THIELE CENTRE
FOR APPLIED MATHEMATICS IN NATURAL SCIENCE

Orthonormal polynomial expansions and lognormal sum densities

Søren Asmussen, Pierre-Olivier Goffard
and Patrick J. Laub

Research Report

No. 01 | February 2016

Orthonormal polynomial expansions and lognormal sum densities

Søren Asmussen, Pierre-Olivier Goffard and Patrick J. Laub

Department of Mathematics
Aarhus University

Abstract

Approximations for an unknown density g in terms of a reference density f_ν and its associated orthonormal polynomials are discussed. The main application is the approximation of the density f of a sum S of lognormals which may have different variances or be dependent. In this setting, g may be f itself or a transformed density, in particular that of $\log S$ or an exponentially tilted density. Choices of reference densities f_ν that are considered include normal, gamma and lognormal densities. For the lognormal case, the orthonormal polynomials are found in closed form and it is shown that they are not dense in $L_2(f_\nu)$, a result that is closely related to the lognormal distribution not being determined by its moments and provides a warning to the most obvious choice of taking f_ν as lognormal. Numerical examples are presented and comparison are made to established approaches such as the Fenton–Wilkinson method and skew-normal approximations. Also extension to density estimation for statistical data sets and non-Gaussian copulas are outlined.

Keywords:: Lognormal distribution, sums of lognormally distributed random variable, orthogonal polynomial, density estimation, Stieltjes moment problem, numerical approximation of functions, exponential tilting, conditional Monte Carlo, Archimedean copula, Gram–Charlier expansion, Hermite polynomial, Laguerre polynomial

1 Introduction

The lognormal distribution arises in a wide variety of disciplines such as engineering, economics, insurance, finance, and across the sciences Aitchison and Brown (1957); Crow and Shimizu (1988); Johnson et al. (1994); Limpert et al. (2001); Dufresne (2009). Therefore, it is natural that sums S of n lognormals come up in a number of contexts. A basic example in finance is the Black–Scholes model, which assumes that security prices are lognormals, and hence the value of a portfolio with n securities has the form S. In insurance, individual claim sizes are often taken as independent lognormals, so the total claim amount in a certain period is again of form 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).

The distribution S is, however, not available in explicit form, and evaluating it numerically or approximating it is considered to be a challenging problem with a long history. The classical approach is to use an approximation with another lognormal distribution. This goes back at least to Fenton (1960b) and it is nowadays known as the *Fenton–Wilkinson method* as according to Marlow (1967) this approximation was already used by Wilkinson in 1934. However, it can be rather inaccurate when the number of summands is rather small, or when the dispersion parameter is too high. Also tail approximations have been extensively discussed, with the right tail being a classical example in subexponential theory,Embrechts et al. (1997), and the study of the left tail being more recent, Asmussen et al. (2015), Gulisashvili and Tankov (2013).

This paper discusses a different method, to approximate the probability density function (p.d.f.) f via polynomials $\{Q_k\}$ which are orthonormal w.r.t. some reference measure ν. In the general formulation, one is interested in approximating a target density g using the density f_ν of ν as reference and g some other density. One then finds a series representation of g/f_ν of the form $\sum_{k=0}^{\infty} a_k Q_k$, and then the approximation of g is

$$\widehat{g}(x) \;=\; g_\nu(x) \sum_{k=0}^{K} a_k Q_k(x), \tag{1.1}$$

for some suitable K. The most obvious connection to the lognormal sum problem is $g = f$, but we shall look also at other possibilities, to take g as the density of $\log S$ and transform back to get the approximation $\widehat{f}(x) = \widehat{g}(\log x)/x$ or to use an exponential tilting. The choice of ν is a crucial step, and three candidates for ν are investigated: the normal, the gamma, and the lognormal distributions.

The form of the Q_k is classical for the normal distribution where it is the Hermite polynomials and for the gamma where it is the Laguerre polynomials, but for the lognormal distributions it does not appear to be in the literature and we give here the functional expression (Theorem 2.3). The Fenton–Wilkinson method may be seen as the $K = 2$ case of f_ν being lognormal of the general scheme, and this choice of f_ν may be the most obvious one. However, we show that in the lognormal case the orthonormal polynomials are not dense in $L_2(f_\nu)$. This result is closely related to the lognormal distribution not being determined by its moments Heyde (1963); Berg et al. (1984) and indicates that a lognormal f_ν is potentially dangerous. For this reason, the rest of the paper concentrates on taking the reference distribution as normal (using the logarithmic transformation) or gamma (using exponential tilting).

After discussing the details of the orthonormal polynomials expansions in Sections 2 and 3, we proceed in Section 4 to show a number of numerical examples. The polynomial expansions are compared to existing methods as Fenton–Wilkinson and a more recent approximation in terms of log skew normal distributions Hcine and Bouallegue (2015), as well as to exact values obtained by numerical quadrature in cases where this is possible or by Monte Carlo density estimation. Section 4 also outlines an extension to statistical data sets and non-Gaussian copulas. Appendices A.1 contains a technical proof and Appendix A.2 some new material on the $\mathcal{SLN}$ Laplace transform.

2 Orthogonal polynomial representation of probability density functions

Let X be a random variable which has a density f with respect to some measure $\lambda \geq 0$ (typically Lebesgue measure on an interval or counting measure on a subset of $\mathbb{Z}$). If f is unknown but the distribution of X is expected to be close to some probability measure ν with p.d.f. f_ν, one may use f_ν as a first approximation to f and next try to improve by invoking suitable correction terms.

In the setting of this paper X is the sum of lognormal r.v.s and the correction terms are obtained by expansions in terms of orthonormal polynomials. Before going into the details of the lognormal example, let us consider the general case.

Assuming all moments of ν to be finite, the standard Gram–Schmidt orthogonalization technique shows the existence of a set of polynomials $\{Q_k\}_{k\in\mathbb{N}_0}$ which are orthonormal in $\mathcal{L}^2(\nu)$ equipped with the usual inner product $\langle g, h\rangle = \int gh \,\mathrm{d}\nu$ and the corresponding norm $\|g\|^2 = \langle g, g\rangle$. That is, the Q_k satisfy

$$\big\langle Q_i, Q_j \big\rangle = \int Q_i(x)Q_j(x)\,\mathrm{d}\nu(x) = \delta_{ij}\,, \qquad i,j \in \mathbb{N}_0, \tag{2.1}$$

where δ_{ij} denotes the Kronecker symbol. If there exists an $\alpha > 0$ such that

$$\int \mathrm{e}^{\alpha|x|}\,\mathrm{d}\nu(x) < \infty\,, \tag{2.2}$$

the set $\{Q_k\}_{k\in\mathbb{N}_0}$ is complete in $\mathcal{L}^2(\nu)$, cf. Chapter 7 of the book by Nagy Nagy (1965). The implication is that if f/f_ν is in $\mathcal{L}^2(\nu)$, that is, if

$$\int \frac{f(x)^2}{f_\nu(x)^2}\,\mathrm{d}\nu(x) = \int \frac{f(x)^2}{f_\nu(x)}\,\mathrm{d}\lambda(x) < \infty\,, \tag{2.3}$$

we may expand f/f_ν as $\sum_{k=0}^\infty a_k Q_k$ where

$$a_k = \big\langle f/f_\nu, Q_k \big\rangle = \int fQ_k \;\mathrm{d}\lambda = \mathbb{E}\left[Q_k(X)\right]\,. \tag{2.4}$$

This suggests that we use (1.1) as an approximation of f in situations where the p.d.f. of X is unknown but the moments are accessible.

Remark 2.1. If the first m moments of X and ν coincide, one has $a_k = 0$ for $k = 1, \ldots, m$. When choosing ν, a possible guideline is therefore to match as many moments as possible. ◊

Due to the Parseval relationship $\sum_{k=0}^\infty a_k^2 = \|f/f_\nu\|^2$, the coefficients of the polynomial expansion, $\{a_k\}_{k\in\mathbb{N}_0}$, tend toward 0 as $k \to \infty$. The accuracy of the approximation (1.1), for a given order of truncation K, depends upon how swiftly the coefficients decay; note that the $\mathcal{L}^2$ loss of the approximation of f/f_ν is $\sum_{K+1}^\infty a_k^2$. Note also that the orthogonal polynomials can be specified recursively (see Thm. 3.2.1 of Szegö (1939)) which allows a reduction of the computing time required for the coefficients' evaluation and makes it feasible to consider rather large K.

2.1 Normal reference distribution

A common choice as a reference distribution is the normal $\mathcal{N}(\mu, \sigma^2)$. The associated orthonormal polynomial are given by

$$Q_k(x) = \frac{1}{k! 2^{k/2}} H_k\left(\frac{x-\mu}{\sigma\sqrt{2}}\right), \tag{2.5}$$

where $\{H_k\}_{k\in\mathbb{N}_0}$ are the Hermite polynomials, defined in Szegö (1939) for instance. If f is continuous, a sufficient (and close to necessary) condition for $f/f_\nu \in \mathcal{L}^2(\nu)$ is

$$f(x) = \mathcal{O}\big(\mathrm{e}^{-ax^2}\big) \qquad \text{as } x \to \pm\infty \text{ with } a > \big(4\sigma^2\big)^{-1}. \tag{2.6}$$

Indeed, we can write the integral in (2.3) as $I_1+I_2+I_3$, the integrals over $(-\infty, -A)$, $[-A, A]$, resp. (A, ∞). Note that $I_2 < \infty$ follows since the integrand is finite by continuity, whereas the finiteness of I_1, I_3 is ensured by the integrands being $\mathcal{O}(\mathrm{e}^{-bx^2})$ where $b = 2a - 1/2\sigma^2 > 0$. Similar arguments apply to conditions (2.9) and (2.12) below.

Remark 2.2. The expansion formed by a standard normal baseline distribution and Hermite polynomials is known in the literature as Gram–Charlier expansion of type A, and the application to a standardised sum is the Edgeworth expansion, cf. Cramer (1999), Barndorff-Nielsen and Cox (1989). ♢

2.2 Gamma reference distribution

If X has support $(0, \infty)$, it is natural to look for a ν with the same property. One of the most apparent possibilities is the gamma distribution, denoted Gamma(r, m) where r is the shape parameter and m the scale parameter. The p.d.f. is

$$f_\nu(x) = \frac{x^{r-1}\mathrm{e}^{-x/m}}{m^r\Gamma(r)}, \qquad x \in \mathbb{R}_+. \tag{2.7}$$

The associated polynomials are given by

$$Q_n(x) = (-1)^n \left[\frac{\Gamma(n+r)}{\Gamma(n+1)\Gamma(r)}\right]^{-1/2} L_n^{r-1}(x/m), \qquad n \in \mathbb{N}_0, \tag{2.8}$$

where $\{L_n^{r-1}\}_{n\in\mathbb{N}_0}$ denote the *generalised Laguerre polynomials*, see Szegö (1939); in Mathematica these are accessible via the `LaguerreL` function. Similarly to (2.6), one has the following condition for $f/f_\nu \in \mathcal{L}^2(\nu)$:

$$\begin{aligned} f(x) &= \mathcal{O}\big(\mathrm{e}^{-\delta x}\big) && \text{as } x \to \infty \text{ with } \delta > 1/2m, \text{ and} \\ f(x) &= \mathcal{O}\big(x^\beta\big) && \text{as } x \to 0 \text{ with } \beta > r/2 - 1. \end{aligned} \tag{2.9}$$

2.3 Lognormal reference distribution

The lognormal distribution $\mathcal{LN}(\mu, \sigma^2)$ is the distribution of e^Y where $Y \sim \mathcal{N}(\mu, \sigma^2)$. It has support on $\mathbb{R}_+$. The polynomials orthogonal to the $\mathcal{LN}(\mu, \sigma^2)$ are given in the following proposition, to be proved in the Appendix:

Theorem 2.3. *The polynomials orthonormal with respect to the lognormal distribution are given by*

$$Q_k(x) = \frac{\mathrm{e}^{-\frac{k^2\sigma^2}{2}}}{\sqrt{[\mathrm{e}^{-\sigma^2}, \mathrm{e}^{-\sigma^2}]_k}} \sum_{i=0}^{k} (-1)^{k+i} \mathrm{e}^{-i\mu - \frac{i^2\sigma^2}{2}} e_{k-i}\big(1, \ldots, \mathrm{e}^{(k-1)\sigma^2}\big) x^i, \tag{2.10}$$

for $k \in \mathbb{N}_0$ *where*

$$e_i\left(X_1, \ldots, X_k\right) = \begin{cases} \sum_{1 \le j_1 < \cdots < j_i \le k} X_{j_1} \cdots X_{j_i}, & \text{for } i \le k, \\ 0, & \text{for } i > k, \end{cases} \tag{2.11}$$

are the elementary symmetric polynomials and $[x, q]_n = \prod_{i=0}^{n-1} \left(1 - xq^i\right)$ *is the Q-Pochhammer symbol.*

Remark 2.4. The result of Theorem 2.3 does not appear to be in the literature; the closest reference seems to be a 1923 paper by Wigert Wigert (1923) who considers the distribution with p.d.f. $\ell \mathrm{e}^{-\ell^2 \ln^2(x)}/\sqrt{\pi}$ (later called the Stieltjes–Wigert distribution). ♢

The equivalent of condition (2.6) for $f/f_\nu \in \mathcal{L}^2(\nu)$ now becomes

$$f(x) = \mathcal{O}\big(\mathrm{e}^{-b \log^2 x}\big) \qquad \text{for } x \to 0 \text{ and } \infty \text{ with } b > \big(4\sigma^2\big)^{-1}, \tag{2.12}$$

which is rather mild. However, a key difficulty in taking the reference distribution as lognormal is the following result related to the fact that the lognormal and the Stieltjes-Wigert distributions are not characterised by their moments, see Heyde (1963); Berg et al. (1984); Chihara (1979); Christiansen (2003). Hence, the orthogonal polynomials associated with the lognormal p.d.f. and the Stieltjes-Wigert p.d.f. are also the orthogonal polynomials for some other distribution.

Proposition 2.5. *The set of orthonormal polynomials in Theorem 2.3 is incomplete in* $\mathcal{L}^2(\nu)$*. That is,* $\operatorname{span}\{Q_k\}_{k \in \mathbb{N}_0}$ *is a proper subset of* $\mathcal{L}^2(\nu)$*.*

Proof. Let Y be a r.v. whose distribution is the given lognormal ν and X a r.v. with a distribution different from Y but with the same moments. According to (Berg et al., 1984, pp. 201–202) such an X can be chosen such that f_X/f_ν is bounded and hence in $\mathcal{L}^2(\nu)$. The projection of f/f_ν onto $\operatorname{span}\{Q_k\}$ is then

$$\begin{aligned} \sum_{k=0}^{\infty} \left\langle f/f_\nu, Q_k \right\rangle Q_k &= \sum_{k=0}^{\infty} \mathbb{E}\left[Q_k(X)\right] Q_k = \sum_{k=0}^{\infty} \mathbb{E}\left[Q_k(Y)\right] Q_k \\ &= Q_0 = 1 \neq f/f_\nu, \end{aligned}$$

where the first step used (2.4) and the second that the moments are the same. This implies $f/f_\nu \in \mathcal{L}^2(\nu) \setminus \operatorname{span}\{Q_k\}$ and the assertion. □

2.4 Convergence of the estimators w.r.t. K

Orthogonal polynomial approximations generally become more accurate as the order of the approximation K increases. Figure 1 shows a specific orthogonal polynomial approximation, $\widehat{f_{\mathcal{N}}}$ (to be described in Section 3.2), converging to the true $\mathcal{SLN}$ density f for increasing K. In this example, we take the $\mathcal{SLN}$ distribution with $\mu = (0, 0, 0)^\top$, $\Sigma_{ii} = 0.1$, and $\rho = -0.1$.

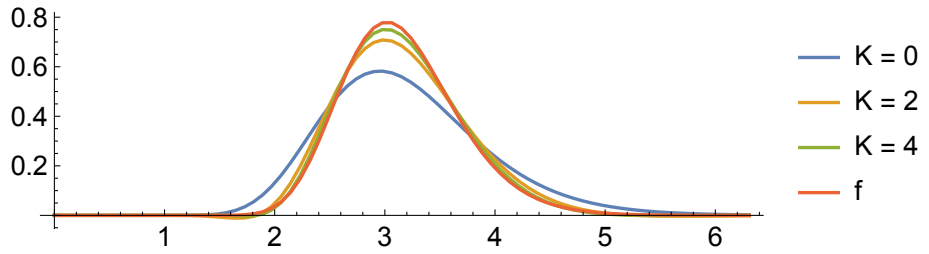


Figure 1: Examples of orthogonal polynomial approximations using a $\mathcal{N}(1.13, 0.23^2)$ reference converging to the target f with increasing K.

Proposition 2.5 implies that orthogonal polynomial approximations with a lognormal reference distribution cannot be relied upon to converge to the desired target density but may have a different limit (the orthogonal projection described there). The next plot, Figure 2, illustrates this phenomenon. The approximation appears to converge, but not to the target density. Our theoretical discussion suggests that this incorrect limit density has the same moments as the target lognormal distribution, and this was verified numerically for the first few moments,

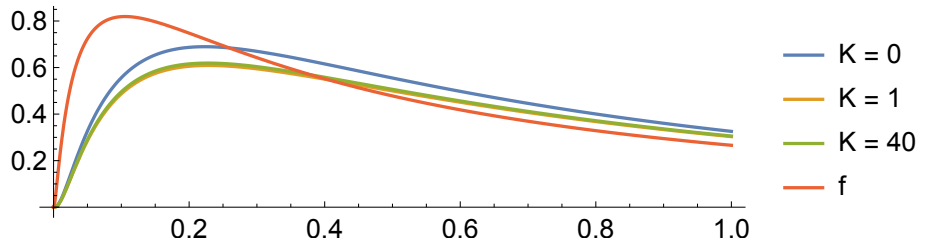


Figure 2: Example of orthogonal polynomial approximations of f using a $\mathcal{LN}(0, 1.22^2)$ reference not converging to the $\mathcal{LN}(0, 1.50^2)$ target.

Lastly, it must be noted that we cannot in practice take K arbitrarily large, due to numerical errors incurred in calculating the $\{a_k\}$ coefficients. Obviously this can be overcome by using infinite precision operations, however this swiftly becomes prohibitively slow. Software tools like `Mathematica` allow for arbitrarily large but finite precision, which gives on the flexibility to choose a desired accuracy/speed trade-off. We use this technology and select $K \leq 40$.

3 Application to lognormal sums

We now turn to our main case of interest where $X = S$ is a lognormal sum. Specifically,

$$S = \mathrm{e}^{X_1} + \cdots + \mathrm{e}^{X_n}, \qquad n \geq 2, \tag{3.1}$$

where the vector $\boldsymbol{X} = (X_1, \ldots, X_n)$ is governed by a multivariate normal distribution $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)^\top$ is the mean vector and $\boldsymbol{\Sigma} = \left(\sigma_{ij}\right)$ the covariance matrix. We write this distribution as $\mathcal{SLN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and hereafter denote its p.d.f. as f. We are interested in computing the p.d.f. when the summands exhibit dependency ($\boldsymbol{\Sigma}$ is non-diagonal). This is an ambitious goal given that the p.d.f. of the sum of two i.i.d lognormally distributed random variables is already unknown.

The validity of the polynomial approximations rely on the $\mathcal{L}^2$ integrability condition (2.3), which is difficult to check because the p.d.f. of S is not available. We will need asymptotic results describing the left and the right tail of the distribution of S, which we collect in the following subsection.

3.1 Tail asymptotics of lognormal sums

The tail asymptotics of $f(x)$ are given in the following lemma, which simply collects the results from Corollary 2 of Gulisashvili and Tankov (2015) and Theorem 1 of Asmussen and Rojas-Nandapaya (2008).

Lemma 3.1. *We have*

$$f(x) = \mathcal{O}(\exp\{-c_1 \ln(x)^2\}) \text{ as } x \to 0 \text{ and} \tag{3.2}$$
$$f(x) = \mathcal{O}(\exp\{-c_2 \ln(x)^2\}) \text{ as } x \to \infty \tag{3.3}$$

where

$$c_1 = \left[2 \min_{\boldsymbol{w} \in \Delta} \boldsymbol{w}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{w}\right]^{-1} \quad \text{and} \quad c_2 = \left[2 \max_{i=1,\dots,n} \sigma_{ii}\right]^{-1},$$

with the notation that $\Delta = \{\boldsymbol{w} \,|\, w_i \in \mathbb{R}_+, \sum_{i=1}^n w_i = 1\}$.

We are also interested in the asymptotic behaviour of $Z = \ln(S)$ later in the paper. Writing the p.d.f. of Z as f_Z we have $f_Z(z) = \mathrm{e}^z f(\mathrm{e}^z)$. Together with L'Hôpital's rule this gives the following results (extending Gao et al. (2009)):

Corollary 3.2. *We have*

$$f_Z(z) = \mathcal{O}(\exp\{-c_1 z^2\}) \text{ as } z \to -\infty \text{ and} \tag{3.4}$$
$$f_Z(z) = \mathcal{O}(\exp\{-c_2 z^2\}) \text{ as } z \to +\infty \tag{3.5}$$

where the constants are as in Lemma 3.1.

3.2 Lognormal sums via a normal reference distribution

Consider transforming S to $Z = \ln(S)$ and expanding this density with orthogonal polynomials using a normal distribution as reference. That is, our approximation to f using a $\mathcal{N}(\mu, \sigma^2)$ reference is

$$\widehat{f}_{\mathcal{N}} = \frac{1}{\sigma x} \widehat{f}_Z \left(\frac{\ln x - \mu}{\sigma} \right) \qquad \text{where } \widehat{f}_Z(z) = \phi(z) \sum_{i=1}^{K} a_i \, Q_i(z),$$

with $\phi(\cdot)$ being the standard normal p.d.f. The following result tells us when the integrability condition $f_Z / f_\nu \in \mathcal{L}^2(\nu)$ is satisfied. It follows immediately by combining (2.6) and Corollary 3.2

Proposition 3.3. *Consider* $Z = \ln(S)$ *where* S *is* $\mathcal{SLN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ *distributed. Let* ν *be the probability measure associated to the normal distribution* $\mathcal{N}(\mu, \sigma^2)$. *We have* $f_Z / f_\nu \in \mathcal{L}^2(\nu)$ *if*

$$2\sigma^2 > (2c_2)^{-1} = \max_{i=1,\dots,n} \Sigma_{ii} \,. \tag{3.6}$$

Computing the $\{\hat{a}_k\}_{k\in\mathbb{N}_0}$ coefficients can be done using Crude Monte Carlo (CMC), as in

$$\widehat{a}_k = \frac{1}{R}\sum_{r=1}^{R} Q_n(S_r)\,, \quad S_1,\ldots,S_R \overset{\text{i.i.d.}}{\sim} \mathcal{SLN}(\boldsymbol{\mu},\boldsymbol{\Sigma})$$

for $k = 0,\ldots,K$. We can use the same $S_1,\ldots,S_R$ for all $\widehat{a}_k$ together with a smoothing technique called *common random numbers* Asmussen and Glynn (2007); Glasserman (2003). Note that a non-trivial amount of computational time is typically spent just constructing the Hermite polynomials. Incorporating the Hermite polynomial's recurrence relation in our calculations achieved a roughly 40× speed-up compared with using Mathematica's `HermiteH`.

3.3 Lognormal sums via a gamma reference distribution

When ν is Gamma(r,m), it makes little sense to expand f in terms of $\{Q_k\}_{k\in\mathbb{N}_0}$ and f_ν as the integrability condition (2.9) fails, $f/f_\nu \notin \mathcal{L}^2(\nu)$. The workaround consists in using orthogonal polynomials to expand the *exponentially tilted* distribution, denoted $\mathcal{SLN}_\theta(\boldsymbol{\mu},\boldsymbol{\Sigma})$. This distribution's p.d.f. is

$$f_\theta(x) = \frac{\mathrm{e}^{-\theta x} f(x)}{\mathscr{L}(\theta)}\,, \qquad \theta \geq 0, \tag{3.7}$$

where $\mathscr{L}(\theta) = \mathbb{E}[\mathrm{e}^{-\theta S}]$ is the Laplace transform of S. Asmussen et al. Asmussen et al. (2015) investigated the use of $f_\theta(x)$ in approximating the survival function of S, and developed asymptotic forms and Monte Carlo estimators of this density.

Remark 3.4. The use of gamma distribution and Laguerre polynomials links our approach to a well established technique called the *Laguerre method*. The expansion is an orthogonal projection onto the basis of Laguerre functions constructed by multiplying Laguerre polynomials and the square root of the exponential distribution with parameter 1. The method is described in Abate et al. (1995). Note also that the damping procedure employed when integrability problems arise is quite similar to considering the exponentially tilted distribution instead of the real one. The use of the gamma distribution as reference is applied to actuarial science in Goffard et al. (2016, 2015). ♢

Using (2.9), we immediately obtain the following result which sheds light on how to tune the parameters of the reference gamma distribution so the integrability condition $f_\theta/f_\nu \in \mathcal{L}^2(\nu)$ is satisfied.

Proposition 3.5. *Consider the r.v. S_θ distributed according to the exponentially-tilted $\mathcal{SLN}_\theta(\boldsymbol{\mu},\boldsymbol{\Sigma})$ distribution. Let ν be the probability measure associated with the Gamma(r,m) distribution. We have $f_\theta/f_\nu \in \mathcal{L}^2(\nu)$ if $m > 1/2\theta$.*

Hereafter we assume that the parameters r and m of $f_\nu \sim$ Gamma(r,m) are chosen to satisfy Proposition 3.5's conditions.

Our approximation—based upon rearranging (3.7)—is of the form

$$\widehat{f}(x) = \mathrm{e}^{\theta x}\mathscr{L}(\theta)\widehat{f}_\theta(x) = \mathrm{e}^{\theta x}\mathscr{L}(\theta)\sum_{k=0}^{K} a_k Q_k(x) f_\nu(x)\,. \tag{3.8}$$

The coefficients $a_k = \mathbb{E}[Q_k(S_\theta)]$ can be estimated in (at least) three different ways: (i) using CMC, (ii) using Monte Carlo with a change of measure so $\theta \to 0$, or (iii) by directly computing the moments $\mathbb{E}[S_\theta^k]$. The first method is nontrivial, as simulating from f_θ likely requires using acceptance-rejection (as in Asmussen et al. (2015)). Options (ii) and (iii) use

$$a_k = \mathbb{E}[Q_k(S_\theta)] =: q_{k0} + q_{k1}\mathbb{E}[S_\theta] + \cdots + q_{kk}\mathbb{E}[S_\theta^k] \tag{3.9}$$

where $\{q_{ki}\}$ are the coefficients in Q_k, and

$$\mathbb{E}[S_\theta^i] = \frac{\mathbb{E}[S^i \mathrm{e}^{-\theta S}]}{\mathscr{L}(\theta)} =: \frac{\mathscr{L}_i(\theta)}{\mathscr{L}(\theta)}\,.$$

The $\mathscr{L}_i(\theta)$ notation was selected to highlight the link between $\mathbb{E}[S_n^i \mathrm{e}^{-\theta S_n}]$ and the ith derivative of $\mathscr{L}(\theta)$.

All three methods require access to the Laplace transform, and method (iii) requires $\mathscr{L}_i(\theta)$, however none of $\mathscr{L}(\theta)$ or $\mathscr{L}_i(\theta)$ are available in closed form. Our approach to circumvent these problems is presented in the Appendix.

4 Numerical illustrations

We take several approximations $\widehat{f}$ and compare them against the benchmark of numerical integration. One form of f particularly useful for numerical integration, in terms of the $\mathcal{LN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ density $f_{\mathcal{LN}}$, is as a surface integral, $f(s) = n^{-\frac{1}{2}} \int_{\Delta_n^s} f_{\mathcal{LN}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$, where $\Delta_n^s = \{\boldsymbol{x} \in \mathbb{R}_+^n : \|\boldsymbol{x}\|_1 = s\}$. Mathematica integrates this within a reasonable time for $n = 2$ to 4 using `NIntegrate` and `ParametricRegion`). For $n > 4$ we qualitatively assess the performance of the estimators by plotting them.

The quantitative error measure used is the $\mathcal{L}^2$ norm of $(\widehat{f} - f)$ restricted to $(0, \mathbb{E}[S])$. We focus on this region as at one hand it is the hardest to approximate (indeed, Lemma 3.1 shows that just a single lognormal is a theoretically justified approximation of the $\mathcal{SLN}$ right tail) and that at the other of high relevance in applications, see for example the introduction of Asmussen et al. (2015) and the references therein.

4.1 The estimators

We will compare the following approximations:

- the Fenton-Wilkinson approximation $\widehat{f}_{\mathrm{FW}}$, cf. Fenton (1960a), consists in approximating the distribution of S by a single lognormal with the same first and second moment;

- the log skew normal approximation $\widehat{f}_{\mathrm{Sk}}$, cf. Hcine and Bouallegue (2015)[1], is a refinement of Fenton–Wilkinson by using a log skew normal as approximation and fitting the left tail in addition to the first and second moment;
- the conditional Monte Carlo approximation $\widehat{f}_{\mathrm{Cond}}$, cf. Example 4.3 on p. 146 of Asmussen and Glynn (2007), uses the representation $f(x) = \mathbb{E}\big[\mathbb{P}(S \in \mathrm{d}x \,|\, Y)\big]$ for some suitable Y (here chosen as one of the normal r.v.s X_i occurring in (3.1)) and simulates the conditional expectation;
- $\widehat{f}_{\mathcal{N}}$ is the approximation described in Section 3.2 using a logarithmic transformation and the Hermite polynomials with a normal reference distribution;
- $\widehat{f}_{\Gamma}$ is the approximation described in Section 3.3 using exponential tilting and the Laguerre polynomials with a gamma reference distribution.

These approximations are all estimators of functions (i.e., not pointwise estimators, such as in Laub et al. (2016)) and they do not take excessive computational effort to construct. The first two, $\widehat{f}_{\mathrm{FW}}$ and $\widehat{f}_{\mathrm{Sk}}$, only need $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ and do not have any Monte Carlo element. Similarly, the estimator $\widehat{f}_{\Gamma}$ when utilising the Gauss–Hermite quadrature described in (6.2) in the Appendix does not use Monte Carlo. For the remaining approximations we utilise the *common random numbers* technique, meaning that the same $R = 10^5$ i.i.d. $\mathcal{SLN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ samples $\boldsymbol{S} = (S_1, \ldots, S_R)^\top$ are given to each algorithm. Lastly, all the estimators except $\widehat{f}_{\Gamma}$ satisfy $\int \widehat{f}(x)\,\mathrm{d}x = 1$. One problem with the orthogonal polynomial estimators is that they can take negative values; this can easily be fixed, but we do not make that adjustment here.

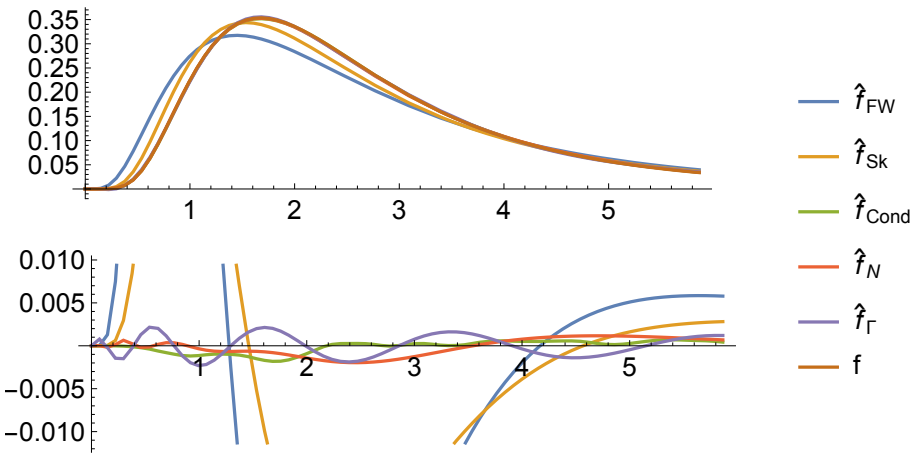
For $\widehat{f}_{\mathcal{N}}$, we take $\mu = \mathbb{E}[Z]$ and $\sigma^2 = \mathbb{V}\mathrm{ar}[Z]$, calculated using numerical integration. The $\widehat{f}_{\Gamma}$ case is more difficult. Equation (3.8) shows that we must impose $\theta m < 1$ to ensure that $\widehat{f}_{\Gamma}(x) \to 0$ as $x \to \infty$. Exploring different parameter selections showed that fixing $\theta = 1$ worked reasonably well. Moment matching f_θ to f_ν leads to the selection of m and r. The moments of f_θ, $\widehat{\mathbb{E}f_\theta} = \widehat{\mathscr{L}_1}(\theta)/\widehat{\mathscr{L}_0}(\theta)$ and $\widehat{\mathbb{V}\mathrm{ar}f_\theta} = \widehat{\mathscr{L}_2}(\theta)/\widehat{\mathscr{L}_0}(\theta) - \widehat{\mathbb{E}f_\theta}^2$ can be approximated using the Gauss–Hermite quadrature of (6.2); for this we use $H = 64$, 32, 16 for $n = 2$, 3, 4 respectively (and CMC for $n > 4$).

With these regimes, parameter selection for the reference distributions is automatic, and the only choice the user must make is in selecting K. In these tests we examined various K from 1 to 40, and show the best approximations found. The source code for these tests is available online at Asmussen et al. (2016), and we invite readers to experiment the effect of modifying K and θ and the parameters of the reference distributions.

4.2 Results

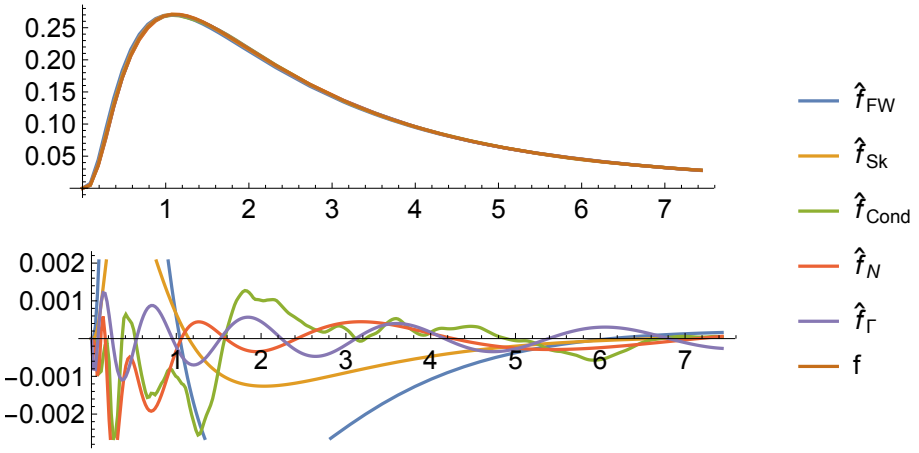
For each test case with $n \leq 4$ we plot the $\widehat{f}(x)$ and $f(x)$ together and then $(\widehat{f}(x) - f(x))$ over $x \in (0, 2\mathbb{E}[S])$. A table then shows the $\mathcal{L}^2$ errors over $(0, \mathbb{E}[S])$.

[1]Note that in Hcine and Bouallegue (2015), the formula for ϵ_{opt} contains an typographic error.



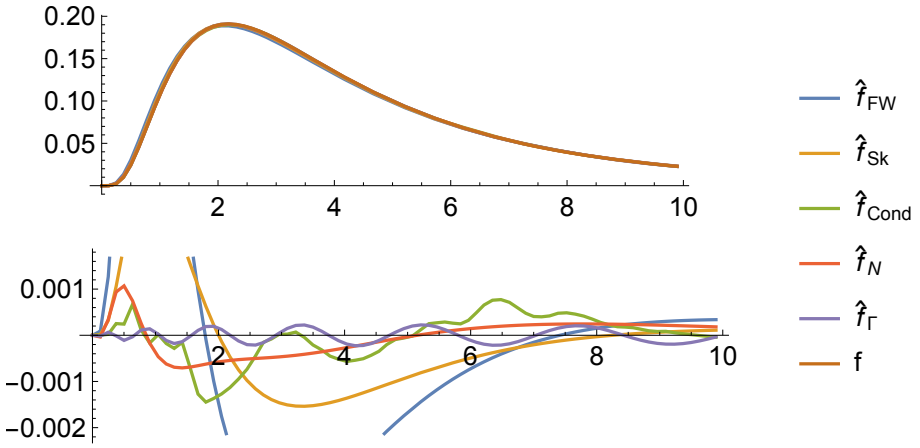
	$\widehat{f}_{\mathrm{FW}}$	$\widehat{f}_{\mathrm{Sk}}$	$\widehat{f}_{\mathrm{Cond}}$	$\widehat{f}_{\mathcal{N}}$	$\widehat{f}_{\Gamma}$
$\mathcal{L}^2$	8.01×10^{-2}	4.00×10^{-2}	1.56×10^{-3}	1.94×10^{-3}	2.28×10^{-3}

Test 1: $\boldsymbol{\mu} = (0, 0)$, $\mathrm{diag}(\boldsymbol{\Sigma}) = (0.5, 1)$, $\rho = -0.2$. Reference distributions used are $\mathcal{N}(0.88, 0.71^2)$ and $\mathrm{Gamma}(2.43, 0.51)$ with $K = 32$, 16 resp.



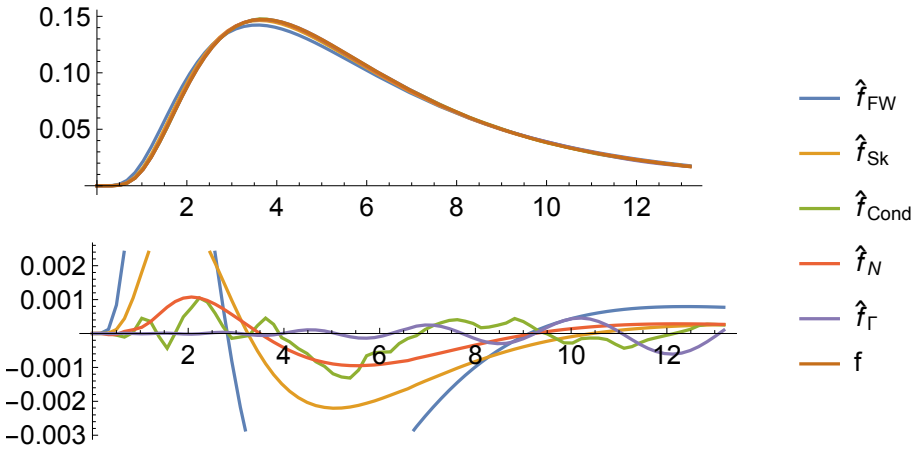
	$\widehat{f}_{\mathrm{FW}}$	$\widehat{f}_{\mathrm{Sk}}$	$\widehat{f}_{\mathrm{Cond}}$	$\widehat{f}_{\mathcal{N}}$	$\widehat{f}_{\Gamma}$
$\mathcal{L}^2$	1.02×10^{-2}	3.49×10^{-3}	1.78×10^{-3}	7.86×10^{-4}	7.24×10^{-4}

Test 2: $\boldsymbol{\mu} = (-0.5, 0.5)$, $\mathrm{diag}(\boldsymbol{\Sigma}) = (1, 1)$, $\rho = 0.5$. Reference distributions used are $\mathcal{N}(0.91, 0.90^2)$ and $\mathrm{Gamma}(2.35, 0.51)$ with $K = 32$, 16 resp.



	$\widehat{f}_{\mathrm{FW}}$	$\widehat{f}_{\mathrm{Sk}}$	$\widehat{f}_{\mathrm{Cond}}$	$\widehat{f}_{\mathcal{N}}$	$\widehat{f}_{\Gamma}$
$\mathcal{L}^2$	9.48×10^{-3}	3.71×10^{-3}	1.60×10^{-3}	1.18×10^{-3}	3.53×10^{-4}

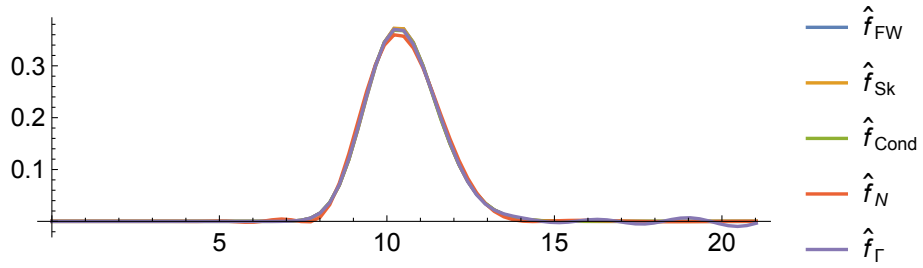
Test 3: $n = 3$, $\mu_i = 0$, $\Sigma_{ii} = 1$, $\rho = 0.25$. Reference distributions used are $\mathcal{N}(1.32, 0.74^2)$ and Gamma$(3, 0.57)$ with $K = 7$, 25 resp.



	$\widehat{f}_{\mathrm{FW}}$	$\widehat{f}_{\mathrm{Sk}}$	$\widehat{f}_{\mathrm{Cond}}$	$\widehat{f}_{\mathcal{N}}$	$\widehat{f}_{\Gamma}$
$\mathcal{L}^2$	1.82×10^{-2}	6.60×10^{-3}	1.90×10^{-3}	1.80×10^{-3}	1.77×10^{-4}

Test 4: $n = 4$, $\mu_i = 0$, $\Sigma_{ii} = 1$, $\rho = 0.1$. Reference distributions used are $\mathcal{N}(1.32, 0.74^2)$ and Gamma$(3.37, 0.51)$ with $K = 18$, 18 resp.

The following test case shows the density approximations for a large n.

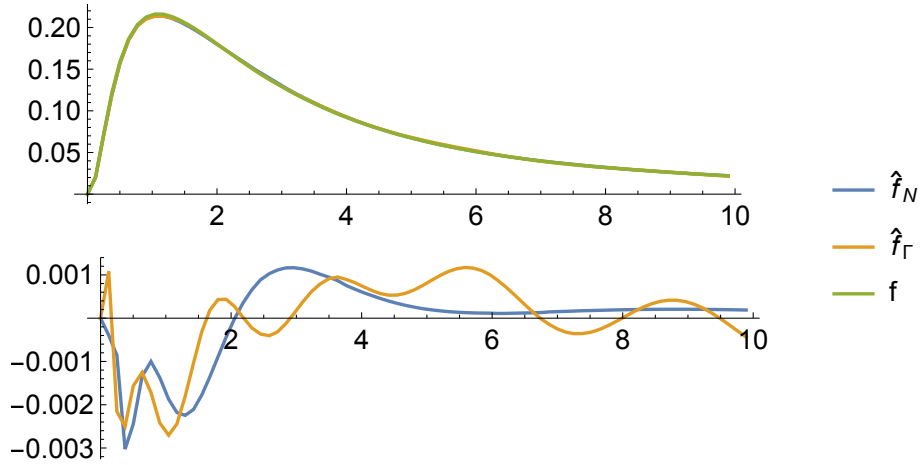


Test 5: Sum of 10 i.i.d. $\mathcal{LN}(0, 0.1)$ r.v.s. Reference distributions used are $\mathcal{N}(2.35, 0.23^2)$ and Gamma$(12.61, 0.25)$ with $K = 18$, 35 resp.

Finally, we fit $\widehat{f}_{\mathcal{N}}$ and $\widehat{f}_{\Gamma}$ to simulated data (10^5 replications) for the sum of lognormals with a non-Gaussian dependence structure. Specifically, we take the sum of $n = 3$ standard lognormal r.v.s with a *Clayton copula*, defined by its distribution function

$$C_\theta^{\mathrm{Cl}}(u_1, \ldots, u_n) = \left(1 - n + \sum_{i=1}^{n} u_i^{-\theta}\right)^{-1/\theta}, \qquad \text{for } \theta > 0\,.$$

The Kendall's tau correlation of the C_θ^{Cl} copula is $\tau = \theta/(\theta+2)$ McNeil et al. (2015).



Test 6: Sum of 3 $\mathcal{LN}(0, 1)$ r.v.s with $C_{10}^{\mathrm{Cl}}(\cdot)$ copula (i.e., $\tau = \frac{5}{6}$). Reference distributions used are $\mathcal{N}(1.46, 0.71^2)$ and Gamma$(8.78, 0.25)$ with $K = 40$. The $\mathcal{L}^2$ errors of $\widehat{f}_{\mathcal{N}}$ and $\widehat{f}_{\Gamma}$ are 2.45×10^{-3} and 2.04×10^{-3} respectively.

Our overall conclusion of the numerical examples is that no single method can be considered as universally superior. Of the methods in the literature, the log skew normal approximations is generally better than Fenton-Wilkinson, which is unsurprising given it is an extension introducing one more parameter. The estimators, $\widehat{f}_{\mathcal{N}}$ and $\widehat{f}_{\Gamma}$, based on orthogonal polynomial approximation techniques, are very flexible. They also display as least as good and sometimes better p.d.f. estimates over the interval $(0, \mathbb{E}[S])$ and their periodic error indicates that they would supply even more accurate c.d.f. estimates. One should note, however, that their performance relies on the tuning of parameters and that somewhat greater effort is involved in

their computation (though this is mitigated through the availability of the software in Asmussen et al. (2016)).

An interesting feature of $\widehat{f}_{\mathcal{N}}$ and $\widehat{f}_{\Gamma}$ is that the Frank copula example indicates some robustness to the dependence structure used. In view of the current interest in financial applications of non-Gaussian dependence this seems a promising line for future research.

Acknowledgements

We are grateful to Jakob Schach Møller for helpful discussions on orthonormal expansions.

5 Proof of Proposition 2.3

Proof. The polynomials orthogonal with respect to the lognormal distribution will be derived using the general formula

$$Q_n(x) = \frac{1}{\sqrt{D_{n-1,n-1}D_{n,n}}} \begin{vmatrix} s_0 & s_1 & \cdots & s_n \\ s_1 & s_2 & \cdots & s_{n+1} \\ \vdots & \vdots & & \vdots \\ s_{n-1} & s_n & \cdots & s_{2n-1} \\ 1 & x & \cdots & x^n \end{vmatrix}, \qquad n \geq 1, \tag{5.1}$$

where $\{s_n\}_{n\in\mathbb{N}_0}$ denotes the moment sequence of the lognormal distribution and $D_{n,n} = |\{s_{k,l}\}_{0\leq k,l\leq n}|$ is a Hankel determinant. The moments of the lognormal distribution are given by $s_n = p^n q^{n^2}$, where $p = \mathrm{e}^{\mu}$ and $q = \mathrm{e}^{\sigma^2/2}$. Consider

$$D_{n,n} = \begin{vmatrix} 1 & pq & \cdots & p^n q^{n^2} \\ pq & p^2q^4 & & p^{n+1}q^{(n+1)^2} \\ \vdots & \vdots & & \vdots \\ p^{n-1}q^{(n-1)^2} & p^n q^{n^2} & \cdots & p^{2n-1}q^{(2n-1)^2} \\ p^n q^{n^2} & p^{n+1}q^{(n+1)^2} & \cdots & p^{2n}q^{(2n)^2} \end{vmatrix}, \qquad n \geq 1, \tag{5.2}$$

and denote by R_k the kth row and by C_ℓ the ℓth column. We apply the elementary operations $R_{k+1} \to p^{-k}q^{-k^2}R_{k+1}$, and $C_{\ell+1} \to p^{-\ell}q^{-\ell^2}C_{\ell+1}$ for $k, \ell = 0, \ldots, n$ to get a Vandermonde type determinant. Thus we have

$$D_{n,n} = \mathrm{e}^{n(n+1)\mu}\mathrm{e}^{\frac{n(n+1)(2n+1)}{3}\sigma^2} \prod_{k=0}^{n-1} \left[\mathrm{e}^{-\sigma^2}; \mathrm{e}^{-\sigma^2}\right]_k \tag{5.3}$$

We expand the determinant in (5.1) with respect to the last row to get

$$Q_n(x) = \frac{1}{\sqrt{D_{n-1,n-1}D_{n,n}}} \sum_{k=0}^{n} (-1)^{n+k} x^k D_{n-1,n}^{-k}, \tag{5.4}$$

where $D_{n-1,n}^{-k}$ is $D_{n,n}$ with the last row and the $(k+1)$th column deleted. We perform on $D_{n-1,n}^{-k}$ the following operations: $R_{j+1} \to p^{-j}q^{-j^2}R_{j+1}$, for $j = 0, \ldots, n-1$, $C_{j+1} \to$

$p^{-j}q^{-j^2}C_{j+1}$, for $j=0,\dots,k-1$, and finally $C_j \to p^{-j}q^{-j^2}C_j$, for $j=k+1,\dots,n$. We obtain

$$D_{n-1,n}^{-k} = p^{n^2-k}q^{\frac{2n^3+n}{3}-k^2}\begin{vmatrix} 1 & \alpha_0 & \cdots & \alpha_0^{k-1} & \alpha_0^{k+1} & \cdots & \alpha_0^n \\ 1 & \alpha_1 & \cdots & \alpha_1^{k-1} & \alpha_1^{k+1} & \cdots & \alpha_1^{n+1} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 1 & \alpha_{n-1} & \cdots & \cdots & \cdots & \cdots & \alpha_{n-1}^n \end{vmatrix},$$

where $\alpha_k = q^{2k}$, for $k=0,\dots,n-1$. Expanding the polynomial $B(X)=\prod_{i=0}^{n-1}(x-\alpha_i)$, we get

$$B(x) = x^n + \beta_{n-1}x^{n-1} + \cdots + \beta_0,$$

where $\beta_k = (-1)^{n-k}e_{n-k}(\alpha_0,\dots,\alpha_{n-1})$, and $e_k(X_1,\dots,X_n)$ denotes the elementary symmetric polynomial, defined previously in (2.11). We apply the elementary operation $C_n \to C_n + \sum_{j=0}^{k-1} a_j C_{j+1} + \sum_{j=k+1}^{n-1} a_j C_j$, followed by $n-k$ cyclic permutations to get

$$D_{n-1,n}^{-k} = p^{n-k}q^{n^2-k^2}e_{n-k}\big(1,\dots,q^{2(n-1)}\big)D_{n-1,n-1}. \tag{5.5}$$

Inserting (5.3) and (5.5) into (5.4) leads to (2.10). □

6 Computing the coefficients of the expansion $\{a_k\}_{k\in\mathbb{N}_0}$ in the gamma case

We extend here the techniques developed in Laub et al. (2016) to construct an approximation for $\mathscr{L}_i(\theta)$. We note that $\mathscr{L}_i(\theta) \propto \int_{\mathbb{R}^n} \exp\{-h_{\theta,i}(\boldsymbol{x})\}\,\mathrm{d}\boldsymbol{x}$ where

$$h_{\theta,i}(\boldsymbol{x}) = -i\ln(\mathbf{1}^\top \mathrm{e}^{\boldsymbol{\mu}+\boldsymbol{x}}) + \theta\mathbf{1}^\top \mathrm{e}^{\boldsymbol{\mu}+\boldsymbol{x}} + \frac{1}{2}\boldsymbol{x}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{x}\,, \qquad i\in\mathbb{N}_0.$$

This uses the notation $\mathbf{e}^{\boldsymbol{x}} = (\mathrm{e}^{x_1},\dots,\mathrm{e}^{x_n})^\top$. Next, define $\boldsymbol{x}^*$ as the minimiser of $h_{\theta,i}$ (calculated numerically), and consider a second order Taylor expansion of $h_{\theta,i}$ about $\boldsymbol{x}^*$. Denote $\widetilde{\mathscr{L}_i}(\theta)$ as the approximation where $h_{\theta,i}$ is replaced by this Taylor expansion in $\mathscr{L}_i(\theta)$. Simplifying yields

$$\widetilde{\mathscr{L}_i}(\theta) = \frac{\exp\{-h_{\theta,i}(\boldsymbol{x}^*)\}}{\sqrt{|\boldsymbol{\Sigma}\boldsymbol{H}|}} \tag{6.1}$$

where $\boldsymbol{H}$, the Hessian of $h_{\theta,i}$ evaluated at $\boldsymbol{x}^*$, is

$$\boldsymbol{H} = i\frac{\mathbf{e}^{\boldsymbol{\mu}+\boldsymbol{x}^*}(\mathbf{e}^{\boldsymbol{\mu}+\boldsymbol{x}^*})^\top}{(\mathbf{1}^\top\mathbf{e}^{\boldsymbol{\mu}+\boldsymbol{x}^*})^2} + \boldsymbol{\Sigma}^{-1} - \mathrm{diag}(\boldsymbol{\Sigma}^{-1}\boldsymbol{x}^*)\,.$$

As $\theta\to\infty$ we have $\widetilde{\mathscr{L}_i}(\theta)\to\mathscr{L}_i(\theta)$. We can rewrite $\mathscr{L}_i(\theta) = \widetilde{\mathscr{L}_i}(\theta)I_i(\theta)$ and estimate $I_i(\theta)$, as in Laub et al. (2016).

Proposition 6.1. *The moments of the exponentially-tilted distribution $\mathcal{SLN}_\theta(\boldsymbol{\mu},\boldsymbol{\Sigma})$, denoted $\mathscr{L}_i(\theta)$, can be written as $\mathscr{L}_i(\theta) = \widetilde{\mathscr{L}_i}(\theta)I_i(\theta)$ where $\widetilde{\mathscr{L}_i}(\theta)$ is in* (6.1) *and*

$$I_i(\theta) = \sqrt{|\boldsymbol{\Sigma}\boldsymbol{H}|}\,v(\mathbf{0})^{-1}\,\mathbb{E}[v(\boldsymbol{\Sigma}^{\frac{1}{2}}Z)]$$

where $Z\sim\mathcal{N}(\mathbf{0},\boldsymbol{I})$, and

$$v(\boldsymbol{z}) = \exp\{i\ln(\mathbf{1}^\top\mathbf{e}^{\boldsymbol{\mu}+\boldsymbol{x}^*+\boldsymbol{z}}) - \theta\mathbf{1}^\top\mathbf{e}^{\boldsymbol{\mu}+\boldsymbol{x}^*+\boldsymbol{z}} - (\boldsymbol{x}^*)^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{z}\}\,.$$

Proof. We begin by substituting $\boldsymbol{x} = \boldsymbol{x}^* + \boldsymbol{H}^{-\frac{1}{2}}\boldsymbol{y}$ into $\mathscr{L}_i(\theta)$, then multiply by $\exp\{\pm \text{ some constants}\}$:

$$\begin{aligned}
\mathscr{L}_i(\theta) &= \int_{\mathbb{R}^n} \frac{(2\pi)^{-\frac{n}{2}}}{\sqrt{|\boldsymbol{\Sigma}|}} \exp\{i \log(\mathbf{1}^\top \mathrm{e}^{\boldsymbol{\mu}+\boldsymbol{x}}) - \theta \mathbf{1}^\top \mathrm{e}^{\boldsymbol{\mu}+\boldsymbol{x}} - \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{x}\} \,\mathrm{d}\boldsymbol{x} \\
&= \int_{\mathbb{R}^n} \frac{(2\pi)^{-\frac{n}{2}}}{\sqrt{|\boldsymbol{\Sigma}\boldsymbol{H}|}} \exp\{i \log(\mathbf{1}^\top \mathrm{e}^{\boldsymbol{\mu}+\boldsymbol{x}^*+\boldsymbol{H}^{-\frac{1}{2}}\boldsymbol{y}}) - \theta \mathbf{1}^\top \mathrm{e}^{\boldsymbol{\mu}+\boldsymbol{x}^*+\boldsymbol{H}^{-\frac{1}{2}}\boldsymbol{y}} \\
&\qquad - \frac{1}{2}(\boldsymbol{x}^* + \boldsymbol{H}^{-\frac{1}{2}}\boldsymbol{y})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}^* + \boldsymbol{H}^{-\frac{1}{2}}\boldsymbol{y})\} \,\mathrm{d}\boldsymbol{y} \\
&= \widetilde{\mathscr{L}_i}(\theta) \exp\{-i \log(\mathbf{1}^\top \mathrm{e}^{\boldsymbol{\mu}+\boldsymbol{x}^*}) + \theta \mathbf{1}^\top \mathrm{e}^{\boldsymbol{\mu}+\boldsymbol{x}^*}\} \\
&\qquad \times \int_{\mathbb{R}^n} (2\pi)^{-\frac{n}{2}} \exp\{i \log(\mathbf{1}^\top \mathrm{e}^{\boldsymbol{\mu}+\boldsymbol{x}^*+\boldsymbol{H}^{-\frac{1}{2}}\boldsymbol{y}}) - \theta \mathbf{1}^\top \mathrm{e}^{\boldsymbol{\mu}+\boldsymbol{x}^*+\boldsymbol{H}^{-\frac{1}{2}}\boldsymbol{y}} \\
&\qquad\qquad - (\boldsymbol{x}^*)^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{H}^{-\frac{1}{2}}\boldsymbol{y} - \frac{1}{2}\boldsymbol{y}^\top (\boldsymbol{\Sigma}\boldsymbol{H})^{-1}\boldsymbol{y}\} \,\mathrm{d}\boldsymbol{y}\,.
\end{aligned}$$

That is, $\mathscr{L}_i(\theta) = \widetilde{\mathscr{L}_i}(\theta) I_i(\theta)$. In $I_i(\theta)$, take the change of variable $\boldsymbol{y} = (\boldsymbol{\Sigma}\boldsymbol{H})^{\frac{1}{2}}\boldsymbol{z}$, and the result follows. □

Remark 6.2. The form of $I_i(\theta)$ naturally suggests evaluation using *Gauss–Hermite* quadrature:

$$\widehat{\mathscr{L}_i}(\theta) = \frac{\exp\{-h_{\theta,i}(\boldsymbol{x}^*)\}}{v(\mathbf{0})\,\pi^{\,n/2}} \sum_{i_1=1}^{H} \cdots \sum_{i_n=1}^{H} v(\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{z}) \prod_{j=1}^{n} w_{i_j} \tag{6.2}$$

where $\boldsymbol{z} = (z_{i_1}, \ldots, z_{i_n})^\top$, the set of weights and nodes $\{(w_i, z_i) : 1 \le i \le H\}$ is specified by the Gauss–Hermite quadrature algorithm, and $H \ge 1$ is the order of the approximation. This approximation is accurate, especially so when the i in $\mathscr{L}_i$ becomes large. Even for $\mathscr{L}$ $(= \mathscr{L}_0)$ this method appears to outperform the quasi-Monte Carlo scheme outlined in Laub et al. (2016). ◊

Thus, with $\widehat{\mathscr{L}_i}(\theta)$ given in (6.2), we can now estimate the coefficients. The three methods correspond to

1. $\widehat{a}_k = R^{-1} \sum_{r=1}^{R} Q_k(S_r)$, for $S_1, \ldots, S_R \overset{\text{i.i.d.}}{\sim} f_\theta(x)$,
2. $\widehat{a}_k = \sum_{j=0}^{k} q_{kj}\, \widehat{\mathbb{E}[S_\theta^j]} = q_{k0} + (R\,\widehat{\mathscr{L}}(\theta))^{-1} \sum_{j=1}^{k} q_{kj} \sum_{r=1}^{R} S_r^j \mathrm{e}^{-\theta S_r}$, from (3.9), where $S_1, \ldots, S_R \overset{\text{i.i.d.}}{\sim} f(x)$,
3. $\widehat{a}_k = q_{k0} + \widehat{\mathscr{L}}(\theta)^{-1} \sum_{j=1}^{k} q_{kj}\, \widehat{\mathscr{L}_j}(\theta)$.

In the numerical illustrations, we switched between using methods (2) and (3) for large and small n respectively. Algorithms for efficient simulation from f_θ is work in progress.

References

J. Abate, G.L. Choudhury, and W. Whitt. On the Laguerre method for numerically inverting Laplace transforms. *INFORMS Journal on Computing*, 8(4):413–427, 1995.

John Aitchison and James AC Brown. *The Lognormal Distribution with Special Reference to its Uses in Economics*. Cambridge University Press, 1957.

S. Asmussen and L. Rojas-Nandapaya. Asymptotics of sums of lognormal random variables with Gaussian copula. *Statistics and Probability Letters*, 78(16):2709–2714, 2008.

Søren Asmussen and Peter W Glynn. *Stochastic Simulation: Algorithms and Analysis*, volume 57 of *Stochastic Modelling and Applied Probability series*. Springer, 2007.

Søren Asmussen, Jens Ledet Jensen, and Leonardo Rojas-Nandayapa. Exponential family techniques in the lognormal left tail. *Scandinavian Journal of Statistics*, 2015. To appear.

Søren Asmussen, Pierre-Olivier Goffard, and Patrick J. Laub. *Online accompaniment for "Orthonormal polynomial expansions and lognormal sum densities"*, 2016. Available at `https://github.com/Pat-Laub/SLNOrthogonalPolynomials`.

Ole E Barndorff-Nielsen and David Roxbee Cox. *Asymptotic Techniques for use in Statistics*. Chapman & Hall, 1989.

Christian Berg, Jens Peter Reus Christensen, and Paul Ressel. *Harmonic Analysis on Semigroups*, volume 100 of *Graduate Texts in Mathematics*. Springer–Verlag, 1984.

T. S. Chihara. On generalized Stieljes-Wigert and related orthogonal polynomials. *Journal of Computational and Applied Mathematics*, 5(4):291–297, 1979.

J. S. Christiansen. The moment problem associated with the Stieljes-Wigert polynomials. *Journal of Mathematical Analysis and Applications*, 277(1):218–245, 2003.

H. Cramer. *Mathematical Methods of Statistics*, volume 9. Princeton University Press, 1999.

Edwin L Crow and Kunio Shimizu. *Lognormal Distributions: Theory and Applications*. Dekker New York, 1988.

Daniel Dufresne. Sums of lognormals. Technical report, Centre for Actuarial Sciences, University of Melbourne, 2009.

P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modeling Extremal Events for Insurance and Finance*. Springer-Verlag, 1997.

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

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

X. Gao, H. Xu, and D. Ye. Asymptotic behavior of tail density for sum of correlated lognormal variables. *International Journal of Mathematics and Mathematical Sciences*, 2009.

Paul Glasserman. *Monte Carlo Methods in Financial Engineering*, volume 53 of *Stochastic Modelling and Applied Probability series*. Springer, 2003.

Pierre-Olivier Goffard, Stéphane Loisel, and Denys Pommeret. Polynomial approximations for bivariate aggregate claim amount probability distributions. *Methodology and Computing in Applied Probability*, pages 1–24, 2015.

Pierre-Olivier Goffard, Stéphane Loisel, and Denys Pommeret. A polynomial expansion to approximate the ultimate ruin probability in the compound poisson ruin model. *Journal of Computational and Applied Mathematics*, 296:499–511, 2016.

John A Gubner. *Probability and Random Processes for Electrical and Computer Engineers*. Cambridge University Press, 2006.

A. Gulisashvili and P. Tankov. Tail behavior of sums and differences of log-normal random variables. *ArXiv preprint arXiv:1309.3057*, 2013.

Archil Gulisashvili and Peter Tankov. Tail behavior of sums and differences of log-normal random variables. *Bernoulli*, 2015. To appear, accessed online on 26th August 2015 at `http://www.e-publications.org/ims/submission/BEJ/user/submissionFile/17119?confirm=ef609013`.

Marwane Ben Hcine and Ridha Bouallegue. Highly accurate log skew normal approximation to the sum of correlated lognormals. *In the Proc. of NeTCoM 2014*, 2015.

C.C. Heyde. On a property of the lognormal distribution. *J. Roy. Statist.*, 5(Ser. B): 392–393, 1963.

Norman L Johnson, Samuel Kotz, and Narayanaswamy Balakrishnan. *Continuous Univariate Distributions*, volume 1. New York: Wiley, 2 edition, 1994.

Patrick J. Laub, Søren Asmussen, Jens L. Jensen, and Leonardo Rojas-Nandayapa. Approximating the Laplace transform of the sum of dependent lognormals. *Volume 48A of Advances in Applied Probability.*, 2016. In Festschrift for Nick Bingham, C. M. Goldie and A. Mijatovic (Eds.), Probability, Analysis and Number Theory. To appear.

Eckhard Limpert, Werner A Stahel, and Markus Abbt. Log-normal distributions across the sciences: Keys and clues. *BioScience*, 51(5):341–352, 2001.

NA Marlow. A normal limit theorem for power sums of independent random variables. *Bell System Technical Journal*, 46(9):2081–2089, 1967.

Alexander J McNeil, Rüdiger Frey, and Paul Embrechts. *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, 2nd edition, 2015.

B. S. Nagy. *Introduction to Real Functions and Orthogonal Expansions*. Akadémiai Kiadó, 1965.

G. Szegö. *Orthogonal Polynomials*, volume XXIII. American Mathematical Society Colloquium Publications, 1939.

Olof Thorin and Nils Wikstad. Calculation of ruin probabilities when the claim distribution is lognormal. *Astin Bulletin*, 9(1-2):231–246, 1977.

S. Wigert. Sur les polynomes orthogonaux et l'approximation des fonctions continues. *Almqvist and Wiksell*, 1923.