OPTIMAL RARE EVENT MONTE CARLO FOR MARKOV MODULATED REGULARLY VARYING RANDOM WALKS

Karthyek R. A. Murthy
Sandeep Juneja

Tata Institute of Fundamental Research
Mumbai, 400005, INDIA

Jose Blanchet

IEOR Department
Columbia University
New York, NY 10027-6699, USA

ABSTRACT

Most of the efficient rare event simulation methodology for heavy-tailed systems has concentrated on processes with stationary and independent increments. Motivated by applications such as insurance risk theory, in this paper we develop importance sampling estimators that are shown to achieve asymptotically vanishing relative error property (and hence are strongly efficient) for the estimation of large deviation probabilities in Markov modulated random walks that possess heavy-tailed increments. Exponential twisting based methods, which are effective in light-tailed settings, are inapplicable even in the simpler case of random walk involving i.i.d. heavy-tailed increments. In this paper we decompose the rare event of interest into a dominant and residual component, and simulate them independently using state-independent changes of measure that are both intuitive and easy to implement.

1 INTRODUCTION

In this paper, we develop efficient importance sampling algorithms that employ simple state-independent changes of measure for computing the large deviation probabilities in random walks with heavy-tailed increments that are modulated by an ergodic Markov chain. To be specific, let $S_n := X_1 + \ldots + X_n$. The increments X_n have regularly varying heavy tails and they are modulated by an independent Markov chain $\xi = (\xi_n : n \geq 1)$ (see Section 3 for precise problem statement). We assume that the steady-state drift of the walk is zero. We are interested in the efficient computation of large deviation probabilities $\mathbb{P}\{S_n > b\}$ when n and b are large. Because of the analytical intractability in arriving at exact expressions for such probabilities, and since the asymptotic approximations might have a large error when computed for a finite n, we develop a fast simulation algorithm that computes estimates which stay within a pre-specified relative error with high probability.

Stochastic models involving random walks with heavy-tailed increments have received substantial attention in queuing and insurance risk theories because of their ability to explain long-range dependence in tele-traffic data, and highly variable claim sizes in insurance (see (Resnick 1997), (Embrechts, Klüppelberg, and Mikosch 1997), and (Adler, Feldman, and Taqqu 1998)). Because of their relative simplicity, random walks with independent increments have been analysed well in the literature to obtain various tail asymptotics and efficient simulation algorithms. On the other hand, it has been observed that the less studied case of random walks with dependent increments, particularly the kind of dependence where we have an underlying stochastic process that modulates the increments, are of importance in the study of risk in changing economic environments (see (Delbaen and Haezendonck 1987), (Paulsen 1993) and (Paulsen and Gjessing 1997)) and in the analysis of certain queuing networks (see (Baccelli, Schlegel, and Schmidt 1999) and (Huang and Sigman 1999)). Algorithms for the estimation of the large deviation probabilities $\mathbb{P}\{S_n > b\}$ are particularly important because they act as building block to many complex rare event problems involving combinations

of renewal processes: for examples in queueing, see (Parekh and Walrand 1989) and in financial credit risk modeling, see (Glasserman and Li 2005) and (Bassamboo, Juneja, and Zeevi 2008).

In light-tailed settings, exponential twisting based importance sampling methods, both state-independent (as in (Bucklew, Ney, and Sadowsky 1990)) and dynamic (as in (Dupuis and Wang 2005)), are shown to work for the simulation of large-deviation probabilities involving additive functionals of Markov chains. Whereas in heavy-tailed settings, exponential twisting is inapplicable because of the absence of a neighborhood around zero where the log-moment generating function of the increments is finite; a list of other difficulties in arriving at an importance sampling measure even for the simple case of i.i.d. increments can be found in (Asmussen, Binswanger, and Højgaard 2000). For the case of i.i.d. increments, we have a parametrized state-dependent sampler for the estimation of $\mathbb{P}\{S_n > b\}$ due to (Blanchet and Liu 2008), and two recent state-independent algorithms in (Murthy and Juneja 2012) and (Murthy, Juneja, and Blanchet 2013). In (Blanchet and Liu 2008), the parameters to sample increments based on the current position of the walk is carefully done in every step by ensuring that a certain Lyapunov inequality holds; here the fact that each increment has zero mean has been used crucially to enforce the Lyapunov inequality. However in the case of modulated walks, increments need not have zero mean, and the generalization of such a method does not seem straightforward. The means of overcoming such difficulties by testing for Lyapunov inequalities in more than one time step in the context of simulation of certain level crossing probabilities has been studied in (Lam 2011).

The algorithm that we develop in this paper is optimal in the sense that it achieves asymptotically vanishing relative error (see Section 2.2 for a definition) uniformly throughout the moderate and large deviations regime. We also allow increments to have infinite variance, a situation which can very well happen when dealing with practical applications demonstrating heavy-tailed behaviour. One more interesting aspect of our importance sampling changes of measure is that it does not involve any change in the transition probabilities of the modulating Markov chain, unlike the case of modulated walks with light-tailed increments where exponential twisting is performed to favour certain states over others. As in (Murthy, Juneja, and Blanchet 2013), we partition the event of interest into a dominant and residual component, and prescribe simple, intuitive sampling measures, that are easy to implement and offer a wide scope for generalizing to complex multi-dimensional settings.

The rest of the paper is organized as follows: After a brief discussion of certain preliminary concepts in Section 2, we explain the simulation problem in Section 3, develop our sampling methodology and prove its efficiency in Section 4. We demonstrate the performance of our simulation methods on certain numerical examples in Section 5. Some of the technical proofs are presented in the appendix.

2 PRELIMINARIES

We briefly review the notion of regular varying tails, the use of importance sampling in rare event simulation for the computation of rare event probabilities, and efficiency notions of associated algorithms. Before that we explain certain notations that recur: For any distribution function $F(\cdot)$, we use $\bar{F}(\cdot) := 1 - F(\cdot)$ to denote the tail of $F(\cdot)$. For any random variable X, $CV[X] := \sqrt{\mathrm{Var}[X]}/\mathbb{E}X$ denotes the coefficient of variation of X. For given functions $f : \mathbb{R}^+ \to \mathbb{R}^+$ and $g : \mathbb{R}^+ \to \mathbb{R}^+$, we say $f(x) = O(g(x))$ if there exists $B > 0$ and x_1 large enough such that $f(x) \leq Bg(x)$ for all $x > x_1$; we use $f(x) = o(g(x))$ if $f(x)/g(x) \to 0$, and $f(x) \sim g(x)$ if $f(x)/g(x) \to 1$, as $x \nearrow \infty$.

2.1 Regularly Varying Tails

A function $L : \mathbb{R}^+ \to \mathbb{R}^+$ is said to be *slowly varying* (at infinity) if $\lim_{x \to \infty} L(tx)/L(x) = 1$ for any $t > 0$. Some examples of slowly varying functions include $\log(x), \log\log(x)$ and $1 - e^{-x}$. A random variable X is said to be *regularly varying* with index $-\alpha$ if $\mathbb{P}\{X > x\} = x^{-\alpha}L(x)$ for some slowly varying function $L(\cdot)$. It can be easily verified that for a random variable X with regularly varying tail, $\mathbb{E}[\exp(\theta X)] = \infty$ for any $\theta > 0$. These distribution functions capture the concept of polynomially decaying tails, and they

form an important class of heavy-tailed distributions. See (Embrechts, Klüppelberg, and Mikosch 1997) or (Borovkov and Borovkov 2008) for a detailed account on regular variation.

2.2 Use of Importance Sampling in Rare Event Simulation

Let $(A_n : n \geq 1)$ be a sequence of events in the probability space $(\Omega, \mathscr{F}, \mathbb{P})$ such that $z_n := \mathbb{P}(A_n) \searrow 0$, as the rarity parameter n grows to infinity. Exact computation of the rare event probabilities motivated by practical applications is often impossible, and one resorts to obtaining approximate estimates via Monte-Carlo methods. So given ε and $\delta > 0$, we aim to obtain estimators $\hat{z}_n$ for z_n such that the relative error $|\hat{z}_n - z_n|/z_n$ is at most ε with probability at least $1 - \delta$.

   In importance sampling for rare event simulation, the general objective is to come up with a sequence of changes of measure $(\mathbb{P}_n(\cdot) : n \geq 1)$ from which samples of the random variables $(Z_n : n \geq 1)$ can be obtained that satisfy $\mathbb{E}_n[Z_n] = z_n$; here $\mathbb{E}_n[\cdot]$ denotes the expectation operator associated with the measure $\mathbb{P}_n(\cdot)$. Such a sequence of measures $(\mathbb{P}_n(\cdot) : n \geq 1)$ and the corresponding unbiased estimators $(Z_n : n \geq 1)$ are collectively referred to as an importance sampling algorithm in this paper. Given an unbiased estimator Z_n for z_n, the following procedure yields estimates $\hat{z}_n$ that satisfy the above specified relative precision criterion: Take N i.i.d. replications of Z_n and take $\hat{z}_n$ as their sample mean. If the number of replications N is larger than

$$N(n) = \frac{CV^2[Z_n]}{\delta \varepsilon^2}, \tag{1}$$

we obtain estimators with desired performance due to Chebyshev's inequality as below:

$$\mathbb{P}\left(\frac{|\hat{z}_n - z_n|}{z_n} > \varepsilon \right) \leq \frac{\text{Var}(\hat{z}_n)}{z_n^2 \varepsilon^2} = \frac{CV^2[Z_n]}{N(n)\varepsilon^2} < \delta.$$

The procedure of drawing samples from the original measure $\mathbb{P}(\cdot)$ and having $Z_n = \mathbb{I}(A_n)$ as the estimator is referred to as naive simulation. From (1), it can be seen that naive simulation requires $z_n^{-1} \nearrow \infty$ replications, and hence is not a practically attractive solution. As is well known, the choice $\mathbb{P}_n^*(\cdot) := \mathbb{P}(\cdot|A_n)$ as an importance sampling measure yields zero variance for the associated estimator $Z_n = z_n \mathbb{I}(A_n)$ (see e.g., (Asmussen and Glynn 2007)). Then, every sample obtained in simulation equals z_n with $\mathbb{P}_n^*(\cdot)$ probability 1. However, the explicit dependence of Z_n on z_n, the quantity which we want to estimate, makes this method impractical. Nevertheless it provides useful guidelines in choosing importance sampling change of measures. Motivated by (1), we have the following efficiency notions that quantify the performance of importance sampling algorithms:

Definition 1 The sequence $(Z_n : n \geq 1)$ of unbiased importance sampling estimators is said to achieve

   (a)   *asymptotically vanishing relative error* if $CV[Z_n] = o(1)$,
   (b)   *strong efficiency* if $CV[Z_n] = O(1)$, and
   (c)   *weak efficiency* if for every $\varepsilon > 0$, $\text{Var}[Z_n]/z_n^{2-\varepsilon} = o(1)$, as $n \nearrow \infty$.

   The significance of these definitions can be seen from (1): if an algorithm is strongly efficient, the number of simulation runs required to guarantee the desired relative precision stays bounded even as $n \nearrow \infty$. If $\text{Var}(Z_n) = o\left(z_n^2\right)$, then $(Z_n : n \geq 1)$ satisfies asymptotically vanishing relative error property. As a result, it is enough to generate $o(\delta^{-1}\varepsilon^{-2})$ i.i.d. replications of the estimator. As is apparent from the definition, all strongly efficient algorithms are weakly efficient, and vanishing relative error is the strongest notion among all three. Also it can be verified that naive simulation is not even weakly efficient.

3 THE RARE EVENT SIMULATION PROBLEM

Consider a discrete-time ergodic Markov chain $\xi = \{\xi_n : n \geq 1\}$ taking values in a finite state space $\chi = \{1, 2, \ldots, s\}$. It has a stationary distribution π; that is, for any $i, j \in \chi$, there exists limits:

$$\lim_{n \nearrow \infty} \mathbb{P}\{\xi_n = j | \xi_1 = i\} = \pi_j > 0, \text{ satisfying } \sum_{j=1}^{s} \pi_j = 1.$$

We consider the random walk $(S_n : n \geq 1)$ in $\mathbb{R}$ with

$$S_0 := 0 \text{ and } S_n := X_1 + \ldots + X_n, \text{ for } n \geq 1$$

satisfying the following properties:

(a) The increments $(X_n : n \geq 1)$ are conditionally independent given the Markov chain ξ,
(b) There exists a family of distribution functions $\{F_k(\cdot) : k \in \chi\}$, whose tail probabilities $\bar{F}_k(x) = 1 - F_k(x)$ are regularly varying at infinity, and they satisfy,

$$\mathbb{P}\{X_n > x | \xi\} = \mathbb{P}\{X_n > x | \xi_n\} = \bar{F}_k(x).$$

We allow the regularly varying functions to have different indices and slowly varying functions: for $k \in \chi$, the tails $\bar{F}_k(\cdot)$ are of the form $\bar{F}_k(x) = x^{-\alpha_k} L_k(x), x \in \mathbb{R}$ for slowly varying functions $L_k(\cdot)$ and indices $\alpha_k > 1$.

For $k \in \chi$, name the expectations $\int_{-\infty}^{\infty} x F_k(dx)$ as μ_k and define the steady-state mean $\mu_\pi := \sum_{j=1}^{s} \mu_j \pi_j$. Under these conditions we have, $S_n/n \to \mu_\pi$ a.s. We are interested in the rare event probabilities $\mathbb{P}\{S_n > n\mu_\pi + b\}$, for large values of n and b. To be precise: Define $\alpha := \min\{\alpha_1, \ldots, \alpha_s\}$ and $\beta := (\alpha \wedge 2)^{-1}$. While studying these probabilities, without loss of generality, we can take μ_π to be zero. From (Borovkov and Borovkov 2008), we have the following asymptotics: for any $b > n^{\beta + \varepsilon}$,

$$\mathbb{P}\{S_n \geq b\} \sim n\bar{F}_\pi(b), \text{ as } n \nearrow \infty, \tag{2}$$

where $\bar{F}_\pi(\cdot) = \pi_1 \bar{F}_1(\cdot) + \ldots + \pi_s \bar{F}_s(\cdot)$ is a regularly function with index $\alpha > 1$. Given any $\varepsilon > 0$, we are interested in developing important sampling algorithms for efficiently computing the moderate and large deviation probabilities $\mathbb{P}\{S_n > b\}$, for values of $b > n^{\beta + \varepsilon}$, as $n \nearrow \infty$. We make the following non-restrictive technical assumption:

Assumption 1 If $\int_{-\infty}^{\infty} x^2 \bar{F}_\pi(dx) = \infty$, then $\overline{\lim}_{x \to \infty} \frac{F_\pi(-x)}{\bar{F}_\pi(x)} < \infty$.

This assumption just encodes that in case of increments having infinite variance, the heaviest of the left tails is lighter than the heaviest of the right tails in the family $\{F_k(\cdot) : k \in \chi\}$. If this is not the case, we just take $-\alpha$ to be the the maximum of the indices of regular variation of all the tails (both left and right) in the family, set $\beta = \alpha^{-1}$, and look for estimating probabilities $\mathbb{P}\{S_n > b\}$ with $b > n^{\beta + \varepsilon}$. For example, if the heaviest tail index in the family is $-3/2$, then we consider values of b larger than $n^{2/3 + \varepsilon}$.

4 SIMULATION METHODOLOGY

In random walks involving heavy tails, it has been well understood that the sum becomes large, most likely because, one of the increments take a large value; that is, the partial sum sequence witnesses a large positive jump. This idea, famously referred to as "the big jump principle" has been a recurrent theme in the simulation of rare events in heavy-tailed random walks. To use this appropriately, we partition the event of interest $\{S_n > b\}$ into two components,

$$A_{\text{dom}}(n, b) := \left\{ S_n > b, \max_{k \leq n} \geq b \right\} \text{ and } A_{\text{res}}(n, b) := \left\{ S_n > b, \max_{k \leq n} < b \right\},$$

and estimate their probabilities separately. (Juneja 2007) considered such a decomposition for the estimation of tail probabilities $\mathbb{P}\{S_n > b\}$ as $b \nearrow \infty$ for a fixed n. We also have, as in (Borovkov and Borovkov 2008), that:

$$\mathbb{P}\{S_n > b\} \sim \mathbb{P}\left\{S_n > b, \max_{k \leq n} \geq b\right\}, \text{ as } n \nearrow \infty.$$

Asymptotically the most likely way for S_n to exceed b is via A_{dom}; thus for large values of n and b, $\mathbb{P}(A_{\text{res}})$ is vanishingly small compared to $\mathbb{P}(A_{\text{dom}})$, which explains naming the events as A_{dom} and A_{res}.

4.1 Simulation of A_{dom}

When one of the increments $X_1, \ldots, X_n$ is large, according to the big jump principle, the sum is also large with high probability. For efficiently simulating the event A_{dom}, we should choose an importance sampling measure that induces large positive jumps in the sum with appropriate probabilities that reflect the large deviations behaviour of S_n. Consider the following sampling procedure:

Algorithm (given n and $b > n^{\beta + \varepsilon}$)
Step 1: Simulate a realization $\{\xi_1, \ldots, \xi_n\}$ according to the original dynamics of the Markov chain ξ
Step 2: Set $q \longleftarrow \sum_{i=1}^n \bar{F}_{\xi_i}(b)$
Step 3: Choose $I \in \{1, 2, \ldots, n\}$ such that $\mathbb{P}\{I = i\} = \bar{F}_{\xi_i}(b)/q$.
Step 4: For $j \in \{1, 2, \ldots, n\} - \{I\}$, generate independent samples X_j from $F_{\xi_j}(\cdot)$
Step 5: Generate sample X_I following distribution $\frac{F_{\xi_I}(dx)}{\bar{F}_{\xi_I}(b)}\mathbf{1}(x \geq b)$
Step 6: Set $S \longleftarrow X_1 + \ldots + X_n$ and COUNT $\longleftarrow \#\{i \leq i \leq n : X_i \geq b\}$
Step 7: RETURN $q \cdot \mathbf{1}(S > b)/\text{COUNT}$

Let $\tilde{\mathbb{P}}(\cdot)$ be the probability measure induced in the path space by the above sampling procedure, and let $\tilde{\mathbb{E}}[\cdot]$ denote the associated expectation operator; for brevity, the dependence on n and b has not been highlighted in the notation. Note that $\mathbb{P}(\cdot)$ is absolutely continuous with respect to the importance sampling measure $\tilde{\mathbb{P}}(\cdot)$ when restricted to A_{dom}. The above algorithm might be surprising in the sense that we simulate the modulating Markov chain according to the original dynamics, and not from some other importance sampling transition probabilities that favour sampling the increments often from the heaviest of the tails. But we will show that sampling according to the original dynamics will suffice. We need to argue that the values returned by the above algorithm are unbiased, and have low variance. The sampling procedure induces the following probabilities conditional on the realization of the modulating Markov chain ξ:

$$\tilde{\mathbb{P}}\{X_1 \in dx_1, \ldots, X_n \in dx_n | \xi_1, \ldots, \xi_n\} = \sum_{i=1}^n \mathbb{P}\{I = i\} \cdot \frac{\prod_{j=1}^n \mathbb{P}\{X_j \in dx_j\}}{\bar{F}_{\xi_i}(b)}\mathbf{1}(x_i > b)$$

$$= \sum_{i=1}^n \frac{\bar{F}_{\xi_i}(b)}{q(\xi_1, \ldots, \xi_n)} \cdot \frac{\mathbb{P}\{X_1 \in dx_1, \ldots, X_n \in dx_n\}}{\bar{F}_{\xi_i}(b)}\mathbf{1}(x_i > b),$$

where $q(\xi_1, \ldots, \xi_n) := \sum_{i=1}^n \bar{F}_{\xi_i}(b)$. Therefore conditional on the Markov chain realization $(\xi_1, \ldots, \xi_n)$, the likelihood ratio of $\mathbb{P}(\cdot)$ with respect to $\tilde{\mathbb{P}}(\cdot)$ on the set A_{dom} is $q(\xi_1, \ldots, \xi_n)/\sum_{i=1}^n \mathbb{I}(X_i > b)$. Thus for given n and $b > n^{\beta + \varepsilon}$, the random variable

$$Z_{\text{dom}}(n, b) := \frac{q(\xi_1, \ldots, \xi_n)}{\#\{1 \leq i \leq n : X_i > b\}}\mathbb{I}(A_{\text{dom}})$$

is an unbiased estimator of $\mathbb{P}(A_{\text{dom}})$ under measure $\tilde{\mathbb{P}}(\cdot)$. Here Lemma 1 establishes that the estimator has low variance.

Lemma 1 Given $\varepsilon > 0$, uniformly for $b > n^{\beta + \varepsilon}$,

$$\text{Var}[Z_{\text{dom}}(n,b)] = o\left(\mathbb{P}\{S_n > b\}^2\right), \quad \text{as } n \nearrow \infty.$$

Proof. For every $k \in \chi$, let $N_k := \#\{1 \leq i \leq n : \xi_i = k\}$. Then $q(\xi_1, \ldots, \xi_n) = \sum_{i=1}^n \bar{F}_{\xi_i}(b)$ can alternatively be written as: $q(\xi_1, \ldots, \xi_n) = \sum_{k=1}^s N_k \bar{F}_k(b)$. Further define,

$$Y_n := \frac{\sum_{k=1}^s \frac{N_k}{n} \bar{F}_k(b)}{\bar{F}_\pi(b)}.$$

Since $\#\{1 \leq i \leq n : X_i > b\}$ is at least 1 on A_{dom}, we have $Z_{\text{dom}}(n,b) \leq q(\xi_1, \ldots, \xi_n)$, and hence,

$$\frac{Z_{\text{dom}}(n,b)}{n\bar{F}_\pi(b)} \leq Y_n. \tag{3}$$

Recall that $\bar{F}_\pi(b) = \sum_{k=1}^s \pi_k \bar{F}_k(b)$. Therefore,

$$Y_n = \frac{\sum_{k=1}^s \frac{N_k}{n} \bar{F}_k(b)}{\sum_{k=1}^s \pi_k \bar{F}_k(b)} \leq \frac{\sum_{k=1}^s \bar{F}_k(b)}{\sum_{k=1}^s \pi_{\min} \bar{F}_k(b)} \leq \frac{1}{\pi_{\min}},$$

where $\pi_{\min} := \min\{k \in \chi : \pi_k\}$, which is non-zero because of the ergodicity of the Markov chain ξ. For the same reason, the occupation measures of the Markov chain converge to the stationary distribution π as in: $N_k/n \to \pi_k$, as $n \nearrow \infty$, for all $k \in \{1, 2, \ldots, s\}$. Thus uniformly for all b, $Y_n \to 1$, as $n \nearrow \infty$; and since $|Y_n|$ is uniformly bounded, because of bounded convergence theorem, we have $\tilde{\mathbb{E}}[Y_n^2] \sim 1$, as $n \nearrow \infty$. Then given $\varepsilon > 0$, from (3), for large enough values of n, we have:

$$\tilde{\mathbb{E}}\left[Z_{\text{dom}}^2(n,b)\right] \leq (1+\varepsilon)(n\bar{F}_\pi(b))^2.$$

Since $\tilde{\mathbb{E}}[Z_{\text{dom}}(n,b)] = \mathbb{P}\{S_n > b\} \sim n\bar{F}_\pi(b)$ because of (2), we have the desired result on variance of $Z_{\text{dom}}(n,b)$:

$$\text{Var}[Z_{\text{dom}}(n,b)] = o\left(\mathbb{P}\{S_n > b\}^2\right), \quad \text{as } n \nearrow \infty.$$

$\square$

4.2 Simulation of A_{res}

All the increments $X_1, \ldots, X_n$ are bounded (by b) in the set A_{res}. So we can apply ideas similar to exponential twisting. Given $b > 0$, for every $k \in \chi$, define $\Lambda_k(\theta) := \log\left(\int_{-\infty}^b \exp(\theta x) F_k(dx)\right)$, for $\theta \geq 0$. Consider the following family of 'truncated' and 'exponentially titled' distributions:

$$\hat{F}_k(dx) = \exp\left(\theta_{n,b} x - \Lambda_k(\theta_{n,b})\right) F_k(dx) \mathbf{1}(x \leq b), \quad x \in \mathbb{R}, k \in \chi,$$

with $\theta_{n,b}$ given by,

$$\theta_{n,b} := \frac{-\log(n\bar{F}_\pi(b))}{b}. \tag{4}$$

The sampling procedure will involve simulation of the Markov chain ξ according to its original dynamics as in the simulation of A_{dom}. Now given the realization $\{\xi_1, \ldots, \xi_n\}$ of the modulating process, generate the increments X_i independently according to the exponentially twisted law $\hat{F}_{\xi_i}(\cdot)$, for $i = 1, \ldots, n$. This will result in the following unbiased estimator for $\mathbb{P}(A_{\text{res}})$:

$$Z_{\text{res}}(n,b) := \exp\left(-\theta_{n,b} S_n + \sum_{i=1}^n \Lambda_{\xi_i}(\theta_{n,b})\right) \mathbb{I}(A_{\text{res}}).$$

Algorithm (given n and $b > n^{\beta+\varepsilon}$)
Step 1: Simulate a realization $\{\xi_1, \ldots, \xi_n\}$ according to the original dynamics of the Markov chain ξ
Step 2: Set $\theta_{n,b} \longleftarrow -\log(n\bar{F}_\pi(b))/b$ and $C \longleftarrow \Lambda_{\xi_1}(\theta_{n,b}) + \ldots + \Lambda_{\xi_n}(\theta_{n,b})$
Step 3: For $j = 1, \ldots, n$, generate independent samples X_j from $\hat{F}_{\xi_j}(\cdot)$
Step 4: Actualize $S \longleftarrow X_1 + \ldots + X_n$
Step 5: RETURN $\exp(-\theta_{n,b}S + C)\mathbf{1}(S \geq b)$

Let $\hat{\mathbb{P}}(\cdot)$ be the measure induced by drawing samples according to the above algorithm and let $\hat{\mathbb{E}}[\cdot]$ be the expectation operator associated with $\hat{\mathbb{P}}(\cdot)$. Lemma 2, which is proved in the appendix, gives an upper bound on the normalizing constants $\exp(\Lambda_k(\theta_{n,b}))$.

Lemma 2 Given $\varepsilon > 0$, uniformly for $b > n^{\beta+\varepsilon}$ and $k \in \chi$,

$$\exp(\Lambda_k(\theta_{n,b})) \leq 1 + \theta_{n,b}\mu_k + \frac{\bar{F}_k(b)}{n\bar{F}_\pi(b)}(1 + o(1)), \text{ as } n \nearrow \infty.$$

Lemma 3 Given $\varepsilon > 0$, uniformly for $b > n^{\beta+\varepsilon}$,

$$\text{Var}[Z_{\text{res}}(n,b)] = o\left(\mathbb{P}\{S_n > b\}^2\right), \text{ as } n \nearrow \infty.$$

Proof. Since $S_n > b$ on A_{res} and $b\theta_{n,b} = -\log(n\bar{F}_\pi(b))$,

$$Z_{\text{res}}(n,b) \leq \exp\left(-\theta_{n,b}b + \sum_{i=1}^n \Lambda_{\xi_i}(\theta_{n,b})\right)\mathbb{I}(A_{\text{res}}) \leq n\bar{F}_\pi(b)\exp\left(\sum_{i=1}^n \Lambda_{\xi_i}(\theta_{n,b})\right)\mathbb{I}(A_{\text{res}}).$$

The second moment can be evaluated as below:

$$\hat{\mathbb{E}}\left[Z_{\text{res}}^2(n,b)\right] = \mathbb{E}\left[Z_{\text{res}}(n,b)\right] \leq n\bar{F}_\pi(b)\mathbb{E}\left[\exp\left(\sum_{i=1}^n \Lambda_{\xi_i}(\theta_{n,b})\right); A_{\text{res}}\right]. \tag{5}$$

As before, define $N_k := \sum_{i=1}^n \mathbb{I}(\xi_i = k)$, for $k \in \chi$. Then $\sum_{i=1}^n \Lambda_{\xi_i}(\theta_{n,b}) = \sum_{k=1}^s N_k\Lambda_k(\theta_{n,b})$. Now due to Jensen's inequality,

$$\exp\left(\sum_{i=1}^n \Lambda_{\xi_i}(\theta_{n,b})\right) = \exp\left(n\sum_{k=1}^s \frac{N_k}{n}\Lambda_k(\theta_{n,b})\right) \leq \left(\sum_{k=1}^s \frac{N_k}{n}\exp(\Lambda_k(\theta_{n,b}))\right)^n.$$

Given $\delta > 0$, from Lemma 2 we have, $\exp(\Lambda_k(\theta_{n,b})) \leq 1 + \theta_{n,b}\mu_k + (1+\delta)\frac{\bar{F}_k(b)}{n\bar{F}_\pi(b)}$, for large values of n, and $k \in \chi$. We have also used that $b\theta_{n,b} = -\log(n\bar{F}_\pi(b))$. Since $N_1 + \ldots + N_s = n$,

$$\sum_{k=1}^s \frac{N_k}{n}\exp(\Lambda_k(\theta_{n,b})) \leq 1 + \theta_{n,b}\sum_{k=1}^s \frac{N_k}{n}\mu_k + \frac{1+\delta}{n\bar{F}_\pi(b)}\sum_{k=1}^s \frac{N_k}{n}\bar{F}_k(b).$$

$$\leq 1 + \theta_{n,b}\sum_{k=1}^s \left(\frac{N_k}{n} - \pi_k\right)\mu_k + \frac{1+\delta}{n\bar{F}_\pi(b)}\sum_{k=1}^s \left(\frac{N_k}{n} - \pi_k\right)\bar{F}_k(b) + \frac{1+\delta}{n},$$

because $\sum_{k=1}^s \pi_k\mu_k = \mu_\pi = 0$ and $\sum_{k=1}^s \pi_k\bar{F}_k(b) =: \bar{F}_\pi(b)$. Since $1 + x \leq \exp(x)$,

$$\left(\sum_{k=1}^s \frac{N_k}{n}\exp(\Lambda_k(\theta_{n,b}))\right)^n \leq \exp\left(n\theta_{n,b}\sum_{k=1}^s \left(\frac{N_k}{n} - \pi_k\right)\mu_k + \frac{1+\delta}{\bar{F}_\pi(b)}\sum_{k=1}^s \left(\frac{N_k}{n} - \pi_k\right)\bar{F}_k(b) + 1 + \delta\right).$$

For $\delta' > 0$, define sets $B_0 := \bigcup_{k \in \chi} \{N_k \leq n\pi_k + n^{1/2+\delta'}\}$ and $B_j := \bigcup_{k \in \chi} \{2^{j-1}n^{1/2+\delta'} < N_k - n\pi_k \leq +2^j n^{1/2+\delta'}\}, j \geq 1$. We also have, $\bar{F}_\pi(b) \geq \pi_{\min} \sum_k \bar{F}_k(b)$. Then from (5),

$$\hat{\mathbb{E}}\left[Z_{\text{res}}^2(n,b)\right] \leq n\bar{F}_\pi(b) \exp\left((1+\delta)\left(1+\frac{1}{\pi_{\min}}\right)\right) \sum_{j\geq 0} \exp\left(2^j n^{1/2+\delta'} \theta_{n,b}\right) \mathbb{P}(B_j \cap A_{\text{res}}). \qquad (6)$$

Since the ergodic chain ξ takes values in a finite state space, we have for any $j \geq 1, \mathbb{P}(B_j) \leq c \exp(-2^{2j-1}n^{2\delta'})$ for some constant c (Hoeffding-type inequality). Since $b > n^{\beta+\varepsilon} \geq n^{1/2+\varepsilon}$, we have $n^{1/2+\delta'} \theta_{n,b} = o(1)$ if $\delta' < \varepsilon$. Therefore,

$$\sum_{j\geq 0} \exp\left(2^j n^{1/2+\delta'} \theta_{n,b}\right) \mathbb{P}(B_j \cap A_{\text{res}}) \leq \exp\left(n^{1/2+\delta'} \theta_{n,b}\right) \mathbb{P}(A_{\text{res}}) + c \sum_{j\geq 1} \exp\left(-2^{2j-1}n^{2\delta'} + 2^j n^{1/2+\delta'} \theta_{n,b}\right)$$

$$\leq \mathbb{P}(A_{\text{res}})(1+o(1)) + O(\exp(-n^{2\delta'})).$$

Combining this with (6) and the asymptotics that $\mathbb{P}\{S_n > b\} \sim \mathbb{P}\{S_n > b, M_n \geq b\} \sim n\bar{F}_\pi(b)$ as $n \nearrow \infty$, we have: $\text{Var}[Z_{\text{res}}(n,b)] = o\left(\mathbb{P}\{S_n > b\}^2\right)$. $\qquad \square$

Now we have Theorem 1 that comments about the efficiency of the overall estimation procedure.

Theorem 1 If the realizations of the estimators Z_{dom} and Z_{res} are generated respectively from the measures $\tilde{\mathbb{P}}(\cdot)$ and $\hat{\mathbb{P}}(\cdot)$, and if we let,
$$Z(n,b) := Z_{\text{dom}}(n,b) + Z_{\text{res}}(n,b),$$
then under Assumption 1, the family of estimators $(Z(n,b) : n \geq 1, b > n^{\beta+\varepsilon})$ achieves asymptotically vanishing relative error for the estimation of $\mathbb{P}\{S_n > b\}$, as $n \nearrow \infty$; that is,

$$\frac{\text{Var}[Z(n,b)]}{(\mathbb{P}\{S_n > b\})^2} = o(1),$$

as $n \nearrow \infty$, uniformly for $b > n^{\beta+\varepsilon}$.

Proof. Since the realizations of Z_{dom} and Z_{res} are generated independent of each other, the variance of Z is just the sum of variances of Z_{dom} and Z_{res}; the proof is now evident from Lemmas 1, 3 and Equation (2). $\qquad \square$

A consequence of the above theorem is that, due to (1), the number of i.i.d. replications of $Z(n,b)$ required to achieve ε-relative precision with probability at least $1 - \delta$ is at most $o(\varepsilon^{-2}\delta^{-1})$, independent of the rarity parameters n and b. In our algorithms, each replication demands $O(n)$ computational effort, thus requiring a overall computational cost of $O(n)$, as $n \nearrow \infty$.

5 NUMERICAL EXPERIMENTS

We present the results of numerical simulations for computing large deviation probabilities $\mathbb{P}\{S_n > b\}$ using our algorithms for the following examples:

Example 1: Consider the time-homogeneous Markov chain ξ taking values in $\{0,1\}$ with transition probabilities given by: $\mathbb{P}\{\xi_{n+1} = 1 - i | \xi_n = i\} = 2/3$, for $i = 0,1$. The tail probabilities of the corresponding increment when the Markov chain is in state i is given by, $\mathbb{P}\{X_n > x | \xi_n = i\} = \bar{F}_i(x) = (17/12 + x)^{-\alpha_i}$, for $x \geq -5/12$, with $\alpha_0 = 3$ and $\alpha_1 = 4$; Consider the random walk $(S_n : n \geq 0)$ with $S_n = X_1 + \ldots + X_n$, whose increments are modulated by the Markov chain ξ. It can be verified that the stationary distribution π of the Markov chain ξ is $\pi_i = 0.5, i \in \{0,1\}$, and that the steady-state drift of the random walk S_n is 0. In Table 1, we present the results of our estimation procedure for $N = 10000$ simulation runs for the computation of $\mathbb{P}\{S_n > n\}$ for values of $n = 100, 500$ and 1000. From 1, it can be seen that the empirically observed

Table 1: Numerical result for Example 1 - here Std. error denotes the standard deviation of the estimator of $\mathbb{P}\{S_n > n\}$ based on 10,000 simulation runs; CV denotes the empirically observed coefficient of variation

n	Asymptotic expression $n\bar{F}_{\pi}(b)$	Proposed estimator $(\hat{z})$ for $\mathbb{P}\{S_n > n\}$	Std. error	CV of $\hat{z}$
100	4.84×10^{-5}	4.64×10^{-5}	7.48×10^{-8}	0.16
500	1.99×10^{-6}	1.95×10^{-6}	2.79×10^{-9}	0.14
1000	4.98×10^{-7}	4.92×10^{-7}	6.04×10^{-10}	0.12

coefficient of variation of the estimates is small, and it decreases as n increases.

Example 2: To highlight the superior numerical performance of our state-independent algorithms, we consider the simple case of i.i.d. increments and compare it with existing algorithms in the literature. Take $X = \Lambda R$, where $\mathbb{P}\{\Lambda > x\} = 1 \wedge x^{-4}, R \sim \text{Laplace}(1)$, and Λ is independent of R. Consider the random walk $(S_n : n \geq 0)$ with independent increments having the same distribution as that of X. We again use $N = 10,000$ simulation runs to estimate $\mathbb{P}\{S_n > n\}$ for $n = 100, 500$ and 1000. In Table 2, we compare the numerical estimates obtained by our simulation procedure with the true values of $\mathbb{P}\{S_n > n\}$ evaluated in (Blanchet and Liu 2008) via inverse transform techniques; further, a comparison of performance of our methodology with Algorithms 1 and 2 in (Blanchet and Liu 2008) (referred to as BL1 and BL2) has also been presented. From the columns CV, CV of BL1, and CV of BL2, it can be inferred that our state-independent simulation procedures yield estimators with substantially lower coefficient of variation throughout the range of values considered.

Table 2: Numerical result for Example 2 - here Std. error denotes the standard deviation of the estimator of $\mathbb{P}\{S_n > n\}$ based on 10,000 simulation runs; CV denotes the empirically observed coefficient of variation

n	$\mathbb{P}\{S_n > n\}$	Proposed Estimator $(\hat{z})$ for $\mathbb{P}\{S_n > n\}$	Std. error	CV of $\hat{z}$	CV of BL1	CV of BL2
100	2.21×10^{-5}	2.17×10^{-5}	4.31×10^{-7}	1.97	10.3	4.7
500	1.04×10^{-7}	1.05×10^{-7}	6.91×10^{-10}	0.66	1.0	4.1
1000	1.25×10^{-8}	1.29×10^{-8}	6.91×10^{-11}	0.53	1.1	3.8

6 Conclusions

In this paper we have presented the first provably efficient rare event simulation estimator for large deviations probabilities in the context of heavy-tailed Markov random walks. In contrast to recent estimators that have been developed, using state-dependent importance sampling, and Lyapunov inequalities, our estimator is based on static mixtures of state-independent importance sampling distributions. These mixtures are obtained by decomposing the large deviations event of interest into components that can be treated easily either by the so-called principle of the big jump or by applying standard exponential tilting to suitably truncated random variables. We believe that the principles developed in this paper have further applicability beyond our current settings.

A APPENDIX

Here we prove Lemma 2 after establishing the following sequence of results.

Lemma 4 The distribution function $F(\cdot)$ is such that $\bar{F}(x) := 1 - F(x) = x^{-\alpha}L(x)$ for some slowly varying function $L(\cdot)$ and $\alpha > 1$. Let $\mu := \int_{-\infty}^{\infty} xF(dx)$. Then for any pair of sequences $\{x_n\}, \{\phi_n\}$ satisfying $x_n \nearrow \infty$

and $\phi_n x_n \nearrow \infty$, the integral,

$$\int_{-\infty}^{x_n} e^{\phi_n x} F(dx) \leq 1 + \phi_n \mu + c\phi_n^\kappa + e^{2\alpha} \bar{F}\left(\frac{2\alpha}{\phi_n}\right) + e^{\phi_n x_n} \bar{F}(x_n)(1 + o(1)), \text{ as } n \nearrow \infty,$$

for any $0 < \kappa < \alpha \wedge 2$, and some constant c which does not depend on n and b.

Proof. Let $I_1 := \int_{-\infty}^{\gamma/\phi_n} e^{\phi_n x} F(dx)$ and $I_2 := \int_{\gamma/\phi_n}^{x_n} e^{\phi_n x} F(dx)$. Then we have,

$$\begin{aligned}
I_1 &\leq \int_{-\infty}^{\gamma/\phi_n} F(dx) + \phi_n \int_{-\infty}^{\gamma/\phi_n} xF(dx) + \phi_n^\kappa \int_{-\infty}^{\gamma/\phi_n} |x|^\kappa e^{\phi_n x} F(dx) \\
&\leq \int_{-\infty}^{\infty} F(dx) + \phi_n \int_{-\infty}^{\infty} xF(dx) + \phi_n^\kappa e^\gamma \int_{-\infty}^{\infty} |x|^\kappa F(dx) \\
&= 1 + \phi_n \mu + c\phi_n^\kappa,
\end{aligned} \tag{7}$$

where $c := e^\gamma \int_{-\infty}^{\infty} |x|^\kappa F(dx) < \infty$ because $\mathbb{E}|X|^\kappa < \infty$, which in turn is true because $\kappa < \alpha$ and from Assumption 1. We have also used $\mathbb{E}X = 0$ to arrive at (7). Integrating by parts for the second integral I_2:

$$\begin{aligned}
I_2 &= -\int_{\gamma/\phi_n}^{x_n} e^{\phi_n x} \bar{F}(dx) = e^{\phi_n \gamma/\phi_n} \bar{F}(\gamma/\phi_n) - e^{\phi_n x_n} \bar{F}(x_n) + \phi_n \int_{\gamma/\phi_n}^{x_n} e^{\phi_n x} \bar{F}(x)dx \\
&\leq e^\gamma \bar{F}(\gamma/\phi_n) + I_2',
\end{aligned} \tag{8}$$

where, $I_2' := \phi_n \int_{\gamma/\phi_n}^{x_n} e^{\phi_n x} \bar{F}(x)dx$. Now changing variable to $u = \phi_n(x_n - x)$ results in:

$$I_2' = e^{\phi_n x_n} \int_0^{\phi_n x_n - \gamma} e^{-u} \bar{F}\left(x_n - \frac{u}{\phi_n}\right) du = e^{\phi_n x_n} \bar{F}(x_n) \int_0^{\phi_n x_n - \gamma} e^{-u} g_n(u) du, \tag{9}$$

where $g_n(u) := \bar{F}(x_n - u/\phi_n)/\bar{F}(x_n) = \bar{F}(x_n(1 - u/\phi_n x_n))/\bar{F}(x_n)$. Since $L(\cdot)$ is slowly varying and $\phi_n x_n \to \infty$, given any $\delta > 0$, for all n large enough we have:

$$(1 - \delta)\left(1 - \frac{u}{\phi_n x_n}\right)^{-\alpha+\delta} \leq g_n(u) \leq (1 + \delta)\left(1 - \frac{u}{\phi_n x_n}\right)^{-\alpha-\delta}.$$

This preliminary fact about slowly varying functions can be found in, e.g., Theorem 1.1.4 of (Borovkov and Borovkov 2008). So for any fixed u, we have $g_n(u) \to 1$ as $n \nearrow \infty$. Now fix $\delta = \frac{\alpha}{2}$. Then for n large enough, $g_n(u) \leq (1 + \alpha/2)(1 - u/\phi_n x_n)^{-\frac{3\alpha}{2}}$. Let $h(u) = (1 - u/\phi_n x_n)^{-\frac{3\alpha}{2}}$. Since $\log h(0) = 0$ and $\frac{d}{du}(\log(h(u)) \leq \frac{3\alpha}{2\gamma}$ for $0 \leq u \leq \phi_n x_n - \gamma$, we have $h(u) \leq e^{\frac{3\alpha u}{2\gamma}}$ on the same interval. Therefore if we choose $\gamma = 2\alpha$, the integrand in I_2' is bounded for large enough n by an integrable function as below:

$$\left|e^{-u} g_n(u) \mathbf{1}_{\{0 \leq u \leq \phi_n x_n - \gamma\}}\right| \leq \left|e^{-u}\left(1 + \frac{\alpha}{2}\right) h(u) \mathbf{1}_{\{0 \leq u \leq \phi_n x_n - \gamma\}}\right| \leq \left(1 + \frac{\alpha}{2}\right) e^{-u + \frac{3\alpha u}{2\gamma}} = \left(1 + \frac{\alpha}{2}\right) e^{-\frac{u}{4}}.$$

Applying dominated convergence theorem, we get $\int_0^{\phi_n x_n - \gamma} e^{-u} g_n(u) du \sim 1$ as $n \nearrow \infty$. Since $\int_{-\infty}^{x_n} e^{\phi_n x} F(dx) = I_1 + I_2$, combining this result with (7), (8) and (9), completes the proof. $\square$

Lemma 5 Given any $\varepsilon > 0$, for $\theta_{n,b}$ as in (4), uniformly for $b > n^{\beta+\varepsilon}$ and $k \in \chi$,

(a) $n\theta_{n,b}^\kappa \searrow 0$ for some $\kappa \in (0, \alpha \wedge 2)$, and (b) $\bar{F}_k\left(\frac{2\alpha}{\theta_{n,b}}\right) = O\left(\frac{1}{n}\right)$, as $n \nearrow \infty$.

Proof. (a) We have $\bar{F}_\pi(x) = x^{-\alpha} L_\pi(x)$, for some slowly varying function $L_\pi(\cdot)$. Given any $\delta > 0$ for sufficiently large values of b, we have $b^{-\delta} \le L_\pi(b) \le b^\delta$ because of the slowly varying nature of $L_\pi(\cdot)$. Therefore we have $L_\pi(b) = b^{o(1)}$ as $b \nearrow \infty$. Further noting that $b > n^{\beta+\varepsilon}$ helps us to write:

$$n\theta_{n,b}^\kappa = \frac{n}{b^\kappa} \log^\kappa \left(\frac{1}{n\bar{F}(b)} \right) \le n^{1-\kappa(\beta+\varepsilon)} \log^\kappa \left(\frac{b^\alpha}{nL_\pi(b)} \right).$$

$$\text{Take } \kappa := \begin{cases} 2, & \text{if } \alpha > 2 \\ (1+\varepsilon)/(\frac{1}{\alpha}+\varepsilon), & \text{if } 1 < \alpha \le 2. \end{cases} \tag{10}$$

Then $\kappa < \alpha$, and $\kappa(\beta+\varepsilon) \ge 1+\varepsilon/2$. Then $n\theta_{n,b}^\kappa \searrow 0$ as $n \nearrow \infty$, uniformly for $b > n^{\beta+\varepsilon}$.

(b) Since $\theta_{n,b} = -\log(n\bar{F}_\pi(b))/b$, and $\bar{F}_k(\cdot)$ is regularly varying, given any $\delta > 0$, for n large enough,

$$\frac{\bar{F}_k\left(\frac{2\alpha}{\theta_{n,b}}\right)}{\bar{F}_k(b)} = \frac{\bar{F}_k\left(\frac{2\alpha b}{-\log(n\bar{F}_\pi(b))}\right)}{\bar{F}_k(b)} \le \left(\frac{-\log(n\bar{F}_\pi(b))}{2\alpha} \right)^{\alpha+\delta}.$$

The above inequality is just an application of Theorem 1.1.4 of (Borovkov and Borovkov 2008). Therefore,

$$n\bar{F}_k\left(\frac{2\alpha}{\theta_{n,b}}\right) \le n\frac{L_k(b)}{b^{\alpha_k}} \left(\frac{-\log(n\bar{F}_\pi(b))}{2\alpha} \right)^{\alpha+\delta} = o(1), \text{ uniformly for } b > n^{\beta+\varepsilon} \text{ as } n \nearrow \infty.$$

Here the convergence to 0 is justified because $\alpha_k \ge \alpha$ and $b > n^{\beta+\varepsilon}$. $\qquad\square$

Proof of Lemma 2

From the definition of $\Lambda_k(\cdot)$ and Lemma 4, we have:

$$\exp(\Lambda_k(\theta_{n,b})) = \int_{-\infty}^b \exp(\theta_{n,b}x) F_k(dx)$$

$$\le 1 + \theta_{n,b}\mu_k + c\theta_{n,b}^\kappa + e^{2\alpha}\bar{F}_k\left(\frac{2\alpha}{\theta_{n,b}}\right) + \exp(\theta_{n,b}b)\bar{F}_k(b)(1+o(1)),$$

for κ as in (10). Usage of Lemma 4 is justified because $\theta_n b = -\log(n\bar{F}_\pi(b)) \nearrow \infty$. From Lemma 5, we have $n\theta_{n,b}^\kappa = o(1)$ and $\bar{F}_k\left(\frac{2\alpha}{\theta_{n,b}}\right) = o\left(\frac{1}{n}\right)$, uniformly for $b > n^{\beta+\varepsilon}$. Therefore,

$$\exp(\Lambda_b(\theta_n)) \le 1 + \theta_{n,b}\mu_k + \frac{\bar{F}_k(b)}{n\bar{F}_\pi(b)}(1+o(1)), \text{ as } n \nearrow \infty,$$

thus proving the result. $\qquad\square$

REFERENCES

Adler, R. J., R. E. Feldman, and M. S. Taqqu. 1998. *A practical guide to heavy tails*. Boston, MA: Birkhäuser Boston Inc. Statistical techniques and applications, Papers from the workshop held in Santa Barbara, CA, December 1995.

Asmussen, S., K. Binswanger, and B. Højgaard. 2000. "Rare events simulation for heavy-tailed distributions". *Bernoulli* 6 (2): 303–322.

Asmussen, S., and P. W. Glynn. 2007. *Stochastic simulation: algorithms and analysis*, Volume 57 of *Stochastic Modelling and Applied Probability*. New York: Springer.

Baccelli, F., S. Schlegel, and V. Schmidt. 1999. "Asymptotics of stochastic networks with subexponential service times". *Queueing Systems* 33 (1-3): 205–232.

Bassamboo, A., S. Juneja, and A. J. Zeevi. 2008. "Portfolio Credit Risk with Extremal Dependence: Asymptotic Analysis and Efficient Simulation". *Operations Research* 56 (3): 593–606.

Blanchet, J. H., and J. Liu. 2008. "State-dependent importance sampling for regularly varying random walks". *Adv. in Appl. Probab.* 40 (4): 1104–1128.

Borovkov, A. A., and K. A. Borovkov. 2008. *Asymptotic analysis of random walks*, Volume 118 of *Encyclopedia of Mathematics and its Applications*. Cambridge: Cambridge University Press. Heavy-tailed distributions, Translated from the Russian by O. B. Borovkova.

Bucklew, J. A., P. Ney, and J. S. Sadowsky. 1990. "Monte Carlo Simulation and Large Deviations Theory for Uniformly Recurrent Markov Chains". *Journal of Applied Probability* 27 (1): pp. 44–59.

Delbaen, F., and J. Haezendonck. 1987. "Classical risk theory in an economic environment". *Insurance: Mathematics and Economics* 6 (2): 85 – 116.

Dupuis, P., and H. Wang. 2005. "Dynamic importance sampling for uniformly recurrent Markov chains". *Ann. Appl. Probab.* 15 (1A): 1–38.

Embrechts, P., C. Klüppelberg, and T. Mikosch. 1997. *Modelling extremal events*, Volume 33 of *Applications of Mathematics (New York)*. Berlin: Springer-Verlag. For insurance and finance.

Glasserman, P., and J. Li. 2005. "Importance Sampling for Portfolio Credit Risk". *Management Science* 51 (11): 1643–1656.

Huang, T., and K. Sigman. 1999. "Steady-state asymptotics for tandem, split-match and other feedforward queues with heavy tailed service". *Queueing Systems* 33 (1-3): 233–259.

Juneja, S. 2007. "Estimating tail probabilities of heavy tailed distributions with asymptotically zero relative error". *Queueing Syst.* 57 (2-3): 115–127.

Lam, H. 2011. *Efficient Monte Carlo estimation and asymptotic analysis for stochastic systems*. Ph. D. thesis, Harvard University.

Murthy, K. R. A., and S. Juneja. 2012, oct.. "State-independent importance sampling for estimating large deviation probabilities in heavy-tailed random walks". In *Performance Evaluation Methodologies and Tools (VALUETOOLS), 2012 6th International Conference on*, 127 –135.

Murthy, K. R. A., S. Juneja, and J. Blanchet. 2013. "State-independent Importance Sampling for Random Walks with Independent Increments". *Preprint*.

Parekh, S., and J. Walrand. 1989. "A quick simulation method for excessive backlogs in networks of queues". *IEEE Trans. Automat. Control* 34 (1): 54–66.

Paulsen, J. 1993. "Risk theory in a stochastic economic environment". *Stochastic Processes and their Applications* 46 (2): 327 – 361.

Paulsen, J., and H. K. Gjessing. 1997. "Ruin Theory with Stochastic Return on Investments". *Advances in Applied Probability* 29 (4): pp. 965–985.

Resnick, S. I. 1997. "Heavy tail modeling and teletraffic data". *Ann. Statist.* 25 (5): 1805–1869. With discussion and a rejoinder by the author.

AUTHOR BIOGRAPHIES

KARTHYEK R. A. MURTHY is a Ph.D. student in the School of Technology and Computer Science at Tata Institute of Fundamental Research, Mumbai, India. He holds a Bachelor of Engineering from College of Engineering Guindy, Chennai. His research interests revolves around Applied Probability and Optimization. His email address is kamurthy@tifr.res.in.

SANDEEP JUNEJA is a Professor in the School of Technology and Computer Science at Tata Institute of Fundamental Research. His Ph.D. was from Department of Operations Research at Stanford University (1993). His research interests include applied probability, computational finance, simulation and rare-event analysis. He serves on the editorial board on Mathematics of Operations Research.

JOSE BLANCHET is a faculty member of the IEOR Department at Columbia University. Jose holds a Ph.D. in Management Science and Engineering from Stanford University. Prior to joining Columbia he was a faculty member in the Statistics Department at Harvard University. Jose is a recipient of the 2009 Best Publication Award given by the INFORMS Applied Probability Society and of the 2010 Erlang Prize. He also received a PECASE award given by NSF in 2010. He worked as an analyst in Protego Financial Advisors, a leading investment bank in Mexico. He has research interests in applied probability and Monte Carlo methods. He serves in the editorial board of Advances in Applied Probability, Journal of Applied Probability, Mathematics of Operations Research and QUESTA.