State-independent Importance Sampling for Estimating Large Deviation Probabilities in Heavy-tailed Random Walks

Karthyek Rajhaa A. M.
Tata Institute of Fundamental Research
Mumbai
Email: kamurthy@tifr.res.in

Sandeep Juneja
Tata Institute of Fundamental Research
Mumbai
Email: juneja@tifr.res.in

***Abstract*—Efficient simulation of rare events involving sums of heavy-tailed random variables has been an active research area in applied probability over the last fifteen years. These rare events arise in many applications including telecommunications, computer and communication networks, insurance and finance. These problems are viewed as challenging, since large deviations theory inspired and exponential twisting based importance sampling distributions that work well for rare events involving sums of light tailed random variables fail in these settings. Moreover, there exist negative results suggesting that state-independent importance sampling methods that work well in light-tailed settings fail for certain rare events involving sums of heavy-tailed random variables. This has led to the development of growing literature for efficiently simulating such events using more nuanced, and in many cases, computationally demanding state-dependent importance sampling methods. In this article we shed new light on this issue by observing that simpler state-independent exponential twisting based importance sampling methods, suitably adjusted in the tails, can provide strongly efficient algorithms to estimate such rare event probabilities. Specifically, we develop strongly efficient state-independent importance sampling algorithms for the classical large deviations probability that sums of independent, identically distributed random variables with regularly varying tails exceed an increasing threshold both in the case where the number of random variables increases to infinity and when it is fixed.**

I. Introduction

In this article we consider the problem of efficient simulation via importance sampling of large deviation probabilities $\mathbb{P}\{S_n > b\}$ for $b \geq \tilde{c}n^{\frac{1}{2}+\epsilon}$ given $\tilde{c}$ and $\epsilon > 0$ as $n \nearrow \infty$. Here $S_n = \sum_{i=1}^n X_i$ is the sum of independent identically distributed (i.i.d) zero mean random variables with regularly varying right tails (roughly speaking, this means that the right tails decay as a power law; see II-A for a precise definition). For an introduction to the use of importance sampling in rare event simulation see, e.g., [1] Chapters V and VI, and [2]. Over the last few years, elegant state-dependent importance sampling algorithms have been proposed in [3], [4] and [5] to efficiently simulate these large-deviation probabilites; state-dependence essentially means that the importance sampling distribution for generating X_i depends on the realized values of $X_1, \ldots, X_{i-1}$ (typically through S_{i-1}); state-independence on the other hand implies that such dependence does not exist and samples of $X_1, \ldots, X_n$ can be drawn independently. State-independent methods enjoy obvious advantages over state-dependent ones in terms of complexity of generating samples and ease of implementation. When the running time requirement is stringent, samples can even be generated offline before hand, which may be difficult for state-dependent methods. Our key contribution is that we show that proposed simpler state-independent importance sampling algorithms can also provide strongly efficient algorithms for computing such large deviation probabilities involving random variables (rv) with regularly varying tails. The general folklore in the rare event simulation community is that exponentially twisted distributions may be used to estimate such probabilities when the right tails of the constituent X_i's are light*, but these methods fail when the underlying tails are instead heavy. Our another contribution is that we show that by suitably adjusting along the right tail, state-independent exponential twisting can be used to efficiently estimate $\mathbb{P}\{S_n > b\}$ when X_is have regularly varying right tails.

The problem of efficient estimation of $\mathbb{P}\{S_n > b\}$ as $n \nearrow \infty$ when X_is are light-tailed as well as heavy-tailed has received considerable attention in the literature (see [6], [7], [8], [9], and [10] for literature on light-tailed random walks, [4] and [5] for the right heavy-tailed ones), mainly because this forms a building block to many more complex rare event problems involving combination of renewal processes. For examples in queueing, see [11], and in financial credit risk modeling, see [12] and [13].

This problem was first addressed in [6] where they considered X_is that were light-tailed. They used an importance sampling measure obtained by exponentially twisting the original distribution of each X_i appropriately to arrive at a weakly efficient algorithm for estimating $P(S_n > na)$ for $a > 0$ (the notions of weakly efficient and strongly efficient estimation methods for rare event probabilities are standard; these are

*A rv X is said to have a heavy right (resp., left) tail if $\mathbb{E}[e^{\theta X}] = \infty$ for all positive (negative) values of θ. Otherwise it is said to have a light right (resp., left) tail. For presentation convenience, in this paper unless otherwise specified, by light-tailed rv we mean right light-tailed rv, and by heavy-tailed rv we mean right heavy-tailed rv.

VALUETOOLS 2012, October 09-12, Cargèse, France

DOI 10.4108/valuetools.2012.250304

reviewed later in Section II-B). This methodology does not carry over to heavy-tailed random variables as arriving at appropriate exponentially twisted distributions assumes the existence of the moment generating function of the original random variable in a right neighborhood of zero. Reference [14] provides an account of why large deviations based methods for approximating zero-variance measure may fail in a heavy-tailed setting.

In the heavy-tailed setting, while state-independent importance sampling distributions have been proposed for estimating probabilities such as $\mathbb{P}(S_n > b)$ in the asymptotic regime where n is fixed and $b \nearrow \infty$ (see [15], [14], [16], [17], [18] and [19]), for problems that involve large number of heavy-tailed random variables, increasingly researchers resort to using state-dependent importance sampling (see, e.g., [18], [20], [4] and [5]). This approach is also supported by [21], where for certain rare probabilities involving heavy-tailed random walks, it is shown that state-independent importance sampling algorithms cannot be weakly efficient. We refer the reader to [22] for an excellent survey of state-dependent importance sampling methods.

As mentioned earlier, our key aim is to show that state-independent importance sampling algorithm based on the proposed 'tail-adjusted' exponential twisting provides strongly efficient algorithms for estimating $\mathbb{P}\{S_n > b\}$ when component X_is have regularly varying right tails, for $b \geq \tilde{c}n^{\frac{1}{2}+\epsilon}$ as $n \nearrow \infty$. We believe that this approach has much wider applicability. To illustrate this, we further show that similar importance sampling distribution can also be used to develop strongly efficient algorithms for estimating $\mathbb{P}\{S_n > b\}$ when n is fixed and b increases to infinity. Note that algorithms proposed in [15], [14] and [16] for this problem were weakly efficient. [17] and [19] exploit alternative representations for $\mathbb{P}(S_n > b)$ to arrive at faster algorithms. Our approach can be potentially adapted to those representations to provide further performance improvement.

The organization of the rest of the paper is as follows: In Section II we discuss the preliminary concepts relevant to the problems addressed. We introduce our state-independent importance sampling algorithms for estimating $\mathbb{P}\{S_n > b\}, b \geq \tilde{c}n^{\frac{1}{2}+\epsilon}$ as $n \nearrow \infty$ in Section III and for $\mathbb{P}\{S_n > b\}$ as $b \nearrow \infty$ with fixed n in Section IV. We discuss a numerical simulation example in Section V, where we compare the performance of the proposed algorithm, which runs in $O(n)$ time with that of the state-dependent importance sampling algorithm in [3], which runs in $O(n^2)$ time; there we note that the observed values of relative error are comparable for large values of n. We end with a brief conclusion in Section VI. Some of the more technical proofs are presented in the appendix.

II. Preliminaries

In this section we first specify the exceedance probabilities considered in this paper for efficient simulation. We then briefly review rare event simulation and the use of importance sampling for estimating rare event probabilities. We also review the relevant efficiency notions that asymptotically quantify the performance improvement offered by successful importance sampling algorithms. We then review the asymptotics in the existing literature for the tail probabilities that we need in proving efficiency results for the proposed algorithms. Finally we review some of the importance sampling techniques proposed in the literature for efficient simulation of large deviation probabilities associated with random walks involving light as well as regularly varying heavy-tailed random variables.

A. Problem setup

Let $\{X_n : n \geq 1\}$ denote a collection of i.i.d. random variables with distribution function F, mean 0 and finite variance σ^2. The right tail of F is taken to be regularly-varying, so that $\bar{F}(x) = 1 - F(x) = L(x)/x^\alpha$ for some slowly varying function L and $\alpha > 2$; as is well known, function L is said to be slowly varying if for all $t > 0$,

$$L(tx)/L(x) \to 1$$

as $x \nearrow \infty$. Regularly varying distributions form an important class of heavy-tailed distributions. The random walk associated with the collection $\{X_n : n \geq 1\}$ is given by $S_0 = 0$,

$$S_n = X_1 + \ldots + X_n$$

for each $n \geq 1$. We let $M(\theta) = \mathbb{E}[e^{\theta X_1}]$ and $\Lambda(\theta) = \ln[M(\theta)]$ denote the moment and cumulant generating functions of X_1, respectively.

The following assumption is imposed on the random variables X_i throughout the paper.

Assumption 1: The left tail of F is light, that is, there exists $\phi > 0$ such that $M(-\phi) < \infty$.

In Sections III and IV, respectively, we provide strongly efficient algorithms for the following two probabilities:

- **P1**: computation of $\mathbb{P}\{S_n > b\}$ as $n \nearrow \infty$, for $b > \tilde{c}n^{\frac{1}{2}+\epsilon}$ given $\tilde{c}, \epsilon > 0$, and
- **P2**: computation of $\mathbb{P}\{S_n > b\}$ as $b \nearrow \infty$ for fixed n.

B. Rare event simulation and importance sampling

Let A denote a rare event on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, i.e., $z := \mathbb{P}(A) > 0$ is small (in our setup A corresponds to the event $\{S_n > b\}$). Suppose that we are interested in obtaining an estimator $\hat{z}$ for z such that the relative error $|\hat{z} - z|/z < \delta$, with probability at least $1 - \epsilon$, for given ϵ and $\delta > 0$.

Naive simulation for estimating z involves drawing N independent samples of the indicator $\mathbb{I}_A$ and taking their sample mean as the estimator. For a different measure $\tilde{\mathbb{P}}$ such that the Radon-Nikodym derivative $\frac{d\mathbb{P}}{d\tilde{\mathbb{P}}}$ is well defined on A, we get:

$$\mathbb{P}(A) = \int_A \frac{d\mathbb{P}}{d\tilde{\mathbb{P}}}(\omega) d\tilde{\mathbb{P}}(\omega) = \tilde{\mathbb{E}}\left[L\mathbb{I}_A\right],$$

where $L := d\mathbb{P}/d\tilde{\mathbb{P}}$ and $\tilde{\mathbb{E}}$ is the expectation associated with $\tilde{\mathbb{P}}$. Define $Z := L\mathbb{I}_A$; then Z is an unbiased estimator of z under measure $\tilde{\mathbb{P}}$. If N i.i.d samples $Z_1, \ldots, Z_N$ of Z can be drawn from $\tilde{\mathbb{P}}$, then by strong law of large numbers we have:

$$\hat{z}_{_N} := \frac{Z_1 + \ldots + Z_N}{N} \to z \text{ a.s.},$$

as $N \nearrow \infty$. This method of generating an estimator is called *importance sampling* (IS). The measure $\tilde{\mathbb{P}}$ is called the importance sampling measure and Z is called an importance sampling estimator.

Using Chebyshev's inequality allows us to find an upper bound on the required number of samples N to achieve the desired relative precision:

$$\mathbb{P}\left(\frac{|\hat{z}_{N} - z|}{z} > \delta\right) \leq \frac{Var(\hat{z}_{N})}{z^2\delta^2} = \frac{CV^2(Z)}{N\delta^2}.$$

Here $CV(Z) = \sqrt{Var(Z)}/z$ is the coefficient of variation of Z. This enables us to conclude that if we generate at least

$$N = \frac{CV^2(Z)}{\epsilon\delta^2} \tag{1}$$

i.i.d. samples of Z, we can guarantee the desired relative precision.

In naive simulation we use the measure $\mathbb{P}$ itself and have $Z = \mathbb{I}_A$ as the estimator; so the number of samples required in (1) grows (roughly proportional to z^{-1}) to infinity if $z \searrow 0$.

As is well known, the choice $\mathbb{P}^*(\cdot) := \mathbb{P}(\cdot|A)$ as an importance sampling measure yields zero variance for the associated estimator $Z = z\mathbb{I}_A$. Then, in simulation every sample equals z with $\mathbb{P}^*$ probability 1. However, the explicit dependence of Z on z, the quantity which we want to estimate makes this method impractical.

Efficiency notions of algorithms: Consider a family of events $\{A_n : n \geq 1\}$ such that $z_n := \mathbb{P}(A_n) \searrow 0$ as $n \nearrow \infty$. For an importance sampling algorithm to compute $\{z_n : n \geq 1\}$, we come up with a sequence of changes of measures $\{\tilde{\mathbb{P}}_n : n \geq 1\}$ and estimators $\{Z_n : n \geq 1\}$ such that $\tilde{\mathbb{E}}_n Z_n = z_n$, where $\tilde{\mathbb{E}}_n$ denotes the expectation operator under $\tilde{\mathbb{P}}_n$.

Definition 1: The sequence $\{Z_n : n \geq 1\}$ of unbiased importance sampling estimators of $\{z_n : n \geq 1\}$, is said to be *strongly efficient* if,

$$\overline{\lim_{n\to\infty}} \frac{\tilde{\mathbb{E}}_n\left(Z_n^2\right)}{z_n^2} < \infty. \tag{2}$$

It is said to be *weakly efficient* if for each $\epsilon > 0$,

$$\overline{\lim_{n\to\infty}} \frac{\tilde{\mathbb{E}}_n\left(Z_n^2\right)}{z_n^{2-\epsilon}} < \infty. \tag{3}$$

From (1), we can see that if an algorithm is strongly efficient, the number of simulation runs required to guarantee desired relative precision stays bounded as $n \nearrow \infty$. It also follows that strong efficiency implies weak efficiency and that naive simulation is not even weakly efficient.

C. Tail asymptotics for i.i.d sums

(a) The well-known asymptotics

$$\mathbb{P}\{S_n > b\} \sim n\bar{F}(b), \text{ as } n \nearrow \infty \tag{4}$$

for $b > \sqrt{n \log n}$ can be found in [23], [24] or [25]†. Let $M_n := \max_{k\leq n} X_k$, then it is easily verified that

$$\mathbb{P}\{M_n > b\} \sim n\bar{F}(b).$$

Additionally, the following asymptotics can be found in [24]: as $n \nearrow \infty$,

$$\begin{aligned}
&\mathbb{P}\{S_n > b, M_n < b\} = o(n\bar{F}(b)) \\
&\sup_{b\geq\sqrt{n\log n}} |\mathbb{P}\{N(n,b) = 1|S_n \geq b\} - 1| \to 0, \\
&\sup_{b\geq\sqrt{n\log n}} |\mathbb{P}\{M_{n-1} \leq b, S_n \geq b|X_n > b\} - 1| \to 0,
\end{aligned}$$

where $N(n,b)$ denotes the cardinality of $\{1 \leq i \leq n : X_i > b\}$. These large deviations asymptotics reveal that with the number of summands growing to infinity, with high probability, the sum becomes large because of one of the components becomes large.

(b) The following tail probability asymptotic of $\mathbb{P}\{S_n > b\}$ for fixed n can be found, e.g., in [26]:

$$\mathbb{P}\{S_n > b\} \sim n\bar{F}(b), \text{ as } b \nearrow \infty. \tag{5}$$

D. Review of existing IS techniques

1) For light-tails: Suppose that $\{Y_i\}$ are i.i.d., zero mean, light-tailed random variables (that is, their moment generating function exists in a neighborhood of zero) with distribution function F_Y. Sadowsky and Bucklew [6] introduce a weakly-efficient method for estimating $\mathbb{P}\{\sum_{i=1}^n Y_i > na\}, a > 0$ that involves generating i.i.d. samples $\{Y_i\}$ using the distribution function obtained by *exponentially twisting* the original distribution function, that is, using

$$F_\theta(dx) = e^{\theta x - \Lambda_Y(\theta)} F_Y(dx), \tag{6}$$

where $\Lambda_Y(\theta) = \ln E[e^{\theta Y_i}]$ denotes the log-moment generating function of Y_i. [6] propose that $\theta > 0$ be chosen such that $\Lambda_Y'(\theta) = a$, (this is assumed to exist). Note that the resulting estimator has the form:

$$Z_n = \exp\left(-\left(\theta\sum_{i=1}^n Y_i - n\Lambda_Y(\theta)\right)\right)\mathbb{I}_{\left\{\sum_{i=1}^n Y_i > na\right\}}$$

2) For heavy-tails: For regularly varying right tailed random variables $\{X_i\}$, the associated log-moment generating function $\Lambda(\theta) = \infty$, for each $\theta > 0$; therefore the above method of exponential twisting is not directly applicable for simulation of $P(S_n > b)$. Blanchet and Liu [4] propose two efficient state-dependent importance sampling algorithms for solving **P1**. The better of the two involves drawing samples for increment X_{k+1} based on S_k from an appropriately parameterized family of density functions. In particular, they recommend using

$$\tilde{\mathbb{P}}\{X_{k+1} \in dx|S_k = s\} = p_n\frac{F(dx)\mathbf{1}_{\{x>a(b-s)\}}}{\bar{F}(a(b-s))} + (1-p_n)\frac{F(dx)\mathbf{1}_{\{x\leq a(b-s)\}}}{F(a(b-s))},$$

†We say that $\{a_n\} \sim \{b_n\}$ if $\frac{a_n}{b_n} \to 1$ as $n \to \infty$

where p_n is the suitably chosen mixture probability and $a \in (0,1)$. Asymptotic forms of $\mathbb{P}\{S_n > b\}$ given in [23] are used to construct a family of functions that satisfy a certain Lyapunov inequality; relevant parameters such as p_n and a are chosen in a way that the Lyapunov inequality holds. The state-dependent scheme used in very recent paper [5] avoids using asymptotic approximations of $\mathbb{P}\{S_n > b\}$ by employing sequential importance sampling and resampling methods.

III. Proposed algorithm for P1

The proposed importance sampling distribution for drawing samples of $X_1, \ldots, X_n$ to estimate $\mathbb{P}\{S_n > b\}$ is:

$$\tilde{F}_n(dx) := c_n e^{\theta_n \cdot (x \wedge b)} F(dx), \tag{7}$$

with $\theta_n := \theta_n(b)$ given by,

$$\theta_n(b) = \frac{1}{b} \log\left(\frac{1}{n\bar{F}(b)}\right). \tag{8}$$

In (7) $c_n := c_n(b)$ denotes the appropriate normalizing constant, which is well-defined because:

$$\frac{1}{c_n(b)} = \int_{-\infty}^{b} e^{\theta_n x} F(dx) + e^{\theta_n b}\bar{F}(b) \leq e^{\theta_n b} = \frac{1}{n\bar{F}(b)}.$$

Though we do not emphasize in notation, it can be seen that all $\tilde{F}_n, \theta_n$ and c_n defined above depend on both n and b. Since we suitably modify the tail and circumvent the problem of normalizing constant c_n becoming 0, we refer to $\{\tilde{F}_n : n \geq 1\}$ as tail-adjusted exponentially twisted distributions corresponding to F. It is easy to check that $\theta_n \searrow 0$ as $n \nearrow \infty$.

Let $\tilde{\mathbb{P}}_n$ be the probability measure induced when $\{X_i : i \geq n\}$ are independently distributed according to $\tilde{F}_n$. Then under the importance sampling change of measure $\tilde{\mathbb{P}}_n$, the unbiased estimator for $\mathbb{P}\{S_n > b\}$ takes the form:

$$Z_1(n,b) = \frac{1}{c_n^n} \exp\left(-\theta_n \sum_{i=1}^{n} (X_i \wedge b)\right) \mathbb{I}_{\{S_n > b\}}. \tag{9}$$

One simulation run of the corresponding state-independent importance sampling algorithm is given below:

Algorithm 1:
Given parameters $b \geq \tilde{c} n^{\frac{1}{2}+\epsilon}, \tilde{c}, \epsilon > 0$
STEP 1: Draw n i.i.d samples $x_1, \ldots, x_n$ from $\tilde{F}_n$
STEP 2: $s \longleftarrow x_1 + \ldots + x_n$
STEP 3: $L \longleftarrow \frac{1}{c_n^n} e^{-\theta_n \sum_{i=1}^n (x_i \wedge b)}$
STEP 4: RETURN $L\mathbf{1}_{\{s > b\}}$
STOP

Doing many independent simulation runs and taking the sample average of the returned values gives an unbiased estimator of $\mathbb{P}\{S_n > b\}$.

Observe that the IS measure $\tilde{\mathbb{P}}_n$ induced by $\tilde{F}_n$ is such that the event $\{X_k > b\}$ receives a considerable mass of c_n/n under $\tilde{\mathbb{P}}_n$. We provide asymptotic bounds for the integral $\int_{-\infty}^{b} e^{\theta_n x} F(dx)$ in the appendix and there we prove the following lemma:

Lemma 1: $\lim_{n \to \infty} c_n = 1$.

We now argue that the proposed measure $\tilde{\mathbb{P}}_n$ assigns similar probability to events $\{X_j > x\}$ for each $x \geq b$ as the zero-variance measure asymptotically as $n \nearrow \infty$. Recall that $\mathbb{P}\{X_j > x | S_n > b\}$ denotes the probability assigned to $\{X_j > x\}$ under the zero-variance measure. We have,

$$\mathbb{P}\{X_j > x | S_n > b\} = \mathbb{P}\{S_n > b | X_j > x\} \frac{\mathbb{P}\{X_j > x\}}{\mathbb{P}\{S_n > b\}}.$$

Since $\mathbb{P}\{S_n > b | X_j > x\} \sim 1$ as $n \nearrow \infty$, we get

$$\mathbb{P}\{X_j > x | S_n > b\} \sim \frac{\bar{F}(x)}{n\bar{F}(b)}.$$

It can be verified that $\tilde{\mathbb{P}}_n\{X_j > x\} = c_n \frac{\bar{F}(x)}{n\bar{F}(b)}$. Therefore from Lemma 1,

$$\mathbb{P}\{X_j > x | S_n > b\} \sim \tilde{\mathbb{P}}\{X_j > x\}, \text{ as } n \nearrow \infty.$$

A. Strong efficiency of Z_1

For proving the strong efficiency of Z_1 we bound its second moment separately over the elements of the partition $\{A_k : 0 \leq k \leq n\}$, with A_k denoting the event of exactly k of $X_1, X_2, \ldots, X_n$ crossing b. Recall that θ_n and c_n depend on both n and b, though in notation we often make only the dependence on n explicit. The following lemma, useful for establishing strong efficiency of the above algorithm, is proved in the appendix.

Lemma 2: Under Assumption 1, for any given $\tilde{c}, \epsilon > 0$, there exists a constant D_1 such that:

(a) $\frac{1}{(c_n(b))^n} \leq D_1$,

(b) $\left(\frac{\tilde{\mathbb{E}}_n\left[\exp\left(-2\theta_n X_1 \mathbb{I}_{\{X_1 < b\}}\right)\right]}{c_n}\right)^n \leq D_1$

are satisfied $\forall n$ and $b \geq \tilde{c} n^{\frac{1}{2}+\epsilon}$.

The strong efficiency of Algorithm 1 is established in the following theorem.

Theorem 1: Under Assumption 1, Algorithm 1 is strongly efficient when $b \geq \tilde{c} n^{\frac{1}{2}+\epsilon}$, that is,

$$\overline{\lim_{n\to\infty}} \frac{\tilde{\mathbb{E}}_n[Z_1^2(n,b)]}{(\mathbb{P}\{S_n > b\})^2} < \infty, \tag{10}$$

given any $b \geq \tilde{c} n^{\frac{1}{2}+\epsilon}, \forall \tilde{c}, \epsilon > 0$.

Proof: Recall that $A_k = \{\omega : \sum_{i=1}^n \mathbb{I}_{\{X_i > b\}} = k\}$. Let $B_k := A_k \cap \{S_n > b\}$, for $k = 0, \ldots, n$ and let

$$I_k(n) := \frac{1}{c_n^{2n}} \tilde{\mathbb{E}}_n \left[\exp\left(-2\theta_n \sum_{i=1}^{n} (X_i \wedge b)\right); B_k \right],$$

then $\tilde{\mathbb{E}}_n[Z_1^2(n,b)] = \sum_{k=0}^n I_k(n)$. On the set B_0, none of the $\{X_1, \ldots, X_n\}$ cross b but $S_n > b$. Therefore,

$$\begin{aligned} c_n^{2n} I_0(n) &= \tilde{\mathbb{E}}_n\left[e^{-2\theta_n \sum_{i=1}^n X_i}; B_0\right] \\ &\leq e^{-2\theta_n b}\tilde{\mathbb{P}}_n(B_0) \\ &\leq e^{2\log(n\bar{F}(b))}. \end{aligned}$$

$$\text{Thus, } I_0(n) \leq \frac{(n\bar{F}(b))^2}{c_n^{2n}}. \tag{11}$$

We observe that $X_i \wedge b = X_i \mathbb{I}_{\{X_i < b\}} + b\mathbb{I}_{\{X_i \geq b\}}$, and on $B_k, k = 1, \ldots, n$ $\sum_{i=1}^n \mathbb{I}_{\{X > b\}} = k$. Therefore,

$$\begin{aligned} c_n^{2n} I_k(n) &= \tilde{\mathbb{E}}_n \left[e^{-2\theta_n \sum_{i=1}^n X_i \mathbb{I}_{\{X_i < b\}}} e^{-2\theta_n \sum_{i=1}^n b \mathbb{I}_{\{X_i \geq b\}}}; B_k \right] \\ &= \tilde{\mathbb{E}}_n \left[e^{-2\theta_n \sum_{i=1}^n X_i \mathbb{I}_{\{X_i < b\}}} e^{-2k\theta_n b}; B_k \right] \\ &\leq (n\bar{F}(b))^{2k} \tilde{\mathbb{E}}_n \left[e^{-2\theta_n \sum_{i=1}^n X_i \mathbb{I}_{\{X_i < b\}}} \right], \end{aligned}$$

where in the last line we have used the relation $\theta_n b = -\log n\bar{F}(b)$, and the fact that $\tilde{\mathbb{E}}_n[Y; B_k] \leq \tilde{\mathbb{E}}_n[Y]$ for $Y \geq 0$. Therefore, for $k = 1, \ldots, n$,

$$I_k(n) \leq \frac{(n\bar{F}(b))^{2k}}{c_n^{2n}} \left(\tilde{\mathbb{E}}_n \left[e^{-2\theta_n X_1 \mathbb{I}_{\{X_1 < b\}})} \right] \right)^n. \tag{12}$$

Now combining the above equation with (11),

$$\frac{\tilde{\mathbb{E}}_n \left[Z_1^2(n,b) \right]}{(n\bar{F}(b))^2} \leq \frac{1}{c_n^{2n}} + \frac{1}{c_n^n} \left(\frac{\tilde{\mathbb{E}}_n \left[e^{-2\theta_n X_1 \mathbb{I}_{\{X_1 < b\}}} \right]}{c_n} \right)^n \times \left(\frac{1 - (n\bar{F}(b))^{2n}}{1 - (n\bar{F}(b))^2} \right)$$

From (4), $\mathbb{P}\{S_n > b\} \sim n\bar{F}(b)$ as $n \nearrow \infty$ uniformly for $b > \sqrt{n \log n}$. Therefore application of Lemma 2 results in:

$$\overline{\lim_{n\to\infty}} \frac{\tilde{\mathbb{E}}[Z_1^2(n,b)]}{(\mathbb{P}\{S_n > b\})^2} \leq D_1^2 + D_1.D_1.1 < \infty,$$

thus establishing the strong efficiency of Algorithm 1. ∎

IV. Proposed algorithm for P2

In this section we consider efficient simulation of the rare probabilities $\mathbb{P}\{S_n > b\}$ for some fixed n as $b \nearrow \infty$. Efficient state-independent algorithms already exist, refer [15], [14], [16], [17] and [19] for some of the related literature. Our aim is to illustrate that the importance sampling measure resulting by doing tail-adjusted exponential twisting as before for **P1**, yields a strongly efficient algorithm for **P2** as well.

The importance sampling distribution $\hat{F}_b$ we recommend for drawing samples for $X_1, \ldots, X_n$ is:

$$\hat{F}_b(dx) := c_b e^{\theta_b \cdot (x \wedge b)} F(dx), \tag{13}$$

where,

$$\theta_b = \frac{1}{b} \log \left(\frac{1}{n\bar{F}(b)} \right), \tag{14}$$

and c_b is the appropriate normalizing constant. Let $\hat{\mathbb{P}}_b$ be the measure induced by following the above importance sampling distribution for drawing independent samples of X_is. Let $\hat{\mathbb{E}}_b$ be the associated expectation operator. The unbiased importance sampling estimator for $\mathbb{P}\{S_n > b\}$ is given by:

$$Z_2(b) = \frac{1}{c_b^n} \exp \left(-\theta_b \sum_{i=1}^n (X_i \wedge b) \right) \mathbb{I}_{\{S_n > b\}}. \tag{15}$$

We provide only the outline of the proof of strong efficiency, as the details are almost identical to that in the proof for Algorithm 1.

Theorem 2: Under Assumption 1, estimator $Z_2(b)$ is strongly efficient under measure $\hat{\mathbb{P}}_b$; that is,

$$\overline{\lim_{b\to\infty}} \frac{\hat{\mathbb{E}}_b \left[Z_2^2(b) \right]}{(\mathbb{P}\{S_n > b\})^2} < \infty. \tag{16}$$

Proof: As before for $k = 1, \ldots, n$, let

$$B_k := \left\{ \omega : \sum_{i=1}^n \mathbb{I}_{\{X_i > b\}} = k \right\} \cap \{S_n > b\}, \text{ and}$$

$$I_k(b) := \frac{1}{c_b^{2n}} \hat{\mathbb{E}}_b \left[e^{-2\theta_b \sum_{i=1}^n (X_i \wedge b)}; B_k \right].$$

So we have $\hat{\mathbb{E}}_b[Z_2^2(b)] = \sum_{k=0}^n I_k(b)$. As in the proof of Lemma 2, we can show that:

$$\overline{\lim_{b\to\infty}} \frac{1}{c_b} \leq 1 + \frac{2}{n}, \tag{17}$$

$$\overline{\lim_{b\to\infty}} \frac{\hat{\mathbb{E}}_b \left[e^{-2\theta_b X_1 \mathbb{I}_{\{X_1 \leq b\}}} \right]}{c_b} \leq \frac{1}{n}. \tag{18}$$

For $I_k, k = 1, \ldots, n$ we just follow the same procedure to get,

$$I_0(b) \leq \frac{(n\bar{F}(b))^2}{c_b^{2n}},$$

$$I_k(b) \leq \frac{(n\bar{F}(b))^{2k}}{c_b^{2n}} \left(\hat{\mathbb{E}}_b \left[e^{-2\theta_b X_1 \mathbb{I}_{\{X_1 < b\}})} \right] \right)^n.$$

Therefore,

$$\frac{\hat{\mathbb{E}}_b Z_2^2(b)}{(n\bar{F}(b))^2} \leq \frac{1}{c_b^{2n}} + \frac{1}{c_b^n} \left(\frac{\hat{\mathbb{E}}_b \left[e^{-2\theta_b X_1 \mathbb{I}_{\{X_1 < b\}}} \right]}{c_b} \right)^n \times \left(\frac{1 - (n\bar{F}(b))^{2n}}{1 - (n\bar{F}(b))^2} \right)$$

Using the asymptotics $\mathbb{P}\{S_n > b\} \sim n\bar{F}(b)$ as $b \nearrow \infty$, (17) and (18) we get:

$$\overline{\lim_{b\to\infty}} \frac{\hat{\mathbb{E}}_b Z_2^2(b)}{(\mathbb{P}\{S_n > b\})^2} \leq \left(1 + \frac{2}{n} \right)^{2n} + \left(1 + \frac{2}{n} \right)^n \frac{1}{n^n} 1 < \infty.$$

This completes the proof. ∎

V. Numerical examples

Here we explain the results of simulation experiments for computing $\mathbb{P}\{S_n > na\}$ for different values of a and n. Suppose that $\{X_n : n \geq 1\}$ have a Pareto tail distribution in the right side and exponential tail distribution in the left. Specifically, the probability distribution function for X_i is:

$$F(x) := \begin{cases} \frac{e^{2x}}{2}, & \text{if } x < 0, \\ 1 - \frac{1}{2(1+x)^3}, & \text{otherwise} \end{cases}$$

The random variable X_i has zero mean. We use $N = 1000$ simulation runs to estimate $\mathbb{P}\{S_n > na\}$ for $n = 50, 75, 100$ and $a = 1, 2, 3$. In the implementation of Algorithm 1, we draw samples for the increments from the density given by (7). For $a = 1$, we take $\theta_{50} = 0.0745, \theta_{75} = 0.0542, \theta_{100} = 0.0431$ as dictated by (8); θ_n is chosen similarly for other choices of a as well. Table I provides a comparison of the

TABLE I

AS IS APPARENT FROM THE VR COLUMN OF ALGORITHM 1, THE PROPOSED IS SCHEME GIVES LARGE VARIANCE REDUCTION COMPARED TO NAIVE SIMULATION; THE RE COLUMN APPEARS TO BE BOUNDED, THUS SUGGESTING STRONG EFFICIENCY; ALSO FOR LARGE n, THE VARIANCE REDUCTION OFFERED BY ALGORITHM 1 IS COMPARABLE WITH THAT OF ALGORITHM BL. RECALL THAT THE COMPUTATIONAL EFFORT FOR SAMPLE GENERATION IN ALGORITHM 1 AND ALGORITHM BL SCALE LIKE $O(n)$ AND $O(n^2)$, RESPECTIVELY

		ALGORITHM 1			ALGORITHM BL		
	n	Estimator (9)	VR	RE	Estimator	VR	RE
a=1	50	1.905×10^{-4}	6.591×10^{1}	2.861	1.972×10^{-4}	2.041×10^{3}	1.536
	75	1.037×10^{-4}	1.145×10^{3}	2.903	1.085×10^{-4}	5.255×10^{3}	1.886
	100	4.887×10^{-5}	3.213×10^{3}	2.529	5.012×10^{-5}	5.494×10^{3}	1.447
a=2	50	2.833×10^{-5}	7.677×10^{3}	2.144	2.621×10^{-5}	6.290×10^{4}	0.810
	75	1.098×10^{-5}	2.336×10^{4}	1.975	1.012×10^{-5}	9.699×10^{4}	1.052
	100	6.663×10^{-6}	4.369×10^{4}	1.853	6.650×10^{-6}	9.059×10^{4}	1.289
a=3	50	6.947×10^{-6}	3.725×10^{3}	1.966	6.835×10^{-6}	2.988×10^{5}	0.705
	75	3.489×10^{-6}	8.575×10^{4}	1.829	3.531×10^{-6}	4.388×10^{5}	0.799
	100	1.785×10^{-6}	1.617×10^{5}	1.861	1.659×10^{-6}	3.723×10^{5}	1.320

performance of Algorithm 1 which runs in $O(n)$ time with that of the $O(n^2)$ algorithm proposed in [3], which is identified as Algorithm BL in the table. In Table I, VR denotes the ratio of the variance of naive simulation estimator to the variance of the estimator corresponding to the specified algorithm. Variance of the naive simulation estimator is taken to be $\hat{\gamma}(1-\hat{\gamma})$, where $\hat{\gamma}$ denotes the estimate of the probability $\mathbb{P}\{S_n > na\}$ obtained using the specified algorithm. Furthermore, RE denotes the ratio of empirical standard deviation to the estimate $\hat{\gamma}$.

The relative error of the estimates under Algorithm 1 appear to be bounded as n increases, thus suggesting strong efficiency. Further it can be noted that the variance reduction by Algorithm 1 is comparable with that of the Algorithm BL, as n increases. Using 1000 simulation runs, for large values of n, it has been observed that Algorithm 1 runs at least 20 times faster than Algorithm BL, thus underscoring the computational advantage offered by the $O(n)$ running time of Algorithm 1 over the $O(n^2)$ running time of Algorithm BL.

VI. CONCLUSION

In this paper we revisited the problem of estimating $P\{S_n > b\}$ in the settings where both n and b increase to infinity (**P1**) as well as when n is fixed and b increases to infinity (**P2**) when the constituent random variables have heavy regularly varying right tails. We showed that state-independent exponential twisting based algorithms can be developed to efficiently estimate these probabilities. The current literature on the other hand focuses on estimating **P1** using more nuanced state-dependent methods. In many settings, the proposed algorithms offer implementation and computational benefits over state-dependent algorithms. Also, our work re-examines the widely held belief that exponential twisting is applicable only when light-tailed random variables are involved.

We believe that the proposed methods are applicable in greater generality. For instance, in our ongoing effort we attempt to efficiently estimate probabilities such as

$$\mathbb{P}\left\{\max_{k\leq n} S_k \geq b\right\}$$

as both n and b increase to infinity, when the constituent variables have a negative mean and heavy right-tails. Such probabilities have wide application in insurance and in queueing (see, e.g. [27] and [28]).

APPENDIX

We aim to provide proofs of Lemmas 1 and 2 making use of Lemmas 3 and 4, which are stated and proved below. The proof of the Lemma 3 below, which gives asymptotic upper bound for the integral $\int_{-\infty}^{b} e^{\theta_n x}F(dx)$, follows the approach in [25] for bounding similar integral while deriving large deviations asymptotics for $\mathbb{P}\{S_n > b\}$.

Lemma 3: For any pair of sequences $\{x_n\}, \{\phi_n\}$ satisfying $x_n \nearrow \infty$ and $\phi_n x_n \nearrow \infty$, the integral,

$$\int_{-\infty}^{x_n} e^{\phi_n x}F(dx) \leq 1 + \frac{e^{2\alpha}\sigma^2\phi_n^2}{2} + e^{2\alpha}\bar{F}\left(\frac{2\alpha}{\phi_n}\right) + e^{\phi_n x_n}\bar{F}(x_n)(1+\epsilon_n),$$

where $\epsilon_n \geq 0$ is such that $\epsilon_n \searrow 0$ as $n \nearrow \infty$

Proof: We split the region of integration into $(-\infty, \gamma/\phi_n]$ and $(\gamma/\phi_n, x_n]$ for some constant $\gamma > 0$; the partition is such that the integrand stays bounded in the former despite its growth to $(-\infty,\infty)$.

Let $I_1 := \int_{-\infty}^{\gamma/\phi_n} e^{\phi_n x}F(dx)$ and $I_2 := \int_{\gamma/\phi_n}^{x_n} e^{\phi_n x}F(dx)$. Since $e^{\phi_n x} \leq 1 + \phi_n x + \frac{\phi_n^2 x^2}{2}e^{\phi_n x}$,

$$\begin{aligned} I_1 &\leq \int_{-\infty}^{\gamma/\phi_n} F(dx) + \phi_n \int_{-\infty}^{\gamma/\phi_n} xF(dx) \\ &\quad + \frac{\phi_n^2}{2}\int_{-\infty}^{\gamma/\phi_n} x^2 e^{\phi_n x}F(dx) \\ &\leq \int_{-\infty}^{\infty} F(dx) + \phi_n\int_{-\infty}^{\infty} xF(dx) + \frac{\phi_n^2 e^{\gamma}}{2}\int_{-\infty}^{\infty} x^2F(dx) \\ &= 1 + \frac{e^{\gamma}\phi_n^2\sigma^2}{2}, \end{aligned} \tag{19}$$

because $\mathbb{E}X = 0$ and $\mathbb{E}X^2 = \sigma^2$. Integrating by parts for the second integral,

$$
\begin{aligned}
I_2 &= -\int_{\gamma/\phi_n}^{x_n} e^{\phi_n x} \bar{F}(dx) \\
&= e^{\phi_n \gamma/\phi_n} \bar{F}(\gamma/\phi_n) - e^{\phi_n x_n} \bar{F}(x_n) + \phi_n \int_{\gamma/\phi_n}^{x_n} e^{\phi_n x} \bar{F}(x)dx \\
&\leq e^{\gamma} \bar{F}(\gamma/\phi_n) + I_2', \qquad (20)
\end{aligned}
$$

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

$$
\begin{aligned}
I_2' &= e^{\phi_n x_n} \int_0^{\phi_n x_n - \gamma} e^{-u} \bar{F}\left(x_n - \frac{u}{\phi_n}\right) du \\
&= e^{\phi_n x_n} \bar{F}(x_n) \int_0^{\phi_n x_n - \gamma} e^{-u} g_n(u) du, \qquad (21)
\end{aligned}
$$

where,

$$
g_n(u) := \frac{\bar{F}\left(x_n - \frac{u}{\phi_n}\right)}{\bar{F}(x_n)} = \frac{\bar{F}\left(x_n \left(1 - \frac{u}{\phi_n x_n}\right)\right)}{\bar{F}(x_n)}.
$$

Since L 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, which is Theorem 1.1.4 of [25], helps in concluding that $g_n(u) \sim 1$ as $n \to \infty$ for any $u \in [0, \phi_n x_n - \gamma]$.
Now fix $\delta = \frac{\alpha}{2}$. Then for n large enough,

$$
g_n(u) \leq \left(1 + \frac{\alpha}{2}\right)\left(1 - \frac{u}{\phi_n x_n}\right)^{-\frac{3\alpha}{2}}. \qquad (22)
$$

To show that $g_n(u)$ is uniformly bounded in n and $u \in [0, \phi_n x_n - \gamma]$ by an integrable function, so that we can apply dominated convergence, we call the term involving u in the upper bound in (22) as $h(u)$,that is $h(u) = (1 - u/\phi_n x_n)^{-\frac{3\alpha}{2}}$. 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:

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

Applying dominated convergence theorem, we get

$$
\int_0^{\phi_n x_n - \gamma} e^{-u} g_n(u) du \sim 1 \text{ as } n \nearrow \infty.
$$

Since $\int_{-\infty}^{x_n} e^{\phi_n x} F(dx) = I_1 + I_2$, combining this result with (19), (20) and (21), completes the proof. ■

Lemma 4: For $b \geq \tilde{c} n^{\frac{1}{2}+\epsilon}$, given $\tilde{c}$ and $\epsilon > 0$, if θ_n is given by (8), then the following hold true as $n \nearrow \infty$:

(a) $\theta_n = o\left(\frac{1}{\sqrt{n}}\right)$,

(b) $\bar{F}\left(\frac{2\alpha}{\theta_n}\right) = o(\frac{1}{n})$.

Proof: (a) We have $\bar{F}(x) = \frac{L(x)}{x^\alpha}$. Since L is slowly varying, given any $\delta > 0$ for sufficiently large values of b, we have $b^{-\delta} \leq L(b) \leq b^{\delta}$, thus yielding $L(b) = b^{o(1)}$ as $b \nearrow \infty$. Since $b \geq \tilde{c} n^{\frac{1}{2}+\epsilon}, n/b^2 \leq \tilde{c}^{-1} n^{-2\epsilon}$. Then,

$$
\begin{aligned}
n\theta_n^2 &= \frac{n}{b^2} \log^2\left(\frac{1}{n\bar{F}(b)}\right) \\
&= \frac{n}{b^2} \log^2\left(\frac{b^\alpha}{nL(b)}\right) \\
&= O\left(\frac{\log^2 n}{n^{2\epsilon}}\right) \to 0, \text{ as } n \nearrow \infty.
\end{aligned}
$$

(b) Since $\bar{F}$ is regularly varying, By Theorem 1.1.4 (iii) of [25], given any $\delta > 0$, for n large enough,

$$
\begin{aligned}
n\bar{F}\left(\frac{2\alpha}{\theta_n}\right) &= n\bar{F}\left(\frac{2\alpha b}{-\log(n\bar{F}(b))}\right) \\
&\leq n\left(\frac{-\log\left(n\bar{F}(b)\right)}{2\alpha}\right)^{\alpha+\delta} \bar{F}(b) \\
&= O\left((\log n)^{\alpha+\delta}\right) \frac{nL(b)}{b^\alpha} \to 0, \text{ as } n \nearrow \infty.
\end{aligned}
$$

because $\alpha > 2$ and L is slowly varying. ■

Now using the above two lemmas, we provide proofs for the results left unproved in previous sections.

Proof of Lemma 1

Since $c_n(b)$ is the normalizing constant in (7),

$$
\begin{aligned}
\frac{1}{c_n(b)} &= \int_{-\infty}^{b} e^{\theta_n x} F(dx) + e^{\theta_n b} \bar{F}(b) \\
&\leq 1 + \frac{e^{2\alpha} \sigma^2 \theta_n^2}{2} + e^{2\alpha} \bar{F}\left(\frac{2\alpha}{\theta_n}\right) + \frac{1}{n}(2 + \epsilon_n), \qquad (23)
\end{aligned}
$$

where we have used the Lemma 3 and the relationship, $\theta_n b = -\log\left(n\bar{F}(b)\right)$. Usage of Lemma 3 is justified because $\theta_n b \nearrow \infty$ as $n \nearrow \infty$. Since the terms involved are of $o(1)$ as noted in Lemma 4, $\underline{\lim}_{n\to\infty} c_n \geq 1$. Now for the other side,

$$
\begin{aligned}
\frac{1}{c_n} &= \int_{-\infty}^{b} e^{\theta_n x} F(dx) + \frac{1}{n} \\
&\geq F(b) e^{\theta_n \int_{-\infty}^{b} \frac{x}{F(b)} F(dx)} + \frac{1}{n},
\end{aligned}
$$

where the last inequality is due to Jensen's inequality. Since $\int_{-\infty}^{b} xF(dx) \to 0, \theta_n \searrow 0$ and $F(b) \nearrow 1$, as $n \nearrow \infty$ we get $\overline{\lim}_{n\to\infty} c_n \leq 1$. ■

Proof of Lemma 2

(a) In (23) if we use $1 + x \leq e^x$ for positive x, then

$$
\begin{aligned}
\frac{1}{c_n^n(b)} &\leq e^{\frac{e^{2\alpha}\sigma^2 n\theta_n^2}{2} + ne^{2\alpha}\bar{F}\left(\frac{2\alpha}{\theta_n}\right) + n\frac{1}{n}(2+\epsilon_n)} \\
&= e^{o(1) + 1.(2+\epsilon_n)} \to e^2,
\end{aligned}
$$

where the final equality is due to application of Lemma 4 for the first two terms in the exponent. Thus we have established that $c_n^{-n}(b)$ stays bounded.

(b) Recollect that Λ is the cumulant generating function of X_1 under $\mathbb{P}$. Now consider,

$$\begin{aligned}\tilde{\mathbb{E}}_n\left[e^{-2\theta_n X_1 \mathbb{I}_{\{X_1<b\}}}\right] &= \int_{-\infty}^{b} e^{-2\theta_n x}\tilde{F}_n(dx) + 1 - \tilde{F}_n(b) \\ &= c_n \int_{-\infty}^{b} e^{-\theta_n x} F(dx) + \frac{c_n}{n} \\ &\le c_n\left(e^{\Lambda(-\theta_n)} + \frac{1}{n}\right), \end{aligned} \tag{24}$$

where the second equality follows from (7) and using $1 - \tilde{F}_n(b) = c_n e^{\theta_n b}\bar{F}(b)$.

Since $\theta_n \searrow 0$, using Assumption 1 we can say that $\Lambda(-\theta_n) < \infty$, for large enough n. Due to the continous differentiability of Λ in $(-\phi, 0)$, Taylor expansion of Λ around 0 gives

$$\Lambda(-\theta_n) = 0 + \frac{\theta_n^2}{2}\Lambda(\tilde{\theta}_n),$$

for some $\tilde{\theta}_n$ between $-\theta_n$ and 0. Since $\Lambda(0) = \sigma^2$, given any $\delta > 0$, we use continuity of Λ'' to write $\Lambda''(\tilde{\theta}_n) \le \sigma^2(1+\delta)$, for all n large enough. Thus we conclude that, given any $\delta > 0$, for sufficiently large values of n,

$$\Lambda(-\theta_n) \le \frac{\theta_n^2\sigma^2(1+\delta)}{2}. \tag{25}$$

Therefore from (24), we can write:

$$\frac{\tilde{\mathbb{E}}_n[e^{-2\theta_n X_1 \mathbb{I}_{\{X_1<b\}}}]}{c_n} \le e^{\theta_n^2 \frac{\sigma^2}{2}(1+\delta)} + \frac{1}{n}.$$

Then,

$$\begin{aligned}\left(\frac{\tilde{\mathbb{E}}_n[e^{-2\theta_n X_1 \mathbb{I}_{\{X_1<b\}}}]}{c_n}\right)^n &\le e^{n\theta_n^2 \frac{\sigma^2}{2}(1+\delta)}\left(1 + \frac{1}{n e^{\theta_n^2 \frac{\sigma^2}{2}(1+\delta)}}\right)^n \\ &\le e^{n\theta_n^2 \frac{\sigma^2}{2}(1+\delta) + e^{-\theta_n^2 \frac{\sigma^2}{2}(1+\delta)}}, \end{aligned} \tag{26}$$

holds true for all n large enough. Here we have used $1 + x \le e^x$ for x positive.
Since we have $n\theta_n^2 = o(1)$ from Lemma 4, (26) results in,

$$\overline{\lim_{n\to\infty}}\left(\frac{\tilde{\mathbb{E}}_n[e^{-2\theta_n X_1 \mathbb{I}_{\{X_1<b\}}}]}{c_n}\right)^n \le e^{0+e^0} = e.$$

Therefore we can find an appropriate constant D_1 such that both the inequalities stated in Lemma 2 hold true. ■

REFERENCES

[1] S. Asmussen and P. W. Glynn, *Stochastic simulation: algorithms and analysis*, ser. Stochastic Modelling and Applied Probability. New York: Springer, 2007, vol. 57.

[2] S. Juneja and P. Shahabuddin, "Rare-event simulation techniques: An introduction and recent advances," in *Simulation*, ser. Handbooks in Operations Research and Management Science, S. G. Henderson and B. L. Nelson, Eds. Elsevier, 2006, vol. 13, pp. 291 – 350. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S092705070613011X

[3] J. Blanchet and J. Liu, "Efficient simulation for large deviation probabilities of sums of heavy-tailed increments," in *Proceedings of the 38th Winter Simulation Conference*, ser. WSC '06. Winter Simulation Conference, 2006, pp. 757–764. [Online]. Available: http://portal.acm.org/citation.cfm?id=1218112.1218253

[4] J. Blanchet and J. C. Liu, "State-dependent importance sampling for regularly varying random walks," *Advances in Applied Probability*, vol. 40, pp. 1104–1128, 2008.

[5] H. Chan, S. Deng, and T. Lai, "Rare-event simulation of heavy-tailed random walks by sequential importance sampling and resampling," *Preprint*, 2012.

[6] J. S. Sadowsky and J. A. Bucklew, "On large deviations theory and asymptotically efficient Monte Carlo estimation," *IEEE Trans. Inform. Theory*, vol. 36, no. 3, pp. 579–588, 1990. [Online]. Available: http://dx.doi.org/10.1109/18.54903

[7] A. B. Dieker and M. Mandjes, "On asymptotically efficient simulation of large deviation probabilities," *Adv. in Appl. Probab.*, vol. 37, no. 2, pp. 539–552, 2005. [Online]. Available: http://dx.doi.org/10.1239/aap/1118858638

[8] J. Blanchet, K. Leder, and P. W. Glynn, "Efficient simulation of light-tailed sums: an old-folk song sung to a faster new tune...," in *Monte Carlo and quasi-Monte Carlo methods 2008*. Berlin: Springer, 2009, pp. 227–248. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-04107-5_13

[9] P. Glasserman and S. Juneja, "Uniformly efficient importance sampling for the tail distribution of sums of random variables," *Math. Oper. Res.*, vol. 33, no. 1, pp. 36–50, 2008.

[10] S. Dey and S. Juneja, "Efficient estimation of density and probability of large deviations of sum of iid random variables," in *Proceedings of the 38th Winter Simulation Conference*, dec. 2011, pp. 3800 –3811.

[11] S. Parekh and J. Walrand, "A quick simulation method for excessive backlogs in networks of queues," *IEEE Trans. Automat. Control*, vol. 34, no. 1, pp. 54–66, 1989. [Online]. Available: http://dx.doi.org/10.1109/9.8649

[12] P. Glasserman and J. Li, "Importance sampling for portfolio credit risk," *Management Science*, vol. 51, no. 11, pp. 1643–1656, 2005.

[13] A. Bassamboo, S. Juneja, and A. J. Zeevi, "Portfolio credit risk with extremal dependence: Asymptotic analysis and efficient simulation," *Operations Research*, vol. 56, no. 3, pp. 593–606, 2008.

[14] S. Asmussen, K. Binswanger, and B. Højgaard, "Rare events simulation for heavy-tailed distributions," *Bernoulli*, vol. 6, no. 2, pp. 303–322, 2000. [Online]. Available: http://dx.doi.org/10.2307/3318578

[15] S. Asmussen and K. Binswanger, "Simulation of ruin probabilities for subexponential claims," *ASTIN Bulletin*, vol. 27, pp. 297–318, 1997.

[16] S. Juneja and P. Shahabuddin, "Simulating heavy tailed processes using delayed hazard rate twisting," *ACM Trans. Model. Comput. Simul.*, vol. 12, no. 2, pp. 94–118, Apr. 2002. [Online]. Available: http://doi.acm.org/10.1145/566392.566394

[17] S. Asmussen and D. P. Kroese, "Improved algorithms for rare event simulation with heavy tails," *Adv. in Appl. Probab.*, vol. 38, no. 2, pp. 545–558, 2006. [Online]. Available: http://dx.doi.org/10.1239/aap/1151337084

[18] P. Dupuis, K. Leder, and H. Wang, "Importance sampling for sums of random variables with regularly varying tails," *ACM Trans. Model. Comput. Simul.*, vol. 17, no. 3, 2007.

[19] S. Juneja, "Estimating tail probabilities of heavy tailed distributions with asymptotically zero relative error," *Queueing Syst.*, vol. 57, no. 2-3, pp. 115–127, 2007.

[20] J. Blanchet and P. Glynn, "Efficient rare-event simulation for the maximum of heavy-tailed random walks," *Ann. Appl. Probab.*, vol. 18, no. 4, pp. 1351–1378, 2008. [Online]. Available: http://dx.doi.org/10.1214/07-AAP485

[21] A. Bassamboo, S. Juneja, and A. Zeevi, "On the inefficiency of state-independent importance sampling in the presence of heavy tails," *Oper. Res. Lett.*, vol. 35, no. 2, pp. 251–260, 2007. [Online]. Available: http://dx.doi.org/10.1016/j.orl.2006.02.002

[22] J. Blanchet and H. Lam, "State-dependent importance sampling for rare-event simulation: An overview and recent advances," *Surveys in Operations Research and Management Science*, vol. 17, no. 1, pp. 38 – 59, 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1876735411000250

[23] L. V. Rozovskiĭ, "Probabilities of large deviations of sums of independent random variables with a common distribution function that belongs to the domain of attraction of a normal law," *Teor. Veroyatnost.*

i Primenen., vol. 34, no. 4, pp. 686–705, 1989. [Online]. Available: http://dx.doi.org/10.1137/1134080

[24] T. Mikosch and A. V. Nagaev, "Large deviations of heavy-tailed sums with applications in insurance," *Extremes*, vol. 1, no. 1, pp. 81–110, 1998. [Online]. Available: http://dx.doi.org/10.1023/A:1009913901219

[25] A. A. Borovkov and K. A. Borovkov, *Asymptotic analysis of random walks*, ser. Encyclopedia of Mathematics and its Applications. Cambridge: Cambridge University Press, 2008, vol. 118, heavy-tailed distributions, Translated from the Russian by O. B. Borovkova. [Online]. Available: http://dx.doi.org/10.1017/CBO9780511721397

[26] P. Embrechts, C. Klüppelberg, and T. Mikosch, *Modelling extremal events*, ser. Applications of Mathematics (New York). Berlin: Springer-Verlag, 1997, vol. 33, for insurance and finance.

[27] S. Asmussen and H. Albrecher, *Ruin probabilities*, 2nd ed., ser. Advanced Series on Statistical Science & Applied Probability, 14. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2010.

[28] S. Foss, D. Korshunov, and S. Zachary, *An introduction to heavy-tailed and subexponential distributions*, ser. Springer Series in Operations Research and Financial Engineering. New York: Springer, 2011. [Online]. Available: http://dx.doi.org/10.1007/978-1-4419-9473-8