# IMPORTANCE SAMPLING APPROXIMATIONS TO VARIOUS PROBABILITIES OF RUIN OF SPECTRALLY NEGATIVE LEVY RISK PROCESSES

Riccardo Gatto

Submitted: February 2014
Revised: May 2014

**Abstract**

This article provides importance sampling algorithms for computing the probabilities of various types ruin of spectrally negative Lévy risk processes, which are ruin over the infinite time horizon, ruin within a finite time horizon and ruin past a finite time horizon. For the special case of the compound Poisson process perturbed by diffusion, algorithms for computing probabilities of ruins by creeping (i.e. induced by the diffusion term) and by jumping (i.e. by a claim amount) are provided. It is shown that these algorithms have either bounded relative error or logarithmic efficiency, as $t, x \to \infty$, where $t > 0$ is the time horizon and $x > 0$ is the starting point of the risk process, with $y = t/x$ held constant and assumed either below or above a certain constant.

**Key words and phrases**

Bounded relative error; exponential tilt; Legendre-Fenchel transform; logarithmic efficiency; Lundberg conjugated measure; ruin due to creeping and to jump; ruin past a finite time horizon, within a finite and the infinite time horizons; saddlepoint approximation; short and long time horizons.

---

The author thanks: the Swiss National Science Foundation, for financial support (grant 200021-121901); the Editor-in-Chief, an Associate Editor and two Referees, for important suggestions which improved the quality of this article; Benjamin Baumgartner, for the numerical results of Section 5; and Ilya Molchanov.

2010 Mathematics Subject Classification: 65C05, 60G51.

Address: Institute of Mathematical Statistics and Actuarial Science, Department of Mathematics and Statistics, University of Bern, Alpeneggstrasse 22, 3012 Bern, Switzerland.

Email: `gatto@stat.unibe.ch`.

# 1 Introduction

Stochastic simulation is a popular technique for computing quantities which admit neither simple closed-form expressions nor simple numerical algorithms. Important applications are for computing probabilities of nontrivial rare events and, in particular, for the probability of ruin of the insurer's risk process. Computing accurately a rare event probability by simulation requires selecting an appropriate sampling distribution. This change of measure is referred as importance sampling. An optimal solution is to sample after an exponential tilt of the original probability measure. This was suggested by Siegmund (1976). Importance sampling for probabilities of ruin was suggested by Asmussen (1985) and by Section X.4 of Asmussen (2000). Collamore (2002) proposed importance sampling techniques for the multidimensional ruin problem for general Markov additive sequences of random vectors. Collamore et al. (2014) proposed rare event simulation for processes generated via stochastic fixed point equations. A general reference on stochastic simulation and rare event simulation is Asmussen and Glynn (2007). This article provides generalizations of importance sampling algorithms for finite and infinite time probabilities of ruin of the compound Poisson risk process, originally proposed by Asmussen (2000), Section X.4, to the class of spectrally negative Lévy processes with light-tailed downwards jumps. It also provides importance sampling formulae for the probability of the ruin past a finite time horizon and, in the case of the compound Poisson process perturbed by diffusion, importance sampling estimators to probabilities of ruin by creeping, i.e. due to the diffusion term, and by jumping, i.e. due to a claim amount. The suggested importance sampling algorithms have either bounded relative error or logarithmic efficiency.

The compound Poisson risk process perturbed by diffusion is the $\mathbb{R}^{[0,\infty)}$-valued process $Y$ is defined by

$$Y_t = x + ct - Z_t + \sigma W_t, \ \forall t \geq 0, \tag{1}$$

where $x \geq 0$ is the initial capital, $c > 0$ is the premium rate, $Z = \{Z_t\}_{t\geq 0}$ is a homogeneous compound Poisson process allowing for positive jumps only and representing the total claim amount, $\sigma \geq 0$ and $W$ is an independent standard Wiener standard process representing various uncertainties. Thus $Y$ represents the evolution of the insurer's financial reserve over the time. The drifted compound Poisson process perturbed by diffusion

$$S_t = Z_t - ct - \sigma W_t, \ \ \forall t \geq 0, \tag{2}$$

is called aggregate loss process. As already mentioned, this article considers the loss process $S$ more generally as a spectrally positive Lévy process, which is precisely defined in Section 2.1. This choice allows to retain the diffusion and jump properties of (2) and it brings the following advantages: substantially higher flexibility, general mathematical analysis and direct modelling of the aggregate claim amount (which bypasses compounding of individual claims). Besides the compound Poisson and the Wiener processes, some other basic examples of Lévy processes are the Gamma, the inverse Gaussian and the $\alpha$-stable processes. Lévy processes are commonly used in risk theory and the related literature has become considerable, see e.g. Furrer et al. (1997), Yang and Zhang (2001), Klüppelberg et al. (2004), Morales (2004), Avram et al. (2007), Kyprianou and Palmowski (2007), Biffis and Morales (2010), etc. Palmowski and Pistorius (2009) propose a limiting approximation for the finite time first passage probability of Lévy processes, with the same kind of asymptotics considered in this article. Lévy processes provide adequate models in various other fields than risk theory, like finance, queuing theory, economics, physics, etc. Some recent publications in these other fields are the following. Lévy processes are used in the context of stochastic resonance, where systems perturbed by white Lévy noises are more informative than systems perturbed by white Gaussian noise, see Dybiec (2009), who shows that double stochastic resonance can be observed in a single-well potential perturbed by Lévy stable noises. Dybiec and Gudowska-Nowak (2009) consider an approach to the analysis of noise-induced effects in stochastic dynamics under the influence of Lévy white noise perturbations, describing interactions of the analyzed system with its complex surroundings. In illicit economic transactions, Perc (2007a) shows that Lévy flights facilitate defection. Finally, in the context of evolutionary game theory, Perc (2007b) analyzes the impact of Lévy distributed stochastic payoff variations on the evolution of cooperation in the spatial prisoner's dilemma game. Lévy Processes have become very important in finance, for example for pricing a stock unit, where large jumps account for extreme market fluctuations and small jumps for instantaneous trading, see e.g. Carr et al. (2002). In the context of queuing theory, Kella and Whitt (1992) introduce a special martingale which can be applied to queues and storage processes driven by a Lévy process. Some general references on Lévy processes are Applebaum (2004), Bertoin (1996), Kyprianou (2006) and Sato (1999).

The financial risk inherent to (1), or to its Lévy generalization, can be represented by the following probabilities of ruin, that we want to approximate by importance sampling. Let us first define the time of ruin as

$$T_x = \begin{cases} \inf\left\{t \in (0,\infty) \mid Y_t \le 0\right\}, & \text{if the infimum exists,} \\ \infty, & \text{otherwise.} \end{cases}$$

The probability of ruin within the finite time horizon $[0,t]$ is defined by

$$\psi(x,t) = \mathsf{P}[T_x \le t], \quad \forall t \in (0,\infty). \tag{3}$$

It is the probability that $\{Y_t\}_{t\ge 0}$ falls below the zero line prior to time $t$. The probability of ruin within the infinite time horizon is defined by

$$\psi(x) = \mathsf{P}[T_x < \infty] = \lim_{t\to\infty} \psi(x,t).$$

It is the probability that $\{Y_t\}_{t\geq 0}$ ever falls below the zero line. Finally, the probability of non-ruin within the finite time horizon $[0,t]$ and ruin past time $t$, given by

$$\bar{\psi}(x,t) = \mathsf{P}[t < T_x < \infty] = \mathsf{P}[T_x < \infty] - \mathsf{P}[T_x \leq t] = \psi(x) - \psi(x,t). \tag{4}$$

The infinite time horizon probability of ruin can be decomposed as $\psi(x) = \psi^{(1)}(x) + \psi^{(2)}(x)$, where $\psi^{(1)}(x) = \mathsf{P}[T_x < \infty \wedge Y_{T_x} = 0]$ is the probability that the null line is first crossed by an oscillation of the path of the risk process, i.e. by creeping, and where $\psi^{(2)}(x) = \mathsf{P}[T_x < \infty \wedge Y_{T_x} < 0]$ is the probability that the null line is first crossed by a jump of the path risk process, i.e. by an individual claim amount. Unless $\sigma = 0$, the regularity of the Wiener part of the process yields $\psi(0) = \psi^{(1)}(0) = 1$ and thus $\psi^{(2)}(0) = 0$. We can define $\psi^{(1)}(x,t)$ and $\psi^{(2)}(x,t)$ in the analogous manner to (3). We can also define $\bar{\psi}^{(1)}(x,t)$ and $\bar{\psi}^{(2)}(x,t)$ in the analogous manner to (4). In all these cases, the upper index 1 refers again to cross by creeping and the upper index 2 refers again to cross by jump, of the level zero. In the practice, ruin due to claim is more important than ruin due to oscillation, because in the first case the deficit at ruin can be substantial, whereas it is equal to zero in the second case, from the Wiener part of the process. Although we are primary interested in the insurer's ruin problem, similar probabilities could also be relevant in the context of pricing financial options.

There exist alternative methods to stochastic simulation for computing probabilities of ruin. The saddlepoint approximation is another large deviations approximation which allows to approximate with high accuracy small probabilities of rare events. The saddlepoint approximation in this context originates from Daniels (1954) and Lugannani and Rice (1980). Gatto and Mosimann (2012) and Gatto and Baumgartner (2014) provide comparisons between importance sampling and the saddlepoint approximation, for probabilities of ruin in finite and infinite time horizons, however only for the compound Poisson process with Wiener perturbation. In this article we consider the more general spectrally negative Lévy risk processes. While the accuracy of saddlepoint approximations is very high and usually comparable to that of importance sampling, its computational burden is considerably smaller. As a matter of fact, saddlepoint approximations are conceptually more sophisticated and by far less popular than Monte Carlo methods.

The remaining part of this article has the following structure. Section 2 gives a compact survey of Lévy processes with important definitions, assumptions and results. Section 3 provides efficient importance sampling estimators of probabilities of ruins occurring over the infinite time horizon, within a finite time horizon and past a finite time horizon, for spectrally negative Lévy risk processes. Proofs of asymptotic efficiencies, as $t, x \to \infty$, with $y \stackrel{\text{def}}{=} t/x$ fixed and either bounded from above or bounded from below, are also given. Section 4 provides some efficient estimators of probabilities of ruin by creeping or by jumping, together with proofs of efficiencies, using same asymptotics as before. Section 5 concludes the article with a short numerical illustration.

# 2 Some important facts on Lévy processes

This section presents some main definitions, assumptions and basic results relating to Lévy processes.

## 2.1 Definitions, assumptions and results

Without restricting to the compensated and perturbed compound Poisson loss process $S$ given by (2), we generally assume $S$ is a spectrally positive Lévy process, which is introduced below. The Laplace exponent of any Lévy process $L$ on $\mathbb{R}^{[0,\infty)}$ is defined as

$$\kappa(v) = \log \mathsf{E}\left[\mathrm{e}^{vL_1}\right], \tag{5}$$

$\forall v \in \mathbb{R}$ s.t. $\kappa(v) < \infty$. Thus $\kappa$ is the cumulant generating function of $L_1$ and in fact $t\kappa$ is the cumulant generating function of $L_t$, $\forall t \geq 0$. The Lévy-Khintchine representation yields

$$\kappa(v) = \gamma v + \frac{1}{2}\sigma^2 v^2 + \int_{\mathbb{R}} (\mathrm{e}^{vx} - 1 - vx\,\mathrm{I}\{|x| < 1\})\mathrm{d}\nu(x), \tag{6}$$

where $\gamma \in \mathbb{R}$, $\sigma > 0$ and $\nu$ is a Lévy measure, i.e. a measure on $(\mathbb{R}\backslash\{0\}, \mathcal{B}(\mathbb{R}\backslash\{0\}))$ which satisfies

$$\int_{\mathbb{R}} (1 \wedge x^2)\mathrm{d}\nu(x) < \infty. \tag{7}$$

One can check that any finite measure is a Lévy measure and also that a Lévy measure is necessarily a $\sigma$-finite. The characteristic triplet associated with (6) or with the Lévy process $L$ is $(\gamma, \sigma^2, \nu)$.

An a.s. nondecreasing Lévy process is called subordinator. It can be seen that a stochastic process in $\mathbb{R}_+^{[0,\infty)}$ is a subordinator, if its Laplace exponent is given by

$$\beta v + \int_{\mathbb{R}_+} (\mathrm{e}^{vx} - 1)\mathrm{d}\nu(x), \tag{8}$$

$\forall v \in \mathbb{R}$ s.t. the integral is finite, for some $\beta \geq 0$ and for some Lévy measure $\nu$ satisfying $\nu[(-\infty, 0)] = 0$ and

$$\int_{\mathbb{R}_+} (1 \wedge x)\mathrm{d}\nu(x) < \infty. \tag{9}$$

Clearly, (9) implies (7), whereas (8) can be obtained from (6) with $\sigma^2 = 0$ and $\beta = \gamma - \int_{(0,1)} x\mathrm{d}\nu(x)$.

The Lévy process $L$ with Lévy measure $\nu$ can be decomposed as the sum of: a constant drift (i.e. a linear function), a Wiener process and a jump process $J$ in $\mathbb{R}^{[0,\infty)}$. The jump process itself is the sum of two processes: a first one displaying infinitely many jumps of vanishing magnitude per unit of time, i.e. infinite activity, plus a second one displaying finitely many jumps of substantial or significant magnitude per unit of time. Also, the jump process $J$ is characterized by the Lévy measure $\nu$, which represents the intensity of the jumps. Given the total Lévy mass $\lambda \stackrel{\text{def}}{=} \nu[\mathbb{R}\backslash\{0\}]$, $\lambda < \infty$ iff $J$ is a compound Poisson process with jump size distribution $\xi \stackrel{\text{def}}{=} \nu/\lambda$. This case is characterized by the absence of the first component of $J$, of infinite activity. The process $L$ is called spectrally positive if it is not a subordinator and $\nu[\mathbb{R}_-] = 0$. Further, $L$ is called spectrally negative if $-L$ is spectrally positive. The jumps of a spectrally positive (negative) Lévy process can only be directed upwards (downwards).

We now consider $S$ with Laplace exponent (6). Suppose that $\kappa$ exists over a neighborhood of zero. Let $v \in \mathbb{R}$, then $\kappa(v) < \infty$ if

$$\int_{(-\infty,-1]} |\mathrm{e}^{vx} - 1| \mathrm{d}\nu(x) + \int_{(-1,1)} |\mathrm{e}^{vx} - 1 - vx| \mathrm{d}\nu(x) + \int_{[1,\infty)} |\mathrm{e}^{vx} - 1| \mathrm{d}\nu(x) < \infty. \tag{10}$$

Consider

$$\chi_1(v) \stackrel{\text{def}}{=} \int_{(-\infty,-1]} \mathrm{e}^{vx} \mathrm{d}\nu(x) \text{ and } \chi_2(v) \stackrel{\text{def}}{=} \int_{[1,\infty)} \mathrm{e}^{vx} \mathrm{d}\nu(x). \tag{11}$$

The following simplifications are consequences of (7). As $|\mathrm{e}^{vx} - 1 - vx| \leq \mathrm{e}v^2x^2/2$, $\forall x \in (-1, 1)$ (from Taylor expansion), the second integral in (10) is always finite. If $v > 0$, then $\chi_2(v) < \infty$ is equivalent to the finiteness of the third integral in (10). If $v < 0$, then $\chi_1(v) < \infty$ is equivalent to the finiteness of the first integral in (10). Therefore, $\kappa(v) < \infty$ is equivalent to $\chi_2(v) < \infty$, if $v > 0$, and to $\chi_1(v) < \infty$, if $v < 0$. From the fact that $S$ is assumed spectrally positive, then $\chi_1(v) = 0$, $\forall v \in \mathbb{R}$, and therefore we have:

if $v < 0$, then $\kappa(v) < \infty$, and

if $v > 0$, then $\kappa(v) < \infty \Leftrightarrow \chi_2(v) < \infty$.

The fact $\chi_2(v) < \infty$, for some $v > 0$, is referred as light-tailness of the upwards jumps of the spectrally positive process. The first assumption on $S$ is $\exists s \in (0, \infty]$ s.t. $\lim_{v \to s, v<s} \kappa(v) = \infty$ and $\kappa(s - \varepsilon) < \infty$, $\forall \varepsilon > 0$, which is referred as steepness of the Laplace exponent. This steepness can be simplified to

$$\exists s \in (0, \infty] \text{ s.t. } \lim_{v \to s, v<s} \chi_2(v) = \infty \text{ and } \chi_2(s - \varepsilon) < \infty,\ \forall \varepsilon > 0. \tag{12}$$

It clearly implies light-tailness of upwards jumps. The second assumption is $\mu \stackrel{\text{def}}{=} \mathsf{E}[S_1] < 0$, i.e.

$$\mu = \gamma + \int_{(-\infty,-1]\cup[1,\infty)} x \mathrm{d}\nu(x) < 0, \tag{13}$$

which is referred as net profit condition.

We can finally note the following important result due to Zolotarev (1964), which could be applied for approximating numerically the infinite time horizon probability of ruin,

$$v \int_0^\infty \mathrm{e}^{-vx} \psi(x) \mathrm{d}x = 1 + \frac{v\mu}{\kappa(-v)},$$

$\forall v$ s.t. $\kappa(-v) < \infty$, i.e. $\forall v > 0$ and $\forall v < 0$ s.t. $\chi_2(-v) < \infty$.

## 2.2 Exponential tilt of the probability measure

We define the spectrally positive Lévy loss process $S$ over the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathsf{P})$. The time of ruin $T_x$ is thus a stopping time of $\{\mathcal{F}_t\}_{t\geq 0}$ and the $\sigma$-algebra at stopping time $\mathcal{F}_{T_x} = \{A \in \mathcal{F} \,|\, A \cap \{T_x \leq t\} \in \mathcal{F}_t, \forall t \geq 0\}$ represents the information accumulated until time $T_x$. Let $\theta \in \mathbb{R}$ s.t. $\kappa(\theta) < \infty$. Assume there exists an equivalent probability measure $\mathsf{P}_\theta$ over $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0})$ which transforms the Laplace exponent (5) to

$$\kappa_\theta(v) \overset{\text{def}}{=} \log \mathsf{E}_\theta \left[\mathrm{e}^{vS_1}\right] = \kappa(\theta + v) - \kappa(\theta), \tag{14}$$

$\forall v \in \mathbb{R}$ s.t. $\kappa(\theta + v) < \infty$, where $\mathsf{E}_\theta$ denotes the expectation under $\mathsf{P}_\theta$. Steepness of the Laplace exponent implies $\exists\, \theta, v > 0$ s.t. $\kappa_\theta(v) < \infty$. The measure $\mathsf{P}_\theta$ is the exponential tilt of $\mathsf{P}$ and it easily seen that the class of Lévy processes is algebraically closed under exponential tilting. Precisely, under $\mathsf{P}_\theta$, $S$ remains a Lévy process and it has characteristic triplet $(\gamma_\theta, \sigma_\theta^2, \nu_\theta)$ given by

$$\gamma_\theta = \gamma + \sigma^2\theta + \int_{(-1,1)} x(\mathrm{e}^{\theta x} - 1)\mathrm{d}\nu(x),\ \sigma_\theta^2 = \sigma^2 \ \text{ and } \ \mathrm{d}\nu_\theta(x) = \mathrm{e}^{\theta x}\mathrm{d}\nu(x). \tag{15}$$

Thus, either from (13) and (15), or from computing $\kappa'_\theta(0) = \kappa'(\theta)$, we obtain

$$\mu_\theta \overset{\text{def}}{=} \mathsf{E}_\theta[S_1] = \gamma + \sigma^2\theta + \int_{\mathbb{R}} x(\mathrm{e}^{\theta x} - \mathrm{I}\{|x| < 1\})\mathrm{d}\nu(x). \tag{16}$$

Let $t \geq 0$. If we restrict $\mathsf{P}$ and $\mathsf{P}_\theta$ to $\mathcal{F}_t$, then the Radon-Nikodym derivative of these restricted measures is

$$\frac{\mathrm{d}\mathsf{P}}{\mathrm{d}\mathsf{P}_\theta} = \exp\{-\theta S_t + t\kappa(\theta)\}.$$

This means that, $\forall A \in \mathcal{F}_t$,

$$\mathsf{P}[A] = \mathsf{E}_\theta[\exp\{-\theta S_t + t\kappa(\theta)\}; A].$$

Further, if $A \in \mathcal{F}_{T_x}$ and $A \subset \{T_x < \infty\}$, then

$$\mathsf{P}[A] = \mathsf{E}_\theta[\exp\{-\theta S_{T_x} + T_x\kappa(\theta)\}; A]. \tag{17}$$

The adjustment coefficient or Lundberg's exponent $r$ is the positive solution in $v$ of

$$\kappa(v) = 0, \tag{18}$$

when it exists, and the exponential tilt with $\theta = r$ is called Lundberg conjugation. If the steepness condition (12) holds, then $r$ does indeed exist. In particular, $\mu_r = \kappa_r'(0) = \kappa'(r) > 0$ implies that $S$ has a positive drift under $\mathsf{P}_r$, whence $\mathsf{P}_r[T_x < \infty] = 1$. Lundberg conjugation leads also to the Cramer-Lundberg approximation

$$\psi(x) \sim a_r \, \mathrm{e}^{-rx}, \text{ as } x \to \infty, \tag{19}$$

where $a_r = -\mathsf{E}[S_1]/\mathsf{E}_r[S_1]$.

# 3 Importance sampling algorithms for various probabilities of ruin

This section gives the major results of this article, which are efficient importance sampling algorithms for probabilities of ruin in finite time horizon, in infinite time horizon and past a finite time horizon. These are probabilities of overall ruin, i.e. obtained by creeping or by jump. Ruin by creeping will be addressed in Section 4. The general spectrally negative Lévy risk process is considered here. Section 3.1 gives the exponential tilt estimators of the three probabilities of ruins. The main result providing efficient estimators probabilities of ruins is given in Section 3.2. These estimators are large deviations methods, in the sense that they are either logarithmic efficient or possess bounded relative error, as $t, x \to \infty$, with $y = t/x$ fixed. Section 3.3 gives the proofs of the efficiencies of the estimators.

As mentioned above, the proposed importance sampling estimators will fulfill one of the two standard criteria for Monte Carlo estimators of a sequence of rare events $\{A(x)\}_{x \geq 0}$. This is a sequence of events which satisfies $\theta(x) \overset{\text{def}}{=} \mathsf{P}[A(x)] \overset{x\to\infty}{\longrightarrow} 0$. First, the Monte Carlo estimator $\Theta(x) \overset{\text{def}}{=} \mathrm{I}_{A(x)}$ of $\theta(x)$, $\forall x \geq 0$, is called logarithmic efficient, if

$$\liminf_{x \to \infty} \frac{|\log \operatorname{var}(\Theta(x))|}{|\log \theta^2(x)|} \geq 1. \tag{20}$$

The concept of logarithmic efficiency or logarithmic asymptotics arises from large deviations theory. Second, the estimator possesses bounded relative error if

$$\limsup_{x \to \infty} \frac{\operatorname{var}(\Theta(x))}{\theta^2(x)} < \infty. \tag{21}$$

Note that (20) can be re-expressed as

$$\forall \varepsilon > 0, \ \limsup_{x \to \infty} \frac{\operatorname{var}(\Theta(x))}{\theta^{2-\varepsilon}(x)} < \infty, \tag{22}$$

which is trivially weaker that (21), see also Asmussen and Glynn (2007), p. 159.

### 3.1 Exponential tilt importance sampling estimators

We define the deficit or overshoot at ruin as $D_x = -Y_{T_x} = S_{T_x} - x \geq 0$, on $\{T_x < \infty\}$. Let $t \geq 0$ and $\theta \in \mathbb{R}$ s.t. $\kappa(\theta) < \infty$, i.e. s.t. $\chi_2(\theta) < \infty$, if $\theta > 0$. Setting $A = \{T_x \leq t\}$ in (17) yields $\psi(x,t) = \mathsf{E}_\theta\left[\Psi(x,t,\theta)\right]$, where

$$\Psi(x,t,\theta) = \mathrm{I}\{T_x \leq t\}\,\mathrm{e}^{-\theta S_{T_x} + T_x\kappa(\theta)} = \mathrm{e}^{-\theta x}\mathrm{I}\{T_x \leq t\}\,\mathrm{e}^{-\theta D_x + T_x\kappa(\theta)}$$

is the Monte Carlo estimator of $\psi(x,t)$ under the probability measure $\mathsf{P}_\theta$. The corresponding Monte Carlo approximation, or algorithm, is then $n^{-1}\sum_{k=1}^n \Psi_k(x,t,\theta)$, where $\Psi_1(x,t,\theta),\ldots,\Psi_n(x,t,\theta)$ are independent generations of $\Psi(x,t,\theta)$ under $\mathsf{P}_\theta$. Setting $A = \{T_x < \infty\}$ in (17) yields $\psi(x) = \mathsf{E}_\theta[\Psi(x,\theta)]$, where

$$\Psi(x,\theta) = \mathrm{e}^{-\theta x}\mathrm{I}\{T_x < \infty\}\,\mathrm{e}^{-\theta D_x + T_x\kappa(\theta)}$$

is the Monte Carlo estimator of $\psi(x)$ under $\mathsf{P}_\theta$. The Monte Carlo approximation is $n^{-1}\sum_{k=1}^n \Psi_k(x,\theta)$, where $\Psi_1(x,\theta),\ldots,\Psi_n(x,\theta)$ are independent generations of $\Psi(x,\theta)$ under $\mathsf{P}_\theta$. Setting $A = \{t < T_x < \infty\}$ in (17) yields $\bar{\psi}(x,t) = \mathsf{E}_\theta[\bar{\Psi}(x,t,\theta)]$, where

$$\bar{\Psi}(x,t,\theta) = \mathrm{e}^{-\theta x}\mathrm{I}\{t < T_x < \infty\}\,\mathrm{e}^{-\theta D_x + T_x\kappa(\theta)} = \Psi(x,\theta) - \Psi(x,t,\theta)$$

is the Monte Carlo estimator of $\bar{\psi}(x,t)$ under $\mathsf{P}_\theta$. The Monte Carlo approximation is $n^{-1}\sum_{k=1}^n \bar{\Psi}_k(x,t,\theta)$, where $\bar{\Psi}_1(x,t,\theta),\ldots,\bar{\Psi}_n(x,t,\theta)$ are independent generations of $\bar{\Psi}(x,t,\theta)$ under $\mathsf{P}_\theta$.

### 3.2 Efficient importance sampling estimators

The optimal choices of $\theta$ for the importance sampling algorithms of Section 3.1 are given in Result 3.1 below. Optimality refers either to logarithmic efficiency or to bounded relative error.

**Result 3.1.** *Assume that the net profit condition (13) and the steepness condition (12) hold. Factorize the finite time horizon as $t = xy$, for $y > 0$ fixed, where $x > 0$ is the initial capital. Let $v_y$ be the solution in $v$ of $\kappa'(v) = 1/y$, i.e. of*

$$y\left(\gamma + \sigma^2 v + \int_{\mathbb{R}} x\left(\mathrm{e}^{vx} - \mathrm{I}\{|x| \leq 1\}\right)\mathrm{d}\nu(x)\right) = 1 \tag{23}$$

*Let*

$$y_r = \frac{1}{\mu_r}, \tag{24}$$

*where $r$ is the adjustment coefficient given by (18) and $\mu_r$ is given by (16). In the finite time horizon, $t < \infty$ and we distinguish the two following cases:*

*the short time horizon, where $t < x/\mu_r \Leftrightarrow y < y_r$, and*

*the long time horizon, where $t > x/\mu_r \Leftrightarrow y > y_r$.*

*The five following importance sampling estimators are at least logarithmic efficient.*

1. *In the short time horizon,*

$$\Psi(x,t,v_y) = \mathrm{e}^{-v_y x}\mathrm{I}\{T_x \le t\}\ \mathrm{e}^{-v_y D_x + T_x \kappa(v_y)}$$

*is a logarithmic efficient estimator of* $\psi(x,t)$*, as* $t, x \to \infty$*, under* $\mathsf{P}_{v_y}$*.*

2. *In the long time horizon,*

$$\Psi(x,t,r) = \mathrm{e}^{-rx}\mathrm{I}\{T_x \le t\}\ \mathrm{e}^{-rD_x}$$

*is an estimator with bounded relative error of* $\psi(x,t)$*, as* $t, x \to \infty$*, under* $\mathsf{P}_r$*.*

3. *In the infinite time horizon,*

$$\Psi(x,r) = \mathrm{e}^{-rx}\mathrm{e}^{-rD_x}$$

*is an estimator with bounded relative error of* $\psi(x)$*, as* $x \to \infty$*, under* $\mathsf{P}_r$*.*

4. *In the short time horizon,*

$$\bar{\Psi}(x,t,r) = \mathrm{e}^{-rx}\mathrm{I}\{t < T_x\}\ \mathrm{e}^{-rD_x}$$

*is an estimator with bounded relative error of* $\bar{\psi}(x,t)$*, as* $t, x \to \infty$*, under* $\mathsf{P}_r$*.*

5. *In the long time horizon,*

$$\bar{\Psi}(x,t,v_y) = \mathrm{e}^{-v_y x}\mathrm{I}\{t < T_x\}\ \mathrm{e}^{-v_y D_x + T_x \kappa(v_y)}$$

*is a logarithmic efficient estimator of* $\bar{\psi}(x,t)$*, as* $t, x \to \infty$*, under* $\mathsf{P}_{v_y}$*.*

In the terminology of asymptotic analysis, $v_y$ is the saddlepoint of the function $t\kappa$ at $x$.

The algorithms for computing any of the quantities mentioned under parts 1-5 of Result 3.1 are obtained by generating a large number of realizations of the corresponding estimator, under the proposed exponentially tilted probability measure. The average of these realizations yields the efficient estimator. The details of this procedure applied to $\psi(x,t)$ in the short time horizon, i.e. to part 1 of Result 3.1, are given in the next algorithm. Similar algorithms can be written for parts 2-5 of Result 3.1.

**Algorithm 3.2.** *Importance sampling for* $\psi(x,t)$ *in the short time horizon consists in the following steps.*

1. *Compute the saddlepoint* $v_y$ *by solving (23).*
2. *Iterate the following steps a large number of times.*
   - *(a) Generate a path of the Lévy process under* $\mathsf{P}_{v_y}$*, that is with the characteristic triplet* $(\gamma_{v_y}, \sigma^2_{v_y}, \nu_{v_y})$ *determined from (15).*
   - *(b) Obtain the associated values of the overshoot* $D_x$ *and time of ruin* $T_x$*.*
   - *(c) Compute and store* $\Psi(x,t,v_y) = \mathrm{e}^{-v_y x}\mathrm{I}\{T_x \le t\}\ \mathrm{e}^{-v_y D_x + T_x \kappa(v_y)}$*.*
3. *Approximate* $\psi(x,t)$ *by the average of the generated values* $\Psi(x,t,v_y)$*.*

## 3.3 Proofs of efficiencies of importance sampling estimators

The following lemmas are required by the proof of Result 3.1, which is given at the end of the section.

**Lemma 3.3.** *Assume that the net profit condition (13) and the steepness condition (12) hold. Define*

$$v_0 = \text{arginf}_{v \in \mathbb{R}} \kappa(v) \quad \textit{and} \tag{25}$$

$$l_y = v_y - \kappa(v_y) y. \tag{26}$$

*Then*

$$\begin{aligned} &v_0 < r < l_y < v_y, \quad \textit{in the short time horizon, i.e. when } \; y < y_r, \quad \textit{and} \\ &v_0 < v_y < r < l_y, \quad \textit{in the long time horizon, i.e. when } \; y > y_r, \end{aligned}$$

*where $r$ is the adjustment coefficient given by (18), $v_y$ is the saddlepoint given by (23) and $y_r$ is given by (24).*

**Remark 3.4.** *The Legendre-Fenchel transform of the cumulant generating function $t\kappa$ is given by*

$$\Lambda_{x,y} = \sup_{v \in (-\infty, s)} vx - t\kappa(v). \tag{27}$$

*where $s$ is the steepness point of the Laplace exponent given in (12). Thus we have*

$$\Lambda_{x,y} = v_y x - xy\kappa(v_y) = x\{v_y - y\kappa(v_y)\} = x l_y.$$

Lemma 3.3 is a direct consequence of the convexity of the cumulant generating function $\kappa$.

**Lemma 3.5.** *Assume that the net profit condition (13) and the steepness condition (12) hold. Let $\theta > v_0$ s.t. $\kappa(\theta) < \infty$, where $v_0$ is defined by (25), and let $\tau(\theta) = \sqrt{\kappa''(\theta)/\mu_\theta^3}$. Then,*

$$\lim_{x \to \infty} \frac{T_x}{x} = \frac{1}{\mu_\theta}, \; \mathsf{P}_\theta \; \textit{a.s.}, \tag{28}$$

$$\frac{T_x - \frac{x}{\mu_\theta}}{\sqrt{x}} \xrightarrow{\text{d}} \mathcal{N}\left(0, \tau^2(\theta)\right), \; \textit{as } x \to \infty, \quad \textit{under } \mathsf{P}_\theta, \tag{29}$$

$$\lim_{x \to \infty} \frac{\mathsf{E}_\theta[T_x]}{x} = \frac{1}{\mu_\theta} \tag{30}$$

*and, given $T_x < \infty$,*

$$\frac{T_x}{x} \xrightarrow{\mathsf{P}} \frac{1}{\mu_r}, \quad \textit{as } x \to \infty, \tag{31}$$

*where $r$ is the adjustment coefficient given by (18).*

*Proof of Lemma 3.5.* The Strong law of large numbers yields

$$\lim_{t\to\infty}\frac{S_t}{t}=\Big(\frac{S_{\lfloor t\rfloor}}{\lfloor t\rfloor}+\underbrace{\frac{S_{\lfloor t\rfloor}-S_t}{\lfloor t\rfloor}}_{\overset{t\to\infty}{\longrightarrow}0}\Big)\underbrace{\frac{\lfloor t\rfloor}{t}}_{\overset{t\to\infty}{\longrightarrow}1}=\mu_\theta,\ \mathsf{P}_\theta\text{-a.s.} \tag{32}$$

From $\mu_\theta>0$ follows that

$$\mathsf{P}_\theta[T_x<\infty]=1 \tag{33}$$

and, because $T_x$ is nondecreasing in $x$ and $\mathsf{P}_\theta$-a.s. unbounded,

$$\mathsf{P}_\theta\left[\lim_{x\to\infty}T_x=\infty\right]=1. \tag{34}$$

From (32), (33), (34) and from $D_x=\mathrm{o}(x)$, as $x\to\infty$, $\mathsf{P}_\theta$-a.s., we find

$$\frac{1}{\mu_\theta}=\lim_{t\to\infty}\frac{t}{S_t}=\lim_{x\to\infty}\frac{T_x}{S_{T_x}}=\lim_{x\to\infty}\frac{T_x}{x+D_x}=\lim_{x\to\infty}\frac{T_x}{x},\ \mathsf{P}_\theta\text{-a.s.},$$

i.e. (28).

From the Central limit theorem,

$$\frac{S_t-t\mu_\theta}{\sqrt{t}}=\Big(\frac{S_{\lfloor t\rfloor}-\lfloor t\rfloor\mu_\theta}{\sqrt{\lfloor t\rfloor}}+\underbrace{\frac{S_t-S_{\lfloor t\rfloor}}{\sqrt{\lfloor t\rfloor}}}_{\overset{t\to\infty}{\longrightarrow}0}-\underbrace{\frac{t-\lfloor t\rfloor}{\sqrt{\lfloor t\rfloor}}\mu_\theta}_{\overset{t\to\infty}{\longrightarrow}0}\Big)\underbrace{\sqrt{\frac{\lfloor t\rfloor}{t}}}_{\overset{t\to\infty}{\longrightarrow}1}\overset{\mathrm{d}}{\longrightarrow}\mathcal{N}\left(0,\kappa''(\theta)\right),\quad\text{as } t\to\infty,$$

under $\mathsf{P}_\theta$, here and thereafter. This asymptotic normality and condition (28) allow to apply Anscombe's theorem, which yields

$$\frac{S_{T_x}-T_x\mu_\theta}{\sqrt{T_x}}\overset{\mathrm{d}}{\longrightarrow}\mathcal{N}\left(0,\kappa''(\theta)\right),\quad\text{as } x\to\infty,$$

i.e.

$$\frac{x+D_x-T_x\mu_\theta}{\sqrt{T_x}}\overset{\mathrm{d}}{\longrightarrow}\mathcal{N}\left(0,\kappa''(\theta)\right),\quad\text{as } x\to\infty.$$

This last result with $D_x=\mathrm{o}(x)$, as $x\to\infty$, a.s., yield

$$\frac{T_x-\frac{x}{\mu_\theta}}{\sqrt{T_x}}\overset{\mathrm{d}}{\longrightarrow}\mathcal{N}\left(0,\frac{\kappa''(\theta)}{\mu_\theta^2}\right),\quad\text{as } x\to\infty,$$

and (29) is obtained with Slutski's theorem.

Regarding (30), from $\mathsf{E}_\theta[D_x]=\mathrm{o}(x)$, as $x\to\infty$, we find

$$\mathsf{E}_\theta[S_{T_x}]=x+\mathsf{E}_\theta[D_x]=x+\mathrm{o}(x),\ \text{as } x\to\infty. \tag{35}$$

Because $\{\tilde{S}_t\}_{t\geq 0}\overset{\mathrm{def}}{=}\{S_t-t\mathsf{E}_\theta[S_1]\}_{t\geq 0}$ is a $(\mathsf{P}_\theta,\{\mathcal{F}_t\}_{t\geq 0})$-martingale and $\mathsf{P}_\theta[T_x<\infty]=1$, the Optional stopping theorem tells $\mathsf{E}_\theta[\tilde{S}_{T_x}]=\mathsf{E}_\theta[\tilde{S}_0]$, which yields

$$\mathsf{E}_\theta[S_{T_x}]=\mathsf{E}_\theta[T_x]\mathsf{E}_\theta[S_1]=\mathsf{E}_\theta[T_x]\mu_\theta. \tag{36}$$

Thus (35) and (36) imply (30).

To show (31), let $\varepsilon > 0$, then

$$\begin{aligned}
\mathsf{P}\left[\left|\frac{T_x}{x}-\frac{1}{\mu_r}\right|>\varepsilon \,\middle|\, T_x<\infty\right] &= \frac{\mathsf{P}\left[\left|\frac{T_x}{x}-\frac{1}{\mu_r}\right|>\varepsilon \wedge T_x<\infty\right]}{\mathsf{P}[T_x<\infty]} \\
&= \frac{\mathrm{e}^{-rx}\mathsf{E}_r\left[\mathrm{e}^{-rD_x}; \left|\frac{T_x}{x}-\frac{1}{\mu_r}\right|>\varepsilon\right]}{\mathrm{e}^{-rx}\mathsf{E}_r[\mathrm{e}^{-rD_x}]} \\
&\leq \frac{\mathsf{P}_r\left[\left|\frac{T_x}{x}-\frac{1}{\mu_r}\right|>\varepsilon\right]}{\mathsf{E}_r[\mathrm{e}^{-rD_x}]} \\
&\overset{x\to\infty}{\longrightarrow} 0,
\end{aligned}$$

from (28). □

**Lemma 3.6.** *Assume that the net profit condition (13) and the steepness condition (12) hold. In the short time horizon, i.e. for fixed $y < y_r$, we have*

$$-\frac{\log\psi(x,xy)}{x} \overset{x\to\infty}{\longrightarrow} l_y, \;\; i.e. \;\; \psi(x,xy) \sim \mathrm{e}^{-\Lambda_{x,y}}, \;\; as\ x\to\infty,$$

*where $l_y$ is given by (26) and $\Lambda_{x,y}$ is given by (27).*

*Proof of Lemma 3.6.* From Lemma 3.3, $r < v_y$, when $y < y_r$, and so $\kappa(v_y) > 0$, where $\kappa$ is given by (6), $y_r$ by (24), $v_y$ by (23) and $r$ by (18). So we have

$$\begin{aligned}
\psi(x,xy) &\geq \mathsf{E}_{v_y}\left[\exp\{-v_y S_{T_x} + T_x\kappa(v_y)\}; xy-\sqrt{x}\tau(v_y) < T_x \leq xy\right] \\
&= \exp\{-v_y x + \kappa(v_y)xy\} \\
&\quad\; \mathsf{E}_{v_y}\left[\exp\{-v_y D_x + (T_x - xy)\kappa(v_y)\}; xy-\sqrt{x}\tau(v_y) < T_x \leq xy\right] \\
&\geq \exp\{-l_y x\}\mathsf{E}_{v_y}\left[\exp\{-v_y D_x - \sqrt{x}\tau(v_y)\kappa(v_y)\}; xy-\sqrt{x}\tau(v_y) < T_x \leq xy\right] \\
&= \exp\{-l_y x - \kappa(v_y)\sqrt{x}\tau(v_y)\}\mathsf{E}_{v_y}\left[\exp\{-v_y D_x\}; -1 < \frac{T_x - \frac{x}{\mu_{v_y}}}{\sqrt{x}\tau(v_y)} \leq 0\right] \\
&\sim \exp\{-l_y x - \kappa(v_y)\sqrt{x}\tau(v_y)\}u(y)\left\{\Phi(1)-\frac{1}{2}\right\}, \quad \text{as } x\to\infty, \qquad (37)
\end{aligned}$$

where

$$u(y) = \lim_{x\to\infty}\mathsf{E}_{v_y}\left[\exp\{-v_y D_x\}\right] \qquad (38)$$

and $\Phi$ denotes the standard normal distribution function. The asymptotic equivalence in (37) is due to Stam's Lemma, which states that $D_x$ and $T_x$ are asymptotically independent, as $x \to \infty$, and to (29) of Lemma 3.5. Thus, from (37),

$$\liminf_{x\to\infty}\frac{\log\psi(x,xy)}{x} \geq -l_y.$$

The analogous result with limsup replacing liminf and reversed inequality can be obtained in a similar way and the lemma is thus proved. □

**Lemma 3.7.** *Assume that the net profit condition (13) and the steepness condition (12) hold. Consider the factorization $t = xy$, with fixed $y > y_r$, i.e. in the long time horizon. Then*

$$\frac{\psi(x,xy)}{\psi(x)} \overset{x\to\infty}{\longrightarrow} \begin{cases} 0, & \textit{if } \ y < y_r, \ \textit{ i.e. in the short time horizon}, \\ 1, & \textit{if } \ y > y_r, \ \textit{ i.e. in the long time horizon}, \end{cases} \tag{39}$$

*where $y_r$ is given by (24).*

*Proof of Lemma 3.7.* Indeed (31) implies that

$$\frac{\psi(x,xy)}{\psi(x)} = \frac{\mathsf{P}[T_x \le xy]}{\mathsf{P}[T_x < \infty]} = \mathsf{P}[T_x \le xy \,|\, T_x < \infty] \overset{x\to\infty}{\longrightarrow} \begin{cases} 0, & \text{if } \ y < y_r, \\ 1, & \text{if } \ y > y_r. \end{cases}$$

□

Note that the part relating to short time horizon of Lemma 3.7 can be refined as follows. From Lemma 3.6 and Cramer-Lundberg's approximation (19), we obtain

$$\frac{\psi(x,xy)}{\psi(x)} \sim a_r^{-1} \mathrm{e}^{(r-l_y)x} \overset{x\to\infty}{\longrightarrow} 0,$$

because Lemma 3.3 tells $r < l_y$.

Lemma 3.8 below, when restricted to the compound Poisson process, is known as Arfwedson's saddlepoint approximation. It is thus the generalization to Lévy processes.

**Lemma 3.8.** *Assume that the net profit condition (13) and the steepness condition (12) hold. Consider the factorization $t = xy$, for some fixed $y > y_r$, with $y_r$ given by (24), and thus the long time horizon. Then*

$$\bar\psi(x,xy) \sim b_y x^{-\frac{1}{2}} \mathrm{e}^{-\Lambda_{x,y}}, \quad \textit{as } x\to\infty,$$

*where*

$$b_y = -\frac{u(y)}{\sqrt{2\pi}\kappa(v_y)\tau(v_y)} > 0$$

*and where $\Lambda_{x,y}$ is given by (27).*

*Proof of Lemma 3.8.* From Stam's Lemma, we obtain

$$\begin{aligned} \bar\psi(x,xy) &= \mathrm{e}^{-v_y x}\mathsf{E}_{v_y}\left[\exp\{-v_y D_x + \kappa(v_y)T_x\}; T_x > xy\right] \\ &\sim \mathrm{e}^{-v_y x} u(y) \mathsf{E}_{v_y}\left[\exp\{\kappa(v_y)T_x\}; T_x > xy\right], \quad \text{as } \ x\to\infty, \end{aligned}$$

where $u(y)$ is given by (38). Let $Z \sim \mathcal{N}(0,1)$. Lemma 3.5 and $\mu_{v_y} = 1/y$ yield, as $x\to\infty$,

$$\begin{aligned} \mathsf{E}_{v_y}\left[\exp\{\kappa(v_y)T_x\}; T_x > xy\right] &\sim \mathsf{E}\left[\exp\{\kappa(v_y)[\sqrt{x}\tau(v_y)Z + xy]\}; Z > 0\right] \\ &= \mathrm{e}^{\kappa(v_y)xy}\int_0^\infty \exp\{-[-\kappa(v_y)\sqrt{x}\tau(v_y)]z\}\mathrm{d}\Phi(z) \\ &= \frac{\mathrm{e}^{\kappa(v_y)xy}}{-\kappa(v_y)\tau(v_y)\sqrt{2\pi x}} \underbrace{\int_0^\infty \mathrm{e}^{-u}\exp\left\{-\frac{u^2}{2\kappa^2(v_y)\tau^2(v_y)x}\right\}\mathrm{d}u}_{\overset{x\to\infty}{\longrightarrow} 1}. \end{aligned}$$

□

We can now give the proof of the main result.

*Proof of Result 3.1.* 1. In the short time horizon $y < y_r$,

$$\begin{aligned}\mathsf{E}_{v_y}\left[\Psi^2(x,xy,v_y)\right] &\le \mathrm{e}^{-2v_y x}\mathsf{E}_{v_y}\left[\exp\{2[xy\kappa(v_y) - v_y D_x]\}; T_x \le xy\right] \\ &\le \mathrm{e}^{-2v_y x}\mathsf{E}_{v_y}\left[\exp\{2[x(v_y - l_y) - v_y D_x]\}\right] \\ &\le \mathrm{e}^{-2l_y x},\end{aligned}$$

from (26). With Lemma 3.6 we now obtain

$$\begin{aligned}\liminf_{x\to\infty}\frac{-\log \operatorname{var}_{v_y}\left(\Psi^2(x,xy,v_y)\right)}{-\log\psi^2(x,xy)} &\ge \frac{1}{2}\liminf_{x\to\infty}\frac{-\log \mathsf{E}_{v_y}\left[\Psi^2(x,xy,v_y)\right]}{l_y x} \\ &\ge \liminf_{x\to\infty}\frac{l_y x}{l_y x} \\ &= 1,\end{aligned}$$

which justifies logarithmic efficiency in the short time horizon.

2. In the long time horizon $y > y_r$, (39) and the Cramer-Lundberg approximation (19) yield

$$\psi(x,xy) \sim \psi(x) \sim a_r \mathrm{e}^{-rx}, \text{ as } x \to \infty.$$

Moreover,

$$\mathsf{E}_r\left[\Psi^2(x,xy,r)\right] \le \mathrm{e}^{-2rx}\mathsf{E}_r\left[\mathrm{e}^{-2rD_x}\right] \le \mathrm{e}^{-2rx}.$$

Thus

$$\limsup_{x\to\infty}\frac{\mathsf{E}_r\left[\Psi^2(x,xy,r)\right]}{\psi^2(x,xy)} \le \limsup_{x\to\infty}\frac{\mathrm{e}^{-2rx}}{(a_r\mathrm{e}^{-rx})^2} = a_r^{-2} < \infty,$$

which justifies bounded relative error in the long time horizon.

3. In the infinite time horizon, we have

$$\mathsf{E}_r\left[\Psi^2(x,r)\right] = \mathrm{e}^{-2rx}\mathsf{E}_r\left[\mathrm{e}^{-2rD_x}\right] \le \mathrm{e}^{-2rx}.$$

From this fact and from the Cramer-Lundberg approximation (19), we have

$$\limsup_{x\to\infty}\frac{\mathsf{E}_r\left[\Psi^2(x,r)\right]}{\psi^2(x)} \le \limsup_{x\to\infty}\frac{\mathrm{e}^{-2rx}}{(a_r\mathrm{e}^{-rx})^2} = a_r^{-2} < \infty,$$

which justifies the bounded relative error in the infinite time horizon.

4. In the short time horizon $y < y_r$, we have

$$\bar{\psi}(x,xy) = \psi(x) - \psi(x,xy) = \psi(x) + \mathrm{o}(\psi(x)) = a_r\mathrm{e}^{-rx}\{1 + \mathrm{o}\,(1)\}, \text{ as } x \to \infty,$$

where the second equality above is due to (39) and the third one to (19). So we obtain

$$
\begin{aligned}
\limsup_{x\to\infty}\frac{\mathsf{E}_r\left[\bar{\Psi}^2(x,xy,r)\right]}{\bar{\psi}^2(x,xy)} &= \limsup_{x\to\infty}\frac{\mathrm{e}^{-2rx}\mathsf{E}_r[\mathrm{e}^{-2rD_x};xy<T_x]}{(a_r\mathrm{e}^{-rx})^2}\\
&\le \limsup_{x\to\infty}\frac{\mathrm{e}^{-2rx}}{(a_r\mathrm{e}^{-rx})^2}\\
&= a_r^{-2}<\infty,
\end{aligned}
$$

which justifies bounded relative error in the short time horizon.

5. In the long time horizon $y>y_r$, Lemma 3.3 yields $\kappa(v_y)<0$ and thus

$$
\begin{aligned}
\mathsf{E}_{v_y}\left[\bar{\Psi}^2(x,xy,v_y)\right] &\le \mathrm{e}^{-2v_yx}\mathsf{E}_{v_y}\left[\exp\{2[xy\kappa(v_y)-v_yD_x]\};T_x>xy\right]\\
&\le \mathrm{e}^{-2v_yx}\mathsf{E}_{v_y}\left[\exp\{2[x(v_y-l_y)-v_yD_x]\}\right]\\
&\le \mathrm{e}^{-2l_yx},
\end{aligned}
$$

from (26). This inequality and Lemma 3.8 lead to

$$
\begin{aligned}
\limsup_{x\to\infty}\frac{\mathsf{E}_{v_y}\left[\bar{\Psi}^2(x,xy,v_y)\right]}{\bar{\psi}^{2-\varepsilon}(x,xy)} &\le \limsup_{x\to\infty}\frac{\mathrm{e}^{-2l_yx}}{b_y^{2-\varepsilon}x^{\frac{\varepsilon}{2}-1}\mathrm{e}^{-(2-\varepsilon)l_yx}}\\
&= b_y^{\varepsilon-2}\lim_{x\to\infty}x^{1-\frac{\varepsilon}{2}}\mathrm{e}^{-\varepsilon l_yx}\\
&=0,\quad \forall\varepsilon>0,
\end{aligned}
$$

because $b_y$ is a fixed constant. This justifies logarithmic efficiency, expressed in terms of (22), in the long time horizon.

□

**Remark 3.9.** Besides the above detailed proof, parts of Result 3.1 can be intuitively understood as follows. We can first note that the sampling measure of the long time horizon of case 2, is the same as the one of the infinite time horizon of case 3, namely Lundberg's conjugated measure $\mathsf{P}_r$, which makes ruin almost sure in the infinite time horizon. The variability of $\exp\{-\theta D_x+T_x\kappa(\theta)\}$ is indeed substantially reduced when $\theta=r$, because $T_x\kappa(r)=0$ and also $D_x=\mathrm{o}(x)$, $x\to\infty$, $\mathsf{P}_r$-a.s.

Regarding the short time horizon of case 1, note first that if $\theta>v_0$, then $\mu_\theta=\kappa_\theta'(0)=\kappa'(\theta)>0$. So given $v_y>v_0$, (30) of Lemma 3.5 yields

$$
\mathsf{E}_{v_y}[T_x]\sim\frac{x}{\mu_{v_y}}=xy=t,\quad \text{as } x\to\infty.
$$

As a consequence, $\mathsf{P}_{v_y}$ is a re-centering of $\mathsf{P}$ at the asymptotic mean of $T_x$, which is a sensible shift for approximating the distribution of $T_x$ by simulation (or even by other ways). Note that just like $\mathsf{P}_r$ does, $\mathsf{P}_{v_y}$ does also make ruin almost sure in the infinite time horizon, although it does not remove the time of ruin $T_x$ from the exponent of the estimator.

# 4 Importance sampling algorithms for various probabilities of ruin by creeping and by jumping

This section is concerned with various probabilities of ruin by creeping and jumping presented at the end of Section 1 and it is restricted to the compound Poisson risk processes perturbed by diffusion (1). Section 4.1 presents some main results and exponential tilt importance sampling estimators to probabilities of ruin by creeping and by jumping. Section 4.2 gives the efficient importance sampling estimators in the infinite time horizon and past a finite time horizon. Proofs of efficiencies are given in Section 4.3.

## 4.1 Generalities and exponential tilt importance sampling estimators

Let $Z$ be a compound Poisson process with claim occurrence rate $\lambda = \int_{\mathbb{R}_+} \mathrm{d}\nu(x) \in (0,\infty)$ and individual claim amount distribution $\xi = \nu/\lambda$, where the Lévy measure $\nu$ is defined on $(\mathbb{R}_+\backslash\{0\}, \mathcal{B}(\mathbb{R}_+\backslash\{0\}))$. So the Lévy process $S$ given in (2) is the drifted and perturbed compound Poisson loss process and its characteristic triplet is given by $\gamma = \lambda \int_0^1 x\mathrm{d}\xi(x) - c$, $\nu = \lambda\xi$ and by $\sigma$ equal to the one of (2). The Laplace exponent of $S$, or cumulant generating function of $S_1$, is given by (6) and simplifies to

$$\kappa(v) = \frac{1}{2}v^2\sigma^2 - cv + \lambda\big\{M_\xi(v) - 1\big\},$$

where $M_\xi(v) = \int_{\mathbb{R}_+} \mathrm{e}^{vx}\mathrm{d}\xi(x)$ is the moment generating function of the individual claim amounts. We assume $\kappa$ steep, which implies that $\kappa$ and thus $M_\xi$ exist over a neighborhood of zero, and we assume $\mu = \mathsf{E}[S_1] = \lambda M'_\xi(0) - c < 0$, which is the net profit condition.

Dufresne and Gerber (1991) provide the Cramer-Lundberg approximation for the probability of ruin by creeping. It has the form

$$\psi^{(1)}(x) \sim a_r^{(1)}\,\mathrm{e}^{-rx}, \text{ as } x \to \infty, \tag{40}$$

for some constant $a_r^{(1)} > 0$ depending on $r$, the adjustment coefficient (18).

Let $t \geq 0$ and $\theta \in \mathbb{R}$ s.t. $M_\xi(\theta) < \infty$. Setting $A = \{T_x \leq t \wedge D_x = 0\}$ in (17) yields $\psi^{(1)}(x,t) = \mathsf{E}_\theta\left[\Psi^{(1)}(x,t,\theta)\right]$, where

$$\Psi^{(1)}(x,t,\theta) = \mathrm{e}^{-\theta x}\mathrm{I}\{T_x \leq t \wedge D_x = 0\}\,\mathrm{e}^{T_x\kappa(\theta)}$$

is the Monte Carlo estimator under $\mathsf{P}_\theta$. Setting $A = \{T_x < \infty \wedge D_x = 0\}$ in (17) yields $\psi^{(1)}(x) = \mathsf{E}_\theta[\Psi^{(1)}(x,\theta)]$, where

$$\Psi^{(1)}(x,\theta) = \mathrm{e}^{-\theta x}\mathrm{I}\{T_x < \infty \wedge D_x = 0\}\,\mathrm{e}^{T_x\kappa(\theta)}$$

is the Monte Carlo estimator under $\mathsf{P}_\theta$. Setting $A = \{t < T_x < \infty \wedge D_x = 0\}$ in (17) yields $\bar{\psi}^{(1)}(x,t) = \mathsf{E}_\theta[\bar{\Psi}^{(1)}(x,t,\theta)]$, where

$$\bar{\Psi}^{(1)}(x,t,\theta) = \mathrm{e}^{-\theta x}\mathrm{I}\{t < T_x < \infty \wedge D_x = 0\}\,\mathrm{e}^{T_x\kappa(\theta)} = \Psi^{(1)}(x,\theta) - \Psi^{(1)}(x,t,\theta)$$

is the Monte Carlo estimator of $\bar{\psi}^{(1)}(x,t)$ under $\mathsf{P}_\theta$. The estimators $\Psi^{(2)}(x,t,\theta)$, $\Psi^{(2)}(x,\theta)$ and $\bar{\Psi}^{(2)}(x,t,\theta)$ are obtained in the same manner, after replacing $D_x = 0$ by $D_x > 0$ in the three events $A$ above. In the practice, the Monte Carlo approximations are based on independent generations of these estimators, with $D_x = 0$ replaced by $D_x < \varepsilon$ and $D_x > 0$ replaced by $D_x \geq \varepsilon$, for a threshold $\varepsilon > 0$ which selected small w.r.t. the discretization unit of the time axis.

## 4.2 Efficient importance sampling estimators

The choices of $\theta$ in the above mentioned importance sampling algorithms which lead to logarithmic efficiency or to bounded relative error are given in Result 4.1 below.

**Result 4.1.** *Consider the compound Poisson risk processes perturbed by diffusion (1). Assume that that the net profit condition $\lambda M'_\xi(0) - c < 0$ and the steepness condition (12) on $\kappa$ hold. Let $v_y$ and $r$ be defined by (23) and (18), respectively. The six following importance sampling estimators have bounded relative error.*

1. *In the infinite time horizon,*

$$\Psi^{(1)}(x,r) = \mathrm{e}^{-rx}\mathrm{I}\{D_x = 0\}$$

*and*

$$\Psi^{(2)}(x,r) = \mathrm{e}^{-rx}\mathrm{I}\{D_x > 0\}\mathrm{e}^{-rD_x}$$

*are estimators with bounded relative error of $\psi^{(1)}(x)$ and $\psi^{(2)}(x)$, as $x \to \infty$, under $\mathsf{P}_r$.*

2. *In the short time horizon,*

$$\bar{\Psi}^{(1)}(x,t,r) = \mathrm{e}^{-rx}\mathrm{I}\{t < T_x \wedge D_x = 0\}$$

*and*

$$\bar{\Psi}^{(2)}(x,t,r) = \mathrm{e}^{-rx}\mathrm{I}\{t < T_x \wedge D_x > 0\}\mathrm{e}^{-rD_x}$$

*are estimators with bounded relative error of $\bar{\psi}^{(1)}(x,t)$ and $\bar{\psi}^{(2)}(x,t)$, as $t, x \to \infty$, under $\mathsf{P}_r$.*

3. *In the long time horizon,*

$$\bar{\Psi}^{(1)}(x,t,v_y) = \mathrm{e}^{-v_y x}\mathrm{I}\{t < T_x \wedge D_x = 0\}\ \mathrm{e}^{T_x \kappa(v_y)}$$

*and*

$$\bar{\Psi}^{(2)}(x,t,v_y) = \mathrm{e}^{-v_y x}\mathrm{I}\{t < T_x \wedge D_x > 0\}\ \mathrm{e}^{-v_y D_x + T_x \kappa(v_y)}$$

*are estimators with bounded relative error of $\bar{\psi}^{(1)}(x,t)$ and $\bar{\psi}^{(2)}(x,t)$, as $t, x \to \infty$, under $\mathsf{P}_{v_y}$.*

### 4.3 Proofs of efficiencies of importance sampling estimators

*Proof of Result 4.1.* We provide detailed justifications for bounded relative error with the estimators of the three probabilities of ruin due to creeping. The justifications for the estimators of the three probabilities of ruin due to jump are similar and can be obtained by minor adaptations of the following arguments.

1. In the infinite time horizon, $\mathsf{E}_r\left[\Psi^{(1)2}(x,r)\right] \leq \mathrm{e}^{-2rx}$ and the Cramer-Lundberg approximation (40) lead to

$$\limsup_{x\to\infty} \frac{\mathsf{E}_r\left[\Psi^{(1)2}(x,r)\right]}{\psi^{(1)2}(x)} \leq \limsup_{x\to\infty} \frac{\mathrm{e}^{-2rx}}{(a_r^{(1)}\mathrm{e}^{-rx})^2} = a_r^{(1)-2} < \infty,$$

   giving bounded relative error in the infinite time horizon.

2. In the short time horizon $y < y_r$, we have

$$\begin{aligned}\bar{\psi}^{(1)}(x,xy) &= \psi^{(1)}(x) - \psi^{(1)}(x,xy) \geq \psi^{(1)}(x) - \psi(x,xy)\\ &= \psi^{(1)}(x) - \mathrm{o}(\psi(x)) = a_r^{(1)}\mathrm{e}^{-rx}\{1+\mathrm{o}\,(1)\}, \quad \text{as } x\to\infty,\end{aligned}$$

   where the second equality is due to (39) and the third one to (40) and (19). So we obtain

$$\begin{aligned}\limsup_{x\to\infty} \frac{\mathsf{E}_r\left[\bar{\Psi}^{(1)2}(x,xy,r)\right]}{\bar{\psi}^{(1)2}(x,xy)} &= \limsup_{x\to\infty} \frac{\mathrm{e}^{-2rx}\mathsf{E}_r[\mathrm{e}^{-2rD_x}; xy < T_x \wedge D_x = 0]}{(a_r^{(1)}\mathrm{e}^{-rx})^2}\\ &\leq \limsup_{x\to\infty} \frac{\mathrm{e}^{-2rx}}{(a_r^{(1)}\mathrm{e}^{-rx})^2}\\ &= a_r^{(1)-2} < \infty,\end{aligned}$$

   which justifies bounded relative error in the short time horizon.

3. In the long time horizon $y > y_r$, Lemma 3.3 yields $\kappa(v_y) < 0$ and thus

$$\begin{aligned}\mathsf{E}_{v_y}\left[\bar{\Psi}^{(1)2}(x,xy,v_y)\right] &\leq \mathrm{e}^{-2v_yx}\mathsf{E}_{v_y}\left[\exp\{2[xy\kappa(v_y) - v_yD_x]\}; T_x > xy \wedge D_x = 0\right]\\ &\leq \mathrm{e}^{-2v_yx}\mathsf{E}_{v_y}\left[\exp\{2[x(v_y - l_y)]\}\right]\\ &= \mathrm{e}^{-2l_yx},\end{aligned}$$

   from (26). From Lemmas 3.3, 3.8 and (40), we obtain

$$\begin{aligned}\bar{\psi}^{(1)}(x,xy) &= \psi^{(1)}(x) - \psi^{(1)}(x,xy)\\ &\geq \psi^{(1)}(x) - \psi(x,xy)\\ &= a_r^{(1)}\mathrm{e}^{-rx}\{1+\mathrm{o}(1)\} - b_y x^{-\frac{1}{2}}\mathrm{e}^{-l_yx}\{1+\mathrm{o}(1)\}\\ &= a_r^{(1)}\mathrm{e}^{-rx}\{1+\mathrm{o}(1)\}, \quad \text{as } x\to\infty.\end{aligned}$$

With these two last results we obtain

$$\limsup_{x\to\infty} \frac{\mathsf{E}_{v_y}\left[\bar{\Psi}^{(1)2}(x,xy,v_y)\right]}{\bar{\psi}^{(1)2}(x,xy)} \leq \limsup_{x\to\infty} \frac{\mathrm{e}^{-2l_y x}}{(a_r^{(1)}\mathrm{e}^{-rx})^2} = a_r^{(1)-2} \lim_{x\to\infty} \mathrm{e}^{-2(l_y-r)x} = 0.$$

Thus we have bounded relative error in the long time horizon.

□

# 5 A numerical illustration

This section provides a short numerical illustration of the accuracy of one of the proposed importance sampling estimators for the following situation. We consider the risk process with diffusion $\{Y_t\}_{t\geq 0}$ given in (1) with the following parameters: the Poisson rate is $\lambda = 1$, the premium rate is $c = 2$ and the variance of the Wiener process is $\sigma^2 = 0.4$. The individual claim amount distribution $\xi$ is hypo-exponential with vector parameter $(\nu_1, \nu_2) = (1, 10)$. Precisely, let $V_1, \ldots, V_n$ be independent random variables having exponential distribution with parameters $\nu_1, \ldots, \nu_n > 0$, respectively, i.e. $\mathsf{P}[V_k > u] = \mathrm{e}^{-\nu_i u}$, $\forall u > 0$ and for $k = 1, \ldots, n$, then the individual claim amount $\Xi = \sum_{k=1}^n V_k$ is hypo-exponentially distributed with parameter $(\nu_1, \ldots, \nu_n)$. The moment generating function is given by $M_\xi(v) = \mathsf{E}[\mathrm{e}^{v\Xi}] = \prod_{k=1}^n (1 - v/\nu_k)^{-1}$, $\forall v < \min\{\nu_1, \ldots, \nu_n\}$, and in our situation $M_\xi(v) = 10/\{(1-v)(10-v)\}$, $\forall v < 1$. The mean claim size is $M'_\xi(0) = 1.1$ and thus the net profit condition holds. We want to assess the accuracy of an importance sampling algorithm of Result 3.1 for $\psi(x, 10)$, for $x$ within 0 and 10. We find $r = 0.4234$ and $\mu_r = \mathsf{E}_r[S_1] = 1.4990$. With $t = 10$, we have that $x < \mu_r t = 14.9898$ corresponds to the long time horizon. We compute directly the relative deviations of importance sampling w.r.t. the saddlepoint approximation of Gatto and Baumgartner (2014), which is another large deviations technique. Similar approximations are thus expected with both methods. Importance sampling is obtained by generating 40 000 paths for each point plotted in Figure 1, i.e. for various initial capitals inbetween 0 and 10. Each point is a relative error obtained by the absolute difference between importance sampling and saddlepoint approximations, divided by importance sampling approximation. The line plotted in Figure 1 is a smoother of the plotted points based on the median. Most relative deviations between importance sampling and saddlepoint approximation are below 5% and this coincidence between these two different techniques gives good evidence of high accuracy for these two alternative methods. More comparisons can be found in Gatto and Baumgartner (2014). A numerical comparison for the infinite time horizon probability of ruin can be found in Gatto and Mosimann (2012), where it is again shown that importance sampling and saddlepoint approximation have very small relative deviations. The computer programs used for this comparison are written in `R` and can be found at `http://cran.r-project.org/web/packages/finiteruinprob`.

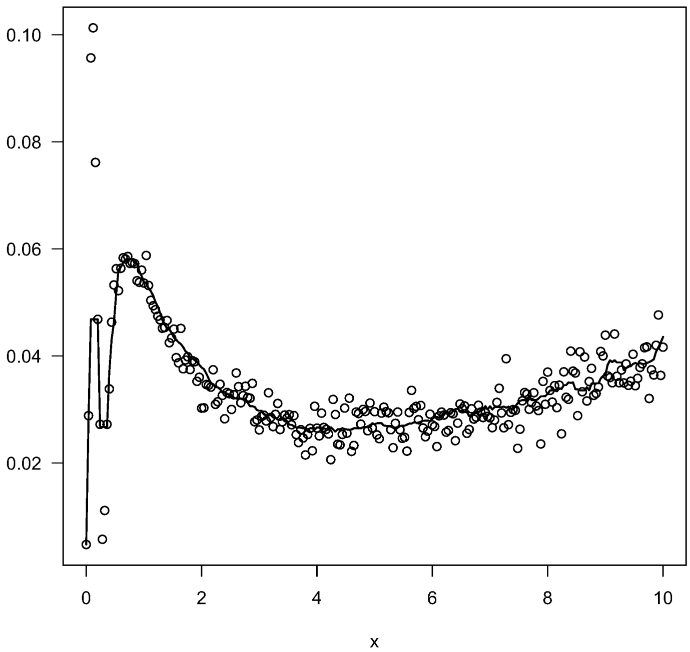


Figure 1: Relative errors between importance sampling and saddlepoint approximation to $\psi(x, 10)$

# 6 References

Applebaum, D. (2004), *Lévy Processes and Stochastic Calculus*, Cambridge.

Asmussen, S. (1985), "Conjugate processes and the simulation of ruin problems", *Stochastic Processes and their Applications*, 20, 213-229.

Asmussen, S. (2000), *Ruin Probabilities*, World Scientific.

Asmussen, S., Glynn, P. W. (2007), *Stochastic Simulation. Algorithms and Analysis*, Springer.

Avram, F., Palmowski, Z., Pistorius, M. R. (2007), "On the optimal dividend problem for a spectrally negative Lévy process", *Annals of Applied Probability*, 17, 156-180.

Bertoin, J. (1996), *Lévy Processes*, Cambridge University Press.

Biffis, E., Morales, M. (2010), "On a generalization of the Gerber-Shiu function to path dependent penalties", *Insurance: Mathematics and Economics*, 46, 92-97.

Carr, P, Geman, H., Madan, D., Yor, M. (2002), "The fine structure of asset returns: an empirical investigation", *Journal of Business*, 75, 305-332.

Collamore, J. F. (2002), "Importance sampling techniques for the multidimensional ruin problem for general markov additive sequences of random vectors", *Annals of Applied*

*Probability*, 12, 382-421.

Collamore, J. F., Diao, G., Vidyashankar, A. N. (2014), "Rare event simulation for processes generated via stochastic fixed point equations", *Annals of Applied Probability*, to appear.

Daniels, H. E. (1954), "Saddlepoint approximations in statistics", *Annals of Mathematical Statistics*, 25, 631-650.

Dybiec, B. (2009), "Lévy noises: double stochastic resonance in a single-well potential", *Physical Review E*, 80, 041111.

Dybiec, B., Gudowska-Nowak, E. (2009), "Lévy stable noise-induced transitions: stochastic resonance, resonant activation and dynamic hysteresis", *Journal of Statistical Mechanics: Theory and Experiment*, P05004.

Dufresne, F., Gerber, H. U. (1991), "Risk theory for the compound Poisson process that is perturbed by diffusion", *Insurance: Mathematics and Economics*, 10, 51-59.

Furrer, H., Michna, Z, Weron, A. (1997), "Stable Lévy motion approximation in collective risk theory", *Insurance: Mathematics and Economics*, 20, 97-114.

Gatto, R., Baumgartner, B. (2014), "Saddlepoint approximations to the probability of ruin in finite time for the compound Poisson risk process perturbed by diffusion", *Methodology and Computing in Applied Probability*, available online.

Gatto, R., Mosimann, M. (2012), "Four approaches to compute the probability of ruin in the compound Poisson risk process with diffusion", *Mathematical and Computer Modelling*, 55, 1169-1185.

Kella, O., Whitt, W. (1992), "Useful martingales for stochastic storage processes with Lévy input", *Journal of Applied Probability*, 29, 396-403.

Klüppelberg, C., Kyprianou, A. E., Maller, R. A. (2004), "Ruin probabilities and overshoots for general Lévy insurance risk processes", *Annals of Applied Probabability*, 14, 1766-1801.

Kyprianou, A. E. (2006), *Introductory Lectures on Fluctuations of Lévy Processes with Applications*, Springer.

Kyprianou, A. E., Palmowski, Z. (2007), *Distributional study of de Finettis dividend problem for a general Lévy insurance risk process*, *Journal of Applied Probability*, 44, 428-443.

Lugannani, R., Rice, S. (1980), "Saddle point approximation for the distribution of the sum of independent random variables", *Advances in Applied Probability*, 12, 475-490.

Morales, M. (2004), "Risk theory with the generalized inverse Gaussian Lévy process", *Astin Bulletin*, 34, 361377.

Palmowski, Z., Pistorius, M. (2009), "Cramer asymptotics for finite time first passage probabilities of general Lévy processes", *Statistics & Probability Letters*, 79, 1752-1758.

Perc, M. (2007a), "Flights towards defection in economic transactions", *Economics Letters*, 97, 5863.

Perc, M. (2007b), "Transition from Gaussian to Lévy distributions of stochastic payoff variations in the spatial prisoners dilemma game", *Physical Review E*, 75, 022101.

Sato, K. (1999), *Lévy Processes and Infinite Divisibility*, Cambridge University Press.

Siegmund, D. (1976), "Importance sampling in the Monte Carlo study of sequential tests", *Annals of Statistics*, 4, 673-684.

Yang, H., Zhang, L. (2001), "Spectrally negative Lévy processes with applications in risk theory", *Advances in Applied Probability*, 33, 281-291.