Modern Stochastics: Theory and Applications 0 (0) (2025) 1–18
https://doi.org/10.15559/25-VMSTA277

Exit times for some nonlinear autoregressive processes

Göran Högnäs, Brita Jung

Faculty of Science and Engineering, Åbo Akademi, FIN-20500 Åbo, Finland

ghognas@abo.fi (G. Högnäs), brita.jung@abo.fi (B. Jung)

Received: 29 August 2024, Revised: 10 January 2025, Accepted: 9 March 2025,
Published online: 25 March 2025

Abstract The expected exit time from the interval $[-1, 1]$ is investigated for an autoregressive process defined recursively by

$$X^{\varepsilon}_{n+1} = f\big(X^{\varepsilon}_n\big) + \varepsilon\xi_{n+1}, \quad n = 0, 1, 2, \dots, \; X_0 = 0.$$

Here, ε is a small positive parameter, $f : \mathbb{R} \mapsto \mathbb{R}$ is usually a contractive function and $\{\xi_n\}_{n\geq 1}$ is a sequence of i.i.d. random variables. In this paper, previous results for a linear function $f(x) = ax$ are extended to more general cases, with the main focus on piecewise linear functions.

Keywords Exit time, first passage time, autoregressive process, large deviation principle
2020 MSC 60G17, 60F10

1 Introduction

Consider a stochastic process $\{X^{\varepsilon}_n\}_{n=0}^{\infty}$ of autoregressive type, defined by

$$X^{\varepsilon}_{n+1} = f\left(X^{\varepsilon}_n\right) + \varepsilon\xi_{n+1}, \quad X_0 = 0, \tag{1}$$

where f is a continuous mapping from $\mathbb{R}$ to itself with a fixed point at the origin, $\{\xi_n\}_{n=1}^{\infty}$ is a sequence of i.i.d. random variables and ε is a small positive parameter. It is a Markov chain, and under suitable assumptions on $\{\xi_n\}_{n=1}^{\infty}$ and f, it is interesting to investigate how long it takes for the process to leave a neighbourhood of the origin.

The original motivation behind our work is to study the time until extinction of a population. A stochastic process that models a population may be positive recurrent and stay at a certain level (or carrying capacity) for a very long time, and when extinction happens, the process first leaves a neighbourhood around that level. Populations can be modeled by, for example, branching processes (such as those treated in [5] and [6]) or models such as the Ricker model which we will use as an example at the end of the paper; it has been studied in [9].

In [12] and the updated version [13], Klebaner and Liptser used the large deviation principle to get an upper bound on the exit time from a set for a process. As an example, they considered the linear autoregressive process, defined as in (1) with $f(x) = ax$, $|a| < 1$, and normally distributed innovations. For this example, they showed that the exit time τ_ε from the interval $(-1, 1)$ satisfies

$$\limsup_{\varepsilon\to 0} \varepsilon^2 \log E_{x_0}(\tau_\varepsilon) \leq \frac{1-a^2}{2}. \tag{2}$$

The corresponding lower bound has been found by other methods [18]. This means that the upper bound on the right-hand side above is the best possible one.

We have studied a corresponding multivariate case [11], and the results were extended to the ARMA model in [15]. Exit times for autoregressive processes with other noise distributions have also been studied, in [8]. Different aspects of the exit time problem for linear autoregressive processes are treated in, for example, [2–4] and [17]. A related exit problem in a different setting was treated in [14]. A case with piecewise linear function $f(x)$ was studied by Anděl et al. in [1].

In this paper we extend the previous results from the linear case to some cases with piecewise linear functions. We will show that the large deviation principle gives explicit (asymptotic) upper bounds on the expectation of the exit time also in these cases. We also apply the methods used to other nonlinear functions, such as quadratic functions and the Ricker model.

In Section 2, we summarize how the large deviation method results in a sum that should be minimized for an upper bound of the exit time. This is based on the methods used in the proofs of Theorems 2.2 and 3.1 in [13]. In Section 3 we study how the minimization can be done over more restricted sets. In Section 4, which contains the main results of this paper, we get the explicit upper bounds in several piecewise linear cases. In Section 5 we explore some other nonlinear cases. In Section 6 we point out a connection between the results and the stationary distribution of the process.

2 Large deviation tools and bounds for exit times

In this section, which is based on work by Klebaner and Liptser in [12] and Jung in [11], we summarize how the large deviation principle (LDP) can be used to get an upper bound of the asymptotics of an exit time from a set for a process. The definition of the LDP used in [12] is as follows (this follows Varadhan's definition in [19], with the addition that the rate of speed $q(\varepsilon)$ is a function of ε such that $q(\varepsilon) \to 0$ as $\varepsilon \to 0$).

Let $\{P_\varepsilon\}$ be a family of probability measures on the Borel subsets of a complete separable metric space Z. The family $\{P_\varepsilon\}$ satisfies the large deviation principle with a rate function I if there is a function I from Z into $[0, \infty]$ that satisfies the following conditions: $0 \leq I(z) \leq \infty \; \forall z \in Z$, I is lower semicontinuous, the set $\{z : I(z) \leq l\}$

is a compact set in Z for all $l < \infty$ and

$$\limsup_{\varepsilon\to 0} q(\varepsilon)\log P_\varepsilon(C) \leq -\inf_{z\in C} I(z) \quad \text{for every closed set } C \subset Z \text{ and}$$
$$\liminf_{\varepsilon\to 0} q(\varepsilon)\log P_\varepsilon(G) \geq -\inf_{z\in G} I(z) \quad \text{for every open set } G \subset Z.$$

In [12] and [13], Klebaner and Liptser considered a family of processes of the type

$$X^\varepsilon_{n+1} = g\big(X^\varepsilon_n, \dots, X^\varepsilon_{n-m+1}, \varepsilon\xi_{n+1}\big), \tag{3}$$

where g is a continuous function on $\mathbb{R}^m$, $\{\xi_n\}_{n=m}^\infty$ is a sequence of i.i.d. random variables and $x_0, \dots, x_{m-1}$ are given starting values. They gave conditions under which the LDP holds for the family $\varepsilon\xi$ (where ξ is a copy of ξ_m) and proved that when $\varepsilon\xi$ obeys an LDP with rate function $I(z)$, it follows that (X^ε_n) obeys an LDP with a rate function $J(\bar{y})$ that can be written explicitly using $I(z)$:

$$J(\bar{y}) = \sum_{k=m}^{\infty} \inf_{v_k : y_k = f(y_{k-1},\dots y_{k-m}, v_k)} I(y_k) \quad \text{when } u_0 = x_0, \dots, u_{m-1} = x_{m-1}, \tag{4}$$

and $J(\bar{y}) = \infty$ otherwise.

Klebaner and Liptser also showed in [12] and [13] how the LDP can be used to get bounds of the asymptotics of the expected exit time of the process. Let the exit time τ_ε of the process be defined as

$$\tau_\varepsilon := \min\big\{t \geq m : X^\varepsilon_n \notin \Omega\big\} \tag{5}$$

for a set Ω. For the expected exit time it holds that

$$E_{x_0,\dots,x_{m-1}}(\tau_\varepsilon) \leq \frac{2M}{\inf_{x_0,\dots,x_{m-1}\in\Omega} P_{x_0,\dots,x_{m-1}}(\tau_\varepsilon \leq M)}$$

for any set of starting points $x_0, \dots, x_{m-1} \in \Omega$ and any integer $M \geq m$ (for details, see [11]). If the infimum in the denominator is attained for the starting points $x^*_0, \dots, x^*_{m-1} \in \Omega$, the inequality above implies that

$$\limsup_{\varepsilon\to 0} q(\varepsilon)\log E_{x_0,\dots,x_{m-1}}(\tau_\varepsilon) \leq -\lim_{\varepsilon\to 0} q(\varepsilon)\log P_{x^*_0,\dots,x^*_{m-1}}(\tau_\varepsilon \leq M), \tag{6}$$

if the right-hand side limit exists. Since

$$P_{x^*_0,\dots,x^*_{m-1}}(\tau_\varepsilon \leq M) = P_{x^*_0,\dots,x^*_{m-1}}\big(X^\varepsilon_t \notin \Omega \text{ for some } t \in \{m, \dots, M\}\big),$$

the limit on the right-hand side in (6) may be calculated if we have a large deviation principle for the family of probability measures induced by $\{X^\varepsilon_t\}_{t\geq 0}$ and if the function f and the set Ω are suitable.

From this point onward in the paper, we consider a process of autoregressive type, where

$$X^\varepsilon_{n+1} = f\big(X^\varepsilon_n\big) + \varepsilon\xi_{n+1}, \quad n = 0, 1, 2, \dots, \tag{7}$$

and $X_0 = 0$. Here f is a continuous function on $\mathbb{R}$ and $\{\xi_n\}_{n\geq 1}$ is a sequence of i.i.d. standard normal random variables. Then $I(z) = \frac{z^2}{2}$ and the function g in (3) is reduced to

$$g(y_{n-1}, \ldots, y_{n-m+1}, z_n) = f(y_{n-1}) + z_n.$$

We consider exit times from the interval $(-1, 1)$, so $\Omega = (-1, 1)$.

Klebaner and Liptser considered this case as an example in [12] and [13] (other examples were Poisson distributed noise, and sums of normally distributed and Poisson distributed random variables), and showed that this family of processes obeys the large deviation principle with $q(\varepsilon) = \varepsilon^2$ and

$$I(y_0, y_1, y_2, \ldots) = \frac{1}{2}\sum_{n=1}^{\infty}\big(y_n - f(y_{n-1})\big)^2. \tag{8}$$

For the exit time

$$\tau_\varepsilon = \min\big\{n \geq 1 : |X_n^\varepsilon| \geq 1\big\}, \tag{9}$$

we then have

$$\begin{aligned}
&\limsup_{\varepsilon\to 0} \varepsilon^2 \log E_{x_0}\tau_\varepsilon \leq -\lim_{\varepsilon\to 0} \varepsilon^2 \log P_{x_0^*}(\tau_\varepsilon \leq M) \\
&= \inf_{\substack{\max_{1\leq n\leq M}|y_n|\geq 1\\ y_0 = x_0^*}} I(y_0, y_1, y_2, \ldots) = \inf_{\substack{\max_{1\leq n\leq M}|y_n|\geq 1\\ y_0 = x_0^*}} \frac{1}{2}\sum_{n=1}^{\infty}\big(y_n - f(y_{n-1})\big)^2 \\
&= \inf_{1\leq N\leq M}\left(\inf_{\substack{|y_N|\geq 1\\ y_0 = x_0^*\\ |y_n|<1, n=1,\ldots,N}} \frac{1}{2}\sum_{n=1}^{N}\big(y_n - f(y_{n-1})\big)^2\right),
\end{aligned} \tag{10}$$

where the last equality holds because one can choose $y_n = f(y_{n-1})$ for all $n \geq N+1$ and get the same infimum.

For the autoregressive case with a linear function $f(x) = ax$, $|a| < 1$, and τ_ε as above, Klebaner and Liptser showed that

$$\limsup_{\varepsilon\to 0} \varepsilon^2 \log E_{x_0}(\tau_\varepsilon) \leq \frac{1-a^2}{2} \tag{11}$$

by minimizing the sum. This was done by considering the telescoping sum

$$\sum_{n=1}^{N} a^{N-n}(y_n - ay_{n-1}) = y_N$$

when $y_0 = x_0 = 0$ and applying the Cauchy–Schwarz inequality to get

$$\sum_{n=1}^{N}(y_n - ay_{n-1})^2 \geq \frac{y_N^2}{\sum_{n=1}^{N} a^{N-n}}. \tag{12}$$

Here, $|y_N| = 1$, and the result in (11) follows since M can be arbitrarily large. Note that if we instead were to study exits from a scaled interval $(-c, c)$, $c > 0$, so that $\tau_\varepsilon^c = \min\{n \geq 1 : |X_n^\varepsilon| \geq c\}$, then $|y_N| = c$ and the upper bound would be

$$\limsup_{\varepsilon \to 0} \varepsilon^2 \log E_{x_0}\left(\tau_\varepsilon^c\right) \leq \frac{(1-a^2)c^2}{2}. \tag{13}$$

3 Minimizing the sum

In the previous section we saw that the asymptotics for the exit times of processes of the type defined in (7) with $N(0, 1)$-normal white noise is determined by the function f through the infimum of the sum of squares in (10). In this section we study some properties of these sums for particular classes of autoregression functions f. Since the infimum of the sum is attained for $|y_N| = 1$, it does not matter how the function f is defined outside of the interval $[-1, 1]$, and we only focus on how it is defined on $[-1, 1]$.

Lemma 1. *If f is increasing on $[-1, 1]$ and $f(0) = 0$, the sum can be minimized separately over positive and negative values:*

$$\inf_{\substack{|y_N|\geq 1\\ y_0=0}} \frac{1}{2}\sum_{n=1}^{N}\left(y_n - f(y_{n-1})\right)^2 \tag{14}$$

$$= \min\left(\inf_{\substack{y_N=1, y_0=0\\ y_n\geq 0, n=1,\ldots,N-1}} \frac{1}{2}\sum_{n=1}^{N}\left(y_n - f(y_{n-1})\right)^2, \right.$$

$$\left. \inf_{\substack{y_N=-1, y_0=0\\ y_n\leq 0, n=1,\ldots,N-1}} \frac{1}{2}\sum_{n=1}^{N}\left(y_n - f(y_{n-1})\right)^2\right). \tag{15}$$

Proof. Assume that the infimum of the sum is attained for $\{y_n^*\}_{n=0}^N$, where $y_0^* = 0$ and $y_N^* = 1$, and let

$$S^* = \sum_{n=1}^{N}\left(y_n^* - f\left(y_{n-1}^*\right)\right)^2.$$

Then $S^* \leq 1$. We will show by induction that $y_n^* \geq 0$ for $n = 1, \ldots, N-1$. We show first that $y_{N-1}^* \geq 0$. If $y_{N-1}^* < 0$, then $f(y_{N-1}^*) \leq 0$ and $(1 - f(y_{N-1}))^2 \geq 1$. Also,

$$S^* \geq \left(1 - f\left(y_{N-1}^*\right)\right)^2 + \left(y_L^*\right)^2,$$

where $L = \min\{i | y_i \neq 0\}$. It follows that $S^* > 1$, which is a contradiction. Thus, $y_{N-1}^* \geq 0$.

Now, assume that $y_N^*, y_{N-1}^*, \ldots, y_{N-K+1}^* \geq 0$ for some $K < N$. Make the contrary assumption that $y_{N-K}^* < 0$. Then $f(y_{N-K}^*) \leq 0$ and

$$S^* = \left(1 - f\left(y_{N-1}^*\right)\right)^2 + \cdots + \left(y_{N-K+2}^* - f\left(y_{N-K+1}^*\right)\right)^2$$

$$
\begin{aligned}
&+\big(y^*_{N-K+1}-f\big(y^*_{N-K}\big)\big)^2+\big(y^*_{N-K}-f\big(y^*_{N-K-1}\big)\big)^2+\dots+y_1^2\\
&\geq \big(1-f\big(y^*_{N-1}\big)\big)^2+\dots+\big(y^*_{N-K+2}-f\big(y^*_{N-K+1}\big)\big)^2\\
&+\big(y^*_{N-K+1}-0)\big)^2+0+\dots+0.
\end{aligned}
$$

In fact, the inequality above is strict, since

$$
\begin{aligned}
S^* \geq{}& \big(1-f\big(y^*_{N-1}\big)\big)^2+\dots+\big(y^*_{N-K+2}-f\big(y^*_{N-K+1}\big)\big)^2\\
&+\big(y^*_{N-K+1}-0)\big)^2+\big(y^*_L\big)^2
\end{aligned}
$$

where $L=\min\{i\leq N-K\,|\,y^*_i\neq 0\}$. Thus, S^* is not the minimal sum, which is a contradiction. It follows that $y^*_{N-K}\geq 0$.

If we assume instead that the infimum on the left-hand side in (14) is attained for a sequence $\{y'_n\}_{n=0}^N$, where $y'_0=0$ and $y'_N=-1$, one can show that $y'_n\leq 0$ for $n=1,\dots,N$ in a similar way. □

Lemma 2. *If f is increasing on $[-1,1]$, $f(0)=0$ and f is odd, so that $f(-x)=-f(x)$ on $[-1,1]$, we can minimize over only positive values:*

$$
\inf_{\substack{|y_N|\geq 1\\ y_0=0}}\frac{1}{2}\sum_{n=1}^{N}\big(y_n-f(y_{n-1})\big)^2=\inf_{\substack{y_N=1,y_0=0\\ y_n\geq 0,n=1,\dots,N-1}}\frac{1}{2}\sum_{n=1}^{N}\big(y_n-f(y_{n-1})\big)^2. \qquad (16)
$$

Proof. This follows immediately from Lemma 1, since the two infima on the right-hand side in (14) have the same value. □

Lemma 3. *If f is increasing, $f(0)=0$ and $|f(x)|<|x|$ on $(-1,1)\backslash\{0\}$, the sum should be minimized separately over positive values and increasing sequences or negative values and decreasing sequences:*

$$
\inf_{\substack{y_N=1,y_0=0\\ y_n\geq 0,n=1,\dots,N-1}}\frac{1}{2}\sum_{n=1}^{N}\big(y_n-f(y_{n-1})\big)^2=\inf_{\substack{y_N=1,y_0=0\\ y_0\leq y_1\leq y_2\leq\dots\leq y_N}}\frac{1}{2}\sum_{n=1}^{N}\big(y_n-f(y_{n-1})\big)^2 \qquad (17)
$$

and

$$
\inf_{\substack{y_N=-1,y_0=0\\ y_n\leq 0,n=1,\dots,N-1}}\frac{1}{2}\sum_{n=1}^{N}\big(y_n-f(y_{n-1})\big)^2=\inf_{\substack{y_N=-1,y_0=0\\ y_0\geq y_1\geq y_2\geq\dots\geq y_N}}\frac{1}{2}\sum_{n=1}^{N}\big(y_n-f(y_{n-1})\big)^2. \qquad (18)
$$

Proof. We prove equality (17). Assume that the sum $\sum_{n=1}^N(y_n-f(y_{n-1}))^2$ is minimized by the sequence $\{y^*_n\}_{n=0}^N$, where $y^*_0=0$, $y^*_N=1$ and $y^*_n\in[0,1]$ for $n=1,\dots,N-1$. We show by induction that $y^*_N\geq y^*_{N-1}\geq\dots\geq y^*_1\geq y^*_0$. It is given that $y^*_N\geq y^*_{N-1}$. Assume that $y^*_N\geq y^*_{N-1}\geq\dots\geq y^*_{N-k+1}$ for some k. We show that $y^*_{N-k+1}\geq y^*_{N-k}$.

If $y^*_{N-k+1}=0$, it is clear that the minimum of the sum is attained for $y^*_{N-k}=y^*_{N-k-1}=\dots=y^*_1=y^*_0=0$. Then $y^*_{N-k+1}\geq y^*_{N-k}$.

If $y^*_{N-k+1} > 0$, make the contrary assumption that $y^*_{N-k+1} < y^*_{N-k}$. Then $y^*_{N-k} \in (y^*_{N-m}, y^*_{N-m+1}]$ for some $m \in \{1, \ldots, k-1\}$. It follows that

$$\begin{aligned}\sum_{n=1}^{N}\big(y_n^* - f(y_{n-1}^*)\big)^2 = {}& \big(1 - f(y^*_{N-1})\big)^2 + \cdots + \big(y^*_{N-m+2} - f(y^*_{N-m+1})\big)^2\\ &+ \big(y^*_{N-m+1} - f(y^*_{N-m})\big)^2 + \big(y^*_{N-m} - f(y^*_{N-m-1})\big)^2\\ &+ \cdots + \big(y^*_{N-k+1} - f(y^*_{N-k})\big)^2 + \big(y^*_{N-k} - f(y^*_{N-k-1})\big)^2\\ &+ \cdots + \big(y^*_2 - f(y^*_1)\big)^2 + \big(y^*_1\big)^2\\ \geq{}& \big(1 - f(y^*_{N-1})\big)^2 + \cdots + \big(y^*_{N-m+2} - f(y^*_{N-m+1})\big)^2 \qquad (19)\\ &+ \big(y^*_{N-m+1} - f(y^*_{N-k})\big)^2 + 0 + \cdots + 0\\ &+ \big(y^*_{N-k} - f(y^*_{N-k-1})\big)^2 + \cdots + \big(y^*_1\big)^2,\end{aligned}$$

because $y^*_{N-m+1} - f(y^*_{N-m}) \geq y^*_{N-m+1} - f(y^*_{N-k})$. Equality is attained if

$$f\big(y^*_{N-m}\big) = f\big(y^*_{N-k}\big) \quad \text{and} \quad y^*_{N-m} = f\big(y^*_{N-m-1}\big), \ldots, y^*_{N-k+1} = f\big(y^*_{N-k}\big). \tag{20}$$

If this is true, f is constant on the interval $[y^*_{N-m}, y^*_{N-k}]$, and $f(y^*_{N-m}) = y^*_{N-k+1}$. Also,

$$f\big(y^*_{N-m}\big) = f\big(f\big(y^*_{N-m-1}\big)\big) = \cdots = f\big(f\big(\ldots\big(f\big(y^*_{N-k+1}\big)\big)\big)\big) < y^*_{N-k+1},$$

because $y^*_{N-k+1} > 0$. Thus, (20) does not hold, which implies that equality is not attained in (19). Then, the sum $\sum_{n=1}^{N}(y_n - f(y_{n-1}))^2$ is not minimized by the sequence $\{y_n^*\}_{n=0}^N$, which is a contradiction. Thus, $y^*_{N-k+1} \geq y^*_{N-k}$. □

Remark 1. If f and g are as in Lemma 3 and $|f(x)| \leq |g(x)|$ on $[-1, 1]$, then

$$\inf_{\substack{|y_N|\geq 1\\ y_0=0}} \frac{1}{2}\sum_{n=1}^{N}\big(y_n - g(y_{n-1})\big)^2 \leq \inf_{\substack{|y_N|\geq 1\\ y_0=0}} \frac{1}{2}\sum_{n=1}^{N}\big(y_n - f(y_{n-1})\big)^2. \tag{21}$$

Proof. By Lemma 3, the minimum on the right-hand side is attained for an increasing sequence $\{y_n\}_{n=0}^N$:

$$\inf_{\substack{y_N=1, y_0=0\\ y_n\geq 0, n=1,\ldots,N-1}} \frac{1}{2}\sum_{n=1}^{N}\big(y_n - f(y_{n-1})\big)^2 = \inf_{\substack{y_N=1, y_0=0\\ y_0\leq y_1\leq\cdots\leq y_N}} \frac{1}{2}\sum_{n=1}^{N}\big(y_n - f(y_{n-1})\big)^2.$$

Now, for each $n = 1, \ldots, N$,

$$y_n - f(y_{n-1}) \geq y_n - y_{n-1} \geq 0,$$

and the same is true when f is replaced by g. Since $f(x) \leq g(x)$ on $[0, 1]$,

$$y_n - f(y_{n-1}) \geq y_n - g(y_{n-1}),$$

and it follows that

$$\big(y_n - f(y_{n-1})\big)^2 \geq \big(y_n - g(y_{n-1})\big)^2.$$

The statement of the lemma then follows. □

4 Piecewise linear functions

Consider the process

$$X_{n+1}^{\varepsilon} = f\left(X_n^{\varepsilon}\right) + \varepsilon \xi_{n+1}, \tag{22}$$

where $X_0 = 0$, $\{\xi_n\}_{n \geq 1}$ is a sequence of i.i.d. standard normal random variables, ε is a small positive parameter and f is a continuous piecewise linear function. We consider exit times from the interval $(-1, 1)$, so $\Omega = (-1, 1)$ and

$$\tau_{\varepsilon} = \min\left\{n \geq 1 : |X_n^{\varepsilon}| \geq 1\right\}. \tag{23}$$

As in Section 3, the definition of f outside of $[-1, 1]$ does not have an impact on the results in this section.

Proposition 1. *Let f be a function on $\mathbb{R}$ that satisfies*

$$f(x) = \begin{cases} a(x+b) & \text{if } -1 \leq x \leq -b, \\ 0 & \text{if } -b < x < b, \\ a(x-b) & \text{if } b \leq x \leq 1, \end{cases}$$

where $0 \leq a < 1$ and $0 \leq b < 1$. Then, for any $M \geq 1$,

$$\limsup_{\varepsilon \to 0} \varepsilon^2 \log E_{x_0} \tau_{\varepsilon} \leq \frac{1}{2} \min\left(1, \inf_{2 \leq N \leq M} \left(\left(1 - a^2\right) \frac{(1 + \frac{a - a^N}{1-a} b)^2}{1 - a^{2N}}\right)\right). \tag{24}$$

If $a = 1$ and $0 \leq b < 1$,

$$\limsup_{\varepsilon \to 0} \varepsilon^2 \log E_{x_0} \tau_{\varepsilon} \leq \frac{1}{2} \inf_{1 \leq N \leq M} \left(\frac{(1 + (N-1)b)^2}{N}\right) \tag{25}$$

for any $M \geq 1$.

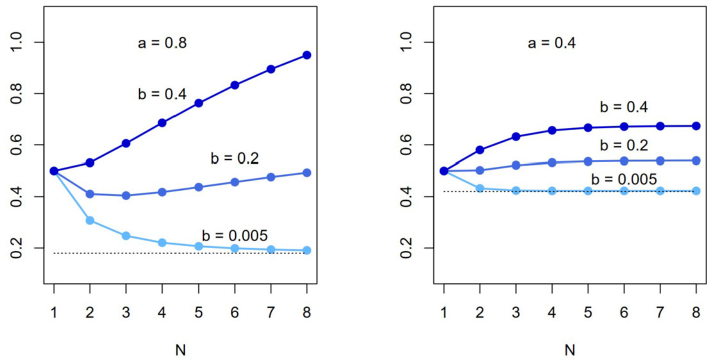

Fig. 1. The upper bound on the right-hand side in (24) illustrated for different values of a and b and $N = 1$ (when the value is 1/2) and $N = 2, \ldots 8$. The dotted line shows the value $(1-a^2)/2$ in each case

Proof. We have

$$\inf_{\substack{|y_N|\geq 1\\ y_0=0}} \sum_{n=1}^{N} \big(y_n - f(y_{n-1})\big)^2 = \inf_{\substack{|y_N|=1\\ |y_0|=b}} \sum_{n=1}^{N} \big(y_n - f(y_{n-1})\big)^2$$
$$= \inf_{\substack{y_N=1\\ y_0=b}} \sum_{n=1}^{N} \big(y_n - f(y_{n-1})\big)^2,$$

because $f(0) = f(b)$ and it is enough to take the infimum over positive values because of Lemma 2. If $y_N = 1$, $y_0 = b$ and $c_n = a^{N-n}$ for $n = 1, \dots, N$, we have the following telescoping sum:

$$\sum_{n=1}^{N} c_n\big(y_n - f(y_{n-1})\big) = \begin{cases} 1 & \text{if } N = 1, \\ 1 + ab\frac{1-a^{N-1}}{1-a} & \text{if } N \geq 2. \end{cases}$$

By the Cauchy–Schwarz inequality,

$$\sum_{n=1}^{N} \big(y_n - f(y_{n-1})\big)^2 \geq \frac{\sum_{n=1}^{N} c_n(y_n - f(y_{n-1}))}{(\sum_{n=1}^{N} c_n^2)},$$

where equality can be attained. Since $\sum_{n=1}^{N} c_n^2 = (1-a^{2N})/(1-a^2)$, it follows that

$$\limsup_{\varepsilon\to 0} \varepsilon^2 \log E_{x_0}\tau_\varepsilon \leq \frac{1}{2}\min\left(1, \inf_{2\leq N\leq M}\left((1-a^2)\frac{(1+\frac{a-a^N}{1-a}b)^2}{1-a^{2N}}\right)\right),$$

for any $M \geq 1$. The value of the infimum, as well as for which N it is attained, depends on the choices of a and b; in some cases the minimum is 1 and in some cases it it less than one.

If we let $a = 1$ and $0 < b < 1$, we can use the same method as above with the telescoping sum. Then $c_n = a^{N-n} = 1$ for $n = 1, \dots, N$,

$$\sum_{n=1}^{N} c_n\big(y_n - f(y_{n-1})\big) = 1 + (N-1)b$$

and the result is

$$\limsup_{\varepsilon\to 0} \varepsilon^2 \log E_{x_0}\tau_\varepsilon \leq \frac{1}{2}\inf_{1\leq N\leq M} \frac{(1+(N-1)b)^2}{N}. \tag{26}$$

Here, the infimum is 1 (which is attained for $N = 1$) if $b \geq 1/3$. For $b < 1/3$, the optimal N is either $\lfloor \frac{1}{b} \rfloor - 1$ or $\lfloor \frac{1}{b} \rfloor$. □

Note that if $b = 0$ in expression (24), the infimum is attained for $N = M$. The inequality holds for any $M \geq 1$ and also as $M \to \infty$. Then the right-hand side in (24) becomes $(1-a^2)/2$, the result for the autoregressive process (11).

Proposition 2. *Let $a \in (0,1)$ and $c \in (0,1]$ and let f be a function that satisfies*

$$f(x) = \begin{cases} -ac & \text{if } -1 \le x < -c, \\ ax & \text{if } -c \le x \le c, \\ ac & \text{if } c < x \le 1. \end{cases}$$

Then

$$\limsup_{\varepsilon \to 0} \varepsilon^2 \log E\tau_\varepsilon \le \begin{cases} \frac{1}{2}((1-ac)^2 + (1-a^2)c^2), & \text{if } c \le a, \\ \frac{1}{2}(1-a^2), & \text{if } c \ge a. \end{cases} \tag{27}$$

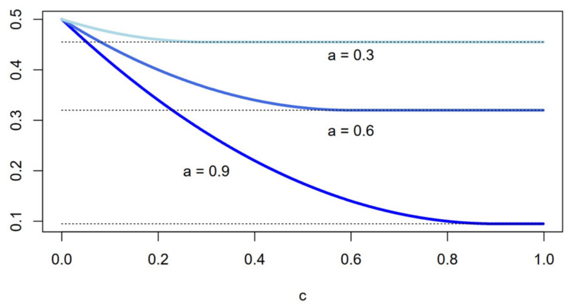

Fig. 2. The upper bound on the right-hand side in (27) drawn as a function of c for some chosen values of a. The dotted lines show the value $(1-a^2)/2$ for each a

Proof. By Lemma 2 and Lemma 3 we only need to minimize over positive and increasing sequences. We determine the infimum of

$$\sum_{n=1}^{N} \big(y_n - f(y_{n-1})\big)^2$$

for $y_0 = 0$, $y_0 \le y_1 \le \cdots \le y_{N-1} \le y_N$ and $y_N = 1$. The set of sequences which we minimize over can be split into two parts: Either $y_{N-1} < c$ or the d last elements in the sequence are larger than c: $y_{N-d} \ge c$. If $y_{N-d} \ge c$,

$$\sum_{n=1}^{N} \big(y_n - f(y_{n-1})\big)^2 \ge (1-ac)^2 + (d-1)(c-ac)^2 + \sum_{n=1}^{N-d} \big(y_n - f(y_{n-1})\big)^2,$$

and this lower bound is attained for $y_{N-d+1} = \cdots = y_{N-1} = c$. Minimizing the sum on the right-hand side above when $0 \le y_1 \le \cdots \le y_{N-d-1} \le c$ and $y_{N-d} \ge c$ is the same minimizing problem as in (13), since $f(y_{n-1}) = ay_{n-1}$ when $y_{n-1} \le c$. We get

$$\inf_{\substack{y_N=1, y_0=0, y_{N-d} \ge c \\ y_0 \le y_1 \le \cdots \le y_N}} \sum_{n=1}^{N} \big(y_n - f(y_{n-1})\big)^2 = (1-ac)^2 + (d-1)(c-ac)^2 + \frac{(1-a^2)c^2}{1-a^{2(N-d)}}.$$

This value is smallest if $d = 1$, and it is then

$$(1-ac)^2 + \frac{(1-a^2)c^2}{1-a^{2(N-1)}}. \tag{28}$$

On the other hand, if $y_{N-1} < c$,

$$\begin{aligned}
&\inf_{\substack{y_N=1, y_0=0, y_{N-1}<c \\ y_0 \le y_1 \le \cdots \le y_N}} \sum_{n=1}^{N} \big(y_n - f(y_{n-1})\big)^2 \\
&= \inf_{\substack{y_N=1, y_0=0, y_{N-1}<c \\ y_0 \le y_1 \le \cdots \le y_N}} \sum_{n=1}^{N} (y_n - a y_{n-1})^2 = \frac{1-a^2}{1-a^{2N}}
\end{aligned}$$

if $a < c$, because the minimizing problem then coincides with the same problem for the autoregressive process. If $a \ge c$,

$$\begin{aligned}
&\inf_{\substack{y_N=1, y_0=0, y_{N-1}<c \\ y_0 \le y_1 \le \cdots \le y_N}} \sum_{n=1}^{N} \big(y_n - f(y_{n-1})\big)^2 \\
&= \inf_{\substack{y_N=1, y_0=0, y_{N-1}<c \\ y_0 \le y_1 \le \cdots \le y_N}} \left((1-a y_{n-1})^2 + \sum_{n=1}^{N-1} \big(y_n - f(y_{n-1})\big)^2 \right)
\end{aligned}$$

where it is optimal to have y_{N-1} close to c, and we get the same infimum as in (28). To summarize,

$$\begin{aligned}
&\inf_{\substack{y_N=1, y_0=0 \\ y_0 \le y_1 \le \cdots \le y_N}} \sum_{n=1}^{N} \big(y_n - f(y_{n-1})\big)^2 \\
&= \min\left\{ (1-ac)^2 + \frac{(1-a^2)c^2}{1-a^{2(N-1)}}, \frac{1-a^2}{1-a^{2N}} \right\}.
\end{aligned}$$

Since

$$\limsup_{\varepsilon \to 0} \varepsilon^2 \log E_{x_0} \tau_\varepsilon \le \inf_{1 \le N \le M} \inf_{\substack{y_N=1, y_0=0 \\ y_0 \le y_1 \le \cdots \le y_N}} \sum_{n=1}^{N} \big(y_n - f(y_{n-1})\big)^2$$

for any positive integer M, we can let M be arbitrarily large. We get

$$\limsup_{\varepsilon \to 0} \varepsilon^2 \log E_{x_0} \tau_\varepsilon \le \begin{cases} \frac{1}{2}\big((1-ac)^2 + (1-a^2)c^2\big), & \text{if } c \le a, \\ \frac{1}{2}(1-a^2), & \text{if } c \ge a. \end{cases}$$

□

Proposition 3. *Let* $0 \le a < 1$ *and let* f *be a function that satisfies*

$$f(x) = \begin{cases} 0 & \text{if } -1 \le x < 0, \\ -ax & \text{if } 0 \le x \le 1. \end{cases}$$

Then

$$\limsup_{\varepsilon \to 0} \varepsilon^2 \log E_{x_0} \tau_\varepsilon \le \frac{1}{2} \frac{1}{1+a^2}. \tag{29}$$

Proof. We study the sum

$$S_N := \sum_{n=1}^{N} \big(y_n - f(y_{n-1})\big)^2.$$

If $y_N = 1$, then $S_N \ge 1$, since $f(x) \le 0$ for all $x \in [-1, 1]$. Also, if $y_N = -1$ and $y_{N-1} < 0$, we have $S_N \ge 1$. In the case when $y_N = -1$ and $y_{N-1} \ge 0$, it is optimal to have $y_{N-2} = \cdots = y_1 = y_0 = 0$. The smallest value that the sum can take is the minimum of $(-1 + ax)^2 + x^2$ for $x \in [0, 1]$, which is $1/(1 + a^2)$. Thus,

$$\inf_{1 \le N \le M} \inf_{\substack{|y_N|=1, y_0=0 \\ y_1,\ldots,y_{N-1} \in (-1,1)}} \sum_{n=1}^{N} \big(y_n - f(y_{n-1})\big)^2 = \frac{1}{1+a^2},$$

and it follows that

$$\limsup_{\varepsilon \to 0} \varepsilon^2 \log E_{x_0} \tau_\varepsilon \le \frac{1}{2} \frac{1}{1+a^2}.$$ □

Note that if $a = 0$ in Proposition 3, we have $f(x) = 0$ on $[-1, 1]$, which is a special case of the autoregressive process with $a = 0$. The upper bound is then just $1/2$ which agrees with the result in the autoregressive case.

Proposition 4. *Let* f *be a function that satisfies*

$$f(x) = \begin{cases} -bx & \text{if } -1 \le x < 0, \\ -ax & \text{if } 0 \le x \le 1, \end{cases}$$

where $0 < a < 1$, $0 < b < 1$. *Then*

$$\limsup_{\varepsilon \to 0} \varepsilon^2 \log E_{x_0} \tau_\varepsilon \le \frac{1}{2} \min\left(\frac{1-(ab)^2}{1+a^2}, \frac{1-(ab)^2}{1+b^2} \right). \tag{30}$$

Proof. Let

$$S = \sum_{n=1}^{N} \big(y_n - f(y_{n-1})\big)^2.$$

Consider the cases $y_N = 1$ and $y_N = -1$ separately. First, let $y_N = -1$. Clearly, the sum S is smallest if the sequence $\{y_n\}_{n=1}^N$ has alternating signs: $y_{N-1} > 0$, $y_{N-2} < 0, \ldots$. Putting the derivative of S with respect to y_n equal to zero gives

$$y_n - f(y_{n-1}) = f'(y_n)\big(y_{n+1} - f(y_n)\big),$$

where $f'(y_n) = -b$ if $y_n < 0$ and $f'(y_n) = -a$ if $y_n > 0$. Let $s_n := y_n - f(y_{n-1})$. Then, if $N = 2M$ so that N is even, it is optimal to have $y_1 > 0$. Then

$$s_{2k} = -a^{-k} b^{-k+1} s_1,$$
$$s_{2k+1} = a^{-k} b^{-k} s_1,$$

and it follows that

$$y_N = -s_1 \big[a^M b^{M-1} + a^{M-2} b^{M-1} + a^{M-2} b^{M-3} + a^{M-4} b^{M-3} + \cdots + a^{-M+2} b^{-M+1} + a^{-M} b^{-M+1} \big].$$

This is a sum of two geometric sums. Since $y_N = -1$,

$$s_1 = \frac{1}{(1+a^2)(ab)^N a^{-2} b^{-1}} \cdot \frac{(ab)^{-2} - 1}{(ab)^{-2N} - 1}.$$

The value of the sum is

$$S = \sum_{n=1}^{N} s_n^2 = \frac{1-(ab)^2}{1+a^2} \cdot \frac{1}{1-(ab)^{2N}}.$$

If N is odd, it is optimal to have $y_1 < 0$. Then the same calculations as above follow, if a is replaced by b and vice versa. It follows that the value of the sum is

$$S = \sum_{n=1}^{N} s_n^2 = \frac{1-(ab)^2}{1+b^2} \cdot \frac{1}{1-(ab)^{2N}}.$$

The cases when $y_N = 1$ and N is odd or even, give the same values of the sum. The smallest values are attained when N is large, and the statement of the proposition follows. □

We note that when $a = b$ in Proposition 4, we have the autoregressive process with $f(x) = -ax$ on $[-1, 1]$. The upper bound in Proposition 4 reduces to $1/(2(1-a^2))$ which is the bound for the autoregressive process.

Remark 2. The TAR(1)-model, where $f(x) = a|x|$ for $|a| < 1$, is another piecewise linear example. For this model,

$$\limsup_{\varepsilon \to 0} \varepsilon^2 \log E\tau_\varepsilon \leq \frac{1-a^2}{2}. \tag{31}$$

Proof. If $0 \leq a < 1$,

$$\inf_{\substack{|y_N| \geq 1 \\ y_0 = 0}} \frac{1}{2} \sum_{n=1}^{N} (y_n - a|y_{n-1}|)^2 = \inf_{\substack{y_N = 1, y_0 = 0 \\ y_n \geq 0, n = 1, \ldots, N-1}} \frac{1}{2} \sum_{n=1}^{N} (y_n - a y_{n-1})^2, \tag{32}$$

and we have the same infimum as in the autoregressive case (which was treated in [12] and [13]). The case $-1 < a \leq 0$ is treated in a similar way. □

The upper bound in (31) is sharp; this was shown in [18], where Novikov's martingale method was used for this process to get the corresponding lower bound (the proof was almost the same as for the autoregressive process).

5 Other nonlinear functions

For a more general function f it is not always possible to minimize the sum and get explicit upper bounds. Numerical calculations may be needed. It is also possible to construct (nonstrict) upper bounds by simply evaluating the sum for some sequence of values instead of actually minimizing it. This is illustrated in the following case. Consider the case when f is quadratic, so we have the process

$$X_{n+1}^{\varepsilon} = f\left(X_n^{\varepsilon}\right) + \varepsilon\xi_{n+1}, \tag{33}$$

where $X_0 = 0$, $\{\xi_n\}_{n\geq 1}$ is a sequence of i.i.d. standard normal random variables, ε is a small positive parameter and $f(x) = ax^2$.

Proposition 5. *When $f(x) = ax^2$ and $0 \leq a \leq 0.5$,*

$$\limsup_{\varepsilon\to 0} \varepsilon^2 \log E\tau_\varepsilon \leq \frac{1}{2}.$$

If $a \geq 0.5$,

$$\limsup_{\varepsilon\to 0} \varepsilon^2 \log E\tau_\varepsilon \leq \frac{1}{2}\left(\frac{1}{a} - \frac{1}{4a^2}\right). \tag{34}$$

Proof. Since $f(x) \geq 0$ and f is even, it is optimal to have $y_N = 1$ and $y_n \geq 0$ $\forall n$. We have

$$\begin{aligned}\sum_{n=1}^{N}\left(y_n - f(y_n)\right)^2 &= \sum_{n=1}^{N}\left(y_n - ay_{n-1}^2\right)^2 \\ &= 1 + y_{n-1}^2(1-2a) + y_{n-2}^2(1-2ay_{n-1}) + \cdots + y_1^2(1-2ay_2) \\ &\quad + a^2\left(y_{n-1}^4 + y_{n-2}^4 + \cdots + y_2^2\right) \geq 1\end{aligned}$$

when $0 \leq a \leq 0.5$, and equality is achieved by putting $y_1 = y_2 = \cdots = y_{N-1} = 0$.

For $a \geq 0.5$,

$$\inf_{1\leq N\leq M}\left(\inf_{\substack{|y_N|\geq 1,\, y_0=0\\ |y_n|<1,\, n=1,2,\ldots,n-1}} \frac{1}{2}\sum_{n=1}^{N}\left(y_n - ay_{n-1}^2\right)^2\right) \leq \inf_{\substack{y_2=1,\, y_0=0\\ |y_1|<1,}} \frac{1}{2}\sum_{n=1}^{2}\left(y_n - ay_{n-1}^2\right)^2,$$

where the infimum on right-hand side is $\frac{1}{2}(\frac{1}{a} - \frac{1}{4a^2})$ (it is attained for $y_1 = \sqrt{\frac{1}{a} - \frac{1}{2a^2}}$). This gives the upper bound in (34). □

This is not necessarily the best upper bound for all $a \geq 0.5$, since we have only calculated the infimum for $N = 2$, but at least we get an upper bound. Numerical calculations suggest that the bound in (34) is the best one for a slightly larger than 0.5, and that the optimal choice of N increases as a increases.

For another example of a nonlinear case, consider the classical deterministic Ricker model defined by

$$x_{n+1} = x_n e^{r-\gamma x_n}, \quad n = 0, 1, 2, \ldots,$$

where x_t represents the size or density of a population at time t, $r > 0$ models the growth rate and $\gamma > 0$ is an environmental factor [9]. By suitably renorming the population, we may take $\gamma = 1$. The model then has a fixed point at $x = r$ (and one at $x = 0$). If we introduce stochasticity in the model by adding normally distributed white noise, and move the fixed point $x = r$ to the origin, we have the process

$$X_{n+1}^{\varepsilon} = f\left(X_n^{\varepsilon}\right) + \varepsilon \xi_{n+1},$$

where $f(x) = (x + r)e^{-x} - r$ and $\{\xi_n\}_{n \geq 1}$ is a sequence of i.i.d. $N(0, 1)$ variables. We can examine the time until the process exits from a suitable neighbourhood of the origin.

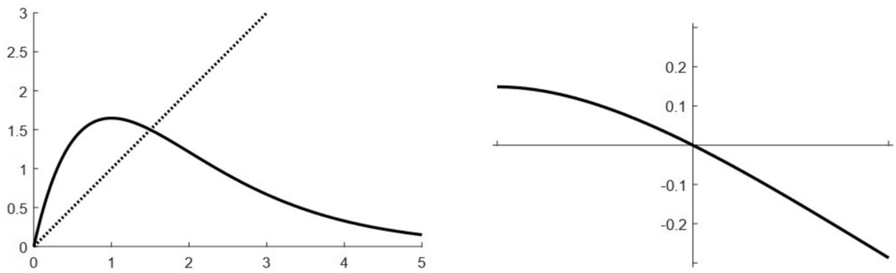

Fig. 3. The function $f(x) = xe^{r-x}$ for $r = 1.5$. We have a fixed point at $x = r$. On the right, we see a part of the plot with the fixed point moved to the origin

If $r = 1.5$, consider for example the time until exit from the interval $[-0.5, 0.5]$. Numerical calculations of the infimum

$$\inf_{1 \leq N \leq M} \left(\inf_{\substack{|y_N| \geq 0.5 \\ y_0 = 0 \\ |y_n| < 0.5, n=1,\ldots,N}} \frac{1}{2} \sum_{n=1}^{N} \left(y_n - f(y_{n-1})\right)^2 \right)$$

give the approximate value 0.09, so that

$$\limsup_{\varepsilon \to 0} \varepsilon^2 \log E\tau_\varepsilon \lessapprox 0.09.$$

The derivative of f at the origin is $1 - r$, so a suitable linear approximation of the function f is $l(x) = -0.5x$. By replacing f by l, we have an autoregressive process for which the upper bound of the exit time is

$$\limsup_{\varepsilon \to 0} \varepsilon^2 \log E\tau_\varepsilon \leq 0.5^2 \frac{1 - 0.5^2}{2} = 0.09375.$$

We note that a linear approximation might give good enough approximations of the upper bounds, in cases when the neighbourhood around the origin is chosen to be rather small.

6 Connection to stationary distribution

Mark Kac proved in 1947 that the mean return time of a discrete Markov chain to a point x is the reciprocal of the invariant probability $\pi(x)$. In [7], we explored this idea by comparing the exit time for the process defined by the stochastic difference equation

$$X_{n+1}^{\varepsilon} = f\left(X_n^{\varepsilon}\right) + \varepsilon \xi_{n+1} \tag{35}$$

and the return time to a certain set for the same process. We get an upper bound for the asymptotics of the exit time, and that bound is the reciprocal of the stationary distribution of the process, evaluated at the point where the level curve of the stationary distribution touches the boundary of the set (or interval in the univariate case) from which the process exits.

In the univariate case with $f(x) = ax$, $x \in \mathbb{R}$, (that is, for the autoregressive process), this method gives the same upper bound as the LDP method:

$$\limsup \varepsilon^2 \log E\tau_{\varepsilon} \leq \frac{1-a^2}{2}.$$

These methods only give an upper bound, but we know that the bound is sharp in the autoregressive case; the corresponding lower bound can be shown by other methods – see [18], where Novikov's martingale method ([16] and [17]) was used, and [11].

In Section 4, we saw several examples of processes of the type defined in (35) where f was a piecewise linear function. For these examples, it is not straightforward to derive expressions for the stationary distributions of the processes. We also note that in Section 4, where the LDP method was used, the definition of f outside of the interval $[-1, 1]$ could be ignored. However, when calculating stationary distributions, the definition of f outside of $[-1, 1]$ matters a great deal.

We observe that there is another piecewise linear case for which the stationary distribution is known, and this is a TAR(1) process with a threshold of 0, studied by Anděl et al. [1], where

$$X_{n+1}^{\varepsilon} = f\left(X_n^{\varepsilon}\right) + \varepsilon \xi_{n+1},$$

with $f(x) = -|ax|$ for $|a| < 1$. In [1], they gave the following formula for the stationary distribution of this process:

$$\frac{1}{\varepsilon}\left(\frac{2(1-a^2)}{\pi}\right)^{1/2} \exp\left(-\frac{1}{2}(1-a^2)\frac{x^2}{\varepsilon^2}\right)\Phi\left(\frac{-ax}{\varepsilon}\right),$$

where Φ is the cumulative distribution function of a $N(0, 1)$ random variable. We note that

$$-\varepsilon^2 \log\left(\frac{1}{\varepsilon}\left(\frac{2(1-a^2)}{\pi}\right)^{1/2} \exp\left(-\frac{1}{2}(1-a^2)\frac{x^2}{\varepsilon^2}\right)\Phi\left(\frac{-ax}{\varepsilon}\right)\right) \to \frac{1-a^2}{2}$$

at the point $x = -1$. This means that in this particular case, the value of the limit of the stationary distribution, evaluated at the point where the process exits the interval, coincides with the upper bound achieved by use of the LDP method in Remark 2.

A preprint of a previous version of this paper has been posted on ArXiv [10].

Acknowledgement

The authors would like to thank the two anonymous referees for their constructive comments which have led to improvement of the paper.

References

[1] Anděl, J., Netuka, I., Zvára, K.: On threshold autoregressive processes. Kybernetika **20**(2), 89–106 (1984). MR0747062

[2] Basak, G.K., Ho, K.-W.R.: Level-crossing probabilities and first-passage times for linear processes. Adv. Appl. Probab. **36**, 643–666 (2004). MR2058153. https://doi.org/10.1239/aap/1086957589

[3] Baumgarten, C.: Survival probabilities of autoregressive processes. ESAIM Probab. Stat. **18**, 145–170 (2014). MR3143737. https://doi.org/10.1051/ps/2013031

[4] Di Nardo, E.: On the first passage time for autoregressive processes. Sci. Math. Jpn. **21**, 61–76 (2008). MR2399804

[5] Hamza, K., Jagers, P., Klebaner, F.C.: On the establishment, persistence and inevitable extinction of populations. J. Math. Biol. **72**, 797–820 (2016). MR3459167. https://doi.org/10.1007/s00285-015-0903-2

[6] Högnäs, G.: On the lifetime of a size-dependent branching process. Stoch. Models **35**, 119–131 (2019). MR3969510. https://doi.org/10.1080/15326349.2019.1578241

[7] Högnäs, G., Jung, B.: Analysis of a stochastic difference equation: Exit times and invariant distributions. Fasc. Math. **44**, 69–74 (2010). MR2722632

[8] Högnäs, G., Jung, B.: Exit times for some autoregressive processes with non-Gaussian noise distributions. Contemp. Math. **668**, 111–117 (2016). MR3536695. https://doi.org/10.1090/conm/668/13399

[9] Högnäs, G.: On the quasi-stationary distribution of a stochastic Ricker model. Stoch. Process. Appl. **70**, 243–263 (1997). MR1475665. https://doi.org/10.1016/S0304-4149(97)00064-1

[10] Högnäs, G., Jung, B.: Exit times for some nonlinear autoregressive processes (2019). arXiv:1912.08514.

[11] Jung, B.: Exit times for multivariate autoregressive processes. Stoch. Process. Appl. **123**, 3052–3063 (2013). MR3062436. https://doi.org/10.1016/j.spa.2013.03.003

[12] Klebaner, F., Liptser, R.: Large deviations for past-dependent recursions. Probl. Inf. Transm. **32**, 23–34 (1996). MR1441520

[13] Klebaner, F., Liptser, R.: Large deviations for past-dependent recursions (2006). arXiv:math/0603407.

[14] Klebaner, F.C., Liptser, R.: Moderate deviations for randomly perturbed dynamical systems. Stoch. Process. Appl. **80**(2), 157–176 (1999). MR1682255. https://doi.org/10.1016/S0304-4149(98)00075-1

[15] Koski, T., Jung, B., Högnäs, G.: Exit times for ARMA processes. Adv. Appl. Probab. **50**(A), 191–195 (2018). MR3905100. https://doi.org/10.1017/apr.2018.79

[16] Novikov, A.A.: On the first passage time of an autoregressive process over a level and an application to a "disorder" problem. Theory Probab. Appl. **35**(2), 269–279 (1991). https://doi.org/10.1137/1135035

[17] Novikov, A., Kordzakhia, N.: Martingales and first passage times of AR(1) sequences. Stochastics **80**(2-3), 197–210 (2008). MR2402164. https://doi.org/10.1080/17442500701840885

[18] Ruths, B.: Exit times for past-dependent systems. Surv. Appl. Ind. Math. (Obozr. Prikl. Prom. Mat.) **15**, 25–30 (2008)

[19] Varadhan, S.R.S.: Large Deviations and Applications. Society for Industrial and Applied Mathematics (1984). https://epubs.siam.org/doi/abs/10.1137/1.9781611970241. MR0758258. https://doi.org/10.1137/1.9781611970241.bm