_* implies that, if $[a,b]$ is an interval that does not intersect the spectrum of the limiting operator A_*, then the expected proportion of eigenvalues of A_n in $[a,b]$ tends to zero as n grows to infinity. Therefore, $\forall \varepsilon > 0$, for n large enough, with high probability,

$$\sigma(A_*) \subset \sigma(A_n) + [-\varepsilon, \varepsilon]. \tag{3}$$

The main object of the paper of Bordenave and Collins (2019) is to establish the reverse inclusion. This is related to the notion of strong convergence, that we introduce in the next section.

3.4. Strong convergence, strong asymptotic freeness, linearization trick

The fact that $(A_n)_{n\geq 1}$ converge in $*$-distribution to A_* does not rule out the possibility of $o(n)$ eigenvalues staying away from the limiting spectrum. These eigenvalues are usually called *outliers*. In our context, there are obvious outliers, coming from the Perron–Frobenius eigenvalues of the operators $S_{i,n}$. For example, one can check that $S_{1,n} + S_{1,n}^* + \cdots + S_{k,n} + S_{k,n}^*$ will always have an eigenvalue equal to $2k$ associated to the constant vector $\mathbf{1}$. This will produce an outlier as long as $k \geq 2$. To get rid of these trivial eigenvalues, we will only consider the operators $S_{i,n}$ restricted to the orthogonal $\mathbf{1}^\perp$ of the constant vector $\mathbf{1}$, or equivalently the operator A_n restricted to $H_0 := (\mathbb{C}^r \otimes \mathbf{1})^\perp$. Their main theorem is the following:

Theorem 3.13. — *Consider a sequence of random operators $(A_n)_{n\geq 1}$ distributed according to the symmetric random permutation model (1). Then the Hausdorff distance between the spectrum $\sigma(A_*)$ of the operator A_* defined in (2) and the spectrum $\sigma((A_n)_{|H_0})$ of the operator A_n restricted to H_0 converges to zero in probability as n goes to infinity. Otherwise stated, $\forall \varepsilon > 0$,*

$$\mathbb{P}\left(\sigma((A_n)_{|H_0}) \subset \sigma(A_*) + [-\varepsilon, \varepsilon]\right) \xrightarrow[n\to\infty]{} 1 \tag{4}$$

$$\mathbb{P}\left(\sigma(A_*) \subset \sigma((A_n)_{|H_0}) + [-\varepsilon, \varepsilon]\right) \xrightarrow[n\to\infty]{} 1 \tag{5}$$

The second point comes directly from (3) and the fact that the spectrum of A_* does not contain the Perron–Frobenius eigenvalue (as the constant vector $\mathbf{1}$ is not in $\ell^2(X_*)$). We will develop in detail the proof of (4). Before that, we want to relate this inclusion with the notion of *strong asymptotic freeness*. With Definition 3.5 in mind, we define *strong convergence* as follows:

Definition 3.14. — *Let $((\mathcal{A}_n, \tau_n))_{n\geq 1}$ be a sequence of C^*-probability spaces and $(\mathcal{A}, \tau)$ a C^*-probability space. If, for any $n \geq 1$, $a_{1,n}, \dots, a_{k,n}$ is a k-tuple of random variables in $(\mathcal{A}_n, \tau_n)$ and if there exist $a_1, \dots, a_k \in \mathcal{A}$ such that, for any non-commutative polynomial P in $2k$ variables, we have*

$$\tau_n(P(a_{1,n}, a_{1,n}^*, \dots, a_{k,n}, a_{k,n}^*)) \xrightarrow[n\to\infty]{} \tau(P(a_1, a_1^*, \dots, a_k, a_k^*)),$$

and in operator norm we have

$$\|P(a_{1,n}, a_{1,n}^*, \dots, a_{k,n}, a_{k,n}^*)\| \xrightarrow[n\to\infty]{} \|P(a_1, a_1^*, \dots, a_k, a_k^*)\|,$$

we say that $(a_{1,n}, \dots, a_{k,n})_{n\geq 1}$ converges strongly *to $(a_1, \dots, a_k)$. If $(a_1, \dots, a_k)$ is a family of random variables that are freely independent (with respect to τ), we say that $(a_{1,n}, \dots, a_{k,n})$ are* strongly asymptotically free*.*

The following proposition, that can be found e.g. in (Collins and Male, 2014), clarifies the link between strong convergence and control of outliers. If h is a self-adjoint element in a C^*-probability space $(\mathcal{A}, \tau)$, its spectral measure μ_h is the unique probability measure on $\mathbb{R}$ such that, for any $k \geq 1$, $\tau(h^k) = \int t^k \mathrm{d}\mu_h(t)$; we denote its support by $\mathrm{supp}(\mu_h)$.

Proposition 3.15. — *Let $((\mathcal{A}_n, \tau_n))_{n\geq 1}$ be a sequence of $*$-probability spaces and $(\mathcal{A}, \tau)$ a $*$-probability space. For any $n \geq 1$, let $a_{1,n}, \dots, a_{k,n}$ be k-tuple of random variables in $(\mathcal{A}_n, \tau_n)$ and $a_1, \dots, a_k \in (\mathcal{A}, \tau)$. Then the two following statements are equivalent :*

- *$(a_{1,n}, \dots, a_{k,n})_{n\geq 1}$ converges strongly to $(a_1, \dots, a_k)$,*
- *for any polynomial P such that $h_n := P(a_{1,n}, \dots, a_{k,n})$ is self-adjoint, $(P(a_{1,n}, \dots, a_{k,n}))_{n\geq 1}$ converges in $*$-distribution to $h := P(a_1, \dots, a_k)$ and $\forall \varepsilon > 0$, for n large enough,*

$$\mathit{supp}(\mu_{h_n}) \subset \mathit{supp}(\mu_h) + [-\varepsilon, \varepsilon],$$

where μ_{h_n} and μ_h are respectively the spectral measures of h_n and h.

The strong convergence result obtained by Bordenave and Collins (2019) can be stated as follows:

Theorem 3.16. — *Let $d \geq 1$ be fixed. For every $n \geq 1$, consider a d-uple $(S_{1,n}, \dots, S_{d,n})$ of random permutations of $[n]$ satisfying $(\mathrm{H}\boldsymbol{\sigma})$. Then, for any non-commutative polynomial P, for any $\varepsilon > 0$,*

$$\mathbb{P}\left(\Big|\|P((S_{1,n})_{|H_0}, \dots, (S_{d,n})_{|H_0})\| - \|P(\lambda(g_1), \dots, \lambda(g_d))\|\Big| > \varepsilon\right) \xrightarrow[n\to\infty]{} 0,$$

where $\|\cdot\|$ *stands for the operator norm (on the space of matrices for the first term and in the* C^**-algebra* $C_{\mathrm{red}}(X_*)$ *for the second term).*

Since $(\lambda(g_1),\dots,\lambda(g_d))$ *are freely independent*[(6)], *we say that* $((S_{1,n})_{|H_0},\dots,(S_{d,n})_{|H_0})$ *are strongly asymptotically free.*

This result may look stronger than Theorem 3.13 as it implies convergence of the norm of any polynomial P in the operators $S_{i,n}$ whereas Theorem 3.13 only deals with polynomials of degree one (with matrix coefficients). In fact, due to the *linearization trick*, Theorem 3.16 can be seen as a corollary of Theorem 3.13. In the context of strong convergence, the argument is due to Haagerup and Thorbjørnsen (2005), although the idea of linearizing polynomial problems by going to matrices is much older and known under different names in different mathematical communities. More precisely, for P a non-commutative polynomial in d variables and $r\in\mathbb{N}^*$, a matrix $\hat{P}:=\begin{pmatrix}0 & U\\ V & Q\end{pmatrix}\in M_r(\mathbb{C})\otimes\mathbb{C}\langle X_1,\dots,X_d\rangle$, where $Q\in M_{r-1}(\mathbb{C})\otimes\mathbb{C}\langle X_1,\dots,X_d\rangle$ and U and V are respectively row and column vectors of size $r-1$ with entries in $\mathbb{C}\langle X_1,\dots,X_d\rangle$, is a *linearization* of P if

$$\hat{P}=b_0\otimes\mathbf{1}+b_1\otimes X_1+\dots+b_d\otimes X_d,$$

and $P=-UQ^{-1}V$.

For example, a possible linearization of a monomial $P=X_{i_1}\dots X_{i_r}$ is

$$\hat{P}=\begin{pmatrix} & & & X_{i_1}\\ & & X_{i_2} & -1\\ & \iddots & \iddots & \\ X_{i_r} & -1 & & \end{pmatrix}$$

One can show that any polynomial P admits a linearization and if P is self-adjoint, the linearization can be chosen self-adjoint.

From there, one can get a criterion which is crucial for proving strong convergence:

Proposition 3.17 (Linearization trick). — *If* $\boldsymbol{u}:=(u_1,\dots,u_d)$ *is a d-uple of elements of a* C^**-algebra* $\mathcal{A}$ *and* $\boldsymbol{v}:=(v_1,\dots,v_d)$ *is a d-uple of elements of a* C^**-algebra* $\mathcal{B}$*, then the following are equivalent:*

- *for any non-commutative polynomial* P *in* d *variables and their adjoints,* $\|P(\boldsymbol{u})\|=\|P(\boldsymbol{v})\|$,
- *for any integer* r*, and* $r\times r$ *matrices* $a_0,\dots,a_d$ *such that* $a_0\otimes 1+a_1\otimes u_1+\dots+a_d\otimes u_d$ *and* $a_0\otimes 1+a_1\otimes v_1+\dots+a_d\otimes v_d$ *are self-adjoint, we have*

$$\|a_0\otimes 1+a_1\otimes u_1+\dots+a_d\otimes u_d\|=\|a_0\otimes 1+a_1\otimes v_1+\dots+a_d\otimes v_d\|.$$

(6) This is a slight abuse of notation: it would be more correct to say that $((\lambda(g_i),\lambda(g_{i+q}))_{i\in[q]},(\lambda(g_i))_{2q+1\le i\le d})$ is a family of freely independent random variables and that the corresponding family for the $S_{i,n}$'s are strongly asymptotically free.

This explains how Theorem 3.16 is a corollary of Theorem 3.13.

In the next section, we will develop the main lines of the proof of Theorem 3.13 and introduce in particular the notion of non-backtracking operators. We will go back to free probability, in particular strong asymptotic freeness, in the last section of these notes, where we will describe in detail some applications of this property.

4. THE USE OF NON-BACKTRACKING OPERATORS

4.1. Ihara–Bass formula and applications

For a matrix $H := (H_{ij})_{i,j\in[n]} \in M_n(\mathbb{C})$ with complex entries, the *non-backtracking matrix* associated to H is $B := (B_{ej})_{e,f\in[n]^2}$ defined, for $e = (i, j)$ and $f = (k, \ell)$, by

$$B_{ef} = H_{k\ell}\delta_{j,k}(1 - \delta_{i,\ell}).$$

The name *non-backtracking* comes from the following interpretation: if H is the adjacency matrix of a graph G, then, for any $\ell \in \mathbb{N}^*$, $(H^m)_{ij}$ is the number of paths from i to j of length m in G, whereas, if $e = (i, j)$ and $f = (k, \ell)$, $(B^m)_{ef}$ is the number of non-backtracking paths of length m starting with the edge e and ending with the edge f. A non-backtracking path may have cycles but cannot immediately go back to the vertex it comes from. Although the non-backtracking operator is non-normal even when H is self-adjoint, it is a powerful tool for the study of spectral properties of random graphs. The Ihara–Bass formula allows to link the spectrum of H with the spectrum of its non-backtracking counterpart. This statement and its proof can be found e.g. in (Benaych-Georges, Bordenave, and Knowles, 2020).

Proposition 4.1 (Ihara–Bass formula). — *Let $\lambda \in \mathbb{C}$ be such that $\lambda^2 \neq H_{ij}H_{ji}$, for $i, j \in [n]$, and define*

$$H_{ij}(\lambda) := \frac{\lambda H_{ij}}{\lambda^2 - H_{ij}H_{ji}} \text{ and } m_i(\lambda) := 1 + \sum_{k=1}^{N} \frac{H_{ik}H_{ki}}{\lambda^2 - H_{ik}H_{ki}}.$$

Then $\lambda \in \sigma(B)$ if and only if $\det(M(\lambda) - H(\lambda)) = 0$, where $M(\lambda)$ is the diagonal matrix with non-zero entries equal to $m_1(\lambda), \dots, m_n(\lambda)$.

In the model of Bordenave and Collins (2019), A_n can be seen as the adjacency matrix of a weighted graph where each vertex carries a loop edge with weight a_0 and for each $(x, i) \in [n] \times [d]$, we draw an edge between x and $\sigma_i(x)$ with weight a_i. The associated non-backtracking operator can be written

$$B_n = \sum_{j\neq i^*} a_j \otimes S_{i,n} \otimes E_{ij}.$$

This means that, if $e = (x, i)$ and $f = (y, j)$,

$$(B_n)_{ef} = a_j\delta_{y,\sigma_i(x)}(1 - \delta_{j,i^*}).$$

Therefore, if $\gamma = (\gamma_1, \dots, \gamma_k)$ is a path with $\gamma_t = (x_t, i_t)$, we denote by $a(\gamma) := \prod_{t=1}^k a_{i_t}$. The path γ is non-backtracking if $\forall t \in [k-1], i_{t+1} \neq i_t^*$. For $e, f \in [n] \times [d]$, let Γ_{ef}^k be the set of non-backtracking paths of length k such that $\gamma_1 = e$ and $\gamma_k = f$, so that

$$(B^k)_{ef} = \sum_{\gamma \in \Gamma_{ef}^{k+1}} a(\gamma) \prod_{t=1}^{k} (S_{i,n})_{x_t x_{t+1}}. \tag{6}$$

The Ihara–Bass formula reads:

Proposition 4.2 (Ihara–Bass - symmetric random permutation model)
Let $\lambda \notin \{\sigma(a_i a_{i^*}), i \in [d]\}$, and define $A_{n,\lambda} := a_0(\lambda) + \sum_{i=1}^d a_i(\lambda) \otimes S_{i,n}$, with

$$a_i(\lambda) = \lambda a_i (\lambda^2 - a_{i^*} a_i)^{-1},$$

and

$$a_0(\lambda) := -1 - \sum_{i=1}^d a_i (\lambda^2 - a_{i^*} a_i)^{-1} a_{i^*}.$$

Then $\lambda \in \sigma(B_n)$ if and only if $0 \in \sigma(A_{n,\lambda})$.

To exploit this relation for the study of the spectrum of A_n, we perform some reverse engineering: for a given μ, we want to construct an operator $A_{n,\mu}$ of the same form as A_n and its non-backtracking operator $B_{n,\mu}$ such that

$$\mu \in \sigma(A_n) \text{ if and only if } 1 \in \sigma(B_{n,\mu}). \tag{7}$$

This is in fact possible if $\mu \notin \mathrm{full}(\sigma(\mathrm{A}_*))$. For a bounded D, $\mathrm{full}(D) = \mathbb{C} \setminus U$, where U is the unique infinite component of $\mathbb{C} \setminus D$[(7)]. For such a μ, we now explain the recipe to construct $A_{n,\mu}$ and therefore $B_{n,\mu}$ satisfying (7). Let $G(\mu) := (\mu - A_*)^{-1}$ be the resolvent operator of A_*. This is also an operator on $\mathbb{C}^r \otimes \ell^2(X_*)$. It can be seen as an infinite block matrix where each block is an $r \times r$ matrix indexed by a pair of indices $x, y \in X_* \times X_*$. We let

$$\hat{a}_i(\mu) := G_{ee}(\mu)^{-1} G_{eg_i}(\mu),$$

with e the neutral element in X_* and $g_1, \dots, g_d$ the generators, and define

$$A_{n,\mu} := \sum_{i=1}^d \hat{a}_i(\mu) \otimes S_{i,n} \text{ and } B_{n,\mu} := \sum_{j \neq i^*} \hat{a}_j(\mu) \otimes S_{i,n} \otimes E_{ij}.$$

Then (7) holds.

We also denote by

$$A_{*,\mu} := \sum_{i=1}^d \hat{a}_i(\mu) \otimes \lambda(g_i) \text{ and } B_{*,\mu} := \sum_{j \neq i^*} \hat{a}_j(\mu) \otimes \lambda(g_i) \otimes E_{ij}.$$

For any operator T, we denote by $\rho(T) := \sup\{|\lambda|, \lambda \in \sigma(T)\}$ its spectral radius. One can show that, for $\mu \notin \mathrm{full}(\sigma(\mathrm{A}_*))$, $\rho(B_{*,\mu}) < 1$. This leads to the following criterion for comparing the spectrum of A_n and A_*:

(7) $\mathrm{full}(D)$ “fills the holes” of D.

Theorem 4.3. — *For any* $\varepsilon > 0$, $\exists \delta > 0$, *such that*

$$\textit{if for all } \mu \in \mathbb{C}, \rho(B_{n,\mu}) < \rho(B_{*,\mu}) + \delta,$$

then $\mathrm{full}(\sigma(A_n))$ *is in an* ε*-neighborhood of a slight modification*[(8)] *of* $\sigma(A_*)$*. Moreover,if we define* $K_0 := (\mathbb{C}^r \otimes \mathbf{1} \otimes \mathbb{C}^d)^\perp = \{f, \sum_x f(x,i) = 0, \forall i \in [d]\}$*, then the same holds true if we replace* $B_{n,\mu}$ *by* $(B_{n,\mu})_{|K_0}$ *and* A_n *by* $(A_n)_{|H_0}$*.*

Therefore, to get (4), it is enough to show the following:

Theorem 4.4. — *For any* $\varepsilon > 0$, *under Assumption 2.5,*

$$\mathbb{P}\left(\forall a_i \textit{ such that } \max(\|a_i\| \vee \|a_i^{-1}\|^{-1}) \leq 1/\varepsilon, \rho((B_n)_{|K_0}) \leq \rho(B_*) + \varepsilon\right) \xrightarrow[n\to\infty]{} 1.$$

4.2. A glimpse of the Füredi–Komlós moment method

The core of the proof to show Theorem 4.4 is to use a moment method. In RMT, moment methods were first used to study the empirical spectral measure. This goes back at least to the pioneering work of Wigner. For a random matrix M_n, with eigenvalues $\lambda_{1,n}, \ldots, \lambda_{n,n}$, we denote by $\hat{\mu}_n := \frac{1}{n}\sum_{i=1}^n \delta_{\lambda_{i,n}}$ the empirical measure of its eigenvalues. The main idea is to rewrite the moments of this spectral empirical measure $m_n(k) := \int x^k \mathrm{d}\hat{\mu}_n$ in terms of traces of powers of the matrix M_n, namely $m_n(k) = \mathrm{tr}_n(M_n^k)$, where tr_n is the normalized trace on $M_n(\mathbb{C})$. We then expand

$$\mathbb{E}\left[\mathrm{tr}_n(M_n^k)\right] = \frac{1}{n}\sum_{i_1,\ldots,i_k \in [n]} \mathbb{E}\left(M_{i_1 i_2} \cdots M_{i_k i_1}\right),$$

look at the sequence $(i_1, \ldots, i_k)$ as a path $\gamma_1 = (i_1, i_2), \ldots, \gamma_k = (i_k, i_1)$ and identify the contributions to the sum according to the geometric or combinatorial properties of the paths. For example, in the case when $M_n \in \mathrm{GUE}(n)$, the paths contributing to the sum are the contours of rooted trees, leading to limiting moments given by Catalan numbers.

Later, the method has been adapted by Füredi and Komlós (1981) to study the spectral radius or the spectral norm of the matrix M_n. The idea is the following: to capture the behavior of the largest eigenvalue, one has to compute moments of order k_n growing with the size n of the matrix. Indeed, for $\rho > 0$, if we want to show that, for any $\varepsilon > 0$,

$$\mathbb{P}(\|M_n\| \leq \rho(1+\varepsilon)) \xrightarrow[n\to\infty]{} 1,$$

we write

$$\|M_n\|^{2k_n} = \|M_n M_n^*\|^{k_n} \leq n\mathrm{tr}_n((M_n M_n^*)^{k_n}),$$

and we might lose a factor n in the last bound. Now, if we choose $k_n \gg \log n$ and show that

$$\mathbb{E}\left[\mathrm{tr}_n((M_n M_n^*)^{k_n})\right] \leq \rho^{2k_n}(1+\varepsilon)^{2k_n},$$

(8) We do not detail it here. For more details, see the definition of $\mathrm{full}(\hat{\sigma}(A_*))$ in Bordenave and Collins (2019)

then, for any $\delta > 0$,

$$\mathbb{P}(\|M_n\| \leq \rho(1+\varepsilon)(1+\delta)) \geq 1 - \frac{n\mathbb{E}\left[\operatorname{tr}_n((M_n M_n^*)^{k_n})\right]}{(1+\delta)^{2k_n}}$$
$$\geq 1 - \exp\left(-2k_n \log(1+\delta) + \log n\right) \xrightarrow[n\to\infty]{} 1.$$

This method has been very successful and has led in particular to the first universality results for the Tracy–Widom distribution by Soshnikov (1999). However, in the symmetric permutation model, it is hopeless to apply directly this method to the operators $(A_n)_{|H_0}$ but one can circumvent this obstacle by applying it to the non-backtracking operators.

4.3. High order moments of the non-backtracking operators

The moment method for the non-backtracking operator B_n is very involved, it is the main technical part and we do not intend here to give a precise account of the proof. We rather want to explain what are the main steps and the main difficulties. In this paragraph, we will omit the n subscript to lighten the notation a bit.

Following the method we have just explained above, to prove Theorem 4.4, we need to bound the spectral radius by the spectral norm of powers of the operator. More precisely, recalling that $K_0 = (\mathbb{C}^r \otimes \mathbf{1} \otimes \mathbb{C}^d)^\perp$, we will use that

$$\rho(B_{|K_0}) = \rho(B^\ell_{|K_0})^{1/\ell} \leq \sup_{g\in K_0, \|g\|_2=1} \|B^\ell g\|_2^{1/\ell}.$$

Restricting B to the subspace K_0 boils down to performing an orthogonal projection of the operators S_i onto $\mathbf{1}^\perp$. More explicitely, if for every $i \in [d]$, $\underline{S}_i := S_i - \frac{1}{n}\mathbf{1} \otimes \mathbf{1}$ is such a projection, we define

$$\underline{B} := \sum_{j \neq i^*} a_j \otimes \underline{S}_i \otimes E_{ij}$$

and one can check that, for any $g \in K_0$, $B^\ell g = \underline{B}^\ell g$.

But the matrix $\underline{B}$ won't be used directly. Before projecting on K_0, we will replace B^ℓ by a matrix $B^{(\ell)}$ that coincide with B^ℓ with high probability but has better properties. This step is known as *removing the tangles*, and will be described right after. We will then project on K_0, that is consider $\underline{B^{(\ell)}}$. However, as K_0 is not necessarily invariant under $B^{(\ell)}$, there will be some remainder term. Let us describe these steps more precisely. We first recall from (6) that, for any $\ell \in \mathbb{N}^*$,

$$(B^\ell)_{ef} = \sum_{\gamma \in \Gamma^{\ell+1}_{ef}} a(\gamma) \prod_{t=1}^{\ell} (S_i)_{x_t x_{t+1}},$$

with $\Gamma^{\ell+1}_{ef}$ the set of non-backtracking paths of length $\ell+1$ such that $\gamma_1 = e$ and $\gamma_{\ell+1} = f$. For a path $\gamma \in \Gamma^{\ell+1}_{ef}$, let G_γ be the graph with vertices $V_\gamma := \{x_t, t \in [\ell+1]\}$ and edges $E_\gamma := \{[x_t, i_t, x_{t+1}], t \in [\ell]\}$, where i_t can be seen as the *color* of the edge $[x_t, i_t, x_{t+1}]$. We can now define the notion of *tangle*.

Definition 4.5 (Tangles). — *A graph H is* tangle-free *if it contains at most one cycle. For any $\ell \in \mathbb{N}^*$, a graph H is* ℓ-tangle-free *if, for every vertex x, $(H,x)_\ell$ contains at most one cycle, where $(H,x)_\ell$ is the subgraph of H restricted to the vertices at distance at most ℓ from x for the graph distance. We say that a path γ is tangle-free if G_γ is.*

We denote by F^k (respectively F^k_{ef}) the subset of tangle-free paths in Γ^k (resp. Γ^k_{ef}). For ℓ fixed, we denote by

$$(B^{(\ell)})_{ef} := \sum_{\gamma \in F^{\ell+1}_{ef}} a(\gamma) \prod_{t=1}^{\ell} (S_i)_{x_t x_{t+1}}.$$

If the permutations $\sigma_1, \ldots, \sigma_d$ satisfy Assumption (H$\boldsymbol{\sigma}$), we denote by $G^{\boldsymbol{\sigma}}$ the graph whose vertex set is $[n]$ and whose edges are $[x,i,y]$ such that $\sigma_i(x) = y$ and $\sigma_{i^*}(y) = x$. For any $\ell \in \mathbb{N}^*$, if $G^{\boldsymbol{\sigma}}$ is ℓ-tangle-free, then $\forall 0 \le k \le 2\ell$, $B^k = B^{(k)}$. As above, if we denote by

$$(\underline{B^{(\ell)}})_{ef} := \sum_{\gamma \in F^{\ell+1}_{ef}} a(\gamma) \prod_{t=1}^{\ell} (\underline{S}_i)_{x_t x_{t+1}},$$

then even if $G^{\boldsymbol{\sigma}}$ is ℓ-tangle-free, we have a priori that $\underline{B}^k \neq \underline{B^{(k)}}$. Let us write down more explicitely the difference between the two quantities. If we set

$$\overline{B} := \sum_{j \neq i^*} a_j \otimes (\mathbf{1} \otimes \mathbf{1}) \otimes E_{ij},$$

one can check that

$$B^{(\ell)} = \underline{B^{(\ell)}} + \frac{1}{n} \sum_{k=1}^{\ell} \underline{B^{(k-1)}}\, \overline{B} B^{(\ell-k)} - \frac{1}{n} \sum_{k=1}^{\ell} R_k^{(\ell)}, \tag{8}$$

where $R_k^{(\ell)}$ can be thought of as an error term equal to

$$(R_k^{(\ell)})_{ef} = \sum_{\gamma \in F^{\ell+1}_{k,ef} \setminus F^{\ell+1}_{ef}} a(\gamma) \Big(\prod_{t=1}^{k-1} (\underline{S}_i)_{x_t x_{t+1}} \Big) \Big(\prod_{t=k+1}^{\ell} (S_i)_{x_t x_{t+1}} \Big),$$

where $F_k^{\ell+1}$ is the set of paths of length $\ell+1$ that can be decomposed in $\gamma' \in F^k, \gamma'' \in F^2$ and $\gamma''' \in F^{\ell-k+1}$ and $F^{\ell+1}_{k,ef} = F^{\ell+1}_k \cap \Gamma^{\ell+1}_{ef}$. Note that the concatenation of three tangle-free paths is not necessarily tangle-free.

Now, if $G^{\boldsymbol{\sigma}}$ is ℓ-tangle-free, then one can check that the second term in (8) cancels on K_0, thus, for any $g \in K_0$,

$$B^\ell g = B^{(\ell)} g = \underline{B^{(\ell)}} g - \frac{1}{n} \sum_{k=1}^{\ell} R_k^{(\ell)} g,$$

so that if $G^{\boldsymbol{\sigma}}$ is ℓ-tangle-free,

$$\rho(B_{|K_0}) \le \Big(\|\underline{B^{(\ell)}}\| + \frac{1}{n} \sum_{k=1}^{\ell} \|R_k^{(\ell)}\| \Big)^{1/\ell}.$$

Remember that, in the Füredi–Komlós method, we need to choose $\ell = \ell_n$ growing with n. The next natural question to ask is: for which values of ℓ_n is the probability that $G^{\boldsymbol{\sigma}}$ is ℓ_n-tangle-free large enough ? This is a nice question on random permutations that boils down to estimating the expected number of cycles of length ℓ in $G^{\boldsymbol{\sigma}}$. Through such an estimate, Bordenave and Collins (2019) got the following:

LEMMA 4.6. — *For random permutations satisfying* (H$\boldsymbol{\sigma}$), *there exists* $c > 0$ *such that* $\forall 1 \leq \ell \leq \sqrt{n}$,

$$\mathbb{P}\left(G^{\boldsymbol{\sigma}} \text{ is } \ell\text{-tangled }\right) \leq c\ell^3 \frac{(d-1)^{4\ell}}{n}.$$

Then, for $\ell = \ell_n$ large enough, we need to show that $\|\underline{B^{(\ell)}}\|$ is close to $\rho(B_*)^\ell$ and that $\|R_k^{(\ell)}\|$ is negligible. The adequate controls can be stated as follows:

PROPOSITION 4.7. — *Let* $(a_i)_{i\in[d]}$ *be fixed, satisfying the symmetry condition* (H$\boldsymbol{a}$) *and assume that* $\max(\|a_i\| \vee \|a_i^{-1}\|^{-1}) \leq 1/\varepsilon$. *Then,* $\exists c, \rho_1 > 0, \forall 1 \leq \ell \leq \log n$,

$$\mathbb{P}\left(\|\underline{B^{(\ell)}}\| \leq (\log n)^{20}(\rho(B_*)+\varepsilon)^\ell\right) \geq 1 - c\exp\Big(-\frac{\ell \log n}{c \log\log n}\Big),$$
$$\mathbb{P}\left(\|R_k^{(\ell)}\| \leq (\log n)^{40}\rho_1^\ell\right) \geq 1 - c\exp\Big(-\frac{\ell \log n}{c \log\log n}\Big).$$

Now, by choosing $\ell_n \sim \frac{\log n}{\kappa}$, with $\kappa > 1$, satisfying $\kappa > \log\left((d-1)^4 \vee \left(\frac{4\rho_1}{\varepsilon}\right)\right)$ and using a net argument on the a_i's that we do not detail here, one can get the required bound (4). This concludes the proof of Theorem 3.13. As explained above, through the linearization trick, we get Theorem 3.16 as a corollary.

5. APPLICATIONS OF STRONG ASYMPTOTIC FREENESS

To summarize, the construction of a sequence of almost-Ramanujan (colored, weighted) graphs for which $(A_n)_{n\geq 1}$ plays the role of (generalized) adjacency operators is equivalent to strong asymptotic freeness for a family of permutation matrices. We conclude this presentation by listing a few other results involving strong convergence or strong asymptotic freeness and their consequences in other domains.

5.1. Reminder on the link between strong asymptotic freeness and outliers of random matrices

In Section 3.4, we have defined strong asymptotic freeness and clarified in Proposition 3.15 its relation to outliers of random matrices. To illustrate this link, we will cite some results of Collins and Male (2014) in which they use strong asymptotic freeness to show the absence of outliers for some ensembles of random matrices.

Proposition 5.1. — *Let $\mathbf{U}_n$ be a p-uple of $n \times n$ independent Haar unitary matrices and $\mathbf{Y}_n$ a q-uple of $n \times n$ matrices that are deterministic or random but independent of $\mathbf{U}_n$. Let $\mathbf{u}$ be a p-uple of Haar unitaries and $\mathbf{y}$ a q-uple of random variables, freely independent from $\mathbf{u}$ in a C^*-algebra $(\mathcal{A}, \tau)$. If $(\mathbf{Y}_n)_{n\geq 1}$ converges strongly to $\mathbf{y}$, then $(\mathbf{U}_n, \mathbf{Y}_n)_{n\geq 1}$ converges strongly to $(\mathbf{u}, \mathbf{y})$.*

As a immediate corollary, they got the following result:

Corollary 5.2. — *Let A_n, B_n be two $n \times n$ independent Hermitian random matrices. Assume that:*

– *the law of one of the matrices is invariant under unitary conjugacy,*
– *almost surely, the empirical spectral measure of A_n (respectively B_n) converges to a compactly supported probability measure μ (respectively ν),*
– *almost surely, for any neighborhood of the support of μ (respectively ν), for n large enough, the eigenvalues of A_n (respectively B_n) belong to the respective neighborhood.*

Then one has that 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[9]*.*

If the third condition in the corollary is not fulfilled, then outliers may appear. This phenomenon has been extensively studied in the RMT framework and most results can be understood through free probability theory, but with different tools. In this direction, we strongly recommend the review paper on deformed models by Capitaine and Donati-Martin (2017).

5.2. $\mathrm{Ext}(C^*_{\mathrm{red}}(F_r))$ is not a group

As mentioned at the beginning of the paper, there has been a constant interplay between RMT and operator algebra. It is in particular striking to see that the first result explicitely involving strong convergence and strong asymptotic freeness was a paper by Haagerup and Thorbjørnsen (2005) entitled *A new application of random matrices:* $\mathrm{Ext}(C^*_{\mathrm{red}}(F_2))$ *is not a group*. This paper is a very important step in the theory for several reasons. One of them is that it introduced the linearization trick that we have presented above and that has been used since then in all the strong convergence results we are aware of. Although it is a bit far from the motivation of Bordenave and Collins (2019), we will describe in detail this result, mainly for historical reasons. We start by giving a definition of $\mathrm{Ext}(C^*_{\mathrm{red}}(F_r))$. We start from the Hilbert space $\mathcal{H} = \ell^2(\mathbb{N})$. Its Calkin algebra is $\mathcal{C}(\mathcal{H}) := \mathcal{B}(\mathcal{H})/\mathcal{K}(\mathcal{H})$, which is quotient of the space $\mathcal{B}(\mathcal{H})$ of bounded operators on $\mathcal{H}$ by the set $\mathcal{K}(\mathcal{H})$ of compact operators, the

[9] We don't want to get into the detail of the definition of this operation on measures. Here, we just need to know that $\mu \boxplus \nu$ is the distribution of $a + b$, where a has distribution μ, b has distribution ν and a and b are freely independent.

quotient map being denoted by $q\colon \mathcal{B}(\mathcal{H}) \to \mathcal{C}(\mathcal{H})$. We also denote by $\mathcal{U}(\mathcal{H})$ the unitary group of $\mathcal{B}(\mathcal{H})$. For $\mathcal{A}$ a C^*-algebra, if π_1 and π_2 are two one-to-one $*$-homomorphisms from $\mathcal{A}$ to $\mathcal{C}(\mathcal{H})$, we define the equivalence relation

$$\pi_1 \sim \pi_2 \iff \exists u \in \mathcal{U}(\mathcal{H}), \forall a \in \mathcal{A}, \pi_2(a) = q(u)\pi_1(a)q(u)^*.$$

Then $\mathrm{Ext}(\mathcal{A})$ is the set of equivalence classes for this equivalence relation. As $\mathcal{H}\oplus\mathcal{H} \simeq \mathcal{H}$, $(\pi_1, \pi_2) \mapsto \pi_1 \oplus \pi_2$ defines a natural semigroup structure on $\mathrm{Ext}(\mathcal{A})$. The question to know under which condition on $\mathcal{A}$ it would be a group was controversial in operator algebra in the nineties. Voiculescu obtained that $\mathrm{Ext}(\mathcal{A})$ was a unital semigroup for all separable unital C^*-algebras $\mathcal{A}$. Anderson (1978) provided the first example for which $\mathrm{Ext}(\mathcal{A})$ is not a group and, if F_r denotes the free group with r generators, the question for $\mathcal{A} = C^*_{\mathrm{red}}(F_r)$ remained open for a long time. Voiculescu (1993) gave the following very useful criterion:

Theorem 5.3. — *If there exists a sequence of unitary representations $\pi_n\colon F_r \to M_n(\mathbb{C})$ such that $\forall h_1, \ldots, h_m \in F_r$ and $c_1, \ldots, c_m \in \mathbb{C}$,*

$$\lim_{n\to\infty}\Big\|\sum_{j=1}^{m} c_j\pi_n(h_j)\Big\| = \Big\|\sum_{j=1}^{m} c_j\lambda(h_j)\Big\|, \tag{9}$$

then $\mathrm{Ext}(C^*_{\mathrm{red}}(F_r))$ is not a group.

Now, let us explain how strong asymptotic freeness for independent GUE matrices (shown in (Haagerup and Thorbjørnsen, 2005)) implies that (9) holds true. The idea is to construct explicitely the sequence $(\pi_n)_{n\geq 1}$ as follows: we let

$$\varphi(t) = \begin{cases} -\pi, & \text{if } t \leq -2 \\ \pi, & \text{if } t \geq 2 \\ \int_0^t \sqrt{4-s^2}\,\mathrm{d}s & \text{if } -2 \leq t \leq 2 \end{cases}$$

and $\psi(t) := \exp(i\varphi(t))$. If $(s_i)_{i\in[r]}$ is a family of semicircular elements that are freely independent and $u_i = \psi(s_i)$, then there is an isomorphism $\Phi\colon C^*_{\mathrm{red}}(F_r) \to C^*((u_i)_{i\in[r]})$ such that $\Phi(\lambda(g_i)) = u_i$. If $(X_{1,n}, \ldots, X_{r,n})$ are independent $\mathrm{GUE}(n)$ matrices, and $\forall i \in [r]$, $U_{i,n}(\omega) = \psi(X_{i,n}(\omega))$, we obtain a sequence of unitary matrices and, for any $\omega \in \Omega$, there exists $\pi_{n,\omega}\colon F_r \to \mathcal{U}(M_n(\mathbb{C}))$ such that $\pi_{n,\omega}(g_i) = U_{i,n}(\omega)$. Then using strong asymptotic freeness, one can check that $\forall \omega \in \Omega$, $\forall h_1, \ldots, h_m \in F_r$ and $c_1, \ldots, c_m \in \mathbb{C}$,

$$\lim_{n\to\infty}\Big\|\sum_{j=1}^{m} c_j\pi_{n,\omega}(h_j)\Big\| = \Big\|\sum_{j=1}^{m} c_j\lambda(h_j)\Big\|,$$

and it is enough to choose $\pi_n = \pi_{n,\omega}$, with ω in the set of probability 1 onto which this last equality holds. Some generalizations of the results of Haagerup and Thorbjornssen have been conjectured, in relation to the Peterson–Thom conjecture (see e.g. the recent work of Hayes (2020)).

5.3. Estimation of the norm of random matrices

In this last part, we explain how strong convergence results can be used to give interesting bounds on the norm of random matrices. We will present asymptotic bounds that are in majority consequences of Proposition 5.1, but also remarkable non-asymptotic bounds that have been recently obtained by Bandeira, Boedihardjo, and van Handel (2021). Indeed, strong convergence implies convergence of the norm of polynomials in random matrices to their free counterpart and there are several examples for which the norm of the free counterpart has been computed. In particular, Akemann and Ostrand (1976) showed that, if $u_1, \ldots, u_p$ are Haar unitaries that are freely independent, then, for any $a_1, \ldots, a_p \in \mathbb{R}$, we have

$$\left\| \sum_{i=1}^{p} a_i u_i \right\| = \min_{t \geq 0} \left\{ 2t + \sum_{i=1}^{p} \left(\sqrt{t^2 + |a_i|^2} - t \right) \right\}. \tag{10}$$

In particular,

$$\left\| \sum_{i=1}^{p} u_i \right\| = 2\sqrt{p-1}.$$

We can therefore deduce that, if $U_{1,n}, \ldots, U_{p,n}$ are independent Haar unitary random matrices, then almost surely,

$$\left\| \sum_{i=1}^{p} U_{i,n} \right\| \xrightarrow[n \to \infty]{} 2\sqrt{p-1},$$

which is like a unitary analogue of the Alon–Boppana bound given in Theorem 2.1. In the same vein, motivated by questions for random walks on the free group, Kesten (1959) showed that

$$\left\| \sum_{i=1}^{p} (u_i + u_i^*) \right\| = 2\sqrt{2p-1},$$

so that

$$\left\| \sum_{i=1}^{p} (U_{i,n} + U_{i,n}^*) \right\| \xrightarrow[n \to \infty]{} 2\sqrt{p-1}.$$

Note that Lehner (1999) also gave a formula for the norm of the operator A_*, which is a generalisation of (10) to the case when $a_0, \ldots, a_d$ are matrices.

As a conclusion, we now present remarkable non-asymptotic bounds, that is concentration inequalities for random matrices, recently obtained Bandeira, Boedihardjo, and van Handel (2021). Their initial motivation is to understand the spectral norm of an arbitrary $d \times d$ self-adjoint random matrix with centered, jointly Gaussian entries. Such a matrix X can be written

$$X := \sum_{i=1}^{p} g_i A_i, \tag{11}$$

where A_i are deterministic self-adjoint $d \times d$ matrices and $(g_i)_{i\in[p]}$ are independent real standard Gaussian variables. It is known that, if we define $\Sigma(X) := \|\sum_{i=1}^p A_i^2\|$, then we have the bound

$$c\Sigma(X) \leq \mathbb{E}\|X\| \leq C\Sigma(X)\sqrt{\log d}.$$

Their goal is to improve the upper bound. To get a free analogue of the matrix (11), a natural idea is to replace the Gaussian variables by semi-circular elements to define

$$X_{\mathrm{free}} := \sum_{i=1}^{p} A_i \otimes s_i,$$

where $(s_i)_{i\in[p]}$ are freely independent semicircular elements. The authors establish a general bound of the form:

$$\mathbb{E}\|X\| \leq \|X_{\mathrm{free}}\| + Cv(X)^{1/2}\Sigma(X)^{1/2}(\log d)^{3/4}, \tag{12}$$

with $v(X)^2 = \|\mathrm{Cov}(X)\|$ being the spectral norm of the covariance matrix of X. As a consequence, they got an inclusion of the spectrum which is reminiscent of (4): with high probability,

$$\sigma(X) \subset \sigma(X_{\mathrm{free}}) + [-\varepsilon, \varepsilon],$$

where ε is of order $v(X)^{1/2}\Sigma(X)^{1/2}(\log d)^{3/4}$. The bound is particularly relevant when ε is small in comparison to $\|X_{\mathrm{free}}\|$. They treat a large variety of examples for which the bound (12) improves on known results (random matrices with independent entries, sparse Wigner matrices etc.) or gives new concentration inequalities (patterned random matrices, independent block matrices etc.) From these non-asymptotic bounds, they can also deduce strong asymptotic freeness for a lot of models, showing that the latter is much more ubiquitous than expected. The interplay between strong asymptotic freeness and random matrix theory is certainly to be continued.

REFERENCES

Charles A. Akemann and Phillip A. Ostrand (1976). "Computing norms in group C^*-algebras", *Amer. J. Math.* **98** (4), pp. 1015–1047.

Noga Alon (1986). "Eigenvalues and expanders", in: vol. 6. 2. Theory of computing (Singer Island, Fla., 1984), pp. 83–96.

Alon Amit and Nathan Linial (2002). "Random graph coverings. I. General theory and graph connectivity", *Combinatorica* **22** (1), pp. 1–18.

Joel Anderson (1978). "A C^*-algebra $\mathcal{A}$ for which $\mathrm{Ext}(\mathcal{A})$ is not a group", *Ann. of Math. (2)* **107** (3), pp. 455–458.

Afonso S. Bandeira, March T. Boedihardjo, and Ramon van Handel (2021). *Matrix Concentration Inequalities and Free Probability*. Version 1. arXiv: 2108.06312.

Florent Benaych-Georges, Charles Bordenave, and Antti Knowles (2020). "Spectral radii of sparse random matrices", *Ann. Inst. Henri Poincaré Probab. Stat.* **56** (3), pp. 2141–2161.

Charles Bordenave (2020). "A new proof of Friedman's second eigenvalue theorem and its extension to random lifts", *Ann. Sci. Éc. Norm. Supér. (4)* **53** (6), pp. 1393–1439.

Charles Bordenave and Benoît Collins (2019). "Eigenvalues of random lifts and polynomials of random permutation matrices", *Ann. of Math. (2)* **190** (3), pp. 811–875.

Mireille Capitaine and Catherine Donati-Martin (2017). "Spectrum of deformed random matrices and free probability", in: *Advanced topics in random matrices*. Vol. 53. Panor. Synthèses. Soc. Math. France, Paris, pp. 151–190.

Benoît Collins and Camille Male (2014). "The strong asymptotic freeness of Haar and deterministic matrices", *Ann. Sci. Éc. Norm. Supér. (4)* **47** (1), pp. 147–163.

Joel Friedman (2008). "A proof of Alon's second eigenvalue conjecture and related problems", *Mem. Amer. Math. Soc.* **195** (910), pp. viii+100.

Zoltan Füredi and János Komlós (1981). "The eigenvalues of random symmetric matrices", *Combinatorica* **1** (3), pp. 233–241.

Yonatan Greenberg (1995). "Spectra of graphs and their covering trees". in Hebrew. PhD thesis. Hebrew University of Jerusalem.

Uffe Haagerup and Steen Thorbjørnsen (2005). "A new application of random matrices: $\mathrm{Ext}(C^*_{\mathrm{red}}(F_2))$ is not a group", *Ann. of Math. (2)* **162** (2), pp. 711–775.

Ben Hayes (2020). *A random matrix approach to the Peterson–Thom conjecture*. Version 2. arXiv: 2008.12287.

Harry Kesten (1959). "Symmetric random walks on groups", *Trans. Amer. Math. Soc.* **92**, pp. 336–354.

Emmanuel Kowalski (2019). *An introduction to expander graphs*. Vol. 26. Cours Spécialisés. Société Mathématique de France, Paris, pp. x+276.

Franz Lehner (1999). "Computing norms of free operators with matrix coefficients", *Amer. J. Math.* **121** (3), pp. 453–486.

Alexander Lubotzky, Ralph Phillips, and Peter Sarnak (1988). "Ramanujan graphs", *Combinatorica* **8** (3), pp. 261–277.

James A. Mingo and Roland Speicher (2017). *Free probability and random matrices*. Vol. 35. Fields Institute Monographs. Springer, New York; Fields Institute for Research in Mathematical Sciences, Toronto, pp. xiv+336.

Alexandru Nica (1993). "Asymptotically free families of random unitaries in symmetric groups", *Pacific J. Math.* **157** (2), pp. 295–310.

Alexander Soshnikov (1999). "Universality at the edge of the spectrum in Wigner random matrices", *Comm. Math. Phys.* **207** (3), pp. 697–733.

Roland Speicher (2019). *Lecture Notes on "Free Probability Theory"*. Version 1. arXiv: 1908.08125.

Dan Voiculescu (1991). "Limit laws for random matrices and free products", *Invent. Math.* **104** (1), pp. 201–220.

——— (1993). "Around quasidiagonal operators", *Integral Equations Operator Theory* **17** (1), pp. 137–149.

Mylène Maïda

Univ. Lille, CNRS, UMR 8524
Laboratoire Paul Painlevé,
F-59000 Lille, France
E-mail : mylene.maida@univ-lille.fr