Bernoulli **21**(3), 2015, 1575–1599
DOI: 10.3150/14-BEJ614

Maxima of long memory stationary symmetric α-stable processes, and self-similar processes with stationary max-increments

TAKASHI OWADA[1] and GENNADY SAMORODNITSKY[2]

[1]*Faculty of Electrical Engineering, Technion, Haifa, Israel 32000. E-mail: takashiowada@ee.technion.ac.il*
[2]*School of Operations Research and Information Engineering, and Department of Statistical Science, Cornell University, Ithaca, NY 14853, USA. E-mail: gs18@cornell.edu*

We derive a functional limit theorem for the partial maxima process based on a long memory stationary α-stable process. The length of memory in the stable process is parameterized by a certain ergodic-theoretical parameter in an integral representation of the process. The limiting process is no longer a classical extremal Fréchet process. It is a self-similar process with α-Fréchet marginals, and it has stationary max-increments, a property which we introduce in this paper. The functional limit theorem is established in the space $D[0,\infty)$ equipped with the Skorohod M_1-topology; in certain special cases the topology can be strengthened to the Skorohod J_1-topology.

Keywords: conservative flow; extreme value theory; pointwise dual ergodicity; sample maxima; stable process

1. Introduction

The asymptotic behaviour of the partial maxima sequence $M_n = \max_{1\le k\le n} X_k$, $n=1,2,\dots$ for an i.i.d. sequence $(X_1, X_2, \dots)$ of random variables is the subject of the classical extreme value theory, dating back to Fisher and Tippett [11]. The basic result of this theory says that only three one-dimensional distributions, the Fréchet distribution, the Weibull distribution and the Gumbel distribution, have a max-domain of attraction. If Y has one of these three distributions, then for a distribution in its domain of attraction, and a sequence of i.i.d. random variables with that distribution,

$$\frac{M_n - b_n}{a_n} \quad \Rightarrow \quad Y \tag{1.1}$$

for properly chosen sequences (a_n), (b_n); see, for example, Chapter 1 in Resnick [27] or Section 1.2 in de Haan and Ferreira [6]. Under the same max-domain of attraction assumption, a functional version of (1.1) was established in Lamperti [18]: with the same sequences (a_n), (b_n) as in (1.1),

$$\left(\frac{M_{\lfloor nt\rfloor} - b_n}{a_n}, t\ge 0\right) \quad \Rightarrow \quad \bigl(Y(t), t\ge 0\bigr) \tag{1.2}$$

for a nondecreasing right continuous process $(Y(t), t \geq 0)$, and the convergence is weak convergence in the Skorohod J_1-topology on $D[0, \infty)$. The limiting process is often called *the extremal process*; its properties were established in Dwass [7,8] and Resnick and Rubinovitch [28].

Much of the more recent research in extreme value theory concentrated on the case when the underlying sequence $(X_1, X_2, \ldots)$ is stationary, but may be dependent. In this case the extrema of the sequence may cluster, and it is natural to expect that the limiting results (1.1) and (1.2) will, in general, have to be different. The extremes of moving average processes have received special attention; see, for example, Rootzén [29], Davis and Resnick [5] and Fasen [10]. The extremes of the GARCH(1, 1) process were investigated in Mikosch and Stărică [21]. The classical work on the extremes of dependent sequences is Leadbetter *et al*. [20]; in some cases this clustering of the extremes can be characterized through the *extremal index* (introduced, originally, in Leadbetter [19]). The latter is a number $0 \leq \theta \leq 1$. Suppose that a stationary sequence $(X_1, X_2, \ldots)$ has this index, and let $(\tilde{X}_1, \tilde{X}_2, \ldots)$ be an i.i.d. sequence with the same one-dimensional marginal distributions as $(X_1, X_2, \ldots)$. If (1.1) and (1.2) hold for the i.i.d. sequence, then the corresponding limits will satisfy $\tilde{Y} \stackrel{d}{=} \tilde{Y}(1)$, but the limit in (1.1) for the dependent sequence $(X_1, X_2, \ldots)$ will satisfy $Y \stackrel{d}{=} \tilde{Y}(\theta)$. In particular, the limit will be equal to zero if the extremal index is equal to zero. This case can be viewed as that of long range dependence in the extremes, and it has been mostly neglected by the extreme value community. Long range dependence is, however, an important phenomenon in its own right, and in this paper we take a step towards understanding how long range dependence affects extremes.

A random variable X is said to have a regularly varying tail with index $-\alpha$ for $\alpha > 0$ if

$$P(X > x) = x^{-\alpha} L(x), \qquad x > 0,$$

where L is a slowly varying at infinity function, and the distribution of any such random variable is in the max-domain of attraction of the Fréchet distribution with the same parameter α; see, for example, Resnick [27]. Recall that the Fréchet law $F_{\alpha,\sigma}$ on $(0, \infty)$ with the tail index α and scale $\sigma > 0$ satisfies

$$F_{\alpha,\sigma}(x) = \exp\{-\sigma^\alpha x^{-\alpha}\}, \qquad x > 0. \tag{1.3}$$

Sometimes the term *α-Fréchet* is used. In this paper, we discuss the case of regularly varying tails and the resulting limits in (1.2). The limits obtained in this paper belong to the family of the so-called *Fréchet processes*, defined below. We would like to emphasize that, even for stationary sequences with regularly varying tails, non-Fréchet limits may appear in (1.2). We are postponing a detailed discussion of this point to a future publication.

A stochastic process $(Y(t), t \in T)$ (on an arbitrary parameter space T) is called a Fréchet process if for all $n \geq 1, a_1, \ldots, a_n > 0$ and $t_1, \ldots, t_n \in T$, the weighted maximum $\max_{1 \leq j \leq n} a_j Y(t_j)$ follows a Fréchet law as in (1.3). The best known Fréchet process is the extremal Fréchet process obtained in the scheme (1.2) starting with an i.i.d. sequence with regularly varying tails. The extremal Fréchet process $(Y(t), t \geq 0)$ has finite-dimensional distributions defined by

$$\begin{aligned}\big(Y(t_1), Y(t_2), \ldots, Y(t_n)\big) \stackrel{d}{=} \big(&X^{(1)}_{\alpha, t_1^{1/\alpha}}, \max\big(X^{(1)}_{\alpha, t_1^{1/\alpha}}, X^{(2)}_{\alpha,(t_2-t_1)^{1/\alpha}}\big), \ldots, \\ &\max\big(X^{(1)}_{\alpha, t_1^{1/\alpha}}, X^{(2)}_{\alpha,(t_2-t_1)^{1/\alpha}}, \ldots, X^{(n)}_{\alpha,(t_n-t_{n-1})^{1/\alpha}}\big)\big)\end{aligned} \tag{1.4}$$

for all n and $0 \le t_1 < t_2 < \cdots < t_n$. The different random variables in the right-hand side of (1.4) are independent, with $X_{\alpha,\sigma}^{(k)}$ having the Fréchet law $F_{\alpha,\sigma}$ in (1.3), for any $k = 1, \dots, n$. The stationarity and independence of the max-increments of the extremal Fréchet processes make it similar to the better known Lévy processes which have stationary and independent sum-increments. The structure of general Fréchet processes has been extensively studied in the last several years. These processes were introduced in Stoev and Taqqu [39], and their representations (as a part of a much more general context) were studied in Kabluchko and Stoev [14]. Stationary Fréchet processes (in particular, their ergodicity and mixing) were discussed in Stoev [38], Kabluchko *et al.* [13] and Wang and Stoev [41].

In this paper, we concentrate on the maxima of stationary α-stable processes with $0 < \alpha < 2$. Recall that a random vector $\mathbf{X}$ in $\mathbb{R}^d$ is called α-stable if for any A and $B > 0$ we have

$$A\mathbf{X}^{(1)} + B\mathbf{X}^{(2)} \stackrel{d}{=} \left(A^\alpha + B^\alpha\right)^{1/\alpha}\mathbf{X} + \mathbf{y},$$

where $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ are i.i.d. copies of $\mathbf{X}$, and $\mathbf{y}$ is a deterministic vector (unless $\mathbf{X}$ is deterministic, necessarily, $0 < \alpha \le 2$). A stochastic process $(X(t), t \in T)$ is called α-stable if all of its finite-dimensional distributions are α-stable. We refer the reader to Samorodnitsky and Taqqu [36] for information on α-stable processes. When $\alpha = 2$, an α-stable process is Gaussian, while in the case $0 < \alpha < 2$, both the left and the right tails of a (nondegenerate) α-stable random variable X are (generally) regularly varying with exponent α. That is,

$$P(X > x) \sim c_+ x^{-\alpha}, \qquad P(X < -x) \sim c_- x^{-\alpha} \qquad \text{as } x \to \infty$$

for some $c_+, c_- \ge 0$, $c_+ + c_- > 0$. That is, if $(X_1, X_2, \dots)$ is an i.i.d. sequence of α-stable random variables, then the i.i.d. sequence $(|X_1|, |X_2|, \dots)$ satisfies (1.1) and (1.2) with $a_n = n^{1/\alpha}$ (and $b_n = 0$), $n \ge 1$. Of course, we are not planning to study the extrema of an i.i.d. α-stable sequence. Instead, we will study the maxima of (the absolute values of) a stationary α-stable process. The reason we have chosen to work with stationary α-stable processes is that their structure is very rich, and is also relatively well understood. This will allow us to study the effect of that structure on the limit theorems (1.1) and (1.2). We are specifically interested in the long range dependent case, corresponding to the zero value of the extremal index.

The structure of stationary symmetric α-stable (SαS) processes has been clarified in the last several years in the works of Jan Rosiński; see, for example, Rosiński [30,31]. The integral representation of such a process can be chosen to have a very special form. The class of stationary SαS processes we will investigate requires a representation slightly more restrictive than the one generally allowed. Specifically, we will consider discrete-time stationary processes of the form

$$X_n = \int_E f \circ T^n(x)\, \mathrm{d}M(x), \qquad n = 1, 2, \dots, \tag{1.5}$$

where M is a SαS random measure on a measurable space $(E, \mathcal{E})$ with a σ-finite *infinite* control measure μ. The map $T : E \to E$ is a measurable map that preserves the measure μ. Further, $f \in L^\alpha(\mu)$. See Samorodnitsky and Taqqu [36] for details on α-stable random measures and integrals with respect to these measures. It is elementary to check that a process with a representation

(1.5) is, automatically, stationary. Recall that any stationary SαS process has a representation of the form:

$$X_n = \int_E f_n(x)\,\mathrm{d}M(x), \qquad n = 1, 2, \dots, \tag{1.6}$$

with

$$f_n(x) = a_n(x)\left(\frac{\mathrm{d}\mu \circ T^n}{\mathrm{d}\mu}(x)\right)^{1/\alpha} f \circ T^n(x), \qquad x \in E \tag{1.7}$$

for $n = 1, 2, \dots$, where $T : E \to E$ is a one-to-one map with both T and T^{-1} measurable, mapping the control measure μ into an equivalent measure, and the sequence (a_n) takes values ± 1 (and has the so-called cocycle property). Here M is SαS (and $f \in L^\alpha(\mu)$). See Rosiński [30].

In (1.5) we assume, however, that map T is measure preserving. The main reason is that the ergodic-theoretical notions we are using have been developed for measure preserving maps. Indeed, it has been observed that the ergodic-theoretical properties of the map T, either in (1.5) or in (1.7), have a major impact on the memory of a stationary α-stable process. See, for example, Surgailis *et al*. [40], Samorodnitsky [33,34], Roy [32], Resnick and Samorodnitsky [25], Owada and Samorodnitsky [24], Owada [23]. The most relevant for this work is the result of Samorodnitsky [33], who proved that, if the map T in (1.5) or in (1.7) is conservative, then using the normalization $a_n = n^{1/\alpha}$ ($b_n = 0$) in (1.1), as indicated by the marginal tails, produces the zero limit, so the partial maxima grow, in this case, strictly slower than at the rate of $n^{1/\alpha}$. On the other hand, if the map T is not conservative, then the normalization $a_n = n^{1/\alpha}$ in (1.1) is the correct one, and it leads to a Fréchet limit (we will survey the ergodic-theoretical notions in the next section). Therefore, the extrema of SαS processes corresponding to conservative flows cluster so much that the sequence of the partial maxima grows at a slower rate than that indicated by the marginal tails. This case can be thought of as indicating long range dependence. It is, clearly, inconsistent with a positive extremal index.

The Fréchet limit obtained in (1.1) by Samorodnitsky [33] remains valid when the map T is conservative (but with the normalization of a smaller order than $n^{1/\alpha}$), as long as the map T satisfies a certain additional assumption. If one views the stationary α-stable process as a natural function of the Poisson points forming the random measure M in (1.6) then, informally, this assumption guarantees that only the largest Poisson point contributes, distributionally, to the asymptotic behaviour of the partial maxima of the process. In this paper, we restrict ourselves to this situation as well. However, we will look at the limits obtained in the much more informative functional scheme (1.2). In this paper, the assumption on the map T will be expressed in terms of the rate of growth of the so-called wandering rate sequence, which we define in the sequel. We would like to emphasize that, when this wandering rate sequence grows at a rate slower than the one assumed in this paper, new phenomena seem to arise. Multiple Poisson points may contribute to the asymptotic distribution of the partial maxima, and non-Fréchet limit may appear in (1.2). We leave a detailed study of this to a subsequent work.

In the next section, we provide the elements of the infinite ergodic theory needed for the rest of the paper. In Section 3 we introduce a new notion, that of a process with stationary max-increments. It turns out that the possible limits in the functional maxima scheme (1.2) (with $b_n = 0$) are self-similar with stationary max-increments. We discuss the general properties of such processes and then specialize to the concrete limiting process we obtain in the main result of the paper, stated and proved in Section 4.

2. Ergodic theoretical notions

In this section, we present some basic notation and notions of, mostly infinite, ergodic theory used in the sequel. The main references are Krengel [15], Aaronson [2], and Zweimüller [43].

Let $(E, \mathcal{E}, \mu)$ be a σ-finite, infinite measure space. We will say that $A = B \bmod \mu$ if $A, B \in \mathcal{E}$ and $\mu(A \triangle B) = 0$. For $f \in L^1(\mu)$ we will often write $\mu(f)$ for the integral $\int f \, d\mu$.

Let $T : E \to E$ be a measurable map preserving the measure μ. The sequence (T^n) of iterates of T is called a *flow*, and the ergodic-theoretical properties of the map and the flow are identified. A map T is called *ergodic* if any T-invariant set A (i.e., a set such that $T^{-1}A = A \bmod \mu$) is trivial, that is, it satisfies $\mu(A) = 0$ or $\mu(A^c) = 0$. A map T is said to be *conservative* if

$$\sum_{n=1}^{\infty} \mathbf{1}_A \circ T^n = \infty \qquad \text{a.e. on } A$$

for any $A \in \mathcal{E}$, $0 < \mu(A) < \infty$; if T is also ergodic, then the restriction "*on* A" is not needed.

The *conservative part* of a measure-preserving T is the largest T-invariant subset C of E such that the restriction of T to C is conservative. The set $D = E \setminus C$ is the *dissipative part* of T (and the decomposition $E = C \cup D$ is called *the Hopf decomposition* of T).

The *dual operator* $\widehat{T} : L^1(\mu) \to L^1(\mu)$ is defined by

$$\widehat{T} f = \frac{d(\nu_f \circ T^{-1})}{d\mu}, \qquad f \in L^1(\mu), \tag{2.1}$$

where ν_f is the signed measure $\nu_f(A) = \int_A f \, d\mu$, $A \in \mathcal{E}$. The dual operator satisfies the duality relation

$$\int_E \widehat{T} f \cdot g \, d\mu = \int_E f \cdot g \circ T \, d\mu \tag{2.2}$$

for $f \in L^1(\mu)$, $g \in L^\infty(\mu)$. Note that (2.1) makes sense for any nonnegative measurable function f on E, and the resulting $\widehat{T} f$ is again a nonnegative measurable function. Furthermore, (2.2) holds for arbitrary nonnegative measurable functions f and g.

A conservative, ergodic and measure preserving map T is said to be *pointwise dual ergodic*, if there exists a normalizing sequence $a_n \nearrow \infty$ such that

$$\frac{1}{a_n} \sum_{k=1}^{n} \widehat{T}^k f \to \mu(f) \qquad \text{a.e. for every } f \in L^1(\mu). \tag{2.3}$$

The property of pointwise dual ergodicity rules out invertibility of the map T. Since the measure μ is infinite, choosing a nonnegative function f and using Fatou's lemma shows that only rates $a_n = o(n)$ are possible in pointwise dual ergodicity. Intuitively, as will be seen in (2.6) below, the longer time it takes the trajectory of a point under the map T to return to a set of a finite positive measure, the smaller is the normalizing sequence (a_n).

Sometimes we require that for some functions the above convergence takes place uniformly on a certain set. A set $A \in \mathcal{E}$ with $0 < \mu(A) < \infty$ is said to be a *uniform set* for a conservative, ergodic and measure preserving map T, if there exist a normalizing sequence $a_n \nearrow \infty$ and a nontrivial nonnegative measurable function $f \in L^1(\mu)$ (nontriviality means that f is different from zero on a set of positive measure) such that

$$\frac{1}{a_n}\sum_{k=1}^{n}\widehat{T}^k f \to \mu(f) \qquad \text{uniformly, a.e. on } A. \tag{2.4}$$

If (2.4) holds for $f = \mathbf{1}_A$, the set A is called *a Darling–Kac set*. A conservative, ergodic and measure preserving map T is pointwise dual ergodic if and only if T admits a uniform set; see Proposition 3.7.5 in Aaronson [2]. In particular, it is legitimate to use the same normalizing sequence (a_n) both in (2.3) and (2.4).

Let $A \in \mathcal{E}$ with $0 < \mu(A) < \infty$. The frequency of visits to the set A along the trajectory $(T^n x)$, $x \in E$, is naturally related to the *wandering rate* sequence

$$w_n = \mu\left(\bigcup_{k=0}^{n-1} T^{-k}A\right). \tag{2.5}$$

If we define the first entrance time to A by

$$\varphi_A(x) = \min\{n \ge 1 : T^n x \in A\}$$

(notice that $\varphi_A < \infty$ a.e. on E since T is conservative and ergodic), then $w_n \sim \mu(\varphi_A < n)$ as $n \to \infty$. Since T is also measure preserving, we have $\mu(A \cap \{\varphi_A > k\}) = \mu(A^c \cap \{\varphi_A = k\})$ for $k \ge 1$ (see, e.g., Zweimüller [43]). Therefore, alternative expressions for the wandering rate sequence are

$$w_n = \mu(A) + \sum_{k=1}^{n-1}\mu\bigl(A^c \cap \{\varphi_A = k\}\bigr) = \sum_{k=0}^{n-1}\mu\bigl(A \cap \{\varphi_A > k\}\bigr).$$

Suppose now that T is a pointwise dual ergodic map, and let A be a uniform set for T. It turns out that, under an assumption of regular variation, there is a precise connection between the wandering rate sequence (w_n) and the normalizing sequence (a_n) in (2.3) and (2.4). Specifically, let RV_γ represent the class of regularly varying at infinity sequences (or functions, depending on the context) of index γ. If either $(w_n) \in RV_\beta$ or $(a_n) \in RV_{1-\beta}$ for some $\beta \in [0, 1]$, then

$$a_n \sim \frac{1}{\Gamma(2-\beta)\Gamma(1+\beta)}\frac{n}{w_n} \qquad \text{as } n \to \infty. \tag{2.6}$$

Proposition 3.8.7 in Aaronson [2] gives one direction of this statement, but the argument is easily reversed. The normalizing sequence (a_n) and the wandering rate sequence (w_n) are both related to the frequency with which a uniform set A is visited along the trajectory $(T^n x)$ that starts in A.

We finish this section with a statement on distributional convergence of the partial maxima for pointwise dual ergodic flows. It will be used repeatedly in the proof of the main theorem. For a measurable function f on E define

$$M_n(f)(x) = \max_{1\le k\le n} \left| f \circ T^k(x)\right|, \qquad x \in E, n \ge 1.$$

The proposition below involves weak convergence in the space $D[0,\infty)$ equipped with two different topologies, the Skorohod J_1-topology and the Skorohod M_1-topology, introduced in Skorohod [37]. The details could be found, for instance, in Billingsley [4] (for the J_1-topology), and in Whitt [42] (for the M_1-topology). See also Remark 2.2.

In the sequel, we will use the convention $\max_{k\in K} b_k = 0$ for a nonnegative sequence (b_n), if $K = \varnothing$.

Proposition 2.1. *Let T be a pointwise dual ergodic map on a σ-finite, infinite, measure space $(E, \mathcal{E}, \mu)$. We assume that the normalizing sequence (a_n) is regularly varying with exponent $1-\beta$ for some $0 < \beta \le 1$. Let $A \in \mathcal{E}$, $0 < \mu(A) < \infty$, be a uniform set for T. Define a probability measure on E by $\mu_n(\cdot) = \mu(\cdot \cap \{\varphi_A \le n\})/\mu(\{\varphi_A \le n\})$. Let $f : E \to \mathbb{R}$ be a measurable bounded function supported by the set A, that is, $\mathrm{supp}(f) \subset A$. Let $\|f\|_\infty = \inf\{M : |f(x)| \le M \text{ a.e. on } A\}$. Then*

$$\begin{aligned}&\big(M_{\lfloor nt\rfloor}(f), 0 \le t \le 1\big)\\ &\quad\Rightarrow \quad \|f\|_\infty \big(\mathbf{1}_{\{V_\beta \le t\}}, 0 \le t \le 1\big) \qquad \text{in the } M_1\text{-topology on } D[0,1],\end{aligned} \tag{2.7}$$

where the law of the left-hand side is computed with respect to μ_n, and V_β is a random variable defined on a probability space $(\Omega', \mathcal{F}', P')$ with $P'(V_\beta \le x) = x^\beta$, $0 < x \le 1$. If $f = \mathbf{1}_A$, then the convergence above takes place in the J_1-topology as well.

Remark 2.2. It is not difficult to see why the weak convergence in (2.7) holds in the J_1-topology for indicator functions, but only in the M_1-topology in general. Indeed, for functions f other than the indicator function, the limiting value of $\|f\|_\infty$ may have an asymptotically non-vanishing probability of being reached in multiple closely placed steps, which precludes the J_1-tightness, since the J_1-modulus does not become small; see, for example, Theorem 13.2 in Billingsley [4]. One can easily construct (very general) examples of situations in which this can be made precise. On the other hand, if $f = \mathbf{1}_A$, then the limiting value is reached by a single jump, matching the single jump in the limiting process, which gives convergence in the J_1-topology.

Proof of Proposition 2.1. For $0 < \varepsilon < 1$, let $A_\varepsilon = \{x \in A : |f(x)| \ge (1-\varepsilon)\|f\|_\infty\}$. Note that each A_ε is uniform since A is uniform. Clearly,

$$(1-\varepsilon)\|f\|_\infty \mathbf{1}_{\{\varphi_{A_\varepsilon}(x)\le nt\}} \le M_{\lfloor nt\rfloor}(f)(x) \le \|f\|_\infty \mathbf{1}_{\{\varphi_A(x)\le nt\}} \qquad \mu\text{-a.e.}$$

for all $n \ge 1$ and $0 \le t \le 1$. Since for monotone functions weak convergence in the M_1-topology is implied by convergence in finite-dimensional distributions (see, e.g., Proposition 2 in Avram and Taqqu [3]), we can use Theorem 3.2 in Billingsley [4] in a finite-dimensional situation. The statement of the proposition will follow once we show that, for a uniform set B (which could be either A or A_ε) the law of φ_B/n under μ_n converges to the law of V_β. Let $(w_n^{(B)})$ be the corresponding wandering rate sequence. Since (2.6) holds for $(w_n^{(B)})$ with the same normalizing constants (a_n), we know that $w_n^{(B)} \sim w_n^{(A)} := w_n$ as $n \to \infty$. Therefore,

$$\mu_n\left(\frac{\varphi_B}{n} \le x\right) = \frac{\mu(\varphi_B \le \lfloor nx \rfloor)}{\mu(\varphi_A \le n)} \sim \frac{w_{\lfloor nx \rfloor}^{(B)}}{w_n} \to x^\beta$$

for all $0 < x \le 1$, because the wandering rate sequence (w_n) is regularly varying with index β by (2.6).

Next, suppose that $f(x) = \mathbf{1}_A(x)$. In this case, $M_{\lfloor nt \rfloor}(\mathbf{1}_A)(x) = \mathbf{1}_{\{\varphi_A(x) \le nt\}}$. An application of the Skorohod embedding theorem tells us that on some common probability space, the time of the jump of the process $\mathbf{1}_{\{\varphi_A(\cdot) \le nt\}}$ converges a.s. to the time of the jump of the process $\mathbf{1}_{\{V_\beta \le t\}}$. This, in turn, implies a.s. convergence of these processes in the space $D[0,1]$ in the J_1-topology, hence their weak convergence in that topology. □

3. Self-similar processes with stationary max-increments

The limiting process obtained in the next section shares with any possible limits in the functional maxima scheme (1.2) (with $b_n = 0$) two very specific properties, one of which is classical, and the other is less so. Recall that a stochastic process $(Y(t), t \ge 0)$ is called self-similar with exponent H of self-similarity if for any $c > 0$

$$\big(Y(ct), t \ge 0\big) \stackrel{d}{=} \big(c^H Y(t), t \ge 0\big)$$

in the sense of equality of finite-dimensional distributions. The best known classes of self-similar processes arise in various versions of a functional central limit theorem for stationary processes, and they have an additional property of stationary increments. Recall that a stochastic process $(Y(t), t \ge 0)$ is said to have stationary increments if for any $r \ge 0$

$$\big(Y(t+r) - Y(r), t \ge 0\big) \stackrel{d}{=} \big(Y(t) - Y(0), t \ge 0\big); \tag{3.1}$$

see, for example, Embrechts and Maejima [9] and Samorodnitsky [35]. In the context of the functional limit theorem for the maxima (1.2), a different property appears.

Definition 3.1. *A stochastic process $(Y(t), t \ge 0)$ is said to have stationary max-increments if for every $r \ge 0$, there exists, perhaps on an enlarged probability space, a stochastic process $(Y^{(r)}(t), t \ge 0)$ such that*

$$\begin{aligned} \big(Y^{(r)}(t), t \ge 0\big) &\stackrel{d}{=} \big(Y(t), t \ge 0\big), \\ \big(Y(t+r), t \ge 0\big) &\stackrel{d}{=} \big(Y(r) \vee Y^{(r)}(t), t \ge 0\big). \end{aligned} \tag{3.2}$$

Notice the analogy between the definition (3.1) of stationary increments (when $Y(0) = 0$) and Definition 3.1. Since the operations of taking the maximum is not invertible (unlike summation), the latter definition, by necessity, is stated in terms of existence of the max-increment process $(Y^{(r)}(t), t \geq 0)$.

Theorem 3.2. *Let $(X_1, X_2, \ldots)$ be a stationary sequence. Assume that for some sequence $a_n \to \infty$, and a stochastic process $(Y(t), t \geq 0)$ such that $P(Y(t) = Y(1)) < 1$ for $t \neq 1$,*

$$\left(\frac{1}{a_n} M_{\lfloor nt \rfloor}, t \geq 0\right) \quad \Rightarrow \quad \big(Y(t), t \geq 0\big)$$

in terms of convergence of finite-dimensional distributions. Then $(Y(t), t \geq 0)$ is self-similar with exponent $H > 0$ of self-similarity, and has stationary max-increments. Furthermore, $(Y(t), t \geq 0)$ is continuous in probability. The sequence (a_n) is regularly varying with index H.

Proof. The facts that the limiting process $(Y(t), t \geq 0)$ is self-similar with exponent $H \geq 0$ of self-similarity, and that the sequence (a_n) is regularly varying with index H, follow from the Lamperti theorem; see Lamperti [17], or Theorem 2.1.1 in Embrechts and Maejima [9]. The case $H = 0$ is ruled out by the assumption that $P(Y(t) = Y(1)) < 1$ for $t \neq 1$. Lamperti's theorem is usually stated and proved in the context of convergence in the situation when the time is scaled by a parameter converging to infinity along the real values, whereas in our situation the time scaling converges to infinity along a discrete sequence of the integers. However, it is easy to check that for maxima of stationary processes convergence along a discrete sequence provides the same information as convergence along all real values. Note, further, that for every $0 \leq t_1 < t_2$ and n large enough,

$$\frac{1}{a_n}(M_{\lfloor nt_2 \rfloor} - M_{\lfloor nt_1 \rfloor}) \leq \frac{1}{a_n} \max_{nt_1 < i \leq nt_2} X_i \stackrel{\text{st}}{\leq} \frac{1}{a_n} M_{\lfloor 2n(t_2 - t_1) \rfloor}$$

by the stationarity. Taking weak limits, we see that the difference $Y(t_2) - Y(t_1)$ is nonnegative and bounded stochastically by $Y(2(t_2 - t_1))$. Therefore, it follows from the self-similarity of $(Y(t), t \geq 0)$ that it is continuous in probability.

We check now the stationarity of the max-increments of the limiting process. Let $r > 0$, and $t_i > 0, i = 1, \ldots, k$, some $k \geq 1$. Write

$$\frac{1}{a_n} M_{\lfloor n(t_i + r) \rfloor} = \frac{1}{a_n} M_{\lfloor nr \rfloor} \bigvee \frac{1}{a_n} \max_{nr < j \leq n(t_i + r)} X_j, \qquad i = 1, \ldots, k. \tag{3.3}$$

By the assumption of the theorem and stationarity of the process $(X_1, X_2, \ldots)$,

$$\frac{1}{a_n} M_{\lfloor nr \rfloor} \quad \Rightarrow \quad Y(r), \qquad \left(\frac{1}{a_n} \max_{nr < j \leq n(t_i + r)} X_j, i = 1, \ldots, k\right) \quad \Rightarrow \quad \big(Y(t_1), \ldots, Y(t_k)\big)$$

as $n \to \infty$. Since every weakly converging sequence is tight, and a sequence with tight marginals is itself tight, we conclude that

$$\left(\frac{1}{a_n} M_{\lfloor nr \rfloor}, \left(\frac{1}{a_n} \max_{nr < j \leq n(t_i + r)} X_j, i = 1, \ldots, k\right)\right)$$

is a tight sequence. This tightness means that for every sequence $n_m \to \infty$ there is a subsequence $n_{m(l)} \to \infty$ and a k-dimensional random vector $(Y^{(r)}(t_1), \ldots, Y^{(r)}(t_k)) \stackrel{d}{=} (Y(t_1), \ldots, Y(t_k))$ such that as $l \to \infty$,

$$\begin{aligned}&\left(\frac{1}{a_{n_{m(l)}}} M_{\lfloor n_{m(l)} r \rfloor}, \left(\frac{1}{a_{n_{m(l)}}} \max_{n_{m(l)} r < j \le n_{m(l)}(t_i + r)} X_j, i = 1, \ldots, k\right)\right)\\ &\quad \Rightarrow \quad \left(Y(r), \left(Y^{(r)}(t_1), \ldots, Y^{(r)}(t_k)\right)\right).\end{aligned}$$

Let now τ_i, $i = 1, 2, \ldots$ be an enumeration of the rational numbers in $[0, \infty)$. A diagonalization argument shows that there is a sequence $n_m \to \infty$ and a stochastic process $(Y^{(r)}(\tau_i), i = 1, 2, \ldots)$ with $(Y^{(r)}(\tau_i), i = 1, 2, \ldots) \stackrel{d}{=} (Y(\tau_i), i = 1, 2, \ldots)$ such that

$$\begin{aligned}&\left(\frac{1}{a_{n_m}} M_{\lfloor n_m r \rfloor}, \left(\frac{1}{a_{n_m}} \max_{n_m r < j \le n_m(\tau_i + r)} X_j, i = 1, 2, \ldots\right)\right)\\ &\quad \Rightarrow \quad \left(Y(r), \left(Y^{(r)}(\tau_i), i = 1, 2, \ldots\right)\right)\end{aligned} \tag{3.4}$$

in finite-dimensional distributions, as $m \to \infty$. We extend the process $Y^{(r)}$ to the entire positive half-line by setting

$$Y^{(r)}(t) = \frac{1}{2}\left(\lim_{\tau \uparrow t,\ \text{rational}} Y^{(r)}(\tau) + \lim_{\tau \downarrow t,\ \text{rational}} Y^{(r)}(\tau)\right), \qquad t \ge 0.$$

The continuity in probability implies that this process is a version of $(Y(t), t \ge 0)$. This continuity in probability, (3.4) and monotonicity imply that as $m \to \infty$,

$$\left(\frac{1}{a_{n_m}} M_{\lfloor n_m r \rfloor}, \left(\frac{1}{a_{n_m}} \max_{n_m r < j \le n_m(t + r)} X_j, t \ge 0\right)\right) \quad \Rightarrow \quad \left(Y(r), \left(Y^{(r)}(t), t \ge 0\right)\right) \tag{3.5}$$

in finite-dimensional distributions. Now the stationarity of max-increments follows from (3.3), (3.5) and continuous mapping theorem. □

Remark 3.3. Self-similar processes with stationary max-increments arising in a functional maxima scheme (1.2) are close in spirit to the stationary self-similar extremal processes of O'Brien *et al*. [22], while extremal processes themselves are defined as random sup measures. A random sup measure is, as its name implies, indexed by sets. They also arise in a limiting maxima scheme similar to (1.2), but with a stronger notion of convergence. Every stationary self-similar extremal processes trivially produces a self-similar process with stationary max-increments via restriction to sets of the type $[0, t]$ for $t \ge 0$, but the connection between the two objects remains unclear. Our limiting process in Theorem 4.1 below can be extended to a stationary self-similar extremal processes, but the extension is highly nontrivial, and will not be pursued here.

It is not our goal in this paper to study in details the properties of self-similar processes with stationary max-increments, so we restrict ourselves to the following basic result.

Proposition 3.4. *Let $(Y(t), t \geq 0)$ be a nonnegative self-similar process with stationary max-increments, and exponent H of self-similarity. Suppose $(Y(t), t \geq 0)$ is not identically zero. Then $H \geq 0$, and the following statements hold.*

(a) *If $H = 0$, then $Y(t) = Y(1)$ a.s. for every $t > 0$.*
(b) *If $0 < EY(1)^p < \infty$ for some $p > 0$, then $H \leq 1/p$.*
(c) *If $H > 0$, $(Y(t), t \geq 0)$ is continuous in probability.*

Proof. By the stationarity of max-increments, $Y(t)$ is stochastically increasing with t. This implies that $H \geq 0$.

If $H = 0$, then $Y(n) \stackrel{d}{=} Y(1)$ for each $n = 1, 2 \ldots.$ We use (3.2) with $r = 1$. Using $t = 1$ we see that, in the right-hand side of (3.2), $Y(1) = Y^{(1)}(1)$ a.s. Since $Y^{(1)}(n) \geq Y^{(1)}(1)$ a.s., we conclude, using $t = n$ in the right-hand side of (3.2), that $Y(1) = Y^{(1)}(n)$ a.s. for each $n = 1, 2, \ldots.$ By monotonicity, we conclude that the process $(Y^{(1)}(t), t \geq 0)$, hence also the process $(Y(t), t \geq 0)$, is a.s. constant on $[1, \infty)$ and then, by self-similarity, also on $(0, \infty)$.

Next, let $p > 0$ be such that $0 < EY(1)^p < \infty$. It follows from (3.2) with $r = 1$ that

$$2^H Y(1) \stackrel{d}{=} Y(2) \stackrel{d}{=} \max\bigl(Y(1), Y^{(1)}(1)\bigr).$$

Therefore,

$$2^{pH} EY(1)^p = EY(2)^p = E\bigl[Y(1)^p \vee Y^{(1)}(1)^p\bigr] \leq 2EY(1)^p.$$

This means that $pH \leq 1$.

Finally, we take arbitrary $0 < s < t$. We use (3.2) with $r = s$. For every $\eta > 0$,

$$\begin{aligned} P\bigl(Y(t) - Y(s) > \eta\bigr) &= P\bigl(Y(s) \vee Y^{(s)}(t-s) - Y(s) > \eta\bigr) \\ &\leq P\bigl(Y^{(s)}(t-s) > \eta\bigr) = P\bigl((t-s)^H Y(1) > \eta\bigr). \end{aligned}$$

Hence, continuity in probability. □

We now introduce a crucial object for the subsequent discussion, which is the limiting process obtained in the main limit theorem of Section 4. It has a somewhat deceptively simple representation that we presently describe.

Let $\alpha > 0$, and consider the extremal Fréchet process $Z_\alpha(t), t \geq 0$, defined in (1.4), with the scale $\sigma = 1$. For $0 < \beta < 1$, we define a new stochastic process by

$$Z_{\alpha,\beta}(t) = Z_\alpha\bigl(t^\beta\bigr), \qquad t \geq 0. \tag{3.6}$$

We will refer to this process as the *time scaled extremal Fréchet process*.

The next proposition places this process in the general framework introduced earlier in this section.

Proposition 3.5. *The process $Z_{\alpha,\beta}$ in (3.6) is self-similar with $H = \beta/\alpha$ and has stationary max-increments.*

Proof. Since the extremal Fréchet process is self-similar with $H = 1/\alpha$, it is immediately seen that the process $Z_{\alpha,\beta}$ is self-similar with $H = \beta/\alpha$.

To show the stationarity of max-increments, we start with a useful representation of the extremal Fréchet process $Z_\alpha(t), t \ge 0$ in terms of the points of a Poisson random measure. Let $((j_k, s_k))$ be the points of a Poisson random measure on $\mathbb{R}_+^2$ with mean measure $\rho_\alpha \times \lambda$, where $\rho_\alpha(x,\infty) = x^{-\alpha}$, $x > 0$ and λ is the Lebesgue measure on $\mathbb{R}_+$. Then an elementary calculation shows that

$$\big(Z_\alpha(t), t \ge 0\big) \stackrel{d}{=} \big(\sup\{j_k : s_k \le t\}, t \ge 0\big).$$

Therefore, $(Z_{\alpha,\beta}(t), t \ge 0) \stackrel{d}{=} (U_{\alpha,\beta}(t), t \ge 0)$, where

$$U_{\alpha,\beta}(t) = \sup\{j_k : s_k \le t^\beta\}, \qquad t \ge 0. \tag{3.7}$$

Given $r > 0$, we define

$$U_{\alpha,\beta}^{(r)}(t) = \sup\{j_k : (t+r)^\beta - t^\beta \le s_k \le (t+r)^\beta\}.$$

Since $0 < \beta < 1$, we have

$$\big((t_1+r)^\beta - t_1^\beta, (t_1+r)^\beta\big) \subset \big((t_2+r)^\beta - t_2^\beta, (t_2+r)^\beta\big)$$

for $0 \le t_1 < t_2$. The nested nature of these sets implies that

$$\big(U_{\alpha,\beta}^{(r)}(t), t \ge 0\big) \stackrel{d}{=} \big(U_{\alpha,\beta}(t), t \ge 0\big),$$

because only the obvious equality of the one-dimensional distributions must be checked. Furthermore, since $(t+r)^\beta - t^\beta \le r^\beta$, we see that

$$U_{\alpha,\beta}(t+r) = U_{\alpha,\beta}(r) \vee U_{\alpha,\beta}^{(r)}(t) \qquad \text{for all } t \ge 0.$$

This means that the process $U_{\alpha,\beta}$ has stationary max-increments and, hence, so does the process $Z_{\alpha,\beta}$. □

Note that the max-increment process $(U_{\alpha,\beta}^{(r)}(t))$ in the proof of Proposition 3.5 is not independent of the random variable $U_{\alpha,\beta}(r)$ if $\beta < 1$. The case $\beta = 1$ corresponds to the extremal Fréchet process, whose max-increments are both stationary and independent.

It is interesting to note that, by part (b) of Proposition 3.4, any H-self-similar process with stationary max-increments and α-Fréchet marginals, must satisfy $H \le 1/\alpha$. The exponent $H = \beta/\alpha$ with $0 < \beta \le 1$ of the process $Z_{\alpha,\beta}$ (with $\beta = 1$ corresponding to the extremal Fréchet process Z_α) covers the entire interval $(0, 1/\alpha]$. Therefore, the upper bound of part (b) of Proposition 3.4 is, in general, the best possible.

We finish this section by mentioning that an immediate conclusion from (3.7) is the following representation of the time scaled extremal Fréchet process $Z_{\alpha,\beta}$ on the interval $[0, 1]$:

$$\big(Z_{\alpha,\beta}(t), 0 \le t \le 1\big) \stackrel{d}{=} \left(\bigvee_{j=1}^{\infty} \Gamma_j^{-1/\alpha} \mathbf{1}_{\{V_j \le t\}}, 0 \le t \le 1\right), \tag{3.8}$$

where Γ_j, $j = 1, 2, \ldots$, are arrival times of a unit rate Poisson process on $(0, \infty)$, and (V_j) are i.i.d. random variables with $P(V_1 \le x) = x^\beta$, $0 < x \le 1$, independent of (Γ_j).

4. A functional limit theorem for partial maxima

In this section, we state and prove our main result, a functional limit theorem for the partial maxima of the discrete-time stationary process $\mathbf{X} = (X_1, X_2, \ldots)$ given in (1.5). Recall that T is a conservative, ergodic and measure preserving map on a σ-finite, infinite, measure space $(E, \mathcal{E}, \mu)$. We will assume that T is a pointwise dual ergodic map with normalizing sequence (a_n) that is regularly varying with exponent $1 - \beta$; equivalently, the wandering sequence (w_n) in (2.5) is assumed to be regularly varying with exponent β. Crucially, we will assume that $1/2 < \beta < 1$. See Remark 4.3 after the proof of Theorem 4.1 below.

Define

$$b_n = \left(\int_E \max_{1 \le k \le n} \left| f \circ T^n(x) \right|^\alpha \mu(\mathrm{d}x) \right)^{1/\alpha}, \qquad n = 1, 2, \ldots. \tag{4.1}$$

The sequence (b_n) is known to play an important role in the rate of growth of partial maxima of an α-stable process of the type (1.5). It also turns out to be a proper normalizing sequence for our functional limit theorem. In Samorodnitsky [33] it was shown that, for a canonical kernel (1.7), if the map T is conservative, then the sequence (b_n) grows at a rate strictly slower than $n^{1/\alpha}$. The extra assumptions imposed in the current paper will guarantee a more precise statement. We will prove that, in fact, $(b_n) \in RV_{\beta/\alpha}$ and, more specifically,

$$\lim_{n \to \infty} \frac{b_n^\alpha}{w_n} = \| f \|_\infty \tag{4.2}$$

(where (w_n) is the wandering sequence). This fact has an interesting message, because it explicitly shows that the rate of growth of the partial maxima is determined both by the heaviness of the marginal tails (through α) and by the length of memory (through β). Such a precise measure of the length of memory is not present in Samorodnitsky [33].

In contrast, if the map T has a nontrivial dissipative component, then the sequence (b_n) grows at the rate $n^{1/\alpha}$, and so do the partial maxima of the stationary SαS process; see Samorodnitsky [33]. This is the limiting case of the setup in the present paper, as β gets closer to 1. Intuitively, the smaller is β, the longer is the memory in the process.

The basic idea in the proof of our main result, Theorem 4.1 below, is similar to the idea in the proof of Theorems 3.1 and 4.1 in Samorodnitsky [33] and is based on a Poisson representation of the process and a "single jump" property; see Remark 4.3.

We recall the tail constant of an α-stable random variable given by

$$C_\alpha = \left(\int_0^\infty x^{-\alpha} \sin x \, \mathrm{d}x \right)^{-1} = \begin{cases} (1 - \alpha)/\big(\Gamma(2 - \alpha) \cos(\pi\alpha/2)\big) & \text{if } \alpha \ne 1, \\ 2/\pi & \text{if } \alpha = 1; \end{cases}$$

see Samorodnitsky and Taqqu [36].

Theorem 4.1. *Let T be a conservative, ergodic and measure preserving map on a σ-finite infinite measure space $(E, \mathcal{E}, \mu)$. Assume that T is a pointwise dual ergodic map with normalizing sequence $(a_n) \in RV_{1-\beta}$, $0 \le \beta \le 1$. Let $f \in L^\alpha(\mu) \cap L^\infty(\mu)$, and assume that f is supported by a uniform set A for T, that is, $\operatorname{supp}(f) \subset A$. Let $\alpha > 0$. Then the sequence (b_n) in (4.1) satisfies (4.2).*

Assume now that $0 < \alpha < 2$ and $1/2 < \beta < 1$. If M is a $S\alpha S$ random measure on $(E, \mathcal{E})$ with control measure μ, then the stationary $S\alpha S$ process $\mathbf{X}$ given in (1.5) satisfies

$$\left(\frac{1}{b_n} \max_{1 \le k \le \lfloor nt \rfloor} |X_k|, t \ge 0\right) \Rightarrow \left(C_\alpha^{1/\alpha} Z_{\alpha,\beta}(t), t \ge 0\right) \quad \text{in } D[0,\infty) \tag{4.3}$$

in the Skorohod M_1-topology. Moreover, if $f = \mathbf{1}_A$, then the above convergence occurs in the Skorohod J_1-topology as well.

Remark 4.2. The functional limit theorem in Theorem 4.1 above, once again, involves weak convergence in two different topologies, that is, the Skorohod J_1-topology and the Skorohod M_1-topology. The issue is similar to that in Proposition 2.1; see Remark 2.2.

Proof of Theorem 4.1. We start with verifying (4.2). Obviously,

$$b_n^\alpha \le \|f\|_\infty \mu(\varphi_A \le n),$$

and, recalling that $w_n \sim \mu(\varphi_A \le n)$, we get the upper bound

$$\limsup_{n\to\infty} \frac{b_n^\alpha}{w_n} \le \|f\|_\infty.$$

On the other hand, take an arbitrary $\epsilon \in (0, \|f\|_\infty)$. The set

$$B_\epsilon = \{x \in A : |f(x)| \ge \|f\|_\infty - \epsilon\}$$

is a uniform set for T. A lower bound for b_n^α is obtained via the obvious inequality

$$b_n^\alpha \ge (\|f\|_\infty - \epsilon)\mu\left(\bigcup_{j=1}^n T^{-j} B_\epsilon\right).$$

Indeed, let $(w_n^{(\epsilon)})$ be the corresponding wandering rate sequence to the set B_ϵ. As argued in Proposition 2.1, we know that $w_n \sim w_n^{(\epsilon)} \sim \mu(\varphi_{B_\epsilon} \le n)$. Therefore,

$$\liminf_{n\to\infty} \frac{b_n^\alpha}{w_n} = \liminf_{n\to\infty} \frac{b_n^\alpha}{\mu(\varphi_{B_\epsilon} \le n)} \ge \|f\|_\infty - \epsilon.$$

Letting $\epsilon \to 0$, we obtain (4.2).

Suppose now that $0 < \alpha < 2$ and $1/2 < \beta < 1$. We continue with proving convergence in the finite-dimensional distributions in (4.3). Since for random elements in $D[0,\infty)$ with nondecreasing sample paths, weak convergence in the M_1-topology is implied by the finite-dimensional weak convergence, this will also establish (4.3) in the sense of weak convergence in the M_1-topology.

Fix $0 = t_0 < t_1 < \cdots < t_d$, $d \ge 1$. We may and will assume that $t_d \le 1$. We use a series representation of the random vector $(X_1, \dots, X_n)$: with $f_k = f \circ T^k$, $k = 1, 2, \dots$,

$$(X_k, k = 1, \dots, n) \stackrel{d}{=} \left(b_n C_\alpha^{1/\alpha} \sum_{j=1}^{\infty} \epsilon_j \Gamma_j^{-1/\alpha} \frac{f_k(U_j^{(n)})}{\max_{1 \le i \le n} |f_i(U_j^{(n)})|}, k = 1, \dots, n \right). \tag{4.4}$$

Here (ϵ_j) are i.i.d. Rademacher random variables (symmetric random variables with values ± 1), (Γ_j) are the arrival times of a unit rate Poisson process on $(0, \infty)$, and $(U_j^{(n)})$ are i.i.d. E-valued random variables with the common law η_n defined by

$$\frac{d\eta_n}{d\mu}(x) = \frac{1}{b_n^\alpha} \max_{1 \le k \le n} \left| f_k(x) \right|^\alpha, \qquad x \in E. \tag{4.5}$$

The sequences (ϵ_j), (Γ_j), and $(U_j^{(n)})$ are taken to be independent. We refer to Section 3.10 of Samorodnitsky and Taqqu [36] for series representations of α-stable random vectors. The representation (4.4) was also used in Samorodnitsky [33], and the argument below is structured similarly to the corresponding argument ibid.

The crucial consequence of the assumption $1/2 < \beta < 1$ is that, in the series representation (4.4), only the largest Poisson jump will play an important role. It is shown in Samorodnitsky [33] that, under the assumptions of Theorem 4.1, for every $\eta > 0$,

$$\varphi_n(\eta) \equiv P\left(\bigcup_{k=1}^{n} \left\{ \Gamma_j^{-1/\alpha} \frac{|f_k(U_j^{(n)})|}{\max_{1 \le i \le n} |f_i(U_j^{(n)})|} > \eta \right. \right. \\ \left. \left. \text{for at least 2 different } j = 1, 2, \dots \right\} \right) \to 0 \tag{4.6}$$

as $n \to \infty$.

We will proceed in two steps. First, we will prove that

$$\left(\bigvee_{j=1}^{\infty} \Gamma_j^{-1/\alpha} \frac{\max_{1 \le k \le \lfloor nt_i \rfloor} |f_k(U_j^{(n)})|}{\max_{1 \le k \le n} |f_k(U_j^{(n)})|}, i = 1, \dots, d \right) \\ \Rightarrow \quad \left(Z_{\alpha,\beta}(t_i), i = 1, \dots, d \right) \qquad \text{in } \mathbb{R}_+^d. \tag{4.7}$$

Next, we will prove that, for fixed $\lambda_1, \dots, \lambda_d > 0$, for every $0 < \delta < 1$,

$$P\left(b_n^{-1} \max_{1 \le k \le \lfloor nt_i \rfloor} |X_k| > \lambda_i, i = 1, \dots, d \right) \\ \le P\left(C_\alpha^{1/\alpha} \bigvee_{j=1}^{\infty} \Gamma_j^{-1/\alpha} \frac{\max_{1 \le k \le \lfloor nt_i \rfloor} |f_k(U_j^{(n)})|}{\max_{1 \le k \le n} |f_k(U_j^{(n)})|} > \lambda_i(1 - \delta), i = 1, \dots, d \right) + o(1) \tag{4.8}$$

and that

$$P\Big(b_n^{-1} \max_{1\le k\le \lfloor nt_i\rfloor} |X_k| > \lambda_i, i=1,\dots,d\Big) \tag{4.9}$$
$$\ge P\left(C_\alpha^{1/\alpha} \bigvee_{j=1}^{\infty} \Gamma_j^{-1/\alpha} \frac{\max_{1\le k\le \lfloor nt_i\rfloor} |f_k(U_j^{(n)})|}{\max_{1\le k\le n} |f_k(U_j^{(n)})|} > \lambda_i(1+\delta), i=1,\dots,d\right) + \mathrm{o}(1).$$

Since the Fréchet distribution is continuous, the weak convergence

$$\Big(b_n^{-1} \max_{1\le k\le \lfloor nt_i\rfloor} |X_k|, i=1,\dots,d\Big) \quad\Rightarrow\quad \big(Z_{\alpha,\beta}(t_i), i=1,\dots,d\big) \qquad \text{in } \mathbb{R}_+^d$$

will follow by taking δ arbitrarily small.

We start with proving (4.7). For $n=1,2,\dots$, $N_n = \sum_{j=1}^{\infty} \delta_{(\Gamma_j, U_j^{(n)})}$ is a Poisson random measure on $(0,\infty)\times \bigcup_{k=1}^n T^{-k}A$ with mean measure $\lambda\times\eta_n$. Define a map $S_n : \mathbb{R}_+ \times \bigcup_{k=1}^n T^{-k}A \to \mathbb{R}_+^d$ by

$$S_n(r,x) = r^{-1/\alpha}\big(M_n(f)(x)\big)^{-1}\big(M_{\lfloor nt_1\rfloor}(f)(x),\dots,M_{\lfloor nt_d\rfloor}(f)(x)\big), \qquad r>0, x\in \bigcup_{k=1}^n T^{-k}A.$$

Then, for $\lambda_1,\dots,\lambda_d > 0$,

$$\begin{aligned}
&P\left(\bigvee_{j=1}^{\infty} \Gamma_j^{-1/\alpha} \frac{\max_{1\le k\le \lfloor nt_i\rfloor} |f_k(U_j^{(n)})|}{\max_{1\le k\le n} |f_k(U_j^{(n)})|} \le \lambda_i, i=1,\dots,d\right)\\
&= P\big[N_n\big(S_n^{-1}\big((0,\lambda_1]\times\cdots\times(0,\lambda_d]\big)^c\big)=0\big]\\
&= \exp\big\{-(\lambda\times\eta_n)\big(S_n^{-1}\big((0,\lambda_1]\times\cdots\times(0,\lambda_d]\big)^c\big)\big\}\\
&= \exp\left\{-(\lambda\times\eta_n)\left\{(r,x): \bigvee_{j=1}^{d} \lambda_j^{-\alpha} \frac{(M_{\lfloor nt_j\rfloor}(f)(x))^\alpha}{(M_n(f)(x))^\alpha} > r\right\}\right\}\\
&= \exp\left\{-b_n^{-\alpha} \int_E \bigvee_{j=1}^{d} \lambda_j^{-\alpha} M_{\lfloor nt_j\rfloor}(f)^\alpha \,\mathrm{d}\mu\right\}.
\end{aligned}$$

We use (4.2) and the weak convergence in Proposition 2.1 to obtain

$$\begin{aligned}
b_n^{-\alpha} \int_E \bigvee_{j=1}^{d} \lambda_j^{-\alpha} M_{\lfloor nt_j\rfloor}(f)^\alpha \,\mathrm{d}\mu &\sim \|f\|_\infty^{-1} \int_E \bigvee_{j=1}^{d} \lambda_j^{-\alpha} M_{\lfloor nt_j\rfloor}(f)^\alpha \,\mathrm{d}\mu_n\\
\to \quad \int_{\Omega'} \bigvee_{j=1}^{d} \lambda_j^{-\alpha} \mathbf{1}_{\{V_\beta\le t_j\}} \,\mathrm{d}P' &= \sum_{i=1}^{d} \big(t_i^\beta - t_{i-1}^\beta\big)\left(\bigwedge_{j=i}^{d} \lambda_j\right)^{-\alpha}.
\end{aligned}$$

Therefore,

$$P\left(\bigvee_{j=1}^{\infty}\Gamma_j^{-1/\alpha}\frac{\max_{1\le k\le\lfloor nt_i\rfloor}|f_k(U_j^{(n)})|}{\max_{1\le k\le n}|f_k(U_j^{(n)})|}\le\lambda_i, i=1,\dots,d\right)$$

$$\to\quad\exp\left\{-\sum_{i=1}^{d}\big(t_i^{\beta}-t_{i-1}^{\beta}\big)\left(\bigwedge_{j=i}^{d}\lambda_j\right)^{-\alpha}\right\}=P\big(Z_{\alpha,\beta}(t_i)\le\lambda_i, i=1,\dots,d\big).$$

The claim (4.7) has, consequently, been proved.

We continue with the statements (4.8) and (4.9). Since the arguments are very similar, we only prove (4.8). Let $K\in\mathbb{N}$ and $0<\epsilon<1$ be constants so that

$$K+1>\frac{4}{\alpha}\quad\text{and}\quad\delta-\epsilon K>0.$$

Then

$$P\Big(b_n^{-1}\max_{1\le k\le\lfloor nt_i\rfloor}|X_k|>\lambda_i, i=1,\dots,d\Big)$$

$$\le P\left(C_\alpha^{1/\alpha}\bigvee_{j=1}^{\infty}\Gamma_j^{-1/\alpha}\frac{\max_{1\le k\le\lfloor nt_i\rfloor}|f_k(U_j^{(n)})|}{\max_{1\le k\le n}|f_k(U_j^{(n)})|}>\lambda_i(1-\delta), i=1,\dots,d\right)$$

$$+\varphi_n\Big(C_\alpha^{-1/\alpha}\epsilon\min_{1\le i\le d}\lambda_i\Big)+\sum_{i=1}^{d}\psi_n(\lambda_i,t_i),$$

where

$$\psi_n(\lambda,t)=P\left(C_\alpha^{1/\alpha}\max_{1\le k\le\lfloor nt\rfloor}\left|\sum_{j=1}^{\infty}\epsilon_j\Gamma_j^{-1/\alpha}\frac{f_k(U_j^{(n)})}{\max_{1\le i\le n}|f_i(U_j^{(n)})|}\right|>\lambda,\right.$$

$$C_\alpha^{1/\alpha}\bigvee_{j=1}^{\infty}\Gamma_j^{-1/\alpha}\frac{\max_{1\le k\le\lfloor nt\rfloor}|f_k(U_j^{(n)})|}{\max_{1\le k\le n}|f_k(U_j^{(n)})|}\le\lambda(1-\delta),\ \text{and for each } m=1,\dots,n,$$

$$\left.C_\alpha^{1/\alpha}\Gamma_j^{-1/\alpha}\frac{|f_m(U_j^{(n)})|}{\max_{1\le i\le n}|f_i(U_j^{(n)})|}>\epsilon\lambda\ \text{for at most one } j=1,2,\dots\right).$$

By (4.6), it is enough to show that

$$\psi_n(\lambda,t)\to 0 \tag{4.10}$$

for all $\lambda>0$ and $0\le t\le 1$.

For every $k=1,2,\dots,n$, the Poisson random measure represented by the points

$$\Big(\epsilon_j\Gamma_j^{-1/\alpha}f_k\big(U_j^{(n)}\big)\Big(\max_{1\le i\le n}\big|f_i\big(U_j^{(n)}\big)\big|\Big)^{-1}, j=1,2,\dots\Big)$$

has the same mean measure as that represented by the points

$$\left(\epsilon_j \Gamma_j^{-1/\alpha} \|f\|_\alpha b_n^{-1},\ j = 1, 2, \ldots\right),$$

where $\|f\|_\alpha = (\int_E |f|^\alpha \, d\mu)^{1/\alpha}$. In fact, the common mean measure assigns the value $x^{-\alpha}\|f\|_\alpha^\alpha/2$ to the sets (x, ∞) and $(-\infty, -x)$ for every $x > 0$. Therefore, these two Poisson random measures coincide distributionally. We conclude that the probability in (4.10) is bounded by

$$\begin{aligned}
&\sum_{k=1}^{\lfloor nt \rfloor} P\Bigg(C_\alpha^{1/\alpha} \left| \sum_{j=1}^{\infty} \epsilon_j \Gamma_j^{-1/\alpha} \frac{f_k(U_j^{(n)})}{\max_{1\le i\le n} |f_i(U_j^{(n)})|} \right| > \lambda, \\
&\qquad C_\alpha^{1/\alpha} \bigvee_{j=1}^{\infty} \Gamma_j^{-1/\alpha} \frac{f_k(U_j^{(n)})}{\max_{1\le i\le n} |f_i(U_j^{(n)})|} \le \lambda(1-\delta), \\
&\qquad C_\alpha^{1/\alpha} \Gamma_j^{-1/\alpha} \frac{|f_k(U_j^{(n)})|}{\max_{1\le i\le n} |f_i(U_j^{(n)})|} > \epsilon\lambda \text{ for at most one } j = 1, 2, \ldots \Bigg) \\
&\quad = \lfloor nt \rfloor P\Bigg(C_\alpha^{1/\alpha} \left| \sum_{j=1}^{\infty} \epsilon_j \Gamma_j^{-1/\alpha} \right| > \lambda \|f\|_\alpha^{-1} b_n, C_\alpha^{1/\alpha} \bigvee_{j=1}^{\infty} \Gamma_j^{-1/\alpha} \le \lambda(1-\delta)\|f\|_\alpha^{-1} b_n, \\
&\qquad\qquad C_\alpha^{1/\alpha} \Gamma_j^{-1/\alpha} > \epsilon\lambda \|f\|_\alpha^{-1} b_n \text{ for at most one } j = 1, 2, \ldots \Bigg) \\
&\quad \le n P\left(C_\alpha^{1/\alpha} \left| \sum_{j=K+1}^{\infty} \epsilon_j \Gamma_j^{-1/\alpha} \right| > (\delta - \epsilon K)\lambda \|f\|_\alpha^{-1} b_n \right) \\
&\quad \le \frac{n \|f\|_\alpha^4 C_\alpha^{4/\alpha}}{(\delta - \epsilon K)^4 \lambda^4 b_n^4} E \left| \sum_{j=K+1}^{\infty} \epsilon_j \Gamma_j^{-1/\alpha} \right|^4.
\end{aligned}$$

Due to the choice $K + 1 > 4/\alpha$,

$$E \left| \sum_{j=K+1}^{\infty} \epsilon_j \Gamma_j^{-1/\alpha} \right|^4 < \infty;$$

see Samorodnitsky [33] for a detailed proof. Since $n/b_n^4 \to 0$ as $n \to \infty$, (4.10) follows.

Suppose now that $f = \mathbf{1}_A$. In that case, the probability measure η_n defined in (4.5) coincides with the probability measure μ_n of Proposition 2.1. In order to prove weak convergence in the J_1-topology, we will use a truncation argument. We may and will restrict ourselves to the space $D[0, 1]$. Let $K = 1, 2, \ldots$. First of all, we show, in the notation of (3.8), the convergence

$$\begin{aligned}&\left(C_\alpha^{1/\alpha}\max_{1\le k\le\lfloor nt\rfloor}\left|\sum_{j=1}^{K}\epsilon_j\Gamma_j^{-1/\alpha}\mathbf{1}_A\circ T^k\big(U_j^{(n)}\big)\right|, 0\le t\le 1\right)\\&\quad\Rightarrow\quad\left(C_\alpha^{1/\alpha}\bigvee_{j=1}^{K}\Gamma_j^{-1/\alpha}\mathbf{1}_{\{V_j\le t\}}, 0\le t\le 1\right)\end{aligned}\tag{4.11}$$

in the J_1-topology on $D[0,1]$. Indeed, by (4.6), outside of an event of asymptotically vanishing probability, the process in the left-hand side of (4.11) is

$$\left(C_\alpha^{1/\alpha}\bigvee_{j=1}^{K}\Gamma_j^{-1/\alpha}\max_{1\le k\le\lfloor nt\rfloor}\mathbf{1}_A\circ T^k\big(U_j^{(n)}\big), 0\le t\le 1\right).\tag{4.12}$$

By Proposition 2.1, we can put all the random variables involved on the same probability space so that the time of the single step in the jth term in (4.12) converges a.s. for each $j=1,\dots,K$ to V_j. Then, trivially, the process in (4.12) converges a.s. in the J_1-topology on $D[0,1]$ to the process in the right-hand side of (4.11). Therefore, the weak convergence in (4.11) follows.

Next, we note that in the J_1-topology on the space $D[0,1]$,

$$\begin{aligned}&\left(C_\alpha^{1/\alpha}\bigvee_{j=1}^{K}\Gamma_j^{-1/\alpha}\mathbf{1}_{\{V_j\le t\}}, 0\le t\le 1\right)\\&\quad\to\quad\left(C_\alpha^{1/\alpha}\bigvee_{j=1}^{\infty}\Gamma_j^{-1/\alpha}\mathbf{1}_{\{V_j\le t_i\}}0\le t\le 1\right)\qquad\text{as } K\to\infty \text{ a.s.}\end{aligned}$$

This is so because, as $K\to\infty$,

$$\begin{aligned}&\sup_{0\le t\le 1}\left(\bigvee_{j=1}^{\infty}\Gamma_j^{-1/\alpha}\mathbf{1}_{\{V_j\le t\}}-\bigvee_{j=1}^{K}\Gamma_j^{-1/\alpha}\mathbf{1}_{\{V_j\le t\}}\right)\\&\quad\le\Gamma_{K+1}^{-1/\alpha}\to 0\qquad\text{a.s.}\end{aligned}$$

According to Theorem 3.2 in Billingsley [4], the J_1-convergence in (4.3) will follow once we show that

$$\lim_{K\to\infty}\limsup_{n\to\infty}P\left(\max_{1\le k\le n}\left|\sum_{j=K+1}^{\infty}\epsilon_j\Gamma_j^{-1/\alpha}\mathbf{1}_A\circ T^k\big(U_j^{(n)}\big)\right|>\epsilon\right)=0$$

for every $\epsilon > 0$. Write

$$P\left(\max_{1\le k\le n}\left|\sum_{j=K+1}^{\infty}\epsilon_j\Gamma_j^{-1/\alpha}\mathbf{1}_A\circ T^k\big(U_j^{(n)}\big)\right|>\epsilon\right)$$

$$\le \int_0^{(\epsilon/2)^{-\alpha}} \mathrm{e}^{-x}\frac{x^{K-1}}{(K-1)!}\,\mathrm{d}x$$

$$+\int_{(\epsilon/2)^{-\alpha}}^{\infty} \mathrm{e}^{-x}\frac{x^{K-1}}{(K-1)!}$$

$$\times P\left(\max_{1\le k\le n}\left|\sum_{j=1}^{\infty}\epsilon_j(\Gamma_j+x)^{-1/\alpha}\mathbf{1}_A\circ T^k\big(U_j^{(n)}\big)\right|>\epsilon\right)\mathrm{d}x.$$

Clearly, the first term vanishes when $K\to\infty$. Therefore, it is sufficient to show that for every $x\ge(\epsilon/2)^{-\alpha}$,

$$P\left(\max_{1\le k\le n}\left|\sum_{j=1}^{\infty}\epsilon_j(\Gamma_j+x)^{-1/\alpha}\mathbf{1}_A\circ T^k\big(U_j^{(n)}\big)\right|>\epsilon\right)\to 0 \tag{4.13}$$

as $n\to\infty$.

To this end, choose $L\in\mathbb{N}$ and $0<\xi<1/2$ so that

$$L+1>\frac{4}{\alpha}\quad\text{and}\quad\frac{1}{2}-\xi L>0. \tag{4.14}$$

By (4.6), we can write

$$P\left(\max_{1\le k\le n}\left|\sum_{j=1}^{\infty}\epsilon_j(\Gamma_j+x)^{-1/\alpha}\mathbf{1}_A\circ T^k\big(U_j^{(n)}\big)\right|>\epsilon\right)$$

$$\le P\left(\max_{1\le k\le n}\left|\sum_{j=1}^{\infty}\epsilon_j(\Gamma_j+x)^{-1/\alpha}\mathbf{1}_A\circ T^k\big(U_j^{(n)}\big)\right|>\epsilon,\ \text{and for each } m=1,\dots,n, \right. \tag{4.15}$$

$$\left.(\Gamma_j+x)^{-1/\alpha}\mathbf{1}_A\circ T^m\big(U_j^{(n)}\big)>\xi\epsilon \text{ for at most one } j=1,2,\dots\right)+\mathrm{o}(1).$$

Notice that for every $k=1,\dots,n$, the Poisson random measure represented by the points

$$\big(\epsilon_j(\Gamma_j+x)^{-1/\alpha}\mathbf{1}_A\circ T^k\big(U_j^{(n)}\big),\ j=1,2,\dots\big)$$

is distributionally equal to the Poisson random measure represented by the points

$$\big(\epsilon_j\big(b_n^\alpha\mu(A)^{-1}\Gamma_j+x\big)^{-1/\alpha},\ j=1,2,\dots\big).$$

Therefore, the first term on the right-hand side of (4.15) can be bounded by

$$\sum_{k=1}^{n} P\Bigg(\Bigg|\sum_{j=1}^{\infty} \epsilon_j (\Gamma_j + x)^{-1/\alpha} \mathbf{1}_A \circ T^k\big(U_j^{(n)}\big)\Bigg| > \epsilon,$$

$$(\Gamma_j + x)^{-1/\alpha} \mathbf{1}_A \circ T^k\big(U_j^{(n)}\big) > \xi\epsilon \text{ for at most one } j = 1, 2, \ldots\Bigg)$$

$$= nP\Bigg(\Bigg|\sum_{j=1}^{\infty} \epsilon_j \big(b_n^{\alpha}\mu(A)^{-1}\Gamma_j + x\big)^{-1/\alpha}\Bigg| > \epsilon,$$

$$\big(b_n^{\alpha}\mu(A)^{-1}\Gamma_j + x\big)^{-1/\alpha} > \xi\epsilon \text{ for at most one } j = 1, 2, \ldots\Bigg)$$

$$\le nP\Bigg(\Bigg|\sum_{j=L+1}^{\infty} \epsilon_j \big(b_n^{\alpha}\mu(A)^{-1}\Gamma_j + x\big)^{-1/\alpha}\Bigg| > \bigg(\frac{1}{2} - \xi L\bigg)\epsilon\Bigg).$$

In the last step we used the fact that, for $x \ge (\epsilon/2)^{-\alpha}$, the magnitude of each term in the infinite sum does not exceed $\epsilon/2$. By the contraction inequality for Rademacher series (see, e.g., Proposition 1.2.1 of Kwapień and Woyczyński [16]),

$$nP\Bigg(\Bigg|\sum_{j=L+1}^{\infty} \epsilon_j \big(b_n^{\alpha}\mu(A)^{-1}\Gamma_j + x\big)^{-1/\alpha}\Bigg| > \bigg(\frac{1}{2} - \xi L\bigg)\epsilon\Bigg)$$

$$\le 2nP\Bigg(\Bigg|\sum_{j=L+1}^{\infty} \epsilon_j \Gamma_j^{-1/\alpha}\Bigg| > \bigg(\frac{1}{2} - \xi L\bigg)\epsilon\mu(A)^{-1/\alpha} b_n\Bigg).$$

As before, by Markov's inequality and using the constraints of the constants $L \in \mathbb{N}$ and $0 < \xi < 1/2$ given in (4.14),

$$2nP\Bigg(\Bigg|\sum_{j=L+1}^{\infty} \epsilon_j \Gamma_j^{-1/\alpha}\Bigg| > \bigg(\frac{1}{2} - \xi L\bigg)\epsilon\mu(A)^{-1/\alpha} b_n\Bigg)$$

$$\le \frac{2n\mu(A)^{4/\alpha}}{(2^{-1} - \xi L)^4 \epsilon^4 b_n^4} E\Bigg|\sum_{j=L+1}^{\infty} \epsilon_j \Gamma_j^{-1/\alpha}\Bigg|^4 \to 0$$

as $n \to \infty$ and, hence, (4.13) follows. □

Remark 4.3. The crucial point in the proof of the theorem is the "single Poisson jump property" (4.6) that shows that, essentially, a single Poisson point of the type $\Gamma_j^{-1/\alpha}$ plays the decisive role in determining the size of a partial maximum. This enabled us to show that the normalized partial maxima converge to the first Poisson point $\Gamma_1^{-1/\alpha}$ (which, of course, has exactly the standard α-Fréchet law). We can guarantee the "single Poisson jump property" in the case $1/2 < \beta < 1$. On the other hand, in the range $0 < \beta < 1/2$, the condition (4.6) is no longer valid. We believe that the limiting process will involve a finite, but random, number of the Poisson points of the type $\Gamma_j^{-1/\alpha}$. This will preclude a limiting Fréchet law. The details of this are still being worked out, and will appear in a future work. In the boundary case $\beta = 1/2$, the statement (4.3) still holds under certain additional conditions. This is the case, for example, for the Markov shift operators presented at the end of the paper. See also Example 5.3 in Samorodnitsky [33].

Remark 4.4. There is no doubt that the convergence result in Theorem 4.1 can be extended to more general infinitely divisible random measures M in (1.5), under appropriate assumptions of regular variation of the Lévy measure of M and integrability of the function f. In particular, regardless of the size of $\alpha > 0$, the time scaled extremal Fréchet processes $Z_{\alpha,\beta}$ are likely to appear in the limit in (4.3). Furthermore, the symmetry of the process $\mathbf{X}$ has very little to do with the limiting distribution of the partial maxima. For example, a straightforward symmetrization argument allows one to extend (4.3) to skewed α-stable processes, at least in the sense of convergence of finite-dimensional distributions. The reason we decided to restrict the presentation to the symmetric stable case had to do with a particularly simple form of the series representation (4.4) available in this case. This has allowed us to avoid certain technicalities that might have otherwise blurred the main message, which is the effect of memory on the functional limit theorem for the partial maxima.

One can obtain concrete examples of the situations in which the result of Theorem 4.1 applies by taking, for instance, one of the variety of pointwise dual ergodic operators provided in Aaronson [1] and Zweimüller [43], and embedding them into the integral form of stationary SαS processes. We conclude the current paper by mentioning the example of a flow generated by a null recurrent Markov chain. This example appears in Samorodnitsky [33], Owada and Samorodnitsky [24], and Owada [23] as well.

Consider an irreducible null recurrent Markov chain $(x_n, n \geq 0)$ defined on an infinite countable state space $\mathbb{S}$ with the transition matrix (p_{ij}). Let $(\pi_i, i \in \mathbb{S})$ be its unique (up to constant multiplication) invariant measure with $\pi_{i_0} = 1$ for some fixed state $i_0 \in \mathbb{S}$. Note that (π_i) is necessarily an infinite measure. Define a σ-finite and infinite measure on $(E, \mathcal{E}) = (\mathbb{S}^{\mathbb{N}}, \mathcal{B}(\mathbb{S}^{\mathbb{N}}))$ by

$$\mu(B) = \sum_{i \in \mathbb{S}} \pi_i P_i(B), \qquad B \subseteq \mathbb{S}^{\mathbb{N}},$$

where $P_i(\cdot)$ denotes the probability law of (x_n) starting in state $i \in \mathbb{S}$. Let

$$T(x_0, x_1, \ldots) = (x_1, x_2, \ldots)$$

be the usual left shift operator on $\mathbb{S}^{\mathbb{N}}$. Then T preserves μ. Since the Markov chain is irreducible and null recurrent, T is conservative and ergodic (see Harris and Robbins [12]).

We consider the set $A = \{x \in \mathbb{S}^{\mathbb{N}} : x_0 = i_0\}$ with the fixed state $i_0 \in \mathbb{S}$ chosen above. Since

$$\widehat{T}^k \mathbf{1}_A(x) = P_{i_0}(x_k = i_0) \qquad \text{for } x \in A$$

is constant on A (see Section 4.5 in Aaronson [2]), we can choose as the normalizing sequence $a_n = \sum_{k=1}^n P_{i_0}(x_k = i_0)$, and see that the expression $a_n^{-1} \sum_{k=1}^n \widehat{T}^k \mathbf{1}_A(x)$ is identically equal to $1 = \mu(A)$ on A. Therefore, the map T is pointwise dual ergodic, and the Darling–Kac set condition, in fact, reduces to a simple identity. Let

$$\varphi_A(x) = \min\{n \geq 1 : x_n \in A\}, \qquad x \in \mathbb{S}^{\mathbb{N}}$$

be the first entrance time, and assume that

$$\sum_{k=1}^n P_{i_0}(\varphi_A \geq k) \in RV_\beta \tag{4.16}$$

for some $\beta \in (1/2, 1)$. Two equivalent conditions to (4.16) are given in Resnick *et al*. [26]. Note that the exponent of regular variation β controls how frequently the Markov chain returns to A. Since $\mu(\varphi_A = k) = P_{i_0}(\varphi_A \geq k)$ for $k \geq 1$ (see Lemma 3.3 in Resnick *et al*. [26]), we have

$$w_n \sim \mu(\varphi_A \leq n) \in RV_\beta.$$

Then all of the assumptions of Theorem 4.1 are satisfied for any $f \in L^\alpha(\mu) \cap L^\infty(\mu)$, supported by A.

Acknowledgements

We are very grateful for the uncommonly detailed and useful comments we have received from two anonymous referees and an anonymous Associate Editor. These comments helped us to introduce a number of improvements to the paper (this includes eliminating a serious error). This research was partially supported by the ARO Grants W911NF-07-1-0078 and W911NF-12-10385, NSF Grant DMS-1005903 and NSA Grant H98230-11-1-0154 at Cornell University.

References

[1] Aaronson, J. (1981). The asymptotic distributional behaviour of transformations preserving infinite measures. *J. Anal. Math.* **39** 203–234. MR0632462

[2] Aaronson, J. (1997). *An Introduction to Infinite Ergodic Theory*. *Mathematical Surveys and Monographs* **50**. Providence, RI: Amer. Math. Soc. MR1450400

[3] Avram, F. and Taqqu, M.S. (1992). Weak convergence of sums of moving averages in the α-stable domain of attraction. *Ann. Probab.* **20** 483–503. MR1143432

[4] Billingsley, P. (1999). *Convergence of Probability Measures*, 2nd ed. *Wiley Series in Probability and Statistics*: *Probability and Statistics*. New York: Wiley. MR1700749

[5] Davis, R. and Resnick, S. (1985). Limit theory for moving averages of random variables with regularly varying tail probabilities. *Ann. Probab.* **13** 179–195. MR0770636

[6] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory*. *An Introduction*. *Springer Series in Operations Research and Financial Engineering*. New York: Springer. MR2234156

[7] Dwass, M. (1964). Extremal processes. *Ann. Math. Statist.* **35** 1718–1725. MR0177440

[8] Dwass, M. (1966). Extremal processes. II. *Illinois J. Math.* **10** 381–391. MR0193661

[9] Embrechts, P. and Maejima, M. (2002). *Selfsimilar Processes*. *Princeton Series in Applied Mathematics*. Princeton, NJ: Princeton Univ. Press. MR1920153

[10] Fasen, V. (2005). Extremes of regularly varying Lévy-driven mixed moving average processes. *Adv. in Appl. Probab.* **37** 993–1014. MR2193993

[11] Fisher, R.A. and Tippett, L. (1928). Limiting forms of the frequency distributions of the largest or smallest member of a sample. *Math. Proc. Cambridge Philos. Soc.* **24** 180–190.

[12] Harris, T.E. and Robbins, H. (1953). Ergodic theory of Markov chains admitting an infinite invariant measure. *Proc. Natl. Acad. Sci. USA* **39** 860–864. MR0056873

[13] Kabluchko, Z., Schlather, M. and de Haan, L. (2009). Stationary max-stable fields associated to negative definite functions. *Ann. Probab.* **37** 2042–2065. MR2561440

[14] Kabluchko, Z. and Stoev, S. (2015). Stochastic integral representations and classification of sum- and max-infinitely divisible processes. *Bernoulli*. To appear. Available at arXiv:1207.4983.

[15] Krengel, U. (1985). *Ergodic Theorems*. *de Gruyter Studies in Mathematics* **6**. Berlin: de Gruyter. MR0797411

[16] Kwapień, S. and Woyczyński, W.A. (1992). *Random Series and Stochastic Integrals*: *Single and Multiple*. *Probability and Its Applications*. Boston, MA: Birkhäuser. MR1167198

[17] Lamperti, J. (1962). Semi-stable stochastic processes. *Trans. Amer. Math. Soc.* **104** 62–78. MR0138128

[18] Lamperti, J. (1964). On extreme order statistics. *Ann. Math. Statist.* **35** 1726–1737. MR0170371

[19] Leadbetter, M.R. (1983). Extremes and local dependence in stationary sequences. *Z. Wahrsch. Verw. Gebiete* **65** 291–306. MR0722133

[20] Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. *Springer Series in Statistics*. New York: Springer. MR0691492

[21] Mikosch, T. and Stărică, C. (2000). Limit theory for the sample autocorrelations and extremes of a GARCH(1, 1) process. *Ann. Statist.* **28** 1427–1451. MR1805791

[22] O'Brien, G.L., Torfs, P.J.J.F. and Vervaat, W. (1990). Stationary self-similar extremal processes. *Probab. Theory Related Fields* **87** 97–119. MR1076958

[23] Owada, T. (2013). Limit theory for the sample autocovariance for heavy tailed stationary infinitely divisible processes generated by conservative flows. Technical report. Available at arXiv:1302.0058.

[24] Owada, T. and Samorodnitsky, G. (2015). Functional central limit theorem for heavy tailed stationary ifinitely divisible processes generated by conservative flows. *Ann. Probab.* **43** 240–285. MR3298473

[25] Resnick, S. and Samorodnitsky, G. (2004). Point processes associated with stationary stable processes. *Stochastic Process. Appl.* **114** 191–209. MR2101240

[26] Resnick, S., Samorodnitsky, G. and Xue, F. (2000). Growth rates of sample covariances of stationary symmetric α-stable processes associated with null recurrent Markov chains. *Stochastic Process. Appl.* **85** 321–339. MR1731029

[27] Resnick, S.I. (1987). *Extreme Values*, *Regular Variation*, *and Point Processes*. *Applied Probability. A Series of the Applied Probability Trust* **4**. New York: Springer. MR0900810

[28] Resnick, S.I. and Rubinovitch, M. (1973). The structure of extremal processes. *Adv. in Appl. Probab.* **5** 287–307. MR0350867

[29] Rootzén, H. (1978). Extremes of moving averages of stable processes. *Ann. Probab.* **6** 847–869. MR0494450

[30] Rosiński, J. (1995). On the structure of stationary stable processes. *Ann. Probab.* **23** 1163–1187. MR1349166

[31] Rosiński, J. (2006). Minimal integral representations of stable processes. *Probab. Math. Statist.* **26** 121–142. MR2301892

[32] Roy, E. (2007). Ergodic properties of Poissonian ID processes. *Ann. Probab.* **35** 551–576. MR2308588

[33] Samorodnitsky, G. (2004). Extreme value theory, ergodic theory and the boundary between short memory and long memory for stationary stable processes. *Ann. Probab.* **32** 1438–1468. MR2060304
[34] Samorodnitsky, G. (2005). Null flows, positive flows and the structure of stationary symmetric stable processes. *Ann. Probab.* **33** 1782–1803. MR2165579
[35] Samorodnitsky, G. (2006). Long range dependence. *Found. Trends Stoch. Syst.* **1** 163–257. MR2379935
[36] Samorodnitsky, G. and Taqqu, M.S. (1994). *Stable Non-Gaussian Random Processes. Stochastic Modeling*. New York: Chapman and Hall. MR1280932
[37] Skorohod, A.V. (1956). Limit theorems for stochastic processes. *Teor. Veroyatn. Primen.* **1** 289–319. MR0084897
[38] Stoev, S.A. (2008). On the ergodicity and mixing of max-stable processes. *Stochastic Process. Appl.* **118** 1679–1705. MR2442375
[39] Stoev, S.A. and Taqqu, M.S. (2005). Extremal stochastic integrals: A parallel between max-stable processes and α-stable processes. *Extremes* **8** 237–266. MR2324891
[40] Surgailis, D., Rosiński, J., Mandrekar, V. and Cambanis, S. (1993). Stable mixed moving averages. *Probab. Theory Related Fields* **97** 543–558. MR1246979
[41] Wang, Y. and Stoev, S.A. (2010). On the structure and representations of max-stable processes. *Adv. in Appl. Probab.* **42** 855–877. MR2779562
[42] Whitt, W. (2002). *Stochastic-Process Limits. An Introduction to Stochastic-Process Limits and Their Application to Queues. Springer Series in Operations Research*. New York: Springer. MR1876437
[43] Zweimüller, R. (2009). Surrey notes on infinite ergodic theory. Lecture notes, Surrey Univ.

Received July 2013 and revised December 2013