Exponential Family Techniques
for the Lognormal Left Tail

Søren Asmussen, Jens Ledet Jensen
and Leonardo Rojas-Nandayapa

14

Exponential Family Techniques
for the Lognormal Left Tail

Søren Asmussen[1], Jens Ledet Jensen[1] and Leonardo Rojas-Nandayapa[2]

[1]Department of Mathematics, Aarhus University
[2]School of Mathematics and Physics, University of Queensland

Abstract

Let X be lognormal(μ, σ^2) with density $f(x)$, let $\theta > 0$ and define $L(\theta) = \mathbb{E}e^{-\theta X}$. We study properties of the exponentially tilted density (Esscher transform) $f_\theta(x) = e^{-\theta x} f(x)/L(\theta)$, in particular its moments, its asymptotic form as $\theta \to \infty$ and asymptotics for the saddlepoint $\theta(x)$ determined by $\mathbb{E}[X e^{-\theta X}]/L(\theta) = x$. The asymptotic formulas involve the Lambert W function. The established relations are used to provide two different numerical methods for evaluating the left tail probability of lognormal sum $S_n = X_1 + \cdots + X_n$: a saddlepoint approximation and an exponential twisting importance sampling estimator. For the latter we demonstrate logarithmic efficiency. Numerical examples for the cdf $F_n(x)$ and the pdf $f_n(x)$ of S_n are given in a range of values of σ^2, n, x motivated from portfolio Value-at-Risk calculations.

Keywords: Lognormal distribution, Cramér function, Esscher transform, exponential change of measure, Laplace transform, Laplace method, saddlepoint approximation, Lambert W function, rare event simulation, importance sampling, VaR.

MSC: 60E05, 60E10, 90-04

1 Introduction

The lognormal distribution arises in a wide variety of disciplines such as engineering, economics, insurance or finance, and is often employed in modeling across the sciences (Aitchison and Brown, 1957; Crow and Shimizu, 1988; Dufresne, 2009; Johnson et al., 1994; Limpert et al., 2001). In consequence, it is natural that sums of lognormals come up in a number of contexts. For instance, a basic example in finance is the Black-Scholes model, which asserts that security prices can be modeled as independent lognormals (equivalently, the logprices are independent normally distributed). This implies that the value of a portfolio with n securities can be conveniently modeled as a sum of lognormals. Another example occurs in the valuation of arithmetic Asian options where the payoff depends on the finite sum of correlated lognormals

(Dufresne, 2004; Milevsky and Posner, 1998). In insurance, individual claim sizes are often modeled as independent lognormals, so the total claim amount in a certain period is a random sum of lognormals (Thorin and Wikstad, 1977). A further example occurs in telecommunications, where the inverse of the signal-to-noise ratio (a measure of performance in wireless systems) can be modeled as a sum of i.i.d. lognormals (Gubner, 2006).

However, the distribution of a sum of n lognormals S_n is not available in explicit form and its numerical approximation is considered to be a challenging problem. In consequence, a number of methods for its evaluation has been developed across several decades, but these can rarely deliver arbitrary precisions in the whole support of the distribution, particularly in the tails. The later case is of key relevance in certain applications which often require to evaluate tail probabilities at very high precisions. For instance, the Value-at-Risk (VaR) is an important measure of market risk defined as an appropriate $(1 - \alpha)$-quantile of the loss distribution, and the standard financial treatise Basel II (2004) asks for calculations of the VaR for so small values as $\alpha = 0.03\%$.

When considering lognormals sums, the literature has sofar concentrated on the right tail (with the exception of the recent paper Gulisashvili and Tankov, 2014). In this paper, our object of study is rather the left tail and certain mathematical problems that naturally come up in this context. To be precise, let Y_i be normal(μ_i, σ_i^2) (we don't at the moment specify the dependence structure), let $X_i = e^{Y_i}$ and $S_n = X_1 + \cdots + X_n$. We then want to compute $\mathbb{P}(S_n \leq z)$ in situations where this probability is small.

An obvious motivation for this problem comes from the VaR problem. Here S_n may represent the future value of the portfolio. If Π is the present value, $\Pi - S_n$ is then the loss, and so calculation of α-quantiles are equivalent to left tail calculations for S_n. A further example occurs in the wireless systems setting, where an outage occurs when the signal-to-noise ratio exceeds a large threshold. The outage probability is therefore obviously related to the left tail probability of a lognormal sum.

The problem of approximating the distribution of a sum of i.i.d. lognormals has as mentioned a long history. The classical approach is to approximate the distribution of a sum of i.i.d. lognormals with another lognormal distribution. This goes back at least to Fenton (1960) and it is nowadays known as the *Fenton-Wilkinson method*; according to Marlow (1967) this approximation was already used by Wilkinson since 1934. However, the Fenton-Wilkinson method, being a central limit type result, can deliver rather inaccurate approximations of the distribution of the lognormal sum when the number of summand is rather small or the dispersion parameter is too high—in particular in the tail regions. Another topic which has been much studied recently is approximations and simulation algorithms for right tail probabilities $\mathbb{P}(S_n \geq y)$ under heavy-tailed assumptions and allowing for dependence, see in particular Asmussen et al. (2011); Asmussen and Rojas-Nandayapa (2008); Blanchet and Rojas-Nandayapa (2011); Foss and Richards (2010); Mitra and Resnick (2009). For further literature surveys, see Gulisashvili and Tankov (2014).

Our approach is to use the saddlepoint approximations and a closely related simulation algorithm based on the same exponential change of measure. This requires

i.i.d. assumptions, in particular $\mu_i \equiv \mu$, $\sigma_i^2 \equiv \sigma^2$. Since μ is just a scaling factor, we will assume $\mu = 0$. The saddlepoint approximation occurs in various (closely related) forms, but all involve the function $\kappa(\theta) = \log L(\theta)$, where

$$L(\theta) = \mathbb{E}e^{-\theta X_i} = \int_0^\infty e^{-\theta x} f(x)\, dx \ \text{ with } \ f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\log^2 x / 2\sigma^2},$$

and its two first derivatives $\kappa'(\theta)$ and $\kappa''(\theta)$ (note that since the right tail of the lognormal distribution is heavy, these quantities are only defined for $\theta \geq 0$). Define the exponentially tilted density $f_\theta(x)$ (Esscher transform) by

$$f_\theta(x) = e^{-\theta x - \kappa(\theta)} f(x), \quad x > 0, \tag{1.1}$$

and let its corresponding cumulative distribution function be F_θ with expectation operator $\mathbb{E}_\theta$. Then

$$\kappa'(\theta) = -\mathbb{E}_\theta X, \quad \kappa''(\theta) = \text{Var}_\theta X \tag{1.2}$$

and one can connect the given distribution of S_n (corresponding to $\theta = 0$) to the $\mathbb{P}_\theta$-distribution by means of the likelihood ratio identity

$$\mathbb{P}(S_n \in A) = \mathbb{E}_\theta\big[\exp\{\theta S_n + n\kappa(\theta)\}; \, S_n \in A\big].$$

The details of the saddlepoint approximation involve writing $z = nx$, defining the *saddlepoint* $\theta(x)$ as the solution of the equation

$$-\kappa'\big(\theta(x)\big) = \mathbb{E}_{\theta(x)}[X] = x, \tag{1.3}$$

and taking $\theta = \theta(x)$. This choice of θ means that $\mathbb{E}_\theta S_n = z$ so that the $\mathbb{P}_\theta$-distribution is centered around z and central limit expansions apply. For a short exposition of the implementation of this program in its simplest form, see p. 355, Asmussen (2003).

The application of saddlepoint approximations to the lognormal left tail seems to have appeared for the first time in the third author's 2008 Dissertation (Rojas-Nandayapa, 2008), but in a more incomplete and preliminary form than the one presented here. A first difficulty is that $\kappa(\theta)$ is not explicitly available for the lognormal distribution. However, approximations with error rates were recently given in the companion paper Asmussen et al. (2014b). The result is in terms of the Lambert W function $\mathcal{W}(a)$ (Corless et al., 1996), defined as the unique solution of $\mathcal{W}(a)e^{\mathcal{W}(a)} = a$ for $a > 0$. The expression for $\kappa(\theta)$ in Asmussen et al. (2014b) is

$$L(\theta) = \frac{\exp\left\{-\dfrac{\mathcal{W}(\theta\sigma^2)^2 + 2\,\mathcal{W}(\theta\sigma^2)}{2\sigma^2}\right\}}{\sqrt{1 + \mathcal{W}(\theta\sigma^2)}} \int_{-\infty}^\infty g_0(z)\, dz, \tag{1.4}$$

where g_0 is a certain function such that $\int g_0(z) dz$ is close to 1 (see Section 2 for more detail; we also give an extension to expectations of the form $\mathbb{E}[X^k e^{-\theta X}]$ there). Note that the Lambert W function is convenient for numerical computations since it is implemented in many software packages.

The paper is organized as follows. In Section 2, we study the exponential family $(F_\theta)_{\theta \geq 0}$. We give an approximation to the derivatives of the Laplace transform, an

approximation to the saddlepoint $\theta(x)$, and discuss various approximations to the tilted density f_θ. The first important application of our results, namely the saddlepoint approximation for $\mathbb{P}(S_n \leq z)$, is given in Section 3. The second is a Monte Carlo estimator for $\mathbb{P}(S_n \leq z)$ given in Section 4.2. It follows a classical route (VI.2, Asmussen and Glynn, 2007) by attempting importance sampling with importance distribution $F_{\theta(x)}$, but the implementation faces the difficulty that neither $\theta(x)$ nor $\kappa(\theta(x))$ are explicitly known. The importance sampling algorithm requires simulation from F_θ, and we suggest an acceptance-rejection (A-R) for this with a Gamma proposal. The Appendix contains a proof that the importance sampling proposed in Section 4.2 has a certain asymptotical efficiency property.

2 The exponential family generated by the lognormal distribution

We let F be the cumulative distribution function of X and adopt the notation $X \sim \mathrm{LN}(0, \sigma^2)$. For convenience, we write f_n and F_n for the pdf and cdf of S_n, respectively.

The exponential tilting scheme in the Introduction is often also referred to as *Esscher transformation*. Note that since $\kappa(\theta)$ is well-defined for all $\theta > 0$, the saddlepoint $\theta(x)$ exists for all $0 < x \leq \mathbb{E}X$ (the relevant case for our left tail problem) and large deviations results can be used. The latter are based on the *Legendre-Fenchel transform* defined as the convex conjugate $\kappa^*(x) = \kappa(\theta(x)) + x\theta(x)$.

We first consider ways of evaluating and approximating derivatives of the Laplace transform given through

$$L_k(\theta) = \mathbb{E}[X^k \mathrm{e}^{-\theta X}] = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-h_k(y)} \mathrm{d}y, \quad \text{with } h_k(y) = -ky + \theta \mathrm{e}^y + \frac{y^2}{2\sigma^2}. \quad (2.1)$$

Define the following quantities:

$$w_k(\theta) = \mathcal{W}(\theta \sigma^2 \mathrm{e}^{k\sigma^2}), \qquad \sigma_k(\theta)^2 = \frac{\sigma^2}{1 + w_k(\theta)},$$

$$L_a(k, \theta) = \frac{\sigma_k(\theta)}{\sigma} \exp\left\{ -\frac{1}{2\sigma^2} w_k(\theta)^2 - \frac{1}{\sigma^2} w_k(\theta) + \tfrac{1}{2} k^2 \sigma^2 \right\}, \qquad (2.2)$$

$$I_k(\theta) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{w_k(\theta)}{\sigma^2} \left(\mathrm{e}^{z\sigma_k(\theta)} - 1 - z\sigma_k(\theta) \right) - \frac{\sigma_k(\theta)^2}{2\sigma^2} z^2 \right\} \mathrm{d}z, \quad (2.3)$$

$$J_k(\theta) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{w_0(\theta)}{\sigma^2} \left(\mathrm{e}^{z\sigma_0(\theta)} - 1 - z\sigma_0(\theta) \right) - \frac{\sigma_0(\theta)^2}{2\sigma^2} z^2 \right.$$
$$\left. + k\sigma_0(\theta)z - \tfrac{1}{2}\sigma_0(\theta)^2 k^2 \right\} \mathrm{d}z.$$

The following proposition extends Proposition 2.1 of Asmussen et al. (2014b). To understand the orders of the different terms one should keep in mind that $w_k(\theta)$ is asymptotically of order $\log(\theta)$ for $\theta \to \infty$. Also we use the fact that $w_k(0) = 0$.

Proposition 2.1. *Let $X \sim \mathrm{LN}\left(0, \sigma^2\right)$, $k \in \mathbb{N}$ and $\theta \geq 0$. Then*

$$L_k(\theta) = L_a(k, \theta)I_k(\theta) \text{ and } L_k(\theta) = L_a(0, \theta) \exp\left\{-kw_0(\theta) + \tfrac{1}{2}\sigma_0(\theta)^2 k^2\right\} J_k(\theta). \quad (2.4)$$

Moreover,

$$I_k(\theta) = \begin{cases} 1 & \theta = 0, \\ 1 - \dfrac{3w_k(\theta)\sigma_k(\theta^4}{24\sigma^2} + \dfrac{5w_k(\theta)^2\sigma_k(\theta)^6}{24\sigma^4} + \mathrm{O}\left(\sigma_k(\theta)^4\right) & \theta \to \infty, \end{cases}$$

and

$$J_k(\theta) = \begin{cases} 1 & \theta = 0, \\ 1 - \dfrac{(1 + 4k)w_0(\theta)\sigma_0(\theta^4}{8\sigma^2} + \dfrac{5w_0(\theta)^2\sigma_0(\theta)^6}{24\sigma^4} + \mathrm{O}\left(\sigma_0(\theta)^4\right) & \theta \to \infty. \end{cases}$$

The proof of Proposition 2.1 is based on the Laplace approximation as the proof of Proposition 2.1 in the companion paper Asmussen et al. (2014b). We give here only a brief sketch of the proof. For the first result involving $L_a(k, \theta)$, the function $h_k(y)$ is expanded around its minimizer $y_k(\theta)$ given as the solution to $\theta e^y + y/\sigma^2 - k = 0$, that is, $y_k(\theta) = k\sigma^2 - w_k(\theta)$. Then the exponential part of $L_a(k, \theta)$ is simply $-h_k\left(y_k(\theta)\right)$, and the exponential part of the integrand in $I_k(\theta)$ is $-\left\{h_k\left(y_k(\theta) + \sigma_k(\theta)z\right) - h_k\left(y_k(\theta)\right)\right\}$. Expansion of the latter gives

$$-\tfrac{1}{2}z^2 - \frac{w_k(\theta)\sigma_k(\theta)^3}{6\sigma^2}z^3 - \frac{w_k(\theta)\sigma_k(\theta)^4}{24\sigma^2}z^4 + \mathrm{O}\left(\sigma_k(\theta)^3|z|^5\right),$$

and expanding the exponential of the last three terms gives the result in the proposition for $I_k(\theta)$. For the alternative formula with $J_k(\theta)$ we expand $h_k(y)$ around $y_0(\theta)$. Then the exponential part of $L_a(0, \theta)$ together with $-kw_0(\theta)$ is simply $-h(y_0(\theta))$, and the exponential part of the integrand in $J_k(\theta)$ is $-\left\{h_k\left(y_0(\theta) + \sigma_0(\theta)z\right) - h_k\left(y_0(\theta)\right)\right\} - \sigma_0(\theta)^2 k^2/2$. Expanding the latter we get

$$-\tfrac{1}{2}\left(z - k\sigma_0(\theta)\right)^2 - \frac{w_0(\theta)\sigma_0(\theta)^3}{6\sigma^2}z^3 - \frac{w_0(\theta)\sigma_0(\theta)^4}{24\sigma^2}z^4 + \mathrm{O}\left(\sigma_0(\theta)^3|z|^5\right),$$

which leads to the result in the proposition.

The results of Proposition 2.1 immediately lead to an approximation to the mean and variance of the exponentially tilted measure. These are denoted by $\mathbb{E}_\theta$ and Var_θ, respectively. Note that, although the results below are for $\theta \to \infty$, the approximations are actually exact for $\theta = 0$ as well.

Corollary 2.2. *Let $X \sim \mathrm{LN}\left(0, \sigma^2\right)$. Then as $\theta \to \infty$*

$$\mathbb{E}_\theta[X] = \exp\left\{-w_0(\theta) + \tfrac{1}{2}\sigma_0(\theta)^2\right\}\left(1 + \mathrm{O}(\sigma_0(\theta)^2)\right), \quad (2.5)$$

$$\mathrm{Var}_\theta[X] = \exp\left\{-2w_0(\theta) + \sigma_0(\theta)^2\right\}\left(e^{\sigma_0(\theta)^2} - 1\right)\left(1 + \mathrm{O}(\sigma_0(\theta)^2)\right). \quad (2.6)$$

Proof. Simply use $\mathbb{E}_\theta[X] = L_1(\theta)/L_0(\theta)$ and $\mathrm{Var}_\theta[X] = L_2(\theta)/L_0(\theta) - (L_1(\theta)/L_0(\theta))^2$ together with the second part of (2.4). $\qquad\square$

5

Interestingly, one could at least at the heuristic level identify the tilted measure F_θ as approximate lognormal with parameters $-w_0(\theta)$ and $\sigma_0(\theta)^2$, $F_\theta \approx \mathrm{LN}(-w_0(\theta),$ $\sigma_0(\theta)^2)$. For $\theta = 0$ we have the correct lognormal distribution and, as Corollary 2.2 shows, the lognormal approximation produces the correct asymptotic mean and variance as $\theta \to \infty$. The approximation is illustrated in Figure 1 with $\sigma = 0.25$, where the true log–density f_θ and the approximating lognormal density are displayed for $\theta = 10, 25, 100$. It is notable how little even such a large values as $\theta = 100$ shifts the distribution towards the origin, which can be explained by the lognormal density decaying only slowly to 0 as $x \downarrow 0$. As can be seen from the figure the approximation is very good, more so as $\theta \to \infty$. We next show that, as $\theta \to \infty$, the limiting centered and scaled tilted density f_θ is a standard normal density. It follows from this that the lognormal approximation becomes exact as $\theta \to \infty$. We will use the result below in the following sections.

Proposition 2.3. *Write the tilted density $f_\theta(x)$ as $\exp\{-m(x) - \kappa(\theta)\}/\sqrt{2\pi\sigma^2}$ with $m(x) = \log(x) + (\log(x))^2/(2\sigma^2) + \theta x$. Furthermore, let $w = w_0(\theta)$ and define $m_0(u) = m\big(e^{-w}(1 + \sigma u/\sqrt{w})\big) - m(e^{-w})$. Then, as $\theta \to \infty$,*

$$m_0(u) = \tfrac{1}{2}u^2 + \mathrm{O}\big((|u| + |u|^3)\sigma/\sqrt{w}\big) \quad \textit{for} \quad \frac{|u|^3}{\sqrt{w}} \le 1,$$

and for θ sufficiently large

$$m_0'(u) > \frac{1}{2}\Big(\frac{\sqrt{w}}{\sigma}\Big)^{1/6} \qquad \textit{for} \quad u > \Big(\frac{\sqrt{w}}{\sigma}\Big)^{1/6},$$

$$m_0'(u) < -\frac{1}{2}\Big(\frac{\sqrt{w}}{\sigma}\Big)^{1/6} \qquad \textit{for} \quad u < -\Big(\frac{\sqrt{w}}{\sigma}\Big)^{1/6}.$$

The above relations imply that the centered and scaled density f_θ converges to a standard normal density, and moments of f_θ converge as well.

Proof. We first note that the lognormal density $f(x)$ is logconcave for $x < e^{1-\sigma^2}$ since

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} \log(f(x)) = -\frac{1}{x^2\sigma^2}\big(-\log(x) + \sigma^2 - 1\big) < 0 \quad \textit{for} \quad x < e^{1-\sigma^2}. \tag{2.7}$$

We rewrite $m_0(u)$ as

$$m_0(u) = \log\Big(1 + \frac{\sigma}{\sqrt{w}}u\Big) + \frac{1}{2\sigma^2}\Big\{\Big[-w + \log\Big(1 + \frac{\sigma}{\sqrt{w}}u\Big)\Big]^2 - w^2\Big\} + \frac{\sqrt{w}}{\sigma}u.$$

Taylor expanding $\log(1 + \sigma u/\sqrt{w})$ we obtain the first result of the proposition. Next, we find the derivative of $m_0(u)$:

$$m_0'(u) = \frac{\sigma/\sqrt{w}}{1 + \sigma u/\sqrt{w}} + \frac{\sqrt{w}}{\sigma} + \frac{1}{\sigma^2}\big[-w + \log(1 + \sigma u/\sqrt{w})\big]\frac{\sigma/\sqrt{w}}{1 + \sigma u/\sqrt{w}}.$$

For $u > (\sqrt{w}/\sigma)^{1/6}$ we get the bound

$$m_0' > \frac{\sqrt{w}}{\sigma}\Big(1 - \frac{1}{1 + \sigma u/\sqrt{w}}\Big) = \frac{(\sqrt{w}/\sigma)^{1/6}}{1 + (\sqrt{w}/\sigma)^{-5/6}} > \tfrac{1}{2}(\sqrt{w}/\sigma)^{1/6},$$

as long as $\sigma/\sqrt{w} < 1$, which is true for $\theta \to \infty$. For $u < -(\sqrt{w}/\sigma)^{1/6}$ we have from the logconcavity that $m_0'(u) < m_0'\big(-(\sqrt{w}/\sigma)^{1/6}\big)$. For the latter we find

$$m_0'\big(-(\sqrt{w}/\sigma)^{1/6}\big) \sim -\Big(\frac{\sqrt{w}}{\sigma}\Big)^{1/6} \quad \text{as} \quad \theta \to \infty.$$
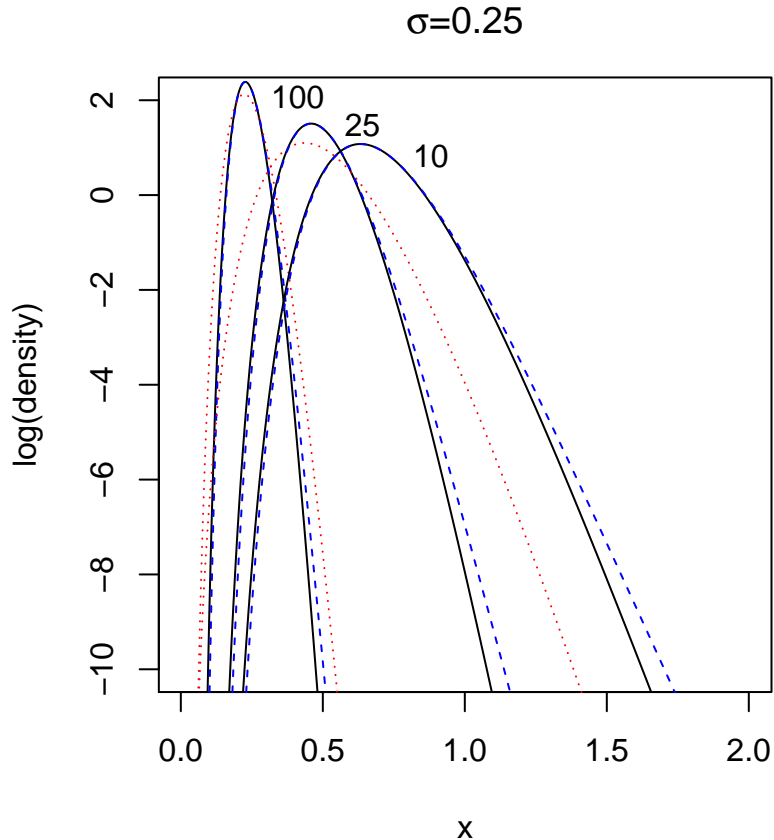
$\square$

Figure 1: Log of tilted density f_θ (full drawn) and lognormal approximation (dashed). The three cases are $\theta = 10, 25, 100$ all with $\sigma = 0.25$. Included is also for $\theta = 25, 100$ the Gamma approximation considered in Section 4.1 (dotted curve).

2.1 The saddlepoint $\theta(x)$

Corollary 2.2 in a natural way leads to an approximation to the saddlepoint $\theta(x)$, the latter being the solution of the equation $L_1(\theta)/L_0(\theta) = x$. We simply let the approximation $\tilde{\theta}(x)$ be the solution of

$$\exp\big\{-w_0(\theta) + \tfrac{1}{2}\sigma_0(\theta)^2\big\} = x.$$

This gives the equation $-w_0(\theta) + \tfrac{1}{2}\sigma^2/(1+w_0(\theta)) = \log(x)$, which leads to a quadratic equation in $w_0(\theta)$. Since $w_0(\theta) \geq 0$, we find

$$w_0(\theta) = \gamma(x) \quad \text{with} \quad \gamma(x) = \tfrac{1}{2}\big(-1 - \log x + \sqrt{(1 - \log x)^2 + 2\sigma^2}\big).$$

7

Then from the definition of $w_0(\theta)$ we obtain

$$\tilde{\theta}(x) = \frac{1}{\sigma^2}\gamma(x)e^{\gamma(x)}. \tag{2.8}$$

The following proposition states the quality of this approximation.

Proposition 2.4. *For $x \to 0$ we have $\tilde{\theta}(x) \sim (-\log x)/(x\sigma^2)$ and*

$$\mathbb{E}_{\tilde{\theta}(x)}[X] = x\Big(1 + O\Big(\frac{1}{|\log(x)|}\Big)\Big),$$

$$\theta(x) = \tilde{\theta}(x)\Big(1 + O\Big(\frac{1}{|\log(x)|}\Big)\Big).$$

Proof. Below we write $\bar{\theta}$ for $\theta(x)$ and $\tilde{\theta}$ for $\tilde{\theta}(x)$. We first note that $\gamma(x) \sim -\log(x)$ for $x \to 0$, which gives that $\tilde{\theta} = \gamma(x)e^{\gamma(x)}/\sigma^2 \sim (-\log x)/(x\sigma^2)$. From the definition of $\tilde{\theta}$ and (2.5) we have $\mathbb{E}_{\tilde{\theta}}[X] = x(1 + O(1/w_0(\tilde{\theta}))$, and the first result follows from $w_0(\tilde{\theta}) = \gamma(x) \sim -\log(x)$. Using that $\mathbb{E}_{\bar{\theta}}[X] = x$ we rewrite the first result as $\mathbb{E}_{\bar{\theta}}[X]/\mathbb{E}_{\tilde{\theta}}[X] = 1 + O(1/w_0(\tilde{\theta}))$. From a Taylor expansion of $\mathbb{E}_\theta[X]$ around $\tilde{\theta}$ we get to first order

$$1 + O\Big(\frac{1}{w_0(\tilde{\theta})}\Big) \approx 1 + \frac{\mathrm{Var}_{\tilde{\theta}}[X]}{\mathbb{E}_{\tilde{\theta}}[X]}\tilde{\theta}\Big(1 - \frac{\bar{\theta}}{\tilde{\theta}}\Big) \approx 1 + \frac{\frac{\sigma^2}{w_0(\tilde{\theta})}e^{-2w_0(\tilde{\theta})}}{e^{-w_0(\tilde{\theta})}}\frac{w_0(\tilde{\theta})}{\sigma^2}e^{w_0(\tilde{\theta})}\Big(1 - \frac{\bar{\theta}}{\tilde{\theta}}\Big)$$

$$\approx 1 + \Big(1 - \frac{\bar{\theta}}{\tilde{\theta}}\Big),$$

from which we conclude that $1 - \bar{\theta}/\tilde{\theta} = O(1/w_0(\tilde{\theta}))$ or $\bar{\theta} = \tilde{\theta}(1 + O(1/|\log(x)|))$. $\quad\square$

In Sections 3 and 4.2 we will employ the results of this section to construct a saddlepoint approximation and a Monte Carlo estimator of the left tail probability of a sum of lognormal random variables. In particular, the asymptotic results derived above will be useful to show that when the approximation $\tilde{\theta}(x)$ used as the tilting parameter of an exponential change of measure the Monte Carlo estimator remains asymptotically efficient as $x \to 0$.

3 Saddlepoint approximation in the left tail of a lognormal sum

Daniels' saddlepoint method produces an approximation of the density function of a sum of i.i.d. random variables which is valid asymptotically on the number of summands. The first and second order approximations are embodied in the formula

$$f_n(nx) \approx \sqrt{\frac{1}{2\pi n\kappa''(\theta(x))}}\exp\big\{n\kappa^*(x)\big\}\big(1 + \tfrac{1}{n}[\zeta_4(\theta(x))/8 - 5\zeta_3(\theta(x))^2/24]\big),$$

where $\kappa^*(x) = \kappa(\theta(x)) + x\theta(x)$ is the convex conjugate of $\kappa(x)$ and

$$\zeta_k(\theta) = \frac{\kappa^{(k)}(\theta)}{\kappa''(\theta)^{k/2}},$$

8

is the standardized cumulant.

The corresponding saddlepoint approximation for the cumulative distribution function is given by Jensen (1995)

$$F_n(nx) = \frac{1}{\lambda_n(x)} \exp\left\{n\kappa^*(x)\right\}\left\{B_0\big(\lambda_n(x)\big)\right.$$
$$\left. + \frac{\zeta_3\big(\theta(x)\big)}{6\sqrt{n}}B_3\big(\lambda_n(x)\big) + \frac{\zeta_4\big(\theta(x)\big)}{24n}B_4\big(\lambda_n(x)\big) + \frac{\zeta_3\big(\theta(x)\big)^2}{72n}B_6\big(\lambda_n(x)\big)\right\},$$

where $\lambda_n(x) = \theta(x)\sqrt{n\kappa''(\theta(x))}$ and

$$B_0(\lambda) = \lambda e^{\lambda^2/2}\Phi(-\lambda),$$
$$B_3(\lambda) = -\left\{\lambda^3 B_0(\lambda) - (\lambda^3 - \lambda)/\sqrt{2\pi}\right\},$$
$$B_4(\lambda) = \lambda^4 B_0(\lambda) - (\lambda^4 - \lambda^2)/\sqrt{2\pi},$$
$$B_6(\lambda) = \lambda^6 B_0(\lambda) - (\lambda^6 - \lambda^4 + 3\lambda^2)/\sqrt{2\pi}.$$

General results for the saddlepoint approximation state that for a fixed x the relative error is $O(1/n)$ for the first order approximation and $O(1/n^2)$ for the second order approximation. More can be said, however, for the case of a lognormal sum. It is simple to see that the density $f(x)$ is logconcave for $x < e^{1-\sigma^2}$, see (2.7), and according to Jensen (1995), we therefore have that the saddlepoint approximations have the stated relative errors uniformly for x in a region around zero. Furthermore, the convergence of the tilted density as $\theta \to \infty$ outlined in Proposition 2.3 implies that the saddlepoint approximation become exact in the same limit.

To calculate the saddlepoint approximation we need to find the Laplace transform and its derivatives numerically. We want to implement the integration in such a way that the relative accuracy of the integration is of the same order irrespective of the argument θ. For $k = 0, 1, 2, 3, 4$ we want to evaluate the integral $L_k(\theta)$ from (2.1). This leads to the integral $I_k(\theta)$ from (2.3). Instead of the scale $\sigma_k(\theta)$ chosen there, we consider another scale τ and the integral

$$I_k(\theta) = \frac{1}{\sqrt{2\pi}}\frac{\tau}{\sigma_k(\theta)}\int_{-\infty}^{\infty}\exp\{-\tilde{h}(z)\}\mathrm{d}z, \quad \tilde{h}(z) = \frac{w_k(\theta)}{\sigma^2}\big(e^{\tau z} - 1 - \tau z\big) + \frac{\tau^2}{2\sigma^2}z^2.$$

Since $\tilde{h}''(z) = \tau^2\big(w_k(\theta)e^{\tau z} + 1\big)/\sigma^2 > 0$ we see that $\tilde{h}$ is convex. Choosing the scale τ such that $2\tilde{h}(-\tau) = 1$ we obtain that $2\tilde{h}(z)$ is a convex function bounded between 0 and 1 for $-1 < z < 0$, is above $-z$ for $z < -1$ and with $\tilde{h}(z) \geq \tilde{h}(-z)$ for $z > 0$. In this way the precision of the numerical integration of $\exp\{-\tilde{h}(z)\}$ will be of the same order irrespective of the value of w and σ^2. In practice we can take τ as

$$\tau = \begin{cases} \sigma_k(\theta), & \sigma_k(\theta) \leq c_0, \\ \sqrt{w_k(\theta)^2 + 2w_k(\theta) + \sigma^2} - w_k(\theta) & \sigma_k(\theta) > c_0, \end{cases}$$

where c_0 is an arbitrary constant. Unless σ^2 is large we can use $\tau = \sigma_k(\theta)$ for all θ.

4 Simulation

4.1 Random variate generation from F_θ

We first consider the problem of generating a random variable from the tilted density f_θ from (1.1).

The obvious naive choice is acceptance-rejection (A-R; Asmussen and Glynn, 2007, II.2), simulating Z from f and rejecting with probaility $e^{-\theta Z}$. This choice produces a very simple algorithm for generating from f_θ and the method is exact even when we do not have an explicit expression for $\kappa(\theta)$.

Algorithm 4.1.

1. Simulate $U \sim U(0,1)$, $Z \sim LN(0, \sigma^2)$.
2. If $U > e^{-\theta Z}$ repeat. Else, return Z.

Ideally, we would like to have a rejection probability p close to 1, but in our case $p = e^{\kappa(\theta)}$, so as the value of θ increases, the probability of acceptance diminishes, and hence the expected number of rejection steps goes to infinity. In consequence, this estimator is very inefficient for large values of θ.

As noted in Proposition 2.3 if X_θ is a random variable with density f_θ, the variable $U = (X_\theta - e^{-w})\sqrt{w}e^w/\sigma$, $w = w_0(\theta)$, has a standard normal distribution in the limit $\theta \to \infty$. However, the limiting normal distribution can not be used as proposal for an A-R algorithm because the right tail is lighter than that of X_θ. Similarly, the lognormal approximation is not applicable because the left tail of $\big(X_\theta + w_0(\theta)\big)/\sigma_0(\theta)$ is lighter.

What we know, however, is that the right tail of X_θ is trivially lighter than $e^{-\theta x}$. This points to the possibility of using a gamma proposal $Z \sim Gamma(\lambda, \theta)$. Rewriting the tilted density f_θ as

$$
\frac{e^{-\kappa(\theta)}}{\sqrt{2\pi\sigma^2}} x^{\lambda-1} \exp\Big\{-\theta x - \frac{1}{2\sigma^2}(\log x)^2 - \lambda \log x\Big\}
$$
$$
= \frac{e^{-\kappa(\theta)+\frac{1}{2}\lambda^2\sigma^2}}{\sqrt{2\pi\sigma^2}} x^{\lambda-1} \exp\Big\{-\theta x - \frac{1}{2\sigma^2}(\lambda\sigma^2 + \log x)^2\Big\}. \tag{4.1}
$$

ewe choose λ such that $\lambda\sigma^2 + \log(\mathbb{E}[Z]) = 0$. Solving for λ we obtain $0 = \lambda\sigma^2 + \log(\lambda/\theta)$ or $\lambda = w_0(\theta)/\sigma^2$. This gives the following A-R algorithm.

Algorithm 4.2.

1. Simulate $U \sim U(0,1)$ and $Z \sim Gamma(w_0(\theta)/\sigma^2 + 1, \theta)$.
2. If $U > \exp\big\{-\big(w_0(\theta) + \log Z\big)/2\sigma^2\big\}$ repeat. Else, return $X_\theta = Z$.

Let us center and scale Z using $\mathbb{E}[Z] = e^{-w}$ and $Var[Z] = \sigma^2 e^{-2w}/w$, $w = w_0(\theta)$. From this we write $Z = e^{-w}(1 + \sigma V/\sqrt{w})$. Then as $\theta \to \infty$ we have that V becomes standard normally distributed, and the acceptance probability becomes

$$
\mathbb{E}\big[\exp\big\{-\big(w_0(\theta) + \log Z\big)/2\sigma^2\big\}\big] = \mathbb{E}\big[\exp\big\{-\big(\log(1 + \sigma V/\sqrt{w}\big)^2/2\sigma^2\big\}\big] \to 1.
$$

Actually, from (4.1) we can give an explicit expression for the acceptance probability as follows:

$$\mathbb{E}[e^{-\frac{1}{2\sigma^2}(w+\log Z)^2}] = \int_{-\infty}^{\infty} \frac{\theta^{w/\sigma^2}}{\Gamma(w/\sigma^2)} x^{\lambda-1} \exp\left\{-\theta x - \frac{1}{2\sigma^2}(\lambda\sigma^2 + \log x)^2\right\} \mathrm{d}x$$

$$= \frac{\theta^{w/\sigma^2}\sqrt{2\pi\sigma^2}}{\Gamma(w/\sigma^2)e^{-\kappa(\theta)+\frac{1}{2}\lambda^2\sigma^2}}.$$

Figure 2 shows the acceptance probability for both Algorithm 4.1 and Algorithm 4.2 for various values of σ. In the figure the acceptance probabilities for Algorithm 4.1 are almost the same for the three values of σ considered. For small values of θ Algorithm 4.1 is better than Algorithm 4.2. Thus, it seems natural to choose between the two algorithms according to which has the highest acceptance probability. Note also, that for small values of σ there is a region of values of θ where neither algorithm performs particular well.
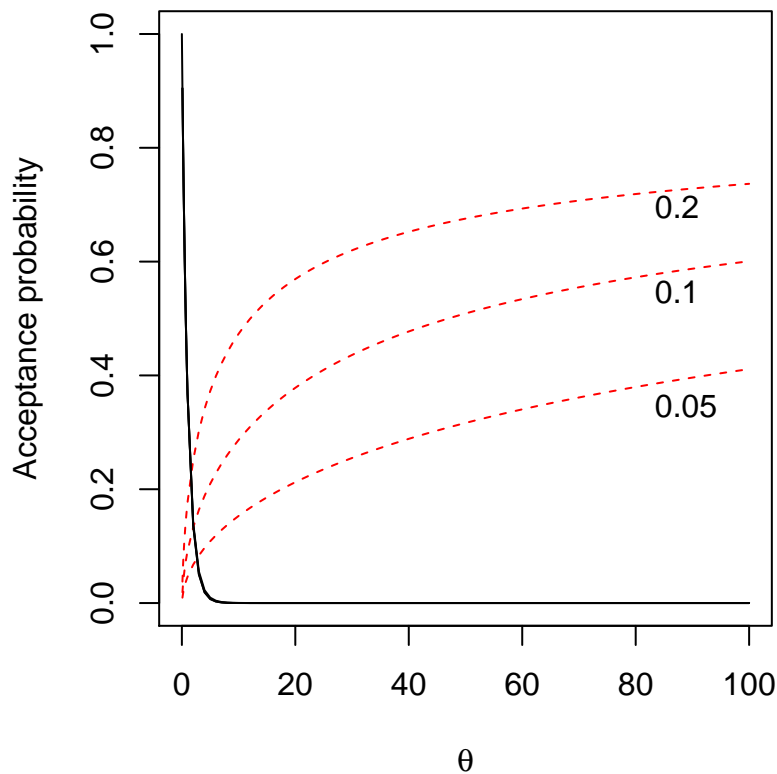

Figure 2: The acceptance probability for Algorithm 4.1 (full drawn lines) and for Algorithm 4.2 (dashed lines) for the three cases $\sigma = 0.05, 0.1, 0.2$. Thee are three full drawn lines that are almost indistinguishable.

4.2 Efficient Monte Carlo for left tails of lognormal sums

In this section we develop an *asymptotically efficient Monte Carlo estimator* $\widehat{\alpha}_n(x)$, for the left tail probability of a lognormal sum $\alpha_n(x) = \mathbb{P}(S_n \leq nx)$ as $x \to 0$.

We start by recalling some standard concepts from rare event simulation (VI.1 Asmussen and Glynn, 2007). In our setting, we say that a Monte Carlo estimator

11

$\widehat{\alpha}_n(x)$ is *strongly efficient* or has *bounded relative error* as $x \to 0$ if

$$\limsup_{x \to 0} \frac{\operatorname{Var} \widehat{\alpha}_n(x)}{\alpha_n^2(x)} < \infty.$$

This efficiency property implies that the number of replications required to estimate $\alpha_n(x)$ with certain fixed relative precision remains bounded as $x \to 0$. A weaker criterion is *logarithmic efficiency* defined as

$$\limsup_{x \to 0} \frac{\operatorname{Var} \widehat{\alpha}_n(x)}{\alpha_n^{2-\epsilon}(x)} = 0, \qquad \forall \epsilon > 0.$$

From a practical point of view, there is no substantial difference between these two criteria. However, it is often easier to prove logarithmic efficiency rather than bounded relative error. Logarithmic efficiency implies that the number of replications needed for achieving certain relative precision grows at rate of order at most $|\log(\alpha_n(x))|$. The efficiency properties can also be studied for $n \to \infty$ instead of $x \to 0$. We mention this situation below also.

An unbiased estimator can be obtained by using the variance reduction technique *importance sampling* (V.1, Asmussen and Glynn, 2007). This method relies on the existence of a Radon-Nikodym derivative with respect to a probability measure, say $\mathbb{Q}$. If we are interested in estimating $\mathbb{E}[h(W)]$, where $\mathbb{E}$ is the expectation operator under the measure $\mathbb{P}$, and $\mathbb{Q}$ is an absolutely continuous measure with respect to $\mathbb{P}$, then it holds that

$$\mathbb{E}[h(W)] = \mathbb{E}^{\mathbb{Q}}[L\, h(W)],$$

where $\mathbb{E}^{\mathbb{Q}}$ is the expectation operator under the measure $\mathbb{Q}$ and $L = \mathrm{d}\mathbb{P}/\mathrm{d}\mathbb{Q}$ is the Radon-Nikodym derivative of $\mathbb{P}$ with respect to $\mathbb{Q}$ (the last also goes under the name *likelihood ratio* in the simulation community). Hence, if X is simulated according to $\mathbb{Q}$, then $L\, h(W)$ serves as an unbiased estimator of the quantity $\mathbb{E}[h(W)]$. The strategy of selecting an importance distribution from the exponential family generated by the lognormal $\{F_\theta : \theta \in \Theta\}$ is often referred to as *exponential twisting, exponential tilting* or simply *exponential change of measure*. Ideally, the twisting parameter θ is selected as the value of the saddlepoint $\theta(\cdot)$ from (1.3) evaluated at x.

Notice, however, that difficulties arise in the right tail if the X_i's are heavy-tailed: then the integral associated with $\mathbb{E}[\mathrm{e}^{-\theta X}]$ diverges for negative values of the argument θ and in consequence, the equation (1.3) has no solution if $x > \mathbb{E}[X_i]$. Further difficulties in the heavy-tailed environment are exposed in Asmussen et al. (2000); Bassamboo et al. (2008). Nevertheless, exponential twisting can be implemented for the left tail probability of a lognormal sum; moreover, it turns out that it is logarithmically efficient.

Theorem 4.3. *Consider $X_1, \ldots, X_n \sim F_{\theta(x)}$ and set $S_n = X_1 + \cdots + X_n$. Define*

$$\beta_n(x) = L(\theta(x))^n \, \mathrm{e}^{\theta(x) S_n} \, \mathbb{I}\{S_n < nx\}.$$

Then $\beta_n(x)$ is a logarithmically efficient and unbiased estimator of $\alpha_n(x)$ as $n \to \infty$.

Proof. The lognormal density is log-concave for small x (see (2.7)) and so the result follows immediately from the proof of Theorem 2.10, Chapter VI in Asmussen and Glynn (2007). $\qquad\square$

The above optimal exponential twisting algorithm requires the value of the Laplace transform $L(\theta)$ and the saddlepoint $\theta(\cdot)$. These can be found by numerical integration. We next consider an alternative estimator based on the approximation (2.8) to the saddlepoint and an unbiased estimator of the Laplace transform. This alternative estimator is unbiased and logarithmic efficient as $x \to 0$

Algorithm 4.4.

1. Use the approximation $\tilde{\theta}(x)$ to the sadlepoint given in (2.8).

2. Obtain n independent unbiased estimates $\widehat{\mathcal{L}}_i(\tilde{\theta}(x)) = L_a(0, \tilde{\theta}(x))V_i$ of the Laplace transform and set $\widehat{\mathcal{L}}^n(\tilde{\theta}(x)) = \prod_{i=1}^n \widehat{\mathcal{L}}_i(\tilde{\theta}(x))$.

3. Simulate $X_1, \ldots, X_n \sim F_{\tilde{\theta}(x)}$ and set $S_n = X_1 + \cdots + X_n$.

4. Return
$$\widehat{\alpha}_n(x) = \mathrm{e}^{\tilde{\theta}(x)S_n} \widehat{\mathcal{L}}^n(\tilde{\theta}(x)) \, \mathbb{I}(S_n < nx). \tag{4.2}$$

The product of n independent copies of an unbiased estimate $\widehat{\mathcal{L}}(\tilde{\theta}(x))$ is needed because $\widehat{\mathcal{L}}^n(\tilde{\theta}(x))$ is not an unbiased estimate of $L^n(\tilde{\theta}(x))$. As suggested in Asmussen et al. (2014b) an unbiased estimator of $L(\theta)$ is obtained as

$$\widehat{\mathcal{L}}(\theta) = L_a(0, \theta)I(Y, \theta) \tag{4.3}$$

where $L_a(0, \theta)$ is given in (2.2), $Y \sim \mathrm{N}(0, \sigma^2)$ and

$$I(Y, \theta) = \exp\left\{ -\frac{w_0(\theta)}{\sigma^2}\left(\mathrm{e}^Y - 1 - Y\right) \right\} \frac{\sigma}{\sigma_0(\theta)}.$$

We next state the properties of the proposed algorithm.

Proposition 4.5. *Let $\widehat{\alpha}_n(x)$ be defined as in (4.2) and assume that $\mathbb{E}[V_i^2] \le c$ as $x \to 0$ for some constant c. Then $\widehat{\alpha}_n(x)$ is an unbiased and logarithmic efficient estimator of $\alpha_n(x)$ as $x \to 0$.*

For the proof, see Appendix A.

4.3 Density estimation

Consider the problem of estimating the density of a lognormal sum via simulation. Following Asmussen and Glynn (2007), Example V.4.3 p. 146, slightly extended, we first note that the conditional density at nx of S_n given

$$S_{n,-i} = X_1 + \cdots + X_{i-1} + X_{i+1} + \cdots + X_n = S_n - X_i$$

is $f(nx - S_{n,-i})$. Hence an unbiased estimator of $f_n(nx)$ is $\sum_1^n f(nx - S_{n,-i})/n$.

However, since we are dealing with values of x far to the left of $\mathbb{E}X$, it is likely that $S_{n,-i} > x$ so that $f(nx - S_{n,-i}) = 0$ and the procedure will come out with a

large number of zeroes. Hence we employ the same importance sampling estimator as used above. That is, we simulate the X_j from $F_{\tilde{\theta}(x)}$ and return the estimator

$$\widehat{f}_n(nx) \;=\; \frac{\exp\{\tilde{\theta}(x)nx + n\kappa\big(\tilde{\theta}(x)\big)\}}{n} \sum_{i=1}^{n} f_{\tilde{\theta}(x)}(nx - S_{n,-i})$$

(in practice to be averaged over R replications). A slight reformulation gives

$$\widehat{f}_n(nx) \;=\; \frac{1}{n} \sum_{i=1}^{n} f(nx - S_{n,-i}) \exp\{\tilde{\theta}(x)S_{n,-i} + (n-1)\kappa\big(\tilde{\theta}(x)\big)\}\,.$$

In Gulisashvili and Tankov (2014), an importance sampling estimator for $F_n(z)$ is suggested and it is written that a parallel estimator for $f_n(z)$ can be constructed in the same way. Nevertheless, we do not follow the details for the construction of that estimator of $f_n(z)$.

5 Numerical examples

In our numerical experiments, we have taken parameter values that we consider realistic from the point of view of financial applications. A yearly volatility of order 0.25 is often argued to be typical. We have considered periods of lengths one year, one quarter, one month and one week, corresponding to $\sigma = 0.25$, $\sigma = 0.25/\sqrt{4} = 0.125$, $\sigma = 0.25/\sqrt{12} = 0.072$, resp. $\sigma = 0.25/\sqrt{52} = 0.035$. Real-life portfolios are often large, even in the thousands; the values we have chosen are $n = 4, 16, 64, 256$.

For each combination of n and σ we have conducted several numerical empirical analyses. In all numerical experiments involving simulation we have employed $R = 100,000$ replications. The complete set of numerical results can be found in Asmussen et al. (2014a). Here we present and discuss an example with $n = 16$ and $\sigma = 0.125$.

Approximation of the Cramér function

We consider the approximation $\tilde{\theta}(x)$ given in equation (2.8) to the saddlepoint $\theta(x)$. The overall result is given in Proposition 2.4. Here we consider a few numerical illustrations. Table 1 gives $\tilde{\theta}(x)$, $\theta(x)$ and the mean under the tilted measure corresponding to $\tilde{\theta}(x)$. We want the latter mean to be close to x. As can be realized from the table the relative error of the latter mean as an approximation to x is less than one percent. Furthermore, when using $\tilde{\theta}(x)$ as the initial value in a Newton-Raphson search for $\theta(x)$, in all cases considered in Table 1 at most four iterations are needed to find $\theta(x)$ to accuracy 10^{-10}.

Left tail of the Lognormal Sum

Next we verify the approximations for the cdf and pdf of the lognormal sum. We have thereby been thinking of a portfolio of n assets with next-period values $Y_1, \ldots, Y_n$ assumed i.i.d. lognormal(μ, σ^2), such that a loss corresponds to a small value x of

Table 1: Evaluation of the approximation $\tilde{\theta}(x)$ from (2.8) to the saddlepoint $\theta(x)$ for the case $\sigma = 0.125$.

x	$\tilde{\theta}(x)$	$\theta(x)$	$\mathbb{E}_{\tilde{\theta}(x)}[X]$
1.0	0.500	0.496	0.9999
0.9	8.048	7.992	0.8994
0.8	18.477	18.360	0.7990
0.7	33.325	33.134	0.6989
0.6	55.322	55.037	0.5989
0.5	89.724	89.312	0.4991
0.4	147.857	147.257	0.3992
0.3	258.516	257.602	0.2994
0.2	517.522	515.977	0.1996
0.1	1478.659	1475.167	0.0998

$S_n = Y_1, \ldots, Y_n$. When choosing x, we have had in mind the recommended VaR values 0.99%–0.9997% of Basel II (2004) and chosen $\mathbb{P}(S_n \leq nx)$ to cover the interval 0.0001–0.0100.

We have proposed two type of approximations: saddlepoint approximations and Monte Carlo estimators. Thus, in Tables 2–3 we included the first and second order saddlepoint approximation (labeled Saddle$_1$ and Saddle$_2$) based on our formulas in Section 3, and a Monte Carlo estimators (MC) based on our algorithms in Section 4. The last is based on the proposed importance sampling estimator where the importance distribution is selected from the exponential family. The general estimator for the CDF of the lognormal sum has the form

$$\widehat{F}_n(nx) = L(\theta)^n \, e^{\theta S_n} \, \mathbb{I}\{S_n < nx\}, \tag{5.1}$$

where $S_n = X_1 + \cdots + X_n$ and $X_1, \ldots, X_n$ is a sample from the exponential family. Similarly, the MC estimator of the pdf of the lognormal sum has the form

$$\widehat{f}_n(nx) = L(\theta)^{n-1} \left[\frac{1}{n} \sum_{i=1}^{n} e^{\theta S_{n,-i}} f(nx - S_{n,-i}). \right].$$

The parameter θ defining the distribution is selected to be equal to the saddlepoint $\theta(\cdot)$ evaluated at x. Table 2 contains the numerical results for the CDF while results for the PDF are given in Table 3.

In the Tables 2–3, $\theta(x)$ corresponds to the solution obtained by using Newton-Raphson and the one used for obtaining the saddlepoint approximations and MC estimators. In the cases considered, the first and second order saddlepoint approximations for both the CDF and PDF are quite close, and the second order approximation agrees with the results from the Monte Carlo simulations. The last column of Table 2 indicates the relative error that one would introduce on replacing the Laplace transform in (5.1) with its approximation $L_a(0, \tilde{\theta})$. For $n = 16$ the relative errors are $(1 + \epsilon)^n - 1$, where ϵ is the entry in the table.

5.1 Comparison with Gulisashvili and Tankov (2014)

For comparison purposes we consider the following example where the asymptotic results of Gulisashvili and Tankov (2014) have been included. As that approximation

Table 2: Approximation of the CDF of a lognormal sum with $n = 16$ and $\sigma = 0.125$. The entry L_{app} is the relative error $L(\tilde{\theta}(x))/L_a(0, \tilde{\theta}(x)) - 1$.

x	$\theta(x)$	$\tilde{\theta}(x)$	Saddle$_1$	Saddle$_2$	MC	L_{app}
0.70	33.13	33.33	$1.755 \cdot 10^{-31}$	$1.761 \cdot 10^{-31}$	$(1.748 \pm 0.124) \cdot 10^{-31}$	$2.12 \cdot 10^{-4}$
0.80	18.36	18.48	$9.752 \cdot 10^{-14}$	$9.807 \cdot 10^{-14}$	$(9.819 \pm 0.171) \cdot 10^{-14}$	$2.04 \cdot 10^{-4}$
0.85	12.74	12.83	$3.009 \cdot 10^{-8}$	$3.031 \cdot 10^{-8}$	$(3.003 \pm 0.045) \cdot 10^{-8}$	$1.83 \cdot 10^{-4}$
0.90	7.99	8.05	$1.615 \cdot 10^{-4}$	$1.632 \cdot 10^{-4}$	$(1.624 \pm 0.098) \cdot 10^{-4}$	$1.48 \cdot 10^{-4}$
0.91	7.13	7.18	$5.892 \cdot 10^{-4}$	$5.956 \cdot 10^{-4}$	$(5.921 \pm 0.069) \cdot 10^{-4}$	$1.38 \cdot 10^{-4}$
0.92	6.30	6.34	$1.890 \cdot 10^{-3}$	$1.912 \cdot 10^{-3}$	$(1.932 \pm 0.021) \cdot 10^{-3}$	$1.28 \cdot 10^{-4}$
0.93	5.49	5.53	$5.358 \cdot 10^{-3}$	$5.424 \cdot 10^{-3}$	$(5.431 \pm 0.056) \cdot 10^{-3}$	$1.17 \cdot 10^{-4}$
0.94	4.71	4.74	$1.350 \cdot 10^{-2}$	$1.368 \cdot 10^{-2}$	$(1.363 \pm 0.013) \cdot 10^{-2}$	$1.06 \cdot 10^{-4}$
0.95	3.95	3.98	$3.039 \cdot 10^{-2}$	$3.081 \cdot 10^{-2}$	$(3.056 \pm 0.028) \cdot 10^{-2}$	$9.29 \cdot 10^{-5}$
0.98	1.82	1.83	$1.872 \cdot 10^{-1}$	$1.901 \cdot 10^{-1}$	$(1.911 \pm 0.014) \cdot 10^{-1}$	$4.92 \cdot 10^{-5}$

Table 3: Approximation of the PDF of a lognormal sum with $n = 16$ and $\sigma = 0.125$. The first and second order saddlepoint approximations are identical to the order stated.

x	$\theta(x)$	$\tilde{\theta}(x)$	Saddle$_1$	Saddle$_2$	MC
0.70	33.13	33.33	$5.873 \cdot 10^{-30}$	$5.873 \cdot 10^{-30}$	$(5.855 \pm 0.050) \cdot 10^{-30}$
0.80	18.36	18.48	$1.829 \cdot 10^{-12}$	$1.829 \cdot 10^{-12}$	$(1.834 \pm 0.016) \cdot 10^{-12}$
0.85	12.74	12.83	$3.975 \cdot 10^{-7}$	$3.975 \cdot 10^{-7}$	$(3.967 \pm 0.034) \cdot 10^{-7}$
0.90	7.99	8.05	$1.388 \cdot 10^{-3}$	$1.388 \cdot 10^{-3}$	$(1.393 \pm 0.012) \cdot 10^{-3}$
0.91	7.13	7.18	$4.576 \cdot 10^{-3}$	$4.577 \cdot 10^{-3}$	$(4.582 \pm 0.039) \cdot 10^{-3}$
0.92	6.30	6.34	$1.318 \cdot 10^{-2}$	$1.319 \cdot 10^{-2}$	$(1.317 \pm 0.011) \cdot 10^{-2}$
0.93	5.49	5.53	$3.332 \cdot 10^{-2}$	$3.332 \cdot 10^{-2}$	$(3.324 \pm 0.029) \cdot 10^{-2}$
0.94	4.71	4.74	$7.415 \cdot 10^{-2}$	$7.416 \cdot 10^{-2}$	$(7.416 \pm 0.064) \cdot 10^{-2}$
0.95	3.95	3.98	$1.459 \cdot 10^{-1}$	$1.460 \cdot 10^{-1}$	$(1.456 \pm 0.013) \cdot 10^{-1}$
0.98	1.82	1.83	$5.520 \cdot 10^{-1}$	$5.520 \cdot 10^{-1}$	$(5.505 \pm 0.047) \cdot 10^{-1}$

is only valid for values $x < 1$ we restrict to this set. The results are summarized in Table 4. We note that the relative error of our MC estimates (not shown) is of the order 10^{-2} in all cases.

However, only for very small values of $\mathbb{P}(S_n \leq z$ in the range of parameters we have considered does the asymptotic expression in Gulisashvili and Tankov (2014) become close to the exact value. It should be noted, however, that Gulisashvili and Tankov (2014) also applies to some specific types of dependence as well as different μ_i, σ_i^2 whereas we are restricted to the i.i.d. case.

Table 4: Approximations of the CDF and PDF of a lognormal sum with $n = 4$ and $\sigma = 0.250$.

x	z	$\theta(x)$	CDF			PDF		
			Saddle$_1$	MC	GT	Saddle$_1$	MC	GT
0.03	0.1	2365.14	$1.02 \cdot 10^{-192}$	$1.03 \cdot 10^{-192}$	$4.77 \cdot 10^{-192}$	$2.42 \cdot 10^{-189}$	$2.43 \cdot 10^{-189}$	$7.03 \cdot 10^{-189}$
0.05	0.2	961.13	$3.93 \cdot 10^{-128}$	$4.01 \cdot 10^{-128}$	$2.88 \cdot 10^{-127}$	$3.80 \cdot 10^{-125}$	$3.80 \cdot 10^{-125}$	$1.49 \cdot 10^{-124}$
0.08	0.3	554.44	$1.60 \cdot 10^{-96}$	$1.62 \cdot 10^{-96}$	$1.79 \cdot 10^{-95}$	$8.92 \cdot 10^{-94}$	$8.93 \cdot 10^{-94}$	$4.59 \cdot 10^{-93}$
0.10	0.4	369.92	$7.40 \cdot 10^{-77}$	$7.50 \cdot 10^{-77}$	$1.29 \cdot 10^{-75}$	$2.76 \cdot 10^{-74}$	$2.76 \cdot 10^{-74}$	$1.89 \cdot 10^{-73}$
0.13	0.5	267.46	$3.53 \cdot 10^{-63}$	$3.53 \cdot 10^{-63}$	$1.01 \cdot 10^{-61}$	$9.53 \cdot 10^{-61}$	$9.51 \cdot 10^{-61}$	$8.92 \cdot 10^{-60}$
0.15	0.6	203.51	$5.03 \cdot 10^{-53}$	$5.16 \cdot 10^{-53}$	$2.57 \cdot 10^{-51}$	$1.04 \cdot 10^{-50}$	$1.04 \cdot 10^{-50}$	$1.40 \cdot 10^{-49}$
0.17	0.7	160.39	$3.72 \cdot 10^{-45}$	$3.69 \cdot 10^{-45}$	$3.96 \cdot 10^{-43}$	$6.05 \cdot 10^{-43}$	$6.04 \cdot 10^{-43}$	$1.29 \cdot 10^{-41}$
0.20	0.8	129.71	$7.21 \cdot 10^{-39}$	$7.28 \cdot 10^{-39}$	$2.13 \cdot 10^{-36}$	$9.50 \cdot 10^{-37}$	$9.51 \cdot 10^{-37}$	$3.80 \cdot 10^{-35}$
0.23	0.9	106.96	$9.91 \cdot 10^{-34}$	$1.01 \cdot 10^{-33}$	$1.66 \cdot 10^{-30}$	$1.08 \cdot 10^{-31}$	$1.08 \cdot 10^{-31}$	$1.24 \cdot 10^{-29}$

References

Aitchison, I. and J. A. C. Brown (1957). *The Lognormal distribution with special reference to its uses in economics.* Cambridge: Cambridge University Press.

Asmussen, S. (2003). *Applied probability and queues.* New York: Springer-Verlag.

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

Asmussen, S., J. Blanchet, S. Juneja, and L. Rojas-Nandayapa (2011). Efficient simulation of tail probabilities of sums of correlated lognormals. *Annals of Operations Research 189*, 5–23.

Asmussen, S. and P. W. Glynn (2007). *Stochastic simulation: algorithms and analysis.* New York: Springer-Verlag.

Asmussen, S., J. L. Jensen, and L. Rojas-Nandayapa (2014a). Numerical examples accompanying the paper "Exponential Family Techniques for the Lognormal Left Tail". http://www.smp.uq.edu.au/people/LeonardoRojas-Nandayapa/NumericalExamples.pdf.

Asmussen, S., J. L. Jensen, and L. Rojas-Nandayapa (2014b). On the Laplace transform of the Lognormal distribution. *Submitted*.

Asmussen, S. and L. Rojas-Nandayapa (2008). Asymptotics of sums of lognormal random variables with Gaussian copula. *Stat. Probab. Lett. 78*, 2709–2714.

Basel II (2004). *Basel II: revised international capital framework.*

Bassamboo, A., S. Juneja, and A. Zeevi (2008). Portfolio credit risk with extremal dependence: Asymptotic analysis and efficient simulation. *Operations Research 56*, 593–606.

Blanchet, J. H. and L. Rojas-Nandayapa (2011). Efficient simulation of tail probabilities of sums of dependent random variables. *Journal of Applied Probability 48A*, 147–164.

Corless, R. M., G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth (1996). On the Lambert W function. *Advances in Computational Mathematics 5*, 329–359.

Crow, E. L. and K. Shimizu (1988). *Lognormal distributions: theory and applications*. New York: Marcel Dekker Inc.

Dufresne, D. (2004). The log-normal approximation in financial and other computations. *Advances in Applied Probability*, 747–773.

Dufresne, D. (2009). Sums of lognormals. Technical report, Centre for Actuarial Sciences, University of Melbourne.

Fenton, L. (1960). The sum of log-normal probability distributions in scatter transmission systems. *IRE Transactions on Communications Systems 8*(1), 57–67.

Foss, S. and A. Richards (2010). On sums of conditionally independent subexponential random variables. *Mathematics of Operations Research 35*, 102–119.

Gubner, J. A. (2006). A new formula for lognormal characteristic functions. *IEEE Transactions on Vehicular Technology 55*(5), 1668–1671.

Gulisashvili, A. and P. Tankov (2014). Tail behavior of sums and differences of log-normal random variables. *arXiv:1309.3057*.

Jensen, J. L. (1995). *Saddlepoint approximations*. Oxford: Oxford Science Publications.

Johnson, N. L., S. Kotz, and N. Balakrishnan (1994). *Continuous univariate distributions*, Volume 1. New York: Wiley.

Limpert, E., W. A. Stahel, and M. Abbt (2001). Log-normal distributions across the sciences: Keys and clues. *BioScience 51*, 341–352.

Marlow, N. A. (1967). A normal limit theorem for power sums of normal random variables. *Bell System Technical Journal 46*, 2081–2089.

Milevsky, M. A. and S. E. Posner (1998). Asian options, the sum of lognormals, and the reciprocal gamma distribution. *The Journal of Financial and Quantitative Analysis*.

Mitra, A. and S. I. Resnick (2009). Aggregation of risks and asymptotic independence. *Advances in Applied Probability 41*, 797–828.

Rojas-Nandayapa, L. (2008). *Risk probabilities: asymptotics and simulation*. Ph. D. thesis, Aarhus University.

Thorin, O. and N. Wikstad (1977). Calculation of ruin probabilities when the claim distribution is lognormal. *Astin Bulletin*, 231–246.

A   Appendix: Proof of Proposition 4.5

Let $\tilde{\theta} = \tilde{\theta}(x)$, $\tilde{\mu} = \mathbb{E}_{\tilde{\theta}}[X]$, $\tilde{\sigma}^2 = \text{Var}_{\tilde{\theta}}[X]$ and $Z_n = (S_n - n\tilde{\mu})/\tilde{\sigma}$, where S_n is based on a random sample from $F_{\tilde{\theta}}$. We want to estimate

$$\alpha_n(x) = \mathbb{P}(S_n \leq nx) = \mathbb{E}_{\tilde{\theta}}[L(\tilde{\theta})^n \mathrm{e}^{\tilde{\theta} S_n} \mathbb{I}\{S_n < nx\}$$

$$= \frac{\mathrm{e}^{n\tilde{\theta}\tilde{\mu}}}{\sqrt{n}\tilde{\theta}\tilde{\sigma}} L(\tilde{\theta})^n \mathbb{E}_{\tilde{\theta}}[\sqrt{n}\tilde{\theta}\tilde{\sigma}\mathrm{e}^{\sqrt{n}\tilde{\theta}\tilde{\sigma}(\xi - Z_n)}\mathbb{I}\{Z_n < \xi\}], \quad \xi = \sqrt{n}(x - \tilde{\mu})/\tilde{\sigma}. \quad \text{(A.1)}$$

We know from Proposition 2.3 that as $x \to 0$, corresponding to $\tilde{\theta} \to \infty$, the distribution of Z_n approaches a standard normal distribution. Furthermore, $\tilde{\theta} \sim |\log x|/(\sigma^2 x)$, $\tilde{\mu} \sim x(1 + \mathrm{O}(1/|\log x|))$, $\tilde{\sigma} \sim x\sigma/\sqrt{(|\log x|)}$, which gives

$$\xi = \frac{\sqrt{n}(x - \tilde{\mu})}{\tilde{\sigma}} = \mathrm{O}\big(\sqrt{n}/\sqrt{(|\log x|)}\big) \to 0,$$

and

$$\sqrt{n}\tilde{\theta}\tilde{\sigma} \sim \frac{\sqrt{|\log x|}}{\sigma} \to \infty.$$

These findings show that the mean value in (A.1) tends to $1/\sqrt{2\pi}$ as $x \to 0$. The same type of argument also gives that

$$\mathbb{E}_{\tilde{\theta}}\big[\big\{\sqrt{n}\tilde{\theta}\tilde{\sigma}\mathrm{e}^{-\sqrt{n}\tilde{\theta}\tilde{\sigma}(xi-Z_n)}\mathbb{I}\{Z_n < \xi\}\big\}^2\big]$$
$$= \tfrac{1}{2}\sqrt{n}\tilde{\theta}\tilde{\sigma}E_{\tilde{\theta}}\big[2\sqrt{n}\tilde{\theta}\tilde{\sigma}\mathrm{e}^{-2\sqrt{n}\tilde{\theta}\tilde{\sigma}(\xi-Z_n)}\mathbb{I}\{Z_n < \xi\}\big]$$
$$\sim \frac{1}{2\sqrt{2\pi}}\sqrt{n}\tilde{\theta}\tilde{\sigma}. \tag{A.2}$$

Consider now the unbiased estimator

$$\hat{\beta}_n(x) = L(\tilde{\theta})^n \mathrm{e}^{\tilde{\theta}S_n}\mathbb{I}\{S_n < nx\},$$

where S_n is based on a sample from the tilted measure $F_{\tilde{\theta}}$. The above calculations show that

$$\frac{\mathrm{Var}[\hat{\beta}_n(x)]}{\beta_n(x)^{2-\epsilon}} = \mathrm{O}\bigg(\frac{\big\{\mathrm{e}^{n\tilde{\theta}\tilde{\mu}}L(\tilde{\theta})^n/(\sqrt{n}\tilde{\theta}\tilde{\sigma})\big\}^2\sqrt{n}\tilde{\theta}\tilde{\sigma}}{\big\{\mathrm{e}^{n\tilde{\theta}\tilde{\mu}}L(\tilde{\theta})^n/(\sqrt{n}\tilde{\theta}\tilde{\sigma})\big\}^{2-\epsilon}}\bigg)$$
$$= \mathrm{O}\Big(\big\{\mathrm{e}^{n\tilde{\theta}\tilde{\mu}}L(\tilde{\theta})^n\big\}^{\epsilon}\big(\sqrt{n}\tilde{\theta}\tilde{\sigma}\big)^{1-\epsilon}\Big)$$
$$= \mathrm{O}\big(\exp\{-\epsilon w_0(\tilde{\theta})/\sigma^2\}w_0(\tilde{\theta})^{(1-\epsilon)/2}\big) \to 0.$$

This shows the logaritmic efficiency of the estimator $\hat{\beta}_n(x)$.

Consider next the unbiased estimator

$$\hat{\alpha}_n(x) = \widehat{\mathcal{L}}^n(\tilde{\theta})\mathrm{e}^{\tilde{\theta}S_n}\mathbb{I}\{S_n < nx\},$$

where $\widehat{\mathcal{L}}^n(\tilde{\theta}) = \prod_{i=1}^n \widehat{\mathcal{L}}_i(\tilde{\theta})$. Here $\widehat{\mathcal{L}}_i(\tilde{\theta})$, $i = 1, \ldots, n$, are independent and $\widehat{\mathcal{L}}_i(\tilde{\theta}) = L_a(0, \tilde{\theta})V_i$ with

$$\mathbb{E}[V_i] = L(\tilde{\theta})/L_a(0, \tilde{\theta}) \quad \text{and} \quad \mathbb{E}[V_i^2] \le c,$$

for some constant c. Instead of (A.2) we have

$$\mathbb{E}_{\tilde{\theta}}\Big[\Big\{\Big(\prod_{i=1}^n V_i\Big)\sqrt{n}\tilde{\theta}\tilde{\sigma}\mathrm{e}^{-\sqrt{n}\tilde{\theta}\tilde{\sigma}(xi-Z_n)}\mathbb{I}\{Z_n < \xi\}\Big\}^2\Big] = \mathrm{O}\big(c^n\sqrt{n}\tilde{\theta}\tilde{\sigma}\big),$$

and

$$\frac{\mathrm{Var}[\hat{\alpha}_n(x)]}{\alpha_n(x)^{2-\epsilon}} = \mathrm{O}\bigg(\frac{\big\{\mathrm{e}^{n\tilde{\theta}\tilde{\mu}}L_a(0, \tilde{\theta})^n/(\sqrt{n}\tilde{\theta}\tilde{\sigma})\big\}^2\sqrt{n}\tilde{\theta}\tilde{\sigma}}{\big\{\mathrm{e}^{n\tilde{\theta}\tilde{\mu}}L_a(0, \tilde{\theta})^n/(\sqrt{n}\tilde{\theta}\tilde{\sigma})\big\}^{2-\epsilon}}\bigg)$$
$$= \mathrm{O}\Big(\big\{\mathrm{e}^{n\tilde{\theta}\tilde{\mu}}L_a(0, \tilde{\theta})^n\big\}^{\epsilon}\big(\sqrt{n}\tilde{\theta}\tilde{\sigma}\big)^{1-\epsilon}\Big)$$
$$= \mathrm{O}\big(\exp\{-\epsilon w_0(\tilde{\theta})/\sigma^2\}w_0(\tilde{\theta})^{(1-\epsilon)/2}\big) \to 0.$$

This shows the logaritmic efficiency of the estimator $\hat{\alpha}_n(x)$.