J. Appl. Prob. Spec. Vol. **48A,** 165–182 (2011)
© *Applied Probability Trust* 2011

NEW FRONTIERS IN APPLIED PROBABILITY

A Festschrift for SØREN ASMUSSEN
Edited by P. GLYNN, T. MIKOSCH and T. ROLSKI

Part 4. Simulation

ON EXACT SAMPLING OF STOCHASTIC PERPETUITIES

JOSE H. BLANCHET, *Columbia University*

Department of Industrial Engineering and Operations Research, Columbia University, S. W. Mudd Building, 500 West 120th Street, New York, NY 10025, USA.

KARL SIGMAN, *Columbia University*

Department of Industrial Engineering and Operations Research, Columbia University, S. W. Mudd Building, 500 West 120th Street, New York, NY 10025, USA.
Email address: karl.sigman@columbia.edu

APPLIED PROBABILITY TRUST
AUGUST 2011

https://doi.org/10.1239/jap/1318940463 Published online by Cambridge University Press

ON EXACT SAMPLING OF STOCHASTIC PERPETUITIES

By JOSE H. BLANCHET and KARL SIGMAN

Abstract

A stochastic perpetuity takes the form $D_\infty = \sum_{n=0}^{\infty} \exp(Y_1 + \cdots + Y_n)B_n$, where $(Y_n : n \geq 0)$ and $(B_n : n \geq 0)$ are two independent sequences of independent and identically distributed random variables (RVs). This is an expression for the stationary distribution of the Markov chain defined recursively by $D_{n+1} = A_n D_n + B_n$, $n \geq 0$, where $A_n = \mathrm{e}^{Y_n}$; D_∞ then satisfies the stochastic fixed-point equation $D_\infty \overset{\mathrm{D}}{=} AD_\infty + B$, where A and B are independent copies of the A_n and B_n (and independent of D_∞ on the right-hand side). In our framework, the quantity B_n, which represents a random reward at time n, is assumed to be positive, unbounded with $\mathrm{E}B_n^p < \infty$ for some $p > 0$, and have a suitably regular continuous positive density. The quantity Y_n is assumed to be light tailed and represents a discount rate from time n to $n-1$. The RV D_∞ then represents the net present value, in a stochastic economic environment, of an infinite stream of stochastic rewards. We provide an exact simulation algorithm for generating samples of D_∞. Our method is a variation of *dominated coupling from the past* and it involves constructing a sequence of dominating processes.

Keywords: Perfect sampling; coupling from the past; Markov chain; stochastic perpetuity

2010 Mathematics Subject Classification: Primary 65C05
Secondary 60J05; 68U20

1. Introduction

Let $(Y_n : n \geq 0)$ and $(B_n : n \geq 0)$ be two independent sequences of independent and identically distributed (i.i.d.) random variables (RVs), with Y and B denoting generic such copies. Suppose that the B_n are positive and denote the amount of reward obtained by running a system at time n, and that the discount rate from time n to time $n-1$ is precisely Y_n, so that the present value of B_n at time 0 is $B_n \exp(Y_n + Y_{n-1} + \cdots + Y_1)$. The net present value obtained by running the system over an infinite time horizon (starting with B_0 at time 0) is then given by the so-called stochastic perpetuity

$$D_\infty = \sum_{n=0}^{\infty} \exp(Y_1 + \cdots + Y_n)B_n.$$

This is an expression for the stationary distribution of the Markov chain defined recursively by

$$D_{n+1} = A_n D_n + B_n, \qquad n \geq 0, \tag{1}$$

where in our setup $A_n = \mathrm{e}^{Y_n}$; the expression for D_∞ is then derived by setting $D_0 = 0$ and iterating recursion (1) out to $n = \infty$ (while reversing the labeling of the RVs). Recursion (1) is an example of an ARCH(1) model, an important time series model used in statistics and econometrics; see [3, p. 469].

© *Applied Probability Trust* 2011

https://doi.org/10.1239/jap/1318940463 Published online by Cambridge University Press

When $\mathrm{E}\log A < 0$ and $\mathrm{E}\log(1+B) < \infty$ (see, for example, [18]), then this Markov chain has a proper stationary distribution limit D that satisfies the stochastic fixed-point equation

$$D \stackrel{\mathrm{D}}{=} AD + B,$$

where A and B are independent (and independent of D on the right-hand side).

We are interested in designing a simulation algorithm that allows us to obtain perfect samples of D_∞ under assumptions that allow us to accommodate a wide range of models of interest given the previous economic interpretation for D_∞. In particular, we assume that Y (not $A = \mathrm{e}^Y$) has a finite moment generating function in a neighborhood of the origin and that $\mathrm{E}Y \in (-\infty, 0)$. (The latter assumption is assumed merely to ensure the finiteness of D_∞.) We assume that B is positive with $\mathrm{E}B^p < \infty$ for some $p > 0$. The most important assumption that we impose concerns the existence of a suitably regular density for B, which is to be positive, continuous on $[0, \infty)$, and have a tail decay that is not too light (see Section 2). The types of example that are of most interest to us include situations in which B has a heavy-tailed distribution, such as a Pareto distribution. However, light-tailed distributions, such as a mixture of exponentials, are also tractable under our framework. We do not consider super-exponential tails, but we believe that methods related to our development here could be adapted to the case in which B has bounded support.

A situation that can be easily handled (hence, left out in the current paper) is that in which $p = \mathrm{P}(Y = -\infty) > 0$ (equivalently, $p = \mathrm{P}(A = 0) > 0$). For then, a simple direct *coupling from the past* (CFTP) algorithm (a general method introduced in [21]) applies as long as we can generate RVs distributed as A and B. Generate i.i.d. A_n, $n \geq 1$, until time $T = \min\{n \geq 0 \colon A_n = 0\}$ (geometric with success probability p), and then, given $T = n$, generate n i.i.d. copies $B_1, \dots, B_n$ and construct D_0 recursively from the past (from time $-n+1$ up to 0), using (1) with $D_{-n+1} = B_n$, and using the $T-1$ values (used to define T) $A_{n-1}, \dots, A_1$ and $B_{n-1}, \dots, B_1$. For example, if $T = 1$ then set $D_0 = B_1$. If $T = 2$ then set $D_0 = B_2 A_1 + B_1$, and so on. Then D_0 is distributed exactly as D_∞.

The above economic interpretation of D_∞ is useful in areas such as pension fund dynamics; in [8] a model was proposed based on stochastic perpetuities for the valuation of the pension fund. In the context of insurance risk theory, it is known that the distribution of D_∞ plays a crucial role in the evaluation of ruin probabilities with investments; see, for instance, [14], [19], and [23]. Explicit expressions for the distribution of D_∞ are, however, very challenging to obtain. Nevertheless, under very specific assumptions on the distributions of B and Y, such explicit expressions have been found in [12] and [24]; see also [20]. We can also view D_∞ as the stationary distribution of a continuous-time 'growth-collapse process' right before collapse epochs (see, for example, [16] and its many references). More general models in tree-like structures are discussed in [1] and [15].

As discussed in [4], D_∞ also plays a key role in applications arising in mathematical physics and finance. Applications to communication systems are given in [22]. Finally, Embrechts and Goldie [9] and Goldie and Grübel [13] mention applications in the complexity analysis of algorithms related to the so-called 'quickselect' algorithm and also in analytic number theory.

Applications to complexity analysis of sorting algorithms motivated the currently existing exact simulation methods for sampling D_∞. However, in those cases, $B = 1$ and $A = U^{1/\beta}$, where $U \sim U[0, 1]$ and $\beta > 0$; these perpetuities are known as Vervaat perpetuities [24]. The existing algorithms in this setting are constructed to sample from Vervaat perpetuities and related models. The first such sampling method based on a density approximation is given in [6]. It presents a sequence of upper and lower bounds for the density of D_∞, and then applies

https://doi.org/10.1239/jap/1318940463 Published online by Cambridge University Press

acceptance rejection; see also [7]. Fill and Huber [11] recently developed a dominated-coupling-from-the-past-based (DCFTP-based) procedure to sample Vervaat perpetuities. DCFTP is used to deal with the problem of sampling the steady-state distribution of unbounded Markov chains; see [17]. In turn, DCFTP was developed after the seminal paper [21], in which the CFTP protocol was introduced. A recent exposition on CFTP is given in [2, p. 120]. A nice summary of DCFTP is given in [11].

A generic class of DCFTP algorithms has been developed in [5] and [17]. There it was shown that, under certain ergodicity assumptions, DCFTP can be constructed using a suitable Foster–Lyapunov function and a suitable subsample scheme. While these procedures are substantially general and are in principle applicable to our setting, there are important limitations that are outlined by Connor and Kendall [5, p. 788]. In particular, in their algorithm they assumed that appropriate information is available in terms of the transition kernel of a Markov chain that is constructed based on k iterates of (1)—which is impractical in our setting. The value of k is found in order to ensure ergodicity of a suitable dominating process which turns out to be the workload system of a suitably defined D/M/1 queue.

In the present paper we will use a variation of DCFTP to generate our exact samples of D_∞. Typically, DCFTP requires the construction of a dominating stationary Markov chain that serves as a stochastic upper bound. Instead, we construct a sequence of upper bounds that does not form a stationary Markov chain per se, but otherwise is used in the same way as the dominating chain in DCFTP. We point out why our sequence of processes does not directly induce a simulatable stationary Markov chain at the end of Section 2. For the construction of our stochastic upper bounds, we also develop a simulation procedure to exactly sample from the steady-state distribution of a suitable GI/G/1 queue, relaxing some of the assumptions of an earlier algorithm given in [10].

The rest of the paper is organized as follows. In Section 2, we introduce our assumptions and our basic strategy which is summarized at the end of the section. The stochastic upper bounds required to implement our strategy are given in Sections 3 and 4. Finally, the remaining proofs to guarantee a finite termination time of our algorithm are given in Section 5.

2. Assumptions and basic strategy

We impose the following assumptions on the distributional properties of Y and B (generic copies of the Y_ns and the B_ns, respectively). (Recall that $Y = \ln(A)$.)

(A) $\psi(\theta) = \log \mathrm{E} \exp(\theta Y) < \infty$ for some $\theta > 0$, where Y is not deterministic. Furthermore, we assume that $\psi'(0) = \mathrm{E}\, Y < 0$; by the nondeterministic assumption we also have $\psi''(0) > 0$.

(B) B is unbounded (i.e. $\mathrm{P}(B > x) > 0,\ x \geq 0$) with continuous positive density $f(\cdot)$ on $(0, \infty)$ for which there exists a $\lambda_\kappa \in (0, \infty)$ such that

$$\lambda_\kappa \leq \inf_{0 \leq z \leq 1,\, y \geq \kappa} \frac{f(y-z)}{f(y-\kappa)} \tag{2}$$

for each $\kappa \geq 1$.

(C) $\mathrm{E} B^p < \infty$ for some $p > 0$.

We assume that the tail, $\bar{F}(\cdot) := \mathrm{P}(B > \cdot) = \int_\cdot^\infty f(s)\, \mathrm{d}s$, is available in closed form. Finally, we assume that, for each $\kappa > 0$, we can find a constant $C(\kappa) \in (0, \infty)$ such that

https://doi.org/10.1239/jap/1318940463 Published online by Cambridge University Press

$\mathrm{E}[(B+\kappa)^p] < C(\kappa)$ for some $p > 1$. Moreover, note that assumptions (A) and (C) imply the existence of a $\kappa_0 \in (0, \infty)$ satisfying

$$\mathrm{E}\log\left(\frac{1+\kappa_0\exp(Y)+B}{1+\kappa_0}\right) < \frac{1}{2}\mathrm{E}\,Y_1 < 0. \tag{3}$$

The previous observation will be used in the construction of a suitable Lyapunov bound to guarantee a finite termination time of our algorithm. In the sequel, we will be interested in the values of κ in (B) that are larger than κ_0.

Remark 1. In (A) we have assumed that Y has exponential moments and is nondeterministic. This is a mild assumption (recall that $A = \mathrm{e}^Y$) because it allows us to accommodate virtually all models of interest rate dynamics used in practice. The case of deterministic Y is substantially easier and can be treated with methods similar to those we discuss here. The most important assumption we impose is on B. The bound in (2) is naturally satisfied in applications in which the distribution of B is known and can be chosen by the modeler. It allows us to accommodate tails that are not too thin (typical tails that decrease at most exponentially fast satisfy (2)), for instance exponential, gamma, Pareto, or Weibull. A tail decreasing like a Gaussian or faster, for example, will typically not satisfy (2). Finally, assumption (C) is, we believe, also very mild and natural if one is concerned, as we are, with rewards that have unbounded support.

As we indicated earlier, our development is based on a slight variation of DCFTP in order to sample from the steady-state distribution of the Markov chain

$$D_{n+1} = \exp(Y_n)D_n + B_n, \tag{4}$$

with $D_0 = 0$. In order to apply these techniques, we need the following elements.

2.1. Elements of DCFTP for a monotone chain

We need to construct the following elements.

(E1) A recursion that preserves the monotonicity implied by recursion (4), but which also allows us to detect coalescence via coupling. This construction will require a suitable minorization condition.

(E2) A sequence of stochastic upper bounds for the steady-state distribution at subsequent deterministic times in the past, assuming that we have simulated the process starting from an arbitrarily long time in the distant past. These stochastic upper bounds need to be constructed jointly.

Typically DCFTP requires constructing a coupling and a suitable Markov chain that dominates the Markov chain of interest. It also requires being able to sample the dominating Markov chain in stationarity and being able to simulate the chain backwards in time. We use a slight variation of DCFTP because we do not construct a dominating chain per se but only a suitable sequence of stochastic upper bounds.

We will now construct (E1) and (E2). Recursion (4) is convenient because it has a useful monotonicity structure. In particular, the mapping

$$\phi_0(d, y, b) = \exp(y)d + b$$

is monotone increasing in d. However, in addition to monotonicity, in our construction we will need to introduce a coupling, as this is the tool that we will use to detect coalescence

https://doi.org/10.1239/jap/1318940463 Published online by Cambridge University Press

(as indicated in (E1)). Therefore, we will take advantage of the following minorization, which is applicable to the RV B.

Condition 1. (Minorization condition.) *Throughout our construction, we will set $\kappa > 0$. The selection of κ will be given according to our running time analysis in the last section, depending on a Lyapunov inequality. Suppose that $z \in [0, \kappa]$. Then, because of (2),*

$$\mathrm{P}(B + z \in y + \mathrm{d}y) = f(y - z)\,\mathbf{1}(y \geq z)\,\mathrm{d}y \geq \lambda_\kappa f(y - \kappa)\,\mathbf{1}(y \geq \kappa)\,\mathrm{d}y.$$

Therefore, we have

$$\mathrm{P}(B + z \in y + \mathrm{d}y) = \lambda_k\,\mathrm{P}(B + \kappa \in y + \mathrm{d}y) + (1 - \lambda_k)R(z, \mathrm{d}y), \tag{5}$$

where

$$R(z, \mathrm{d}y) = \frac{f(y - z)\,\mathbf{1}(y \geq z)\,\mathrm{d}y - \lambda_k f(y - \kappa)\,\mathbf{1}(y \geq \kappa)\,\mathrm{d}y}{1 - \lambda_k}.$$

Note that

$$\bar{H}_z(t) := \int_t^\infty R(z, \mathrm{d}y) = \frac{\bar{F}(t - z) - \lambda_k \bar{F}(t - \kappa)}{1 - \lambda_k}.$$

As a check, note that $\bar{H}_z(z) = 1$. In fact, in terms of the tail of the distribution of $B + z$ we see that the splitting in (5) yields the obvious identity

$$\begin{aligned}\mathrm{P}(B + z > t) &= \lambda_k \bar{F}(t - \kappa) + (1 - \lambda_k)\frac{\bar{F}(t - z) - \lambda_k \bar{F}(t - \kappa)}{1 - \lambda_k} \\ &= \lambda_k \bar{F}(t - \kappa) + (1 - \lambda_k)\bar{H}_z(t).\end{aligned}$$

We now consider the Markov chain representation based on the splitting induced by the minorization condition. If we let $z = \exp(y)d$, representation (5) induces a mapping $\phi_1(z, U, B)$ such that

$$\phi_1(z, U, B) \stackrel{\mathrm{D}}{=} \phi_0(d, y, B), \tag{6}$$

where U and B are independent with U uniformly distributed over $[0, 1]$. In particular, we let

$$\phi_1(z, U, B) = \begin{cases} B + z & \text{if } z := \exp(y)d > \kappa, \\ \mathbf{1}(U \leq \lambda_k)(B + \kappa) + \bar{H}_z^{-1}(\bar{F}(B))\,\mathbf{1}(U > \lambda_k) & \text{if } z := \exp(y)d \leq \kappa, \end{cases}$$

and define $B'(z, U, B) = \phi_1(z, U, B) - z$. We can then write

$$\phi_1(z, U, B) = z + B'(z) = \exp(y)d + B'(\exp(y)d).$$

The previous representation will allow us to deal with (E1). In order to see this, we need to introduce some notation and verify monotonicity properties of the mapping $\phi_1(\cdot, U, B)$. Assume that the RVs $\{(U_n, B_n) : n \geq 0\}$ (with the i.i.d. uniforms (U_n) independent of all else) are given, together with the sequence $\{Y_n : n \geq 0\}$. In fact, we will require to set the B_n according to (8) below, and the Y_ns will be simulated according to a suitable random walk

https://doi.org/10.1239/jap/1318940463 Published online by Cambridge University Press

construction as explained in Section 4. For $j = 0, 1, \dots, n$, set

$$\begin{aligned} D_0(n, w) &= w, \\ D_1(n, w) &= \phi_1(\exp(Y_n) D_0(n, w), U_{n-1}, B_{n-1}) \\ &= \exp(Y_n) D_0(n, w) + B'_{n-1}(\exp(Y_n) w, U_{n-1}, B_{n-1}), \\ D_2(n, w) &= \phi_1(\exp(Y_{n-1}) D_1(n, w), U_{n-1}, B_{n-1}) \\ &= \exp(Y_{n-1}) D_1(n, w) + B'_{n-2}(\exp(Y_{n-1}) D_1(n, w), U_{n-2}, B_{n-2}), \\ &\vdots \\ D_j(n, w) &= \exp(Y_{n-j+1}) D_{j-1}(n, w) + B'_{n-j}(\exp(Y_{n-j+1}) D_{j-1}(n, w), U_{n-j}, B_{n-j}). \end{aligned}$$

In simple words, the previous recursions are interpreted as follows. The value $D_0(n, w) = w$ indicates an initial value equal to w at n units of time in the past. Then, $D_j(n, w)$ represents the value at $n - j$ units of time in the past given that at n units of time in the past the position was w. Note that the value at $n - j$ units of time in the past given the position at $n - j + 1$ units of time in the past depends only on U_{n-j} and B_{n-j}, and that the driving sequence $\{(U_i, B_i) : i \geq 0\}$ is kept fixed even if we start the iterations at arbitrarily long times n in the past. Clearly, if $n \geq n_0$ and $0 \leq j \leq n_0$,

$$D_{n-j}(n, w) = D_{n_0 - j}(n_0, D_{n-n_0}(n, w)).$$

We now show the following useful monotonicity property.

Proposition 1. *If $w \leq v$ then*

$$D_j(n, w) \leq D_j(n, v)$$

for $j = 0, \dots, n$. Moreover,

$$z + B'(z, U, B) \leq (B + \kappa) + z.$$

Proof. Initially, we verify that if $z_0 \leq z_1$ then

$$\phi_1(z_0, U, B) \leq \phi_1(z_1, U, B).$$

First, if $z_0 \leq z_1 \leq \kappa$ or $\kappa \leq z_0 \leq z_1$, the result is clear because $\bar{H}_{z_0}(t) \leq \bar{H}_{z_1}(t)$ for all $t \geq 0$. Now, if $z_0 \leq \kappa \leq z_1$ then we have two cases. If $U \leq \lambda_k$ then clearly the inequality holds because $B + z_1 \geq B + \kappa$. If $U > \lambda_k$ then we need to show that

$$B + z_1 \geq \bar{H}_{z_0}^{-1}(\bar{F}(B)).$$

Now, since $\bar{H}_{z_0}(\cdot)$ is decreasing, $\bar{H}_{z_0}^{-1}(\cdot)$ is also decreasing and, therefore,

$$\bar{H}_{z_0}(B + z_1) \leq \bar{F}(B).$$

In addition, we note that $\bar{H}_{z_0}(t) \leq \bar{H}_{z_1}(t)$. Therefore, we have

$$\bar{H}_{z_0}(B + z_1) \leq \bar{H}_1(B + z_1) \leq \bar{H}_1(B + \kappa) = \frac{\bar{F}(B) - \lambda_k}{1 - \lambda_k} \leq \bar{F}(B).$$

Thus, the claim holds true. Now we proceed with the statement of the proposition using induction. For $j = 0$, the claim holds by definition. Assume that the inequality holds for $j - 1$.

https://doi.org/10.1239/jap/1318940463 Published online by Cambridge University Press

Then, by induction and monotonicity of $\phi_1(\cdot, U_{n-j}, B_{n-j})$, we obtain

$$\begin{aligned} D_j(n, w) &= \exp(Y_{n-j+1})D_{j-1}(n, w) + B'_{n-j}(D_{j-1}(n, w)) \\ &= \phi_1(\exp(Y_{n-j+1})D_{j-1}(n, w), U_{n-j}, B_{n-j}) \\ &\leq \phi_1(\exp(Y_{n-j+1})D_{j-1}(n, v), U_{n-j}, B_{n-j}) \\ &= \exp(Y_{n-j+1})D_{j-1}(n, v) + B'_{n-j}(D_{j-1}(n, v)), \end{aligned}$$

verifying the claim for j. The second part of the proposition follows similar steps. If $z \leq \kappa$ then $z + B'(z, U, B) \leq B + \kappa$ and $z \geq \kappa$ implies that $z + B'(z, U, B) = B + z$. In any case, $z + B'(z, U, B) \leq (B + \kappa) + z$.

To complete the construction of the basic elements behind our algorithm, let us write $S_n = Y_1 + \cdots + Y_n$ ($S_0 = 0$). It follows from (6) that, for any w,

$$D_n(n, w) \stackrel{\mathrm{D}}{=} \exp(S_n)w + \sum_{j=0}^{n-1} \exp(S_j)B_j,$$

and, therefore, since $\exp(S_n)w \to 0$ almost surely as $n \nearrow \infty$,

$$X \stackrel{\mathrm{D}}{=} \lim_{n \longrightarrow \infty} D_n(n, w),$$

where

$$X = \sum_{n=0}^{\infty} \exp(S_n)B_n.$$

Now let us define

$$\begin{aligned} B_n^+ &= B_n + \kappa, \\ W_n^+ &= B_n^+ + \exp(Y_{n+1})B_{n+1}^+ + \exp(Y_{n+1} + Y_{n+2})B_{n+2}^+ + \cdots, \\ W_n &= B_n + \exp(Y_{n+1})B_{n+1} + \exp(Y_{n+1} + Y_{n+2})B_{n+2} + \cdots. \end{aligned}$$

We actually have

$$X \stackrel{\mathrm{D}}{=} D_n(n, W_n).$$

Now assume that we can find a sequence of RVs $(V_k^+ : k \geq 0)$ such that $V_k^+ \geq W_k^+$; this is precisely (E2). The next basic result allows us to detect coalescence.

Proposition 2. *If there exists an $N_0 < \infty$ with probability 1 such that, for some $1 \leq j \leq N_0$,*

$$\exp(Y_j)D_{N_0-j}(N_0, V_{N_0}^+) \leq 1 \quad \textit{and} \quad U_{j-1} \leq \lambda_k,$$

then $D_n(n, W_n^+) = D_{N_0}(N_0, V_{N_0}^+)$ for all $n \geq N_0$. Moreover, $D_{N_0}(N_0, V_{N_0}^+) \stackrel{\mathrm{D}}{=} X$.

Proof. Note that, if $n \geq N_0$,

$$D_{n-N_0}(n, W_n^+) \leq D_{n-N_0}^+(n, W_n^+) = W_{N_0}^+ \leq V_{N_0}^+.$$

https://doi.org/10.1239/jap/1318940463 Published online by Cambridge University Press

Therefore, for each $1 \le j \le N_0$, we have

$$\begin{aligned} D_{n-j}(n, W_n^+) &= D_{N_0-j}(N_0, D_{n-N_0}(n, W_n^+)) \\ &\le D_{N_0-j}(N_0, D_{n-N_0}^+(n, W_n^+)) \\ &= D_{N_0-j}(N_0, W_{N_0}^+) \\ &\le D_{N_0-j}(N_0, V_{N_0}^+). \end{aligned}$$

So we have

$$\exp(Y_j) D_{n-j}(n, W_n^+) \le \exp(Y_j) D_{N_0-j}(N_0, V_{N_0}^+) \le 1$$

and $U_{j-1} \le 1$. This implies that the coalescence (coupling) occurs and, therefore, we must have $D_n(n, W_n^+) = D_{N_0}(N_0, V_{N_0}^+)$. To show that indeed $D_{N_0}(N_0, V_{N_0}^+) \stackrel{\mathrm{D}}{=} X$, we simply observe that

$$D_n(n, W_n^+) \stackrel{\mathrm{D}}{=} D_n(n, X^+) \stackrel{\mathrm{D}}{=} \exp(S_n) X^+ + \sum_{j=0}^{n-1} \exp(S_j) B_j,$$

where X^+ is a copy of W_n^+ which is independent of all the B_js and Y_js. The right-hand side converges to X almost surely as $n \nearrow \infty$ and the left-hand side equals $D_{N_0}(N_0, V_{N_0}^+)$ for $n \ge N_0$; the result follows.

The previous result is not very useful unless we are able to find a sequence of stochastic upper bounds (the V_k^+s) and ensure that $\mathrm{P}(N_0 < \infty) = 1$. The construction of these upper bounds will require first dealing with the maximum of an appropriate random walk and then simulating the B_ns in a suitable fashion. We will study the construction of the V_k^+s in the next sections. Assuming that such a construction is in place and that $N_0 < \infty$, the basic algorithm takes the following form. The proof that $\mathrm{P}(N_0 < \infty) = 1$ will be given in the last section of the paper.

Algorithm 1. (*Exact simulation of* X.) Set $\kappa > \kappa_0$, where κ_0 satisfies (3).

Step 1. At iteration $l \ge 1$ set $k = 2l$. Sample V_k^+, and let $D_0(k, V_k^+) = V_k^+$. (The definition of V_k^+ is given in (9) below.)

Step 2. Obtain $D_j(k, V_k^+)$ for $j = 1, 2, \dots, k$.

Step 3. If there exists a j such that $\exp(Y_j) D_{N_0-j}(N_0, V_{N_0}^+) \le \kappa$ and $U_{j-1} \le \lambda_k$, then let $X = D_k(k, V_k^+)$ and stop, otherwise let $m \leftarrow m + 1$ and go to step 1.

Remark 2. Note that we are using a Markov chain that has the same structure as D_n in order to construct our DCFTP-type algorithm, namely, one in which B_n is replaced by B_n^+. The standard application of DCFTP would therefore involve simulating a stationary version of the dominating Markov chain. This problem, however, is basically equivalent to the original problem. Nevertheless, we note that to carry over the basic ideas behind DCFTP, all we need is the construction of a stochastic upper bound for the steady distribution of our dominating chain; this is the role played by the V_k^+s and this is why our processes do not directly induce a single stationary Markov chain.

3. Simulatable stochastic upper bounds for the steady-state distribution

For any $a \in (0, 1)$, let us define

$$S_n(a) = Y_1(a) + \cdots + Y_n(a), \quad \text{where} \quad Y_j(a) = Y_j + a,$$

https://doi.org/10.1239/jap/1318940463 Published online by Cambridge University Press

and write

$$W_k^+ = \sum_{n=k}^{\infty} \exp(S_n(a) - S_k(a)) \exp(-(n-k)a) B_n^+$$
$$\le \exp(ka) \exp(M_k(a)) \sum_{n=k}^{\infty} \exp(-na) B_n^+, \tag{7}$$

where

$$M_k(a) = -S_k(a) + \max_{n \ge k} S_n(a).$$

Our strategy for constructing and simulating a suitable upper bound V_k^+ takes advantage of representation (7). We need to explain how to simulate subsequent elements of the sequence $\{M_k(a) : k \ge 0\}$. We also need to sample an upper bound for the infinite sum that is present on the right-hand side of (7); we first deal with this infinite sum and discuss the simulation of the $M_k(a)$s in the next section.

A useful observation is that, by assumption, for each $\alpha \in (0, a/2)$,

$$\mathrm{P}(B_n^+ > \exp(n\alpha)) \le C(\kappa) \exp(-n\alpha p).$$

Since $p > 1$, the Borel–Cantelli lemma ensures that the event $\{B_n^+ > n\alpha\}$ occurs just finitely many times. Let us define $T_0 = 0$ and $T_j = \inf\{n > T_{j-1} : B_n + \kappa > \exp(n\alpha)\}$ for $j = 1, 2, \dots$. We note that if $J = \max\{j \ge 0 : T_j < \infty\}$ then, as indicated earlier, $J < \infty$ almost surely and $1 \le \chi := \max\{n \ge 0 : B_n + \kappa > \exp(n\alpha)\} < \infty = T_J$.

We will explain how to simulate the B_ns jointly with the T_js. To do this, first define two sequences of independent RVs, namely $(\hat{B}_n : n \ge 0)$ and $(\bar{B}_n : n \ge 0)$. The corresponding distributions are as follows: $\hat{B}_n$ follows the distribution of B_n given that $B_n + \kappa > \exp(n\alpha)$, and $\bar{B}_n$ follows the distribution of B_n given that $B_n + \kappa \le \exp(n\alpha)$. We also define $(I_n : n \ge 0)$ to be a sequence of independent Bernoulli RVs (independent of the $\hat{B}_n$s and the $\bar{B}_n$s) such that $\mathrm{P}(I_n = 1) = p(n) := \mathrm{P}(B_n + \kappa \ge \exp(\alpha n))$. We can then write

$$B_n = \hat{B}_n I_n + \bar{B}_n (1 - I_n), \tag{8}$$

and $T_j = \inf\{n > T_{j-1} : I_n = 1\}$ for $j = 1, 2, \dots$ with $T_0 = 0$. Moreover, we have

$$\sum_{n=k}^{\infty} \exp(-na) B_n^+ \le \sum_{n=k}^{\infty} \exp(-na) \hat{B}_n I_n + \frac{\exp(-ka/2)}{1 - \exp(-a/2)}$$
$$= \sum_{n=k}^{\chi} \exp(-na) \hat{B}_n I_n + \frac{\exp(-ka/2)}{1 - \exp(-a/2)}.$$

Therefore, if we define

$$V_k^+ = \exp(ka) \exp(M_k(a)) \sum_{n=k}^{\chi} \exp(-na) \hat{B}_n I_n + \frac{\exp(M_k(a) + ka/2)}{1 - \exp(-a/2)} \tag{9}$$

then it clearly follows from (7) that $V_k^+ \ge W_k^+$.

https://doi.org/10.1239/jap/1318940463 Published online by Cambridge University Press

We now explain how to sample $T_1, T_2, T_3 \ldots$. We will consider T_1 as only sampling T_j, since the sampling of T_j given T_{j-1} is entirely analogous. Note that $T_1 > T_0 = 0$ and

$$P(T_1 = k) = p(k) \prod_{j=1}^{k-1} (1 - p(j))$$

for $k \geq 1$. Moreover, we have $P(T_1 = \infty) = \prod_{j=1}^{\infty}(1 - p(j)) \in (0, 1)$. We note that, by assumption, using Chebyshev's inequality,

$$p(n) \leq \min\{C(\kappa) \exp(-n\alpha p), 1\}.$$

Therefore, if m is such that $\frac{1}{2} > C(k) \exp(-map)$,

$$\begin{aligned}
\prod_{j=1}^{m-1}(1 - p(j)) \geq P(T_1 = \infty) \\
&= \prod_{j=1}^{\infty}(1 - p(j)) \\
&\geq \prod_{j=1}^{m-1}(1 - p(j)) \exp\biggl(-\sum_{j=m}^{\infty} \frac{2C(\kappa)}{\exp(m\alpha p)}\biggr) \\
&= \prod_{j=0}^{m-1}(1 - p(j)) \exp\biggl(-\frac{2C(\kappa)}{\exp(map)(1 - \exp(\alpha p))}\biggr). \qquad (10)
\end{aligned}$$

Consequently, in order to sample a Bernoulli RV Z with parameter $P(T_1 = \infty)$, we can simply let

$$Z = \mathbf{1}(U \leq P(T_1 = \infty)),$$

where U is uniformly distributed in $[0, 1]$. Note that, with probability 1, the condition that $U \leq P(T_1 = \infty)$ can be obtained from the bounds of (10) by making m sufficiently large without computing the infinite product in the definition of $P(T_1 = \infty)$.

Now, if $T_1 < \infty$, we need to simulate an RV with probability mass function

$$\begin{aligned}
P(T_1 = k \mid T_1 < \infty) &= p(k) \frac{\prod_{j=1}^{k-1}(1 - p(j))}{\prod_{j=1}^{\infty}(1 - p(j))} \\
&\leq \frac{1}{\prod_{j=0}^{\infty}(1 - p(j))} \min\{C(\kappa) \exp(-k\alpha p), 1\}.
\end{aligned}$$

Once again, we apply an acceptance-rejection procedure. A suitable proposal RV K, with probability mass function

$$P(K = k) = \exp(-[k - 1]\alpha p)(1 - \exp(\alpha p))$$

for $k \geq 1$ works in this case. This type of procedure allows us to simulate the sequence $(I_n : n \geq 0)$. Simulating the sequences $(\hat{B}_n : n \geq 0)$ and $(\bar{B}_n : n \geq 0)$ is immediate.

https://doi.org/10.1239/jap/1318940463 Published online by Cambridge University Press

4. The maxima of a negative drift random walk

Our goal is to simulate the $M_k(a)$s jointly with the random walk $(S_n(a): n \le k)$. The design of our algorithm is based on importance sampling. We first need the next lemma, which follows easily from the strict convexity of $\psi(\cdot)$ and so its proof is omitted.

Lemma 1. *Suppose that the moment generating function of the nondegenerate RV Y is finite in a neighborhood of the origin, so that $\psi'(0) < 0$ and $\psi''(0) > 0$. Define $\psi_a(\theta) = \log \mathrm{E} \exp(\theta Y(a)) = \psi(\theta) + a\theta$. Then we can always find $a > 0$ and $\eta = \eta(a) > 0$ such that $\psi_a(\eta) = 0$.*

Remark 3. In the so-called Cramer case, that is, when there exists $\theta_* > 0$ such that $\psi(\theta_*) > 0$, then, for any $a \in (0, |\psi'(0)|/2)$, we can find the required $\eta(a)$.

Lemma 1 guarantees that there exist $a > 0$ and $\eta > 0$ such that $\psi_a(\eta) = 0$, $\psi_a'(0) < 0$, and $\psi_a'(\eta) > 0$. The root η allows us to define a convenient change of measure which we will use repeatedly in our sampling strategy. In particular, if $L_n = \exp(\eta S_n(a))$ then $(L_n : n \ge 0)$ is a positive martingale and induces a probability measure P_η defined for each $A \in \sigma(S_j(a): j \le k)$ (the σ-field generated by $S_1(a), \dots, S_k(a)$) via $\mathrm{P}_\eta(A) = \mathrm{E}[\exp(-\eta S_k(a)); A]$. It is well known that, under $\mathrm{P}_\eta(\cdot)$, the random walk has positive drift equal to $\psi_a'(0) > 0$. In fact, if we let $\xi > 0$ and set $T_\xi = \inf\{n \ge 0: S_n(a) > \xi\}$, then we have

$$\mathrm{P}(T_\xi < \infty) = \mathrm{E}_\eta[\exp(-\eta S_{T_\xi})].$$

Moreover, if $\xi_1 > \xi_0$ then

$$\mathrm{P}(T_{\xi_0} < \infty,\ T_{\xi_1} = \infty) = \mathrm{E}_\eta[\exp(-\eta S_{T_{\xi_0}})\, \mathrm{P}_{S_{T_{\xi_0}}}(T_{\xi_1} = \infty)].$$

If all we wanted was to simulate $M_0(a)$, we could take advantage of the following idea of Ensor and Glynn [10]. They observed that if an artificial RV τ, exponentially distributed with unit mean and independent of the random walk under P_η, is introduced then

$$\mathrm{P}(M_0(a) > x) = \mathrm{E}_\eta[\exp(-\eta S_{T_x})] = \mathrm{P}_\eta\left(\frac{\tau}{\eta} > S_{T_x}\right).$$

Then, if we define the (random) function $G(u) = \inf\{x \ge 0: S_{T_x} > u\}$, we obtain $G(S_{T_x}) = x$ for almost every x with respect to the Lebesgue measure and, therefore, we conclude that

$$\mathrm{P}(M_0(a) > x) = \mathrm{P}_\eta\left(G\left(\frac{\tau}{\eta}\right) > x\right). \tag{11}$$

In other words, we have $M_0(a) \overset{\mathrm{D}}{=} G(\tau/\eta)$, and, therefore, we can sample $M_0(a)$ in finite time by sampling τ and then computing $G(\tau/\eta)$, which requires simulating $S_1(a), \dots, S_{T_{\tau/\eta}}(a)$ under $\mathrm{P}_\eta(\cdot)$. Our problem, however, is to jointly simulate the $M_k(a)$s and the underlying random walk and for this reason we will require a sequential procedure.

Fix $m \ge 1$ so that $\exp(-3\eta m) < \frac{1}{2}$. This is a technical constraint on m whose nature will become evident momentarily. Define the sequence of times $\Delta_1 = \inf\{n \ge 0: S_n(a) < -2m\}$, $\Gamma_1 = \inf\{n \ge \Delta_1: S_n(a) - S_{\Delta_1}(a) > m\}$, and, for $j \ge 2$, $\Delta_j = \inf\{n \ge \Gamma_{j-1}\mathbf{1}(\Gamma_{j-1} < \infty) \vee \Delta_{j-1}: S_n < S_{\Delta_{j-1}} - 2m\}$ and $\Gamma_j = \inf\{n \ge \Delta_j: S_n - S_{\Delta_j} > m\}$. We use the convention that if $\Gamma_{j-1} = \infty$ then $\Gamma_{j-1}\mathbf{1}(\Gamma_{j-1} < \infty) = 0$, so we have $\Gamma_{j-1}\mathbf{1}(\Gamma_{j-1} < \infty) > \Delta_{j-1}$ if and only if $\Gamma_{j-1} < \infty$. We will sequentially simulate the random walk at the times $\Delta_1, \Delta_2, \dots$

https://doi.org/10.1239/jap/1318940463 Published online by Cambridge University Press

jointly with the sequence $\Gamma_1, \Gamma_2, \ldots$. Note that $\mathrm{P}(\Gamma_1 = \infty \mid S_{\Delta_1}(a)) > 0$, so simulating Δ_2 sequentially given Γ_1 requires being able to simulate the random walk conditional on $\Gamma_1 = \infty$, and similarly for subsequent Δ_js. We will explain how to simulate the Δ_js and Γ_js sequentially jointly with the underlying random walk. However, first we note that indeed this sequential simulation procedure is all that is needed to simulate the $M_k(a)$s jointly with the random walk $(S_n(a) \colon n \geq 0)$.

Proposition 3. *We have $\Delta_n < \infty$ with probability 1 for each $n \geq 1$ and $\Delta_n \nearrow \infty$ as $n \nearrow \infty$ also with probability 1. Furthermore, if $\exp(-3\eta m) < \frac{1}{2}$ then the event $\mathrm{P}(\Gamma_n = \infty$ infinitely often$) = 1$. Consequently, for each $k \geq 0$, we can find $N_0(k) = \inf\{n \geq 1 \colon \Delta_n \geq k\}$ and $\mathcal{T}(k) = \inf\{j \geq N_0(k) + 1 \colon \Gamma_j = \infty\}$, both finite RVs such that $M_k(a) = -S_k(a) + \max_{k \leq n \leq \Delta_{\mathcal{T}(k)}} S_n(a)$.*

Proof. The first statement of the proposition follows easily from the law of large numbers since $\mathrm{E} Y_1(a) < 0$. Now we show that $\mathrm{P}(\Gamma_n = \infty$ infinitely often$) = 1$. First it follows from the definition of Γ_1 that $\mathrm{P}(\Gamma_1 = \infty \mid S_{\Delta_1}(a)) = \mathrm{P}(T_m = \infty) > 0$. We claim that, for $j \geq 2$, we can find a $\delta > 0$ such that

$$\mathrm{P}(\Gamma_j = \infty \mid S_1, \ldots, S_{\Delta_j}, \Gamma_1, \ldots, \Gamma_{j-1}) \geq \delta > 0.$$

To see this, first suppose that $\Gamma_l < \infty$ for each $l = 1, 2, \ldots, j-1$. Then, by the strong Markov property we have

$$\mathrm{P}(\Gamma_j = \infty \mid S_1, \ldots, S_{\Delta_j}, \Gamma_1, \ldots, \Gamma_{j-1}) = \mathrm{P}(T_m = \infty) > 0.$$

Now suppose that $\Gamma_l = \infty$ for some $l \leq j-1$, and let $l^* = \max\{l \leq j-1 \colon \Gamma_l = \infty\}$. Define $K = S_{\Delta_{l^*}} + m - S_{\Delta_j} \geq 3m$, and note that

$$\mathrm{P}(\Gamma_j < \infty \mid S_1, \ldots, S_{\Delta_j}, \Gamma_1, \ldots, \Gamma_{j-1}) = \mathrm{P}(T_m < \infty \mid T_K = \infty).$$

Keep in mind that in the conditional probability that appears on the right-hand side we regard K as a deterministic constant. Now we have

$$\mathrm{P}(T_m < \infty \mid T_K < \infty) = \frac{\mathrm{P}(T_m < \infty,\ T_K = \infty)}{1 - \mathrm{P}(T_K < \infty)} = \frac{\mathrm{E}_\eta[\exp(-\eta S_{T_m})\,\mathrm{P}_{S_{T_m}}(T_K = \infty)]}{1 - \mathrm{P}(T_K < \infty)}.$$

Since $K \geq 3m$, we have $\mathrm{P}(T_K = \infty) = 1 - \mathrm{P}(T_K < \infty) \geq 1 - \exp(-3\eta m)$. Therefore, the previous expression implies that

$$\mathrm{P}(\Gamma_j = \infty \mid S_1, \ldots, S_{\Delta_j}, \Gamma_1, \ldots, \Gamma_{j-1}) \geq 1 - \frac{\exp(-3\eta m)}{1 - \exp(-3\eta m)}.$$

The right-hand side is strictly positive if $\exp(-3\eta m) < \frac{1}{2}$. Since the right-hand side is nonrandom, it follows from the Borel–Cantelli lemma that $\mathrm{P}(\Gamma_n = \infty$ infinitely often$) = 1$. Finally, the fact that $M_k(a) = -S_k(a) + \max_{k \leq n \leq \Delta_{\mathcal{T}(k)}} S_n(a)$ follows easily by construction. Note that it is important, however, to define $\mathcal{T}(k) \geq N_0(k) + 1$, so that $\Delta_{N_0(k)+1}$ is computed first and we can make sure that the maximum of the sequence $\{S_n(a) \colon n \geq k\}$ is achieved between k and $\Delta_{\mathcal{T}(k)}$.

The next three lemmas provide the basis to simulate the random walk $(S_n(a) \colon n \geq 0)$ jointly with the Δ_js and the Γ_js. First, in Lemma 2 we provide a representation that allows us to simulate a Bernoulli RV with success parameter $\mathrm{P}(T_{\xi_0} < \infty \mid T_{\xi_1} = \infty)$. The result is straightforward and so the proof is omitted.

https://doi.org/10.1239/jap/1318940463 Published online by Cambridge University Press

Lemma 2. *Let* $0 < \xi_0 < \xi_1 \leq \infty$. *Then we have*

$$\mathrm{P}(M_0(a) > \xi_0 \mid M_0(a) \leq \xi_1) = \mathrm{P}(T_{\xi_0} < \infty \mid T_{\xi_1} = \infty).$$

In particular, we can simulate a Bernoulli RV with parameter $\mathrm{P}(T_{\xi_0} < \infty \mid T_{\xi_1} = \infty)$ *if we just sample* $M_0(a)$ *given that* $M_0(a) \leq \xi_1$ *and then output* $\mathbf{1}(M_0(a) > \xi_0)$.

Next we describe how to simulate the random walk conditional on $T_{\xi_0} < \infty$ and $T_{\xi_1} = \infty$.

Lemma 3. *Let* $0 < \xi_0 < \xi_1 \leq \infty$, *consider any sequence of bounded positive measurable functions* $f_{k+1}\colon \mathbb{R}^{k+1} \to [0, \infty)$, *and define* $\zeta(\xi_0, \xi_1) = \exp(-\eta S_{T_{\xi_0}}) \mathrm{P}_{S_{T_{\xi_0}}}(T_{\xi_1} = \infty)$. *Then we obtain*

$$\mathrm{E}[f_{T_{\xi_0}}(S_0(a), \dots, S_{T_{\xi_0}}(a)) \mid T_{\xi_0} < \infty,\ T_{\xi_1} = \infty] = \frac{\mathrm{E}_\eta[f_{T_{\xi_0}}(S_0(a), \dots, S_{T_{\xi_0}}(a))\zeta(\xi_0, \xi_1)]}{\mathrm{E}_\eta[\zeta(\xi_0, \xi_1)]}.$$

So, if $\mathrm{P}^*(\cdot) = \mathrm{P}(\cdot \mid T_{\xi_0} < \infty,\ T_{\xi_1} = \infty)$, *we conclude that*

$$\frac{\mathrm{dP}^*}{\mathrm{dP}_\eta} = \frac{\zeta(\xi_0, \xi_1)}{\mathrm{E}_\eta[\zeta(\xi_0, \xi_1)]} \leq \frac{1}{\mathrm{E}_\eta[\zeta(\xi_0, \xi_1)]}. \tag{12}$$

Consequently, we can apply acceptance rejection. In particular, we propose a sample $S_1(a), \dots, S_{T_{\xi_0}}(a)$ *from* $\mathrm{P}_\eta(\cdot)$ *and accept with probability*

$$\exp(-\eta S_{T_{\xi_0}}) \mathrm{P}_{S_{T_{\xi_0}}}(T_{\xi_1} = \infty).$$

Finally, we observe that acceptance occurs with probability precisely equal to $\mathrm{P}(T_{\xi_0} < \infty, T_{\xi_1} = \infty)$. *In particular, if* $\xi_1 = \infty$, *we have* $\mathrm{P}_{S_{T_{\xi_0}}}(T_{\xi_1} = \infty) = 1$ *and* $\mathrm{P}(T_{\xi_0} < \infty,\ T_{\xi_1} = \infty) = \mathrm{P}(T_{\xi_0} < \infty)$, *so in this case the acceptance step yields a Bernoulli with parameter* $\mathrm{P}(T_{\xi_0} < \infty)$, *and if the Bernoulli is successful, the sample path follows the law* $S_1(a), \dots, S_{T_{\xi_0}}(a)$ *given that* $T_{\xi_0} < \infty$.

Proof. Martingale theory and the strong Markov property yield

$$\begin{aligned}&\mathrm{E}[f_{T_{\xi_0}}(S_0(a), \dots, S_{T_{\xi_0}}(a)),\ T_{\xi_0} < \infty,\ T_{\xi_1} = \infty] \\ &\quad = \mathrm{E}_\eta[f_{T_{\xi_0}}(S_0(a), \dots, S_{T_{\xi_0}}(a)) \exp(-\eta S_{T_{\xi_0}}) \mathrm{P}_{S_{T_{\xi_0}}}(T_{\xi_1} = \infty)].\end{aligned}$$

Letting $f_k = 1$ we conclude that

$$\mathrm{P}(T_{\xi_0} < \infty,\ T_{\xi_1} = \infty) = \mathrm{E}_\eta[\zeta(\xi_0, \xi_1)],$$

and, therefore, we arrive at the likelihood ratio in (12). The rest of the proof follows using standard results for acceptance-rejection algorithms.

Finally, given $\xi_0 \in (0, \infty)$, define $T_{-\xi_0} = \inf\{n \geq 0\colon S_n < -\xi_0\}$. We will explain how to simulate a path up to time $T_{-\xi_0}$ conditional on $T_{\xi_1} = \infty$ for any $\xi_1 \in (0, \infty]$.

Lemma 4. *Let* $0 < \xi_0 < \xi_1 \leq \infty$, *and consider any sequence of bounded positive measurable functions* $f_{k+1}\colon \mathbb{R}^{k+1} \to [0, \infty)$. *Then*

$$\mathrm{E}[f_{T_{\xi_0}}(S_0(a), \dots, S_{T_{-\xi_0}}(a)) \mid T_{\xi_1} = \infty] = \frac{\mathrm{E}[f_{T_{\xi_0}}(S_0(a), \dots, S_{T_{\xi_0}}(a)) \mathrm{P}_{S_{T_{-\xi_0}}}(T_{\xi_1} = \infty)]}{\mathrm{P}(T_{\xi_1} = \infty)}.$$

https://doi.org/10.1239/jap/1318940463 Published online by Cambridge University Press

So, if $\mathrm{P}^*(\cdot) = \mathrm{P}(\cdot \mid T_{\xi_1} = \infty)$, *we conclude that*

$$\frac{\mathrm{dP}^*}{\mathrm{dP}} = \frac{\mathrm{P}_{S_{T_{-\xi_0}}}(T_{\xi_1} = \infty)}{\mathrm{P}(T_{\xi_1} = \infty)} \leq \frac{1}{\mathrm{P}(T_{\xi_1} = \infty)}.$$

Consequently, we can apply acceptance rejection to sample $S_0(a), \dots, S_{T_{-\xi_0}}(a)$ *given* $T_{\xi_1} = \infty$. *In particular, we propose a sample* $S_1(a), \dots, S_{T_{-\xi_0}}(a)$ *from* $S_0(a), \dots, S_{T_{-\xi_0}}(a)$ *under the nominal (unconditional probability) and accept the path with probability* $\mathrm{P}_{S_{T_{-\xi_0}}(a)}(T_{\xi_1} = \infty)$.

Proof. The result follows directly from the strong Markov property and basic facts of acceptance rejection.

We conclude the section with a summary of the sequential procedure for the random walk.

Algorithm 2. (*Sequential simulation of the random walk.*) Select m such that $\exp(-3\eta m) < \frac{1}{2}$. Set $s_0 = s = 0$, $\xi_1 = \infty$, and $t = 0$.

Output the random walk $(s_0, s_1, \dots)$ and the times $(\Delta_1, \Delta_2, \dots)$, $(\Gamma_1, \Gamma_2, \dots)$.

At iteration $k \geq 1$ proceed as follows.

Step 1. Sample $S_1(a), \dots, S_{T_{-2m}}(a)$, T_{-2m} given that $T_{\xi_1} = \infty$ and $S_0(a) = 0$ (apply Lemma 4 and (11)).

Step 2. Let $\Delta_k = t + T_{-2m}$, $s_{t+1} = s + S_1(a), \dots, s_{\Delta_k} = s + S_{T_{-2m}}(a)$, $t \leftarrow \Delta_k$, and $s \leftarrow s_{\Delta_k}$.

Step 3. Simulate a Bernoulli RV J with parameter $\mathrm{P}(T_m < \infty \mid T_{\xi_1} = \infty)$ (apply Lemma 2).

Step 4. **If** $J = 1$ then sample $S_1(a), \dots, S_{T_m}(a)$, T_m given that $T_m < \infty$ and $T_{\xi_1} = \infty$ (apply Lemma 3). Let $\Gamma_k = t + T_m$, $s_{t+1} = s + S_1(a), \dots, s_{\Gamma_k} = s + S_{T_m}(a)$, $t \leftarrow \Delta_k$, and $s \leftarrow s_{\Gamma_k}$. **Else** $(J = 0)$ let $\xi_1 = s + m$ and $\Gamma_k = \infty$.

Step 5. Set $k \leftarrow k + 1$ and go to step 1.

5. Finite termination time

All the elements of our algorithm for the exact simulation of X are in place now. The simulation of the B_ns has been discussed in Section 3, how to simulate $Y_n = Y_n(a) - a$ for $n \geq 1$ has been discussed in Section 4, and how to construct and simulate the V_n^+s have also been discussed in these two sections. The only remaining issue is to make sure that $N_0 < \infty$ with probability 1.

Let ϱ be the number of iterations that are required to terminate the algorithm. In order to show that $\mathrm{E}\varrho < \infty$, we will take advantage of the following bound:

$$\begin{aligned} V_k^+ &= \exp(ka)\exp(M_k(a)) \sum_{n=k}^{\chi} \exp(-na)\hat{B}_n I_n + \frac{\exp(M_k(a) + ka/2)}{1 - \exp(-a/2)} \\ &\leq D^+ \\ &:= \exp(\chi a)\exp(M_0(2a)) \sum_{n=0}^{\chi} \exp(-n\alpha)\hat{B}_n + \frac{\exp(M_0(2a))}{1 - \exp(-a/2)}. \end{aligned}$$

The following lemma allows us to provide a bound on $\mathrm{E}\varrho$ based on a bound on the mean of $\Theta_\kappa = \inf\{n \geq 0 \colon D_n < \kappa\}$, given that D_0 is selected as an independent copy of D^+. The fact that $N_0 < \infty$ with probability 1 clearly follows from the fact that $\mathrm{E}\varrho < \infty$.

https://doi.org/10.1239/jap/1318940463 Published online by Cambridge University Press

Lemma 5. *Let* $g_\kappa(d) = \mathrm{E}_d\Theta_\kappa$. *Then*

$$\mathrm{E}g_\kappa(D^+) \le \frac{1}{1-\lambda_\kappa}\Big\{\mathrm{E}g_\kappa(D^+) + \sup_{0\le x\le 1} \mathrm{E}_x[g_\kappa(\exp(Y)x + B)]\Big\}.$$

Proof. First, for convenience, let us extend the construction $\{D_j(2l, V_{2l}^+)\colon 0 \le j \le 2l\}$ to values of $j > 2l$ using the dynamics of the Markov chain with an independent sequence of B_ns, Y_ns, and U_ns. Now let $A_{2l}(1, V_{2l}^+) = \inf\{j \ge 0\colon D_j(2l, V_{2l}^+) \le 1\}$, and set

$$A_{2l}(i+1, V_{2l}^+) = \inf\{j > A_{2l}(i, V_{2l}^+)\colon D_j(2l, V_{2l}^+) \le 1\}.$$

We then define $N(k, V_{2l}^+) = \max\{m \ge 0\colon A_{2l}(m, V_{2l}^+) \le k\}$. We note that

$$\mathrm{P}(\varrho > l) \le \mathrm{E}[(1-\lambda)^{N(2l, V_{2l}^+)}].$$

The previous inequality simply says that $\varrho > l$ implies that the process $\{D_j(2l, V_{2l}^+)\colon 0 \le j \le 2l\}$ either does not visit the interval $[0, \kappa]$ or, when it does visit $[0, \kappa]$, a successful coupling does not occur. Now we introduce an artificial RV τ, geometrically distributed with success parameter λ. Then

$$\mathrm{E}[(1-\lambda)^{N(2l, V_{2l}^+)}] = \mathrm{P}(\tau > N(2l, V_{2l}^+)) \le \mathrm{P}(A_{2l}(\tau, V_{2l}^+) > 2l).$$

Consequently,

$$\mathrm{E}\varrho \le \mathrm{E}A_{2l}(\tau, V_{2l}^+).$$

By monotonicity and a submartingale argument, we obtain

$$\mathrm{E}A_{2l}(\tau, V_{2l}^+) \le \frac{1}{1-\lambda_\kappa}\Big\{\mathrm{E}g_\kappa(D^+) + \sup_{0\le x\le 1} \mathrm{E}_x[g_\kappa(\exp(Y)x + B)]\Big\},$$

thereby concluding the result.

In order to obtain a bound on $g_\kappa(d)$, we will take advantage of the following well-known Foster–Lyapunov criterion.

Proposition 4. *Suppose that we can find a nonnegative function* $h_\kappa(\cdot)$ *such that*

$$\mathrm{E}[h_\kappa(d\exp(Y_1) + B_1)] - h_\kappa(d) \le -1$$

for all $d \ge \kappa$. *Then* $g_\kappa(d) \le h_\kappa(d)$.

The following lemma provides the construction of a suitable Lyapunov function from the previous proposition.

Lemma 6. *Given* $\kappa > \kappa_0$, *we can find a* $c \in (0, \infty)$ *such that* $h(x) = c\log(1+x)$ *is an appropriate Lyapunov function. In particular, for each* $d \ge \kappa$, *we have* $\mathrm{E}_d\Theta_\kappa \le h_\kappa(d)$.

Proof. Note that

$$0 \le \log\bigg(\frac{1 + d\exp(Y_1) + B_1}{1+d}\bigg) \le \log(1 + B_1 + \exp(Y_1)).$$

In addition, for each $\delta > 0$, we can find a $C_\delta \in (0, \infty)$ such that $\log(1+r) \le C_\delta r^\delta$. Therefore,

$$\log\left(\frac{1 + B_1 + \exp(Y_1)}{1 + B_1}\right) = \log\left(1 + \frac{\exp(Y_1)}{1 + B_1}\right) \le C_\delta \exp(\delta Y_1),$$

and we conclude that

$$0 \le \log\left(\frac{1 + d\exp(Y_1) + B_1}{1 + d}\right) \le \log(1 + B_1) + C_\delta \exp(\delta Y_1).$$

We then recall bound (3) and the fact that $\kappa > \kappa_0$ to conclude that if $c > 2/|\mathrm{E} Y_1|$ and $d \le \kappa$, then

$$\mathrm{E}[h_\kappa(d\exp(Y_1) + B_1)] - h_\kappa(d) = c\,\mathrm{E}\log\left(\frac{1 + d\exp(Y_1) + B_1}{1 + d}\right) \le \frac{c\,\mathrm{E}\,Y_1}{2} \le -1,$$

thereby concluding the result.

Lemma 7. *We have*

$$\mathrm{E} g_\kappa(D^+) < \infty \quad \textit{and} \quad \sup_{0 \le x \le 1} \mathrm{E}_x[g_\kappa(\exp(Y)x + B)] < \infty.$$

Therefore, $\mathrm{E}\varrho < \infty$ *and, thus, the algorithm terminates in finite time with probability* 1.

Proof. It suffices to show that $\mathrm{E}\log(1 + D^+) < \infty$ to show the first part of the result. The second part is straightforward, following a similar argument as in the proof of Lemma 6. Note that

$$\begin{aligned}
&\mathrm{E}\log(1 + D^+) \\
&\quad\le \mathrm{E} M_0(2a) + \mathrm{E}\left\{\log\left[1 + \exp(\chi a)\sum_{n=0}^{\chi}\exp(-n\alpha)\hat{B}_n + \frac{1}{1 - \exp(-a/2)}\right]\right\} \\
&\quad\le \mathrm{E} M_0(2a) + \log\left(\frac{3}{1 - \exp(-a/2)}\right) + a\,\mathrm{E}\,\chi + \mathrm{E}\log\left(\chi \max_{n=0,\dots,\chi}\exp(-n\alpha)\hat{B}_n\right).
\end{aligned}$$

It is well know that $M_0(a)$ has exponentially decaying tails, so, in particular, $\mathrm{E} M_0(a) < \infty$. Now we have

$$\mathrm{P}(\chi = k) \le \mathrm{P}(\kappa + B_k \ge \exp(k\alpha)).$$

Since we are assuming that $\mathrm{E} B_k^p < \infty$ for some $p > 0$ we clearly obtain $\mathrm{E}\chi < \infty$. Observe that

$$\mathrm{E}\log\left(\max_{n=0,\dots,\chi}\exp(-n\alpha)\hat{B}_n\right) \le \mathrm{E}\left[\sum_{n=0}^{\chi}\{-n\alpha + \log(\hat{B}_n)\}\right],$$

and also note that, for each $\delta > 0$, there exists a constant $C(\delta, n)$ such that $C(\delta, n) \to 0$ as $n \nearrow \infty$. Then

$$\begin{aligned}
\mathrm{E}[-n\alpha + \log(B_n + \kappa) \mid B_n + \kappa \ge \exp(na)] &= \int_{na}^{\infty}\frac{\mathrm{P}(\log(B_n + \kappa) > t)}{\mathrm{P}(\log(B_n + \kappa) > n\alpha)}\,\mathrm{d}t \\
&\le C(\delta, n)\,\mathrm{P}(\log(B_n + \kappa) > n\alpha)^{-\delta}.
\end{aligned}$$

https://doi.org/10.1239/jap/1318940463 Published online by Cambridge University Press

Therefore, if $\alpha \in (0, a/2)$,

$$\mathrm{E}\left[\sum_{n=0}^{\chi}\{-n\alpha + \log(\hat{B}_n)\}\right] \le \mathrm{E}\{\chi\, \mathrm{P}(\log(B_n+\kappa) > \chi)^{-\delta}\}.$$

We then conclude that

$$\mathrm{E}[\chi\, \mathrm{P}(\log(B_n+\kappa) > \chi \mid \chi)^{-\delta}] \le \sum_{k=1}^{\infty} k\, \mathrm{P}(\log(B_n+\kappa) > k)^{1-\delta} < \infty,$$

where the last inequality follows from Chebyshev's bound again using the fact that $\mathrm{E}(B_k+\kappa)^p < \infty$ for some $p > 0$. The result then follows from Lemma 5.

Acknowledgements

We are grateful to the referee and the Associate Editors for their comments and suggestions. Thanks to Jing Dong and the Technical Editor for their helpful comments. The first author acknowledges support from the NSF under grants DMS-0902075 and CMMI-0846816.

References

[1] Aldous, D. J. and Bandyopadhyay, A. (2005). A survey of max-type recursive distributional equations. *Ann. Appl. Prob.* **15,** 1047–1110.
[2] Asmussen, S. and Glynn, P. W. (2008). *Stochastic Simulation: Algorithms and Analysis*. Springer, New York.
[3] Campbell, J. Y., Lo, A. W. and MacKinlay, C. (1999). *The Econometrics of Financial Markets*. Princeton University Press, Princeton, NJ.
[4] Carmona, P., Petit, F. and Yor, M. (2001). Exponential functionals of Lévy processes. In *Lévy Processes*, eds O. Barndorff-Nielsen *et al.*, Birkhäuser, Boston, MA, pp. 41–55.
[5] Connor, S. B. and Kendall, W. S. (2007). Perfect simulation for a class of positive recurrent Markov chains. *Ann. Appl. Prob.* **3,** 781–808.
[6] Devroye, L. (2001). Simulating perpetuities. *Methodology Comput. Appl. Prob.* **3,** 97–115.
[7] Devroye, L. and Neininger, R. (2002). Density approximation and exact simulation of random variables that are the solutions of fixed-point equations. *Adv. Appl. Prob.* **34,** 441–468.
[8] Dufresne, D. (1990). The distribution of a perpetuity, with applications to risk theory and pension funding. *Scand. Actuarial J.* **1990,** 39–79.
[9] Embrechts, P. and Goldie, C. M. (1994). Perpetuities and random equations. In *Asymptotic Statistics* (Proc. 5th Prague Symp.), Physica, Heidelberg, pp. 75–86.
[10] Ensor, K. B. and Glynn, P. W. (2000). Simulating the maximum of a random walk. *J. Statist. Planning Infer.* **85,** 127–135.
[11] Fill, J. A. and Huber, M. L. (2010). Perfect simulation of Vervaat perpetuities. *Electron. J. Prob.* **15,** 96–109.
[12] Gjessing, H. K. and Paulsen, J. (1997). Present value distributions with applications to ruin theory and stochastic equations. *Stoch. Process. Appl.* **71,** 123–144.
[13] Goldie, C. M. and Grübel, R. (1996). Perpetuities with thin tails. *Adv. Appl. Prob.* **28,** 463–480.
[14] Harrison, J. M. (1977). Ruin problems with compounding assets. *Stoch. Process. Appl.* **5,** 67–79.
[15] Jelenkovic, P. R. and Olvera-Cravioto, M. (2010). Implicit renewal theory and power tails on trees. Preprint. Available at http://arxiv.org/abs/1006.3295v3.
[16] Kella, O. (2009). On growth-collapse processes with stationary structure and their shot-noise counterparts. *J. Appl. Prob.* **46,** 363–371.
[17] Kendall, W. S. (2004). Geometric ergodicity and perfect simulation. *Electron. Commun. Prob.* **9,** 140–151.
[18] Kesten, H. (1973). Random difference equations and renewal theory for products of random matrices. *Acta Math.* **131,** 207–248.
[19] Paulsen, J. (1998). Sharp conditions for certain ruin in a risk process with stochastic return on investments. *Stoch. Process. Appl.* **75,** 135–148.
[20] Pollack, M. and Siegmund, D. (1985). A diffusion and its applications to detecting a change in the drift of Brownian motion. *Biometrika* **72,** 267–280.
[21] Propp, J. G. and Wilson, D. B. (1996). Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures Algorithms* **9,** 223–252.

https://doi.org/10.1239/jap/1318940463 Published online by Cambridge University Press

[22] MAULIK, K. AND ZWART, B. (2006). Tail asymptotics of exponential functionals of Lévy processes. *Stoch. Process. Appl.* **116,** 156–177.
[23] NYRHINEN, H. (2001). Finite and infinite time ruin probabilities in a stochastic economic environment. *Stoch. Process. Appl.* **92,** 265–285.
[24] VERVAAT, W. (1977). On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. *Adv. Appl. Prob.* **11,** 750–783.

JOSE H. BLANCHET, *Columbia University*

Department of Industrial Engineering and Operations Research, Columbia University, S. W. Mudd Building, 500 West 120th Street, New York, NY 10025, USA.

KARL SIGMAN, *Columbia University*

Department of Industrial Engineering and Operations Research, Columbia University, S. W. Mudd Building, 500 West 120th Street, New York, NY 10025, USA. Email address: karl.sigman@columbia.edu

https://doi.org/10.1239/jap/1318940463 Published online by Cambridge University Press