The Annals of Applied Probability
1998, Vol. 8, No. 2, 408–437

THE EXTREMAL INDEX OF A HIGHER-ORDER STATIONARY MARKOV CHAIN

BY SEOKHOON YUN

University of Suwon

The paper presents a method of computing the extremal index of a real-valued, higher-order (kth-order, $k \geq 1$) stationary Markov chain $\{X_n\}$. The method is based on the assumption that the joint distribution of $k+1$ consecutive variables is in the domain of attraction of some multivariate extreme value distribution. We introduce limiting distributions of some rescaled stationary transition kernels, which are used to define a new $(k-1)$th-order Markov chain $\{Y_n\}$, say. Then, the kth-order Markov chain $\{Z_n\}$ defined by $Z_n = Y_1 + \cdots + Y_n$ is used to derive a representation for the extremal index of $\{X_n\}$. We further establish convergence in distribution of multilevel exceedance point processes for $\{X_n\}$ in terms of $\{Z_n\}$. The representations for the extremal index and for quantities characterizing the distributional limits are well suited for Monte Carlo simulation.

1. Introduction. The effort of characterizing the tail of the distribution of the maximum of a large number of observations has long been made by several statisticians. Under stationarity of observations, this is usually achieved by incorporating both the tail of the marginal distribution and the so-called extremal index. The extremal index is therefore a key parameter for studying the extremal behavior of a stationary sequence of random variables (r.v.'s). See Loynes (1965), O'Brien (1974) and Leadbetter, Lindgren and Rootzén (1983). In this paper, we present a method of computing the extremal index of a real-valued, kth-order ($k \geq 1$) stationary Markov chain. The method is based on the assumption that the joint distribution of $k+1$ consecutive variables is in the domain of attraction of some multivariate extreme value distribution. To give a more complete description of the extremal behavior of the chain, we also establish convergence in distribution of multilevel exceedance point processes and give representations for quantities characterizing the distributional limits.

Let $\{X_n\colon n \geq 1\}$ be a (strictly) stationary sequence of r.v.'s with marginal distribution function (d.f.) F. Then, for all sufficiently large n, it is typically the case that

$$\mathbf{P}\{M_n \leq u_n\} \approx F^{n\theta}(u_n), \tag{1.1}$$

where $M_n := \max\{X_1, \ldots, X_n\}$, u_n is any high level such that $n(1-F(u_n))$ converges to a positive number as $n \to \infty$ and θ is a fixed number in $[0,1]$. The θ is called the extremal index of the sequence $\{X_n\}$ (see Section 3 for a rigorous definition) and measures the strength of the dependence of $\{X_n\}$. If

Received February 1995; revised June 1997.

AMS 1991 *subject classifications*. Primary 60G70, 60J05; secondary 60G10, 60G55.

Key words and phrases. Extremal index, multivariate extreme value distributions, exceedance point processes, stationary Markov chains.

$X_1, \ldots, X_n$ are independent, then $\theta = 1$; if $X_1, \ldots, X_n$ are very highly dependent, then $\theta \approx 0$. It is also noted that, when $\theta > 0$, θ^{-1} may be interpreted as the asymptotic mean number of exceedances of a high level in a cluster of the sequence $\{X_n\}$ under a suitable mixing condition, which was justified by Leadbetter (1983).

There are basically two directions of modern research in handling the extremal index θ. One is to find good estimators for θ and study their probabilistic properties. In view of the simple relation (1.1), it is enough to estimate θ and the tail of the marginal d.f. F in order to estimate the tail of the distribution of the maximum M_n. The estimation of θ was discussed by Nandagopalan (1990), Hsing (1991, 1993) and Smith and Weissman (1994). For the estimation of the tail of a probability distribution, see Pickands (1975), Smith (1987) and Dekkers and de Haan (1989).

The other direction is to compute θ when the dependence structure of $\{X_n\}$ is given. It is unfortunately not known how to compute θ analytically except for some special cases. For instance, Berman (1964) showed that a stationary standard Gaussian sequence $\{X_n\}$ has $\theta = 1$ if the autocovariance function $r_n = \mathrm{Cov}(X_1, X_{1+n})$ satisfies $r_n \log n \to 0$ as $n \to \infty$. Later on, Rootzén (1978) computed θ for a class of moving averages of stable processes and Chernick (1981) considered a particular stationary first-order autoregressive sequence to compute θ. For a general stationary sequence, O'Brien (1987) and Rootzén (1988) found similar characterizations for θ, which seem to be suitable for applications to Markov chains but still intractable for computational purpose.

The characterizations were used by Smith (1992a) to find a technique for computing θ of a stationary Markov chain. Assuming the standard Gumbel marginals and the joint distribution of two successive variables being in the domain of attraction of a bivariate extreme value distribution, he showed that the tails of the chain behave like a random walk. As a result, θ is found as the solution to a Wiener-Hopf integral equation. Perfekt (1994) extended this to more general margins. Instead of the standard Gumbel marginals, he assumed that the marginal distribution is in the domain of attraction of a univariate extreme value distribution. Assuming also the existence of a limiting distribution for the transition probabilities which is in fact another interpretation of Smith's assumption for the joint distribution of two successive variables, he found an expression for θ and further characterized the extremal properties of the chain in terms of exceedance point processes.

This paper extends the results of Smith and Perfekt to a real-valued, kth-order stationary Markov chain $\{X_n\}$. The kth-order Markov chains are useful and flexible tools for modeling local dependence in time series. We assume that the joint distribution of $k+1$ consecutive variables belongs to the domain of attraction of some multivariate extreme value distribution, and introduce limiting distributions of some rescaled stationary transition kernels. These limiting transition kernels are used to define a new $(k-1)$th-order Markov chain $\{Y_n\}$, say, from which another kth-order Markov chain $\{Z_n\}$ is defined as $Z_n = Y_1 + \cdots + Y_n$. The chain $\{Z_n\}$ is then effectively used to give a representation for the extremal index of the original chain $\{X_n\}$. The representation is well suited for Monte Carlo simulation in computing the extremal index. For a more complete description of the extremal behavior of the chain $\{X_n\}$, we further discuss convergence in distribution of multilevel exceedance point processes and give representations for quantities characterizing the distributional limits in terms of the chain $\{Z_n\}$.

The rest of the paper is organized as follows. Section 2 discusses the domain of attraction of a multivariate extreme value distribution and constructs the limiting distributions generating $\{Y_n\}$ and thus $\{Z_n\}$. Section 3 gives a representation for the extremal index of $\{X_n\}$ in terms of $\{Z_n\}$ and Section 4 establishes convergence in distribution of multilevel exceedance point processes for $\{X_n\}$. Finally, Section 5 contains four examples which show how our results are applied.

2. Domains of attraction and limiting transition kernels. In this section we consider criteria for domains of attraction of multivariate extreme value distributions and introduce limiting distributions of some rescaled stationary transition kernels of a kth-order stationary Markov chain. These limiting transition kernels are used to define a new $(k-1)$th-order Markov chain which determines the extremal behavior of the original chain.

For $\mathbf{a}, \mathbf{b}, \mathbf{x} \in \Re^p$ $(p \geq 1)$ and for $\alpha, \beta \in \Re$, we write

$$\mathbf{a}\mathbf{x} + \mathbf{b} = (a_1 x_1 + b_1, \dots, a_p x_p + b_p),$$
$$\alpha\mathbf{x} + \beta = (\alpha x_1 + \beta, \dots, \alpha x_p + \beta).$$

For a p-dimensional (dim.) d.f. F_p, if there exist p-dim. vectors $\mathbf{a}_n > \mathbf{0}$ (with componentwise ordering) and $\mathbf{b}_n \in \Re^p$, $n = 1, 2, \dots,$ such that

$$F_p^n(\mathbf{a}_n\mathbf{x} + \mathbf{b}_n) \to_w G_p(\mathbf{x}) \quad \text{as } n \to \infty \tag{2.1}$$

for some p-dim. nondegenerate d.f. G_p, where $\to_w$ denotes weak convergence, then F_p is said to be in the domain of attraction of G_p, written $F_p \in \mathscr{D}(G_p)$, and G_p is called a p-dim. extreme value distribution. Galambos (1987) and Resnick (1987) have good reviews on multivariate extreme value theory.

For $p = 1$, $F_1 \in \mathscr{D}(G_1)$ is equivalent to the condition that there exists a $\xi \in \Re$ such that

$$\lim_{u \uparrow x_{F_1}} \frac{1 - F_1(u + g(u)x)}{1 - F_1(u)} = (1 + \xi x)^{-1/\xi}, \qquad 1 + \xi x > 0, \tag{2.2}$$

where $x_{F_1} := \sup\{x\colon F_1(x) < 1\}$, and

$$\begin{aligned} &x_{F_1} = \infty \text{ and } g(u) = \xi u, && \text{if } \xi > 0,\\ &g(u) \text{ is some strictly positive function}, && \text{if } \xi = 0,\\ &x_{F_1} < \infty \text{ and } g(u) = -\xi(x_{F_1} - u), && \text{if } \xi < 0. \end{aligned} \tag{2.3}$$

In this case, if we take $b_n = \inf\{x\colon F_1(x) \geq 1 - 1/n\}$ and $a_n = g(b_n)$, we have

$$F_1^n(a_n x + b_n) \to \Omega_\xi(x) := \exp\{-(1 + \xi x)^{-1/\xi}\}, \qquad 1 + \xi x > 0, \quad \text{as } n \to \infty. \tag{2.4}$$

Therefore, one may take $G_1 = \Omega_\xi$. Throughout the paper, the case $\xi = 0$ is always interpreted as the limit $\xi \to 0$, that is, $\Omega_0(x) = \exp(-e^{-x})$, $x \in \Re$, the standard Gumbel distribution. When $\xi = 0$, the function g is unique up to asymptotic equivalence, that is, if there is another $\tilde{g}$ satisfying (2.2), then $\tilde{g}(u) \sim g(u)$ as $u \uparrow x_{F_1}$. Condition (2.2) is in fact a reformulation of Theorem 1.6.2 of Leadbetter, Lindgren and Rootzén (1983).

The auxiliary function g in (2.3) is assumed to satisfy the following properties: as $u \uparrow x_{F_1}$,

$$
\begin{aligned}
&u + g(u)x \to x_{F_1} \quad \text{for any } x \text{ with } 1 + \xi x > 0,\\
&g(u + g(u)x)/g(u) \to 1 + \xi x \quad \text{locally uniformly in } x \text{ with } 1 + \xi x > 0, \qquad (2.5)\\
&(x_{F_1} - u)/g(u) \to 1/|\xi| \quad \text{if } x_{F_1} < \infty.
\end{aligned}
$$

This is obviously true for $\xi \neq 0$. When $\xi = 0$, this is of no loss of generality because, if (2.2) holds for some g which does not satisfy (2.5), then there exists another $\tilde{g}$ which satisfies (2.2) and (2.5), with $\tilde{g}(u) \sim g(u)$ as $u \uparrow x_{F_1}$. For details, see Lemmas 1.2 and 1.3 and Proposition 1.4 of Resnick (1987).

For $p \geq 2$, G_p in (2.1) has no such finite-parameter representation as in (2.4) for G_1. A higher-dimensional extension of (2.2) is, however, possible due to Marshall and Olkin (1983). For simplicity, F_p is assumed to have equal univariate marginals. That is, let $F_p(\mathbf{x})$ be a p-dim. d.f. with equal univariate marginals $F_1(x)$, and let $G_p(\mathbf{x})$ be a p-dim. extreme value distribution with equal univariate marginals $G_1(x) = \Omega_\xi(x)$ for some $\xi \in \Re$. Then $F_p \in \mathscr{D}(G_p)$ if and only if

$$
\lim_{u \uparrow x_{F_1}} \frac{1 - F_p(u + g(u)\mathbf{x})}{1 - F_1(u)} = -\log G_p(\mathbf{x}), \qquad 1 + \xi \mathbf{x} > \mathbf{0}, \tag{2.6}
$$

where x_{F_1} and g satisfy (2.3) and (2.5). This is a reformulation of Propositions 3.1–3.3 of Marshall and Olkin (1983). As a representation of G_p, Pickands (1981) showed that there exists a finite positive measure Q_p on the $(p-1)$-dim. unit simplex $S_p = \{\mathbf{w} \in \Re^p \colon \mathbf{w} \geq \mathbf{0},\ \sum_{i=1}^p w_i = 1\}$ satisfying $\int_{S_p} w_i \, dQ_p(\mathbf{w}) = 1$, $i = 1, \dots, p$, such that

$$
G_p(\mathbf{x}) = \exp\left[-\int_{S_p} \max_{1 \le i \le p} \{w_i(1 + \xi x_i)^{-1/\xi}\} \, dQ_p(\mathbf{w})\right], \qquad 1 + \xi \mathbf{x} > \mathbf{0}. \tag{2.7}
$$

Now, let $\{X_n \colon n \geq 1\}$ denote a real-valued, kth-order stationary Markov chain. Here, $\{X_n\}$ being a kth-order Markov chain means that, for every n, the conditional distribution of X_n given the past depends only on the k immediate past values. It is clear from the stationarity of $\{X_n\}$ that the distribution of the whole chain is determined by the joint distribution F_{k+1}, say, of $(X_n, \dots, X_{n+k})$ via its successive transition kernels. Assuming F_{k+1} is absolutely continuous, we use the following notation: for $i = 1, \dots, k + 1$, we write

$$
\begin{aligned}
\mathbf{x}_i &= (x_1, \dots, x_i) \in \Re^i, \qquad \mathbf{X}_i = (X_1, \dots, X_i),\\
F_i(\mathbf{x}_i) &= F_i(x_1, \dots, x_i) = \mathbf{P}\{X_{n+1} \leq x_1, \dots, X_{n+i} \leq x_i\},\\
f_i(\mathbf{x}_i) &= \partial^i F_i(\mathbf{x}_i)/(\partial x_1 \dots \partial x_i),
\end{aligned}
$$

and, for $j = 1, \dots, k$, we write

$$\mathbf{x}_{j+1} = (x_1, \dots, x_j, x_{j+1}) = (\mathbf{x}_j, x_{j+1}) \in \Re^{j+1},$$
$$F_{j+1}(x_{j+1}|\mathbf{x}_j) = \mathbf{P}\{X_{n+j+1} \le x_{j+1}|(X_{n+1}, \dots, X_{n+j}) = (x_1, \dots, x_j)\},$$
$$f_{j+1}(x_{j+1}|\mathbf{x}_j) = \partial F_{j+1}(x_{j+1}|\mathbf{x}_j)/\partial x_{j+1} = f_{j+1}(\mathbf{x}_{j+1})/f_j(\mathbf{x}_j).$$

We now assume that $F_{k+1} \in \mathscr{D}(G_{k+1})$ for some $(k+1)$-dim. extreme value distribution G_{k+1} with equal univariate marginals $G_1 = \Omega_\xi$ for some $\xi \in \Re$. Then this implies that $F_i \in \mathscr{D}(G_i)$, $i = 1, \dots, k+1$, where $G_i(\mathbf{x}_i) := G_{k+1}(\mathbf{x}_i, x_{G_1}, \dots, x_{G_1})$, an i-dim. marginal of G_{k+1}. We therefore have, from (2.2) and (2.6), that, for each $i = 1, \dots, k+1$,

$$\lim_{u \uparrow x_{F_1}} \frac{1 - F_i(u + g(u)\mathbf{x}_i)}{1 - F_1(u)} = -\log G_i(\mathbf{x}_i), \qquad 1 + \xi\mathbf{x}_i > \mathbf{0}, \tag{2.8}$$

where x_{F_1} and g satisfy (2.3) and (2.5). Taking partial derivatives in these convergences gives a plausibility of the following essential assumption under which the theory of this paper is developed.

ASSUMPTION A. Suppose that

$$\lim_{u \uparrow x_{F_1}} \frac{g(u) f_1(u + g(u)x_1)}{1 - F_1(u)} = (1 + \xi x_1)^{-1/\xi - 1}, \qquad 1 + \xi x_1 > 0,$$

and that, for each $j = 1, \dots, k$,

$$l_j(x_{j+1}; \mathbf{x}_j) := \lim_{u \uparrow x_{F_1}} g(u) f_{j+1}(u + g(u)x_{j+1}|u + g(u)\mathbf{x}_j), \qquad 1 + \xi\mathbf{x}_{j+1} > \mathbf{0},$$

exist finite.

LEMMA 2.1. *For each $j = 1, \dots, k$, $(1 + \xi x_{j+1}) l_j(x_{j+1}; \mathbf{x}_j)$ must be a function of*

$$\nabla\mathbf{x}_{j+1} := \left(\frac{1}{\xi}\log\left(\frac{1 + \xi x_2}{1 + \xi x_1}\right), \dots, \frac{1}{\xi}\log\left(\frac{1 + \xi x_{j+1}}{1 + \xi x_j}\right)\right),$$

provided that f_{k+1} is continuous.

PROOF. It follows from the properties (2.5) of g that

$$\begin{aligned} &l_j(0; (1 + \xi x_{j+1})^{-1}(\mathbf{x}_j - x_{j+1})) \\ &\quad = \lim_{u \uparrow x_{F_1}} g(u) f_{j+1}(u|u + g(u)(1 + \xi x_{j+1})^{-1}(\mathbf{x}_j - x_{j+1})) \\ &\quad = (1 + \xi x_{j+1}) l_j(x_{j+1}; \mathbf{x}_j). \end{aligned}$$

Therefore, $(1 + \xi x_{j+1}) l_j(x_{j+1}; \mathbf{x}_j)$ is a function of $(1 + \xi x_{j+1})^{-1}(\mathbf{x}_j - x_{j+1})$ and thus a function of $\nabla\mathbf{x}_{j+1}$. $\square$

It is noted that, when $\xi = 0$, $\triangledown \mathbf{x}_{j+1}$ is interpreted as $(x_2 - x_1, \dots, x_{j+1} - x_j)$. The $l_j(x_{j+1}; \mathbf{x}_j)$ is usually (but not always) given by

$$\left(\frac{\partial^{j+1} \log G_{j+1}(\mathbf{x}_{j+1})}{\partial x_1 \dots \partial x_{j+1}}\right) \Big/ \left(\frac{\partial^j \log G_j(\mathbf{x}_j)}{\partial x_1 \dots \partial x_j}\right).$$

Assuming f_{k+1} is continuous, we define a function $h_j(y_j; y_1, \dots, y_{j-1})$ on $\Re^j$ by letting

$$h_j\left(\frac{1}{\xi}\log\left(\frac{1+\xi x_{j+1}}{1+\xi x_j}\right); \triangledown \mathbf{x}_j\right) := (1+\xi x_{j+1}) l_j(x_{j+1}; \mathbf{x}_j), \qquad 1+\xi \mathbf{x}_{j+1} > \mathbf{0},$$

which is well defined by Lemma 2.1. We also define, for $y_1, \dots, y_{j-1} \in \Re$,

$$H_j(y; y_1, \dots, y_{j-1}) := 1 - \int_y^\infty h_j(t; y_1, \dots, y_{j-1})\, dt, \qquad y \in \{-\infty\} \cup \Re.$$

Then, for every fixed $\mathbf{y}_{j-1} \in \Re^{j-1}$, $H_j(y; \mathbf{y}_{j-1})$ may be considered as a d.f. of y on $\{-\infty\} \cup \Re$, where it has a mass $1 - \int_{-\infty}^\infty h_j(t; \mathbf{y}_{j-1})\, dt$ at $y = -\infty$, which is possibly positive, and is absolutely continuous on $\Re$, having $h_j(y; \mathbf{y}_{j-1})$ as its density. It is noted that the following are equivalent:

$$\text{(2.9)}\quad \begin{array}{lll} \text{(i)} & \forall \mathbf{y}_{j-1} \in \Re^{j-1}, & H_j(-\infty; \mathbf{y}_{j-1}) = 0; \\ \text{(ii)} & \forall \mathbf{y}_{j-1} \in \Re^{j-1}, & h_j(\,\cdot\,; \mathbf{y}_{j-1}) \text{ is a p.d.f. on } \Re; \\ \text{(iii)} & \forall \mathbf{x}_j \text{ with } 1+\xi\mathbf{x}_j > \mathbf{0}, & l_j(\,\cdot\,; \mathbf{x}_j) \text{ is a p.d.f. on } \{x\colon 1+\xi x > 0\}. \end{array}$$

By p.d.f. we mean probability density function. In fact, these equivalent conditions imply that, as $u \uparrow x_{F_1}$,

$$F_{j+1}(u + g(u)x_{j+1} | u + g(u)\mathbf{x}_j) \to H_j\left(\frac{1}{\xi}\log\left(\frac{1+\xi x_{j+1}}{1+\xi x_j}\right); \triangledown \mathbf{x}_j\right), \qquad 1+\xi\mathbf{x}_{j+1} > \mathbf{0}.$$

The limiting distributions H_j, $j = 1, \dots, k$, constructed above can generate a $\{-\infty\} \cup \Re$-valued, $(k-1)$th-order Markov chain $\{Y_n\colon n \ge 1\}$ as follows:

1. $Y_1 \sim H_1(y_1)$.
2. For $j = 2, \dots, k$, $Y_j | (Y_1, \dots, Y_{j-1}) \sim H_j(y_j; Y_1, \dots, Y_{j-1})$ if $Y_1, \dots, Y_{j-1} > -\infty$; put $Y_j = -\infty$, otherwise.
3. For $j \ge k+1$, $Y_j | (Y_{j-k+1}, \dots, Y_{j-1}) \sim H_k(y_j; Y_{j-k+1}, \dots, Y_{j-1})$ if $Y_{j-k+1}, \dots, Y_{j-1} > -\infty$; put $Y_j = -\infty$, otherwise.

When $k = 1$, $\{Y_n\}$ becomes a sequence of i.i.d. r.v.'s with d.f. H_1. We also remark that, when $k \ge 2$, the state $-\infty$ is an absorbing state of the chain $\{Y_n\}$, that is, once the chain visits the state $-\infty$, it must stay there forever. Finally, define

$$\text{(2.10)}\qquad Z_n := Y_1 + \dots + Y_n, \qquad n = 1, 2, \dots,$$

so that $\{Z_n\colon n \ge 1\}$ is a $\{-\infty\} \cup \Re$-valued, kth-order Markov chain. This chain plays an important role in examining the extremal behavior of the original chain $\{X_n\}$ (see Theorems 3.1 and 4.1).

Now, suppose that we only know $F_1 \in \mathscr{D}(\Omega_\xi)$ for some $\xi \in \Re$, with auxiliary function g satisfying (2.3) and (2.5), and that Assumption A holds. Then it can be shown that $F_{k+1} \in \mathscr{D}(G_{k+1})$ for some $(k+1)$-dim. extreme value distribution G_{k+1} with equal univariate marginals $G_1 = \Omega_\xi$ if, for each $j = 1, \dots, k$, (2.9) holds [see Yun (1994) for details]. This situation gives a clear description of the tail behavior of $\{X_n\}$ as follows. Define $\phi(x) := \xi^{-1}(e^{\xi x} - 1)$ and $x^*_{F_1} := \inf\{x\colon F_1(x) > 0\}$.

LEMMA 2.2. *Let $\{X_n\colon n \ge 1\}$ be a kth-order stationary Markov chain, and let F_{k+1} be the d.f. of $(X_1, \dots, X_{k+1})$ having a continuous p.d.f. f_{k+1}. Suppose that $F_1 \in \mathscr{D}(\Omega_\xi)$ for some $\xi \in \Re$, with auxiliary function g satisfying* (2.3) *and* (2.5), *and that Assumption A holds. Let $\{Z_n\colon n \ge 1\}$ be the kth-order Markov chain defined by* (2.10), *and let T be an* Exp(1)*-distributed r.v. which is independent of $\{Z_n\}$. If, for each $j = 1, \dots, k$,* (2.9) *holds, then, for every $p \ge 2$, as $u \uparrow x_{F_1}$,*

$$\begin{aligned}&\mathbf{P}\big\{((X_2 - u)/g(u), \dots, (X_p - u)/g(u)) \in \cdot | X_1 > u\big\}\\ &\qquad \to_w \mathbf{P}\big\{(\phi(Z_1 + T), \dots, \phi(Z_{p-1} + T)) \in \cdot\big\}.\end{aligned}$$

PROOF. We use the notation

$$f_1^{(u)}(x_1) := \frac{g(u) f_1(u + g(u)x_1)}{1 - F_1(u)};$$

$$f_{j+1}^{(u)}(x_{j+1} | \mathbf{x}_j) := g(u) f_{j+1}(u + g(u)x_{j+1} | u + g(u)\mathbf{x}_j), \qquad 1 \le j \le k;$$

$$\psi_p^{(u)}(\mathbf{x}_p) := \begin{cases} \displaystyle\prod_{j=1}^{p-1} f_{j+1}^{(u)}(x_{j+1} | \mathbf{x}_j), \quad 2 \le p \le k, \\ \displaystyle\left(\prod_{j=1}^{k-1} f_{j+1}^{(u)}(x_{j+1} | \mathbf{x}_j)\right)\left(\prod_{j=k}^{p-1} f_{k+1}^{(u)}(x_{j+1} | x_{j-k+1}, \dots, x_j)\right), \\ \hfill p \ge k+1; \end{cases}$$

$$\psi_{p-1}(\mathbf{x}_p) := \begin{cases} \displaystyle\prod_{j=1}^{p-1} l_j(x_{j+1}; \mathbf{x}_j), \quad 2 \le p \le k, \\ \displaystyle\left(\prod_{j=1}^{k-1} l_j(x_{j+1}; \mathbf{x}_j)\right)\left(\prod_{j=k}^{p-1} l_k(x_{j+1}; x_{j-k+1}, \dots, x_j)\right), \\ \hfill p \ge k+1; \end{cases}$$

$$\eta_{p-1}(\mathbf{y}_{p-1}) := \begin{cases} \displaystyle\prod_{j=1}^{p-1} h_j(y_j; \mathbf{y}_{j-1}), \quad 2 \le p \le k, \\ \displaystyle\left(\prod_{j=1}^{k-1} h_j(y_j; \mathbf{y}_{j-1})\right)\left(\prod_{j=k}^{p-1} h_k(y_j; y_{j-k+1}, \dots, y_{j-1})\right), \\ \hfill p \ge k+1. \end{cases}$$

Then, for every $p \geq 2$, since $\lim_{u \uparrow x_{F_1}} f_1^{(u)}(x_1)\psi_p^{(u)}(\mathbf{x}_p) = (1+\xi x_1)^{-1/\xi-1}\psi_{p-1}(\mathbf{x}_p)$, $1 + \xi\mathbf{x}_p > \mathbf{0}$, and since $(1+\xi x_1)^{-1/\xi-1}\psi_{p-1}(\mathbf{x}_p)$ is a p.d.f. on $(0, x_{\Omega_\xi}) \times (x^*_{\Omega_\xi}, x_{\Omega_\xi})^{p-1}$, we have, by Scheffé's theorem [Billingsley (1985), Theorem 16.11], for any $\mathbf{w}_{p-1}$ with $1 + \xi\mathbf{w}_{p-1} > \mathbf{0}$, as $u \uparrow x_{F_1}$,

$$\begin{aligned}
&\mathbf{P}\{(X_2 - u)/g(u) \leq w_1, \ldots, (X_p - u)/g(u) \leq w_{p-1} | X_1 > u\} \\
&\quad = \int_0^\infty \int_{-\infty}^{w_1} \cdots \int_{-\infty}^{w_{p-1}} f_1^{(u)}(x_1)\psi_p^{(u)}(\mathbf{x}_p)\, dx_p \ldots dx_2\, dx_1 \\
&\quad \to \int_0^{x_{\Omega_\xi}} \int_{x^*_{\Omega_\xi}}^{w_1} \cdots \int_{x^*_{\Omega_\xi}}^{w_{p-1}} (1+\xi x_1)^{-1/\xi-1}\psi_{p-1}(\mathbf{x}_p)\, dx_p \ldots dx_2\, dx_1.
\end{aligned}$$

This, if we put $x_1 = \phi(t)$ and $x_j = \phi(t + y_1 + \cdots + y_{j-1})$, $j = 2, \ldots, p$, is equal to

$$\begin{aligned}
&\int_0^\infty \int_{-\infty}^{\phi^{-1}(w_1)-t} \int_{-\infty}^{\phi^{-1}(w_2)-t-y_1} \cdots \int_{-\infty}^{\phi^{-1}(w_{p-1})-t-\sum_{j=1}^{p-2} y_j} \\
&\qquad\qquad\qquad \times e^{-t}\eta_{p-1}(\mathbf{y}_{p-1})\, dy_{p-1} \ldots dy_2\, dy_1\, dt \\
&\quad = \mathbf{P}\Bigg\{Y_1 \leq \phi^{-1}(w_1) - T, Y_2 \leq \phi^{-1}(w_2) - T - Y_1, \ldots, \\
&\qquad\qquad Y_{p-1} \leq \phi^{-1}(w_{p-1}) - T - \sum_{j=1}^{p-2} Y_j\Bigg\} \\
&\quad = \mathbf{P}\{\phi(Z_1 + T) \leq w_1, \ldots, \phi(Z_{p-1} + T) \leq w_{p-1}\},
\end{aligned}$$

where $\{Y_n\}$ is the chain in (2.10). This completes the proof. □

Although Lemma 2.2 is useful for obtaining a representation of the extremal index of $\{X_n\}$ (see Section 3), it cannot be applied when (2.9) fails for some $j = 1, \ldots, k$. In Section 3, we shall handle this case by imposing Assumption B below in addition to Assumption A. For this, we need the following definition.

DEFINITION 2.3. A class of real-valued functions defined on $\Re$ is said to be locally uniformly integrable over an unbounded interval if the class is uniformly integrable over any compact subset of that interval.

For example, real-valued functions defined on $\Re$ are locally uniformly integrable if they are dominated by a continuous function.

ASSUMPTION B. Suppose that there exists a $u^* < x_{F_1}$ such that the class

$$\left\{\frac{g(u)f_1(u + g(u)x_1)}{1 - F_1(u)} \colon u^* \leq u < x_{F_1}\right\}$$

of functions of x_1 is locally uniformly integrable over $(x_{\Omega_\xi}, x^*_{\Omega_\xi})$, and that, for each $j = 1, \ldots, k$ and for every fixed $\mathbf{x}_j$ with $1 + \xi\mathbf{x}_j > \mathbf{0}$, there exists a

$u_j^*(\mathbf{x}_j) < x_{F_1}$ such that the class

$$\{g(u)f_{j+1}(u + g(u)x_{j+1}|u + g(u)\mathbf{x}_j) \colon u_j^*(\mathbf{x}_j) \le u < x_{F_1}\}$$

of functions of x_{j+1} is also locally uniformly integrable over $(x_{\Omega_\xi}, x^*_{\Omega_\xi})$.

Assumptions A and B are satisfied by a large class of models for F_{k+1}. In particular, the following lemma shows that Assumptions A and B hold automatically if F_{k+1} is itself a multivariate extreme value distribution. We here remark that, when $\xi > 0$, the auxiliary function $g(u) = \xi u$ in (2.3) can be replaced by $g(u) = 1 + \xi u$ without any distortion in developing the whole story of this paper. For convenience of proof, this replacement is adopted in the following lemma.

LEMMA 2.4. *Let $\{X_n \colon n \ge 1\}$ be a kth-order stationary Markov chain, and let F_{k+1} be the d.f. of $(X_1, \dots, X_{k+1})$ having a continuous p.d.f. f_{k+1}. If F_{k+1} is a multivariate extreme value distribution with $F_1 = \Omega_\xi$ for some $\xi \in \Re$, then Assumptions A and B hold and the limit $l_j(x_{j+1}; \mathbf{x}_j)$ in Assumption A is given by*

$$\left(\frac{\partial^{j+1} \log F_{j+1}(\mathbf{x}_{j+1})}{\partial x_1 \dots \partial x_{j+1}}\right) \Big/ \left(\frac{\partial^j \log F_j(\mathbf{x}_j)}{\partial x_1 \dots \partial x_j}\right), \tag{2.11}$$

provided that $\partial^j \log F_j(\mathbf{x}_j)/(\partial x_1 \dots \partial x_j)$ is not zero, where $g(u) = 1 + \xi u$.

PROOF. When $\xi = 0$, the validity of the choice $g(u) = 1$ follows from the fact that $F_1 = \Omega_0$. First, we check Assumption A. The first statement is elementary and so the details are omitted here. For the second statement, for each $i = 1, \dots, k+1$, let $V_i(\mathbf{x}_i) := -\log F_i(\mathbf{x}_i)$ and consider

$$\begin{aligned} f_i(\mathbf{x}_i) &= \frac{\partial^i \exp\{-V_i(\mathbf{x}_i)\}}{\partial x_1 \dots \partial x_i} \\ &= \exp\{-V_i(\mathbf{x}_i)\}\left[-\frac{\partial^i V_i(\mathbf{x}_i)}{\partial x_1 \dots \partial x_i} + \dots + (-1)^i \prod_{s=1}^{i} \frac{\partial V_i(\mathbf{x}_i)}{\partial x_s}\right] \\ &= \exp\{-V_i(\mathbf{x}_i)\} \sum_{m=1}^{i} b_{i,m}(\mathbf{x}_i), \qquad 1 + \xi \mathbf{x}_i > \mathbf{0}, \end{aligned}$$

where

$$b_{i,m}(\mathbf{x}_i) := (-1)^m \sum_{\{D_{i,1}, \dots, D_{i,m}\}} \prod_{r=1}^{m} \frac{\partial^{|D_{i,r}|} V_i(\mathbf{x}_i)}{\langle \partial x_s \rangle_{s \in D_{i,r}}},$$

the $\{D_{i,1}, \dots, D_{i,m}\}$ varying on the partitions of $\{1, \dots, i\}$ such that each of $D_{i,1}, \dots, D_{i,m}$ contains at least one element. By $\langle \partial x_s \rangle_{s \in D_{i,r}}$ we mean $\partial x_{i_1} \cdots \partial x_{i_p}$ if $D_{i,r} = \{i_1 < \dots < i_p\}$. Thus, for any u with $g(u) = 1 + \xi u > 0$,

$$f_i(u + g(u)\mathbf{x}_i) = \exp\{-V_i(u + g(u)\mathbf{x}_i)\} \sum_{m=1}^{i} b_{i,m}(u + g(u)\mathbf{x}_i).$$

Here, since F_i is an i-dim. extreme value distribution, we see, from (2.7) with p, G_p replaced by i, F_i respectively, that, for any $t > 0$, $V_i(\xi^{-1}(t^\xi - 1) + t^\xi \mathbf{x}_i) = t^{-1}V_i(\mathbf{x}_i)$, and so, by taking $t = (g(u))^{1/\xi} = (1 + \xi u)^{1/\xi} > 0$, we have $V_i(u + g(u)\mathbf{x}_i) = V_i(\xi^{-1}(t^\xi - 1) + t^\xi \mathbf{x}_i) = t^{-1}V_i(\mathbf{x}_i) = (g(u))^{-1/\xi}V_i(\mathbf{x}_i)$. Further, since, for each $r = 1, \dots, m$,

$$\frac{\partial^{|D_{i,r}|}V_i(u + g(u)\mathbf{x}_i)}{\langle \partial(u + g(u)x_s)\rangle_{s\in D_{i,r}}} = (g(u))^{-|D_{i,r}|}\frac{\partial^{|D_{i,r}|}V_i(u + g(u)\mathbf{x}_i)}{\langle \partial x_s\rangle_{s\in D_{i,r}}}$$

and $|D_{i,1}| + \cdots + |D_{i,m}| = i$, we have $b_{i,m}(u + g(u)\mathbf{x}_i) = (g(u))^{-i-m/\xi}b_{i,m}(\mathbf{x}_i)$. Therefore, for any u with $g(u) = 1 + \xi u > 0$,

$$f_i(u + g(u)\mathbf{x}_i) = (g(u))^{-i}\exp\{-(g(u))^{-1/\xi}V_i(\mathbf{x}_i)\}\sum_{m=1}^{i}(g(u))^{-m/\xi}b_{i,m}(\mathbf{x}_i).$$

This enables us to write that, for each $j = 1, \dots, k$ and for any $u, \mathbf{x}_{j+1}$ with $g(u) = 1 + \xi u > 0$ and $1 + \xi\mathbf{x}_{j+1} > \mathbf{0}$,

$$\begin{aligned} &g(u)f_{j+1}(u + g(u)x_{j+1}|u + g(u)\mathbf{x}_j)\\ &\quad = g(u)f_{j+1}(u + g(u)\mathbf{x}_{j+1})/f_j(u + g(u)\mathbf{x}_j)\\ &\quad = \exp\{-(g(u))^{-1/\xi}(V_{j+1}(\mathbf{x}_{j+1}) - V_j(\mathbf{x}_j))\}\\ &\qquad \times \Bigg(\sum_{m=1}^{j+1}(g(u))^{-m/\xi}b_{j+1,m}(\mathbf{x}_{j+1})\Bigg)\Bigg/\Bigg(\sum_{m=1}^{j}(g(u))^{-m/\xi}b_{j,m}(\mathbf{x}_j)\Bigg). \end{aligned} \tag{2.12}$$

Since $(g(u))^{-1/\xi} \to 0$ as $u \uparrow x_{F_1} = x_{\Omega_\xi}$, the functional form of (2.12) guarantees the existence of the limit $l_j(x_{j+1}; \mathbf{x}_j)$ in Assumption A. Moreover, if $\partial^j V_j(\mathbf{x}_j)/(\partial x_1 \dots \partial x_j)$ is not zero, then it is obvious from (2.12) that $l_j(x_{j+1}; \mathbf{x}_j)$ is given by (2.11).

Second, we check Assumption B. For the first statement, observe that, for any u with $g(u) = 1 + \xi u > 0$,

$$\frac{g(u)f_1(u + g(u)x_1)}{1 - F_1(u)} = a(u)(1 + \xi x_1)^{-1/\xi-1}\exp\{-(g(u))^{-1/\xi}(1 + \xi x_1)^{-1/\xi}\},$$

$1 + \xi x_1 > 0$, where $a(u) = (g(u))^{-1/\xi}/[1 - \exp\{-(g(u))^{-1/\xi}\}]$. Here, since $a(u) \to 1$ as $u \uparrow x_{F_1}$, there exists a u^* such that $a(u) \le 3/2$ whenever $u^* \le u < x_{F_1}$. Thus, whenever $u^* \le u < x_{F_1}$,

$$\frac{g(u)f_1(u + g(u)x_1)}{1 - F_1(u)} \le \frac{3}{2}(1 + \xi x_1)^{-1/\xi-1}, \qquad 1 + \xi x_1 > 0,$$

where the right-hand side is a continuous function of x_1. For the second statement, consider (2.12) again. Here, fix $\mathbf{x}_j$ with $1 + \xi\mathbf{x}_j > \mathbf{0}$ and let m^* denote the smallest m, $1 \le m \le j$, such that $b_{j,m}(\mathbf{x}_j) \ne 0$, that is, $b_{j,m}(\mathbf{x}_j) = 0$ for $1 \le m \le m^* - 1$ and $b_{j,m^*}(\mathbf{x}_j) \ne 0$. Then, for $1 \le m \le m^* - 1$, $b_{j+1,m}(\mathbf{x}_{j+1}) = 0$

also for any x_{j+1} with $1+\xi x_{j+1} > 0$ since, otherwise, $l_j(x_{j+1}; \mathbf{x}_j)$ cannot be finite. Thus (2.12) is reduced to

$$\begin{aligned} g(u) f_{j+1}(u + g(u)x_{j+1} | u + g(u)\mathbf{x}_j) &\\ = \exp\{-(g(u))^{-1/\xi}(V_{j+1}(\mathbf{x}_{j+1}) - V_j(\mathbf{x}_j))\} &\\ \times \left(\sum_{r=0}^{j+1-m^*} (g(u))^{-r/\xi} b_{j+1,\, m^*+r}(\mathbf{x}_{j+1})\right) \bigg/ \left(\sum_{r=0}^{j-m^*} (g(u))^{-r/\xi} b_{j,\, m^*+r}(\mathbf{x}_j)\right). \end{aligned}$$

Here, the denominator of the right-hand side converges to $b_{j,\, m^*}(\mathbf{x}_j)$ as $u \uparrow x_{F_1}$, which implies $b_{j,\, m^*}(\mathbf{x}_j) > 0$ since, otherwise, that is, if $b_{j,\, m^*}(\mathbf{x}_j) < 0$, then $f_j(u+g(u)\mathbf{x}_j) < 0$ for all u sufficiently close to x_{F_1}. Also, there exists a $u_j^*(\mathbf{x}_j)$ such that $\sum_{r=0}^{j-m^*} (g(u))^{-r/\xi} b_{j,\, m^*+r}(\mathbf{x}_j) \geq (1/2) b_{j,\, m^*}(\mathbf{x}_j)$ whenever $u_j^*(\mathbf{x}_j) \leq u < x_{F_1}$. Therefore, since $V_{j+1}(\mathbf{x}_{j+1}) \geq V_j(\mathbf{x}_j)$, we conclude that, whenever $u_j^*(\mathbf{x}_j) \leq u < x_{F_1}$,

$$\begin{aligned} g(u) f_{j+1}(u + g(u)x_{j+1} | u + g(u)\mathbf{x}_j) &\\ \leq \frac{2}{b_{j,\, m^*}(\mathbf{x}_j)} \sum_{r=0}^{j+1-m^*} (g(u_j^*(\mathbf{x}_j)))^{-r/\xi} |b_{j+1,\, m^*+r}(\mathbf{x}_{j+1})|, \end{aligned}$$

$1+\xi x_{j+1} > 0$, where the right-hand side is a continuous function of x_{j+1}. The proof is complete. □

3. Representation of extremal index. In this section, we derive a representation for the extremal index of a kth-order stationary Markov chain using the chain defined by (2.10). This turns out to be very useful in computing the extremal index by means of Monte Carlo simulation.

Let $\{X_n : n \geq 1\}$ be a stationary sequence of r.v.'s with marginal d.f. F_1. For a sequence $\{u_n : n \geq 1\}$ of real numbers, we write: for $1 \leq s \leq t \leq n$, define $\mathscr{A}_{s,\,t}^{(n)}$ to be the class of all intersections of the events $\{X_i \leq u_n\}$, $s \leq i \leq t$, and, for $q = 1, \ldots, n-1$, write

$$\alpha_{n,\,q} = \max\Big\{|\mathbf{P}(A \cap B) - \mathbf{P}(A)\mathbf{P}(B)| : A \in \mathscr{A}_{1,\,s}^{(n)},\ B \in \mathscr{A}_{s+q,\,n}^{(n)}, 1 \leq s \leq n-q\Big\}. \tag{3.1}$$

If $\alpha_{n,\,q_n} \to 0$ as $n \to \infty$ for some sequence $q_n = o(n)$, then the condition $D(u_n)$ is said to hold for $\{X_n\}$. The condition $D(u_n)$, which is weaker than the usual strong mixing, was introduced by Leadbetter (1974) and is a widely used mixing condition to restrict long-range dependence of the events $\{X_i \leq u_n\}$.

Assume that, for each $\tau > 0$, there exists a sequence $\{u_n(\tau) : n \geq 1\}$ of real numbers such that

$$n(1 - F_1(u_n(\tau))) \to \tau \quad \text{as } n \to \infty. \tag{3.2}$$

The sequence $\{X_n\}$ is said to have extremal index θ, $0 \leq \theta \leq 1$, if, for each $\tau > 0$, $\mathbf{P}\{M_n \leq u_n(\tau)\} \to e^{-\theta\tau}$ as $n \to \infty$. It is noted that, by Theorem 1.7.13

of Leadbetter, Lindgren and Rootzén (1983), if F_1 is continuous, then, for each $\tau > 0$, there always exists a sequence $\{u_n(\tau)\}$ satisfying (3.2). Now, assume that, for each $\tau > 0$, $D(u_n(\tau))$ holds for $\{X_n\}$. Then, by O'Brien (1987) [see also Rootzén (1988)], the sequence $\{X_n\}$ has extremal index θ if and only if, for some $\tau_0 > 0$,

$$\mathbf{P}\{X_i \le u_n(\tau_0),\ 2 \le i \le p_n | X_1 > u_n(\tau_0)\} \to \theta \quad \text{as } n \to \infty, \tag{3.3}$$

where $\{p_n: n \ge 1\}$ is a sequence of positive integers such that

$$p_n = o(n), \qquad n\alpha_{n,q_n} = o(p_n), \qquad q_n = o(p_n), \tag{3.4}$$

the $\alpha_{n,q}$ being given in (3.1) with $u_n = u_n(\tau_0)$ and the $q_n = o(n)$ being such that $\alpha_{n,q_n} \to 0$ (e.g., one may take p_n as the integer part of $\max\{n\alpha_{n,q_n}^{1/2}, (nq_n)^{1/2}\}$). Following Smith (1992a) [see also Perfekt (1994)], if we further assume that

$$\lim_{p\to\infty} \overline{\lim_{n\to\infty}} \sum_{i=p}^{p_n} \mathbf{P}\{X_i > u_n | X_1 > u_n\} = 0 \tag{3.5}$$

with $u_n = u_n(\tau_0)$, then (3.3) is equivalent to the existence of

$$\theta = \lim_{p\to\infty} \lim_{u\uparrow x_{F_1}} \mathbf{P}\{X_i \le u,\ 2 \le i \le p | X_1 > u\}. \tag{3.6}$$

In case of a Markov chain, (3.6) can be reduced to a more appealing form from a computational point of view. Specifically, let $\{X_n: n \ge 1\}$ be a kth-order stationary Markov chain. Then, Lemma 2.2 says that, under the same assumptions of the lemma, (3.6) is reduced to

$$\theta = \lim_{p\to\infty} \mathbf{P}\{Z_1 \le -T, \dots, Z_p \le -T\} = \mathbf{P}\Big\{\sup_{i\ge 1} Z_i \le -T\Big\}, \tag{3.7}$$

where $\{Z_n: n \ge 1\}$ is the kth-order Markov chain defined by (2.10) and T an Exp(1)-distributed r.v. which is independent of $\{Z_n\}$. Representation (3.7) is particularly helpful in computing θ by simulation. However, it should be noted that (2.9) was required to hold for each $j = 1, \dots, k$ in Lemma 2.2. The following theorem handles the general case where (2.9) may fail possibly for some $j = 1, \dots, k$.

THEOREM 3.1. *Let $\{X_n: n \ge 1\}$ be a kth-order stationary Markov chain, and let F_{k+1} be the d.f. of $(X_1, \dots, X_{k+1})$ having a continuous p.d.f. f_{k+1} such that $F_{k+1} \in \mathscr{D}(G_{k+1})$ with auxiliary function g satisfying* (2.3) *and* (2.5), *where G_{k+1} is some $(k+1)$-dim. extreme value distribution with equal univariate marginals $G_1 = \Omega_\xi$ for some $\xi \in \Re$. Suppose that Assumptions A and B hold and that*

$$\lim_{M\to\infty} \overline{\lim_{u\uparrow x_{F_1}}} \sup\Big\{\mathbf{P}\{X_{k+1} > u | \mathbf{X}_k = \mathbf{x}_k\}: \min_{1\le i\le k} x_i \le u - g(u)(1 - M^{-\xi})/\xi\Big\} = 0. \tag{3.8}$$

Let $\{u_n(\tau)\}$, $\tau > 0$, be sequences satisfying (3.2). *Assume that $D(u_n(\tau))$ holds for each $\tau > 0$ and that, for some $\tau_0 > 0$,* (3.5) *holds with $u_n = u_n(\tau_0)$, where $\{p_n\}$ is a sequence satisfying* (3.4). *Then, $\{X_n\}$ has extremal index θ given by*

$$\theta = \log(G_k(\mathbf{0})/G_{k+1}(\mathbf{0})) - \mathbf{P}\Big\{\max_{1\le i\le k} Z_i \le -T,\ \sup_{i\ge k+1} Z_i > -T\Big\}, \tag{3.9}$$

where $\{Z_n: n \ge 1\}$ is the kth-order Markov chain defined by (2.10) *and T an* Exp(1)*-distributed r.v. which is independent of $\{Z_n\}$.*

PROOF. We begin with (3.6). Letting $\theta_p^{(u)} = \mathbf{P}\{X_i \le u,\ X_p > u,\ 2 \le i \le p-1 | X_1 > u\}$, $p \ge k+2$, and using (2.8), (3.6) can be rewritten as

$$\begin{aligned}\theta &= \lim_{u\uparrow x_{F_1}} \mathbf{P}\{X_i \le u, 2 \le i \le k+1 | X_1 > u\} - \lim_{p\to\infty}\lim_{u\uparrow x_{F_1}} \sum_{j=k+2}^{p} \theta_j^{(u)}\\ &= \log(G_k(\mathbf{0})/G_{k+1}(\mathbf{0})) - \sum_{p=k+2}^{\infty} \lim_{u\uparrow x_{F_1}} \theta_p^{(u)},\end{aligned}$$

and thus it is enough to show that, for each $p \ge k+2$,

$$\lim_{u\uparrow x_{F_1}} \theta_p^{(u)} = \mathbf{P}\{Z_i \le -T,\ Z_{p-1} > -T,\ 1 \le i \le p-2\}.$$

For $M > 1$, let $c_\xi^{(u)}(M) := u - g(u)(1-M^{-\xi})/\xi$, $d_\xi^{(u)}(M) := u + g(u)(M^\xi - 1)/\xi$, and, for fixed $p \ge k+2$, define the events

$$\begin{aligned}A(u,M) &= \Big\{u < X_1 \le d_\xi^{(u)}(M),\ X_i \le u,\ u < X_p \le d_\xi^{(u)}(M),\ 2 \le i \le p-1\Big\},\\ B(u,M) &= \Big\{u < X_1 \le d_\xi^{(u)}(M),\ c_\xi^{(u)}(M) < X_i \le u,\ u < X_p \le d_\xi^{(u)}(M),\\ &\qquad\qquad 2 \le i \le p-1\Big\},\\ C_j(u,M) &= \Big\{u < X_1 \le d_\xi^{(u)}(M),\ c_\xi^{(u)}(M) < X_i \le u,\ X_j \le c_\xi^{(u)}(M),\\ &\qquad X_m \le u,\ u < X_p \le d_\xi^{(u)}(M),\ 2 \le i \le j-1,\ j+1 \le m \le p-1\Big\},\\ &\qquad\qquad j = 2, \dots, p-1.\end{aligned}$$

Then, since $\lim_{M\to\infty}\overline{\lim}_{u\uparrow x_{F_1}} \mathbf{P}\{X_s > d_\xi^{(u)}(M) | X_1 > u\} = 0$ from (2.2) for $s \ge 1$, we have $\lim_{u\uparrow x_{F_1}} \theta_p^{(u)} = \lim_{M\to\infty}\lim_{u\uparrow x_{F_1}} \mathbf{P}\{A(u,M)|X_1 > u\}$. Also note that

$$\begin{aligned}\mathbf{P}\{A(u,M)|X_1 > u\} &= \mathbf{P}\{B(u,M)|X_1 > u\}\\ &\qquad + \sum_{j=2}^{p-1} \mathbf{P}\{C_j(u,M)|X_1 > u\}.\end{aligned} \tag{3.10}$$

We now show that, for $j = 2, \dots, p-1$,

$$\lim_{M\to\infty} \overline{\lim_{u\uparrow x_{F_1}}} \mathbf{P}\Big\{u < X_1 \le d_\xi^{(u)}(M),\ X_j \le c_\xi^{(u)}(M),\ X_p > u | X_1 > u\Big\} = 0, \tag{3.11}$$

and thus $\lim_{M\to\infty} \overline{\lim}_{u\uparrow x_{F_1}} \mathbf{P}\{C_j(u, M)|X_1 > u\} = 0$. For this, let r be the integer part of $(p - j - 1)/k$ and consider

$$\begin{aligned}
&\mathbf{P}\Big\{X_j \le c_\xi^{(u)}(M),\ X_p > u|X_1 = u + g(u)x_1\Big\}\\
&\quad\le \sum_{s=0}^{r-1} \mathbf{P}\Big\{X_{j+sk} \le c_\xi^{(u)}(M^{(r+1-s)/(r+1)}),\\
&\qquad\qquad X_{j+(s+1)k} > c_\xi^{(u)}(M^{(r-s)/(r+1)})|X_1 = u + g(u)x_1\Big\}\\
&\qquad + \mathbf{P}\Big\{X_{j+rk} \le c_\xi^{(u)}(M^{1/(r+1)}), X_p > u|X_1 = u + g(u)x_1\Big\}\\
&\quad\le \sum_{s=0}^{r} \sup\Big\{\mathbf{P}\{X_{k+1} > c_\xi^{(u)}(M^{(r-s)/(r+1)})|\mathbf{X}_k = \mathbf{y}_k\}:\\
&\qquad\qquad \min_{1\le i\le k} y_i \le c_\xi^{(u)}(M^{(r+1-s)/(r+1)})\Big\},
\end{aligned}$$

which is because $\{X_n\}$ is a kth-order stationary Markov chain. Thus, using the properties (2.5) of g, it can be seen that (3.8) implies

$$\lim_{M\to\infty} \overline{\lim_{u\uparrow x_{F_1}}} \mathbf{P}\Big\{X_j \le c_\xi^{(u)}(M),\ X_p > u|X_1 = u + g(u)x_1\Big\} = 0 \quad \text{uniformly in } x_1.$$

This, together with the first statement of Assumption A, shows that (3.11) holds. We therefore have $\lim_{u\uparrow x_{F_1}} \theta_p^{(u)} = \lim_{M\to\infty} \lim_{u\uparrow x_{F_1}} \mathbf{P}\{B(u, M)|X_1 > u\}$ from (3.10) and (3.11). Here, using the same notation as in the proof of Lemma 2.2, $\mathbf{P}\{B(u, M)|X_1 > u\}$ can be rewritten as

$$\int_0^{(M^\xi-1)/\xi} \int_{(M^{-\xi}-1)/\xi}^0 \cdots \int_{(M^{-\xi}-1)/\xi}^0 \int_0^{(M^\xi-1)/\xi} \times f_1^{(u)}(x_1)\psi_p^{(u)}(\mathbf{x}_p)\, dx_p dx_{p-1} \dots dx_2\, dx_1,$$

which converges to

$$\int_0^{(M^\xi-1)/\xi} \int_{(M^{-\xi}-1)/\xi}^0 \cdots \int_{(M^{-\xi}-1)/\xi}^0 \int_0^{(M^\xi-1)/\xi} \times (1+\xi x_1)^{-1/\xi-1}\psi_{p-1}(\mathbf{x}_p)\, dx_p\, dx_{p-1} \dots dx_2\, dx_1$$

as $u \uparrow x_{F_1}$ by applying Assumptions A and B successively. Finally, since $(M^{\xi} - 1)/\xi \uparrow x_{\Omega_\xi}$ and $(M^{-\xi} - 1)/\xi \downarrow x^*_{\Omega_\xi}$ as $M \to \infty$, we have, by letting $M \to \infty$,

$$\lim_{u \uparrow x_{F_1}} \theta_p^{(u)} = \int_0^{x_{\Omega_\xi}} \int_{x^*_{\Omega_\xi}}^0 \cdots \int_{x^*_{\Omega_\xi}}^0 \int_0^{x_{\Omega_\xi}} \times (1 + \xi x_1)^{-1/\xi - 1} \psi_{p-1}(\mathbf{x}_p)\, dx_p\, dx_{p-1} \ldots dx_2\, dx_1,$$

which, if we put $x_1 = \phi(t)$ and $x_j = \phi(t + y_1 + \cdots + y_{j-1})$, $j = 2, \ldots, p$, is equal to

$$\begin{aligned} &\int_0^{\infty} \int_{-\infty}^{-t} \int_{-\infty}^{-t-y_1} \cdots \int_{-\infty}^{-t-\sum_{j=1}^{p-3} y_j} \int_{-t-\sum_{j=1}^{p-2} y_j}^{\infty} \\ &\qquad \times e^{-t} \eta_{p-1}(\mathbf{y}_{p-1})\, dy_{p-1}\, dy_{p-2} \ldots dy_2\, dy_1\, dt \\ &\quad = \mathbf{P}\Bigg\{ -\infty < Y_1 \le -T, \ldots, -\infty < Y_{p-2} \le -T - \sum_{j=1}^{p-3} Y_j, \\ &\qquad\qquad Y_{p-1} > -T - \sum_{j=1}^{p-2} Y_j \Bigg\} \\ &\quad = \mathbf{P}\{Z_1 \le -T, \ldots, Z_{p-2} \le -T,\ Z_{p-1} > -T\}, \end{aligned}$$

where $\{Y_n\}$ is the chain in (2.10). This completes the proof. □

REMARK 3.2.

1. When $k = 1$, representation (3.9) coincides with (3.7) since

$$\begin{aligned} \log(G_1(0)/G_2(\mathbf{0})) &= \lim_{u \uparrow x_{F_1}} \mathbf{P}\{X_2 \le u | X_1 > u\} \\ &= 1 - \lim_{u \uparrow x_{F_1}} \mathbf{P}\{X_2 > u | X_1 > u\} \\ &= 1 - \mathbf{P}\{Z_1 > -T\} = \mathbf{P}\{Z_1 \le -T\}. \end{aligned}$$

   Therefore, (3.9) can be viewed as a direct extension of Smith (1992a) [see also Perfekt (1994)]. However, when $k \ge 2$, (3.9) may not coincide with (3.7) since, for $i \ge 3$, $\lim_{u \uparrow x_{F_1}} \mathbf{P}\{X_i > u | X_1 > u\}$ need not be equal to $\mathbf{P}\{Z_{i-1} > -T\}$ without further assumptions (see Corollary 3.3 below).
2. Recently, Perfekt (1997) extended the results of Perfekt (1994) for a Markov chain to a higher-order case under a weaker condition than (3.8), but with a stronger assumption on the limiting transition kernel. Translated to our setting, he let the sup in (3.8) be taken over $\max_{1 \le i \le k} x_i \le u - g(u)(1 - M^{-\xi})/\xi$ and assumed the existence of a limiting transition kernel so that the resulting chain can visit and leave the state $-\infty$ unless there occurred k successive visits of that state before, which is not allowed in our framework. Recall that the chain $\{Z_n\}$ defined in (2.10) cannot leave the state $-\infty$ once it is visited. In view of the increase of strength of dependence of variables as k increases, we believe condition (3.8) is not so strong in practice.

Sometimes the G_{k+1} in Theorem 3.1 may be infeasible to find out explicitly or the quantity $\log(G_k(\mathbf{0})/G_{k+1}(\mathbf{0}))$ in (3.9) may be intractable to compute. The following corollary strengthens condition (3.8) to remove these possible difficulties.

COROLLARY 3.3. *For $k \geq 2$ in Theorem 3.1, drop the assumptions $F_{k+1} \in \mathscr{D}(G_{k+1})$ and (3.8), and suppose instead that $F_1 \in \mathscr{D}(\Omega_\xi)$ for some $\xi \in \Re$, with auxiliary function g satisfying (2.3) and (2.5), and that, for each $j = 2, \dots, k$,*

$$\lim_{M\to\infty} \varlimsup_{u\uparrow x_{F_1}} \sup\Big\{\mathbf{P}\{X_{j+1} > u | \mathbf{X}_j = \mathbf{x}_j\}: \\ \min_{1\leq i\leq j} x_i \leq u - g(u)(1 - M^{-\xi})/\xi\Big\} = 0. \tag{3.12}$$

Then, $\{X_n\}$ has extremal index θ given by (3.7).

PROOF. Letting $\theta_p^{(u)} = \mathbf{P}\{X_i \leq u,\ X_p > u,\ 2 \leq i \leq p-1 | X_1 > u\}$, $p \geq 3$, (3.6) is rewritten as

$$\theta = 1 - \lim_{u\uparrow x_{F_1}} \mathbf{P}\{X_2 > u | X_1 > u\} - \sum_{p=3}^{\infty} \lim_{u\uparrow x_{F_1}} \theta_p^{(u)}.$$

By the same argument as in the proof of Theorem 3.1, it can be shown from (3.12) that, for each $p \geq 3$, $\lim_{u\uparrow x_{F_1}} \theta_p^{(u)} = \mathbf{P}\{Z_i \leq -T, Z_{p-1} > -T, 1 \leq i \leq p-2\}$. Since $\lim_{u\uparrow x_{F_1}} \mathbf{P}\{X_2 > u | X_1 > u\} = \mathbf{P}\{Z_1 > -T\}$, the proof is complete. □

4. Exceedance point processes. In this section, we establish convergence in distribution of multilevel exceedance point processes for a kth-order stationary Markov chain and give representations for quantities characterizing the distributional limits in terms of the extremal index of the original chain and the chain defined by (2.10).

Let $\{X_n: n \geq 1\}$ be a stationary sequence of r.v.'s with marginal d.f. F_1. Assume that, for each $\tau > 0$, there exists a sequence $\{u_n(\tau): n \geq 1\}$ of real numbers satisfying (3.2). For each $\tau > 0$, let $N_n^{(\tau)}$ be the point process on $(0, \infty)$ defined by

$$N_n^{(\tau)}(B) := \sum_{i=1}^{\infty} \delta_{i/n}(B) 1_{\{X_i > u_n(\tau)\}} \tag{4.1}$$

for Borel sets $B \subset (0, \infty)$, where $\delta_{i/n}(\cdot)$ denotes the Dirac measure with mass 1 at i/n. Thus, $N_n^{(\tau)}$ is the point process of time-normalized exceedances of the level $u_n(\tau)$ by the r.v.'s $X_1, X_2, \dots$. The $N_n^{(\tau)}$ gives a more complete description of the extremal behavior of $\{X_n\}$ than M_n since, for instance, $\{M_n \leq u_n(\tau)\} = \{N_n^{(\tau)}((0, 1]) = 0\}$.

For a fixed $r \geq 1$, let $\tau_1 > \cdots > \tau_r > 0$ and write $\mathbf{N}_n := (N_n^{(\tau_1)}, \dots, N_n^{(\tau_r)})$, a multilevel exceedance point process for $\{X_n\}$. Then, by (3.2), $u_n(\tau_1) < \cdots <$

$u_n(\tau_r)$ for all sufficiently large n. For convergence in distribution of $\mathbf{N}_n$ as $n \to \infty$, we need a slightly stronger mixing condition than Leadbetter's one. Specifically, for sequences $\{u_n^{(j)}: n \geq 1\}$, $j = 1, \ldots, r$, of real numbers, we write: for $1 \leq s \leq t \leq n$, define $\mathscr{B}_{s,t}^{(n)}$ to be the σ-field generated by the events $\{X_i \leq u_n^{(j)}\}$, $s \leq i \leq t$, $1 \leq j \leq r$, and, for $q = 1, \ldots, n-1$, write

$$\begin{aligned}\beta_{n,q} = \max\Big\{&|\mathbf{P}(A \cap B) - \mathbf{P}(A)\mathbf{P}(B)|: \\ &A \in \mathscr{B}_{1,s}^{(n)},\ B \in \mathscr{B}_{s+q,n}^{(n)},\ 1 \leq s \leq n-q\Big\}.\end{aligned} \tag{4.2}$$

If $\beta_{n,q_n} \to 0$ as $n \to \infty$ for some sequence $q_n = o(n)$, then the condition $\triangle(u_n^{(1)}, \ldots, u_n^{(r)})$ is said to hold for $\{X_n\}$. This mixing condition was introduced by Hsing (1987) and is weaker than the strong mixing.

Now, assume that the sequence $\{X_n\}$ has extremal index $\theta > 0$ and that, for each $\sigma > 0$, $\triangle(u_n(\sigma\tau_1), \ldots, u_n(\sigma\tau_r))$ holds for $\{X_n\}$. We use the notation: let $\mathscr{I} := \{\mathbf{i} \in \{0, 1, \ldots\}^r: i_1 \geq \cdots \geq i_r,\ i_1 \geq 1\}$, and for each $\mathbf{i} \in \mathscr{I}$, write $\mathbf{i}_{(j)} := (i_1 - 1, \ldots, i_j - 1, i_{j+1}, \ldots, i_r)$, $j = 0, 1, \ldots, r$ (i.e., $\mathbf{i}_{(0)} = \mathbf{i}$); let $\mathbf{P}_n^{(j)}\{\cdot\} := \mathbf{P}\{\cdot | X_1 > u_n(\tau_j)\}$, $j = 1, \ldots, r$, and we mean $(N_n^{(\tau_1)}(B), \ldots, N_n^{(\tau_r)}(B))$ by $\mathbf{N}_n(B)$ for Borel sets $B \subset (0, \infty)$. Then, by Perfekt (1994) [see also Hsing (1984) and Rootzén (1988)], $\mathbf{N}_n = (N_n^{(\tau_1)}, \ldots, N_n^{(\tau_r)})$ converges in distribution to some point process $(N^{(\tau_1)}, \ldots, N^{(\tau_r)})$ as $n \to \infty$ if and only if, for each $\mathbf{i} \in \mathscr{I}$, there exists a constant $\pi(\mathbf{i})$ such that

$$\begin{aligned}\sum_{j=1}^{r} \frac{\tau_j}{\tau_1} \Big[&\mathbf{P}_n^{(j)}\{\mathbf{N}_n((1/n, p_n/n]) = \mathbf{i}_{(j)}\} \\ &- \mathbf{P}_n^{(j)}\{\mathbf{N}_n((1/n, p_n/n]) = \mathbf{i}_{(j-1)}\}\Big] \to \theta\pi(\mathbf{i})\end{aligned} \tag{4.3}$$

as $n \to \infty$, where $\{p_n: n \geq 1\}$ is a sequence of positive integers such that

$$p_n = o(n), \qquad n\beta_{n,q_n} = o(p_n), \qquad q_n = o(p_n), \tag{4.4}$$

the $\beta_{n,q}$ being given in (4.2) with $u_n^{(j)} = u_n(\tau_j)$, $1 \leq j \leq r$, and the $q_n = o(n)$ being such that $\beta_{n,q_n} \to 0$. In this case, π is a probability measure on $\mathscr{I}$ and, moreover, for each $\sigma > 0$, $(N_n^{(\sigma\tau_1)}, \ldots, N_n^{(\sigma\tau_r)})$ converges in distribution to a point process $(N^{(\sigma\tau_1)}, \ldots, N^{(\sigma\tau_r)})$ with Laplace transform

$$\begin{aligned}&\mathbf{E}\left[\exp\left(-\sum_{j=1}^{r} \int_0^\infty g_j\, dN^{(\sigma\tau_j)}\right)\right] \\ &\quad = \exp\left[-\theta\sigma\tau_1 \int_0^\infty (1 - L(g_1(x), \ldots, g_r(x)))\, dx\right], \qquad g_j \geq 0,\end{aligned} \tag{4.5}$$

where L is the Laplace transform of π; thus, each marginal $N^{(\sigma\tau_j)}$ is a compound Poisson process with intensity $\theta\sigma\tau_j$.

Returning to our proposed model, let $\{X_n: n \geq 1\}$ be a kth-order stationary Markov chain. Then (4.3) can be characterized in terms of the chain defined by (2.10). Specifically, for each $\tau > 0$, let $S^{(\tau)}$ be the point process on $(0, \infty)$ defined by

$$S^{(\tau)}(B) := \sum_{i=1}^{\infty} \delta_i(B) 1_{\{Z_i + T > -\log \tau\}} \tag{4.6}$$

for Borel sets $B \subset (0, \infty)$, where $\{Z_n: n \geq 1\}$ is the kth-order Markov chain defined by (2.10) and T an Exp(1)-distributed r.v. which is independent of $\{Z_n\}$. Finally, for each $j = 1, \dots, r$, writing

$$\mathbf{S}^{(j)} := (S^{(\tau_1/\tau_j)}, \dots, S^{(\tau_r/\tau_j)}), \tag{4.7}$$

we have the following.

THEOREM 4.1. *Let $\{X_n: n \geq 1\}$ be a kth-order stationary Markov chain, and let F_{k+1} be the d.f. of $(X_1, \dots, X_{k+1})$ having a continuous p.d.f. f_{k+1}. Suppose that $F_1 \in \mathscr{D}(\Omega_\xi)$ for some $\xi \in \Re$, with auxiliary function g satisfying* (2.3) *and* (2.5), *and that Assumptions A and B hold. Further, suppose that, when $k = 1$,* (3.12) *holds for $j = 1$; when $k \geq 2$,* (3.12) *holds for each $j = 2, \dots, k$. Let $\{u_n(\tau)\}$, $\tau > 0$, be sequences satisfying* (3.2), *and let $N_n^{(\tau)}$, $\tau > 0$, be the point processes defined by* (4.1). *Assume that $\{X_n\}$ has extremal index $\theta > 0$ and that, for some fixed $\tau_1 > \cdots > \tau_r > 0$, $\triangle(u_n(\sigma\tau_1), \dots, u_n(\sigma\tau_r))$ holds for each $\sigma > 0$. Assume further that* (3.5) *holds with $u_n = u_n(\tau_1)$, where $\{p_n\}$ is a sequence satisfying* (4.4). *Then, for each $\sigma > 0$, $(N_n^{(\sigma\tau_1)}, \dots, N_n^{(\sigma\tau_r)})$ converges in distribution to the point process $(N^{(\sigma\tau_1)}, \dots, N^{(\sigma\tau_r)})$ characterized by* (4.5), *with π given by*

$$\pi(\mathbf{i}) = \theta^{-1} \sum_{j=1}^{r} \frac{\tau_j}{\tau_1} \big[\mathbf{P}\{\mathbf{S}^{(j)}((0, \infty)) = \mathbf{i}_{(j)}\} - \mathbf{P}\big\{\mathbf{S}^{(j)}((0, \infty)) = \mathbf{i}_{(j-1)}\big\}\big], \qquad \mathbf{i} \in \mathscr{I},$$

where $\mathbf{S}^{(j)}$, $j = 1, \dots, r$, are the point processes defined by (4.7).

PROOF. Using (3.5), it follows from (4.3) that it is enough to show that, for each $j = 1, \dots, r$,

$$\lim_{p\to\infty} \lim_{n\to\infty} \mathbf{P}_n^{(j)}\{\mathbf{N}_n((1/n, p/n]) = \mathbf{i}\} = \mathbf{P}\{\mathbf{S}^{(j)}((0, \infty)) = \mathbf{i}\}$$

for any $\mathbf{i} = (i_1, \dots, i_r)$ with $i_1 \geq \cdots \geq i_r \geq 0$. Again, this follows if we show that, for each $p \geq 2$,

$$\mathbf{P}_n^{(j)}\{\mathbf{N}_n((1/n, p/n]) = \mathbf{i}\} \to \mathbf{P}\{\mathbf{S}^{(j)}((0, p-1]) = \mathbf{i}\} \quad \text{as } n \to \infty.$$

Here, the left-hand side can be expressed as a linear combination of 1 and probabilities of the form

$$\mathbf{P}_n^{(j)}\{X_{m_i} \leq u_n(\tau_{v_i}),\ X_{m_s} > u_n(\tau_{v_s}),\ i = 1, \dots, s-1\}$$

for $s = 1, \dots, p$, $2 \le m_1 < \cdots < m_s \le p$ and $v_1, \dots, v_s \in \{1, \dots, r\}$; thus, writing $u_n^{(v)} = u_n(\tau_v)$, $v = 1, \dots, r$, and $A(n) = \{X_{m_i} \le u_n^{(v_i)},\ X_{m_s} > u_n^{(v_s)},\ 1 \le i \le s-1\}$, it suffices to show that

$$\begin{aligned} \mathbf{P}_n^{(j)}(A(n)) \to \mathbf{P}\big\{ & Z_{m_i - 1} + T \le -\log(\tau_{v_i}/\tau_j), \\ & Z_{m_s - 1} + T > -\log(\tau_{v_s}/\tau_j),\ 1 \le i \le s-1 \big\} \end{aligned} \tag{4.8}$$

as $n \to \infty$, where $\{Z_n\}$ and T are the same as in (4.6). To show this, for $M > 1$ and $v = 1, \dots, r$, let $c_\xi^{(n)}(M, v) := u_n^{(v)} - g(u_n^{(j)})(\tau_v/\tau_j)^{-\xi}(1 - M^{-\xi})/\xi$ and $d_\xi^{(n)}(M, v) := u_n^{(v)} + g(u_n^{(j)})(\tau_v/\tau_j)^{-\xi}(M^{\xi} - 1)/\xi$. Then, from (2.2), (2.5) and (3.2), we have

$$\begin{aligned} (u_n^{(v)} - u_n^{(j)})/g(u_n^{(j)}) &\to ((\tau_v/\tau_j)^{-\xi} - 1)/\xi, \\ (c_\xi^{(n)}(M, v) - u_n^{(j)})/g(u_n^{(j)}) &\to (M^{-\xi}(\tau_v/\tau_j)^{-\xi} - 1)/\xi, \\ (d_\xi^{(n)}(M, v) - u_n^{(j)})/g(u_n^{(j)}) &\to (M^{\xi}(\tau_v/\tau_j)^{-\xi} - 1)/\xi \end{aligned} \tag{4.9}$$

as $n \to \infty$ for $v = 1, \dots, r$. If $m_s = 2$, then (4.8) can be shown similarly as below without using condition (3.12) and so the details are omitted. We now assume that $m_s \ge 3$, and define the events

$$\begin{aligned} B(n, M) &= \big\{X_l > c_\xi^{(n)}(M, v_s),\ 2 \le l \le m_s - 1\big\}, \\ C_t(n, M) &= \big\{X_l > c_\xi^{(n)}(M, v_s),\ X_t \le c_\xi^{(n)}(M, v_s),\ 2 \le l \le t-1\big\}, \\ & \qquad t = 2, \dots, m_s - 1. \end{aligned}$$

Then, for any $M > 1$, we have

$$\mathbf{P}_n^{(j)}(A(n)) = \mathbf{P}_n^{(j)}(A(n) \cap B(n, M)) + \sum_{t=2}^{m_s - 1} \mathbf{P}_n^{(j)}(A(n) \cap C_t(n, M)).$$

Here, using (2.5), (3.12) and (4.9), it can be shown by a similar argument as in showing (3.11) that, for $t = 2, \dots, m_s - 1$,

$$\lim_{M\to\infty} \varlimsup_{n\to\infty} \mathbf{P}_n^{(j)}\Big\{u_n^{(j)} < X_1 \le d_\xi^{(n)}(M, j),\ X_t \le c_\xi^{(n)}(M, v_s),\ X_{m_s} > u_n^{(v_s)}\Big\} = 0,$$

and thus $\lim_{M\to\infty} \varlimsup_{n\to\infty} \mathbf{P}_n^{(j)}(A(n) \cap C_t(n, M)) = 0$, since

$$\lim_{M\to\infty} \varlimsup_{n\to\infty} \mathbf{P}_n^{(j)}\{X_h > d_\xi^{(n)}(M, v)\} = 0$$

from (2.2) for $h \ge 1$ and $v = 1, \dots, r$. Therefore, if we further define the event

$$\begin{aligned} D(n, M) = \Big\{ & u_n^{(j)} < X_1 \le d_\xi^{(n)}(M, j),\ c_\xi^{(n)}(M, v_s) < X_l \le d_\xi^{(n)}(M, j), \\ & c_\xi^{(n)}(M, v_s) < X_{m_i} \le u_n^{(v_i)},\ u_n^{(v_s)} < X_{m_s} \le d_\xi^{(n)}(M, v_s), \\ & \qquad 2 \le l \le m_s - 1,\ 1 \le i \le s-1 \Big\}, \end{aligned}$$

then we have

$$\begin{aligned}\lim_{n\to\infty} \mathbf{P}_n^{(j)}(A(n)) &= \lim_{M\to\infty}\lim_{n\to\infty} \mathbf{P}_n^{(j)}(A(n)\cap B(n,M))\\ &= \lim_{M\to\infty}\lim_{n\to\infty} \mathbf{P}_n^{(j)}(D(n,M)).\end{aligned}$$

Finally, using the fact that $(M^{\xi}(\tau_v/\tau_j)^{-\xi}-1)/\xi \uparrow x_{\Omega_\xi}$ and $(M^{-\xi}(\tau_v/\tau_j)^{-\xi}-1)/\xi \downarrow x^*_{\Omega_\xi}$ as $M\to\infty$ for $v=1,\dots,r$, it can be shown by applying Assumptions A and B successively that

$$\begin{aligned}\lim_{M\to\infty}\lim_{n\to\infty} \mathbf{P}_n^{(j)}(D(n,M)) = \mathbf{P}\big\{&Z_{m_i-1}+T\le -\log(\tau_{v_i}/\tau_j),\\ &Z_{m_s-1}+T > -\log(\tau_{v_s}/\tau_j),\ 1\le i\le s-1\big\}.\end{aligned}$$

The details are similar to those in the proof of Theorem 3.1 and so are omitted. This completes the proof. □

REMARK 4.2. In Theorem 4.1, if (2.9) holds for each $j=1,\dots,k$ as in Lemma 2.2, then Assumption B and condition (3.12) are, of course, unnecessary.

5. Examples. Let $\{X_n\colon n\ge 1\}$ denote a kth-order stationary Markov chain, and let F_{k+1}, f_{k+1} denote the d.f., p.d.f. of $(X_1,\dots,X_{k+1})$, respectively. In this section, we consider four examples of F_{k+1} to compute the extremal index of $\{X_n\}$. Some of these examples were practically used in modeling multivariate extremes in environmental data [see Coles and Tawn (1991) and Smith, Tawn and Coles (1997)].

Since we focus on only the absolutely continuous case for F_{k+1}, there always exist sequences $\{u_n(\tau)\}$, $\tau>0$, satisfying (3.2). On the other hand, it is not easy to show directly from the functional form of F_{k+1} (or f_{k+1}) that $D(u_n(\tau))$ holds for each $\tau>0$. We follow instead a general method used by O'Brien (1987), Rootzén (1988) and Smith (1992a), which involves the concept of Harris chain. For a detailed consideration of the Harris chain, the reader is referred to Nummelin (1984) and Asmussen (1987). We begin with a Markov chain $\{J_n\colon n\ge 1\}$ on $(E,\mathscr{E})$, where E is a general state space and $\mathscr{E}$ a σ-field in E. We call $R\in\mathscr{E}$ recurrent if $\mathbf{P}\{J_n\in R \text{ infinitely often} | J_1=x\}=1$ for every $x\in E$, and further, we call $R\in\mathscr{E}$ a regeneration set if R is recurrent and if, for some $n_0\ge 2$, the $\mathbf{P}\{J_{n_0}\in\cdot|J_1=x\}$, $x\in R$, contain a common component, that is, there exist $0<\varepsilon<1$ and some probability measure λ on $(E,\mathscr{E})$ such that $\mathbf{P}\{J_{n_0}\in A|J_1=x\}\ge\varepsilon\lambda(A)$, $A\in\mathscr{E}$, $x\in R$. A Markov chain with a regeneration set is called a Harris chain, and if in addition the regeneration set has a finite mean recurrence time, then it is called a positive recurrent Harris chain. An aperiodic positive recurrent Harris chain is also called a Harris ergodic chain. According to O'Brien (1987), it is known that, if $f\colon E\to\Re$ is a measurable function and $\{J_n\}$ a stationary Harris ergodic chain, then the stationary sequence $\{f(J_n)\}$ is strongly mixing.

Returning to the kth-order stationary Markov chain $\{X_n\}$, we take $J_n = (X_n, X_{n+1}, \dots, X_{n+k-1})$, $n = 1, 2, \dots,$ so that $\{J_n\}$ is a stationary Markov chain on $(E, \mathscr{E})$, where $E = \Re^k$ [or $(0,\infty)^k$] and $\mathscr{E} = \mathscr{B}(E)$, the Borel σ-field in E. Then, since f_{k+1} in our examples is positive, continuous and bounded on E and since, for any $B \in \mathscr{E}$,

$$\begin{aligned}\mathbf{P}\{J_{k+1} \in B | J_1 = (x_1, \dots, x_k)\} &\\ = \int_B f_{k+1}(y_1|x_1, \dots, x_k) f_{k+1}(y_2|x_2, \dots, x_k, y_1)\cdots&\\ \times f_{k+1}(y_k|x_k, y_1, \dots, y_{k-1})\, d\mathbf{y}_k,&\end{aligned}$$

it can be seen that $\{J_n\}$ is a stationary Harris chain, using the same type of arguments as in Example 3.1 of Asmussen (1987). Moreover, Smith (1992b) showed that $\{J_n\}$ for $k = 1$ in Example 5.2 below is, in fact, geometrically ergodic. Similar arguments can be applied to $\{J_n\}$ for $k \geq 2$ and $\{J_n\}$ in other examples below to show that the corresponding $\{J_n\}$ is geometrically ergodic. Geometric ergodicity is a stronger concept than ergodicity. Therefore, choosing the first coordinate function for $f\colon E \to \Re$, that is, $f(x_1, \dots, x_k) = x_1$, makes $\{X_n\} = \{f(J_n)\}$ strongly mixing.

EXAMPLE 5.1 (Multivariate gamma distribution). Let $\{X_n\}$ be a kth-order stationary Markov chain in which F_{k+1} is a multivariate gamma distribution with parameter $\alpha > 0$, that is, F_{k+1} is defined by its density f_{k+1} as [see Johnson and Kotz (1972), page 217]:

$$f_{k+1}(\mathbf{x}_{k+1}) = \frac{1}{\Gamma(\alpha)}\left\{\exp\left(-\sum_{s=1}^{k+1} x_s\right)\right\}\int_0^{\tilde{x}_{k+1}} t^{\alpha-1}e^{kt}\,dt, \qquad \mathbf{x}_{k+1} > \mathbf{0}, \tag{5.1}$$

where $\tilde{x}_{k+1} := \min\{x_1, \dots, x_{k+1}\}$. We are going to apply Theorem 3.1 to find the extremal index θ of $\{X_n\}$. It is readily checked that $F_{k+1} \in \mathscr{D}(G_{k+1})$ with auxiliary function $g(u) = 1$, where $G_{k+1}(\mathbf{x}_{k+1}) = \exp(-\sum_{i=1}^{k+1}\exp(-x_i))$, $\mathbf{x}_{k+1} \in \Re^{k+1}$. Obviously, $G_1 = \Omega_0$. Since every lower-dimensional marginal density function $f_i(\mathbf{x}_i)$, $i = 1, \dots, k$, is of the same form as (5.1), it can be easily shown that, for each $j = 1, \dots, k$,

$$\begin{aligned}l_j(x_{j+1}; \mathbf{x}_j) &= \lim_{u\to\infty} f_{j+1}(u + x_{j+1}|u + \mathbf{x}_j)\\ &= (1 - 1/j)\exp(j\tilde{x}_{j+1} - x_{j+1} - (j-1)\tilde{x}_j), \qquad \mathbf{x}_{j+1} \in \Re^{j+1},\end{aligned}$$

which shows that Assumption A holds. Notice here that (2.9) holds for each $j = 2, \dots, k$, but that (2.9) fails for $j = 1$. Assumption B is elementary. For condition (3.8), observe that

$$f_{k+1}(x_{k+1}|\mathbf{x}_k) = \frac{\exp(-x_{k+1})\int_0^{\tilde{x}_{k+1}} t^{\alpha-1}e^{kt}\,dt}{\int_0^{\tilde{x}_k} t^{\alpha-1}e^{(k-1)t}\,dt} \leq \exp(\tilde{x}_{k+1} - x_{k+1}), \qquad \mathbf{x}_{k+1} > \mathbf{0}.$$

Thus, for $\tilde{x}_k \leq u - \log M$,

$$\mathbf{P}\{X_{k+1} > u|\mathbf{X}_k = \mathbf{x}_k\} \leq \int_u^\infty \exp(\tilde{x}_{k+1} - x_{k+1})\,dx_{k+1} = \exp(\tilde{x}_k - u) \leq \frac{1}{M},$$

from which (3.8) follows. Therefore, if we verify the basic condition (3.5), then (3.9) implies $\theta = \log(G_k(\mathbf{0})/G_{k+1}(\mathbf{0})) = 1$, since $l_1(x_2; x_1) = 0$ so that $H_1(-\infty) = 1$ and thus that $Z_n = -\infty$, $n \geq 1$.

To verify (3.5), observe first that, for each $j = 1, \dots, k$ and for any $\mathbf{y}_j \in \Re^j$,

$$\begin{aligned}
\lim_{y_0 \to \infty} & f_{j+1}(y_0 + y_1 + \cdots + y_j | y_0, y_0 + y_1, \dots, y_0 + y_1 + \cdots + y_{j-1}) \\
&= h_j(y_j; \mathbf{y}_{j-1}) \\
&= \begin{cases} (1 - 1/j)\exp[(j-1)(y_j + \gamma(\mathbf{y}_{j-1}))], & \text{if } y_j < -\gamma(\mathbf{y}_{j-1}), \\ (1 - 1/j)\exp[-y_j - \gamma(\mathbf{y}_{j-1})], & \text{if } y_j \geq -\gamma(\mathbf{y}_{j-1}), \end{cases}
\end{aligned}$$

where $\gamma(\mathbf{y}_{j-1}) = \max\{y_1 + \cdots + y_{j-1}, y_2 + \cdots + y_{j-1}, \dots, y_{j-1}, 0\}$. Here, notice that, for each $j = 2, \dots, k$, $\int_{-\infty}^{\infty} y_j h_j(y_j; \mathbf{y}_{j-1})\, dy_j = (j-2)/(j-1) - \gamma(\mathbf{y}_{j-1})$, which is obviously negative for large $y_1, \dots, y_{j-1}$, and that $h_1(y_1) = 0$. Therefore, there exist $0 < t < 1$, $y^* \in \Re$ such that, for each $j = 1, \dots, k$, $\int_{-\infty}^{\infty} \exp(t y_j) h_j(y_j; \mathbf{y}_{j-1})\, dy_j < 1$ whenever $y_1, \dots, y_{j-1} > y^*$. Using this, it can be seen that there exist $0 < \eta < 1$, $x^* > 0$ such that, for each $j = 1, \dots, k$,

$$\mathbf{E}\{\exp[t(X_{n+j} - X_{n+j-1})] | (X_n, X_{n+1}, \dots, X_{n+j-1}) = \mathbf{x}_j\} \leq \eta \tag{5.2}$$

whenever $x_1, \dots, x_j > x^*$. Let $\{u_n\}$ be a sequence satisfying (3.2) with $u_n = u_n(\tau)$ for some $\tau > 0$. Then inequalities (5.2) can be successively used to show that, for each $i \geq 2$,

$$\begin{aligned}
\mathbf{P}&\{X_2 > x^*, \dots, X_{i-1} > x^*,\ X_i > u_n | X_1 > u_n\} \\
&= \mathbf{P}\{X_2 > x^*, \dots, X_{i-1} > x^*, \exp[t(X_i - u_n)] > 1 | X_1 > u_n\} \\
&\leq \eta^{i-1} \exp(-t u_n) \frac{\int_{u_n}^{\infty} \exp(t x_1) f_1(x_1)\, dx_1}{1 - F_1(u_n)} \to \frac{\eta^{i-1}}{1-t} \quad \text{as } n \to \infty.
\end{aligned} \tag{5.3}$$

On the other hand, since $\sup\{\mathbf{P}\{X_i > u_n | \mathbf{X}_j = \mathbf{x}_j\}\colon \tilde{x}_j \leq x^*\} = O(n^{-1})$ for $i \geq k+1$ and $j = 1, \dots, k$, it follows that, for each $i \geq k+1$,

$$\begin{aligned}
\sum_{s=2}^{i-1} & \mathbf{P}\{X_2 > x^*, \dots, X_{s-1} > x^*,\ X_s \leq x^*,\ X_i > u_n | X_1 > u_n\} \\
&= O(n^{-1}) \sum_{s=2}^{i-1} \mathbf{P}\{X_2 > x^*, \dots, X_{s-1} > x^*,\ X_s \leq x^* | X_1 > u_n\} \\
&= O(n^{-1}).
\end{aligned} \tag{5.4}$$

From (5.3) and (5.4), it therefore follows that (3.5) holds for any $p_n = o(n)$.

EXAMPLE 5.2 (Logistic model). Let $\{X_n\}$ be a kth-order stationary Markov chain in which F_{k+1} follows the law of the logistic model with two parameters $\xi \in \Re$ and $r \geq 1$, that is,

$$F_{k+1}(\mathbf{x}_{k+1}) = \exp\left[-\left\{\sum_{s=1}^{k+1} (1 + \xi x_s)^{-r/\xi}\right\}^{1/r}\right], \qquad 1 + \xi \mathbf{x}_{k+1} > \mathbf{0}. \tag{5.5}$$

The bivariate case for this model was used by several statisticians [e.g., see Tawn (1988) and Smith (1992a)]. This model may be considered as a basic model for modeling multivariate extremes because of its simple structure, yet it gives independence of variables through complete dependence. The F_{k+1} is itself a multivariate extreme value distribution with $F_1 = \Omega_\xi$ and the lower-dimensional marginal d.f.'s $F_i(\mathbf{x}_i)$, $i = 1, \dots, k$, are also of form (5.5). We remove the case $r = 1$ which corresponds to the independence of $X_n, \dots, X_{n+k}$. Although this model has two parameters, the extremal index θ of $\{X_n\}$ is invariant under the choice of ξ. Therefore, we consider only the case $\xi = 0$ and, of course, $g(u) = 1$. Then, for each $i = 1, \dots, k+1$, since $\log F_i(\mathbf{x}_i) = -(\sum_{s=1}^{i} \exp(-rx_s))^{1/r}$, we apply Lemma 2.4 to have, for each $j = 1, \dots, k$,

$$l_j(x_{j+1}; \mathbf{x}_j) = \lim_{u\to\infty} f_{j+1}(u + x_{j+1} | u + \mathbf{x}_j)$$
$$= \frac{jr-1}{\sum_{s=1}^{j} \exp[r(x_{j+1} - x_s)]} \left\{ 1 + \frac{1}{\sum_{s=1}^{j} \exp[r(x_{j+1} - x_s)]} \right\}^{1/r-j-1}, \qquad \mathbf{x}_{j+1} \in \Re^{j+1}.$$

Notice here that (2.9) holds for each $j = 1, \dots, k$. Therefore, from Lemma 2.2, the extremal index θ of $\{X_n\}$ is given by (3.7), provided that the basic condition (3.5) is justified. For $k = 1$, (3.5) was verified by Smith (1992a). A similar argument as in Example 5.1 can be used here for general $k \geq 1$. Specifically, observe first that, for each $j = 1, \dots, k$,

$$\lim_{y_0\to\infty} F_{j+1}(y_0 + y_1 + \dots + y_j | y_0, y_0 + y_1, \dots, y_0 + y_1 + \dots + y_{j-1})$$
$$(5.6) \qquad = H_j(y_j; \mathbf{y}_{j-1})$$
$$= \left\{ 1 + \frac{\exp(-ry_j)}{1 + \sum_{s=1}^{j-1} \exp[r(y_s + \dots + y_{j-1})]} \right\}^{1/r-j}, \qquad \mathbf{y}_j \in \Re^j.$$

Here, notice that H_1 has a negative mean and that, for each $j = 2, \dots, k$, there exists a $\mathbf{y}_{j-1}^* = (y_1^*, \dots, y_{j-1}^*)$ of large values such that $H_j(\,\cdot\,; \mathbf{y}_{j-1}^*)$ has a negative mean. But, since $H_j(y_j; \mathbf{y}_{j-1}) \geq H_j(y_j; \mathbf{y}_{j-1}^*)$, $y_j \in \Re$, whenever $\mathbf{y}_{j-1} \geq \mathbf{y}_{j-1}^*$, $H_j(\,\cdot\,; \mathbf{y}_{j-1})$ has also a negative mean whenever $\mathbf{y}_{j-1} \geq \mathbf{y}_{j-1}^*$. Therefore, there exist $0 < t < 1$, $y^* \in \Re$ such that, for each $j = 1, \dots, k$, $\int_{-\infty}^{\infty} \exp(ty_j) H_j(dy_j; \mathbf{y}_{j-1}) < 1$ whenever $y_1, \dots, y_{j-1} > y^*$. This again implies that there exist $0 < \eta < 1$, $x^* \in \Re$ such that, for each $j = 1, \dots, k$, (5.2) holds whenever $x_1, \dots, x_j > x^*$. The remaining part of the proof of (3.5) is exactly the same as in Example 5.1.

Representation (3.7) can be effectively used to compute the extremal index θ of $\{X_n\}$ using simulation. The algorithm to generate $\{Z_n\}$ based on H_j, $j = 1, \dots, k$, with the inverse transform method is as follows.

ALGORITHM 1.

1. Generate independent, $\mathscr{U}(0, 1)$-distributed random numbers $U_1, U_2, \dots$.
2. $Z_0 \leftarrow 0$.

3. For $1 \le j \le k$, $Z_j \leftarrow -r^{-1}\{\log(U_j^{r/(1-rj)} - 1) + \log(\sum_{s=1}^{j} \exp(-rZ_{s-1}))\}$.
4. For $j \ge k+1$, $Z_j \leftarrow -r^{-1}\{\log(U_j^{r/(1-rk)} - 1) + \log(\sum_{s=j-k+1}^{j} \exp(-rZ_{s-1}))\}$.

For each fixed $k = 1, 2, 3, 4, 5, 10$, I computed

$$\theta_p = \mathbf{P}\{Z_1 \le -T, \dots, Z_p \le -T\} \tag{5.7}$$

based on 10^5 simulations of the process $\{Z_n\}$ and T, varying $p = 10, 20, 50, 100, 200, 500$ for every $1/r = 0.1, 0.2, \dots, 0.9$. It is noted that the worst-case standard error of each computed value is about 0.0016. According to the simulations, for lower k, θ_p becomes stable very quickly over all possible values of r. For higher k, the convergence of θ_p is somewhat slow, particularly for middle range of $1/r$. However, the simulation results suggest that $\theta \approx \theta_{500}$ can be used as a good approximation in practice unless $k > 10$. For $k > 10$, it is recommended that higher values of p be used. The final results for the extremal index θ based on θ_{500} are summarized in Table 1 and plotted in Figure 1 to give visualization of the overall trend. As expected intuitively, θ is decreasing as the order k of the chain $\{X_n\}$ grows, which is simply because the clusters of high-level exceedances tend to be widened as k increases.

EXAMPLE 5.3 (Mixture model). Let $\{X_n\}$ be a kth-order stationary Markov chain in which F_{k+1} is a multivariate extreme value distribution defined by

$$F_{k+1}(\mathbf{x}_{k+1}) = \exp\left[-\alpha\left\{\sum_{s=1}^{k+1}(1+\xi x_s)^{-r/\xi}\right\}^{1/r} - (1-\alpha)\sum_{s=1}^{k+1}(1+\xi x_s)^{-1/\xi}\right], \qquad 1 + \xi\mathbf{x}_{k+1} > \mathbf{0},$$

where $0 < \alpha < 1$, $r > 1$ and $\xi \in \Re$. We call this a mixture model since $-\log F_{k+1}(\mathbf{x}_{k+1})$ is given as a convex combination of those of the logistic model and the independence model. This is, in fact, a special case of the asymmetric logistic model developed by Tawn (1990) and considered here since it yields a

TABLE 1
Extremal index in logistic model

	k					
1/r	**1**	**2**	**3**	**4**	**5**	**10**
0.1	0.017	0.005	0.002	0.002	0.001	0.001
0.2	0.060	0.024	0.013	0.009	0.006	0.002
0.3	0.131	0.056	0.033	0.024	0.017	0.009
0.4	0.222	0.110	0.068	0.051	0.039	0.018
0.5	0.328	0.184	0.126	0.096	0.076	0.042
0.6	0.450	0.284	0.213	0.170	0.143	0.085
0.7	0.582	0.414	0.332	0.284	0.251	0.167
0.8	0.718	0.574	0.499	0.448	0.410	0.314
0.9	0.861	0.772	0.715	0.677	0.650	0.568

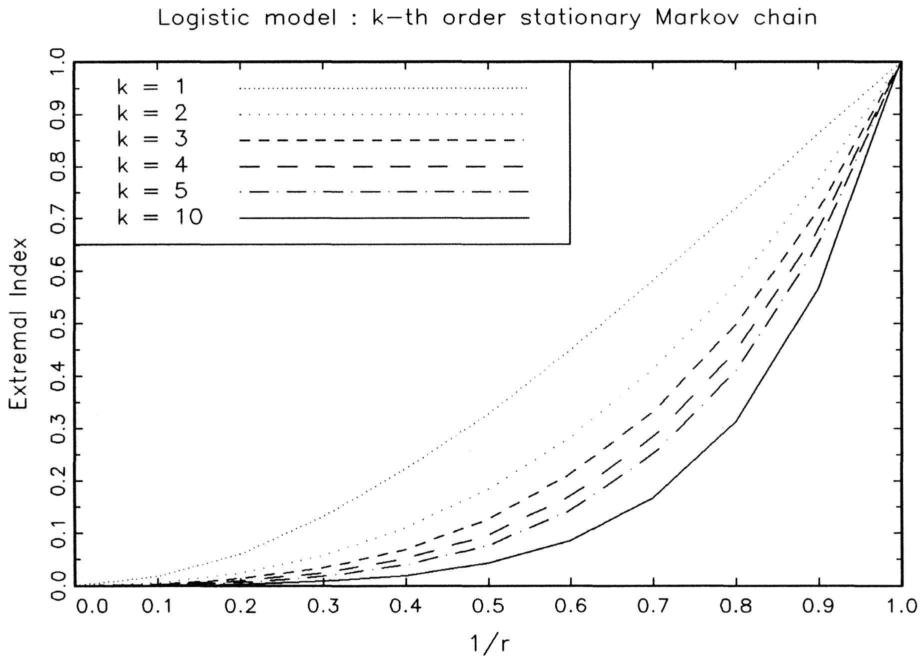

FIG. 1. *Plot of extremal index in logistic model.*

nontrivial extremal index θ where (2.9) does not hold for all $j = 1, \dots, k$. As before, θ is invariant under the choice of ξ so that we consider only the case $\xi = 0$ and, of course, $g(u) = 1$. First, using Lemma 2.4, it can be seen that $H_1(y_1) = 1 - \alpha + \alpha(1 + \exp(-ry_1))^{1/r-1}$, $y_1 \in \{-\infty\} \cup \Re$, and that, for each $j = 2, \dots, k$, $H_j(y_j; \mathbf{y}_{j-1})$ is given by (5.6). Here, since $H_1(-\infty) = 1 - \alpha > 0$, we apply Theorem 3.1 to compute θ. Verification of condition (3.8) is tedious and omitted here [see Yun (1994) for details]. For the basic condition (3.5), a similar method as in Example 5.2 can be used. Therefore, the extremal index θ of $\{X_n\}$ is given by (3.9) with $G_k(\mathbf{0})$, $G_{k+1}(\mathbf{0})$ replaced by $F_k(\mathbf{0})$, $F_{k+1}(\mathbf{0})$, respectively, that is, $\theta = \lim_{p\to\infty} \theta_p$, where

$$\theta_p = \alpha\{(k+1)^{1/r} - k^{1/r}\} + 1 - \alpha - \mathbf{P}\Big\{\max_{1\le i\le k} Z_i \le -T,\ \max_{k+1\le i\le p} Z_i > -T\Big\}.$$

The algorithm to generate $\{Z_n\}$ based on H_j, $j = 1, \dots, k$, is as follows.

ALGORITHM 2.

I. When $k = 1$:
  1. Generate independent, $\mathscr{U}(0, 1)$-distributed random numbers $U_1, U_2, \dots$.
  2. $Z_0 \leftarrow 0$.
  3. For $j \ge 1$:
     (i) If $U_{2j-1} > \alpha$ or $Z_{j-1} = -\infty$, then $Z_j \leftarrow -\infty$.
     (ii) If $U_{2j-1} \le \alpha$ and $Z_{j-1} > -\infty$, then $Z_j \leftarrow Z_{j-1} - r^{-1}\log(U_{2j}^{r/(1-r)} - 1)$.

II. When $k \geq 2$:
1. Generate independent, $\mathscr{U}(0,1)$-distributed random numbers $U_0, U_1, U_2, \dots$.
2. $Z_0 \leftarrow 0$.
3. If $U_0 > \alpha$, then $Z_1 = Z_2 = \cdots \leftarrow -\infty$.
4. If $U_0 \leq \alpha$, then use steps 3 and 4 of Algorithm 1.

Since this model still contains two parameters $0 < \alpha < 1$ and $r > 1$, my simulation study was limited to the case where $\alpha = 1/r$. For each fixed $k = 1, 2, 3, 4, 5, 10$, I calculated θ_p based on 10^5 simulations of the process $\{Z_n\}$ and T, varying $p = 10, 20, 50, 100, 200, 500$ for every $\alpha = 1/r = 0.1, 0.2, \dots, 0.9$. According to these simulations, the convergence rate of θ_p turns out to be very similar to that of θ_p in the logistic model. Therefore, $\theta \approx \theta_{500}$ can be used as a good approximation unless $k > 10$. The final results for the extremal index θ based on θ_{500} are plotted in Figure 2. One thing interesting is that θ is smoothly decreasing from 1 as $\alpha = 1/r$ increases and then rapidly increasing toward 1 again after a certain point. This is not surprising because the model approaches the independence model as $\alpha = 1/r$ approaches either endpoint 0 or 1.

EXAMPLE 5.4 (Multivariate F-distribution). Let $\{X_n\}$ be a kth-order stationary Markov chain in which F_{k+1} is a multivariate F-distribution with

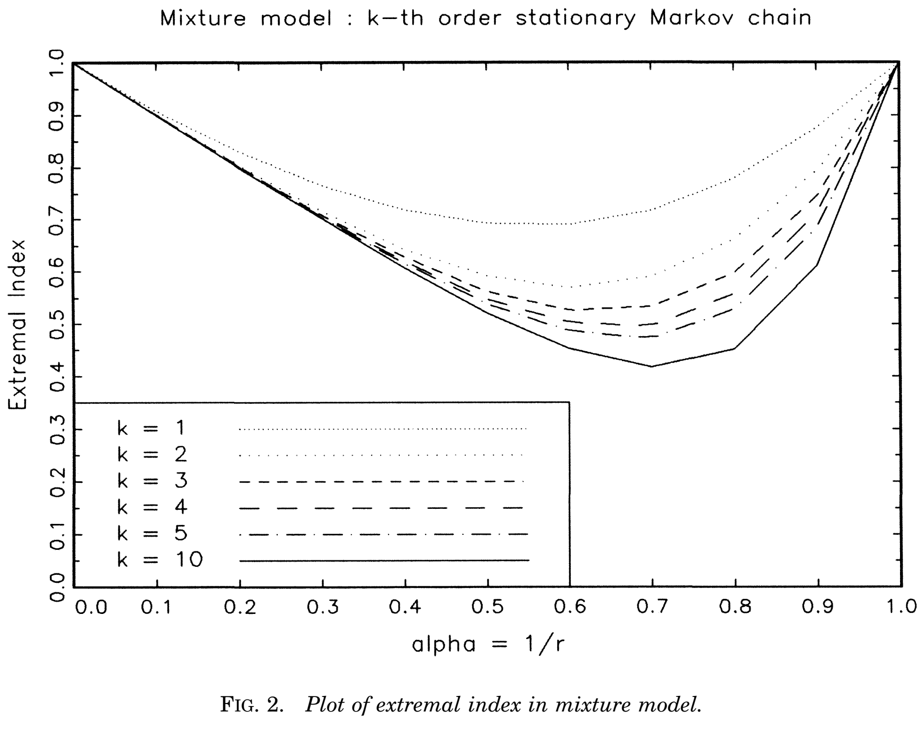

FIG. 2. *Plot of extremal index in mixture model.*

parameter ν (ν: a positive integer), which is defined by its density f_{k+1} as [see Johnson and Kotz (1972), page 240]:

$$f_{k+1}(\mathbf{x}_{k+1}) = \frac{\Gamma((1+k/2)\nu)}{(\Gamma(\nu/2))^{k+2}} \cdot \frac{\prod_{s=1}^{k+1} x_s^{\nu/2-1}}{(1+\sum_{s=1}^{k+1} x_s)^{(1+k/2)\nu}}, \qquad \mathbf{x}_{k+1} > \mathbf{0}. \tag{5.8}$$

Here, the univariate marginal F_1 is the $\mathrm{F}(\nu,\nu)$-distribution so that $F_1 \in \mathscr{D}(\Omega_{2/\nu})$. This model is therefore considered here as an example for which $\xi = 2/\nu$ is positive. It is readily checked that $F_{k+1} \in \mathscr{D}(G_{k+1})$ with auxiliary function $g(u) = 2\nu^{-1}u$, where $G_{k+1}(\mathbf{x}_{k+1})$, $1 + 2\nu^{-1}\mathbf{x}_{k+1} > \mathbf{0}$, is given by

$$-\log G_{k+1}(\mathbf{x}_{k+1}) = \sum_{i=1}^{k+1}(1+2\nu^{-1}x_i)^{-\nu/2} + \sum_{\substack{D\subset\{1,\dots,k+1\}\\ |D|\ge 2}} (-1)^{|D|-1} J_{k+1}(D;\mathbf{x}_{k+1}),$$

the $J_{k+1}(D;\mathbf{x}_{k+1})$ being defined by

$$\frac{(2/\nu)^{r-1}\Gamma((r+1)\nu/2)}{\Gamma(\nu)(\Gamma(\nu/2))^{r-1}} \int_{x_{i_1}}^{\infty} \cdots \int_{x_{i_r}}^{\infty} \frac{\prod_{s=1}^{r}(1+2\nu^{-1}y_s)^{\nu/2-1}}{\{\sum_{s=1}^{r}(1+2\nu^{-1}y_s)\}^{(r+1)\nu/2}}\, d\mathbf{y}_r$$

for $D = \{i_1 < \cdots < i_r\} \subset \{1,\dots,k+1\}$ with $r \ge 2$. Using the fact that every lower-dimensional marginal density function $f_i(\mathbf{x}_i)$, $i = 1,\dots,k$, is of the same form as (5.8), it can be easily seen that, for each $j = 1,\dots,k$,

$$\begin{aligned} l_j(x_{j+1};\mathbf{x}_j) &= \lim_{u\to\infty} g(u) f_{j+1}(u+g(u)x_{j+1} \mid u+g(u)\mathbf{x}_j) \\ &= \frac{\Gamma((1+j/2)\nu)}{\Gamma(\nu/2)\Gamma((j+1)\nu/2)} \cdot \frac{2/\nu}{1+2\nu^{-1}x_{j+1}} \left\{\frac{1+2\nu^{-1}x_{j+1}}{\sum_{s=1}^{j+1}(1+2\nu^{-1}x_s)}\right\}^{\nu/2} \\ &\quad \times \left\{\frac{\sum_{s=1}^{j}(1+2\nu^{-1}x_s)}{\sum_{s=1}^{j+1}(1+2\nu^{-1}x_s)}\right\}^{(j+1)\nu/2}, \qquad x_1,\dots,x_{j+1} > -\nu/2. \end{aligned}$$

Notice here that each $l_j(x;\mathbf{x}_j)$ is the p.d.f. of the r.v. $X = \{\sum_{s=1}^{j}(x_s + \nu/2)\}/(1/B - 1) - \nu/2$, where $B \sim \mathrm{Beta}(\nu/2, (j+1)\nu/2)$. Therefore, from Lemma 2.2, the extremal index θ of $\{X_n\}$ is given by (3.7). Here, the basic condition (3.5) is assumed to hold.

The algorithm to generate $\{Z_n\}$ is as follows.

ALGORITHM 3.

1. $Z_0 \leftarrow 0$.
2. For $1 \le j \le k$, generate B_j from $\mathrm{Beta}(\nu/2, (j+1)\nu/2)$ and $Z_j \leftarrow 2^{-1}\nu\{\log(\sum_{s=1}^{j}\exp(2\nu^{-1}Z_{s-1})) - \log(1/B_j - 1)\}$.

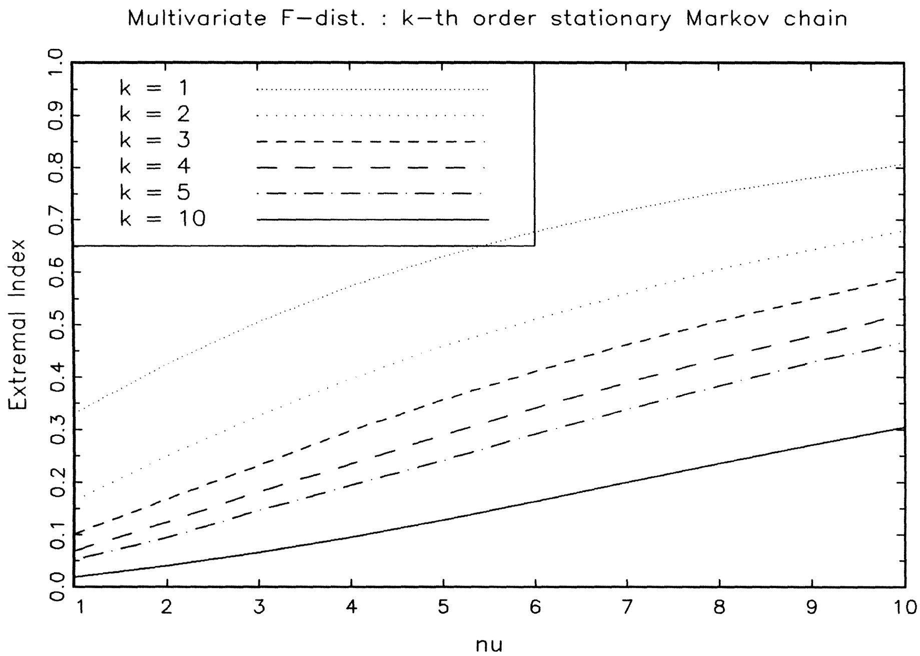

FIG. 3. *Plot of extremal index in multivariate F-distribution.*

3. For $j \geq k+1$, generate B_j from Beta$(\nu/2, (k+1)\nu/2)$ and $Z_j \leftarrow 2^{-1}\nu\{\log(\sum_{s=j-k+1}^{j} \exp(2\nu^{-1} Z_{s-1})) - \log(1/B_j - 1)\}$.

For each fixed $k = 1, 2, 3, 4, 5, 10$, I computed θ_p defined by (5.7), based on 10^5 simulations of the process $\{Z_n\}$ and T, varying $p = 10, 20, 50, 100, 200, 500$ for every $\nu = 1, 2, \ldots, 10$. The convergence rate of θ_p turns out to be faster than that in the logistic model. The simulation results for the extremal index θ based on θ_{500} are plotted in Figure 3.

Acknowledgments. This paper is based on the author's Ph.D. dissertation. He would like to thank Professor R. L. Smith for suggesting the problem and thereafter for continuing guidance, encouragement and helpful discussions during the course of his Ph.D. dissertation at the University of North Carolina, Chapel Hill. He also would like to thank an anonymous referee for detailed valuable comments which led to considerable improvements in both content and shape.

REFERENCES

ASMUSSEN, S. (1987). *Applied Probability and Queues*. Wiley, Chichester.

BERMAN, S. M. (1964). Limit theorems for the maximum term in stationary sequences. *Ann. Math. Statist.* **35** 502–516.

BILLINGSLEY, P. (1985). *Probability and Measure*, 2nd ed. Wiley, New York.

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

COLES, S. G. and TAWN, J. A. (1991). Modelling extreme multivariate events. *J. Roy. Statist. Soc. Ser. B* **53** 377–392.

DEKKERS, A. L. M. and DE HAAN, L. (1989). On the estimation of the extreme-value index and large quantile estimation. *Ann. Statist.* **17** 1795–1832.

GALAMBOS, J. (1987). *The Asymptotic Theory of Extreme Order Statistics*, 2nd ed. Krieger, Florida. (1st ed. published 1978 by Wiley, New York.)

HSING, T. (1984). Point processes associated with extreme value theory. Ph.D. dissertation, Dept. Statistics, Univ. North Carolina, Chapel Hill.

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

HSING, T. (1991). Estimating the parameters of rare events. *Stochastic Process. Appl.* **37** 117–139.

HSING, T. (1993). Extremal index estimation for a weakly dependent stationary sequence. *Ann. Statist.* **21** 2043–2071.

JOHNSON, N. L. and KOTZ, S. (1972). *Distributions in Statistics: Continuous Multivariate Distributions*. Wiley, New York.

LEADBETTER, M. R. (1974). On extreme values in stationary sequences. *Z. Wahrsch. Verw. Gebiete* **28** 289–303.

LEADBETTER, M. R. (1983). Extremes and local dependence in stationary sequences. *Z. Wahrsch. Verw. Gebiete* **65** 291–306.

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

LOYNES, R. M. (1965). Extreme values in uniformly mixing stationary stochastic processes. *Ann. Math. Statist.* **36** 993–999.

MARSHALL, A. W. and OLKIN, I. (1983). Domains of attraction of multivariate extreme value distributions. *Ann. Probab.* **11** 168–177.

NANDAGOPALAN, S. (1990). Multivariate extremes and estimation of the extremal index. Ph.D. dissertation, Dept. Statistics, Univ. North Carolina, Chapel Hill.

NUMMELIN, E. (1984). *General Irreducible Markov Chains and Non-Negative Operators*. Cambridge Univ. Press.

O'BRIEN, G. L. (1974). The maximum term of uniformly mixing stationary processes. *Z. Wahrsch. Verw. Gebiete* **30** 57–63.

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

PERFEKT, R. (1994). Extremal behaviour of stationary Markov chains with applications. *Ann. Appl. Probab.* **4** 529–548.

PERFEKT, R. (1997). Extreme value theory for a class of Markov chains with values in $\Re^d$. *Adv. in Appl. Probab.* **29** 138–164.

PICKANDS, J. (1975). Statistical inference using extreme order statistics. *Ann. Statist.* **3** 119–131.

PICKANDS, J. (1981). Multivariate extreme value distributions. In *Proceedings of the 43rd Session of the International Statistical Institute*, Buenos Aires, Argentina **2** 859–878.

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

ROOTZÉN, H. (1978). Extremes of moving averages of stable processes. *Ann. Probab.* **6** 847–869.

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

SMITH, R. L. (1987). Estimating tails of probability distributions. *Ann. Statist.* **15** 1174–1207.

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

SMITH, R. L. (1992b). On an approximation formula for the extremal index in a Markov chain. Mimeo Series 2087, Institute of Statistics, Dept. Statistics, Univ. North Carolina, Chapel Hill.

SMITH, R. L. and WEISSMAN, I. (1994). Estimating the extremal index. *J. Roy. Statist. Soc. Ser. B* **56** 515–528.

SMITH, R. L., TAWN, J. A. and COLES, S. G. (1997). Markov chain models for threshold exceedances. *Biometrika* **84** 249–268.

TAWN, J. A. (1988). Bivariate extreme value theory: models and estimation. *Biometrika* **75** 397–415.

TAWN, J. A. (1990). Modelling multivariate extreme value distributions. *Biometrika* **77** 245–253.

YUN, S. (1994). Extremes and threshold exceedances in higher order Markov chains with applications to ground-level ozone. Ph.D. dissertation, Dept. Statistics, Univ. North Carolina, Chapel Hill.

DEPARTMENT OF APPLIED STATISTICS
UNIVERSITY OF SUWON
SUWON, KYONGGI-DO 445-743
SOUTH KOREA
E-MAIL: syun@stat.suwon.ac.kr