A MATRIX INTERPOLATION BETWEEN CLASSICAL AND FREE MAX OPERATIONS.
I. THE UNIVARIATE CASE

FLORENT BENAYCH-GEORGES AND THIERRY CABANAL-DUVILLARD

Abstract. Recently, Ben Arous and Voiculescu considered taking the maximum of two free random variables and brought to light a deep analogy with the operation of taking the maximum of two independent random variables. We present here a new insight on this analogy: its concrete realization based on random matrices giving an interpolation between classical and free settings.

Contents

1. Introduction

The free probability theory has been a very active field in mathematics over the last two decades, constructed in a deep analogy with classical probability theory. Nowadays, there is an unofficial dictionary

Date: February 23, 2009.
MSC 2000 subject classifications. 15A52, 46L54
Key words. random matrices, free probability, max-stable laws

of concepts in both theories: many fundamental notions or results of classical probability theory, such as Law of Large Numbers, Central Limit Theorem, Gaussian distribution, convolution, cumulants, infinite divisibility have their precise counterpart in free probability theory.

Recently, Ben Arous and Voiculescu [2] have added a new item in this dictionary, bringing to light the fact that the operation called the *classical upper extremal convolution*, which associates the distribution of the supremum $X \vee Y$ to the distributions of two independent random variables X and Y, has an analogue in free probability theory, called the *free upper extremal convolution*, which associates the distribution of the supremum $X \vee Y$ to the distributions of two free random variables X and Y. From the point of view of cumulative distribution functions[1], both of these operations have a concise interpretation: the classical (resp. free) upper extremal convolution of two probability measures with distribution functions F, G is the probability measure with distribution function FG (resp. $\max(0, F+G-1)$). The purpose of this paper is to construct random matrix models providing a better understanding of the free upper extremal convolution and a new insight on the relation between classical and free upper extremal convolutions.

Let us first notice the following fact (which is an immediate consequence of remark 2.1): for μ, ν probability measures on the real line, for $(X_i)_{i\geq 1}$ and $(Y_i)_{i\geq 1}$ two independent families of independent identically distributed random variables, with respective distributions μ, ν, if for each $N \geq 1$, one denotes by $Z_{N,1} \geq Z_{N,2} \geq \cdots \geq Z_{N,N}$ the N largest elements of the multiset $(X_1, \ldots, X_N, Y_1, \ldots, Y_N)$, then the empirical probability measure $\frac{1}{N}\sum_{i=1}^{N} \delta_{Z_{N,i}}$ converges almost surely to the free upper extremal convolution of μ and ν as N tends to infinity. This result may appear as a coincidence. But, if we consider $(X_1, \ldots, X_N)$ and $(Y_1, \ldots, Y_N)$ as the eigenvalues of two independent random matrices $X^{(N)}$ and $Y^{(N)}$ which are invariant in law under conjugation by any unitary matrix, then $(Z_{N,1}, \ldots, Z_{N,N})$ are the eigenvalues of the supremum[2] of $X^{(N)}$ and $Y^{(N)}$ with respect to the spectral order introduced by Olson [9] and used by Ben Arous and Voiculescu in their paper. Since $X^{(N)}$ and $Y^{(N)}$ are known to be asymptotically free, this gives actually a first interpretation of the free upper extremal convolution.

We would like to give a deeper insight on the analogy between the classical and free settings. There is a morphism $\Lambda^{\vee}$ from the set of probability measures endowed with the classical upper convolution to the same set endowed with the free upper convolution. It maps any probability measure with distribution function F to the one with distribution function $\max(0, 1 + \log F)$. Moreover, Ben Arous and Voiculescu have established that $\Lambda^{\vee}$ provides a remarkable correspondence between classical max-stable laws and free max-stable laws, which preserves the domains of attraction. In this paper, we shall give a concrete realization of this morphism via a random matrix model.

More precisely, for each probability measure μ on the real line and each positive integer N, we shall define a law $\Lambda^{\vee}_N(\mu)$ on the set of N by N Hermitian matrices such that:

a) a $\Lambda^{\vee}_N(\mu)$-distributed random matrix is invariant under conjugation by any unitary matrix,
b) for any pair μ_1, μ_2 of probability measures on the real line and any pair M_1, M_2 of independent random matrices distributed respectively with respect to $\Lambda^{\vee}_N(\mu_1)$ and $\Lambda^{\vee}_N(\mu_2)$, the law of the supremum $M_1 \vee M_2$ of M_1 and M_2 is $\Lambda^{\vee}_N(\mu)$, where μ is the classical upper extremal convolution of μ_1 and μ_2,
c) if for all N, M_N is a random matrix with law $\Lambda^{\vee}_N(\mu)$, then the empirical spectral law of M_N tends almost surely to $\Lambda^{\vee}(\mu)$ as N tends to infinity.

Such a model produces a new clue with regards to the relevance of the analogy between the two max-operations, and it may be used to provide more intuitive proofs of some of Ben Arous and Voiculescu's results. For instance, we have already mentioned that unitarily-invariant random matrices behave asymptotically as free random variables. Therefore one can re-prove immediately, as a consequence of a), b) and c), the fact that $\Lambda^{\vee}$ is a morphism between the classical and the free upper extremal convolutions (and thus the formula $F_{X\vee Y} = \max(0, F_X + F_Y - 1)$ for free random variables X, Y). We shall see later

[1] The *cumulative distribution function* of a random variable X with distribution μ is the function, denoted either by F_X or F_μ, is defined by $F_\mu(x) = \mu((-\infty, x])$, for all $x \in \mathbb{R}$.

[2] The *supremum* of two Hermitian matrices is defined in section 2.3.

that Ben Arous and Voiculescu's results about max-stable laws and domains of attraction can also be proved with our random matrix model.

In a forthcoming paper, we will show how this approach is also appropriate to study the multivariate counterpart of the free upper extremal convolution.

The paper is organized as follows. First, we shall focus on the upper extremal convolution for classical, free and Hermitian random variables. Then, we shall introduce our models, via a strategy which is close to the one of our previous papers [3] and [5], and set the main theorems. The last part will be dedicated to the proofs.

2. The upper extremal convolutions

2.1. **For independent real random variables.** Let X and Y be real independent random variables, with cumulative distribution functions F_X and F_Y, and let us denote $X \vee Y = \max(X, Y)$. The cumulative distribution function of $X \vee Y$ is then equal to $F_{X\vee Y} = F_X F_Y$.

2.2. **For free random variables.** Let (M, τ) be a tracial W^*-probability space, that is a von Neumann algebra M endowed with an ultraweakly continuous faithful trace-state τ. The *maximum* $X \vee Y$ *of two self-adjoints elements* X and Y in M is defined with respect to the spectral order introduced by Olson in 1971 (cf [9]; see also [1, 2]):

- If p and q are two self-adjoint projectors, then $q \vee q$ is the Hermitian projector on $\overline{\Im(p) + \Im(q)}$.
- If a and b are two self-adjoint elements, with resolutions of identity E_a and E_b, then $a \vee b$ is the self-adjoint element with resolution of identity given by $E_a \vee E_b$: $a \vee b$ is the only self-adjoint element h such that for all real number t,

$$\chi_{(t,\infty)}(h) = \chi_{(t,\infty)}(a) \vee \chi_{(t,\infty)}(b),$$

where $\chi_{(t,\infty)}$ denotes the indicator function of the interval (t, ∞) and is applied *via* the functional calculus.

In 2006, in [2], Ben Arous and Voiculescu have determined the cumulative distribution function of $X \vee Y$ whenever X, Y are free:

$$\forall t \in \mathbb{R}, \quad F_{X\vee Y}(t) = \max(0, F_X(t) + F_Y(t) - 1) =: F_X \boxtimes\!\!\!\!\vee\, F_Y(t). \tag{1}$$

It is easy to note that the function $\Lambda^\vee(u) = \max(0, 1 + \log u)$ is a kind of morphism between classical and free upper extremal convolutions:

$$\Lambda^\vee(F_X F_Y) = \Lambda^\vee(F_X) \boxtimes\!\!\!\!\vee\, \Lambda^\vee(F_Y).$$

Notations: • In order to avoid the use of too many notations, we shall define the binary operation $\boxtimes\!\!\!\!\vee$ on the set of probability measures on the real line with the same symbol as the operation $\boxtimes\!\!\!\!\vee$ on the set of cumulative distribution functions. For all pair μ, ν of probability measures on the real line, $\mu \boxtimes\!\!\!\!\vee\, \nu$ is defined by:

$$F_{\mu \boxtimes\!\!\!\!\vee\, \nu} = F_\mu \boxtimes\!\!\!\!\vee\, F_\nu \;\; (= \max(0, F_\mu(t) + F_\nu(t) - 1) \;).$$

• We shall also define the operator $\Lambda^\vee$ on the set of probability measures on the real line with the same symbol as the function $\Lambda^\vee(x) = \max(0, 1 + \log x)$ on $[0, +\infty)$: for all probability measure μ on the real line, $\Lambda^\vee(\mu)$ is the probability measure with cumulative distribution function:

$$F_{\Lambda^\vee(\mu)}(t) = \Lambda^\vee(F_\mu(t)) \;\; (= \max(0, 1 + \log F_\mu(t)) \;).$$

Remark 2.1. *The free upper extremal convolution also appears in a much simpler situation than the maximum of two free operators. Let* $x_1, \ldots, x_N, y_1, \ldots, y_N \in \mathbb{R}$*, and* $z_1 \geq \ldots \geq z_N$ *the* N *largest elements of the multiset* $(x_1, \ldots, x_N, y_1, \ldots, y_N)$*. Then* $\frac{1}{N}\sum_{i=1}^N \delta_{z_i}$ *is the free upper extremal convolution of* $\frac{1}{N}\sum_{i=1}^N \delta_{x_i}$ *and* $\frac{1}{N}\sum_{i=1}^N \delta_{y_i}$.

2.3. **For independent Hermitian random matrices.** Let N be a positive integer. The maximum operation introduced by Olson is also defined on the space of $N\times N$ Hermitian matrices, as a particular case of the definition given in section 2.2.

Concretely, for A, B Hermitian matrices with the same size, $A\vee B$ can be defined in the following way. Let $\{\lambda_1>\lambda_2>\cdots>\lambda_p\}$ be the union of the spectrums of A and of B. For $i=0,\ldots,p$, let us define $E_i=\sum_{j=1}^i(\ker(A-\lambda_jI)+\ker(B-\lambda_jI))$. Then $A\vee B$ is the Hermitian matrix with eigenvalues the λ_j's such that $E_j\neq E_{j-1}$, with associated eigenspace $E_j\cap(E_{j-1}^\perp)$. In other words, if, for any subspace F, P_F designs the orthogonal projector onto F, we have

$$A\vee B=\lambda_1P_{E_1}+\lambda_2P_{E_2\cap(E_1^\perp)}+\lambda_3P_{E_3\cap(E_2^\perp)}+\cdots+\lambda_pP_{E_p\cap(E_{p-1}^\perp)}.$$

The following proposition resumes the characteristics of the operation $\vee$ on unitarily invariant random matrices. Recall that the *empirical spectral law* of a matrix is the uniform measure on the multiset of its eigenvalues counted by multiplicity.

Proposition 2.2. *If A and B are two $N\times N$ Hermitian random independent matrices, whose laws are invariant by conjugation by any unitary matrix, then*

- *$A\vee B$ is an Hermitian random matrix, whose law is invariant under the conjugation by any unitary matrix;*
- *the N eigenvalues of $A\vee B$ are the N largest of the $2N$ eigenvalues of A and B (counted with multiplicity);*
- *the cumulative distribution functions F_A, F_B and $F_{A\vee B}$ of the respective empirical spectral laws of A, B and $A\vee B$ are almost surely linked by the relation*

$$F_{A\vee B}=\max(0,F_A+F_B-1). \tag{2}$$

Remark 2.3. *It is noteworthy that formulas* (1) *and* (2) *are identical, showing that the matricial and free upper extremal convolutions are somehow the same operation. As free random variables can be approximated by independent unitarily invariant random matrices, the rather trivial result* (2) *provides an intuitive proof of formula* (1).

3. The matricial interpolation

The aim of this section is to introduce random matrices giving quite a natural interpretation of the mapping $\Lambda^\vee$. We simply follow the same approach as we did for the Bercovici-Pata bijection in [3] and [5].

Let F_μ denote its cumulative distribution function for μ probability measure on the real line. For all integer $k\geq 1$, $F_\mu^{\frac{1}{k}}$ is the cumulative distribution function of a probability measure on the real line, which will be denoted by $\mu^{\frac{1}{k}}$. It is clear that if $X_1,\ldots,X_k$ are independent random variables distributed according to $\mu^{\frac{1}{k}}$, $\max\{X_1,\ldots,X_k\}$ should be distributed according to μ.

Let us consider $N,k\geq 1$. Let $X_1,\ldots,X_N$ be independent random variables distributed according to $\mu^{\frac{1}{k}}$ and U an unitary, Haar distributed, $N\times N$ random matrix independent of $X_1,\ldots,X_N$. Define the random matrix M by

$$M=U\begin{pmatrix}X_1 & & \\ & \ddots & \\ & & X_N\end{pmatrix}U^* \tag{3}$$

Let $M_1,\ldots,M_k$ be independent replicas of M.

Theorem 3.1. *As k tends to infinity, the distribution of $M_1\vee\cdots\vee M_k$ converges to a probability measure $\Lambda_N^\vee(\mu)$ on the space of $N\times N$ Hermitian matrices. This measure is invariant under the action of the* *unitary group by conjugation and in the case where μ has no atom, under $\Lambda_N^\vee(\mu)$, the joint distribution of the ranked eigenvalues is absolutely continuous with respect to $\mu^{\otimes N}$, with density*

$$N^N 1_{t_1\geq\cdots\geq t_N}\prod_{i=1}^N \frac{F_\mu(t_N)}{F_\mu(t_i)} \tag{4}$$

at any point $(t_1,\ldots,t_N)\in\mathbb{R}^N$ such that for all $i=1,\ldots,N$, $F_\mu(t_i)\neq 0$ (the density being set to zero anywhere else).

Moreover, if ν is another probability measure on the real line and A, B are independent random matrices respectively distributed according to $\Lambda_N^\vee(\mu)$, $\Lambda_N^\vee(\nu)$, then $A\vee B$ is distributed according to $\Lambda_N^\vee(\rho)$, where ρ is the probability measure on the real line such that $F_\rho = F_\mu F_\nu$.

Corollary 3.2. *With the preceding notations, we have*

(i) $\Lambda_N^\vee(\mu)$ is max-infinitely divisible,

(ii) if μ is max-stable, then so is $\Lambda_N^\vee(\mu)$.

In the following theorem, we are interested in the limit of the spectral measure of a $\Lambda_N^\vee(\mu)$-distributed random matrix, when its dimension N tends to infinity.

Theorem 3.3. *Let μ be a probability measure on the real line. For all $N\geq 1$, let M_N a random matrix with law $\Lambda_N^\vee(\mu)$, $\lambda_1,\ldots,\lambda_N$ its eigenvalues and $\hat\mu_N = \frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i}$ its empirical spectral law. Then, when N goes to infinity, $\hat\mu_N$ converges weakly almost surely to $\Lambda^\vee(\mu)$.*

The previous theorem may be used to provide new, more intuitive proofs of some of the Ben Arous and Voiculescu's results in [2].

A first example has been given in the introduction: it is a new derivation of the formula (1) defining the free upper extremal convolution in terms of cumulative distribution functions.

For second example, the fact that $\Lambda^\vee$ maps any classical max-stable law to a free one is a direct consequence of (ii) of Corollary 3.2 and of formulas (1) and (2).

Finally, let us consider the preservation of the domains of attraction. Ben Arous and Voiculescu have proved that for any classical max-stable law μ, $\Lambda^\vee(\mu)$ is freely max-stable and that for any cumulative distribution function F, for any sequences $a_k>0, b_k\in\mathbb{R}$, we have

$$F(a_k\cdot+b_k)^k \underset{k\to\infty}{\longrightarrow} F_\mu(\cdot) \Longrightarrow F(a_k\cdot+b_k)^{\boxtimes k} \underset{k\to\infty}{\longrightarrow} \Lambda^\vee(F_\mu(\cdot)) \tag{5}$$

(to state the reciprocal implication, one needs to choose a right inverse to $\Lambda^\vee$, which we shall not do here). This result can be generalized as follows: for all sequence (F_k) of cumulative distribution functions, if F is a cumulative distribution function, we have

$$F_k(\cdot)^k \underset{k\to\infty}{\longrightarrow} F(\cdot) \Longrightarrow F_k(\cdot)^{\boxtimes k} \underset{k\to\infty}{\longrightarrow} \Lambda^\vee(F(\cdot)) \tag{6}$$

(the proof of (6) is given in the last section of the paper).

Let us now explain how our random matrix model gives an heuristic explanation of (6). For sake of simplicity, let F_k be equal to $F^{\frac{1}{k}}$, with F the cumulative distribution function of μ (for more general F_k, a slight modification of our random matrix model is enough to apply what follows).

For each $N\geq 1$ and for each $k\geq 1$, let us consider a family $M_1(N,k),\ldots,M_k(N,k)$ of independent replicas of the random matrix M of (3), i.e.

$$M = U\begin{pmatrix} X_1 & & \\ & \ddots & \\ & & X_N\end{pmatrix}U^*,$$

where $X_1, \ldots, X_N$ are independent random variables with cumulative distribution function $F^{\frac{1}{k}}$ and U an unitary, Haar distributed, $N \times N$ random matrix independent of $X_1, \ldots, X_N$.

- Let k tend to infinity. Following the definition of $\Lambda_N^{\vee}(\mu)$ given in theorem 3.1, we get:

$$M_1(N,k) \vee \cdots \vee M_k(N,k) \underset{k\to\infty}{\longrightarrow} M_N \text{ in distribution,} \tag{7}$$

  where for all N, M_N is a $\Lambda_N^{\vee}(\mu)$-distributed random matrix.
  Let now N tend to infinity. From theorem 3.3, we deduce:

$$\text{Empirical Spectral Law}(M_N) \underset{N\to\infty}{\longrightarrow} \Lambda^{\vee}(\mu). \tag{8}$$

- Now, let N tend to infinity before k does. According to the law of large numbers, the empirical spectral law of $M_i(N,k)$ converges almost surely, for each k and $i = 1, \ldots, k$, to the law with distribution function $F^{\frac{1}{k}}$, which shall be denoted by $\mu^{\frac{1}{k}}$. Therefore, we deduce from the asymptotic freeness of independent randomly rotated random matrices (more precisely by the equality between formulas (1) and (2)), the following almost sure convergence:

$$\text{Empirical Spectral Law}(M_1(N,k) \vee \cdots \vee M_k(N,k)) \underset{N\to\infty}{\longrightarrow} \mu^{\frac{1}{k}} \Box\!\!\!\!\vee \cdots \Box\!\!\!\!\vee\, \mu^{\frac{1}{k}}. \tag{9}$$

Joining (7), (8), (9), we get the following diagram

$$\begin{array}{ccc}
M_1(d,k) \vee \cdots \vee M_k(d,k) & \xrightarrow{k\to\infty} & \Lambda_N^{\vee}(\mu) \\
| & & | \\
N \to \infty & & N \to \infty \\
\downarrow & & \downarrow \\
\text{Empirical Spectral Law:} & & \text{Empirical Spectral Law:} \\
\mu^{\frac{1}{k}} \Box\!\!\!\!\vee \cdots \Box\!\!\!\!\vee\, \mu^{\frac{1}{k}} & & \Lambda^{\vee}(\mu)
\end{array}$$

The right-hand side of (6), i.e. the fact that $(\mu^{\frac{1}{k}})^{\Box\!\!\!\!\vee k} \underset{k\to\infty}{\longrightarrow} \Lambda^{\vee}(\mu)$, only means that one can add an edge $\underset{k\to\infty}{\longrightarrow}$ between both bottom vertices of the diagram, i.e. that the operations $k \to \infty$ and $N \to \infty$ are commutative, which is quite expected.

4. Proofs

4.1. **Preliminary results.** Both of the theorems will be proved first for measures which are absolutely continuous with respect to the Lebesgue measure, and then extended to all measures by approximation. In this context, the appropriate approximation tool is given by the following lemma.

Lemma 4.1. *Let us define, for μ probability measure on the real line, the function $F_\mu^{<-1>}$ on $(0,1)$ by $F_\mu^{<-1>}(u) = \min\{x \in \mathbb{R}\,;\, F_\mu(x) \geq u\}$.*

(i) For μ, ν probability measures on the real line, we have the following equalities (between quantities which can be infinite)

$$\inf\{\varepsilon > 0\,;\, F_\nu(\bullet - \varepsilon) \leq F_\mu(\bullet) \leq F_\nu(\bullet + \varepsilon)\} = ||F_\mu^{<-1>} - F_\nu^{<-1>}||_\infty = \inf ||X - Y||_\infty,$$

where the infimum in the third term is taken on pairs (X,Y) of random variables defined on a same probability space with respective distributions μ, ν.

(ii) For any probability measure μ on the real line, for any $\varepsilon > 0$, there exists a probability measure μ_ε on the real line such that F_{μ_ε} is smooth and

$$F_{\mu_\varepsilon}(\bullet - \varepsilon) \leq F_\mu(\bullet) \leq F_{\mu_\varepsilon}(\bullet + \varepsilon).$$

Proof. (i) We prove these equalities by cyclic majorizations.

- Let us prove first

$$\inf\{\varepsilon > 0\,;\, \forall x \in \mathbb{R}, F_\nu(x - \varepsilon) \leq F_\mu(x) \leq F_\nu(x + \varepsilon)\} \geq ||F_\mu^{<-1>} - F_\nu^{<-1>}||_\infty.$$

Let us consider $\varepsilon > 0$ such that $F_\nu(\bullet - \varepsilon) \le F_\mu(\bullet) \le F_\nu(\bullet + \varepsilon)$. Since this is equivalent to $F_\mu(\bullet - \varepsilon) \le F_\nu(\bullet) \le F_\mu(\bullet + \varepsilon)$, this inequation is symmetric in μ and ν. Hence, it suffices to prove that $F_\nu^{<-1>}(u) - F_\mu^{<-1>}(u) \le \varepsilon$ for all $u \in (0,1)$. We have:

$$F_\nu(F_\mu^{<-1>}(u) + \varepsilon) \ge F_\mu(F_\mu^{<-1>}(u)) \ge u,$$

hence, $F_\nu^{<-1>}(u) \le F_\mu^{<-1>}(u) + \varepsilon$.

- The inequality

$$||F_\mu^{<-1>} - F_\nu^{<-1>}||_\infty \ge \inf ||X - Y||_\infty$$

is due to the fact that for U random variable with uniform distribution on $(0,1)$, $X := F_\mu^{<-1>}(U), Y := F_\nu^{<-1>}(U)$ are respectively distributed according to μ, ν.

- To conclude, it suffices to prove

$$\inf ||X - Y||_\infty \ge \inf\{\varepsilon > 0\,;\, F_\nu(\bullet - \varepsilon) \le F_\mu(\bullet) \le F_\nu(\bullet + \varepsilon)\}.$$

So let us consider a pair (X, Y) of random variables defined on the same probability space with respective distributions μ, ν. Consider $\varepsilon > 0$ such that $|X - Y| \le \varepsilon$ uniformly on the probability space. Then, for all real number x,

$$\mathbb{P}(Y \le x - \varepsilon) \le \mathbb{P}(X \le x) \le \mathbb{P}(Y \le x + \varepsilon).$$

(ii) Let $\varepsilon > 0$, and F a smooth non decreasing function on the real line such that for all $k \in \mathbb{Z}$, $F(k\varepsilon) = F_\mu(k\varepsilon)$. Such a function exists obviously, as an example, one can consider $F(x) = \int_{-\infty}^x f(t)\,\mathrm{d}t$, for

$$f(t) = \sum_{k\in\mathbb{Z}} [F(k\varepsilon) - F((k-1)\varepsilon)]\varphi(t - k\varepsilon)$$

with φ smooth non negative function on the real line, with support contained in $[-\varepsilon, 0]$ satisfying $\int_{-\varepsilon}^0 \varphi(t)\,\mathrm{d}t = 1$. Hence F is a càdlàg non decreasing function on the real line such that $F(x)$ tends to zero (resp. one) as x tends to $-\infty$ (resp. $+\infty$). It follows that $F = F_{\mu_\varepsilon}$ for a certain probability measure μ_ε on the real line. Moreover, for all real number x, if $k \in \mathbb{Z}$ is such that $k\varepsilon \le x < (k+1)\varepsilon$, then, one has:

$$F_{\mu_\varepsilon}(x - \varepsilon) \le F_{\mu_\varepsilon}(k\varepsilon) = F_\mu(k\varepsilon) \le F_\mu(x) \le F_\mu((k+1)\varepsilon) = F_{\mu_\varepsilon}((k+1)\varepsilon) \le F_{\mu_\varepsilon}(x + \varepsilon).$$

□

For any positive integer N, let us define $\max_N$ to be the function from $\cup_{n\ge N}\mathbb{R}^n$ to $\mathbb{R}^N$ which maps any vector x of $\mathbb{R}^n$, $n \ge N$, to the vector of the N largest coordinates of x ranked in decreasing order. The following property is a basic result, see for instance [7].

Proposition 4.2. *Let F be the cumulative distribution function of a probability measure μ with no atom on the real line. Let, for $n \ge N$, $X_1, \dots, X_n$ be independent random variables with law μ. The distribution of $\max_N(X_1, \dots, X_n)$ has density*

$$(t_1, \dots, t_N) \mapsto \frac{n!}{(n-N)!} 1_{t_1 \ge \cdots \ge t_N} F(t_N)^{n-N}$$

with respect to $\mu^{\otimes N}$.

Before stating the next proposition, we shall recall that for μ be a probability measure on the real line and k positive integer, $\mu^{\frac{1}{k}}$ is the probability measure on the real line with cumulative distribution function $F_\mu^{\frac{1}{k}}$, i.e. the law m such that for $X_1, \dots, X_k$ independent random variables with law m, $\max(X_1, \dots, X_k)$ has law μ.

Proposition 4.3. *(i) Let μ be a probability measure on the real line. For all integer $N \ge 1$, the push forward of the probability measure $(\mu^{\frac{1}{k}})^{\otimes kN}$ on $\mathbb{R}^{kN}$ by the function $\max_N$ converges weakly, as the integer k tends to infinity, to a probability measure on $\mathbb{R}^N$ denoted by μ_N.*

(ii) When μ has no atom, μ_N is absolutely continuous with respect to $\mu^{\otimes N}$, with density

$$N^N 1_{t_1 \geq \cdots \geq t_N} \prod_{i=1}^{N} \frac{F_\mu(t_N)}{F_\mu(t_i)}$$

at any point $(t_1, \ldots, t_N) \in \mathbb{R}^N$ such that for all $i = 1, \ldots, N$, $F_\mu(t_i) \neq 0$.

(iii) Let us endow $\mathbb{R}^N$ with the norm $||x|| = \max_i |x_i|$. Then, for any probability measures μ, ν on the real line,

$$\inf\{||V - W||_\infty\,;\, V, W \text{ random vectors defined on the same space with respective laws } \mu_N, \nu_N\}$$
$$\leq \inf\{||X - Y||_\infty\,;\, X, Y \text{ random variables defined on the same space with respective laws } \mu, \nu\}.$$

(iv) For any pair μ, ν of probability measures on the real line, if ρ is the probability measure on the real line such that $F_\rho = F_\mu F_\nu$, for all $N \geq 1$, ρ_N is the push-forward, by the function $\max_N$, of the probability measure $\mu_N \otimes \nu_N$ on $\mathbb{R}^{2N}$.

Proof. Let $k, N \geq 1$, and μ a probability measure with no atom. From property 4.2, we infer first that $\mu^{\frac{1}{k}} = \frac{1}{k} F_\mu^{\frac{1}{k}-1}\, \mathrm{d}\mu$, and then that the distribution $\max_N((\mu^{\frac{1}{k}})^{\otimes kN})$ has a density with respect to $\mu^{\otimes N}$ equal to

$$(t_1, \ldots, t_N) \mapsto \frac{(dk)!}{k^N (dk - N)!} 1_{t_1 \geq \cdots \geq t_N} \prod_{i=1}^{N} \frac{F_\mu(t_N)^{1-\frac{1}{k}}}{F_\mu(t_i)^{1-\frac{1}{k}}}$$

at any point $(t_1, \ldots, t_N) \in \mathbb{R}^N$ such that for all $i = 1, \ldots, N$, $F(t_i) \neq 0$ (the density can obviously be set to zero anywhere else). As k tends to infinity, this density stays uniformly bounded and converges pointwise to

$$N^N 1_{t_1 \geq \cdots \geq t_N} \prod_{i=1}^{N} \frac{F_\mu(t_N)}{F_\mu(t_i)},$$

hence (i) for probability measures with no atom and (ii) are proved.

Now, let us complete the proof of (i). Let us consider a probability measure μ and, for all positive integer k, a family $X(k, 1), \ldots, X(k, kN)$ of independent random variables with law $\mu^{\frac{1}{k}}$. To prove (i), by Theorem 1.12.4 of [10], it suffices to prove that for any real bounded Lipschitz function f on $\mathbb{R}^N$, the sequence $\mathbb{E}(f(\max_N(X(k, 1), \ldots, X(k, kN))))$ converges as k tends to infinity. So let us fix such a function. We shall prove that the previous sequence is a Cauchy sequence. Let us fix $\varepsilon > 0$. Let us consider μ_ε as in (ii) of lemma 4.1. Note that for all positive integers k, we also have:

$$F_{\mu_\varepsilon}^{\frac{1}{k}}(\bullet - \varepsilon) \leq F_\mu^{\frac{1}{k}}(\bullet) \leq F_{\mu_\varepsilon}^{\frac{1}{k}}(\bullet + \varepsilon).$$

So, based on (i) of lemma 4.1, we may suppose that for all $k \geq 1$, on the probability space where the $X(k, i)$'s are defined, there is a family

$$Y(k, 1), \ldots, Y(k, kN)$$

of random variables such that

(a) the $Y(k, i)$'s are independent and distributed according to $\mu_\varepsilon^{\frac{1}{k}}$,

(b) for all $i = 1, \ldots, kN$, $|X(k, i) - Y(k, i)| \leq 2\varepsilon$ almost surely.

Note that μ_ε has no atom, hence by (a) and what we just proved, the sequence

$$\mathbb{E}(f(\max_N(Y(k, 1), \ldots, Y(k, kN))))$$

is a Cauchy sequence. Note also that by (b), if C is a Lipschitz constant for f with respect to the norm $||x|| = \max_i |x_i|$, then for all k,

$$|\mathbb{E}(f(\max_N(X(k, 1), \ldots, X(k, kN)))) - \mathbb{E}(f(\max_N(Y(k, 1), \ldots, Y(k, kN))))| \leq 2C\varepsilon.$$

Hence there is $k_0 \geq 1$ such that for all $k, k' \geq k_0$,

$$|\mathbb{E}(f(\max_N(X(k,1),\dots,X(k,kN)))) - \mathbb{E}(f(\max_N(X(k',1),\dots,X(k',k'd))))| \leq (4C+1)\varepsilon.$$

Thus we have proved (i) for any probability measure μ.

Now, let us prove (iii). Consider μ, ν probability measures on the real line. Note that by part (i) of lemma 4.1, it suffices to prove that for all positive ε such that

$$F_\nu(\bullet - \varepsilon) \leq F_\mu(\bullet) \leq F_\nu(\bullet + \varepsilon),$$

for all $\alpha > \varepsilon$, there exists a pair V, W of random vectors defined on the same space with respective laws μ_N, ν_N such that $||V - W||_\infty \leq \alpha$. Let us consider such a positive ε and $\alpha > \varepsilon$. For all $k \geq 1$, we also have:

$$F_\nu^{\frac{1}{k}}(\bullet - \varepsilon) \leq F_\mu^{\frac{1}{k}}(\bullet) \leq F_\nu^{\frac{1}{k}}(\bullet + \varepsilon).$$

So, according to part (i) of lemma 4.1, we shall consider, for all $k \geq 1$, a family

$$X(k,1),\dots,X(k,kN),Y(k,1),\dots,Y(k,kN)$$

of random variables defined on the same space such that

(a) the $X(k,i)$'s are independent and distributed according to $\mu^{\frac{1}{k}}$,

(b) the $Y(k,i)$'s are independent and distributed according to $\nu^{\frac{1}{k}}$,

(c) for all $i = 1,\dots,kN$, $|X(k,i) - Y(k,i)| \leq \alpha$ almost surely.

Let, for all $k \geq 1$, τ_k be the joint law, on $\mathbb{R}^{2N}$, of the random vector

$$(\max_N(X(k,1),\dots,X(k,kN)), \max_N(Y(k,1),\dots,Y(k,kN)).$$

The law of $\max_N(X(k,1),\dots,X(k,kN))$ (resp. of $\max_N(Y(k,1),\dots,Y(k,kN))$) converges weakly to μ_N (resp. to ν_N) as k tends to infinity. It follows that the sequence τ_k is tight, i.e. relatively compact for the topology of weak convergence (Theorem 6.1 of [4]). Let τ be the weak limit of a subsequence of τ_k. Let (V, W) be a τ-distributed random vector of $\mathbb{R}^N \times \mathbb{R}^N$. Then the law of V (resp. of W) is μ_N (resp. ν_N). Moreover, it is easy to notice that for all k, τ_k is supported by $\{(x,y)\,;\, x \in \mathbb{R}^N, y \in \mathbb{R}^N, ||x-y|| \leq \alpha\}$. Indeed, let $\sigma, \tau \in S_n$ such that $X(k,\sigma(1)) \geq \cdots \geq X(k,\sigma(kN))$ and $Y(k,\tau(1)) \geq \cdots \geq Y(k,\tau(kN))$; then for any $i = 1,\dots,kN$,

$$\begin{aligned} X(k,\sigma(i)) &= \max_{V \subset \{1,\dots,kN\}, \#V = i} \min(X(k,j), j \in V) \\ &\geq \max_{V \subset \{1,\dots,kN\}, \#V = i} \min(Y(k,j) - \alpha, j \in V) \\ &= Y(k,\tau(i)) - \alpha \quad \text{a.s.} \end{aligned}$$

By symmetry, this proves $|X(k,\sigma(i)) - Y(k,\tau(i))| \leq \alpha$ a.s., and this implies

$$\begin{aligned} \|\max_N(X(k,1),\dots,X(k,kN)) - \max_N(Y(k,1),\dots,Y(k,kN))\| &= \max_{i=1,\dots,N} |X(k,\sigma(i)) - Y(k,\tau(i))| \\ &\leq \alpha \quad \text{a.s.} \end{aligned}$$

Letting k go to infinity, this establishes the inequality $\|V - W\| \leq \alpha$ almost surely.

To prove (iv), consider for any $k \geq 1$,

$$X(k,1),\dots,X(k,kN),Y(k,1),\dots,Y(k,kN)$$

is a family of independent random variables such that for all $i = 1,\dots,kN$, $X(k,i)$ (resp. $Y(k,i)$) is distributed according to $\mu^{\frac{1}{k}}$ (resp. $\nu^{\frac{1}{k}}$). Then $X(k,1) \vee Y(k,1),\dots,X(k,kN) \vee Y(k,kN)$ are i.i.d. with law $\rho^{\frac{1}{k}}$, and consequently μ_N, ν_N and ρ_N are the respective weak limit distributions, as k tends to infinity, of

$$\max_N[X(k,1),\dots,X(k,kN)], \quad \max_N[Y(k,1),\dots,Y(k,kN)]$$

$$\text{and} \quad \max_N[X(k,1) \vee Y(k,1),\dots,X(k,kN) \vee Y(k,kN)]. \tag{10}$$

Hence, by continuity and commutativity of the maximum operations, $\max_N(\mu_N \otimes \nu_N)$ is the weak limit, as k tends to infinity, of the distribution of

$$\max_N[X(k,1),\ldots,X(k,kN),Y(k,1),\ldots,Y(k,kN)]. \tag{11}$$

To conclude, we shall establish that the distributions of the vectors of (10) and (11) have the same weak limits as k tends to infinity. Obviously, it reduces to prove that the probability of the event

$$\{\max_N[X(k,1)\vee Y(k,1),\ldots,X(k,kN)\vee Y(k,kN)]$$
$$\neq \max_N[X(k,1),\ldots,X(k,kN),Y(k,1),\ldots,Y(k,kN)]\} \tag{12}$$

tends to zero as k tends to infinity. Note that this event is equivalent to the fact that there is $i \in \{1,\ldots,kN\}$ such that $X(k,i)$ and $Y(k,i)$ are both coordinates of

$$\max_N[X(k,1),\ldots,X(k,kN),Y(k,1),\ldots,Y(k,kN)]$$

and are not equal to any of the others $X(k,j)$'s and $Y(k,l)$'s. Hence, if one denotes by $I(k)$ (resp. $J(k)$) the set of i's in $\{1,\ldots,kN\}$ such that $X(k,i)$ (resp. $Y(k,i)$) is one of the coordinates of

$$\max_N[X(k,1),\ldots,X(k,kN)] \quad \text{(resp. of } \max_N[Y(k,1),\ldots,Y(k,kN)]),$$

the event of (12) is contained in the event $I(k)\cap J(k)\neq\emptyset$, which probability is equal to $\binom{kN}{N}^{-1}$, since we may choose for $I(k)$ and $J(k)$ to be independent and to have uniform distribution on the set of subsets of $\{1,\ldots,kN\}$ with cardinality N. Thus, (iv) is proved. □

4.2. **Proof of proposition 2.2.** The first statement is immediate. The third one follows immediately form the second one. To establish the second statement, it suffices to prove that almost surely, for any $t\in\mathbb{R}$, $\mathrm{Im}(\chi_{(t,\infty)}(A))+\mathrm{Im}(\chi_{(t,\infty)}(B))$ has the maximal dimension conditionally to the dimensions of $\mathrm{Im}(\chi_{(t,\infty)}(A)$ and $\mathrm{Im}(\chi_{(t,\infty)}(B)$, i.e. that

$$\dim[\mathrm{Im}(\chi_{(t,\infty)}(A))+\mathrm{Im}(\chi_{(t,\infty)}(B))]=\min\{N,\dim[\mathrm{Im}(\chi_{(t,\infty)}(A))]+\dim[\mathrm{Im}(\chi_{(t,\infty)}(B)]\}\quad \text{a.s.} \tag{13}$$

Let $p=\dim[\mathrm{Im}(\chi_{(t,\infty)}(A))]$, $q=\dim[\mathrm{Im}(\chi_{(t,\infty)}(B))]$. Due to the unitarily invariance of A (resp. B), the law of $\mathrm{Im}(\chi_{(t,\infty)}(A))$ (resp. $\mathrm{Im}(\chi_{(t,\infty)}(B))$) is also invariant under the action of any unitary matrix, hence is uniform on the set of all subspaces of dimension p (resp. q). Moreover, $\mathrm{Im}(\chi_{(t,\infty)}(A))$ and $\mathrm{Im}(\chi_{(t,\infty)}(B))$ are independent.

Let us consider a family $(g_{i,j}, i\in\{1,\ldots,N\}, j\geq 1)$ of independent complex standard Gaussian random variables. The matrix $(g_{i,j}, i,j\in\{1,\ldots,N\})$ is distributed w.r.t. the Circular Unitary Ensemble. Hence, it is unitarily invariant and a.s. inversible. Therefore, the subspace generated by vectors $(g_{i,j}, i\in\{1,\ldots,N\})_{j=1}^{p}$ (resp. $(g_{i,j}, i\in\{1,\ldots,N\})_{j=p+1}^{p+q}$, $(g_{i,j}, i\in\{1,\ldots,N\})_{j=1}^{p+q}$) has dimension p (resp. q, $\min(N,p+q)$) a.s., and its law is uniform on the set of all subspaces of same dimension. This proves (13).

Remark 4.4. *Note that the proof of the second statement could also have been deduced from Theorem 2.2 of* [6].

4.3. **Proof of theorem 3.1.** Following proposition 2.2, if we denote $M^{(k)}=M_1\vee\cdots\vee M_k$, then $M^{(k)}$ is an Hermitian random matrix, whose law is invariant under conjugation with any unitary matrix and whose eigenvalues $\lambda_1^{(k)}\geq\ldots\geq\lambda_N^{(k)}$ are equal to the N largest of the Nk eigenvalues of $M_1,\ldots,M_k$ which are all independent with same distribution $\mu^{1/k}$. Therefore, by theorem 4.3.5 of [8], there exists U unitary, Haar distributed and independent of the vector $(\lambda_1^{(k)},\ldots,\lambda_N^{(k)})$, such that:

$$M^{(k)}=U\begin{pmatrix}\lambda_1^{(k)} & & \\ & \ddots & \\ & & \lambda_N^{(k)}\end{pmatrix}U^*.$$

Note that the distribution of the vector $(\lambda_1^{(k)}, \dots, \lambda_N^{(k)})$ is the push forward of the probability measure $(\mu^{\frac{1}{k}})^{\otimes kN}$ on $\mathbb{R}^{kN}$ by the function $\max_N$. Hence, the first part of the theorem follows from (i) and (ii) of proposition 4.3, and the second part follows from (iv) of the same proposition.

4.4. **Proof of corollary 3.2.** Both assertions are easy consequences of the last part of theorem 3.1.

(i) Let $p \geq 1$. By definition, $F_\mu = \left(F_{\mu^{\frac{1}{p}}}\right)^p$. Therefore $\Lambda_N^\vee(F_\mu)$ is the distribution of $H_1 \vee \cdots \vee H_p$, with $H_1, \dots, H_p$ i.i.d. with law $\Lambda_N^\vee(F_{\mu^{\frac{1}{p}}})$. This means that $\Lambda_N^\vee(F_\mu)$ is max-infinitely divisible.

(ii) Let $p \geq 1$. Since μ is max-stable, there exists $a_p > 0$, $b_p \in \mathbb{R}$ such that μ is the distribution of $(a_p X_1 + b_p) \vee \cdots \vee (a_p X_p + b_p)$, where $X_1, \dots, X_p$ are i.i.d. with law μ. Let $H_1, \dots, H_p$ be i.i.d. with law $\Lambda_N^\vee(\mu)$. It is quite immediate from the construction that the image of the distribution of $a_p X_i + b_p$ is the law of $a_p H_i + b_b I_N$ where I_N denotes the identity matrix. Therefore, $\Lambda_N^\vee(\mu)$ is the distribution of $(a_p H_1 + b_p I_N) \vee \cdots \vee (a_p H_p + b_p I_N)$. This proves that $\Lambda_N^\vee(\mu)$ is max-stable.

4.5. **Proof of theorem 3.3.** Suppose first that F_μ is smooth on the real line. It implies that μ is absolutely continuous with respect to the Lebesgue measure and that $\mathrm{d}\mu(x) = F'_\mu(x)\,\mathrm{d}x$. Since F_μ is continuous on the real line, the almost sure weak convergence of the sequence of random probability measures $\hat{\mu}_N$ to $\Lambda^\vee(\mu)$ is equivalent to the fact the for all rational number t, $\hat{\mu}_N\left((-\infty, t]\right)$ converges almost surely to $\max(0, 1 + \log F_\mu(t))$. We shall prove that for all $t \in \mathbb{R}$, $\hat{\mu}_N\left((-\infty, t]\right)$ converges almost surely to $\max(0, 1 + \log F_\mu(t))$. Let us fix a real number t.

The first step is to compute $\mathbb{E}\left[e^{\lambda \hat{\mu}_N((-\infty,t])}\right]$: with the notations $\lambda_0 = +\infty$, $\lambda_{N+1} = -\infty$, we have

$$\begin{aligned}
\mathbb{E}\left[e^{\lambda \hat{\mu}_N((-\infty,t])}\right] &= \mathbb{E}\left[\sum_{p=0}^{N} e^{\frac{\lambda(N-p)}{N}} 1_{\lambda_{p+1} \leq t < \lambda_p}\right] \\
&= \sum_{p=0}^{N} e^{\frac{\lambda(N-p)}{N}} \mathbb{P}\left(\lambda_{p+1} \leq t < \lambda_p\right)
\end{aligned}$$

Now, note that for all real number T such that $F(T) > 0$, for all integer non negative m,

$$\int_{(T,\infty)} (\log F_\mu(t))^m \frac{\mathrm{d}\mu(t)}{F_\mu(t)} = -\frac{(\log F_\mu(T))^{m+1}}{m+1} \quad \text{and} \quad \int_{(-\infty,T]} F_\mu(t)^m \mathrm{d}\mu(t) = \frac{F_\mu(T)^{m+1}}{m+1}.$$

Hence for $p = 0, \dots, N-1$,

$$\mathbb{P}\left(\lambda_{p+1} \leq t < \lambda_p\right) = \int_{t \in \mathbb{R}^N} N^N 1_{t \geq t_{p+1} \geq \cdots \geq t_N} 1_{t < t_p \leq \cdots \leq t_1} \prod_{i=1}^{N} \frac{F_\mu(t_N)}{F_\mu(t_i)} \,\mathrm{d}\mu(t_i) = \frac{N^p}{p!} F_\mu(t)^N (-\log F_\mu(t))^p.$$

And for $p = N$:

$$\mathbb{P}\left(t < \lambda_N\right) = \frac{N^N}{(N-1)!} \int_t^{+\infty} F_\mu(t_N)^N \left(-\log F_\mu(t_N)\right)^{N-1} \mathrm{d}\mu(t_N) = \int_0^{-\log F_\mu(t)} \frac{N^N}{(N-1)!} u^{N-1} e^{-Nu}\,\mathrm{d}u.$$

Thus, using the fact that $\sum_{p=0}^{N-1}\frac{x^p}{p!}=e^x\int_x^{+\infty}\frac{1}{(N-1)!}u^{N-1}e^{-u}\,\mathrm{d}u$, we get:

$$\begin{aligned}
&\mathbb{E}\left[e^{\lambda\hat{\mu}_N((-\infty,t])}\right]\\
&= e^{\lambda}F_\mu(t)^N\sum_{p=0}^{N-1}\frac{(-Ne^{-\frac{\lambda}{N}}\log F_\mu(t))^p}{p!}+\int_0^{-\log F_\mu(t)}\frac{N^N}{(N-1)!}u^{N-1}e^{-Nu}\,\mathrm{d}u\\
&= e^{\lambda}F_\mu(t)^N e^{-Ne^{-\frac{\lambda}{N}}\log F_\mu(t)}\int_{-Ne^{-\frac{\lambda}{N}}\log F_\mu(t)}^{+\infty}\frac{1}{(N-1)!}u^{N-1}e^{-u}\,\mathrm{d}u+\int_0^{-\log F_\mu(t)}\frac{N^N}{(N-1)!}u^{N-1}e^{-Nu}\,\mathrm{d}u\\
&= e^{\lambda+N\left(1-e^{-\frac{\lambda}{N}}\right)\log F_\mu(t)}\int_{-e^{-\frac{\lambda}{N}}\log F_\mu(t)}^{+\infty}\frac{N^N}{(N-1)!}u^{N-1}e^{-Nu}\,\mathrm{d}u+\int_0^{-\log F_\mu(t)}\frac{N^N}{(N-1)!}u^{N-1}e^{-Nu}\,\mathrm{d}u.
\end{aligned}$$

The following inequalities will be useful, and they can be established from the Chernov inequality or proved directly:

(14)
$$\forall x\in[0,1),\ \int_0^x\frac{N^N}{(N-1)!}u^{N-1}e^{-Nu}\,\mathrm{d}u\le x^Ne^{N(1-x)},\quad \forall x>1,\ \int_x^{+\infty}\frac{N^N}{(N-1)!}u^{N-1}e^{-Nu}\,\mathrm{d}u\le x^Ne^{N(1-x)}.$$

We have three cases to consider:

Case 1: $-\log F_\mu(t)>1$. If moreover $\lambda\in[0,N\log(-\log F_\mu(t)))$, then $-e^{-\frac{\lambda}{N}}\log F_\mu(t)>1$, and by (14),

$$\begin{aligned}
& e^{\lambda+N\left(1-e^{-\frac{\lambda}{N}}\right)\log F_\mu(t)}\int_{-e^{-\frac{\lambda}{N}}\log F_\mu(t)}^{+\infty}\frac{N^N}{(N-1)!}u^{N-1}e^{-Nu}\,\mathrm{d}u\\
\le\ & e^{\lambda+N\left(1-e^{-\frac{\lambda}{N}}\right)\log F_\mu(t)}\left(-e^{-\frac{\lambda}{N}}\log F_\mu(t)\right)^N e^{N\left(1+e^{-\frac{\lambda}{N}}\log F_\mu(t)\right)}\\
=\ & e^{N(1+\log F_\mu(t)+\log(-\log F_\mu(t)))}<1
\end{aligned}$$

since $1+\log F_\mu(t)+\log(-\log F_\mu(t))<0$. Hence for all $N\ge1$, $\mathbb{E}\left[e^{\lambda\hat{\mu}_N((-\infty,t])}\right]\le 2$, and from the Chernov inequality, we get then: for all $\varepsilon>0$,

$$\mathbb{P}\left\{\hat{\mu}_N((-\infty,t])>\varepsilon\right\}\le\inf_{\lambda\in[0,N\log(-\log F_\mu(t)))}2e^{-\varepsilon\lambda}=2e^{-\varepsilon N\log(-\log F_\mu(t)))}$$

This is enough, with Borel-Cantelli lemma, to prove the almost-sure convergence of $\hat{\mu}_N((-\infty,t])$ to $0=\max(0,1+\log F(t))$.

Case 2: $-\log F_\mu(t)<1$. Then by (14),

$$\int_0^{-\log F_\mu(t)}\frac{N^N}{(N-1)!}u^{N-1}e^{-Nu}\,\mathrm{d}u\le(-\log F_\mu(t))^N e^{N(1+\log F_\mu(t))}=e^{N\delta_t}$$

with $\delta_t=1+\log F_\mu(t)+\log(-\log F_\mu(t))<0$.

- Moreover, we suppose $\lambda\ge0$. Then

$$\begin{aligned}
\mathbb{E}\left[e^{\lambda[\hat{\mu}_N((-\infty,t])-(1+\log F_\mu(t))]}\right] &\le \left(e^{\lambda}F_\mu(t)^Ne^{-Ne^{-\frac{\lambda}{N}}\log F_\mu(t)}+e^{N\delta_t}\right)e^{-\lambda(1+\log F_\mu(t))}\\
&= e^{-N\log F_\mu(t)\left(\frac{\lambda}{N}-(1-e^{-\frac{\lambda}{N}})\right)}+e^{N\left(\delta_t-\frac{\lambda}{N}(1+\log F_\mu(t))\right)}
\end{aligned}$$

From the Chernov inequality, we infer then: for all $\varepsilon>0$

$$\begin{aligned}
\mathbb{P}\left\{\hat{\mu}_N((-\infty,t])-(1+\log F_\mu(t))>\varepsilon\right\} &\le \inf_{\lambda\ge0}\left(e^{-N\log F_\mu(t)\left(\frac{\lambda}{N}-(1-e^{-\frac{\lambda}{N}})\right)}+e^{N\left(\delta_t-\frac{\lambda}{N}(1+\log F_\mu(t))\right)}\right)e^{-\lambda\varepsilon}\\
&= \inf_{\lambda\ge0}\left(e^{-N\left(\log F_\mu(t)\left(\lambda-(1-e^{-\lambda})\right)+\lambda\varepsilon\right)}+e^{N(\delta_t-\lambda(1+\log F_\mu(t)+\varepsilon))}\right)\\
&\le 2e^{-NC}
\end{aligned}$$

for some constant $C > 0$. This gives, using Borel-Cantelli lemma:

$$\limsup_{N\longrightarrow+\infty} \hat{\mu}_N((-\infty,t]) \le 1+\log F_\mu(t) \text{ a.s.}$$

- We suppose now $\lambda \le 0$. Using the same trick, we get

$$\begin{aligned}\mathbb{P}\left\{\hat{\mu}_N((-\infty,t]) - (1+\log F_\mu(t)) < -\varepsilon\right\} &\le \inf_{\lambda\le 0}\left(e^{-N\left(\log F_\mu(t)\left(\lambda-(1-e^{-\lambda})\right)-\lambda\varepsilon\right)} + e^{N(\delta_t-\lambda(1+\log F_\mu(t)-\varepsilon))}\right)\\ &\le 2e^{-NC'}\end{aligned}$$

for some constant $C' > 0$. This gives:

$$\liminf_{N\longrightarrow+\infty} \hat{\mu}_N((-\infty,t]) \ge 1+\log F_\mu(t) \qquad \text{a.s.}$$

And the result follows:

$$\lim_{N\longrightarrow+\infty} \hat{\mu}_N((-\infty,t]) = 1+\log F_\mu(t) = \max(0, 1+\log F_\mu(t)) \qquad \text{a.s.}$$

Case 3: $\log F_\mu(t) = -1$. We know that $F_\mu(u)$ tends to 0 as u tends to $-\infty$. Hence, there is $u < t$ such that $\log F_\mu(u) < -1$. For all N, $\hat{\mu}_N((-\infty,t]) \ge \hat{\mu}_N((-\infty,u])$, hence,

$$\liminf_{N\longrightarrow+\infty} \hat{\mu}_N((-\infty,t]) \ge \liminf_{N\longrightarrow+\infty} \hat{\mu}_N((-\infty,u]) = 0 = 1+\log F_\mu(t) \qquad \text{a.s.}$$

Moreover, for any positive ε, by continuity of F_μ and by the fact that $F_\mu(u)$ tends to 1 as u tends to $+\infty$, there is $u > t$ such that $1+\log F_\mu(t) < 1+\log F_\mu(u) < 1+\log F_\mu(t)+\varepsilon$. Hence,

$$\limsup_{N\longrightarrow+\infty} \hat{\mu}_N((-\infty,t]) \le \limsup_{N\longrightarrow+\infty} \hat{\mu}_N((-\infty,u]) = 1+\log F_\mu(u) < 1+\log F_\mu(t)+\varepsilon \text{ a.s.}$$

Therfore, $\lim_{N\longrightarrow+\infty} \hat{\mu}_N((-\infty,t]) = 1+\log F_\mu(t)$ almost surely.

The theorem is proved in the case where F_μ is smooth. Now, we consider a probability measure μ on the real line without making any assumption about F_μ. By Theorem 1.12.4 of [10], the distance

$$\mathrm{d}(m,m') := \sup\left|\int f\,\mathrm{d}m - \int f\,\mathrm{d}m'\right|,$$

where the sup is taken on the set BL_1 bounded Lipschitz functions f on the real line with Lipschitz constant ≤ 1 and such that $||f||_\infty \le 1$, is a distance which defines the weak topology on the set of probability measures on the real line.

Thus we have to prove that almost surely, as N tends to infinity,

$$\lim \mathrm{d}(\hat{\mu}_N, \Lambda^\vee(\mu)) = 0,$$

i.e. that for all $\varepsilon > 0$, almost surely, for N large enough,

$$\mathrm{d}(\hat{\mu}_N, \Lambda^\vee(\mu)) \le \varepsilon. \tag{15}$$

So let us fix $\varepsilon > 0$. By part (ii) of lemma 4.1, there exists a probability measure ν on the real line such that F_ν is smooth and

$$F_\nu(\bullet - \varepsilon/6) \le F_\mu(\bullet) \le F_\nu(\bullet + \varepsilon/6).$$

Note that the same obviously holds if one replaces μ by $\Lambda^\vee(\mu)$ and ν by $\Lambda^\vee(\nu)$, since for all probability measure m, $F_{\Lambda^\vee(m)} = \max(0, 1+\log F_m)$. Hence, according to part (i) of lemma 4.1,

$$\mathrm{d}(\Lambda^\vee(\mu), \Lambda^\vee(\nu)) \le \varepsilon/6. \tag{16}$$

Note also that by definition of $\Lambda^\vee_N(\mu)$, the vector $(X_1 \ge \cdots \ge X_N)$ of ranked eigenvalues of M_N has law μ_N. Following part (iii) of proposition 4.3, one can suppose that for all N, on the same space as M_N, there is a random matrix N_N with law $\Lambda^\vee_N(\nu)$ with ranked eigenvalues $(Y_1 \ge \cdots \ge Y_k)$ such that for all $i = 1, \ldots, N$, $|X_i - Y_i| \le \varepsilon/3$ almost surely. If one denotes the spectral law of N_N by $\hat{\nu}_N$, it implies that

$$\forall N \ge 1, \mathrm{d}(\hat{\mu}_N, \hat{\nu}_N) \le \varepsilon/3 \qquad \text{almost surely.} \tag{17}$$

Please note finally that by the first part of the proof,

$$\text{almost surely, for } N \text{ large enough,} \qquad \mathrm{d}(\hat{\nu}_N, \Lambda^\vee(\nu)) \leq \varepsilon/3. \tag{18}$$

Equations (16), (17), (18), together, imply (15). Thus, the theorem is proved.

4.6. **Proof of the implication** (6)**.** Let us end the paper with the proof of the implication (6). We consider a sequence F_k of cumulative distribution functions and a cumulative distribution function F such that at any x where F is continuous, $F_k(x)^k \underset{k\to\infty}{\longrightarrow} F(x)$. Let us prove that for any such x, $F_k(x)^{\boxtimes k} \underset{k\to\infty}{\longrightarrow} \Lambda^\vee(F(x))$. Let us denote, for $y \in \mathbb{R}$, $y^+ = \max(0, y)$. Note that since for any $a, b \in \mathbb{R}$ such that $b \leq 0$, we have $(a^+ + b)^+ = (a+b)^+$, by induction on k, we prove easily that for all $G_1, \ldots, G_k$ distribution functions,

$$G_1 \boxtimes \cdots \boxtimes G_k = (G_1 + \cdots + G_k - k + 1)^+.$$

Thus we have to prove

$$(1 + k(F_k(x) - 1))^+ \underset{k\to\infty}{\longrightarrow} (1 + \log F(x))^+, \quad \text{i.e.} \quad k(F_k(x) - 1) \underset{k\to\infty}{\longrightarrow} \log F(x),$$

which follows directly from the hypothesis $F_k(x)^k \underset{k\to\infty}{\longrightarrow} F(x)$.

References

[1] T. Ando. Majorization, doubly stochastic matrices and comparison of eigenvalues. *Linear Algebra Appl.*, 18:163–248, 1989.

[2] G. Ben Arous & D.V. Voiculescu. Free extreme values. *Ann. Probab.*, 34(5):2037–2059, 2006.

[3] F. Benaych-Georges. Classical and free infinitely divisible distributions and random matrices. *Ann. Probab.*, 33(3):1134–1170, 2005.

[4] P. Billingsley. *Convergence of probability measures*. John Wiley & Sons Inc., New York, 1968.

[5] T. Cabanal-Duvillard. A matrix representation of the Bercovici-Pata bijection. *Electron. J. Probab.*, 10(18):632–661, 2005.

[6] B. Collins. Product of random projections, Jacobi ensembles and universality problems arising from free probability. *Probab. Theory Related Fields*, 133(3):315–344, 2005.

[7] H. A. David & H. N. Nagaraja. *Order statistics. Third edition*. Wiley Series in Probability and Statistics. Wiley-Interscience, 2003.

[8] F. Hiai & D. Petz. *The semicircle law, free random variables and entropy*, volume 77 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2000.

[9] M. P. Olson. The selfadjoint operators of a von Neumann algebra form a conditionally complete lattice. *Proc. Amer. Math. Soc.*, 28:537–544, 1971.

[10] A. W. van der Vaart & J. A. Wellner. *Weak convergence and empirical processes*. Springer Series in Statistics. Springer-Verlag, New York, 1996. With applications to statistics.

Florent Benaych-Georges: LPMA, UPMC Univ Paris 6, Case courier 188, 4, Place Jussieu, 75252 Paris Cedex 05, France, and CMAP, École Polytechnique, route de Saclay, 91128 Palaiseau Cedex, France

E-mail address: florent.benaych@gmail.com

Thierry Cabanal-Duvillard: MAP 5, UMR CNRS 8145 - Université Paris Descartes, 45 rue des Saints-Pères 75270 Paris cedex 6, France

E-mail address: Thierry.Cabanal-Duvillard@mi.parisdescartes.fr