*The Annals of Applied Probability*
1994, Vol. 4, No. 2, 529–548

# EXTREMAL BEHAVIOUR OF STATIONARY MARKOV CHAINS WITH APPLICATIONS

BY ROLAND PERFEKT

*University of Lund*

In this paper the extremal behaviour of real-valued, stationary Markov chains is studied under fairly general assumptions. Conditions are obtained for convergence in distribution of multilevel exceedance point processes associated with suitable families of high levels. Although applicable to general stationary sequences, these conditions are tailored for Markov chains and are seen to hold for a large class of chains. The extra assumptions used are that the marginal distributions belong to the domain of attraction of some extreme value law together with rather weak conditions on the transition probabilities. Also, a complete convergence result is given. The results are applied to an AR(1) process with uniform margins and to solutions of a first order stochastic difference equation with random coefficients.

**1. Introduction.** Markov models are valuable in a very wide variety of applied problems. Examples are random coefficient autoregressive models which find widespread use in nonlinear time series analysis and as the ARCH processes in econometrics (see, e.g., [21] and [6]), the Lindley process in queueing theory and the max-autoregressive sequences introduced in [2]. In many of these situations extreme values are of vital importance; for example, in economics they may correspond to bankruptcy or to a forced exchange rate realignment, and in queueing theory an excessive queue length may cause a breakdown of the entire system. Although the extremal properties of the specific models mentioned can be examined by a direct approach (see [9], [19] and [2], respectively), all three fit well into a more general setting considered here.

Let $\{X_k;\ k \geq 0\}$ be a not necessarily Markovian stationary sequence with marginal distribution $F$. Assume that for each $\tau > 0$ there exists a sequence $\{u_n(\tau)\}$ of real numbers such that

$$n\big(1 - F(u_n(\tau))\big) \to \tau \quad \text{as } n \to \infty. \tag{1.1}$$

Our main concern is the asymptotic distribution of the point processes $N_n^{(\tau)}$, $\tau > 0$, of time-normalized exceedances of $u_n(\tau)$ by $\{X_k\}$, defined by

$$N_n^{(\tau)}(A) = \#\{k;\, k/n \in A,\, X_k > u_n(\tau)\}$$

for Borel sets $A \subseteq [0, \infty)$.

Received December 1991; revised September 1993.
*AMS* 1991 *subject classifications*. Primary 60G70, 60J05; secondary 60G10, 60G55.
*Key words and phrases.* Extreme values, exceedance point processes, stationary Markov chains.

Fix an integer $r \geq 1$ and $\tau_1 > \cdots > \tau_r > 0$. Write $\mathbf{N}_n$ for the vector of point processes $(N_n^{(\tau_1)}, \dots, N_n^{(\tau_r)})$ and, with a slight abuse of notation, let $\mathbf{N}_n(I)$ denote $(N_n^{(\tau_1)}(I), \dots, N_n^{(\tau_r)}(I))$, $I$ being an interval in $[0, \infty)$. Suppose that $\{X_k\}$ satisfies a mild long range dependence restriction and let $\{p_n\}$ be a sequence of positive integers that is increasing with $n$ at a suitable rate; see Section 2 for details. Hsing [11], [13] shows that if $\{X_k\}$ has extremal index $\theta > 0$ (see Section 2), then a necessary and sufficient condition for convergence in distribution of $\mathbf{N}_n$ as $n \to \infty$ is convergence of the conditional distribution of $\mathbf{N}_n((0, p_n/n])$ given that there is at least one exceedance of $u_n(\tau_1)$ in $[1, p_n]$. He further characterizes the limit $(N^{(\tau_1)}, \dots, N^{(\tau_r)})$. In particular it is seen that the marginals $N^{(\tau_i)}$ are compound Poisson processes with intensity $\theta\tau_i$ for the Poisson events and compounding probabilities independent of $\tau$. Rootzén [19] gives conditions for convergence of $N_n$ for $r = 1$ which are equivalent to those of Hsing. These conditions are easier to verify for Markov chains.

The extremal index, which measures "the clustering of extremes," is important for the extremal properties of stationary sequences. O'Brien [17] and Rootzén [19] give similar characterizations of the extremal index which are tailored for applications to Markov chains. In a recent paper, Smith [20] assumes standard Gumbel marginals and the existence of a well behaved transition density, and shows that the chain looks like a random walk when the initial value is chosen large enough. More precisely, given $X_0 = u$ tending to $\infty$, $X_1 - X_0, X_2 - X_1, \dots, X_p - X_{p-1}$ are asymptotically i.i.d. random variables (r.v.'s). As a consequence, the extremal index is found as the solution to a Wiener–Hopf integral equation.

In the present paper we introduce the notion of a *tail chain* which determines the extremal characteristics of (a stationary version of) the original Markov sequence. It is shown that the assumptions of absolutely continuous marginals and the existence of a transition density are not needed for the result of [20]. In fact we only require that the marginal distribution of the chain is in the domain of attraction of some extreme value law together with rather weak conditions on the transition probabilities. Furthermore we establish convergence in distribution of the multilevel exceedance point processes, $\mathbf{N}_n$, and thus give a rather complete description of the extremal properties. Expressions are given for the extremal index and for other quantities characterizing the distributional limits. Explicit calculation is seldom possible, but the expressions are well suited for simulation. The results give, for instance, asymptotic joint distributions for the $k$th largest maxima under appropriate normalizations.

In Section 2 we consider criteria for weak convergence of $\mathbf{N}_n$ and make extensions of the one-level result of [19] to the multilevel case and to complete convergence; see Theorems 2.5 and 2.6, respectively. Extremes of stationary Markov chains are considered in Section 3, whereas Section 4 contains two examples. The first, concerning an AR(1) process with uniform marginals, is mainly of theoretical interest. The second one finds the extremal behaviour of solutions of first order stochastic difference equations with random coefficients, generalizing results of [9].

**2. Exceedance point processes for general stationary sequences.** In this section we find conditions for convergence in distribution of the point process $\mathbf{N}_n = (N_n^{(\tau_1)}, \dots, N_n^{(\tau_r)})$ defined in Section 1. The conditions are convenient for applications to Markov chains, a subject that will be further developed in Section 3. Each marginal, $N_n^{(\tau_i)}$, of $\mathbf{N}_n$ is a point process on $[0, \infty)$. Weak convergence of $N_n^{(\tau_i)}$ is as defined in [14] and convergence of the vector $\mathbf{N}_n$ is with respect to the product topology. Thus, $\mathbf{N}_n$ converges to some $\mathbf{N} = (N^{(\tau_1)}, \dots, N^{(\tau_r)})$ if and only if $(N_n^{(\tau_i)}(B_{i,j});\ i = 1, \dots, r,\ j = 1, \dots, k_i)$ converges in distribution to $(N^{(\tau_i)}(B_{i,j});\ i = 1, \dots, r,\ j = 1, \dots, k_i)$ for $k_i = 1, 2, \dots$ and $B_{i,j}$ bounded Borel sets with $\mathbf{P}\{N^{(\tau_i)}(\partial B_{i,j}) > 0\} = 0$, where $\partial B$ is the boundary of $B$.

Assume, as in Section 1, that $\{X_k;\ k \ge 0\}$ is stationary with marginal distribution $F$ and that for each $\tau > 0$ there is a sequence $\{u_n(\tau)\}$ of real numbers satisfying (1.1). The sequence $\{X_k\}$ will be required to satisfy certain long range dependence restrictions. Loosely speaking, extreme events situated far apart are assumed to be almost independent.

Definition 2.1. *Let $\{u_n^{(m)};\ n \ge 1\}$, $m = 1, 2, \dots, r$, be sequences of real numbers. For $1 \le i \le j \le n$, define $\mathscr{B}_i^j = \mathscr{B}_i^j(u_n^{(1)}, \dots, u_n^{(r)})$ to be the $\sigma$-field generated by the events $\{X_k \le u_n^{(m)}\}$, $i \le k \le j$, $1 \le m \le r$, and let $\mathscr{A}_i^j$ be the subclass of $\mathscr{B}_i^j$ consisting of intersections of events in the generating class. For $1 \le l \le n-1$, write*

$$\beta_{n,l} = \max\left\{|\mathbf{P}(A \cap B) - \mathbf{P}(A)\mathbf{P}(B)|\colon A \in \mathscr{B}_1^j,\ B \in \mathscr{B}_{j+l}^n,\ 1 \le j \le n-l\right\}$$

*and let $\alpha_{n,l}$ $(\le \beta_{n,l})$ be the corresponding maximum when the $A$'s and $B$'s are restricted to $\mathscr{A}_1^j$ and $\mathscr{A}_{j+l}^n$, respectively. The condition $\Delta(u_n^{(1)}, \dots, u_n^{(r)})$ is said to hold for $\{X_k\}$ if $\beta_{n,l_n} \to 0$ as $n \to \infty$ for some sequence $l_n = o(n)$. Similarly, the condition $D(u_n^{(1)}, \dots, u_n^{(r)})$ holds for $\{X_k\}$ if $\alpha_{n,l_n} \to 0$ as $n \to \infty$ for some $l_n = o(n)$.*

The condition $D(u_n^{(1)}, \dots, u_n^{(r)})$ is a frequently used mixing condition due to Leadbetter, whereas the slightly stronger condition $\Delta(u_n^{(1)}, \dots, u_n^{(r)})$ was introduced by Hsing [11]. Both conditions are implied by strong mixing.

Typically, exceedances of a given level $u_n$ appear in clusters. To identify these we divide $\{1, 2, \dots\}$ into intervals of equal length, depending on $n$, and assign exceedances from different intervals to separate clusters. The following definition makes precise the interval length, $p_n$, needed to accomplish asymptotic independence of extreme events in separate intervals (cf. [11]).

Definition 2.2. *Suppose that $\{X_k\}$ satisfies the condition $\Delta(u_n^{(1)}, \dots, u_n^{(r)})$. A sequence of positive integers $\{p_n;\ n \ge 1\}$ is said to be $\Delta(u_n^{(1)}, \dots, u_n^{(r)})$-separating if, as $n \to \infty$, $p_n/n \to 0$ and there exists a sequence $\{l_n\}$ such that $l_n/p_n \to 0$ and $n\beta_{n,l_n}/p_n \to 0$, where $\beta_{n,l}$ is given in Definition 2.1.*

Of course the $l_n$ in the definition must satisfy $\beta_{n,l_n} \to 0$. Conversely, any $l_n = o(n)$ satisfying $\beta_{n,l_n} \to 0$ can be used in constructing a $\Delta(\cdot)$-separating sequence $p_n$ by taking $p_n$ to the integer part of $\max(n\beta_{n,l_n}^{1/2}, (nl_n)^{1/2})$. A $D(u_n^{(1)}, \dots, u_n^{(r)})$-separating sequence is defined in the same manner using the $\{\alpha_{n,l}\}$ instead.

Let $M_{i,j} = \max\{X_{i+1}, \dots, X_j\}$ for $i < j$ and write $M_j$ for $M_{0,j}$. The sequence $\{X_k\}$ is said to have extremal index $\theta$ $(0 \le \theta \le 1)$ if $\mathbf{P}\{M_n \le u_n(\tau)\} \to e^{-\theta\tau}$ or equivalently if $\mathbf{P}\{M_n \le u_n(\tau)\} \sim F^{n\theta}(u_n(\tau))$ as $n \to \infty$, for each $\tau > 0$. For easy reference we next state O'Brien's [17] result on the extremal index for a stationary sequence (cf. also Rootzén [19]). Here $D(u_n(\tau))$ is the version of $D(u_n^{(1)}, \dots, u_n^{(r)})$ with $r = 1$ and $u_n^{(1)} = u_n(\tau)$.

PROPOSITION 2.3. *Suppose $D(u_n(\tau))$ holds for $\{X_k\}$, for each $\tau > 0$. Then $\{X_k\}$ has extremal index $\theta$ if and only if for some $\tau_0 > 0$ there is a $D(u_n(\tau_0))$-separating sequence $\{p_n\}$ such that*

$$\mathbf{P}\left\{M_{p_n} \le u_n(\tau_0) | X_0 > u_n(\tau_0)\right\} \to \theta \quad \text{as } n \to \infty. \tag{2.1}$$

We remark that if $\{X_k\}$ has extremal index $\theta$ and $\{u_n(\tau)\}$ is any sequence such that $n(1 - F(u_n(\tau))) \to \tau > 0$ and $D(u_n(\tau))$ holds, then it can be seen that (2.1) holds with $u_n(\tau_0)$ replaced by $u_n(\tau)$ for every $D(u_n(\tau))$-separating sequence.

Now, fix an integer $r \ge 1$ and let $\tau_1 > \cdots > \tau_r > 0$. Then, by (1.1), we have $u_n(\tau_1) < \cdots < u_n(\tau_r)$ for all sufficiently large $n$. It will here be convenient to work with a simple transformation of the exceedance processes. Specifically, define $\tilde{N}_n^{(j)}$ to be the point process of exceedances of $u_n(\tau_j)$ that do not exceed $u_n(\tau_{j+1})$, that is, $\tilde{N}_n^{(j)} = N_n^{(\tau_j)} - N_n^{(\tau_{j+1})}$, $j = 1, \dots, r-1$, and $\tilde{N}_n^{(r)} = N_n^{(\tau_r)}$, and let $\tilde{\mathbf{N}}_n := (\tilde{N}_n^{(1)}, \dots, \tilde{N}_n^{(r)})$. Further, for fixed $\mathbf{i} = (i_1, \dots, i_r) \in \mathscr{Z}_+^r := \{0, 1, \dots\}^r$, let $A_j = A_j(\mathbf{i})$ denote the set $\{\mathbf{l} \in \mathscr{Z}^r;\ l_j = i_j - 1,\ l_k \ge i_k,\ k \ne j\}$ and put $A_0 = A_0(\mathbf{i}) = \{\mathbf{l} \in \mathscr{Z}_+^r;\ l_k \ge i_k,\ k = 1, 2, \dots, r\}$. The following lemma, which will be the key part in proving Theorem 2.5, relates the unconditional law of the processes to the conditional law given $X_0$. In the proof we make use of the following observation: For events $E_1, \dots, E_n$ such that every intersection consisting of more than $m$ $(\le n)$ different $E_i$'s has zero probability, it holds that

$$\sum_{i=1}^{n} \mathbf{P}(E_i) = \sum_{j=1}^{m} \mathbf{P}\left\{\bigcup_{i_1 < \cdots < i_j} \left(E_{i_1} \cap \cdots \cap E_{i_j}\right)\right\}.$$

LEMMA 2.4. *Suppose that $\{X_k\}$ satisfies $D(u_n(\tau_1))$ and let $i_j$, $j = 1, \dots, r$, be nonnegative integers with at least one distinct from zero. Then, with $\tau_{r+1} = 0$ and $u_n(\tau_{r+1}) = \infty$,*

$$\begin{aligned} \sum_{j=1}^{r} \left(\frac{\tau_j - \tau_{j+1}}{\tau_1}\right) \mathbf{P}\left\{\tilde{\mathbf{N}}_n\left(\left(0, \frac{p_n}{n}\right]\right) \in A_j | X_0 \in (u_n(\tau_j), u_n(\tau_{j+1})]\right\} \\ - n\mathbf{P}\left\{\tilde{\mathbf{N}}_n\left(\left(0, \frac{p_n}{n}\right]\right) \in A_0\right\} \Big/ (p_n\tau_1) \to 0 \quad \text{as } n \to \infty \end{aligned} \tag{2.2}$$

*for every $D(u_n(\tau_1))$-separating sequence $\{p_n\}$.*

PROOF. Assume first that each $i_j$ is strictly positive. Let $\{p_n\}$ be $D(u_n(\tau_1))$-separating and let $\nu_j(m) = \max\{l \le p_n + m;\ \tilde{N}_n^{(j)}([l/n, (p_n + m)/n]) = i_j\}$ for $m \ge 0$, $j = 1, \dots, r$; that is, $\nu_j(m)$ is the last time before $p_n + m$ for which the interval $[l, p_n + m]$ contains $i_j$ exceedances of $u_n(\tau_j)$ not exceeding $u_n(\tau_{j+1})$. With $c = \sum_{j=1}^r i_j$ and using stationarity, we have

$$
\begin{aligned}
&\mathbf{P}\{\tilde{\mathbf{N}}_n((0, p_n/n]) \in A_0\} \\
&\quad = \sum_{j=1}^{r} \sum_{k=1}^{p_n - c + 1} \mathbf{P}\{\nu_j(0) = k,\ \nu_l(0) > \nu_j(0);\ l \ne j\} \\
&\quad = \sum_{j=1}^{r} \sum_{k=1}^{p_n - c + 1} \mathbf{P}\Big\{X_k \in (u_n(\tau_j), u_n(\tau_{j+1})], \\
&\qquad\qquad \tilde{\mathbf{N}}_n((k/n, p_n/n]) \in A_j\Big\} \\
&\quad \ge \sum_{j=1}^{r} \sum_{k=1}^{p_n - c + 1} \mathbf{P}\Big\{X_k \in (u_n(\tau_j), u_n(\tau_{j+1})], \\
&\qquad\qquad \tilde{\mathbf{N}}_n((k/n, (p_n + k)/n]) \in A_j\Big\} \\
&\qquad - \sum_{j=1}^{r} \sum_{k=1}^{p_n - c + 1} \mathbf{P}\{\nu_j(k) = k,\ M_{p_n, 2p_n} > u_n(\tau_1)\} \\
&\quad \ge \sum_{j=1}^{r} (p_n - c + 1)\mathbf{P}\Big\{X_0 \in (u_n(\tau_j), u_n(\tau_{j+1})], \\
&\qquad\qquad \tilde{\mathbf{N}}_n((0, p_n/n]) \in A_j\Big\} \\
&\qquad - c\mathbf{P}\{M_{p_n} > u_n(\tau_1),\ M_{p_n, 2p_n} > u_n(\tau_1)\}.
\end{aligned} \tag{2.3}
$$

In the third step we used $\mathbf{P}(A) \ge \mathbf{P}(B) - \mathbf{P}(B \cap A^c)$ for each of the summands. The last inequality follows from the discussion just preceeding the lemma, since at most $i_j$ of the events $\{\nu_j(k) = k\}$, $k \le p_n - c + 1$, can occur at the same time.

For an inequality in the other direction, partition the sample space according to whether there is an exceedance of $u_n(\tau_1)$ or not in the interval $(p_n, 2p_n]$. In the same manner as before, we get

$$
\begin{aligned}
&\mathbf{P}\{\tilde{\mathbf{N}}_n((0, p_n/n]) \in A_0\} \\
&\quad \le \sum_{j=1}^{r} (p_n - c + 1)\mathbf{P}\Big\{X_0 \in (u_n(\tau_j), u_n(\tau_{j+1})], \\
&\qquad\qquad \tilde{\mathbf{N}}_n((0, p_n/n]) \in A_j\Big\} \\
&\qquad + \mathbf{P}\{M_{p_n} > u_n(\tau_1),\ M_{p_n, 2p_n} > u_n(\tau_1)\}.
\end{aligned} \tag{2.4}
$$

Now, combining (2.3) and (2.4), (2.2) will follow from (1.1) if we show that $n\mathbf{P}\{M_{p_n} > u_n(\tau_1),\ M_{p_n, 2p_n} > u_n(\tau_1)\}/p_n \to 0$ as $n \to \infty$. However, this follows using standard arguments; see, for example, the proof of Lemma 3.3.1 in [16], since $D(u_n(\tau_1))$ holds and $\{p_n\}$ was assumed $D(u_n(\tau_1))$-separating.

The case when some $i_j$'s equal zero is treated similarly. □

Before stating the main result of this section, we introduce some further notation. For a fixed integer $r \geq 1$, let $\mathscr{I} = \{\mathbf{i} \in \mathscr{Z}_+^r;\ i_1 \geq \cdots \geq i_r,\ i_1 \geq 1\}$. Further, for fixed $\mathbf{i} \in \mathscr{I}$ write $\mathbf{i}_j$ for the $r$-vector with $k$th component equal to $i_k - 1$ if $k \leq j$ and $i_k$ otherwise, $j = 0, 1, \ldots, r$ (i.e., $\mathbf{i}_0 = \mathbf{i}$). Finally, let $P_j$ be the conditional distribution of $\{X_k\}$ given $X_0 > u_n(\tau_j)$, $j = 1, \ldots, r$. We have the following multilevel version of Theorem 4.1(ii) in [19].

THEOREM 2.5. *Suppose the stationary sequence $\{X_k;\ k \geq 0\}$ has extremal index $\theta > 0$ and that $\Delta(u_n(\sigma\tau_1), \ldots, u_n(\sigma\tau_r))$ holds for $\{X_k\}$ for each $\sigma > 0$ and some $\tau_1 > \cdots > \tau_r > 0$. Then $\mathbf{N}_n = (N_n^{(\tau_1)}, \ldots, N_n^{(\tau_r)})$ converges in distribution to some point process $(N^{(\tau_1)}, \ldots, N^{(\tau_r)})$ as $n \to \infty$ if and only if for all $\mathbf{i} \in \mathscr{I}$ there exists a constant $\pi(\mathbf{i})$ such that*

$$(2.5)\qquad \sum_{j=1}^{r} \frac{\tau_j}{\tau_1}\left(P_j\left\{\mathbf{N}_n\left(\left(0, \frac{p_n}{n}\right]\right) = \mathbf{i}_j\right\} - P_j\left\{\mathbf{N}_n\left(\left(0, \frac{p_n}{n}\right]\right) = \mathbf{i}_{j-1}\right\}\right) \to \theta\pi(\mathbf{i})$$

*as $n \to \infty$, for some $\Delta(u_n(\tau_1), \ldots, u_n(\tau_r))$-separating sequence $\{p_n\}$. In this case, $\pi$ is a probability measure on $\mathscr{I}$ and $(N_n^{(\sigma\tau_1)}, \ldots, N_n^{(\sigma\tau_r)})$ converges in distribution to a point process with Laplace transform*

$$(2.6)\qquad \begin{aligned} &\mathbf{E}\exp\left(-\sum_{j=1}^{r}\int_0^\infty f_j\, dN^{(\sigma\tau_j)}\right) \\ &\quad = \exp\left(-\theta\sigma\tau_1\int_0^\infty (1 - L(f_1(t), \ldots, f_r(t)))\, dt\right) \end{aligned}$$

*for each $\sigma > 0$, where $N^{(\sigma\tau_j)}$ is the jth marginal of the limiting process, $f_j \geq 0$ and $L$ is the Laplace transform of $\pi$.*

PROOF. It follows from [11], Theorems 4.2.3 and 4.2.4, that it is enough to show that (2.5) is equivalent to

$$\mathbf{P}\left\{\mathbf{N}_n((0, p_n/n]) = \mathbf{i} \mid N_n^{(\tau_1)}((0, p_n/n]) > 0\right\} \to \pi(\mathbf{i}) \quad \text{as } n \to \infty.$$

Using Lemma 2.4 with $r = 1$, $i_1 = 1$, together with Proposition 2.3 and the subsequent remark, we see that $n\mathbf{P}\{N_n^{(\tau_1)}((0, p_n/n]) > 0\}/(p_n\tau_1) \to \theta$ as $n \to \infty$, for every $D(u_n(\tau_1))$-separating sequence $\{p_n\}$. Since any $\Delta(u_n(\tau_1), \ldots, u_n(\tau_r))$-separating sequence is also $D(u_n(\tau_1))$-separating, we therefore only have to show that

$$\sum_{j=1}^{r} \frac{\tau_j}{\tau_1}\left(P_j\left\{\mathbf{N}_n\left(\left(0, \frac{p_n}{n}\right]\right) = \mathbf{i}_j\right\} - P_j\left\{\mathbf{N}_n\left(\left(0, \frac{p_n}{n}\right]\right) = \mathbf{i}_{j-1}\right\}\right)$$
$$- n\mathbf{P}\left\{\mathbf{N}_n\left(\left(0, \frac{p_n}{n}\right]\right) = \mathbf{i}\right\}\Big/(p_n\tau_1) \to 0 \quad \text{as } n \to \infty.$$

However, this follows from Lemma 2.4 by a straightforward although somewhat tedious calculation; we omit the details. □

The Laplace transform $L$ appearing in (2.6) is independent of $\sigma$ (see the proof of Theorem 4.2.3 in [11]) and hence the probability measure $\pi$ appearing in the theorem depends on $\tau_1, \tau_2, \dots, \tau_r$ only through $\tau_2/\tau_1, \tau_3/\tau_1, \dots, \tau_r/\tau_1$ and is independent of $\tau_1$ if $r = 1$. In the sequel the $\pi(\mathbf{i})$'s will be referred to as the compounding probabilities. Also, it follows from (2.6) that each marginal $N^{(\sigma\tau_i)}$ of the limiting process is a compound Poisson process with intensity $\theta\sigma\tau_i$ for the Poisson events. If we only assume that $\Delta(u_n(\sigma\tau_1), \dots, u_n(\sigma\tau_r))$ holds when $\sigma = 1$, the conclusion of the theorem is still valid for the restrictions of the $N_n^{(\tau_i)}$'s and $N^{(\tau_i)}$'s to $(0,1]$ with, in (2.6), $\sigma = 1$ and the integration being over $(0,1]$; see [11].

We remark that it is straightforward, using Theorem 2.5, to give a multilevel version of Corollary 4.2(ii) in [19], which concerns extremes of stationary and *regenerative* sequences. This, in turn, can be used to derive an extended description of the extremal characteristics of the Lindley process as given in [19], Section 6(i). The Lindley process $\{W_n;\ n \geq 0\}$ is a Markov chain constructed from an i.i.d. sequence $\{D_n\}$ by the recursion $W_{n+1} = (W_n + D_{n+1})_+$, $n = 0, 1, \dots,$ where $x_+ := \max(0, x)$. In the tail, $\{W_n\}$ will certainly look like a random walk and in Section 3 it will be seen that a similar tail behaviour, determined by functions of i.i.d. sequences, is shown by a large class of Markov chains. In fact, the extremal behaviour of the Lindley process may equally well be derived from the results in Section 3.

By slightly strengthening the assumptions, it is possible to state a complete convergence result for $\{X_k\}$ (cf. [12]). Specifically, in addition we require $u_n(\tau)$ to be continuous and strictly decreasing in $\tau > 0$ for each $n$ so that the inverse function $u_n^{-1}$ is well defined. Further, assume $\Delta(u_n(\tau_1), \dots, u_n(\tau_r))$ to hold for $\{X_k\}$ for each $r = 1, 2, \dots$ and each choice of $\tau_1, \tau_2, \dots, \tau_r > 0$. For convenience we call this slightly stronger mixing condition the $\Delta$ condition. Let $\mathbf{N}_n^*$ denote the point process of $[0,\infty) \times (0,\infty)$ with points at $(k/n, u_n^{-1}(X_k))$, $k \geq 0$.

THEOREM 2.6. *Suppose that the stationary sequence $\{X_k;\ k \geq 0\}$ has extremal index $\theta > 0$ and satisfies the condition $\Delta$. Then $\mathbf{N}_n^*$ converges in distribution to some point process $\mathbf{N}^*$ on $[0,\infty) \times (0,\infty)$ if and only if for each $r \geq 1$ and each choice of $\tau_1 > \cdots > \tau_r > 0$ there exist constants $\pi(\mathbf{i})$, $\mathbf{i} \in \mathscr{I}$,* *satisfying* (2.5) *for some* $\Delta(u_n(\tau_1), \ldots, u_n(\tau_r))$*-separating sequence* $\{p_n\}$. *In this case*, $\mathbf{N}^*$ *has the representation*

$$\sum_{i=1}^{\infty} \sum_{j=1}^{K_i} \delta_{(S_i, T_i Y_{ij})}, \tag{2.7}$$

*where* $(S_i, T_i)$, $i \geq 1$, *are the points of a homogenous Poisson process on* $[0, \infty) \times (0, \infty)$ *with mean* $\theta$, $Y_{ij}$, $1 \leq j \leq K_i$, *are the points of a point process* $\eta_i$ *on* $[1, \infty)$ *with* 1 *as an atom*, $\eta_1, \eta_2, \ldots$ *being i.i.d. and independent of the Poisson process, and* $\delta_{(x,y)}(A) = 1$ *if* $(x, y) \in A$ *and zero otherwise.*

PROOF. This follows at once by combining Theorem 2.5 with Theorem 4.4.2 and Corollary 4.5.8 of [11]; compare also Section 4 of [12]. □

The point processes $\eta_i$ describe the local dependence structure of the sequence asymptotically. Write $\eta$ for an arbitrary $\eta_i$. Using [11], the second part of Theorem 4.4.2 and the representation of the limiting process of Theorem 2.6, it is seen from the structure of $\mathbf{N}^*$ that the following relation between the distribution of $\eta$ and the compounding probabilities holds:

$$\pi(\mathbf{i}) = \int_{\tau_1/\tau_p}^{\tau_1/\tau_{p+1}} \mathbf{P}\big\{\eta([1, x\tau_j/\tau_1)) = i_j; j = 1, 2, \ldots, p\big\} x^{-2}\, dx, \tag{2.8}$$

where $p = \max\{j \geq 1;\ i_j \geq 1\}$ and $\tau_1/\tau_{r+1} = \infty$. For an example where the structure of $\mathbf{N}^*$ may be explicitly written down, see Section 4, Example 4.1.

## 3. Exceedance point processes for stationary Markov chains.

In this section we investigate extremal properties of a fairly large class of stationary Markov chains using Proposition 2.3 and Theorem 2.5 together with ideas from Smith [20]. Smith concentrates on computing the extremal index in the absolutely continuous case. However, his observation concerning tail behaviour of the chains applies also in more general circumstances. Following [20] we will use the basic assumption

$$\lim_{p \to \infty} \limsup_{n \to \infty} \mathbf{P}\big\{M_{p, p_n} > u_n | X_0 > u_n\big\} = 0 \tag{3.1}$$

for suitable sequences $\{p_n\}$ and $\{u_n\}$.

Let $\{X_k;\ k \geq 0\}$ be a Markov chain with state space $E \subseteq \mathfrak{R}$. Write $x_E = \sup E$ and for a stationary chain with marginal distribution function (d.f.) $F$ let $x_F := \sup\{x; F(x) < 1\} \leq \infty$. If $x_F < \infty$, we assume that $E \cap [x_F, \infty)$ is empty, implying $x_E = x_F$ regardless of whether the right endpoint of $F$ is finite or not. The stationary chains under consideration will be restricted to those for which $F$ belongs to the domain of attraction of some extreme value law. This guarantees that the conditional law of $X_0$, properly normalized, given that $X_0 > u$, converges to a generalized Pareto distribution as $u \uparrow x_F$; see Proposition 3.1. Also, whenever the extremal index exists and is strictly positive, there is a linear normalization of the maximum, $M_n$, which converges in distribution as $n \to \infty$ to some nondegenerate r.v., necessarily having an extreme value type distribution.

For ease of reference we supply the following result, which is a reformulation of Theorem 1.6.2 in [16].

PROPOSITION 3.1. *A d.f. $F$ is in the domain of attraction of some extreme value law if and only if for some $\gamma$, $-\infty < \gamma < \infty$,*

$$(3.2)\quad \lim_{u \uparrow x_F} [1 - F(u + g(u)x)]/[1 - F(u)] = (1 - \gamma x)_+^{1/\gamma}, \quad -\infty < x < \infty,$$

*where $y_+ := \max\{0, y\}$ and*

$$x_F = \infty \text{ and } g(u) = -\gamma u \qquad \text{if } \gamma < 0,$$
$$x_F < \infty \text{ and } g(u) = \gamma(x_F - u) \qquad \text{if } \gamma > 0.$$

*If $\gamma = 0$, then the auxiliary function $g$ is unique up to asymptotic equivalence and strictly positive on $(x_0, x_F)$ for some $x_0 < x_F$. Hence, if (3.2) holds, there exist $a_n > 0$, $b_n$ such that*

$$F^n(a_n x + b_n) \to \exp\left(-(1 - \gamma x)_+^{1/\gamma}\right) \quad \text{as } n \to \infty.$$

Consider the Lindley process $\{W_n\}$ mentioned in Section 2. Under assumptions on the distribution of $D_1$, a stationary version exists satisfying $\mathbf{P}\{W_0 > u\} \sim C\exp(-\alpha u)$ as $u \to \infty$ for some constants $C$, $\alpha > 0$; see, for example, [3]. Hence (3.2) holds for the marginal distribution with $\gamma = 0$, $g(u) \equiv \alpha^{-1}$ and $x_F = \infty$. Further, the transition probabilities satisfy

$$\mathbf{P}\{W_1 - W_0 \le x | W_0 = u\} \to \mathbf{P}\{D_1 \le x\} \quad \text{as } u \to \infty.$$

Aldous [1] calls Markov processes with this behaviour additive. Using a heuristic argument he approximates the extremal index by $\mathbf{P}\{Z + M < 0\}$, where $Z$ and $M$ are independent, $Z$ has an exponential distribution and $M$ is the supremum of a random walk generated by an i.i.d. sequence $\{D_n\}$. According to [19], Section 6(i), this is an exact result in the case of the Lindley process.

Suppose now that $\{X_k;\, k \ge 0\}$ is stationary with marginal d.f. $F$ such that (3.2) holds for some $\gamma$, $-\infty < \gamma < \infty$. Define $\psi(x) := (1 - \gamma x)_+^{-1/\gamma}$ and let $\to_w$ denote weak convergence. Further, suppose the transition probabilities satisfy

$$(3.3)\quad \mathbf{P}\{\psi([X_1 - u]/g(u)) \le x | X_0 = u\} \to_w H(x) \quad \text{as } u \uparrow x_E,\ u \in E,$$

for some d.f. $H$ on $[0, \infty)$. If $H(\{0\}) > 0$, suppose in addition that

$$(3.4)\quad \lim_{\delta \downarrow 0} \limsup_{u \uparrow x_E} \sup_{\substack{x \in E \\ \psi([x-u]/g(u)) \le \delta}} \mathbf{P}\{X_1 > u | X_0 = x\} = 0.$$

Observe that (3.3) is equivalent to $\mathbf{P}\{X_1/X_0 \le x | X_0 = u\} \to_w H(x^{-1/\gamma})$ and $\mathbf{P}\{(x_E - X_0)/(x_E - X_1) \le x | X_0 = u\} \to_w H(x^{1/\gamma})$ for $\gamma < 0$ and $\gamma > 0$, respectively, and $x \geq 0$. When $\gamma = 0$, (3.3) is the same as $\mathbf{P}\{(X_1 - X_0)/g(X_0) \leq x | X_0 = u\} \to_w H(e^x)$, $-\infty \leq x < \infty$. Similarly, (3.4) can be stated in three different versions with the sup taken over $x \in E$ such that $x \leq u\delta$, $x \leq u + g(u)\log\delta$ and $x \leq x_E - (x_E - u)/\delta$ when $\gamma < 0$, $\gamma = 0$ and $\gamma > 0$, respectively. The Lindley process satisfies (3.3) with $H(x) = \mathbf{P}\{\exp(\alpha D_1) \leq x\}$; some other examples are discussed in Section 4.

It should be pointed out that the case $H(\{0\}) > 0$ is possible; for instance, asymptotic independence of $X_1$ and $X_0$ in the tails corresponds to $H(\{0\}) = 1$. Condition (3.4) is a generalization of condition (2.9)(ii) in [20], which is assumed to hold throughout [20]. Berman [4], Theorem 11.4.1, uses a similar condition, which implies (3.4), when proving a limit result for the maximum of a not necessarily stationary Markov sequence satisfying Doeblin's hypothesis.

Now, let $\{A_n;\ n \geq 1\}$ be an i.i.d. sequence with marginal distribution $H$ and let $Y_0$ be a r.v. independent of $\{A_n\}$. Define the *tail chain* $\{Y_n;\ n \geq 0\}$ by setting $Y_n = A_n Y_{n-1}$ for $n \geq 1$ and denote by $P^\mu$ the law of $\{Y_n\}$ when $Y_0$ has distribution $\mu$. The use of the word "tail" refers to the fact that the behaviour of the original chain is determined by $\{Y_n\}$ if the initial value, $X_0$, is sufficiently far out in the right tail of the marginal distribution, $F$; see Theorem 3.2 and Lemma 3.3. Further, let $R(x)$ be the number of times the tail chain exceeds $x(> 0)$ for $n \geq 1$; that is,

$$R(x) = \#\{n \geq 1; Y_n > x\}, \qquad x > 0. \tag{3.5}$$

Finally, for fixed $\tau_1 > \tau_2 > \cdots > \tau_r > 0$, let the $r$-vector $\mathbf{\Lambda}_j$ be given by

$$\mathbf{\Lambda}_j = \big(R(\tau_j/\tau_1), \ldots, R(\tau_j/\tau_r)\big), \qquad j = 1, \ldots, r.$$

We now formulate the main result concerning the extremal properties of the class of Markov chains under study.

THEOREM 3.2. *Suppose $\{X_k;\ k \geq 0\}$ is a stationary Markov chain which satisfies* (3.2) *and* (3.3) *for some $\gamma$, $-\infty < \gamma < \infty$. Also, if the limit $H$ in* (3.3) *puts positive mass at* 0, *assume in addition that* (3.4) *holds. Let the initial distribution of the tail chain be given by $\mu(dx) = x^{-2}\,dx$, $x > 1$, and let $\{u_n(\tau)\}$, $\tau > 0$, be sequences satisfying* (1.1).

(i) *Assume $D(u_n(\tau))$ holds for each $\tau > 0$. If for some $\tau_0 > 0$ there is a $D(u_n(\tau_0))$-separating sequence $\{p_n\}$ such that* (3.1) *holds with $u_n = u_n(\tau_0)$, then $\{X_k\}$ has extremal index $\theta$ given by*

$$\theta = P^\mu\{R(1) = 0\}. \tag{3.6}$$

(ii) *Suppose $\{X_k\}$ has extremal index $\theta > 0$ and, for some $\tau_1 > \cdots > \tau_r > 0$, satisfies $\Delta(u_n(\sigma\tau_1), \ldots, u_n(\sigma\tau_r))$ for each $\sigma > 0$. Suppose further there is a $\Delta(u_n(\tau_1), \ldots, u_n(\tau_r))$-separating sequence $\{p_n\}$ such that* (3.1) *holds with $u_n = u_n(\tau_1)$. Then, for each $\sigma > 0$, $(N_n^{(\sigma\tau_1)}, \ldots, N_n^{(\sigma\tau_r)})$ converges in distribution to* *the point process* $(N^{(\sigma\tau_1)},\dots,N^{(\sigma\tau_r)})$ *characterized in Theorem* 2.5, *with compounding probabilities* $\pi(\mathbf{i})$ *given by*

$$(3.7)\qquad \pi(\mathbf{i}) = \theta^{-1}\sum_{j=1}^{r}\frac{\tau_j}{\tau_1}\left[P^{\mu}\{\boldsymbol{\Lambda}_j=\mathbf{i}_j\} - P^{\mu}\{\boldsymbol{\Lambda}_j=\mathbf{i}_{j-1}\}\right],\qquad \mathbf{i}\in\mathcal{J}.$$

For the proof of the theorem we observe that one may without loss of generality assume that the auxiliary function $g$ has the following properties when $u\uparrow x_F$:

$$(3.8)(\mathrm{i})\qquad g(u)/u \to -\gamma \quad \text{if } x_F=\infty,$$

$$(3.8)(\mathrm{ii})\qquad g(u)/(x_F-u)\to\gamma \quad \text{if } x_F<\infty,$$

$$(3.8)(\mathrm{iii})\qquad yg(u)+u\to x_F \quad \text{if } y\gamma<1,$$

$$(3.8)(\mathrm{iv})\qquad g(yg(u)+u)/g(u)\to 1-y\gamma$$

locally uniformly for $y$ with $y\gamma<1$.

This is trivial for $\gamma\neq 0$ and follows when $\gamma=0$ from [18], Propositions 1.1 and 1.4 and Lemmas 1.2 and 1.3, which show that if (3.2) holds for some auxiliary function $g$ which does not have the properties (3.8), then there exists another auxiliary function $\tilde g$ which satisfies (3.8) and (3.2), with $g$ replaced by $\tilde g$ and with $\tilde g\sim g$ as $u\uparrow x_F$.

The following lemma has an interest in its own right since it describes the local behaviour of the chain when it is started from "a high level."

LEMMA 3.3. *Let* $\{X_k;\ k\ge 0\}$ *be a not necessarily stationary Markov chain. Suppose the transition probabilities satisfy* (3.3) *with* $H(\{0\})=0$, *for some* $\gamma$, $-\infty<\gamma<\infty$, *and some function* $g(u)$ *which is strictly positive in some interval* $(x_0, x_E)$ *and satisfies* (3.8) *with* $x_F$ *replaced by* $x_E$.

*Then, for every* $p\ge 1$, *the conditional law of* $(\psi([X_k-u]/g(u));\ k=1,\dots,p)$ *given that* $X_0=u$ *converges to the law of* $(\prod_1^k A_i;\ k=1,\dots,p\}$ *as* $u\uparrow x_E$, $u\in E$, *where* $\{A_n;\ n\ge 1\}$ *is an i.i.d. sequence with marginal distribution* $H$.

PROOF. Assume $\gamma<0$. Then using (3.8), $x_E=\infty$ and (3.3) is equivalent to $\overline{H}_u(x) := \mathbf{P}\{X_1/X_0\le x\mid X_0=u\}\to_w H(x^{-1/\gamma})=\overline{H}(x)$, say, for $x\ge 0$. Let $C(\overline{H})$ be the set of continuity points of $\overline{H}$ and choose $0<x_1'<x_1$, $0\le x_2$, such that $x_1', x_1, x_2\in C(\overline{H})$. Then, using the Markov property,

$$\begin{aligned}
&\left|\mathbf{P}\{X_2/X_1\le x_2,\ x_1'<X_1/X_0\le x_1\mid X_0=u\} - \overline{H}(x_2)\left[\overline{H}(x_1)-\overline{H}(x_1')\right]\right|\\
&\quad\le \int_{(x_1',x_1]}\left|\mathbf{P}\{X_2/X_1\le x_2\mid X_1=uy\}-\overline{H}(x_2)\right| 1_{E/u}(y)\,\overline{H}_u(dy)\\
&\qquad + |\overline{H}_u(x_1)-\overline{H}(x_1)| + |\overline{H}_u(x_1')-\overline{H}(x_1')|\\
&\quad\le \sup_{y\in[x_1',\infty)\cap E/u}\left|\mathbf{P}\{X_2/X_1\le x_2\mid X_1=uy\}-\overline{H}(x_2)\right| + o(1)\\
&\quad\to 0 \quad \text{as } u\to\infty,\ u\in E.
\end{aligned}$$

Since $\overline{H}(\{0\}) = 0$ by assumption, it follows that

$$\mathbf{P}\{X_2/X_1 \le x_2,\, X_1/X_0 \le x_1 | X_0 = u,\, X_1 > 0\} \to_w \overline{H}(x_2)\overline{H}(x_1).$$

With the same type of arguments it is straightforward to show inductively, that for any $p \ge 1$,

$$\mathbf{P}\{X_i > 0;\, i = 1,\dots,p-1 | X_0 = u\} \to 1$$

and

$$\mathbf{P}\{X_i/X_{i-1} \le x_i;\, i = 1,\dots,p | X_0 = u,\, X_i > 0;\, i = 1,\dots,p-1\}$$
$$\to_w \prod_{i=1}^{p} \overline{H}(x_i) \quad \text{as } u \to \infty,\, u \in E.$$

The result for $\gamma < 0$ now follows from the continuous mapping theorem.

The cases $\gamma = 0$ and $\gamma > 0$ can be treated similarly; we omit the details. □

We remark that when $\gamma = 0$, the lemma says that the conditional law of $((X_k - u)/g(u);\, k = 1,\dots,p)$ converges to the law of $(\Sigma_1^k \log A_i;\, k = 1,\dots,p)$, which is the random walk behaviour observed by Smith [20].

PROOF OF THEOREM 3.2. Fix $r \ge 1$ and $\tau_1 > \tau_2 > \cdots > \tau_r > 0$. Write $u_n^{(j)} = u_n(\tau_j)$, $j = 1,\dots,r$, and recall that $P_j$ denotes the conditional law of $\{X_k\}$ given that $X_0 > u_n^{(j)}$. Using (3.1) it follows from Proposition 2.3 and Theorem 2.5 that it is enough to show

$$\lim_{p\to\infty} \lim_{n\to\infty} P_j\{\mathbf{N}_n((0, p/n]) = \mathbf{i}\} = P^{\mu}\{\mathbf{\Lambda}_j = \mathbf{i}\}$$

for $\mathbf{i} = (i_1,\dots,i_r)$, $i_1 \ge \cdots \ge i_r \ge 0$ and $j = 1,\dots,r$.

Let $\mathbf{\Lambda}_j^{(p)}$ be the $r$-vector with $k$th component

$$\Lambda_{j,k}^{(p)} = \#\{1 \le m \le p;\, Y_m > \tau_j/\tau_k\}, \qquad k, j = 1,\dots,r,$$

where $\{Y_n\}$ is the tail chain. Since $P^{\mu}\{\mathbf{\Lambda}_j^{(p)} \le \mathbf{i}\} \downarrow P^{\mu}\{\mathbf{\Lambda}_j \le \mathbf{i}\}$ when $p \to \infty$ it suffices to prove that

$$P_j\{\mathbf{N}_n((0, p/n]) = \mathbf{i}\} \to P^{\mu}\{\mathbf{\Lambda}_j^{(p)} = \mathbf{i}\} \quad \text{as } n \to \infty.$$

However, this follows if we show that

$$P_j\{X_{m_s} \le u_n^{(l_s)};\, s = 1,\dots,t\} \to P^{\mu}\{Y_{m_s} \le \tau_j/\tau_{l_s};\, s = 1,\dots,t\} \tag{3.9}$$

as $n \to \infty$, for $t = 1,\dots,p$, $1 \le m_1 < \cdots < m_t \le p$ and $l_1,\dots,l_t \in \{1,\dots,r\}$.

To this end, let $\kappa_n$ denote the conditional law of $(X_0 - u_n^{(j)})/g(u_n^{(j)})$ given that $X_0 > u_n^{(j)}$. With $I_\gamma := \{x > 0;\, x\gamma < 1\}$ and $E_n := I_\gamma \cap ([E - u_n^{(j)}]/g(u_n^{(j)}))$ we have $\kappa_n(E_n) = 1$. Also, from (3.2) $\kappa_n \to_w \kappa$, where $\kappa(x) := 1 - (\psi(x))^{-1} = 1 - (1 - \gamma x)_+^{1/\gamma}$, $x \in I_\gamma$.

Assume first that $H(\{0\}) = 0$. For $x \in I_\gamma$ put $a_n(x) = u_n^{(j)} + xg(u_n^{(j)}) \to x_F$ as $n \to \infty$ and define

$$h_n(x) = \mathbf{P}\left\{\psi\left(\frac{X_{m_s} - a_n(x)}{g(a_n(x))}\right) \le \psi\left(\frac{u_n^{(l_s)} - a_n(x)}{g(a_n(x))}\right); 1 \le s \le t | X_0 = a_n(x)\right\}$$

if $a_n(x) \in E$ and equal to 0 otherwise,

$$h(x) = \mathbf{P}\left\{\prod_{i=1}^{m_s} A_i \le (\psi(x))^{-1}\tau_j/\tau_{l_s}; 1 \le s \le t\right\}.$$

Observe that from (1.1), (3.2) and (3.8) we have

$$\psi\big(\big[u_n^{(l_s)} - a_n(x)\big]/g(a_n(x))\big) \to (\psi(x))^{-1}\tau_j/\tau_{l_s} \tag{3.10}$$

locally uniformly in $x$ as $n \to \infty$ for $s = 1, \dots, t$.

Formulating (3.9) in a different notation, we have to show $\int_{I_\gamma} h_n(x)\kappa_n(dx) \to \int_{I_\gamma} h(x)\kappa(dx)$. This will follow from a slightly extended version of Theorem 5.5 in [5] if we can show that the set

$$C = \{x \in I_\gamma; \exists \{x_n\}, x_n \in E_n \text{ with } x_n \to x \text{ but } h_n(x_n) \nrightarrow h(x)\}$$

consists of an at most countable number of elements and hence has zero probability under the limiting measure $\kappa$. Let $D$ denote the set of discontinuities of the d.f. of $(\prod_1^{m_1} A_i, \dots, \prod_1^{m_t} A_i)$ on the line segment $\{(\psi(x))^{-1}(\tau_j/\tau_{l_1}, \dots, \tau_j/\tau_{l_t}); x \in I_\gamma\}$ in $\mathfrak{R}^t$. Since $x \in E_n$ implies $a_n(x) \in E$, it follows from Lemma 3.3 and (3.10) that $C \subseteq D$, thereby completing the proof when $H(\{0\}) = 0$.

Next, suppose (3.3) holds with $H(\{0\}) > 0$ and that (3.4) holds as well. Using the Markov property, we see that for every $p \ge 1$,

$$\sup_{\substack{x \in E \\ \psi([x-u]/g(u)) \le \delta}} \mathbf{P}\{X_p > u | X_0 = x\}$$
$$\le \sum_{j=1}^{p} \sup_{\substack{x \in E \\ \psi([x-u]/g(u)) \le \delta^{(p-j+1)/p}}} \mathbf{P}\{X_1 > u + g(u)\psi^{-1}(\delta^{(p-j)/p}) | X_0 = x\},$$

where $\psi^{-1}(y) = (1 - y^{-\gamma})/\gamma$, $y \in (0, \infty)$, is the inverse function of $\psi$ on $I_\gamma$. From this it is straightforward to show, using (3.8), that (3.4) implies

$$\lim_{\delta \downarrow 0} \limsup_{u \uparrow x_F} \sup_{\substack{x \in E \\ \psi([x-u]/g(u)) \le \delta}} \mathbf{P}\{X_p > u | X_0 = x\} = 0 \text{ for every } p \ge 1. \tag{3.11}$$

For fixed $t = 1, \dots, p$, $1 \le m_1 < \cdots < m_t \le p$ and $l_1, \dots, l_t \in \{1, \dots, r\}$, define $\nu_n = \min\{1 \le k \le t; X_{m_k} > u_n^{(l_k)}\}$ and $\nu = \min\{1 \le k \le t; Y_{m_k} > \tau_j/\tau_{l_k}\}$. To show (3.9), we this time find it more convenient to prove

$$P_j\{\nu_n = s\} \to P^\mu\{\nu = s\} \quad \text{as } n \to \infty. \tag{3.12}$$

Since (3.12) implies (3.9) this will be sufficient to complete the proof.

If $m_s = 1$ the proof of (3.12) is completely analogous to the case when $H(\{0\}) = 0$, so we can assume that $m_s > 1$. Proceeding as in the proof of Lemma 3.3, we get

$$\mathbf{P}\left\{\psi\left(\frac{X_i - X_{i-1}}{g(X_{i-1})}\right) \in (a_i, b_i]; i = 1,\dots,p | X_0 = u\right\} \to \prod_{i=1}^{p} (H(b_i) - H(a_i)) \tag{3.13}$$

as $u \uparrow x_F$, $u \in E$, provided $0 < a_i < b_i \le \infty$ for $i = 1,\dots,p-1$, $-\infty \le a_p < b_p \le \infty$ and $a_i, b_i$ are, when finite, continuity points of $H$. Choose $\delta > 0$ to be a continuity point of $H$ such that $H(\delta) < 1$ if $H(\{0\}) < 1$. For $k = 0, 1, \dots$ let $B_k := \{g(X_i) > 0;\, i = 0,\dots,k\}$ and define

$$D_{k+1} = B_k \cap \left\{\psi\left(\frac{X_{k+1} - X_k}{g(X_k)}\right) \le \delta,\, \psi\left(\frac{X_i - X_{i-1}}{g(X_{i-1})}\right) > \delta;\, i = 1,\dots,k\right\}$$

and

$$D^*_{k+1} = B_k \cap \left\{\psi\left(\frac{X_i - X_{i-1}}{g(X_{i-1})}\right) > \delta;\, i = 1,\dots,k+1\right\}.$$

From (3.8) one finds that $\{X_0 > u_n^{(j)}\} \subseteq \bigcup_{i=1}^{m_s-1} D_i \cup D^*_{m_s-1}$ for all $n$ sufficiently large. Thus, for such $n$,

$$\begin{aligned} P_j\{\nu_n = s\} &= \sum_{i=1}^{m_s-1} P_j(\{\nu_n = s\} \cap D_i) + P_j(\{\nu_n = s\} \cap D^*_{m_s-1}) \\ &= S_n(\delta) + T_n(\delta), \quad \text{say.} \end{aligned}$$

It is straightforward, but somewhat lengthy, to show from (3.11) and (3.13) that $\lim_{\delta \downarrow 0} \limsup_{n\to\infty} S_n(\delta) = 0$; we omit the details. Furthermore, similar reasoning as for the case when $H(\{0\}) = 0$ shows that

$$T_n(\delta) \to P^{\mu}\{\nu = s,\, A_k > \delta;\, 1 \le k \le m_s - 1\} \quad \text{as } n \to \infty.$$

Since $\delta$ can be taken arbitrarily small and $P^{\mu}\{\nu = s,\, A_k = 0 \text{ some } k \le m_s - 1\} = 0$ the proof is complete. □

The following corollary points out some situations in which the one-level compounding distribution $\{\pi(i);\, i \ge 1\}$ has a simple form, being geometric with the extremal index, $\theta$, as parameter; that is, $\pi(i) = \theta(1-\theta)^{i-1}$, $i \ge 1$.

COROLLARY 3.4. *Suppose $\{X_k;\, k \ge 0\}$ is a Markov chain satisfying the conditions in Theorem* 3.2.

(i) *If $A_1 < 1$ a.s., then $\theta = 1 - \mathbf{E}A_1$.*
(ii) *If $A_1 = 0$ w.p. $p > 0$ and $A_1 \ge 1$ w.p. $1-p$, then $\theta = \mathbf{P}\{A_1 = 0\}$.*

*In either case the one-level compounding distribution $\{\pi(i)\}$ is geometric with parameter $\theta$.*

PROOF. This follows from straightforward calculations using the expressions for $\theta$ and $\{\pi(i)\}$ given in (3.6) and (3.7) of Theorem 3.2, respectively. □

*An extension of the theorem when $\gamma < 0$.* Suppose (3.2) holds with $\gamma < 0$, that $\inf\{x;\ F(x) > 0\} = -\infty$ and that

$$\mathbf{P}\{X_1/u \le x | X_0 = u\} \to_w \overline{H}(x) \quad \text{as } |u| \to \infty,\ u \in E, \tag{3.14}$$

for some d.f. $\overline{H}$ on $(-\infty, \infty)$, where we allow $\overline{H}((-\infty, 0)) > 0$. If $\overline{H}(\{0\}) > 0$, assume in addition that (3.4) holds or that

$$\lim_{\delta \downarrow 0} \limsup_{u \to \infty} \sup_{\substack{x \in E \\ |x| \le \delta u}} \mathbf{P}\{|X_1| > u | X_0 = x\} = 0. \tag{3.15}$$

Let $\{\overline{A}_n;\ n \ge 1\}$ be an i.i.d. sequence with $\mathbf{P}\{\overline{A}_1 \le x\} = \overline{H}(x)$ and define the tail chain through $Y_n = \overline{A}_n Y_{n-1}$, $n \ge 1$, $Y_0$ being independent of $\{\overline{A}_n\}$. Then, if $\{X_n\}$ satisfies the conditions in (i) and (ii) of Theorem 3.2, the result of the theorem holds with the initial distribution $\mu$ replaced by $\overline{\mu}(dx) := |\gamma|^{-1} x^{1/\gamma}\, dx$, for $x > 1$ and $\boldsymbol{\Lambda}_j$ replaced by $\overline{\boldsymbol{\Lambda}}_j := (R((\tau_j/\tau_1)^{-\gamma}), \dots, R((\tau_j/\tau_r)^{-\gamma}))$, $j = 1, \dots, r$, where $R(\cdot)$ is given in (3.5). Also, if $|\overline{A}_1| < 1$ a.s., then

$$\theta = 1 - \left[\mathbf{E}|\overline{A}_1|^{-1/\gamma} 1_{\{\overline{A}_1 > 0\}} + \left(\mathbf{E}|\overline{A}_1|^{-1/\gamma} 1_{\{\overline{A}_1 < 0\}}\right)^2 \sum_{i=0}^{\infty} \left(\mathbf{E}|\overline{A}_1|^{-1/\gamma} 1_{\{\overline{A}_1 > 0\}}\right)^i\right].$$

Specifically, if $-1 < \overline{A}_1 \le 0$, then $\theta = 1 - (\mathbf{E}|\overline{A}_1|^{-1/\gamma})^2$ and also if $\overline{A}_1 = 0$ w.p. $p > 0$ and $\overline{A}_1 \le -1$ w.p. $1 - p$, then $\theta = 1 - (\mathbf{P}\{\overline{A}_1 \le -1\})^2$. In both of these cases, $\{\pi(i)\}$ is geometric with parameter $\theta$. This extension can be proven in exactly the same way as Theorem 3.2 and Corollary 3.4.

**4. Examples.** In this section we apply the results of Section 3 to two examples. The first one illustrates the ideas in a setting where it is possible to write down the extremal characteristics of the involved random sequence explicitly. The second example, which is more interesting from an applied point of view, concerns first order stochastic difference equations with random coefficients. Extremal behaviour was studied in [9] for the one-level case and under the assumption that all variables are nonnegative. Here we consider both multiple levels and the possibility of negative process values.

EXAMPLE 4.1. Let $X_0$ be uniformly distributed on $E = (0, 1)$ and for $k \ge 1$ let $X_k$ be given by the first-order autoregressive scheme

$$X_k = \frac{1}{r} X_{k-1} + \varepsilon_k, \tag{4.1}$$

where $r \ge 2$ is an integer, $\{\varepsilon_k\}$ are i.i.d. and uniformly distributed on $\{0, 1/r, \dots, (r-1)/r\}$ and $\{\varepsilon_k\}$ is independent of $X_0$. The stationary sequence $\{X_k\}$ was studied by Chernick [7] as an example where clustering of high values occurs. He showed by direct argument that $\{X_k\}$ has extremal index $\theta = (r-1)/r$. We will here use Theorem 3.2 to find both the extremal index and the limiting compounding probabilities associated with exceedances of one and two levels. Furthermore, we will explicitly describe the limiting process $\mathbf{N}^*$ of the complete convergence result in Theorem 2.6.

Let $F$ be the d.f. of $X_0$. Then (3.2) holds for $F$ with $\gamma = 1$ and using (4.1) it is easily seen that the transition probabilities satisfy (3.3) with

$$H(x) = \begin{cases} (r-1)/r, & 0 \le x < r, \\ 1, & r \le x. \end{cases} \tag{4.2}$$

Since $H(\{0\}) = (r-1)/r > 0$, we also have to verify (3.4). However, for any $\delta$ and $u$ with $0 < \delta \le 1/r$, $u > 1-\delta$, we have $\sup_{x \le 1-(1-u)/\delta} \mathbf{P}\{X_1 > u | X_0 = x\} = 0$ and (3.4) follows immediately.

Next, defining $u_n(\tau) = 1 - \tau/n$ for $\tau > 0$, we verify $\Delta(u_n(\tau_1), \dots, u_n(\tau_r))$ for $r \ge 1$, $\tau_1 > \cdots > \tau_r > 0$. Note that any event in $\mathscr{B}_i^j = \sigma(\{X_k \le u_n(\tau_m)\};\ i \le k \le j,\ 1 \le m \le r)$ might be written as a finite, disjoint union of events of type

$$\{u_n(\tau_{m_k}) < X_k \le u_n(\tau_{m_k+1});\ i \le k \le j\},$$

where $m_k \in \{0, 1, \dots, r\}$, $u_n(\tau_0) = -\infty$ and $u_n(\tau_{r+1}) = \infty$. Let $\varepsilon > 0$ be given. Then for fixed $\nu$ with $0 < \nu < 1$, using the same type of arguments as in the proof of Theorem 2.1 in [9], it is seen that

$$\begin{aligned} \beta_{n,[n\nu]} \le n \sum_{i=1}^{r} \mathbf{P}\{1 - (\tau_i + \varepsilon)/n < X_0 \le 1 - (\tau_i - \varepsilon)/n\} \\ + 2n\mathbf{P}\{X_0/r^{[n\nu]} > \varepsilon/(2n)\}. \end{aligned}$$

Since obviously $n\mathbf{P}\{X_0 > \varepsilon r^{[n\nu]}/(2n)\} \to 0$ as $n \to \infty$ it follows that $\beta_{n,[n\nu]} \to 0$ as $n \to \infty$ and $\Delta(u_n(\tau_1), \dots, u_n(\tau_r))$ holds.

Now the results of parts (i) and (ii) of Theorem 3.2 both follow if we show that

$$\lim_{p\to\infty} \limsup_{n\to\infty} \mathbf{P}\{M_{p,p_n} > u_n(\tau) | X_0 > u_n(\tau)\} = 0 \tag{4.3}$$

for every $\tau > 0$ and every sequence $\{p_n\}$ with $p_n \to \infty$, $p_n = o(n)$. However,

$$\begin{aligned} &\mathbf{P}\{X_j > u_n(\tau) | X_0 > u_n(\tau)\} \\ &\quad = \mathbf{P}\{X_j > u_n(\tau),\ \varepsilon_i > 0;\ i = 1, 2, \dots, j-1 | X_0 > u_n(\tau)\} \\ &\qquad + \mathbf{P}\{X_j > u_n(\tau),\ \varepsilon_i = 0, \text{some } i = 1, 2, \dots, j-1 | X_0 > u_n(\tau)\} \\ &\quad = (A) + (B), \quad \text{say}. \end{aligned} \tag{4.4}$$

Here $(A) \le ((r-1)/r)^{j-1}$ and

$$\begin{aligned} (B) &= \sum_{i=1}^{j-1} \mathbf{P}\{X_j > u_n(\tau),\ \varepsilon_i = 0,\ \varepsilon_l > 0;\ l \le i-1 | X_0 > u_n(\tau)\} \\ &\le \sum_{i=1}^{j-1} \mathbf{P}\{X_j > u_n(\tau) | X_i = 1/r\}(r-1)^{i-1}/r^i \\ &= \sum_{i=1}^{j-1} \mathbf{P}\{X_{j-i} > u_n(\tau) | X_0 = 1/r\}(r-1)^{i-1}/r^i, \end{aligned} \tag{4.5}$$

where we used the Markov property and the fact that $\mathbf{P}\{X_j > u_n(\tau)|X_i = x\}$, $0 < x < 1$, $i < j$, is a nondecreasing function of $x$. The conditional distribution of $X_p$ given $X_0 = 1/r$ is uniform on $\{1/r^{p+1} + i/r^p;\ i = 0, \dots, r^p - 1\}$ and therefore

$$\mathbf{P}\{X_p > u_n(\tau)|X_0 = 1/r\} \le \mathbf{P}\{X_p > u_n(\tau)\} + 1/r^{p+1}, \qquad p \ge 1.$$

Hence, it follows from (4.5) that

$$(B) \le \mathbf{P}\{X_0 > u_n(\tau)\} + (j-1)((r-1)/r)^{j-1}. \tag{4.6}$$

Finally, combining (4.4) and (4.6) gives

$$\mathbf{P}\{M_{p,p_n} > u_n(\tau)|X_0 > u_n(\tau)\} \le p_n \mathbf{P}\{X_0 > u_n(\tau)\} + \sum_{j=p+1}^{\infty} j((r-1)/r)^{j-1},$$

and (4.3) follows provided $p_n = o(n)$. Compare also the derivation of the local dependence restriction $D^{(2)}(u_n)$ for $\{X_k\}$ in [8].

Thus the conditions of Corollary 3.4(ii) hold and $\{X_k\}$ has extremal index $\theta = (r-1)/r$ and the one-level compounding probabilities are $\pi(i) = (r-1)r^{-i}$, $i \ge 1$.

In particular by Theorem 6.1 in [13], if $M_n^{(k)}$ is the $k$th largest among $X_1, \dots, X_n$, then

$$\begin{aligned}\lim_{n\to\infty} \mathbf{P}\left\{M_n^{(k)} \le \frac{x}{n} + 1\right\} &= \exp\left(\frac{r-1}{r}x\right) \\ &\times \left(1 + \sum_{j=1}^{k-1}\sum_{i=j}^{k-1} \frac{(-((r-1)/r)x)^j}{j!}\binom{i-1}{j-1}\left(\frac{r-1}{r}\right)^j\left(\frac{1}{r}\right)^{i-j}\right),\end{aligned}$$

for $x < 0$ and $k \ge 1$.

Next consider exceedances of two different levels, $u_n(\tau_1)$ and $u_n(\tau_2)$ with $\tau_1 > \tau_2 > 0$. Let $x = \tau_1/\tau_2 > 1$ and let $m = m(x)$ denote the integer part of $\log x/\log r$. It is straightforward to show that the two-level compounding probabilities $\{\pi(i,j) = \pi(i,j;x);\ i \ge j \ge 0,\ i \ge 1\}$ are given by

$$\pi(i,j;x) = \begin{cases} (r-1)r^{-i}, & i \le m, j = 0, \\ r^{1-i} - x^{-1}, & i = m+1; j = 0, \\ x^{-1}r^{1-j} - r^{-i}, & i = m+j, j \ge 1, \\ r^{1-i} - x^{-1}r^{-j}, & i = m+j+1, j \ge 1, \\ 0, & \text{otherwise}. \end{cases} \tag{4.7}$$

The distribution (4.7) may be used to obtain the joint asymptotic distribution for any two $M_m^{(k)}$, $M_n^{(l)}$, $k \ne l$; compare [11], Theorem 4.3.2.

Finally consider the point process $\mathbf{N}_n^*$ on $[0,\infty) \times (0,\infty)$ with points at $(k/n, n(1-X_k))$, $k \ge 0$. With $S_i(x) = \sum_{j=1}^{i} \pi(j,0;x)$, for $x = \tau_1/\tau_2 > 1$, $i \ge 1$, it follows from (2.8) and $S_i(x) = \int_1^x \mathbf{P}\{\eta([1, y)) \le i\} y^{-2}\,dy$. However, according to (4.7),

$$S_i(x) = \begin{cases} 1 - x^{-1}, & 1 < x \le r^i, \\ 1 - r^{-i}, & r^i < x, \end{cases}$$

and hence we see that $\mathbf{P}\{\eta([1, x)) \le i\} = 1$ if $1 < x \le r^i$ and zero otherwise. Therefore, the representation (2.7) holds for $\mathbf{N}^*$ with $Y_{ij} = r^{j-1}$ a.s. $j \ge 1$; that is, $\mathbf{N}^*$ is distributed as $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\delta_{(S_i, T_i r^{j-1})}$, where $\{S_i, T_i\}$ are the points of a homogenous Poisson process on $[0, \infty) \times (0, \infty)$ with mean $(r-1)/r$.

EXAMPLE 4.2. Let $Y_0$ be a r.v. and let $\{(A_n, B_n); n \ge 1\}$ be i.i.d. $\Re^2$-valued r.v.'s independent of $Y_0$. Define $Y_n$, $n \ge 1$, by means of the stochastic difference equation

$$Y_n = A_n Y_{n-1} + B_n. \tag{4.8}$$

Extremal behaviour of solutions of (4.8) was studied in [9] for the case where both the $A_n$'s and the $B_n$'s are nonnegative a.s. The results, including expressions for the extremal index and the one-level limiting compounding probabilities, were applied to a first order ARCH process of interest in econometric modelling. Here we will extend the results in [9] by allowing $\mathbf{P}\{A_1 < 0\} > 0$ and establishing a multilevel result.

Suppose that $B_1/(1 - A_1)$ is nondegenerate and that the conditional distribution of $\log|A_1|$ given $A_1 \ne 0$ is nonlattice. Further, suppose that there is an $\alpha > 0$ such that $\mathbf{E}|A_1|^\alpha = 1$, $\mathbf{E}|A_1|^\alpha \log^+|A_1| < \infty$ and $0 < \mathbf{E}|B_1|^\alpha < \infty$. Under these assumptions, results of Vervaat [22] and Kesten [15] show that the equation $Y_\infty =_d A_1 Y_\infty + B_1$, where $Y_\infty$ is independent of $(A_1, B_1)$, has a solution unique in distribution, given by

$$Y_\infty =_d \sum_{j=1}^{\infty} B_j \prod_{i=1}^{j-1} A_i.$$

If $Y_0 =_d Y_\infty$, then $\{Y_n\}$ is stationary and $Y_n$ converges in distribution to $Y_\infty$ whatever the distribution of $Y_0$. Furthermore, there exist nonnegative constants $c_+$ and $c_-$, with $c_+ + c_- > 0$, such that

$$\mathbf{P}\{Y_\infty > t\} \sim c_+ t^{-\alpha} \quad \text{and} \quad \mathbf{P}\{Y_\infty < -t\} \sim c_- t^{-\alpha} \quad \text{as } t \to \infty. \tag{4.9}$$

Let $Y_0 =_d Y_\infty$, so that $\{Y_n\}$ is stationary. We will henceforth restrict our attention to the case when $c_+ > 0$. This is obviously satisfied when $\{(A_n, B_n)\}$ are $\Re_+^2$-valued and furthermore, using a result by Goldie ([10], Theorem 3.1), we see that $c_+ > 0$ also when $\mathbf{P}\{A_1 < 0\} > 0$.

Define $u_n(x) = xn^{1/\alpha}$, $x > 0$. To verify the conditions in Theorem 3.2, we first show that

$$\limsup_{n\to\infty} \mathbf{P}\left\{M_{p,p_n} > xn^{1/\alpha} | Y_0 > xn^{1/\alpha}\right\} \to 0 \quad \text{as } p \to \infty \tag{4.10}$$

for any $x > 0$ and any sequence $\{p_n\}$ with $p_n \to \infty$, $p_n = o(n)$. As in the proof of Theorem 2.1 in [9], introduce the auxiliary process

$$Y_n^\# = Y_0 \prod_{j=1}^{n} A_j,\ n \ge 1,\ Y_0^\# = Y_0,$$

and write $\Delta_n = Y_n - Y_n^{\#}$, $n \ge 0$. Since $\{\Delta_n\}$ is independent of $Y_0$, we have

$$\begin{aligned}
&\mathbf{P}\{M_{p,p_n} > xn^{1/\alpha} | Y_0 > xn^{1/\alpha}\} \\
&\quad \le \mathbf{P}\left\{ \bigvee_{j=p+1}^{p_n} Y_j^{\#} > \frac{xn^{1/\alpha}}{2} \middle| Y_0 > xn^{1/\alpha} \right\} + \mathbf{P}\left\{ \bigvee_{j=p+1}^{p_n} \Delta_j > \frac{xn^{1/\alpha}}{2} \right\} \\
&\quad = (A) + (B), \quad \text{say.}
\end{aligned}$$

Writing $Q_n$ for the conditional distribution of $Y_0/(xn^{1/\alpha})$ given $Y_0 > xn^{1/\alpha}$, it follows from (4.9) that $Q_n$ converges weakly to $Q$, where $Q(dy) = \alpha y^{-\alpha-1}\,dy$, $y > 1$. Recall that $Y_0 =_d Y_\infty$ and that $c_+$ was assumed strictly positive.

Now, for given $\varepsilon > 0$ choose $\tilde{x}$ such that $Q\{(\tilde{x}, \infty)\} < \varepsilon$. We then have

$$\begin{aligned}
(A) &= \int_{(1,\infty)} \mathbf{P}\left\{ \bigvee_{j=p+1}^{p_n} \prod_{i=1}^{j} A_i > 1/(2y) \right\} Q_n(dy) \\
&\le \int_{(1,\tilde{x}]} \mathbf{P}\left\{ \bigvee_{j=p+1}^{p_n} \prod_{i=1}^{j} A_i > 1/(2y) \right\} Q_n(dy) + Q_n\{(\tilde{x},\infty)\}.
\end{aligned} \tag{4.11}$$

It follows from the assumptions made that $\mathbf{E}\log|A_1| < 0$; see, for example, [10], Lemma 2.2, and therefore there exists a $t > 0$ with $\mathbf{E}|A_1|^t < 1$; compare [9], Section 3. Since the first term in (4.11) is bounded above by $\sum_{j=p+1}^{p_n}(\mathbf{E}|A_1|^t)^j/(1/(2\tilde{x}))^t$, we get $\lim_{p\to\infty}\limsup_{n\to\infty}(A) = 0$. Further,

$$\begin{aligned}
(B) &\le \mathbf{P}\left\{ \bigvee_{j=p+1}^{p_n} |\Delta_j| > \frac{xn^{1/\alpha}}{2} \right\} \\
&\le \mathbf{P}\left\{ \bigvee_{j=p+1}^{p_n} |Y_0| \prod_{i=1}^{j} |A_i| > \frac{xn^{1/\alpha}}{4} \right\} + \mathbf{P}\left\{ \bigvee_{j=p+1}^{p_n} |Y_j| > \frac{xn^{1/\alpha}}{4} \right\} \\
&\le \mathbf{P}\left\{ \bigvee_{j=p+1}^{\infty} |Y_0| \prod_{i=1}^{j} |A_i| > \frac{xn^{1/\alpha}}{4} \right\} + p_n \mathbf{P}\left\{ |Y_0| > \frac{xn^{1/\alpha}}{4} \right\}.
\end{aligned} \tag{4.12}$$

The first term in (4.12) converges to zero as $n \to \infty$ since $\mathbf{E}\log|A_1| < 0$ implies $\bigvee_{j=1}^{\infty}\prod_{i=1}^{j}|A_i| < \infty$ a.s. and the second term also tends to zero, by (4.9), provided $p_n = o(n)$. Hence (4.10) follows.

Next, it follows along similar lines as in Example 4.1 that the mixing condition $\Delta(u_n(x_1), \dots, u_n(x_r))$ holds for $\{Y_n\}$.

By (4.9) the d.f. of $Y_0$ satisfies (3.2) with $\gamma = -\alpha^{-1}$. Further, it is easily seen both that (3.14) holds with $\overline{H}(x) = \mathbf{P}\{A_1 \le x\}$ and that (3.15) is satisfied in case $\mathbf{P}\{A_1 = 0\}$ is nonzero. Hence the results of Theorem 3.2 hold for $\{Y_n\}$, with $\mu$ and $\mathbf{\Lambda}_j$ replaced by $\overline{\mu}$ and $\overline{\mathbf{\Lambda}}_j$, respectively, as given in the paragraph following Corollary 3.4.

We finally remark that the assumption that $\{Y_n\}$ is stationary is not crucial; the point process convergence result holds regardless of the distribution of $Y_0$.

**Acknowledgments.** I wish to express my gratitude to Holger Rootzén for his constant support and advice during the preparation of this paper. Thanks are also due to a referee for suggesting the use of the von Mises parametrization of the extreme value laws.

## REFERENCES

[1] ALDOUS, D. (1989). *Approximations via the Poisson Clumping Heuristic.* Springer, New York.

[2] ALPUIM, M. T. (1989). An extremal Markovian sequence. *J. Appl. Probab.* **26** 219–232.

[3] ASMUSSEN, S. (1987). *Applied Probability and Queues.* Wiley, New York.

[4] BERMAN, S. M. (1992). *Sojourns and Extremes of Stochastic Processes.* Wadsworth & Brooks/Cole, Belmont, CA.

[5] BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

[6] BOLLERSLEV, T. and ENGLE, R. F. (1986). Modelling the persistence of conditional variances. *Econom. Rev.* **27** 1–50.

[7] CHERNICK, M. R. (1981). A limit theorem for the maximum of autoregressive processes with uniform marginal distributions. *Ann. Probab.* **9** 145–149.

[8] CHERNICK, M. R., HSING, T. and MCCORMICK, W. P. (1991). Calculating the extremal index for a class of stochastic processes. *Adv. in Appl. Probab.* **23** 835–850.

[9] DE HAAN, L., RESNICK, S. I., ROOTZÉN, H. and DE VRIES, C. G. (1989). Extremal behaviour of solutions to a stochastic difference equation with applications to ARCH processes. *Stochastic Process. Appl.* **32** 213–224.

[10] GOLDIE, C. M. (1991). Implicit renewal theory and tails of solutions of random equations. *Ann. Appl. Probab.* **1** 126–166.

[11] HSING, T. (1984). Point processes associated with extreme value theory. Dissertation, Technical Report 83, Center for Stochastic Processes, Univ. North Carolina.

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

[13] 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.

[14] KALLENBERG, O. (1976). *Random Measures.* Academic, New York.

[15] KESTEN, H. (1973). Random difference equations and renewal theory for products of random matrices. *Acta Math.* **131** 207–248.

[16] LEADBETTER, M. R., LINDGREN, G. and ROOTZÉN, H. (1983). *Extremes and Related Properties of Random Sequences and Processes.* Springer, New York.

[17] O'BRIEN, G. L. (1987). Extreme values for stationary and Markov sequences. *Ann. Probab.* **15** 281–291.

[18] RESNICK, S. I. (1987). *Extreme Values, Regular Variation and Point Processes.* Springer, New York.

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

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

[21] TONG, H. (1990). *Time Series; A Dynamical System Approach.* Oxford Univ. Press.

[22] VERVAAT, W. (1979). On a stochastic difference equation and a representation of nonnegative infinitely divisible random variables. *Adv. in Appl. Probab.* **11** 750–783.

DEPARTMENT OF MATHEMATICAL STATISTICS
UNIVERSITY OF LUND
BOX 118
S-221 00 LUND
SWEDEN