

# Importance sampling of heavy-tailed iterated random functions

By Bohan Chen, Chang-Han Rhee & Bert Zwart

Centrum Wiskunde & Informatica  
Science Park 123, 1098 XG Amsterdam, Netherlands

September 13, 2016

## Abstract

We consider a stochastic recurrence equation of the form  $Z_{n+1} = A_{n+1}Z_n + B_{n+1}$ , where  $\mathbb{E}[\log A_1] < 0$ ,  $\mathbb{E}[\log^+ B_1] < \infty$  and  $\{(A_n, B_n)\}_{n \in \mathbb{N}}$  is an i.i.d. sequence of positive random vectors. The stationary distribution of this Markov chain can be represented as the distribution of the random variable  $Z \triangleq \sum_{n=0}^{\infty} B_{n+1} \prod_{k=1}^n A_k$ . Such random variables can be found in the analysis of probabilistic algorithms or financial mathematics, where  $Z$  would be called a stochastic perpetuity. If one interprets  $-\log A_n$  as the interest rate at time  $n$ , then  $Z$  is the present value of a bond that generates  $B_n$  unit of money at each time point  $n$ . We are interested in estimating the probability of the rare event  $\{Z > x\}$ , when  $x$  is large; we provide a consistent simulation estimator using state-dependent importance sampling for the case, where  $\log A_1$  is heavy-tailed and the so-called Cramér condition is not satisfied. Our algorithm leads to an estimator for  $P(Z > x)$ . We show that under natural conditions, our estimator is strongly efficient. Furthermore, we extend our method to the case, where  $\{Z_n\}_{n \in \mathbb{N}}$  is defined via the recursive formula  $Z_{n+1} = \Psi_{n+1}(Z_n)$  and  $\{\Psi_n\}_{n \in \mathbb{N}}$  is a sequence of i.i.d. random Lipschitz functions.

## 1 Introduction

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

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

where  $\{\Psi_n\}_{n \in \mathbb{N}}$  is a sequence of i.i.d. positive random Lipschitz functions and  $Z_0 \in \mathbb{R}$  is arbitrary but independent of the sequence  $\{\Psi_n\}_{n \in \mathbb{N}}$ . For  $k < n$ , define the backward iteration as

$$\Psi_{k:n}(z) \triangleq \Psi_k \circ \Psi_{k+1} \circ \cdots \circ \Psi_n(z).$$

Define  $Z^{(n)}(z_0) \triangleq \Psi_{1:n}(z_0)$ . Under some mild conditions (cf. Assumption 2 below), the sequence  $\{Z^{(n)}(z_0)\}_{n \in \mathbb{N}}$  converges a.s. to some random variable  $Z$ . Moreover, this limit does not depend on the choice of the initial condition  $z_0$  (cf. Dyszewski, 2016, Theorem 3.1) and has the same distribution as the stationary solution to (1). For simplicity we set

$$Z^{(n)} \triangleq Z^{(n)}(0). \quad (2)$$

We assume that  $\Psi_n$  is such that  $Z^{(n)}$  is increasing in  $n$ . Define  $T(x) = \inf\{n \geq 0 : Z^{(n)} > x\}$ . This paper develops efficient simulation methods for estimating the tail probability of  $Z$ , i.e. we are interested in computing

$$\mathbb{P}(Z > x) = \mathbb{P}(T(x) < \infty),$$

when  $x$  is large.

The main example we have in mind, is the so called stochastic perpetuity (also known as infinite horizon discounted reward). More precisely, we consider the random difference equation, where  $\Psi_n$  is an affine transformation, that is  $\Psi_n(z) = A_n z + B_n$ . The formula (1) can be written as

$$Z_{n+1} = A_{n+1} Z_n + B_{n+1}, \quad (3)$$

for  $n \in \mathbb{N}$ . It is well known that if  $\mathbb{E}[\log A_1] < 0$  and  $\mathbb{E}[\log^+ B_1] < \infty$ , then the Markov chain given by (3) has a unique stationary distribution, which has the same distribution as the random variable

$$Z \triangleq \sum_{n=0}^{\infty} B_{n+1} e^{S_n},$$

where  $S_n \triangleq S_{n-1} + X_n$ ,  $S_0 \triangleq 0$  and  $X_n \triangleq \log A_n$ . Such random variables can be found in the analysis of probabilistic algorithms or financial mathematics, where  $Z$  would be called a stochastic perpetuity. If one interprets  $-\log A_n$  as the interest rate at time  $n$ , then  $Z$  is the present value of a bond that generates  $B_n$  unit of money at each time point  $n$ . Perpetuities also occur in the context of ruin problems with investments, in the study of financial time series such as ARCH-type processes (cf. e.g. Embrechts, Klüppelberg, & Mikosch, 1997), in tail asymptotics for exponential functionals of Lévy processes (see e.g. Maulik & Zwart, 2006), etc. A book devoted to (3) is Buraczewski, Damek, and Mikosch (2016). Although some particular cases exist that allow for an explicit analysis (see e.g. Vervaat, 1979), it is hard to come up with 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 where  $x$  in  $\mathbb{P}(T(x) < \infty)$  is large, which is the focus of this paper.

In this paper we develop a state-dependent importance sampling algorithm that can be proved to be strongly efficient. By state-dependent, we mean that the importance sampling distribution for generating  $Z^{(n+1)}$  is dependent on the current state  $Z^{(n)}$ . We say that an estimator  $L(x)$  for  $\mathbb{P}(T(x) < \infty)$  is strongly efficient (for a discussion of efficiency in rare-event simulation, see e.g. Asmussen & Glynn, 2007) if

$$\sup_{x>1} \frac{\mathbb{E}L^2(x)}{\mathbb{P}(T(x) < \infty)^2} < \infty. \quad (4)$$

To explain the idea behind our algorithm, consider a stochastic perpetuity, where  $B_n = 1$ . One difficulty that arises in our setting—where  $\log A_1$  is heavy-tailed—is that the Cramér condition is not satisfied (a study of the Cramér case can be found in Blanchet, Lam, & Zwart, 2012), and hence, standard techniques such as exponential change of measure cannot be used. The algorithm we provide in the present paper is based on the fact that the stochastic perpetuity is closely related to the maximum of a random walk. More precisely, for  $\gamma \in (0, -\mathbb{E}X_1)$  we observe that

$$Z = \sum_{n=0}^{\infty} \exp(S_n) = \sum_{n=0}^{\infty} \exp\{S_n + n\gamma\} \exp(-n\gamma) \leq \exp\left\{\max_{n \geq 0} (S_n + n\gamma)\right\} \frac{1}{1 - e^{-\gamma}}. \quad (5)$$

For a general sequence  $\{\Psi_n\}_{n \in \mathbb{N}}$ , we construct a slightly more involved upper bounding random walk and use it to construct a coupling. This allows us to leverage an importance sampling algorithm designed for random walks, in Blanchet and Glynn (2008). We can extend this idea to a general Markov chain given by (1). The tail asymptotics in this case have been derived by Dyszewski (2016). Our extension of (5) leads to a shorter proof of the asymptotic upper bound given in that paper.

Note that  $Z$  is defined over an infinite horizon, and hence, requires an infinite amount of computational effort for generating each sample; a natural approach to address such an issue is to work with approximations by finite-time truncation. In this paper we study the bias introduced by such approximations and show that our estimator has a vanishing relative bias as  $x \rightarrow \infty$ . We also study the asymptotic behavior of the bias with respect to the truncation time. We show that the relative bias converges to 0 at a polynomial rate, which depends on the moment condition of  $\log A_1$ . It should be mentioned that such a convergence rate is due to the heavy-tailed nature of  $\log A_1$ ; in case Cramér condition holds for  $\log A_1$ , geometric convergence rates typically ensue, i.e, the relative bias converges exponentially to 0 (cf. e.g. Basrak, Davis, & Mikosch, 2002, Theorem 2.8). By identifying such a convergence rate and proving a uniform bound on the relative moment of the associated estimator (slightly

stronger statement than the strong efficiency proved in Blanchet & Glynn, 2008), we show that one can apply the bias elimination technique studied in Rhee and Glynn (2015) to construct strongly efficient and unbiased estimators.

The rest of the paper is organized as follows. In Section 2 we briefly review an efficient simulation algorithm for the maximum of heavy-tailed random walks proposed by Blanchet and Glynn (2008). In Section 3 we introduce the random objects that can be handled by our algorithm. In Section 4 and 5 we use the result from Section 3 to prove an asymptotic upper bound of  $\mathbb{P}(Z > x)$ . Moreover, we construct an efficient algorithm for estimating  $\mathbb{P}(Z > x)$  and analyze its relative bias. The main result of this paper is given as Theorem 6. In Section 6 we analyze the asymptotic behavior of the bias with respect to the truncation time, based on a simple example; we apply the method studied in Rhee and Glynn (2015) to obtain an unbiased estimator. In Section 7 we present our computational results.

## 2 Notations and Preliminary Results

In this section we will first introduce several notations, then we will recall some well known preliminary results.

Let  $x^+ = \max(x, 0)$  denote the positive part of  $x$  and let  $\log^+(x) = \max(\log x, 0) = \log(\max(x, 1))$ . We first recall the following lemma, which will be very useful in validating our new estimator.

**Lemma 1** (Glynn, 2012). *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 non-negative martingale that is adapted to  $\{Y_n\}_{n \in \mathbb{N}}$  for which  $\mathbb{E}M_0 = 1$ . Let  $T$  be a stopping time adapted to  $\{Y_n\}_{n \in \mathbb{N}}$ . Define a sequence of probability measures as  $\mathbb{P}_n(A') = \mathbb{E}\mathbb{1}_{A'}M_n$ , for  $A' \in \mathcal{F}$ . Then there exists a probability measure  $\tilde{\mathbb{P}}$ , such that  $\tilde{\mathbb{P}}(A') = \mathbb{P}_n(A')$ , for  $A' \in \mathcal{F}$  and  $n \in \mathbb{N}$ . Furthermore, we have that  $\mathbb{E}\mathbb{1}_{\{T < \infty\}} = \tilde{\mathbb{E}}\mathbb{1}_{\{T < \infty\}}M_T^{-1}$ .*

Our goal is to find a suitable martingale  $M_n$  such that the strong efficiency criterion in (4) is satisfied. Here we consider first a useful example proposed by Blanchet and Glynn (2008), where the authors develop 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 example, we introduce the following definition.

**Definition 1.** A random variable  $X$  is said to possess a long tail, if for every  $c \in \mathbb{R}$ , we have that  $\mathbb{P}(X > t + c) \sim \mathbb{P}(X > t)$  as  $t \rightarrow \infty$ .  $X$  is called subexponential if  $\mathbb{P}(X_1^+ + X_2^+ > t) \sim 2\mathbb{P}(X^+ > t)$  as  $t \rightarrow \infty$ , where  $X_1$  and  $X_2$  are independent copies of  $X$ . Moreover,  $X$  is said to belong to the family  $S^*$  if the following holds

$$2\mathbb{E}X^+\mathbb{P}(X > t) \sim \int_0^t \mathbb{P}(X > t - s)\mathbb{P}(X > s)ds$$

as  $t \rightarrow \infty$ .

If  $X_1$  belongs to  $S^*$ , then both the distribution of  $X_1$  and its integrated tail are subexponential (cf. Klüppelberg, 1988, Theorem 3.2) and, in particular, long tailed. Furthermore, the Pakes-Veraverbeke's theorem (cf. e.g. Veraverbeke, 1977 and Zachary, 2004) says

$$\mathbb{P}\left(\max_{n \geq 0} S_n > x\right) \sim -\frac{1}{\mathbb{E}X_1} \int_x^\infty \mathbb{P}(X_1 > t)dt, \quad (6)$$

as  $x \rightarrow \infty$ , where  $S_n = S_{n-1} + X_n$ .

**Example 1.** 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.,  $S_n = S_{n-1} + X_n$ ,  $S_0 = 0$ . Assume that  $\mathbb{E}X_1 < 0$  and  $X_1$  belongs to  $S^*$ . We are interested in estimating  $\mathbb{P}(\tau(x) < \infty) = \mathbb{P}(\max_{n \geq 0} S_n > x)$ , where  $\tau(x) = \inf\{n \geq 0 : S_n > x\}$ . Let  $v(z)$  be a positive function on  $(-\infty, x)$ , let  $P(y, dz)$  denote the transition kernel of the random walk. Instead of  $P(y, dz)$ , one can simulate the random walk via another transition kernel

$$Q(y, dz) = P(y, dz) \frac{v(z)}{w(y)}, \quad \forall y \in (-\infty, x], z \in \mathbb{R}, \quad (7)$$

where  $w(y)$  is the normalization constant that is given by  $w(y) = \int_{\mathbb{R}} v(z)P(y, dz)$ . Choosing

$$M_n^{-1} = \prod_{k=1}^n \frac{w(S_{k-1})}{v(S_k)}$$

and applying Lemma 1, this yields a potential candidate of the estimator, which has the form  $L(x) = \mathbb{1}_{\{\tau(x) < \infty\}} M_{\tau(x)}^{-1}$ . It is “a potential candidate” because for each  $a \leq 0$ , any  $M_n^{-1}$  constructed by the pair  $w(\cdot + a)$  and  $v(\cdot + a)$  is also a possible choice. Define a non-negative random variable  $W$  that is independent of  $\{X_n\}_{n \in \mathbb{N}}$  with tail probability

$$\mathbb{P}(W > t) \triangleq \min \left[ 1, -\frac{1}{\mathbb{E}X_1} \int_t^\infty \mathbb{P}(X_1 > s) ds \right]. \quad (8)$$

Blanchet and Glynn (2008) propose to choose

$$v(z) \triangleq \mathbb{P}(W > -(z - x)), \quad (9)$$

and

$$w(y) \triangleq \mathbb{P}(X_1 + W > -(y - x)). \quad (10)$$

From (6) we can see that the choice above of  $v(\cdot)$  and  $w(\cdot)$  leads to a good approximation of the so-called zero-variance importance distribution, which involves sampling from the conditional distribution of the random walk given  $\{\tau(x) < \infty\}$  (cf. Blanchet & Glynn, 2008, Theorem 1). By showing

$$w(y) - v(y) = o(\mathbb{P}(X_1 > -y)), \quad \text{as } y \rightarrow -\infty, \quad (11)$$

(for details see Blanchet & Glynn, 2008, Proposition 3) and, for each  $\delta \in (0, 1)$ , the existence of a constant  $a_* = a_*(\delta) \in (-\infty, 0]$  such that

$$-\delta \leq \frac{v^2(y) - w^2(y)}{\mathbb{P}(X_1 > -y)w(y)}, \quad \forall y \leq x + a_*, \quad (12)$$

the authors were able to control the second moment of the estimator via a Lyapunov bound (for details see Blanchet & Glynn, 2008, Theorem 2, Proposition 2, Proposition 3). We summarize one of their results in the next theorem, which will prove to be useful in our context.

**Theorem 2.** (Blanchet & Glynn, 2008, Theorem 3) *Suppose that  $\mathbb{E}X_1 < 0$  and  $X_1$  belongs to  $S^*$ . Let  $v$  and  $w$  be defined as in (9) and (10). For fixed  $\delta \in (0, 1)$ , set  $a_* = a_*(\delta) \leq 0$  satisfying (12). Let an unbiased estimator of  $\mathbb{P}(\max_{n \in \mathbb{N}} S_n > x)$  be given by*

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

Then

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

where  $\mathbb{E}^{Q_{a_*}}$  denotes the expectation w.r.t. the random process  $\{S_n\}_{n \in \mathbb{N}}$  having a one-step transition kernel

$$\begin{aligned} Q_{a_*}(y, dz) &= P(y, dz) \frac{v(z + a_*)}{w(y + a_*)} \\ &= \frac{\mathbb{P}(y + X_1 \in dz) v(z + a_*)}{w(y + a_*)} \\ &= \mathbb{P}(y + X_1 \in dz \mid X_1 + W > -(y - x) - a_*). \end{aligned}$$

Theorem 2 implies that the following algorithm is strongly efficient.

**Algorithm 1.**

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

**STEP 1.** Initialize  $s \leftarrow 0$  and  $L \leftarrow 1$ .

**STEP 2.** Set  $s' \leftarrow s$ , generate a random variable  $Y$  with the distribution given by

$$\mathbb{P}(Y \in dz) = \mathbb{P}(s' + X_1 \in dz \mid X_1 + W > -(s' - x) - a_*),$$

where  $W$  is defined as in (8). Update  $s \leftarrow s' + Y$  and

$$L \leftarrow \frac{w(s' + a_*)}{v(s + a_*)} L.$$

**STEP 3.** If  $s > x$  then return  $L$ . Otherwise, go to STEP 2.

### 3 Stochastic Perpetuity and Iterated Random Functions

In this section we specify the random object that can be handled by our algorithm. We start with an example of stochastic perpetuity and construct a stochastic upper bound that can be written as a functional of a suitable random walk  $S_n(\gamma)$ , which will be defined in three different levels of generality—Example 2, general stochastic perpetuities, and stochastic recursions of the form (19). Furthermore, using the upper bound we can define crossing levels  $s(x)$  (which will be defined in three different levels of generality as well) and a stopping time

$$\tau_\gamma(x) = \inf\{n \geq 0 : S_n(\gamma) > s(x)\},$$

such that

$$\{Z > x\} \subseteq \{\max_{n \geq 0} S_n(\gamma) > s(x)\} \quad \text{and} \quad \tau_\gamma(x) \leq T(x). \quad (13)$$

Since the change of measure proposed in Blanchet and Glynn (2008) 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  $Z^{(m)}$ , until the stopping time  $\tau_\gamma(x)$ . By doing this, 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 Example 1. Then we simulate the path of the random walk after  $\tau_\gamma(x)$  under the original measure. In the second step we extend the method to the general case. Other properties such as efficiency will be discussed in Section 4 and 5.

#### 3.1 Stochastic Perpetuity

To illustrate our idea, let us consider a simple example, namely, a stochastic perpetuity that generates exact one unit of money at each time point  $n$ .

**Example 2.** Consider the Markov chain defined via the random difference equations

$$Z_{n+1} = A_{n+1}Z_n + 1, \quad t \in \mathbb{N}, \quad (14)$$

where  $\mathbb{E}[\log A_1] < 0$  and  $\{A_n\}_{n \in \mathbb{N}}$  is an i.i.d. sequence of positive random variables, which is independent of  $Z_0$ . The unique stationary distribution of this Markov chain has the same distribution as the random variable

$$Z \triangleq \sum_{n=0}^{\infty} e^{S_n},$$

where  $S_n \triangleq S_{n-1} + X_n$ ,  $S_0 \triangleq 0$  and  $X_n \triangleq \log A_n$ . Let  $\gamma_1 \in (0, -\mathbb{E}X_1)$  be fixed. For the stochastic perpetuity  $Z$ , we observe that

$$Z = \sum_{n=0}^{\infty} \exp(S_n) = \sum_{n=0}^{\infty} \exp\{S_n + n\gamma_1\} \exp(-n\gamma_1) \leq \exp\left\{\max_{n \geq 0}(S_n + n\gamma_1)\right\} \frac{1}{1 - e^{-\gamma_1}}. \quad (15)$$

Using (15) we can define  $s(x) \triangleq \log x + \log(1 - e^{-\gamma_1})$  and  $\tau_\gamma(x) \triangleq \inf\{n \geq 0 : S_n + n\gamma_1 > s(x)\}$ , such that (13) holds. To see  $\tau_\gamma(x) \leq T(x)$ , suppose  $\tau_\gamma(x) = N_\tau$ , we have that  $\max_{n \leq N_\tau - 1} S_n(\gamma) \leq s(x)$ . Combining this with the fact that

$$\begin{aligned} \sum_{n=0}^{N_\tau-1} \exp(S_n) &= \sum_{n=0}^{N_\tau-1} \exp\{S_n + n\gamma_1\} \exp(-n\gamma_1) \\ &\leq \exp\left\{\max_{n \leq N_\tau-1} S_n + n\gamma_1\right\} \frac{1 - e^{N_\tau}}{1 - e^{-\gamma_1}} \\ &\leq \exp(s(x)) \frac{1}{1 - e^{-\gamma_1}} \\ &= x, \end{aligned}$$

we can conclude that  $T(x) \geq N_\tau = \tau_\gamma(x)$ .

Example 2 shows that we can bound  $Z$  by a functional of the mean-shifted random walk, if we assume that there is no randomness in  $B_n$ . However, extending the idea to general stochastic perpetuities is not straightforward. The reason is that we can not deal  $B_n$  separately due to the potential dependence structure between  $\{A_n\}_{n \in \mathbb{N}}$  and  $\{B_n\}_{n \in \mathbb{N}}$ . To be precise, consider the Markov chain defined via the stochastic difference equations

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

where  $\mathbb{E}[\log A_1] < 0$ ,  $\mathbb{E}[\log^+ B_1] < \infty$  and  $\{(A_n, B_n)\}_{n \in \mathbb{N}}$  is an i.i.d. sequence of positive random vectors that are independent of  $Z_0$ . First we need to assume the following.

*Assumption 1.* Assume that  $(A_1, B_1)$  satisfies the conditions as follows.

- a) Let  $A_1, B_1 > 0$  a.s,  $\mathbb{E}[\log A_1] < 0$  and  $\mathbb{E}[\log^+ B_1] < \infty$ .
- b)  $\mathbb{E}[(\log^+(\max(A_1, B_1)))^{1+\eta}] < \infty$ , for some  $\eta > 0$ .
- c) For  $(A_1, B_1)$  we have that

$$\mathbb{P}(A_1 > x, B_1 \leq -x) = o(\mathbb{P}(\max(A_1, B_1) > x)).$$

Under Assumption 1, it is well known (cf. Buraczewski et al., 2016 and Dyszewski, 2016) that the unique stationary distribution of this Markov chain exists, has right-unbounded support and has the same distribution as the random variable

$$Z \triangleq \sum_{n=0}^{\infty} B_{n+1} e^{S_n},$$

where  $S_n \triangleq S_{n-1} + X_n$ ,  $S_0 \triangleq 0$  and  $X_n \triangleq \log A_n$ . In the next step we want to construct an upper bound for  $Z$ , which can be written as a functional of a suitable random walk  $S_n(\gamma)$ . We have the following Lemma.

**Lemma 3.** *Under Assumption 1, there exists a constant  $\gamma_2$  such that*

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

*Moreover, there exists a constant  $\gamma_1 \in (0, -\mathbb{E} \max(\log^+ B_1 - \gamma_2, \log A_1))$  such that*

$$Z \leq e^{\gamma_2} \sum_{n=0}^{\infty} e^{S'_n} \leq \exp\left\{\max_{n \geq 0} S_n(\gamma)\right\} \frac{e^{\gamma_2}}{1 - e^{-\gamma_1}}, \quad (16)$$

*where  $\gamma = (\gamma_1, \gamma_2)$ ,  $S'_n = S'_{n-1} + \max(\log^+ B_n - \gamma_2, \log A_n)$  and  $S_n(\gamma) = S'_n + n\gamma_1$ .*

*Proof.* Under Assumption 1 a) we obtain that

$$\lim_{\gamma'_2 \rightarrow \infty} \max(\log^+ B_1 - \gamma'_2, \log A_1) = \log A_1, \text{ a.s.}$$

since  $\log^+ B_1$  and  $\log A_1$  are finite a.s. Furthermore, we have that

$$|\max(\log^+ B_1 - \gamma'_2, \log A_1)| \leq |\max(\log^+ B_1, \log A_1)| + |\log A_1|. \quad (17)$$

Using Assumption 1 b) we have that  $\mathbb{E} \log^+ A_1 \leq \mathbb{E} \log^+ \max(A_1, B_1) < \infty$ . Since  $\mathbb{E} \log A_1 < 0$ , we conclude that  $\mathbb{E} |\log A_1| < \infty$ . Combining this with (17) we have that for any  $\gamma_1 > 0$ ,

$$\begin{aligned} \mathbb{E} |\max(\log^+ B_1 - \gamma'_1, \log A_1)| &\leq \mathbb{E} |\max(\log^+ B_1, \log A_1)| + \mathbb{E} |\log A_1| \\ &= \mathbb{E} |\log^+(\max(B_1, A_1))| + \mathbb{E} |\log A_1| < \infty. \end{aligned}$$

Using the dominated convergence theorem we obtain the existence of  $\gamma_2$ . For the stochastic perpetuity  $Z$  and the constant  $\gamma_2 > 0$ , we have that

$$Z \leq \sum_{n=0}^{\infty} \max(B_{n+1}, 1) e^{S_n} = e^{\gamma_2} \sum_{n=0}^{\infty} e^{(\log^+ B_{n+1} - \gamma_2) + S_n} \leq e^{\gamma_2} \sum_{n=0}^{\infty} e^{S'_n}, \quad (18)$$

where  $S'_n = S'_{n-1} + \max(\log^+ B_n - \gamma_2, \log A_n)$ . To see the last inequality, comparing  $S'_{n+1}$  with  $(\log^+ B_{n+1} - \gamma_2) + S_n$  component wise, we have that

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

Now let  $\gamma_1 \in (0, -\mathbb{E} \max(\log^+ B_1 - \gamma_2, \log A_1))$  be fixed. From (18), we observe that

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

where  $\gamma = (\gamma_1, \gamma_2)$  and  $S_n(\gamma) \triangleq S_{n-1}(\gamma) + \max(\log^+ B_n - \gamma_2, \log A_n) + \gamma_1 = S'_n + n\gamma_1$ .  $\square$

Now from (16) we can define  $s(x) \triangleq \log x - \gamma_2 + \log(1 - e^{-\gamma_1})$  and  $\tau_\gamma(x) \triangleq \inf\{n \geq 0 : S_n(\gamma) > s(x)\}$ , such that (13) holds.

### 3.2 Iterated Random Functions

As we indicated in the introduction, stochastic perpetuities can be considered as a special case of (1) with  $\Psi_n$  being an affine transformation. Thus, it is natural to consider the Markov chain given by (1), where for each  $\Psi_n$  there exists a random vector  $(A_n, B_n, D_n)$  satisfying

$$A_n z + B_n - D_n \leq \Psi_n(t) \leq A_n z^+ + B_n^+ + D_n, \quad (19)$$

for  $z \in \mathbb{R}$ . We assume that  $\{(A_n, B_n, D_n)\}_{n \in \mathbb{N}}$  are i.i.d. and  $\{\Psi_n\}_{n \in \mathbb{N}}$  is a sequence of i.i.d. positive random Lipschitz functions with

$$\text{Lip}(\Psi_n) \triangleq \sup_{t_1 \neq t_2} \left| \frac{\Psi_n(t_1) - \Psi_n(t_2)}{t_1 - t_2} \right|.$$

Analogous to Assumption 1, we assume the following holds.

*Assumption 2.* Assume that  $\Psi_n$  is such that  $Z^{(n)}$  defined as in (2) is increasing in  $n$ . Moreover, assume that (19) hold, and that  $(A_1, B_1, D_1)$  satisfies the following conditions:

- a)  $A_1, D_1 > 0$  a.s,  $\mathbb{E}[\log A_1] > -\infty$  and  $\mathbb{E}[\log \text{Lip}(\Psi_1)] < 0$ . Moreover,  $\mathbb{E}[\log^+ |B_1^+ + D_1|] < \infty$  and  $\mathbb{E}[\log^+ |B_1 - D_1|] < \infty$ .
- b)  $\mathbb{E}[(\log^+(\max(A_1, B_1)))^{1+\eta}] < \infty$ , for some  $\eta > 0$ .
- c) The following tail behaviors

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

and

$$\mathbb{P}(A_1 > x, B_1 - D_1 \leq -x) = o(\mathbb{P}(\max(A_1, B_1) > x))$$

hold.

*Remark 1.* We want to mention that there are interesting examples that satisfy the increasing property of the backward iteration. For instance, consider the stochastic equation given by

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

This corresponds to a second order random polynomial equation, which is studied by Goldie (1991).

Define  $Z \triangleq \lim_{n \rightarrow \infty} \Psi_{1:n}(0)$ . Recall that (cf. Dyszewski, 2016), under Assumption 2, the unique stationary solution to (1) exists, has the same distribution as  $Z$ , has right-unbounded support and can be bounded from above with a stochastic perpetuity  $\bar{Z}$ , which is given by

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

where  $\bar{B}_n \triangleq \max(B_n^+ + D_n, 1)$ ,  $S_n \triangleq S_{n-1} + \log(A_n)$  and  $S_0 \triangleq 0$  (cf. Dyszewski, 2016). Analogous to the previous section, our goal should be to 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  $S_n(\gamma)$ . First we claim the following lemma.

**Lemma 4.** *Under Assumption 2, there exists a constant  $\gamma_2$  such that*

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

Moreover, there exists a constant  $\gamma_1 \in (0, -\mathbb{E} \max(\log^+ B_1 - \gamma_2, \log A_1))$  such that

$$Z \leq e^{\gamma_2} \sum_{n=0}^{\infty} e^{S'_n} \leq \exp\left\{\max_{n \geq 0} S_n(\gamma)\right\} \frac{e^{\gamma_2}}{1 - e^{-\gamma_1}}, \quad (20)$$

where  $\gamma = (\gamma_1, \gamma_2)$ ,  $S'_n = S'_{n-1} + \max(\log^+(B_n^+ + D_n) - \gamma_2, \log A_n)$  and  $S_n(\gamma) = S'_n + n\gamma_1$ .

*Proof.* Analogous to the proof of Lemma 3. □

Now from (20) we can define  $s(x) \triangleq \log x - \gamma_2 + \log(1 - e^{-\gamma_1})$  and  $\tau_\gamma(x) \triangleq \inf\{n \geq 0 : S_n(\gamma) > s(x)\}$ , such that (13) holds.

## 4 Asymptotic Upper Bound and Strong Efficiency

Recall that in Section 3 we have developed a stochastic upper bound for each of our random objects. Moreover, using these upper bounds we also have defined a crossing level  $s(x)$  and a stopping time

$$\tau_\gamma(x) = \inf\{n \geq 0 : S_n(\gamma) > s(x)\},$$

such that  $\{Z > x\} \subseteq \{\max_{n \geq 0} S_n(\gamma) > s(x)\}$  and  $\tau_\gamma(x) \leq T(x)$ , for each case respectively.

It turns out that the upper bounds we derived in the previous section are not only helpful for constructing our algorithm, they can also be used to derive an asymptotic upper bound for  $\mathbb{P}(Z > x)$  as  $x \rightarrow \infty$ . In this section, we first analyze the asymptotic behavior of  $\mathbb{P}(Z > x)$ . After that, we show that our estimator is strongly efficient.



## 4.1 Asymptotic Upper Bound

In the theory of large deviations, one is interested in results such as the asymptotic behavior of  $\mathbb{P}(Z > x)$  as  $x \rightarrow \infty$ . More precisely, we are interested in finding a function  $f(x)$  such that  $\mathbb{P}(Z > x) \sim f(x)$  as  $x \rightarrow \infty$ . Usually, obtaining an asymptotic lower bound is easier than obtaining an upper bound. It turns out that the stochastic upper bounds we derived for  $Z$  can be used in deriving the asymptotic upper bound of  $\mathbb{P}(Z > x)$ . To illustrate this, let us get back to our example.

**Example 2** (continued). Consider a stochastic perpetuity with  $B_n = 1$ , i.e.  $Z = \sum_{n=0}^{\infty} e^{S_n}$ . Moreover, assume that the integrated tail of  $X_1$  is subexponential. On the one hand, using (18), we obtain that

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

Since the integrated tail of  $X_1 + \gamma_1$  is also subexponential, applying the Pakes-Veraverbekes Theorem we have that

$$\limsup_{x \rightarrow \infty} \frac{\mathbb{P}(Z > x)}{\int_{\log x}^{\infty} \mathbb{P}(X_1 > t) dt} \leq \limsup_{x \rightarrow \infty} \frac{\mathbb{P}\left(\max_{n \geq 0} S_n(\gamma) > s(x)\right)}{\int_{\log x}^{\infty} \mathbb{P}(X_1 > t) dt} = -\frac{1}{\mathbb{E}X_1 + \gamma_1}. \quad (21)$$

Letting  $\gamma_1 \rightarrow 0$  we conclude that

$$\limsup_{x \rightarrow \infty} \frac{\mathbb{P}(Z > x)}{\int_{\log x}^{\infty} \mathbb{P}(X_1 > t) dt} \leq -\frac{1}{\mathbb{E}X_1}.$$

On the other hand we observe that  $Z \geq \exp(\max_{n \geq 0} S_n)$ . Therefore applying again the Pakes-Veraverbekes Theorem we obtain that

$$\liminf_{x \rightarrow \infty} \frac{\mathbb{P}(Z > x)}{\int_{\log x}^{\infty} \mathbb{P}(X_1 > t) dt} \geq \liminf_{x \rightarrow \infty} \frac{\mathbb{P}\left(\max_{n \geq 0} S_n > \log x\right)}{\int_{\log x}^{\infty} \mathbb{P}(X_1 > t) dt} = -\frac{1}{\mathbb{E}X_1}. \quad (22)$$

Combining (21) and (22) we conclude that

$$\mathbb{P}(Z > x) \sim -\frac{1}{\mathbb{E}X_1} \int_{\log x}^{\infty} \mathbb{P}(X_1 > t) dt.$$

Consider the stationary distribution of the Markov chain given by (1). As we indicated earlier, Dyszewski (2016) shows that under subexponential assumptions on the random variable  $\log^+ \max(A_1, B_1)$  the tail asymptotics can be described using the integrated tail function of  $\log^+ \max(A_1, B_1)$ . However, the upper bound we derived in Section 3.2 yields us a shorter proof for the asymptotic upper bound in Dyszewski (2016, Theorem 3.1).

**Lemma 5.** (A shorter proof of the asymptotic upper bound in Dyszewski, 2016, Theorem 3.1) *Let Assumption 2 hold. Furthermore, assume that the integrated tail of  $\log \max(A_1, B_1)$  is subexponential. Then we have that*

$$\limsup_{x \rightarrow \infty} \frac{\mathbb{P}(Z > x)}{\bar{F}_I(\log(x))} \leq -\frac{1}{\mathbb{E}[\log(A_1)]},$$

where  $\bar{F}_I$  denotes the integrated tail of  $\log \max(A_1, B_1)$ .

*Proof.* From the upper bound we constructed in Section 3.2, we know that

$$\mathbb{P}(Z > x) \leq \mathbb{P}\left(\max_{n \geq 0} S_n(\gamma) > s(x)\right). \quad (23)$$

Due to Assumption 2 c) we know that the integrated tail of  $\log(\max(A_1, B_1^+ + D_1, 1))$  is also subexponential. Moreover, we have the following inequality:

$$\log(\max(A_1, B_1^+ + D_1, 1)) - \gamma_2 \leq \log(\max(A_1, e^{-\gamma_2}(B_1^+ + D_1), e^{-\gamma_2})) \leq \log(\max(A_1, B_1^+ + D_1, 1)).$$

The increments of the random walk  $S_n(\gamma)$  have a subexponential integrated tail. Using the Pakes-Veraverbeke theorem we get the following relationship for the RHS of (23), namely

$$\mathbb{P}\left(\max_{n \geq 0} S_n(\gamma) > s(x)\right) \sim -\frac{1}{\mathbb{E}[\max(\bar{B}_1 - \gamma_2, \log A_1)] + \gamma_1} \bar{F}_I(\log(x)). \quad (24)$$

Now, letting  $\gamma_2 \rightarrow \infty$  and  $\gamma_1 \rightarrow 0$  yields the result.  $\square$

*Remark 2.* Note that the assumption of an increasing backward iteration of  $\Psi_n$  is not needed in the proof of Lemma 5. Therefore, it provides indeed a shorter proof of the asymptotic upper bound given in Dyszewski (2016, Theorem 3.1).

## 4.2 Strong Efficiency

Given the asymptotic behavior of  $\mathbb{P}(Z > x)$ , we are able to show the strong efficiency of our estimator. Recall that, based on (13), our algorithm constructs a state-dependent change of measure according to the methods introduced by Blanchet and Glynn (2008). Define the following elements, which are needed for the change of measure, via

$$\mathbb{P}(W_\gamma > t) \triangleq \min\left[1, \frac{1}{\mathbb{E}S_1(\gamma)} \int_t^\infty \mathbb{P}(S_1(\gamma) > s) ds\right],$$

$$v_\gamma(z) \triangleq \mathbb{P}(W_\gamma > -(z - s(x))), \quad (25)$$

and

$$w_\gamma(y) \triangleq \mathbb{P}(S_1(\gamma) + W_\gamma > -(y - s(x))). \quad (26)$$

We propose an estimator and show its strong efficiency in the following theorem.

**Theorem 6.** *Suppose that  $\mathbb{E}S_1(\gamma) < 0$  and  $S_1(\gamma)$  belongs to  $S^*$ . Let  $v_\gamma$  and  $w_\gamma$  be defined as in (25) and (26). For fixed  $\delta \in (0, 1)$ , one can choose  $a_* = a_*(\delta) \leq 0$  so that*

$$-\delta \leq \frac{v_\gamma^2(y) - w_\gamma^2(y)}{\mathbb{P}(X_1 > -y)w_\gamma(y)}, \quad \forall y \leq s(x) + a_*.$$

Let

$$L_T(x) \triangleq \mathbb{1}_{\{T(x) < \infty\}} \prod_{k=1}^{\tau_\gamma(x)} \frac{w_\gamma(S_{k-1}(\gamma) + a_*)}{v_\gamma(S_k(\gamma) + a_*)}. \quad (27)$$

Then  $L_T(x)$  is an unbiased estimator of  $\mathbb{P}(Z > x)$  and

$$\sup_{x > 1} \frac{\mathbb{E}^{Q_{a_*}^\gamma} L_T^2(x)}{\mathbb{P}(Z > x)^2} < \infty,$$

where  $\mathbb{E}^{Q_{a_*}^\gamma}$  denotes the expectation w.r.t. the Markov chain  $\{S_n(\gamma)\}_{n \in \mathbb{N}}$  having a one-step transition kernel

$$Q_{a_*}^\gamma(y, dz) = P(y, dz) \frac{v_\gamma(z + a_*)}{w_\gamma(y + a_*)}.$$

*Proof.* 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  $\tau_\gamma(x) \leq T(x)$ , applying Lemma 1, we can conclude that

$$\mathbb{E}^{Q_{a_*}^\gamma} L_T(x) = \mathbb{P}(T(x) < \infty) = \mathbb{P}(Z > x).$$

For the strong efficiency we have that

$$\begin{aligned}
\frac{\mathbb{E}^{Q_{a_*}^\gamma} L_T^2(x)}{\mathbb{P}(Z > x)^2} &= \frac{\mathbb{E}^{Q_{a_*}^\gamma} \left[ \mathbb{1}_{\{Z > x\}} M_{\tau_\gamma}^{-2}(x) \right]}{\mathbb{P}(Z > x)^2} \\
&\leq \frac{\mathbb{E}^{Q_{a_*}^\gamma} \left[ \mathbb{1}_{\{\max_{n \geq 0} S_n(\gamma) > s(x)\}} M_{\tau_\gamma}^{-2}(x) \right]}{\mathbb{P}(Z > x)^2} \\
&= \frac{\mathbb{E}^{Q_{a_*}^\gamma} \left[ \mathbb{1}_{\{\max_{n \geq 0} S_n(\gamma) > s(x)\}} M_{\tau_\gamma}^{-2}(x) \right]}{\mathbb{P} \left( \max_{n \geq 0} S_n(\gamma) > s(x) \right)^2} \left( \frac{\mathbb{P} \left( \max_{n \geq 0} S_n(\gamma) > s(x) \right)}{\mathbb{P}(Z > x)} \right)^2,
\end{aligned}$$

where the first term in the last equation is guaranteed to be bounded over  $x \in (1, \infty)$  due to Theorem 2. Hence, only the latter term remains to be analyzed. Define

$$\chi(x) = \frac{\mathbb{P} \left( \max_{n \geq 0} S_n(\gamma) > s(x) \right)}{\mathbb{P}(Z > x)}.$$

From Dyszewski (2016, Theorem 3.1) we have that

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

Since by assumption the integrated tail  $\bar{F}_I$  is subexponential, it is in particular long tailed. Combining (24) and (28) we obtain that

$$\limsup_{x \rightarrow \infty} \chi(x) \leq \frac{\mathbb{E} \log A_1}{\mathbb{E}[\max(\bar{B}_1 - \gamma_2, \log A_1)] + \gamma_1}. \quad (29)$$

Using the fact that  $\chi(x)$  is bounded over a compact interval, we obtain that

$$\sup_{x > 1} \frac{\mathbb{P} \left( \max_{n \geq 0} S_n(\gamma) > s(x) \right)}{\mathbb{P}(Z > x)} < \infty.$$

□

## 5 Asymptotic Unbiasedness

The estimator derived in Theorem 6 requires the computation of  $\mathbb{1}_{\{Z > x\}}$ , 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, although there is plenty of literature on this topic; see, for example, Blanchet and Wallwater (2015) and Blanchet and Sigman (2011). Conditional on  $\{\tau_\gamma(x) < \infty\}$ , using the strong Markov property, we have that

$$Z = \Psi_{1:\tau_\gamma(x)}(Z'),$$

where  $Z' \triangleq \lim_{M \rightarrow \infty} \Psi_{\tau_\gamma(x)+1:\tau_\gamma(x)+M}(0) \stackrel{d}{=} Z$  is independent of  $\Psi_{1:\tau_\gamma(x)}$ . Therefore, a natural choice for approximating the distribution of  $Z'$  is a truncated sum. More precisely, letting  $M \in \mathbb{N}$  be fixed; our modified estimator takes the form

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

where  $Z'^{(M)} \triangleq \Psi_{\tau_\gamma(x)+1:\tau_\gamma(x)+M}(0)$ . To illustrate this, let us consider an extension of Example 2.

**Example 2** (continued). Consider a stochastic perpetuity with  $B_n = 1$ . Moreover, for  $\alpha > 0$ , assume that  $X_1$  is regularly varying with index  $\alpha + 1$ ; i.e, for the tail distribution of  $X_1$ , we have that  $\bar{F}(x) \sim x^{-\alpha-1}L(x)$  with  $L$  being a slowly varying function. Let  $a(x)$  denote the auxiliary function of  $X_1$  (c.f, e.g. Asmussen & Klüppelberg, 1996) that is given by  $a(x) = x/\alpha$  in the regularly varying case. On the set  $\{\tau_\gamma(x) < \infty\}$ , using the strong Markov property, we have that

$$Z = \sum_{n=0}^{\infty} e^{S_n} = A'_x Z' + B'_x,$$

where  $A'_x = e^{S_{\tau_\gamma(x)}}$ ,  $B'_x = \sum_{n=0}^{\tau_\gamma(x)-1} e^{S_n}$ , and  $Z' = \sum_{n=0}^{\infty} e^{S_{\tau_\gamma(x)+n} - S_{\tau_\gamma(x)}}$  is a random variable that is independent of  $(A'_x, B'_x)$ . Our modified estimator takes the form

$$L_T^\Delta(x, M) = \mathbb{1}_{\{\tau_\gamma(x) < \infty, A'_x Z'^{(M)} + B'_x > x\}} \prod_{k=1}^{\tau_\gamma(x)} \frac{w_\gamma(S_{k-1}(\gamma) + a_*)}{v_\gamma(S_k(\gamma) + a_*)},$$

where  $Z'^{(M)} = \sum_{n=0}^M e^{S_{\tau_\gamma(x)+n} - S_{\tau_\gamma(x)}}$ . The relative bias is defined as

$$\Delta(x, M) = \left| \frac{\mathbb{E}^{Q_{a_*}^\gamma} L_T^\Delta(x, M) - \mathbb{P}(T(x) < \infty)}{\mathbb{P}(T(x) < \infty)} \right|.$$

From  $\{\tau_\gamma(x) < \infty, \sum_{n=0}^{\tau_\gamma(x)+M} e^{S_n} > x\} \subseteq \{T(x) < \infty\} \subseteq \{\tau_\gamma(x) < \infty\}$  we obtain that

$$\Delta(x, M) = \Theta(x, M) \frac{\mathbb{P}\left(\max_{n \geq 0} S_n(\gamma) > s(x)\right)}{\mathbb{P}(Z > x)}, \quad (30)$$

where

$$\begin{aligned} \Theta(x, M) &= \mathbb{P}\left(A'_x Z'^{(M)} + B'_x \leq x, A'_x Z' + B'_x > x \mid \tau_\gamma(x) < \infty\right) \\ &= \mathbb{P}\left(\frac{\log Z'^{(M)}}{a(\log x)} \leq \frac{\log(x - B'_x)}{a(\log x)} - \frac{\log(A'_x)}{a(\log x)} < \frac{\log Z'}{a(\log x)} \mid \tau_\gamma(x) < \infty\right) \\ &= \mathbb{P}\left(\frac{\log Z'^{(M)}}{a(\log x)} \leq \frac{\log(x - B'_x)}{a(\log x)} - \frac{S_{\tau_\gamma(x)}(\gamma) - \tau_\gamma(x)\gamma}{a(\log x)} < \frac{\log Z'}{a(\log x)} \mid \tau_\gamma(x) < \infty\right). \end{aligned}$$

Note that we have seen in the proof of Theorem 6, the latter term in the RHS of (30) is bounded. Therefore, to show that the relative bias vanishes as  $x \rightarrow \infty$ , it is enough to show that  $\Theta(x, M) \rightarrow 0$  as  $x \rightarrow \infty$ . In the following corollary we derive the limiting distribution of  $\xi_x$  conditional on  $\{\tau_\gamma(x) < \infty\}$ , where

$$\xi_x \triangleq \frac{\log(x - B'_x)}{a(\log x)} - \frac{S_{\tau_\gamma(x)}(\gamma) - \tau_\gamma(x)\gamma}{a(\log x)} = \frac{\log\left(1 - \frac{B'_x}{x}\right)}{a(\log x)} - \frac{S_{\tau_\gamma(x)}(\gamma) - \log x}{a(\log x)} + \frac{\tau_\gamma(x)\gamma}{a(\log x)}.$$

**Proposition 7.** Let  $\mu \triangleq -\mathbb{E}X_1$ . Conditional on  $\{\tau_\gamma(x) < \infty\}$ ,  $\xi_x$  converges in distribution to  $\xi \stackrel{d}{=} \gamma V_\alpha / (\mu - \gamma) - T_\alpha$  as  $x \rightarrow \infty$ , where  $V_\alpha$  is a positive random variable and its tail is given by

$$\bar{G}_\alpha(x) = \mathbb{P}(V_\alpha > x) = (1 + x/\alpha)^{-\alpha}, \quad x > 0,$$

and  $T_\alpha$  is defined on the same probability space, such that  $\mathbb{P}(V_\alpha > x, T_\alpha > y) = \bar{G}_\alpha(x + y)$ . Moreover, the density of  $\xi$  is given by

$$f_\xi(y) = \frac{\mu - \gamma}{\mu} \left(1 - \frac{y}{\alpha}\right)^{-\alpha-1} \mathbb{1}_{\{y < 0\}} + \frac{\mu - \gamma}{\mu} \left(1 + \frac{(\mu - \gamma)y}{\alpha\gamma}\right)^{-\alpha-1} \mathbb{1}_{\{y \geq 0\}}.$$

Since both random variables  $\log Z'^{(M)}/a(\log x)$  and  $\log Z'/a(\log x)$  converge in probability to 0 as  $x \rightarrow \infty$ , the relative bias vanishes.

It should be noted that the result from the example above remains valid, if we assume that  $X_1$  belongs to the maximum domain of attraction of the Gumbel distribution. However, we want to show the result of vanishing relative bias in a general context as described in Section 3.2. Moreover, we are only assuming that the integrated tail of  $\log(\max(A_n, B_n))$  is subexponential. We need the following Lemma, of which the proof uses a similar technique as the proof of Theorem 1 in Palmowski and Zwart (2007).

**Lemma 8.** *Let  $\mu \triangleq -\mathbb{E}X_1$  and  $\mu_\gamma \triangleq -\mathbb{E}S_1(\gamma)$ . For  $\nu, K > 0$  consider the sets*

$$E_n^{(1)} = E_n^{(1)}(K, \nu) = \{S_j \in (-j(\mu + \nu) - K, -j(\mu - \nu) + K), j \leq n\},$$

$$E_n^{(2)} = E_n^{(2)}(K, \nu) = \{S_j(\gamma) \in (-j(\mu_\gamma + \nu) - K, -j(\mu_\gamma - \nu) + K), j \leq n\},$$

and

$$E_n^{(3)} = E_n^{(3)}(K, \nu) = \{|\underline{B}_j| \leq e^{\nu j + K}, j \leq n\},$$

where  $\underline{B}_j = B_j - D_j$ . Then, for  $\nu, \epsilon > 0$ , there exists  $K > 0$ , such that

$$\mathbb{P} \left( \bigcap_{n \geq 1} \left( E_n^{(1)} \cap E_n^{(2)} \cap E_n^{(3)} \right) \right) \geq 1 - \epsilon.$$

*Proof.* In the proof of Theorem 1 in Palmowski and Zwart (2007), the authors state 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

$$\mathbb{P}(|Y_j| \leq e^{\nu j + K}, j \leq n) \rightarrow 1,$$

as  $K \rightarrow \infty$  uniformly with respect to  $n$ . Using this argument we conclude that  $\mathbb{P}(E_n^{(3)}) \rightarrow 1$  as  $K \rightarrow \infty$  uniformly with respect to  $n$ . Further, combining this fact with the SLLN for  $\{S_n\}_{n \geq 0}$  and  $\{S_n(\gamma)\}_{n \geq 0}$  (for details see eg. Asmussen, Schmidli, & Schmidt, 1999, Lemma 3.1), we can always take  $K$  large enough such that

$$\mathbb{P}(E_n^{(1)} \cap E_n^{(2)} \cap E_n^{(3)}) \geq 1 - \epsilon,$$

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.  $\square$

Due to the fact that  $\left\{ \tau_\gamma(x) < \infty, \Psi_{1:\tau_\gamma(x)}(Z'^{(M)}) > x \right\} \subseteq \{T(x) < \infty\}$ , in order to prove the vanishing relative bias result, it is sufficient to show that

$$\liminf_{x \rightarrow \infty} \frac{\mathbb{P} \left( \tau_\gamma(x) < \infty, \Psi_{1:\tau_\gamma(x)}(Z'^{(M)}) > x \right)}{\mathbb{P}(T(x) < \infty)} \geq 1. \quad (31)$$

In the following theorem we use a similar proof technique as in the proof of Theorem 3.1 in Dyszewski (2016).

**Theorem 9.** *Let Assumption 2 hold. Moreover, we are assuming that the integrated tail of  $\log(\max(A_1, B_1))$  is subexponential. Then (31) holds.*

*Proof.* Define

$$E_n = E_n^{(1)} \cap E_n^{(2)} \cap E_n^{(3)} \cap \left\{ \max(A_{n+1}, \underline{B}_{n+1}) > x e^{n(\mu+\nu)+L+K}, \underline{B}_{n+1} \geq -x e^{n(\mu-\nu)-K} \right\} \cap \{Z'^{(1)} > \nu\},$$

where  $L > 0$  is chosen to be large enough, such that the sets  $\{E_n\}_{n \geq 0}$  are disjoint. Moreover, we can show that  $E_n \subseteq \{\tau_\gamma(x) = n + 1, \Psi_{1:\tau_\gamma(x)}(Z^{(1)}) > x\} \subseteq \{\tau_\gamma(x) < \infty, \Psi_{1:\tau_\gamma(x)}(Z^{(M)}) > x\}$ . To see this, on  $E_n$  we have that

$$\begin{aligned}
\Psi_{1:n+1}(Z^{(1)}) &\geq \sum_{k=0}^{n-1} \underline{B}_{k+1} \prod_{j=1}^k A_j + \left( \underline{B}_{n+1} + Z^{(1)} A_{n+1} \right) \prod_{j=1}^n A_j \\
&\geq - \sum_{k=0}^{n-1} |\underline{B}_{k+1}| \prod_{j=1}^k A_j + \left( \underline{B}_{n+1} + x e^{n(\mu-\nu)-K} + Z^{(1)} A_{n+1} \right) \prod_{j=1}^n A_j - x e^{n(\mu-\nu)-K} \prod_{j=1}^n A_j \\
&\geq - \frac{e^{2K}}{1 - e^{-\mu+2\nu}} + \min(\nu, 1) \max \left( A_{n+1}, \underline{B}_{n+1} + x e^{n(\mu-\nu)-K} \right) e^{-n(\mu+\nu)-K} - x \\
&\geq - \frac{e^{2K}}{1 - e^{-\mu+2\nu}} + \min(\nu, 1) \max \left( A_{n+1}, \underline{B}_{n+1} \right) e^{-n(\mu+\nu)-K} - x \\
&\geq - \frac{e^{2K}}{1 - e^{-\mu+2\nu}} + \min(\nu, 1) x e^L - x > x,
\end{aligned}$$

for sufficiently large  $L$  that does not depend on  $x$ . Since  $\{S_j(\gamma)\}_{j \leq n}$  is bounded by  $K$ ,  $\mu > \mu_\gamma$  and

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

for sufficiently large  $L$  that does not depend on  $x$ , we can also conclude that  $\tau_\gamma(x) = n + 1 < \infty$  by taking  $x$  sufficiently large. This implies that

$$\begin{aligned}
\mathbb{P} \left( \tau_\gamma(x) < \infty, \Psi_{1:\tau_\gamma(x)}(Z^{(M)}) > x \right) &\geq \sum_{n \geq 0} \mathbb{P}(E_n) \\
&\geq (1 - \epsilon) \mathbb{P}(Z^{(1)} > \nu) \sum_{n \geq 0} \left\{ \mathbb{P} \left( \max(A_1, \underline{B}_1) > x e^{n(\mu+\nu)+L+K} \right) \right. \\
&\quad \left. - \mathbb{P} \left( A_1 > x e^{n(\mu+\nu)+L+K}, \underline{B}_1 < -x e^{n(\mu-\nu)+K} \right) \right\}. \tag{32}
\end{aligned}$$

From Assumption 2c) we conclude that, for any  $\epsilon' > 0$ , by taking sufficiently large  $x$ , the following holds

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

Combining this with (32), we obtain that

$$\begin{aligned}
&\mathbb{P} \left( \tau_\gamma(x) < \infty, \Psi_{1:\tau_\gamma(x)}(Z^{(M)}) > x \right) \\
&\geq (1 - \epsilon) \mathbb{P}(Z^{(1)} > \nu) \sum_{n \geq 0} \left\{ \mathbb{P} \left( \max(A_1, \underline{B}_1) > x e^{n(\mu+\nu)+L+K} \right) \right. \\
&\quad \left. - \epsilon' \mathbb{P} \left( \max(A_1, \underline{B}_1) > x e^{n(\mu-\nu)+K} \right) \right\}. \tag{33}
\end{aligned}$$

Since  $\mathbb{P}(\max(A_1, \underline{B}_1) > y)$  is decreasing in  $y$ , we observe that, for any  $n$ ,

$$\mathbb{P} \left( \max(A_1, \underline{B}_1) > x e^{n(\mu+\nu)+L+K} \right) \geq \frac{1}{\mu + \nu} \int_{\log x + L + K + n(\mu+\nu)}^{\log x + L + K + (n+1)(\mu+\nu)} \mathbb{P}(\log \max(A_1, \underline{B}_1) > y) dy,$$

and that

$$\mathbb{P}\left(\max(A_1, \underline{B}_1) > xe^{n(\mu-\nu)+K}\right) \leq \frac{1}{\mu-\nu} \int_{\log x+K+(n-1)(\mu-\nu)}^{\log x+K+n(\mu-\nu)} \mathbb{P}(\log \max(A_1, \underline{B}_1) > y) dy.$$

Moreover, using the fact that  $\bar{F}_I$  is long tailed, we obtain from (33) that

$$\begin{aligned} & \mathbb{P}\left(\tau_\gamma(x) < \infty, \Psi_{1:\tau_\gamma(x)}(Z'^{(M)}) > x\right) \\ & \geq (1-\epsilon)\mathbb{P}(Z'^{(1)} > \nu) \left( \frac{1}{\mu+\nu} \bar{F}_I(\log x + L + K) - \frac{\epsilon'}{\mu-\nu} \bar{F}_I(\log x + L + K - (\mu-\nu)) \right) \\ & \sim (1-\epsilon)\mathbb{P}(Z'^{(1)} > \nu) \left( \frac{1}{\mu+\nu} - \frac{\epsilon'}{\mu-\nu} \right) \bar{F}_I(\log x) \\ & \sim \mu(1-\epsilon)\mathbb{P}(Z'^{(1)} > \nu) \left( \frac{1}{\mu+\nu} - \frac{\epsilon'}{\mu-\nu} \right) \mathbb{P}(T(x) < \infty). \end{aligned} \quad (34)$$

where in (34) we use Dyszewski (2016, Theorem 3.1). Letting  $\epsilon, \epsilon', \nu \rightarrow 0$  we obtain the result. This result implies that the relative bias converge to 0, since the numerator in (31) is always smaller than the denominator.  $\square$

Let the conditions in Theorem 6 and Theorem 9 be satisfied. The following algorithm for estimating  $\mathbb{P}(Z > x)$  has bounded relative error and vanishing relative bias as  $x \rightarrow \infty$ .

**Algorithm 2.**

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

**STEP 1.** Initialize  $s \leftarrow 0$ ,  $z \leftarrow 1$  and  $L \leftarrow 1$ .

**STEP 2.** Set  $s' \leftarrow s$  and  $z' \leftarrow z$ . Run Algorithm 1 until the random walk  $S_n(\gamma)$  crosses  $s(x)$ . Meanwhile, update  $s$  and  $L$  according to STEP 2 of Algorithm 1, then update  $z$  via the backward iteration.

**STEP 3.** Set  $i \leftarrow 0$ ,  $s' \leftarrow s$  and  $z' \leftarrow z$ . While  $i < M$ , update  $z$  via the backward iteration,  $z' \leftarrow z$  and  $i \leftarrow i + 1$ .

**STEP 4.** If  $z > x$  then return  $L$ . Otherwise, return 0.

## 6 Truncation Index and Unbiased Estimator

In this section we analyze the asymptotic behavior of the relative bias as  $M \rightarrow \infty$  for fixed  $x$ , based on which we propose an unbiased estimator for  $\mathbb{P}(Z > x)$  using the technique studied in Rhee and Glynn (2015). We first go back to the Example 2—i.e., the stochastic perpetuity example with  $B_n = 1$ .

**Example 2** (continued). Consider the Markov chain given by (14). From (30) we know that in order to analyze the relative bias we need to consider the term  $\Theta(x, M)$ , which is given by

$$\Theta(x, M) = \mathbb{P}\left(A'_x Z'^{(M)} + B'_x \leq x, A'_x Z' + B'_x > x \mid \tau_\gamma(x) < \infty\right).$$

Let  $\mathbb{P}^{(x)}(\cdot)$  denote the conditional probability of  $\mathbb{P}(\cdot \mid \tau_\gamma(x) < \infty)$  and  $\mathbb{E}^{(x)}$  the corresponding expectation operator. Note that

$$\begin{aligned} \Theta(x, M) &= \int \mathbb{1}_{\left\{Z'^{(M)} \leq \frac{x-B'_x}{A'_x}, Z' > \frac{x-B'_x}{A'_x}\right\}} d\mathbb{P}^{(x)} \\ &= \int \mathbb{P}^{(x)}\left(Z'^{(M)} \leq y, Z' > y\right) \mathbb{P}_{\xi_x}^{(x)}(dy) \\ &= \int \left\{ \mathbb{P}^{(x)}(Z' > y) - \mathbb{P}^{(x)}\left(Z'^{(M)} > y\right) \right\} \mathbb{P}_{\xi_x}^{(x)}(dy), \end{aligned} \quad (35)$$

where (35) follows from the fact that  $\{Z'^{(M)} > y\} \subseteq \{Z' > y\}$ . Using the strong Markov property we have that  $Z'^{(M)} \stackrel{d}{=} Z^{(M)}$  and  $Z' \stackrel{d}{=} Z$  under  $\mathbb{P}^{(x)}$ . Therefore, we can write (35) as

$$\Theta(x, M) = \int \left\{ \mathbb{P}(Z > y) - \mathbb{P}(Z^{(M)} > y) \right\} \mathbb{P}_{\xi_x}^{(x)}(dy).$$

Combining this with the fact that the backward iteration  $Z^{(M)}$  has the same distribution as  $Z_M$ , we obtain that

$$\Theta(x, M) = \int \left\{ \mathbb{P}(Z > y) - \mathbb{P}(Z_M > y) \right\} \mathbb{P}_{\xi_x}^{(x)}(dy) \leq d_{TV}(Z_M, Z) \quad (36)$$

where  $d_{TV}$  denotes the total variation distance. To get a handle on this quantity, we apply the Lyapunov criterion in Jarner and Roberts (2002, Theorem 3.6), which implies a polynomial convergence rate of the  $M$ -step transition kernel to the invariant distribution in the total variation norm. We assume that the Markov chain  $\{Z_n\}_{n \in \mathbb{N}}$  given by (14) is irreducible and aperiodic; this is the case, for example, if  $A_1$  has a Lebesgue density (Buraczewski et al., 2016, Lemma 2.2.2). Moreover, assume that there exists an integer  $q \geq 2$  such that  $\mathbb{E}|X_1|^q < \infty$ . In order to establish the Lyapunov condition, let  $V(x) = 1 \vee (\log x)^q$ . Note that  $V(x) = (\log x)^q \mathbb{1}_{\{x > e\}} + \mathbb{1}_{\{x \leq e\}}$  and hence the binomial expansion gives

$$\begin{aligned} PV(x) &= \mathbb{E} \left[ (\log(A_1x + 1))^q \mathbb{1}_{\{A_1x+1 > e\}} + \mathbb{1}_{\{A_1x+1 \leq e\}} \right] \\ &= \mathbb{E} \left[ \left( \log \frac{A_1x+1}{x} + \log x \right)^q \mathbb{1}_{\{A_1x+1 > e\}} + \mathbb{1}_{\{A_1x+1 \leq e\}} \right] \\ &= \mathbb{E} \left[ (\log x)^q \mathbb{1}_{\{A_1x+1 > e\}} + \sum_{i=1}^q \binom{q}{i} (\log x)^{q-i} \left( \log \frac{A_1x+1}{x} \right)^i \mathbb{1}_{\{A_1x+1 > e\}} + \mathbb{1}_{\{A_1x+1 \leq e\}} \right] \\ &= V(x) + \mathbb{E}(\log x)^q (\mathbb{1}_{\{A_1x+1 > e\}} - \mathbb{1}_{\{x > e\}}) + \mathbb{E}(\mathbb{1}_{\{A_1x+1 \leq e\}} - \mathbb{1}_{\{x \leq e\}}) \\ &\quad + q(\log x)^{q-1} \mathbb{E} \left( \log \frac{A_1x+1}{x} \mathbb{1}_{\{A_1x+1 > e\}} \right) + \mathbb{E} \sum_{i=2}^q \binom{q}{i} (\log x)^{q-i} \left( \log \frac{A_1x+1}{x} \right)^i \mathbb{1}_{\{A_1x+1 > e\}} \end{aligned}$$

For  $x > e$ ,

$$\begin{aligned} PV(x) &\leq V(x) + \mathbb{P}(A_1x + 1 \leq e) + q(\log x)^{q-1} \mathbb{E} \left( \log \frac{A_1x+1}{x} \mathbb{1}_{\{A_1x+1 > e\}} \right) \\ &\quad + \sum_{i=2}^q \binom{q}{i} (\log x)^{q-i} \mathbb{E} \left[ \left( \log \frac{A_1x+1}{x} \right)^i \mathbb{1}_{\{A_1x+1 > e\}} \right]. \end{aligned}$$

Note that  $\left| \log \frac{A_1x+1}{x} \mathbb{1}_{\{A_1x+1 > e\}} \right| \leq |\log A_1| + C$  for some constant  $C$  and noting that the right-hand-side doesn't depend on  $x$  and has finite  $q$ -th moment, there has to be  $c_i$ 's such that

$$\sum_{i=2}^q \binom{q}{i} (\log x)^{q-i} \mathbb{E} \left[ \left( \log \frac{A_1x+1}{x} \right)^i \mathbb{1}_{\{A_1x+1 > e\}} \right] \leq \sum_{i=0}^{q-2} c_i (\log x)^i \leq \epsilon (\log x)^{q-1}$$

for sufficiently large  $x$ . On the other hand, note that  $\log \frac{A_1x+1}{x} \mathbb{1}_{\{A_1x+1 > e\}}$  converges to  $X_1 = \log A_1$  almost surely as  $x \rightarrow \infty$  and hence by dominated convergence  $\mathbb{E} \left( \log \frac{A_1x+1}{x} \mathbb{1}_{\{A_1x+1 > e\}} \right) \rightarrow \mathbb{E} \log A_1 < 0$ . Therefore, for any fixed  $\epsilon > 0$ ,

$$q(\log x)^{q-1} \mathbb{E} \left( \log \frac{A_1x+1}{x} \mathbb{1}_{\{A_1x+1 > e\}} \right) \leq (q\mathbb{E}X_1 + \epsilon)(\log x)^{q-1}$$

for sufficiently large  $x$ . Choosing  $\epsilon$  so that  $q\mathbb{E}X_1 + 3\epsilon < 0$  and noting that  $\mathbb{P}(A_1x + 1 \leq e) \rightarrow 0$  as  $x \rightarrow \infty$ , as well as  $(\log x)^{q-1} = ((\log x)^q \mathbb{1}_{\{x > e\}} + \mathbb{1}_{\{x \leq e\}})^{\frac{q-1}{q}}$  for  $x > e$ , we conclude that there exists  $K$  such that

$$\begin{aligned} PV(x) &\leq V(x) + \epsilon (\log x)^{q-1} + (q\mathbb{E}X_1 + \epsilon)(\log x)^{q-1} + \epsilon (\log x)^{q-1} \\ &\leq V(x) - cV^{(q-1)/q}(x) \end{aligned}$$



for  $x > K$ , where  $c = -(q\mathbb{E}X_1 + 3\epsilon) > 0$ . Finally, since  $PV(x)$ ,  $V(x)$  and  $V^{(q-1)/q}(x)$  are bounded on  $C = [0, K]$ , there exists a constant  $b$  such that

$$PV(x) \leq V(x) - cV^{(q-1)/q}(x) + b\mathbb{1}_C,$$

which is the sufficient condition in Jarner and Roberts (2002, Theorem 3.6) for polynomial ergodicity; we conclude that the  $M$ -step transition kernel converges to the stationary distribution in the total variation norm at a polynomial rate with order  $q - 1$ , i.e., there exists a constant  $\kappa'$  satisfying

$$\Theta(x, M) \leq d_{TV}(Z_M, Z) < \kappa M^{-(q-1)}, \quad (37)$$

for all  $M \in \mathbb{N}$ . It should be noted that an exact expression of the constant  $\kappa$  can be obtained in a few special cases—for example, see e.g. Douc, Moulines, and Soulier (2007), Kalashnikov and Tsitsiashvili (1999) and the references therein. However, applying the method studied in Rhee and Glynn (2015) we can get rid of this constant altogether and obtain an unbiased, strongly efficient estimator. In order to apply the method, a sufficient condition is to bound

$$\frac{\mathbb{E}^{Q_{a^*}^\gamma} (L_T^\Delta(x, M) - L_T(x))^2}{\mathbb{P}(Z > x)^2} \quad (38)$$

by a decreasing function of  $M$  independent of  $x$ . Once we can have such a bound, we can construct an unbiased estimator that is given by

$$L_T^{\text{RG}}(x) \triangleq \sum_{i=0}^N \frac{L_T^\Delta(x, 2^i) - L_T^\Delta(x, 2^{i-1})}{\mathbb{P}(N \geq i)}, \quad (39)$$

whose second moment is

$$\sum_{i=0}^{\infty} \frac{\mathbb{E}^{Q_{a^*}^\gamma} (L_T^\Delta(x, 2^{i-1}) - L_T(x))^2 - \mathbb{E}^{Q_{a^*}^\gamma} (L_T^\Delta(x, 2^i) - L_T(x))^2}{\mathbb{P}(N \geq i)},$$

where  $N$  is a random truncation index independent of everything else and  $L_T^\Delta(x, 2^i)$  is interpreted as 0 if  $i < 0$  (for details see Rhee & Glynn, 2015, Theorem 1). It turns out that such a bound on (38) can be derived easily, if

$$\sup_{x>0} \frac{\mathbb{E}^{Q_{a^*}^\gamma} L_T^{2+\epsilon}(x)}{\mathbb{P}(Z > x)^{2+\epsilon}} < \infty \quad (40)$$

holds for some  $\epsilon > 0$ . In case this is possible for some positive  $\epsilon > 0$ , we can proceed as follows: let  $\mathcal{E}_x(i) = A'_x Z'^{(2^i)} + B'_x$  and  $\mathcal{E}_x = A'_x Z' + B'_x$ . For  $\beta \in (0, 1)$ , using the Hölder's inequality we get that

$$\begin{aligned} \frac{\mathbb{E}^{Q_{a^*}^\gamma} [(L_T^\Delta(x, 2^i) - L_T(x))^2]}{P(Z > x)^2} &= \frac{\mathbb{E}^{Q_{a^*}^\gamma} \left[ \mathbb{1}_{\{\tau_\gamma(x) < \infty, \mathcal{E}_x(i) \leq x, \mathcal{E}_x > x\}} (M_{\tau_\gamma}^{-1}(x))^2 \right]}{\mathbb{P}(Z > x)^2} \\ &= \frac{\mathbb{E}^{Q_{a^*}^\gamma} \left[ \left( \mathbb{1}_{\{\tau_\gamma(x) < \infty, \mathcal{E}_x(i) \leq x, \mathcal{E}_x > x\}} M_{\tau_\gamma}^{-1}(x) \right)^\beta \left( \mathbb{1}_{\{T(x) < \infty\}} M_{\tau_\gamma}^{-1}(x) \right)^{2-\beta} \right]}{\mathbb{P}(Z > x)^2} \\ &\leq \frac{\mathbb{E}^{Q_{a^*}^\gamma} \left[ \mathbb{1}_{\{\tau_\gamma(x) < \infty, \mathcal{E}_x(i) \leq x, \mathcal{E}_x > x\}} M_{\tau_\gamma}^{-1}(x) \right]^\beta}{\mathbb{P}(Z > x)^\beta} \frac{\mathbb{E}^{Q_{a^*}^\gamma} \left[ \mathbb{1}_{\{T(x) < \infty\}} M_{\tau_\gamma}^{-1}(x)^{\frac{2-\beta}{1-\beta}} \right]^{1-\beta}}{\mathbb{P}(Z > x)^{2-\beta}} \\ &= \left[ \frac{\mathbb{E}^{Q_{a^*}^\gamma} \mathbb{1}_{\{\tau_\gamma(x) < \infty, \mathcal{E}_x(i) \leq x, \mathcal{E}_x > x\}} M_{\tau_\gamma}^{-1}(x)}{\mathbb{P}(Z > x)} \right]^\beta \left[ \frac{\mathbb{E}^{Q_{a^*}^\gamma} L_T^{\frac{2-\beta}{1-\beta}}(x)}{\mathbb{P}(Z > x)^{\frac{2-\beta}{1-\beta}}} \right]^{1-\beta}. \quad (41) \end{aligned}$$

The first term in (41) is bounded by  $(\kappa 2^{-i(q-1)})^\beta$  due to (37), and the second term is uniformly bounded w.r.t.  $x$ . Therefore, the second moment relative to  $P(Z > x)^2$  can be bounded uniformly

w.r.t.  $x$ , if we choose  $\beta$  and  $N$  in a suitable way. For example, setting  $\beta \triangleq \frac{1+2\epsilon}{q-1}$  where  $\epsilon > 0$  is small enough to ensure that  $\beta < 1$ , and then  $N$  such that  $\mathbb{P}(N \geq i) = 2^{-i(1+\epsilon)}$ , one can achieve the purpose. Finally, we are left with verifying (40), which turns out to be possible by adapting the idea as described in Blanchet and Glynn (2008). Using the same argument as in the proof of Theorem 6, it is sufficient to analyze the term that is given by

$$\frac{\mathbb{E}^{Q_{a_*}^\gamma} L_{\tau_\gamma}^{2+\epsilon}(x)}{\mathbb{P}(\tau_\gamma(x) < \infty)^{2+\epsilon}}. \quad (42)$$

The following Lemma, which can be considered as an extension of Theorem 2, proves that the estimator  $L_T^{RG}(x)$  defined in (39) is an unbiased estimator of  $\mathbb{P}(Z > x)$  and is strongly efficient.

**Lemma 10.** *Suppose that  $\mathbb{E}X_1 < 0$  and  $X_1$  belongs to  $S^*$ . Fix  $\gamma \in (0, -\mathbb{E}X_1)$ . Let  $v_\gamma$  and  $w_\gamma$  be defined as in (25) and (26). Let  $\epsilon > 0$ . For fixed  $\delta \in (0, 1)$ , one can choose  $a_* = a_*(\delta) \leq 0$  so that*

$$-\delta \leq \frac{v_\gamma^{2+\epsilon}(y) - w_\gamma^{2+\epsilon}(y)}{\mathbb{P}(X_1 > -y)w_\gamma^{1+\epsilon}(y)}, \quad \forall y \leq s(x) + a_*.$$

Let

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

Then  $L_{\tau_\gamma}(x)$  is an unbiased estimator of  $\mathbb{P}(\max_{n \in \mathbb{N}} S_n > s(x))$  under  $\mathbb{P}^{Q_{a_*}}$ , and

$$\sup_{x > 0} \frac{\mathbb{E}^{Q_{a_*}} L_{\tau_\gamma}^{2+\epsilon}(x)}{\mathbb{P}\left(\max_{n \geq 0} S_n > s(x)\right)^{2+\epsilon}} < \infty,$$

where  $\mathbb{P}^{Q_{a_*}}$  denotes the probability measure associated with the random walk  $\{S_n\}_{n \in \mathbb{N}}$  having the one-step transition kernel

$$Q_{a_*}(y, dz) = P(y, dz) \frac{v_\gamma(z + a_*)}{w_\gamma(y + a_*)},$$

and  $\mathbb{E}^{Q_{a_*}}$  denotes the corresponding expectation operator.

*Proof.* The proof including the existence of  $a_*$  can be found in Appendix B.  $\square$

Note that the above discussion can easily be extended to cover the general stochastic recursion. We conclude this section with the following theorem.

**Theorem 11.** *Suppose that  $\mathbb{E}S_1(\gamma) < 0$ ,  $S_1(\gamma)$  belongs to  $S^*$  and Assumption 2 holds. Let  $v_\gamma$  and  $w_\gamma$  be defined as in (25) and (26). For fixed  $\delta \in (0, 1)$  and  $\beta \in (0, 1)$ , set  $a_* = a_*(\delta) \leq 0$  satisfying*

$$-\delta \leq \frac{v_\gamma^{\frac{2-\beta}{1-\beta}}(y) - w_\gamma^{\frac{2-\beta}{1-\beta}}(y)}{\mathbb{P}(X_1 > -y)w_\gamma^{\frac{1}{1-\beta}}(y)}, \quad \forall y \leq s(x) + a_*.$$

Moreover, assume that  $\mathbb{E}|\log A_1|^q + \mathbb{E}|\log \bar{B}_1|^q < \infty$ . Then, it is possible to choose  $N$  independently of  $x$ , such that

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

converges, and hence, the estimator  $L_T^{RG}(x)$  defined in (39) is unbiased and strongly efficient.

*Remark 3.* Note that, in general, the bias elimination scheme of Rhee and Glynn (2015) is not guaranteed to produce non-negative 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  $L_T^{RG}(x)$  is always non-negative.

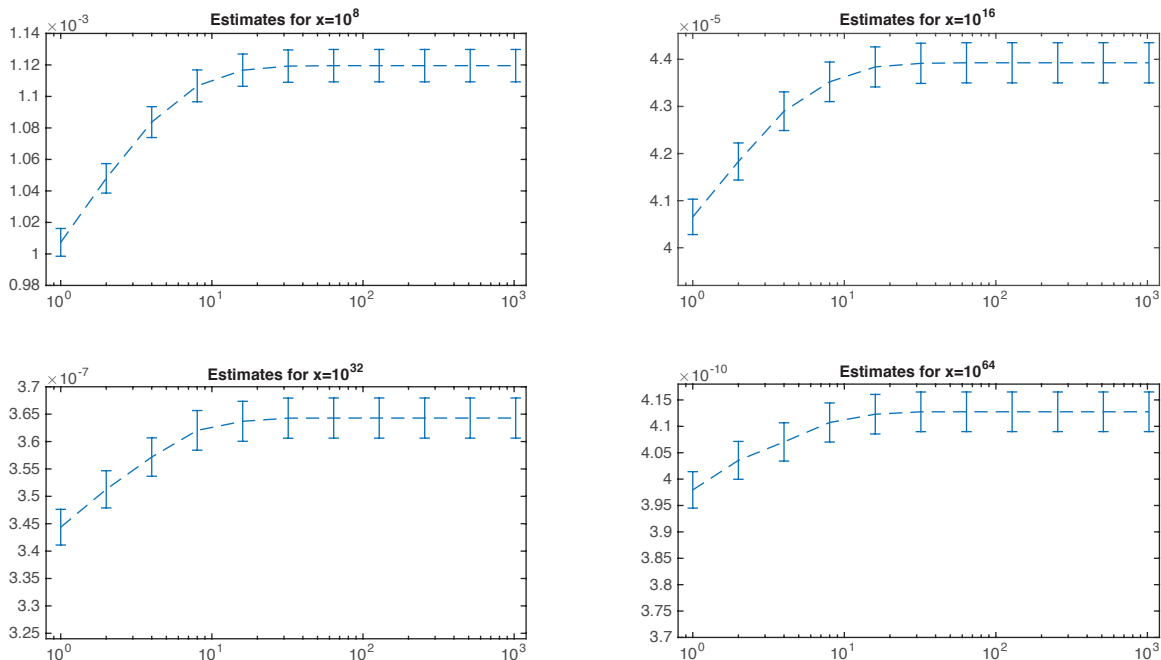


Figure 1: Estimated probabilities for changing values of  $M$ . The  $y$ -axis values indicate the estimated rare-event probabilities and the vertical bars indicate the 95% confidence intervals. The  $x$ -axis values indicate the truncation index  $M$ . In each of the subplot, we can see that as  $M$  increases, the estimated probability converges to a fixed value, which suggests that our estimator is consistent with respect to  $M$ . Comparing the four subplots for different value of  $x$ , one can see that the relative bias for small  $M$  decreases as  $x$  grows.

## 7 Numerical Results

Here we investigate our algorithm numerically based on a stochastic perpetuity with  $B_n = 1$ . We consider the increment  $X_n \stackrel{d}{=} \mathcal{W} - 3/2$  where  $\mathcal{W}$  is a random variable with Weibull distribution:

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

For the algorithmic parameters, we chose  $a_* = -10$ ,  $\gamma = 0.5$ . Moreover, we use a geometric distributed random truncation index with parameter 0.5. Figure 1 shows the change of estimated probability with respect to the different choice of  $M$  for 4 different values of  $x = 10^8$ ,  $x = 10^{16}$ ,  $x = 10^{32}$ , and  $x = 10^{64}$  in each of the four plots. One can see that the estimated probability stabilizes as  $M$  grows, which suggests that our estimator is consistent as  $M \rightarrow \infty$ . Comparing the four plots, one can also tell that the initial bias for small  $M$  decrease as  $x$  increases, which is consistent with the conclusion of Theorem 9 (vanishing relative bias). Table 1 reports the estimated probabilities, their 95%-confidence intervals and the estimated coefficients of variation, that is, the estimated standard deviation divided by the sample mean (based on 200000 samples), for different values of  $x$  and  $M$ . In the last column, we present the results produced with the unbiased algorithm as introduced in Section 6. 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 fix truncation method tend to converge to the estimated probability produced with the unbiased algorithm as  $M$  grows. These observations are consistent with the strong efficiency—predicted in Theorem 6 and Theorem 11—of our estimators.

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

Table 1: Estimated rare-event probability and 95% confidence intervals. One can see that, on the one hand the ratio between the standard deviation and the estimated probability stays roughly constant for different combinations of  $x$  and  $M$ , on the other hand, as  $M$  grows, the estimated probability produced with the fix truncation method tends to converge to the estimated probability produced with the unbiased algorithm, which suggests the consistency and the strong efficiency of our estimators predicted by Theorem 6 and Theorem 11.

## Appendix

### A Proof of Proposition 7

First we claim that, conditional on  $\{\tau_\gamma(x) < \infty\}$ , the first term in  $\xi_x$  converges to 0 in probability. Let  $\epsilon > 0$  be arbitrary, we have that

$$\mathbb{P}\left(\frac{\log\left(1 - \frac{B'_x}{x}\right)}{a(\log x)} \leq -\epsilon \mid \tau_\gamma(x) < \infty\right) = \mathbb{P}\left(\frac{B'_x}{x} \geq 1 - x^{-\frac{\epsilon}{\alpha}} \mid \tau_\gamma(x) < \infty\right).$$

Let  $M_{\tau_\gamma(x)}$  denote the maximum of  $\{S_i(\gamma)\}_{i \leq \tau_\gamma(x)-1}$ . It is well known that  $M_{\tau_\gamma(x)} = \mathcal{O}(1)$  (cf. the proof of Theorem 1.1 in Asmussen & Klüppelberg, 1996). Moreover,  $B'_x$  is bounded by

$$B'_x = \sum_{n=0}^{\tau_\gamma(x)-1} e^{S_n} = \sum_{n=0}^{\tau_\gamma(x)-1} e^{S_n(\gamma) - n\gamma} \leq e^{M_{\tau_\gamma(x)}} \sum_{n=0}^{\tau_\gamma(x)-1} e^{-n\gamma} \leq e^{M_{\tau_\gamma(x)}} (1 - e^{-\gamma})^{-1}.$$

Therefore, for  $\alpha > 0$ , we have that

$$\begin{aligned} \mathbb{P}\left(\frac{B'_x}{x} \geq 1 - x^{-\frac{\epsilon}{\alpha}} \mid \tau_\gamma(x) < \infty\right) &\leq \mathbb{P}\left(M_{\tau_\gamma(x)} - \log x - \log(1 - e^{-\gamma}) \geq \log(1 - x^{-\frac{\epsilon}{\alpha}}) \mid \tau_\gamma(x) < \infty\right) \\ &= \mathbb{P}\left(\frac{M_{\tau_\gamma(x)}}{a(\log x)} - \frac{\log x}{a(\log x)} - \frac{\log(1 - e^{-\gamma})}{a(\log x)} \geq \frac{\log(1 - x^{-\frac{\epsilon}{\alpha}})}{a(\log x)} \mid \tau_\gamma(x) < \infty\right) \\ &\rightarrow 0, \end{aligned} \tag{43}$$

as  $x \rightarrow \infty$ , since  $\log x/a(\log x) = \alpha$ . The convergence of the last two terms in  $\xi_x$  is an application of Asmussen and Klüppelberg (1996, Theorem 1.1).

## B Proof of Lemma 10

First notice that the numerator in (42) is equal to  $\mathbb{E}L_{\tau_\gamma}^{1+\epsilon}(x)$ . Analogous to Blanchet and Glynn (2008, Theorem 2), we can derive a similar result for  $\mathbb{E}L_{\tau_\gamma}^{1+\epsilon}(x)$  simply by replacing  $r$  with  $r^{1+\epsilon}$ , where

$$r(y, z) = \frac{v(z)}{w(y)},$$

and  $v, w$  are defined in the corresponding way as in (7). Based on this observation we claim the following Proposition (cf. Blanchet & Glynn, 2008, Proposition 2):

**Proposition 12.** *Let  $P(y, dz)$  denote the transition kernel of the random walk  $\{S_n\}_{n \in \mathbb{N}}$ . Assume that  $w(y) > 0$  for  $y \leq s(x)$  and that there exists a finite-valued function  $h : \mathbb{R} \rightarrow [\delta_1, \infty)$  satisfying*

$$w^{1+\epsilon}(y) \int v(z)h(z)P(y, dz) \leq h(y)v^{2+\epsilon}(y), \quad (44)$$

for  $y \leq s(x)$ . If  $h(z) \geq 1$  for  $z > s(x)$  and  $v(z) \geq \delta_2 > 0$  for  $z > s(x)$ , then we have that

$$\mathbb{E}L_{\tau_\gamma}^{1+\epsilon}(x) \leq \delta_1^{-1} \delta_2^{-(2+\epsilon)} v^{2+\epsilon}(y) h(y).$$

Recall that  $v_\gamma$  defined as in (25) depends implicitly on  $x$ . The Pakes-Veraverbeke's Theorem implies that  $\mathbb{P}(\tau_\gamma(x) < \infty) \sim v_\gamma(y)$  for every fixed  $y$  as  $x \rightarrow \infty$ . This observation gives us a way to prove Lemma 10: first find a suitable function  $h$  such that (44) holds with  $v$  and  $w$  being defined as in (21) and (22), then the result can be yielded using Proposition 12. Define

$$h(y) = \mathbb{1}_{(-\infty, s(x)-a_*]}(y) + (1 - \delta) \mathbb{1}_{(s(x)-a_*, \infty)}(y).$$

Now (44) is equivalent to (cf. the proof of Theorem 3 in Blanchet & Glynn, 2008)

$$-\delta \leq \frac{v_\gamma^{2+\epsilon}(y + a_*) - w_\gamma^{2+\epsilon}(y + a_*)}{\mathbb{P}(X_1 > -y - a_*) w_\gamma^{1+\epsilon}(y + a_*)}, \quad \forall y \leq s(x).$$

Using the definition of  $w_\gamma$  and noticing the non-negativity of  $W_\gamma$ , (11) implies particularly that  $w(y) - v(y) = o(w(y))$ , as  $y \rightarrow -\infty$ . Therefore, such  $a_*$  can be found.

## References

- Asmussen, S., & Glynn, P. (2007). *Stochastic simulation: Algorithms and analysis* (1st ed., Vol. 57). Springer-Verlag New York.
- Asmussen, S., & Klüppelberg, C. (1996). Large deviations results for subexponential tails, with applications to insurance risk. *Stochastic Processes and their Applications*, 64(1), 103-125.
- Asmussen, S., Schmidli, H., & Schmidt, V. (1999, June). Tail probabilities for non-standard risk and queueing processes with subexponential jumps. *Advances in Applied Probability*, 31(2), 422-447.
- Basrak, B., Davis, R., A., & Mikosch, T. (2002). Regular variation of GARCH processes. *Stochastic Processes and their Applications*, 99(1), 95-115.
- Blanchet, J., & Glynn, P. (2008, August). Efficient rare-event simulation for the maximum of heavy-tailed random walks. *The Annals of Applied Probability*, 18(4), 1351-1378.
- Blanchet, J., Lam, H., & Zwart, B. (2012). Efficient rare-event simulation for perpetuities. *Stochastic Processes and their Applications*, 122, 3361 - 3392.
- Blanchet, J., & Sigman, K. (2011, August). On exact sampling of stochastic perpetuities. *Journal of Applied Probability*, 48A, 165-182.
- Blanchet, J., & Wallwater, A. (2015, November). Exact sampling of stationary and time-reversed queues. *ACM Transactions on Modeling and Computer Simulation*, 25(4), 26:1-26:27.
- Buraczewski, D., Damek, E., & Mikosch, T. (2016). *Stochastic models with power-law tails: The equation  $x=ax+b$*  (1st ed.). Springer International Publishing.

- Douc, R., Moulines, E., & Soulier, P. (2007, August). Computable convergence rates for sub-geometric ergodic Markov chains. *Bernoulli*, 13(3), 831–848.
- Dyszewski, P. (2016). Iterated random functions and slowly varying tails. *Stochastic Processes and their Applications*, 126(2), 392–413.
- Embrechts, P., Klüppelberg, C., & Mikosch, T. (1997). *Modelling extremal events for insurance and finance* (1st ed., Vol. 33). Springer-Verlag Berlin Heidelberg.
- Glynn, P. (2012). Lecture notes for MS&E Stochastic Calculus and Control. *Stanford University*.
- Goldie, M., Charles. (1991). Implicit renewal theory and tails of solutions of random equations. *The Annals of Applied Probability*, 1(1), 126-166.
- Jarner, S. F., & Roberts, G. O. (2002, 02). Polynomial convergence rates of Markov chains. *The Annals of Applied Probability*, 12(1), 224–247.
- Kalashnikov, V., & Tsitsiashvili, G. (1999). Tails of waiting times and their bounds. *Queueing Systems*, 32(1), 257–283.
- Klüppelberg, C. (1988, March). Subexponential distributions and integrated tails. *Journal of Applied Probability*, 25(1), 132-141.
- Maulik, K., & Zwart, B. (2006). Tail asymptotics for exponential functionals of Lévy processes. *Stochastic Processes and their Applications*, 116, 156-177.
- Palmowski, Z., & Zwart, B. (2007). Tail asymptotics of the supremum of a regenerative process. *Journal of Applied Probability*, 44(2), 349-365.
- Rhee, C.-H., & Glynn, P. (2015). Unbiased estimation with square root convergence for SDE models. *Articles in Advance*, 1-18.
- Veraverbeke, N. (1977). Asymptotic behaviour of Wiener-Hopf factors of a random walk. *Stochastic Processes and their Applications*, 5(1), 27–37.
- Vervaat, W. (1979). On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. *Advances in Applied Probability*, 11(4), 750-783.
- Zachary, S. (2004, January). A note on Veraverbeke’s theorem. *Queueing Systems*, 46(1-2), 9-14.