Adv. Appl. Prob. **50,** 805–832 (2018)
*doi:*10.1017/apr.2018.37

IMPORTANCE SAMPLING OF HEAVY-TAILED ITERATED RANDOM FUNCTIONS

BOHAN CHEN,* **
CHANG-HAN RHEE *** AND
BERT ZWART,* **** *Centrum Wiskunde & Informatica*

Abstract

We consider the stationary solution Z of the Markov chain $\{Z_n\}_{n\in\mathbb{N}}$ defined by $Z_{n+1} = \Psi_{n+1}(Z_n)$, where $\{\Psi_n\}_{n\in\mathbb{N}}$ is a sequence of independent and identically distributed random Lipschitz functions. We estimate the probability of the event $\{Z > x\}$ when x is large, and develop a state-dependent importance sampling estimator under a set of assumptions on Ψ_n such that, for large x, the event $\{Z > x\}$ is governed by a single large jump. Under natural conditions, we show that our estimator is strongly efficient. Special attention is paid to a class of perpetuities with heavy tails.

Keywords: State-dependent importance sampling; heavy-tailed distribution; iterated random function; perpetuities

2010 Mathematics Subject Classification: Primary 65C05
Secondary 60J05

1. Introduction

We consider an $\mathbb{R}$-valued Markov chain $\{Z_n\}_{n\in\mathbb{N}}$ defined by

$$Z_{n+1} = \Psi_{n+1}(Z_n), \tag{1}$$

where $\{\Psi_n\}_{n\in\mathbb{N}}$ is a sequence of independent and identically distributed (i.i.d.) positive random Lipschitz functions (see (14)) and $Z_0 \in \mathbb{R}$ is independent of the sequence $\{\Psi_n\}_{n\in\mathbb{N}}$. Under mild conditions (see Assumption B1), the stationary solution to (1) has the same distribution as the almost-sure limit Z of the sequence $\{\Psi_1 \circ \cdots \circ \Psi_n(Z_0)\}_{n\in\mathbb{N}}$; see [12] for details. We assume that Ψ_n is such that $\Psi_1 \circ \cdots \circ \Psi_n(Z_0)$ is increasing in n. In this paper we develop efficient simulation methods for estimating the tail probability of Z, i.e. we are interested in computing $\mathbb{P}(Z > x) = \mathbb{P}(T(x) < \infty)$ for large x, where $T(x) = \inf\{n \geq 0 \colon \Psi_1 \circ \cdots \circ \Psi_n(Z_0) > x\}$.

Two examples are of particular interest in the above setting. The first example is the so-called stochastic perpetuity. More precisely, consider the random difference equation $\Psi_n(z) = A_n z + B_n$. Then recursion (1) becomes

$$Z_{n+1} = A_{n+1} Z_n + B_{n+1}, \tag{2}$$

Received 15 September 2016; revision received 6 June 2018.
* Postal address: Stochastics Group, Centrum Wiskunde & Informatica, Science Park 123, 1098 XG, Amsterdam, The Netherlands.
** Email address: b.chen@cwi.nl
*** Current address: Industrial Engineering and Management Sciences, 2145 Sheridan Road, Tech C150, Evanston, Illinois, IL 60208-3109, USA. Email address: chang-han.rhee@northwestern.edu
**** Email address: bert.zwart@cwi.nl

where $\{(A_n, B_n)\}_{n\geq 0}$ is a sequence of i.i.d. $\mathbb{R}$-valued random vectors, independent of the initial random variable Z_0. It is well known (see, e.g. [8, Chapter 2]) that if $\mathbb{E}[\log A_1] < 0$ and $\mathbb{E}[\log^+ B_1] < \infty$ then the Markov chain given by (2) has a unique stationary distribution, which has the same distribution as $Z = \sum_{n=0}^{\infty} B_{n+1}\mathrm{e}^{S_n}$, where $S_n = \sum_{i=1}^{n} \log A_i$. Moreover, noting that $T(x) = \inf\{n \geq 0\colon \sum_{k=0}^{n} B_{k+1}\mathrm{e}^{S_k} > x\}$, our objective is to show that $\mathbb{P}(Z > x) = \mathbb{P}(T(x) < \infty)$ in the positive B_n case. Perpetuities occur, e.g. in the context of ruin problems with investments, in the study of financial time series, such as ARCH-type processes (see, e.g. [13]), and in tail asymptotics for exponential functionals of Lévy processes (see, e.g. [24]). Although some particular cases exist that allow for an explicit analysis (see, e.g. [30]), it is difficult to obtain exact results for the distribution of Z in general. Thus, Monte Carlo simulation arises as a natural approach to deal with the analysis of stochastic perpetuities, including the large deviations regime in which x in $\mathbb{P}(Z > x)$ is large, the focus of this paper.

Another example of (1) is the Lindley recursion which describes the waiting time of a customer in a single-server queue. More precisely, we consider (1) with $\Psi_n(z) = \max\{0, z + X_n\}$, where X_n, $n \in \mathbb{N}$, is a sequence of i.i.d. $\mathbb{R}$-valued random variables. It is well known (see, e.g. [15]) that the stationary solution of the Markov chain $Z_{n+1} = \max\{0, Z_n + X_{n+1}\}$, $n \in \mathbb{N}$, represents the all-time maximum of a random walk, denoted by $\max_{n\geq 0} S_n$, where $S_n = \sum_{i=1}^{n} X_i$. A similar connection holds, of course, between e^{Z_n} and $\mathrm{e}^{Z_{n+1}}$ in this context. The exponentiated form of the Lindley recursion is actually more suitable for our purposes: by developing a connection between iterated random functions (perpetuities) and the maximum of a random walk, we utilize rare-event simulation techniques to estimate $\mathbb{P}(\max_{n\geq 0} S_n > x)$ and construct efficient simulation algorithms for computing the tail probability of the stationary solutions to (1) and (2) under a heavy-tailed setup.

Before we state a more precise description of our results, we first mention some related works. In the more general context of iterated random functions, Goldie [16] studied the tail behavior of Z under sufficient conditions on Ψ_n for which $\mathbb{P}(Z > x)$ behaves as a power-law distribution. We refer the reader to [25] for a more recent study. A related study to the present work was carried out by Dyszewski [12], who showed that the tail of Z is slowly varying under the same conditions (see Assumptions A1, A2, B1, and B2) applied in this work. Dyszewski's result was an extension of a classical result of Pakes [26] and Veraverbeke [29]. for the maximum of random walks.

Turning to the special case of stochastic perpetuities, sufficient conditions for $\mathbb{P}(Z > x)$ to decay at an exponential rate were established by Goldie and Grübel [17], who assumed that $|A_1|$ is bounded by 1 and that the moment generating function of B_1 exists in a neighborhood of the origin. By assuming that $\mathbb{E}|A_1|^{\alpha} = 1$ and $\mathbb{E}|B_1|^{\alpha} < \infty$ for some $\alpha > 0$, Kesten [22] and later Goldie [16] proved that Z has a power-law distribution with exponent α. Moreover, a result due to Grincevicius [19] and later generalized by Grey [18] states that the tail of Z is regularly varying with some index, say $-\alpha$, if B_1 is regularly varying with the same index $-\alpha$ and $\mathbb{E}A_1^{\alpha} < 1$. For a more extensive overview of the literature in this area, see, e.g. [8] and the references therein.

A study on rare-event simulation of primary interest to this work is that of Blanchet and Glynn [5], who designed an algorithm for estimating the tail probability of the all-time maximum of heavy-tailed random walks. A major contribution of the present paper is the extension of the algorithm in [5] to the more general setting of [12]. Asmussen and Nielsen [3] also studied rare-event simulation for perpetuities and iterated random functions considering deterministic interest rates. Blanchet and Zwart [7] estimated the tail probability of perpetuities with deterministic premiums (B_n). Later, Blanchet *et al.* [6] developed simulation algorithms for perpetuities when the discount factor and premium are modeled by a Markov chain. Furthermore, Collamore *et al.* [10] provide simulation estimators for the tail distribution of Z as in (1) with $\Psi_n(z) = A_n \max\{z, D_n\} + B_n$.

The contributions of the present paper are as follows. For stochastic perpetuities, we propose a strongly efficient simulation algorithm for estimating $\mathbb{P}(Z > x)$. To this end, we make several assumptions; see Assumptions A1–A5. We illustrate the generality of these assumptions through examples and by providing sufficient conditions in Remarks 3, 4, 6, and 7. We construct an upper bounding random walk for the stochastic perpetuity, which leads to an asymptotic result for the tail probability of Z under a heavy-tailed assumption on $\log \max\{A_1, B_1\}$. Note that Z is defined over an infinite horizon and, hence, requires an infinite amount of computational effort to generate each sample when using a crude Monte Carlo sampling approach. A natural approach to address such an issue is to obtain approximations by means of finite-time truncation. We study the bias introduced by such approximations and show that our estimator has a vanishing relative bias as $x \to \infty$. By making a slightly stronger, but not restrictive, assumption (see Assumption A5), we are able to identify the rate at which the bias decays with respect to the truncation time. Applying the bias elimination technique studied in [28], we then propose strongly efficient and *unbiased* estimators for $\mathbb{P}(Z > x)$. Finally, we extend these results to the more general setting of (1). In Section 3.2 we make a couple of extra assumptions (see Assumptions B0 and B1) on Ψ_n. Our setting is almost identical to that of [9], [12], and [16]. In Remark 9 we give examples that satisfy our assumptions.

The most important aspect of this work is that we connect our class of iterated random functions to the maximum of a random walk. To illustrate this, consider (2) with $B_n = 1$. We connect the stochastic perpetuity with the maximum of the random walk by observing that, for $\gamma \in (0, -\mathbb{E}S_1)$,

$$Z = \sum_{n=0}^{\infty} \exp(S_n) = \sum_{n=0}^{\infty} \exp(S_n + n\gamma)\exp(-n\gamma) \le \exp\Big(\max_{n\ge 0}(S_n + n\gamma)\Big)\frac{1}{1-\mathrm{e}^{-\gamma}}, \qquad (3)$$

where $S_n = \sum_{i=1}^{n} \log A_i$. The upper bounding random walk constructed in (3) allows us to construct a coupling, and leverage the importance sampling algorithm designed by Blanchet and Glynn [5]. It turns out that we can extend this idea to the general setting of (1) by constructing a slightly more involved upper bounding random walk. Note that our extension of (3) leads to a shorter proof of the asymptotic upper bound given in [12], which we believe to be of intrinsic interest.

The rest of the paper is organized as follows. In Section 2 we introduce our notation and basic background information. The main results are stated in Section 3, first in the context of stochastic perpetuity and then in the context of the iterated functions setting. Numerical results are presented in Section 4. All proofs can be found in Section 5.

2. Notation and preliminary results

In this section we introduce the notation use throughout the paper, and recall some preliminary results from the literature.

For $(x, y) \in \mathbb{R}^2$, let $x \wedge y \triangleq \min\{x, y\}$ and $x \vee y \triangleq \max\{x, y\}$. For $x \in \mathbb{R}$, let $x^+ = x \vee 0$ denote the positive part of x and let $\log^+ x = 0 \vee \log x = \log(x \vee 1)$. Let $c \in \mathbb{R} \cup \{\pm\infty\}$, and let $f(x)$ and $g(x)$ be nonnegative real-valued functions. We respectively write $f(x) \sim g(x)$, $f(x) = o(g(x))$, and $f(x) = \mathcal{O}(g(x))$ as $x \to c$, if $\lim_{x\to c} f(x)/g(x) = 1$, $\lim_{x\to c} f(x)/g(x) = 0$, and $\limsup_{x\to c} f(x)/g(x) < \infty$.

To describe the efficiency of a rare-event simulation algorithm, we adopt a widely applied criterion (for a discussion of efficiency in rare-event simulation, see, e.g. [2]). Suppose that we are interested in a sequence of rare events $\mathcal{E}(x)$ that become more rare as $x \to \infty$. Let $L(x)$ be an unbiased estimator of the rare-event probability $\mathbb{P}(\mathcal{E}(x))$. We say that $L(x)$ is strongly efficient if $\mathbb{E}L(x)^2 = \mathcal{O}(\mathbb{P}(\mathcal{E}(x))^2)$ as $x \to \infty$. In particular, strong efficiency implies that the number of simulation runs required to estimate the target probability to a given relative accuracy is bounded with respect to (w.r.t.) x.

As we mentioned in the introduction, a state-dependent importance sampling scheme will be used in this paper. We recall the following result that will be very useful in validating our new estimator.

Result 1. (Asmussen [1, Proposition 3.1 and Theorem 3.2].) *Let Y_n, $n \in \mathbb{N}$, be a sequence of random variables on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let M_n, $n \in \mathbb{N}$, be a nonnegative martingale that is adapted to Y_n for which $\mathbb{E}M_0 = 1$. Let Γ be a stopping time adapted to Y_n. Define a sequence of probability measures as $\mathbb{P}_n(A') = \mathbb{E}\mathbf{1}_{A'}M_n$ for $A' \in \mathcal{F}_n \triangleq \sigma(Y_1, \dots, Y_n)$, where $\mathbf{1}$ is the indicator function. Then there exists a probability measure $\tilde{\mathbb{P}}$ such that $\tilde{\mathbb{P}}(A') = \mathbb{P}_n(A')$ for $A' \in \mathcal{F}_n$ and $n \in \mathbb{N}$. Furthermore, we have $\mathbb{E}\mathbf{1}_{\{\Gamma<\infty\}} = \tilde{\mathbb{E}}\mathbf{1}_{\{\Gamma<\infty\}}M_\Gamma^{-1}$.*

Next, we recall a simulation algorithm proposed in [5], where the authors developed an efficient state-dependent importance sampling strategy for estimating the tail probability of a random walk crossing a certain level. Before we go through the details of the simulation algorithm, we introduce the following definition.

Definition 1. Let Y be a random variable on $\mathbb{R}$. Let the integrated tail of Y, as a function of x, be defined by

$$x \mapsto 1 \wedge \int_x^\infty \mathbb{P}(Y > t)\,\mathrm{d}t.$$

We say that Y is long tailed if, for every $c \in \mathbb{R}$, we have

$$\mathbb{P}(Y > t + c) \sim \mathbb{P}(Y > t) \quad \text{as } t \to \infty.$$

We say that Y is subexponential, if

$$\mathbb{P}(Y_{(1)}^+ + Y_{(2)}^+ > t) \sim 2\mathbb{P}(Y^+ > t) \quad \text{as } t \to \infty,$$

where $Y_{(1)}$ and $Y_{(2)}$ are independent copies of Y. Moreover, we say that Y is strongly subexponential, or Y belongs to the class S^*, if

$$2\mathbb{E}Y^+\mathbb{P}(Y > t) \sim \int_0^t \mathbb{P}(Y > t - s)\mathbb{P}(Y > s)\,\mathrm{d}s \quad \text{as } t \to \infty.$$

Remark 1. Note that the integrated tail function defines a probability distribution. Moreover, if Y belongs to S^*, both the distribution of Y and its integrated tail are subexponential (see [23, Theorem 3.2]) and, in particular, long tailed.

Consider a random walk $\{S_n\}_{n\in\mathbb{N}}$ generated by a sequence of i.i.d. random variables $\{X_n\}_{n\in\mathbb{N}}$, i.e. $S_n = \sum_{i=1}^n X_i$. Assume that $\mathbb{E}X_1 < 0$, and X_1 belongs to S^*. Let $P(y, \mathrm{d}z)$ denote the transition kernel of the random walk $\{S_n\}_{n\in\mathbb{N}}$. Define a nonnegative random variable W that is independent of $\{X_n\}_{n\in\mathbb{N}}$ with tail probability

$$\mathbb{P}(W > y) \triangleq \min\biggl[1, -\frac{1}{\mathbb{E}X_1}\int_y^\infty \mathbb{P}(X_1 > t)\,\mathrm{d}t\biggr].$$

Fix $x > 0$. To estimate $\mathbb{P}(\max_{n\geq 0} S_n > x) = \mathbb{P}(\tau(x) < \infty)$, where

$$\tau(x) = \inf\{n \geq 0\colon S_n > x\},$$

Blanchet and Glynn [5] suggested simulating the random walk via another transition kernel

$$Q_{a_*}(y,\, \mathrm{d}z) \triangleq P(y,\, \mathrm{d}z)\frac{v(z + a_*)}{w(y + a_*)} \quad \text{for all } y \in (-\infty, x],\ z \in \mathbb{R}, \tag{4}$$

where

$$v(z) \triangleq \mathbb{P}(W > -(z - x)), \qquad w(y) \triangleq \mathbb{P}(X_1 + W > -(y - x)), \tag{5}$$

and a_* is such that, for fixed $\delta \in (0, 1)$,

$$-\delta \leq \frac{v^2(y) - w^2(y)}{\mathbb{P}(X_1 > -y)w(y)} \quad \text{for all } y \leq x + a_*. \tag{6}$$

Let $\mathbb{P}^{Q_{a*}}$ and $\mathbb{E}^{Q_{a*}}$ denote respectively the probability measure and the expectation w.r.t. the random process $\{S_n\}_{n\in\mathbb{N}}$ having a one-step transition kernel $Q_{a_*}(y, \mathrm{d}z)$ as in (4). In the following theorem, we state the simulation estimator proposed in [5], which will prove to be useful in our context.

Result 2. (Blanchet and Glynn [5, Theorem 3].) *Suppose that* $\mathbb{E}X_1 < 0$*, and* X_1 *belongs to* S^**. Let* v *and* w *be defined as in (5). For fixed* $\delta \in (0, 1)$*, there exists an* $a_* = a_*(\delta) \leq 0$ *such that (6) holds. Then*

$$L_\tau(x) = \mathbf{1}_{\{\tau(x)<\infty\}} \prod_{k=1}^{\tau(x)} \frac{w(S_{k-1} + a_*)}{v(S_k + a_*)}$$

is an unbiased estimator of $\mathbb{P}(\max_{n\geq 0} S_n > x)$ *under* $\mathbb{P}^{Q_{a*}}$*; moreover, it is strongly efficient, i.e.*

$$\sup_{x>0} \frac{\mathbb{E}^{Q_{a*}} L_\tau^2(x)}{\mathbb{P}(\max_{n\geq 0} S_n > x)^2} < \infty.$$

Remark 2. The existence of such an a_* as in Result 2 is guaranteed by the fact that (see [5, Proposition 3] for details)

$$w(y) - v(y) = o(\mathbb{P}(X_1 > -y)) \quad \text{as } y \to -\infty. \tag{7}$$

We will extend this algorithm to the setting of (1). Unfortunately, it is not straightforward to generate our estimator, say L, such that $\mathbb{E}L = \mathbb{P}(Z > x)$ in finite computation time. However, there exists a sequence L_n, $n \in \mathbb{N}$, of $\mathcal{L}^2$ approximations (i.e. $\mathbb{E}[(L_n - L)^2] \to 0$ as $n \to \infty$) that can be generated exactly in finite time. Rhee and Glynn [28] considered this situation and we recall one of their results which will prove to be crucial for our purposes.

Result 3. (Rhee and Glynn [28, Theorem 2].) *Let* L_n *and* L *be such that* $\mathbb{E}[(L_n - L)^2] \to 0$ *as* $n \to \infty$*. Let* N *be a nonnegative integer-valued random variable, independent of* L_n, $n \in \mathbb{N}$*, such that* $\mathbb{P}(N \geq n) > 0$ *for all* $n \geq 0$*. If*

$$\sum_{n=1}^{\infty} \frac{\mathbb{E}[(L_n - L)^2]}{\mathbb{P}(N \geq n)} < \infty,$$

https://doi.org/10.1017/apr.2018.37 Published online by Cambridge University Press

then $\bar{L}$ defined by

$$\bar{L} \triangleq \sum_{n=0}^{N} \frac{L_n - L_{n-1}}{\mathbb{P}(N \geq n)}$$

(with $L_{-1} = 0$) is an unbiased estimator of $\mathbb{E}L$, and

$$\mathbb{E}[\bar{L}^2] = \sum_{n=0}^{\infty} \frac{\mathbb{E}[(L_{n-1} - L)^2] - \mathbb{E}[(L_n - L)^2]}{\mathbb{P}(N \geq n)} < \infty.$$

In order to apply Result 3 in our context, we conclude this section with the following extension of Result 2, which means that the algorithm proposed in [5] can be used to yield an estimator with a bounded relative $(2+\varepsilon)$th moment for some $\varepsilon > 0$. The proofs of this lemma together with other results presented in this paper can be found in Section 5.

Lemma 1. *Let $S_n = \sum_{i=1}^{n} X_i$ be a random walk. Suppose that $\mathbb{E}X_1 < 0$, and X_1 belongs to $\mathcal{S}^*$. Let v and w be defined as in (5). For any fixed $\varepsilon > 0$ and $\delta \in (0, 1)$, there exists an $a_* = a_*(\varepsilon, \delta) \leq 0$ such that*

$$-\delta \leq \frac{v^{2+\varepsilon}(y) - w^{2+\varepsilon}(y)}{\mathbb{P}(X_1 > -y + x) w^{1+\varepsilon}(y)} \quad \text{for all } y \leq x + a_*.$$

Let

$$L_\tau(x) \triangleq \mathbf{1}_{\{\tau(x) < \infty\}} \prod_{k=1}^{\tau(x)} \frac{w(S_{k-1} + a_*)}{v(S_k + a_*)}.$$

Then $\mathbb{E}^{Q_{a_*}} L_\tau(x) = \mathbb{P}(\max_{n \geq 0} S_n > x)$ and

$$\sup_{x > 0} \frac{\mathbb{E}^{Q_{a_*}} L_\tau^{2+\varepsilon}(x)}{\mathbb{P}(\max_{n \geq 0} S_n > x)^{2+\varepsilon}} < \infty.$$

3. Main results

This section contains our main results. In Section 3.1 we consider the stochastic perpetuity as in (2). Recall that Z_n, $n \in \mathbb{N}$, is defined by

$$Z_{n+1} = A_{n+1} Z_n + B_{n+1} \quad \text{for } n \in \mathbb{N}.$$

Recalling that $Z = \sum_{n=0}^{\infty} B_{n+1} \mathrm{e}^{S_n}$ and $S_n = \sum_{i=1}^{n} \log A_i$, we are interested in estimating $\mathbb{P}(Z > x)$, where x is large. For this, we make several assumptions; see Assumptions A1–A5. We discuss the generality of these assumptions by giving examples as well as sufficient conditions in Remarks 3, 4, 6, and 7. To construct our simulation estimator, we construct a stochastic upper bound that can be written as a functional of a suitable random walk $S_n(\gamma)$. Then, using this upper bound, we define a crossing level $s(x)$ and a stopping time $\tau_\gamma(x) = \inf\{n \geq 0 \colon S_n(\gamma) > s(x)\}$ such that $\{Z > x\} \subseteq \{\tau_\gamma(x) < \infty\}$. Since the change of measure proposed by Blanchet and Glynn [5] is strongly efficient for estimating the tail probability of the maximum of heavy-tailed random walks, a natural strategy is to keep track of the random process $S_n(\gamma)$, $n \in \mathbb{N}$, while simulating the sequence $\sum_{k=0}^{n} B_{k+1} \mathrm{e}^{S_k}$, $n \in \mathbb{N}$, until the stopping time $\tau_\gamma(x)$. In doing so, we can construct a state-dependent change of measure using the path of the random walk until $\tau_\gamma(x)$ according to the method introduced in Section 2.

https://doi.org/10.1017/apr.2018.37 Published online by Cambridge University Press

Then we continue to simulate the path of the random walk after $\tau_\gamma(x)$ under the original measure. Based on this idea, we propose a simulation algorithm for estimating $\mathbb{P}(Z > x)$ and discuss its properties such as strong efficiency in the rest of Section 3.1. In Section 3.2 we extend the results of Section 3.1 to the general setting, where $Z_n,\ n \in \mathbb{N}$, was defined in (1) as

$$Z_{n+1} = \Psi_{n+1}(Z_n) \quad \text{for } n \in \mathbb{N},$$

and $\{\Psi_n\}_{n\in\mathbb{N}}$ is a sequence of i.i.d. random functions that is independent of Z_0. Note that all the proofs of the results in this section can be found in Section 5.

3.1. Stochastic perpetuity

We consider the Markov chain $Z_n,\ n \in \mathbb{N}$, as in (2). To guarantee the positive recurrence of $\{Z_n\}_{n\in\mathbb{N}}$, we assume the following.

Assumption A1. *We have the following assumptions:*

(i) $A_1 > 0$ *almost surely (a.s.),* $\mathbb{E}\log A_1 < 0$*, and* $\mathbb{E}\log^+ |B_1| < \infty$*;*

(ii) $\mathbb{E}\log^+(A_1 \vee B_1) < \infty$*;*

(iii) $\mathbb{P}(A_1 > x, B_1 \le -x) = o(\mathbb{P}(A_1 \vee B_1 > x))$.

Recall that, under Assumption A1, the unique stationary distribution of this Markov chain exists, has right-unbounded support, and has the same distribution as the random variable $Z \triangleq \sum_{n=0}^{\infty} B_{n+1}\mathrm{e}^{S_n}$, where $S_n = \sum_{i=1}^{n} \log A_i$; see, e.g. [8, Chapter 2] or [16] for more details. As mentioned at the beginning of Section 3, we start by developing a connection between perpetuities and the maximum of a random walk. More precisely, we construct an upper bound for Z that can be written as a functional of a suitable random walk $S_n(\gamma)$. We formulate the result in the following lemma.

Lemma 2. *Let Assumption A1 hold. There exists a constant* γ_2 *such that*

$$\mathbb{E}[(\log^+ B_1^+ - \gamma_2) \vee \log A_1] < 0.$$

Moreover, there exists a constant $\gamma_1 \in (0, -\mathbb{E}[\log A_i \vee (\log^+ B_1^+ - \gamma_2)])$ *such that*

$$Z \le \exp\Big(\max_{n\ge 0} S_n(\gamma)\Big)\frac{\mathrm{e}^{\gamma_2}}{1-\mathrm{e}^{-\gamma_1}} < \infty, \tag{8}$$

where $S_n(\gamma) = S_n(\gamma_1, \gamma_2) = \sum_{i=1}^{n}[\log A_i \vee (\log^+ B_i^+ - \gamma_2) + \gamma_1]$ *and* $\mathbb{E}S_1(\gamma) < 0$.

Now from (8) we define $s(x) \triangleq \log x - \gamma_2 + \log(1-\mathrm{e}^{-\gamma_1})$ and $\tau_\gamma(x) \triangleq \inf\{n \ge 0\colon S_n(\gamma) > s(x)\}$ such that the following holds:

$$\{Z > x\} \subseteq \Big\{\max_{n\ge 0} S_n(\gamma) > s(x)\Big\}. \tag{9}$$

As we will see in the proof of Theorem 2, the asymptotic behavior of $\mathbb{P}(Z > x)$ as $x \to \infty$ will be useful in establishing the strong efficiency of our estimator. Thus, we derive a tail estimate for Z in Theorem 1 below. To be precise, we are interested in finding a function $f(x)$ such that $\mathbb{P}(Z > x) = \mathcal{O}(f(x))$ as $x \to \infty$. Moreover, we focus on the case where the following assumption holds.

Assumption A2. *The integrated tail (see Definition 1) of* $\log(A_1 \vee B_1)$*, denoted by* $\bar{F}_I$*, is subexponential.*

https://doi.org/10.1017/apr.2018.37 Published online by Cambridge University Press

Remark 3. As mentioned in the introduction, the focus of this paper is to propose Monte Carlo estimators for $\mathbb{P}(Z > x)$, which is slowly varying as $x \to \infty$. Indeed, $\mathbb{P}(Z > x)$ is slowly varying under Assumptions A1 and A2. A proof can be found in, e.g. [12]; in Theorem 1 (and Theorem 5) we provide an independent proof for the asymptotic upper bound of $\mathbb{P}(Z > x)$ (under a general setting; see Assumptions B1 and B2).

Theorem 1. *If Assumptions A1 and A2 hold, we have*

$$\limsup_{x\to\infty} \frac{\mathbb{P}(Z > x)}{\bar{F}_I(\log(x))} \le -\frac{1}{\mathbb{E}\log A_1}.$$

By constructing the upper bound as in Lemma 2, we establish a connection between perpetuities and the maximum of a random walk. This connection will allow us to utilize rare-event simulation techniques for estimating $\mathbb{P}(\max_{n\ge 0} S_n > x)$ in designing an efficient simulation estimator for $\mathbb{P}(Z > x)$. To construct the simulation estimator of $\mathbb{P}(Z > x)$, define a nonnegative random variable W_γ that is independent of $\{(A_n, B_n)\}_{n\in\mathbb{N}}$ with tail probability

$$\mathbb{P}(W_\gamma > t) \triangleq \min\biggl[1, -\frac{1}{\mathbb{E}S_1(\gamma)} \int_t^\infty \mathbb{P}(S_1(\gamma) > s)\,\mathrm{d}s\biggr],$$

and define

$$v_\gamma(z) \triangleq \mathbb{P}(W_\gamma > -(z - s(x))) \quad \text{and} \quad w_\gamma(y) \triangleq \mathbb{P}(S_1(\gamma) + W_\gamma > -(y - s(x))). \tag{10}$$

Let $P^\gamma(y,\,\mathrm{d}z)$ denote the transition kernel of the random walk $\{S_n(\gamma)\}_{n\in\mathbb{N}}$. For fixed a_*, let $\mathbb{E}^{Q^\gamma_{a_*}}$ denote the expectation w.r.t. the stochastic process $\{S_n(\gamma)\}_{n\in\mathbb{N}}$ having a one-step transition kernel

$$Q^\gamma_{a_*}(y,\,\mathrm{d}z) = \begin{cases} P^\gamma(y,\,\mathrm{d}z)v_\gamma(z + a_*)w_\gamma(y + a_*)^{-1} & \text{for } n \le \tau_\gamma(x), \\ P^\gamma(y,\,\mathrm{d}z) & \text{for } n > \tau_\gamma(x). \end{cases} \tag{11}$$

We propose an estimator and show its strong efficiency in Theorem 2. We make a slightly stronger assumption on the tail asymptotics of $\log(A_1 \vee B_1)$ as follows.

Assumption A3. *The distribution of* $\log(A_1 \vee B_1)$ *belongs to the class* S^*.

Remark 4. Assumption A3 is not restrictive in the sense that the class S^* of strongly subexponential random variables includes regularly varying, lognormal, and Weibull-type distributions, among many others. For more properties of strongly subexponential distributions, we refer the reader to [14, Section 3.4].

Theorem 2. *Let Assumptions A1 and A3 hold. Let* v_γ *and* w_γ *be defined as in (10). For fixed* $\delta \in (0, 1)$, *there exists an* $a_* = a_*(\delta) \le 0$ *such that*

$$-\delta \le \frac{v_\gamma^2(y) - w_\gamma^2(y)}{\mathbb{P}(X_1 > -y + s(x))w_\gamma(y)} \quad \textit{for all } y \le s(x) + a_*. \tag{12}$$

Let

$$L_T(x) \triangleq \mathbf{1}_{\{T(x)<\infty\}} \prod_{k=1}^{\tau_\gamma(x)} \frac{w_\gamma(S_{k-1}(\gamma) + a_*)}{v_\gamma(S_k(\gamma) + a_*)}.$$

https://doi.org/10.1017/apr.2018.37 Published online by Cambridge University Press

Then $L_T(x)$ is an unbiased and strongly efficient estimator of $\mathbb{P}(Z > x)$, i.e.

$$\sup_{x>1} \frac{\mathbb{E}^{Q^{\gamma}_{a_*}} L^2_T(x)}{\mathbb{P}(Z > x)^2} < \infty.$$

The estimator derived in Theorem 2 requires the computation of $\mathbf{1}_{\{Z>x\}}$ and, hence, is unbiased only if we can generate Z in finite time. Generating a perfect sample from Z in our current setting is not straightforward. To address this issue, we apply the bias elimination technique introduced by Rhee and Glynn [28]. The plan for the rest of this section is as follows. First, we propose a family of simulation algorithms by approximating the path $\{Z_n\}_{n>\tau_\gamma(x)}$ with $\{Z_n\}_{\tau(x)<n\le\tau_\gamma(x)+M}$ for a fixed and sufficiently large M; we show that the latter family of simulation algorithms yields biased estimators with vanishing relative bias as $x \to \infty$. Consequently, we are able to apply the bias elimination technique and obtain an unbiased estimator that is strongly efficient and runs in finite time. To begin with, note that

$$Z = \sum_{n=0}^{\tau_\gamma(x)} B_{n+1}\mathrm{e}^{S_n} + \mathrm{e}^{S_{\tau_\gamma(x)}} \underbrace{\sum_{n=\tau_\gamma(x)+1}^{\infty} B_{n+1}\mathrm{e}^{S_n - S_{\tau_\gamma(x)}}}_{\triangleq Z'},$$

where Z' is independent of $\sum_{n=0}^{\tau_\gamma(x)} B_{n+1}\mathrm{e}^{S_n}$ and $\mathrm{e}^{S_{\tau_\gamma(x)}}$, and has the same distribution as Z. A natural choice for approximating the distribution of Z' is a truncated sum. More precisely, let $M \in \mathbb{N}$ be fixed, and our modified estimator takes the form

$$L^{\Delta}_T(x, M) = \mathbf{1}_{\{\tau_\gamma(x)<\infty,\ \sum_{n=0}^{\tau_\gamma(x)+M} B_{n+1}\mathrm{e}^{S_n}>x\}} \prod_{k=1}^{\tau_\gamma(x)} \frac{w_\gamma(S_{k-1}(\gamma)+a_*)}{v_\gamma(S_k(\gamma)+a_*)}. \tag{13}$$

We state a simulation algorithm for generating one sample of $L^{\Delta}_T(x, M)$.

Algorithm 1. The algorithm comprises five steps.

Step 1. For fixed $\delta \in (0, 1)$, set $a_* \leftarrow a_*(\delta) \le 0$ satisfying (12).

Step 2. Set $m \leftarrow 1$, $n \leftarrow 0$, $Z \leftarrow 0$, $S_n(\gamma) \leftarrow 0$, and $L_T(x) \leftarrow 1$.

Step 3. While $n < \tau_\gamma(x)$:

(i) update $S_{n+1}(\gamma) \leftarrow S_n(\gamma) + X_{n+1}(\gamma)$ by sampling the increment

$$X_{n+1}(\gamma) = (\log^+ B^+_{n+1} - \gamma_2) \vee \log A_{n+1} + \gamma_1$$

conditional on $X_{n+1}(\gamma) + W > s(x) - S_n(\gamma) - a_*$;

(ii) sample (A_{n+1}, B_{n+1}) conditional on the value of $X_{n+1}(\gamma)$;

(iii) update $L_T(x) \leftarrow L_T(x) w_\gamma(S_n(\gamma) + a_*) v_\gamma(S_{n+1}(\gamma) + a_*)^{-1}$ and $Z \leftarrow Z + B_{n+1} \prod_{i=1}^{n} A_i$;

(iv) update $n \leftarrow n + 1$.

Step 4. While $m \le M$:

(i) sample $(A_{\tau_\gamma(x)+m}, B_{\tau_\gamma(x)+m})$ under the original measure;

https://doi.org/10.1017/apr.2018.37 Published online by Cambridge University Press

(ii) update $Z \leftarrow Z + B_{\tau_\gamma(x)+m} \prod_{i=1}^{\tau_\gamma(x)+m} A_i$;

(iii) update $m \leftarrow m + 1$.

Step 5. If $Z > x$ then return $L_T^\Delta(x, M)$; otherwise, return 0.

Remark 5. In step 3(ii) of Algorithm 1, sampling (A_{n+1}, B_{n+1}) conditional on $X_{n+1}(\gamma)$ under the change of measure is equivalent to sampling it conditional on $X_{n+1}(\gamma)$ under the original measure. To see this, note that for $n + 1 \le \tau_\gamma(x)$ and any measurable set $C \subseteq \mathbb{R}^2$,

$$\begin{aligned}
&\mathbb{E}^{Q_{a_*}^\gamma}[\mathbf{1}_{(A_{n+1},\, B_{n+1})\in C} \mid S_i(\gamma), i \le n, X_{n+1}(\gamma)] \\
&\quad = \mathbb{E}\Bigg[\Bigg(\prod_{i=1}^{n+1} \frac{w_\gamma(S_{i-1}(\gamma) + a_*)}{v_\gamma(S_i(\gamma) + a_*)}\Bigg) \,\Bigg|\, S_i(\gamma),\ i \le n, X_{n+1}(\gamma)\Bigg]^{-1} \\
&\qquad \times \mathbb{E}\Bigg[\mathbf{1}_{(A_{n+1},\, B_{n+1})\in C}\Bigg(\prod_{i=1}^{n+1} \frac{w_\gamma(S_{i-1}(\gamma) + a_*)}{v_\gamma(S_i(\gamma) + a_*)}\Bigg) \,\Bigg|\, S_i(\gamma),\ i \le n, X_{n+1}(\gamma)\Bigg] \\
&\quad = \Bigg(\prod_{i=1}^{n+1} \frac{w_\gamma(S_{i-1}(\gamma) + a_*)}{v_\gamma(S_i(\gamma) + a_*)}\Bigg)^{-1} \\
&\qquad \times \Bigg(\prod_{i=1}^{n+1} \frac{w_\gamma(S_{i-1}(\gamma) + a_*)}{v_\gamma(S_i(\gamma) + a_*)}\Bigg)\mathbb{E}[\mathbf{1}_{(A_{n+1},\, B_{n+1})\in C} \mid S_i(\gamma),\ i \le n, X_{n+1}(\gamma)] \\
&\quad = \mathbb{E}[\mathbf{1}_{(A_{n+1},\, B_{n+1})\in C} \mid S_i(\gamma),\ i \le n, X_{n+1}(\gamma)] \\
&\quad = \mathbb{P}((A_{n+1},\ B_{n+1}) \in C \mid X_{n+1}(\gamma)).
\end{aligned}$$

Next we will analyze the performance of our modified estimator. In Theorem 3 we show that, under the following assumption, $\mathbb{E}^{Q_{a_*}^\gamma} L_T^\Delta(x, M)/\mathbb{P}(T(x) < \infty)$ converges to 1 as $x \to \infty$, establishing that the relative bias of L_T^Δ vanishes.

Assumption A4. *We assume that* $B_1 \ge 0$ *a.s.*

Remark 6. Under Assumption A4, Assumption A1(iii) is redundant.

Theorem 3. *Under Assumptions A1, A3, and A4,* $L_T^\Delta(x, M)$ *as in (13) is asymptotically unbiased as* $x \to \infty$*, i.e. we have*

$$\lim_{x\to\infty} \frac{\mathbb{E}^{Q_{a_*}^\gamma} L_T^\Delta(x, M)}{\mathbb{P}(T(x) < \infty)} = 1 \quad \textit{uniformly in } M \in \mathbb{N}.$$

We are now ready to apply the bias elimination technique in Result 3 to the estimators proposed in (13) as mentioned in the paragraph above Algorithm 1. By analyzing the asymptotic behavior of the relative bias as $M \to \infty$ for fixed x (see Lemma 5), we are able to apply the bias elimination technique and obtain an unbiased estimator for $\mathbb{P}(Z > x)$. We introduce the following assumption; the unbiased estimator is then given in Theorem 4.

Assumption A5. *(i) The Markov chain* $\{Z_n\}_{n\in\mathbb{N}}$ *given by (2) is irreducible and aperiodic.*

(ii) There exists $q \ge 2$ *such that* $\mathbb{E}|\log A_1|^q + \mathbb{E}|\log B_1^+|^q < \infty$.

Remark 7. Assumption A5(i) is satisfied, e.g. if (A_1, B_1) has a Lebesgue density; see [8, Lemma 2.2.2].

Theorem 4. *Let Assumptions A1 and A3–A5 hold. Let v_γ and w_γ be defined as in (10). For fixed $\delta \in (0, 1)$ and $\beta \in (0, 1)$, there exists an $a_* = a_*(\delta) \le 0$ satisfying*

$$-\delta \le \frac{v_\gamma^{(2-\beta)/(1-\beta)}(y) - w_\gamma^{(2-\beta)/(1-\beta)}(y)}{\mathbb{P}(X_1 > -y + s(x)) w_\gamma^{1/(1-\beta)}(y)} \quad \text{for all } y \le s(x) + a_*.$$

Moreover, it is possible to construct a random variable N independent of x such that

$$\sum_{n=0}^{\infty} \frac{\mathbb{E}^{Q_{a_*}^{\gamma}}(L_T^{\Delta}(x, 2^{n-1}) - L_T(x))^2}{\mathbb{P}(Z > x)^2 \mathbb{P}(N \ge n)} < \infty$$

and, hence, the estimator (see Rhee and Glynn [28])

$$L_T^{\mathrm{RG}}(x) \triangleq \sum_{n=0}^{N} \frac{L_T^{\Delta}(x, 2^n) - L_T^{\Delta}(x, 2^{n-1})}{\mathbb{P}(N \ge n)}$$

with L_T^{Δ} as in (13) is unbiased and strongly efficient.

Remark 8. As we will see in the proof of Theorem 4, one possible choice is to sample N with $\mathbb{P}(N \le n) = 1 - (1-p)^n$ for $n \ge 1$, where $p < 1 - 2^{-(q-1)}$ and q is as in Assumption A5(ii). In general, the bias elimination scheme of [28] is not guaranteed to produce nonnegative estimators, which might not be ideal in the context of estimating (rare event) probabilities. However, in our case $L_T^{\Delta}(x, M)$ increases w.r.t. M and, hence, the resulting unbiased estimator $L_T^{\mathrm{RG}}(x)$ is always nonnegative.

3.2. Iterated random functions

We consider the Markov chain $\{Z_n\}_{n\ge 0}$ with $Z_{n+1} = \Psi_n(Z_n)$, where Ψ_n satisfies the following assumption. For similar settings in which Markov chains generated by iterated random functions are analyzed, see, e.g. [9], [12], and [16].

Assumption B0. *We assume that $\{\Psi_n\}_{n\in\mathbb{N}}$ is a sequence of i.i.d. random Lipschitz functions with*

$$\mathrm{Lip}(\Psi_n) \triangleq \sup_{z_1 \ne z_2} \left| \frac{\Psi_n(z_1) - \Psi_n(z_2)}{z_1 - z_2} \right|. \tag{14}$$

Moreover, there exists a sequence of i.i.d. random vectors $\{(A_n, B_n, D_n)\}_{n\in\mathbb{N}}$ such that

$$A_n z + B_n - D_n \le \Psi_n(z) \le A_n z^+ + B_n^+ + D_n \quad \text{for all } z \in \mathbb{R}. \tag{15}$$

In addition, we can sample Ψ_n from the conditional distribution, given $(\log^+(B_n^+ + D_n) - \gamma_2) \vee \log A_n$ for γ_2 as in Lemma 3 below.

The goal of this section is to extend the results in Section 3.1 to the setting as described above. To achieve this, we introduce a list of additional assumptions that are extensions of Assumptions A1–A5. To begin with, we consider an extension of Assumption A1.

Assumption B1. *Assume that (15) holds and (A_1, B_1, D_1) satisfies the following conditions.*

(i) *$A_1, D_1 > 0$ a.s., $\mathbb{E}\log A_1 > -\infty$, and $\mathbb{E}\log \mathrm{Lip}(\Psi_1) < 0$. Moreover, $\mathbb{E}\log^+|B_1 + D_1| < \infty$ and $\mathbb{E}\log^+|B_1 - D_1| < \infty$.*

(ii) $\mathbb{E}\log^+(A_1 \vee B_1) < \infty$.

https://doi.org/10.1017/apr.2018.37 Published online by Cambridge University Press

(iii) *Let the following tail behaviors hold:*

$$\begin{aligned}\mathbb{P}(\max(A_1, B_1 + D_1) > x) &\sim \mathbb{P}(\max(A_1, B_1) > x),\\ \mathbb{P}(\max(A_1, B_1 - D_1) > x) &\sim \mathbb{P}(\max(A_1, B_1) > x),\\ \mathbb{P}(A_1 > x, B_1 - D_1 \le -x) &= o(\mathbb{P}(\max(A_1, B_1) > x)).\end{aligned}$$

Define $\Psi_{1:n}(z) \triangleq \Psi_1 \circ \Psi_2 \circ \cdots \circ \Psi_n(z)$ for each $z \in \mathbb{R}$, and $Z \triangleq \lim_{n\to\infty} \Psi_{1:n}(Z_0)$. Recall that (see [12, Theorem 3.1]) under Assumption B1, the unique stationary solution to (1) exists, is finite, has the same distribution as Z, and has right-unbounded support. Moreover, the distribution of Z does not depend on the initial condition Z_0. Thus, without loss of generality we set $Z_0 = 0$. Note that Z can be bounded from above by a stochastic perpetuity

$$\bar{Z} \triangleq \sum_{n=0}^{\infty} \bar{B}_{n+1} \mathrm{e}^{S_n},$$

where $\bar{B}_n \triangleq \max(B_n^+ + D_n, 1)$ and $S_n = \sum_{i=1}^n \log A_i$. Analogous to the previous section, we construct an upper bound for $\bar{Z}$ (and thus for Z) that can be written as a functional of the maximum of a suitable random walk $S_n(\gamma)$.

Lemma 3. *Under Assumption B1, there exists a constant γ_2 such that*

$$\mathbb{E}[\max(\log^+(B_1^+ + D_1) - \gamma_2, \log A_1)] < 0.$$

Moreover, there exists a constant $\gamma_1 \in (0, -\mathbb{E}[\log A_i \vee (\log^+ B_1 - \gamma_2)])$ such that

$$Z \le \exp\Bigl(\max_{n\ge 0} S_n(\gamma)\Bigr) \frac{\mathrm{e}^{\gamma_2}}{1 - \mathrm{e}^{-\gamma_1}} < \infty, \tag{16}$$

where $S_n(\gamma) = S_n(\gamma_1, \gamma_2) = \sum_{i=1}^n [\log A_i \vee (\log^+ (B_i^+ + D_i) - \gamma_2) + \gamma_1]$.

Let $S_n(\gamma)$ be as in Lemma 3. Now from (16) we can define $s(x) \triangleq \log x - \gamma_2 + \log(1 - \mathrm{e}^{-\gamma_1})$ and $\tau_\gamma(x) \triangleq \inf\{n \ge 0 \colon S_n(\gamma) > s(x)\}$ such that (9) holds. Thanks to [12], under subexponential assumptions on the random variable $\log(A_1 \vee B_1)$, the tail asymptotics can be described using the integrated tail function of $\log(A_1 \vee B_1)$. However, the upper bound derived in Lemma 3 yields a shorter proof for the asymptotic upper bound in [12, Theorem 3.1].

Assumption B2. *The integrated tail of* $\log(A_1 \vee B_1)$*, denoted by $\bar{F}_I$, is subexponential.*

Theorem 5. *If Assumptions B1 and B2 hold, we have*

$$\limsup_{x\to\infty} \frac{\mathbb{P}(Z > x)}{\bar{F}_I(\log(x))} \le -\frac{1}{\mathbb{E}\log A_1}.$$

For fixed $a_* \le 0$, and v_γ and w_γ as in (10), recall that $\mathbb{P}^{Q^\gamma_{a_*}}$ and $\mathbb{E}^{Q^\gamma_{a_*}}$ denote respectively the probability measure and the expectation w.r.t. the stochastic process $\{S_n(\gamma)\}_{n\in\mathbb{N}}$ having a one-step transition kernel $Q^\gamma_{a_*}$ as in (11). Given the asymptotic behavior of $\mathbb{P}(Z > x)$, in Theorem 6 we show the strong efficiency (under $\mathbb{P}^{Q^\gamma_{a_*}}$) of our estimator.

Assumption B3. *The distribution of* $\log(A_1 \vee B_1)$ *belongs to the class S^*.*

Theorem 6. *Let Assumptions B1 and B3 hold, and v_γ and w_γ be as in (10). For fixed $\delta \in (0, 1)$, we can choose $a_* = a_*(\delta) \le 0$ such that (12) holds. Let*

$$L_T(x) \triangleq \mathbf{1}_{\{T(x)<\infty\}} \prod_{k=1}^{\tau_\gamma(x)} \frac{w_\gamma(S_{k-1}(\gamma) + a_*)}{v_\gamma(S_k(\gamma) + a_*)}.$$

Then $L_T(x)$ is a strongly efficient estimator of $\mathbb{P}(Z > x)$ under $\mathbb{P}^{Q^\gamma_{a_*}}$.

Recall that $\Psi_{1:n}(z) \triangleq \Psi_1 \circ \Psi_2 \circ \cdots \circ \Psi_n(z)$ for each $z \in \mathbb{R}$. Define

$$L^\Delta_T(x, M) \triangleq \mathbf{1}_{\{\tau_\gamma(x)<\infty,\, \Psi_{1:\tau_\gamma(x)+M}(Z_0)>x\}} \prod_{k=1}^{\tau_\gamma(x)} \frac{w_\gamma(S_{k-1}(\gamma) + a_*)}{v_\gamma(S_k(\gamma) + a_*)}. \tag{17}$$

As in Section 3.1, we approximate $L_T(x)$ by $L^\Delta_T(x, M)$. Analogous to Theorem 3, in Theorem 7 below we show that the estimator as in (17) is asymptotically unbiased.

Assumption B4. *For each z, $\Psi_{1:n}(z)$ is increasing in n.*

Theorem 7. *Under Assumptions B1, B3, and B4, $L^\Delta_T(x, M)$ in (17) is asymptotically unbiased as $x \to \infty$, i.e. we have*

$$\lim_{x\to\infty} \frac{\mathbb{E}^{Q_{a_*}} L^\Delta_T(x, M)}{\mathbb{P}(T(x) < \infty)} = 1 \quad \textit{uniformly in } M \in \mathbb{N}.$$

Applying again Result 3, in Theorem 8 below we construct an unbiased estimator for estimating $\mathbb{P}(Z > x)$. To do this, we need the following assumptions.

Assumption B5. *(i) The Markov chain $\{Z_n\}_{n\in\mathbb{N}}$ given by (1) is irreducible and aperiodic.*

(ii) There exists $q \ge 2$ such that $\mathbb{E}|\log A_1|^q + \mathbb{E}|\log B_1^+|^q + \mathbb{E}|\log D_1|^q < \infty$.

Assumption B6. *There exists $\underline{z}$ such that $\Psi_n([\underline{z}, \infty)) \subseteq [\underline{z}, \infty)$ and Ψ_n is bijective on $[\underline{z}, \infty)$ a.s.*

Remark 9. Assumptions B4 and B6 are satisfied if, for instance, the stochastic equation is

$$Z_{n+1} = \sqrt{A_{n+1} Z_n^2 + B_{n+1} Z_n + C_{n+1}}.$$

This corresponds to a second-order random polynomial equation, [16]. Other examples are, for instance, $\Psi_n(z) = \max\{A_n z, B_n\}$ and $\Psi_n(z) = A_n \max\{z, B_n\} + C_n$.

Theorem 8. *Let Assumptions B1 and B3–B6 hold. Let v_γ and w_γ be as in (10). For fixed $\delta \in (0, 1)$ and $\beta \in (0, 1)$, there exists an $a_* = a_*(\delta) \le 0$ satisfying*

$$-\delta \le \frac{v_\gamma^{(2-\beta)/(1-\beta)}(y) - w_\gamma^{(2-\beta)/(1-\beta)}(y)}{\mathbb{P}(X_1 > -y + s(x)) w_\gamma^{1/(1-\beta)}(y)} \quad \textit{for all } y \le s(x) + a_*.$$

Then it is possible to construct a random variable N independent of x such that

$$\sum_{n=0}^{\infty} \frac{\mathbb{E}^{Q^\gamma_{a_*}} (L^\Delta_T(x, 2^{n-1}) - L_T(x))^2}{\mathbb{P}(Z > x)^2 \mathbb{P}(N \ge n)} < \infty$$

https://doi.org/10.1017/apr.2018.37 Published online by Cambridge University Press

and, hence, the estimator $L_T^{\mathrm{RG}}(x)$ *defined by*

$$L_T^{\mathrm{RG}}(x) \triangleq \sum_{n=0}^{N} \frac{L_T^{\Delta}(x, 2^n) - L_T^{\Delta}(x, 2^{n-1})}{\mathbb{P}(N \geq n)}$$

with $L_T^{\Delta}(x, M)$ *as in (17) is unbiased and strongly efficient.*

Remark 10. As in Remark 8, N can be chosen such that $\mathbb{P}(N \leq n) = 1 - (1 - p)^n$ for $n \geq 1$, where $p < 1 - 2^{-(q-1)}$ and q is as in Assumption B5(ii).

4. Numerical results

Here we investigate our algorithm numerically, based on a stochastic perpetuity with $B_n = 1$. We consider the increment $\log A_n \overset{\mathrm{D}}{=} \mathcal{W} - \frac{3}{2}$, where $\mathcal{W}$ is a random variable with Weibull distribution, i.e.

$$\mathbb{P}(\mathcal{W} > t) = \exp(-2t^{1/2}).$$

For the algorithmic parameters, we choose $a_* = -10$ and $\gamma = \frac{1}{2}$. Moreover, we use a geometric distributed random truncation index with parameter $\frac{1}{2}$. In Figure 1 we present the change of the estimated probability w.r.t the different choices of M for the four different values $x = 10^8$, $x = 10^{16}$, $x = 10^{32}$, and $x = 10^{64}$ in each of the four plots. We see that the estimated probability stabilizes as M grows, which confirms that our estimator is consistent as $M \to \infty$. Comparing the four plots, we see that the initial bias for small M decreases as x increases, which is consistent with the conclusion of Theorem 3 (vanishing relative bias). In Table 1 we present the estimated probabilities, their 95% confidence intervals, and the estimated coefficients of variation, i.e. the estimated standard deviation divided by the sample mean (based on 200 000

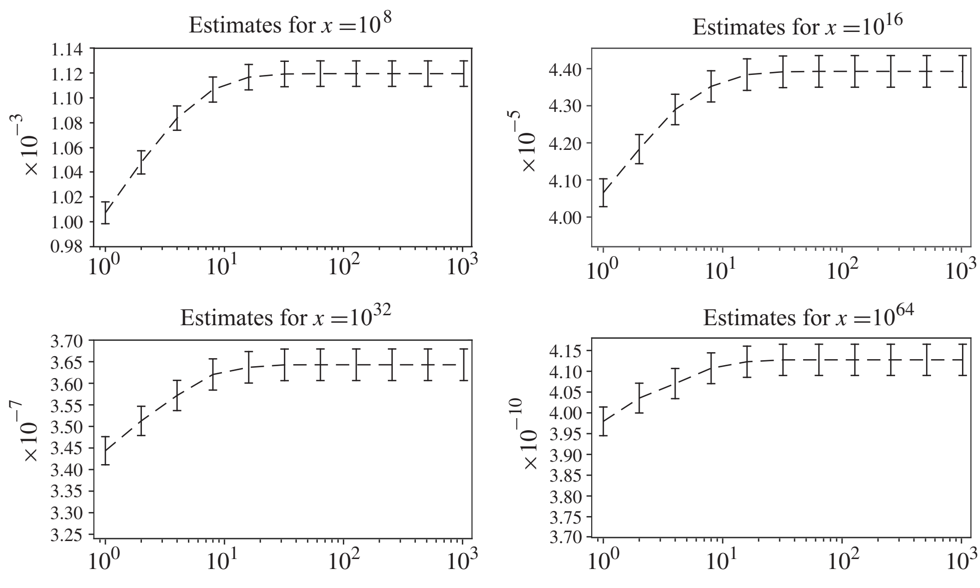

FIGURE 1: Estimated probabilities for changing values of M. The y-axis values indicate the estimated rare-event probabilities and the vertical bars indicate the 95% confidence intervals. The x-axis values indicate the truncation index M.

https://doi.org/10.1017/apr.2018.37 Published online by Cambridge University Press

TABLE 1: Estimated rare-event probability (Est), 95% confidence intervals (CI), and the estimated coefficients of variation (CV).

M	$x = 10^8$			$x = 10^{16}$		
	Est	CI	CV	Est	CI	CV
2^2	1.083×10^{-3}	$\pm 0.009 \times 10^{-3}$	2.06	4.271×10^{-5}	$\pm 0.041 \times 10^{-5}$	2.17
2^4	1.117×10^{-3}	$\pm 0.010 \times 10^{-3}$	2.10	4.373×10^{-5}	$\pm 0.042 \times 10^{-5}$	2.22
2^6	1.120×10^{-3}	$\pm 0.010 \times 10^{-3}$	2.10	4.383×10^{-5}	$\pm 0.043 \times 10^{-5}$	2.22
2^8	1.120×10^{-3}	$\pm 0.010 \times 10^{-3}$	2.10	4.383×10^{-5}	$\pm 0.043 \times 10^{-5}$	2.22
RG	1.119×10^{-3}	$\pm 0.013 \times 10^{-3}$	2.70	4.375×10^{-5}	$\pm 0.053 \times 10^{-5}$	2.76
M	$x = 10^{32}$			$x = 10^{64}$		
	Est	CI	CV	Est	CI	CV
2^2	3.583×10^{-7}	$\pm 0.035 \times 10^{-7}$	2.25	4.079×10^{-10}	$\pm 0.037 \times 10^{-10}$	2.05
2^4	3.646×10^{-7}	$\pm 0.037 \times 10^{-7}$	2.28	4.120×10^{-10}	$\pm 0.037 \times 10^{-10}$	2.06
2^6	3.650×10^{-7}	$\pm 0.037 \times 10^{-7}$	2.29	4.123×10^{-10}	$\pm 0.038 \times 10^{-10}$	2.06
2^8	3.650×10^{-7}	$\pm 0.037 \times 10^{-7}$	2.29	4.123×10^{-10}	$\pm 0.038 \times 10^{-10}$	2.06
RG	3.663×10^{-7}	$\pm 0.045 \times 10^{-7}$	2.81	4.115×10^{-10}	$\pm 0.041 \times 10^{-10}$	2.27

samples) for different values of x and M. In the last column, we present the results produced with the unbiased algorithm as introduced in Theorem 4. We can see that on the one hand, the ratio between the estimated probability and the standard deviation stays roughly constant over a range of x values and M values; on the other hand, the estimated probability using the fixed truncation method tends to converge to the estimated probability produced with the unbiased algorithm as M grows. These observations illustrate the strong efficiency (Theorems 2 and 4) of our estimators.

5. Proofs

In this section we provide proofs of the results presented in this paper. Let

$$\tilde{v}(z) \triangleq \mathbb{P}(W > -z), \qquad \tilde{w}(z) \triangleq \mathbb{P}(X_1 + W > -z), \qquad \tilde{Q}(y, \mathrm{d}z) \triangleq \frac{P(y, \mathrm{d}z)\tilde{v}(z)}{\tilde{w}(y)}.$$

For $y \leq 0$, let $\mathbb{E}_y^{\tilde{Q}}$ denote the expectation operator associated with $\tilde{S}_n \triangleq y + S_n$ having the transition kernel $\tilde{Q}$, conditional on $\tilde{S}_0 = y$. Let $\Gamma = \inf\{n \geq 0 \colon \tilde{S}_n > 0\}$.

Lemma 4. *Let $\varepsilon > 0$ be given. Suppose that there exist constants $\delta_1, \delta_2 > 0$ and a finite-valued function $h \colon \mathbb{R} \longrightarrow [\delta_1, \infty)$ such that*

$$\tilde{w}^{1+\varepsilon}(y) \int \tilde{v}(z) h(z) P(y, \mathrm{d}z) \leq h(y) \tilde{v}^{2+\varepsilon}(y) \quad \textit{for } y \leq 0. \tag{18}$$

If $h(z) \geq 1$ for $z > 0$ and $\tilde{v}(z) \geq \delta_2 > 0$ for $z > 0$, we have

$$\mathbb{E}_y^{Q'} \mathbf{1}_{\{\Gamma < \infty\}} \prod_{k=1}^{\Gamma} \frac{\tilde{w}^{2+\varepsilon}(\tilde{S}_{k-1})}{\tilde{v}^{2+\varepsilon}(\tilde{S}_k)} \leq \delta_1^{-1} \delta_2^{-(2+\varepsilon)} \tilde{v}^{2+\varepsilon}(y) h(y) \quad \textit{for } y \leq 0.$$

https://doi.org/10.1017/apr.2018.37 Published online by Cambridge University Press

Proof. Let $\mathbb{E}_y$ denote the expectation operator associated with $\{S_n\}_{n\geq 0}$ having the transition kernel P, conditional on $\tilde{S}_0 = y$. Recall [5, Theorem 2(iii)], where it was proved that if there exists a finite-valued nonnegative function $\tilde{h}$ such that

$$(K\tilde{h})(y) \leq \tilde{h}(y) - \eta(y) \quad \text{for } y \leq 0,$$

where

$$(K\tilde{h})(y) = \int_{(-\infty,0]} \tilde{h}(z)\frac{\tilde{w}^{1+\varepsilon}(y)}{\tilde{v}^{1+\varepsilon}(z)}P(y,\, \mathrm{d}z) \quad \text{and} \quad \eta(y) = \int_{(0,\infty)} \frac{\tilde{w}^{1+\varepsilon}(y)}{\tilde{v}^{1+\varepsilon}(z)}P(y,\, \mathrm{d}z),$$

then

$$\mathbb{E}_y \mathbf{1}_{\{\Gamma<\infty\}} \prod_{k=1}^{\Gamma} \frac{\tilde{w}^{1+\varepsilon}(\tilde{S}_{k-1})}{\tilde{v}^{1+\varepsilon}(\tilde{S}_k)} \leq \tilde{h}(y) \quad \text{for } y \leq 0. \tag{19}$$

Define $\tilde{h}(\cdot) = \delta_1^{-1}\delta_2^{-(2+\varepsilon)}h(\cdot)v^{2+\varepsilon}(\cdot)$. Note that

$$\begin{aligned}
&\delta_1^{-1}\delta_2^{-(2+\varepsilon)}\tilde{w}^{1+\varepsilon}(y)\int \tilde{v}(z)h(z)P'(y, \mathrm{d}z)\\
&\quad = \delta_1^{-1}\delta_2^{-(2+\varepsilon)}\tilde{w}^{1+\varepsilon}(y)\bigg(\int_{(-\infty,0]} + \int_{(0,\infty)}\bigg)\tilde{v}(z)h(z)P(y, \mathrm{d}z)\\
&\quad = (K\tilde{h})(y) + \delta_1^{-1}\delta_2^{-(2+\varepsilon)}\tilde{w}^{1+\varepsilon}(y)\int_{(0,\infty)} \tilde{v}(z)h(z)P(y, \mathrm{d}z).
\end{aligned}$$

Thus, (18) is equivalent to

$$(K\tilde{h})(y) \leq \tilde{h}(y) - \delta_1^{-1}\delta_2^{-(2+\varepsilon)}\tilde{w}^{1+\varepsilon}(y)\int_{(0,\infty)} \tilde{v}(z)h(z)P(y, \mathrm{d}z). \tag{20}$$

On the other hand, we have

$$\begin{aligned}
\eta(y) &= \int_{(0,\infty)} \frac{\tilde{w}^{1+\varepsilon}(y)}{\tilde{v}^{1+\varepsilon}(z)}P(y, \mathrm{d}z)\\
&\leq \delta_2^{-(2+\varepsilon)}\tilde{w}^{1+\varepsilon}(y)\int_{(0,\infty)} \tilde{v}(z)P(y, \mathrm{d}z)\\
&\leq \delta_1^{-1}\delta_2^{-(2+\varepsilon)}\tilde{w}^{1+\varepsilon}(y)\int_{(0,\infty)} h(z)\tilde{v}(z)P(y, \mathrm{d}z).
\end{aligned} \tag{21}$$

Using (20) and (21), (18) implies that

$$(K\tilde{h})(y) \leq \tilde{h}(y) - \eta(y).$$

From [5, Theorem 2(iii)], we conclude that (19) holds, as required. □

Proof of Lemma 1. We first find h that satisfies (18) in Lemma 4. Define

$$h(y) = 1 - \delta\mathbf{1}_{\{y>-a_*\}}.$$

https://doi.org/10.1017/apr.2018.37 Published online by Cambridge University Press

We will find a suitable $a_* \le 0$ later. Note that (18) in this case becomes

$$\tilde{w}(y+a_*)^{-1}\mathbb{E}\tilde{v}(X_1+y+a_*)h(X+y) \le h(y)\biggl(\frac{\tilde{v}(y+a_*)}{\tilde{w}(y+a_*)}\biggr)^{2+\varepsilon} \le \biggl(\frac{\tilde{v}(y+a_*)}{\tilde{w}(y+a_*)}\biggr)^{2+\varepsilon} \quad \text{for } y \le 0. \tag{22}$$

Here we use the fact that $h(y)=1$ for $y \le 0$. By the definition of $\tilde{v}$,

$$\begin{aligned}
&\mathbb{E}\tilde{v}(X_1+y+a_*)h(X_1+y) \\
&\quad = \mathbb{E}[\mathbb{E}[\mathbf{1}_{\{W>-X_1-y-a_*\}} \mid X_1]] - \delta\mathbb{E}[\mathbb{E}[\mathbf{1}_{\{W>-X_1-y-a_*\}} \mid X_1]\mathbf{1}_{\{X_1+y>-a_*\}}] \\
&\quad = \mathbb{P}(W+X_1>-y-a_*) - \delta\mathbb{P}(W+X_1>-y-a_*,\, X_1+y>-a_*).
\end{aligned}$$

Since $\tilde{w}(y+a_*) = \mathbb{P}(X_1+W>-y-a_*)$, (22) is equivalent to

$$1-\delta\mathbb{P}(X_1+y>-a_* \mid W+X_1>-y-a_*) \le \frac{\tilde{v}^{2+\varepsilon}(y+a_*)}{\tilde{w}^{2+\varepsilon}(y+a_*)},$$

which is equivalent to

$$\begin{aligned}
&-\delta \le \frac{\tilde{v}^{2+\varepsilon}(y+a_*)-\tilde{w}^{2+\varepsilon}(y+a_*)}{\mathbb{P}(X_1>-y-a_*)\tilde{w}^{1+\varepsilon}(y+a_*)} \quad \text{for all } y \le 0, \\
&\iff \quad -\delta \le \frac{v^{2+\varepsilon}(x+y+a_*)-w^{2+\varepsilon}(x+y+a_*)}{\mathbb{P}(X_1>-y-a_*)w^{1+\varepsilon}(x+y+a_*)} \quad \text{for all } y \le 0, \\
&\iff \quad -\delta \le \frac{v^{2+\varepsilon}(y)-w^{2+\varepsilon}(y)}{\mathbb{P}(X_1>-y+x)w^{1+\varepsilon}(y)} \quad \text{for all } y \le x+a_*.
\end{aligned} \tag{23}$$

Using the definition of w and the nonnegativity of W, (7) implies that $w(y)-v(y)=o(w(y))$ and, hence, $v(y) \sim w(y)$ as $y \to -\infty$. Therefore, there exists an a_* satisfying (23) and, hence, (18). Since $\inf_{z\ge 0}\tilde{v}(z+a_*) = \mathbb{P}(W>-a_*)$, Lemma 4 applies and we have

$$\mathbb{E}\mathbf{1}_{\{\tilde{\tau}(0)<\infty\}}\prod_{k=1}^{\tilde{\tau}(0)}\frac{\tilde{w}^{1+\varepsilon}(\tilde{S}_{k-1}+a_*)}{\tilde{v}^{1+\varepsilon}(\tilde{S}_k+a_*)} \le \delta^{-1}\mathbb{P}(W>-a_*)^{-(2+\varepsilon)}\tilde{v}^{2+\varepsilon}(y) \quad \text{for } y \le 0.$$

Recall the Pakes–Veraverbeke theorem (see [29] and [31]), i.e.

$$\mathbb{P}\Bigl(\max_{n\ge 0}\tilde{S}_n>0\Bigr) \sim -\frac{1}{\mathbb{E}X_1}\int_{-y}^{\infty}\mathbb{P}(X_1>t)\,\mathrm{d}t \quad \text{as } y \to -\infty.$$

This implies that for any fixed y,

$$\mathbb{P}\Bigl(\max_{n\ge 0}\tilde{S}_n>0\Bigr) \sim \tilde{v}(y) \quad \text{as } y \to -\infty.$$

Combining this with the fact that $\mathbb{P}(\max_{n\ge 0}\tilde{S}_n>0)/\tilde{v}(y)$ is bounded as a function of y on compact sets, we obtain

$$\sup_{y<0}\mathbb{P}\Bigl(\max_{n\ge 0}\tilde{S}_n>0\Bigr)^{-2}\mathbb{E}\mathbf{1}_{\{\tilde{\tau}(0)<\infty\}}\prod_{k=1}^{\tilde{\tau}(0)}\frac{\tilde{w}^{1+\varepsilon}(\tilde{S}_{k-1}+a_*)}{\tilde{v}^{1+\varepsilon}(\tilde{S}_k+a_*)} < \infty,$$

https://doi.org/10.1017/apr.2018.37 Published online by Cambridge University Press

which is equivalent to

$$\sup_{x>0} \frac{\mathbb{E}^{Q_{a*}} L_\tau^{2+\varepsilon}(x)}{\mathbb{P}(\max_{n\geq 0} S_n > x)^{2+\varepsilon}} < \infty. \qquad \square$$

Proof of Lemma 2. Note that $\max\{((\log^+ B_1 - \gamma_2') \vee \log A_1), 0\} \leq |\log^+ B_1 \vee \log A_1|$, and $\min\{((\log^+ B_1 - \gamma_2') \vee \log A_1), 0\}$ is bounded from above and nonincreasing w.r.t. γ_2'. Since $(\log^+ B_1 - \gamma_2') \vee \log A_1 = \max\{((\log^+ B_1 - \gamma_2') \vee \log A_1), 0\} + \min\{((\log^+ B_1 - \gamma_2') \vee \log A_1), 0\}$, we can apply bounded convergence for the maximum and monotone convergence for minimum to obtain

$$\lim_{\gamma_2' \to \infty} \mathbb{E}[(\log^+ B_1 - \gamma_2') \vee \log A_1] = \mathbb{E} \lim_{\gamma_2' \to \infty} (\log^+ B_1 - \gamma_2') \vee \log A_1 = \mathbb{E} \log A_1 < 0.$$

Therefore, there exists γ_2 such that $\mathbb{E}[(\log^+ B_1 - \gamma_2') \vee \log A_1] < 0$.

Now we have

$$Z \leq \sum_{n=0}^{\infty} \max(B_{n+1}, 1) \mathrm{e}^{S_n} = \mathrm{e}^{\gamma_2} \sum_{n=0}^{\infty} \mathrm{e}^{(\log^+ B_{n+1} - \gamma_2) + S_n} \leq \mathrm{e}^{\gamma_2} \sum_{n=0}^{\infty} \mathrm{e}^{S_n'}, \tag{24}$$

where $S_n' = S_{n-1}' + (\log^+ B_n - \gamma_2) \vee \log A_n$. Note that the last inequality can be checked by comparing S_{n+1}' with $(\log^+ B_{n+1} - \gamma_2) + S_n$ for each n, i.e.

$$\begin{aligned}(\log^+ B_{n+1} - \gamma_2) + S_n &= (\log^+ B_{n+1} - \gamma_2) + \sum_{k=1}^{n} \log A_k \\ &\leq (\log^+ B_{n+1} - \gamma_2) \vee \log A_{n+1} + \sum_{k=1}^{n} (\log^+ B_k - \gamma_2) \vee \log A_k \\ &= S_{n+1}'.\end{aligned}$$

Now fix $\gamma_1 \in (0, -\mathbb{E}[(\log^+ B_1 - \gamma_2) \vee \log A_1])$. From (24), we see that

$$Z \leq \mathrm{e}^{\gamma_2} \sum_{n=0}^{\infty} \mathrm{e}^{S_n'} = \mathrm{e}^{\gamma_2} \sum_{n=0}^{\infty} \mathrm{e}^{S_n' + n\gamma_1} \mathrm{e}^{-n\gamma_1} \leq \exp\Big(\max_{n\geq 0} S_n(\gamma)\Big) \frac{\mathrm{e}^{\gamma_2}}{1 - \mathrm{e}^{-\gamma_1}},$$

where $\gamma = (\gamma_1, \gamma_2)$, $S_n' = S_{n-1}' + (\log^+ B_n - \gamma_2) \vee \log A_n$, and $S_n(\gamma) = S_n' + n\gamma_1$. Note that $\mathbb{E} S_1(\gamma) < 0$ by the choice of γ_1. Hence, $\max_{n\geq 0} S_n(\gamma)$ is finite a.s. $\square$

Proof of Theorem 1. From the upper bound constructed in Lemma 2, we know that

$$\mathbb{P}(Z > x) \leq \mathbb{P}\Big(\max_{n\geq 0} S_n(\gamma) > s(x)\Big). \tag{25}$$

Due to Assumption A2, the integrated tail of $\log^+(A_1 \vee B_1^+)$ is also subexponential. Moreover, it is straightforward to check that

$$\log^+(A_1 \vee B_1^+) - \gamma_2 \leq \log(\max\{A_1, \mathrm{e}^{-\gamma_2} B_1^+, \mathrm{e}^{-\gamma_2}\}) \leq \log^+(A_1 \vee B_1^+).$$

Therefore, the increments of the random walk $S_n(\gamma)$ have a subexponential integrated tail. Using the Pakes–Veraverbeke theorem, we obtain the relationship for the right-hand side of (25), i.e.

$$\mathbb{P}\Big(\max_{n\geq 0} S_n(\gamma) > s(x)\Big) \sim -\frac{1}{\mathbb{E}[(\log^+ B_1^+ - \gamma_2) \vee \log A_1] + \gamma_1} \bar{F}_I(\log(x)). \tag{26}$$

https://doi.org/10.1017/apr.2018.37 Published online by Cambridge University Press

Thus,

$$\limsup_{x\to\infty} \frac{\mathbb{P}(Z > x)}{\bar{F}_I(\log x)} \le -\frac{1}{\mathbb{E}[(\log^+ B_1^+ - \gamma_2) \vee \log A_1] + \gamma_1}.$$

Now letting $\gamma_2 \to \infty$ and $\gamma_1 \to 0$, the result follows. □

Proof of Theorem 2. Let

$$M_n^{-1} = \prod_{k=1}^{n} \frac{w_\gamma(S_{k-1}(\gamma) + a_*)}{v_\gamma(S_k(\gamma) + a_*)}.$$

Obviously, $\{M_n\}_{n\in\mathbb{N}}$ is a martingale and, therefore, $\{M_{n\wedge\tau_\gamma(x)}\}_{n\in\mathbb{N}}$ is also a martingale. Since $\tau_\gamma(x) \le T(x)$ we can apply Lemma 1 to obtain

$$\mathbb{E}^{Q^\gamma_{a*}} L_T(x) = \mathbb{P}(T(x) < \infty) = \mathbb{P}(Z > x).$$

For the strong efficiency, we have

$$\begin{aligned}
\frac{\mathbb{E}^{Q^\gamma_{a*}} L_T^2(x)}{\mathbb{P}(Z > x)^2} &= \frac{\mathbb{E}^{Q^\gamma_{a*}} \mathbf{1}_{\{Z>x\}} M_{\tau_\gamma}^{-2}(x)}{\mathbb{P}(Z > x)^2} \\
&\le \frac{\mathbb{E}^{Q^\gamma_{a*}} \mathbf{1}_{\{\max_{n\ge 0} S_n(\gamma) > s(x)\}} M_{\tau_\gamma}^{-2}(x)}{\mathbb{P}(Z > x)^2} \\
&= \frac{\mathbb{E}^{Q^\gamma_{a*}} \mathbf{1}_{\{\max_{n\ge 0} S_n(\gamma) > s(x)\}} M_{\tau_\gamma}^{-2}(x)}{\mathbb{P}(\max_{n\ge 0} S_n(\gamma) > s(x))^2} \left(\frac{\mathbb{P}(\max_{n\ge 0} S_n(\gamma) > s(x))}{\mathbb{P}(Z > x)}\right)^2, \qquad (27)
\end{aligned}$$

where the first term in the last equation is guaranteed to be bounded over $x \in (1, \infty)$ due to Result 2. Hence, only the latter term remains to be analyzed. From [12, Theorem 3.1], we have

$$\liminf_{x\to\infty} \frac{\mathbb{P}(Z > x)}{\bar{F}_I(\log(x))} \ge -\frac{1}{\mathbb{E}\log A_1}. \qquad (28)$$

Since, by assumption, the integrated tail $\bar{F}_I$ is subexponential, it is, in particular, long tailed. Combining (26) and (28), we obtain

$$\limsup_{x\to\infty} \frac{\mathbb{P}(\max_{n\ge 0} S_n(\gamma) > s(x))}{\mathbb{P}(Z > x)} \le \frac{\mathbb{E}\log A_1}{\mathbb{E}[\max(\bar{B}_1 - \gamma_2, \log A_1)] + \gamma_1}. \qquad (29)$$

Using the fact that the left-hand side of (29) is bounded over a compact interval, we obtain the result. □

Proof of Theorem 3. Note that the Markov chain we considered in Section 3.1 is a special case of (1). Thus, we refer the reader to the proof of Theorem 7 for details. □

Proof of Theorem 4. We wish to apply Result 3 to the sequence $\{L_T^\Delta(x, 2^n)\}_{n\in\mathbb{N}}$. Therefore, we need to check the existence of a random variable N such that

$$\sum_{n=0}^{\infty} \frac{\mathbb{E}^{Q^\gamma_{a*}}[(L_T^\Delta(x, 2^{n-1}) - L_T(x))^2] - \mathbb{E}^{Q^\gamma_{a*}}[(L_T^\Delta(x, 2^n) - L_T(x))^2]}{\mathbb{P}(N \ge n)\mathbb{P}(Z > x)^2} < \infty. \qquad (30)$$

https://doi.org/10.1017/apr.2018.37 Published online by Cambridge University Press

We bound

$$\frac{\mathbb{E}^{Q^{\gamma}_{a*}}(L^{\Delta}_{T}(x, 2^n) - L_T(x))^2}{\mathbb{P}(Z > x)^2}$$

by a geometrically decreasing function of n that does not depend on x. For $\beta \in (0, 1)$, using Hölder's inequality, we obtain

$$\begin{aligned}
&\frac{\mathbb{E}^{Q^{\gamma}_{a*}}(L^{\Delta}_{T}(x, 2^n) - L_T(x))^2}{P(Z > x)^2} \\
&= \frac{\mathbb{E}^{Q^{\gamma}_{a*}}\mathbf{1}_{\{\tau_\gamma(x)<\infty,\ \sum_{k=0}^{\tau_\gamma(x)+2^n} B_{k+1}\mathrm{e}^{S_k}\le x,\ Z>x\}}(M^{-1}_{\tau_\gamma}(x))^2}{\mathbb{P}(Z > x)^2} \\
&= \frac{\mathbb{E}^{Q^{\gamma}_{a*}}(\mathbf{1}_{\{\tau_\gamma(x)<\infty,\ \sum_{k=0}^{\tau_\gamma(x)+2^n} B_{k+1}\mathrm{e}^{S_k}\le x,\ Z>x\}}M^{-1}_{\tau_\gamma}(x))^{\beta}(\mathbf{1}_{\{T(x)<\infty\}}M^{-1}_{\tau_\gamma}(x))^{2-\beta}}{\mathbb{P}(Z > x)^2} \\
&\le \frac{(\mathbb{E}^{Q^{\gamma}_{a*}}\mathbf{1}_{\{\tau_\gamma(x)<\infty,\ \sum_{k=0}^{\tau_\gamma(x)+2^n} B_{k+1}\mathrm{e}^{S_k}\le x,\ Z>x\}}M^{-1}_{\tau_\gamma}(x))^{\beta}}{\mathbb{P}(Z > x)^{\beta}} \\
&\quad\times \frac{(\mathbb{E}^{Q^{\gamma}_{a*}}\mathbf{1}_{\{T(x)<\infty\}}M^{-1}_{\tau_\gamma}(x)^{(2-\beta)/(1-\beta)})^{1-\beta}}{\mathbb{P}(Z > x)^{2-\beta}} \\
&= \underbrace{\bigg(\frac{\mathbb{E}^{Q^{\gamma}_{a*}}\mathbf{1}_{\{\tau_\gamma(x)<\infty,\ \sum_{k=0}^{\tau_\gamma(x)+2^n} B_{k+1}\mathrm{e}^{S_k}\le x,\ Z>x\}}M^{-1}_{\tau_\gamma}(x)}{\mathbb{P}(Z > x)}\bigg)^{\beta}}_{\mathrm{I}_2} \underbrace{\bigg(\frac{\mathbb{E}^{Q^{\gamma}_{a*}}L_T^{(2-\beta)/(1-\beta)}(x)}{\mathbb{P}(Z > x)^{(2-\beta)/(1-\beta)}}\bigg)^{1-\beta}}_{\mathrm{I}_1}.
\end{aligned}$$

Now we analyze terms I_1 and I_2. Using the same argument around (27) in the proof of Theorem 2, to bound I_1 it is sufficient to analyze $\mathbb{P}(\tau_\gamma(x) < \infty)^{-(2+\varepsilon)}\mathbb{E}^{Q^{\gamma}_{a*}}L^{2+\varepsilon}_{\tau_\gamma}(x)$. From Lemma 1, we see that $\mathbb{P}(\tau_\gamma(x) < \infty)^{-(2+\varepsilon)}\mathbb{E}^{Q^{\gamma}_{a*}}L^{2+\varepsilon}_{\tau_\gamma}(x)$ is bounded w.r.t. x; therefore, I_1 is also bounded w.r.t. x. Turning to I_2, we claim that it can be bounded by $\kappa 2^{-n(q-1)}$ for some constant $\kappa > 0$. To see this, note that

$$\begin{aligned}
\mathrm{I}_2 &= \frac{\mathbb{P}(\tau_\gamma(x) < \infty,\ \sum_{k=0}^{\tau_\gamma(x)+2^n} B_{k+1}\mathrm{e}^{S_k} \le x,\ Z > x)}{\mathbb{P}(Z > x)} \\
&= \underbrace{\mathbb{P}^{(x)}\bigg(\sum_{k=0}^{\tau_\gamma(x)+2^n} B_{k+1}\mathrm{e}^{S_k} \le x,\ Z > x\bigg)}_{\mathrm{I}_3}\frac{\mathbb{P}(\tau_\gamma(x) < \infty)}{\mathbb{P}(Z > x)},
\end{aligned}$$

where $\mathbb{P}^{(x)}(\cdot)$ denotes the conditional distribution $\mathbb{P}(\cdot \mid \tau_\gamma(x) < \infty)$. Hence, it is sufficient to analyze the behavior of I_3 w.r.t. M. Note that

$$Z = \underbrace{\sum_{k=0}^{\tau_\gamma(x)} B_{k+1}\mathrm{e}^{S_k}}_{B'_x} + \underbrace{\mathrm{e}^{S_{\tau_\gamma(x)}}}_{A'_x}\underbrace{\sum_{k=1}^{\infty} B_{\tau_\gamma(x)+k}\mathrm{e}^{S_{\tau_\gamma(x)+k}-S_{\tau_\gamma(x)}}}_{Z'} \tag{31}$$

$$\sum_{k=0}^{\tau_\gamma(x)+M} B_{k+1}\mathrm{e}^{S_k} = B'_x + A'_x\underbrace{\sum_{k=1}^{M} B_{\tau_\gamma(x)+k}\mathrm{e}^{S_{\tau_\gamma(x)+k}-S_{\tau_\gamma(x)}}}_{Z'(M)}. \tag{32}$$

Combining (31) and (32) with Assumption A4, we obtain

$$\begin{aligned}\mathrm{I}_3 &= \int \mathbf{1}_{\{Z'^{(M)} \le (x-B'_x)/A'_x,\, Z' > (x-B'_x)/A'_x\}}\, \mathrm{d}\mathbb{P}^{(x)} \\ &= \int \mathbb{P}^{(x)}(Z'^{(M)} \le y, Z' > y)\, \mathrm{d}\mathbb{P}^{(x)}\biggl(\frac{x-B'_x}{A'_x} \le y\biggr) \\ &= \int \{\mathbb{P}^{(x)}(Z' > y) - \mathbb{P}^{(x)}(Z'^{(M)} > y)\}\, \mathrm{d}\mathbb{P}^{(x)}\biggl(\frac{x-B'_x}{A'_x} \le y\biggr).\end{aligned}$$

Using the strong Markov property, it follows that $Z'^{(M)} \overset{\mathrm{D}}{=} \sum_{k=0}^{M} B_{k+1}\mathrm{e}^{S_k}$ and $Z' \overset{\mathrm{D}}{=} Z$ under $\mathbb{P}^{(x)}$. Hence, we have

$$\mathrm{I}_3 = \int \biggl\{\mathbb{P}^{(x)}(Z > y) - \mathbb{P}^{(x)}\biggl(\sum_{k=0}^{M} B_{k+1}\mathrm{e}^{S_k} > y\biggr)\biggr\}\, \mathrm{d}\mathbb{P}^{(x)}\biggl(\frac{x-B'_x}{A'_x} \le y\biggr).$$

Combining this with the fact that the backward iteration $\sum_{k=0}^{M} B_{k+1}\mathrm{e}^{S_k}$ has the same distribution as Z_M defined in (2), we obtain

$$\mathrm{I}_3 = \int \{\mathbb{P}(Z > y) - \mathbb{P}(Z_M > y)\}\, \mathrm{d}\mathbb{P}^{(x)}\biggl(\frac{x-B'_x}{A'_x} \le y\biggr) \le d_{\mathrm{TV}}(Z_M, Z),$$

where d_{TV} denotes the total variation distance. To understand this quantity, we apply the Lyapunov criterion from [20, Theorem 3.6], which implies a polynomial convergence rate of the M-step transition kernel to the invariant distribution in the total variation norm. In view of Lemma 5, there exists a constant κ such that $(\mathrm{I}_3) \le d_{\mathrm{TV}}(Z_M, Z) \le \kappa M^{-(q-1)}$ for all $M \in \mathbb{N}$. It should be noted that an exact expression of the constant κ can be obtained in a few special cases; see, e.g. [11] and [21] and the references therein. By choosing N such that $\mathbb{P}(N \le n) = 1 - (1-p)^n$ for $n \ge 1$ with $p < 1 - 2^{-(q-1)}$, we conclude that the left-hand side of (30) is bounded; hence, by applying Result 3, we can remove this constant and obtain an unbiased, strongly efficient estimator. □

Lemma 5. *Let Z_n be a Markov chain as in (2) such that Assumption A5 holds. Then there exists a constant κ such that $d_{\mathrm{TV}}(Z_n, Z) \le \kappa n^{-(q-1)}$.*

Proof. We wish to apply [20, Theorem 3.6]. In order to establish the Lyapunov condition as in (34) below, let $V(x) = 1 \vee (\log x)^q$ and $PV(x) \triangleq \mathbb{E}V(A_1x + B_1)$. Note that $V(x) = (\log x)^q \mathbf{1}_{\{x>\mathrm{e}\}} + \mathbf{1}_{\{x \le \mathrm{e}\}}$ and, hence, the binomial expansion yields

$$\begin{aligned}PV(x) &= \mathbb{E}[\log(A_1x + B_1)]^q \mathbf{1}_{\{A_1x+B_1>\mathrm{e}\}} + \mathbb{P}(A_1x + B_1 \le \mathrm{e}) \\ &\le \mathbb{E}[\log(A_1x + B_1^+)]^q \mathbf{1}_{\{A_1x+B_1>\mathrm{e}\}} + \mathbb{P}(A_1x + B_1 \le \mathrm{e}) \\ &= \mathbb{E}\biggl[\log\frac{A_1x + B_1^+}{x} + \log x\biggr]^q \mathbf{1}_{\{A_1x+B_1>\mathrm{e}\}} + \mathbb{P}(A_1x + B_1 \le \mathrm{e}) \\ &= \mathbb{E}\biggl[(\log x)^q \mathbf{1}_{\{A_1x+B_1>\mathrm{e}\}} + \sum_{i=1}^{q}\binom{q}{i}(\log x)^{q-i}\biggl(\log\frac{A_1x + B_1^+}{x}\biggr)^i \mathbf{1}_{\{A_1x+B_1>\mathrm{e}\}}\biggr] \\ &\quad + \mathbb{P}(A_1x + B_1 \le \mathrm{e})\end{aligned}$$

https://doi.org/10.1017/apr.2018.37 Published online by Cambridge University Press

$$= V(x) + \mathbb{E}[\log x]^q(\mathbf{1}_{\{A_1x+B_1>\mathrm{e}\}} - \mathbf{1}_{\{x>\mathrm{e}\}}) + \mathbb{P}(A_1x + B_1 \le \mathrm{e}) - \mathbf{1}_{\{x\le\mathrm{e}\}}$$
$$+ q(\log x)^{q-1}\mathbb{E}\bigg[\log \frac{A_1x + B_1^+}{x}\bigg]\mathbf{1}_{\{A_1x+B_1>\mathrm{e}\}}$$
$$+ \sum_{i=2}^{q}\binom{q}{i}(\log x)^{q-i}\mathbb{E}\bigg[\bigg(\log \frac{A_1x + B_1^+}{x}\bigg)^i \mathbf{1}_{\{A_1x+B_1>\mathrm{e}\}}\bigg].$$

For $x > \mathrm{e}$,

$$PV(x) \le V(x) + \mathbb{P}(A_1x + B_1 \le \mathrm{e}) + q(\log x)^{q-1}\mathbb{E}\bigg[\log \frac{A_1x + B_1^+}{x}\mathbf{1}_{\{A_1x+B_1>\mathrm{e}\}}\bigg]$$
$$+ \sum_{i=2}^{q}\binom{q}{i}(\log x)^{q-i}\mathbb{E}\bigg[\bigg(\log \frac{A_1x + B_1^+}{x}\bigg)^i \mathbf{1}_{\{A_1x+B_1>\mathrm{e}\}}\bigg].$$

Note that

$$\begin{aligned}\log A_1 &\le \log \frac{A_1x + B_1^+}{x}\\ &\le \log(A_1 + B_1^+)\\ &\le \log(2(A_1 \vee B_1^+))\\ &= \log(A_1 \vee B_1^+) + \log 2\\ &= (\log A_1) \vee (\log B_1^+) + \log 2\\ &\le |\log A_1| + |\log B_1^+| + \log 2\end{aligned}$$

and, hence,

$$\bigg|\log \frac{A_1x + B_1^+}{x}\bigg| \le |\log A_1| + |\log B_1^+| + \log 2. \tag{33}$$

Moreover, the right-hand side of (33) does not depend on x and has finite qth moment. Thus, there are constants c_i, $i \ge 1$, such that

$$\sum_{i=2}^{q}\binom{q}{i}(\log x)^{q-i}\mathbb{E}\bigg[\bigg(\log \frac{A_1x + B_1^+}{x}\bigg)^i \mathbf{1}_{\{A_1x+B_1>\mathrm{e}\}}\bigg] \le \sum_{i=0}^{q-2} c_i(\log x)^i \le \varepsilon(\log x)^{q-1}$$

for sufficiently large x. On the other hand, note that $\log((A_1x + B_1^+)/x)\mathbf{1}_{\{A_1x+B_1>\mathrm{e}\}}$ converges to $\log A_1$ a.s. as $x \to \infty$ and, hence, by dominated convergence

$$\mathbb{E}\bigg[\log \frac{A_1x + B_1^+}{x}\bigg]\mathbf{1}_{\{A_1x+B_1>\mathrm{e}\}} \to \mathbb{E}\log A_1 < 0.$$

Therefore, for any fixed $\varepsilon > 0$,

$$q(\log x)^{q-1}\mathbb{E}\bigg[\log \frac{A_1x + B_1^+}{x}\mathbf{1}_{\{A_1x+B_1>\mathrm{e}\}}\bigg] \le (q\mathbb{E}\log A_1 + \varepsilon)(\log x)^{q-1}$$

for sufficiently large x. Choosing ε so that $q\mathbb{E}\log A_1 + 3\varepsilon < 0$ and noting that

$$\mathbb{P}(A_1x + B_1 \le \mathrm{e}) \to 0 \quad \text{as } x \to \infty,$$

as well as $(\log x)^{q-1} = ((\log x)^q \mathbf{1}_{\{x>\mathrm{e}\}} + \mathbf{1}_{\{x\le \mathrm{e}\}})^{(q-1)/q}$ for $x > \mathrm{e}$, we conclude that there exists K such that

$$\begin{aligned} PV(x) &\le V(x) + \varepsilon(\log x)^{q-1} + (q\mathbb{E}\log A_1 + \varepsilon)(\log x)^{q-1} + \varepsilon(\log x)^{q-1} \\ &\le V(x) - cV^{(q-1)/q}(x) \quad \text{for } x > K, \end{aligned}$$

where $c = -(q\mathbb{E}\log A_1 + 3\varepsilon) > 0$. Finally, since $PV(x)$, $V(x)$, and $V^{(q-1)/q}(x)$ are bounded on $[0, K]$, there exists a constant b such that

$$PV(x) \le V(x) - cV^{(q-1)/q}(x) + b\mathbf{1}_{[0,K]}, \tag{34}$$

which is the sufficient condition of [20, Theorem 3.6] for polynomial ergodicity. Thus, we obtain the result. □

Proof of Lemma 3. Recall that Z can be bounded by a stochastic perpetuity

$$\bar{Z} \triangleq \sum_{n=0}^{\infty} \bar{B}_{n+1}\mathrm{e}^{S_n},$$

where $\bar{B}_n \triangleq \max(B_n^+ + D_n, 1)$ and $S_n = \sum_{i=1}^n \log A_i$. Applying the same technique as in the proof of Lemma 2 to $\bar{Z}$, we obtain the result. □

Proof of Theorems 5 and 6. We omit the details here, since they can be proved analogously as in Theorems 1 and 2. □

Proof of Theorem 7. Recall that $\Psi_{1:n}(Z_0) \triangleq \Psi_1 \circ \Psi_2 \circ \cdots \circ \Psi_n(Z_0)$. Due to the fact that $\{\tau_\gamma(x) < \infty, \Psi_{1:\,\tau_\gamma(x)+M}(Z_0) > x\} \subseteq \{T(x) < \infty\}$, in order to prove the vanishing relative bias result, it is sufficient to show that

$$\liminf_{x\to\infty} \frac{\mathbb{P}(\tau_\gamma(x) < \infty, \Psi_{1:\,\tau_\gamma(x)+M}(Z_0) > x)}{\mathbb{P}(T(x) < \infty)} \ge 1. \tag{35}$$

Recall that $S_n = \sum_{i=1}^n \log A_i$ and $S_n(\gamma) = n\gamma_1 + \sum_{i=1}^n [(\log^+(B_i^+ + D_i) - \gamma_2) \vee \log A_i]$. Let $\mu \triangleq -\mathbb{E}S_1$ and $\mu_\gamma \triangleq -\mathbb{E}S_1(\gamma)$. For $\nu, K > 0$ consider the sets

$$\begin{aligned} E_n^{(1)} &= E_n^{(1)}(K,\nu) = \{S_j \in (-j(\mu+\nu) - K, -j(\mu-\nu) + K),\ j \le n\}, \\ E_n^{(2)} &= E_n^{(2)}(K,\nu) = \{S_j(\gamma) \in (-j(\mu_\gamma+\nu) - K, -j(\mu_\gamma-\nu) + K),\ j \le n\}, \\ E_n^{(3)} &= E_n^{(3)}(K,\nu) = \{|\underline{B}_j| \le \mathrm{e}^{\nu j + K},\ j \le n\}, \end{aligned}$$

where $\underline{B}_j = B_j - D_j$. Define

$$\begin{aligned} E_n = E_n^{(1)} \cap E_n^{(2)} \cap E_n^{(3)} \cap \{\Psi_{n+2}(Z_0) > \nu\} \\ \cap \{\max(A_{n+1}, \underline{B}_{n+1}) > x\mathrm{e}^{n(\mu+\nu)+L+K}, \underline{B}_{n+1} \ge -x\mathrm{e}^{n(\mu-\nu)-K}\}, \end{aligned}$$

where $L > 0$ is chosen to be large enough such that the sets $\{E_n\}_{n\ge1}$ are disjoint. The existence of such an L is guaranteed by the fact that $E_n \subseteq \{\tau_\gamma(x) = n+1\}$ (see below). Now we show that $E_n \subseteq \{\tau_\gamma(x) = n+1, \Psi_{1:\,\tau_\gamma(x)+1}(Z_0) > x\}$. To see that $E_n \subseteq \{\tau_\gamma(x) = n+1\}$, note that

$\{S_j(\gamma)\}_{j\le n}$ is bounded by K, $\mu > \mu_\gamma$ (due to the fact that $S_1 \le S_1(\gamma)$) and

$$\begin{aligned}
S_{n+1}(\gamma) &= S_n(\gamma) + \log(\max(\bar{B}_{n+1}\mathrm{e}^{-\gamma_2}, A_{n+1})) + \gamma_1 \\
&> -n(\mu_\gamma + \nu) - K + \log(\max(\underline{B}_{n+1}, A_{n+1})) - \gamma_2 + \gamma_1 \\
&> \log x + n(\mu - \mu_\gamma) + L - \gamma_2 + \gamma_1 \\
&> \log x + L - \gamma_2 + \gamma_1 \\
&> s(x)
\end{aligned}$$

for sufficiently large L that does not depend on x. Thus, we conclude that $\tau_\gamma(x) = n+1 < \infty$ by taking sufficiently large x. To see that $E_n \subseteq \{\tau_\gamma(x) = n+1, \Psi_{1:\,\tau_\gamma(x)+1}(Z_0) > x\}$, note that $\Psi_{1:\,n}(z) \ge \sum_{k=0}^{n-1} \underline{B}_{k+1} \prod_{j=1}^{k} A_j + z\prod_{j=1}^{n} A_j$ from (15) and Assumption B4. Moreover,

$$|\underline{B}_{k+1}| \prod_{j=1}^{k} A_j = |\underline{B}_{k+1}|\mathrm{e}^{S_k} \le \mathrm{e}^{\nu(k+1)+K}\mathrm{e}^{-k(\mu-\nu)+K} = \mathrm{e}^{-k(\mu-2\nu)+2K+\nu} \quad \text{on } E_n$$

and, hence,

$$\begin{aligned}
\Psi_{1:\,\tau_\gamma(x)+1}(Z_0) &= \Psi_{1:\,n+2}(Z_0) \\
&= \Psi_{1:\,n}(\Psi_{n+1:\,n+2}(Z_0)) \\
&\ge \sum_{k=0}^{n-1} \underline{B}_{k+1} \prod_{j=1}^{k} A_j + (\underline{B}_{n+1} + \Psi_{n+2}(Z_0)A_{n+1}) \prod_{j=1}^{n} A_j \\
&\ge -\sum_{k=0}^{n-1} |\underline{B}_{k+1}| \prod_{j=1}^{k} A_j + (\underline{B}_{n+1} + x\mathrm{e}^{n(\mu-\nu)-K} + \Psi_{n+2}(Z_0)A_{n+1}) \prod_{j=1}^{n} A_j \\
&\qquad - x\mathrm{e}^{n(\mu-\nu)-K} \prod_{j=1}^{n} A_j \\
&\ge -\frac{\mathrm{e}^{2K+\nu}}{1-\mathrm{e}^{-\mu+2\nu}} + (\underline{B}_{n+1} + x\mathrm{e}^{n(\mu-\nu)-K} + \nu A_{n+1}) \prod_{j=1}^{n} A_j - x \\
&\ge -\frac{\mathrm{e}^{2K+\nu}}{1-\mathrm{e}^{-\mu+2\nu}} + \min(\nu, 1)\max(A_{n+1}, \underline{B}_{n+1} + x\mathrm{e}^{n(\mu-\nu)-K}) \\
&\qquad\qquad \times \mathrm{e}^{-n(\mu+\nu)-K} - x \\
&\ge -\frac{\mathrm{e}^{2K+\nu}}{1-\mathrm{e}^{-\mu+2\nu}} + \min(\nu, 1)\max(A_{n+1}, \underline{B}_{n+1})\mathrm{e}^{-n(\mu+\nu)-K} - x \\
&\ge -\frac{\mathrm{e}^{2K+\nu}}{1-\mathrm{e}^{-\mu+2\nu}} + \min(\nu, 1)x\mathrm{e}^{L} - x \\
&> x
\end{aligned}$$

for sufficiently large L that does not depend on x. Note that from Lemma 6 below,

$$\begin{aligned}
\mathbb{P}(E_n) &= \mathbb{P}(E_n^{(1)} \cap E_n^{(2)} \cap E_n^{(3)})\mathbb{P}(\Psi_{n+2}(Z_0) > \nu) \\
&\qquad \times \mathbb{P}(\max(A_{n+1}, \underline{B}_{n+1}) > x\mathrm{e}^{n(\mu+\nu)+L+K}, \underline{B}_{n+1} \ge -x\mathrm{e}^{n(\mu-\nu)-K}) \\
&= \mathbb{P}(E_n^{(1)} \cap E_n^{(2)} \cap E_n^{(3)})\mathbb{P}(\Psi_1(Z_0) > \nu) \\
&\qquad \times \mathbb{P}(\max(A_1, \underline{B}_1) > x\mathrm{e}^{n(\mu+\nu)+L+K}, \underline{B}_1 \ge -x\mathrm{e}^{n(\mu-\nu)-K})
\end{aligned}$$

$$
\begin{aligned}
&\geq (1-\varepsilon)\mathbb{P}(\Psi_1(Z_0) > \nu) \\
&\quad \times \{\mathbb{P}(\max(A_1, \underline{B}_1) > x\mathrm{e}^{n(\mu+\nu)+L+K}) \\
&\qquad - \mathbb{P}(\max(A_1, \underline{B}_1) > x\mathrm{e}^{n(\mu+\nu)+L+K}, \underline{B}_1 < -x\mathrm{e}^{n(\mu-\nu)-K})\} \\
&\geq (1-\varepsilon)\mathbb{P}(\Psi_1(Z_0) > \nu)\{\mathbb{P}(\max(A_1, \underline{B}_1) > x\mathrm{e}^{n(\mu+\nu)+L+K}) \\
&\qquad\qquad - \mathbb{P}(A_1 > x\mathrm{e}^{n(\mu+\nu)+L+K}, \underline{B}_1 < -x\mathrm{e}^{n(\mu-\nu)-K})\}.
\end{aligned}
$$

Moreover, since $E_n \subseteq \{\tau_\gamma(x) < \infty, \Psi_{1\colon \tau_\gamma(x)+M}(Z_0) > x\}$ and $E_n, n \geq 1$ are disjoint, it follows that

$$
\begin{aligned}
&\mathbb{P}(\tau_\gamma(x) < \infty, \Psi_{1\colon \tau_\gamma(x)+M}(Z_0) > x) \\
&\quad \geq \sum_{n\geq 0} \mathbb{P}(E_n) \\
&\quad \geq (1-\varepsilon)\mathbb{P}(\Psi_1(Z_0) > \nu) \sum_{n\geq 0}\{\mathbb{P}(\max(A_1, \underline{B}_1) > x\mathrm{e}^{n(\mu+\nu)+L+K}) \\
&\qquad\qquad - \mathbb{P}(A_1 > x\mathrm{e}^{n(\mu+\nu)+L+K}, \underline{B}_1 < -x\mathrm{e}^{n(\mu-\nu)-K})\}. \quad (36)
\end{aligned}
$$

From Assumption B1(iii), we conclude that for any $\varepsilon' > 0$ and taking sufficiently large x, the following holds:

$$
\begin{aligned}
\mathbb{P}(A_1 > x\mathrm{e}^{n(\mu+\nu)+L+K}, \underline{B}_1 < -x\mathrm{e}^{n(\mu-\nu)-K}) &\leq \mathbb{P}(A_1 > x\mathrm{e}^{n(\mu-\nu)-K}, \underline{B}_1 < -x\mathrm{e}^{n(\mu-\nu)-K}) \\
&\leq \varepsilon'\mathbb{P}(\max(A_1, \underline{B}_1) > x\mathrm{e}^{n(\mu-\nu)-K}).
\end{aligned}
$$

Combining this with (36), we obtain

$$
\begin{aligned}
&\mathbb{P}(\tau_\gamma(x) < \infty, \Psi_{1\colon \tau_\gamma(x)+M}(Z_0) > x) \\
&\quad \geq (1-\varepsilon)\mathbb{P}(\Psi_1(Z_0) > \nu) \sum_{n\geq 0}\{\mathbb{P}(\max(A_1, \underline{B}_1) > x\mathrm{e}^{n(\mu+\nu)+L+K}) \\
&\qquad\qquad - \varepsilon'\mathbb{P}(\max(A_1, \underline{B}_1) > x\mathrm{e}^{n(\mu-\nu)-K})\}. \qquad (37)
\end{aligned}
$$

For a given $\varepsilon'' > 0$, let x be sufficiently large so that

$$
\left| 1 - \frac{\mathbb{P}(\log\max(A_1, \underline{B}_1) > y)}{\mathbb{P}(\log\max(A_1, B_1) > y)} \right| \leq \varepsilon''.
$$

Since $\mathbb{P}(\max(A_1, \underline{B}_1) > y)$ is decreasing in y,

$$
\begin{aligned}
&\sum_{n\geq 0} \mathbb{P}(\max(A_1, \underline{B}_1) > x\mathrm{e}^{n(\mu+\nu)+L+K}) \\
&\quad \geq \sum_{n\geq 0} \frac{1}{\mu+\nu} \int_{\log x+L+K+n(\mu+\nu)}^{\log x+L+K+(n+1)(\mu+\nu)} \mathbb{P}(\log\max(A_1, \underline{B}_1) > y)\,\mathrm{d}y \\
&\quad \geq \sum_{n\geq 0} \frac{1-\varepsilon''}{\mu+\nu} \int_{\log x+L+K+n(\mu+\nu)}^{\log x+L+K+(n+1)(\mu+\nu)} \mathbb{P}(\log\max(A_1, B_1) > y)\,\mathrm{d}y \\
&\quad = \frac{1-\varepsilon''}{\mu+\nu} \bar{F}_I(\log x + L + K),
\end{aligned}
$$

and

$$\begin{aligned}
&\sum_{n\geq 0}\mathbb{P}(\max(A_1,\underline{B}_1) > x\mathrm{e}^{n(\mu-\nu)-K})\\
&\quad\leq \sum_{n\geq 0}\frac{1+\varepsilon''}{\mu-\nu}\int_{\log x-K+(n-1)(\mu-\nu)}^{\log x-K+n(\mu-\nu)}\mathbb{P}(\log\max(A_1,\underline{B}_1) > y)\,\mathrm{d}y\\
&\quad\leq \sum_{n\geq 0}\frac{1+\varepsilon''}{\mu-\nu}\int_{\log x-K+(n-1)(\mu-\nu)}^{\log x-K+n(\mu-\nu)}\mathbb{P}(\log\max(A_1,\underline{B}_1) > y)\,\mathrm{d}y\\
&\quad= \frac{1+\varepsilon''}{\mu-\nu}\bar{F}_I(\log x-K-\mu+\nu).
\end{aligned}$$

Moreover, using the fact that $\bar{F}_I$ is long tailed, we obtain, from (37),

$$\begin{aligned}
&\mathbb{P}(\tau_\gamma(x)<\infty,\Psi_{1:\,\tau_\gamma(x)}(\Psi_{\tau_\gamma(x)+1:\,\tau_\gamma(x)+M}(Z_0)) > x)\\
&\quad\geq (1-\varepsilon)\mathbb{P}(\Psi_1(Z_0)>\nu)\bigg(\frac{1-\varepsilon''}{\mu+\nu}\bar{F}_I(\log x+L+K)\\
&\qquad\qquad -\frac{\varepsilon'(1+\varepsilon'')}{\mu-\nu}\bar{F}_I(\log x+L+K-\mu+\nu)\bigg)\\
&\quad\sim (1-\varepsilon)\mathbb{P}(\Psi_1(Z_0)>\nu)\bigg(\frac{1-\varepsilon''}{\mu+\nu}-\frac{\varepsilon'(1+\varepsilon'')}{\mu-\nu}\bigg)\bar{F}_I(\log x)\\
&\quad\sim \mu(1-\varepsilon)\mathbb{P}(\Psi_1(Z_0)>\nu)\bigg(\frac{1-\varepsilon''}{\mu+\nu}-\frac{\varepsilon'(1+\varepsilon'')}{\mu-\nu}\bigg)\mathbb{P}(T(x)<\infty), \qquad (38)
\end{aligned}$$

where we use [12, Theorem 3.1] in the final step. Recall that the distribution of the stationary solution to (1) does not depend on the initial condition Z_0 and, hence, without loss of generality we can set $Z_0=0$. Noting that $\Psi_1(0)\geq 0$ and, hence, $\mathbb{P}(\Psi_1(Z_0)>\nu)\to 1$ as $\nu\to 0$, we let $\varepsilon,\varepsilon',\varepsilon'',\nu\to 0$ to obtain (35). This implies that the relative bias converge to 0 since the numerator in (35) is always smaller than the denominator. □

Lemma 6. *Consider the sets $E_n^{(1)}$, $E_n^{(2)}$, and $E_n^{(3)}$ as in the proof of Theorem 7. Then, for $\nu,\varepsilon>0$, there exists $K>0$ such that*

$$\mathbb{P}\bigg(\bigcap_{n\geq 1}(E_n^{(1)}\cap E_n^{(2)}\cap E_n^{(3)})\bigg)\geq 1-\varepsilon.$$

Proof. In the proof of [27, Theorem 1], the authors stated that for any $\nu>0$ and any i.i.d. sequence $\{Y_n\}_{n\geq 0}$ with $\mathbb{E}[\log^+|Y_1|]<\infty$, it holds that

$$\mathbb{P}(|Y_j|\leq \mathrm{e}^{\nu j+K},\ j\leq n)\to 1\quad\text{as } K\to\infty \text{ uniformly w.r.t. } n.$$

Using this argument, we conclude that $\mathbb{P}(E_n^{(3)})\to 1$ as $K\to\infty$ uniformly w.r.t. n. Further, combining this fact with the strong law of large numbers for $\{S_n\}_{n\geq 0}$ and $\{S_n(\gamma)\}_{n\geq 0}$ (see, e.g. [4, Lemma 3.1]), we can always take large enough K such that

$$\mathbb{P}(E_n^{(1)}\cap E_n^{(2)}\cap E_n^{(3)})\geq 1-\varepsilon\quad\text{for all } n\in\mathbb{N}.$$

Finally, since the sequence of sets $E_n^{(1)}\cap E_n^{(2)}\cap E_n^{(3)}$, $n\geq 0$, is decreasing in the sense of inclusion, we obtain the result. □

Proof of Theorem 8. Note that this result can be proved by following similar arguments as in the proof of Theorem 4, hence, we provide only a sketch of the proof. Recall that in the context of iterated random functions, we have the estimator

$$L_T^{\Delta}(x, M) \triangleq \mathbf{1}_{\{\tau_\gamma(x)<\infty,\ \Psi_{1:\,\tau_\gamma(x)+M}(Z_0)>x\}} \prod_{k=1}^{\tau_\gamma(x)} \frac{w_\gamma(S_{k-1}(\gamma)+a_*)}{v_\gamma(S_k(\gamma)+a_*)}.$$

Analogous to the proof of Theorem 4, we wish to bound

$$\frac{\mathbb{E}^{Q_{a*}^{\gamma}}(L_T^{\Delta}(x, 2^n) - L_T(x))^2}{\mathbb{P}(Z > x)^2}$$

by a decreasing function of n independent of x. Again, by using Hölder's inequality it is sufficient to bound

$$\begin{aligned}(\mathrm{I}_1') &\triangleq \frac{\mathbb{E}^{Q_{a*}^{\gamma}} \mathbf{1}_{\{\tau_\gamma(x)<\infty,\ \Psi_{1:\,\tau_\gamma(x)+M}(Z_0)\le x,\ Z>x\}} M_{\tau_\gamma}^{-1}(x)}{\mathbb{P}(Z > x)} \\ &= \underbrace{\mathbb{P}^{(x)}(\Psi_{1:\,\tau_\gamma(x)}(\Psi_{\tau_\gamma(x)+1:\,\tau_\gamma(x)+M}(Z_0)) \le x,\ \Psi_{1:\,\tau_\gamma(x)}(Z') > x)}_{\mathrm{I}_2'} \frac{\mathbb{P}(\tau_\gamma(x) < \infty)}{\mathbb{P}(Z > x)},\end{aligned}$$

where $Z' \triangleq \lim_{M\to\infty} \Psi_{\tau_\gamma(x)+1:\,\tau_\gamma(x)+M}(Z_0) \overset{\mathrm{D}}{=} Z$ and $\mathbb{P}^{(x)}(\cdot)$ denotes the conditional distribution $\mathbb{P}(\cdot \mid \tau_\gamma(x) < \infty)$. Since Ψ_n is Lipschitz and bijective, $\Psi_{1:\,\tau_\gamma(x)}^{-1}$ is either strictly increasing or strictly decreasing. Without loss of generality, we assume that $\Psi_{1:\,\tau_\gamma(x)}^{-1}$ is strictly increasing, since the case of $\Psi_{1:\,\tau_\gamma(x)}^{-1}$ being strictly decreasing can be dealt with similarly. Using the strong Markov property we obtain

$$\mathrm{I}_2' = \int \{\mathrm{P}(\mathrm{Z} > \mathrm{y}) - \mathrm{P}(\mathrm{Z_M} > \mathrm{y})\}\, \mathrm{d}\mathbb{P}^{(\mathrm{x})}(\Psi_{1:\,\tau_\gamma(\mathrm{x})}^{-1}(\mathrm{x}) \le \mathrm{y}) \le \mathrm{d_{TV}}(\mathrm{Z_M}, \mathrm{Z}).$$

By Lemma 7 below we obtain the result. □

Lemma 7. *Let Z_n be a Markov chain as in (1) such that Assumption B5 hold. Then there exists a constant κ such that $d_{\mathrm{TV}}(Z_n, Z) \le \kappa n^{-(q-1)}$.*

Proof. Let $V(x) = 1 \vee (\log x)^q$ and $PV(x) \triangleq \mathbb{E}V(\Psi_1(x))$. By noting that

$$\begin{aligned}PV(x) &\triangleq \mathbb{E}[\log(\Psi_1(x))]^q \mathbf{1}_{\{\Psi_1(x)>\mathrm{e}\}} + \mathbb{P}(\Psi_1(x) \le \mathrm{e}) \\ &\le \mathbb{E}[\log(A_1 x + B_1^+ + D_1)]^q \mathbf{1}_{\{\Psi_1(x)>\mathrm{e}\}} + \mathbb{P}(\Psi_1(x) \le \mathrm{e}),\end{aligned}$$

the result follows immediately from similar arguments as in the proof of Lemma 5. □

Acknowledgement

The authors gratefully acknowledge the support from the Netherlands Organization for Scientific Research (NWO) through the Vici grant 639.033.413.

References

[1] Asmussen, S. (2003). *Applied Probability and Queues.* Springer, New York.
[2] Asmussen, S. and Glynn, P. W. (2007). *Stochastic Simulation: Algorithms and Analysis*. Springer, New York.

[3] Asmussen, S. and Nielsen, H. M. (1995). Ruin probabilities via local adjustment coefficients. *J. Appl. Prob.* **32,** 736–755.

[4] Asmussen, S., Schmidli, H. and Schmidt, V. (1999). Tail probabilities for non-standard risk and queueing processes with subexponential jumps. *Adv. Appl. Prob.* **31,** 422–447.

[5] Blanchet, J. and Glynn, P. (2008). Efficient rare-event simulation for the maximum of heavy-tailed random walks. *Ann. Appl. Prob.* **18,** 1351–1378.

[6] Blanchet, J., Lam, H. and Zwart, B. (2012). Efficient rare-event simulation for perpetuities. *Stoch. Process. Appl.* **122,** 3361–3392.

[7] Blanchet, J. and Zwart, B. (2007). Importance sampling of compounding processes. In *Proc. 2007 Winter Simulation Conference*, IEEE, pp. 372–379.

[8] Buraczewski, D., Damek, E. and Mikosch, T. (2016). *Stochastic Models with Power-Law Tails: The Equation $X = AX + B$*. Springer, Cham.

[9] Collamore, J. F. and Vidyashankar, A. N. (2013). Tail estimates for stochastic fixed point equations via nonlinear renewal theory. *Stoch. Process. Appl.* **123,** 3378–3429.

[10] Collamore, J. F., Diao, G. and Vidyashankar, A. N. (2014). Rare event simulation for processes generated via stochastic fixed point equations. *Ann. Appl. Prob.* **24,** 2143–2175.

[11] Douc, R., Moulines, E. and Soulier, P. (2007). Computable convergence rates for sub-geometric ergodic Markov chains. *Bernoulli* **13,** 831–848.

[12] Dyszewski, P. (2016). Iterated random functions and slowly varying tails. *Stoch. Process. Appl.* **126,** 392–413.

[13] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events: For Insurance and Finance*. Springer, Berlin.

[14] Foss, S., Korshunov, D. and Zachary, S. (2013). *An Introduction to Heavy-Tailed and Subexponential Distributions*, 2nd edn. Springer, New York.

[15] Ganesh, A., O'Connell, N. and Wischik, D. (2004). *Big Queues*. Springer, Berlin.

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

[17] Goldie, C. M. and Grübel, R. (1996). Perpetuities with thin tails. *Adv. Appl. Prob.* **28,** 463–480.

[18] Grey, D. R. (1994). Regular variation in the tail behaviour of solutions of random difference equations. *Ann. Appl. Prob.* **4,** 169–183.

[19] Grincevičius, A. K. (1975). Limit theorems for products of random linear transformations on the line. *Lithuanian Math. J.* **15,** 568–579.

[20] Jarner, S. F. and Roberts, G. O. (2002). Polynomial convergence rates of Markov chains. *Ann. Appl. Prob.* **12,** 224–247.

[21] Kalashnikov, V. and Tsitsiashvili, G. (1999). Tails of waiting times and their bounds. *Queueing Systems Theory Appl.* **32,** 257–283.

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

[23] Klüppelberg, C. (1988). Subexponential distributions and integrated tails. *J. Appl. Prob.* **25,** 132–141.

[24] Maulik, K. and Zwart, B. (2006). Tail asymptotics for exponential functionals of Lévy processes. *Stoch. Process. Appl.* **116,** 156–177.

[25] Mirek, M. (2011). Heavy tail phenomenon and convergence to stable laws for iterated Lipschitz maps. *Prob. Theory Relat. Fields* **151,** 705–734.

[26] Pakes, A. G. (1975). On the tails of waiting-time distributions. *J. Appl. Prob.* **12,** 555–564.

[27] Palmowski, Z. and Zwart, B. (2007). Tail asymptotics of the supremum of a regenerative process. *J. Appl. Prob.* **44,** 349–365.

[28] Rhee, C.-H. and Glynn, P. W. (2015). Unbiased estimation with square root convergence for SDE models. *Operat. Res.* **63,** 1026–1043.

[29] Veraverbeke, N. (1977). Asymptotic behaviour of Wiener-Hopf factors of a random walk. *Stoch. Process. Appl.* **5,** 27–37.

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

[31] Zachary, S. (2004). A note on Veraverbeke's theorem. *Queueing Systems* **46,** 9–14.