Extremal Indices, Geometric Ergodicity Of Markov Chains, and MCMC *

by

Gareth O. Roberts
Department of Mathematics and Statistics
Lancaster University

and

Jeffrey S. Rosenthal
Department of Statistics
University of Toronto

and

Bruno Sousa
Departamento de Matemática para a Ciência e Tecnologia
Universidade do Minho

Technical Report No. 0405 July 7, 2004

TECHNICAL REPORT SERIES

University of Toronto

Department of Statistics

Extremal Indices, Geometric Ergodicity of Markov Chains, and MCMC *

Gareth O. Roberts † Jeffrey S. Rosenthal ‡ Bruno Sousa §

July 7, 2004

Abstract

We consider the relationship between extremal indices on the one hand, and geometric ergodicity of Markov chains on the other hand. We point out an example of an ergodic Markov chain where the extremal index is zero. A theoretical result is developed to assess when the extremal index is positive under the assumption of geometric ergodicity. Its application is analyzed through some examples arising from MCMC algorithms.

Keywords: Extremal Index; Geometric Ergodocity; Markov Chains, MCMC.

1 Introduction

*Project supported by a Premier's Research Excellence Award of Ontario.

†Gareth O. Roberts: Department of mathematics and Statistics, Fylde College, Lancaster University, Lancaster, LA1 4YF, England; g.o.robert@lancaster.ac.uk

‡Jeffrey S. Rosenthal: Department of Statistics, University of Toronto, Toronto, Ontario, Canada M5S 3G3; jeff@math.toronto.edu; supported in part by NSERC of Canada.

§Bruno Sousa: Departamento de Matemática para a Ciência e Tecnologia, Universidade do Minho, 4800-058 Guimarães, Portugal; bruno@mct.uminho.pt

It is well-established that maxima of stationary processes with given fixed stationary distributions, is affected by the dependence structure of the process. This dependence is effectively captured by the *extremal index.* However little is known about how the extremal index is related to more general mixing properties of stationary processes. The aim of this paper is to relate the extremal index to the concept of *geometric ergodicity.*

The extremal index, written as θ, takes values in $[0, 1]$, and can be interpreted as an indicator of extremal dependence, with $\theta = 1$ indicating asymptotic independence of extreme events. On the other hand $\theta = 0$ represents the case where we can expect strong clusterings of extreme events. In this case it is natural to expect excursions from 'moderate' values to have heavy-tailed distributions. For Markov chains this behaviour is characteristic of non-geometrically ergodic Markov chains. Thus a natural question to ask is whether $\theta = 0$ is related to non-geometric ergodicity. In this paper we shall see that the answer to this question is no in general, but that under certain conditions the two conditions are equivalent.

We begin with a brief overview of the extremal index in section 2. In section 3 an example of an ergodic Markov chain where $\theta = 0$ is given, suggesting that such cases can be of interest in certain frameworks. Results which imply that $\theta > 0$ given geometric ergodicity and additional conditions, are presented in section 4. Section 5 gives examples to illustrate the application of our main result. Throughout the paper we have a particular interest in Markov chains which are produced by MCMC algorithms, and most of our examples come from simple MCMC problems.

2 The Extremal Index

One of the main results in classical extreme value theory is due to Fisher and Tippett in 1928 [3], and proved rigorously in 1943 by Gnedenko [4], and is related to the asymptotic behavior of the maximum of an independent identical distributed (i.i.d.) sequence of random variables, $\{X_n, n \geq 1\}$. The result is as follows:

Theorem 2.1 *Let $M_n = \max_{1 \leq i \leq n} (X_i)$, where X_i's are i.i.d. random variables. If*

$$P[a_n(M_n - b_n) \leq x] \xrightarrow{\mathcal{D}} G(x), \text{ as } n \to \infty \tag{2.1}$$

holds for some constants $a_n > 0$, b_n and some nondegenerate G, then G must have one of the following forms:

Type I: $G(x) = \exp(-e^{-x})$, $x \in \mathbb{R}$,

Type II: $G(x) = \begin{cases} 0, & x \leq 0 \\ \exp(-x^{-\alpha}), & x > 0 \text{ and some } \alpha > 0, \end{cases}$

Type III: $G(x) = \begin{cases} \exp(-(-x)^{\alpha}), & x \leq 0 \text{ and some } \alpha > 0 \\ 1, & x > 0. \end{cases}$

Conversely, any such distribution function G may appear as a limit in (2.1) *and in fact does so when G is itself the distribution function of each X_i.*

The distribution functions in Theorem 1.1 are called the *standard extreme value distributions.* An important result in the study of the limit distribution of the maximum of a sample $(X_1, \ldots, X_n)$ or, in other words, the distribution of the random variable $M_n = \max_{1 \leq i \leq n} (X_i)$, is the following theorem

Theorem 2.2 *Let $\{u_n, n \geq 1\}$ be a sequence of real numbers and $0 \leq \tau \leq \infty$. If $X_1, X_2, \ldots$ is a sequence of i.i.d. random variables with distribution function F, then*

$$P[M_n \leq u_n] \to e^{-\tau}, \tag{2.2}$$

if and only if

$$n(1 - F(u_n)) \to \tau. \tag{2.3}$$

Notice that equation (2.1) is a special case of (2.2) with $\tau = -\log G(x), u_n = a_n^{-1}x + b_n$.

Let $\{X_n, n \geq 1\}$ be a strictly stationary sequence, i.e. $(X_{i_1}, \ldots, X_{i_n}) \stackrel{\mathcal{D}}{=} (X_{i_1+k}, \ldots, X_{i_n+k})$ for any choice of indices $i_1 < \ldots < i_n$ and $k \in \mathbb{N}$. For simplicity, we will use the word "stationary" to refer to a strictly stationary sequence. The extremal index of a stationary sequence is defined as

Definition 2.1 *The stationary process $\{X_n, n \geq 1\}$ has extremal index θ, with $0 \leq \theta \leq 1$, if for each $\tau > 0$*

(i) *there exists a sequence $u_n(\tau)$ such that $n(1 - F(u_n(\tau))) \to \tau$ and*

(ii) $P[M_n \leq u_n(\tau)] \to e^{-\theta\tau}$,

as $n \to \infty$.

From Leadbetter in [7] we have that if there exists a sequence $u_n(\tau)$ such that condition (i) in the previous definition holds for a fixed $\tau > 0$, then there exists such a sequence $u_n(\tau)$ for all $\tau > 0$. As an example, Leadbetter in [7] considered the case that if $u_n(1)$ satisfies condition (i) in Definition 2.1 with $\tau = 1$, then define $u_n(\tau) = u_{[n/\tau]}(1)$.

Some dependence structure needs to be assumed to obtain an extremal types result as Theorem 1.1. Consider any integers $i_1, \ldots, i_n$, and let $F_{i_1 \ldots i_n}(u)$ represents $F_{i_1 \ldots i_n}(u, \ldots, u)$, where $F_{i_1 \ldots i_n}(x_1, \ldots, x_n)$ denotes the joint distribution function of $(X_{i_1}, \ldots, X_{i_n})$. Define

$$\alpha_{n,l} = \max_{\substack{1 \leq i_1 < \ldots < i_p < j_1 < \ldots < j_{p'} \leq n \\ j_1 - i_p \geq l}} \left\{ \left| F_{i_1,\ldots,i_p,j_1,\ldots,j_{p'}}(u_n) - F_{i_1,\ldots,i_p}(u_n) F_{j_1,\ldots,j_{p'}}(u_n) \right| \right\} \tag{2.4}$$

Condition $D(u_n)$ is said to hold if $\alpha_{n,l_n} \to 0$ as $n \to \infty$ for some sequence $l_n = o(n)$.

By Theorem 3.7.1 in Leadbetter *et al* [8], if condition *(i)* of Definition 2.1 holds and $D(u_n)$ is satisfied for each τ, and if $P[M_n \leq u_n]$ converges for some $\tau > 0$, then condition *(ii)* of Definition 2.1 holds with some θ, $0 \leq \theta \leq 1$, for all $\tau > 0$ and thus the process $\{X_n\}$ has an extremal index.

Leadbetter, Lindgren and Rootzén in [8] showed the following extremal types theorem for stationary sequences.

Theorem 2.3 *Let $\{X_n\}$ be a stationary sequence such that $M_n = \max_{1\leq i\leq n}(X_i)$ has a non-degenerate limiting distribution G as in (2.1). Suppose that $D(u_n)$ holds for each u_n of the form $u_n = a_n^{-1}x + b_n$, for each x with $0 < G(x) < 1$. Then G is one of the three classical extremal types.*

Although condition $D(u_n)$ might be sufficient to guarantee the existence of the extremal index, it is not a necessary condition. This is pointed out by Leadbetter in [7] through an example that first appeared in Davis [2]. In this case, condition $D(u_n)$ does not hold, but the extremal index is $1/2$.

In the context of point processes of clusters, Leadbetter in 1983 [7] introduces an equivalent way of determining the extremal index. His main result is as follows:

Theorem 2.4 *Let the stationary sequence $\{X_n\}$ satisfy $D(u_n)$ for each $\tau > 0$ where u_n satisfies (i) in Definition 2.1. Let k_n be chosen to satisfy $k_n\alpha_{n,l_n} \to 0$, as $n \to \infty$ and $k_n l_n = o(n)$ and let $\{X_n\}$ have extremal index θ, $0 \leq \theta \leq 1$. Then the point process N_n of cluster positions for exceedances of u_n converges in distribution to a Poisson Process N on $(0, 1]$ with intensity parameter $\theta\tau$.*

The extremal index can be seen as the inverse of the mean cluster size or, in other words, it can be interpreted as the limiting mean number of exceedances in an interval of length $r_n = o(n)$, given that at least one exceedance occurred in that interval. Let N_{r_n} be the number of exceedances of u_n in an interval of length r_n, we have

Proposition 2.1 *Choose k_n such that $k_n\alpha_{n,l_n} \to 0$, as $n \to \infty$ and $k_n l_n = o(n)$, and let $r_n = [n/k_n]$. If the extremal index exists, it can be determined as*

$$\theta^{-1} = \lim_{n\to\infty} \mathbf{E}\left(N_{r_n}(u_n) | N_{r_n}(u_n) \geq 1\right).$$

The proof of these result can be seen in Leadbetter [7].

3 An Example with Extremal Index Zero

$\theta = 0$ occurs naturally in many contexts of interest. We have particular interest in MCMC algorithms so we give a simple example here in that context.

In the past, interest in the case where the extremal index is equal to zero has been limited to a more theoretical and academic discussion. However, consider the random walk Metropolis (RWM) algorithm on a standard Cauchy density which is stationary with $\pi(x) = \frac{1}{\pi(1+x^2)}$ for $x \in \mathbb{R}$. Let the proposal density $q(x, y)$ be Uniform on the interval $(x - \delta, x + \delta)$. In other words, given X_n, a proposed value Y_{n+1} is generated from a uniform $(X_n - \delta, X_n + \delta)$, i.e. $Y_{n+1} = X_n + U_{n+1}$ with U_{n+1} independent

Uniform $(-\delta, \delta)$ random variables. Then $X_{n+1} = Y_{n+1}$ with probability $\alpha(X_n, Y_{n+1})$, or $X_{n+1} = X_n$ with probability $1 - \alpha(X_n, Y_{n+1})$. The acceptance probability is defined as

$$\alpha(X_n, Y_{n+1}) = \min\left(1, \frac{\pi(Y_{n+1})}{\pi(X_n)}\right) = \min\left(1, \frac{1+X_n^2}{1+Y_{n+1}^2}\right).$$

Theorem 3.1 *Consider $\{X_n\}$ the random walk induced by the Metropolis algorithm on a standard Cauchy density. Then the extremal index of the chain is zero ($\theta = 0$).*

We will need some auxiliary results and definitions to prove this theorem. Consider the new sequence $\{Z_n\}$ defined as follows:

$$Z_n = \begin{cases} Z_{n-1} + U_n & \text{if } Z_{n-1} + U_n \geq \frac{\delta}{2}, \\ \frac{\delta}{2} & \text{o.w.} \end{cases} = \max\left(\frac{\delta}{2}, Z_{n-1} + U_n\right), \tag{3.1}$$

where $\{U_n, n = 1, 2, \ldots\}$ is a collection of i.i.d. Uniform $(-\delta, \delta)$ random variables, as defined above.

Lemma 3.1 *Consider the two processes $\{X_n\}$ and $\{Z_n\}$ defined in the previous paragraphs. If $X_0 \leq Z_0$, then $X_n \leq Z_n$ for all $n \in \mathbb{N}$.*

Proof: Assume that $X_n \leq Z_n$. It needs to be proven that $X_{n+1} \leq Z_{n+1}$. Therefore, consider the four different possible cases:

1. If $X_n \leq 0$ and $U_{n+1} \leq 0$, then $X_{n+1} \leq X_n < 0$. Since $Z_n \geq \frac{\delta}{2}$ for $n \in \mathbb{N}$, the inequality $X_{n+1} \leq Z_{n+1}$ follows immediately.

2. If $X_n \leq 0$ and $U_{n+1} > 0$ such that $X_{n+1} < 0$, then $X_{n+1} \leq Z_{n+1}$ as explained in the previous paragraph. If X_{n+1} is non-negative, then it implies that $U_{n+1} \geq -X_n > 0$, which means that $Z_{n+1} = Z_n + U_{n+1} \geq X_n + U_{n+1} \geq X_{n+1}$.

3. If $X_n > 0$ and $U_{n+1} > 0$, then $Z_{n+1} = Z_n + U_{n+1} \geq X_n + U_{n+1} \geq X_{n+1}$.

4. For $X_n > 0$ and $-2X_n \leq U_{n+1} \leq 0$ the acceptance probability $\alpha(X_n, Y_{n+1}) = 1$, which means that $X_{n+1} = X_n + U_{n+1}$. Therefore, $X_{n+1} \leq Z_n + U_{n+1} \leq Z_{n+1}$. We need to consider two different scenarios: $X_n \geq \frac{\delta}{2}$ and $0 \leq X_n < \frac{\delta}{2}$. In the first case, $-2X_n \leq -\delta$ and since the random walk increments are bounded by $(-\delta, \delta)$, we always accept $X_{n+1} = X_n + U_{n+1}$. Consequently, $X_{n+1} \leq Z_{n+1}$. For the second case, $0 \leq X_n < \frac{\delta}{2}$, then $X_{n+1} = X_n \in (0, \delta/2)$ with probability $1 - \alpha(X_n, Y_{n+1}) < 1$ whenever $U_{n+1} \in (-\delta, -2X_n)$. Since $Z_n \geq \frac{\delta}{2}$ for all $n \in \mathbb{N}$, we can say that $X_{n+1} \leq Z_{n+1}$.

Hence, the RWM process $\{X_n\}$ is bounded from above by the process Z_n, for all $n \in \mathbb{N}$. ∎

Now, consider the following lemma:

Lemma 3.2 *Consider the process $\{Z_n\}$ defined above and the random walk $\{S_n\}$ with increments $\{U_n\}$ i.i.d. Uniform $(-\delta, \delta)$ and $S_0 = 0$. If $Z_0 \geq 0$, then for $n \in \mathbb{N}$*

$$Z_n = \max\left(Z_0 + S_n, S_n - S_1, \ldots, S_n - S_{n-1}, \frac{\delta}{2}\right).$$

Proof: The process $\{Z_n\}$ has increments larger or equal to the ones of the random walk $\{S_n\}$, i.e.

$$Z_n - Z_{n-k} \geq S_n - S_{n-k},$$

for $k = 0, 1, \ldots, n$. For $k = n$ we obtain $Z_n \geq Z_0 + S_n$. Since $Z_{n-k} \geq 0$ for all $k = 0, 1, \ldots, n$, we have $Z_n \geq S_n - S_{n-k}$. Hence,

$$Z_n \geq \max\left(Z_0 + S_n, S_n - S_1, \ldots, S_n - S_{n-1}, \frac{\delta}{2}\right). \tag{3.2}$$

If $Z_0 + S_k \geq \frac{\delta}{2}$ for every $1 \leq k \leq n$ then

$$Z_n = Z_0 + S_n \leq \max\left(Z_0 + S_n, S_n - S_1, \ldots, S_n - S_{n-1}, \frac{\delta}{2}\right). \tag{3.3}$$

Otherwise, there will be a $Z_m = \frac{\delta}{2}$ for some $m \leq n$. Take $k = \max\left\{m : Z_m = \frac{\delta}{2}\right\}$, then

$$Z_n = S_n - S_{n-k} \leq \max\left(Z_0 + S_n, S_n - S_1, \ldots, S_n - S_{n-1}, \frac{\delta}{2}\right). \tag{3.4}$$

Thus, from (3.2), (3.3) and (3.4) we have

$$Z_n = \max\left(Z_0 + S_n, S_n - S_1, \ldots, S_n - S_{n-1}, \frac{\delta}{2}\right),$$

for $n \in \mathbb{N}$. ∎

We will need one last lemma before proving Theorem 3.1.

Lemma 3.3 *Consider the process $\{Z_n\}$ defined in (3.1). If $Z_0 = 0$ and $x > \frac{\delta}{2}$ then*

$$P(\max_{1 \leq j \leq n} Z_j > x) \leq n^2 P(\max_{1 \leq i \leq n} S_i > x), \tag{3.5}$$

with $\{S_n\}$ a random walk with increments $\{U_n\}$ i.i.d. Uniform $(-\delta, \delta)$.

Proof: If $Z_0 = 0$, then from Lemma 3.2

$$P(Z_n > x) = P\left[\max\left(S_n, S_n - S_1, \ldots, S_n - S_{n-1}, \frac{\delta}{2}\right) > x\right].$$

If $x > \frac{\delta}{2}$ then $P(Z_n > x) = P\left[\cup_{i=1}^{n} (U_n + \cdots + U_i) > x\right]$. Therefore,

$$\begin{aligned}
P(\max_{1\leq j\leq n} Z_j > x) &= P\left[\cup_{j=1}^{n}\{Z_j > x\}\right] \leq \sum_{j=1}^{n} P(Z_j > x) \\
&\leq \sum_{j=1}^{n}\sum_{i=1}^{j} P(U_j + \cdots + U_i > x) = \sum_{j=1}^{n}\sum_{i=1}^{j} P(S_{j-i+1} > x) \\
&\leq \sum_{j=1}^{n} j\, P(\max_{1\leq i\leq j} S_{j-i+1} > x) \leq n \sum_{j=1}^{n} P(\max_{1\leq i\leq j} S_{j-i+1} > x) \\
&\leq n^2 P(\max_{1\leq i\leq n} S_i > x).
\end{aligned}$$

■

Let us now prove Theorem 3.1.

Proof: (Of Theorem 3.1) Under stationarity, condition (i) in Definition 2.1 is satisfied for $u_n = \frac{n}{\tau}$. This comes from the fact that $1 - F(x) = \frac{1}{2} - \frac{1}{\pi}\arctan x \sim x^{-1}$, where the symbol $f(x) \sim g(x)$ means that $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$. Hence, with $u_n = \frac{n}{\tau}$, we have $\lim_{n\to\infty} n[1 - F(n/\tau)] = \lim_{n\to\infty} n\frac{\tau}{n} = \tau$.

To verify condition (ii) from Definition 2.1, i.e. $P[M_n \leq u_n] \to e^{-\theta\tau}$ as $n \to \infty$, where the random variable M_n is defined as $\max(X_1, \ldots, X_n)$, consider the sequence $\{Z_n\}$ defined above. Using Lemma 3.1, we can write

$$P[M_n > n/\tau] \;\leq\; P\left[\max_{1\leq i\leq n} Z_i \geq n/\tau\right]. \tag{3.6}$$

Therefore, letting $x = \frac{n}{\tau}$ and using Lemma 3.3, the equation above becomes

$$P[M_n > n/\tau] \;\leq\; P\left[\max_{1\leq i\leq n} Z_i \geq n/\tau\right] \leq n^2 P(\max_{1\leq i\leq n} S_i > n/\tau). \tag{3.7}$$

Using Theorem 1.1 of Gut [5] we obtain

$$\sum_{n=1}^{\infty} n^2 P\left(\max_{1\leq i\leq n} |S_k| > n\varepsilon\right) < \infty \quad \text{for all} \quad \varepsilon > 0.$$

Since $\sum_{n=1}^{\infty} n^2 P\left(\max_{1\leq i\leq n} S_k > n\varepsilon\right) \leq \sum_{n=1}^{\infty} n^2 P\left(\max_{1\leq i\leq n} |S_k| > n\varepsilon\right) < \infty$, and taking $\varepsilon = \tau^{-1} > 0$ we obtain that

$$n^2 P(\max_{1\leq i\leq n} S_i > n/\tau) \to 0 \text{ as } n \to \infty.$$

Therefore, from (3.7) we conclude $P\left(M_n > \frac{n}{\tau}\right) \to 0$ as $n \to \infty$, which is to say that condition (ii) of Definition 2.1 is satisfied with $\theta = 0$. The RWM algorithm on a standard Cauchy density has extremal index equal to 0. ■

The example we have given here uses random walk Metropolis on a heavy tailed target density. In this context, the algorithms are well-known to converge slower than geometrically (see [9]). This strongly suggests a relationship between geometric ergodicity and the extremal index. We would not expect an absolute connection between the concepts since geometric ergodicity considers stability properties related to stationarity whereas the extremal index measures a property of excursions from stationarity. However the next section will show that connections between the concepts can still be made.

In future work we will try to look into the case where $\theta = 0$ and what type of rates of convergence can be expected for the distribution of the maximum. In the next section, we restrict ourselves to $\theta > 0$ and in particular its relationship with geometric ergodicity of a Markov chain.

4 Extremal Index and Geometric Ergodicity

We shall give here some essential background to geometric ergodicity. In fact since we wish to allow our methodology to extend to non-Markov stationary processes, we shall work directly with so-called *drift* conditions which are equivalent to geometric ergodicity in the Markov case.

A set C is small if for some $\delta > 0$, $n > 0$, and some probability measure ν concentrated on C, we have $P^n(x, .) \geq \delta\nu(.)$, for all $x \in C$ and $P^n(x, .)$ the n-step transition probability kernel.

For the following results, let u be any real number, σ_u be the first time the chain hits $[u, \infty)$, and τ_u the first time after σ_u at which the chain hits $(-\infty, u)$.

Lemma 4.1 *If $PV(x) \leq V(x) - c$ for $x \geq u$, some function $V \geq 1$, and a constant $c > 0$, then*

$$\mathbf{E}(\tau_u - \sigma_u \,|\, \mathcal{F}_{\sigma_u}) \leq \frac{V(X_{\sigma_u})}{c}, \tag{4.1}$$

where $\mathcal{F}_{\sigma_u}$ is the information contained up to the stopping time σ_u.

Proof: The result follows from the fact that $Z_n = V(X_n) + cn$ is a supermartingale for $X_n \geq u$, i.e.

$$\begin{aligned} \mathbf{E}(Z_{n+1} | \, Z_n) &= \mathbf{E}\left(V\left(X_{n+1}\right) | \, X_n\right) + c(n+1) = PV\left(X_n\right) + c(n+1) \\ &\leq V\left(X_n\right) - c + c(n+1) = V\left(X_n\right) + cn \\ &= Z_n. \end{aligned}$$

Since Z_n is a supermartingale and $\tau_u > \sigma_u$ we obtain $\mathbf{E}\left(Z_{\tau_u} | \, \mathcal{F}_{\sigma_u}\right) \leq \mathbf{E}\left(Z_{\sigma_u} | \, \mathcal{F}_{\sigma_u}\right)$. By the definition of Z_n we can write the inequality as $\mathbf{E}\left(V\left(X_{\tau_u}\right) | \, \mathcal{F}_{\sigma_u}\right) + c\,\mathbf{E}\left(\tau_u | \, \mathcal{F}_{\sigma_u}\right) \leq V\left(X_{\sigma_u}\right) + c\mathbf{E}\left(\sigma_u | \, \mathcal{F}_{\sigma_u}\right)$. Equivalently, we have $c\,\mathbf{E}\left(\tau_u - \sigma_u | \, \mathcal{F}_{\sigma_u}\right) \leq V\left(X_{\sigma_u}\right) - \mathbf{E}\left(V\left(X_{\tau_u}\right) | \, \mathcal{F}_{\sigma_u}\right)$. Since the function V is nonnegative, the result in (4.1) follows immediately, i.e. $\mathbf{E}\left(\tau_u - \sigma_u | \, \mathcal{F}_{\sigma_u}\right) \leq \frac{V(X_{\sigma_u})}{c}$. ■

Lemma 4.2 *If $PV(x) \leq \lambda V(x)$ for $x \in C^c$, with C a small set such that $C \subseteq (-\infty, u)$, some non-decreasing function $V \geq 1$ and $\lambda \in (0,1)$, then*

$$\mathbf{E}(\tau_u - \sigma_u | \mathcal{F}_{\sigma_u}) \leq \frac{V(X_{\sigma_u})}{(1-\lambda)V(u)}, \tag{4.2}$$

where $\mathcal{F}_{\sigma_u}$ is the information contained up to the stopping time σ_u.

Proof: The condition $PV \leq \lambda V$ can be written as $PV \leq V - (1-\lambda)V$ with $\lambda \in (0,1)$. Since V in non-decreasing, and for $x \geq u$ we have $PV(x) \leq V(x) - (1-\lambda)V(x) \leq V(x) - (1-\lambda)V(u)$. Apply the previous lemma with $c = (1-\lambda)V(u)$ and the result (4.2) follows. ∎

Lemma 4.3 *If $PV(x) \leq \lambda V(x)$ for $x \in C^c$, with C a small set such that $C \subseteq (-\infty, u)$, some non-decreasing function $V \geq 1$ and $\lambda \in (0,1)$, then*

$$\mathbf{E}\left(\tau_u - \sigma_u | \sigma_u\right) \leq \frac{\mathbf{E}[V(X_{\sigma_u})]}{(1-\lambda)V(u)}.$$

Proof: It follows immediately from Lemma 4.2, i.e.

$$\mathbf{E}(\tau_u - \sigma_u | \sigma_u) = \mathbf{E}[\mathbf{E}(\tau_u - \sigma_u | \sigma_u, X_{\sigma_u})] \leq \frac{\mathbf{E}[V(X_{\sigma_u})]}{(1-\lambda)V(u)}.$$

∎

Lemma 4.4 *Let $\{X_n\}$ be a stationary process such that $PV(x) \leq \lambda V(x)$ for $x \in C^c$, with C a small set such that $C \subseteq (-\infty, u)$, and some non-decreasing function $V \geq 1$. Then, for any positive constant r we obtain*

$$\mathbf{E}\left(N_r(u) | N_r(u) \geq 1\right) \leq \frac{\mathbf{E}[V(X_{\sigma_u})]}{(1-\lambda)V(u)} + \frac{\lambda}{1-\lambda} + r\frac{A}{V(u)},$$

where $N_r(u)$ is the number of exceedances of u in an interval of length r, and $A = \sup_{x \in C} PV(x) < \infty$.

Proof: Define $\tau_C = \inf\{n > \sigma_u : X_n \in C\}$, the first time after σ_u at which the chain hits the set C, and let the $\sum_{k=1}^{0}(.) = 0$. Therefore,

$$\begin{aligned}
\mathbf{E}\left[N_r(u) | N_r(u) \geq 1\right] \leq\ & \mathbf{E}[\tau_u - \sigma_u | \sigma_u \leq r] & (4.3)\\
& + \mathbf{E}[\#\text{ of exceedances in the interval } (\tau_u, \tau_C)] & (4.4)\\
& + \mathbf{E}[\#\text{ of exceedances in the interval } (\tau_C, r)]. & (4.5)
\end{aligned}$$

From Lemma 4.3, we know that (4.3) is bounded from above by $\frac{\mathbf{E}[V(X_{\sigma_u})]}{(1-\lambda)V(u)}$.

Consider $\mu = \mathcal{L}(X_{\tau_u})$. Thus, the expected value in (4.4) becomes

$$
\begin{aligned}
&\mathbf{E}[\#\text{ of exceedances in the interval } (\tau_u, \tau_C)] = \\
&= \mathbf{E}\left[\mathbf{E}[\#\text{ of exceedances in the interval } (\tau_u, \tau_C) | X_{\tau_u}]\right] \\
&= \int \mathbf{E}\left(\sum_{k=1}^{\tau_C - \tau_u} \mathbf{1}_{[u,\infty)}(X_{\tau_u+k}) | X_{\tau_u} = y\right) \mu(dy) \mathbf{1}_{C^c}(X_{\tau_u}, \ldots, X_{\tau_C - 1}) \\
&\leq \int \mathbf{E}\left(\sum_{k=1}^{\infty} \mathbf{1}_{[u,\infty)}(X_{\tau_u+k}) | X_{\tau_u} = y\right) \mu(dy) \mathbf{1}_{C^c}(X_{\tau_u}, \ldots, X_{\tau_C - 1}) \\
&= \int \sum_{k=1}^{\infty} \mathbf{E}\left(\mathbf{1}_{[u,\infty)}(X_{\tau_u+k}) | X_{\tau_u} = y\right) \mu(dy) \mathbf{1}_{C^c}(X_{\tau_u}, \ldots, X_{\tau_C - 1}) \\
&= \sum_{k=1}^{\infty} \int P\left(X_{\tau_u+k} \geq u | X_{\tau_u} = y\right) \mu(dy) \mathbf{1}_{C^c}(X_{\tau_u}, \ldots, X_{\tau_C - 1}), \\
&= \sum_{k=1}^{\infty} \int P(X_k \geq u | X_0 = y) \mu(dy) \mathbf{1}_{C^c}(X_0, \ldots, X_{k-1}) \\
&= \sum_{k=1}^{\infty} \int P(V(X_k) \geq V(u) | X_0 = y) \mu(dy) \mathbf{1}_{C^c}(X_0, \ldots, X_{k-1}) \text{ (V is non-decreasing)}
\end{aligned}
$$

$$
\begin{aligned}
&\leq \sum_{k=1}^{\infty} \int \frac{\mathbf{E}_y(V(X_k))}{V(u)} \mu(dy) \mathbf{1}_{C^c}(X_0, \ldots, X_{k-1}) \text{ (Markov's inequality)} \\
&= \sum_{k=1}^{\infty} \int \frac{P^k V(y)}{V(u)} \mu(dy) \mathbf{1}_{C^c}(X_0, \ldots, X_{k-1}) \\
&\leq \sum_{k=1}^{\infty} \int \frac{\lambda^k V(u)}{V(u)} \mu(dy) \text{ (Drift condition and } y < u) \\
&= \frac{\lambda}{1-\lambda}.
\end{aligned} \tag{4.6}
$$

Let $\nu = \mathcal{L}(X_{\rho_k})$, where $\rho_k = \min\{s : X_m \in C^c \text{ for } s < m \leq k\}$ is the last time before k that the chain visits the set C. The expected value in (4.5) is equal to

$$\begin{aligned}
\mathbf{E}\left[\#\text{ of exceedances in the interval } (\tau_C, r]\right] &= \sum_{k=1}^{r} P\left(X_k \geq u, \tau_C \leq k\right) \\
&= \sum_{k=1}^{r} P\left(X_k \geq u \middle| \tau_C \leq k\right) P\left(\tau_C \leq k\right) \\
&= \sum_{k=1}^{r} \sum_{l} P\left(X_k \geq u, \rho_k = l \middle| \tau_C \leq k\right) P\left(\tau_C \leq k\right) \\
&= \sum_{k=1}^{r} \sum_{l} P\left(X_k \geq u \middle| \rho_k = l, \tau_C \leq k\right) P\left(\rho_k = l \middle| \tau_C \leq k\right) P\left(\tau_C \leq k\right) \\
&= \sum_{k=1}^{r} \sum_{l} \int P\left(X_k \geq u \middle| X_{\rho_k} = x, \rho_k = l, \tau_C \leq k\right) P\left(\rho_k = l,\, \tau_C \leq k\right) \nu(dx) \\
&\leq \sum_{k=1}^{r} \sum_{l} \int P\left(V(X_k) \geq V(u) \middle| X_{\rho_k} = x, \rho_k = l, \tau_C \leq k\right) P\left(\rho_k = l\right) \nu(dx) \\
&\leq \sum_{k=1}^{r} \sum_{l} \int \frac{\mathbf{E}\left[V(X_k) \middle| X_{\rho_k} = x, \rho_k = l, \tau_C \leq k\right]}{V(u)} P\left(\rho_k = l\right) \nu(dx) \\
&\leq \frac{1}{V(u)} \sum_{k=1}^{r} \sum_{l} \int \lambda^{k-\rho_k-1} PV(x) P\left(\rho_k = l\right) \nu(dx) \text{ (Drift condition)} \\
&\leq \frac{1}{V(u)} \sum_{k=1}^{r} \int PV(x) \nu(dx) \\
&\leq \frac{1}{V(u)} \sum_{k=1}^{r} \int \sup_x PV(x) \nu(dx) \\
&\leq r \frac{A}{V(u)}. \qquad (4.7)
\end{aligned}$$

Putting the above together, we obtain

$$\mathbf{E}\left(N_r(u) \middle| N_r(u) \geq 1\right) \leq \frac{\mathbf{E}[V(X_{\sigma_u})]}{(1-\lambda)V(u)} + \frac{\lambda}{1-\lambda} + r\frac{A}{V(u)}.$$

■

Consider $\{u_n\}$ a sequence of real numbers.

Theorem 4.1 *Let $\{X_n\}$ be a stationary process satisfying the drift condition $PV \leq \lambda V + b\mathbf{1}_C$ for a small set C bounded from above and, for some $n_0 \in \mathbb{N}$, $C \subseteq (-\infty, u_n)$ $\forall_{n \geq n_0}$, and a non-decreasing function V. Suppose further that condition $D(u_n)$ holds for some sequence $l_n = o(n)$ and that the process has an extremal index in the sense of Definition 2.1. Then the extremal index of the process is positive, $\theta > 0$, if* (*i*) *the sequence u_n satisfies condition (i) in Definition 2.1,*

(*ii*) $\frac{\mathbf{E}[V(X_{\sigma_{u_n}})]}{V(u_n)}$ *is bounded in n, and*

(*iii*) $\frac{r_n}{V(u_n)}$ *is bounded,*

for some positive sequence r_n.

Proof: Since condition $D(u_n)$ holds for some sequence $l_n = o(n)$, choose $k_n = \min\left(\alpha_{n,l_n}^{-1/2}, \left(\frac{n}{l_n}\right)^{1/2}\right)$. Since $k_n \leq \alpha_{n,l_n}^{-1/2}$, then $k_n\alpha_{n,l_n} \leq \sqrt{\alpha_{n,l_n}} \to 0$ as $n \to \infty$. On the other hand, $\frac{k_n l_n}{n} \leq \sqrt{\frac{n}{l_n}\frac{l_n}{n}} = \sqrt{\frac{l_n}{n}} \to 0$ as $n \to \infty$. Therefore, with this choice of k_n, the conditions $k_n\alpha_{n,l_n} \to 0$, as $n \to \infty$ and $k_n l_n = o(n)$ from Theorem 3.4 in [7] are satisfied. Let $r_n = \left[\frac{n}{k_n}\right]$ and, since $\{X_n\}$ has an extremal index, it follows:

$$\theta^{-1} = \lim_{n\to\infty} \mathbf{E}\left(N_{r_n}(u_n) | N_{r_n}(u_n) \geq 1\right). \tag{4.8}$$

Since the set C is bounded from above, let's say by $M < \infty$, and V is a non-decreasing function, then $\sup_{x\in C} PV(x) \leq \lambda \sup_{x\in C} V(x) + b \leq \lambda V(M) + b < \infty$. Hence, from Lemma 4.4, we have that for all $n \geq n_0$

$$\mathbf{E}\left(N_{r_n}(u_n) | N_{r_n}(u_n) \geq 1\right) \leq \frac{\mathbf{E}[V(X_{\sigma_{u_n}})]}{(1-\lambda)V(u_n)} + \frac{\lambda}{1-\lambda} + r_n\frac{A}{V(u_n)}, \tag{4.9}$$

with σ_{u_n} the first time the chain hits $[u_n, \infty)$, and $A = \lambda V(M) + b < \infty$. Conditions (ii), (iii) imply that

$$\mathbf{E}\left(N_{r_n}(u_n) | N_{r_n}(u_n) \geq 1\right) < \infty,$$

for all $n \geq n_0$.

Letting $n \to \infty$ and considering (4.8), it follows that $\theta^{-1} < \infty$, i.e. $\theta > 0$. ∎

In what follows, we are particularly interested in the dynamics between the extremal index and geometric ergodic Markov chains defined on a state space $\mathcal{X}$, with transition probabilities $P(x,.)$ and stationary distribution $\pi(.)$. A fundamental result by Meyn and Tweedie [10] says that the chain is geometric ergodic if and only if it satisfies the geometric drift condition given by

$$PV(x) \equiv \int V(y)P(x,dy) \leq \lambda V(x) + b\mathbf{1}_C(x), \quad x \in \mathcal{X}, \tag{4.10}$$

for a small set $C \subseteq \mathcal{X}$, a $\pi - a.e.$ finite function $V : \mathcal{X} \to [1,\infty]$, and constants $\lambda < 1$ and $b < \infty$. It also follows from Meyn and Tweedie (1993) that we can always choose V and C such that $\sup_{x\in C} V(x) < \infty$.

Consider some auxiliary results.

Lemma 4.5 *Let $\{X_n\}$ be a stationary Markov chain, and consider $\alpha_{n,l}$ defined in* (2.4). *Then*

$$\alpha_{n,l} \leq \max\left|P(X_{j_1} \leq u_n | X_{i_p} \leq u_n) - P(X_{j_1} \leq u_n)\right|, \tag{4.11}$$

where the maximum is taken over all positive integers $i_p < j_1$ such that $j_1 - i_p \geq l$.

Proof: Consider the integers $i_1, \dots, i_p$ and $j_1, \dots, j_{p'}$, such that $1 \le i_1 < \dots < i_p < j_1 < \dots < j_{p'} \le n$ and $j_1 - i_p \ge l$. Let $F_{i_1,\dots,i_p,j_1,\dots,j_{p'}}(u_n) = P(X_{i_1} \le u_n, \dots, X_{j_{p'}} \le u_n)$. From the Markov property it follows that

$$\begin{aligned} F_{i_1,\dots,i_p,j_1,\dots,j_{p'}}(u_n) &= P(X_{i_1} \le u_n) P(X_{i_2} \le u_n | X_{i_1} \le u_n) \dots P(X_{j_1} \le u_n | X_{i_p} \le u_n) \\ & \quad P(X_{j_2} \le u_n | X_{j_1} \le u_n) \dots P(X_{j_{p'}} \le u_n | X_{j_{p'}-1}). \end{aligned} \tag{4.12}$$

Similarly,

$$F_{i_1,\dots,i_p}(u_n) = P(X_{i_1} \le u_n) P(X_{i_2} \le u_n | X_{i_1} \le u_n) \dots P(X_{i_p} \le u_n | X_{i_p-1} \le u_n), \tag{4.13}$$

and

$$F_{j_1,\dots,j_{p'}}(u_n) = P(X_{j_1} \le u_n) P(X_{j_2} \le u_n | X_{j_1} \le u_n) \dots P(X_{j_{p'}} \le u_n | X_{j_{p'}-1}). \tag{4.14}$$

From equations (4.12)-(4.14), we can write $\alpha_{n,l}$ in (2.4) as follows:

$$\begin{aligned} \alpha_{n,l} &= \max \Big| P(X_{i_1} \le u_n) P(X_{i_2} \le u_n | X_{i_1} \le u_n) \dots P(X_{i_p} \le u_n | X_{i_p-1} \le u_n) \times \\ & \quad \times P(X_{j_1} \le u_n | X_{i_p} \le u_n) P(X_{j_2} \le u_n | X_{j_1} \le u_n) \dots P(X_{j_{p'}} \le u_n | X_{j_{p'}-1}) - \\ & \quad - P(X_{i_1} \le u_n) P(X_{i_2} \le u_n | X_{i_1} \le u_n) \dots P(X_{i_p} \le u_n | X_{i_p-1} \le u_n) \times \\ & \quad \times P(X_{j_1} \le u_n) P(X_{j_2} \le u_n | X_{j_1} \le u_n) \dots P(X_{j_{p'}} \le u_n | X_{j_{p'}-1}) \Big| \\ &= \max \Big| P(X_{i_1} \le u_n) P(X_{i_2} \le u_n | X_{i_1} \le u_n) \dots P(X_{i_p} \le u_n | X_{i_p-1} \le u_n) \times \\ & \quad \times P(X_{j_2} \le u_n | X_{j_1} \le u_n) \dots P(X_{j_{p'}} \le u_n | X_{j_{p'}-1}) \times \\ & \quad \times [P(X_{j_1} \le u_n | X_{i_p} \le u_n) - P(X_{j_1} \le u_n)] \Big|, \end{aligned}$$

where the maximum is taken over all integers $i_1, \dots, i_p$ and $j_1, \dots, j_{p'}$, such that $1 \le i_1 < \dots < i_p < j_1 < \dots < j_{p'} \le n$ and $j_1 - i_p \ge l$.

Since the $P(X_{i_1} \le u_n) P(X_{i_2} \le u_n | X_{i_1} \le u_n) \dots P(X_{i_p} \le u_n | X_{i_p-1} \le u_n) P(X_{j_2} \le u_n | X_{j_1} \le u_n) \dots P(X_{j_{p'}} \le u_n | X_{j_{p'}-1}) \le 1$, we obtain

$$\alpha_{n,l} \le \max \left| P(X_{j_1} \le u_n | X_{i_p} \le u_n) - P(X_{j_1} \le u_n) \right|,$$

where the maximum is taken over all positive integers $i_p < j_1$ such that $j_1 - i_p \ge l$. ■

Lemma 4.6 *Let $\{X_n\}$ be a geometric ergodic Markov chain. Then $\alpha_{n,l_n} \to 0$ as $n \to \infty$ for any sequence $l_n = o(n)$ such that $l_n \to \infty$ as $n \to \infty$.*

Proof: If the Markov chain is geometric ergodic then there exists some $\rho < 1$ and $R < \infty$ such that, for all $n \in \mathbb{N}$ and all $A \subseteq \mathcal{X}$

$$|P^n(x, A) - \pi(A)| \le RV(x)\rho^n, \tag{4.15}$$

where π is the stationary distribution and V is any solution of the drift condition

$$PV \le \lambda V + b\mathbf{1}_C,$$

for $\lambda < 1$, $b < \infty$ and C a small set. Applying Theorem 14.3.7 of Meyn and Tweedie [10] with $f(x) = (1-\lambda)V(x)$ and $s(x) = b\mathbf{1}_C(x)$ we obtain that

$$\pi(V) = \int V(x)\pi(dx) \leq \frac{b}{1-\lambda} < \infty. \tag{4.16}$$

Under stationarity, it follows from Lemma 4.5 and equation (4.15) that, for any positive integers $i_p < j_1$ such that $j_1 - i_p \geq l$,

$$\begin{aligned} \alpha_{n,l_n} &\leq \max_{i_p,j_1} \left| P(X_{j_1} \leq u_n | X_{i_p} \leq u_n) - P(X_{j_1} \leq u_n) \right| \\ &\leq \max_{i_p,j_1} \frac{\int \left| P^{j_1-i_p}(x,(-\infty,u_n]) - \pi((-\infty,u_n]) \right| \pi(dx)\mathbf{1}_{(-\infty,u_n]}(x)}{\pi((-\infty,u_n])} \\ &\leq \max_{i_p,j_1} \frac{\int RV(x)\rho^{j_1-i_p}\pi(dx)\mathbf{1}_{(-\infty,u_n]}(x)}{\pi((-\infty,u_n])} \\ &\leq \frac{R\pi(V)}{\pi((-\infty,u_n])} \max_{i_p,j_1} \rho^{j_1-i_p}. \end{aligned} \tag{4.17}$$

From the definition of α_{n,l_n}, $j_1 - i_p \geq l_n$. Therefore, equation (4.17) can be bounded from above as follows:

$$\alpha_{n,l_n} \leq \frac{R\pi(V)}{\pi((-\infty,u_n])} \rho^{l_n}. \tag{4.18}$$

Equation (4.16) implies that $\frac{R\pi(V)}{\pi((-\infty,u_n])} < \infty$. Now, choose a sequence $l_n = o(n)$ such that $l_n \to \infty$. Then, since $\rho < 1$, $\alpha_{n,l_n} \to 0$ as $n \to \infty$ for any such sequence l_n. ∎

Theorem 4.2 *Let $\{X_n\}$ be a stationary Markov chain which is geometrically ergodic. Suppose it satisfies $PV \leq \lambda V + b\mathbf{1}_C$, for a small set C bounded from above and, for some $n_0 \in \mathbb{N}$, $C \subseteq (-\infty, u_n)\ \forall_{n \geq n_0}$, and where the drift function V is non-decreasing. Suppose further that the Markov chain has extremal index θ in the sense of Definition 2.1. Let $l_n = o(n)$ such that $l_n \to \infty$ as $n \to \infty$ and consider α_{n,l_n} defined in (2.4). Then the extremal index of the chain is positive, $\theta > 0$, if*

- (i) *the sequence u_n satisfies condition (i) in Definition 2.1,*
- (ii) $\frac{\mathbf{E}[V(X_{\sigma_{u_n}})]}{V(u_n)}$ *is bounded in n, and*
- (iii) $\frac{r_n}{V(u_n)}$ *is bounded,*

for some positive sequence r_n.

Proof: From Lemma 4.6, $\alpha_{n,l_n} \to 0$ as $n \to \infty$ for any sequence $l_n = o(n)$ such that $l_n \to \infty$ as $n \to \infty$. Therefore, condition $D(u_n)$ holds and the remainder of the proof is identical to the proof of Theorem 4.1. ∎

In what follows, we present several simplifications of the conditions in Theorem 4.1 and 4.2, but first, consider the following definitions:

Definition 4.1 *The process $\{X_n\}$ with density $dP(x,y)$ is* strongly stochastic monotone *(SSM) if for $x_1 < x_2$, $\frac{dP(x_2,y)}{dP(x_1,y)}$ is non-decreasing on y.*

Definition 4.2 *The process $\{X_n\}$ with density $dP(x,y)$ has* strongly stochastic decreasing increments *(SSDI) if for $x_1 < x_2$, $\frac{dP(x_2,y+x_2)}{dP(x_1,y+x_1)}$ is non-increasing on y.*

Corollary 4.1 *If $\{X_n\}$ is a* strongly stochastic monotone *Markov chain with strongly stochastic decreasing increments and V is log-Lipschitz then condition (ii) of Theorems 4.1 and 4.2 holds.*

Proof: To show that (ii) from Theorem 4.1 and 4.2 holds we only need to prove that $\frac{\mathbf{E}[V(X_{\sigma_u})]}{V(u)} \leq c < \infty$, for any real u and some constant c.

$$\begin{aligned}
\frac{\mathbf{E}[V(X_{\sigma_u})]}{V(u)} &= \mathbf{E}[\exp\{\log V(X_{\sigma_u}) - \log V(u)\}] \\
&\leq \mathbf{E}[\exp\{\alpha(X_{\sigma_u} - u)\}] \qquad \text{(V is log-Lipschitz)} \\
&= \mathbf{E}[\exp\{\alpha(X_{\sigma_u} - u)\} | \sigma_u = 0]P(\sigma_u = 0) + \mathbf{E}[\exp\{\alpha(X_{\sigma_u} - u)\} | \sigma_u > 0]P(\sigma_u > 0) \\
&\leq \mathbf{E}[\exp\{\alpha(X_o - u)\} | X_o \geq u]P(X_o \geq u) + \mathbf{E}_u[\exp\{\alpha(X_{\sigma_u} - u)\}] \qquad \text{(by SSM)} \\
&= \mathbf{E}[\exp\{\alpha(X_o - u)\}] + \mathbf{E}_u[\exp\{\alpha(X_{\sigma_u} - u)\}] \\
&\leq \mathbf{E}[\exp\{\alpha(X_o - u)\}] + \mathbf{E}_{u_o}[\exp\{\alpha(X_{\sigma_{u_o}} - u_o)\}] = c < \infty \qquad \text{(by SSDI)},
\end{aligned}$$

for all $u_o \leq u$. ∎

Corollary 4.2 *Theorems 4.1 and 4.2 still holds if we replace condition (ii) by*

(ii') V is log-Lipschitz and the chain has bounded increments.

Proof: Since V is log-Lipschitz and the chain has bounded increments it follows

$$\left|\log \frac{V(X_{\sigma_u})}{V(u)}\right| \leq k|X_{\sigma_u} - u| \leq K, \tag{4.19}$$

for some $0 < k, K < \infty$. Therefore, condition (ii) of the previous theorem can be bounded as follows

$$\frac{\mathbf{E}[V(X_{\sigma_u})]}{(1-\lambda)V(u)} \leq \frac{e^K}{1-\lambda} < \infty, \tag{4.20}$$

and the result follows. ∎

5 Some Examples

5.1 Random walk Metropolis-Hastings algorithm on a standard exponential density

Consider the random walk Metropolis-Hastings (RWMH) algorithm on a standard exponential density, $\pi(x) = e^{-x}$ for $x > 0$. Let the proposal density $dP(x,y)$ be Uniform on the interval $(x-1, x+1)$.

Then we will prove the following result.

Theorem 5.1 *Consider $\{X_n\}$ the random walk induced by the Metropolis-Hastings algorithm on a standard exponential density. Then the extremal index of the chain is positive ($\theta > 0$).*

Proof: If we take $V(x) = \pi(x)^{-\alpha} = e^{\alpha x}$, for $\alpha, x > 0$, the process is geometric ergodic. Notice that $PV(x) = \frac{e^{\alpha}-e^{-\alpha}}{2\alpha}V(x)$. Therefore, $\lambda = \frac{e^{\alpha}-e^{-\alpha}}{2\alpha}$ and $b \geq 0$ will satisfy the drift condition $PV \leq \lambda V + b1_C$ for any set C.

Condition (i) of Theorem 4.2 is satisfy for $u_n = \log \frac{n}{\tau}$, i.e. $n(1-\pi(u_n)) = ne^{-u_n} = \tau$.

For condition (ii) we will use Corollary 4.1. Let $x < y$, then $|\log V(y) - \log V(x)| = \alpha|y-x|$ which means that the function V is log-Lipschitz. On the other hand we have for $x_1 < x_2$, $\frac{dP(x_2,y)}{dP(x_1,y)} = 1$ for $x_2 - 1 < y < x1 + 1$. Therefore the process $\{X_n\}$ is strongly stochastic monotone. The last condition such that (ii) holds in Corollary 4.1 is that $\{X_n\}$has strongly stochastic decreasing increments. This is true since for $x_1 < x_2$ and $-1 < y < 1$ we obtain, $\frac{dP(x_2,y+x_2)}{dP(x_1,y+x_1)} = 1$. we can now conclude that condition (ii) of Theorem 4.2 is satisfied.

For the last condition in Theorem 4.2 to be satisfied we need to have $\frac{r_n}{V(u_n)}$ bounded. Since $r_n = \left\lfloor \frac{n}{k_n} \right\rfloor$, we have $r_n \leq \max\left(n\alpha_{n,l_n}^{1/2}, n^{1/2}\, l_n^{1/2}\right)$. From the proof of Lemma 4.6 we have that $\alpha_{n,l_n} \leq c\rho^{l_n}$ for some $c < \infty$ and $\rho < 1$. Taking $l_n = \log n$ we obtain

$$\frac{r_n}{V(u_n)} \leq \max\left(\frac{nc^{1/2}\rho^{\frac{1}{2}\log n}}{\tau^{-\alpha}n^{\alpha}}, \frac{n^{1/2}(\log n)^{1/2}}{\tau^{-\alpha}n^{\alpha}}\right) \tag{5.1}$$

The first term in (5.1) goes to zero for $\alpha > 1 + \frac{1}{2}\log\rho$ and the second term also converges to zero for $\alpha > \frac{1}{2}$. Therefore, for $\alpha > \max\left(\frac{1}{2}, 1 + \frac{1}{2}\log\rho\right)$ we have $\frac{r_n}{V(u_n)} \to 0$ as $n \to \infty$.

From Theorem 4.2 we have that the extremal index of the RWMH on a standard exponential density is positive ($\theta > 0$). ■

5.2 Gaussian AR Example

We shall consider the following simple Gaussian AR model.

Theorem 5.2 *Let $\{X_k\}$ be a Markov chain defined on the one-dimensional real line by*

$$\mathcal{L}(X_k|\, X_{k-1} = x) = N\left(\frac{x}{2}, \frac{3}{4}\right),$$

where $N(\mu, \sigma^2)$ is a normal distribution with mean μ and variance σ^2. Then the extremal index of this chain is positive.

Letting $C = [-\sqrt{10}, \sqrt{10}]$, the process is geometric ergodic with the drift function defined as $V(x) = \pi(x)^{-\alpha}$, $\lambda = \frac{1}{\sqrt{1-\alpha\sigma^2}} e^{5\alpha\left(\frac{1}{4(1-\alpha\sigma^2)}-1\right)}$ and $b = \frac{(2\pi)^{\frac{\alpha}{2}}}{\sqrt{1-\alpha\sigma^2}} \left[e^{\frac{5\alpha}{4-3\alpha}} - e^{\frac{15\alpha(\alpha-1)}{4-3\alpha}} \right]$ (see [12]). In this example σ^2 is equal to $\frac{3}{4}$. Writing $X_k = \frac{X_{k-1}}{2} + \sqrt{\frac{3}{4}} Z_k$ with $\{Z_K\}$ a sequence of i.i.d. $N(0,1)$ random variables, it is clear that if the distribution of X_{k-1} is $N(0,1)$, so is the distribution of X_k. In other words the stationary distribution of this process $\pi(.) = \Phi(.)$ the standard normal distribution. In results that follow, we will assume these choices for the drift function and the stationary distribution.

Lemma 5.1 *Under stationarity, condition (i) of Theorem 4.2 holds with $u_n = \Phi^{-1}\left(1 - \frac{\tau}{n}\right)$ for the process defined in Theorem 5.2.*

Proof: Taking $u_n = \Phi^{-1}\left(1 - \frac{\tau}{n}\right)$ we obtain $n(1 - F(u_n)) = n\left(1 - \Phi\left(\Phi^{-1}\left(1 - \frac{\tau}{n}\right)\right)\right) = \tau$. ∎

We will need the following auxiliary results.

Lemma 5.2 *The Markov chain $\{X_k\}$ defined in Theorem 5.2 is strongly stochastic monotone.*

Proof: For $x_1 < x_2$ we have

$$\frac{dP(x_2, y)}{dP(x_1, y)} = \frac{e^{-\frac{1}{2\sigma^2}(y-x_2)^2}}{e^{-\frac{1}{2\sigma^2}(y-x_1)^2}} = e^{\frac{1}{2\sigma^2}[2(x_2-x_1)y + (x_1^2 - x_2^2)]},$$

which is an increasing function with y. ∎

Lemma 5.3 *Consider the Markov chain $\{X_k\}$ defined in Theorem 5.2. Then*

$$\mathbf{E}\left(V(X_{\sigma_{u_n}})\right) \leq \frac{8\sigma^2 (2\pi)^{\alpha/2}}{1 - 2\alpha\sigma^2} \frac{e^{\frac{\alpha}{2}u_n^2}}{u_n(\sqrt{16\,\sigma^2 + u_n^2} - u_n)}$$

for u_n sufficiently large and $\alpha < \frac{2}{3}$.

Proof: Consider $\mathbf{E}\left(V(X_{\sigma_{u_n}})\right) = (2\pi)^{\alpha/2} \mathbf{E}\left[e^{\frac{\alpha}{2} X_{\sigma_{u_n}}}\right]$. Now,

$$\mathbf{E}\left[e^{\frac{\alpha}{2} X_{\sigma_{u_n}}}\right] = \max_{X_{k-1} < u_n} \mathbf{E}\left[e^{\frac{\alpha}{2} X_k} |\, X_{k-1}, X_k \geq u_n\right].$$

By strongly stochastic monotonicity, we obtain that

$$\mathbf{E}\left[e^{\frac{\alpha}{2} X_{\sigma_{u_n}}}\right] \leq \mathbf{E}\left[e^{\frac{\alpha}{2} X_k} |\, X_{k-1} = u_n, X_k \geq u_n\right].$$

The probability density function of the random variable $\{X_k |\, X_{k-1} = u_n,\ X_k \geq u_n\}$ is given by

$$f_{X_k|\, X_{k-1}=u_n, X_k \geq u_n}(y) = \frac{f(y)}{1 - F(u_n)} \mathbf{1}_{\{y \geq u_n\}},$$

with f and F, the pdf and the cdf of a $N\left(\frac{u_n}{2}, \sigma^2\right)$ distribution with $\sigma^2 = \frac{3}{4}$, respectively. Therefore,

$$
\begin{aligned}
\mathbf{E}\left[e^{\frac{\alpha}{2}X_k} | X_{k-1}=u_n, X_k \geq u_n\right] &\leq \frac{1}{1-\Phi\left(\frac{u_n - u_n/2}{\sigma}\right)} \int_{u_n}^{\infty} e^{\frac{\alpha}{2}y^2} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-u_n/2)^2}{2\sigma^2}} dy \\
&= \frac{1}{1-\Phi\left(\frac{u_n}{2\sigma}\right)} \frac{1}{\sqrt{1-\alpha\sigma^2}} \exp\left\{\frac{\alpha\, u_n^2}{8(1-\alpha\sigma^2)}\right\} \int_{u_n}^{\infty} \frac{1}{\sqrt{2\pi}\frac{\sigma}{\sqrt{1-\alpha\sigma^2}}} e^{-\frac{1}{2\frac{\sigma^2}{1-\alpha\sigma^2}}\left(y - \frac{u_n}{2(1-\alpha\sigma^2)}\right)} dy \\
&= \frac{1}{\sqrt{1-\alpha\sigma^2}} \exp\left\{\frac{\alpha\, u_n^2}{8(1-\alpha\sigma^2)}\right\} \frac{1-\Phi\left(\frac{u_n - \frac{u_n}{2(1-\alpha\sigma^2)}}{\frac{\sigma}{\sqrt{1-\alpha\sigma^2}}}\right)}{1-\Phi\left(\frac{u_n}{2\sigma}\right)} \\
&= \frac{1}{\sqrt{1-\alpha\sigma^2}} \exp\left\{\frac{\alpha\, u_n^2}{8(1-\alpha\sigma^2)}\right\} \frac{1-\Phi\left(\frac{1-2\alpha\sigma^2}{2\sigma\sqrt{1-\alpha\sigma^2}}\, u_n\right)}{1-\Phi\left(\frac{u_n}{2\sigma}\right)},
\end{aligned}
$$

for $1-\alpha\sigma^2 > 0$ and $1-2\alpha\sigma^2 > 0 \Rightarrow \alpha < \frac{2}{3}$.

From Abramowitz and Stegun [1] we have that $1-\Phi(x) \leq \frac{\phi(x)}{x}$ for $x \geq 2.2$, and $1-\Phi(x) \geq \frac{\sqrt{4+x^2}-x}{2}\phi(x)$ for $x > 1.4$. Therefore, for sufficiently large u_n it follows

$$
\begin{aligned}
\mathbf{E}\left[e^{\frac{\alpha}{2}X_k} | X_{k-1}=u_n, X_k \geq u_n\right] &\leq \frac{1}{\sqrt{1-\alpha\sigma^2}} \exp\left\{\frac{\alpha\, u_n^2}{8(1-\alpha\sigma^2)}\right\} \frac{\frac{\exp\left\{-\frac{1}{2}\frac{(1-2\alpha\sigma^2)^2}{4\sigma^2(1-\alpha\sigma^2)}u_n^2\right\}}{\frac{1-2\alpha\sigma^2}{2\sigma\sqrt{1-\alpha\sigma^2}}u_n}}{\frac{\sqrt{16\sigma^2+u_n^2}-u_n}{4\sigma}\exp\left\{-\frac{1}{2}\frac{u_n^2}{4\sigma^2}\right\}} \\
&= \frac{8\sigma^2}{1-2\alpha\sigma^2} \frac{e^{\frac{\alpha}{2}u_n^2}}{u_n(\sqrt{16\sigma^2+u_n^2}-u_n)}.
\end{aligned}
$$

Therefore, we have for sufficiently large u_n and $\alpha < \frac{2}{3}$ that

$$
\mathbf{E}\left(V(X_{\sigma_{u_n}})\right) \leq \frac{8\sigma^2(2\pi)^{\alpha/2}}{1-2\alpha\sigma^2} \frac{e^{\frac{\alpha}{2}u_n^2}}{u_n(\sqrt{16\sigma^2+u_n^2}-u_n)}.
$$

■

Lemma 5.4 *Consider the Markov chain $\{X_k\}$ defined in Theorem 5.2. Then, for u_n any real number sequence such that $u_n \to \infty$ as $n \to \infty$ and $\alpha < \frac{2}{3}$, condition (ii) of Theorem 4.2 holds, i.e.*

$$
\frac{\mathbf{E}\left(V(X_{\sigma_{u_n}})\right)}{V(u_n)} < \infty.
$$

Proof: From the previous lemma it follows:

$$
\begin{aligned}
\frac{\mathbf{E}\left(V(X_{\sigma_{u_n}})\right)}{V(u_n)} &\leq \frac{\frac{8\sigma^2(2\pi)^{\alpha/2}}{1-2\alpha\sigma^2}\frac{e^{\frac{\alpha}{2}u_n^2}}{u_n(\sqrt{16\sigma^2+u_n^2}-u_n)}}{(2\pi)^{\alpha/2}e^{\frac{\alpha}{2}u_n^2}} \\
&= \frac{1}{2(1-2\alpha\sigma^2)} \frac{\sqrt{16\sigma^2+u_n^2}+u_n}{u_n}.
\end{aligned}
$$

Since $u_n \to \infty$ as $n \to \infty$, we obtain from above $\frac{\mathbf{E}\left(V(X_{\sigma_{u_n}})\right)}{V(u_n)} \to \frac{1}{1-2\alpha\sigma^2}$ as $n \to \infty$. ■

Lemma 5.5 *Condition (iii) of Theorem 4.2 holds for the Markov chain defined in Theorem 5.2 with u_n any real number sequence such that $u_n \to \infty$ as $n \to \infty$ and* $\max\left(\frac{1}{2}, 1 + \frac{1}{2}\log\rho\right) < \alpha < 1$.

Proof: Identically to the previous example we have $r_n \le \max\left(nc^{1/2}\rho^{\frac{1}{2}\log n}, n^{1/2}(\log n)^{1/2}\right)$, with $c < \infty$ and $\rho < 1$. Hence,

$$\frac{r_n}{V(u_n)} \le \max\left(\frac{c^{\frac{1}{2}}}{(2\pi)^{\alpha/2}}\frac{n\rho^{\frac{1}{2}\log n}}{e^{\frac{\alpha}{2}u_n^2}}, \frac{1}{(2\pi)^{\alpha/2}}\frac{n^{1/2}(\log n)^{1/2}}{e^{\frac{\alpha}{2}u_n^2}}\right). \tag{5.2}$$

Recall that $u_n = \Phi^{-1}\left(1 - \frac{\tau}{n}\right)$ and let $\tau = e^{-x}$ with $x \in \mathbb{R}$. Then

$$\frac{u_n^2}{2} = x + \log n - \frac{1}{2}\log(4\pi) - \frac{1}{2}\log\log n + o(1).$$

Hence, the first term in (5.2) goes to zero if $\alpha > 1 + \frac{1}{2}\log\rho$, as can be seen below.

$$\begin{aligned} &\lim_{n\to\infty}\frac{n\rho^{\frac{1}{2}\log n}}{e^{\frac{\alpha}{2}u_n^2}} = \frac{n\rho^{\frac{1}{2}\log n}}{e^{\alpha x}(4\pi)^{-\alpha/2}e^{\alpha\, o(1)}n^{\alpha}(\log n)^{-\alpha/2}} \\ \propto\quad &\lim_{n\to\infty}\frac{n\rho^{\frac{1}{2}\log n}}{n^{\alpha}(\log n)^{-\alpha/2}} = \exp\left\{\left(1-\alpha+\frac{1}{2}\log\rho\right)\log n + \frac{\alpha}{2}\log\log n\right\} = 0. \end{aligned}$$

In what follows, we see that the second term in (5.2) also goes to zero if $\alpha - \frac{1}{2} > \frac{\alpha-1}{2} \Leftrightarrow \alpha > 0$.

$$\begin{aligned} &\lim_{n\to\infty}\frac{n^{1/2}(\log n)^{1/2}}{e^{\frac{\alpha}{2}u_n^2}} = \frac{n^{1/2}(\log n)^{1/2}}{e^{\alpha x}(4\pi)^{-\alpha/2}e^{\alpha\, o(1)}n^{\alpha}(\log n)^{-\alpha/2}} \\ \propto\quad &\lim_{n\to\infty}\frac{(\log n)^{\frac{\alpha+1}{2}}}{n^{\alpha-\frac{1}{2}}} = \lim_{n\to\infty}\frac{\frac{1}{n}(\log n)^{\frac{\alpha-1}{2}}}{\left(\alpha-\frac{1}{2}\right)n^{\alpha-\frac{1}{2}-1}} = 0. \end{aligned}$$

Therefore we conclude that $\frac{r_n}{V(u_n)} \to 0$ as $n \to \infty$ if $\max\left(0, 1 + \frac{1}{2}\log\rho\right) < \alpha < 1$. ■

Let's now prove Theorem 5.2.

Proof: (Of Theorem 5.2) In the beginning of the example, we chose $\lambda = \frac{1}{\sqrt{1-\alpha\sigma^2}}e^{5\alpha\left(\frac{1}{4(1-\alpha\sigma^2)}-1\right)}$ and $b = \frac{(2\pi)^{\frac{\alpha}{2}}}{\sqrt{1-\alpha\sigma^2}}\left[e^{\frac{5\alpha}{4-3\alpha}} - e^{\frac{15\alpha(\alpha-1)}{4-3\alpha}}\right]$. We also proved in Lemma 5.2 that the Markov chain here defined is strongly stochastic monotone. Hence, from Theorem 4.1 (ii) in [9], we have $\|P^n(x,.) - \pi\|_V \le$

$\frac{2}{1-\rho^{-1}\lambda}\left[V(x)+\frac{b}{1-\lambda}\right]\rho^n$, for any $\rho \geq \lambda$. Therefore, choosing $\alpha = 0.6$, we have $\lambda = 0.2626$ and $b = 1.5626$. Taking $\rho = 0.3$, Lemmas 5.1, 5.4 and 5.5 hold, and we conclude that conditions $(i) - (iii)$ of Theorem 4.2 are satisfied. Hence, the extremal index for this Markov chain is positive ($\theta > 0$). ∎

Acknowledgments: We thank Jon Tawn and Neil Madras for helpful discussions.

References

[1] Abramowitz, M. and Stegun, I. A. (1986), *Handbook of Mathematical Functions*, Dover.

[2] Davis, R.A. (1982), Limit Laws for the Maximum and Minimum of Stationary Sequences, *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, **61**, 31-42.

[3] Fisher, R.A. and Tippett, L.H.C.(1928), Limiting forms of the frequency distribution of the largest or smallest of a sample, *Proc. Cambridge Philos. Soc.*, **24**, 180-190.

[4] Gnedenko, B.V. (1943), Sur la distribution limite du terme maximum d'une séeri aléatoire, *Ann. Math.*, **44**, 423-453.

[5] Gut, Allan (1983), Complete Convergence and Convergence Rates for Randomly Indexed Partial Sums with an Application to some First Passage Times, *Acta Math. Hung.*, **42**, 225-232.

[6] Gut, Allan (1988), *Stopped Random Walks, Limit Theorems and Applications*, Springer-Verlag, London.

[7] Leadbetter, M.R. (1983), Extremes and Local Dependence in Stationary Sequences, *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, **65**, 291-306.

[8] Leadbetter, M.R., Lindgren, G. and Rootzén (1983), *Extremes and Related Properties of Random Sequences and Processes*, Springer-Verlag, New York.

[9] Mengersen, K.L. and Tweedie, R.L. (1996), Rates of Convergence of the Hastings and Metropolis Algorithms, *The Annals of Statistics*, **24**, 101-121.

[10] Meyn, S.P. and Tweedie, R.L. (1993), *Markov chains and stochastic stability*, Springer-Verlag, London.

[11] O'Brien, G.L. (1987), Extreme Values for Stationary and Markov Sequences, *The Annals of Probability*, **15**, 281-291.

[12] Roberts, G.O. and Rosenthal, J.S. (1997), Shift-Coupling and Convergence Rates of Ergodic Averages, *Comm. in Stat. - Stoch. Models*, **13**, 147-165.

[13] Smith, R.L. (1992), The extremal index for Markov chains, *J. Appl. Prob.*, **29**, 37-45.