Probab. Theory Relat. Fields (2010) 147:161–183
DOI 10.1007/s00440-009-0204-z

Free point processes and free extreme values

G. Ben Arous · V. Kargin

Received: 4 August 2008 / Revised: 15 January 2009 / Published online: 5 March 2009

Abstract We continue here the study of free extreme values begun in Ben Arous and Voiculescu (Ann Probab 34:2037–2059, 2006). We study the convergence of the free point processes associated with free extreme values to a free Poisson random measure (Voiculescu in Lecture notes in mathematics. Springer, Heidelberg, pp. 279–349, 1998; Barndorff-Nielsen and Thorbjornsen in Probab Theory Relat Fields 131:197–228, 2005). We relate this convergence to the free extremal laws introduced in Ben Arous and Voiculescu (Ann Probab 34:2037–2059, 2006) and give the limit laws for free order statistics.

Mathematics Subject Classification (2000) 46L54 · 62G32

1 Introduction

In classical probability theory, the theory of extreme values for i.i.d. random variables is elementary and well understood. Recently, a similar theory has been introduced in the context of free probability theory, in which the role of independent random variables is played by freely independent operators in a Hilbert space [3]. The asymptotic behavior of the maximum of N free operators is given in [3], where the maximum is taken for the spectral order relation on operators [1,13]. The theory emerging is then

Research supported in part by NSF Grants DMS 0808180, and OISE 0730136.

G. Ben Arous (✉)
Courant Institute of Mathematical Sciences, 251 Mercer Street, New York, NY 10012, USA
e-mail: benarous@cims.nyu.edu

V. Kargin
Department of Mathematics, Stanford University, Palo Alto, CA 94305, USA
e-mail: kargin@stanford.edu

parallel to the classical theory for maxima of i.i.d. random variables. In this paper, we make the next step in developing this parallel picture. We study the behavior of the full point process of normalized free extreme values. We show that it converges to a free Poisson random measure as soon as the normalized free maximum converges. One should notice that the notion of "free order statistics" is not readily available. Indeed, the notion of a "second largest statistic" is not at all clear. This difficulty is mirrored in the nature of the limiting object. Free Poisson random variables are not discrete. We will see (in Theorem 3) that our main convergence theorem (Theorem 1) leads to results with no classical analogs for order statistics.

The basic element in both classical and free theory of extremes is a probability measure μ. In the classical case, we take a sequence of i.i.d. random variables X_i, distributed according to μ, and introduce their order statistics, i.e., order them in non increasing order:

$$X^{(0)} \geq X^{(1)} \geq X^{(2)} \geq \cdots \geq X^{(n-1)},$$

so that $X^{(0)}$ is the maximum of the n-sample, $X^{(1)}$ the second largest value and so on. The basic question is to describe the asymptotic behavior of the distribution of these order statistics once properly normalized, when n tends to ∞.

Let $F_{n,k}$ denotes the distribution function of the normalized order statistics $\frac{X^{(k)}-b_n}{a_n}$, for well chosen normalization constants a_n and b_n

$$F_{n,k}(t) = P\left[\frac{X^{(k)} - b_n}{a_n} \leq t\right].$$

The first question addresses the behavior of the maximum, i.e., the asymptotic behavior of $F_{n,0}$. It was shown in the classical works by [7,8], and [9] that there are only three types of possible limit laws, to which $F_{n,0}$ can weakly converge. These laws (Weibull, Frechet or Gumbel) are called "extreme value distributions":

$$\text{Type I}: \quad G(x) = \exp\left(-e^{-x}\right), \qquad -\infty < x < \infty;$$

$$\text{Type II}: \quad G(x) = \begin{cases} 0, & x \leq 0, \\ \exp\left(-x^{-\alpha}\right), \quad \text{for some } \alpha > 0, & x > 0; \end{cases}$$

$$\text{Type III}: \quad G(x) = \begin{cases} \exp\left(-\left(-x\right)^{\alpha}\right), \quad \text{for some } \alpha > 0, & x \leq 0, \\ 1, & x > 0. \end{cases}$$

Moreover, the nature of the max-domain of attraction of these extreme value distributions is well known as well as the possible choices for normalization constants [11,14].

In the free probability context, a sequence of free self-adjoint operators X_i is taken, such that each of X_i has the spectral probability distribution μ. In recent work [3], a maximum operation was defined which maps any n-tuple of self-adjoint operators to another self-adjoint operator, which is called their maximum. The definition is based on the so-called spectral order for self-adjoint operators: $A \preceq B$ iff all spectral projections $1_{(-\infty,t]}(A)$ are greater than or equal to the corresponding spectral projections $1_{(-\infty,t]}(B)$.

The spectral order is stronger than the usual order on operators, according to which $A \leq B$ iff $B - A$ is non-negative definite. The main benefit of the spectral order is that the set of all self-adjoint operators forms a lattice with respect to this order. In particular, if S is the set of all operators C such that $A_i \preceq C$ for each of $A_1, \ldots, A_n$, then S has a unique minimal element which is called $\max\{A_1, \ldots, A_n\}$. This property does not hold if self-adjoint operators are considered with respect to the usual order on operators. Note, however, that the lattice of self-adjoint operators with respect to the spectral order is not a vector lattice in the sense that $A - B \succeq 0$ does not imply that $A \succeq B$. For a counter-example and other information about the spectral order, see [13].

By analogy with the classical case, the sequence of normalized maxima is defined as

$$\max_{1 \leq i \leq n} \{(X_i - b_n I)/a_n\}$$

where the maximum here is understood with respect to the spectral order. Then, $F_{n,0}^{\text{free}}(x)$ is defined as the spectral distribution function of this normalized maximum.

In [3] the following question is solved: When does the sequence of $F_{n,0}^{\text{free}}$ converges weakly?

The answer to this question is very similar to the answer in the classical case: There are only three possible types of limit laws, and for a given μ, the distributions $F_{n,0}^{\text{free}}$ can converge to only one of them:

$$\text{Type I}: \quad G^{\text{free}}(x) = \begin{cases} 0, & x \leq 0, \\ 1 - e^{-x}, & x > 0; \end{cases}$$

$$\text{Type II}: \quad G^{\text{free}}(x) = \begin{cases} 0, & x \leq 1, \\ 1 - x^{-\alpha}, \quad \text{for some } \alpha > 0, & x > 1; \end{cases}$$

$$\text{Type III}: \quad G^{\text{free}}(x) = \begin{cases} 0, & x \leq -1 \\ 1 - (-x)^{\alpha}, \quad \text{for some } \alpha > 0, & -1 < x \leq 0, \\ 1, & x > 0. \end{cases}$$

As in the classical case, this allows defining domains of attraction of the free limit laws. Similar to the results about sums of free operators [6], an important fact is that, even though the limit laws are different in the classical and free cases, the domains of attraction are the same as well as the normalization constants! More precisely $F_{n,0}$ converges weakly to the extreme value distribution $G(x)$ iff $F_{n,0}^{\text{free}}$ converges weakly to G^{free} of the same type as $G(x)$.

This rigid link between classical and free probability theory for extreme values is thus exactly similar to the analogous results for sums of i.i.d. random variables, as developed in [6].

In order to investigate this situation further, let us return to the classical case and consider the random point process

$$N_n = \sum_{i=1}^{n} \delta_{(X_i - b_n)/a_n}.$$

The next question of classical extreme value theory is to understand the convergence of this point process. This question is naturally related to the convergence of the distributions $F_{n,k}$. If μ is in the domain of attraction of a classical extreme value distribution $G(x)$, or equivalently if $F_{n,0}$ converges to $G(x)$ for some choice of normalization constants a_n and b_n, then the point process N_n weakly converges to a Poisson random measure with intensity measure $\lambda\,(dx)$ with $\lambda(x, \infty) = -\log G(x)$. Conversely, if N_n weakly converges to a Poisson random measure with the intensity measure $\lambda\,(dx)$, then the distribution of any order statistics $F_{n,k}$ converges to a limit law $G_{(k)}$ which is easily computable from $\lambda(dx)$ or equivalently from $G(x)$, see below or [14].

What is the free analogue of the point process N_n? To motivate our definition, note that we can think about N_n as a linear functional on the space of bounded measurable functions: $\langle N_n, f\rangle =: \sum_{i=1}^{n} f\left((X_i - b_n)/a_n\right)$. This functional takes values in the space of bounded random variables. We will define a free point process analogously. We begin with a slightly greater generality and associate a free random process to any triangular array of free random variables.

Let $\overline{\mathcal{A}}$ be the set of densely-defined closed operators affiliated with a von Neumann algebra $\mathcal{A}$, and let $\mathcal{B}_\infty(\mathbb{R})$ denote the set of all bounded, Borel measurable functions $f : \mathbb{R} \to \mathbb{R}$.

Definition 1 Let $X_{i,n} \in \overline{\mathcal{A}}$, $(i = 1, \ldots, n; n = 1, \ldots)$ be a triangular array of freely independent self-adjoint variables. Then the free point process M_n associated with the array $X_{i,n}$ is the sequence of $\mathcal{A}$-valued functionals on $\mathcal{B}_\infty(\mathbb{R})$, defined by the following formula:

$$\langle M_n, f\rangle := \sum_{i=1}^{n} f\left(X_{i,n}\right).$$

The triangular array of free variables that we use in applications to free extremes is, of course, $X_{i,n} = (X_i - b_n)/a_n$, where X_i is a sequence of free self-adjoint variables.

We can also define the concept of weak convergence of a free point process as a weak-$*$ convergence of the corresponding functionals. In the classical case, after a suitable scaling, the point process N_n converges to a Poisson random measure. It turns out that in the non-commutative case the free point process converges to a *free Poisson random measure*, which was recently defined in [15] and [2]. The following three theorems are the main results of our paper.

Theorem 1 *Let $G(x)$ be a classical extreme value distribution, i.e. a Gumbel, Frechet or Weibull distribution. Let $\overline{x} = \inf\{x \,:\, G(x) > 0\}$ and define a measure $\lambda\,(dx)$ on $[\overline{x}, \infty)$ by the equality $\lambda\left((x, \infty)\right) = -\log G(x)$. The following statements are equivalent:*

(i) *μ belongs to the domain of attraction of the classical extremal limit law $G(x)$, i.e., for some constants a_n and b_n the distribution $F_{n,0}$ converges weakly to $G(x)$;*
(ii) *μ belongs to the domain of attraction of the free extremal limit law G^{free}, i.e., for some constants a_n and b_n the spectral distribution of the normalized free maximum, $F_{n,0}^{\text{free}}$ converges weakly to G^{free};*
(iii) *For some a_n and b_n, the point process N_n weakly converges on $(\overline{x}, \infty)$ to the Poisson random measure with intensity $\lambda(dx)$;*
(iv) *For some a_n and b_n, the free point process M_n weakly converges on $(\overline{x}, \infty)$ to the free Poisson random measure with intensity $\lambda(dx)$.*

In case one of the equivalent conditions in Theorem 1 is satisfied, then all the normalization constants a_n and b_n can be taken to be the same in all four statements.

The equivalences of (i) and (iii) follows from the results in [14] (see, e.g., Section 4.2.2 on page 209), and the equivalence of (i) and (ii) was proved in [3]. Thus, we only need to prove the equivalence of (i) and (iv).

The equivalence of (i) and (iv) will be seen, in Sect. 3, as a consequence of the following more general result about convergence of free point processes. Recall that a measure is called Radon if $\mu(K) < \infty$ for every compact K.

Theorem 2 *Let $X_{i,n}$ be a triangular array of free, self-adjoint random variables and let the spectral probability measure of $X_{i,n}$ be μ_n. Let λ be a Radon measure on $D \subseteq \mathbb{R}$. The free point process M_n associated with the array $X_{i,n}$ converges weakly on D to a free Poisson random measure M with the intensity measure λ if and only if*

$$n\mu_n(A) \to \lambda(A) \tag{1}$$

for every Borel set $A \subseteq D$.

We now want to show what Theorem 1 implies for free order statistics. We begin by recalling basic facts about the classical theory of extreme values. If the measure μ is in the domain of attraction of the extreme value distribution $G(x)$, then as mentioned above, the convergence of the point process N_n implies easily the convergence of order statistics. Indeed with the notations introduced above, it is easy to relate the distribution $F_{n,k}$ of the normalized kth order statistics to the point process N_n, through the basic identity:

$$F_{n,k}(t) = P\left[\frac{X^{(k)} - b_n}{a_n} \le t\right] = P\left[N_n(t, \infty) \le k\right] = E\left[1_{[0,k]}(\langle N_n, 1_{(t,\infty)}\rangle)\right].$$

This implies easily that the distribution $F_{n,k}$ of the properly normalized order statistics weakly converges to the distribution

$$G_{(k)}(t) = \sum_{j=0}^{k} e^{-\lambda(t,\infty)} \frac{\lambda(t,\infty)^j}{j!}.$$

We now want to see how this translates in the free context. More precisely, let $X_1, \ldots, X_n$ be freely independent self-adjoint variables with (possibly different) distribution functions F_i. Consider M_n the free point process associated with the sequence X_i and let

$$Y_n(t) := \langle M_n, 1_{(t,\infty)} \rangle = \sum_{i=1}^{n} 1_{(t,\infty)}(X_i).$$

Definition 2 For every real $k \geq 0$, we say that $F_{n,k}^{\mathrm{free}}(t) := E\left[1_{[0,k]}(Y_n(t))\right]$ is the *distribution function of the* kth *order statistic* of the sequence $X_1, \ldots, X_n$, and that it is the kth *order free extremal convolution* of the spectral distribution functions F_i.

Note that the definition is valid not only for all integer k but also for all non-negative real k.

One question that immediately arises is whether we can define an operator, for which the distribution $F_{n,k}^{\mathrm{free}}(t)$ would be a spectral distribution function? The answer to this question is positive. The condition $t' \geq t$ implies that $Y_n(t') \leq Y_n(t)$ and $1_{[0,k]}(Y_n(t')) \geq 1_{[0,k]}(Y_n(t))$. Therefore, as t grows, the operators $1_{[0,k]}(Y_n(t))$ form an increasing family of projections and we can use this family to construct the required operator by the spectral resolution theorem.

Definition 3 For every real $k \geq 0$, let

$$Z^{(k)} = \int t \, d1_{[0,k]}(Y_n(t)).$$

We call $Z^{(k)}$ the kth *order statistic* of the family X_i.

From the construction it is clear that $F_{n,k}(t)$ is the spectral distribution function of the operator $Z^{(k)}$.

In complete analogy with the classical case the limits of these free extremal convolutions can be computed using the limits of free point measures. If $G(x)$ is one of the classical limit laws, then we use $G^{(-1)}(x)$ to denote the functional inverse of $G(x)$. Let

$$t_-(k) = G^{(-1)}\left(\exp\left[-\left(1+\sqrt{k}\right)^2\right]\right),$$
$$t_0(k) = G^{(-1)}\left(\frac{1}{e}\right),$$
$$t_+(k) = G^{(-1)}\left(\exp\left[-\left(1-\sqrt{k}\right)^2\right]\right).$$

Let $\lambda(t) = -\log G(t)$ and $p_t(\xi) = (2\pi\xi)^{-1}\sqrt{4\xi - (1-\lambda(t)+\xi)^2}$.

Theorem 3 *Suppose that measure μ belongs to the domain of attraction of a (classical) limit law $G(x)$ and a_n, b_n are the corresponding norming constants. Assume* *that X_i are free self-adjoint variables with the spectral probability measure μ and let $F_{n,k}^{\text{free}}(t)$ denote the distribution of the kth order statistic of the family $(X_i - b_n)/a_n$, where $i = 1, \ldots, n$. Then, as $n \to \infty$, the distribution $F_{n,k}^{\text{free}}(t)$ converges to a limit, $F_{(k)}(t)$, which is given by the following formula:*

$$F_{(k)}(t) = \begin{cases} 0, & \text{if } t < t_-, \\ \int_{(1-\sqrt{\lambda(t)})^2}^{k} p_t(\xi)\, d\xi, & \text{if } t \in [t_-, t_0], \\ 1 - \lambda_t + \int_{(1-\sqrt{\lambda(t)})^2}^{k} p_t(\xi)\, d\xi, & \text{if } (t_0, t_+], \\ 1 - \lambda(t)\, 1_{[0,1)}(k), & \text{if } t > t_+. \end{cases}$$

It turns out that in the particular case of the 0-order free extremal convolutions, their limits coincide with the limits discovered in [3] (see Definition 6.8 and Theorems 6.9 and 6.11):

$$F_{(0)}^{I}(t) = \left(1 - e^{-t}\right) 1_{(0,\infty)}(t);$$
$$F_{(0)}^{II}(t) = \left(1 - \frac{1}{t^{\alpha}}\right) 1_{(1,\infty)}(t); \text{ and}$$
$$F_{(0)}^{III}(t) = \left(1 - |t|^{\alpha}\right) 1_{(-1,0)}(t) + 1_{[0,\infty)}(t),$$

where α is a positive parameter.

While we were mainly motivated by trying to extend the classical probabilistic phenomena to the setting of free probability, it is worth mentioning that the theory of free extreme values is directly related to natural operations on random matrices (see the recent preprint [4]). The results of this paper can easily be translated in the context of [4].

The rest of the paper is organized as follows. Section 2 gives a brief introduction to free probability theory. Section 3 proves Theorem 1 using Theorem 2. Section 4 details the definition of the convergence of free point process and proves Theorem 2. And Sect. 5 proves Theorem 3.

2 Preliminaries

2.1 Free independence

Definition 4 A W^**-probability space* is a pair $(\mathcal{A}, E)$, where $\mathcal{A}$ is a von Neumann algebra of bounded linear operators acting on elements of a complex separable Hilbert space and E is a faithful normal trace that satisfies the condition $E(I) = 1$. Operators affiliated with algebra $\mathcal{A}$ are called *non-commutative random variables,* or simply *random variables*, and the functional E is called the *expectation.*

If $P(d\lambda)$ is the spectral resolution associated with a normal operator A, then we can define a measure $\mu(d\lambda) = E(P(d\lambda))$. It is easy to check that μ is a probability measure supported on the spectrum of A. We call this measure, μ, the *spectral probability measure associated with operator A and expectation E.*

The most important concept in free probability theory is that of free independence of non-commutative random variables. Let a set of r.v. $A_1, \ldots, A_n$ be given. With each of them we can associate an algebra $\mathcal{A}_i$, which is generated by A_i and A_i^*; that is, it is the weak topology closure of all polynomials in variables A_i and A_i^*. Let $\overline{A_i}$ denote an arbitrary element of algebra $\mathcal{A}_i$.

Definition 5 The algebras $\mathcal{A}_1, \ldots, \mathcal{A}_n$ (and variables $A_1, \ldots, A_n$ that generate them) are said to be *freely independent* or *free*, if the following condition holds:

$$E\left(\overline{A}_{i(1)} \ldots \overline{A}_{i(m)}\right) = 0,$$

provided that $E\left(\overline{A}_{i(s)}\right) = 0$ and $i(s+1) \neq i(s)$ for every s.

For more information about non-commutative probability spaces and free operators we refer the reader to Sections 2.2–2.5 in the book [16] by Voiculescu, Dykema and Nica.

If X and Y are two free self-adjoint random variables with spectral probabilities measures μ and ν respectively, then we denote the spectral probability measure of $X + Y$ as $\mu \boxplus \nu$, and call it the *free additive convolution* of μ and ν.

2.2 Free Poisson random variables

Let X be a self-adjoint operator that has the so-called *free Poisson distribution* with parameter ("intensity") λ. The continuous part of this distribution is supported on the interval $[(1-\sqrt{\lambda})^2, (1+\sqrt{\lambda})^2]$ and the density is

$$p_\lambda(x) = \frac{\sqrt{4x - (1 - \lambda + x)^2}}{2\pi x}.$$

In addition, if $\lambda < 1$, then there is also an atom at zero with the probability weight $1 - \lambda$. We call such an operator X a (non-commutative) *Poisson random variable* with intensity λ and size 1.

The sum of two freely independent Poisson random variables of intensities λ_1 and λ_2 is again a Poisson random variable of intensity $\lambda_1 + \lambda_2$ (see, for example, a remark on page103 in [10]).

If we scale a non-commutative Poisson random variable by a, then we get a variable, which we call a *scaled (non-commutative) Poisson random variable* of intensity λ and size a.

Non-commutative Poisson random variables arise when we convolve a large number, N, of Bernoulli distributions that put probability λ/N on 1 and probability $1-\lambda/N$ on 0. The following result is well-known, see [10,12], or [15].

Proposition 1 *Suppose* μ_n, $(n = 1, 2, \ldots)$ *is a sequence of Bernoulli distributions, such that* $\mu_n(\{1\}) \sim \lambda/n$ *and* $\mu_n(\{0\}) = 1 - \mu_n(\{1\})$. *Define* ν_n *as follows:*

$$\nu_n = \underbrace{\mu_n \boxplus \cdots \boxplus \mu_n}_{n \text{times}}.$$

Then ν_n *weakly converges to the free Poisson distribution with intensity* λ *and size* 1.

2.3 Free Poisson random measure

Definition 6 Let $(\Theta, \mathcal{B}, \nu)$ be a measure space, and put

$$\mathcal{B}_0 = \{B \in \mathcal{B} : \nu(B) < \infty\}.$$

Let further $(\mathcal{A}, E)$ be a W^*-probability space, and let $\mathcal{A}_+$ denote the cone of positive operators in $\mathcal{A}$. Then a *free Poisson random measure* (*fPrm*) on $(\Theta, \mathcal{B}, \nu)$ with values in $(\mathcal{A}, E)$ is a mapping $M : \mathcal{B}_0 \to \mathcal{A}_+$, with the following properties:

(i) For any set B in $\mathcal{B}_0$, $M(B)$ is a free Poisson variable with parameter $\nu(B)$.
(ii) If $r \in \mathbb{N}$, and $B_1, \ldots, B_r \in \mathcal{B}_0$ are disjoint, then $M(B_1), \ldots, M(B_r)$ are free.
(iii) If $r \in \mathbb{N}$, and $B_1, \ldots, B_r \in \mathcal{B}_0$ are disjoint, then $M(\cup_{j=1}^r B_j) = \sum_{j=1}^r M(B_j)$.

The existence of a free Poisson measure for arbitrary spaces $(\Theta, \mathcal{B}, \nu)$ and $(\mathcal{A}, E)$ was shown in [15] and a different proof was given in [2].

Let f be a real-valued simple function in $L^1(\Theta, \mathcal{B}, \nu)$, i.e, suppose that it can be written as

$$f = \sum_{i=1}^r a_i 1_{B_i},$$

for a system of disjoint $B_i \in \mathcal{B}_0$. Then we define the integral of f with respect to a Poisson random measure M as follows:

$$\int_\Theta f\, dM = \sum_{i=1}^r a_i M(B_i).$$

It is possible to check that this definition is consistent. Moreover, as it is shown in [2], this concept can be extended to a larger class of functions:

Proposition 2 *Let f be a real-valued function in $L^1(\Theta, \mathcal{B}, \nu)$ and suppose that s_n is a sequence of real valued simple $\mathcal{B}$-measurable functions, satisfying the condition that there exists a positive ν-integrable function $h(\theta)$, such that $|s_n(\theta)| \leq h(\theta)$ for all n and θ. Suppose also that $\lim_{n\to\infty} s_n(\theta) = f(\theta)$ for all θ Then integrals $\int_\Theta s_n\, dM$ are well-defined and converge in probability to a self-adjoint (possibly unbounded) operator $I(f)$ affiliated with $\mathcal{A}$. Furthermore, the limit $I(f)$ is independent of the choice of approximating sequence s_n of simple functions.*

The resulting functional $I(f)$ is defined for all real valued functions f in $L^1(\Theta, \mathcal{B}, \nu)$ and is called the *integral with respect to the free Poisson random measure M*. It possesses all the usual properties of the integral: additivity, linear scaling, continuity, etc.

3 Proof of Theorem 1

As was noted in Introduction, only the equivalence of (i) and (iv) needs a proof. The equivalence of (i) and (iv) can be reduced to a problem about convergence of free point processes. Indeed, let $\mu_n(A) = \mu(a_n A + b_n)$. Then (i) is equivalent to the statement that $n\mu_n(A) \to \lambda(A)$ for all Borel sets $A \subset (\overline{x}, \infty)$.

Indeed, suppose that μ is in the domain of attraction of $G(x)$, and let $F(x)$ denote the distribution function of the measure μ. Then

$$F^n(a_n x + b_n) \to G(x),$$

For every $x \in (\overline{x}, \infty)$, $G(x)$ is positive, hence we can take logarithms and get

$$n \log F(a_n x + b_n) \to \log G(x),$$

which is equivalent to

$$n(1 - F(a_n x + b_n)) \to -\log G(x) \equiv \lambda((x, \infty)).$$

Consequently,

$$n\mu_n((x, \infty)) \to \lambda((x, \infty)),$$

from which we conclude that $n\mu_n(A) \to \lambda(A)$ for all Borel sets $A \subset (\overline{x}, \infty)$.

By reversing the steps of this argument we obtain the reverse implication: If $n\mu_n(A) \to \lambda(A)$ for all Borel sets $A \subset (\overline{x}, \infty)$, then μ is in the domain of attraction of $G(x)$, and (i) holds.

Therefore the equivalence of (i) and (iv) follows from Theorem 2 if we take $(X_i - a_n)/b_n$ as the triangular array $X_{i,n}$.

4 Proof of Theorem 2

4.1 Weak convergence

In this section, we define precisely the mode of convergence of free point measures that we use. It corresponds to the weak convergence of point processes in the classical case.

Let D be a Borel subset of $\mathbb{R}$ and let $\mathcal{F}_K^\infty(D)$ denote the space of bounded, Borel measurable functions that have compact support on D.

Definition 7 We say that a free point process M_n *converges weakly on* D to a free Poisson random measure M, which is defined on $(D, \mathcal{B}, \lambda)$ and takes values in $\mathcal{A}$, if for every function $f \in \mathcal{F}_K^\infty(D)$ the following convergence holds:

$$\langle M_n, f \rangle \xrightarrow{d} \int_{\mathbb{R}} f \, dM.$$

Sometimes we also need to speak about convergence with respect to a class of functions, which is different from $\mathcal{F}_K^\infty(D)$.

Definition 8 We say that a free point process M_n *converges weakly with respect to a class of functions* $\mathcal{F}$ to a free Poisson random measure M, if for every function $f \in \mathcal{F}$ the following convergence holds:

$$\langle M_n, f \rangle \xrightarrow{d} \int_{\mathbb{R}} f \, dM.$$

We will prove Theorem 2 by considering initially the convergence of free point processes M_n with respect to the class of simple functions (i.e., finite sums of indicator functions), and then approximating functions from a more general class by simple functions.

4.2 Convergence with respect to simple functions

Let $\mathcal{S}(D)$ be the class of simple functions on $D \subset \mathbb{R}$, i.e., the class of finite sums of indicator functions of Borel sets belonging to D.

Proposition 3 *Let $X_{i,n}$ be a triangular array of free, self-adjoint random variables and let the spectral probability measure of $X_{i,n}$ be μ_n. Let λ be a Radon measure on $D \subseteq \mathbb{R}$. If*

$$n\mu_n(A) \to \lambda(A)$$

for each Borel set $A \subset D$, then the free point process M_n associated with the array $X_{i,n}$ converges weakly with respect to $\mathcal{S}(D)$ to a free Poisson random measure M with the intensity measure λ.

Before proving this proposition, we derive some auxiliary results.

Lemma 1 *Suppose $X_{i,n}$ is an array of free and identically distributed random variables with the spectral measure μ_n. Let $n\mu_n(A) \to \lambda(A) < \infty$ as $n \to \infty$. Let $Z_{i,n} = 1_A(X_{i,n})$. Then as $n \to \infty$, the sum $S_n = \sum_{i=1}^n Z_{i,n}$ converges in distribution to a free Poisson random variable with intensity $\lambda(A)$.*

Proof Note that $Z_{i,n}$ are projections with expectation $\mu_n(A)$ and they are free. Therefore, $\sum_{i=1}^n Z_{i,n}$ is the sum of free projections and we can use Proposition 1 to infer the claim of the lemma. □

As the next step to the proof of Proposition 3 we need to check that if Borel sets A_k are disjoint, then the sums $S_k = \sum_{i=1}^n 1_{A_k}\left(X_{i,n}\right)$ are asymptotically free with respect to growing n.

Recall the definition of the asymptotic freeness: Let $(\mathcal{A}_i, E_i)$ be a sequence of non-commutative probability spaces and let X_i and Y_i be two random variables in $\mathcal{A}_i$. Let also x and y be two free operators in a non-commutative probability space $(\mathcal{A}, E)$.

Definition 9 The sequences X_i and Y_i are called *asymptotically free* if the sequence of pairs (X_i, Y_i) converges in distribution to the pair (x, y). That is, for every $\varepsilon > 0$ and every sequence of k -tuples $(n_1, \ldots, n_k)$ with non-negative integers n_j, there exists such i_0 that for $i \geq i_0$, the following inequality holds:

$$\left| E_i \left(X_i^{n_1} Y_i^{n_2}, \ldots, X_i^{n_{k-1}} Y_i^{n_k} \right) - E \left(x^{n_1} y^{n_2}, \ldots, x^{n_{k-1}} y^{n_k} \right) \right| \leq \varepsilon.$$

At the cost of more complicated notation, this definition can be generalized to the case of more than two variables.

Lemma 2 *Let $P_{i,n}^{(k)}$, (where $n = 1, 2, \ldots; i = 1, \ldots, n$, and $k = 1, \ldots, r$) be projections of dimension $\lambda^{(k)}/n$. Assume that for each n, algebras $\mathcal{A}_i$ generated by sets $\left\{ P_{i,n}^{(k)} \right\}_{k=1}^{r}$ are free. Also assume that for each n and i, the projections $P_{i,n}^{(k)}$ are orthogonal to each other, i.e., $P_{i,n}^{(k)} P_{i,n}^{(k')} = 0$ for every pair $k \neq k'$. Let $S_n^{(k)} = \sum_{i=1}^{n} P_{i,n}^{(k)}$. Then as $n \to \infty$, the sequences $S_n^{(k)}$ converge in distribution to freely independent variables $S^{(k)}$ that have free Poisson distributions with parameters $\lambda^{(k)}$, respectively. In particular, the sequences $S_n^{(k)}$ are asymptotically free with respect to growth in n.*

Proof The fact that each of the sequences $S_n^{(k)}$ converge in distribution to a variable $S^{(k)}$ that has a free Poisson distribution is clear from Proposition 1. The essential part is to prove that asymptotic freeness holds. This claim is a direct consequence of Speicher's multidimensional limit theorem (see, for example, Theorem 13.1 in the book [12] by Nica and Speicher). Indeed, we need to prove that all mixed free cumulants of the limit are zero. By Speicher's theorem, this is equivalent to the statement that the following limits are zero:

$$\lim_{n \to \infty} nE \left(P_{1,n}^{(k_1)} P_{1,n}^{(k_2)} \ldots P_{1,n}^{(k_s)} \right) = 0.$$

Here $k_1, \ldots, k_s$ is an arbitrary s-tuple with the property that it has a pair of distinct coordinates, i.e., $k_i \neq k_j$. However, the fact that these limits are zero is clear from the assumption that the projections $P_{i,n}^{(k)}$ are orthogonal to each other. □

Now we can proceed to the proof of Proposition 3.

Proof Let $f = \sum_{k=1}^{r} c_k 1_{A_k}(x)$, where A_k are disjoint Borel sets. Using the assumption that $n\mu_n(A_k) \to \lambda(A_k)$ and Lemma 1, we can find a free Poisson random measure M such that

$$\sum_{i=1}^{n} 1_{A_k}\left(X_{i,n}\right) \xrightarrow{d} M(A_k) = \int_{\mathbb{R}} 1_{A_k}(x)\, M(dx)$$

as $n \to \infty$. Indeed, it is enough to take a Poisson random measure M with the intensity measure λ.

In addition, by Lemma 2, sums $S_k = \sum_{i=1}^{n} 1_{A_k}(X_{i,n})$ become asymptotically free for different k as n grows. Since $M(A_k)$ are free by the definition of the free Poisson measure, this implies that

$$\sum_{k=1}^{r} c_k \sum_{i=1}^{n} 1_{A_k}\left(X_{i,n}\right) \xrightarrow{d} \sum_{k=1}^{r} c_k M(A_k) = \sum_{k=1}^{r} c_k \int_{\mathbb{R}} 1_{A_k}(x)\, M(dx).$$

as $n \to \infty$. Therefore,

$$\sum_{i=1}^{n} f\left(X_{i,n}\right) \xrightarrow{d} \int_{\mathbb{R}} f(x)\, M(dx),$$

where we used the additivity property of the integral with respect to a free Poisson random measure (see [2], Remark 4.2(b)). □

4.3 Convergence with respect to bounded, Borel measurable functions with compact support

The goal of this section is to prove our main Theorem 2.

Consider a bounded, Borel measurable, compactly supported function $f : D \to \mathbb{R}$, such that $0 \le f \le 1$. (A more general case of a function f, which satisfies $C_1 \le f \le C_2$, can be treated similarly.) For positive integers $N = 1, 2, \ldots,$ and $k = 1, \ldots, N$, define the set

$$A_k^{(N)} = \left\{ x \in \operatorname{supp}(f) : \frac{k-1}{N} < f(x) \le \frac{k}{N} \right\}.$$

The sets $A_k^{(N)}$ are disjoint, measurable, and have finite λ-measure. Their union is D.

We define lower and upper approximations to the function f as follows:

$$l^N(x) = \sum_{k=1}^{N} \frac{k-1}{N} 1_{A_k^{(N)}}(x),$$

and

$$u^N(x) = \sum_{k=1}^{N} \frac{k}{N} 1_{A_k^{(N)}}(x),$$

We note that:

(i) $l^N(x) \le u^N(x)$;
(ii) $l^N(x)$ is an increasing sequence of functions;
(iii) $u^N(x)$ is a decreasing sequence of functions, and

(iv) $\lim_{N\to\infty} l^N(x) - u^N(x) = 0$ uniformly in x.

The functions $l^N(x)$ and $u^N(x)$ are simple: $l^N(x) = \sum_{i=1}^N c_k^{(N)} 1_{A_k^{(N)}}(x)$ and $u^N(x) = \sum_{i=1}^N d_k^{(N)} 1_{A_k^{(N)}}(x)$. Note also that $\sup_k (d_k^{(N)} - c_k^{(N)}) = 1/N$ converges to zero as $N \to \infty$.

Let us drop for convenience the superscript N when we consider it as fixed, and simply write $l(x) = \sum_{i=1}^N c_k 1_{A_k}(x)$ and $u(x) = \sum_{i=1}^N d_k 1_{A_k}(x)$, where A_k are disjoint Borel-measurable sets. By Proposition 3, as $n \to \infty$,

$$\sum_{i=1}^{n} l\left(X_{i,n}\right) \xrightarrow{d} \sum_{k=1}^{N} c_k M_k,$$

where M_k are freely independent Poisson random variables with intensities $\lambda_k = \lambda(A_k)$. Let $F_l(x)$ denote the distribution function of $\sum_{k=1}^N c_k M_k$.

Similarly,

$$\sum_{i=1}^{n} u\left(X_{i,n}\right) \xrightarrow{d} \sum_{k=1}^{N} d_k M_k,$$

and we denote the distribution function of $\sum_{k=1}^N d_k M_k$ as $F_u(x)$.

Let $F_{f,n}$ denote the distribution function of $\sum_{i=1}^n f(X_{i,n})$ and let F_f be one of the limit points of this sequence of distribution functions.

Proposition 4 *F_f is a distribution function and $F_u(x) \le F_f(x) \le F_l(x)$ for every x.*

Proof We will infer this from Lemma 3 below and its Corollary. This lemma is a particular case of Weyl's eigenvalue inequalities for operators in a von Neumann algebra of type II_1. If $F_A(x)$ is the spectral distribution function of a self-adjoint operator A, then we define the eigenvalue function $\theta_A(t) = \inf\{x : F_A(x) \ge 1 - t\}$. The function $\theta_A(t)$ is non-increasing and right-continuous. Intuitively, it can be thought of as a "sequence of eigenvalues" of A, indexed in decreasing order by parameter t.

Let us use notation $\theta_A(t-0)$ to denote $\lim_{\varepsilon \downarrow 0} \theta_A(t-\varepsilon)$. Then the following generalization of Weyl inequalities holds:

Lemma 3 *If A and B are two bounded self-adjoint operators from a W^*-probability space $\mathcal{A}$ and if B is non-negative definite, then*

$$\theta_A(t) \le \theta_{A+B}(t) \le \theta_A(t) + \|B\|, \text{ and}$$
$$\theta_A(t-0) \le \theta_{A+B}(t-0) \le \theta_A(t-0) + \|B\|.$$

Corollary 1 *If $B \ge 0$, then $\mu_{A+B} \gg \mu_A$, that is, $F_{A+B}(x) \le F_A(x)$ for each x.*

Proof of Lemma 3 These results easily follow from an inequality in [5] which states that if $(a-\varepsilon, a) \subset [0,1]$, $(b-\varepsilon, b) \subset [0,1]$, and $a+b \leq 1$, then

$$\int_{a+b-\varepsilon}^{a+b} \theta_{A+B}(t)\,dt \leq \int_{a-\varepsilon}^{a} \theta_A(t)\,dt + \int_{b-\varepsilon}^{b} \theta_B(t)\,dt. \tag{2}$$

□

By Corollary 1, for each n the distribution $F_{f,n}$ is between the distribution functions of $\sum_{i=1}^{n} u(X_{i,n})$ and $\sum_{i=1}^{n} l(X_{i,n})$. As n grows, these two sequences of distribution functions approach $F_u(x)$ and $F_l(x)$, respectively. Therefore, every limit point of $F_{f,n}$ is between F_u and F_l. The claim that F_f is a distribution function follows from the fact that both F_u and F_l are distribution functions. □

Now we want to show that $F_u^{(N)}(x)$ approaches $F_l^{(N)}(x)$ as N grows.

Recall that the *Levy distance* between two distribution functions is defined as follows:

$$d_L(F_A, F_B) = \sup_x \inf\{s \geq 0 : F_B(x-s) - s \leq F_A(x) \leq F_B(x+s) + s\}.$$

We can interpret this distance geometrically. Let Γ_A be the graph of function F_A, and at the points of discontinuity let us connect the left and right limits by a (vertical) straight line interval. Call the resulting curve $\widetilde{\Gamma}_A$. Similarly define $\widetilde{\Gamma}_B$. Let d be the maximum distance between $\widetilde{\Gamma}_A$ and $\widetilde{\Gamma}_B$ in the direction from the south-east to the north-west, i.e., in the direction which is obtained by rotating the vertical direction by $\pi/4$ counter-clockwise. Then $d_L(F_A, F_B) = d/\sqrt{2}$.

Proposition 5 *Let K be the sum of intensities of freely independent Poisson random variables M_k and let $F_l(x)$ and $F_u(x)$ be distribution functions of $\sum_{k=1}^{N} c_k M_k$ and $\sum_{k=1}^{N} d_k M_k$ Then*

$$d_L(F_l, F_u) \leq \left(2K + 3\sqrt{K} + 1\right) \sup_{1\leq k\leq N} (d_k - c_k).$$

Remark In the proof of Theorem 2, the finiteness of K will be ensured by the assumptions that measure λ is Radon and that f has a compact support.

For the proof of this proposition we need two lemmas. Lemma 4 provides a bound on the norm of the sum of scaled Poisson random variables in terms of the sizes of these variables, and Lemma 5 relates the Levy distance between two random variables to the norm of their difference.

Lemma 4 *Let M_i, $(i = 1, \ldots, r)$ be freely independent Poisson random variables, which have intensities λ_i, and let b_i be non-negative real numbers. Assume that*

$\sum_{i=1}^{r}\lambda_i \leq K$ *and let* $b = \sup_{1\leq i\leq r} b_i$. *Then*

$$\left\|\sum_{i=1}^{r} b_i M_i\right\| \leq b\left(2K + 3\sqrt{K} + 1\right).$$

Proof Let X_i be free self-adjoint random variables that have zero mean. Then by an inequality from [17]:

$$\left\|\sum_{i=1}^{r} X_i\right\| \leq \max_{1\leq i\leq r} \|X_i\| + \sqrt{\sum_{i=1}^{r} \mathrm{Var}\,(X_i)}.$$

If Y_i are free self-adjoint random variables with non-zero mean, and $X_i = Y_i - E\,(Y_i)$, then the previous inequality implies that

$$\begin{aligned}\left\|\sum_{i=1}^{r} Y_i\right\| &\leq \left|\sum_{i=1}^{r} E\,(Y_i)\right| + \left\|\sum_{i=1}^{r} X_i\right\| \\ &\leq \left|\sum_{i=1}^{r} E\,(Y_i)\right| + \max_{1\leq i\leq r} d\,(Y_i) + \sqrt{\sum_{i=1}^{r} \mathrm{Var}\,(Y_i)}, \qquad (3)\end{aligned}$$

where $d\,(Y_i)$ is the diameter of the support of Y_i.

We will apply this inequality to $Y_i = b_i M_i$ and estimate each of the three terms on the right-hand side of (3) in turn:

(1) Since $E\,(M_i) = \lambda_i$, and $\sum \lambda_i \leq K$, therefore $\sum_{i=1}^{r} b_i E\,(M_i) \leq bK$.
(2) The diameter of the support of $b_i M_i$ is less or equal to $b_i(1 + \sqrt{\lambda_i})^2 \leq b(1 + 2\sqrt{K} + K)$.
(3) Since $\mathrm{Var}\,(M_i) = \lambda_i$, therefore $\sqrt{\sum_{i=1}^{r} \mathrm{Var}\,(b_i M_i)} \leq b\sqrt{K}$.

In sum, $\left\|\sum_{i=1}^{r} b_i M_i\right\| \leq b(2K + 3\sqrt{K} + 1)$. □

Lemma 5 *Let A and B be two bounded self-adjoint operators from a* W^**-probability space* $\mathcal{A}$ *and assume that* $B - A \geq 0$. *Then*

$$d_L\,(F_A, F_B) \leq \|B - A\|.$$

Proof Let F_A and F_B be distribution functions, and θ_A and θ_B be the corresponding eigenvalue functions. Then we claim that

$$d_L\,(F_A, F_B) \leq \sup_{0\leq t\leq 1} |\theta_A\,(t) - \theta_B\,(t)|. \qquad (4)$$

Indeed, let the graphs of functions θ_A and θ_B be denoted as Λ_A and Λ_B, respectively. Connecting the left and right limits at the points of discontinuity gives us the curves $\widetilde{\Lambda}_A$ and $\widetilde{\Lambda}_B$. It is easy to see that these curves can be obtained from curves $\widetilde{\Gamma}_A$ and

$\widetilde{\Gamma}_B$ (i.e., the graphs of $F_A(x)$ and $F_B(x)$ with connected limits at the points of discontinuity) by rotating them around the point (0, 1) counter-clockwise by the angle $\pi/2$ and then shifting the result of the rotation by vector $(0, -1)$. It follows that the distance d, which was used in the definition of the Levy distance can also be defined as the maximum distance between $\widetilde{\Lambda}_A$ and $\widetilde{\Lambda}_B$ in the direction from the south-west to the north-east, i.e., in the direction which is obtained by rotating the vertical direction by $\pi/4$ clockwise.

Since $\theta_A(t)$ and $\theta_B(t)$ are non-increasing functions, therefore

$$d \leq \sqrt{2} \sup_{0 \leq t \leq 1} |\theta_A(t) - \theta_B(t)|.$$

This implies $d_L(F_A, F_B) \leq \sup_{0 \leq t \leq 1} |\theta_A(t) - \theta_B(t)|$.

Inequality (4) and Lemma 3 imply the statement of the lemma. □

Now we can prove Proposition 5:

Proof of Proposition 5 Let $X = \sum_{k=1}^{N} (d_k - c_k) M_k$. By Lemma 4, $\|X\| \leq b(2K + 3\sqrt{K} + 1)$, where $b = \sup_{1 \leq k \leq N} (d_k - c_k)$ and K is the sum of the intensities of M_k. By Lemma 5, this implies that $d_L(F_l, F_u) \leq b(2K + 3\sqrt{K} + 1)$. □

Using Proposition 5, we can proceed to the proof of Theorem 2. By Proposition 3, we know that if N is fixed and $n \to \infty$, then

$$\sum_{i=1}^{n} l^N(X_{i,n}) \xrightarrow{d} \sum_{k=1}^{N} c_i^{(N)} M\left(A_k^{(N)}\right),$$

and

$$\sum_{i=1}^{n} u^N(X_{i,n}) \xrightarrow{d} \sum_{k=1}^{N} d_i^{(N)} M\left(A_k^{(N)}\right),$$

where M is a free Poisson random measure with intensity $\lambda(dx)$. Let the distributions of the right-hand sides be denoted as F_{l^N} and F_{u^N}.

By Corollary 1 (p. 15), F_{l^N} is a decreasing sequence and F_{u^N} is an increasing sequence of distribution functions. In addition, $F_{l^N}(x) \geq F_{u^N}(x)$ for every N and x. Since the sum of intensities of variables $M\left(A_k^{(N)}\right)$ is less than $\lambda(D) < \infty$ by assumption, therefore Proposition 5 is applicable and we can conclude that the Levy distance between F_{l^N} and F_{u^N} converges to zero as $N \to \infty$. Consequently, these two distributions (weakly) converge to a limit distribution function as $N \to \infty$.

Moreover, by the definition of the integral with respect to a free Poisson random measure, this limit equals the distribution function of $\int f(x) M(dx)$.

In addition, by Proposition 4 every limit point of the sequence of $F_{f,n}$ is between F_{l^N} and F_{u^N} for every N, and therefore the sequence of $F_{f,n}$ also converges to the distribution function of $\int f(x) M(dx)$. □

This completes the proof of Theorem 2.

5 Proof of Theorem 3

Recall that we defined the distribution of a free order statistic in the following way. Let $X_1, \ldots, X_n$ be freely independent self-adjoint random variables and let X_i have the spectral distribution functions F_i. Let

$$Y_n(t) = \sum_{i=1}^{n} 1_{(t,\infty)}(X_i).$$

Definition 2 For every real $k \geq 0$, we say that $F_{n,k}(t) =: E\left[1_{[0,k]}(Y_n(t))\right]$ is the distribution function of kth order statistic of the sequence $X_1, \ldots, X_n$, and that it is the kth order free extremal convolution of distribution functions F_i.

It is straightforward to check that in the case of commutative random variables, this definition gives the distribution function of the usual $(\lfloor k \rfloor + 1)$-order statistic.

In the non-commutative case, we need to check that this is a consistent definition, and that $F_{n,k}(t)$ is indeed a probability distribution function for each $k \geq 0$.

It is easy to see that $F_{n,k}(t)$ is non-decreasing in t. Indeed, let $t' \geq t$. Then for each i, $1_{(t',\infty)}(X_i) \leq 1_{(t,\infty)}(X_i)$, and therefore, $Y\left(t'\right) \leq Y(t)$. It follows that $1_{[0,k]}\left(Y\left(t'\right)\right) \geq 1_{[0,k]}(Y(t))$, and therefore $F_{n,k}\left(t'\right) \geq F_{n,k}(t)$.

This function is also right-continuous in t. Consider a sequence $t_m \downarrow t$. First, note that $1_{(t_m,\infty)}(X_i) \stackrel{d}{\to} 1_{(t,\infty)}(X_i)$. Second, since operators $1_{(t,\infty)}(X_i)$ are freely independent for different i, this implies that $Y(t_m) \stackrel{d}{\to} Y(t)$ as $t_m \downarrow t$. Indeed, the operators $Y(t_m)$ and $Y(t)$ are uniformly bounded ($\|Y(t_m)\| \leq n$ and $\|Y(t)\| \leq n$), and the moments of the distribution of $Y(t_m)$ converge to the corresponding moments of the distribution of $Y(t)$.

Third, let the spectral probability distribution functions of $Y(t_m)$ and $Y(t)$ be denoted as $G_m(x)$ and $G(x)$, respectively. Then $E\left[1_{[0,k]}(Y(t_m))\right] = G_m(k)$ and $E\left[1_{[0,k]}(Y(t))\right] = G(k)$. Since $G_m(k) \equiv F_{n,k}(t_m)$, and $G(k) \equiv F_{n,k}(t)$, therefore we aim to prove that $G_m(k) \to G(k)$ as $m \to \infty$ for all k. The convergence $Y(t_m) \stackrel{d}{\to} Y(t)$ means the convergence of the moments of the spectral probability measures of operators $Y(t_m)$ and $Y(t)$, which implies weak convergence of these measures because the measures have uniformly bounded support. This implies that $G_m(k) \to G(k)$ as $m \to \infty$, for all points k at which the probability distribution function $G(k)$ is continuous. We will prove that, moreover, even if $G(x)$ has a jump at $x = k$, then the sequence $G_m(k)$ still converges to $G(k)$. At this point of the argument, it is essential that t_m converges to t from above and therefore $G_m(k) \geq G(k)$.

Indeed, by seeking a contradiction, suppose that $G_m(k)$ does not converge to $G(k)$. Then, take ε such that (i) $G_m(k) - G(k) > \varepsilon$ for all m, and take $k' > k$ such that (ii) k' is a point of continuity of $G(x)$, and (iii) $G\left(k'\right) - G(k) < \varepsilon/2$. Such k' exists because $G(x)$ is a spectral probability distribution function and therefore it is right-continuous. Since $G_m(k)$ is increasing, we conclude from (i), (ii), and (iii) that $G_m\left(k'\right) - G\left(k'\right) > \varepsilon/2$ for all m. But this means that $G_m(x)$ does not converge to $G(x)$ at a point of continuity of $G(x)$, namely, at k'. This is a contradiction, and we conclude that $G_m(k)$ converges to $G(k)$ for all k. This means that $F_k(t)$ is right-continuous in t.

Finally, as $t \to \infty$, $1_{(t,\infty)}(X_i) \xrightarrow{d} 0$. Therefore $Y(t) \xrightarrow{d} 0$, and $1_{[0,k]}(Y(t)) \xrightarrow{d} I$. Hence $F_{n,k}(t) \to 1$ as $t \to \infty$, and we conclude that $F_{n,k}(t)$ is a valid distribution function.

Consider now the special case when $k = 0$. In this case $F_{n,0}(t)$ is the dimension of the nill-space of $Y_n(t)$, which equals to the dimension of the intersection of the nill-spaces of $1_{(t,\infty)}(X_i)$. It is easy to see that this coincides with the definition of the free extremal convolution of the distribution F, which was introduced in [3].

Now let us investigate the question of the limiting behavior of the distributions $F_{n,k}(t)$ when $n \to \infty$. The limits are described in Theorem 3.

Proof of Theorem 3 For each n we re-define:

$$Y_n(t) = \sum_{i=1}^{n} 1_{(t,\infty)}\left(\frac{X_i - b_n}{a_n}\right) = \left\langle M_n, 1_{(t,\infty)}\right\rangle,$$

where M_n is the free point process associated with the triangular array $(X_i - b_n)/a_n$.

The bracket $\langle M_n, 1_{(t,\infty)}\rangle$ converges in distribution to a random variable C_t, which is a free Poisson random variable with the intensity $\lambda(t) = -\log G(t)$. Then, in order to calculate the limit of $F_{n,k}(t)$ for $n \to \infty$, we only need to calculate $E1_{[0,k]}(C_t)$, that is, the distribution function of C_t at k. Let us denote the distribution function of C_t as $H_t(x)$,

For $k < 0$, we have $H_t(k) = 0$. For $k = 0$,

$$H_t(0) = \begin{cases} 1 - \lambda(t), & \text{if } \lambda(t) \le 1, \\ 0, & \text{if } \lambda(t) > 1. \end{cases}$$

For $k > 0$,

$$H_t(k) = \begin{cases} H_t(0), & \text{if } k < \left(1 - \sqrt{\lambda(t)}\right)^2, \\ H_t(0) + \int_{(1-\sqrt{\lambda(t)})^2}^{k} p_t(\xi)\, d\xi, & \text{if } k \in \left[\left(1 - \sqrt{\lambda(t)}\right)^2, \left(1 + \sqrt{\lambda(t)}\right)^2\right], \\ 1 & \text{if } k > \left(1 + \sqrt{\lambda(t)}\right)^2. \end{cases}$$

where

$$p_t(\xi) = \frac{\sqrt{4\xi - (1 - \lambda(t) + \xi)^2}}{2\pi\xi}.$$

Then, we need to compute $F_{(k)}(t)$, which is $H_t(k)$ considered a function of t for a fixed k. Let $\lambda^{-1}(x)$ denote the solution of the equation $\lambda(t) = x$. (That is, if $G^{(-1)}(x)$ is the functional inversion of the limit distribution function $G(t)$, then $\lambda^{-1}(x) = G^{(-1)}(e^{-x})$.)

Then, for $k = 0$:

$$F_{(k)}(t) = \begin{cases} 0, & \text{if } t \leq \lambda^{-1}(1), \\ 1 - \lambda(t), & \text{if } t > \lambda^{-1}(1). \end{cases}$$

For $k \in (0, 1)$:

$$F_{(k)}(t) = \begin{cases} 0, & \text{if } t < \lambda^{-1}\left(\left(1+\sqrt{k}\right)^2\right), \\ \int_{\left(1-\sqrt{\lambda(t)}\right)^2}^{k} p_t(\xi)\, d\xi, & \text{if } t \in \left[\lambda^{-1}\left(\left(1+\sqrt{k}\right)^2\right), \lambda^{-1}(1)\right], \\ 1-\lambda(t)+\int_{\left(1-\sqrt{\lambda(t)}\right)^2}^{k} p_t(\xi)\, d\xi, & \text{if } \left(t \in \lambda^{-1}(1), \lambda^{-1}\left(\left(1-\sqrt{k}\right)^2\right)\right], \\ 1-\lambda(t), & \text{if } t > \lambda^{-1}\left(\left(1-\sqrt{k}\right)^2\right). \end{cases}$$

For $k \geq 1$, we have:

$$F_{(k)}(t) = \begin{cases} 0, & \text{if } t < \lambda^{-1}\left(\left(1+\sqrt{k}\right)^2\right), \\ \int_{\left(1-\sqrt{\lambda(t)}\right)^2}^{k} p_t(\xi)\, d\xi, & \text{if } t \in \left[\lambda^{-1}\left(\left(1+\sqrt{k}\right)^2\right), \lambda^{-1}(1)\right], \\ 1-\lambda(t)+\int_{\left(1-\sqrt{\lambda(t)}\right)^2}^{k} p_t(\xi)\, d\xi, & \text{if } \left(t \in \lambda^{-1}(1), \lambda^{-1}\left(\left(1-\sqrt{k}\right)^2\right)\right], \\ 1, & \text{if } t > \lambda^{-1}\left(\left(1-\sqrt{k}\right)^2\right). \end{cases}$$

Combining these cases, we obtain the following equation:

$$F_{(k)}(t) = \begin{cases} 0, & \text{if } t < \lambda^{-1}\left(\left(1+\sqrt{k}\right)^2\right), \\ \int_{\left(1-\sqrt{\lambda(t)}\right)^2}^{k} p_t(\xi)\, d\xi, & \text{if } t \in \left[\lambda^{-1}\left(\left(1+\sqrt{k}\right)^2\right), \lambda^{-1}(1)\right], \\ 1-\lambda(t)+\int_{\left(1-\sqrt{\lambda(t)}\right)^2}^{k} p_t(\xi)\, d\xi, & \text{if } \left(t \in \lambda^{-1}(1), \lambda^{-1}\left(\left(1-\sqrt{k}\right)^2\right)\right], \\ 1-\lambda(t)\, 1_{[0,1)}(k), & \text{if } t > \lambda^{-1}\left(\left(1-\sqrt{k}\right)^2\right). \end{cases}$$

□

Example Distributions from the domain of attraction of Type II extremal value law

Consider the case of convergence to the Type II extremal value law, when the constants a_n and b_n are chosen in such a way, that the limit law is $G(x) = \exp\left(-x^{-\nu}\right)$ for $x > 0$.Then we can conclude that the limit distribution of the k order statistic is

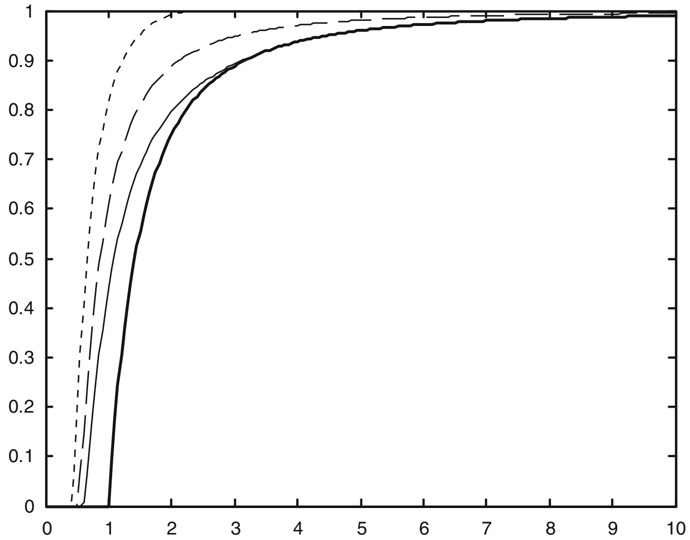

Fig. 1 Distributions of free kth order extreme statistics. The *horizontal axis* is for t, the *vertical axis* is for corresponding probabilities $F(t)$. The figure was computed for the extreme value distribution of type II with the parameter $v = 2$. The *bold solid line* is for $k = 0$, the *thin solid line* is for $k = 0.5$, the *dashed line* is for $k = 1$, and the *dotted line* is for $k = 2$

given as follows:

$$F_{(k)}(t) = \begin{cases} 0, & \text{if } t < \left(1+\sqrt{k}\right)^{-2/\nu}, \\ \int_{(1-t^{-\nu/2})^2}^{k} p_t(\xi)\, d\xi, & \text{if } t \in \left[\left(1+\sqrt{k}\right)^{-2/\nu}, 1\right], \\ 1 - t^{-\nu} + \int_{(1-t^{-\nu/2})^2}^{k} p_t(\xi)\, d\xi, & \text{if } t \in \left(1, \left(\left(1-\sqrt{k}\right)^2\right)^{-1/\nu}\right], \\ 1 - t^{-\nu} 1_{[0,1)}(k), & \text{if } t > \left(\left(1-\sqrt{k}\right)^2\right)^{-1/\nu}, \end{cases}$$

where

$$p_t(\xi) = \frac{\sqrt{4\xi - \left(1 - t^{-\nu} + \xi\right)^2}}{2\pi\xi}.$$

We illustrate this result for some particular values of ν and k.
Consider $k = 0$. Then

$$F_{(0)}(t) = \begin{cases} 0, & \text{if } t < 1, \\ 1 - t^{-\nu}, & \text{if } t \geq 1. \end{cases}$$

This is the Type 2 (“Pareto”) limit distribution in Definition 6.8 of [3].

The distributions of k order statistics for different values of k are illustrated in Fig. 1.

It is interesting to note that if $k > 1$, then for all sufficiently large t, $F_{(k)}(t) = 1$. This can be interpreted as saying that the scaled k order statistic is guaranteed to be less than t_0 for a sufficiently large t_0. In another interpretation, this result means that for our choice of scaling parameters a_n and b_n and for every $k > 1$, if t is sufficiently large, then

$$\left\| \sum_{i=1}^{n} 1_{(a_n t + b_n, \infty)}(X_i) \right\| < k$$

for all large n.

A similar situation occurs in the classical case if the initial distribution (i.e. the distribution of X_i) is bounded from above. In this case the limit distribution is also bounded from above. In contrast, in the free probability case this situation occurs even if the initial distribution is unbounded from above. Our previous example shows that this situation occurs even if the initial distribution has heavy tails.

Acknowledgments The authors would like to thank two very diligent referees for their thorough reports. G. Ben Arous would like to thank D. Voiculescu for useful discussions.

References

1. Ando, T.: Majorization, doubly stochastic matrices, and comparison of eigenvalues. Linear Algebra Appl. **118**, 163–248 (1989)
2. Barndorff-Nielsen, O.E., Thorbjornsen, S.: The Levy-Ito decomposition in free probability. Probab. Theory Relat. Fields **131**, 197–228 (2005)
3. Ben Arous, G., Voiculescu, D.: Free extreme values. Ann. Probab. **34**, 2037–2059 (2006)
4. Benaych-Georges, F., Cabanal-Duvillard, T.: A matrix interpolation between classical and free max operations. I. The Univariate Case (2008). http://www.arxiv.org/abs/0806.3686v2
5. Bercovici, H., Li, W.S.: Inequalities for eigenvalues of sums in a von Neumann algebra. In: Kerchy, L., Foias, C., Gohberg, I., Langer, H. (eds.) Recent Advances in Operator Theory and Related Topics. Operator Theory Advances and Applications, vol. 127, pp. 113–126. Birkhauser, Basel (2001)
6. Bercovici, H., Pata, V., Biane, P.: Stable laws and domains of attraction in free probability theory. Ann. Math. **149**:1023–1060 (1999) (with an appendix by Philippe Biane)
7. Fisher, R.A., Tippett, L.H.C.: Limiting forms of the frequency distribution of the largest or smallest member of a sample. Proc. Camb. Philos. Soc. **24**, 180–190 (1928)
8. Frechet, M.: Sur la loi de probabilite de l'ecart maximum. Ann. Soc. Pol. Math. **6**, 93 (1927)
9. Gnedenko, B.: Sur la distribution limite du terme maximum d'une serie aleatoire. Ann. Math. **44**, 423–453 (1943)
10. Hiai, F., Petz, D.: The semicircle law, free random variables and entropy. Mathematical Surveys and Monographs, 1 edn, vol. 77. American Mathematical Society (2000)
11. Leadbetter, M.R., Lindgren, G., Rootzen, H.: Extremes and Related properties of random sequences and processes. Springer Series in Statistics. Springer, Heidelberg (1983)
12. Nica, A., Speicher, R.: Lectures on the combinatorics of free probability. London Mathematical Society Lecture Note Series, vol. 335. Cambridge University Press, London (2006)
13. Olson, M.P.: The self-adjoint operators of a von Neumann algebra form a conditionally complete lattice. Proc. Am. Math. Soc. **28**, 537–544 (1971)
14. Resnick, S.I.: Extreme values, regular variation, and point processes. Applied Probability, vol. 4. Springer, New York (1987)

15. Voiculescu, D.: Lectures on free probability theory. In: Bernard, P. (ed.) Lectures on probability theory and statistics. Lecture Notes in Mathematics, vol. 1738, pp. 279–349. Springer, Heidelberg (1998)
16. Voiculescu, D., Dykema, K., Nica, A.: Free random variables. CRM Monograph Series, No. 1. A.M.S. Providence, RI (1992)
17. Voiculescu, D.: Addition of certain non-commuting random variables. J. Funct. Anal. **66**, 323–346 (1986)