RARE EVENT SIMULATION FOR STOCHASTIC FIXED POINT EQUATIONS RELATED TO THE SMOOTHING TRANSFORMATION

Jeffrey F. Collamore

Department of Mathematical Sciences
University of Copenhagen
Universitetsparken 5
DK-2100 Copenhagen, Denmark

Anand N. Vidyashankar

Department of Statistics
Volgeneau School of Engineering
Georgia Mason University
Fairfax, VA 22030, U.S.A.

Jie Xu

Systems Engineering & Operations Research Department
Volgeneau School of Engineering
Georgia Mason University
Fairfax, VA 22030, U.S.A.

ABSTRACT

In several applications arising in computer science, cascade theory, and finance, it is of interest to evaluate $\mathbf{P}(V > u)$ for a random variable V satisfying a distributional fixed point equation of the form $V \stackrel{d}{=} \sum_{i=1}^{N} A_i V_i + B$, where $V_1, V_2, \cdots$ are i.i.d. copies of V, independent of $\{A_i\}$ and of B, and N is an integer-valued random variable. Recently, methods have been developed for the linear recursion $V \stackrel{d}{=} AV + B$ and some of its extensions, yielding tail estimates and importance sampling methods for these recursions. However, such methods do not routinely generalize to the non-homogeneous recursions described in the above fixed point equation. Drawing on the techniques from the weighted branching process literature, we describe a modification of the importance sampling algorithm of Collamore et al. (2011) which yields an estimate for the tail probability $\mathbf{P}(V > u)$ which is consistent and also strongly efficient, exhibiting bounded relative error.

1 INTRODUCTION

This paper is concerned with rare event simulation related to the non-homogeneous stochastic fixed point equations of the form

$$V \stackrel{d}{=} \sum_{i=1}^{N} A_i V_i + B, \tag{1.1}$$

where $\mathscr{V} \equiv \{V, V_i : i \geq 1\}$ is a collection of independent and identically distributed (i.i.d.) random variables; $\mathscr{A} \equiv \{A_i : i \geq 1\}$ is a collection of non-negative random variables and B a real valued random variable, both independent of $\mathscr{V}$; and N is an integer-valued random variable, independent of $\mathscr{V}$, $\mathscr{A}$, and B. When $B = 0$, (1.1) is referred to as a homogeneous stochastic fixed point equation (SFPE). These SFPEs arise in a variety of examples; for instance: (i) the Quicksort algorithm, where V represents the stationary solution

to the normalized key comparisons needed to sort a random permutation of length n; (ii) the Hausdorff dimension of Cantor sets; (iii) the stochastic approximation of Google's page rank algorithm; and (iv) the study of martingale limits of Mandelbrot's cascades and of branching random walk.

The solution to (1.1) with $N = \infty$ has received much attention, starting with the work of Durrett and Liggett (1983) and culminating in the work of Alsmeyer et al. (2012). We note that (1.1) describes a fixed point of a mapping $\mathfrak{T}$—referred to as the smoothing transform—on the space of probability measures. In this paper, we use a weighted branching process to describe the solution to (1.1) and focus on estimating $\mathbf{P}(V > u)$ for large values of u. It is well-known that under suitable moment and regularity conditions,

$$\lim_{u \to \infty} u^{\xi} \mathbf{P}(V > u) = C, \tag{1.2}$$

where C is a finite positive constant and $\xi > 0$ is the solution to the equation $\mathbf{E}\left[\sum_{i=1}^{N} A_i^{\xi}\right] = 1$ (see Liu (2000), Iksanov and NegadaÄlov (2007), Jelenković and Olvera-Cravioto (2012)). The constant C is typically defined in terms of the SFPE (1.1) and hence is not very useful from a computational perspective. Recently, in a related problem, Collamore and Vidyashankar (2013b) developed a general method to provide a probabilistic characterization of the constant C in the case where $V \overset{d}{=} AV + B$ and for certain extensions of the form

$$V \overset{d}{=} f(V), \tag{1.3}$$

where $f(\cdot)$ is a random Lipschitz function satisfying appropriate regularity conditions. It is natural to ask whether such an extension is feasible for the SFPE in (1.1). A key ingredient to the characterization in Collamore and Vidyashankar (2013b) involves representing the SFPE as a stationary limit of the forward iterates of a corresponding Markov chain. However, in the context of (1.1), a representation of this type is not available and, for this reason, we have developed an alternative approach (Collamore and Vidyashankar (2013a)), studying an associated *linear* SFPE and representing the tails of V in terms of the stationary solution to this associated SFPE. This approach yields a probabilistic characterization of C in (1.2).

While (1.2) provides an asymptotic description of the rare event probability $\mathbf{P}(V > u)$ as $u \to \infty$, it is also of interest to obtain exact values for $\mathbf{P}(V > u)$ for fixed values of u. When u is large, it is well-known that direct Monte Carlo (MC) methods are not useful, since

$$\limsup_{u \to \infty} \frac{\mathrm{Var}(\mathscr{E}_u)}{\mathbf{P}(V > u)^2} = \infty, \tag{1.4}$$

where $\mathscr{E}_u$ (defined below) represents an unbiased estimator of $\mathbf{P}(V > u)$. In various contexts, alternative methods based on importance sampling—utilizing an exponential change of measure—successfully address the inadequacy of the direct MC approach. Specifically, in the context of (1.3), Collamore et al. (2011) have developed a dynamic importance sampling algorithm, involving a dual change of measure, to compute the tail probabilities. However, this technique does not routinely carry over to more general recursions, such as those described in (1.1).

In this article, we utilize the relationship between the SFPE and a weighted branching process to derive an associated linear SFPE, $V^* \overset{d}{=} A^* V^* + B^*$, for a specific choice of $A^\star$ and $B^\star$ (dependent on $\{A_i\}$, N, B and $\{V_i\}$). We then represent $\mathbf{P}(V > u)$ as a functional of $\mathbf{P}(V^\star > u)$, and adopt the algorithm of Collamore et al. (2011) to estimate $\mathbf{P}(V > u)$ in (1.2) after approximating the associated SFPE with $V^\star \overset{d}{=} A^\star V^\star + B_n^\star$, where $B_n^\star$ converges to B with probability one. Implementation of this algorithm, however, leads to several subtle issues concerning the simulation of size-biased distributions, exponentially shifted size-biased distributions, convergence of the solutions of an approximation to the associated SFPE, and some issues concerning nested simulation.

The rest of the article is structured as follows: Section 2 contains a description of the weighted branching processes and branching random walk, while Section 3 is devoted to a description of an associated linear

SFPE. In Section 4, we present the importance sampling algorithm and some implementation details and numerical results.

2 WEIGHTED BRANCHING PROCESSES AND BRANCHING RANDOM WALK

We begin with a brief description of the Galton-Watson (GW) process. (For further details, see Athreya and Ney (1972).) The process starts with a single ancestor at time 0. This ancestor lives one unit of time and reproduces according to a probability distribution $\{p_j : j \geq 0\}$. Let Z_1 denote the size of the first generation. Then

$$\mathbf{P}(Z_1 = j) = p_j, \quad j = 0, 1, 2, \cdots, \tag{2.1}$$

referred to as the offspring distribution. Each individual of the first generation, in turn, lives one unit of time and reproduces, independent of one another and of the ancestors, according to the same offspring distribution. Let Z_n denote the propulation size of the nth generation, and let $m = \mathbf{E}[Z_1]$. It is well-known (Athreya and Ney (1972)) that the process Z_n either diverges or it converges to zero. Let $\mathscr{S} = \{Z_n \to \infty\}$. Then $\mathscr{S}$ is called the survival set of the process, and $\mathbf{P}(\mathscr{S}) > 0$ if and only if $m > 1$. In this paper, we will assume that $p_0 = 0$, implying that the process survives with probability one. Now, the weighted process is the same as the GW process except that each individual carries a random weight, obtained as a product of the weight of the parent and a random factor introduced at the edge adjoining the parent and its offspring. The weight of the initial ancestor is taken to be one.

To provide a precise description of the weighted branching processes, let ϕ denote the initial ancestor and denote the members of the kth generation, $k \geq 1$, by $(v_1, v_2, \cdots, v_k)$. These elements are then members of the Ulam-Harris tree $\mathbb{T}$, defined as follows. Let $\mathbb{N}$ denote the set of natural numbers with discrete topology, and let

$$\mathbb{T} = \bigcup_{k=0}^{\infty} \mathbb{N}^k, \quad \text{where} \quad \mathbb{N}^0 := \{\phi\}. \tag{2.2}$$

For $v \in \mathbb{T}$, let $|v|$ denote the length (*i.e.*, generation) of v.

Let $X = (X_1, X_2, \cdots)$ be an infinite vector of positive random variables whose components have expectation one, possessing an arbitrary dependence structure. We now associate with every vertex $v \in \mathbb{T}$ an i.i.d. copy of X and denote it by $X(v)$. These vectors represent the random factors described in the previous paragraph. In particular, the vector $(X_1(\phi), X_2(\phi), \cdots)$ represents the weight factor associated with the root vertex, while the weight factor associated with offspring 1 of this ancestor (corresponding to $v = 1$) is $X(1) = (X_1(1), X_2(1), \cdots)$. We stress here that $X(1)$ is an i.i.d. copy of $X(\phi)$. We now define the weights of the vertices $v \in \mathbb{T}$ as follows. First set $L(\phi) = 1$. Then for $v \in \mathbb{N}^n \subset \mathbb{T}$, define

$$L(v) = X_{v_1}(\phi) \prod_{i=1}^{n-1} X_{v_{i+1}}(v_1, v_2, \cdots, v_i). \tag{2.3}$$

Let

$$Z_n = \sum_{|v|=n} L(v) \quad \text{and} \quad W_n = \frac{Z_n}{m^n}, \tag{2.4}$$

where we recall that $m := \mathbf{E}[Z_1]$. Then Z_n can be interpreted as the total weight of the nth generation, while W_n is the normalized total weight. It is easy to see that $\mathbf{E}[W_n] = 1$, as in the GW case, since we assume the weight factors satisfy $\mathbf{E}[L(v)] = 1$. Thus, $\{W_n\}$ is positive martingale sequence converging to random variable W_∞ almost surely. Under the assumption that $\mathbf{E}[W_1 \log W_1] < \infty$, Rösler et al. (2000) establish the L_1 convergence of W_n to W_∞. This yields $\mathbf{E}[W_\infty] = 1$.

2.1 FORWARD AND BACKWARD RECURSIONS

To relate a weighted branching process to an SFPE, it is convenient to consider the forward and backward recursions associated with a weighted branching process. To this end, let N be a positive integer-valued random variable with $\mathbf{E}[N] > 1$. Let $\{N_v : v \in \mathbb{T}\}$ denote a collection of i.i.d. random variables having the same distribution as N. In particular, N_ϕ represents the number of direct descendants (offspring) of the initial ancestor ϕ. To obtain alternate expressions for Z_n, note that by conditioning on the first generation,

$$Z_n = \sum_{i=1}^{N_\phi} X_i(\phi) Z_{n-1,i}, \tag{2.5}$$

where $\{Z_{n-1,i} : 1 \leq i \leq N_\phi\}$ are i.i.d. copies of Z_{n-1}. This is called the backward recursion. Dividing (2.5) by m^n and letting $n \to \infty$, it follows that

$$W_\infty = \sum_{i=1}^{N_\phi} \frac{X_i(\phi)}{m} W_{\infty,i}, \tag{2.6}$$

where $W_{\infty,i}$ are i.i.d. copies of W_∞ and independent of $\mathscr{F}_1 = \sigma(X_1(\phi), X_2(\phi), \cdots)$. This yields the homogeneous SFPE obtained in (1.1) when $B = 0$. Alternatively, conditioning on the $(n-1)^{\text{th}}$ generation, it can be seen that

$$Z_n = \sum_{|v|=(n-1)} L(v) \sum_{i=1}^{N_v} X_i(v). \tag{2.7}$$

This is referred to as the forward recursion. We point out here that, while the backward recursion yields the SFPE, the forward recursion is helpful in designing the algorithm.

2.2 BRANCHING RANDOM WALK

A closely related model to the weighted branching process is the branching random walk (BRW) model studied in Biggins (1977), Biggins (1992), Kingman (1975), and several recent papers. In this model, we start with a single ancestor located at the origin at time 0. The offspring of this ancestor form the first generation, and their positions are described by a point process $\mathbb{Z}^1$ on $\mathbb{R}$. The individuals in the nth generation reproduce independent of each other and of the preceding generations to form the $(n+1)^{\text{th}}$ generation. The displacements of offspring of a parent have the same distribution as $\mathbb{Z}^1$. The Laplace transform of the point process describing a BRW can be seen to be closely related to a weighted branching process. To describe this correspondence, let $Y_{n,r}$ denote the positions of the nth generation population, and define

$$W_n(\theta) = (m(\theta))^{-n} \sum_r \exp(-\theta Y_{n,r}), \tag{2.8}$$

where $m(\theta) = \mathbf{E}\left[\sum_r \exp(-\theta Y_{1,r})\right]$. Alternatively,

$$W_n(\theta) = (m(\theta))^{-n} \sum_{|v|=n} L(v), \tag{2.9}$$

where $L(v)$ is as in (2.3) and $X_i(v) = \exp(-\theta Y_{1,i}) I_{\{\mathbb{Z}^{(1)}(\mathbb{R}) \geq i\}}, i \geq 1$. Thus, it is clear that the BRW is closely related to the weighted branching process. However, in a general weighted branching process, it is sometimes allowed for $v \in \mathbb{T}$ that $N_v = \infty$ with positive probability, while it is typically assumed that $N_v < \infty$ w.p.1 in the case of a branching random walk.

It is well-known that, in the case of a branching random walk, $\{W_n(\theta)\}$ is a non-negative martingale and hence converges almost surely to a random variable $W(\theta)$. Under the assumption that $\mathbf{E}[W_1(\theta) \log^+ W_1(\theta)] < \infty$ and the additional assumption that $\theta \in \mathfrak{D}$, where

$$\mathfrak{D} = \{\theta \in \mathbb{R} : m(\theta) < \infty\} \tag{2.10}$$

is a convex subset of $\mathbb{R}$, Biggins (1977, 1992) established the L_1 convergence of $\{W_n(\theta)\}$ to $W_\infty(\theta)$, which implies $\mathbf{E}[(W(\theta)] = 1$. Similar to the case of a weighted branching process, one can use the branching property to verify that $W_\infty(\theta)$ satisfies the SFPE

$$W_\infty(\theta) = m(\theta)^{-1} \sum_{i=1}^{N_\phi} X_i(\phi) W_{\infty,i}(\theta), \tag{2.11}$$

where $\{W_{\infty,i}(\theta) : 1 \leq i \leq N_\phi\}$ consists of i.i.d. copies of W_∞. In the remainder of this section, by a slight abuse of notation, we write $X_i(\phi)$ also to denote $X_i(\phi)/m(\theta)$. Then (2.11) becomes

$$W_\infty(\theta) = \sum_{i=1}^{N_\phi} X_i(\phi) W_{\infty,i}(\theta), \tag{2.12}$$

which is the same as the SFPE (1.1) with $B = 0$; that is, the homogeneous SFPE. A slightly modified construction, incorporating an immigration term into each of the terms $X_i(\phi)$ in (2.6), yields the SFPE (1.1) with $B \neq 0$; that is, the non-homogeneous SFPE. We observe that in the above discussion, one can obtain the BRW and weighted branching process from one another; the purpose of including both of these descriptions is to emphasize that they are the same, although they have been used in different contexts historically.

3   THE ASSOCIATED SFPE

While the SFPE described in (2.11) is complex, one can develop an alternative linear SFPE by considering the expectations of the martingale limits described in Sections 2.1 and 2.2. Recalling that $\mathbf{E}[W_\infty] = 1$, it is known (Liu (2000), Iksanov and NegadaÄlov (2007)) that if $V^\star$ is a random variable with distribution $Q(\cdot)$, where

$$Q(\cdot) \overset{\text{def}}{=} \mathbf{E}\left[W_\infty I_{\{W_\infty \in \cdot\}}\right], \tag{3.1}$$

then V^* satisfies the linear SFPE

$$V^* \overset{d}{=} A^* V^* + B^*, \tag{3.2}$$

where the distributions of A^* and B^* are determined by the equation

$$\mathbf{E}[g(A^*, B^*)] := \mathbf{E}\left[\sum_{i=1}^{N_\phi} X_i(\phi) g\left(X_i(\phi), \sum_{j \neq i}^{N_\phi} X_j(\phi) Z_{\infty,j}\right)\right], \tag{3.3}$$

for all bounded Borel functions $g(\cdot, \cdot)$ of two variables. In particular, taking $g(x,y) = f(x)$, it follows that

$$\mathbf{E}[f(A^*)] := \mathbf{E}\left[\sum_{i=1}^{N_\phi} X_i(\phi) f(X_i(\phi))\right]. \tag{3.4}$$

Then the tail probabilities of V can be obtained from those of $V^\star$, since, if $F(\cdot)$ represents the distribution of W_∞, then

$$\begin{aligned}
\mathbf{P}(V > u) &= \int_u^\infty dF(v) \\
&= \int_u^\infty \frac{1}{v} dQ(v) \\
&= \mathbf{E}\left[(V^\star)^{-1} I_{\{V^\star > u\}}\right]. \tag{3.5}
\end{aligned}$$

One can now modify the algorithm in Collamore et al. (2011) to derive an estimator for $\mathbf{P}(V > u)$. To this end, we consider the forward recursion

$$V_{n+1}^* = A_{n+1}^* V_n^* + B_{n+1}^*, \tag{3.6}$$

where $\{(A_n^*, B_n^*)\}$ is a collection of i.i.d. random variables with distribution described in (3.3). Then, using (3.4) together with the representation for ξ in (1.2), one can show that there exists $\alpha > 0$ such that

$$\mathbf{E}[(A^*)^\alpha] = 1, \tag{3.7}$$

where necessarily $\alpha = \xi - 1$. Now, to apply the algorithm of Collamore et al. (2011), we notice that V_n^* is a positive recurrent Markov chain with stationary distribution $\pi(\cdot)$ and C-set $\mathscr{C} = [0, M]$ for some $M \geq 0$. Then, as in Collamore et al. (2011), we perform a dual change of measure; namely, simulate from the $\alpha-$shifted distribution of (A^*, B^*) until V_n^* exceeds u, then revert back to the original distribution thereafter. We stop when the process returns to $\mathscr{C}$. Let T_u denote the time it takes the process to first reach the level u, that is,

$$T_u = \inf\{j \geq 1 : V_j^* > u\}, \tag{3.8}$$

and let K denote the first return time to $\mathscr{C}$; that is, $K = \inf\{j \geq 1 : V_j^* \in \mathscr{C}\}$. Then our method can be succinctly written as an algorithm, as stated in the next section.

4   THE ALGORITHM

The following is the rare event simulation algorithm that can be used for calculating the tail probabilities of the SFPE (1.1) with $B = 0$. In the description below, we denote by μ the distribution of $(\log A^*, B^*)$, and set

$$d\mu_\alpha(x, y) = \frac{e^{\alpha x}}{\lambda(\alpha)} d\mu(x, y), \quad \alpha \in \mathbb{R},$$

where $\lambda(\alpha) := \mathbf{E}[(A^*)^\alpha]$. Set $S_n = \sum_{k=1}^n \log A_k^*$. Moreover, let γ denote the stationary distribution of $\{V_n^*\}$ restricted to $\mathscr{C}$; namely, $\pi(E)/\pi(\mathscr{C})$, for any Borel set $E \subset \mathscr{C}$. Then, as $\mathscr{C}$ is near the center of the distribution of V_n^*, this stationary distribution may easily be approximated by a direct MC procedure (as described in more detail in Collamore et al. (2011)), leading to an approximation $\hat{\gamma}_k$ for γ.

Let $\{\mathscr{E}_u^{(j)} : 1 \leq j \leq N\}$ denote N samples obtained from the above algorithm, and let $\hat{\mathscr{E}}_u \equiv \frac{1}{N} \sum_{j=1}^N \mathscr{E}_u^{(j)}$. Then, following Collamore et al. (2011), estimate $\mathbf{P}(V > u)$ by setting it to be equal to $\hat{\mathscr{E}}_u \hat{\pi}(\mathscr{C})$, where $\hat{\pi}(\mathscr{C})$ is an estimate of the stationary distribution of $\mathscr{C}$. Using techniques from Collamore et al. (2011), one can verify that this estimator is consistent and exhibits bounded relative error (under the moment and regularity conditions stated there); in particular,

$$\sup_{u \geq 0} u^{2\xi} \mathbf{E}[\mathscr{E}_u^2] < \infty. \tag{4.1}$$

Additionally, under further regularity conditions, it can be shown that the running of the algorithm grows as a constant multiple of $\log u$ as $u \to \infty$. These results are analogues of Theorems 2.1, 2.2, and 2.3 of Collamore et al. (2011).

4.1 IMPLEMENTATION OF THE ALGORITHM

We implemented the above algorithm with $B = 0$ in (1.1). For the sake of concreteness, we simulated the SFPE using the recursions of the BRW model described in Section 2.2. The random variable N_ϕ was assumed to be distributed as $1 + \tilde{N}$, where $\tilde{N}$ is a Poisson random variable with parameter 0.25; hence $\mathbf{E}[N_\phi] = 1.25$. Conditioned on N_ϕ, we took the positions of the offspring, $\{Y_{1,r} : 1 \leq r \leq N_\phi\}$, to be i.i.d. Normally distributed with mean zero and variance $\sigma^2 = 0.14876$. We took $\theta = 1$ in (2.8). A simple

Algorithm 1 *

$V_0 \sim \hat{\gamma}_k, \ s = 0$

repeat

    $s \leftarrow s + 1$

    $V_s^* = A_s^* V_{s-1}^* + B_s^*, \quad (\log A_s^*, B_s^*) \sim \mu_\alpha$

until $V_s^* > u$ or $V_s^* \in \mathscr{C}$

if $V_s^* > u$ **then**

    repeat

        $s \leftarrow s + 1$

        $V_s^* = A_s^* V_{s-1}^* + B_s^*, \quad (\log A_s^*, B_s^*) \sim \mu$

    until $V_s \in \mathscr{C}$

    $\mathscr{E}_u = \sum_{r=0}^{K-1} \frac{1}{V_r^*} \mathbf{1}_{\{V_r^* > u\}} e^{-\alpha S_{T_u}}$

else

    $\mathscr{E}_u = 0$

end if

Rare Event Simulation Algorithm using Forward Iterations of the linear associated SFPE.

---

calculation then yields $\xi = 3$ and $\alpha = 2$. Hence, based on (1.2), it follows that $P(W_\infty > u) \sim C u^{-3}$ as $u \to \infty$.

To express the above quantities in terms of the weighted branching process, observe that $m(\theta) = \mathbf{E}\left[\sum_{i=1}^{N_\phi} e^{\theta Y_{1,i}}\right] = \mathbf{E}\left[N_\phi\right] \mathbf{E}\left[e^{\theta Y_{1,1}}\right] = 1.25 \times e^{0.14876/2} = 1.3465$ for $\theta = 1$. Thus $X_i(\phi) \sim LN(0, \sigma^2)/m(1) \sim LN(0, 0.14876)/1.3465$. To implement the algorithm, we first have to simulate from the associated SFPE, whose forward iterates are obtained by sampling from A^* and B^* in their shifted distributions. Since the expression for B^* involves the random variable W_∞, we approximate W_∞ by W_{n*} for some $n^* \in \mathbb{Z}_+$. In our implementation, we chose $n^* = 20$. This choice was determined by an exploratory analysis.

We first observe that, under the assumptions on the SFPE, it can be shown that the random variables A^* and B^* are independent. To simulate the i.i.d. copies of A^*, we notice that in Equation (3.4), taking $f(x) = I_{\{x \in \mathscr{F}\}}$ for some Borel set F and conditioning on N_ϕ, we have

$$P(A^* \in F) = \mathbf{E}\left[\sum_{i=1}^{N_\phi} X_i(\phi) I_{\{X_i(\phi) \in F\}}\right] = \mathbf{E}\left[N_\phi\right] \mathbf{E}\left[X_i(\phi) I_{\{X_i(\phi) \in F\}}\right].$$

Noticing that the second term represents a sized-biased distribution, it is easy to see (*cf.* Patil and Rao (1978)) that $A^* \sim LN(0.14876, 0.14876)/1.3465$. Using a similar calculation, one can show that the distribution of the α-shifted random variable A^* is given by $LN(0.44628, 0.14876)/1.3465$.

We now turn to the problem of simulating B^*. A standard calculation shows that for any Borel set G,

$$\mathbf{P}(B^* \in G) = \mathbf{E}\left[N^* I_{\left\{\sum_{j \neq 1} X_j(\phi) W_{\infty,j} \in G\right\}}\right]. \tag{4.2}$$

Thus, to simulate B^*, we adopt a variant of the acceptance-rejection algorithm. The following steps describe this algorithm:

1. Fix q, a "high" quantile of the distribution of N^*.
2. Sample $N^* \sim f(\cdot)$ and $U \sim \text{Unif}(0,1)$.
3. If $\frac{X}{q} > U$, return $L = \left\{\sum_{j \neq 1} X_j(\phi) W_{\infty,j}\right\}$. Else go to Step 2.

Table 1: Importance sampling estimation for the tail probability $\mathbf{P}(V > u)$.

u	$\mathbf{P}(V > u)$	RE	T_u	K	CMC
10	6.06e-3	1.28e3	13.8	54.1	3.89e-3
50	6.06e-5	4.21e2	23.8	75.1	NA
100	8.06e-6	6.07e2	28.3	84.3	NA
500	6.19e-8	4.84e2	39.0	106	NA
1000	8.23e-9	6.86e2	43.7	115	NA

We note that this choice of the high quantile of the distribution of X enables one to produce approximate samples from the distribution with density $xf(x)$. If the distribution of X has bounded support, one can replace q by the maximum possible value of this random variable.

The results of our implementation are summarized in Table 1. The variable CMC represents the crude Monte Carlo estimate. The importance sampling estimates are based on 10^7 simulations, while the CMC estimate is based on 10^8 simulations. For large u, CMC is not applicable (N/A). We notice from the table that the relative error remains bounded and the running time, T_u, and the return time, K, grow roughly logarithmically in u. Additionally, simple calculations illustrate that the constant C in (1.2) is approximately 8. Perhaps more importantly, the ratio of the probabilities for different values of u in Table 1 exhibit the same scaling as described in (1.2). This provides partial evidence that the algorithm described in this paper works in practice. A key difficulty in implementing the algorithm is the challenge of simulating from the size-biased distributions. This problem was partly resolved in our implementation, since the lognormal family is closed, in the sense that the size-biased distribution of a lognormal is also lognormal. Several other distributions belonging to the Pearson type III family of distributions also possess this property, and hence other choices for the distribution of the weight factor in a weighted branching process (or the positions of the offspring in a BRW) are possible.

5 CONCLUSION

In this article, we developed an algorithm for estimating the tails of the non-homogeneous SFPE (1.1). In particular, we showed that the methods of Collamore et al. (2011) can be extended to this setting, yielding an efficient importance sampling algorithm. While the algorithm is seen to work very efficiently, a difficulty in the implementation arises from the implicit expressions for the distributions of (A^*, B^*), which can be made explicit in a wide class of examples. In a more comprehensive study, the authors are also considering and comparing other related approaches to this problem, where these quantities can be obtained more explicitly. For the method described in this paper, efficiency and sharp asymptotics for the running time follow along the lines of Collamore et al. (2011), where similar results have been recently obtained for the related linear recursions.

References

Alsmeyer, G., J. D. Biggins, and M. Meiners. 2012. "The functional equation of the smoothing transform". *The Annals of Probability* 40 (5): 2069–2105.

Athreya, K. B., and P. E. Ney. 1972. *Branching Processes*. Springer-Verlag Berlin.

Biggins, J. D. 1977. "Martingale convergence in the branching random walk". *Journal of Applied Probability* 14 (1): 25–37.

Biggins, J. D. 1992. "Uniform convergence of martingales in the branching random walk". *The Annals of Probability* 20 (1): 137–151.

Collamore, J. F., G. Diao, and A. N. Vidyashankar. 2011. "Rare event simulation for processes generated via stochastic fixed point equations". Submitted (to appear in *The Annals of Applied Probability)*.

Collamore, J. F., and A. N. Vidyashankar. 2013a. "Probabilistic representation of constants in non-homogeneous recursions". In preparation.

Collamore, J. F., and A. N. Vidyashankar. 2013b. "Tail estimates for stochastic fixed point equations via nonlinear renewal theory". *Stochastic Processes and their Applications* 123 (9): 3378–3429.

Durrett, R., and T. M. Liggett. 1983. "Fixed points of the smoothing transformation". *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 64 (3): 275–301.

Iksanov, O., and P. NegadaÄlov. 2007. "The supremum of a martingale related to a branching random walk". *Theory of Probability and Mathematical Statistics* 74:49–57.

Jelenković, P., and M. Olvera-Cravioto. 2012. "Implicit renewal theory and power laws on trees". *Advances in Applied Probability* 44 (2): 528–561.

Kingman, J. F. C. 1975. "The first birth problem for an age-dependent branching process". *The Annals of Probability* 3 (5): 790–801.

Liu, Q. 2000. "On generalized multiplicative cascades". *Stochastic Processes and their Applications* 86 (2): 263–286.

Patil, G. P., and C. R. Rao. 1978. "Weighted distributions and size-biased sampling with applications to wildlife populations and human families". *Biometrics* 34 (2): 179–189.

Rösler, U., V. Topchii, and V. Vatutin. 2000. "Convergence conditions for weighted branching processes". *Diskr. Mat.* 12 (1): 7–23.