Total variation approximations and conditional limit theorems for multivariate regularly varying random walks conditioned on ruin

JOSE BLANCHET[1] and JINGCHEN LIU[2]

[1]*IEOR Department, Columbia University, 500 West 120th Street, 340 W. Mudd Building, New York, NY 10027, USA*
[2]*Department of Statistics, Columbia University, 1255 Amsterdam Ave, New York, NY 10027, USA.*
E-mail: jcliu@stat.columbia.edu

We study a new technique for the asymptotic analysis of heavy-tailed systems conditioned on large deviations events. We illustrate our approach in the context of ruin events of multidimensional regularly varying random walks. Our approach is to study the Markov process described by the random walk conditioned on hitting a rare target set. We construct a Markov chain whose transition kernel can be evaluated directly from the increment distribution of the associated random walk. This process is shown to approximate the conditional process of interest in total variation. Then, by analyzing the approximating process, we are able to obtain asymptotic conditional joint distributions and a conditional functional central limit theorem of several objects such as the time until ruin, the whole random walk prior to ruin, and the overshoot on the target set. These types of joint conditional limit theorems have been obtained previously in the literature only in the one dimensional case. In addition to using different techniques, our results include features that are qualitatively different from the one dimensional case. For instance, the asymptotic conditional law of the time to ruin is no longer purely Pareto as in the multidimensional case.

Keywords: conditional distribution; heavy-tail; multivariate regularly variation; random walk

1. Introduction

The focus of this paper is the development of a precise asymptotic description of the distribution of a multidimensional regularly varying random walk (the precise meaning of which is given in Section 2) conditioned on hitting a rare target set represented as the union of half spaces. In particular, we develop tractable total variation approximations (in the sample path space), based on change-of-measure techniques, for such conditional stochastic processes. Using these approximations we are able to obtain, as a corollary, joint conditional limit theorems of specific objects such as the time until ruin, a Brownian approximation up to (just before) the time of ruin, the "overshoot", and the "undershoot". This is the first paper, as far as we know, that develops refined conditional limit theorems in a multidimensional ruin setting; results in one dimensional settings include, for instance, [5,12,17]; see also [4] for extensions concerning regenerative processes.

The techniques developed to obtain our results are also different from those prevalent in the literature and interesting qualitative features arise in multidimensional settings. For instance,

surprisingly, the asymptotic conditional time to ruin is no longer purely Pareto, as in the one dimensional case. A slowly-varying correction is needed in the multidimensional setting.

Other results in the case of multidimensional regularly varying random walks have been obtained by using a weak convergence approach (see [15]). In contrast to the weak convergence approach, which has been applied to non-Markovian settings [16], the techniques that we present here appear to be suited primarily to Markovian settings. On the other hand, the approach that we shall demonstrate allows to obtain finer approximations to conditional objects, for instance, total variation approximations, and conditional central limit theorems. Moreover, the present approach has also been applied to non-regularly varying heavy-tailed settings, at least in one dimension (see [10]). In addition, if there is a need to improve upon the quality of the approximations the method that we advocate readily provides Monte Carlo algorithms that can be shown to be optimal in a sense of controlling the relative mean squared error uniformly in the underlying large deviations parameter (see [9]). Standard approximation techniques for heavy-tailed large deviations cannot be easily translated into efficient Monte Carlo algorithms (see [3]).

The multidimensional problem that we consider here is a natural extension of the classical one dimensional random ruin problem in a so-called renewal risk model (cf. the texts of [1,2]). We consider a d-dimensional regularly varying random walk $S = (S_n : n \geq 1)$ with $S_0 = 0$ and drift $\eta \in \mathbb{R}^d$ so that $E S_n = n\eta \neq 0$. Define

$$T_{bA^*} = \inf\{n \geq 0 : S_n \in bA^*\},$$

where A^* is the union of half spaces and η points to the interior of some open cone that does not intersect A^*. The paper [14] notes that $P(T_{bA^*} < \infty)$ corresponds to the ruin probabilities for insurance companies with several lines of business. Using natural budget constraints related to the amount of money that can be transferred from one business line to another, it turns out that the target set takes precisely the form of the union of half spaces as we consider here.

Our goal is to illustrate new techniques that can be used to describe very precisely the conditional distribution of the heavy-tailed processes given $T_{bA^*} < \infty$. Our approximations allow to obtain, in a relatively easy way, extensions of previous results in the literature that apply only in the one dimensional settings. Asmussen and Klüppelberg [5] provide conditional limit theorems for the overshoot, the undershoot, and the time until ruin given the eventual occurrence of ruin. Similar results have been obtained recently in the context of Levy processes (see [17] and references there in). We apply our results here to obtain multidimensional analogues of their conditional limit theorems and additional refinements, such as conditional central limit theorems.

Our general strategy is based on the use of a suitable change of measure and later on coupling arguments. The idea is to approximate the conditional distribution of the underlying process step-by-step, in a Markovian way, using a mixture of a large increment that makes the random walk hit the target set and an increment that follows the nominal (original) distribution. The mixture probability is chosen depending on the current position of the random walk. Intuitively, given the current position, the selection of the mixture probability must correspond to the conditional probability of reaching the target set in the immediate next step given that one will eventually reach the target set. The conditional distribution itself is also governed by a Markov process, so the likelihood ratio (or Radon–Nikodym derivative) between the conditional distribution and our approximating distribution can be explicitly written in terms of the ratios of the corresponding

Markov kernels. By showing that the second moment of the likelihood ratio between the conditional distribution and our approximating distribution approaches unity as the rarity parameter $b \to \infty$, we can reach the desired total variation approximation (cf. Lemma 3). A crucial portion of our strategy involves precisely obtaining a good upper bound on the second moment of the likelihood ratio. The likelihood ratio is obtained out of a Markovian representation of two measures. It is natural to develop a Lyapunov-type criterion for the analysis of the second moment. This approach is pursued in Section 3, where we introduce our Lyapunov criterion and develop the construction of the associated Lyapunov function which allows us to bound the second moment of the likelihood ratio of interest as a function of the initial state of the process.

There are several interesting methodological aspects of our techniques that are worth emphasizing. First, the application of change-of-measure ideas is common in the light-tailed settings. However, it is not at all standard in heavy-tailed settings. A second interesting methodological aspect of our technique is the construction of the associated Lyapunov function. This step often requires a substantial amount of ingenuity. In the heavy-tailed setting, as we explain in Section 3, we can take advantage of the fluid heuristics and asymptotic approximations for this construction. This approach was introduced in [6] and has been further studied in [7,9] and [11]. In fact, many of our ideas are borrowed from the rare-event simulation literature, which is not surprising given that our strategy involves precisely the construction of a suitable change of measure and these types of constructions, in turn, lie at the heart of importance sampling techniques. The particular change of measure that we use is inspired by the work of [13] who applied it to the setting of one dimensional (finite) sums of regularly varying increments.

The approximation constructed in this paper is tractable in the sense that it is given by a Markovian description which is easy to describe and is explicit in terms of the increment distribution of the associated random walk (see (13) and Theorem 1). This tractability property has useful consequences both in terms of the methodological techniques and practical use. From a methodological standpoint, given the change-of-measure that we use to construct our Markovian description (basically following (6)), the result of convergence in total variation provides a very precise justification of the intuitive mechanism which drives the ruin in the heavy-tailed situations. In addition, the result allows to directly use this intuitive mechanism to provide functional probabilistic descriptions that, while less precise than total variation approximations, emphasize the most important elements that are present at the temporal scales at which ruin is expected to occur (if it happens at all). These functional results are given in Theorem 2. Our total variation approximation (Theorem 1) allows to construct a very natural coupling which makes the functional probabilistic descriptions given in Theorem 2 relatively straightforward in view of standard strong Brownian approximation results for random walks.

The tractability of our total variation approximation allows for a deep study of the random walk conditioned on bankruptcy, as mentioned earlier, via efficient Monte Carlo simulation, at scales that are finer than those provided by the existing functional limit theorems (cf. [15]). Using the techniques that we pursue here, the results in [10] establish *necessary and sufficient* conditions for optimal estimation of conditional expectations given bankruptcy using importance sampling; surprisingly, one can ensure finite expected termination time and asymptotically optimal relative variance control even when the zero variance change of measure has infinite expected termination time.

As mentioned earlier, we believe that the techniques that we consider here can find potential applications in Markovian settings beyond random walks. This is a research avenue that we are

currently exploring; see, for instance, [8], where we apply similar techniques to multi-queues. Additional results will be reported in the future.

The rest of the paper is organized as follows. In Section 2 we explain our assumptions, describe our approximating process, and state our main results. The estimates showing total variation approximation, which are based on the use of Lyapunov inequalities, are given in Section 3. Finally, the development of conditional functional central limit theorems is given in Section 4.

2. Problem setup and main results

2.1. Problem setup

Let $(X_n: n \geq 1)$ be a sequence of independent and identically distributed (i.i.d.) regularly varying random vectors taking values in $\mathbb{R}^d$. Let X be a generic random variable equal in distribution to X_i. The random vector X is said to have a multivariate regularly varying distribution if there exists a sequence $\{a_n: n \geq 1\}$, $0 < a_n \uparrow \infty$, and a non-null Random measure μ on the compactified and punctured space $\overline{\mathbb{R}}^d \setminus \{0\}$ with $\mu(\overline{\mathbb{R}}^d \setminus \mathbb{R}^d) = 0$ such that, as $n \to \infty$

$$nP\big(a_n^{-1}X \in \cdot\big) \overset{v}{\to} \mu(\cdot), \tag{1}$$

where "$\overset{v}{\to}$" refers to vague convergence. It can be shown that as $b \to \infty$,

$$\frac{P(X \in b\cdot)}{P(\|X\|_2 > b)} \overset{v}{\to} c\mu(\cdot)$$

for some $c > 0$ ([15], Remark 1.1). To simplify notation, a_n is chosen such that $nP(\|X\|_2 > a_n) \to 1$ and with this choice of a_n we have that $c = 1$.

The random vector X has a relatively very small probability of jumping into sets for which $\mu(B) = 0$. If $P(\|X\|_2 > b) = b^{-\alpha}L(b)$ for some $\alpha > 0$ and a slowly varying function $L(\cdot)$ (i.e., $L(tb)/L(b) \longrightarrow 1$ as $b \uparrow \infty$ for each $t > 0$), then we say that $\mu(\cdot)$ has (regularly varying) index α. For further information on multivariate regular variation see [21]; the definition provided above corresponds to the representation in Theorem 6.1, page 173, in [21]. Additional properties that we shall use in our development are discussed in the Appendix.

Define $S_n = X_1 + \cdots + X_n + S_0$ for $n \geq 1$. Throughout the rest of the paper, we shall use the notation $P_s(\cdot)$ for the probability measure on the path-space of the process $S = (S_n: n \geq 0)$ given that $S_0 = s$. Let $v_1^*, \ldots, v_m^* \in \mathbb{R}^d$ and $a_1^*, \ldots, a_m^* \in \mathbb{R}^+$. We define

$$A^* = \bigcup_{j=1}^{m} \{y: y^T v_j^* > a_j^*\} = \Big\{y: \max_{j=1}^{m}(y^T v_j^* - a_j^*) > 0\Big\}. \tag{2}$$

We set $T_{A^*} = \inf\{n \geq 0: S_n \in A^*\}$ and write $bA^* = \{y: y = bx, x \in A^*\}$. Note that

$$bA^* = \big\{z: r_b^*(z) > 0\big\},$$

where

$$r_b^*(z) \triangleq \max_{j=1}^{m}\big(z^T v_j^* - a_j^* b\big).$$

Finally, put

$$u_b^*(s) = P_s(T_{bA^*} < \infty).$$

We are concerned with the asymptotic conditional distribution of $(S_n\colon n \leq T_{bA^*})$ given that $T_{bA^*} < \infty$ as $b \nearrow \infty$. Throughout this paper, we impose the following two assumptions.

Assumption 1. X_n *has a continuous regularly varying distribution with index* $\alpha > 1$ *and* $EX_n = \eta = -\mathbf{1}$, *where* $\mathbf{1} = (1, \ldots, 1)^T \in \mathbb{R}^d$.

Assumption 2. *For each* j, $\eta^T v_j^* = -1$ *and* $\mu(A^*) > 0$.

Remark 1. Assumption 1 indicates that X_n has a continuous distribution. This can be dispensed by applying a smoothing kernel in the definition of the function $H(\cdot)$ introduced later. The assumption that $\eta = -\mathbf{1}$ is equivalent (up to a rotation and a multiplication of a scaling factor) to $EX_n = \eta \neq 0$, so nothing has been lost by imposing this condition. We also assume that $\eta^T v_j^* = -1$ for each j; this, again, can always be achieved without altering the problem structure by multiplying the vector v_j^* and a_j^* by a positive factor as long as $\eta^T v_j^* < 0$. Now, given that the random walk has drift η, it is not difficult to see geometrically that some conditions must be imposed on the v_i^*'s in order to have a meaningful large deviations situation (i.e., $u_b^*(0) \to 0$ as $b \to \infty$). In particular, we must have that A^* does not intersect the ray $\{t\eta\colon t > 0\}$. Otherwise, the Law of Large Numbers might eventually let the process hit the set bA^*. However, avoiding intersection with the ray $\{t\eta\colon t > 0\}$ is not enough to rule out some degenerate situations. For instance, suppose that $A^* = \{y\colon y^T v^* > 1\}$ with $\eta^T v^* = 0$ (i.e., the face of A^* is parallel to η); in this case Central Limit Theorem-type fluctuations might eventually make the random walk hit the target set. Therefore, in order to rule out these types of degenerate situations one requires $\eta^T v_j^* < 0$.

As mentioned earlier, we are interested in providing a tractable asymptotic description of the conditional distribution of $(S_n\colon n \leq T_{bA^*})$ given that $T_{bA^*} < \infty$ as $b \nearrow \infty$. It is well known that, given $T_{bA^*} < \infty$, the process $(S_n\colon n \leq T_{bA^*})$ is Markovian with transition kernel given by

$$K^*(s_0, ds_1) = P(X + s_0 \in ds_1)\frac{u_b^*(s_1)}{u_b^*(s_0)}.$$

Note that $K^*(\cdot)$ is a well defined Markov transition kernel because of the (harmonic) relationship

$$0 < u_b^*(s) = E_s\big[P_s(T_{bA^*} < \infty | X_1)\big] = E\big[u_b^*(s + X_1)\big].$$

The transition kernel $K^*(\cdot)$ is the Doob's h-transform of the original random walk kernel (the name h-transform is given after the harmonic property of the positive function $u_b^*(\cdot)$). The Markov kernel $K^*(\cdot)$ generates a measure $P_s^*(\cdot)$ on the σ-field $\mathcal{F}_{T_{bA^*}}$ generated by the X_k's up to time T_{bA^*}.

Our goal is to construct a *tractable* measure $\hat{P}_s$ on $\mathcal{F}_{T_{bA^*}}$ such that for each fixed s,

$$\lim_{b \to \infty} \sup_{B \in \mathcal{F}_{T_{bA^*}}} \big|\hat{P}_s(B) - P_s^*(B)\big| = 0. \tag{3}$$

Remark 2. The measures $\hat{P}_s$ and P_s^* certainly depend on the specific rarity parameter b. To simplify notation, we omit the index b in the notation of $\hat{P}_s$ and P_s^* when it does not cause confusion.

By tractability we mean that the $\hat{P}_s(\cdot)$ is constructed by a Markov transition kernel that can, in principle, be computed directly from the increment distribution of the random walk. The transition kernel associated to $\hat{P}_s(\cdot)$ will be relatively easy to manipulate and it will be explicit in terms of the associated increment distribution. Together with the strong mode of convergence implied in (3), we will be able to provide refinements to common results that are found in the literature concerning conditional limit theorems of specific quantities of interest (cf. [5,15]).

2.2. Elements of the approximations and main results

A natural strategy that one might pursue in constructing $\hat{P}_s(\cdot)$ consists in taking advantage of the approximations that are available for $u_b^*(s)$. Define

$$
\begin{aligned}
v_b^*(s) &= \int_0^\infty P_s\big(X + t\eta + s \in bA^*\big)\,dt \\
&= \int_0^\infty P_s\Big(\max_{i=1}^m\big[(s + X + t\eta)^T v_i^* - a_i^* b\big] > 0\Big)\,dt \qquad (4) \\
&= \int_0^\infty P_s\big(r_b^*(s + X) > t\big)\,dt = E\big(r_b^*(s + X)^+\big).
\end{aligned}
$$

Note that in the third equality above we have used that $\eta^T v_i^* = -1$. If Assumptions 1 and 2 are in place, it is well known that (see [14] and [15])

$$
u_b^*(s) = v_b^*(s)\big(1 + o(1)\big) \qquad (5)
$$

as $b \nearrow \infty$, uniformly over s in compact sets. This result, together with the form of the h-transform, suggests defining the kernel

$$
K_v(s_0, ds_1) = P(X + s_0 \in ds_1)\frac{v_b^*(s_1)}{w_b^*(s_0)},
$$

where

$$
w_b^*(s_0) = E v_b^*(s_0 + X)
$$

is introduced to make $K_v(\cdot)$ a well defined Markov transition kernel. It is reasonable to expect that $K_v(\cdot)$ and the corresponding probability measure on the sample path space, which we denote by $P_s^{(v)}(\cdot)$, will provide good approximations to both $K^*(\cdot)$ and $P_s^*(\cdot)$. This approach is natural and it has been successfully applied in the one dimensional setting in the context of subexponential increment distributions in [6]. However, in the multidimensional setting it is not entirely

straightforward to evaluate and manipulate either $v_b^*(\cdot)$ or $w_b^*(\cdot)$. Therefore, we shall follow a somewhat different approach.

Our strategy is inspired by the way in which ruin is intuitively expected to occur in the context of heavy-tailed increments; namely, the underlying random walk proceeds according to its nominal dynamics and all of a sudden a large jump occurs which causes ruin. This intuition is made more precise by the form of the kernel

$$\tilde{K}(s_0, ds_1) = p_b(s_0) P(X + s_0 \in ds_1) \frac{I(X + s_0 \in bA^*)}{P(X + s_0 \in bA^*)} + (1 - p_b(s_0)) P(X + s_0 \in ds_1)$$
$$= P(X + s_0 \in ds_1) \left\{ p_b(s_0) \frac{I(X + s_0 \in bA^*)}{P(X + s_0 \in bA^*)} + (1 - p_b(s_0)) \right\}, \tag{6}$$

where $p_b(s_0)$ will be suitably chosen so that

$$p_b(s_0) \approx P_{s_0}(X_1 + s_0 \in bA | T_{bA^*} < \infty). \tag{7}$$

In other words, $\tilde{K}(\cdot)$ is a mixture involving both the ruinous and the regular components. The mixture probability is chosen to capture the appropriate contribution of the ruinous component at every step.

We shall construct $\hat{P}_s(\cdot)$ by studying a family of transition kernels $\hat{K}(\cdot)$ that are very close to $\tilde{K}(\cdot)$. We will not work directly with $\tilde{K}(\cdot)$ to avoid some uniform integrability issues that arise in testing the Lyapunov bound to be described later in Lemma 3. The definition of $\hat{K}(\cdot)$ requires a modification of the target set. This modification will be convenient because of two reasons: first, to localize the analysis of our Lyapunov functions only in a suitable compact region that scales according to the parameter b; second, to apply a Taylor expansion in combination with the dominated convergence theorem. The Taylor expansion will be applied to a mollified version of the function $r_b^*(s)$ in the verification of the Lyapunov bound.

2.2.1. Enlargement procedure of the target region

First, given any $\delta \in (0, 1)$ we define $v_j^*(\delta) = (v_j^* + \delta \eta / \|\eta\|_2^2)/(1 - \delta)$, and observe that $\eta^T v_j^*(\delta) = -1$. We then write, given $\beta > 0$,

$$A = A^* \cup \left(\bigcup_{j=1}^{m} \{y: y^T v_j^*(\delta) > a_j^*\} \right) \cup \left(\bigcup_{i=1}^{d} \{y: y_i \geq \beta\} \right). \tag{8}$$

To simplify the notation let $e_i \in \mathbb{R}^d$ be the vector whose ith component is equal to one and the rest of the components are zero, and express A in the same form as we do for A^*. We write

$$v_j = \begin{cases} v_j^*, & 1 \leq j \leq m, \\ v_j^*(\delta), & m+1 \leq j \leq 2m, \\ e_i, & 2m+1 \leq j \leq 2m+d, \end{cases} \qquad a_j = \begin{cases} a_j^*, & 1 \leq j \leq m, \\ a_{j-m}^*, & m+1 \leq j \leq 2m, \\ \beta, & 2m+1 \leq j \leq 2m+d. \end{cases}$$

We then have that $A = \bigcup_{j=1}^{2m+d} \{y: y^T v_j > a_j\}$. Analogous approximations such as (4) and (5) are applicable. The addition of the vectors v_j for $j \geq m+1$ will be convenient in order to analyze

a certain Lyapunov inequality in a compact set. Now, note that if $u_b(s) = P(T_{bA} < \infty)$, then

$$u_b(s)\big(1 + o(1)\big) = v_b(s) \triangleq \int_0^\infty P(X + t\eta + s \in bA)\, \mathrm{d}t. \tag{9}$$

Moreover, note that

$$u_b^*(s) \le u_b(s)$$

and that, for each fixed s,

$$\lim_{\delta \to 0} \lim_{\beta \to \infty} \lim_{b \to \infty} \frac{u_b^*(s)}{u_b(s)} = 1. \tag{10}$$

Our strategy consists in first obtaining results for the event that $T_{bA} < \infty$. Then, thanks to (10), we can select β arbitrarily large and δ arbitrarily small to obtain our stated results for the conditional distribution of the walk given $T_{bA^*} < \infty$, which is our event of interest.

Now, define

$$r_b(z) \triangleq \max_{j=1}^{2m+d} \big\{\big(z^T v_j - a_j b\big)\big\},$$

and just as we obtained for (4), we can conclude that

$$v_b(s) = E\big(r_b(s + X)^+\big). \tag{11}$$

Given $a \in (0, 1)$ the enlarged region takes the form

$$A_{b,a}(s_0) = \Big\{s_1 \colon \max_{j=1}^{2m+d}\big[(s_1 - s_0)^T v_j - a\big(a_j b - s_0^T v_j\big)\big] > 0\Big\}. \tag{12}$$

2.2.2. *The family of transition kernels $\hat{K}(\cdot)$*

We now describe our proposed approximating kernel $\hat{K}(\cdot)$ based on the enlarged target region. Given $a \in (0, 1)$ we put

$$\hat{K}(s_0, \mathrm{d}s_1) = P(X + s_0 \in \mathrm{d}s_1)\bigg\{\frac{p_b(s_0)I(s_1 \in A_{b,a}(s_0) > 0)}{P(s_0 + X \in A_{b,a}(s_0) > 0)} + \big(1 - p_b(s_0)\big)\bigg\}. \tag{13}$$

The final ingredient in the description of our approximating kernel, and therefore of our approximating probability measure $\hat{P}_s(\cdot)$, corresponds to the precise description of $p_b(s_0)$ and the specification of $a \in (0, 1)$. The scalar a eventually will be chosen arbitrarily close to 1. As we indicated before, in order to follow the intuitive description of the most likely way in which large deviations occur in heavy-tailed settings, we should guide the selection of $p_b(s_0)$ via (7).

Using (9) and our considerations about $v_b(s_0)$, we have that as $b \to \infty$

$$P_{s_0}(X_1 + s_0 \in bA | T_{bA} < \infty) = \frac{P(r_b(s_0 + X) > 0)}{v_b(s_0)}\big(1 + o(1)\big). \tag{14}$$

The underlying approximation (14) deteriorates when $r_b(s_0)$ is not too big or, equivalently, as s_0 approaches the target set bA. In this situation, it is not unlikely that the nominal dynamics will make the process hit bA. Therefore, when s_0 is close enough to bA, we would prefer to select $p_b(s_0) \approx 0$. Due to these considerations and given the form of the jump set specified in $\hat{K}(\cdot)$ above we define

$$p_b(s) = \min\left(\frac{\theta P(s_0 + X \in A_{b,a}(s_0))}{v_b(s_0)}, 1\right) I\left(r_b(s) \leq -\delta_2 b\right) \tag{15}$$

for $\delta_2 > 0$ chosen small enough and $\theta, a \in (0, 1)$ chosen close enough to 1. The selection of all these constants will be done in our development.

2.2.3. *The statement of our main results*

Before we state our main result, we need to describe the probability measure in path space that we will use to approximate

$$P_s^*(S \in \cdot) \triangleq P(S \in \cdot | T_{bA^*} < \infty, S_0 = s),$$

in total variation. Given $\gamma > 0$ define

$$\Gamma = \left\{y\colon y^T \eta \geq \gamma\right\}$$

and set $T_{bA} = \inf\{n \geq 0\colon S_n \in bA\}$, and $T_{b\Gamma} = \inf\{n \geq 0\colon S_n \in b\Gamma\}$. We now define the change of measure that we shall use to approximate $P_s^*(\cdot)$ in total variation.

Definition 1. *Let $\hat{P}_s(\cdot)$ be defined as the measure generated by transitions according to $\hat{K}(\cdot)$ up to time $T_{bA} \wedge T_{b\Gamma}$, with mixture probability as defined in* (15), *and transitions according to $K(\cdot)$ for the increments $T_{bA} \wedge T_{b\Gamma} + 1$ up to infinity.*

We now state our main result.

Theorem 1. *For every $\varepsilon > 0$ there exists $\theta, a, \delta, \delta_2 \in (0, 1)$ (θ, a sufficiently close to 1 and δ, δ_2 sufficiently close to zero), and $\beta, \gamma, b_0 > 0$ sufficiently large so that if $b \geq b_0$*

$$\sup_{B \in \mathcal{F}} \left|\hat{P}_0(B) - P_0^*(B)\right| \leq \varepsilon,$$

where $\mathcal{F} = \sigma(\bigcup_{n=0}^{\infty} \sigma\{S_k\colon 0 \leq k \leq n\})$.

In order to illustrate an application of the previous result, we have the next theorem which follows without much additional effort, as a corollary to Theorem 1. The statement of the theorem, however, requires some definitions that we now present.

Because of regular variation, we can define for any $a_1^*, \ldots, a_m^* > 0$

$$\lim_{b \to \infty} \frac{P(\max_{j=1}^m [X^T v_j^* - a_j^* b] > 0)}{P(\|X\|_2 > b)} = \kappa^*\left(a_1^*, \ldots, a_m^*\right), \tag{16}$$

where for any $t \geq 0$

$$\kappa\left(a_1^* + t, \ldots, a_m^* + t\right) \triangleq \mu\left(\left\{y: \max_{j=1}^m \left(y^T v_j^* - a_j^*\right) > t\right\}\right).$$

Using this representation, we obtain that

$$
\begin{aligned}
v_b^*(s) &= \int_0^\infty P\left(\max_{j=1}^m \left[(s+X)^T v_j^* - a_j^* b\right] > t\right) dt \\
&= \left(1 + o(1)\right) P\left(\|X\|_2 > b\right) \\
&\quad \times \int_0^\infty \kappa\left(a_1^* - b^{-1} s^T v_1^* + b^{-1} t, \ldots, a_m^* - b^{-1} s^T v_m^* + b^{-1} t\right) dt \\
&= \left(1 + o(1)\right) b P\left(\|X\|_2 > b\right) \int_0^\infty \kappa\left(a_1^* - b^{-1} s^T v_1^* + t, \ldots, a_m^* - b^{-1} s^T v_m^* + t\right) dt \\
&= \left(1 + o(1)\right) b P\left(\|X\|_2 > b\right) \int_0^\infty \kappa\left(a_1^* + t, \ldots, a_m^* + t\right) dt
\end{aligned}
\tag{17}
$$

as $b \to \infty$ uniformly over s in a compact set. Actually, an extension to approximation (17) to the case in which $s = \mathrm{O}(b)$ is given in the Appendix; see Lemma 12. To further simplify the notation, we write

$$\kappa_{\mathbf{a}^*}(t) = \kappa\left(a_1^* + t, \ldots, a_m^* + t\right), \tag{18}$$

where $\mathbf{a}^* = (a_1^*, \ldots, a_m^*)$.

Theorem 2. *For each $z > 0$ let $Y^*(z)$ be a random variable with distribution given by*

$$P\left(Y^*(z) \in B\right) = \frac{\mu\left(B \cap \{y: \max_{j=1}^m [y^T v_j^* - a_j^*] \geq z\}\right)}{\mu\left(\{y: \max_{j=1}^m [y^T v_j^* - a_j^*] \geq z\}\right)}.$$

In addition, let Z^* be a positive random variable following distribution

$$P\left(Z^* > t\right) = \exp\left\{-\int_0^t \frac{\kappa_{\mathbf{a}^*}(s)}{\int_s^\infty \kappa_{\mathbf{a}^*}(u)\, du}\, ds\right\}$$

for $t \geq 0$ where $\kappa_{\mathbf{a}^*}(\cdot)$ is as defined in (18). Then if $S_0 = 0$ and $\alpha > 2$, we have that

$$\left(\frac{T_{bA^*}}{b}, \frac{S_{u T_{bA^*}} - u T_{bA^*} \eta}{\sqrt{T_{bA^*}}}, \frac{X_{T_{bA^*}}}{b}\right) \Rightarrow \left(Z^*, CB\left(u Z^*\right), Y^*\left(Z^*\right)\right)$$

in $\mathbb{R} \times D[0,1) \times \mathbb{R}^d$, where $CC^T = \mathrm{Var}(X)$, $B(\cdot)$ is a d-dimensional Brownian motion with identity covariance matrix, $B(\cdot)$ is independent of Z^* and $Y^*(Z^*)$.

Remark 3. The random variable Z^* (multiplied by a factor of b) corresponds to the asymptotic time to ruin. In the one dimensional setting, Z^* follows a Pareto distribution with index $\alpha - 1$. The reason is that in the one dimensional case

$$\kappa_{\mathbf{a}^*}(s) = \frac{\alpha - 1}{s} \int_s^\infty \kappa_{\mathbf{a}^*}(u) \, \mathrm{d}u.$$

This no longer can be ensured in the multidimensional case. Nevertheless, Z^* is still regularly varying with index $\alpha - 1$.

Remark 4. If $\alpha \in (1, 2]$, then our analysis allows to conclude that

$$\left(\frac{T_{bA^*}}{b}, \frac{S_{uT_{bA^*}}}{T_{bA^*}}, \frac{X_{T_{bA^*}}}{b} \right) \Rightarrow \left(Z^*, u\eta, Y^*(Z^*) \right)$$

in $\mathbb{R} \times D[0, 1) \times \mathbb{R}^d$ as $b \to \infty$.

3. Total variation approximations and Lyapunov inequalities

In this section, we provide the proof of Theorem 1. First, it is useful to summarize some of the notation that has been introduced so far.

1. The set A^* be the target set, $v_b^*(s)$ be the approximation of $P_s(T_{bA^*} < \infty)$, and $P_s^*(\cdot) = P(\cdot|T_{bA^*} < \infty)$ be the corresponding conditional distribution.
2. The set A is an enlargement of A^* and depends on δ and β; $v_b(s)$ be the approximation of $P_s(T_{bA} < \infty)$.
3. The set $\Gamma = \{y\colon y^T \eta \geq \gamma\}$ will be used to define an auxiliary conditional distribution below.
4. Under the distribution $\hat{P}_s(\cdot)$ in path space increments follow the transition kernel $\hat{K}$ up to time $T_{bA} \wedge T_{b\Gamma}$.

Now, we shall outline the program that will allow us to proof Theorem 1. The program contains three parts. The first part consists in introducing an auxiliary conditional distribution involving a finite horizon. To this end, we define

$$P_s^{\&}(\cdot) \triangleq P(S \in \cdot | T_{bA} \leq T_{b\Gamma}, S_0 = s). \tag{19}$$

Eventually, as we shall explain, we will select $\delta > 0$ is small enough, β and γ are large enough. The second part consists in showing that $P_s^{\&}$ and P_s^* are close in total variation; this will be done in Lemma 1. Finally, in the third part we show that $P_s^{\&}$ can be approximated by $\hat{P}_s$ in total variation; this will be done in Proposition 1. Theorem 1 then follows directly by combining Lemma 1 and Proposition 1.

In order to carry out the third part of our program, namely, approximating $P_s^{\&}$ by $\hat{P}_s$ in total variation. A natural approach, which we shall follow, is to argue that $\mathrm{d}P_s^{\&}/\mathrm{d}\hat{P}_s$ is close to unity. We define

$$\beta(s) \triangleq \hat{E}_s\left((\mathrm{d}P_s/\mathrm{d}\hat{P}_s)^2 I(T_{bA} \leq T_{b\Gamma}) \right) = P_s(T_{bA} \leq T_{b\Gamma})^2 \hat{E}_s\left((\mathrm{d}P_s^{\&}/\mathrm{d}\hat{P}_s)^2 \right),$$

note that $\beta(s) \geq P_s(T_{bA} \leq T_{b\Gamma})^2$ (by Jensen's inequality). As we shall verify in Lemma 2, if we are able to show that $\beta(0) \leq P_0(T_{bA} \leq T_{b\Gamma})^2(1 + \varepsilon)$ (thus showing that $dP_s^\&/d\hat{P}_s$ is close to unity) then we will be able to claim that $P_0^\&$ and $\hat{P}_0$ are close in total variation.

Obtaining a useful bound for $\beta(s)$ is the most demanding part of the whole program. The strategy relies on the so-called Lyapunov inequalities. The idea is to find a function $g(\cdot)$, which is called a Lyapunov function, satisfying certain criteria specified in Lemma 3 in order to ensure that $g(s) \geq \beta(s)$. Now, constructing Lyapunov functions is not easy, however, we eventually wish to enforce an upper bound corresponding to the behavior of $P_s(T_{bA} \leq T_{b\Gamma})^2$, so it makes sense to use $v_b(\cdot)^2$ (recall equation (5) and (11)) as starting template for $g(\cdot)$. In the process of verifying that criteria in Lemma 3 it is useful to ensure some smoothness properties of a candidate Lyapunov function. So, a mollification procedure is performed to the template suggested by $v_b(\cdot)^2$. The verification of the criteria in Lemma 3 is pursued in Section 3.2.

Now we start executing the first part of the previous program. Before we start it is useful to remark that in our definition of $P_s^\&(\cdot)$ (see (19)), we write $T_{bA} \leq T_{b\Gamma}$ rather than $T_{bA} < T_{b\Gamma}$. This distinction is important in the proof of the next result because, due to the geometry of the sets A and A^*, on the set $T_{bA} > T_{b\Gamma}$ we can guarantee that $S_{T_{b\Gamma}}$ is sufficiently far away from the set bA^*.

Lemma 1. *For each $\varepsilon > 0$ we can find $\delta, \beta, \gamma > 0$ such that*

$$\overline{\lim_{b \to \infty}} \left| \frac{P_0(T_{bA^*} < \infty)}{P_0(T_{bA} \leq T_{b\Gamma})} - 1 \right| \leq \varepsilon;$$

moreover, for b sufficiently large

$$\sup_{B \in \mathcal{F}} \left| P_0^*(B) - P_0^\&(B) \right| \leq \varepsilon.$$

Proof. We prove the first part of the lemma by establishing an upper and lower bound of $P_0(T_{bA^*} < \infty)$, respectively.

Upper bound. Observe that

$$P_0(T_{bA^*} < \infty) = P_0(T_{bA^*} < \infty, T_{bA} \leq T_{b\Gamma}) + P_0(T_{bA^*} < \infty, T_{bA} > T_{b\Gamma})$$

$$\leq P_0(T_{bA} \leq T_{b\Gamma}) + \sup_{\{s: \, s \in b\Gamma, s \notin bA\}} P_s(T_{bA^*} < \infty).$$

Note that if $s \in b\Gamma$ and $s \notin bA$ then $s^T \eta \geq \gamma b$, $s^T v_i = s^T v_i^* + \delta s^T \eta / \|\eta\|_2^2 \leq a_i b(1 - \delta)$ for $i \in \{1, \ldots, m\}$. Therefore,

$$s^T v_i^* \leq a_i b(1 - \delta) - \delta s^T \eta / \|\eta\|_2^2 \leq a_i b - \delta(\gamma + a_i \|\eta\|_2^2)b / \|\eta\|_2^2.$$

Thus, $\sup_{s \in b\Gamma \setminus bA} s^T v_i^* - a_i b \leq -\delta(\gamma + a_i \|\eta\|_2^2)b / \|\eta\|_2^2$. Given $\varepsilon > 0$, after choosing $\delta > 0$, we can select γ large enough so that

$$\sup_{\{s: \, s \in b\Gamma, s \notin bA\}} P_s(T_{bA^*} < \infty) \leq \varepsilon P_0(T_{bA^*} < \infty).$$

Therefore, we have that given $\varepsilon > 0$, we can select $\delta > 0$ sufficiently small and γ large enough so that for all b large enough

$$P_0(T_{bA^*} < \infty) \leq P_0(T_{bA} \leq T_{b\Gamma}) + \varepsilon P_0(T_{bA^*} < \infty),$$

which yields an upper bound of the form

$$P_0(T_{bA^*} < \infty)(1 - \varepsilon) \leq P_0(T_{bA} \leq T_{b\Gamma}).$$

Lower bound. Notice that $P(T_{b\Gamma} < \infty) = 1$. Then, we have that

$$P_0(T_{bA} \leq T_{b\Gamma}) \leq P_0(T_{bA} < \infty) = v_b(0)\big(1 + \mathrm{o}(1)\big) \tag{20}$$

as $b \to \infty$. In addition, by the asymptotic approximation of the first passage time probability, Lemmas 11 and 12 in the Appendix, given $\varepsilon > 0$ we can select $\delta, \beta > 0$ such that

$$1 \leq \varlimsup_{b \to \infty} \frac{v_b(0)}{v_b^*(0)} = \varlimsup_{b \to \infty} \frac{P_0(T_{bA} < \infty)}{P_0(T_{bA^*} < \infty)} \leq 1 + \varepsilon. \tag{21}$$

Thus, we obtain that as $b \to \infty$

$$P_0(T_{bA} \leq T_{b\Gamma}) \leq \big(1 + \varepsilon + \mathrm{o}(1)\big) P_0(T_{bA^*} < \infty).$$

We conclude the lower bound and the first part of the lemma.

The second part of the lemma follows as an easy consequence of the first part. $\qquad\square$

Throughout our development, we then concentrate on approximating in total variation $P_s^{\&}(\cdot)$. We will first prove the following result.

Proposition 1. *For all $\varepsilon, \delta, \beta > 0$ there exists θ, a sufficiently close to 1 from below, δ_2 sufficiently small, and γ, b sufficiently large such that*

$$\sup_{B \in \mathcal{F}} \big|\hat{P}_0(B) - P_0^{\&}(B)\big| \leq \varepsilon,$$

where $\mathcal{F} = \sigma\big(\bigcup_{n=0}^{\infty} \sigma\{S_k : 0 \leq k \leq n\}\big)$.

As noted earlier Proposition 1 combined with Lemma 1 yields the proof of Theorem 1. To provide the proof of Proposition 1, we will take advantage of the following simple yet powerful observation (see also [19]).

Lemma 2. *Let Q_0 and Q_1 be probability measures defined on the same σ-field $\mathcal{G}$ and such that $\mathrm{d}Q_1 = M^{-1}\,\mathrm{d}Q_0$ for a positive r.v. $M > 0$. Suppose that for some $\varepsilon > 0$, $E^{Q_1}(M^2) = E^{Q_0}M \leq 1 + \varepsilon$. Then*

$$\sup_{B \in \mathcal{G}} \big|Q_1(B) - Q_0(B)\big| \leq \varepsilon^{1/2}.$$

Proof. Note that

$$\left|Q_1(B) - Q_0(B)\right| = \left|E^{Q_1}(1 - M; B)\right|$$

$$\leq E^{Q_1}\left(|M - 1|\right) \leq E^{Q_1}\left[(M-1)^2\right]^{1/2} = \left(E^{Q_1}M^2 - 1\right)^{1/2} \leq \varepsilon^{1/2}. \quad \square$$

Lemma 2 will be used to prove Proposition 1. In particular, we will first apply Lemma 2 by letting $\mathcal{G} = \mathcal{F}_{T_{bA} \wedge T_{b\Gamma}}$, $Q_1 = \hat{P}_s$, and $Q_0 = P_s^\&(\cdot)$. The program proceeds as follows. With $\hat{K}$ defined as in (13), we let

$$\hat{k}(s_0, s_1) = \frac{P(X + s_0 \in ds_1)}{\hat{K}(s_0, ds_1)}$$

$$= \left\{\frac{p_b(s_0)I(s_1 \in A_{b,a}(s_0))}{P(s_0 + X \in A_{b,a}(s_0))} + \left(1 - p_b(s_0)\right)\right\}^{-1}$$

$$= \frac{P(s_0 + X \in A_{b,a}(s_0))I(s_1 \in A_{b,a}(s_0))}{p_b(s_0) + (1 - p_b(s_0))P(s_0 + X \in A_{b,a}(s_0))} \tag{22}$$

$$+ \frac{1}{(1 - p_b(s_0))}I\left(s_1 \notin A_{b,a}(s_0)\right).$$

Observe that on $\mathcal{G}$ we have

$$\frac{dP_s^\&}{d\hat{P}_s} = \frac{I(T_{bA} \leq T_{b\Gamma})}{P_s(T_{bA} \leq T_{b\Gamma})} \times \frac{dP_s}{d\hat{P}_s} = \frac{I(T_{bA} \leq T_{b\Gamma})}{P_s(T_{bA} \leq T_{b\Gamma})} \prod_{j=0}^{T_{bA}-1} \hat{k}(S_j, S_{j+1}).$$

Therefore, according to Lemma 2, it suffices to show that

$$\hat{E}_s\left(\left(\frac{dP_s^\&}{d\hat{P}_s}\right)^2\right)$$

$$= \frac{1}{P_s(T_{bA} \leq T_{b\Gamma})^2}\hat{E}_s\left(\left(\frac{dP_s}{d\hat{P}_s}\right)^2 I(T_{bA} \leq T_{b\Gamma})\right)$$

$$= \frac{1}{P_s(T_{bA} \leq T_{b\Gamma})^2}E_s\left(\prod_{j=0}^{T_{bA}-1} \hat{k}(S_j, S_{j+1})I(T_{bA} \leq T_{b\Gamma})\right) \leq 1 + \varepsilon$$

for all b sufficiently large. Consequently, we must be able to provide a good upper bound for the function

$$\hat{\beta}_b(s) \triangleq E_s\left(\prod_{j=0}^{T_{bA}-1} \hat{k}(S_j, S_{j+1})I(T_{bA} \leq T_{b\Gamma})\right).$$

In order to find an upper bound for $\hat{\beta}_b(s)$ we will construct an appropriate Lyapunov inequality based on the following lemma, which follows as in Theorem 2 part (iii) of [6]; see also Theorem 2.6 in [18].

Lemma 3. *Suppose that C and B are given sets. Let $\tau_B = \inf\{n\colon S_n \in B\}$ and $\tau_C = \inf\{n\colon S_n \in C\}$ be the first passage times. Assume that $g(\cdot)$ is a non-negative function satisfying*

$$g(s) \geq E_s\big(\hat{k}(s, S_1)g(S_1)\big) \tag{23}$$

for $s \notin C \cup B$. Then,

$$g(s) \geq E_s\left(g(S_{\tau_C}) \prod_{j=0}^{\tau_C - 1} \hat{k}(S_j, S_{j+1}) I(\tau_C \leq \tau_B, \tau_C < \infty)\right).$$

Furthermore, let $h(\cdot)$ be any non-negative function and consider the expectation

$$\hat{\beta}_b^h(s) = E_s\left(h(S_{\tau_C}) \prod_{j=0}^{\tau_C - 1} \hat{k}(S_j, S_{j+1}) I(\tau_C \leq \tau_B, \tau_C < \infty)\right).$$

If in addition to (23) we have that $g(s) \geq h(s)$ for $s \in C$, then we conclude that

$$g(s) \geq \hat{\beta}_b^h(s).$$

We will construct our Lyapunov function in order to show that for any given $\varepsilon > 0$

$$\varlimsup_{b \to \infty} \frac{\hat{\beta}_b^h(0)}{P_0(T_{bA} \leq T_{b\Gamma})^2} \leq 1 + \varepsilon, \tag{24}$$

with $h \approx 1$. Note that given any $\varepsilon > 0$ we can select $\gamma > 0$ sufficiently large so that for all b large enough

$$P_0(T_{bA} \leq T_{b\Gamma}) \leq P_0(T_{bA} < \infty) \leq (1 + \varepsilon) P_0(T_{bA} \leq T_{b\Gamma}).$$

Thus, using a completely analogous line of thought leading to (4) we would like to construct a Lyapunov function

$$g_b(s) \approx v_b^2(s) \triangleq E\big(r_b(s + X)^+\big)^2. \tag{25}$$

If such a selection of a Lyapunov function is applicable then, given $\varepsilon > 0$, if γ is chosen sufficiently large, we would be able to conclude the bound (24).

3.1. Mollification of $r_b(s)$ and proposal of $g_b(s)$

In order to verify the Lyapunov inequality from Lemma 3 it will be useful to perform a Taylor expansion of the function $g_b(\cdot)$. Since the function $r_b(s + X)^+$ is not smooth in s, we will first perform a mollification procedure. Given $c_0 > 0$ define

$$\varrho_b(s) = c_0 \log\left(\sum_{j=1}^{2m+d} \exp([s^T v_j - a_j b]/c_0)\right)$$

and note that

$$r_b(s) \le \varrho_b(s) \le r_b(s) + c_0 \log(2m + d). \tag{26}$$

Then, for $\delta_0 > 0$ let

$$d(x) = \begin{cases} 0, & x \le -\delta_0, \\ (x + \delta_0)^2/(4\delta_0), & |x| \le \delta_0, \\ x, & x \ge \delta_0. \end{cases}$$

Further note that $d(\cdot)$ is continuously differentiable with derivative $d'(\cdot)$ given by the function

$$d'(x) = \frac{x + \delta_0}{2\delta_0} I\big(|x| \le \delta_0\big) + I(x \ge \delta_0) \le I(x \ge -\delta_0)$$

and that

$$x^+ \le d(x) \le (x + \delta_0)^+. \tag{27}$$

Note that the functions $\varrho_b(\cdot)$ and $d(\cdot)$ depend on c_0 and δ_0, respectively; and that we have chosen to drop this dependence in our notation. The selection of δ_0 is quite flexible given that we will eventually send $b \to \infty$. We choose

$$c_0 \triangleq c_0(b) = \max\big(b^{(3-\alpha)/2}, \tilde{c}_0\big) \tag{28}$$

for some constant $\tilde{c}_0 > 0$ chosen sufficiently large. Note that we have

$$b^{2-\alpha} = \mathrm{o}\big(c_0(b)\big), \qquad c_0(b) = \mathrm{o}(b).$$

Given c_0 and δ_0 first we define

$$H_b(s) = E\big[d\big(\varrho_b(s + X)\big)\big]. \tag{29}$$

We will see that the asymptotics of $H_b(s)$ as b goes to infinity are independent of c_0 and δ_0. Indeed, observe, using inequality (27) that

$$E\big[\varrho_b^+(s + X)\big] = \int_0^\infty P\big(\varrho_b(s + X) > t\big)\,\mathrm{d}t \le H_b(s)$$

$$\le E\big[\big(\varrho_b(s + X) + \delta_0\big)^+\big] = \int_0^\infty P\big(\varrho_b(s + X) > t - \delta_0\big)\,\mathrm{d}t. \tag{30}$$

Therefore, using (26), the inequalities (30) and basic properties of regularly varying functions (i.e., that regularly varying functions possess long tails), we obtain that for any given $\delta_0 > 0$

$$H_b(s) = \big(1 + \mathrm{o}(1)\big) \int_0^\infty P\big(\varrho_b(s + X) > t\big)\,\mathrm{d}t$$

$$= \big(1 + \mathrm{o}(1)\big) \int_0^\infty P\big(\varrho_b(s + X) > t - \delta_0\big)\,\mathrm{d}t \tag{31}$$

$$= \left(1 + o(1)\right) \int_0^\infty P\left(\max_{i=1}^{2m+d}\left\{(s+X)^T v_i - a_i b\right\} > t + o(b)\right) dt$$

$$= \left(1 + o(1)\right) v_b(s) = \left(1 + o(1)\right) P_s\left(T_{bA} < \infty\right)$$

as $b \to \infty$ uniformly over s in any compact set.

Finally, $H_b(\cdot)$ is *twice* continuously differentiable. This is a desirable property for the verification of our Lyapunov inequality. Therefore, *we define the candidate of Lyapunov function, $g_b(\cdot)$ via*

$$g_b(s) = \min\left(c_1 H_b(s)^2, 1\right),$$

where $c_1 \in (1, \infty)$ will be chosen arbitrarily close to 1 if b is sufficiently large. The intuition behind the previous selection of $g_b(s)$ has been explained in the argument leading to (25).

3.2. Verification of the Lyapunov inequality

We first establish the Lyapunov inequality (23) on the region $s \notin b\Gamma$ and $r_b(s) \leq -\delta_2 b$ for $\delta_2 > 0$ suitably small and b sufficiently large. If $r_b(s) \leq -\delta_2 b$ for b large enough so that $g_b(s) < 1$ then, using expression (22), we note that inequality (23) is equivalent to

$$J_1 + J_2 \leq 1, \tag{32}$$

where

$$J_1 = E\left(\frac{g_b(s+X)}{g_b(s)}; s+X \in A_{b,a}(s)\right) \times \frac{P(s+X \in A_{b,a}(s))}{p_b(s) + (1 - p_b(s))P(s+X \in A_{b,a}(s))}$$

and

$$\left(1 - p_b(s)\right) J_2 = E\left(\frac{g_b(s+X)}{g_b(s)}; s+X \notin A_{b,a}(s)\right).$$

In order to verify (32) on the region $s \notin b\Gamma$, $r_b(s) \leq -\delta_2 b$ we will apply a Taylor expansion to the function $g_b(\cdot)$. This Taylor expansion will be particularly useful for the analysis of J_2. The following result, which summarizes useful properties of the derivatives of $\varrho_b(\cdot)$ and $d(\cdot)$, will be useful.

Lemma 4. *Let*

$$w_j(s) = \frac{\exp((s^T v_j - a_j b)/c_0)}{\sum_{i=1}^{2m+d} \exp((s^T v_i - a_i b)/c_0)}.$$

Then,

(i) $\nabla \varrho_b(s) = \sum_{j=1}^{2m+d} v_j w_j(s)$,

(ii) $\Delta \varrho_b(s) = \sum_{j=1}^{2m+d} w_j(s)(1 - w_j(s)) v_j v_j^T / c_0$.

Proof. Item (i) follows from basic calculus. Part (ii) is obtained by noting that

$$\nabla w_j(s) = w_j(s)\nabla \log w_j(s)$$
$$= w_j(s)\big(v_j^T/c_0 - w_j(s)v_j^T/c_0\big) = w_j(s)\big(1 - w_j(s)\big)v_j^T/c_0.$$

Therefore,

$$\Delta \varrho_b(s) = \sum_{j=1}^{2m+d} w_j(s)\big(1 - w_j(s)\big)v_j v_j^T/c_0$$

and the result follows. □

Using the previous lemma it is a routine application of the dominated convergence theorem to show that

$$\nabla H_b(s) = E\big(d'\big(\varrho_b(s + X)\big)\nabla \varrho_b(s + X)\big), \tag{33}$$

where $d'(\cdot)$ denotes the derivative of $d(\cdot)$, and the gradient $\nabla \varrho_b(\cdot)$ is encoded as a column vector. Similarly, the Hessian matrix of $H_b(\cdot)$ is given by

$$\Delta H_b(s) = E\big(\Delta \varrho_b(s + X)d'\big(\varrho_b(s + X)\big) + d''\big(\varrho_b(s + X)\big)\nabla \varrho_b(s + X)\nabla \varrho_b(s + X)^T\big). \tag{34}$$

This will be useful in our technical development.

We now are ready to provide an estimate for the term J_2.

Lemma 5. *For every* $\varepsilon \in (0, 1/2)$, $\delta_2 \in (0, 1)$, $c_1 \in (1, \infty)$ *there exists* $a \in (0, 1)$ *and* $b_0 > 0$ *(depending on* ε, δ, v_i's, γ, δ_0) *such that if* $b \geq b_0$, $s \notin b\Gamma \cup bA$ *and* $r_b(s) \leq -\delta_2 b$ *then*

$$J_2\big(1 - p_b(s)\big) \leq 1 - (2 - 3\varepsilon)\frac{P(\exists j:\ X^T v_j \geq a(a_j b - s^T v_j))}{H_b(s)}.$$

Proof. Recall that

$$\big(1 - p_b(s)\big)J_2 = E\left(\frac{g_b(s + X)}{g_b(s)}; s + X \notin A_{b,a}(s)\right)$$

$$= \int_{\mathbb{R}^d} \frac{g_b(s + x)}{g_b(s)} I\big(s + x \notin A_{b,a}(s)\big)P(X \in \mathrm{d}x) \tag{35}$$

$$= \int_{\mathbb{R}^d} \frac{g_b(s + x)}{g_b(s)} I\big(s + x \notin A_{b,a}(s), \|x\|_2 < b\big)P(X \in \mathrm{d}x) + \mathrm{o}\big(b^{-1}\big).$$

We will analyze the integrand above. Keep in mind that $s + x \notin A_{b,a}(s)$. Given $c_1 > 0$ fixed, note that $g_b(s) < 1$ whenever $r_b(s) \leq -\delta_2 b$ assuming that b is sufficiently large. Therefore,

$$\frac{g_b(s + x)}{g_b(s)} \leq \frac{H_b(s + x)^2}{H_b(s)^2} = 1 + \frac{H_b(s + x) + H_b(s)}{H_b(s)} \times \frac{H_b(s + x) - H_b(s)}{H_b(s)}. \tag{36}$$

In addition, applying a Taylor expansion, we obtain for each $x \in \mathbb{R}^d$

$$
\begin{aligned}
H_b(s + x) - H_b(s) &= \int_0^1 x^T \nabla H_b(s + ux) \, du \\
&= \int_0^1 x^T \left(\nabla H_b(s) + u \int_0^1 \Delta H_b(s + u'ux)x \, du' \right) du.
\end{aligned}
\tag{37}
$$

Observe that the previous expression can be written as

$$
H_b(s + x) - H_b(s) = x^T \nabla H_b(s) + E\big(x^T U \Delta H_b(s + U'Ux)x\big),
\tag{38}
$$

where U and U' are i.i.d. $U(0, 1)$. In what follows, we consider the linear term $x^T \nabla H_b(s)$ and the quadratic term $x^T u \Delta H_b(s + u'ux)x$, respectively.

The linear term. We use the results in (33) and Lemma 4. We shall first start with the term involving $x^T \nabla H_b(s)$. Define

$$
R = \big\{ s: \, r_b(s) \leq -\delta_2 b, \, s \notin b\Gamma \big\}.
\tag{39}
$$

Observe that, uniformly over $s \in R$ we have

$$
E\big(X^T \nabla H_b(s) I\big(s + X \notin A_{b,a}(s), \|x\|_2 < b\big)\big) = \big(\eta^T + o(b^{-\alpha+1})\big) \nabla H_b(s).
$$

Furthermore,

$$
\begin{aligned}
\eta^T \nabla H_b(s) &= E\big(\eta^T d'\big(\varrho_b(s + X)\big) \nabla \varrho_b(s + X)\big) \\
&= E\left(\sum_{j=1}^{2m+d} w_j(s + X)\eta^T v_j^* d'\big(\varrho_b(s + X)\big) \right) \\
&= -E\big(d'\big(\varrho_b(s + X)\big)\big).
\end{aligned}
$$

The last step is due to the fact that $\eta^T v_j = -1$ and $\sum_{j=1}^{2m+d} w_j(s + X) = 1$. We have noted that

$$
\begin{aligned}
P\big(r_b(s + X) \geq \delta_0\big) &\leq E\big(d'\big(\varrho_b(s + X)\big)\big) \\
&\leq P\big(r_b(s + X) \geq -\delta_0 - c_0 \log(2m + d)\big)
\end{aligned}
\tag{40}
$$

and therefore, if $s \in R$, following a reasoning similar to Lemmas 11 and 12, for sufficiently large b_0 (depending on δ_2, γ) and a sufficiently close to 1 we have that if $b \geq b_0$

$$
a^{2\alpha} \leq \frac{P(\exists j: \, X^T v_j \geq a(a_j b - s^T v_j))}{E(d'(\varrho_b(s + X)))} \leq a^{-2\alpha},
$$

which implies

$$
\eta^T \nabla H_b(s) \leq -a^{2\alpha} P\big(\exists j: \, X^T v_j \geq a\big(a_j b - s^T v_j\big)\big).
\tag{41}
$$

Now, for x taking values on any fixed compact set, we have uniformly over $s \in R$ and $s + x \notin A_{b,a}(s)$ that

$$\lim_{b \to \infty} \frac{H_b(s + x) + H_b(s)}{H_b(s)} = 2; \tag{42}$$

this follows easily using the bounds in (30) and the representation in Lemma 12 in the Appendix. We obtain that as $b \to \infty$,

$$\sup_{s \in R} \left| E\left(\left(\frac{H_b(s + X) + H_b(s)}{H_b(s)} - 2 \right) \times X^T \frac{\nabla H_b(s)}{|\nabla H_b(s)|} I\left(s + X \notin A_{b,a}(s), \|x\|_2 < b \right) \right) \right|$$

$$\leq E\left(\sup_{s \in R} \left| \frac{H_b(s + X) + H_b(s)}{H_b(s)} - 2 \right| \times \frac{|X^T \nabla H_b(s)|}{|\nabla H_b(s)|} I\left(s + X \notin A_{b,a}(s), \|x\|_2 < b \right) \right)$$

$$= o(1),$$

where the last step is thanks to the uniform convergence in (42) and the dominated convergence theorem. Similar to the derivation of (40), we have that

$$\left| \nabla H_b(s) \right| = O\left(P\left(\exists j : X^T v_j \geq a(a_j b - s^T v_j) \right) \right).$$

Thus, we obtain that

$$E\left(\frac{H_b(s + X) + H_b(s)}{H_b(s)} \times X^T \nabla H_b(s) I\left(s + X \notin A_{b,a}(s), \|x\|_2 < b \right) \right)$$

$$\leq \left(-2a^{2\alpha} + o(1) \right) P\left(\exists j : X^T v_j \geq a(a_j b - s^T v_j) \right), \tag{43}$$

where the convergence corresponding to the term $o(1)$ is uniform over $s \in R$ as $b \to \infty$.

The quadratic term. We proceed to consider $E(x^T U \Delta H_b(s + U'Ux)x)$ in (38), which in turn feeds into (36).

We take advantage of representation (38). However, to avoid confusion with the random variable X, we introduce an independent copy $\tilde{X}$ of X. Using (34), we then obtain that for $u, u' \in (0, 1)$,

$$\Delta H_b(s + u'ux)$$

$$= E\left(\Delta \varrho_b(s + u'ux + \tilde{X}) d'(\varrho_b(s + u'ux + \tilde{X})) \right)$$

$$\quad + E\left(d''(\varrho_b(s + u'ux + \tilde{X})) \nabla \varrho_b(s + u'ux + \tilde{X}) \nabla \varrho_b(s + u'ux + \tilde{X})^T \right)$$

$$= E\left(\frac{\sum_{j=1}^{2m+d} w_j (1 - w_j) v_j v_j^T}{c_0} d'(\varrho_b(s + u'ux + \tilde{X})) \right)$$

$$\quad + E\left(\frac{I(|\varrho_b(s + u'ux + \tilde{X})| \leq \delta_0)}{2\delta_0} \sum_{j=1}^{2m+d} v_j w_j \left(\sum_{j=1}^{2m+d} v_j w_j \right)^T \right).$$

Since $w_j \in (0, 1)$, we have each of the element of $\Delta H_b(s + u'ux)$ is bounded by

$$\frac{\sum_{j=1}^{2m+d} \|v_j\|_2^2}{c_0} P\big(\varrho_b(s + u'ux + \tilde{X}) \geq -\delta_0\big) \tag{44}$$

$$+ \frac{1}{2\delta_0} \left(\sum_{j=1}^{2m+d} \|v_j\|_2^2 \right)^2 P\big(|\varrho_b(s + u'ux + \tilde{X})| \leq \delta_0\big). \tag{45}$$

We will proceed to analyze each of the terms (44) and (45) separately. Let us start with (44). Note that

$$\varrho_b(s + u'ux + \tilde{X}) > -\delta_0 \tag{46}$$

implies that

$$r_b(s + u'ux + \tilde{X}) \geq -\delta_0 - c_0 \log(2m + d).$$

Equivalently, there exists j such that

$$\tilde{X}^T v_j \geq a_j b - s^T v_j - u'ux^T v_j - \delta_0 - c_0 \log(2m + d).$$

On the other hand, recall that $s + x \notin A_{b,a}(s)$ and therefore, for every j

$$x^T v_j \leq a\big(a_j b - s^T v_j\big).$$

Since $u, u' \in (0, 1)$, we have from our previous observations that inequality (46) implies that for some j

$$\tilde{X}^T v_j \geq (1 - a)\big(a_j b - s^T v_j\big) - \delta_0 - c_0 \log(2m + d). \tag{47}$$

Now, the inequalities

$$r_b(s) = \max_{j=1}^{2m+d} \big[s^T v_j - a_j b\big]$$

$$= - \min_{j=1}^{2m+d} \big[a_j b - s^T v_j\big] \leq -\delta_2 b \leq -2\big(\delta_0 + c_0 \log(2m + d)\big)/(1 - a)$$

together with (47) imply that for some j

$$\tilde{X}^T v_j \geq \tfrac{1}{2}(1 - a)\big(a_j b - s^T v_j\big) \geq \tfrac{1}{2}(1 - a)\delta_2 b. \tag{48}$$

Therefore, we conclude that if $\delta_2 b \geq 2(\delta_0 + c_0 \log(2m + d))/(1 - a)$, by regular variation there exists c' such that for all b sufficiently large

$$P\big(\varrho_b(s + u'ux + \tilde{X}) > -\delta_0\big)$$

$$\leq P\big(\exists j\colon \tilde{X}^T v_j \geq \tfrac{1}{2}(1 - a)\delta_2 b\big) \leq c' P\big(\exists j\colon \tilde{X}^T v_j \geq a\big(a_j b - s^T v_j\big)\big)$$

uniformly over $s \in R$. Note that in the last inequality we use the fact that if $s \notin b\Gamma \cup bA$ then $\|s\|_2 \leq cb$ for some constant c depending on β and γ. Consequently, we conclude that (44) is bounded by

$$c' \frac{\sum_{j=1}^{2m+d} \|v_j\|_2^2}{c_0} P\big(\exists j\colon X^T v_j^* \geq a\big(a_j b - s^T v_j^*\big)\big). \tag{49}$$

Now we proceed with term (45). Note that

$$P\big(\big|\varrho_b\big(s + u'ux + \tilde{X}\big)\big| \leq \delta_0\big) \leq P\big(\varrho_b\big(s + u'ux + \tilde{X}\big) \geq -\delta_0\big),$$

so our previous development leading to the term (49) implies that given any selection of δ_0, a, c_0 we have

$$\frac{P(|\varrho_b(s + u'ux + \tilde{X})| \leq \delta_0)}{P(\exists j\colon X^T v_j \geq a(a_j b - s^T v_j))} = \mathrm{O}(1) \tag{50}$$

uniformly over $s + x \notin A_{b,a}(s)$ and $s \in R$ as $b \to \infty$. Further, we have that as $b \to \infty$

$$\begin{aligned}
\Delta_{s,b}\big(x, u, u'\big) &\triangleq \frac{P(|\varrho_b(s + u'ux + \tilde{X})| \leq \delta_0) I(s + x \notin A_{b,a}(s))}{P(\exists j\colon X^T v_j \geq a(a_j b - s^T v_j))} \\
&= I\big(s + x \notin A_{b,a}(s)\big) \\
&\quad \times \frac{P(\varrho_b(s + u'ux + \tilde{X}) \geq \delta_0) - P(\varrho_b(s + u'ux + \tilde{X}) \geq -\delta_0)}{P(\exists j\colon X^T v_j \geq a(a_j b - s^T v_j))} \\
&= \mathrm{O}\big(b^{-1}\big).
\end{aligned}$$

In addition, the above convergence is uniform on the set $s \in R$, and $s + x \notin A_{b,a}(s)$, and on $u, u' \in (0, 1)$. We now consider two cases: $1 < \alpha \leq 2$ and $\alpha > 2$.

Case: $\alpha > 2$. The increment has finite second moment. Consider the quadratic term,

$$\begin{aligned}
&\left| \int_{\mathbb{R}^d} I\big(s + x \notin A_{b,a}(s), \|x\|_2 < b\big) \frac{H_b(s+x) + H_b(s)}{H_b(s)} E\big(x^T U \Delta H_b\big(s + U'Ux\big)x\big) P(X \in \mathrm{d}x) \right| \\
&\leq \frac{c'}{c_0} E\|X\|_2^2 \sum_{j=1}^{2m+d} \|v_j\|_2^2 P\big(\exists j\colon X^T v_j \geq a\big(a_j b - s^T v_j\big)\big) \\
&\quad + P\big(\exists j\colon X^T v_j \geq a\big(a_j b - s^T v_j\big)\big) \\
&\qquad \times \int_{\mathbb{R}^d} I\big(s + x \notin A_{b,a}(s), \|x\|_2 < b\big) \frac{H_b(s+x) + H_b(s)}{H_b(s)} \\
&\qquad \times E\big|x^T \Delta_{s,b}\big(x, U, U'\big)x\big| P(X \in \mathrm{d}x) \\
&= \left(\frac{c'}{c_0} E\|X\|_2^2 \sum_{j=1}^{2m+d} \|v_j\|_2^2 + \mathrm{o}(1) \right) \times P\big(\exists j\colon X^T v_j \geq a\big(a_j b - s^T v_j\big)\big),
\end{aligned} \tag{51}$$

where the first step is thanks to the analysis results of (44) and (45) and the last step is thanks to the uniform convergence of $\Delta_{s,b}(x, u, u')$ and dominated convergence theorem.

Case: $1 < \alpha \le 2$. The increment has infinite second moment. Note that for $s \in R$

$$\int_{\mathbb{R}^d} I(s + x \notin A_{b,a}(s), \|x\|_2 < b) x^T x \, dx = O(b^{2-\alpha}).$$

Given the choice of c_0 as in (28), the first term in the second step of (51) is

$$\frac{c'}{c_0} \sum_{j=1}^{2m+d} \|v_j\|_2^2 P(\exists j: X^T v_j \ge a(a_j b - s^T v_j)) \int_{\mathbb{R}^d} I(s + x \notin A_{b,a}(s), \|x\|_2 < b) x^T x \, dx$$

$$= o(1) P(\exists j: X^T v_j \ge a(a_j b - s^T v_j)).$$

In addition, given the convergence rate of $\Delta_{s,b}$, the second term in (51) is

$$\int_{\mathbb{R}^d} I(s + x \notin A_{b,a}(s), \|x\|_2 < b) \frac{H_b(s+x) + H_b(s)}{H_b(s)} E|x^T \Delta_{s,b}(x, U, U')x| P(X \in dx)$$

$$= O(b^{1-\alpha}).$$

Given that $c_0 \ge b^{(3-\alpha)/2}$ and b is large enough, (51) is bounded by

$$\varepsilon P(\exists j: X^T v_j \ge a(a_j b - s^T v_j))$$

for all $\alpha > 1$. Therefore, for all $\alpha > 1$, (51) is bounded by

$$2\varepsilon P(\exists j: X^T v_j \ge a(a_j b - s^T v_j)).$$

Summary. We summarize the analysis of the linear and quadratic term by inserting bounds in (43), (51), and expansions (36) and (38) into (35) and obtain that for $r_b(s) \le -\delta_2 b$ and $s \notin b\Gamma$

$$J_2(1 - p_b(s)) = \int \frac{g_b(s+x)}{g_b(s)} I(s + x \notin A_{b,a}(s)) P(X \in dx)$$

$$\le 1 - (2a^{2\alpha} + o(1) + 2\varepsilon) \frac{P(\exists j: X^T v_j \ge a(a_j b - s^T v_j))}{H_b(s)}.$$

We conclude the proof of the lemma by choosing a sufficiently close to 1 and b sufficiently large. $\qquad\square$

The subsequent lemmas provide convenient bounds for $p_b(s)$, J_1 and J_2.

Lemma 6. *For each $\varepsilon > 0$ and any given selection of $\tilde{c}_0$ (feeding into c_0), $\delta_0, \gamma, \delta_2 > 0$ there exists $b_0 \ge 0$ such that if $b \ge b_0$, $r_b(s) \le -\delta_2 b$, and $s \notin b\Gamma$ then*

(i)

$$p_b(s) \le \varepsilon.$$

(ii)

$$1 \leq \frac{H_b(s)}{v_b(s)} \leq 1 + \varepsilon.$$

Proof. Note that if $r_b(s) = \max_{j=1}^{2m+d}[s^T v_j - a_j b] \leq -\delta_2 b$ and there exists j such that $X^T v_j \geq a(a_j b - s^T v_j)$, we must also have that there exists j such that

$$X^T v_j \geq a(a_j b - s^T v_j) \geq a\delta_2 b.$$

On the other hand, $s \notin b\Gamma \cup bA$ then $\|s\|_2 = O(b)$, together with $c_0 = o(b)$, there is a constant c (depending on β and γ) such that

$$H_b(s) \geq E\left[\varrho_b^+(s + X)\right] = \int_0^\infty P\left(\varrho_b(s + X) > t\right) dt$$

$$\geq \int_0^\infty P\left(r_b(s + X) > t - c_0 \log(2m + d)\right) dt$$

$$= \int_0^\infty P\left(\exists j\colon X^T v_j \geq a(a_j b - s^T v_j) + t - c_0 \log(2m + d)\right) dt$$

$$\geq \int_0^\infty P\left(\exists j\colon X^T v_j \geq cb + t\right) dt$$

$$= b \int_0^\infty P\left(\exists j\colon X^T v_j \geq cb + ub\right) du \geq \delta' b P\left(\exists j\colon X^T v_j \geq (c + 1)b\right)$$

for some $\delta' > 0$ small. Therefore,

$$p_b(s) \leq \frac{\theta P(\exists j\colon X^T v_j \geq a\delta_2 b)}{\delta' b P(\exists j\colon X^T v_j \geq (c + 1)b)} = O(1/b)$$

as $b \to \infty$; this bound immediately yields (i). Part (ii) is straightforward from the estimates in (26), (30), and eventually (31). □

Lemma 7. *For any $\varepsilon > 0$, we can choose b sufficiently large so that on the set $r_b(s) \leq -\delta_2 b$,*

$$J_1 \leq (1 + \varepsilon) \frac{P(\exists j\colon X^T v_j \geq a(a_j b - s^T v_j))}{c_1 \theta H_b(s)}.$$

Proof. Choose b large enough such that $g_b(s) < 1$ whenever $r_b(s) \leq -\delta_2 b$. With $p_b(s)$ defined as in (15) and the fact that $g_b(s + X) \leq 1$, it follows easily that

$$J_1 \leq \frac{P(\exists j\colon X^T v_j \geq a(a_j b - s^T v_j))^2}{c_1 H_b(s)^2(p_b(s) + (1 - p_b(s))P(\exists j\colon X^T v_j \geq a(a_j b - s^T v_j)))}$$

$$\leq \frac{P(\exists j\colon X^T v_j \geq a(a_j b - s^T v_j))^2}{c_1 H_b(s)^2 p_b(s)} = (1 + \varepsilon) \frac{P(\exists j\colon X^T v_j \geq a(a_j b - s^T v_j))}{c_1 \theta H_b(s)}.$$

The last step uses the approximation (31) and the definition of $p_b(s)$ in (15).   □

We summarize our verification of the validity of $g_b(s)$ in the next result.

Proposition 2. *Given $\varepsilon, \delta_2 > 0$, let a be as chosen in Lemma 5. We choose $\theta = 1/(1+\varepsilon)^2$ and $c_1 = (1+\varepsilon)^3(1+4\varepsilon)$ such that on the set $r_b(s) \leq -\delta_2 b$ and $s \notin b\Gamma$ we have*

$$J_1 + J_2 \leq 1.$$

Proof. Combining our bounds for $p_b(s)$, J_1 and J_2 given in Lemmas 5, 6, and 7, we obtain

$$J_1 + J_2$$
$$\leq 1 + p_b(s) + \varepsilon p_b(s) + (1+\varepsilon)\frac{P(\exists j\colon X^T v_j \geq a(a_j b - s^T v_j))}{c_1 \theta H_b(s)}$$
$$- 2\frac{P(\exists j\colon X^T v_j \geq a(a_j b - s^T v_j))}{H_b(s)}(1 - 3\varepsilon/2)$$
$$\leq 1 + \frac{P(\exists j\colon X^T v_j \geq a(a_j b - s^T v_j))}{H_b(s)}\left(\theta(1+\varepsilon)^2 + \frac{1+\varepsilon}{c_1\theta}\right)$$
$$- 2\frac{P(\exists j\colon X^T v_j \geq a(a_j b - s^T v_j))}{H_b(s)}(1 - 3\varepsilon/2).$$

We arrive at

$$J_1 + J_2$$
$$\leq 1 + \frac{P(\exists j\colon X^T v_j \geq a(a_j b - s^T v_j))}{H_b(s)}\left(\theta(1+\varepsilon)^2 + \frac{1+\varepsilon}{c_1\theta} - 2(1 - 3\varepsilon/2)\right).$$

We then select $\theta = 1/(1+\varepsilon)^2$ and $c_1 = (1+\varepsilon)^3(1+4\varepsilon)$ and conclude that $J_1 + J_2 \leq 1$ as required for if $\varepsilon \in (0, 1/2)$.   □

3.3. Proofs of Theorem 1 and Proposition 1

We now are ready to provide the proof of Proposition 1. As noted earlier, Theorem 1 follows immediately as a consequence of Proposition 1, combined with Lemma 1.

Proof of Proposition 1. For each $\lambda \in (-\infty, \infty)$, define $\tau(\lambda) = \inf\{n \geq 0\colon r_b(S_n) \geq \lambda\}$. Proposition 2 together with Lemma 3 implies

$$g_b(s) \geq E_s\left(g_b(S_{\tau(-\delta_2 b)})\prod_{j=0}^{\tau(-\delta_2 b)-1}\hat{k}(S_j, S_{j+1})I(\tau(-\delta_2 b) \leq T_{b\Gamma})\right).$$

Given $c_1 > 0$, there exists $\kappa \in (0, \infty)$ such that $r_b(s) \geq \kappa$ implies $g(s) = 1$. So, we have that

$$
\begin{aligned}
g_b(s) &\geq E_s\left(g_b(S_{\tau(-\delta_2 b)}) \prod_{j=0}^{\tau(-\delta_2 b)-1} \hat{k}(S_j, S_{j+1}) I\left(\tau(-\delta_2 b) \leq T_{b\Gamma}\right) \right) \\
&\geq E_s\left(\prod_{j=0}^{\tau(-\delta_2 b)-1} \hat{k}(S_j, S_{j+1}) I\left(\tau(-\delta_2 b) \leq T_{b\Gamma}, r_b(S_{\tau(-\delta_2 b)}) \geq \kappa\right) \right).
\end{aligned}
\tag{52}
$$

We prove the theorem in two steps. The first step is to show that $\hat{P}_0(\cdot)$ is a good approximation, in total variation, to $P_0(\cdot | \tau(-\delta_2 b) < T_{b\Gamma}, r_b(S_{\tau(-\delta_2 b)}) \geq \kappa)$. In step 2, we show that $P_0(\cdot | \tau(-\delta_2 b) < T_{b\Gamma}, r_b(S_{\tau(-\delta_2 b)}) \geq \kappa)$ approximates $P_0(\cdot | T_{bA} < T_{b\Gamma})$ well.

Step 1. Applying Lemmas 2, 3, and a similar argument to the one below Lemma 2, it is sufficient to show that for any $\varepsilon > 0$ we can pick $\delta_2, \gamma, \theta, a, c_1$ such that for all b large enough

$$
g_b(0) \leq (1 + \varepsilon) P_0^2\left(\tau(-\delta_2 b) < T_{b\Gamma}, r_b(S_{\tau(-\delta_2 b)}) \geq \kappa\right).
$$

First, note that

$$
P_0\left(\tau(-\delta_2 b) < \infty\right) \geq H_b(0).
$$

Therefore,

$$
\begin{aligned}
&P_0\left(\tau(-\delta_2 b) < \infty, r_b(S_{\tau(-\delta_2 b)}) \geq \kappa\right) \\
&\quad = P_0\left(\tau(-\delta_2 b) < \infty\right) \times P_0\left(r_b(S_{\tau(-\delta_2 b)}) \geq \kappa | \tau(-\delta_2 b) < \infty\right) \\
&\quad \geq H_b(0) \times P_0\left(r_b(S_{\tau(-\delta_2 b)}) \geq \kappa | \tau(-\delta_2 b) < \infty\right).
\end{aligned}
$$

Then, given $c_1 = (1 + \varepsilon)$, we show that one can pick $\delta_2 > 0$ small enough, depending on $\varepsilon > 0$, such that for b sufficiently large

$$
P_0\left(r_b(S_{\tau(-\delta_2 b)}) \geq \kappa | \tau(-\delta_2 b) < \infty\right) \geq 1/(1 + \varepsilon).
\tag{53}
$$

In order to do this we will use results from one dimensional regularly varying random walks. Define

$$
\tau_i(-\delta_2 b) = \inf\{n \geq 0 : v_i^T S_n - a_i b \geq -\delta_2 b\}
$$

for $i = 1, \ldots, 2m + d$ and let

$$
\mathcal{I} = \left\{ 1 \leq i \leq 2m + d : \varlimsup_{b \to \infty} P(T_{bA} < \infty)/P\left(\tau_i(0) < \infty\right) < \infty \right\}.
$$

Observe that

$$
\max_{i=1}^{2m+d} P\left(\tau_i(0) < \infty\right) \leq P(T_{bA} < \infty) \leq (2m + d) \max_{i=1}^{2m+d} P\left(\tau_i(0) < \infty\right),
$$

so the set $\mathcal{I}$ contains the half-spaces that have substantial probability of being reached given that $T_{bA} < \infty$. We have that

$$
\frac{P_0(r_b(S_{\tau(-\delta_2 b)}) \leq \kappa, \tau(-\delta_2 b) < \infty)}{P_0\big(\tau(-\delta_2 b) < \infty\big)}
$$

$$
\leq \sum_{i=1}^{2m+d} \frac{P_0(r_b(S_{\tau(-\delta_2 b)}) \leq \kappa, \tau_i(-\delta_2 b) = \tau(-\delta_2 b), \tau(-\delta_2 b) < \infty)}{P_0\big(\tau(-\delta_2 b) < \infty\big)}
$$

$$
\leq \sum_{i=1}^{2m+d} \frac{P_0(r_b(S_{\tau_i(-\delta_2 b)}) \leq \kappa, \tau_i(-\delta_2 b) < \infty)}{P_0(\tau(-\delta_2 b) < \infty)}
$$

$$
\leq \sum_{i \in \mathcal{I}} \frac{P_0(v_i^T S_{\tau_i(-\delta_2 b)} - a_i b \leq \kappa, \tau_i(-\delta_2 b) < \infty)}{P_0(\tau(-\delta_2 b) < \infty)} + o(1).
$$

(54)

Now, $i \in \mathcal{I}$ implies that $v_i^T X$ is regularly varying with index α and therefore (see [5]) there exists a constant $c_i' > 0$ (independent of δ_2) such that for all b large enough

$$
P_0\big(v_i^T S_{\tau_i(-\delta_2 b)} - (a_i - \delta_2) b > 2\delta_2 b | \tau_i(-\delta_2 b) < \infty\big) \geq \frac{1}{(1 + 2c_i' \delta_2)^{\alpha-1}}.
$$

Consequently,

$$
\frac{P_0(v_i^T S_{\tau_i(-\delta_2 b)} - a_i b \leq \kappa, \tau_i(-\delta_2 b) < \infty)}{P_0\big(\tau(-\delta_2 b) < \infty\big)}
$$

$$
\leq \big(1 - \big(1 + 2c_i' \delta_2\big)^{1-\alpha}\big) \frac{P_0(\tau_i(-\delta_2 b) < \infty)}{P_0\big(\tau(-\delta_2 b) < \infty\big)}.
$$

(55)

Bound of (55) together with (54) implies that one can select δ_2 small enough and b large enough to satisfy (53) and thus

$$
P_0\big(\tau(-\delta_2 b) < \infty, r_b(S_{\tau(-\delta_2 b)}) \geq \kappa\big) \geq (1 - \varepsilon) H_b(0).
$$

Furthermore, one can choose γ large enough so that

$$
P_0\big(\tau(-\delta_2 b) < T_{b\Gamma}, r_b(S_{\tau(-\delta_2 b)}) \geq \kappa\big) \geq (1 - \varepsilon) H_b(0).
$$

Thanks to Lemma 2, we obtain that

$$
\varlimsup_{b \to \infty} \sup_B \big| \hat{P}_0(B) - P_0\big(B | \tau(-\delta_2 b) < T_{b\Gamma}, r_b(S_{\tau(-\delta_2 b)}) \geq \kappa\big)\big| \leq \varepsilon.
$$

(56)

This completes step 1.

Step 2. With an analogous development for (53), we can establish that with δ_2 chosen small enough and γ large enough

$$\varlimsup_{b\to\infty} \sup_B \left| P_0\big(B|\tau(-\delta_2 b) < T_{b\Gamma}\big) - P_0\big(B|\tau(-\delta_2 b) < T_{b\Gamma}, r_b(S_{\tau(-\delta_2 b)}) \geq \kappa\big)\right| \leq \varepsilon. \qquad (57)$$

In addition, $\{T_{bA} < T_{b\Gamma}\} \subset \{\tau(-\delta_2 b) < T_{b\Gamma}\}$. Using the same type of reasoning leading to (55) we have that for any $\varepsilon > 0$, δ_2 can be chosen so that

$$P_0(T_{bA} < T_{b\Gamma}) \geq \frac{1}{1+\varepsilon} P_0\big(\tau(-\delta_2 b) < T_{b\Gamma}\big)$$

for b and γ large enough. Therefore, we have

$$\varlimsup_{b\to\infty} \sup_B \left| P_0\big(B|\tau(-\delta_2 b) < T_{b\Gamma}\big) - P_0(B|T_{bA} < T_{b\Gamma})\right| \leq \varepsilon. \qquad (58)$$

Combining (57) and (58), we obtain that with γ large enough

$$\varlimsup_{b\to\infty} \sup_B \left| P_0\big(B|\tau(-\delta_2 b) < T_{b\Gamma}, r_b(S_{\tau(-\delta_2 b)}) \geq \kappa\big) - P_0(B|T_{bA} < T_{b\Gamma})\right| \leq \varepsilon. \qquad (59)$$

Then, we put together (56), (59) and conclude that

$$\varlimsup_{b\to\infty} \sup_B \left| \hat{P}_0(B) - P_0(B|T_{bA} < T_{b\Gamma})\right| \leq 2\varepsilon. \qquad \square$$

4. A conditional central limit theorem

The goal of this section is to provide a proof of Theorem 2. The proof is a consequence of the following proposition. First, write for any $t \geq 0$

$$\kappa_{\mathbf{a}}(t) \triangleq \mu\left(\left\{y: \max_{j=1}^{2m+d}\big(y^T v_j - a_j\big) > t\right\}\right), \qquad (60)$$

where $\mathbf{a} = (a_1, \ldots, a_{2m+d})$. Recall that A is defined as in (8).

Proposition 3. *For each $z > 0$ let $Y(z)$ be a random variable with distribution given by*

$$P\big(Y(z) \in B\big) = \frac{\mu\big(B \cap \{y: \max_{j=1}^{2m+d}[y^T v_j - a_j] \geq z\}\big)}{\mu\big(\{y: \max_{j=1}^{2m+d}[y^T v_j - a_j] \geq z\}\big)}.$$

In addition, let Z be a positive random variable following distribution

$$P(Z > t) = \exp\left\{-\int_0^t \frac{\kappa_{\mathbf{a}}(s)}{\int_s^\infty \kappa_{\mathbf{a}}(u)\,du}\,ds\right\}.$$

Let $\hat{S}_0 = 0$ and $\hat{S}_n$ evolve according to the transition kernel (13) *associated with the approximation in Proposition* 1. *Define $T_{bA} = \inf\{n: \hat{S}_n \in bA\}$. Then, as $b \to \infty$, we have that*

$$\left(\frac{T_{bA}}{b}, \frac{\hat{S}_{uT_{bA}} - uT_{bA}\eta}{\sqrt{T_{bA}}}, \frac{\hat{X}_{T_{bA}}}{b} \right) \Rightarrow \left(Z, CB(uZ), Y(Z) \right)$$

in $\mathbb{R} \times D[0, 1) \times \mathbb{R}^d$, where $CC^T = \mathrm{Var}(X)$, $B(\cdot)$ is a d-dimensional Brownian motion with identity covariance matrix, $B(\cdot)$ is independent of Z and $Y(Z)$.

The strategy to prove Proposition 3 is to create a coupling of two processes S and $\hat{S}$ on the same probability space with measure P. For simplicity, we shall assume that $\hat{S}_0 = S_0 = 0$. The process $S = \{S_n: n \geq 0\}$ follows its original law, that is, $S_n = X_1 + \cdots + X_n$ where X_i's are i.i.d. The process $\hat{S}$ evolves according to the transition kernel

$$P(\hat{S}_{n+1} \in ds_{n+1}|\hat{S}_n = s_n) = \hat{K}(s_n, ds_{n+1})$$

obtained in Theorem 1. Now we explain how the process S and $\hat{S}$ are coupled. A transition at time j, given $\hat{S}_{j-1} = \hat{s}_{j-1}$, is constructed as follows. First, we construct a Bernoulli random variable I_j with success parameter $p_b(\hat{s}_{j-1})$. If $I_j = 1$ then we consider X (a generic random variable following the nominal/original distribution) given that $\hat{s}_{j-1} + X \in A_{b,a}(\hat{s}_{j-1})$ and let $\hat{X}_j = X$ (recall that $A_{b,a}(s)$ is defined as in (12)); otherwise if $I_j = 0$ we let $\hat{X}_j = X_j$, that is, we let $\hat{X}_j$ be equal to the jth increment of S. We then define $N_b = \inf\{n \geq 1: I_n = 1\}$ and observe that $\hat{S}_j = S_j$ for $j < N_b$. The increments of processes S and $\hat{S}$ at times $j > N_b$ are independent.

We will first show that $N_b = T_{bA}$ with high probability as $b \nearrow \infty$, once this result has been shown the rest of the argument basically follows from functional central limit theorem for standard random walk. We need to start by arguing that whenever a jump occurs (i.e., $I_n = 1$) the walk reaches the target set with high probability. This is the purpose of the following lemma.

Lemma 8. *For every $\varepsilon, \delta_2, \gamma > 0$ there exists $a, b_0 > 0$ such*

$$P\left(r_b(s + X) > 0|s + X \in A_{b,a}(s)\right) \geq 1 - \varepsilon$$

for all $r_b(s) \leq -\delta_2 b$, $s \notin b\Gamma$, and $b > b_0$.

Proof. Set $s = b \cdot u \in \mathbb{R}^d$ and note that

$$P\left(\max_{j=1}^{2m+d} \left((s + X)^T v_j - a_j b \right) > 0 | s + X \in A_{b,a}(s) \right)$$

$$= P\left(\exists j: X^T v_j \geq b(a_j - u^T v_j) | \exists j: X^T v_j \geq ab(a_j - u^T v_j) \right)$$

$$= \frac{P(\exists j: X^T v_j \geq b(a_j - u^T v_j))}{P(\exists j: X^T v_j \geq ab(a_j - u^T v_j))}.$$

For each fixed u, we have that

$$\lim_{b \to \infty} \frac{P(\exists j: X^T v_j \geq b(a_j - u^T v_j))}{P(\exists j: X^T v_j \geq ab(a_j - u^T v_j))} = \frac{\mu\{y: \exists j: y^T v_j \geq (a_j - u^T v_j)\}}{\mu\{y: \exists j: y^T v_j \geq a(a_j - u^T v_j)\}}.$$

The convergence occurs uniformly over the set of u's such that $r_b(ub)/b \leq -\delta_2$ and $u \notin \Gamma$, which is a compact set. The result then follows by continuity of the radial component in the polar representation of $\mu(\cdot)$ (Lemma 11) as $a \to 1$. $\qquad\square$

Now we prove that $T_{bA} = N_b$ occurs with high probability as $b \nearrow \infty$.

Lemma 9. *For any $\varepsilon > 0$, we can select $\gamma > 0$ sufficiently large so that*

$$\lim_{b \to \infty} P_0(N_b < \infty) \geq 1 - \varepsilon.$$

Proof. Define $T_{b\Gamma} = \inf\{n: \ \hat{S}_n \in b\Gamma\}$. Choose ε' positive and γ large enough so that for all $t < (1 - \varepsilon')\gamma/d$

$$P_0(N_b > tb) \leq P(N_b > tb, T_{b\Gamma} > tb) + o(1)$$

$$= E_0\left[\prod_{k \leq tb}(1 - p_b(\hat{S}_k)); T_{b\Gamma} > tb\right] + o(1)$$

$$\leq E_0\left[\prod_{k \leq tb}(1 - p_b(\hat{S}_k))I\big(\|\hat{S}_k - \eta k\|_2 < \varepsilon' \max\{k, b\}\big)\right]$$

$$+ P_0\left(\sup_{k \leq tb}\big|\|\hat{S}_k - \eta k\|_2 - \varepsilon' \max\{k, b\}\big| > 0\right) + o(1).$$

In the last step, we drop the condition $T_{b\Gamma} > tb$ on the set $\{\|\hat{S}_k - \eta k\|_2 < \varepsilon' \max\{k, b\}\}$. The second term in the last step vanishes as $b \to \infty$ for any $\varepsilon' > 0$.

We claim that we can find constants $\delta', c' > 0$ such that for all $1 \leq k \leq c'\gamma b$

$$\inf_{\{s: \ \|\eta k - s\|_2 \leq \varepsilon' \max\{k, b\}, s \notin b\Gamma\}} p_b(s) \geq \frac{\delta'}{k + b}. \tag{61}$$

To see this, recall that if $r_b(s) \leq -\delta_2 b$ then

$$p_b(s) = \frac{\theta P(\exists j: \ X^T v_j > a(a_j b - s^T v_j))}{\int_0^\infty P(\exists j: \ X^T v_j > a_j b - s^T v_j + t)\, dt}.$$

Now, if $\|\eta k - s\|_2 \leq \varepsilon' \max\{k, b\}$ then, letting $\lambda_+ = \max_{j=1}^{2m+d} \|v_j\|_2$ we obtain, by the Cauchy–Schwarz inequality

$$\big|s^T \eta - kd\big| \leq d^{1/2}\varepsilon' \max\{k, b\}, \tag{62}$$

$$\big|s^T v_j - k\eta^T v_j\big| \leq \lambda\varepsilon' \max\{k, b\}. \tag{63}$$

Inequality (62) implies that

$$s^T \eta \leq d^{1/2}\varepsilon' \max\{k, b\} + kd \leq k\big(\varepsilon'd^{1/2} + d\big) + \varepsilon'd^{1/2}b.$$

We choose $\varepsilon' d^{1/2} < \gamma/2$. Then $k \le \gamma b/2(\varepsilon' d^{1/2} + d)$ implies $s^T \eta < \gamma b$. We shall select

$$c' = \frac{1}{2(\varepsilon' d^{1/2} + d)}.$$

Inequality (63) implies that

$$
\begin{aligned}
p_b(s) &\ge \frac{\theta P(\exists j: \ X^T v_j > a(a_j b + k + \lambda\varepsilon' \max\{k, b\}))}{\int_0^\infty P(\exists j: \ X^T v_j > a_j b + k + t - \lambda\varepsilon' \max\{k, b\}) \, \mathrm{d}t} \\
&\ge \frac{\theta P(\exists j: \ X^T v_j > a((a_j + \lambda\varepsilon')b + k(1 + \lambda\varepsilon')))}{\int_0^\infty P(\exists j: \ X^T v_j > (a_j - \lambda\varepsilon')b + k(1 - \lambda\varepsilon') + t) \, \mathrm{d}t} \\
&\ge \frac{\theta P(\exists j: \ X^T v_j > a((a_j + \lambda\varepsilon')b + k(1 + \lambda\varepsilon')))}{\int_0^\infty P(\exists j: \ X^T v_j > (\min_j a_j - \lambda\varepsilon')b + k(1 - \lambda\varepsilon') + t) \, \mathrm{d}t}.
\end{aligned}
$$

We introduce the change of variables $t = s[(\min_j a_j - \lambda\varepsilon')b + k(1 - \lambda\varepsilon')]$ for the integral in the denominator and obtain that

$$
\int_0^\infty P\left(\exists j: \ X^T v_j > \left(\min_j a_j - \lambda\varepsilon'\right)b + k\left(1 - \lambda\varepsilon'\right) + t\right) \mathrm{d}t
$$

$$
= \left[\left(\min_j a_j - \lambda\varepsilon'\right)b + k\left(1 - \lambda\varepsilon'\right)\right]
$$

$$
\times \int_0^\infty P\left(\exists j: \ X^T v_j > \left[\left(\min_j a_j - \lambda\varepsilon'\right)b + k\left(1 - \lambda\varepsilon'\right)\right](1 + s)\right) \mathrm{d}s.
$$

Notice that for b sufficiently large there exists c'' so that

$$
\int_0^\infty P\left(\exists j: \ X^T v_j > \left[\left(\min_j a_j - \lambda\varepsilon'\right)b + k\left(1 - \lambda\varepsilon'\right)\right](1 + s)\right) \mathrm{d}s
$$

$$
\le c'' \theta P\left(\exists j: \ X^T v_j > a\left((a_j + \lambda\varepsilon')b + k\left(1 + \lambda\varepsilon'\right)\right)\right).
$$

Therefore, we have

$$p_b(s) \ge \frac{1}{c''((\min_j a_j - \lambda\varepsilon)b + k(1 - \lambda\varepsilon))}.$$

The possibility of selecting $\delta' > 0$ to satisfy (61) follows from the previous inequality.

Consequently, having (61) in hand, if $t < c'\gamma$

$$
E_0\left(\prod_{k \le tb}(1 - p_b(\hat{S}_k)) I\left(\|\hat{S}_k - \eta j\|_2 < \varepsilon \max\{j, b\}\right)\right) \le \exp\left(-\sum_{1 \le j \le tb} \frac{\delta'}{j + b}\right).
$$

We then conclude that if

$$
\varliminf_{b \to \infty} P_0(N_b < \infty) \ge \varliminf_{b \to \infty} P_0\left(N_b < c'\gamma b\right) \ge 1 - \frac{1}{(c'\gamma + 1)^{\delta'}} + o(1) \tag{64}
$$

and this implies the statement of the result.   □

Lemma 10. *For any $\varepsilon > 0$ we can select $a, \gamma > 0$ such that*

$$\lim_{b \to \infty} P_0(T_{bA} = N_b) \geq 1 - \varepsilon.$$

Proof. Define

$$\tau(-\delta_2 b) = \inf\{n \geq 0 : r_b(S_n) \geq -\delta_2 b\}.$$

Let c' be as chosen in the proof of Lemma 9. We then have

$$P_0(T_{bA} = N_b, N_b < \infty)$$

$$= \sum_{k \leq c' \gamma b} P_0\Big(N_b = k, \max_{j=1}^{2m+d} [S_k^T v_j - a_i b] \geq 0, T_{bA} > k - 1\Big)$$

$$\geq \sum_{k \leq c' \gamma b} P_0\Big(N_b = k, \max_{j=1}^{2m+d} [\hat{S}_k^T v_j - a_i b] \geq 0, \tau(-\delta_2 b) > k - 1\Big)$$

$$\geq (1 - \varepsilon) \sum_{k \leq c' \gamma b} P_0\big(N_b = k, \tau(-\delta_2 b) > k - 1\big).$$

In the last inequality, we have used Lemma 8. Further,

$$\sum_{k \leq c' \gamma b} P_0\big(N_b = k, \tau(-\delta_2 b) > k - 1\big) \geq P_0\big(N_b \leq c' \gamma b, \tau(-\delta_2 b) > N_b - 1\big)$$

$$\geq P_0\big(N_b \leq c' \gamma b, \tau(-\delta_2 b) = \infty\big)$$

$$\geq P\big(N_b \leq c' \gamma b\big) - P_0\big(\tau(-\delta_2 b) < \infty\big).$$

Since $P_0(\tau(-\delta_2 b) < \infty) \to 0$ as $b \to \infty$, for γ sufficiently large, we conclude

$$\lim_{b \to \infty} P_0(T_{bA} = N_b) \geq \lim_{b \to \infty} P_0(T_{bA} = N_b, N_b < \infty)$$

$$\geq (1 - \varepsilon) \lim_{b \to \infty} P\big(N_b \leq c' \gamma b\big) \geq (1 - \varepsilon)^2,$$

where the last equality follows from (64) in the proof of Lemma 9. We conclude the proof.   □

Proposition 3 will follow as a consequence of the next result.

Proposition 4. *By possibly enlarging the probability space, we have the following three coupling results.*

(i) Let $\kappa_{\mathbf{a}}$ be as defined in (60). There exists a family of sets $(B_b : b > 0)$ such that $P(B_b) \to 1$ as $b \nearrow \infty$ and with the property that if $t \leq \frac{\gamma}{2d}$

$$P(N_b > tb | S) = P(Z_{a,\theta} > t)\big(1 + o(1)\big)$$

as $b \to \infty$ uniformly over $S \in B_b$, where

$$P(Z_{a,\theta} > t) = \exp\left(-\theta \int_0^t \frac{\kappa_{a\mathbf{a}}(as)}{\int_s^\infty \kappa_{\mathbf{a}}(u)\,\mathrm{d}u}\,\mathrm{d}s\right).$$

(ii) *We can embed the random walk $S = \{S_n\colon n \geq 0\}$ and a uniform random variable U on $(0, 1)$ independent of S in a probability space such that $N_b \triangleq N_b(S, U)$ (a function of S and U) and $Z_{a,\theta} \triangleq Z_{a,\theta}(U)$ (a function of U) for all $S \in B_b$ such that*

$$\left|\frac{N_b(S, U)}{b} - Z_{a,\theta}(U)\right| \to 0 \tag{65}$$

as $b \to \infty$ for almost every $U \leq P(Z_{a,\theta} \leq \frac{\gamma}{2d})$. Furthermore, one can construct a d-dimensional Brownian motion $B(t)$ so that

$$S_{\lfloor t \rfloor} = t\eta + CB(t) + e(t), \tag{66}$$

where $CC^T = \mathrm{Var}(X)$ is the covariance matrix of an increment X and $e(\cdot)$ is a (random) function such that

$$\frac{|e(xt)|}{t^{1/2}} \to 0$$

with probability one, uniformly on compact sets on $x \geq 0$ as $t \to \infty$.

(iii) *Finally, we can also embed a family of random variables $\{\hat{X}(a, b, s)\}$, independent of S and U, distributed as X conditioned on $s + X \in A_{b,a}(s)$, coupled with a random variable $Y(z)$ (also independent of S and U) so that for each Borel set B*

$$P\big(Y(z) \in B\big) = \frac{\mu(B \cap \{y\colon \max_{j=1}^{2m+d}(y^T v_j - a_i) \geq z\})}{\mu(\{y\colon \max_{j=1}^{2m+d}(y^T v_j - a_i) \geq z\})},$$

and with the property

$$\lim_{a \to 1, b \to \infty}\left|\frac{\hat{X}(a, b, \eta(z + \xi_b))}{b} - Y(z)\right| \to 0$$

with probability 1 as long as $\xi_b \to 0$.

Proof. We start with the proof of (i). Note that

$$P_0(N_b > tb | S) = \prod_{0 \leq k \leq \lfloor tb \rfloor - 1}\big(1 - p_b(S_j)\big),$$

where, by convention, a product indexed by an empty subset is equal to unity. Now, let $\delta_b = 1/\log b$, $\gamma_{k,\delta_b} = \max(1/\delta_b^2, \delta_b k)$ and

$$B_b = \big\{S\colon \|S_k - k\eta\|_2 \leq \gamma_{k,\delta_b}\big\}$$

for $k \geq 1$ for all $k \leq tb$. It follows easily that $P(B_b) \to 1$ as $b \to \infty$. Recall that if $r_b(s) \leq -\delta_2 b$ and $s^T \eta \leq \gamma b$

$$p_b(s) = \frac{\theta P(\exists j \colon X^T v_j > a(a_j b - s^T v_j))}{\int_0^\infty P(\exists j \colon X^T v_j > a_j b - s^T v_j + t)\, dt}.$$

We will find upper and lower bounds on the numerator and denominator on the set B_b so that we can use regular variation properties to our advantage. Suppose that $S_k = s$. Define $\lambda = \max_j \|v_j\|_2$ and observe that if $\|s - k\eta\|_2 \leq \gamma_{k,\delta_b}$ then for all $1 \leq j \leq 2m + d$ and all $k \geq 0$ we have that

$$\left| s^T v_j - k\eta^T v_j \right| = \left| s^T v_j + k \right| \leq \lambda \gamma_{k,\delta_b}.$$

In addition,

$$s^T \eta \leq kd + d^{1/2} \gamma_{k,\delta_b} \leq \gamma b$$

if $k \leq tb$ and $t \leq \frac{\gamma}{2d}$, for all b large enough. Therefore, if $X^T v_j > a(a_j b - s^T v_j)$ and $\|s - k\eta\|_2 \leq \gamma_{k,\delta_b}$,

$$X^T v_j > a(a_j b - s^T v_j) \Rightarrow X^T v_j > a a_j b + ak - \lambda \gamma_{k,\delta_b}.$$

Consequently, if $\|s - k\eta\|_2 \leq \gamma_{k,\delta_b}$

$$P(\exists j \colon X^T v_j > a(a_j b - s^T v_j))$$
$$\leq P(\exists j \colon X^T v_j > a a_j b + ak - a\lambda \gamma_{k,\delta_b}).$$

Similarly,

$$P(\exists j \colon X^T v_j > a(a_j b - s^T v_j))$$
$$\geq P(\exists j \colon X^T v_j > a_j b + ak + a\lambda \gamma_{k,\delta_b}).$$

Following analogous steps, we obtain that for $\|s - k\eta\|_2 \leq \gamma_{k,\delta_b}$

$$v_b(s) \leq \int_0^\infty P(\exists j \colon X^T v_j > a_j b + k + t - \lambda \gamma_{k,\delta_b})\, dt$$

and

$$v_b(s) \geq \int_0^\infty P(\exists j \colon X^T v_j > a_j b + k + t + \lambda \gamma_{k,\delta_b})\, dt.$$

Our previous upper and lower bounds indicate that on $\|S_k - k\eta\|_2 \leq \gamma_{k,\delta_b}$,

$$v_b(S_k) = (1 + o(1)) \int_0^\infty P(\exists j \colon X^T v_j > a_j b + k + t)\, dt$$

on the set B_b, uniformly as $b \to \infty$ and then

$$
\begin{aligned}
p_b(S_k) &= \theta\big(1 + \mathrm{o}(1)\big) \frac{P(\exists j \colon X^T v_j > a a_j b + a k)}{\int_0^\infty P(\exists j \colon X^T v_j > a_j b + k + t)\, \mathrm{d}t} \\
&= \theta\big(1 + \mathrm{o}(1)\big) \frac{P(\exists j \colon X^T v_j > a a_j b + a k)}{\int_k^\infty P(\exists j \colon X^T v_j > a_j b + u)\, \mathrm{d}u} \\
&= \theta\big(1 + \mathrm{o}(1)\big) \frac{P(\exists j \colon X^T v_j > a a_j b + a k)}{b \int_{k/b}^\infty P(\exists j \colon X^T v_j > a_j b + t b)\, \mathrm{d}t}.
\end{aligned}
$$

By definition of regular variation there exists a slowly varying function $L(\cdot)$ such that

$$
P\big(\exists j \colon X^T v_j - a a_j b > z\big) = L(b) b^{-\alpha} \kappa_{a\mathbf{a}}(z/b).
$$

Therefore, by Karamata's theorem (see [20]) we then have that if $S \in B_b$ and $k \le tb$ with $t \le \gamma/2d$

$$
p_b(S_k) = \theta\big(1 + \mathrm{o}(1)\big) \frac{\kappa_{a\mathbf{a}}(ak/b)}{b \int_{k/b}^\infty \kappa_{\mathbf{a}}(t)\, \mathrm{d}t}
$$

as $b \to \infty$ uniformly over S in B_b. Therefore,

$$
P_0(N_b > tb \mid S) = \prod_{0 \le j \le \lfloor tb \rfloor - 1} \big(1 - p_b(S_j)\big) = \exp\left(-\big(\theta + \mathrm{o}(1)\big) \sum_{j=0}^{\lfloor tb \rfloor - 1} \frac{\kappa_{a\mathbf{a}}(ak/b)}{b \int_{k/b}^\infty \kappa_{\mathbf{a}}(t)\, \mathrm{d}t}\right)
$$

uniformly as $b \nearrow \infty$ over $S \in B_b$ and

$$
\sum_{j=0}^{\lfloor tb \rfloor - 1} \frac{\kappa_{a\mathbf{a}}(ak/b)}{b \int_{k/b}^\infty \kappa_{\mathbf{a}}(t)\, \mathrm{d}t} \to \int_0^t \frac{\kappa_{a\mathbf{a}}(as)}{\int_s^\infty \kappa_{\mathbf{a}}(u)\, \mathrm{d}u}\, \mathrm{d}s,
$$

which implies that as long as $t \le \frac{\gamma}{2d}$

$$
P_0(N_b > tb \mid S) \to P(Z_{a,\theta} > t) = \exp\left(-\theta \int_0^t \frac{\kappa_{a\mathbf{a}}(as)}{\int_s^\infty \kappa_{\mathbf{a}}(u)\, \mathrm{d}u}\, \mathrm{d}s\right).
$$

Part (ii) is straightforward from (i) by Skorokhod embedding for random walks as follows. First, construct a probability space in which the random walk S is strongly approximated by a Brownian motion as in (66) and include a uniform random variable U independent of S. Construct N_b/b applying the generalized inverse cumulative distribution function of N_b/b given S to U. Then, apply the same uniform U to generate $Z_{a,\theta}$ by inversion. Because of our estimates in part (i) we have that (65) holds as long as $U \le P(Z_{a,\theta} \le \frac{\gamma}{2d})$. Part (iii) follows using a similar argument. $\qquad\square$

Now we are ready to provide the proof of Proposition 3.

Proof of Proposition 3. According to Proposition 1, the distribution of $\hat{S}$ approximates the conditional random walk up to time $T_{bA} - 1$ in total variation. By virtue of Lemmas 9 and 10, we can replace $(T_{bA}, X_{T_{bA}})$ by $(N_b, \hat{X}_{N_b})$. Then, it suffices to show weak convergence of

$$\left(\frac{N_b}{b}, \frac{S_{uN_b} - uN_b\eta}{\sqrt{N_b}}, \frac{\hat{X}_{N_b}}{b} \right),$$

given that θ and a can be chosen arbitrarily close to 1. By using the embedding in Proposition 4 we consider $S \in B_b$ so that on $U \leq P(Z_{a,\theta} \leq \gamma/2)$

$$\left(\frac{N_b}{b}, \frac{S_{uN_b} - uN_b\eta}{\sqrt{N_b}}, \frac{\hat{X}_{N_b}}{b} \right)$$
$$= \left(Z_{a,\theta} + \xi_b, \frac{CB(ubZ_{a,\theta} + ub\xi_b) + e(uN_b)}{\sqrt{bZ_{a,\theta} + \xi_b}}, Y_a(Z_{a,\theta} + \xi_b) + \chi_b \right),$$

where $\xi_b, \chi_b \to 0$ as $b \to \infty$ and $\sup_{u \in [0,1]} |e(uN_b)/\sqrt{N_b}| \to 0$ as $N_b \to \infty$. Also, note that as $b \to \infty$ we choose a and θ close 1 and γ sufficiently large. From Lemmas 9 and 10, we must verify that for each $z > 0$

$$\sup_{0 \leq u \leq 1} \left| \frac{B(ubz + ub\xi_b) - B(ubz)}{b^{1/2}} \right| \longrightarrow 0.$$

Given $z > 0$ select b large enough so that $\xi_b \leq \varepsilon$ for each $S \in B_b$. Then, it suffices to bound the quantity

$$\sup_{u,s \in (0,1):\ |u-s| \leq \varepsilon} \left| \frac{B(ubz) - B(sbz)}{b^{1/2}} \right|.$$

However, by the invariance principle, the previous quantity equals in distribution to

$$z^{1/2} \sup_{u,s \in (0,1):\ |u-s| \leq \varepsilon} \left| B(u) - B(s) \right|,$$

which is precisely the modulus of continuity of Brownian motion evaluated ε, which (by continuity of Brownian motion) goes to zero almost surely as $\varepsilon \to 0$. From Proposition 4, we have that $Z_{a,\theta} \to Z$ as $a, \theta \to 1$, where Z is defined as in the statement of the proposition. Since γ can be chosen arbitrarily large as $b \to \infty$, we complete the proof. $\square$

Finally, we give the proof of Theorem 2.

Proof of Theorem 2. It suffices to exhibit a coupling under which

$$\left(Z, CB(uZ), Y(Z) \right) - \left(Z^*, CB(uZ^*), Y^*(Z^*) \right) \longrightarrow 0$$

almost surely as $\beta, \gamma \to \infty$ and $\delta \to 0$, but this is immediate from continuity of the Brownian motion and of the radial component of the measure $\mu(\cdot)$ (Lemmas 11 and 12). $\square$

Appendix: Some properties of regularly varying distributions

In this Appendix, we summarize some important properties of the regularly varying distribution of X, which satisfies the assumptions stated in Section 2.1. We are mostly concerned with some continuity properties of the limiting measure $\mu(\cdot)$ (defined in equation (1)).

The measure $\mu(\cdot)$ can be represented as a *product measure* corresponding to the angular component and the radial component [21]. The angular component

$$\Phi(\cdot) = \frac{\mu(\{x\colon \|x\|_2 > 1, x/\|x\|_2 \in \cdot\})}{\mu(\{x\colon \|x\|_2 > 1\})}$$

corresponds to a probability measure on the $(d-1)$-dimensional sphere in $\mathbb{R}^d$. The radial component $\vartheta(dr)$ is a measure that is absolutely continuous with respect to the Lebesgue measure. Moreover, $\vartheta(dr) = cr^{-\alpha-1}\,dr$ for some constant $c > 0$. Then, the measure μ can be written as the product of ϑ and Φ. We then obtain the following lemma.

Lemma 11. *Let* $\Gamma = \{y\colon \eta^T y > \gamma\}$ *and* $R_2 = \{y\colon \max_{i=1}^{2m+d}(y^T v_i - a_i) \leq -\delta_2\}$ *for* $\gamma > 0$ *large but fixed and* $\delta_2 > 0$ *small. Let* $K = \Gamma^c \cup R_2$ *and define* $\kappa(\cdot)$ *for each* $t > 0$, $z \in K$ *and* $a' = (a'_1, \ldots, a'_{2m+d})$ *in a small neighborhood of* $a = (a_1, \ldots, a_{2m+d})$ *via*

$$\kappa_{\mathbf{a}'}(t, z) = \mu\left(\left\{y\colon \max_{j=1}^{2m+d}\left(y^T v_j + z^T v_j - a'_j\right) > t\right\}\right).$$

Then,

$$\kappa_{\mathbf{a}'}(t, z) = \int_{\mathcal{S}_d} \frac{c}{\alpha} \max_{j=1}^{2m+d}\left(\frac{\cos(\theta, v_j)^+}{-z^T v + a'_j + t}\right)^\alpha \Phi(d\theta). \tag{67}$$

Proof. The result follows immediately from the representation of $\mu(\cdot)$ in polar coordinates. We shall sketch the details. For each j, let $\mathcal{H}_j(a'_j, t) = \{y\colon y^T v_j + z^T v_j - a'_j > t\}$ and note that in polar coordinates, we can represent $\mathcal{H}_j(a'_j, t)$ as

$$\{(\theta, r)\colon r > 0 \text{ and } r\|v_j\|_2 \cos(\theta, v_j) \geq -z^T v + a'_j + t\}.$$

Note that $-z^T v + a'_j + t > 0$ for all $t \geq 0$, $j \in \{1, \ldots, 2m+d\}$ and $z \in K$. Therefore,

$$\mu\left(\left\{y\colon \max_{j=1}^{2m+d}\left(y^T v_j + z^T v_j - a'_j\right) > t\right\}\right)$$

$$= \int_{\mathcal{S}_d} \int_0^\infty cr^{-\alpha-1} I\left(r \geq \min_j\{(-z^T v_j + a'_j + t)/\cos(\theta, v_j)^+\}\right) dr\, \Phi(d\theta)$$

$$= \int_{\mathcal{S}_d} \frac{c}{\alpha} \max_{j=1}^{2m+d}\left(\frac{\cos(\theta, v_j)^+}{-z^T v_j + a'_j + t}\right)^\alpha \Phi(d\theta),$$

and the result follows.   □

Lemma 12. *Assume that $K = \Gamma^c \cup R_2$ is defined as in the previous lemma and note that K is a non-empty compact set. Suppose that $z \in K$, define $s = zb$ and write*

$$v_b(zb) = \int_0^\infty P\Big(\max_{i=1}^{2m+d} \big(v_i^T X + z^T v_i b - a_i b\big) > t \Big)\, dt,$$

$$\kappa_{\mathbf{a}}(t, z) = \mu\Big(\Big\{ y\colon \max_{j=1}^{2m+d} \big(y^T v_j + z^T v_j - a_j\big) > t \Big\}\Big).$$

Then,

$$\lim_{b\to\infty}\sup_{z\in K}\left| \frac{v_b(zb)}{b P(\|X\|_2 > b) \int_0^\infty \kappa_{\mathbf{a}}(t, z)\, dt} - 1 \right| \longrightarrow 0$$

as $b \to \infty$.

Proof. We first write

$$v_b(zb) = b\int_0^\infty P\Big(\max_{i=1}^{2m+d} \big(v_i^T X + bz^T v_i - a_i b\big) > ub \Big)\, du,$$

and define

$$\mathcal{L}_b(u, z) = P\Big(\max_{i=1}^{2m+d} \big(v_i^T X + bz^T v_i - a_i b\big) > ub \Big).$$

Set $\varepsilon \in (0, \delta_2)$ arbitrarily small but fixed and let $M = \max_{i\leq 2m+d} \|v_i\|_2 < 0$. Define $\varepsilon' = \varepsilon/M$ and consider an open cover of the set K by balls with radius ε' centered at points $\mathcal{C}_{\varepsilon'} = \{w_1, \ldots, w_{m'}\} \subset K$. We then have that for every $z \in K$ there exists $w_k \triangleq w_k(z) \in \mathcal{C}_{\varepsilon'}$ such that $\|z - w_k\|_2 \leq \varepsilon'$. Note that for each $z \in K$

$$\mathcal{L}_b(u + \varepsilon, w_k) \leq \mathcal{L}_b(u, z) \leq \mathcal{L}_b(u - \varepsilon, w_k).$$

Consequently,

$$\int_0^\infty \mathcal{L}_b(u + \varepsilon, w_k)\, du \leq v_b(zb) \leq \int_0^\infty \mathcal{L}_b(u - \varepsilon, w_k)\, du.$$

Now we claim that

$$\lim_{b\to\infty} \frac{\int_0^\infty \mathcal{L}_b(u + \varepsilon, w_k)\, du}{b P(\|X\|_2 > b) \int_0^\infty \kappa_{\mathbf{a}}(t + \varepsilon, w_k)\, dt} = 1. \tag{68}$$

The previous limit follows from dominated convergence as follows. First, we have that

$$\frac{\mathcal{L}_b(u + \varepsilon, w_k)}{b P(\|X\|_2 > b)} \longrightarrow \kappa_{\mathbf{a}}(u + \varepsilon, w_k)$$

for every u fixed by the definition regular variation. Then, if $u \in (0, C)$ for any $C > 0$ we conclude that

$$\frac{\mathcal{L}_b(u + \varepsilon, w_k)}{b P(\|X\|_2 > b)} \leq \frac{\mathcal{L}_b(0, w_k)}{b P(\|X\|_2 > b)} = O(1)$$

as $b \to \infty$ and therefore by the bounded convergence theorem, we conclude that

$$\lim_{b \to \infty} \frac{\int_0^C \mathcal{L}_b(u + \varepsilon, w_k) \, du}{b P(\|X\|_2 > b)} = \int_0^C \kappa_{\mathbf{a}}(t + \varepsilon, w_k) \, dt.$$

On the set $u \geq C$ we have that if $c_{i,k} = a_i + |w_k^T v_i|$, then

$$\mathcal{L}_b(u + \varepsilon, w_k) \leq \sum_{i=1}^{2m+d} P\big(\|X\|_2 \|v_i\|_2 > (u + \varepsilon)b - c_{i,k}b\big).$$

By Karamata's theorem for one dimensional regularly varying random variables, it follows that if $\alpha > 1$, then for $C \geq \max_{i,k} c_{i,k}$ the functions

$$\frac{P(\|X\|_2 \|v_i\|_2 > (\cdot + \varepsilon)b - c_{i,k}b)}{b P(\|X\|_2 > b)}$$

are uniformly integrable with respect to the Lebesgue measure on $[C, \infty)$ and therefore we conclude that

$$\lim_{b \to \infty} \frac{\int_C^\infty \mathcal{L}_b(u + \varepsilon, w_k) \, du}{b P(\|X\|_2 > b)}$$

$$= \int_C^\infty \kappa_{\mathbf{a}}(t + \varepsilon, w_k) \, dt$$

and therefore the limit (68) holds. Thus, we have that

$$\sup_{\{z \in \|z - w_k\|_2 \leq \varepsilon'\}} \left| \frac{\int_0^\infty \mathcal{L}_b(u + \varepsilon, w_k) \, du}{b P(\|X\|_2 > b) \int_0^\infty \kappa_{\mathbf{a}}(t, z) \, dt} - 1 \right|$$

$$= \sup_{\{z \in \|z - w_k\|_2 \leq \varepsilon'\}} \left| \frac{\int_0^\infty \kappa_{\mathbf{a}}(t + \varepsilon, z) \, dt}{\int_0^\infty \kappa_{\mathbf{a}}(t, z) \, dt} - 1 \right| + o(1)$$

as $b \to \infty$. Observe from representation (67) and Assumption 2 in Section 2.1 we have that

$$\kappa_{\mathbf{a}}(t, z) > \mu\left(\left\{ y \colon \max_{j=1}^{2m+d} y^T v_j > \delta_2 \right\}\right) > 0.$$

Moreover, it also follows as an easy application of the dominated convergence theorem and our representation in (67) that

$$\lim_{\varepsilon \to 0} \sup_{\{z \in \|z - w_k\|_2 \leq \varepsilon'\}} \left| \frac{\int_0^\infty \kappa_{\mathbf{a}}(t + \varepsilon, z) \, dt}{\int_0^\infty \kappa_{\mathbf{a}}(t, z) \, dt} - 1 \right| = 0.$$

We then conclude that

$$
\varlimsup_{b \to \infty} \sup_{z \in K} \left| \frac{\int_0^\infty \mathcal{L}_b(u, z)\, \mathrm{d}u}{b\, P(\|X\|_2 > b) \int_0^\infty \kappa_{\mathbf{a}}(t, z)\, \mathrm{d}t} - 1 \right|
$$

$$
\leq \max_k \varlimsup_{b \to \infty} \sup_{\{z \in \|z - w_k\|_2 \leq \varepsilon'\}} \left| \frac{\int_0^\infty \mathcal{L}_b(u + \varepsilon, w_k)\, \mathrm{d}u}{b\, P(\|X\|_2 > b) \int_0^\infty \kappa_{\mathbf{a}}(t, z)\, \mathrm{d}t} - 1 \right|
$$

$$
+ \max_k \varlimsup_{b \to \infty} \sup_{\{z \in \|z - w_k\|_2 \leq \varepsilon'\}} \left| \frac{\int_0^\infty \mathcal{L}_b(u + \varepsilon, w_k)\, \mathrm{d}u - \int_0^\infty \mathcal{L}_b(u - \varepsilon, w_k)\, \mathrm{d}u}{b\, P(\|X\|_2 > b) \int_0^\infty \kappa_{\mathbf{a}}(t, z)\, \mathrm{d}t} \right|
$$

$$
\leq \max_k \left| \frac{\int_0^\infty \kappa_{\mathbf{a}}(t + \varepsilon, w_k)\, \mathrm{d}t}{\int_0^\infty \kappa_{\mathbf{a}}(t, z)\, \mathrm{d}t} - 1 \right| + \max_k \left| \frac{\int_0^\infty \kappa_{\mathbf{a}}(t + \varepsilon, w_k)\, \mathrm{d}t - \int_0^\infty \kappa_{\mathbf{a}}(t - \varepsilon, w_k)\, \mathrm{d}t}{\int_0^\infty \kappa_{\mathbf{a}}(t, z)\, \mathrm{d}t} \right|.
$$

Once again use the representation in Lemma 11 and the dominated convergence theorem to conclude that the right-hand side of the previous inequality can be made arbitrarily small as $\varepsilon \to 0$, thereby concluding our result. $\qquad\square$

Acknowledgements

References

[1] Asmussen, S. (2000). *Ruin Probabilities. Advanced Series on Statistical Science & Applied Probability* **2**. River Edge, NJ: World Scientific. MR1794582

[2] Asmussen, S. (2003). *Applied Probability and Queues*, 2nd ed. *Applications of Mathematics* (*New York*) **51**. New York: Springer. MR1978607

[3] Asmussen, S. and Binswanger, K. (1997). Simulation of ruin probabilities for subexponential claims. *Astin Bull.* **27** 297–318.

[4] Asmussen, S. and Foss, S. (2012). On exceedance times for some processes with dependent increments. Available at arXiv:1205.5793.

[5] Asmussen, S. and Klüppelberg, C. (1996). Large deviations results for subexponential tails, with applications to insurance risk. *Stochastic Process. Appl.* **64** 103–125. MR1419495

[6] Blanchet, J. and Glynn, P. (2008). Efficient rare-event simulation for the maximum of heavy-tailed random walks. *Ann. Appl. Probab.* **18** 1351–1378. MR2434174

[7] Blanchet, J., Glynn, P. and Liu, J.C. (2007). Fluid heuristics, Lyapunov bounds and efficient importance sampling for a heavy-tailed $G/G/1$ queue. *Queueing Syst.* **57** 99–113. MR2358076

[8] Blanchet, J., Glynn, P. and Liu, J.C. (2012). Efficient rare-event simulation for the steady-state waiting time of a heavy-tailed $G/G/2$ queue. Technical report, Columbia University, New York, NY.

[9] Blanchet, J. and Liu, J. (2010). Efficient importance sampling in ruin problems for multidimensional regularly varying random walks. *J. Appl. Probab.* **47** 301–322. MR2668490

[10] Blanchet, J. and Liu, J. (2012). Efficient simulation and conditional functional limit theorems for ruinous heavy-tailed random walks. *Stochastic Process*. *Appl*. **122** 2994–3031. MR2931349

[11] Blanchet, J.H. and Liu, J. (2008). State-dependent importance sampling for regularly varying random walks. *Adv. in Appl*. *Probab*. **40** 1104–1128. MR2488534

[12] Doney, R.A. and Kyprianou, A.E. (2006). Overshoots and undershoots of Lévy processes. *Ann*. *Appl*. *Probab*. **16** 91–106. MR2209337

[13] Dupuis, P., Leder, K. and Wang, H. (2007). Importance sampling for sums of random variables with regularly varying tails. *ACM Transactions on Modeling and Computer Simulation – TOMACS* **17** 1–21.

[14] Hult, H. and Lindskog, F. (2006). Heavy-tailed insurance portfolios: Buffer capital and ruin probabilities. Technical Report 1441, Cornell Univ., Ithaca, NY.

[15] Hult, H., Lindskog, F., Mikosch, T. and Samorodnitsky, G. (2005). Functional large deviations for multivariate regularly varying random walks. *Ann. Appl. Probab.* **15** 2651–2680. MR2187307

[16] Hult, H. and Samorodnitsky, G. (2010). Large deviations for point processes based on stationary sequences with heavy tails. *J. Appl*. *Probab*. **47** 1–40. MR2654756

[17] Klüppelberg, C., Kyprianou, A.E. and Maller, R.A. (2004). Ruin probabilities and overshoots for general Lévy insurance risk processes. *Ann. Appl. Probab.* **14** 1766–1801. MR2099651

[18] Kontoyiannis, I. and Meyn, S.P. (2005). Large deviations asymptotics and the spectral theory of multiplicatively regular Markov processes. *Electron*. *J. Probab*. **10** 61–123 (electronic). MR2120240

[19] L'Ecuyer, P., Blanchet, J., Tuffin, B. and Glynn, P. (2010). Asymptotic robustness of estimators in rare-event simulation. *ACM Transactions on Modeling and Computer Simulation – TOMACS* **20** 1–41.

[20] Resnick, S.I. (1987). *Extreme Values*, *Regular Variation*, *and Point Processes. Applied Probability. A Series of the Applied Probability Trust* **4**. New York: Springer. MR0900810

[21] Resnick, S.I. (2007). *Heavy-Tail Phenomena*: *Probabilistic and Statistical Modeling. Springer Series in Operations Research and Financial Engineering*. New York: Springer. MR2271424