PROCESSES OF rTH LARGEST

BORIS BUCHMANN, ROSS MALLER, AND SIDNEY I. RESNICK

Abstract. For integers $n \geq r$, we treat the rth largest of a sample of size n as an $\mathbb{R}^\infty$-valued stochastic process in r which we denote $\boldsymbol{M}^{(r)}$. We show that the sequence regarded in this way satisfies the Markov property. We go on to study the asymptotic behaviour of $\boldsymbol{M}^{(r)}$ as $r \to \infty$, and, borrowing from classical extreme value theory, show that left-tail domain of attraction conditions on the underlying distribution of the sample guarantee weak limits for both the range of $\boldsymbol{M}^{(r)}$ and $\boldsymbol{M}^{(r)}$ itself, after norming and centering. In continuous time, an analogous process $\boldsymbol{Y}^{(r)}$ based on a two-dimensional Poisson process on $\mathbb{R}_+ \times \mathbb{R}$ is treated similarly, but we find that the continuous time problems have a distinctive additional feature: there are always infinitely many points below the rth highest point up to time t for any $t > 0$. This necessitates a different approach to the asymptotics in this case.

1. Introduction

In this paper we consider Markovian and other properties of the order statistics of iid random variables in discrete time, and of extremal processes in continuous time. Although venerable these are important issues and research continues to throw up significant new aspects. As a starting point let $M_n^{(r)}$ be the rth largest among iid random variables $X_1, \dots, X_n$ with cdf F. (Precise specifications of the order statistics will be given later.) It is known that the *finite* sequence $(M_n^{(r)})_{r=1,2,\dots,n}$ is Markov if and only if F is continuous on (ℓ_F, r_F), where ℓ_F and r_F are the left and right extremes of F (see [1]). This is a result concerning the first r order statistics in a finite sample. We proceed to investigate the infinitely many order statistics $(M_n^{(r)}, n \geq r)$ beyond the rth, and further, derive properties of the whole collection $\{\boldsymbol{M}^{(r)} = (M_n^{(r)}, n \geq r), r \geq 1\}$, considered as an $\mathbb{R}^\infty$-valued stochastic process. Apart from their intrinsic interest the properties we derive bring together a number of areas and techniques.

Thus, we begin in Section 2 by setting up the notation required for, then proving, the Markovian property, that the conditional distribution of the infinite sequence $(M_{r+1}^{(r+1)}, M_{r+2}^{(r+1)}, \dots)$, knowing all values $(M_1^{(1)}, M_2^{(1)}, \dots)$, $(M_2^{(2)}, M_3^{(2)}, \dots)$, $\dots$, $(M_r^{(r)}, M_{r+1}^{(r)}, \dots)$, is the same as the conditional distribution knowing only $(M_r^{(r)}, M_{r+1}^{(r)}, \dots)$. No continuity assumptions on F are required for this.

Key words and phrases. extremes, domain of attraction, Markov property, extremal process.

This research was initiated and partially supported by ARC grants DP1092502 and DP160104737. S. Resnick also received significant support from US Army MURI grant W911NF-12-1-0385 to Cornell University; Resnick gratefully acknowledges hospitality, administrative support and space during several visits to the Research School of Finance, Actuarial Studies & Statistics, Australian National University.

In Section 3 we turn to an investigation of asymptotic properties of the collection $\boldsymbol{M}^{(r)}$, for large values of r. The weak convergence of $\boldsymbol{M}^{(r)}$, after norming and centering, is related to domain of attraction theory for the *minimum* of an iid sequence of rvs. A key tool in these proofs is Ignatov's [12] theorem showing that the r-records of an iid sequence are points of a Poisson random measure.

This study is continued in Section 4 for continuous time rth-order extremal processes. Some notable differences between the discrete and continuous time situations emerge here. In particular, unlike in the discrete case, in the continuous time case there are always infinitely many points below the currently considered order statistic, and thus the convergence criterion has to be modified. Section 5 concludes the paper with some modest final thoughts and open problems.

We conclude the present section by mentioning previous and related work. For alternative proofs and other background on Ignatov's (1977) theorem see [9, 10, 12, 21, 26]. Other treatments of the Markov structure of the finite sequence $(M_n^{(r)})_{r=1,2,\dots,n}$ are in [11], [23] and [2]. The latter two papers show that $(M_n^{(r)})_{r=1,2,\dots,n}$ is Markov if information on tied values is incorporated into the sequence. For background on continuous time extremal processes we refer to [18, 19, 21, 22]. Additional references are given throughout the text.

2. Markov Property of Higher Order Extremal Processes with Discrete Indexing

2.1. **Indexing.** Our analysis requires that we keep track of infinite sequences indexed by r where the first members are being moved further out as r increases. To cope with this we use the idea of shifted sequences, with first members replaced by $-\infty$. To see how this works, we start with the sequence space $\mathbb{R}_{-\infty}^{\mathbb{N}} := \{\boldsymbol{x} = (x_n) : x_n \in \mathbb{R}_{-\infty}\,, n \in \mathbb{N}\}$ endowed with the Borel field associated with the product topology. (We employ the notations $\mathbb{N} = \{1, 2, 3, \dots\}$, $\mathbb{R}_{-\infty} := \mathbb{R} \cup \{-\infty\} = [-\infty, \infty)$, and conventions $\sum_\emptyset = 0$, $\prod_\emptyset := 1$, $\pm\infty \times 0 = 0$. Also $\mathbb{R}_{-\infty}^{\mathbb{N},\uparrow} = \{\boldsymbol{x} = (x_n) \in \mathbb{R}_{-\infty}^{\mathbb{N}} :\, x_n \le x_{n+1}\,, n \in \mathbb{N}\}$ denotes the subset of nondecreasing sequences.) The partial maxima operator $\bigvee : \mathbb{R}_{-\infty}^{\mathbb{N}} \mapsto \mathbb{R}_{-\infty}^{\mathbb{N},\uparrow}$ maps a given sequence $\boldsymbol{x} = (x_n)_n \in \mathbb{R}_{-\infty}^{\mathbb{N}}$ to its associated sequence of partial maxima $\bigvee \boldsymbol{x} := (\vee\{x_1, \dots, x_n\})_n$. (In the statistical language R, this is known as *cummax*.)

For a given sequence $\boldsymbol{x} \in \mathbb{R}_{-\infty}^{\mathbb{N}}$ and $r \in \mathbb{N}$, $n \ge r$, let $m_n^{(r)}$ be the rth largest of $x_1, \dots, x_n$, arranged in lexicographical order in case of ties. Then set

$$x_n^{(r)} = \begin{cases} -\infty, & \text{if } n < r; \\ m_n^{(r)}, & \text{if } n \ge r. \end{cases}$$

The *extremal sequence of order r associated with $\boldsymbol{x}$* is the sequence $\boldsymbol{x}^{(r)} \in \mathbb{R}_{\infty}^{\mathbb{N},\uparrow}$, with finite elements $x_n^{(r)}$ augmented with $-\infty$ as follows:

$$\boldsymbol{x}^{(r)} = \big(\underbrace{-\infty, \dots, -\infty,}_{r-1 \text{ entries}} m_n^{(r)}, n \ge r\big). \tag{2.1}$$

Write $\boldsymbol{x}^{(0)} := \boldsymbol{x}$ for the extremal sequence of zero order. The extremal sequence of unit order equals the partial maximum sequence: $\boldsymbol{x}^{(1)} = \bigvee \boldsymbol{x}$.

For a sequence $\boldsymbol{x} = (x_n)_n \in \mathbb{R}^{\mathbb{N}}_{-\infty}$ the *shifted sequence* $\boldsymbol{x}_{\mathcal{R}}$ is $\boldsymbol{x}_{\mathcal{R}} = (-\infty, \boldsymbol{x}) \in \mathbb{R}^{\mathbb{N}}_{-\infty}$. For two sequences $\boldsymbol{x} = (x_n)_n$, $\boldsymbol{y} = (y_n)_n \in \mathbb{R}^{\mathbb{N}}_{-\infty}$, let

$$\boldsymbol{x}_{\mathcal{R}} \wedge \boldsymbol{y} := \{(-\infty)\mathbf{1}_{n=1} + (x_{n-1} \wedge y_n)\mathbf{1}_{n>1}\}_n \in \mathbb{R}^{\mathbb{N}}_{-\infty}$$

be the componentwise minimum of $\boldsymbol{x}$ and $\boldsymbol{y}$, taken after shifting $\boldsymbol{x}$ to the right with proper augmentation with $-\infty$. Thus, componentwise, when $\mathbf{x} = (x_1, x_2, \ldots)$ and $\mathbf{y} = (y_1, y_2, \ldots)$, we have

$$\boldsymbol{x}_{\mathcal{R}} = (-\infty, x_1, x_2, \ldots) \quad \text{and} \quad \boldsymbol{x}_{\mathcal{R}} \wedge \boldsymbol{y} = (-\infty, x_1 \wedge y_2, x_2 \wedge y_3, \ldots).$$

For $n \in \mathbb{N}$, $y_n^{(1)} \geq y_n^{(2)} \geq \cdots \geq y_n^{(n)}$ denotes the order statistics associated with (possibly extended) real numbers $y_1, \ldots, y_n \in \mathbb{R}_{-\infty}$.

In Theorem 2.1, we will show a Markov property for the rth largest of an iid sequence, and since recursions are an effective tool for proving a sequence of random elements is Markovian, we first prove a preliminary result focussing on properties of the shifted sequences.

Proposition 2.1. *For $r \in \mathbb{N}$, we have the identity,*

$$\boldsymbol{x}^{(r+1)} = \bigvee (\boldsymbol{x}^{(r)}{}_{\mathcal{R}} \wedge \boldsymbol{x}) \tag{2.2}$$

or in component form,

$$x_n^{(r+1)} = \bigvee_{j=r+1}^{n} \Big(x_{j-1}^{(r)} \wedge x_j\Big), \quad r \in \mathbb{N},\, n \geq r+1. \tag{2.3}$$

Proof. Fix an integer r and we prove (2.3) by induction on n. The base of the induction is $n = r+1$ and the left side of (2.3) is $x_{r+1}^{(r+1)} = \wedge_{i=1}^{r+1} x_i$. The right side is $x_r^{(r)} \wedge x_{r+1} = \wedge_{i=1}^{r+1} x_i$. So (2.3) is proved for $n = r+1$.

As an induction hypothesis, assume (2.3) is true for $n = r+p$ for $p \geq 1$ and we verify (2.3) is true for $n = r+p+1$. The left side of (2.3) for $n = r+p+1$ is $x_{r+p+1}^{(r+1)} = LHS$. The right side is

$$RHS = \bigvee_{j=r+1}^{r+p+1} \big(x_{j-1}^{(r)} \wedge x_j\big) = \bigvee_{j=r+1}^{r+p} \big(x_{j-1}^{(r)} \wedge x_j\big) \bigvee \big(x_{r+p}^{(r)} \wedge x_{r+p+1}\big)$$

and from the induction hypothesis this is equal to

$$x_{r+p}^{(r+1)} \bigvee \big(x_{r+p}^{(r)} \wedge x_{r+p+1}\big). \tag{2.4}$$

Now consider cases:

Case (a) $x_{r+p+1} > x_{r+p}^{(r)}$**:** For this case, increasing the sample size from $r+p$ to $r+p+1$ means $x_{r+p}^{(r)}$ becomes $x_{r+p+1}^{(r+1)}$. So $RHS = x_{r+p}^{(r+1)} \bigvee x_{r+p}^{(r)} = x_{r+p}^{(r)} = LHS$.

Case (b) $x_{r+p}^{(r+1)} \leq x_{r+p+1} \leq x_{r+p}^{(r)}$**:** The term in parentheses on the right side of (2.4) is

$$x_{r+p}^{(r)} \wedge x_{r+p+1} = x_{r+p+1} = x_{r+p+1}^{(r+1)}$$

and thus

$$RHS = x_{r+p}^{(r+1)} \vee x_{r+p+1}^{(r+1)} = x_{r+p+1}^{(r+1)} = LHS.$$

Case (c) $x_{r+p+1} < x^{(r+1)}_{r+p}$**.:** We have

$$RHS = x^{(r+1)}_{r+p} \vee \left(x^{(r)}_{r+p} \wedge x_{r+p+1}\right) = x^{(r+1)}_{r+p} \vee x_{r+p+1} = x^{(r+1)}_{r+p}$$

and because of where the added point x_{r+p+1} is located, when the sample size increases from $r+p$ to $r+p+1$, the above equals $x^{(r+1)}_{r+p+1} = LHS$.

The three cases exhaust the possibilities and this completes the induction argument. □

2.2. **The iid setting.** Now we add the randomness. Let $\boldsymbol{X} = (X_n)_n \in \mathbb{R}^{\mathbb{N}}$ be an iid sequence of rvs in $\mathbb{R}$ with cdf F and set $\boldsymbol{X}^{(0)} = \boldsymbol{X}$. Then for $r \in \mathbb{N}$ the *r-th order extremal process* is the augmented sequence $\boldsymbol{X}^{(r)} = (X^{(r)}_n)_{n\in\mathbb{N}}$ in $\mathbb{R}^{\mathbb{N}}_{-\infty}$ constructed as in (2.1); specifically,

$$\boldsymbol{X}^{(r)} = \big(\underbrace{-\infty, \dots, -\infty,}_{r-1 \text{ entries}} M^{(r)}_n, n \geq r\big), \tag{2.5}$$

where the $M^{(r)}_n$ are the order statistics of $X_1, X_2, \dots, X_n$ defined lexicographically as for the $m^{(r)}_n$ in (2.1). Note that $\boldsymbol{X}^{(1)} = \bigvee \boldsymbol{X}^{(0)} = \bigvee \boldsymbol{X}$ is the sequence of partial maxima associated with $\boldsymbol{X}$.

To think about the Markov property for $(\boldsymbol{X}^{(r)}, r \geq 1)$, we imagine conditioning on the monotone sequence $\boldsymbol{X}^{(r)} = \boldsymbol{x}^{(r)}$. For indices where the sequence $\boldsymbol{x}^{(r)}$ is a constant, say x, the structure of $\boldsymbol{X}^{(r+1)}$ should be as if we construct the maximum sequence from repeated observations from the conditional distribution of $(X_1 | X_1 \leq x)$. The following construction make this precise.

Let $\boldsymbol{U} = (U_{r,n})_{n,r\in\mathbb{N}}$ be an iid array of uniform r.v.'s in $(0,1)$. Assume $\boldsymbol{X} = \boldsymbol{X}^{(0)}$ and $\boldsymbol{U}$ are independent random elements. For $m \in \mathbb{R}$ with $F(m) > 0$ the left-continuous inverse $u \mapsto F^{\leftarrow}(u|m)$ of the conditional cdf $x \mapsto F(x|m) := P(X_1 \leq x | X_1 \leq m)$ is well-defined; otherwise, if $F(m) = 0$ set $F^{\leftarrow}(u|m) = \mathbf{1}_{m>0}$ with $F^{\leftarrow}(u| -\infty) \equiv 0$.

For $r \in \mathbb{N} = \{1, 2, \dots\}$ introduce two sequences $\widehat{\boldsymbol{X}}_{(r+1)} = (\widehat{X}_{(r+1),n})_n$ and $\widetilde{\boldsymbol{X}}_{(r+1)} = (\widetilde{X}_{(r+1),n})_n$. For the first, we have for $n = 1$ that $\widehat{X}_{(r+1),1} := X^{(1)}_1 = X_1$ and, for $n \geq 2$,

$$\widehat{X}_{(r+1),n} := \begin{cases} F^{\leftarrow}(U_{r,n} | X^{(r)}_n) \prod_{1\leq k\leq r} \mathbf{1}_{X^{(k)}_n = X^{(k)}_{n-1}}, & \text{if } X^{(r)}_{n-1} = X^{(r)}_n \\ \sum_{k=1}^r X^{(k)}_n \mathbf{1}_{X^{(k)}_n > X^{(k)}_{n-1}} \prod_{1\leq l<k} \mathbf{1}_{X^{(l)}_n = X^{(l)}_{n-1}}, & \text{if } X^{(r)}_{n-1} < X^{(r)}_n \end{cases}$$

so if there is no jump in the rth order maximum process we sample from the conditional distribution and if there is a jump, we note the new value that caused the jump. For the second sequence we have $\widetilde{X}_{(r+1),n} := -\infty$ if $n \leq r$ and if $n > r$

$$\widetilde{X}_{(r+1),n} := \begin{cases} X^{(r)}_{n-1}, & \text{if } X^{(r)}_n > X^{(r)}_{n-1}, \\ F^{\leftarrow}(U_{r,n} | X^{(r)}_n), & \text{if } X^{(r)}_n = X^{(r)}_{n-1}, \end{cases}$$

so if there is no jump in the rth order maxima at n, we sample from the conditional distribution and if there is a jump at index n we note the smaller value at $n-1$ that the process jumps from. The sequences $\widetilde{\boldsymbol{X}}_{r+1}$ and $\widehat{\boldsymbol{X}}_{(r+1)}$ depend on $\boldsymbol{X}, \boldsymbol{X}^{(1)}, \dots, \boldsymbol{X}^{(r)}$ only via $\boldsymbol{X}^{(r)}$ and $\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \dots, \boldsymbol{X}^{(r)}$, respectively.

2.3. **Identities in Law and the Markov Property.** Next we provide some identities in law which will show that the sequence $\boldsymbol{X}^{(r)}, r \geq 1$ of extremal processes is a sequence-valued Markov chain.

Theorem 2.1. *For $r \in \mathbb{N}$ the following random variables are equal in distribution as random elements in $(\mathbb{R}_{-\infty}^{\mathbb{N}})^{(r+1)}$ (and hence in $(\mathbb{R}_{-\infty}^{\mathbb{N}})^{\mathbb{N}}$),*

$$(\boldsymbol{X}^{(0)}, \ldots, \boldsymbol{X}^{(r)}) \stackrel{d}{=} (\widehat{\boldsymbol{X}}_{(r+1)}, \boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(r)}), \tag{2.6}$$

and

$$\left(\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(r+1)}\right) \stackrel{d}{=} \left(\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(r)}, \bigvee \widetilde{\boldsymbol{X}}_{(r+1)}\right). \tag{2.7}$$

In particular, $\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)} \ldots$ is a Markov chain with state space $\mathbb{R}_{-\infty}^{\mathbb{N},\uparrow}$, with its conditional distributions satisfying

$$\left(\boldsymbol{X}^{(r+1)} \middle| \boldsymbol{X}^{(r)}, \ldots, \boldsymbol{X}^{(1)}\right) \stackrel{d}{=} \left(\bigvee \widetilde{\boldsymbol{X}}_{(r+1)} \middle| \boldsymbol{X}^{(r)}\right), \quad r \in \mathbb{N}. \tag{2.8}$$

Proof. Indeed, (2.7) follows from (2.6) because

$$\begin{aligned}
\left(\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(r+1)}\right) &= \left(\boldsymbol{X}^{(1)}, \ldots, \bigvee (\boldsymbol{X}^{(r)}{}_{\mathcal{R}} \wedge \boldsymbol{X}^{(0)})\right) && \text{(Proposition 2.1)},\\
&\stackrel{d}{=} \left(\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(r)}, \bigvee (\boldsymbol{X}^{(r)}{}_{\mathcal{R}} \wedge \widehat{\boldsymbol{X}}_{(r+1)})\right) && \text{(from (2.6))}\\
&= \left(\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(r)}, \bigvee \widetilde{\boldsymbol{X}}_{(r+1)}\right) && \text{(definitions)}.
\end{aligned}$$

In (2.7) $\widetilde{\boldsymbol{X}}_{r+1}$ depends on $\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(r)}$ only through $\boldsymbol{X}^{(r)}$, and this holds for all $r \in \mathbb{N}$. In particular, (2.8) holds, and $\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \ldots$ must be a Markov chain.

It remains to show (2.6). For $r \in \mathbb{N}$ let $\mathbb{R}_{-\infty}^{r,\downarrow} := \{\boldsymbol{m} = (m_1, \ldots, m_r) \in \mathbb{R}_{-\infty}^r : m_1 \geq \ldots \geq m_r\}$ be the space of r-tuples with nonincreasing $\mathbb{R}_{-\infty}$-valued components, and introduce a smooth truncation mapping $\boldsymbol{\mu}_r = (\mu_{r,1}, \ldots, \mu_{r,r}) : \mathbb{R}_{-\infty}^{r,\downarrow} \times \mathbb{R} \mapsto \mathbb{R}_{-\infty}^{r,\downarrow}$, by setting $\mu_{r,1}(m, x) := x \vee m_1$, and, for $2 \leq k \leq r$,

$$\mu_{r,k}(\boldsymbol{m}, x) = m_{k-1} \mathbf{1}_{x > m_{k-1}} + m_k \mathbf{1}_{x \leq m_k} + x \mathbf{1}_{m_k < x \leq m_{k-1}},$$

when $x \in \mathbb{R}$ and $\boldsymbol{m} = (m_1, \ldots, m_r) \in \mathbb{R}_{-\infty}^{r,\downarrow}$. Note that

$$\mu_{r,k}(\boldsymbol{m}, x) \geq m_k \text{ for } \boldsymbol{m} \in \mathbb{R}_{-\infty}^{r,\downarrow},\ x \in \mathbb{R},\ 1 \leq k \leq r. \tag{2.9}$$

Also, define mappings $\widetilde{\boldsymbol{\mu}}_r = (\widetilde{\mu}_{r,0}, \ldots, \widetilde{\mu}_{r,r}) : \mathbb{R}_{-\infty}^{r,\downarrow} \times \mathbb{R} \mapsto \mathbb{R} \times \mathbb{R}_{-\infty}^{r,\downarrow}$ and $\widehat{\boldsymbol{\mu}}_r = (\widehat{\mu}_{r,0}, \ldots, \widehat{\mu}_{r,r}) : \mathbb{R}_{-\infty}^{r,\downarrow} \times \mathbb{R} \times (0,1) \mapsto \mathbb{R} \times \mathbb{R}_{-\infty}^{r,\downarrow}$, by setting

$$\widetilde{\mu}_{r,k}(\boldsymbol{m}, x) := \widehat{\mu}_{r,k}(\boldsymbol{m}, x, u) := \mu_{r,k}(\boldsymbol{m}, x), \quad 1 \leq k \leq r,$$

and with $k = 0$, $\widetilde{\mu}_{r,0}(\boldsymbol{m}, x) := x$ and

$$\begin{aligned}
\widehat{\mu}_{r,0}(\boldsymbol{m}, x, u) = {} & F^{\leftarrow}(u | \mu_{r,r}(\boldsymbol{m}, x)) \prod_{1 \leq k \leq r} \mathbf{1}_{\mu_{r,k}(\boldsymbol{m},x) = m_k}\\
& + \sum_{k=1}^{r} \mu_{r,k}(\boldsymbol{m}, x) \mathbf{1}_{\mu_{r,k}(\boldsymbol{m},x) > m_k} \prod_{1 \leq l < k} \mathbf{1}_{\mu_{r,l}(\boldsymbol{m},x) = m_l}
\end{aligned}$$

$$
\begin{aligned}
&= F^{\leftarrow}(u|m_r)\prod_{1\le k\le r}\mathbf{1}_{\mu_{r,k}(\boldsymbol{m},x)=m_k}+x\sum_{k=1}^{r}\mathbf{1}_{\mu_{r,k}(\boldsymbol{m},x)>m_k}\prod_{1\le l<k}\mathbf{1}_{\mu_{r,l}(\boldsymbol{m},x)=m_l}\\
&=\begin{cases}F^{\leftarrow}(u|m_r), & \text{if } x\le m_r,\\ x, & \text{if } x>m_r.\end{cases}
\end{aligned}
\tag{2.10}
$$

for $\boldsymbol{m}=(m_1,\dots,m_r)\in\mathbb{R}^{r,\downarrow}_{-\infty}$, $x\in\mathbb{R}$ and $u\in(0,1)$.

One can check that the component form of the left and right sides of (2.6) is for $n\ge 2$,

$$\Big((X_n,X_n^{(1)},\dots,X_n^{(r)}),n\ge 2\Big)=\Big(\widetilde{\boldsymbol{\mu}}_r(X_{n-1}^{(1)},\dots,X_{n-1}^{(r)},X_n),n\ge 2\Big)\tag{2.11}$$

$$\Big((\hat{X}_{(r+1),n},X_n^{(1)},\dots,X_n^{(r)}),n\ge 2\Big)=\Big(\widehat{\boldsymbol{\mu}}_r(X_{n-1}^{(1)},\dots,X_{n-1}^{(r)},X_n,U_{r,n}),n\ge 2\Big),\tag{2.12}$$

where $X_n \perp\!\!\!\perp (X_{n-1}^{(1)},\dots,X_{n-1}^{(r)})$ and $U_{r,n}\perp\!\!\!\perp(X_{n-1}^{(1)},\dots,X_{n-1}^{(r)},X_n,)$ since we assumed that $\boldsymbol{X}$ and $\boldsymbol{U}$ are independent arrays of iid rv's. The right sides of (2.11) and (2.12) are Markov chains with stationary transition probabilities in the index n (new value is a function of the previous value and an independent quantity) and for $n=1$, the left sides of (2.11) and (2.12) have common initial value $(X_1,X_1,-\infty,\dots,-\infty)\in\mathbb{R}\times\mathbb{R}^{r,\downarrow}_{-\infty}$. Therefore, to prove equality in distribution in (2.6), it suffices to prove both chains have a common transition kernel.

To see this, let $X'\stackrel{d}{=}X_1\sim F$ and $U'\stackrel{d}{=}U_{1,1}\in(0,1)$ be independent rv's. For $x,y\in\mathbb{R}$ with $F(y)>0$ note

$$
\begin{aligned}
P(X'\le y,F^{\leftarrow}(U'|y)\le x)&=P(X'\le y)P(X'\le x|X'\le y)\\
&=F(y)F(x|y)=F(x\wedge y),
\end{aligned}
\tag{2.13}
$$

Consequently, for $\boldsymbol{m}=(m_1,\dots,m_r)$, $\boldsymbol{m}'=(m_1',\dots,m_r')\in\mathbb{R}^{r,\downarrow}_{-\infty}$ with $F(m_k')>0$ for $1\le k\le r$, setting $m_0':=\infty$, we have for the transition probability,

$$
\begin{aligned}
&P\Big((\widehat{X}_{(r+1),n+1},X_{n+1}^{(1)},\dots,X_{n+1}^{(r)})\in(-\infty,x]\times\prod_{k=1}^{r}[-\infty,m_k]\\
&\qquad\qquad\Big|\widehat{X}_{(r+1),n}=y,(X_n^{(1)},\dots,X_n^{(r)})=\boldsymbol{m}'\Big)\\
&=P\Big(\widehat{\boldsymbol{\mu}}_r(\boldsymbol{m}',X',U')\in(-\infty,x]\times\prod_{k=1}^{r}[-\infty,m_k]\Big)\\
&=P\Big(\widehat{\boldsymbol{\mu}}_r(\boldsymbol{m}',X',U')\in(-\infty,x]\times\prod_{k=1}^{r}[-\infty,m_k],X'\le m_r'\Big)\\
&\quad+\sum_{k=1}^{r}P\Big(\widehat{\boldsymbol{\mu}}_r(\boldsymbol{m}',X',U')\in(-\infty,x]\times\prod_{k=1}^{r}[-\infty,m_k],X'\in(m_k',m_{k-1}']\Big)\\
&=A+B.
\end{aligned}
$$

Consider

$$A=P\Big(F^{\leftarrow}(U'|m_r')\le x,\mu_{rk}(\boldsymbol{m}',X',U')\le m_k,k=1,\dots,r;X'\le m_r'\Big).$$

If $m'_k > m_k$ for some $k = 1, \dots, r$, then because of (2.9), the probability A is 0. So assume for $k = 1, \dots, r$, that $m'_k \le m_k$. Then the condition $X' \le m'_r$ in A implies $X' \le m'_k \le m_k$ for $k = 1, \dots, r$ and using (2.13), A reduces to

$$A = P(F^{\leftarrow}(U'|m'_r) \le x, X' \le m'_r) = F(x \wedge m'_r) \prod_{1 \le k \le r} \mathbf{1}_{m'_k \le m_k}.$$

For B we use (2.10) and get

$$B = \sum_{k=1}^{r} P(X' \in (m'_k, m'_{k-1}], X' \le x, \underbrace{\mu_{rl}(\boldsymbol{m}', X')}_{\ge m'_l} \le m_l; l = 1, \dots, r)$$

Fix k and suppose $l > k$. Then the interval $(m'_l, m'_{l-1}]$ is to the left of $(m'_k, m'_{k-1}]$ where X' is located and $\mu_{rl}(\boldsymbol{m}', X') = m'_{l-1}$. The probability is then 0 unless $m_l \ge m'_{l-1}$. If $l < k$, the order of the intervals is reversed, $\mu_{rl}(\boldsymbol{m}', X') = m'_l$, and the probability is 0 unless $m'_l \le m_l$. Thus, B becomes

$$B = \sum_{k=1}^{r} P(m'_k < X' \le x \wedge m_k \wedge m'_{k-1}) \prod_{1 \le l < k} \mathbf{1}_{m'_l \le m_l} \prod_{k < l \le r} \mathbf{1}_{m'_{l-1} \le m_l}.$$

On the other hand, from the left sides of (2.6) and (2.11),

$$\begin{aligned}
P\Big(&(X_{n+1}, X^{(1)}_{n+1}, \dots, X^{(r)}_{n+1}) \in (-\infty, x] \times \prod_{k=1}^{r} [-\infty, m_k] \\
&\qquad \Big| X_n = y, (X_n, X^{(1)}_n, \dots, X^{(r)}_n) = \boldsymbol{m}'\Big) \\
&= P(\widetilde{\boldsymbol{\mu}}_r(\boldsymbol{m}', X') \in (-\infty, x] \times \prod_{k=1}^{r} [-\infty, m_k]) \\
&= P\Big((X', \mu_{rl}(\boldsymbol{m}', X'), l = 1, \dots, r) \in (-\infty, x] \times \prod_{k=1}^{r} [-\infty, m_k]\Big) \\
&= P\Big(X' \le x, X' \le m'_r, \mu_{rl}(\boldsymbol{m}', X') \le m_l, l = 1, \dots, r) \\
&\qquad + \sum_{k=1}^{r} P(X' \le x, X' \in (m_k, m_{k-1}], \mu_{rl}(\boldsymbol{m}', X') \le m_l, l = 1, \dots, r) \\
&= A + B.
\end{aligned}$$

This completes the proof of (2.6) and of Theorem 2.1. $\square$

3. Asymptotic Behaviour of the Discrete Time Process $\boldsymbol{M}^{(r)}$ for large r

In this section we consider asymptotic behaviour as $r \to \infty$ of $\{\boldsymbol{M}^{(r)} := (M^{(r)}_n, n \ge r), r \ge 1\}$ as an $\mathbb{R}^\infty$-valued stochastic process. As r increases we are pushing into values far from the largest, so limit behaviour for both the range of $\boldsymbol{M}^{(r)}$ and $\boldsymbol{M}^{(r)}$ itself, depend critically on left tail behavior of the distribution of X_1. Appropriate left-tail conditions related to domain of attraction conditions in classical extreme value theory make the range and the sequence of rth order maxima converge weakly.

Throughout this section we will assume F is continuous, so the records of $\{X_n\}$ are Poisson with mean measure R [21, page 166] which we denote PRM(R). The assumption of continuity could be relaxed as in [9, 24, 25] but results are most striking and elegant when F is continuous and we proceed in this setting.

3.1. **rth maximum and r-records.** Let $\{X_n, n \geq 1\}$ be iid random variables with common distribution function $F(x)$. Assume $F(x) < 1$ and set $R(x) = -\log(1 - F(x)) = -\log \bar{F}(x)$. Define

$$\begin{aligned} R_n =& \sum_{j=1}^{n} 1_{[X_j \geq X_n]} = \text{relative rank of } X_n \text{ among } X_1, \dots, X_n \\ =& \text{rank of } X_n \text{ at ``birth''}. \end{aligned}$$

It is known [15] that $\{R_n\}$ are independent random variables and R_n is uniformly distributed on $\{1, \dots, n\}$; that is,

$$P[R_n = i] = 1/n, \quad i = 1, \dots, n.$$

Considering $\{\boldsymbol{M}^{(r)}, r \geq 1\}$ as an $\mathbb{R}^\infty$-valued stochastic process, we ask what is the asymptotic behavior of $\boldsymbol{M}^{(r)}$ and its range as a function of r as $r \to \infty$?

Define the r-record times of $\{X_n\}$ by

$$L_0^{(r)} = 0, \quad L_{n+1}^{(r)} = \inf\{j > L_n^{(r)} : R_j = r\}.$$

The r-records are then $\{X_{L_n^{(r)}}, n \geq 1\}$, which for each r, are points of PRM($R(dx)$) by Ignatov's theorem.

We list some initial facts about $\boldsymbol{M}^{(r)}$ and its range.

- For fixed r, $\boldsymbol{M}^{(r)} = \{M_n^{(r)}, n \geq r\}$ jumps at index $k \geq r$ iff
$$R_k \in \{1, \dots, r\},$$
so
$$\{[\boldsymbol{M}^{(r)} \text{ jumps at index } k], k \geq r\}$$
are independent events over k and
$$P[\boldsymbol{M}^{(r)} \text{ jumps at } k] = \frac{r}{k}.$$

Remark 3.1. This has the implication that if we re-index and set $k = r + l$ for $l \geq 0$, then for any fixed l,

$$P[\boldsymbol{M}^{(r)} \text{ jumps at } r + l] = \frac{r}{r + l} \to 1, \quad (r \to \infty).$$

So for large r, $\boldsymbol{M}^{(r)}$ jumps at almost every integer. Define the jump indices

$$\{\tau_l^{(r)}, l \geq 0\} = \{j \geq 1 : M_{r+j}^{(r)} > M_{r+j-1}^{(r)}\} \cup \{0\}.$$

Then in $\mathbb{R}_+^\infty$,

$$\{\tau_l^{(r)}, l \geq 0\} \Rightarrow \{0, 1, 2, \dots\}.$$

- For fixed r, let $\mathcal{R}_r$ be the range of $\boldsymbol{M}^{(r)}$; that is, the distinct points without repetition hit by $\{M_n^{(r)}, n \geq r\}$. Then,

$$\mathcal{R}_r := \bigcup_{p=1}^{r} \{X_{L_n^{(p)}}, n \geq 1\}, \tag{3.1}$$

By Ignatov's theorem [9, 10, 12, 21, 26], this is a sum of r independent PRM(R) processes and therefore the range of $\boldsymbol{M}^{(r)}$ is PRM(rR).

To prove (3.1), suppose $M_n^{(r)} = x$, for some $n \geq r$. Suppose the rth largest of $X_1, \ldots, X_n$ occurs at $X_i = x$ for $i \leq n$. If the rank of X_i were $> r$, it could not be the case that $M_n^{(r)} = x$. This shows that

$$\text{range of } \boldsymbol{M}^{(r)} \subset \bigcup_{p=1}^{r} \{X_{L_n^{(p)}}, n \geq 1\}.$$

Conversely, suppose $X_{L_n^{(p)}} = x$, so at time $L_n^{(p)}$, the rank of $X_{L_n^{(p)}}$ is p. Wait until $r-p$ additional X's have been observed that exceed x and then the rth largest will equal x.

3.2. **Limits for the range $\mathcal{R}_r$ of $\boldsymbol{M}^{(r)}$.** Although our primary interest is in the behavior of $\{\boldsymbol{M}^{(r)}, r \geq 1\}$ as an $\mathbb{R}^\infty$-valued random sequence, it is instructive and helpful to discuss the behavior of the range $\mathcal{R}_r$ of $\boldsymbol{M}^{(r)}$.

As a basic result we derive a deterministic limit for $\mathcal{R}_r$. Let $\mathcal{R}$ be the support of the measure $R(\cdot)$ which corresponds to the monotone function $R(x) = -\log(1-F(x))$. Note $\mathcal{R}$ is also the support of F.

Proposition 3.2. *As $r \to \infty$, $\mathcal{R}_r$, the range of $\boldsymbol{M}^{(r)}$, converges as a random closed set in the Fell topology [13, 14, 27] to the non-random limit $\mathcal{R}$:*

$$\mathcal{R}_r \Rightarrow \mathcal{R}. \tag{3.2}$$

Proof. Since $\mathcal{R}_r \subset \mathcal{R}$, it suffices to show for any open G with $\mathcal{R} \cap G \neq \emptyset$, that

$$P[\mathcal{R}_r \cap G \neq \emptyset] \to 1.$$

However, $\mathcal{R} \cap G \neq \emptyset$ implies $R(G) > 0$ and therefore,

$$\begin{aligned} P[\mathcal{R}_r \cap G \neq \emptyset] =& 1 - P[\text{PRM}(rR(G)) = 0] \\ =& 1 - e^{-rR(G)} \to 1, \quad (r \to \infty) \end{aligned}$$

since $R(G) > 0$. □

The set convergence in (3.2) is to a deterministic limit. Since $\mathcal{R}_r$ is a PRM(rR) point process, we can get a random limit if we center and scale the $\{X_n\}$ so that the mean measure rR converges to a Radon measure. Recall $R(x) = -\log \bar{F}(x)$.

Assume there exist $a_r > 0$ and $b_r \in \mathbb{R}$ and a non-decreasing limit function $g(x)$ with more than one point of increase such that

$$rR(a_r x - b_r) \to g(x), \qquad (r \to \infty). \tag{3.3}$$

For x such that $g(x) > 0$, to counteract $r \to \infty$, we must have $R(a_r x - b_r) \to 0$ and $a_r x - b_r$ converging to the left endpoint of F (and R).

We now explain why e^{-g} is related to an extreme value distribution. Remembering that $e^{-R} = \bar{F}$, equation (3.3) is equivalent to

$$(\bar{F}(a_r x - b_r))^r = \exp\{-rR(a_r x - b_r)\} \to e^{-g(x)}$$

or

$$P\Big[\frac{\bigwedge_{i=1}^r X_i + b_r}{a_r} > x\Big] \to e^{-g(x)}. \tag{3.4}$$

So we recognize e^{-g} as the survivor function of an extreme value distribution of minima of iid random variables. Expressing this in terms of maxima by setting $Y_i = -X_i$ we get (3.4) equivalent to

$$P\Big[\frac{\bigvee_{i=1}^r Y_i - b_r}{a_r} \le -x\Big] \to e^{-g(x)} = G_\gamma(-x), \tag{3.5}$$

for some $\gamma \in \mathbb{R}$, where $G_\gamma(x) = \exp\{-(1+\gamma x)^{-1/\gamma}\}$, $1+\gamma x > 0$ is the shape parameter family of extreme value distributions for maxima [3, 21]. So in (3.3), $g(x) = g_\gamma(x) = -\log G_\gamma(-x)$. The usual way to write (3.5) is

$$rP[Y_1 > a_r(-x) + b_r] \to g(x), \qquad \forall x \text{ s.t. } g(x) > 0,$$

and (3.3) is the same as

$$rF(a_r x - b_r) \to g(x), \qquad \forall x \text{ s.t. } g(x) > 0. \tag{3.6}$$

In particular, apart from centering, we have the cases:

(1) *Gumbel case*: $\gamma = 0$. Then

$$g_0(x) = e^x, \quad x \in \mathbb{R}.$$

(2) *Reverse Weibull case*: $\gamma < 0$: Then $1 + \gamma(-x) > 0$ iff $x > -1/|\gamma|$ and

$$g_\gamma(x) = (1 + |\gamma| x)^{1/|\gamma|}, \quad x > -1/|\gamma|.$$

Adjusting the centering and scaling by taking $b_r = 0$, we find R is regularly varying at 0 and

$$rR(a_r x) \to x^{1/|\gamma|}, \quad x > 0.$$

(3) *Frechét case*: $\gamma > 0$. Then $1 + \gamma(-x) > 0$ iff $x < 1/\gamma$ and

$$g_\gamma(x) = (1 - \gamma x)^{-1/\gamma}, \quad x < 1/\gamma.$$

Adjusting the centering and scaling so the support is $(-\infty, 0)$ we get

$$rR(a_r x) \to |x|^{-1/\gamma}, \quad x < 0,$$

which is regular variation at 0 from the left.

We can apply this analysis to get convergence of $\mathcal{R}_r$ after centering and scaling. Recall $\mathcal{R}_r$ is PRM(rR). A family of Poisson point measures converges weakly iff the mean measures converge (eg. [20]). So replacing

$$X_i \mapsto \frac{X_i + b_r}{a_r}$$

rescales the points of the range to be Poisson with mean measure given by the left side of (3.3). Let

$$\text{supp}_\gamma = \{x : 1 - \gamma x > 0\} \tag{3.7}$$

and $m_\gamma(\cdot)$ be the measure with density $g'_\gamma(x)$, $x \in \text{supp}_\gamma$. Let $M_+(\text{supp}_\gamma)$ be the space of Radon measures on supp_γ, topologized by vague convergence. Then (3.3) implies the vague convergence

$$rR\big(a_r(\cdot) - b_r\big) \xrightarrow{v} m_\gamma(\cdot)$$

in $M_+(\text{supp}_\gamma)$, and thus on $M_+(\text{supp}_\gamma)$ we have

$$(\mathcal{R}_r + b_r)/a_r \Rightarrow PRM(m_\gamma). \tag{3.8}$$

We may realize PRM(m_γ) as follows: Let $\Gamma_i = \sum_{j=1}^{i} E_j$ be a sum of iid standard exponential random variables. The $\{\Gamma_i\}$ are points of a homogeneous Poisson process rate 1 on $[0, \infty)$. The measure m_γ has distribution

$$g_\gamma : \text{supp}_\gamma \mapsto (0, \infty),$$

with inverse

$$g_\gamma^{\leftarrow} : (0, \infty) \mapsto \text{supp}_\gamma.$$

The transformation theory for Poisson processes (eg. [20, Section 5.1]) means $\sum_{i=1}^{\infty} \epsilon_{g_\gamma^{\leftarrow}(\Gamma_i)}$ is PRM(m_γ) on supp_γ. For instance, if $\gamma = 0$, $\text{supp}_0 = \mathbb{R}$, $g_0(x) = e^x$, $x \in \mathbb{R}$, and $g_0^{\leftarrow}(y) = \log y$, $y > 0$, and PRM(m_0) $= \sum_i \epsilon_{\log \Gamma_i}$.

3.3. **Weak convergence of the rth maxima sequence $\boldsymbol{M}^{(r)}$.** Having understood how to get the range $\mathcal{R}_r$ of $\boldsymbol{M}^{(r)}$ to converge, we turn to convergence of $\boldsymbol{M}^{(r)}$ itself. We continue to suppose the minimum domain of attraction condition, so that R satisfies (3.3), and recall $M_+(\text{supp}_\gamma)$ is the space of Radon measures on supp_γ, topologized by vague convergence. Point measures in $M_+(\text{supp}_\gamma)$ are denoted by $\sum_i \epsilon_{x_i}(\cdot)$ where $\epsilon_x(\cdot)$ is the Dirac measure placing mass 1 at x.

We start with a preliminary result on the empirical measures generated by $\{X_i\}$ that will be needed to study the weak convergence of $\{\boldsymbol{M}^{(r)}\}$.

Proposition 3.3. *Assume* (3.3). *If N is a random element of $M_+(\text{supp}_\gamma)$ which is $PRM(m_\gamma)$, then for any $j \geq 0$,*

$$\sum_{i=1}^{r+j} \epsilon_{(X_i + b_r)/a_r} \Rightarrow N = \sum_{i=1}^{\infty} \epsilon_{g_\gamma^{\leftarrow}(\Gamma_i)} = PRM(m_\gamma), \tag{3.9}$$

in $M_+(supp_\gamma)$ and, in fact, jointly for any $k \geq 0$,

$$\Big(\sum_{i=1}^{r+j} \epsilon_{(X_i+b_r)/a_r}; 0 \leq j \leq k\Big) \Rightarrow (N, \dots, N) \tag{3.10}$$

in $M_+(supp_\gamma) \times \cdots \times M_+(supp_\gamma)$.

Proof. We have (3.10) following from (3.9) since with respect to the vague distance $d(\cdot,\cdot)$ on $M_+(\mathrm{supp}_\gamma)$ (see, eg. [20, page 51])

$$d\Big(\sum_{i=1}^{r} \epsilon_{(X_i+b_r)/a_r}, \sum_{i=1}^{r+j} \epsilon_{(X_i+b_r)/a_r}\Big) \Rightarrow 0$$

for any $j \geq 0$. To verify this, let f be positive and continuous with compact support on supp_γ and from equation 3.14 of [20, page 51], it suffices to show

$$E\Big|\sum_{i=1}^{r} f\big((X_i+b_r)/a_r\big) - \sum_{i=1}^{r+j} f\big((X_i+b_r)/a_r\big)\Big| \to 0.$$

The difference is

$$E\sum_{i=r+1}^{r+j} f\big((X_i+b_r)/a_r\big) = E\sum_{i=1}^{j} f\big((X_i+b_r)/a_r\big)$$

and assuming the support of f is a compact set K in supp_γ, this is bounded above by

$$\sup_{x\geq 0} f(x) j P[X_1 \in a_r K - b_r] \to 0,$$

since for $x \in K$, $a_r x - b_r$ converges to the left endpoint of F and under (3.3), there cannot be an atom at this left endpoint.

The result in (3.9) follows by a small modification of the proof of Theorem 5.3 in [20, page 138] since (3.3) is the same as (3.6). □

Now we turn to $\mathbb{R}^\infty$-convergence of the rth maximum sequence. Continue to suppose (3.3). Without normalization, the sequence $\boldsymbol{M}^{(r)}$ converges to a sequence all of whose entries are the left endpoint of F. In order to get $\boldsymbol{M}^{(r)}$ to converge, we must have $M_r^{(r)} = \wedge_{i=1}^r X_i$ converge and this helps explain why a domain of attraction condition for minima is relevant. The condition (3.3) produces a non-trivial limit.

Proposition 3.4. *Suppose the domain of attraction condition* (3.3) *holds. Then in $\mathbb{R}^\infty$,*

$$\frac{\boldsymbol{M}^{(r)} + b_r}{a_r} = \Big(\frac{M_{r+j}^{(r)} + b_r}{a_r}, j \geq 0\Big) \Rightarrow \Big(g_\gamma^{\leftarrow}(\Gamma_l), l \geq 1\Big) \qquad (r \to \infty), \tag{3.11}$$

where $\{\Gamma_l, l \geq 1\}$ are the points of a homogeneous Poisson process on $\mathbb{R}_+$.

Proof. Fix $j \geq 0$ and observe for $x \in \mathrm{supp}_\gamma$,

$$\Big[\frac{M_{r+j}^{(r)} + b_r}{a_r} > x\Big] = \Big[\sum_{i=1}^{r+j} \epsilon_{(X_i+b_r)/a_r}(x,\infty) \geq r\Big] = \Big[\sum_{i=1}^{r+j} \epsilon_{(X_i+b_r)/a_r}((-\infty, x]) \leq j\Big]$$

and therefore

$$\Big[\frac{M^{(r)}_{r+j}+b_r}{a_r}\leq x\Big] = \Big[\sum_{i=1}^{r+j}\epsilon_{(X_i+b_r)/a_r}((-\infty,x])>j\Big].$$

For a non-decreasing sequence $\{x_j\}$ of real numbers in supp_γ,

$$P\Big\{\bigcap_{j=0}^{k}\Big[\frac{M^{(r)}_{r+j}+b_r}{a_r}\leq x_j\Big]\Big\} = P\Big\{\bigcap_{j=0}^{k}\Big[\sum_{i=1}^{r+j}\epsilon_{(X_i+b_r)/a_r}([0,x_j])>j\Big]\Big\}$$

and applying (3.10) yields

$$\begin{aligned}\to& P\{\bigcap_{j=0}^{k}[N((-\infty,x_j])>j]\} = P[\sum_{i=1}^{\infty}\epsilon_{g^{\leftarrow}_\gamma(\Gamma_i)}(-\infty,x_j]>j;\ j=0,\dots,k]\\ =&P[g^{\leftarrow}_\gamma(\Gamma_{j+1})\leq x_j; j=0,\dots,k].\end{aligned}$$

This yields the announced result (3.11). □

4. Continuous time rth-order extremal processes

This section transitions to continuous time problems. The treatment is parallel to what we gave for discretely indexed processes but here the processes are generated by two-dimensional Poisson processes on $\mathbb{R}_+\times\mathbb{R}$ and correspond to rth order extremal processes. One example of an rth order extremal process is obtained by taking the rth largest jump of a Lévy process up to time $t>0$.

The continuous time case differs from the discrete index case, in that there are always infinitely many values *below* your present position. This necessitates differences in treatment. In continuous time we obtain modifications of Brownian motion limits whereas in discrete time we obtain Poisson limits for the rth order extremes.

The setup is as follows. For some numbers $-\infty\leq x_l<x_r\leq\infty$, and an infinite measure Π on (x_l,x_r) satisfying $\Pi(x_l,x_r)=\infty$ and $Q(x):=\Pi(x,x_r)<\infty$ for $x_l<x<x_r$, let

$$N=\sum_k \epsilon_{(t_k,j_k)}, \tag{4.1}$$

be Poisson random measure on $[0,\infty)\times(x_l,x_r)$, with mean measure $Leb\times\Pi$. The notation $\epsilon_{(t,x)}(\cdot)$ denotes a Dirac measure with mass 1 at the point (t,x). Sometimes we write $(t_k,j_k)\in\mathrm{supp}(N)$ to indicate the point (t_k,j_k) is charged by N. We assume x_l and x_r are not atoms of Π and in fact, to make results most elegant we assume $\Pi(\cdot)$ is atomless. (Otherwise, results would be stated in terms of simplifications of point processes; see [9].) Our assumptions mean that

(1) The function $Q(x)$ satisfies $Q(x_r)=0$ and $Q(x_l)=\infty$ so $Q:(x_l,x_r)\mapsto(0,\infty)$ and $Q(x)$ is non-increasing.
(2) For any $t>0$ and $x_r\geq x>x_l: N\big([0,t]\times(x,x_r)\big)<\infty$ almost surely.
(3) For any $t>0$ and $x_r\geq x>x_l: N\big([0,t]\times(x_l,x]\big)=\infty$ almost surely.

Traditionally, the (first-order) extremal process is defined by ([4–8, 18, 19, 21, 22, 25, 28]),

$$Y(t) = Y^{(1)}(t) = \bigvee_{t_k \le t} j_k, \quad 0 < t < \infty,$$

the largest j_k whose t_k coordinate is at or before time t. Alternatively we may write

$$Y(t) = \inf\{x > x_l : N\big([0,t] \times (x, x_r)\big) = 0\} = \inf\{x > x_l : N\big([0,t] \times (x, x_r)\big) < 1\}.$$

We develop the analogues of Propositions 3.2 and 3.4 as $r \to \infty$ for the continuous time rth order extremal process $\boldsymbol{Y}^{(r)} := \{Y^{(r)}(t), 0 < t < \infty\}$ defined as,

$$Y^{(r)}(t) := \inf\{x > x_l : N\big([0,t] \times (x, x_r)\big) < r\}, \quad t > 0. \tag{4.2}$$

This means for $t > 0$, $x_r \ge x > x_l$,

$$[Y^{(r)}(t) > x] = [N\big([0,t] \times (x, x_r)\big) \ge r],$$

and therefore,

$$[Y^{(r)}(t) \le x] = [N\big([0,t] \times (x, x_r)\big) < r]. \tag{4.3}$$

Alternative ways of considering $\boldsymbol{Y}^{(r)}$ are in [9].

What is the behavior of $\{\boldsymbol{Y}^{(r)}, r \ge 1\}$, considered as a sequence of random elements of càdlàg space $D(x_l, x_r)$, as $r \to \infty$? Unlike in Section 3.3, here there are always infinitely many points below your current position and thus the left tail condition (3.6) used for $\boldsymbol{M}^{(r)}$ must be different when considering $\boldsymbol{Y}^{(r)}$. We analyze of the range of $\boldsymbol{Y}^{(r)}$ and for the weak limit behavior of $\boldsymbol{Y}^{(r)}$, instead of relying on Poisson behavior, we rely on asymptotic normality.

4.1. **The range $\mathcal{R}_r$ of $\boldsymbol{Y}^{(r)}$.** Let $\mathcal{R}_r$ be the unique values in the set $\{Y^{(r)}(t), t > 0\}$. As in the discrete time case (3.1), we have

$$\mathcal{R}_r = \bigcup_{p=1}^{r} \{j_k : (t_k, j_k) \in \operatorname{supp}(N),\ N([0,t_k] \times [j_k, \infty)) = p\}. \tag{4.4}$$

To verify (4.4) suppose $x \in \mathcal{R}_r$. There exists $t > 0$ such that $Y^{(r)}(t) = x$, and therefore there exists $(t_k, x) \in \operatorname{supp}(N)$ such that $t_k \le t$. If $N([0,t_k] \times [x, \infty)) > r$, then $Y^{(r)}(t) > x$, giving a contradiction. Thus x is in the right side of (4.4). Conversely, suppose j_k satisfies that there exists t_k such that $(t_k, j_k) \in \operatorname{supp}(N)$ and $N([0,t_k] \times [j_k, \infty)) = p$ for some $p \le r$. Then there exists $t > t_k$ such that $N(t_k, t]] \times [j_k, \infty)) = r - p$ and thus $Y^{(r)}(t) = j_k$. Therefore, j_k belongs to the left side of (4.4). □

When Π is atomless, the range of $Y(t) = Y^{(1)}(t)$ is known to be a Poisson process with mean measure determined by the monotone function $S(x) := -\log \Pi(x, \infty), x > x_l$. This is discussed, for example, in [21, page 183]. In fact, from [9, Theorem 6.2, page 234], the p-records of N are iid in p, and each sequence of p-records forms PRM(S). (A p-record of N is a point j_k such that there exists t_k making $(t_k, j_k) \in \operatorname{supp}(N)$ and $N([0,t_k] \times [j_k, \infty)) = p$.) This and (4.4) allow us to conclude that $\mathcal{R}_r$ is a Poisson process with mean measure $rS(\cdot)$.

This achieves the continuous time analogue of the discrete time discussion at the beginning of Subsection 3.2, and without any normalization we have

$$\mathcal{R}_r \Rightarrow \operatorname{supp}(S), \quad (r \to \infty),$$

in the Fell topology of closed subsets of (x_l, x_r).

Paralleling the discrete time analysis, we proceed to obtain a non-degenerate limit for $\mathcal{R}_r$. We have to be more careful in the continuous case. The reason is that $\mathcal{R}_r$ is PRM with mean measure $rS(\cdot)$ and S is Radon on (x_l, x_r), and it may allocate infinite mass to a neighbourhood of both x_l and x_r. Recall $S(x) := -\log \Pi(x, x_l]$ satisfies $S(x_l) = -\infty$ and $S(x_r) = \infty$.

Assume without loss of generality that $x_l < 0 < x_r$. (If this is not the case, choose an arbitrary point between x_l and x_r.) We make a treatment parallel to the discrete one by splitting the Poisson points of $\mathcal{R}_r$ into those above 0 and those below. So write

$$\mathcal{R}_r = \mathcal{R}_r^+ \bigcup \mathcal{R}_r^-$$

where $\mathcal{R}_r^+$ are the positive Poisson points of $\mathcal{R}_r$ and $\mathcal{R}_r^-$ are the negative points of $\mathcal{R}_r$. The two Poisson processes $\mathcal{R}_r^\pm$ are independent because their points are in disjoint regions. Define the two non-decreasing functions on $\mathbb{R}_+$,

$$S^+(x) = S(0, x] = S(x) - S(0), \qquad 0 < x \le x_r \tag{4.5}$$

$$S^-(x) = S[-x, 0) = S(0) - S(-x), \qquad 0 < x \le -x_l. \tag{4.6}$$

Assume there exist $a^\pm(t) > 0$, $b^\pm(t) \in \mathbb{R}$ and infinite Radon measures $S_\infty^\pm$ on $\mathbb{R}_+$ such that as $t \to \infty$,

$$tS^+(a^+(t)x - b^+(t)) \to S_\infty^+(x), \tag{4.7}$$

$$tS^-(a^-(t)x - b^-(t)) \to S_\infty^-(x). \tag{4.8}$$

The form of $S_\infty^\pm$ is determined by defining probability distribution tails $\bar{H}^\pm(x)$ by

$$\bar{H}^+(x) = e^{-S^+(x)}, \quad 0 < x < x_r, \tag{4.9}$$

$$\bar{H}^-(x) = e^{-S^-(x)}, \quad 0 < x < -x_l. \tag{4.10}$$

Note $\bar{H}^\pm(0) = e^{-S^\pm(0)} = e^{-0} = 1$ and $\bar{H}^+(x_r) = e^{-S^+(x_r)} = e^{-\infty} = 0$ and $\bar{H}^-(-x_l) = 0$, similarly. Then, as in the discussion following (3.3), we find for $\gamma^\pm \in \mathbb{R}$ that

$$e^{-S^\pm(x)} = G_{\gamma^\pm}(-x),$$

where $G_\gamma(x)$ has a form given after (3.5). Note, if we want

$$a^+(t) = a^-(t) \quad \text{and} \quad b^+(t) = b^-(t)$$

up to convergence of types, we would need [16], $-x_l = x_r$ and

$$\bar{H}^+(x) \sim \bar{H}^-(x) \quad (x \to x_r).$$

We now summarize.

Theorem 4.1. *The two Poisson processes $\mathcal{R}_r^{\pm}$ are independent with $\mathcal{R}_r = \mathcal{R}_r^{+} \cup \mathcal{R}_r^{-}$ where $\mathcal{R}_r^{+}$ has mean measure rS^{+} on $\mathbb{R}_+$ and $-\mathcal{R}_r^{-}$ has mean measure rS^{-} on $\mathbb{R}_+$ so that $\mathcal{R}_r^{-}$ are points on $(-\infty, 0)$. As $r \to \infty$, the range centered and scaled converges to a limiting Poisson process,*

$$\Big(\frac{\mathcal{R}_r^{+} + b^{+}(r)}{a^{+}(r)}, \frac{-\mathcal{R}_r^{+} + b^{-}(r)}{a^{-}(r)}\Big) \Rightarrow \Big(\mathcal{R}_\infty^{+}, -\mathcal{R}_\infty^{-}\Big),$$

where the limits are independent Poisson processes on $\mathbb{R}_+$ with mean measures $S_\infty^{\pm}$. So if (4.7) and (4.8) hold, centering positive and negative range points appropriately leads to a limiting Poisson process such that positive points have mean measure $S_\infty^{+}(\cdot)$ and negative range points made positive by taking absolute values have mean measure $S_\infty^{-}(\cdot)$.

4.2. **Finite dimensional convergence of $\boldsymbol{Y}^{(r)}$ as random elements of** $D(x_l, x_r)$**.** In this subsection, we give a left-tail condition on $\Pi(\cdot)$ guaranteeing finite dimensional convergence of $\boldsymbol{Y}^{(r)}$ to a transformed Brownian motion.

Suppose there exist normalizing functions $a(r) > 0$, $b(r) \in \mathbb{R}$, and a non-decreasing limit function $h(x) \in \mathbb{R}$ with at least two points of increase such that for $a(r)x + b(r) \in (x_l, x_r)$,

$$\lim_{r\to\infty} \frac{r - Q\big(a(r)x + b(r)\big)}{\sqrt{r}} = h(x). \tag{4.11}$$

Implications:

(1) If we divide in (4.11) by r instead of $\sqrt{r}$, the limit will be 0 and therefore,

$$Q\big(a(r)x + b(r)\big) \sim r, \quad (r \to \infty). \tag{4.12}$$

Therefore, since $r \to \infty$, we must have that $Q\big(a(r)x + b(r)\big) \to \infty$ and $(x_l, x_r) \ni a(r)x + b(r) \to x_l$.

(2) For any $t > 0$,

$$\begin{aligned} \frac{r - tQ\big(a(r/t)x + b(r/t)\big)}{\sqrt{r}} =& t\Big(\frac{r/t - Q\big(a(r/t)x + b(r/t)\big)}{\sqrt{r/t}\sqrt{t}}\Big) \\ \to& \sqrt{t}h(x), \quad (r \to \infty). \end{aligned} \tag{4.13}$$

(3) If we write $r - Q = (\sqrt{r} - \sqrt{Q})(\sqrt{r} + \sqrt{Q})$ and use (4.12), we get

$$\sqrt{r} - \sqrt{Q(a(r)x + b(r))} \to \frac{1}{2}h(x). \tag{4.14}$$

Remember that Q is decreasing and define a probability distribution function $G(x)$ by $G(x) := \exp\{-\sqrt{Q(x)}\}$ so that G concentrates on (x_l, x_r). Then exponentiate in (4.14) to get

$$e^{\sqrt{r}} e^{-\sqrt{Q(a(r)x+b(r))}} \to e^{\frac{1}{2}h(x)}, \quad (r \to \infty)$$

or after a change of variables $s = e^{\sqrt{r}}$,

$$sG\big(a((\log s)^2)x + b((\log s)^2)\big) = se^{-\sqrt{Q(a((\log s)^2)x + b((\log s)^2))}} \to e^{\frac{1}{2}h(x)}, \tag{4.15}$$

as $s \to \infty$. So we conclude that $G(x) := e^{-\sqrt{Q(x)}}$ is in a domain of attraction of an extreme value distribution for minima. This technique is essentially the same as the one used to study limit laws for record values in [17] or [21].

(4) Form of $h(x)$: As we saw following (3.6), if $\exp\{\frac{1}{2}h(x)\}$ plays the role of $g(x)$ then $h(x)$ must be of the form

$$e^{\frac{1}{2}h(x)} = -\log G_\gamma(-x),$$

where G_γ is an extreme value distribution for maxima of the form

$$G_\gamma(x) = \exp\{-(1+\gamma x)^{-1/\gamma}\}, \quad \gamma \in \mathbb{R},\, 1+\gamma x > 0.$$

So

$$\frac{1}{2}h(x) = \begin{cases} -\frac{1}{\gamma}\log(1-\gamma x), & \text{if } \gamma \neq 0,\, 1-\gamma x > 0, \\ x, & \text{if } \gamma = 0,\, x \in \mathbb{R}. \end{cases} \tag{4.16}$$

Observe that $h : \text{supp}_\gamma \mapsto \mathbb{R}$ and $h^{\leftarrow} : \mathbb{R} \mapsto \text{supp}_\gamma$. Recalling the definition of supp_γ from (3.7), we have

$$\text{supp}_\gamma = \{x \in \mathbb{R} : 1-\gamma x > 0\} = \begin{cases} (-\frac{1}{|\gamma|}, \infty), & \text{if } \gamma < 0, \\ (-\infty, \frac{1}{|\gamma|}, & \text{if } \gamma > 0, \\ \mathbb{R}, & \text{if } \gamma = 0. \end{cases}$$

We apply these findings to obtain a marginal limit distribution for $Y^{(r)}(t)$ under the left tail condition. Assume (4.11). We show that, for fixed t, $Y^{(r)}(t)$ has a limit distribution as $r \to \infty$, after centering and norming. This relies on an elementary fact: if $\{N_n\}$ is a family of Poisson random variables with $E(N_n) \to \infty$ then

$$\frac{N_n - E(N_n)}{\sqrt{\text{Var}(N_n)}} \Rightarrow N(0,1), \quad (n \to \infty). \tag{4.17}$$

From (4.3), we have

$$\begin{aligned} P\Big[\frac{Y^{(r)}(t) - b(r/t)}{a(r/t)} \le x\Big] &= P[N([0,t] \times (a(r/t)x + b(r/t), \infty)) < r] \\ &= P[\frac{N([0,t] \times (a(r/t)x + b(r/t), \infty)) - tQ(a(r/t)x + b(r/t))}{\sqrt{r}} < \frac{r - tQ(a(r/t)x + b(r/t))}{\sqrt{r}}]. \end{aligned}$$

From (4.12), $\sqrt{r}$ is asymptotic to the standard deviation of the Poisson random variable and so the left side random variable converges to a $N(0,1)$ random variable. Using (4.13), the right side converges to $\sqrt{t}h(x)$. We therefore conclude that under the left tail condition (4.11), for any fixed $t > 0$,

$$\lim_{r\to\infty} P\Big[\frac{Y^{(r)}(t) - b(r/t)}{a(r/t)} \le x\Big] = \Phi\big(\sqrt{t}h(x)\big), \quad x \in \text{supp}_\gamma, \tag{4.18}$$

where $\Phi(x)$ is the standard normal cdf.

Now we can prove the following finite dimensional convergence.

Proposition 4.2. *Assume (4.11) holds with $h(x)$ given in (4.16). Let $\{B(t), t \geq 0\}$ be standard Brownian motion. Then as $r \to \infty$,*

$$\frac{Y^{(r)}(t) - b(r/t)}{a(r/t)} \Rightarrow h^{\leftarrow}\Big(\frac{B(t)}{t}\Big), \tag{4.19}$$

in the sense of convergence of finite dimensional distributions for $t > 0$.

Proof. We illustrate the proof by showing bivariate pairs converge for two values of t. So suppose $0 < t_1 < t_2$ and $x_1 < x_2$ are in supp_γ and we show as $r \to \infty$,

$$\begin{aligned} P\Big[\frac{Y^{(r)}(t_i) - b(r/t_i)}{a(r/t_i)} \leq x_i;\, i = 1, 2\Big] &\to P\Big[h^{\leftarrow}\Big(\frac{B(t_i)}{t_i}\Big) \leq x_i;\, i = 1, 2\Big] \\ &= P\big[B(t_i) \leq t_i h(x_i);\, i = 1, 2\big]. \end{aligned} \tag{4.20}$$

We express the statements about $\boldsymbol{Y}^{(r)}$ in terms of the Poisson counting measure and consider:

$$\begin{aligned} &\begin{pmatrix} N\big([0, t_1] \times (a(r/t_1)x_1 + b(r/t_1), \infty)\big) \\ N\big([0, t_2] \times (a(r/t_2)x_2 + b(r/t_2), \infty)\big) \end{pmatrix} \\ &= \begin{pmatrix} N\big([0, t_1] \times (a(r/t_1)(x_1, x_2] + b(r/t_1), \infty)\big) + N\big([0, t_1] \times (a(r/t_1)x_2 + b(r/t_1), \infty)\big) \\ N\big([0, t_1] \times (a(r/t_2)x_2 + b(r/t_2), \infty)\big) + N\big((t_1, t_2] \times (a(r/t_2)x_2 + b(r/t_2), \infty)\big) \end{pmatrix} \\ &= \begin{pmatrix} N_1 + N_2 \\ N_3 + N_4 \end{pmatrix}. \end{aligned}$$

Consider the four terms N_i, $i = 1, \ldots, 4$ in turn.

(1) The term N_1 appropriately normed converges to 0,

$$\frac{N\big([0, t_1] \times (a(r/t_1)(x_1, x_2] + b(r/t_1), \infty)\big) - t_1\Pi(a(r/t_1)(x_1, x_2] + b(r/t_1))}{\sqrt{r}} \Rightarrow 0. \tag{4.21}$$

The reason is that the centering is

$$\begin{aligned} \frac{t_1\Pi(a(r/t_1)(x_1, x_2])}{\sqrt{r}} &= \frac{t_1Q(ax_1 + b) - t_1Q(ax_2 + b)}{\sqrt{r}} \\ &= \frac{r - t_1Q(ax_2 + b)}{\sqrt{r}} - \frac{r - t_1Q(ax_1 + b)}{\sqrt{r}} \\ &\to \sqrt{t_1}(h(x_2) - h(x_1)) > 0. \end{aligned}$$

So the left side of (4.21) is of the form $(N_r - \lambda_r)/\sqrt{r}$ where $\lambda_r/\sqrt{r} \to c > 0$ and thus

$$\text{Var}\Big((N_r - \lambda_r)/\sqrt{r}\Big) = \lambda_r/r \to 0,$$

which verifies the convergence to 0 in (4.21).

(2) The term N_2 becomes asymptotically normal. Let Z_1 be a standard normal random variable and apply (4.17) and (4.12) to get

$$\frac{N\big([0, t_1] \times (a(r/t_1)x_2 + b(r/t_1), \infty)\big) - t_1Q(a(r/t_1)x_2 + b(r/t_1))}{\sqrt{r}} \Rightarrow \sqrt{t_1}Z_1.$$

(3) For N_3, despite its dependence on the variable t_2, we also find
$$\frac{N\big([0,t_1]\times(a(r/t_2)x_2+b(r/t_2),\infty)\big)-t_1Q(a(r/t_2)x_2+b(r/t_2))}{\sqrt{r}}\Rightarrow\sqrt{t_1}Z_1.$$
This result uses a combination of the reasoning that was used for N_1, N_2.

(4) The term N_4 is independent of N_1, N_2, N_3 so there is a standard normal variable $Z_2 \perp\!\!\!\perp Z_1$ and
$$\frac{N\big((t_1,t_2]\times(a(r/t_2)x_2+b(r/t_2),\infty)\big)-(t_2-t_1)Q(a(r/t_2)x_2+b(r/t_2))}{\sqrt{r}}\Rightarrow\sqrt{t_2-t_1}Z_2.$$

We conclude from this carving that
$$\begin{pmatrix}\dfrac{N\big([0,t_1]\times(a(r/t_1)x_1+b(r/t_1),\infty)\big)-t_1Q(a(r/t_1)x_1+b(r/t_1))}{\sqrt{r}}\\ \dfrac{N\big([0,t_2]\times(a(r/t_2)x_2+b(r/t_2),\infty)\big)-t_2Q(a(r/t_2)x_2+b(r/t_2))}{\sqrt{r}}\end{pmatrix}$$
$$\Rightarrow\begin{pmatrix}\sqrt{t_1}Z_1\\ \sqrt{t_1}Z_1+\sqrt{t_2-t_1}Z_2\end{pmatrix},$$
as $r\to\infty$. Use (4.3) to write,
$$P\left[\begin{pmatrix}\dfrac{Y^{(r)}(t_1)-a(r/t_1)}{b(r/t_1)}\\ \dfrac{Y^{(r)}(t_2)-a(r/t_2)}{b(r/t_2)}\end{pmatrix}\le\begin{pmatrix}x_1\\ x_2\end{pmatrix}\right]$$
$$=P\left[\begin{pmatrix}\dfrac{N\big([0,t_1]\times(a(r/t_1)x_1+b(r/t_1),\infty)\big)-t_1Q(a(r/t_1)x_1+b(r/t_1))}{\sqrt{r}}\\ \dfrac{N\big([0,t_2]\times(a(r/t_2)x_2+b(r/t_2),\infty)\big)-t_2Q(a(r/t_2)x_2+b(r/t_2))}{\sqrt{r}}\end{pmatrix}\right.$$
$$\left.<\begin{pmatrix}\dfrac{r-t_1Q(a(r/t_1)x_1+b(r/t_1))}{\sqrt{r}}\\ \dfrac{r-t_2Q(a(r/t_2)x_2+b(r/t_2))}{\sqrt{r}}\end{pmatrix}\right]$$
$$\begin{aligned}&\to P[\sqrt{t_1}Z_1\le t_1h(x_1),\sqrt{t_1}Z_1+\sqrt{t_2-t_1}Z_2\le t_2h(x_2)]\quad(\text{as } r\to\infty)\\ &=P\big[\frac{B(t_1)}{t_1}\le h(x_1),\frac{B(t_2)}{t_2}\le h(x_2)\big]\\ &=P\big[h^{\leftarrow}\big(\frac{B(t_1)}{t_1}\big)\le x_1,h^{\leftarrow}\big(\frac{B(t_2)}{t_2}\big)\le x_2\big].\end{aligned}$$
This verifies (4.20). $\square$

5. Final thoughts

The results of this paper suggest some obvious questions the answers to which have so far eluded us. Is there a jump process limit – presumably some sort of extremal process – in (4.19) corresponding to some sort of Poisson limit regime as opposed to the Brownian motion limit regime? In Proposition 4.2 is a stronger form of convergence – say in the J_1-topology – possible? And so far, the mathematics of proving in a nice way that $\{\boldsymbol{Y}^{(r)}, r \geq 1\}$ is Markov in the càdlàg space $D(0,\infty)$ has not cooperated.

References

[1] B. C. Arnold, A. Becker, U. Gather, and H. Zahedi. On the Markov property of order statistics. *J. Statist. Plann. Inference*, 9(2):147–154, 1984.

[2] E. Cramer and T.H. Tran. Generalized order statistics from arbitrary distributions and the Markov chain property. *J. Statist. Plann. Inference*, 139(12):4064–4071, 2009.

[3] L. de Haan and A. Ferreira. *Extreme Value Theory: An Introduction*. Springer-Verlag, New York, 2006.

[4] P. Deheuvels. A construction of extremal processes. In *Probability and Statistical Inference (Bad Tatzmannsdorf, 1981)*, pages 53–57. Reidel, Dordrecht, 1982.

[5] P. Deheuvels. The strong approximation of extremal processes. II. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 62(1):7–15, 1983.

[6] M. Dwass. Extremal processes. *Ann. Math. Statist*, 35:1718–1725, 1964.

[7] M. Dwass. Extremal processes. II. *Illinois J. Math.*, 10:381–391, 1966.

[8] M. Dwass. Extremal processes. III. *Bull. Inst. Math. Acad. Sinica*, 2:255–265, 1974. Collection of articles in celebration of the sixtieth birthday of Ky Fan.

[9] R. Engelen, P. Tommassen, and W. Vervaat. Ignatov's theorem: a new and short proof. *J. Appl. Probab.*, Special Vol. 25A:229–236, 1988. A celebration of applied probability.

[10] C. M. Goldie and L. C. G. Rogers. The k-record processes are i.i.d. *Z. Wahrsch. Verw. Gebiete*, 67(2):197–211, 1984.

[11] C.M. Goldie and R.A. Maller. Generalized densities of order statistics. *Statist. Neerlandica*, 53(2):222–246, 1999.

[12] Z. Ignatov. Ein von der Variationsreihe erzeugter Poissonscher Punktprozeß. *Annuaire Univ. Sofia Fac. Math. Méc.*, 71(2):79–94 (1986), 1976/77.

[13] G. Matheron. *Random Sets and Integral Geometry*. John Wiley & Sons, New York-London-Sydney, 1975. With a foreword by G.S. Watson, Wiley Series in Probability and Mathematical Statistics.

[14] I. Molchanov. *Theory of Random Sets*. Probability and its Applications (New York). Springer-Verlag London Ltd., London, 2005.

[15] A. Rényi. Théorie des éléments saillants d'une suite d'observations. *Ann. Fac. Sci. Univ. Clermont-Ferrand No.*, 8:7–13, 1962.

[16] S.I. Resnick. Tail equivalence and its applications. *J. Appl. Probab.*, 8:136–156, 1971.

[17] S.I. Resnick. Limit laws for record values. *Stochastic Processes Appl.*, 1:67–82, 1973.

[18] S.I. Resnick. Inverses of extremal processes. *Adv. in Appl. Probab.*, 6:392–406, 1974.

[19] S.I. Resnick. Weak convergence to extremal processes. *Ann. Probab.*, 3(6):951–960, 1975.

[20] S.I. Resnick. *Heavy Tail Phenomena: Probabilistic and Statistical Modeling*. Springer Series in Operations Research and Financial Engineering. Springer-Verlag, New York, 2007. ISBN: 0-387-24272-4.

[21] S.I. Resnick. *Extreme Values, Regular Variation and Point Processes*. Springer, New York, 2008. Reprint of the 1987 original.

[22] S.I. Resnick and M. Rubinovitch. The structure of extremal processes. *Adv. in Appl. Probability*, 5:287–307, 1973.

[23] L. Rüschendorf. Two remarks on order statistics. *J. Statist. Plann. Inference*, 11(1):71–74, 1985.

[24] R. W. Shorrock. On discrete time extremal processes. *Adv. in Appl. Probab.*, 6:580–592, 1974.

[25] R. W. Shorrock. Extremal processes and random measures. *J. Appl. Probability*, 12:316–323, 1975.

[26] A. J. Stam. Independent Poisson processes generated by record values and inter-record times. *Stochastic Process. Appl.*, 19(2):315–325, 1985.

[27] W. Vervaat and H. Holwerda, editors. *Probability and lattices*, volume 110 of *CWI Tract*. Stichting Mathematisch Centrum, Centrum voor Wiskunde en Informatica, Amsterdam, 1997.

[28] I. Weissman. Extremal processes generated by independent nonidentically distributed random variables. *Ann. Probab.*, 3:172–177, 1975.

Dr. Boris Buchmann, Research School of Finance, Actuarial Studies & Statistics,, Australian National University, ACT 0200, Australia

E-mail address: Boris.Buchmann@anu.edu.au

Prof. Ross Maller,, Research School of Finance, Actuarial Studies & Statistics,, Australian National University, Canberra, ACT,Australia,

E-mail address: Ross.Maller@anu.edu.au

Prof. Sidney Resnick, School of Operations Research and Information Engineering, Cornell University, Ithaca, NY 14853 USA

E-mail address: sir1@cornell.edu