Bernoulli **19**(4), 2013, 1419–1448
DOI: 10.3150/12-BEJSP08

Clustering of Markov chain exceedances

SIDNEY I. RESNICK[1] and DAVID ZEBER[2]

[1]*School of ORIE, Rhodes Hall 284, Cornell University, Ithaca, NY 14853, USA. E-mail: sir1@cornell.edu*
[2]*Department of Statistical Science, 301 Malott Hall, Cornell University, Ithaca, NY 14853, USA. E-mail: dsz5@cornell.edu*

The tail chain of a Markov chain can be used to model the dependence between extreme observations. For a positive recurrent Markov chain, the tail chain aids in describing the limit of a sequence of point processes $\{N_n, n \geq 1\}$, consisting of normalized observations plotted against scaled time points. Under fairly general conditions on extremal behaviour, $\{N_n\}$ converges to a cluster Poisson process. Our technique decomposes the sample path of the chain into i.i.d. regenerative cycles rather than using blocking argument typically employed in the context of stationarity with mixing.

1. Introduction

One of the effects of dependence in a time series is that extremes tend to cluster. This has applied implications to risk contagion over time but is also mathematically interesting and the challenge is to precisely relate the dependence structure to the clustering. For Markov dependence, how do we describe exceedance clusters?

Point processes powerfully describe extremal behaviour of certain time series. Under appropriate conditions on marginal distributions and rapid decay of dependence as a function of time lag for the process $\{X_j : j \geq 0\}$, the exceedance point process N_n defined by

$$N_n\big([0, s] \times (a, \infty]\big) = \#\{j \leq sn : X_j > ab_n\} \tag{1.1}$$

converges weakly to a Poisson limit as $n \to \infty$, where $b_n \to \infty$ is a threshold sequence. This leads to a number of results on asymptotic distributions of large order statistics and exceedances of an extreme level. Such results have been developed in a variety of contexts by [3,6,10–12,14,21]. More specific results exist for regularly varying processes [4,7], regenerative processes [1,27], and Markov chains [22,35]. Distributions of functionals of such point processes have been considered in [28,29,36].

For stationary processes, the dependence structure causes extremes to occur in clusters. The clustering is often summarized using the extremal index θ introduced by Leadbetter et al. [14], which is related to the asymptotic mean cluster size. To obtain a point process convergence result, authors often employ the big block/little block technique and mixing conditions, such as Leadbetter's $D(u_n)$ (see [14]), to split the process into approximately independent and identically distributed blocks. With an appropriate choice of block size, extremes within one such block belong asymptotically to the same cluster. Under an assumption controlling the extremal behaviour within each block, such as via the distribution of the number of exceedances, N_n generally converges to a limiting compound Poisson process, where the compounding at each time point approximates the clustering within each block. For Markov chains, the *tail chain* is an

1350-7265

asymptotic process that models behaviour upon reaching an extreme state; see [22,23,26,30–32]. Point process results for stationary Markov chains employ the tail chain to specify the compounding in the limit process. Under Markov dependence, the within-block behaviour is determined merely by conditions on the marginal distribution and the transition kernel. Basrak and Segers [4] extended the tail chain model to general multivariate regularly varying stationary processes.

Rootzén [27] focuses on regenerative processes, which split naturally into *cycles*. In this case, the within-block condition is replaced by an assumption on the extremal behaviour over a cycle. The main difference is that the cycles are of random but finite length, whereas the block size increases deterministically with n. In particular, Rootzén shows that convergence of the sequence of processes counting the number of exceedances depends on the asymptotics of the distribution of the cycle maximum as well as the marginal distribution.

We combine these two approaches to derive the weak limit of $\{N_n\}$ when $\{X_n\}$ is a positive recurrent Markov chain. Such chains display a regenerative structure and in the limit, N_n is approximated by a process consisting of clusters of points stacked above common time points, each corresponding to a separate regenerative cycle. The heights of the points in each cluster are determined by an independent run of the tail chain. This paper requires some distributional results for the tail chain process that were derived in [26] and Section 2 offers a summary of necessary facts. We focus on the case of heavy-tailed marginals, but believe our approach could be extended to accommodate more general marginal distributions.

1.1. Notation and conventions

We review notation and relevant concepts. In general, bold symbols represent vectors or sequences and for $\boldsymbol{x} = (x_1, x_2, \ldots)$, write $\boldsymbol{x}_m := (x_1, \ldots, x_m)$.

$f^{\leftarrow}$	the left-continuous inverse of a monotone function f, i.e., $f^{\leftarrow}(x) = \inf\{y : f(y) \geq x\}$.
RV_ρ	the class of regularly varying functions with index ρ.
$D[0,\infty)$	the space of real-valued càdlàg functions on $[0,\infty)$ endowed with the Skorohod topology.
$D_{\mathrm{left}}[0,\infty)$	left continuous functions on $[0,\infty)$ with finite right hand limits and metrized by the Skorohod metric.
$D^{\uparrow}[0,\infty)$	the subspace of $D_{\mathrm{left}}[0,\infty)$ consisting of non-decreasing functions f with $f(0) = 0$ and $\lim_{x\to\infty} f(x) = \infty$.
$\mathcal{K}(\mathbb{E})$	the collection of compact subsets of $\mathbb{E}$.
$\mathcal{C}(\mathbb{E})$	the space of real-valued continuous, bounded functions on $\mathbb{E}$.
$\mathcal{C}_K^+(\mathbb{E})$	the space of non-negative continuous functions on $\mathbb{E}$ with compact support.
$\mathbb{M}_+(\mathbb{E})$	the space of non-negative Radon measures on $\mathbb{E}$.
$\mathbb{M}_p(\mathbb{E})$	the space of Radon point measures on $\mathbb{E}$.
$\mathbb{LEB}$	Lebesgue measure on $\mathbb{R}$.
$\mathrm{PRM}(\mu)$	Poisson random measure on $\mathbb{E}$ with mean measure μ.
$\epsilon_x(\cdot)$	point measure at x, i.e., $\epsilon_x(A) = \mathbf{1}_A(x)$.
ν_α	a measure on $(0,\infty]$ given by $\nu_\alpha(x,\infty] = x^{-\alpha}$ for $x > 0, \alpha > 0$.
$\Rightarrow$	weak convergence of probability measures [5].

For a space $\mathbb{E}$ which is locally compact with countable base (for example, a subset of $[-\infty,\infty]^d$), a sequence of measures $\{\mu_n\} \subset \mathbb{M}_+(\mathbb{E})$ converges vaguely to $\mu \in \mathbb{M}_+(\mathbb{E})$ (written $\mu_n \xrightarrow{v} \mu$) if $\int_{\mathbb{E}} f\,d\mu_n \to \int_{\mathbb{E}} f\,d\mu$ as $n \to \infty$ for any $f \in \mathcal{C}_K^+(\mathbb{E})$. The vague topology on $\mathbb{M}_+(\mathbb{E})$ is metrizable by the vague metric, d_v, i.e., $d_v(\mu_n,\mu) \to 0$ iff $\mu_n \xrightarrow{v} \mu$. See [13,20,25] for further details. A distribution F on $[0,\infty)$ has a regularly varying tail with index $\alpha > 0$, denoted $1 - F \in \mathrm{RV}_{-\alpha}$, if there exists $b(t) \to \infty$ such that

$$tF\bigl(b(t)\cdot\bigr) \xrightarrow{v} \nu_\alpha(\cdot) \qquad \text{in } \mathbb{M}_+(0,\infty] \qquad \text{as } t \to \infty,$$

where $\nu_\alpha(x,\infty] = x^{-\alpha}$ for $x > 0$. The function $b(\cdot)$ is called a *scaling function*.

If $\boldsymbol{X} = (X_0, X_1, X_2, \ldots)$ is a (homogeneous) Markov chain and K is a Markov transition kernel, we write $\boldsymbol{X} \sim K$ to mean that the dependence structure of $\boldsymbol{X}$ is specified by K, i.e.,

$$\mathsf{P}[X_{n+1} \in \cdot \mid X_n = x] = K(x,\cdot), \qquad n = 0, 1, \ldots.$$

Also, $\mathsf{P}_\mu[\boldsymbol{X} \in \cdot]$ specifies the initial distribution $\mathsf{P}[X_0 \in \cdot] = \mu$ (abbreviate $\mathsf{P}_x := \mathsf{P}_{\epsilon_x}$), and E_μ denotes expectation with respect to P_μ.

2. Extremal component and tail chain approximation

Let $\boldsymbol{X} = (X_0, X_1, \ldots)$ be a Markov chain on $[0,\infty)$ with transition kernel K. The tail chain is a finite-dimensional approximation to the chain $\boldsymbol{X}$ used to study the limit of $\{N_n\}$ given by (1.1). Building on theory developed in [23,30,31], [26] presents the tail chain approximation in terms of a related process known as the *extremal component* of $\boldsymbol{X}$, an approach we follow here.

2.1. Tail chain approximation

Suppose the transition kernel K is in the domain of attraction of a distribution G (denoted $K \in D(G)$) which means [26],

$$K(t, t\cdot) \Rightarrow G(\cdot) \qquad \text{on } [0,\infty] \qquad \text{as } t \to \infty.$$

Taking $\xi_1, \xi_2, \ldots$ i.i.d. random variables with common distribution G, set $\xi(n) = \prod_{j=1}^n \xi_j$, $n \geq 1$ with $\xi(0) = 1$ and write $\boldsymbol{\xi} = \{\xi(n), n \geq 0\}$. The *tail chain associated with* G [22,26,31] is $\boldsymbol{T} = (T_0, T_1, \ldots)$ given by

$$T_n = T_0 \xi_1 \cdots \xi_n = T_0 \xi(n), \qquad n \geq 0. \tag{2.1}$$

Thus $\boldsymbol{T}$ is a multiplicative random walk and $\{0\}$ is an absorbing barrier for $\boldsymbol{T}$, accessible if $G(\{0\}) > 0$.

An *extremal boundary* for $\boldsymbol{X}$ is a function $y(t)$ satisfying $0 \leq y(t) \to 0$, such that

$$K\bigl(tu_t, t\bigl[0, y(t)\bigr]\bigr) \longrightarrow G\bigl(\{0\}\bigr) \qquad \text{as } t \to \infty, \tag{2.2}$$

for any non-negative function $u_t = u(t) \to u > 0$. Such a function always exists if $K \in D(G)$ [26, Section 3.2]. If $y(t)$ is an extremal boundary, any function $0 \le \tilde{y}(t) \to 0$ with $\tilde{y}(t) \ge y(t)$ for $t \ge t_0$ is also an extremal boundary. If $G(\{0\}) = 0$, then $y(t) \equiv 0$ is a convenient choice. Given an extremal boundary for K, the *extremal component* of $\boldsymbol{X}$ is the process $\boldsymbol{X}$ prior to $\boldsymbol{X}$ crossing below the scaled extremal boundary and identically $\mathbf{0}$ afterwards. The first downcrossing occurs at time

$$\tau(t) = \inf\bigl\{n \ge 0 : X_n \le ty(t)\bigr\}, \tag{2.3}$$

and the extremal component is the process $\boldsymbol{X}^{(t)} = (X_0^{(t)}, X_1^{(t)}, \ldots)$ defined by

$$X_n^{(t)} = X_n \cdot \mathbf{1}_{\{n<\tau(t)\}}, \qquad n = 0, 1, \ldots.$$

Starting from an extreme level $X_0 = t$, the extremal boundary separates extreme states from non-extreme states for the scaled process $t^{-1}\boldsymbol{X}$.

The *tail chain approximation* is the following [26, Theorem 3.3]. For $K \in D(G)$, $u > 0$, $m \ge 1$,

$$\mathsf{P}_{tu}\bigl[t^{-1}\bigl(X_1^{(t)}, \ldots, X_m^{(t)}\bigr) \in \cdot\bigr] \Rightarrow \mathsf{P}_u\bigl[(T_1, \ldots, T_m) \in \cdot\bigr] \qquad \text{on } [0,\infty]^m \qquad \text{as } t \to \infty. \tag{2.4}$$

So, the tail chain maps extreme states onto $(0, \infty)$ and contracts non-extreme states to the point $\{0\}$. Note $\tau(t) = \inf\{n \ge 0 : X_n^{(t)} = 0\}$.

If the finite-dimensional extremal behaviour of $\boldsymbol{X}$ is completely accounted for by the extremal component, then the tail chain approximation (2.4) extends from $\boldsymbol{X}^{(t)}$ to $\boldsymbol{X}$. When is this the case? Say that K satisfies the *regularity condition* if for any non-negative function $u_t = u(t) \to 0$,

$$K(tu_t, t\cdot) \Rightarrow \epsilon_0(\cdot) \qquad \text{on } [0,\infty] \qquad \text{as } t \to \infty. \tag{2.5}$$

Equivalent forms of (2.5) exist in [26, Section 4], and a relatively easy-to-check sufficient condition is given in terms of update functions. If either (a) $y(t) \equiv 0$ is an extremal boundary; or (b) K satisfies the regularity condition (2.5), then for $u > 0$, we strengthen (2.4) to [26, Theorem 4.1],

$$\mathsf{P}_{tu}\bigl[t^{-1}(X_1, \ldots, X_m) \in \cdot\bigr] \Rightarrow \mathsf{P}_u\bigl[(T_1, \ldots, T_m) \in \cdot\bigr] \qquad \text{on } [0,\infty]^m \qquad \text{as } t \to \infty. \tag{2.6}$$

2.2. Finite-dimensional convergence

The conditional approximation (2.4) requires that the initial state become extreme. Combining (2.4) with a heavy tailed initial distribution makes $\boldsymbol{X}^{(t)}$ have an unconditional distribution that is regularly varying (in a sense to be discussed) with a limit measure determined by the tail chain. Depending on assumptions, convergences take place on $\mathbb{E}_\sqsupset := (0,\infty] \times [0,\infty]^m$ or the larger space $\mathbb{E}^* := [0,\infty]^{m+1} \setminus \{\mathbf{0}\}$.

Theorem 2.1 ([26, Proposition 5.1(b), Theorem 5.1]). *Let $\boldsymbol{X}$ be a Markov chain on $[0,\infty)$ with $K \in D(G)$, and suppose $X_0 \sim H$, with $1 - H \in \mathrm{RV}_{-\alpha}$ with scaling function $b(\cdot)$. On $\mathbb{E}_\sqsupset$ define the measure*

$$\mu(dx_0, d\boldsymbol{x}_m) = \nu_\alpha(dx_0)\mathsf{P}_{x_0}\bigl[(T_1, \ldots, T_m) \in d\boldsymbol{x}_m\bigr], \tag{2.7}$$

and extend this to a measure μ^* *on* $\mathbb{E}^*$ *by defining* $\mu^*(\cdot \cap \mathbb{E}_\sqsupset) = \mu(\cdot)$ *and* $\mu^*(\mathbb{E}^* \setminus \mathbb{E}_\sqsupset) = 0$. *For any* $m \geq 1$, *the following convergences take place as* $t \to \infty$.

(a) *In* $\mathbb{M}_+((0,\infty]^m \times [0,\infty])$,

$$P\big[(X_0,\dots,X_m)/b(t) \in (\cdot) \cap (0,\infty]^m \times [0,\infty]\big] \xrightarrow{v} \mu\big((\cdot) \cap (0,\infty]^m \times [0,\infty]\big),$$

and in $\mathbb{M}_+(\mathbb{E}_\sqsupset)$

$$t\mathsf{P}\big[\big(X_0^{(b(t))},\dots,X_m^{(b(t))}\big)/b(t) \in \cdot\big] \xrightarrow{v} \mu(\cdot). \tag{2.8}$$

If either (i) $G(\{0\}) = 0$; (ii) $y(t) \equiv 0$ *is an extremal boundary*; *or* (iii) K *satisfies the regularity condition* (2.5), *then* (2.8) *can be strengthened to*

$$t\mathsf{P}\big[(X_0,\dots,X_m)/b(t) \in \cdot\big] \xrightarrow{v} \mu(\cdot), \qquad \textit{in } \mathbb{M}_+(\mathbb{E}_\sqsupset). \tag{2.9}$$

(b) *In the bigger space* $\mathbb{E}^*$, *we have*

$$t\mathsf{P}\big[\big(X_0^{(b(t))},\dots,X_m^{(b(t))}\big)/b(t) \in \cdot\big] \xrightarrow{v} \mu^*(\cdot) \qquad \textit{in } \mathbb{M}_+\big(\mathbb{E}^*\big) \tag{2.10}$$

if and only if

$$\begin{aligned} &\mathsf{E}\xi_1^\alpha < \infty \quad \textit{and} \quad t\mathsf{P}\big[X_j^{(b(t))}/b(t) \in \cdot\big] \xrightarrow{v} \big(\mathsf{E}\xi_1^\alpha\big)^j \nu_\alpha(\cdot) \\ &\textit{in } \mathbb{M}_+(0,\infty], \qquad j = 1,\dots,m. \end{aligned} \tag{2.11}$$

Part (a) requires that the first observation is large and with added conditions, part (b) removes this requirement. Regardless of whether (2.11) holds, the limit is always a lower bound on the tail weight of $X_j^{(t)}$, since

$$\liminf_{t\to\infty} t\mathsf{P}\big[X_j^{(b(t))}/b(t) > x\big] \geq \mu\big((0,\infty] \times [0,\infty]^{j-1} \times (x,\infty] \times [0,\infty]^{m-j}\big) = \big(\mathsf{E}\xi_1^\alpha\big)^j x^{-\alpha}$$

by (2.8) and Lemma 2.1 below. Markov's inequality, (2.8) and a moment condition:

$$\exists \varepsilon > 0 \text{ such that } \lim_{\delta\downarrow 0} \limsup_{t\to\infty} t\mathsf{E}\big[\big(X_j^{(b(t))}/b(t)\big)^\varepsilon \mathbf{1}_{\{X_0 \leq \delta b(t)\}}\big] = 0, \qquad j = 1,\dots,m, \tag{2.12}$$

imply (2.11). See [16].

Here is a formula that helps evaluate μ in (2.7) on certain sets.

Lemma 2.1. *For a random variable* Y, *define the measure* $\nu(dx,dy) = \nu_\alpha(dx)\mathsf{P}[xY \in dy]$ *on* $[0,\infty]^2 \setminus \{(0,0)\}$. *We compute*

$$\nu\big([0,x] \times (y,\infty]\big) = y^{-\alpha}\mathsf{E}\big[Y^\alpha \mathbf{1}_{\{Y > yx^{-1}\}}\big] - x^{-\alpha}\mathsf{P}\big[Y > yx^{-1}\big]. \tag{2.13}$$

In particular, $\nu([0,\infty] \times (y,\infty]) = y^{-\alpha}\mathsf{E}Y^\alpha$.

Proof. We obtain

$$\nu\bigl([0,x]\times(y,\infty]\bigr)=\int_{[0,x]}\nu_\alpha(ds)\mathsf{P}\bigl[Y>ys^{-1}\bigr]=\int_{[x^{-\alpha},\infty]}ds\,\mathsf{P}\bigl[Y>ys^{1/\alpha}\bigr]$$
$$=\int_{[x^{-\alpha},\infty]}ds\,\mathsf{P}\bigl[y^{-\alpha}Y^\alpha>s\bigr]$$

by change of variables. Applying Fubini's theorem, this becomes

$$\int_{(x^{-\alpha},\infty]}\bigl(s-x^{-\alpha}\bigr)\mathsf{P}\bigl[y^{-\alpha}Y^\alpha\in ds\bigr]=y^{-\alpha}\mathsf{E}\bigl[Y^\alpha\mathbf{1}_{\{Y>yx^{-1}\}}\bigr]-x^{-\alpha}\mathsf{P}\bigl[Y>yx^{-1}\bigr].$$

Letting $x\to\infty$, this quantity converges to $y^{-\alpha}\mathsf{E}Y^\alpha$ by monotone convergence. □

2.3. Maximum of the extremal component

We give conditions on the extremal component which enable an informative point process limit result by controlling the positive portion of the extremal component, the random vector of random length $\{X_j^{(t)}:j=0,\dots,\tau(t)-1\}=\{X_j:j=0,\dots,\tau(t)-1\}$. The conditions imply restrictions on the behaviour of the tail chain $\boldsymbol{T}$.

We study a positive recurrent chain $\boldsymbol{X}$ by splitting it into regenerative cycles and analyzing its extremal properties via the extremal components of the cycles. For regenerative processes, Asmussen [1] and Rootzén [27] point out the connection between point process convergence and the asymptotic distribution of cycle maxima. Informed by this approach, we consider when the distribution of the maximum over the extremal component has a regular variation property. The limit measure of this regular variation can be used to compute an extremal index for $\boldsymbol{X}$ [27].

Here is a condition that controls the persistence of non-zero values of the extremal component:

$$\lim_{m\to\infty}\limsup_{t\to\infty}\mathsf{P}\Bigl[\sup_{j\ge m}X_j^{(b(t))}/b(t)>a\Bigm|X_0>\delta b(t)\Bigr]=0\qquad\text{for all }a,\delta>0.\tag{2.14}$$

Note $\sup_{j\ge m}X_j^{(b(t))}=(\sup_{m\le j<\tau(b(t))}X_j)\mathbf{1}_{\{\tau(b(t))>m\}}$. Compare this condition with [4, Condition 4.1] and [22, Equation (3.1)], which are formulated in terms of block sizes. Condition (2.14) is a tightness condition that complements the finite-dimensional convergences (2.8). Section 3 gives simpler sufficient conditions depending on whether $G(\{0\})>0$ or $G(\{0\})=0$. Condition (2.14) requires that the chain drift back to the non-extreme states after visiting an extreme state and makes non-extreme states recurrent and the tail chain transient.

Proposition 2.1. *Let $\boldsymbol{X}\sim K\in D(G)$ be Markov on $[0,\infty)$ with initial distribution H satisfying $1-H\in\mathrm{RV}_{-\alpha}$ with scaling function $b(t)$, so that (2.8) holds. If $\boldsymbol{X}$ satisfies Condition (2.14), then*

$$\lim_{m\to\infty}\mathsf{P}\Bigl[\sup_{j\ge m}\xi(j)>a\Bigr]=0,\qquad a>0,\tag{2.15}$$

and $\xi(n) \to 0$ *as* $n \to \infty$ *in probability and almost surely and therefore,*

$$\mathsf{P}[T_n \to 0] = 1. \tag{2.16}$$

So the tail chain is transient and the additive random walk $\{\log T_n\}_{n\geq 0}$ satisfies $\log T_m \to -\infty$. The tail chain $\boldsymbol{T}$ and $\boldsymbol{X}$ live on the same state space $[0,\infty)$ but for $\boldsymbol{T}$, $\{0\}$ is a special boundary state which represents the collection of non-extreme states of $\boldsymbol{X}$ under the tail chain approximation.

Proof of Proposition 2.1. Observe from (2.8), as $t \to \infty$,

$$\begin{aligned} t\mathsf{P}\Big[X_0 > b(t), \sup_{m\le j\le r} X_j^{(b(t))} > b(t)\Big] &\longrightarrow \int_{(1,\infty]} \nu_\alpha(dx)\mathsf{P}_x\Big[\sup_{m\le j\le r} T_j > 1\Big] \\ &= \int_{(1,\infty]} \nu_\alpha(dx)\mathsf{P}\Big[\sup_{m\le j\le r} \xi(j) > x^{-1}\Big]. \end{aligned}$$

Therefore, by monotonicity and then monotone convergence,

$$\begin{aligned} \limsup_{t\to\infty} t\mathsf{P}\Big[X_0 > b(t), \sup_{j\ge m} X_j^{(b(t))} > b(t)\Big] &\ge \lim_{r\to\infty} \int_{(1,\infty]} \nu_\alpha(dx)\mathsf{P}\Big[\sup_{m\le j\le r} \xi(j) > x^{-1}\Big] \\ &= \int_{(1,\infty]} \nu_\alpha(dx)\mathsf{P}\Big[\sup_{j\ge m} \xi(j) > x^{-1}\Big] \\ &=: \int_{(1,\infty]} \nu_\alpha(dx) f_m(x). \end{aligned}$$

Condition (2.14) implies that $\int_{(1,\infty]} \nu_\alpha(dx) f_m(x) \to 0$ as $m \to \infty$. We claim that $f_m(x) \to 0$ for any $x > 0$. Suppose instead that $\inf_m f_m(x_0) \ge c > 0$ for some x_0. Since the f_m are all increasing in x, we have $\inf_m f_m(x) \ge c$ for $x \ge x_0$. But this implies that

$$\liminf_{m\to\infty} \int_{(1,\infty]} \nu_\alpha(dx) f_m(x) \ge \liminf_{m\to\infty} \int_{(1\vee x_0,\infty]} \nu_\alpha(dx) f_m(x) \ge c\nu_\alpha(1\vee x_0, \infty] > 0$$

by Fatou's Lemma, contradicting Condition 2.14. Therefore, $\mathsf{P}[\sup_{j\ge m} \xi(j) > x^{-1}] \to 0$ as $m \to \infty$ for all $x > 0$, establishing (2.15). □

Condition (2.14) assumes the first observation exceeds $\delta b(t)$ which is in the spirit of (2.8). For translating the stronger convergence of unconditional distributions (2.10) in the bigger space $\mathbb{E}^* := [0,\infty]^{m+1} \setminus \{\mathbf{0}\}$ to point process convergence, we will require an additional assumption:

$$\begin{aligned} &\exists m_0 \ge 1 \text{ such that} \\ &\lim_{\delta\downarrow 0} \limsup_{t\to\infty} t\mathsf{P}\Big[X_0/b(t) \le \delta, \sup_{j\ge m_0} X_j^{(b(t))}/b(t) > a\Big] = 0 \qquad \text{for all } a > 0. \end{aligned} \tag{2.17}$$

Analogously to (2.12), by Markov's inequality a moment condition is sufficient for Condition (2.17):

$$\lim_{\delta\downarrow 0}\limsup_{t\to\infty} t\mathsf{E}\Big[\Big(\sup_{j\ge m_0} X_j^{(b(t))}/b(t)\Big)^{\varepsilon}\mathbf{1}_{\{X_0\le\delta b(t)\}}\Big]=0, \qquad \text{for some } \varepsilon>0.$$

Condition (2.17) implies a uniform bound on the αth moment of the tail chain states.

Proposition 2.2. *Let $\boldsymbol{X}\sim K\in D(G)$ be a Markov chain on $[0,\infty)$ with initial distribution H satisfying $1-H\in \mathrm{RV}_{-\alpha}$, whose extremal component satisfies (2.10) on $\mathbb{M}_+(\mathbb{E}^*)$. If $\boldsymbol{X}$ satisfies Condition (2.17) then,*

$$\mathsf{E}\Big(\sup_{j\ge 1}\xi(j)^{\alpha}\Big)<\infty. \tag{2.18}$$

Remark. Under (2.18), we necessarily have $\mathsf{E}\xi_1^{\alpha}\le 1$ since

$$\sup_{j\ge 1}\big(\mathsf{E}\xi_1^{\alpha}\big)^j=\sup_{j\ge 1}\mathsf{E}\xi(j)^{\alpha}\le \mathsf{E}\Big(\sup_{j\ge 1}\xi(j)^{\alpha}\Big)<\infty.$$

Recalling (2.11), the marginal tails of the extremal component cannot be heavier than the tail of H.

Proof of Proposition 2.2. Under (2.10),

$$t\mathsf{P}\Big[\sup_{m_0\le j\le r} X_j^{(b(t))}>b(t)\Big]\longrightarrow \int_{(0,\infty]}\nu_\alpha(dx)\mathsf{P}\Big[\sup_{m_0\le j\le r}\xi(j)>x^{-1}\Big]=\mathsf{E}\Big(\sup_{m_0\le j\le r}\xi(j)^{\alpha}\Big)$$

by Lemma 2.1. Therefore,

$$\limsup_{t\to\infty} t\mathsf{P}\Big[\sup_{j\ge m_0} X_j^{(b(t))}>b(t)\Big]\ge \lim_{r\to\infty}\mathsf{E}\Big(\sup_{m_0\le j\le r}\xi(j)^{\alpha}\Big)=\mathsf{E}\Big(\sup_{j\ge m_0}\xi(j)^{\alpha}\Big).$$

Furthermore, by Condition (2.17), for some $\delta>0$,

$$\begin{aligned}&\limsup_{t\to\infty} t\mathsf{P}\Big[\sup_{j\ge m_0} X_j^{(b(t))}>b(t)\Big]\\&\quad\le \limsup_{t\to\infty} t\mathsf{P}\Big[X_0\le\delta b(t),\ \sup_{j\ge m_0} X_j^{(b(t))}>b(t)\Big]+\limsup_{t\to\infty} t\mathsf{P}\big[X_0>\delta b(t)\big]<\infty\end{aligned}$$

showing that $\mathsf{E}(\sup_{j\ge m_0}\xi(j)^{\alpha})<\infty$ which is enough for (2.18). □

Under both Conditions (2.14) and (2.17), we derive the tail behaviour of the maximum of the extremal component of $\boldsymbol{X}$.

Proposition 2.3. *Let $\boldsymbol{X}\sim K\in D(G)$ be a Markov chain on $[0,\infty)$ with initial distribution H satisfying $1-H\in\mathrm{RV}_{-\alpha}$, whose extremal component satisfies both (2.10) on $\mathbb{M}_+(\mathbb{E}^*)$. If $\boldsymbol{X}$*

satisfies Conditions (2.14) *and* (2.17), *then*

$$t\mathsf{P}\Big[\sup_{0\le j<\tau(b(t))} X_j/b(t)\in\cdot\Big]\xrightarrow{v} c\cdot\nu_\alpha(\cdot)\qquad \textit{in } \mathbb{M}_+(0,\infty],\tag{2.19}$$

where $c=\mathsf{P}[\sup_{j\ge1}\xi(j)\le 1]+\mathsf{E}[\sup_{j\ge1}\xi(j)^\alpha\mathbf{1}_{\{\sup_{j\ge1}\xi(j)>1\}}]$.

Proof. For $x>0$, we have $[\sup_{j<\tau(b(t))}X_j/b(t)>x]=[\sup_{j\ge0}X_j^{(b(t))}/b(t)>x]$. For $m\ge1$, we have on the one hand, by (2.10),

$$\begin{aligned}\liminf_{t\to\infty} t\mathsf{P}\Big[\sup_{0\le j<\tau(b(t))}X_j/b(t)>x\Big]&\ge\lim_{t\to\infty}t\mathsf{P}\Big[\sup_{0\le j<m}X_j^{(b(t))}/b(t)>x\Big]\\&=x^{-\alpha}+\int_{[0,x)}\nu_\alpha(du)\mathsf{P}_u\Big[\sup_{1\le j<m}T_j>x\Big],\end{aligned}$$

from which, letting $m\to\infty$,

$$\liminf_{t\to\infty} t\mathsf{P}\Big[\sup_{0\le j<\tau(b(t))}X_j/b(t)>x\Big]\ge x^{-\alpha}+\int_{[0,x)}\nu_\alpha(du)\mathsf{P}_u\Big[\sup_{j\ge1}T_j>x\Big].\tag{2.20}$$

On the other hand, for $\delta>0$ we have

$$t\mathsf{P}\bigg[\sup_{j\ge m}\frac{X_j^{(b(t))}}{b(t)}>x\bigg]\le t\mathsf{P}\bigg[\frac{X_0}{b(t)}>\delta,\ \sup_{j\ge m}\frac{X_j^{(b(t))}}{b(t)}>x\bigg]+t\mathsf{P}\bigg[\frac{X_0}{b(t)}\le\delta,\ \sup_{j\ge m}\frac{X_j^{(b(t))}}{b(t)}>x\bigg].$$

Given $\varepsilon>0$, by Condition (2.17), we may choose δ small enough that

$$\limsup_{t\to\infty}t\mathsf{P}\Big[X_0\le\delta b(t),\ \sup_{j\ge m_0}X_j^{(b(t))}>b(t)x\Big]<\varepsilon/2,$$

where m_0 is from Condition (2.17). Condition (2.14) permits the choice $m_1\ge m_0$ so large that

$$\limsup_{t\to\infty}t\mathsf{P}\Big[X_0>\delta b(t),\ \sup_{j\ge m_1}X_j^{(b(t))}>b(t)x\Big]<\varepsilon/2.$$

Therefore, $\limsup_{t\to\infty}t\mathsf{P}[\sup_{j\ge m}X_j^{(b(t))}/b(t)>x]<\varepsilon$ for $m\ge m_1$, and so

$$\begin{aligned}\limsup_{t\to\infty} t\mathsf{P}\Big[\sup_{0\le j<\tau(b(t))}X_j/b(t)>x\Big]&<\lim_{m\to\infty}\limsup_{t\to\infty}t\mathsf{P}\Big[\sup_{0\le j<m}X_j^{(b(t))}/b(t)>x\Big]+\varepsilon\\&=x^{-\alpha}+\int_{[0,x)}\nu_\alpha(du)\mathsf{P}_u\Big[\sup_{j\ge1}T_j>x\Big]+\varepsilon.\end{aligned}$$

Combine this with (2.20), and apply formula (2.13) for $\nu([0,x]\times(x,\infty])$ to complete the proof. $\square$

3. Point process convergence for Markov chains

We now derive the limit of the exceedance point process N_n defined in (1.1), where $\boldsymbol{X} = (X_0, X_1, \ldots)$ is a Markov chain on $[0, \infty)$ with transition kernel $K \in D(G)$. We write

$$N_n = \sum_{j=0}^{\infty} \epsilon_{(j/n, X_j/b_n)}, \tag{3.1}$$

using the notation ϵ_x to denote the measure assigning unit mass at the point x and N_n is a random element of $\mathbb{M}_p([0, \infty) \times (0, \infty])$, the space of Radon point measures on $[0, \infty) \times (0, \infty]$, endowed with the topology of vague convergence [13,20,25].

If $\boldsymbol{X}$ is positive recurrent, it is a regenerative process ([2, Section VII.3], [17]) so the sample path of $\boldsymbol{X}$ splits into identically distributed cycles between visits to certain set. The extremal properties of $\boldsymbol{X}$ are determined by extremal behaviour of the individual cycles. This approach has been developed for Markov chains by Rootzén [27], as well as for queues by Asmussen [1]. Our approach introduces the tail chain approximation to describe the extremal behaviour of the regenerative cycles using their extremal component.

3.1. Cycle decomposition

Consider the case where $\boldsymbol{X}$ has a positive recurrent atom A. For positive recurrent chains, atoms can be constructed by several methods if no natural atom exists. See, e.g., [8, Chapter 6] or [18, Chapter I.5]. An atom is a set such that for a probability distribution H on $[0, \infty)$,

$$K(y, \cdot) = H(\cdot) \qquad \text{for all } y \in A \quad \text{and} \quad \mathsf{P}_y[\tau_A < \infty] = 1 \qquad \text{for } y \geq 0, \tag{3.2}$$

and $\tau_A = \inf\{n \geq 0 : X_n \in A\}$ is the first hitting time of A. Positive recurrence means that

$$\mathsf{E}_H \tau_A < \infty, \tag{3.3}$$

where E_H denotes expectation with respect to H considered as the initial distribution of X_0.

Under (3.2), the sample path of $\boldsymbol{X}$ splits into i.i.d. cycles between visits to A, as follows. Define the times $\{S_k\}$, $\{\tau_k^A\}$ recursively according to

$$\begin{aligned} \tau_0^A &= \tau_A, & S_0 &= \tau_0^A + 1; \\ \tau_k^A &= \inf\{n \geq 0 : X_{S_{k-1}+n} \in A\}, & S_k &= S_{k-1} + \tau_k^A + 1, k \geq 1. \end{aligned} \tag{3.4}$$

Thus, the sequence $0 \leq S_0 - 1 < S_1 - 1 < S_2 - 1 < \cdots$ gives the indices when $\boldsymbol{X}$ is in A, and $X_{S_k} \sim H$ for $k \geq 0$. The values $\tau_k^A \geq 0$ are the number of steps $\boldsymbol{X}$ takes outside of A between visits to A. The cycles end by visits to A; cycles are the random elements

$$C_0 = (X_0, X_1, \ldots, \underset{\substack{\in \\ A}}{X_{\tau_0^A}}) \quad \text{and} \quad C_k = (\underset{\substack{\wr \\ H}}{X_{S_{k-1}}}, \ldots, \underset{\substack{\in \\ A}}{X_{S_{k-1}+\tau_k^A}}), \qquad k \geq 1$$

of the space of finite sequences $\mathcal{S} = \bigcup_{m=1}^{\infty} \mathbb{R}^m$. The strong Markov property implies $C_0, C_1, \ldots$ are independent, and $C_1, C_2, \ldots$ are identically distributed. In particular, for $k \geq 1$,

$$\mathsf{P}\big[\{C_k; \tau_k^A\} \in \cdot\big] = \mathsf{P}\big[\{(X_{S_{k-1}}, \ldots, X_{S_{k-1}+\tau_k^A}); \tau_k^A\} \in \cdot\big] = \mathsf{P}_H\big[\{(X_0, \ldots, X_{\tau_A}); \tau_A\} \in \cdot\big].$$

Furthermore, $0 < S_0 < S_1 < S_2 < \cdots$ is a renewal process, with

$$q = \mathsf{E}(S_1 - S_0) = \mathsf{E}_H \tau_A + 1 < \infty \tag{3.5}$$

by (3.3). Applying the cycle decomposition, we may now write (3.1) as

$$N_n = \sum_{j=0}^{\infty} \epsilon_{(j/n, X_j/b_n)} = \sum_{j=0}^{S_0-1} \epsilon_{(j/n, X_j/b_n)} + \sum_{k=0}^{\infty} \sum_{j=0}^{\tau_{k+1}^A} \epsilon_{((S_k+j)/n, X_{S_k+j}/b_n)} = \chi_n^0 + \chi_n^*. \tag{3.6}$$

As a family of random elements in $\mathbb{M}_p([0,\infty) \times (0,\infty])$, $\{\chi_n^0\}$ is asymptotically negligible.

Lemma 3.1. *Assuming* (3.3) *and* $b_n \to \infty$, $\chi_n^0 \Rightarrow 0$, *the null measure*, *in* $\mathbb{M}_p([0,\infty) \times (0,\infty])$.

Proof. Let $f \in \mathcal{C}_K^+([0,\infty) \times (0,\infty])$ with support in $[0, R] \times [M, \infty]$ for integers R, M. It is sufficient to verify that $\mathsf{P}[\chi_n^0(f) > \gamma] \to 0$, for any $\gamma > 0$. We have as $n \to \infty$,

$$\begin{aligned}\mathsf{P}\big[\chi_n^0(f) > \gamma\big] &= \mathsf{P}\left[\sum_{j=0}^{S_0-1} f\left(\frac{j}{n}, \frac{X_j}{b_n}\right) > \gamma\right] \\ &\leq \sum_{m=0}^{r} \mathsf{P}\left[\sum_{j=0}^{m} f\left(\frac{j}{n}, \frac{X_j}{b_n}\right) > \gamma, \tau_A = m\right] + \mathsf{P}[\tau_A > r],\end{aligned}$$

and

$$\begin{aligned}\sum_{m=0}^{r} \mathsf{P}\left[\sum_{j=0}^{m} f\left(\frac{j}{n}, \frac{X_j}{b_n}\right) > \gamma, \tau_A = m\right] &\leq (r+1)\mathsf{P}\left[\sum_{j=0}^{r} f\left(\frac{j}{n}, \frac{X_j}{b_n}\right) > \gamma\right] \\ &\leq (r+1)\mathsf{P}\Big[\sup_{0 \leq j \leq r} X_j \geq b_n M\Big] \longrightarrow 0.\end{aligned}$$

Choosing r to make $\mathsf{P}[\tau_A > r]$ arbitrarily small, the result follows. □

3.2. Point process convergence

Lemma 3.1 and Slutsky's theorem means that the asymptotic behavior of N_n and χ_n^* are the same. We obtain a weak limit for χ_n^* using the tail chain approximation discussed in Section 2, provided that a cycle's extremal behaviour is adequately described by its extremal component.

As usual, assume $K \in D(G)$, $1 - H \in \mathrm{RV}_{-\alpha}$ and suppose $y(t)$ is an extremal boundary for $\boldsymbol{X}$. We require a mild assumption that the atom A be a bounded subset of the state space $[0, \infty)$,

$$\sup A < \infty, \tag{3.7}$$

as would usually be the case. Fix $k \geq 1$. The number of steps needed by the scaled process in the k-th cycle to cross below the extremal boundary is

$$\tau_k(t) = \inf\bigl\{n \geq 0 : X_{S_{k-1}+n} \leq t y(t)\bigr\}.$$

The extremal component of the k-th cycle is $C_k(t) := \{X_{S_{k-1}+j} : j = 0, \dots, \tau_k(t) - 1\}$.

Without loss of generality, we suppose the extremal component of a cycle is a subset of the complete cycle. To see this, observe from the definition of $\tau(t)$ and τ_A, without loss of generality,

$$\mathsf{P}\bigl[\tau(t) \leq \tau_A, \forall t > 0\bigr] = 1. \tag{3.8}$$

Indeed, (3.7) implies that $A \subset [0, c]$ some c. Define $\tau_c = \inf\{n \geq 0 : X_n \leq c\}$ and $\mathsf{P}[\tau_c \leq \tau_A] = 1$; we claim further that we may suppose $\mathsf{P}[\tau(t) \leq \tau_c, \forall t > 0] = 1$. If $y(t) \geq c/t$ for all $t > 0$, then this follows directly. Otherwise, verify that $\tilde{y}(t) = y(t) \vee c/t$ is also an extremal boundary for K (see the remarks after (2.2)), and the corresponding downcrossing time satisfies $\mathsf{P}[\tilde{\tau}(t) \leq \tau_c, \forall t > 0] = 1$.

Therefore, for $k \geq 1$,

$$\mathsf{P}\bigl[\bigl\{(X_{S_{k-1}}, \dots, X_{S_{k-1}+\tau_k(t)-1}); \tau_k(t), \tau_k^A\bigr\} \in \cdot\bigr] = \mathsf{P}_H\bigl[\bigl\{(X_0, \dots, X_{\tau(t)-1}); \tau(t), \tau_A\bigr\} \in \cdot\bigr] \tag{3.9}$$

and $\{(C_k(t); \tau_k(t), \tau_k^A) : k \geq 1\}$ are independent, since each is a function of $\{C_k; \tau_k^A\}$. These facts suggest we approximate χ_n^* by a point process whose observations consist of the extremal components of iid copies of the chain $\boldsymbol{X}$ started from $X_0 \sim H$. This approximation is facilitated by additional notation. Let $\{\boldsymbol{X}, \boldsymbol{X}_k = (X_{kj}, j \geq 0) : k \geq 0\}$ be i.i.d. copies of the Markov chain $\boldsymbol{X} \sim K$ with respect to $\mathsf{P}_H(\cdot)$; that is the initial distribution of each chain is fixed to be H. Define

$$\tilde{\tau}_{k+1}(t) = \inf\bigl\{j \geq 0 : X_{kj} \leq t y(t)\bigr\}, \qquad k = 0, 1, \dots,$$

and for $k \geq 0$, form the extremal component $\boldsymbol{X}_k^{(t)} = \{X_{kj} \cdot \mathbf{1}_{\{j < \tilde{\tau}_{k+1}(t)\}}, j \geq 0\}$ of the kth chain. Thus with respect to $\mathsf{P}_H(\cdot)$, $(\boldsymbol{X}_k^{(t)}, \tilde{\tau}_{k+1}(t)) \stackrel{d}{=} (\boldsymbol{X}^{(t)}, \tau(t))$ for $k \geq 0$, with the tilde differentiating the times $\tilde{\tau}_k(t)$ defined on the kth process $\boldsymbol{X}_k$ from the cycle times $\tau_k(t)$ defined on $\boldsymbol{X}$. Recall $\tau(t)$ is also defined on $\boldsymbol{X}$.

Next, generate an i.i.d. family of tail chains by letting $\{\boldsymbol{\xi}, \boldsymbol{\xi}_k = (\xi_k(n), n \geq 0) : k \geq 0\}$ be i.i.d. copies of the process $\boldsymbol{\xi} = (\xi(n), n \geq 0)$, recalling the notation around (2.1). Additionally, put $\tau_{k+1}^* = \inf\{j \geq 0 : \xi_k(j) = 0\}$, the first time the kth tail chain hits 0. Use the convention $\inf \varnothing = \infty$; for example, $\tau_{k+1}^* = \infty$ a.s., $k \geq 0$, if $G(\{0\}) = 0$. Finally, let

$$\zeta = \sum_{k=0}^{\infty} \epsilon_{(t_k, i_k)} \sim \mathrm{PRM}(\mathbb{LEB} \times \nu_\alpha),$$

be a Poisson random measure on $\mathbb{M}_p([0,\infty)\times(0,\infty])$, independent of the $\{\boldsymbol{\xi}_k\}$, with mean measure a product of Lebesgue measure on the time axis $[0,\infty)$ and Pareto measure ν_α (given by $\nu_\alpha(x,\infty]=x^{-\alpha}$) on the observation axis $(0,\infty]$. Recall α is the tail index of $\bar{H}$.

The point process consisting of the observations $\boldsymbol{X}_k^{(b_n)}$, spaced in time according to the renewal times $\{S_k\}$, converges to a cluster Poisson process which is basically ζ with a time scaling and compounded in the second coordinate according to the i.i.d. tail chains $\{\boldsymbol{\xi}_k\}$. This result is basic to analyzing the asymptotic behavior of N_n in (3.1).

Proposition 3.1. *Let X be a Markov chain on $[0,\infty)$ with transition kernel $K\in D(G)$, and initial distribution H, such that $tH(b(t)\cdot)\xrightarrow{v}\nu_\alpha(\cdot)$ in $\mathbb{M}_+(0,\infty]$, where $b(t)\to\infty$ and $b_n=b(n)$. The renewal process $\{S_k\}$ is defined in (3.4), with mean interarrival time q given by (3.5). With the notation introduced in the previous paragraphs, we have the following with respect to P_H.*

(a) *If $\boldsymbol{X}$ satisfies Condition (2.14), then given $\delta>0$, in $\mathbb{M}_p([0,\infty)\times(0,\infty])$, as $n\to\infty$,*

$$\eta_n:=\sum_{k=0}^{\infty}\sum_{j=0}^{\tilde{\tau}_{k+1}(b_n)-1}\epsilon_{\left(\frac{S_k+j}{n},\frac{X_{kj}}{b_n}\right)}\mathbf{1}_{\{X_{k0}\geq\delta b_n\}}\Rightarrow\sum_{k=0}^{\infty}\sum_{j=0}^{\tau_{k+1}^*-1}\epsilon_{(qt_k,i_k\xi_k(j))}\mathbf{1}_{\{i_k\geq\delta\}}=:\eta. \tag{3.10}$$

(b) *Suppose $\boldsymbol{X}$ satisfies (2.10) as well as both Conditions (2.14) and (2.17). Then in $\mathbb{M}_p([0,\infty)\times(0,\infty])$, as $n\to\infty$,*

$$\eta_n^*:=\sum_{k=0}^{\infty}\sum_{j=0}^{\tilde{\tau}_{k+1}(b_n)-1}\epsilon_{\left(\frac{S_k+j}{n},\frac{X_{kj}}{b_n}\right)}\Rightarrow\sum_{k=0}^{\infty}\sum_{j=0}^{\tau_{k+1}^*-1}\epsilon_{(qt_k,i_k\xi_k(j))}=:\eta^*. \tag{3.11}$$

Section 4 (p. 1440) contains the proof. Paralleling the discussion in Section 2, we have two results depending on the strength of the conditions. The weaker assumptions of part (a) yield a result that selects cycles starting from an exceedance. Part (b) does not have to do such cycle selection.

The points of the limit process are arranged in stacks above common time points qt_k. The heights of the points in each stack are specified by an independent run of the tail chain starting from i_k. If $G(\{0\})>0$, then the τ_k^* are i.i.d. Geometric random variables with parameter $G(\{0\})$, so all stacks have finite length. If $G(\{0\})=0$, then $\mathsf{P}[\tau_k^*=\infty]=1$ for each k. In this case, Condition (2.14) is necessary to ensure that η^* is Radon, by forcing the tail chain to drift towards 0 as in (2.16). The process η retains only those stacks of η^* whose initial value exceeds the threshold δ. Because there are an infinite number of i_k in any neighbourhood of 0, dispensing with the restriction in δ requires that not too many of the $\xi_k(j)$ are large. This translates to the condition $\mathsf{E}\xi_1^\alpha\leq 1$, provided by Condition (2.17).

To analyze N_n in (3.1), we approximate χ_n^* in (3.6) by η_n^* in (3.11), provided the extremal component adequately describes extremal behaviour within each cycle. If the extremal boundary is not identically zero, behavior between the end of the extremal component and the end of the cycle is not be captured by the tail chain and we require that such observations do not significantly influence extremal properties. To guarantee a result analogous to Part (a) above, we require,

$$\lim_{t\to\infty} \mathsf{P}\Big[\sup_{\tau(b(t))<j<\tau_A} X_j/b(t) > a \Big| X_0 > \delta b(t)\Big] = 0 \qquad \text{for all } a, \delta > 0, \tag{3.12}$$

and for a result analogous to Part (b) above, we require,

$$\lim_{t\to\infty} t\mathsf{P}\Big[\sup_{\tau(b(t))<j<\tau_A} X_j/b(t) > a\Big] = 0 \qquad \text{for all } a > 0. \tag{3.13}$$

With these conditions, the point process N_n converges to the limit η^*, and the distribution of the cycle maximum behaves as if it has a regularly varying tail.

Theorem 3.1. *Let $\boldsymbol{X}$ be a Markov chain on $[0,\infty)$ with transition kernel $K \in D(G)$. Suppose that K has a positive recurrent bounded atom in the sense of (3.2), (3.3), and (3.7). Define the renewal process $\{S_k\}$ with mean interarrival time q as in (3.4) and (3.5) and assume further that $tH(b(t)\cdot) \xrightarrow{v} \nu_\alpha(\cdot)$ in $\mathbb{M}_+(0,\infty]$, where $b(t)\to\infty$. With respect to P_H, the following hold.*

(a) *If $\boldsymbol{X}$ satisfies Conditions (2.14) and (3.12), then given $\delta > 0$,*

$$\tilde{N}_n := \sum_{0\le j<S_0} \epsilon_{(\frac{j}{n},\frac{X_j}{b_n})} + \sum_{k=1}^{\infty} \sum_{S_{k-1}\le j<S_k} \mathbf{1}_{\{X_{S_{k-1}}\ge\delta b_n\}} \epsilon_{(\frac{j}{n},\frac{X_j}{b_n})} \Rightarrow \eta \tag{3.14}$$

in $\mathbb{M}_p([0,\infty)\times(0,\infty])$, as $n\to\infty$, where η is defined in (3.10).

(b) *Suppose additionally that $\boldsymbol{X}$ satisfies (2.10) as well as Conditions (2.14), (2.17) and (3.13). Recall η^* from (3.11). Then, in $\mathbb{M}_p([0,\infty)\times(0,\infty])$, as $n\to\infty$,*

$$N_n \Rightarrow \eta^*, \tag{3.15}$$

and furthermore, the distribution of the cycle maximum has a regularly varying tail,

$$t\mathsf{P}_H\Big[b(t)^{-1}\sup_{0\le j<\tau_A} X_j \in \cdot\Big] \xrightarrow{v} c\cdot\nu_\alpha(\cdot) \qquad \text{in } \mathbb{M}_+(0,\infty], \tag{3.16}$$

where

$$c = \mathsf{P}\Big[\sup_{j\ge1}\xi(j)\le 1\Big] + \mathsf{E}\Big[\sup_{j\ge1}\xi(j)^\alpha \mathbf{1}_{\{\sup_{j\ge1}\xi(j)>1\}}\Big]. \tag{3.17}$$

Proof. (a) First, note that $\tilde{N}_n = \chi_n^0 + \chi_n$, where

$$\chi_n = \sum_{k=0}^{\infty}\sum_{j=0}^{\tau_{k+1}^A} \epsilon_{(\frac{S_k+j}{n},\frac{X_{S_k+j}}{b_n})} \mathbf{1}_{\{X_{S_k}\ge\delta b_n\}}.$$

Hence, by Lemma 3.1 it remains to show that $\chi_n \Rightarrow \eta$. Split χ_n according to the times $\{\tau_k(b_n)\}$:

$$\chi_n = \sum_{k=0}^{\infty} \sum_{j=0}^{\tau_{k+1}(b_n)-1} \epsilon_{\left(\frac{S_k+j}{n}, \frac{X_{S_k+j}}{b_n}\right)} \mathbf{1}_{\{X_{S_k} \geq \delta b_n\}} + \sum_{k=0}^{\infty} \sum_{j=\tau_{k+1}(b_n)}^{\tau_{k+1}^A} \epsilon_{\left(\frac{S_k+j}{n}, \frac{X_{S_k+j}}{b_n}\right)} \mathbf{1}_{\{X_{S_k} \geq \delta b_n\}} = \chi_n' + \chi_n''.$$

The equality holds on the set $\{\tau_k(b_n) \leq \tau_k^A; n \geq 1, k \geq 1\}$, which has probability 1 by (3.8). Because of (3.9) and the independence of the $(C_k(t), \tau_k(t))$, we have $\chi_n' \stackrel{d}{=} \eta_n$ for each n, and $\eta_n \Rightarrow \eta$ by Proposition 3.1(a). By Slutsky's theorem, the result follows if $\chi_n'' \Rightarrow 0$, so we show

$$\mathsf{P}\left[\chi_n''(f) > \gamma\right] = \mathsf{P}\left[\sum_{k=0}^{\infty} \sum_{j=\tau_{k+1}(b_n)}^{\tau_{k+1}^A} f\left(\frac{S_k+j}{n}, \frac{X_{S_k+j}}{b_n}\right) \mathbf{1}_{\{X_{S_k} \geq \delta b_n\}} > \gamma\right] \longrightarrow 0,$$

for any $f \in C_K^+([0,\infty) \times (0,\infty])$ and $\gamma > 0$. Let f have support in $[0, R] \times [M, \infty]$ for integers R, M. The previous probability is bounded by

$$\mathsf{P}\left[\sum_{k=0}^{2Rn-1} \sum_{j=\tau_{k+1}(b_n)}^{\tau_{k+1}^A} f\left(\frac{S_k+j}{n}, \frac{X_{S_k+j}}{b_n}\right) \mathbf{1}_{\{X_{S_k} \geq \delta b_n\}} > 0\right]$$
$$+ \mathsf{P}\left[\sum_{k=2Rn}^{\infty} \sum_{j=\tau_{k+1}(b_n)}^{\tau_{k+1}^A} f\left(\frac{S_k+j}{n}, \frac{X_{S_k+j}}{b_n}\right) \mathbf{1}_{\{X_{S_k} \geq \delta b_n\}} > 0\right].$$

Observe that the second term is at most $\mathsf{P}[S_{2Rn}/n \leq R] = \mathsf{P}[S_{2Rn}/2Rn \leq 1/2] \to 0$ as $n \to \infty$, since $S_n/n \to q$ a.s., and $q \geq 1$ by (3.5). The first term is bounded by

$$\mathsf{P}\left[\bigcup_{k=0}^{2Rn-1} \left(\left\{\frac{X_{S_k}}{b_n} \geq \delta\right\} \cap \bigcup_{j=\tau_{k+1}(b_n)}^{\tau_{k+1}^A} \left\{\frac{X_{S_k+j}}{b_n} \geq M\right\}\right)\right] \leq 2Rn\mathsf{P}_H\left[\frac{X_0}{b_n} \geq \delta, \sup_{\tau(b_n) < j < \tau_A} \frac{X_j}{b_n} \geq M\right],$$

which vanishes as $n \to \infty$ by Condition (3.12).

(b) Recalling the decomposition (3.6), by Lemma 3.1 it is sufficient to show that $\chi_n^* \Rightarrow \eta^*$. This follows by a similar argument as in part (a). Write

$$\chi_n^* = \sum_{k=0}^{\infty} \sum_{j=0}^{\tau_{k+1}(b_n)-1} \epsilon_{\left(\frac{S_k+j}{n}, \frac{X_{S_k+j}}{b_n}\right)} + \sum_{k=0}^{\infty} \sum_{j=\tau_{k+1}(b_n)}^{\tau_{k+1}^A} \epsilon_{\left(\frac{S_k+j}{n}, \frac{X_{S_k+j}}{b_n}\right)} = \chi_n^{*\prime} + \chi_n^{*\prime\prime}.$$

Then $\chi_n^{*\prime} \stackrel{d}{=} \eta_n^* \Rightarrow \eta^*$ by Proposition 3.1(b), and Condition (3.13) implies that $\chi_n^{*\prime\prime} \Rightarrow 0$.

Next, we show (3.16). In light of (3.8), we have

$$0 \leq t\mathsf{P}_H\left[\sup_{0 \leq j < \tau_A} \frac{X_j}{b(t)} > x\right] - t\mathsf{P}_H\left[\sup_{0 \leq j < \tau(b(t))} \frac{X_j}{b(t)} > x\right] \leq t\mathsf{P}_H\left[\sup_{\tau(b(t)) < j < \tau_A} \frac{X_j}{b(t)} > x\right] \longrightarrow 0$$

under Condition (3.13). Recalling that

$$t\mathsf{P}_H\left[\sup_{0\le j<\tau(b(t))}\frac{X_j}{b(t)}>x\right]\longrightarrow cx^{-\alpha}$$

as $t\to\infty$ by Proposition 2.3 (p. 1426), where c is as in (3.17), completes the proof. □

Setting $M_n=\bigvee_{0\le j\le n}X_j$, Rootzén shows [27, Theorem 3.2] that (3.16) implies

$$\mathsf{P}[M_n\le b_nx]\longrightarrow \exp\bigl(-cq^{-1}x^{-\alpha}\bigr),\qquad x>0,$$

where c is given by (3.17), and q is the mean interarrival time (3.5). Hence, in the stationary case, $\theta=c/q$ is the extremal index of the process $\boldsymbol{X}$ ([15, Section 2.2], [14]). On the other hand, for stationary regularly varying Markov chains with $K\in D(G)$ satisfying a condition analogous to Condition (2.14), it is known [4, Remark 4.7],

$$\theta=\mathsf{P}\Bigl[\sup_{j\ge1}Y\xi(j)\le1\Bigr]=\mathsf{P}\Bigl[\sup_{j\ge1}\xi(j)\le1\Bigr]-\mathsf{E}\Bigl[\sup_{j\ge1}\xi(j)^{\alpha}\mathbf{1}_{\{\sup_{j\ge1}\xi(j)\le1\}}\Bigr]=c-\mathsf{E}\Bigl(\sup_{j\ge1}\xi(j)^{\alpha}\Bigr),$$

where $Y\sim\text{Pareto}(\alpha)$ supported on $[1,\infty)$, independent of $\{\xi(j)\}$. Hence, for a stationary Markov chain $\boldsymbol{X}$ satisfying the assumptions of Theorem 3.1(b), the extremal index is given by

$$\theta=\frac{1}{q-1}\mathsf{E}\Bigl(\sup_{j\ge1}\xi(j)^{\alpha}\Bigr)=\frac{\mathsf{E}(\sup_{j\ge1}\xi(j)^{\alpha})}{\mathsf{E}_H\tau_A}.$$

3.3. Discussion of conditions

We now consider simplifications of the above conditions.

3.3.1. *Cases where $G(\{0\})=0$*

If $G(\{0\})=0$, we can replace $\boldsymbol{X}^{(b(t))}$ with $\boldsymbol{X}$ in the finite-dimensional convergence (2.8) when H has a regularly varying tail, meaning that the tail chain approximation completely describes the extremes of the chain $\boldsymbol{X}$ in a finite dimensional sense. However, $G(\{0\})=0$ also implies that for any $m>0$, as $t\to\infty$,

$$\mathsf{P}_t\bigl[m<\tau(t)\le\tau_A\bigr]\longrightarrow1 \tag{3.18}$$

(see [26, Proposition 5.1(d)]) meaning that, as the initial observation becomes more extreme, it takes longer for $\boldsymbol{X}$ to return to A to complete the cycle. Hence, for Condition (2.14) to hold, we need a condition that ensures $\boldsymbol{X}$ eventually drifts away from extreme states:

$$\lim_{m\to\infty}\limsup_{t\to\infty}\mathsf{P}_t\Bigl[\sup_{m\le j<\tau_A}X_j>ta\Bigr]=0\qquad\text{for all } a>0. \tag{3.19}$$

Proposition 3.2. *Suppose $\boldsymbol{X} \sim K \in D(G)$ with $G(\{0\}) = 0$ and positive recurrent bounded atom A, and $X_0 \sim H$ with $1 - H \in \mathrm{RV}_{-\alpha}$. If $\boldsymbol{X}$ satisfies Condition (3.19), both Conditions (2.14) and (3.12) hold and consequently, the convergence (3.14) takes place.*

Proof. We first show that

$$\lim_{m\to\infty} \limsup_{t\to\infty} t\mathsf{P}_H\Big[X_0/b(t) > \delta, \sup_{m\le j<\tau_A} X_j/b(t) > a\Big] = 0 \qquad \text{for all } a, \delta > 0. \tag{3.20}$$

Indeed, for $c > \delta$, we have,

$$\begin{aligned} t\mathsf{P}_H\left[\frac{X_0}{b(t)} > \delta, \sup_{m\le j<\tau_A} \frac{X_j}{b(t)} > a\right] \le{}& \int_{[\delta,c]} t\mathsf{P}_H\left[\frac{X_0}{b(t)} \in du\right]\mathsf{P}_{b(t)u}\left[\sup_{m\le j<\tau_A} \frac{X_j}{b(t)} > a\right] \\ &+ t\mathsf{P}_H\left[\frac{X_0}{b(t)} > c\right]. \end{aligned}$$

Furthermore, for $\delta \le u \le c$,

$$\mathsf{P}_{b(t)u}\left[\sup_{m\le j<\tau_A} \frac{X_j}{b(t)} > a\right] \le \mathsf{P}_{b(t)u}\left[\sup_{m\le j<\tau_A} \frac{X_j}{b(t)u} > \frac{a}{c}\right] \le \sup_{s\ge b(t)\delta} \mathsf{P}_s\left[\sup_{m\le j<\tau_A} \frac{X_j}{s} > \frac{a}{c}\right].$$

Hence, by Condition (3.19),

$$\lim_{m\to\infty} \limsup_{t\to\infty} t\mathsf{P}_H\left[\frac{X_0}{b(t)} > \delta, \sup_{m\le j<\tau_A} \frac{X_j}{b(t)} > a\right] \le \nu_\alpha[\delta, c] \cdot 0 + \nu_\alpha(c, \infty] = c^{-\alpha}.$$

Letting $c \to \infty$ establishes (3.20). As (3.8) implies that $\sup_{m\le j<\tau(b(t))} X_j \le \sup_{m\le j<\tau_A} X_j$, Condition (2.14) follows. To verify Condition (3.12), argue that

$$\begin{aligned} \mathsf{P}_H\left[\frac{X_0}{b(t)} > \delta, \bigvee_{j=\tau(b(t))}^{\tau_A-1} \frac{X_j}{b(t)} > a\right] \le{}& \mathsf{P}_H\left[\frac{X_0}{b(t)} > \delta, \tau\big(b(t)\big) \le m-1\right] \\ &+ \mathsf{P}_H\left[\frac{X_0}{b(t)} > \delta, \bigvee_{j=m}^{\tau_A-1} \frac{X_j}{b(t)} > a\right], \end{aligned}$$

of which the first term vanishes as $t \to \infty$ because of (3.18) (see [26, Proposition 5.1(d)]). Appeal to (3.20) and let $m \to \infty$ to complete the proof. □

Condition (3.19) is a condition on the transition kernel K; this is best discussed by recalling (see [26, p. 5] for discussion) that a transition kernel $K \in D(G)$ has an update function ψ of the form

$$\psi\big(x, (Z, W)\big) = Zx + \phi(x, W), \tag{3.21}$$

where $Z \sim G$ and $t^{-1}\phi(t, w) \to 0$ for $w \in C$ with $\mathsf{P}[W \in C] = 1$ and we can represent K as

$$K(x, B) = \mathsf{P}\big[\psi\big(x, (Z, W)\big) \in B\big].$$

Take $V_r = (Z_r, W_r)$, i.i.d. copies of $V = (Z, W)$, and write $\boldsymbol{V}_r = (V_1, \dots, V_r)$. For $r \geq 1$ let $\psi^r(x, \boldsymbol{V}_r)$ denote the r-step update function, i.e., $K^r(x, B) = \mathsf{P}[\psi^r(x, \boldsymbol{V}_r) \in B]$, and $\psi^0(x) = x$. By iteration,

$$\psi^r(x, \boldsymbol{V}_r) = \left(\prod_{j=1}^{r} Z_j\right)x + \sum_{\ell=1}^{r-1}\left(\prod_{j=\ell+1}^{r} Z_j\right)\phi\big(\psi^{\ell-1}(x, \boldsymbol{V}_{\ell-1}), W_\ell\big) + \phi\big(\psi^{r-1}(x, \boldsymbol{V}_{r-1}), W_r\big).$$

Thus Condition (3.19) requires both $Z_m \to 0$ as in (2.16), and also an asymptotic stochastic boundedness condition on $\phi(\cdot, W)$. Alternately, one could give criteria for Condition (3.19) using mean drift conditions for $\boldsymbol{X}$ or $\log \boldsymbol{X}$ [18, p. 229].

3.3.2. *Cases where $G(\{0\}) > 0$*

In this case, (2.4) implies

$$\mathsf{P}_{tu}\big[\tau(t) = m\big] \longrightarrow \mathsf{P}\big[\tau^* = m\big], \qquad m \geq 1,$$

where τ^* is a Geometric random variable with parameter $G(\{0\})$. Hence, the tail chain terminates after a finite number of steps. If either $y_0(t) \equiv 0$ is an extremal boundary, or K satisfies the regularity condition (2.5) (p. 1422), Theorem 2.1 assures us that convergence (2.8) holds for $\boldsymbol{X}$ with respect to P_H, and Condition (2.14) follows directly since

$$\begin{aligned}\limsup_{t\to\infty} t\mathsf{P}\Big[\sup_{j\geq m} X_j^{(b(t))}/b(t) > a, X_0 > \delta b(t)\Big] &\leq \limsup_{t\to\infty} t\mathsf{P}\big[X_0 > \delta b(t), \tau\big(b(t)\big) \geq m\big] \\ &= \int_\delta^\infty \nu_\alpha(dx)\mathsf{P}\big[x\xi(m) > 0\big] \to 0 \qquad (m \to \infty).\end{aligned}$$

The regularity condition (2.5) extends to any finite number of steps; that is, iterates of K also satisfy the condition. However, unless $y_0(t) \equiv 0$ is an extremal boundary, we need the regularity condition to hold uniformly over the whole cycle of random length τ_A to prevent $\boldsymbol{X}$ from returning to an extreme state within the same cycle, after crossing below the extremal boundary. Condition (3.22) given next accomplishes this. (Note that even if $y_0(t) \equiv 0$ is an extremal boundary for K, we are using an extremal boundary $y(t)$ chosen to satisfy (3.8).)

$$\lim_{t\to\infty} \mathsf{P}_{tu_t}\Big[\sup_{1\leq j<\tau_A} X_j > ta\Big] = 0 \qquad \text{whenever } u_t = u(t) \to 0, \qquad a > 0. \tag{3.22}$$

Recalling the update function form (3.21), the regularity condition (2.5) holds if the function $\phi(\cdot, w)$ is bounded near 0 for each w in a set of probability 1 [26, Proposition 4.1]. Condition

(3.22) is a stronger boundedness restriction on $\phi(\cdot, w)$ near 0. Alternatively, when K satisfies the regularity condition (2.5), Condition (3.22) may be viewed as a restriction on τ_A, since then

$$\lim_{m\to\infty} \limsup_{t\to\infty} \mathsf{P}_{tu_t}[\tau_A > m] = 0 \qquad \text{whenever } u_t = u(t) \to 0, \tag{3.23}$$

it is sufficient for (3.22). This follows from the decomposition

$$\mathsf{P}_{tu_t}\Big[\sup_{1\le j<\tau_A} X_j > ta\Big] \le \mathsf{P}_{tu_t}[\tau_A > m] + \mathsf{P}_{tu_t}\Big[\sup_{1\le j\le m} X_j > ta\Big],$$

with (3.23) controlling the first right-hand term and (2.5) controlling the second.

Proposition 3.3. *Suppose $\boldsymbol{X} \sim K \in D(G)$ with $G(\{0\}) > 0$ and $1 - H \in \mathrm{RV}_{-\alpha}$ and $\boldsymbol{X}$ has a positive recurrent, bounded atom A. Then $\boldsymbol{X}$ satisfies (2.14) with respect to P_H. Moreover, if either*

(i) *$y_0(t) \equiv 0$ is an extremal boundary for K,*

or

(ii) *Condition (3.22) holds,*

then Condition (3.12) holds with respect to P_H and thus convergence (3.14) takes place.

Proof. First, note that by [26, Proposition 5.1(d)], as $t \to \infty$,

$$\begin{aligned} t\mathsf{P}_H\bigg[\frac{X_0}{b(t)} > \delta, \sup_{j\ge m} \frac{{X_j}^{(b(t))}}{b(t)} > a\bigg] &\le t\mathsf{P}_H\bigg[\frac{X_0}{b(t)} > \delta, \tau\big(b(t)\big) > m\bigg] \\ &\to \delta^{-\alpha}\big(1 - G(\{0\})\big)^m. \end{aligned} \tag{3.24}$$

Since $G(\{0\}) > 0$, the right side of (3.24) vanishes as $m \to \infty$, establishing Condition (2.14). Next, to analyze Condition (3.12), consider the case where $y_0(t) \equiv 0$ is an extremal boundary, and write $\tau_0 = \inf\{n \ge 0 : X_n = 0\}$. For any m,

$$\begin{aligned} t\mathsf{P}_H\Big[X_0/b(t) > \delta, \sup_{\tau(b(t))<j<\tau_A} X_j/b(t) > a\Big] &\le \sum_{r=1}^{m} t\mathsf{P}_H\big[X_0/b(t) > \delta, \tau\big(b(t)\big) = r, \tau_0 > r\big] \\ &\quad + t\mathsf{P}_H\big[X_0/b(t) > \delta, \tau\big(b(t)\big) > m\big], \end{aligned} \tag{3.25}$$

which is obtained by splitting according to whether $\tau(b(t)) \le m$ or the complement and using the fact that $\tau(b(t)) = r$ and $\sup_{\tau(b(t))<j<\tau_A} X_j/b(t) > a$ implies $\tau_0 > r$. For a typical term in the sum,

$$\begin{aligned} t\mathsf{P}_H\big[X_0/b(t) > \delta, \tau\big(b(t)\big) = r, \tau_0 > r\big] &\le t\mathsf{P}_H\big[X_0/b(t) > \delta, \tau_0 > r\big] \\ &\quad - t\mathsf{P}_H\big[X_0/b(t) > \delta, \tau\big(b(t)\big) > r\big] \\ &\longrightarrow \delta^{-\alpha}\big(1 - G(\{0\})\big)^r - \delta^{-\alpha}\big(1 - G(\{0\})\big)^r = 0, \end{aligned}$$

the convergence following from [26, Proposition 5.1(d)], since both $y(t)$ and $y_0(t)$ are extremal boundaries. The right most term in (3.25) is handled as in (3.24).

Finally analyze Condition (3.12) when Condition (3.22) holds. For any m, we have

$$\begin{aligned} &t\mathsf{P}_H\Big[X_0/b(t) > \delta, \sup_{\tau(b(t))<j<\tau_A} X_j/b(t) > a\Big] \\ &\quad \le \sum_{r=1}^{m} t\mathsf{P}_H\Big[X_0/b(t) > \delta, \sup_{r<j<\tau_A} X_j/b(t) > a, \tau\big(b(t)\big) = r\Big] \\ &\qquad + t\mathsf{P}_H\big[X_0/b(t) > \delta, \tau\big(b(t)\big) > m\big], \end{aligned}$$

and

$$\begin{aligned} &t\mathsf{P}_H\Big[X_0/b(t) > \delta, \sup_{r<j<\tau_A} X_j/b(t) > a, \tau\big(b(t)\big) = r\Big] \\ &\quad = \int_{(\delta,\infty]\times(y(b(t)),\infty]^{r-1}\times[0,\infty]} t\mathsf{P}_H\big[(X_0, \boldsymbol{X}_r)/b(t) \in d(x_0, \boldsymbol{x}_r)\big] h_t(x_r), \end{aligned} \tag{3.26}$$

where

$$h_t(x) = \mathbf{1}_{\{[0,y(b(t))]\}}(x)\mathsf{P}_{b(t)x}\Big[\sup_{1\le j<\tau_A} X_j/b(t) > a\Big].$$

We claim that $h_t(u_t) \to 0$ whenever $u_t \to u \ge 0$. Indeed, if $u > 0$, then $h_t(u_t) = 0$ for large t such that $y(b(t)) < u$. Otherwise, $u_t \to 0$, and $h_t(u_t) \to 0$ by Condition (3.22). Therefore, the integral converges to 0 by combining Lemmas 8.2 and 8.4 with Theorem 3.2 from [26]. Applying (3.24) completes the proof. □

3.4. Weak convergence to a cluster process

If the finite-dimensional distributions of $\boldsymbol{X}$ are jointly regularly varying (in the sense of (2.10) with $\boldsymbol{X}$ replacing $\boldsymbol{X}^{(b(t))}$), we obtain a point process limit for $\boldsymbol{X}$ under a condition analogous to Condition (2.17): *There exists* $m_0' \ge 1$ *such that*

$$\lim_{\delta\downarrow 0}\limsup_{t\to\infty} t\mathsf{P}_H\Big[X_0/b(t) \le \delta, \sup_{m_0'\le j<\tau_A} X_j/b(t) > a\Big] = 0 \qquad \text{for all } a > 0. \tag{3.27}$$

Proposition 3.4. *Suppose $\boldsymbol{X} \sim K \in D(G)$ has a positive recurrent, bounded atom A, and $1 - H \in \mathrm{RV}_{-\alpha}$. Assume further that, with respect to P_H, $\boldsymbol{X}$ is regularly varying in the sense of (2.10), with $\boldsymbol{X}$ replacing $\boldsymbol{X}^{(b(t))}$, and satisfies Condition (3.12). Under Condition (3.27), both Conditions (2.17) and (3.13) hold with respect to P_H.*

Proof. Recalling $\sup_{m\le j<\tau(b(t))} X_j \le \sup_{m\le j<\tau_A} X_j$ yields (2.17). Next, given $\delta>0$, write

$$\mathsf{P}_H\Big[\sup_{\tau(b(t))<j<\tau_A}\frac{X_j}{b(t)}>a\Big]\le \mathsf{P}_H\Big[\frac{X_0}{b(t)}>\delta,\ \sup_{\tau(b(t))<j<\tau_A}\frac{X_j}{b(t)}>a\Big] + \mathsf{P}_H\Big[\frac{X_0}{b(t)}\le\delta,\ \sup_{1\le j<\tau_A}\frac{X_j}{b(t)}>a\Big].$$

Condition (3.12) makes the first right side term go to 0 as $t\to\infty$ and for the second term we have,

$$\begin{aligned}
&\limsup_{t\to\infty}\mathsf{P}_H\Big[\frac{X_0}{b(t)}\le\delta,\ \sup_{1\le j<\tau_A}\frac{X_j}{b(t)}>a\Big]\\
&\quad\le \limsup_{t\to\infty}\mathsf{P}_H\Big[\frac{X_0}{b(t)}\le\delta,\ \sup_{1\le j<m_0'}\frac{X_j}{b(t)}>a\Big]+\limsup_{t\to\infty}\mathsf{P}_H\Big[\frac{X_0}{b(t)}\le\delta,\ \sup_{m_0'\le j<\tau_A}\frac{X_j}{b(t)}>a\Big]\\
&\quad=\mu^*\big([0,\delta]\times[\mathbf{0},\boldsymbol{a}]^{c}\big)+\limsup_{t\to\infty}\mathsf{P}_H\Big[\frac{X_0}{b(t)}\le\delta,\ \sup_{m_0'\le j<\tau_A}\frac{X_j}{b(t)}>a\Big]
\end{aligned}$$

where $\boldsymbol{a}=(a,\dots,a)$. Letting $\delta\downarrow 0$, the first term vanishes by (2.10), since $\mu^*(\mathbb{E}^*_{\sqsupset}\setminus\mathbb{E}_{\sqsupset})=0$ [26, Theorem 5.1]. The second term is taken care of by Condition (3.27). □

We now rephrase Theorem 3.1 in terms of our new conditions.

Theorem 3.2. *Let $\boldsymbol{X}$ be a Markov chain on $[0,\infty)$ with transition kernel $K\in D(G)$ such that K has a positive recurrent bounded atom in the sense of* (3.2), (3.3), *and* (3.7). *The initial distribution H has a regularly varying tail and satisfies $tH(b(t)\cdot)\xrightarrow{v}\nu_\alpha(\cdot)$ in $\mathbb{M}_+(0,\infty]$. Assume for any $m\ge 0$ that $(X_0,\dots,X_m)$ is regularly varying in $\mathbb{M}_+([0,\infty]^{m+1}\setminus\{\mathbf{0}\})$,*

$$t\mathsf{P}_H\big[(X_0,\dots,X_m)/b(t)\in(dx_0,d\boldsymbol{x}_m)\big]\xrightarrow{v}\nu_\alpha(dx_0)\mathsf{P}_{x_0}\big[(T_1,\dots,T_m)\in d\boldsymbol{x}_m\big] \tag{3.28}$$

and that Condition (3.27) *holds with respect to P_H.*

(a) *If $G(\{0\})=0$, and K satisfies Condition* (3.19), *then*

$$\sum_{j=0}^{\infty}\epsilon_{(\frac{j}{n},\frac{X_j}{b_n})}\Rightarrow\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\epsilon_{(qt_k,i_k\xi_k(j))}\qquad \text{in } \mathbb{M}_p\big([0,\infty)\times(0,\infty]\big)\qquad \text{as } n\to\infty.$$

(b) *If $G(\{0\})>0$, and either $y_0(t)\equiv 0$ is an extremal boundary for K, or K satisfies Condition* (3.22), *then*

$$\sum_{j=0}^{\infty}\epsilon_{(\frac{j}{n},\frac{X_j}{b_n})}\Rightarrow\sum_{k=0}^{\infty}\sum_{j=0}^{\tau^*_{k+1}-1}\epsilon_{(qt_k,i_k\xi_k(j))}\qquad \text{in } \mathbb{M}_p\big([0,\infty)\times(0,\infty]\big)\qquad \text{as } n\to\infty,$$

where the $\{\tau^*_k\}$ are i.i.d. Geometric random variables with parameter $G(\{0\})$.

4. Proof of Proposition 3.1

Recall that $\{\boldsymbol{X}, \boldsymbol{X}_k, k \geq 0\}$ are i.i.d. copies of a Markov chain $\boldsymbol{X} \sim K$ with heavy tailed initial distribution H satisfying $tH((b(t)\cdot) \xrightarrow{v} \nu_\alpha(\cdot)$ in $\mathbb{M}_+(0,\infty]$ where $b(t) \to \infty$. The extremal boundary downcrossing time by $\boldsymbol{X}_k$ is $\{\tilde{\tau}_k(t)\}$. Let $\{\boldsymbol{\xi}, \boldsymbol{\xi}_k, k \geq 0\}$ be i.i.d. copies of the multiplicative random walk $\boldsymbol{\xi} = \{\xi(m), m \geq 0\}$. The hitting time of 0 by $\boldsymbol{\xi}_k$ is τ_k^*. A PRM on $[0,\infty) \times (0,\infty]$ with mean measure $\mathbb{LEB} \times \nu_\alpha$, independent of the $\{\boldsymbol{\xi}, \boldsymbol{\xi}_k, k \geq 0\}$ is $\zeta = \sum \epsilon_{(t_k, i_k)}$ and $\{S_k\}$ is the renewal process given by (3.4) with finite mean interarrival time q. For convenience, write $\boldsymbol{X}_{k,m}^{(t)} = (X_{k0}^{(t)}, \ldots, X_{km}^{(t)})$ and $\boldsymbol{\xi}_{k,m} = (\xi_k(0), \ldots, \xi_k(m))$.

Proof of Proposition 3.1. (a) First, recall that, under our assumptions $K \in D(G)$ and H having a regularly varying tail, the convergence (2.8) takes place for the chain $\boldsymbol{X}$ on the space $\mathbb{E}_\sqsupset := (0,\infty] \times [0,\infty]^m$, with limit measure μ given by (2.7). This implies that [24, Corollary 6.1, p. 183],

$$\sum_{k=0}^{\infty} \epsilon_{(k/n, \boldsymbol{X}_{k,m}^{(b_n)}/b_n)} = \sum_{k=0}^{\infty} \epsilon_{(k/n, (X_{k0}^{(b_n)}, \ldots, X_{km}^{(b_n)})/b_n)} \Rightarrow \mathrm{PRM}(\mathbb{LEB} \times \mu), \tag{4.1}$$

in $\mathbb{M}_p([0,\infty) \times \mathbb{E}_\sqsupset)$. Since $\sum_{k=0}^\infty \epsilon_{(t_k, i_k, \boldsymbol{\xi}_{k,m})}$ is PRM on $[0,\infty) \times \mathbb{E}_\sqsupset$ ([24, Proposition 5.3, p. 123]), a mapping argument ([24, Proposition 5.2, p. 121]) implies

$$\sum_{k=0}^{\infty} \epsilon_{(t_k, i_k \boldsymbol{\xi}_{k,m})} = \sum_{k=0}^{\infty} \epsilon_{(t_k, i_k, i_k \xi_k(1), \ldots, i_k \xi_k(m))} \sim \mathrm{PRM}(\mathbb{LEB} \times \mu),$$

in $\mathbb{M}_p([0,\infty) \times \mathbb{E}_\sqsupset)$, by (2.7). So we can rewrite (4.1) in $\mathbb{M}_p([0,\infty) \times \mathbb{E}_\sqsupset)$,

$$\vartheta_n = \sum_{k=0}^{\infty} \epsilon_{(k/n, \boldsymbol{X}_{k,m}^{(b_n)}/b_n)} \Rightarrow \sum_{k=0}^{\infty} \epsilon_{(t_k, i_k \boldsymbol{\xi}_{k,m})} = \vartheta. \tag{4.2}$$

Second, we rescale the time axis to place points at the epochs S_k. (See [19].) The counting function for the points $\{S_k\}$ is $N(t) = \sum_k \epsilon_{S_k}[0,t]$ and $N^{\leftarrow}(t) = \inf\{s : N(s) \geq t\} = S_{[t]}$ is the left continuous inverse process. Define $\Theta_n(\cdot) = n^{-1} N^{\leftarrow}(n\cdot)$, so that $S_k/n = \Theta_n(k/n)$ and Θ_n is a random element of $D^\uparrow[0,\infty)$, the subspace of non-decreasing elements of $D_{\text{left}}[0,\infty)$. By the Strong Law of Large Numbers, with probability 1,

$$\Theta_n(t) = \frac{[nt]}{n} \frac{S_{[nt]}}{[nt]} \longrightarrow t \cdot q, \qquad t \geq 0,$$

so $\Theta_n(\cdot) \to q(\cdot)$ in $D^\uparrow[0,\infty)$. We transform time points using the mapping $T_1 : D^\uparrow[0,\infty) \times \mathbb{M}_+([0,\infty) \times \mathbb{E}_\sqsupset) \mapsto \mathbb{M}_+([0,\infty) \times \mathbb{E}_\sqsupset)$ given by

$$T_1 m(f) = \iint f\bigl(x(u), v\bigr) m(du, dv), \qquad f \in \mathcal{C}_K^+\bigl([0,\infty) \times \mathbb{E}_\sqsupset\bigr). \tag{4.3}$$

Applying [24, Proposition 3.1, p. 57] to (4.2), we have $(\Theta_n(\cdot), \vartheta_n) \Rightarrow (q(\cdot), \vartheta)$ in $D^{\uparrow}[0,\infty) \times \mathbb{M}_p([0,\infty) \times \mathbb{E}_{\sqsupset})$. Since T_1 is a.s. continuous at $(q(\cdot), \vartheta)$ (Lemma 4.1, p. 1446), the Continuous Mapping Theorem gives in $\mathbb{M}_p([0,\infty) \times \mathbb{E}_{\sqsupset})$,

$$\eta_n' = \sum_{k=0}^{\infty} \epsilon_{(S_k/n, \boldsymbol{X}_{k,m}^{(b_n)}/b_n)} = T_1(\Theta_n, \vartheta_n) \Rightarrow T_1\big(q(\cdot), \vartheta\big) = \sum_{k=0}^{\infty} \epsilon_{(qt_k, i_k \boldsymbol{\xi}_{k,m})} = \eta'. \tag{4.4}$$

Now stack the components of $\boldsymbol{X}_{k,m}^{(b_n)}$ above the time point S_k/n. To make functionals continuous, it is necessary to compactify the state space by letting $\Lambda_\delta := [\delta,\infty] \times [0,\infty]^m$. Define the restriction functional $T_2 : \mathbb{M}_p([0,\infty) \times \mathbb{E}_{\sqsupset}) \to \mathbb{M}_p([0,\infty) \times \Lambda_\delta)$ by $T_2 m = m(\cdot \cap ([0,\infty) \times \Lambda_\delta))$. From [9, Proposition 3.3], T_2 is almost surely continuous at η' provided $\mathsf{P}[\eta'(\partial([0,\infty) \times \Lambda_\delta)) = 0] = 1$ and since $\mathsf{E}(\eta'(\partial([0,\infty) \times \Lambda_\delta))) = 0$ due to $\nu_\alpha(\{\delta\}) = 0$, the a.s. continuity is verified. Therefore, in $\mathbb{M}_p([0,\infty) \times \Lambda_\delta)$, the restricted version of (4.4) is $\eta_n'' := T_2(\eta_n') \Rightarrow T_2(\eta') =: \eta''$. Define the stacking functional $T_3 : \mathbb{M}_p([0,\infty) \times \Lambda_\delta) \to \mathbb{M}_p([0,\infty) \times [0,\infty])$ by

$$T_3\left(\sum_k \epsilon_{(t_k, y_k(0),\ldots,y_k(m))}\right) = \sum_k \sum_{j=0}^{m} \epsilon_{(t_k, y_k(j))}$$

or for $m \in \mathbb{M}_p([0,\infty) \times \Lambda_\delta)$, $f \in \mathcal{C}_K^+([0,\infty) \times [0,\infty])$, $T_3 m(f) = \iint \{\sum_{j=0}^m f(u, v_j)\} m(du, d\boldsymbol{v})$. Given such f with support in $[0,R] \times [0,\infty]$, R a positive integer, the function $\varphi(u,\boldsymbol{v}) := \sum_{j=0}^m f(u, v_j) \in \mathcal{C}_K^+([0,\infty) \times \Lambda_\delta)$, since it is clearly non-negative continuous, and $\varphi = 0$ outside of $[0,R] \times \Lambda_\delta$. The continuity of T_3 is clear: given $m_n \xrightarrow{v} m$ in $\mathbb{M}_p([0,\infty) \times \Lambda_\delta)$, we have $T_3 m_n(f) = m_n(\varphi) \to m(\varphi) = T_3 m(f)$. Consequently, in $\mathbb{M}_p([0,\infty) \times [0,\infty])$,

$$\hat{\eta}_n = \sum_{k=0}^{\infty} \sum_{j=0}^{m} \epsilon_{(\frac{S_k}{n}, \frac{X_{kj}^{(b_n)}}{b_n})} \mathbf{1}_{\{\frac{X_{k0}}{b_n} \geq \delta\}} = T_3\big(\eta_n''\big) \Rightarrow T_3\big(\eta''\big) = \sum_{k=0}^{\infty} \sum_{j=0}^{m} \epsilon_{(qt_k, i_k \xi_k(j))} \mathbf{1}_{\{i_k \geq \delta\}} = \hat{\eta}. \tag{4.5}$$

Now adjust the sum over j to replace $X_{kj}^{(b_n)}$ with X_{kj}. From (4.5), we readily get,

$$\hat{\eta}_n\big(\cdot \cap \big([0,\infty) \times (0,\infty]\big)\big) \Rightarrow \hat{\eta}\big(\cdot \cap \big([0,\infty) \times (0,\infty]\big)\big) \qquad \text{in } \mathbb{M}_p\big([0,\infty) \times (0,\infty]\big) \tag{4.6}$$

by noting that any $f \in \mathcal{C}_K^+([0,\infty) \times (0,\infty])$ extends to $\bar{f} \in \mathcal{C}_K^+([0,\infty) \times [0,\infty])$ with $\bar{f}(s,0) = 0$ for $s \geq 0$. Moreover, recalling $\{\tilde{\tau}_{k+1}(t)\}$ and $\{\tau_{k+1}^*\}$, the first hitting times of 0 by $\{\boldsymbol{X}_k^{(t)}\}$ and $\{\boldsymbol{\xi}_k\}$ respectively, put

$$\sigma_{k+1}(t) = \tilde{\tau}_{k+1}(t) \wedge (m+1) \quad \text{and} \quad \sigma_{k+1}^* = \tau_{k+1}^* \wedge (m+1), \qquad k \geq 0.$$

Using this notation, the convergence (4.6) becomes, in $\mathbb{M}_p([0,\infty) \times (0,\infty])$,

$$\tilde{\eta}_n = \sum_{k=0}^{\infty} \sum_{j=0}^{\sigma_{k+1}(b_n)-1} \epsilon_{(\frac{S_k}{n}, \frac{X_{kj}}{b_n})} \mathbf{1}_{\{\frac{X_{k0}}{b_n} \geq \delta\}} \Rightarrow \sum_{k=0}^{\infty} \sum_{j=0}^{\sigma_{k+1}^*-1} \epsilon_{(qt_k, i_k \xi_k(j))} \mathbf{1}_{\{i_k \geq \delta\}} = \tilde{\eta}. \tag{4.7}$$

Equation (4.7) allows spreading the stacks of $\tilde{\eta}_n$ in time and we verify

$$\tilde{\eta}_n^* = \sum_{k=0}^{\infty} \sum_{j=0}^{\sigma_{k+1}(b_n)-1} \epsilon_{(\frac{S_k+j}{n}, \frac{X_{kj}}{b_n})} \mathbf{1}_{\{\frac{X_{k0}}{b_n} \geq \delta\}} \Rightarrow \tilde{\eta} \qquad \text{in } \mathbb{M}_p\big([0,\infty) \times (0,\infty]\big). \tag{4.8}$$

This follows from Slutsky's theorem if $d_v(\tilde{\eta}_n^*, \tilde{\eta}_n) \xrightarrow{P} 0$, where d_v is the vague metric on $\mathbb{M}_p([0,\infty) \times (0,\infty])$; it suffices to show that $\mathsf{P}[|\tilde{\eta}_n^*(f) - \tilde{\eta}_n(f)| > \gamma] \to 0$ for any $f \in \mathcal{C}_K^+([0,\infty) \times (0,\infty])$ and $\gamma > 0$. For such f with support in $[0,R] \times [M,\infty]$, for R, M positive integers, we have

$$\begin{aligned}
&\mathsf{P}\big[\big|\tilde{\eta}_n^*(f) - \tilde{\eta}_n(f)\big| > \gamma\big] \\
&\quad = \mathsf{P}\left[\left|\sum_{k=0}^{\infty} \sum_{j=0}^{\sigma_{k+1}(b_n)-1} f\left(\frac{S_k+j}{n}, \frac{X_{kj}}{b_n}\right)\mathbf{1}_{\{\frac{X_{k0}}{b_n} \geq \delta\}} - \sum_{k=0}^{\infty} \sum_{j=0}^{\sigma_{k+1}(b_n)-1} f\left(\frac{S_k}{n}, \frac{X_{kj}}{b_n}\right)\mathbf{1}_{\{\frac{X_{k0}}{b_n} \geq \delta\}}\right| > \gamma\right] \\
&\quad \leq \mathsf{P}\left[\sum_{k=0}^{\infty} \sum_{j=1}^{\sigma_{k+1}(b_n)-1} \left|f\left(\frac{S_k+j}{n}, \frac{X_{kj}}{b_n}\right) - f\left(\frac{S_k}{n}, \frac{X_{kj}}{b_n}\right)\right|\right. \\
&\qquad\qquad\qquad \left. \times\, \epsilon_{(\frac{S_k}{n}, \frac{X_{kj}}{b_n})}\big([0,R] \times [M,\infty]\big)\mathbf{1}_{\{\frac{X_{k0}}{b_n} \geq \delta\}} > \gamma\right].
\end{aligned}$$

Since f is uniformly continuous, given $\rho > 0$, there exists $v > 0$ such that $|f(x) - f(y)| < \rho$ whenever $\|x - y\| < v$. For n so large that $m/n < v$, we have

$$\begin{aligned}
&\sum_{k=0}^{\infty} \sum_{j=1}^{\sigma_{k+1}(b_n)-1} \left|f\left(\frac{S_k+j}{n}, \frac{X_{kj}}{b_n}\right) - f\left(\frac{S_k}{n}, \frac{X_{kj}}{b_n}\right)\right| \epsilon_{(\frac{S_k}{n}, \frac{X_{kj}}{b_n})}\big([0,R] \times [M,\infty]\big)\mathbf{1}_{\{\frac{X_{k0}}{b_n} \geq \delta\}} \\
&\quad < \rho \cdot \tilde{\eta}_n\big([0,R] \times [M,\infty]\big),
\end{aligned}$$

implying that

$$\begin{aligned}
\limsup_{n\to\infty} \mathsf{P}\big[\big|\tilde{\eta}_n^*(f) - \tilde{\eta}_n(f)\big| > \gamma\big] &\leq \limsup_{n\to\infty} \mathsf{P}\big[\tilde{\eta}_n\big([0,R] \times [M,\infty]\big) \geq \gamma\rho^{-1}\big] \\
&\leq \mathsf{P}\big[\tilde{\eta}\big([0,R] \times [M,\infty]\big) \geq \gamma\rho^{-1}\big]
\end{aligned}$$

by (4.7). So (4.8) follows by letting $\rho \to 0$.

Finally, we remove the restriction in m on the stacks. Recall the definitions of η_n and η from Proposition 3.1. To apply a Slutsky argument (e.g., [24], Theorem 3.5, p. 56), we show, for $\gamma > 0$,

$$\lim_{m\to\infty} \mathsf{P}\big[d_v(\tilde{\eta}, \eta) > \gamma\big] = 0 \quad \text{and} \quad \lim_{m\to\infty} \limsup_{n\to\infty} \mathsf{P}\big[d_v\big(\tilde{\eta}_n^*, \eta_n\big) > \gamma\big] = 0. \tag{4.9}$$

Let $f \in \mathcal{C}_K^+([0,\infty) \times (0,\infty])$ with support $[0,R] \times [M,\infty]$. Taking $\delta < a < \infty$, we write

$$\left|\tilde{\eta}(f) - \eta(f)\right| = \sum_{k=0}^{\infty} \sum_{j=m+1}^{\infty} f\big(qt_k, i_k\xi_k(j)\big) \cdot (\mathbf{1}_{\{\delta \le i_k < a\}} + \mathbf{1}_{\{i_k \ge a\}}).$$

Hence,

$$\begin{aligned}
\mathsf{P}\Big[\left|\tilde{\eta}(f) - \eta(f)\right| > \gamma\Big] &\le \mathsf{P}\Bigg[\sum_{k=0}^{\infty} \sum_{j=m+1}^{\infty} f\big(qt_k, i_k\xi_k(j)\big)\mathbf{1}_{\{\delta \le i_k < a\}} > \gamma/2\Bigg] \\
&\quad + \mathsf{P}\Bigg[\sum_{k=0}^{\infty} \sum_{j=m+1}^{\infty} f\big(qt_k, i_k\xi_k(j)\big)\mathbf{1}_{\{i_k \ge a\}} > \gamma/2\Bigg] \\
&= A + B.
\end{aligned}$$

Writing $\xi_k^*(m) = \sup_{j \ge m+1} \xi_k(j)$ for $k \ge 0$, term A is bounded by

$$\begin{aligned}
&\mathsf{P}\Bigg[\sum_{k=0}^{\infty} \Bigg\{\epsilon_{(t_k,i_k)}\big([0,R/q] \times [\delta,a)\big) \sum_{j=m+1}^{\infty} \mathbf{1}_{\{\xi_k(j) > \frac{M}{a}\}}\Bigg\} > 0\Bigg] \\
&\quad \le \mathsf{P}\Big[\zeta_m'\big([0,R/q] \times [\delta,\infty] \times (M/a,\infty)\big) > 0\Big],
\end{aligned}$$

where, since $\{\boldsymbol{\xi}_k\}$ are i.i.d. and independent of ζ, in $\mathbb{M}_p([0,\infty) \times (0,\infty] \times [0,\infty])$,

$$\zeta_m' = \sum_{k=0}^{\infty} \epsilon_{(t_k,i_k,\xi_k^*(m))} \sim \mathrm{PRM}\Big(\mathbb{LEB} \times \nu_\alpha \times \mathsf{P}\Big[\sup_{j \ge m+1} \xi(j) \in \cdot\Big]\Big).$$

Therefore, $\mathsf{P}[\zeta_m'([0,R/q] \times [\delta,\infty] \times (Ma^{-1},\infty)) > 0] = 1 - \exp\{-\lambda\}$, where

$$\begin{aligned}
\lambda &= \mathbb{LEB}[0,R/q] \cdot \nu_\alpha[\delta,\infty] \cdot \mathsf{P}\Big[\sup_{j \ge m+1} \xi(j) > Ma^{-1}\Big] \\
&= Rq^{-1}\delta^{-\alpha}\mathsf{P}\Big[\sup_{j \ge m+1} \xi(j) > Ma^{-1}\Big] \\
&\longrightarrow 0
\end{aligned}$$

as $m \to \infty$ by (2.15), a consequence of Condition (2.14). For term B, we have the bound

$$\mathsf{P}\big[\zeta\big([0,R/q] \times [a,\infty]\big) > 0\big] = 1 - \exp\big\{-\mathsf{E}\zeta\big([0,R/q] \times [a,\infty]\big)\big\} = 1 - \exp\big\{-Rq^{-1}a^{-\alpha}\big\}.$$

Letting $a \to \infty$ establishes the first limit in (4.9).

To prove the second limit in (4.9), observe that

$$\begin{aligned}\mathsf{P}\bigl[\bigl|\tilde{\eta}_n^*(f)-\eta_n(f)\bigr|>\gamma\bigr] &= \mathsf{P}\Biggl[\sum_{k=0}^{\infty}\sum_{j=m+1}^{\tilde{\tau}_{k+1}(b_n)-1} f\biggl(\frac{S_k+j}{n},\frac{X_{kj}}{b_n}\biggr)\mathbf{1}_{\{\frac{X_{k0}}{b_n}\geq\delta\}}>\gamma\Biggr]\\ &\leq \mathsf{P}\Biggl[\sum_{k=0}^{2Rn-1}\sum_{j=m+1}^{\infty} f\biggl(\frac{S_k+j}{n},\frac{X_{kj}^{(b_n)}}{b_n}\biggr)\mathbf{1}_{\{\frac{X_{k0}}{b_n}\geq\delta\}}>0\Biggr]\\ &\quad+\mathsf{P}\Biggl[\sum_{k=2Rn}^{\infty}\sum_{j=m+1}^{\infty} f\biggl(\frac{S_k+j}{n},\frac{X_{kj}^{(b_n)}}{b_n}\biggr)\mathbf{1}_{\{\frac{X_{k0}}{b_n}\geq\delta\}}>0\Biggr].\end{aligned}$$

The first term is bounded by

$$\mathsf{P}\Biggl[\bigcup_{k=0}^{2Rn-1}\Biggl(\biggl\{\frac{X_{k0}}{b_n}\geq\delta\biggr\}\cap\bigcup_{j=m+1}^{\infty}\biggl\{\frac{X_{kj}^{(b_n)}}{b_n}\geq M\biggr\}\Biggr)\Biggr]\leq 2Rn\mathsf{P}\biggl[\frac{X_0}{b_n}\geq\delta,\ \sup_{j\geq m+1}\frac{X_j{}^{(b_n)}}{b_n}\geq M\biggr],$$

and

$$\lim_{m\to\infty}\limsup_{n\to\infty} n\mathsf{P}\biggl[\frac{X_0}{b_n}\geq\delta,\ \sup_{j\geq m+1}\frac{X_j{}^{(b_n)}}{b_n}\geq M\biggr]=0$$

by Condition (2.14). The second term is at most $\mathsf{P}[S_{2Rn}/n\leq R]=\mathsf{P}[S_{2Rn}/2Rn\leq 1/2]\to 0$ as $n\to\infty$, since $S_n/n\to q$ a.s., and $q\geq 1$ by (3.5). This establishes (4.9), completing the proof of part (a).

(b) This amounts to removing the restrictions in δ, under the additional assumptions (2.10) and Condition (2.17). We proceed via a Slutsky argument showing that for any $\gamma>0$,

$$\lim_{\delta\to 0}\mathsf{P}\bigl[d_v\bigl(\eta,\eta^*\bigr)>\gamma\bigr]=0\quad\text{and}\quad\lim_{\delta\to 0}\limsup_{n\to\infty}\mathsf{P}\bigl[d_v\bigl(\eta_n,\eta_n^*\bigr)>\gamma\bigr]=0.\tag{4.10}$$

Let $f\in\mathcal{C}_K^+([0,\infty)\times(0,\infty])$ with support $[0,R]\times[M,\infty]$, and note that

$$\bigl|\eta(f)-\eta^*(f)\bigr|=\sum_{k=0}^{\infty}\sum_{j=0}^{\tau_{k+1}^*-1} f\bigl(qt_k,i_k\xi_k(j)\bigr)\mathbf{1}_{\{i_k<\delta\}}.$$

Hence, writing $\xi_k^*=\sup_{j\geq 1}\xi_k(j)$, and $\zeta'=\sum_{k=0}^{\infty}\epsilon_{(t_k,i_k,i_k\xi_k^*)}$, we have

$$\begin{aligned}\mathsf{P}\bigl[\bigl|\eta(f)-\eta^*(f)\bigr|>\gamma\bigr] &\leq \mathsf{P}\Biggl[\sum_{k=0}^{\infty}\epsilon_{(t_k,i_k)}[0,R/q]\times(0,\delta)\sum_{j=0}^{\infty}\mathbf{1}_{\{i_k\xi_k(j)\geq M\}}>0\Biggr]\\ &\leq \mathsf{P}\bigl[\zeta'\bigl([0,R/q]\times(0,\delta)\times[M,\infty]\bigr)>0\bigr].\end{aligned}$$

The $\{\boldsymbol{\xi}_k\}$ are i.i.d. and independent of ζ, so $\zeta' \sim \mathrm{PRM}(\mu')$ on $\mathbb{M}_p[0,\infty) \times (0,\infty] \times [0,\infty]$ with

$$\mu'(ds, dx, dy) = \mathbb{LEB}(dx) \cdot \nu_\alpha(dx) \cdot \mathsf{P}\Big[\sup_{j\geq 1} \xi(j) \in x^{-1} dy\Big]$$

by [24, Proposition 5.6, p. 144]. Therefore, $\mathsf{P}[\zeta'([0, R/q] \times (0,\delta) \times [M,\infty)) > 0] = 1 - \exp\{-\lambda\}$, where by Lemma 2.1,

$$\lambda = Rq^{-1} \int_{(0,\delta)} \nu_\alpha(dx) \mathsf{P}\Big[\sup_{j\geq 1} \xi(j) \geq Mx^{-1}\Big] \leq Rq^{-1} M^{-\alpha} \mathsf{E}\Big[\sup_{j\geq 1} \xi(j)^\alpha \cdot \mathbf{1}_{\{\sup_{j\geq 1}\xi(j) > M\delta^{-1}\}}\Big].$$

Apply (2.18) and dominated convergence as $\delta \downarrow 0$ to get $\lambda \to 0$ and hence the first limit in (4.10).

For the second limit in (4.10), we have

$$\begin{aligned}
\mathsf{P}\big[\big|\eta_n(f) - \eta_n^*(f)\big| > \gamma\big] &= \mathsf{P}\Bigg[\sum_{k=0}^{\infty} \sum_{j=0}^{\tilde{\tau}_{k+1}(b_n)-1} f\bigg(\frac{S_k + j}{n}, \frac{X_{kj}}{b_n}\bigg) \mathbf{1}_{\{\frac{X_{k0}^{(b_n)}}{b_n} < \delta\}} > \gamma\Bigg] \\
&\leq \mathsf{P}\Bigg[\sum_{k=0}^{2Rn-1} \sum_{j=0}^{\tilde{\tau}_{k+1}(b_n)-1} f\bigg(\frac{S_k + j}{n}, \frac{X_{kj}}{b_n}\bigg) \mathbf{1}_{\{\frac{X_{k0}^{(b_n)}}{b_n} < \delta\}} > 0\Bigg] \\
&\quad + \mathsf{P}\Bigg[\sum_{k=2Rn}^{\infty} \sum_{j=0}^{\tilde{\tau}_{k+1}(b_n)-1} f\bigg(\frac{S_k + j}{n}, \frac{X_{kj}}{b_n}\bigg) \mathbf{1}_{\{\frac{X_{k0}^{(b_n)}}{b_n} < \delta\}} > 0\Bigg].
\end{aligned}$$

As above, the second term is at most $\mathsf{P}[S_{2Rn}/n \leq R] \to 0$ as $n \to \infty$. The first term is bounded by

$$\begin{aligned}
&\mathsf{P}\Bigg[\bigcup_{k=0}^{2Rn-1} \Bigg(\bigg\{\frac{X_{k0}^{(b_n)}}{b_n} < \delta\bigg\} \cap \bigcup_{j=1}^{\tilde{\tau}_{k+1}(b_n)-1} \bigg\{\frac{X_{kj}}{b_n} \geq M\bigg\}\Bigg)\Bigg] \\
&\leq 2Rn\mathsf{P}\Bigg[\frac{X_0^{(b_n)}}{b_n} < \delta, \sup_{j\geq 1} \frac{X_j^{(b_n)}}{b_n} \geq M\Bigg] \\
&\leq 2Rn\mathsf{P}\Bigg[\frac{X_0^{(b_n)}}{b_n} < \delta, \sup_{1\leq j\leq m_0} \frac{X_j^{(b_n)}}{b_n} \geq M\Bigg] + 2Rn\mathsf{P}\Bigg[\frac{X_0^{(b_n)}}{b_n} < \delta, \sup_{j\geq m_0+1} \frac{X_j^{(b_n)}}{b_n} \geq M\Bigg] \\
&= A_n(\delta) + B_n(\delta),
\end{aligned}$$

with m_0 as in Condition (2.17). For $A_n(\delta)$, we have by (2.10),

$$\lim_{\delta\downarrow 0} \limsup_{n\to\infty} A_n(\delta) = \lim_{\delta\downarrow 0} 2R\mu^*\big([0,\delta] \times \big([0, M)^{m_0}\big)^{\mathrm{c}}\big) = 0.$$

For the second, $\lim_{\delta\downarrow 0} \limsup_{n\to\infty} B_n(\delta) = 0$ by Condition (2.17). This establishes (4.10). □

For completeness, the following lemma notes the continuity of the map T_1 defined in (4.3). See also [33,34].

Lemma 4.1. *The mapping* $T_1 : D^{\uparrow}[0,\infty) \times \mathbb{M}_+([0,\infty)\times\mathbb{E}) \to \mathbb{M}_+([0,\infty)\times\mathbb{E})$ *given by*

$$T_1 m(f) = \iint f\bigl(x(u), v\bigr) m(du, dv), \qquad f \in \mathcal{C}_K^+\bigl([0,\infty)\times\mathbb{E}\bigr),$$

is continuous at (x, m) *whenever the function* $x(\cdot)$ *is continuous.*

Proof. (a) Suppose $x_n \to x_0$ in $D^{\uparrow}[0,\infty)$ (with respect to the Skorohod topology), where x_0 is continuous, and $m_n \stackrel{v}{\to} m_0$ in $\mathbb{M}_+([0,\infty)\times\mathbb{E})$. Let $f \in \mathcal{C}_K^+([0,\infty)\times\mathbb{E})$ with support contained in $[0,R]\times B$. We show that $T_1 m_n(f) \to T_1 m_0(f)$. For $n \ge 0$, write $f_n(u,v) = f(x_n(u), v)$. The f_n are supported on $x_n^{-1}([0,R]) \times B$, and $x_n^{-1}([0,R]) = [0, x_n^{\leftarrow}(R)]$, where $x_n^{\leftarrow}$ is the right-continuous inverse of x_n. We now argue that the f_n, $n \ge 0$, have a common compact support. Indeed, we have $x_n^{\leftarrow} \to x_0^{\leftarrow}$ pointwise, so $x_n^{\leftarrow}(R) \to x_0^{\leftarrow}(R)$. Thus, for large n, $[0, x_n^{\leftarrow}(R)] \times B \subset [0, x_0^{\leftarrow}(R)+1] \times B$; without loss of generality, $m_0(\partial([0, x_0^{\leftarrow}(R)+1]\times B)) = 0$. Furthermore, $f_n \to f_0$ uniformly: suppose $(u_n, v_n) \to (u_0, v_0) \in [0,\infty)\times\mathbb{E}$. Then $x_n(u_n) \to x_0(u_0)$ since x_0 is continuous, and so $f(x_n(u_n), v_n) \to f(x_0(u_0), v_0)$ by the continuity of f. Consequently, $\tilde{m}_n(f) \to \tilde{m}_0(f)$ by [26, Lemma 8.2(b)]. □

Acknowledgements

S.I. Resnick and D. Zeber were partially supported by ARO Contract W911NF-10-1-0289 and NSA Grant H98230-11-1-0193 at Cornell University.

References

[1] Asmussen, S. (1998). Extreme value theory for queues via cycle maxima. *Extremes* **1** 137–168. MR1814621

[2] Asmussen, S. (2003). *Applied Probability and Queues*, 2nd ed. *Applications of Mathematics* (*New York*) **51**. New York: Springer. MR1978607

[3] Balan, R.M. and Louhichi, S. (2009). Convergence of point processes with weakly dependent points. *J. Theoret. Probab.* **22** 955–982. MR2558660

[4] Basrak, B. and Segers, J. (2009). Regularly varying multivariate time series. *Stochastic Process. Appl.* **119** 1055–1080. MR2508565

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

[6] Borkovec, M. (2000). Extremal behavior of the autoregressive process with ARCH(1) errors. *Stochastic Process. Appl.* **85** 189–207. MR1731021

[7] Davis, R.A. and Hsing, T. (1995). Point process and partial sum convergence for weakly dependent random variables with infinite variance. *Ann. Probab.* **23** 879–917. MR1334176

[8] Durrett, R. (1991). *Probability*: *Theory and Examples*. *The Wadsworth & Brooks/Cole Statistics/Probability Series*. Pacific Grove, CA: Wadsworth & Brooks/Cole Advanced Books & Software. MR1068527

[9] Feigin, P.D., Kratz, M.F. and Resnick, S.I. (1996). Parameter estimation for moving averages with positive innovations. *Ann. Appl. Probab.* **6** 1157–1190. MR1422981

[10] Hsing, T. (1987). On the characterization of certain point processes. *Stochastic Process. Appl.* **26** 297–316. MR0923111

[11] Hsing, T. (1989). Extreme value theory for multivariate stationary sequences. *J. Multivariate Anal.* **29** 274–291. MR1004339

[12] Hsing, T., Hüsler, J. and Leadbetter, M.R. (1988). On the exceedance point process for a stationary sequence. *Probab. Theory Related Fields* **78** 97–112. MR0940870

[13] Kallenberg, O. (1975). *Random Measures*. Berlin: Akademie-Verlag. Schriftenreihe des Zentralinstituts für Mathematik und Mechanik bei der Akademie der Wissenschaften der DDR, Heft 23. MR0431372

[14] 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

[15] Leadbetter, M.R. and Rootzén, H. (1988). Extremal theory for stochastic processes. *Ann. Probab.* **16** 431–478. MR0929071

[16] Maulik, K., Resnick, S. and Rootzén, H. (2002). Asymptotic independence and a network traffic model. *J. Appl. Probab.* **39** 671–699. MR1938164

[17] Meyn, S. and Tweedie, R.L. (2009). *Markov Chains and Stochastic Stability*, 2nd ed. Cambridge: Cambridge Univ. Press. With a prologue by Peter W. Glynn. MR2509253

[18] Meyn, S.P. and Tweedie, R.L. (1993). *Markov Chains and Stochastic Stability*. *Communications and Control Engineering Series*. London: Springer London Ltd. MR1287609

[19] Mikosch, T. and Resnick, S. (2006). Activity rates with very heavy tails. *Stochastic Process. Appl.* **116** 131–155. MR2197971

[20] Neveu, J. (1977). Processus ponctuels. In *École D'Été de Probabilités de Saint-Flour, VI—1976*. *Lecture Notes in Math.* **598** 249–445. Berlin: Springer. MR0474493

[21] Novak, S.Y. (2002). Multilevel clustering of extremes. *Stochastic Process. Appl.* **97** 59–75. MR1870720

[22] Perfekt, R. (1994). Extremal behaviour of stationary Markov chains with applications. *Ann. Appl. Probab.* **4** 529–548. MR1272738

[23] Perfekt, R. (1997). Extreme value theory for a class of Markov chains with values in $\mathbf{R}^d$. *Adv. in Appl. Probab.* **29** 138–164. MR1432934

[24] Resnick, S.I. (2007). *Heavy-Tail Phenomena*: *Probabilistic and Statistical Modeling*. *Springer Series in Operations Research and Financial Engineering*. New York: Springer. MR2271424

[25] Resnick, S.I. (2008). *Extreme Values*, *Regular Variation and Point Processes*. *Springer Series in Operations Research and Financial Engineering*. New York: Springer. Reprint of the 1987 original. MR2364939

[26] Resnick, S.I. and Zeber, D. (2013). Asymptotics of Markov kernels and the tail chain. *Adv. in Appl. Probab.* **45** 186–213.

[27] Rootzén, H. (1988). Maxima and exceedances of stationary Markov chains. *Adv. in Appl. Probab.* **20** 371–390. MR0938151

[28] Segers, J. (2003). Functionals of clusters of extremes. *Adv. in Appl. Probab.* **35** 1028–1045. MR2014268

[29] Segers, J. (2005). Approximate distributions of clusters of extremes. *Statist. Probab. Lett.* **74** 330–336. MR2186477

[30] Segers, J. (2007). Multivariate regular variation of heavy-tailed Markov chains. Preprint. Available at arXiv:math/0701411.

[31] Smith, R.L. (1992). The extremal index for a Markov chain. *J. Appl. Probab.* **29** 37–45. MR1147765

[32] Smith, R.L., Tawn, J.A. and Coles, S.G. (1997). Markov chain models for threshold exceedances. *Biometrika* **84** 249–268. MR1467045

[33] Whitt, W. (1980). Some useful functions for functional limit theorems. *Math. Oper. Res.* **5** 67–85. MR0561155

[34] 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

[35] Yun, S. (1998). The extremal index of a higher-order stationary Markov chain. *Ann. Appl. Probab.* **8** 408–437. MR1624945

[36] Yun, S. (2000). The distributions of cluster functionals of extreme events in a dth-order Markov chain. *J. Appl. Probab.* **37** 29–44. MR1761659