Bernoulli **22**(1), 2016, 444–493
DOI: 10.3150/14-BEJ665

Tail behavior of sums and differences of log-normal random variables

ARCHIL GULISASHVILI[1] and PETER TANKOV[2]

This article is dedicated to the memory of Peter Laurence.

[1]*Department of Mathematics, Ohio University, USA. E-mail: gulisash@ohio.edu*
[2]*Laboratoire de Probabilités et Modèles Aléatoires, Université Paris Diderot – Paris 7, France. E-mail: tankov@math.univ-paris-diderot.fr*

We present sharp tail asymptotics for the density and the distribution function of linear combinations of correlated log-normal random variables, that is, exponentials of components of a correlated Gaussian vector. The asymptotic behavior turns out to depend on the correlation between the components, and the explicit solution is found by solving a tractable quadratic optimization problem. These results can be used either to approximate the probability of tail events directly, or to construct variance reduction procedures to estimate these probabilities by Monte Carlo methods. In particular, we propose an efficient importance sampling estimator for the left tail of the distribution function of the sum of log-normal variables. As a corollary of the tail asymptotics, we compute the asymptotics of the conditional law of a Gaussian random vector given a linear combination of exponentials of its components. In risk management applications, this finding can be used for the systematic construction of stress tests, which the financial institutions are required to conduct by the regulators. We also characterize the asymptotic behavior of the Value at Risk for log-normal portfolios in the case where the confidence level tends to one.

Keywords: importance sampling; Laplace's method; Monte Carlo method; multidimensional Black–Scholes model; multidimensional log-normal distribution; risk management; stress testing; tail-behavior

1. Introduction

The multivariate log-normal distribution is a widely used stochastic model in natural and social sciences, and linear combinations of correlated log-normal random variables arise in many applications. For example, in wireless communications, the distribution of the total interference power coming from several sources is often described by a sum of log-normal variables, and in financial risk management, a linear combination of correlated log-normal variables may represent the value of a portfolio of assets. Since the distribution of such linear combinations is not known in explicit form, a considerable effort has been devoted to developing asymptotic approximations. In particular, Asmussen and Rojas-Nandayapa [6] characterized the behavior of the right tail of the distribution function of the sum of correlated log-normals. Their results can also be deduced from the more recent studies of the tail behavior of sums of dependent subexponential random variables [19,21]. On the other hand, Gao, Xu and Ye [20] computed the asymptotics of the left tail of the sum of *two* correlated log-normal variables. However, beyond these two cases, the tail behavior of linear combinations of log-normal variables is not well understood so far.

1350-7265

In this paper, we present an explicit characterization of the tail asymptotics of the density and the distribution function of arbitrary linear combinations of correlated log-normal random variables. We find new dependence patterns, very different from those which have been established for the right tail of the log-normal sum and more generally for the right tail of the sum of subexponential random variables. The principle of a "single big jump" does not hold: the asymptotic behavior is no longer determined by the single component with the fattest tail but depends on the correlation between the components.

Our paper contains two types of results. Firstly, we compute the tail asymptotics of the distribution function and the density of a linear combination of log-normal variables. These results can be used either to estimate the probability of tail events directly, or to construct efficient variance reduction procedures to estimate these probabilities by Monte Carlo method. In particular, we propose an importance sampling estimator for the left tail of the distribution function of the log-normal sum, which is *logarithmically efficient* in the sense of Asmussen and Glynn [5]. In risk management applications, our asymptotic formulas can be used to evaluate the probability of large portfolio losses within the multidimensional Black–Scholes model. Secondly, as a corollary of the tail asymptotics, we compute the asymptotic law of a Gaussian vector conditional on a linear combination of exponentials of its components. This finding can be used for the systematic construction of stress tests, which the financial institutions are required to conduct by the regulators.

In the present paper we focus on the multidimensional log-normal distribution in view of its importance for applications. However, we expect that our findings and the techniques we develop will stimulate further studies of the multidimensional stochastic models and settings in which the tail behavior is determined by the entire dependence structure rather than by a single component.

Review of relevant literature

The history and the applications of log-normal distributions are reviewed in [1,15,16,22,26]. Sums and integrals of log-normal variables and processes also play an important role in theoretical probability and theoretical physics in relation to the Gaussian multiplicative chaos [23,28]. A considerable effort has been devoted to the numerical approximations of the distribution function of the sum of log-normal variables. In [9,10], the authors find approximations to the density of the sum of log-normals based on approximations of the characteristic function of the univariate log-normal distribution. A similar path is taken in the papers [30] and [33] of Senarante and Tellambura to develop deterministic numerical techniques for the computation of the distribution function of the log-normal sum.

A related approach is to bound the density of the log-normal sum from above and below with some more or less easily computable expressions. Tellambura [32] provides bounds for the distribution function of a sum of 2 or 3 correlated log-normals, and also for the sum of any number of equally-correlated log-normals. The paper [34] of Vanduffel *et al.* is devoted to approximations of the distribution function of the log-normal sum by the distribution function of the conditional expectation of such a sum with respect to an auxiliary conditioning random variable. Another stream of literature discusses bounds for tails of functions of general random vectors with fixed marginals (see [17] and references therein).

Motivated, in particular, by the applications in risk management, several authors have studied the tail behavior of sums of log-normal variables. As already mentioned, Asmussen and Rojas-Nandayapa [6] (see also the dissertation of Rojas-Nandayapa [29]) characterized the behavior of the right tail of the distribution function of the sum of log-normals. These results were used in Asmussen *et al.* [4], to construct importance sampling Monte Carlo estimators for the distribution function of the sum of log-normals in the right tail. Understanding the left tail of the log-normal sum turned out to be considerably more difficult. Szyszkowicz and Yanikomeroglu [31] propose to approximate the left tail of the sum of uncorrelated log-normal variables by a one-dimensional log-normal distribution function. More recently, an important progress was made in the article by Gao, Xu and Ye [20] where explicit asymptotics of the left tail of the distribution of the sum of two correlated log-normals are presented. For a subclass of covariance matrices (see Remark 5 below) these authors also characterize the asymptotic behavior of the left tail of the density of the sum of an arbitrary number of log-normal variables.

The vast majority of publications discuss the sum of log-normal variables. Linear combinations of such variables with coefficients of different signs have received relatively little attention despite their importance for applications, for instance, to spread option pricing in finance (see Carmona and Durrleman [14]). One of the few exceptions is the paper by Lo [27] who considers the distribution of the sum and the difference of two log-normal processes (geometric Brownian motions) and presents a small-time approximation to these distributions. Small-time and small-noise asymptotics of sums and differences of geometric Brownian motions are also discussed in several papers dealing with basket and spread option pricing [7,11,12].

The log-normal distribution is an example of a subexponential distribution (see [18] for the definition of subexponentiality). Numerous publications were devoted to tail estimates for sums or more general functions of dependent sub-exponential random variables (see [3,19,21,24,25, 35] and the references therein). The right-tail behavior of the difference of a positive subexponential random variable and a dependent positive random variable is studied in [2]. Our paper focuses on the cases which cannot be dealt with using the theory of subexponential distributions, such as the *left* tail of the sum of log-normals, or the right tail of the weighted sum with weights of different signs.

Main notation

Throughout the paper, we use boldface for denoting vectors. In particular, $\mathbf{1}$ denotes the vector of suitable dimension, all components of which are equal to 1. A strictly positive random vector $\mathbf{X} = (X_1, \dots, X_n)$ such that the vector $\mathbf{Y} = (Y_1, \dots, Y_n)$ with $Y_i = \log X_i$, $1 \le i \le n$, has an n-dimensional normal distribution with mean vector $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)$ and covariance matrix $\mathfrak{B}$, is called an n-dimensional log-normal vector with parameters $\boldsymbol{\mu}$ and $\mathfrak{B}$. The elements of the matrix $\mathfrak{B}$ will be denoted by b_{ij}. The distribution of $\mathbf{X}$ is called the n-dimensional log-normal distribution and denoted by $\Lambda(\boldsymbol{\mu}, \mathfrak{B})$. In the present paper, we make the standing assumption that $|\mathfrak{B}| > 0$. The inverse matrix of the covariance matrix will be denoted by $\mathfrak{B}^{-1}$, its elements will be denoted by a_{ij}, and we put $A_k = \sum_{j=1}^{n} a_{kj}$, $1 \le k \le n$. The log-normal distribution $\Lambda(\boldsymbol{\mu}, \mathfrak{B})$

admits a density defined by

$$d^{\log}(x_1, \dots, x_n) = \frac{1}{(2\pi)^{n/2}\sqrt{|\mathfrak{B}|}x_1 \cdots x_n} \exp\left\{-\frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(\log x_i - \mu_i)(\log x_j - \mu_j)\right\}, \tag{1}$$

where $x_i > 0$ for $1 \le i \le n$. In particular, the one-dimensional log-normal density with mean $\mu \in \mathbb{R}$ and variance σ^2 is given by

$$d^{\log}(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left\{-\frac{1}{2\sigma^2}(\log x - \mu)^2\right\}, \qquad x > 0. \tag{2}$$

For every integer m with $1 \le m \le n$, we consider the random variable

$$X^{(m)} = \sum_{k=1}^{m} e^{Y_k} - \sum_{k=m+1}^{n} e^{Y_k}. \tag{3}$$

The support of $X^{(m)}$ is equal to $\mathbb{R}$ for $m = 1, \dots, n-1$, and to $\mathbb{R}_+$ for $m = n$. For $m = n$, the variable $X^{(n)}$ is denoted simply by X. The symbols $p^{(m)}$ and p will stand for the density of $X^{(m)}$ and X, respectively.

Our main goal in this paper is to characterize the tail behavior of the distribution of the random variable $X^{(m)}$. We are mainly interested in the asymptotic behavior of the right tail of the variables $X^{(m)}$, $1 \le m \le n-1$, that is, the behavior of $\mathbb{P}[X^{(m)} > x]$ and $p^{(m)}(x)$ as $x \to \infty$, as well as the behavior of the left tail of X (as $x \to 0$). The right tail behavior of the distribution function of X was completely characterized in [6], while in [20], a similar characterization was obtained for the density. The left tail behavior of $X^{(m)}$, $1 \le m \le n-1$, can be deduced from that of the right tail by exchanging the signs of the variables. Since positive coefficients can be incorporated into the mean vector of Y, our results provide a complete characterization of the tail behavior of a linear combination $\sum_{i=1}^{n} \lambda_i e^{Y_i}$, $\lambda_i \in \mathbb{R}$, of components of a log-normal random vector $(e^{Y_1}, \dots, e^{Y_n})$.

Overview of the paper

Section 2 deals with the left tail asymptotics of sums of log-normal variables. This section is split into Section 2.1, where we formulate and discuss some of our main results, and Section 2.2, which contains the proofs.

In Section 2.1, we formulate asymptotic formulas for the distribution function and the distribution density in the general case, under rather mild nondegeneracy conditions (Theorem 1 and Corollary 2). This is done by relating the tail asymptotics to the quadratic optimization problem

$$\min_{\mathbf{w} \in \Delta_n} \mathbf{w}^{\perp} \mathfrak{B} \mathbf{w}, \tag{4}$$

where Δ_n is the set of vectors in $\mathbb{R}^n$ whose components are all non-negative and sum up to one. In particular, the leading term in the asymptotics for both the distribution function and the density is given by

$$\exp\left\{-\frac{\log^2 x}{2\min_{\mathbf{w}\in\Delta_n}\mathbf{w}^\perp\mathfrak{B}\mathbf{w}}\right\}. \tag{5}$$

This is in sharp contrast with the findings of [6] for the asymptotics of the *right tail* of the sum of log-normals, where the leading term is

$$\exp\left\{-\frac{\log^2 x}{2\max_{i=1,\dots,n}\mathfrak{B}_{ii}}\right\}.$$

As an application of Theorem 1 and Corollary 2, we characterize the asymptotic behavior of conditional Laplace transforms of multidimensional Gaussian vectors (see Corollaries 1 and 3). These estimates are used in Section 5, which deals with stress testing of log-normal portfolios. Section 2.2 contains the proofs of the results formulated in Section 2.1. We first establish asymptotic formulas for the distribution function and the density in the special case (see Lemma 1), when the row sums of the inverse covariance matrix are all strictly positive. Here we can apply Laplace's method to the integral, characterizing the distribution density of the log-normal sum. Lemma 1 is used in the proof of Theorem 1.

In Section 3, the results obtained in Section 2 are extended to the case of the difference of two log-normal sums (see the random variable in (3)). These extensions are not trivial, and they are new even in simple cases. We find sharp asymptotic formulas for the right tails of distributions of such differences. The proofs of the results obtained in Section 3 are similar to those in Section 2, but the details are more complicated. In Section 3.1, we formulate several assertions concerning the tail asymptotics of log-normal differences. Theorem 2 characterizes the asymptotic behavior of the right tail of a log-normal difference. In addition, the asymptotic behavior of the conditional Laplace transform is characterized in Corollary 4. Section 3.2 is devoted to the proofs of these results. Here we start with a special case, where Laplace's method can be applied directly (see Lemma 2), and reduce the general case to the special one using quadratic programming methods.

In Section 4, we analyze the performance of our asymptotic formulas via numerical examples, by comparing the theoretical results with Monte Carlo computations. The convergence turns out to be quite slow, which is consistent with logarithmic error bounds in our main results. However, the asymptotic formulas provide a good order of magnitude approximation for a wide range of values of x. This fact enables us to design an importance sampling technique for evaluation of the tail event probabilities by Monte Carlo method, which is logarithmically efficient in the sense of Asmussen and Glynn [5].

The last part of the paper (Section 5) considers applications of our asymptotic formulas to risk management in the context of the multidimensional Black–Scholes model. This model, which represents stock prices as exponentials of correlated Brownian motions, remains widely used for the analysis of large portfolios. Our asymptotic theory provides two types of insights. First, it allows to quantify the tail behavior of portfolios of log-normal stocks. For example, for portfolios with positive weights, the leading term of the probability of a large downside move is given by (5). This means that (4) measures the risk of the portfolio in a downturn. Second, it provides better understanding of the behavior of individual assets under various adverse scenarios. For instance, we consider a typical stress scenario when the normalized value of a benchmark portfolio (or index) drops to x with x small. Theorem 4 characterizes the asymptotic behavior of conditional expectations of the individual assets in the original portfolio under such an adverse scenario, when the benchmark portfolio has only positive weights. This theorem shows that the assets in a market can be categorized into two classes: those assets for which conditional expectations decay proportionally to x as $x \to 0$ (safe assets), and those assets, for which conditional expectations decay much faster than x (dangerous assets). The safe assets are exactly those which have strictly positive weights in the solution to the quadratic programming problem (4). Results such as Theorem 4 may be employed for systematic construction of stress tests, which banks and investment firms are required to conduct by the regulatory bodies. Finally, in Section 5.2, we characterize the asymptotic behavior of the Value at Risk for a log-normal portfolio as the confidence level tends to one (see Theorem 6).

2. Asymptotic behavior of the left tail of a log-normal sum

The present section studies the left tail asymptotics of the random variable $X = \sum_{i=1}^{n} \mathrm{e}^{Y_i}$.

2.1. Left tail of a log-normal sum. Results and discussions

Denote by Δ_n the n-dimensional simplex defined by

$$\Delta_n := \left\{ \mathbf{w} \in \mathbb{R}^n \colon\ w_i \geq 0, i = 1, \dots, n, \text{ and } \sum_{i=1}^{n} w_i = 1 \right\}$$

and let $\bar{\mathbf{w}} \in \Delta_n$ be the unique vector such that

$$\bar{\mathbf{w}}^{\perp} \mathfrak{B} \bar{\mathbf{w}} = \min_{\mathbf{w} \in \Delta_n} \mathbf{w}^{\perp} \mathfrak{B} \mathbf{w}. \tag{6}$$

The existence and uniqueness of $\bar{\mathbf{w}}$ follows from the non-degeneracy of the matrix $\mathfrak{B}$. In the case where $A_k > 0$ for $k = 1, \dots, n$,

$$\bar{\mathbf{w}} = \frac{\mathfrak{B}^{-1}\mathbf{1}}{\mathbf{1}^{\perp}\mathfrak{B}^{-1}\mathbf{1}} \qquad \Longleftrightarrow \qquad \bar{w}_k = \frac{A_k}{\sum_{i=1}^{n} A_i}, \qquad k = 1, \dots, n, \tag{7}$$

which means that $\bar{w}_k > 0$ for $k = 1, \dots, n$. In the general case, we let

$$\begin{aligned} &\bar{n} := \operatorname{Card}\{i = 1, \dots, n \colon\ \bar{w}_i \neq 0\}, \\ &\bar{I} := \{i = 1, \dots, n \colon\ \bar{w}_i \neq 0\} := \{\bar{k}(1), \dots, \bar{k}(\bar{n})\}, \end{aligned} \tag{8}$$

$\bar{\boldsymbol{\mu}} \in \mathbb{R}^{\bar{n}}$ with $\bar{\mu}_i = \mu_{\bar{k}(i)}$, and $\bar{\mathfrak{B}} \in M_{\bar{n}}(\mathbb{R})$ with $\bar{\mathfrak{B}}_{ij} = \mathfrak{B}_{\bar{k}(i),\bar{k}(j)}$. The inverse matrix of $\bar{\mathfrak{B}}$ is denoted by $\bar{\mathfrak{B}}^{-1}$ and its elements and row sums by $\bar{a}_{ij}$ and $\bar{A}_k := \sum_{j=1}^{\bar{n}} \bar{a}_{kj}$.

In the present subsection, we are only dealing with the sum of the exponentials of the random variables $Y_1, \dots, Y_n$. Since these variables are exchangeable, we can assume with no loss of generality that for the covariance matrix $\mathfrak{B}$, $\bar{I} = \{1, \dots, \bar{n}\}$ with $\bar{n} \le n$. By the strict convexity of the objective function, the minimizer of $\min_{\mathbf{w} \in \Delta_{\bar{n}}} \mathbf{w}^{\perp} \bar{\mathfrak{B}} \mathbf{w}$ coincides with the first $\bar{n}$ components of $\bar{\mathbf{w}}$ and therefore belongs to the interior of the set $\mathbb{R}^{\bar{n}}_{+}$. The minimizer over $\Delta_{\bar{n}}$ then coincides with the minimizer over the set $\{\mathbf{w} \in \mathbb{R}^{\bar{n}}\colon \sum_{i=1}^{\bar{n}} w_i = 1\}$, which means that

$$(\bar{w}_i)_{i=1,\dots,\bar{n}} = \frac{\bar{\mathfrak{B}}^{-1}\mathbf{1}}{\mathbf{1}^{\perp}\bar{\mathfrak{B}}^{-1}\mathbf{1}},$$

or, equivalently,

$$\bar{w}_k = \frac{\bar{A}_k}{\sum_{i=1}^{\bar{n}} \bar{A}_i}, \qquad k = 1, \dots, \bar{n}. \tag{9}$$

Since $\sum_{i=1}^{\bar{n}} \bar{A}_i > 0$ (the matrix $\bar{\mathfrak{B}}^{-1}$ is positive definite), this implies that $\bar{A}_k > 0$ for $k = 1, \dots, \bar{n}$. Equation (9) also leads to the following useful formula:

$$\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}} = \frac{1}{\mathbf{1}^{\perp}\bar{\mathfrak{B}}^{-1}\mathbf{1}} = \frac{1}{\sum_{i=1}^{\bar{n}} \bar{A}_i}.$$

The following assumption will be used in the sequel:

($\mathcal{A}$) For every $i \in \{1, \dots, n\} \setminus \bar{I}$,

$$\left(\mathbf{e}^i - \bar{\mathbf{w}}\right)^{\perp}\mathfrak{B}\bar{\mathbf{w}} \neq 0,$$

where $\mathbf{e}^i \in \mathbb{R}^n$ satisfies $e^i_j = 1$ if $i = j$ and $e^i_j = 0$ otherwise.

Remark 1. Assumption ($\mathcal{A}$) is equivalent to the following:

$$\left(\mathbf{e}^i - \bar{\mathbf{w}}\right)^{\perp}\mathfrak{B}\bar{\mathbf{w}} > 0,$$

for every $i \in \{1, \dots, n\} \setminus \bar{I}$. Indeed, the gradient of the minimization functional $\frac{1}{2}\mathbf{w}^{\perp}\mathfrak{B}\mathbf{w}$ at the point $\bar{\mathbf{w}}$ is given by $\mathfrak{B}\bar{\mathbf{w}}$, and for $\varepsilon > 0$ small enough, $\bar{\mathbf{w}} + (\mathbf{e}^i - \bar{\mathbf{w}})\varepsilon$ clearly belongs to Δ_n. Therefore $(\mathbf{e}^i - \bar{\mathbf{w}})^{\perp}\mathfrak{B}\bar{\mathbf{w}} < 0$ would contradict the fact that $\bar{\mathbf{w}}$ is the minimizer.

Assumption ($\mathcal{A}$) is a natural nondegeneracy condition for our problem. The following straightforward equality gives a relation between the optimization problem in (6) and a similar problem without the normalization constraint:

$$\inf_{\mathbf{w} \in \Delta_n, r \ge 0} \frac{r^2}{2}\mathbf{w}^{\perp}\mathfrak{B}\mathbf{w} - r = \inf_{\mathbf{v} \in \mathbb{R}^n\colon v_i \ge 0, i=1,\dots,n} \frac{1}{2}\mathbf{v}^{\perp}\mathfrak{B}\mathbf{v} - \mathbf{1}^{\perp}\mathbf{v}. \tag{10}$$

A minimizer $\bar{\mathbf{v}}$ of the right-hand side can therefore be constructed from the minimizer $\bar{\mathbf{w}}$ of (6) as follows:

$$\bar{\mathbf{v}} = \frac{\bar{\mathbf{w}}}{\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}}.$$

Now, introducing the vector $\boldsymbol{\lambda} \in \mathbb{R}^n$ of Lagrange multipliers for the positivity constraints on the right-hand side of (10), we get the Lagrangian

$$\frac{1}{2}\mathbf{v}^{\perp}\mathfrak{B}\mathbf{v} - \mathbf{1}^{\perp}\mathbf{v} - \boldsymbol{\lambda}^{\perp}\mathbf{v}.$$

At the extremum therefore, $\mathfrak{B}\bar{\mathbf{v}} = \mathbf{1} + \boldsymbol{\lambda}$, or in other words,

$$\frac{\mathfrak{B}\bar{\mathbf{w}}}{\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}} = \mathbf{1} + \boldsymbol{\lambda}.$$

Therefore, assumption ($\mathcal{A}$) simply states that for the constraints, which are saturated, the Lagrange multipliers are not equal to zero (since the constraints are inequalities, this is equivalent to the strict positivity for the multipliers). This is generally true, except when the solution of the unconstrained problem belongs to the boundary of the domain defined by the constraints.

The next assertion provides a sharp asymptotic formula with an error estimate for the distribution function of the random variable X, under assumption ($\mathcal{A}$). A similar formula for the distribution density of X will be formulated below (see Corollary 2).

Theorem 1. *Suppose assumption* ($\mathcal{A}$) *holds. Then, as* $x \to 0$,

$$\begin{aligned}\mathbb{P}[X \le x] = {} & \frac{C}{\bar{A}_1 + \cdots + \bar{A}_{\bar{n}}}\left(\log\frac{1}{x}\right)^{-(1+\bar{n})/2} x^{\sum_{k=1}^{\bar{n}} \bar{A}_k(\log(\bar{A}_1+\cdots+\bar{A}_n)/\bar{A}_k + \bar{\mu}_k)} \\ & \times \exp\left\{-\frac{1}{2}(\bar{A}_1 + \cdots + \bar{A}_{\bar{n}})\log^2\frac{1}{x}\right\}\left(1 + \mathrm{O}\left(\left(\log\frac{1}{x}\right)^{-1}\right)\right),\end{aligned} \tag{11}$$

where

$$\begin{aligned}C = {} & \frac{1}{\sqrt{2\pi}\sqrt{|\bar{\mathfrak{B}}|}}\frac{\sqrt{\bar{A}_1 + \cdots + \bar{A}_{\bar{n}}}}{\sqrt{\bar{A}_1 \cdots \bar{A}_{\bar{n}}}} \\ & \times \exp\left\{-\frac{1}{2}\sum_{i,j=1}^{\bar{n}} \bar{a}_{ij}\left(\log\frac{\bar{A}_1 + \cdots + \bar{A}_{\bar{n}}}{\bar{A}_i} + \bar{\mu}_i\right)\left(\log\frac{\bar{A}_1 + \cdots + \bar{A}_{\bar{n}}}{\bar{A}_j} + \bar{\mu}_j\right)\right\}.\end{aligned} \tag{12}$$

Remark 2. Formula (11) can be rewritten in terms of the solution $\bar{\mathbf{w}}$ to the quadratic programming problem in (6) as follows:

$$\begin{aligned}\mathbb{P}[X \le x] = {} & \widetilde{C}\left(\log\frac{1}{x}\right)^{-(1+\bar{n})/2}\exp\left\{-\frac{(\log x - \bar{\mathbf{w}}^{\perp}\boldsymbol{\mu} - \mathcal{E}(\bar{\mathbf{w}}))^2}{2\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}}\right\} \\ & \times \left(1 + \mathrm{O}\left(|\log x|^{-1}\right)\right)\end{aligned} \tag{13}$$

as $x \to 0$, where $\mathcal{E}(\bar{\mathbf{w}}) = -\sum_{i=1}^{n} \bar{w}_i \log \bar{w}_i$ and

$$\widetilde{C} = C\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}\exp\left(\frac{\mathcal{E}(\bar{\mathbf{w}})^2}{2\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}}\right).$$

The asymptotic behavior of the left tail of the sum of log-normal random variables with positive coefficients is thus intimately related to the quadratic programming problem formulated in (6). In particular, this problem determines which components of the random vector influence the tail behavior.

Theorem 1 and Corollary 2 below allow us to estimate various conditional expectations. The next assertion provides a characterization of the limiting conditional law of the Laplace transform of $Y_1, \ldots, Y_n$, given that $X \le x$.

Corollary 1. *Suppose assumption* $(\mathcal{A})$ *holds. Then, as* $x \to 0$, *for any* $u \in \mathbb{R}^n$,

$$\mathbb{E}\big[\mathrm{e}^{\sum_{i=1}^n u_i Y_i} | X \le x\big] = x^{\sum_{i=1}^n u_i \sum_{j=1}^{\bar{n}} \bar{A}_j b_{ij}} \times \exp\left\{\sum_{i=1}^n u_i \left(\mu_i - \sum_{p,q=1}^{\bar{n}} b_{pi}\bar{a}_{pq}\left(\log \frac{\bar{A}_1 + \cdots + \bar{A}_{\bar{n}}}{\bar{A}_q} + \bar{\mu}_q\right)\right)\right\} \times \exp\left\{\frac{1}{2}\sum_{i,j\notin \bar{I}} u_i u_j \left(b_{ij} - \sum_{p,q=1}^{\bar{n}} \bar{a}_{pq} b_{pi} b_{qj}\right)\right\}\left(1 + \mathrm{O}\left(\left(\log \frac{1}{x}\right)^{-1}\right)\right).$$

Remark 3. Note that

$$\sum_{j=1}^{\bar{n}} \bar{A}_j b_{ij} = \frac{[\mathfrak{B}\bar{\mathbf{w}}]_i}{\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}} \begin{cases} = 1, & i \in \bar{I}, \\ > 1, & i \notin \bar{I} \end{cases}$$

by assumption $(\mathcal{A})$. Let

$$\bar{\lambda}_i = \frac{[\mathfrak{B}\bar{\mathbf{w}}]_i}{\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}} - 1. \tag{14}$$

Then, Corollary 1 implies that the conditional distribution of the vector

$$\mathbf{Y} - (\mathbf{1} + \bar{\boldsymbol{\lambda}}) \log x,$$

given $X \le x$, converges weakly to the (degenerate) Gaussian law with mean

$$\mu_i' = \mu_i - \sum_{p,q=1}^{\bar{n}} b_{pi}\bar{a}_{pq}\left(\log \frac{\bar{A}_1 + \cdots + \bar{A}_{\bar{n}}}{\bar{A}_q} + \bar{\mu}_q\right)$$

and covariance matrix $\mathfrak{B}' = (b_{ij}')$ where

$$b_{ij}' = \left\{b_{ij} - \sum_{p,q=1}^{\bar{n}} \bar{a}_{pq} b_{pi} b_{qj}\right\}\mathbf{1}_{i,j\notin \bar{I}}.$$

Note that for $i \in \bar{I}$, the expression for μ_i' simplifies to

$$\mu_i' = \log \frac{\bar{A}_i}{\bar{A}_1 + \cdots + \bar{A}_{\bar{n}}} = \log \bar{w}_i.$$

The next statement concerns the asymptotics of the distribution density p of the random variable X.

Corollary 2. *Suppose assumption* $(\mathcal{A})$ *holds. Then, as* $x \to 0$,

$$\begin{aligned} p(x) = C\left(\log \frac{1}{x}\right)^{(1-\bar{n})/2} & x^{-1+\sum_{k=1}^{\bar{n}} \bar{A}_k(\log(\bar{A}_1+\cdots+\bar{A}_{\bar{n}})/\bar{A}_k+\bar{\mu}_k)} \\ & \times \exp\left\{-\frac{1}{2}(\bar{A}_1 + \cdots + \bar{A}_{\bar{n}}) \log^2 \frac{1}{x}\right\}\left(1 + \mathrm{O}\left(\left(\log \frac{1}{x}\right)^{-1}\right)\right), \end{aligned} \tag{15}$$

where the constant C *is given by* (12).

Corollary 2 implies that the conditional expectation in Corollary 1 can be taken with respect to the event $\{X = x\}$.

Corollary 3. *Suppose assumption* $(\mathcal{A})$ *holds. Then, as* $x \to 0$, *for any* $u \in \mathbb{R}^n$,

$$\begin{aligned} \mathbb{E}\left[\mathrm{e}^{\sum_{i=1}^n u_i Y_i} | X = x\right] = {} & x^{\sum_{i=1}^n u_i \sum_{j=1}^{\bar{n}} \bar{A}_j b_{ij}} \\ & \times \exp\left(\sum_{i=1}^n u_i \left\{\mu_i - \sum_{p,q=1}^{\bar{n}} b_{pi} \bar{a}_{pq}\left(\log \frac{\bar{A}_1 + \cdots + \bar{A}_{\bar{n}}}{\bar{A}_q} + \bar{\mu}_q\right)\right\}\right) \\ & \times \exp\left(\frac{1}{2} \sum_{i,j \notin \bar{I}} u_i u_j \left\{b_{ij} - \sum_{p,q=1}^{\bar{n}} \bar{a}_{pq} b_{pi} b_{qj}\right\}\right)\left(1 + \mathrm{O}\left(\left(\log \frac{1}{x}\right)^{-1}\right)\right). \end{aligned} \tag{16}$$

Example: the sum of two log-normal variables

Let $n = 2$, and denote the elements of the matrix $\mathfrak{B}$ by $b_{11} = \sigma_1^2$, $b_{22} = \sigma_2^2$ and $b_{12} = b_{21} = \rho\sigma_1\sigma_2$. To fix the ideas we assume $\sigma_1 \geq \sigma_2$. Then, $\mathbf{w} = (v, 1 - v)^{\perp}$ and

$$\mathbf{w}^{\perp}\mathfrak{B}\mathbf{w} = \sigma_1^2 v^2 + \sigma_2^2(1 - v)^2 + 2\rho\sigma_1\sigma_2 v(1 - v).$$

Therefore, the solution to problem (6) is given by

$$\bar{\mathbf{w}} = (\bar{v}, 1 - \bar{v})^{\perp} \qquad \text{with } \bar{v} = \frac{\sigma_2(\sigma_2 - \rho\sigma_1)}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} \vee 0$$

and we have the following three cases:

- If $\rho < \frac{\sigma_2}{\sigma_1}$, then both weights are strictly positive, assumption $(\mathcal{A})$ holds, and the asymptotic behavior of the density p is as follows:

$$\begin{aligned} p(z) = {} & \frac{C}{z\sqrt{|\log z|}} \\ & \times \exp\Big\{ -\frac{1}{2}\big(\mu_1 + x^* - \log z, \mu_2 + y^* - \log z\big) \\ & \qquad \times \mathfrak{B}^{-1}\big(\mu_1 + x^* - \log z, \mu_2 + y^* - \log z\big)^{\perp} \Big\} \\ & \times \left(1 + \mathrm{O}\left(\frac{1}{|\log z|}\right)\right) \end{aligned}$$

as $z \to 0$, with

$$C = \sqrt{\frac{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}{2\pi(\sigma_2^2 - \rho\sigma_1\sigma_2)(\sigma_1^2 - \rho\sigma_1\sigma_2)}},$$

$$x^* = \log \frac{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}{\sigma_2^2 - \rho\sigma_1\sigma_2} \quad \text{and} \quad y^* = \log \frac{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}{\sigma_1^2 - \rho\sigma_1\sigma_2}.$$

- If $\rho > \frac{\sigma_2}{\sigma_1}$, then $\bar{\mathbf{w}} = (0,1)^{\perp}$, assumption $(\mathcal{A})$ holds, and the asymptotic behavior of the density is characterized by

$$p(z) = \frac{1}{z\sigma_2\sqrt{2\pi}} \mathrm{e}^{-(\log z - \mu_2)^2/(2\sigma_2^2)} \left(1 + \mathrm{O}\left(\frac{1}{|\log z|}\right)\right)$$

as $z \to 0$. Note that in this case the asymptotic behavior of p is determined by the second component only.
- The case, where $\rho = \frac{\sigma_2}{\sigma_1}$, is exceptional. Here we have $\bar{\mathbf{w}} = (0,1)^{\perp}$, but assumption $(\mathcal{A})$ does not hold. Thus, Theorem 1 can not be applied.

In [20], Gao, Xu, and Ye characterize the left tail behavior of the sum of two log-normal variables in all the three cases described above. It follows from the results established in [20] that the asymptotic behavior of the density p in the exceptional case is qualitatively different from the behavior of p in the cases where $\rho > \frac{\sigma_2}{\sigma_1}$ or $\rho < \frac{\sigma_2}{\sigma_1}$. This shows that one can not relax assumption $(\mathcal{A})$ without changing the form of the asymptotics, which means that, in a sense, assumption $(\mathcal{A})$ is optimal.

Remark 4. In the second case of the above example, the variance of the second component is so small, that it completely dominates the asymptotic behavior of the left tail of the log-normal sum, so that, in a way, we recover the law of one jump. When this is the case, the asymptotic behavior of the distribution function can be characterized using an elementary argument given in the following proposition. In the text of the proposition and its proof, we denote $\sigma_i = \sqrt{\mathfrak{B}_{ii}}$ and

$\rho_{ij} = \frac{\mathfrak{B}_{ij}}{\sigma_i \sigma_j}$ for $1 \le i, j \le n$. We thank the anonymous referee for bringing this argument to our attention.

Proposition 1. *Assume that for some* i, $\sigma_i < \rho_{ij}\sigma_j$, $\forall j \neq i$. *Then, as* $x \to 0$,

$$\mathbb{P}[X \le x] = \frac{\sigma_i}{\log(1/x)\sqrt{2\pi}} \mathrm{e}^{-(\log x - \mu_i)^2/(2\sigma_i^2)} \left(1 + \mathrm{O}\left(\frac{1}{\sqrt{\log 1/x}}\right)\right). \tag{17}$$

Proof. By Jensen's inequality,

$$\mathbb{E}\big[(x - X)^+\big] \ge \mathbb{E}\big[\big(x - \mathbb{E}[X|Y_i]\big)^+\big].$$

An easy computation shows that

$$\mathbb{E}[X|Y_i] = \mathrm{e}^{Y_i} + \sum_{j \neq i} \exp\left(\mu_j + \sigma_j \rho_{ij}(Y_i - \mu_i)/\sigma_i + \frac{1}{2}\sigma_j^2\big(1 - \rho_{ij}^2\big)\right) \le \mathrm{e}^{Y_i} + c\mathrm{e}^{\alpha Y_i}$$

for some constants $c > 0$ and $\alpha > 1$. Combining this with a simple monotonicity argument, we get

$$\mathbb{E}\big[\big(x - \mathrm{e}^{Y_i} - c\mathrm{e}^{\alpha Y_i}\big)^+\big] \le \mathbb{E}\big[(x - X)^+\big] \le \mathbb{E}\big[\big(x - \mathrm{e}^{Y_i}\big)^+\big].$$

Since, on the event $\{x > \mathrm{e}^{Y_i} + c\mathrm{e}^{\alpha Y_i}\}$, $Y_i < \log x$, we also have,

$$\mathbb{E}\big[\big(x - cx^\alpha - \mathrm{e}^{Y_i}\big)^+\big] \le \mathbb{E}\big[(x - X)^+\big] \le \mathbb{E}\big[\big(x - \mathrm{e}^{Y_i}\big)^+\big]. \tag{18}$$

By standard arguments,

$$\begin{aligned}\mathbb{E}\big[\big(x - \mathrm{e}^{Y_i}\big)^+\big] &= -\mathrm{e}^{\mu_i + \sigma_i^2/2} N\left(-\frac{\mu_i + \sigma_i^2 - \log x}{\sigma_i}\right) + xN\left(-\frac{\mu_i - \log x}{\sigma_i}\right) \\ &= x\sigma_i n\left(-\frac{\mu_i - \log x}{\sigma_i}\right)\left(\frac{\sigma_i^2}{\log^2 1/x} + \mathrm{O}\left(\frac{1}{\log^3 1/x}\right)\right), \quad \text{as } x \to 0,\end{aligned}$$

where N is the standard normal distribution function and n is the standard normal density. From the previous estimate and (18), we deduce that

$$\mathbb{E}\big[(x - X)^+\big] = \frac{x\sigma_i^3}{\log^2(1/x)\sqrt{2\pi}} \mathrm{e}^{-(\mu_i - \log x)^2/(2\sigma_i^2)} \left(1 + \mathrm{O}\left(\frac{1}{\log 1/x}\right)\right) \quad \text{as } x \to 0.$$

To obtain the asymptotics for the distribution function, we use the following simple bound: for $\delta \in (0, x)$,

$$\frac{\mathbb{E}[(x - X)^+] - \mathbb{E}[(x - \delta - X)^+]}{\delta} \le \mathbb{P}[X \le x] \le \frac{\mathbb{E}[(x + \delta - X)^+] - \mathbb{E}[(x - X)^+]}{\delta}.$$

Taking $\delta = \frac{x}{(\log 1/x)^{3/2}}$, we then obtain formula (17) after some straightforwad computations. $\square$

2.2. Left tail of a log-normal sum. Proofs

The following preliminary result characterizes the asymptotic behavior of the distribution function and the density of the random variable X in the tail regime in the special case where $A_k > 0$ for all $k = 1, \dots, n$, or, equivalently, $\bar{n} = n$.

Lemma 1. *Assume that $A_k > 0, k = 1, \dots, n$. Then, as $x \to 0$,*

$$\begin{aligned} p(x) = C\left(\log\frac{1}{x}\right)^{(1-n)/2} & x^{-1+\sum_{k=1}^n A_k(\log(A_1+\cdots+A_n)/A_k+\mu_k)} \\ & \times \exp\left\{-\frac{1}{2}(A_1+\cdots+A_n)\log^2\frac{1}{x}\right\}\left(1+\mathrm{O}\left(\left(\log\frac{1}{x}\right)^{-1}\right)\right), \end{aligned} \tag{19}$$

and

$$\begin{aligned} \mathbb{P}[X \le x] = \frac{C}{A_1+\cdots+A_n}\left(\log\frac{1}{x}\right)^{-(1+n)/2} & x^{\sum_{k=1}^n A_k(\log(A_1+\cdots+A_n)/A_k+\mu_k)} \\ & \times \exp\left\{-\frac{1}{2}(A_1+\cdots+A_n)\log^2\frac{1}{x}\right\}\left(1+\mathrm{O}\left(\left(\log\frac{1}{x}\right)^{-1}\right)\right), \end{aligned} \tag{20}$$

where

$$\begin{aligned} C = & \frac{1}{\sqrt{2\pi}\sqrt{|\mathfrak{B}|}}\frac{\sqrt{A_1+\cdots+A_n}}{\sqrt{A_1\cdots A_n}} \\ & \times \exp\left\{-\frac{1}{2}\sum_{i,j=1}^n a_{ij}\left(\log\frac{A_1+\cdots+A_n}{A_i}+\mu_i\right)\left(\log\frac{A_1+\cdots+A_n}{A_j}+\mu_j\right)\right\}. \end{aligned} \tag{21}$$

Remark 5. Formula (19) for the density of the sum of log-normal random variables under the assumption $A_k > 0$ for $k = 1, \dots, n$ was given in the paper [20], but with a very different notation and without error estimates.

Proof of Lemma 1. We will first prove the formula in (19). The distribution function of X is given by

$$\mathbb{P}[X < x] = \int_0^x \mathrm{d}y_1 \int_0^{x-y_1} \mathrm{d}y_2 \cdots \int_0^{x-y_1-y_2-\cdots-y_{n-1}} \mathrm{d}^{\log}(y_1,\dots,y_n)\,\mathrm{d}y_n.$$

Differentiating the previous formula with respect to x and making the change of variables

$$\begin{aligned} x_1 = y_1/x, \qquad x_2 = y_2/x, \qquad \dots, \\ x_{n-1} = y_{n-1}/x, \qquad x_n = 1-(y_1+y_2+\cdots+y_{n-1})/x, \end{aligned} \tag{22}$$

we see that the density p of the random variable X can be represented as follows:

$$p(x) = x^{n-1} \int_0^1 \mathrm{d}x_1 \int_0^{1-x_1} \mathrm{d}x_2 \cdots \int_0^{1-x_1-\cdots-x_{n-2}} \mathrm{d}^{\log}(xx_1, \dots, xx_n)\, \mathrm{d}x_{n-1}. \tag{23}$$

Remark that x_n is not an independent variable but a function of $x_1, \dots, x_{n-1}$. Now, taking into account (1) and (23), we see that for every $x > 0$,

$$\begin{aligned} p(x) = {} & \frac{1}{(2\pi)^{n/2}\sqrt{|\mathfrak{B}|}x} \int_0^1 \mathrm{d}x_1 \int_0^{1-x_1} \mathrm{d}x_2 \cdots \\ & \times \int_0^{1-x_1-\cdots-x_{n-2}} \frac{1}{x_1 \cdots x_n} \exp\left\{-\frac{1}{2}\sum_{i,j=1}^n a_{ij}\big(\log(xx_i)-\mu_i\big)\big(\log(xx_j)-\mu_j\big)\right\} \mathrm{d}x_{n-1}. \end{aligned} \tag{24}$$

In the tail regime ($x \to 0$), we can isolate the effect of x in formula (24):

$$\begin{aligned} p(x) = {} & \frac{1}{(2\pi)^{n/2}\sqrt{|\mathfrak{B}|}x} \exp\left\{-\frac{1}{2}(A_1+\cdots+A_n)\log^2\frac{1}{x}\right\} \int_0^1 \mathrm{d}x_1 \int_0^{1-x_1} \mathrm{d}x_2 \cdots \\ & \times \int_0^{1-x_1-\cdots-x_{n-2}} \Phi(x_1,\dots,x_{n-1}) \exp\left\{-\log\frac{1}{x}\Psi(x_1,\dots,x_{n-1})\right\} \mathrm{d}x_{n-1}, \end{aligned} \tag{25}$$

where

$$\Phi(x_1,\dots,x_{n-1}) = \frac{1}{x_1\cdots x_n} \exp\left\{-\frac{1}{2}\sum_{i,j=1}^n a_{ij}\left(\log\frac{1}{x_i}+\mu_i\right)\left(\log\frac{1}{x_j}+\mu_j\right)\right\} \tag{26}$$

and

$$\Psi(x_1,\dots,x_{n-1}) = \sum_{k=1}^n A_k\left(\log\frac{1}{x_k}+\mu_k\right). \tag{27}$$

It is clear from formula (25) that it suffices to characterize the asymptotic behavior as $\theta \to \infty$ of the integral

$$\begin{aligned} I(\theta) = {} & \int_0^1 \mathrm{d}x_1 \int_0^{1-x_1} \mathrm{d}x_2 \cdots \\ & \times \int_0^{1-x_1-\cdots-x_{n-2}} \Phi(x_1,\dots,x_{n-1}) \exp\big\{-\theta\Psi(x_1,\dots,x_{n-1})\big\}\, \mathrm{d}x_{n-1}. \end{aligned} \tag{28}$$

We will use the higher-dimensional extension of Laplace's method in the proof.

Recall that $A_k > 0$ for all $1 \le k \le n$. We have

$$\frac{\partial\Psi}{\partial x_k} = -\frac{A_k}{x_k} + \frac{A_n}{1-x_1-\cdots-x_k-\cdots-x_{n-1}}$$

for all $1 \le k \le n-1$. It follows that the function Ψ has a unique critical point $x^* = (x_1^*, \ldots, x_{n-1}^*)$ where

$$x_k^* = \frac{A_k}{A_1 + A_2 + \cdots + A_n}$$

for $1 \le k \le n-1$. Note that the critical point x^* belongs to the interior of the integration set in (28), and moreover, this point is the global minimum point of the function Ψ. Next, using formula (8.3.50) in [13], we obtain

$$\begin{aligned} I\left(\log \frac{1}{x}\right) &= \frac{1}{\sqrt{\det(H(x^*))}} \left(\frac{2\pi}{\log 1/x}\right)^{(n-1)/2} \Phi\big(x_1^*, \ldots, x_{n-1}^*\big) \\ &\quad \times \exp\left\{-\log \frac{1}{x} \Psi\big(x_1^*, \ldots, x_{n-1}^*\big)\right\} \left(1 + O\left(\left(\log \frac{1}{x}\right)^{-1}\right)\right) \end{aligned} \tag{29}$$

as $x \to 0$, where $H(x^*)$ is the Hessian matrix of the function Ψ evaluated at the critical point x^*. Note that

$$\begin{aligned} &\Phi\big(x_1^*, \ldots, x_{n-1}^*\big) \\ &\quad = \frac{(A_1 + \cdots + A_n)^n}{A_1 \cdots A_n} \\ &\qquad \times \exp\left\{-\frac{1}{2} \sum_{k=1}^n a_{kk} \left(\log \frac{A_1 + \cdots + A_n}{A_k} + \mu_k\right)^2 \right. \\ &\qquad\qquad \left. - \sum_{1 \le i < j \le n} a_{ij} \left(\log \frac{A_1 + \cdots + A_n}{A_i} + \mu_i\right)\left(\log \frac{A_1 + \cdots + A_n}{A_j} + \mu_j\right)\right\} \end{aligned} \tag{30}$$

and

$$\Psi\big(x_1^*, \ldots, x_{n-1}^*\big) = \sum_{k=1}^n A_k \left(\log \frac{A_1 + \cdots + A_n}{A_k} + \mu_k\right). \tag{31}$$

Moreover, since

$$\frac{\partial^2 \Psi}{\partial x_k^2}\big(x^*\big) = (A_1 + \cdots + A_n)^2 \left(\frac{1}{A_k} + \frac{1}{A_n}\right), \qquad 1 \le k \le n-1,$$

and

$$\frac{\partial^2 \Psi}{\partial x_i \, \partial x_j}\big(x^*\big) = (A_1 + \cdots + A_n)^2 \frac{1}{A_n}, \qquad 1 \le i < j \le n-1,$$

we have

$$\det\big(H\big(x^*\big)\big) = (A_1 + \cdots + A_n)^{2n-2} \times \begin{vmatrix} A_1^{-1} + A_n^{-1} & A_n^{-1} & \cdots & A_n^{-1} \\ A_n^{-1} & A_2^{-1} + A_n^{-1} & \cdots & A_n^{-1} \\ \cdots & \cdots & \cdots & \cdots \\ A_n^{-1} & A_n^{-1} & \cdots & A_{n-1}^{-1} + A_n^{-1} \end{vmatrix}. \tag{32}$$

Next, using (32) and making long and tedious computations, we get the following equality:

$$\det\big(H\big(x^*\big)\big) = \frac{(A_1 + A_2 + \cdots + A_n)^{2n-1}}{A_1 A_2 \cdots A_n}. \tag{33}$$

Finally, taking into account (25), (29), and (30), (31), and (33), we complete the proof of formula (19) in Lemma 1. Formula (20) can be derived by integrating formula (19), or we can prove (20) directly by employing the same methods as those used in the proof of (19). □

Proof of Theorem 1. Let $k \in \{\bar{n}, \ldots, n-1\}$, $z \in (0, \frac{1}{2})$, and a, b be such that $\mathrm{e}^a + \mathrm{e}^b = z$. Then,

$$\begin{aligned} \mathbb{P}\big[\mathrm{e}^{Y_1} + \cdots + \mathrm{e}^{Y_{k+1}} \le z\big] &\ge \mathbb{P}\big[\mathrm{e}^{Y_1} + \cdots + \mathrm{e}^{Y_k} \le \mathrm{e}^a, Y_{k+1} \le b\big] \\ &= \mathbb{P}\big[\mathrm{e}^{Y_1} + \cdots + \mathrm{e}^{Y_k} \le \mathrm{e}^a\big] - \mathbb{P}\big[\mathrm{e}^{Y_1} + \cdots + \mathrm{e}^{Y_k} \le \mathrm{e}^a, Y_{k+1} > b\big]. \end{aligned}$$

Note that $k \ge \bar{n}$ implies that $\sum_{i=1}^k \bar{w}_i = 1$. The second term in the above formula can be estimated as follows:

$$\begin{aligned} \mathbb{P}\big[\mathrm{e}^{Y_1} + \cdots + \mathrm{e}^{Y_k} \le \mathrm{e}^a, Y_{k+1} > b\big] &\le \mathbb{P}[\bar{w}_1 Y_1 + \cdots + \bar{w}_k Y_k \le a, Y_{k+1} > b] \\ &\le \mathbb{P}\big[\bar{w}_1 Y_1 + \cdots + \bar{w}_k Y_k - a \le \alpha(Y_{k+1} - b)\big] \\ &= N\left(\frac{a - \alpha b - \mathbb{E}[\bar{\mathbf{w}}^\perp \mathbf{Y} - \alpha Y_{k+1}]}{\sqrt{\operatorname{Var}[\bar{\mathbf{w}}^\perp \mathbf{Y} - \alpha Y_{k+1}]}}\right) \end{aligned}$$

for every $\alpha > 0$. Now, let

$$x_{k+1} = \frac{(\mathbf{e}^{k+1})^\perp \mathfrak{B} \bar{\mathbf{w}}}{\bar{\mathbf{w}}^\perp \mathfrak{B} \bar{\mathbf{w}}} > 1$$

(the inequality follows from assumption $(\mathcal{A})$) and choose

$$b = \big(1 + (x_{k+1} - 1)/2\big) \log z \quad \Rightarrow \quad a = \log z + \log\big(1 - z^{(x_{k+1}-1)/2}\big).$$

Noting that

$$\operatorname{Var}\big[\bar{\mathbf{w}}^\perp \mathbf{Y} - \alpha Y_{k+1}\big] = \bar{\mathbf{w}}^\perp \mathfrak{B} \bar{\mathbf{w}} (1 - 2\alpha x_{k+1}) + \alpha^2 \mathfrak{B}_{k+1,k+1},$$

and making the above substitutions, we obtain, for α small enough,

$$\begin{aligned}
&\frac{a-\alpha b-\mathbb{E}[\bar{\mathbf{w}}^{\perp}\mathbf{Y}-\alpha Y_{k+1}]}{\sqrt{\mathrm{Var}[\bar{\mathbf{w}}^{\perp}\mathbf{Y}-\alpha Y_{k+1}]}}\\
&\quad=\frac{(1-\alpha(x_{k+1}+1)/2)\log z+\log(1-z^{(x_{k+1}-1)/2})-\bar{\mathbf{w}}^{\perp}\boldsymbol{\mu}+\alpha\mu_{k+1}}{\sqrt{\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}}\sqrt{1-2\alpha x_{k+1}+\alpha^2\mathfrak{B}_{k+1,k+1}/(\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}})}}\\
&\quad\le\frac{\log z}{\sqrt{\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}}}\frac{1-\alpha(x_{k+1}+1)/2}{\sqrt{1-2\alpha x_{k+1}+\alpha^2\mathfrak{B}_{k+1,k+1}/(\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}})}}+C_{k+1},
\end{aligned}$$

where C_{k+1} is a constant which does not depend on z. Next, for α small enough,

$$\begin{aligned}
\frac{1-\alpha(x_{k+1}+1)/2}{\sqrt{1-2\alpha x_{k+1}+\alpha^2\mathfrak{B}_{k+1,k+1}/(\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}})}}&\ge\left(1-\alpha\frac{x_{k+1}+1}{2}\right)\sqrt{1+2\alpha x_{k+1}-\frac{\alpha^2\mathfrak{B}_{k+1,k+1}}{\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}}}\\
&\ge\left(1-\alpha\frac{x_{k+1}+1}{2}\right)\left(1+\alpha x_{k+1}-\frac{\alpha^2\mathfrak{B}_{k+1,k+1}}{2\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}}\right)\\
&=1+\alpha\frac{x_{k+1}-1}{2}+\mathrm{O}\big(\alpha^2\big),\qquad \alpha\to 0.
\end{aligned}$$

Now it is easy to see that by choosing α small enough, one can always find $\varepsilon_{k+1}>0$ such that

$$\frac{a-\alpha b-\mathbb{E}[\bar{\mathbf{w}}^{\perp}\mathbf{Y}-\alpha Y_{k+1}]}{\sqrt{\mathrm{Var}[\bar{\mathbf{w}}^{\perp}\mathbf{Y}-\alpha Y_{k+1}]}}\le\frac{\log z}{\sqrt{\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}}}(1+\varepsilon_{k+1})+C_{k+1}.$$

We conclude that

$$\begin{aligned}
\mathbb{P}\big[\mathrm{e}^{Y_1}+\cdots+\mathrm{e}^{Y_{k+1}}\le z\big]\ge\mathbb{P}\big[\mathrm{e}^{Y_1}+\cdots+\mathrm{e}^{Y_k}\le z\big(1-z^{(x_{k+1}-1)/2}\big)\big]&\\
-N\left(\frac{\log z}{\sqrt{\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}}}(1+\varepsilon_{k+1})+C_{k+1}\right)&.
\end{aligned}$$

Let us first apply this formula for $k=\bar{n}$ and $z=x$. Since

$$N(y)=\mathrm{O}\left(\frac{\mathrm{e}^{-y^2/2}}{|y|}\right)$$

as $y\to-\infty$, and

$$\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}=\frac{1}{\sum_{i=1}^{\bar{n}}\bar{A}_i},$$

using Lemma 1 to compute the asymptotics of

$$\mathbb{P}\big[\mathrm{e}^{Y_1}+\cdots+\mathrm{e}^{Y_{\bar{n}}}\le x\big]$$

we have that

$$\frac{N(\log x/\sqrt{\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}}(1+\varepsilon_{\bar{n}+1})+C_{\bar{n}+1})}{\mathbb{P}[\mathrm{e}^{Y_1}+\cdots+\mathrm{e}^{Y_{\bar{n}}}\leq x]}=\mathrm{O}\left(\left(\log\frac{1}{x}\right)^{-1}\right)$$

as $x\to 0$. On the other hand, by Lemma 1, for $\delta>0$

$$\begin{aligned}
&\mathbb{P}\big[\mathrm{e}^{Y_1}+\cdots+\mathrm{e}^{Y_{\bar{n}}}\leq x\big(1-x^{\delta}\big)\big]\\
&\quad=\frac{C}{\bar{A}_1+\cdots+\bar{A}_{\bar{n}}}\left(\log\frac{1}{x}-\log\big(1-x^{\delta}\big)\right)^{-\bar{n}/2}\big(x\big(1-x^{\delta}\big)\big)^{\sum_{k=1}^{\bar{n}}\bar{A}_k(\log(\bar{A}_1+\cdots+\bar{A}_{\bar{n}})/\bar{A}_k+\bar{\mu}_k)}\\
&\qquad\times\exp\left(-\frac{1}{2}(\bar{A}_1+\cdots+\bar{A}_{\bar{n}})\left(\log\frac{1}{x}-\log\big(1-x^{\delta}\big)\right)^2\right)\left(1+\mathrm{O}\left(\left(\log\frac{1}{x}\right)^{-1}\right)\right)\\
&\quad=\frac{C}{\bar{A}_1+\cdots+\bar{A}_{\bar{n}}}\left(\log\frac{1}{x}\right)^{-\bar{n}/2}x^{\sum_{k=1}^{\bar{n}}\bar{A}_k(\log(\bar{A}_1+\cdots+\bar{A}_{\bar{n}})/\bar{A}_k+\bar{\mu}_k)}\\
&\qquad\times\exp\left(-\frac{1}{2}(\bar{A}_1+\cdots+\bar{A}_{\bar{n}})\log^2\frac{1}{x}\right)\left(1+\mathrm{O}\left(\left(\log\frac{1}{x}\right)^{-1}\right)\right)
\end{aligned}$$

as $x\to 0$. Therefore, we have shown that

$$\mathbb{P}\big[\mathrm{e}^{Y_1}+\cdots+\mathrm{e}^{Y_{\bar{n}+1}}\leq z\big]\geq\mathbb{P}\big[\mathrm{e}^{Y_1}+\cdots+\mathrm{e}^{Y_{\bar{n}}}\leq z\big]\left(1+\mathrm{O}\left(\left(\log\frac{1}{x}\right)^{-1}\right)\right)$$

as $x\to 0$, and since clearly,

$$\mathbb{P}\big[\mathrm{e}^{Y_1}+\cdots+\mathrm{e}^{Y_{\bar{n}+1}}\leq z\big]\leq\mathbb{P}\big[\mathrm{e}^{Y_1}+\cdots+\mathrm{e}^{Y_{\bar{n}}}\leq z\big],$$

we also get

$$\mathbb{P}\big[\mathrm{e}^{Y_1}+\cdots+\mathrm{e}^{Y_{\bar{n}+1}}\leq z\big]=\mathbb{P}\big[\mathrm{e}^{Y_1}+\cdots+\mathrm{e}^{Y_{\bar{n}}}\leq z\big]\left(1+\mathrm{O}\left(\left(\log\frac{1}{x}\right)^{-1}\right)\right)$$

as $x\to 0$. Iterating this procedure $n-\bar{n}$ times using the induction argument, we finally get that

$$\mathbb{P}\big[\mathrm{e}^{Y_1}+\cdots+\mathrm{e}^{Y_n}\leq z\big]=\mathbb{P}\big[\mathrm{e}^{Y_1}+\cdots+\mathrm{e}^{Y_{\bar{n}}}\leq z\big]\left(1+\mathrm{O}\left(\left(\log\frac{1}{x}\right)^{-1}\right)\right)$$

as $x\to 0$, which completes the proof of the theorem, since the asymptotics for

$$\mathbb{P}\big[\mathrm{e}^{Y_1}+\cdots+\mathrm{e}^{Y_{\bar{n}}}\leq z\big]$$

can be computed using Lemma 1. □

Proof of Corollary 1. For any $u\in\mathbb{R}^n$, we have

$$\mathbb{E}\big[\mathrm{e}^{\sum_{i=1}^n u_iY_i}|X\leq x\big]=\frac{\mathbb{E}[\mathrm{e}^{\sum_{i=1}^n u_iY_i}\mathbf{1}_{X\leq x}]}{\mathbb{P}[X\leq x]}=\mathbb{E}\big[\mathrm{e}^{\sum_{i=1}^n u_iY_i}\big]\frac{\tilde{\mathbb{P}}[X\leq x]}{\mathbb{P}[X\leq x]},\tag{34}$$

where the symbol $\tilde{\mathbb{P}}$ stands for a new probability determined from

$$\frac{\mathrm{d}\tilde{\mathbb{P}}}{\mathrm{d}\mathbb{P}} = \frac{\mathrm{e}^{\sum_{i=1}^n u_i Y_i}}{\mathbb{E}[\mathrm{e}^{\sum_{i=1}^n u_i Y_i}]}.$$

Under this probability, $\mathbf{Y} \sim N(\boldsymbol{\mu} + \mathfrak{B}\mathbf{u}, \mathfrak{B})$. Applying Theorem 1 to the numerator and the denominator of the fraction in (34), and making cancellations, we get the following:

$$\begin{aligned}
&\mathbb{E}\big[\mathrm{e}^{\sum_{i=1}^n u_i Y_i} | X \le x\big] \\
&\quad = x^{\sum_{i=1}^n u_i \sum_{j=1}^{\bar{n}} \bar{A}_j b_{ij}} \mathrm{e}^{\boldsymbol{\mu}^\perp \mathbf{u} + (1/2)\mathbf{u}^\perp \mathfrak{B}\mathbf{u}} \frac{C_{\boldsymbol{\mu}+\mathfrak{B}\mathbf{u},\mathfrak{B}}}{C_{\boldsymbol{\mu},\mathfrak{B}}} \left(1 + \mathrm{O}\left(\left(\log\frac{1}{x}\right)^{-1}\right)\right) \\
&\quad = x^{\sum_{i=1}^n u_i \sum_{j=1}^{\bar{n}} \bar{A}_j b_{ij}} \exp\left\{\frac{1}{2}\sum_{i,j=1}^n u_i u_j \left(b_{ij} - \sum_{p,q=1}^{\bar{n}} \bar{a}_{pq} b_{pi} b_{qj}\right)\right\} \\
&\qquad \times \exp\left\{\sum_{i=1}^n u_i \left(\mu_i - \sum_{p,q=1}^{\bar{n}} \bar{a}_{pq} b_{pi} \left(\log\frac{\bar{A}_1 + \cdots + \bar{A}_{\bar{n}}}{\bar{A}_q} + \bar{\mu}_q\right)\right)\right\} \\
&\qquad \times \left(1 + \mathrm{O}\left(\left(\log\frac{1}{x}\right)^{-1}\right)\right)
\end{aligned}$$

as $x \to 0$. The symbols $C_{\boldsymbol{\mu}+\mathfrak{B}\mathbf{u},\mathfrak{B}}$ and $C_{\boldsymbol{\mu},\mathfrak{B}}$, appearing in the previous estimates, stand for the constant C in Theorem 1, evaluated for the log-normal variables, associated with the pairs $(\boldsymbol{\mu} + \mathfrak{B}\mathbf{u}, \mathfrak{B})$ and $(\boldsymbol{\mu}, \mathfrak{B})$, respectively. It is easy to check that when $i \in \bar{I}$ or $j \in \bar{I}$, necessarily

$$b_{ij} = \sum_{p,q=1}^{\bar{n}} \bar{a}_{pq} b_{pi} b_{qj}.$$

□

Proof of Corollary 2. Recall the formula for the distribution function of X:

$$\begin{aligned}
\mathbb{P}[X \le x] &= \int_{\mathrm{e}^{y_1} + \cdots + \mathrm{e}^{y_n} \le x} \frac{1}{(2\pi)^{n/2}\sqrt{|\mathfrak{B}|}} \\
&\qquad \times \exp\left\{-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu})^\perp \mathfrak{B}^{-1}(\mathbf{y} - \boldsymbol{\mu})\right\} \mathrm{d}y_1 \cdots \mathrm{d}y_n \\
&= \int_{\mathrm{e}^{z_1} + \cdots + \mathrm{e}^{z_n} \le 1} \frac{1}{(2\pi)^{n/2}\sqrt{|\mathfrak{B}|}} \\
&\qquad \times \exp\left\{-\frac{1}{2}(\mathbf{z} + \mathbf{1}\log x - \boldsymbol{\mu})^\perp \mathfrak{B}^{-1}(\mathbf{z} + \mathbf{1}\log x - \boldsymbol{\mu})\right\} \mathrm{d}z_1 \cdots \mathrm{d}z_n.
\end{aligned}$$

Differentiating with respect to x, we obtain an alternative representation for the density:

$$
\begin{aligned}
p(x) = \int_{\mathrm{e}^{z_1}+\cdots+\mathrm{e}^{z_n}\le 1} & \frac{-\mathbf{1}^{\perp}\mathfrak{B}^{-1}(\mathbf{z}+\mathbf{1}\log x-\boldsymbol{\mu})}{x(2\pi)^{n/2}\sqrt{|\mathfrak{B}|}} \\
& \times \exp\left\{-\frac{1}{2}(\mathbf{z}+\mathbf{1}\log x-\boldsymbol{\mu})^{\perp}\mathfrak{B}^{-1}(\mathbf{z}+\mathbf{1}\log x-\boldsymbol{\mu})\right\}\mathrm{d}z_1\cdots\mathrm{d}z_n \\
= -\frac{1}{x}\mathbb{E}\big[\mathbf{1}^{\perp}\mathfrak{B}^{-1}(\mathbf{Y}-\boldsymbol{\mu})\mathbf{1}_{X\le x}\big] &= -\frac{1}{x}\mathbb{E}\left[\sum_{i=1}^{n}A_i(Y_i-\mu_i)\mathbf{1}_{X\le x}\right].
\end{aligned}
$$

Next, we make a transformation inspired by Corollary 1. Taking into account Remark 3, we see that

$$
\begin{aligned}
p(x) &= -\frac{1}{x}\mathbb{E}\left[\sum_{i=1}^{n}A_i\left(Y_i-\log x\sum_{j=1}^{\bar{n}}\bar{A}_j b_{ij}-\mu_i\right)\mathbf{1}_{X\le x}\right] \\
&\quad -\frac{\log x}{x}\mathbb{E}\left[\sum_{i=1}^{n}A_i\sum_{j=1}^{\bar{n}}\bar{A}_j b_{ij}\mathbf{1}_{X\le x}\right] \\
&= \frac{\mathbb{P}[X\le x]}{x}\sum_{i=1}^{n}A_i\mu_i-\frac{1}{x}\sum_{i=1}^{n}A_i\mathbb{E}\big[\big(Y_i-(1+\bar{\lambda}_i)\log x\big)\mathbf{1}_{X\le x}\big] \\
&\quad -\frac{\log x}{x}\sum_{j=1}^{\bar{n}}\bar{A}_j\mathbb{P}[X\le x] \\
&= -\frac{\log x}{x}\sum_{j=1}^{\bar{n}}\bar{A}_j\mathbb{P}[X\le x]+\mathrm{O}\left(\frac{\mathbb{P}[X\le x]}{x}\right)
\end{aligned}
$$

as $x\to 0$. Here the constant $\bar{\lambda}_i$ is defined by (14). In the reasoning above, we used the following estimate, which can be derived from Corollary 1. For every i, as $x\to 0$,

$$
\begin{aligned}
&\big|\mathbb{E}\big[\big(Y_i-(1+\bar{\lambda}_i)\log x\big)\mathbf{1}_{X\le x}\big]\big| \\
&\quad\le \mathbb{P}[X\le x]\big(\mathbb{E}\big[\mathrm{e}^{Y_i-(1+\bar{\lambda}_i)\log x}|X\le x\big]+\mathbb{E}\big[\mathrm{e}^{-(Y_i-(1+\bar{\lambda}_i)\log x)}|X\le x\big]\big) \\
&\quad= C\mathbb{P}[X\le x]\left(1+\mathrm{O}\left(\left(\log\frac{1}{x}\right)^{-1}\right)\right)
\end{aligned}
$$

for some constant C. □

3. Asymptotic behavior of the right tail of a log-normal difference

In this section, we analyze the asymptotic behavior of the distribution function and the density of the random variable $X^{(m)}$ as $x \to +\infty$, assuming with no loss of generality that $m \geq 1$. If $m = 0$, then the support of the distribution of $X^{(m)}$ is $(-\infty, 0)$, and the tail behavior at 0 follows from the results obtained in Section 2.

3.1. Right tail of a log-normal difference. Results and discussions

Let us first consider, for every p with $1 \leq p \leq m$, the random variable

$$X_p^{(1)} = \mathrm{e}^{Y_p} - \sum_{k=m+1}^{n} \mathrm{e}^{Y_k}. \tag{35}$$

Let $\Delta_{m,n}^p$ be the set of weights $\mathbf{w} \in \mathbb{R}^n$ with $w_i = 0$ for $i = 1, \ldots, m$, $i \neq p$; $w_p \geq 0$; $w_i \leq 0$ for $i = m+1, \ldots, n$; and $\sum w_i = 1$. Let $\bar{\mathbf{w}}_p \in \Delta_{m,n}^p$ be the unique point such that

$$\bar{\mathbf{w}}_p \mathfrak{B} \bar{\mathbf{w}}_p = \min_{\mathbf{w} \in \Delta_{m,n}^p} \mathbf{w}^{\perp} \mathfrak{B} \mathbf{w}, \tag{36}$$

and define $\bar{n}^{(p)}$, $\bar{I}^{(p)}$, $\bar{k}^{(p)}$, $\bar{\boldsymbol{\mu}}^{(p)}$, $\bar{\mathfrak{B}}^{(p)}$, $\bar{a}_{ij}^{(p)}$, $\bar{A}^{(p)}$ as in equation (8) and below. We will say that assumption $(\mathcal{A}_1^p)$ holds if for every $i \in \{m+1, \ldots, n\} \setminus \bar{I}^{(p)}$,

$$\big(\mathbf{e}^i - \bar{\mathbf{w}}_p\big)^{\perp} \mathfrak{B} \bar{\mathbf{w}}_p \neq 0.$$

It follows from Theorem 3 (see Section 3.2), that if assumption $(\mathcal{A}_1^p)$ is satisfied, then

$$\mathbb{P}\big[X_p^{(1)} \geq x\big] = \delta_{1,p} (\log x)^{\delta_{2,p}} x^{\delta_{3,p}} \exp\left\{-\frac{\log^2 x}{2\delta_{4,p}}\right\} \big(1 + \mathrm{O}\big((\log x)^{-1}\big)\big) \tag{37}$$

as $x \to \infty$, where

$$\delta_{1,p} = \frac{C^{(p)}}{\sum_{j=1}^{\bar{n}^{(p)}} \bar{A}_j^{(p)}}, \qquad \delta_{2,p} = -\frac{1 + \bar{n}^{(p)}}{2},$$

$$\delta_{3,p} = \sum_{i=1}^{\bar{n}_p} \bar{A}_i^{(p)} \left(\log \frac{\sum_{j=1}^{\bar{n}^{(p)}} \bar{A}_j^{(p)}}{|\bar{A}_i^{(p)}|} + \bar{\mu}_i^{(p)}\right), \quad \text{and} \quad \delta_{4,p} = \left(\sum_{j=1}^{\bar{n}^{(p)}} \bar{A}_j^{(p)}\right)^{-1}.$$

In other words, the exponential rate of decay in the leading term of the asymptotics of $\mathbb{P}[X_p^{(1)} \geq x]$ is determined by the quantity

$$\delta_{4,p} = \min_{\mathbf{w} \in \Delta_{m,n}^p} \mathbf{w}^{\perp} \mathfrak{B} \mathbf{w}.$$

Depending on the covariance matrix $\mathfrak{B}$, this rate may either be equal to b_{pp}, the inverse of the variance of Y_p, in which case the asymptotic behavior of $X_p^{(1)}$ is determined by Y_p only, or be greater than b_{pp}, in which case the asymptotic behavior of $X_p^{(1)}$ is determined by more than one component of the vector $(Y_1, \dots, Y_n)$. One may therefore call the number $\delta_{4,p}$ the relative asymptotic variance of Y_p with respect to $Y_{m+1}, \dots, Y_n$.

The next assertion, which is one of the main results of the present paper, shows that the asymptotic behavior of the distribution function of the random variable $X^{(m)}$ is dominated by one (or several similar) of the random variables $X_p^{(1)}$. We will need the following parameters to describe the above-mentioned domination:

$$\delta_4 = \max_{1 \le p \le m} \delta_{4,p}, \qquad \mathcal{P}_4 = \{p\colon\ 1 \le p \le m, \delta_{4,p} = \delta_4\}, \tag{38}$$

$$\delta_3 = \max_{p \in \mathcal{P}_4} \delta_{3,p}, \qquad \mathcal{P}_3 = \{p \in \mathcal{P}_4\colon\ \delta_{3,p} = \delta_3\}, \tag{39}$$

$$\delta_2 = \max_{p \in \mathcal{P}_3} \delta_{2,p}, \qquad \mathcal{P}_2 = \{p \in \mathcal{P}_3\colon\ \delta_{2,p} = \delta_2\}, \tag{40}$$

and finally

$$\delta_1 = \sum_{p \in \mathcal{P}_2} \delta_{1,p}.$$

Theorem 2. *Let assumption $(\mathcal{A}_1^p)$ hold for every $p = 1, \dots, m$. Then the distribution function of the random variable $X^{(m)}$ defined by (3) satisfies*

$$\mathbb{P}\big[X^{(m)} \ge x\big] = \delta_1 (\log x)^{\delta_2} x^{\delta_3} \exp\left\{-\frac{\log^2 x}{2\delta_4}\right\}\big(1 + \mathrm{O}\big((\log x)^{-1/2}\big)\big) \tag{41}$$

as $x \to \infty$.

Remark 6. When $m = n$, the variables $X_p^{(1)}$ are one-dimensional and log-normal, and the result in Theorem 2 reduces to Theorem 1 of [6] which shows that the asymptotic behavior of the right tail of $\mathrm{e}^{Y_1} + \cdots + \mathrm{e}^{Y_n}$ is determined by the components of $(Y_1, \dots, Y_n)$ which have the largest variance. For other values of m, Theorem 2 extends Theorem 1 of [6] by showing that the asymptotic behavior of the right tail of $X^{(m)}$ is determined by the components of $(Y_1, \dots, Y_m)$, which have the largest relative asymptotic variance with respect to $(Y_{m+1}, \dots, Y_n)$.

As in the case of Theorem 1, several useful corollaries can be derived from Theorem 2. We omit the proofs of those corollaries, since they are very similar to those given in Section 2.1.

Corollary 4. *Suppose that assumption $(\mathcal{A}_1^p)$ holds for every $p = 1, \dots, m$, and that the set $\mathcal{P}_4$*

is a singleton, $\mathcal{P}_4=\{p\}$. Then, as $x\to\infty$, for any $\mathbf{u}\in\mathbb{R}^n$,

$$\begin{aligned}
&\mathbb{E}\big[\mathrm{e}^{\sum_{i=1}^n u_i Y_i}|X\ge x\big]\\
&\quad=\mathbb{E}\big[\mathrm{e}^{\sum_{i=1}^n u_i Y_i}|X=x\big]\\
&\quad=x^{\sum_{j=1}^n u_j\sum_{i=1}^{\bar n^{(p)}}\bar A_i^{(p)} b_{\bar k^{(p)}(i),j}}\\
&\qquad\times\exp\Bigg\{\sum_{j=1}^n u_j\Bigg(\mu_j-\sum_{i,k=1}^{\bar n^{(p)}}\bar a_{ik}^{(p)} b_{\bar k^{(p)}(i),j}\Bigg(\log\frac{\sum_{j=1}^{\bar n^{(p)}}\bar A_j^{(p)}}{|\bar A_k^{(p)}|}+\mu_{\bar k^{(p)}(k)}\Bigg)\Bigg)\Bigg\}\\
&\qquad\times\exp\Bigg\{+\frac{1}{2}\sum_{j,l\notin\bar I^{(p)}}^n u_j u_l\Bigg(b_{jl}-\sum_{i,k=1}^{\bar n^{(p)}}\bar a_{ik}^{(p)} b_{\bar k^{(p)}(i),j} b_{\bar k^{(p)}(k),l}\Bigg)\Bigg\}\\
&\qquad\times\left(1+\mathrm{O}\left(\left(\log\frac{1}{x}\right)^{-1}\right)\right).
\end{aligned}\tag{42}$$

In the next statement, we use the notation introduced before the formulation of Theorem 2.

Corollary 5. *Suppose assumption $(\mathcal{A}_1^p)$ holds for every $p=1,\dots,m$. Then, as $x\to\infty$, the density $p^{(m)}$ of the random variable $X^{(m)}$ satisfies*

$$p^{(m)}(x)=\frac{\delta_1}{\delta_4}(\log x)^{\delta_2+1}x^{\delta_3-1}\exp\left\{-\frac{\log^2 x}{2\delta_4}\right\}\big(1+\mathrm{O}\big((\log x)^{-1/2}\big)\big).\tag{43}$$

3.2. Right tail of a log-normal difference. Proofs

Let us first consider the case when $m=1$. Define

$$\Delta_{1,n}:=\left\{\mathbf{w}\in\mathbb{R}^n\colon\ w_1\ge 0, w_i\le 0, i=2,\dots,n,\text{ and }\sum_{i=1}^n w_i=1\right\},\tag{44}$$

and introduce $\bar{\mathbf{w}}\in\Delta_{1,n}$ as the unique vector such that

$$\bar{\mathbf{w}}^\perp\mathfrak{B}\bar{\mathbf{w}}=\min_{\mathbf{w}\in\Delta_{1,n}}\mathbf{w}^\perp\mathfrak{B}\mathbf{w}.$$

The existence and uniqueness of $\bar{\mathbf{w}}$ follows from the non-degeneracy of $\mathfrak{B}$. When $A_1>0$ and $A_k<0$ for $k=2,\dots,n$, $\bar{\mathbf{w}}$ is given by (7). In the general case, we define $\bar n$, $\bar I$, $\bar{\boldsymbol{\mu}}$, $\bar{\mathfrak{B}}$, $\bar a_{ij}$ and $\bar A$ as in equation (8) and below.

Since by the definition of $\Delta_{1,n}$, $\bar w_1>0$, and moreover the variables $Y_2,\dots,Y_n$ are exchangeable in the definition of $X^{(1)}$, we shall assume with no loss of generality that $\bar I=\{1,\dots,\bar n\}$. This has already been done in Section 2.1.

Observe that the minimum in $\min_{\mathbf{w}\in\Delta_{1,\bar{n}}} \mathbf{w}^{\perp}\bar{\mathfrak{B}}\mathbf{w}$ is attained in the interior of $\mathbb{R}_{+}^{\bar{n}}$. This implies that $\bar{A}_1 > 0$ and $\bar{A}_k < 0$ for $k = 2, \ldots, \bar{n}$ (see a similar reasoning in Section 2.1).

The following preliminary lemma concerns the case where $\bar{n} = n$.

Lemma 2. *Let $A_1 > 0$, $A_2 < 0, \ldots, A_n < 0$. Then the following formulas hold*:

$$\begin{aligned} p^{(1)}(x) &= C(\log x)^{(1-n)/2} x^{-1+\sum_{i=1}^{n} A_i(\log\sum_{j=1}^{n} A_j/|A_i|+\mu_i)} \\ &\quad \times \exp\left\{-\frac{1}{2}\log^2 x\sum_{j=1}^{n} A_j\right\}\left(1+\mathrm{O}\left((\log x)^{-1}\right)\right), \end{aligned} \tag{45}$$

$$\begin{aligned} \mathbb{P}\left[X^{(1)} \geq x\right] &= \frac{C}{A_1+\cdots+A_n}(\log x)^{-(1+n)/2} x^{\sum_{i=1}^{n} A_i(\log\sum_{j=1}^{n} A_j/|A_i|+\mu_i)} \\ &\quad \times \exp\left\{-\frac{1}{2}\left(\log^2 x\right)\sum_{j=1}^{n} A_j\right\}\left(1+\mathrm{O}\left((\log x)^{-1}\right)\right) \end{aligned} \tag{46}$$

as $x \to \infty$. The constant C in (45) is given by

$$\begin{aligned} C &= \exp\left\{-\frac{1}{2}\sum_{i,m=1}^{n} a_{im}\left(\log\frac{\sum_{j=1}^{n} A_j}{|A_i|}+\mu_i\right)\left(\log\frac{\sum_{j=1}^{n} A_j}{|A_m|}+\mu_m\right)\right\} \\ &\quad \times \frac{1}{\sqrt{2\pi|\mathfrak{B}|}}\sqrt{\frac{\sum_{j=1}^{n} A_j}{\prod_{i=1}^{n}|A_i|}}. \end{aligned}$$

Proof. Differentiating the distribution function, we obtain the following representation of the density $p^{(1)}$ of $X^{(1)}$:

$$\begin{aligned} p^{(1)}(x) &= \frac{1}{(2\pi)^{n/2}\sqrt{|\mathfrak{B}|}}\exp\left\{-\frac{1}{2}\log^2 x\sum_{j=1}^{n} A_j\right\} \\ &\quad \times \int_{D_1^{1,n-1}} \widetilde{\Phi}(x_1,\ldots,x_{n-1})\exp\left\{-\log x\,\widetilde{\Psi}(x_1,\ldots,x_{n-1})\right\}\mathrm{d}x_1\cdots\mathrm{d}x_{n-1} \end{aligned} \tag{47}$$

with

$$\begin{aligned} \widetilde{\Phi}(x_1,\ldots,x_{n-1}) &= \Phi(x_1,\ldots,x_{n-1},x_1-x_2-\cdots-x_{n-1}-1), \\ \widetilde{\Psi}(x_1,\ldots,x_{n-1}) &= \Psi(x_1,\ldots,x_{n-1},x_1-x_2-\cdots-x_{n-1}-1), \\ \Phi(x_1,\ldots,x_n) &= \frac{1}{x_1\cdots x_n}\exp\left\{-\frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(\log x_i-\mu_i)(\log x_j-\mu_j)\right\}, \\ \Psi(x_1,\ldots,x_n) &= \sum_{i=1}^{n} A_i(\log x_i-\mu_i) \end{aligned}$$

and

$$D_1^{1,n-1} = \{(x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}\colon\ x_i \geq 0, 1 \leq i \leq n-1; x_1 - x_2 - \cdots - x_{n-1} > 1\}.$$

We have for all $1 \leq i \leq n-1$,

$$\frac{\partial \widetilde{\Psi}}{\partial x_i} = \frac{A_i}{x_i} + \frac{A_n}{x_1 - x_2 - \cdots - x_{n-1} - 1} s_i, \tag{48}$$

where $s_1 = 1$ and $s_i = -1$ for $1 < i \leq n-1$. Now, it is easy to see that the solution $\mathbf{x}^*$ to the system of equations $\frac{\partial \widetilde{\Psi}}{\partial x_i} = 0$, $1 \leq i \leq n-1$, is given by

$$x_i^* = \frac{A_i}{\sum_{j=1}^n A_j} s_i, \qquad 1 \leq i \leq n-1. \tag{49}$$

Under the assumptions in the formulation of the lemma, it is clear that $\mathbf{x}^*$ belongs to the interior of the set $D_1^{1,n-1}$. We will next apply Laplace's method to the integral in (47). However, first we need to check that the Hessian matrix of the function $\widetilde{\Psi}$ at the point $\mathbf{x}^*$, that is, the matrix $H(\mathbf{x}^*) := [h_{im}]_{i,m=1,\dots,n-1}$, is positive-definite. It is not hard to see that

$$h_{im} = -\frac{(\sum_{j=1}^n A_j)^2}{A_n} \qquad \text{if } i \neq m,$$

and

$$h_{ii} = -\left(\sum_{j=1}^n A_j\right)^2 \left(\frac{1}{A_i} + \frac{1}{A_n}\right).$$

Therefore,

$$H(\mathbf{x}^*) = -\frac{(\sum_{j=1}^n A_j)^2}{A_n} J, \tag{50}$$

where J is the $(n-1) \times (n-1)$-matrix with the entries $1 + \frac{A_n}{A_1}, \dots, 1 + \frac{A_n}{A_{n-1}}$ along the main diagonal, and all the entries outside the main diagonal equal to 1. It is an easy exercise in linear algebra to show that the leading principal minor of order p of the matrix J is equal to

$$A_n^{p-1} \left(\prod_{i=1}^p \frac{1}{A_i}\right) \left(\sum_{m=1}^p A_m + A_n\right). \tag{51}$$

Recall that it is assumed in Lemma 2 that

$$A_1 > 0, \qquad A_2 < 0, \qquad \dots, \qquad A_n < 0. \tag{52}$$

Under this assumption, the numbers in (51) are positive for $p = 1, \ldots, n-1$. For instance, if $p = 1$, then we have

$$\frac{1}{A_1}(A_1 + A_n) = \frac{1}{A_1}\left[\sum_{m=1}^{n} A_m - \sum_{m=2}^{n-1} A_m\right] > 0.$$

The previous inequality follows from the positive-definiteness of the matrix $\mathfrak{B}^{-1}$ and condition (52). It follows that the matrix J is positive definite, and hence the matrix $H(\mathbf{x}^*)$ is also positive definite. Moreover, the determinant of $H(\mathbf{x}^*)$ is given by

$$\frac{(\sum_{j=1}^{n} A_j)^{2n-1}}{\prod_{i=1}^{n} |A_i|}.$$

Next, taking in the account what was said above, we see that Laplace's method can be applied to the integral in (47). Similarly to (29), we get the following formula:

$$\begin{aligned}
p^{(1)}(x) &= \frac{1}{(2\pi)^{n/2} x \sqrt{|\mathfrak{B}|}} \exp\left\{-\frac{1}{2}\log^2 x \sum_{j=1}^{n} A_j\right\} \frac{1}{\sqrt{\det(H(\mathbf{x}^*))}} \left(\frac{2\pi}{\log x}\right)^{(n-1)/2} \\
&\quad \times \widetilde{\Phi}(x_1^*, \ldots, x_{n-1}^*) \exp\{-\log x\, \widetilde{\Psi}(x_1^*, \ldots, x_{n-1}^*)\}\big(1 + \mathrm{O}((\log x)^{-1})\big) \\
&= (\log x)^{(1-n)/2} x^{-1+\sum_{i=1}^{n} A_i (\log \sum_{j=1}^{n} A_j/|A_i| + \mu_i)} \exp\left\{-\frac{1}{2}\log^2 x \sum_{j=1}^{n} A_j\right\} \\
&\quad \times \exp\left\{-\frac{1}{2}\sum_{i,m=1}^{n} a_{im}\left(\log\frac{\sum_{j=1}^{n} A_j}{|A_i|} + \mu_i\right)\left(\log\frac{\sum_{j=1}^{n} A_j}{|A_m|} + \mu_m\right)\right\} \\
&\quad \times \frac{1}{\sqrt{2\pi|\mathfrak{B}|}} \sqrt{\frac{\sum_{j=1}^{n} A_j}{\prod_{i=1}^{n} |A_i|}} \big(1 + \mathrm{O}((\log x)^{-1})\big)
\end{aligned}$$

as $x \to \infty$. The asymptotic behavior of the distribution function can be characterized by integrating the asymptotic formula for the density. $\square$

We will next focus on the case where m is still equal to one, but the equality $\bar{n} = n$ may not hold. Our next result requires the following assumption:

($\mathcal{A}_1$) For every $i \in \{1, \ldots, n\} \setminus \bar{I}$,

$$(\mathbf{e}^i - \bar{\mathbf{w}})^{\perp} \mathfrak{B} \bar{\mathbf{w}} \neq 0.$$

Remark 7. Assumption ($\mathcal{A}_1$), although it has the same form as assumption ($\mathcal{A}$) above, is a different assumption on the covariance matrix $\mathfrak{B}$, because the weight vector $\bar{\mathbf{w}}$ is computed differently now (with $\Delta_{1,n}$ instead of Δ). Assumption ($\mathcal{A}_1$) is equivalent to the following: For every $i \in \{1, \dots, n\} \setminus \bar{I}$,

$$(\mathbf{e}^i - \bar{\mathbf{w}})^\perp \mathfrak{B} \bar{\mathbf{w}} < 0. \tag{53}$$

Indeed, since $1 \in \bar{I}$, for every $i \in \{1, \dots, n\} \setminus \bar{I}$ and for $\varepsilon > 0$ small enough, $\bar{\mathbf{w}} - \varepsilon(\mathbf{e}^i - \bar{\mathbf{w}})$ belongs to $\Delta_{1,n}$. Therefore, the inequality opposite to that in (53) would lead to a contradiction to the fact that $\bar{\mathbf{w}}$ is a minimizer.

Theorem 3. *Let assumption* $(\mathcal{A}_1)$ *hold true. Then, as* $x \to +\infty$,

$$\begin{aligned}\mathbb{P}\big[X^{(1)} \ge x\big] = {} & \frac{C}{\bar{A}_1 + \cdots + \bar{A}_{\bar{n}}} (\log x)^{-(1+\bar{n})/2} x^{\sum_{i=1}^{\bar{n}} \bar{A}_i (\log \sum_{j=1}^{\bar{n}} \bar{A}_j / |\bar{A}_i| + \bar{\mu}_i)} \\ & \times \exp\left\{-\frac{1}{2}(\log^2 x) \sum_{j=1}^{\bar{n}} \bar{A}_j\right\} \big(1 + \mathrm{O}\big((\log x)^{-1}\big)\big),\end{aligned} \tag{54}$$

where

$$C = \exp\left\{-\frac{1}{2} \sum_{i,m=1}^{\bar{n}} \bar{a}_{im} \left(\log \frac{\sum_{j=1}^{\bar{n}} \bar{A}_j}{|\bar{A}_i|} + \bar{\mu}_i\right)\left(\log \frac{\sum_{j=1}^{\bar{n}} \bar{A}_j}{|\bar{A}_m|} + \bar{\mu}_m\right)\right\} \times \frac{1}{\sqrt{2\pi|\bar{\mathfrak{B}}|}} \sqrt{\frac{\sum_{j=1}^{\bar{n}} \bar{A}_j}{\prod_{i=1}^{\bar{n}} |\bar{A}_i|}}.$$

Proof. We will only sketch the proof, which is very similar to that of Theorem 1. If $\bar{n} = n$, the result follows from Lemma 2. Next, assume that $k \in \{\bar{n}, \dots, n-1\}$, $x > 1$, and let a, b be such that $x = \mathrm{e}^a - \mathrm{e}^b$. On the one hand, clearly,

$$\mathbb{P}\big[\mathrm{e}^{Y_1} \ge \mathrm{e}^{Y_2} + \cdots + \mathrm{e}^{Y_n} + x\big] \le \mathbb{P}\big[\mathrm{e}^{Y_1} \ge \mathrm{e}^{Y_2} + \cdots + \mathrm{e}^{Y_{n-1}} + x\big].$$

On the other hand,

$$\begin{aligned}&\mathbb{P}\big[\mathrm{e}^{Y_1} \ge \mathrm{e}^{Y_2} + \cdots + \mathrm{e}^{Y_{k+1}} + x\big] \\ &\quad \ge \mathbb{P}\big[\mathrm{e}^{Y_2} + \cdots + \mathrm{e}^{Y_k} \le \mathrm{e}^{Y_1} - \mathrm{e}^a, Y_{k+1} \le b\big] \\ &\quad = \mathbb{P}\big[\mathrm{e}^{Y_2} + \cdots + \mathrm{e}^{Y_k} \le \mathrm{e}^{Y_1} - \mathrm{e}^a\big] - \mathbb{P}\big[\mathrm{e}^{Y_2} + \cdots + \mathrm{e}^{Y_k} \le \mathrm{e}^{Y_1} - \mathrm{e}^a, Y_{k+1} > b\big].\end{aligned}$$

Since $k \ge \bar{n}$, $\sum_{i=1}^k \bar{w}_i = 1$. Moreover, since $\bar{w}_1 > 0$, we may define $\tilde{w}_1 = \frac{1}{w_1}$ and $\tilde{w}_i = -\frac{w_i}{w_1}$ for $i = 1, \dots, k$; these weights are positive and satisfy $\sum_{i=1}^k \tilde{w}_i = 1$.

We have

$$\begin{aligned}&\mathbb{P}\big[\mathrm{e}^{Y_2} + \cdots + \mathrm{e}^{Y_k} \le \mathrm{e}^{Y_1} - \mathrm{e}^a, Y_{k+1} > b\big] \\ &\quad \le \mathbb{P}\big[\tilde{w}_1 \mathrm{e}^a + \tilde{w}_2 \mathrm{e}^{Y_2} + \cdots + \tilde{w}_k \mathrm{e}^{Y_k} \le \mathrm{e}^{Y_1}, Y_{k+1} > b\big] \\ &\quad \le \mathbb{P}[\tilde{w}_1 a + \tilde{w}_2 Y_2 + \cdots + \tilde{w}_k Y_k \le Y_1, Y_{k+1} > b] \\ &\quad = \mathbb{P}[a \le \bar{w}_1 Y_1 + \bar{w}_2 Y_2 + \cdots + \bar{w}_k Y_k, Y_{k+1} > b]\end{aligned}$$

$$
\begin{aligned}
&\le \mathbb{P}\left[a - \sum_{i=1}^{k} \bar{w}_i Y_i \le \alpha(Y_{k+1} - b)\right] \\
&= \mathbb{P}\left[\sum_{i=1}^{k} \bar{w}_i Y_i + \alpha Y_{k+1} \ge a + \alpha b\right] \\
&= N\left(\frac{-a - \alpha b + \mathbb{E}[\bar{\mathbf{w}}^{\perp}\mathbf{Y} + \alpha Y_{k+1}]}{\sqrt{\operatorname{Var}[\bar{\mathbf{w}}^{\perp}\mathbf{Y} + \alpha Y_{k+1}]}}\right)
\end{aligned}
$$

for all $\alpha > 0$.

Next, reasoning as in the proof of Theorem 1, let

$$
x_{k+1} = \frac{(\mathbf{e}^{k+1})^{\perp}\mathfrak{B}\bar{\mathbf{w}}}{\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}} < 1.
$$

Then,

$$
N\left(\frac{-a - \alpha b + \mathbb{E}[\bar{\mathbf{w}}^{\perp}\mathbf{Y} + \alpha Y_{k+1}]}{\sqrt{\operatorname{Var}[\bar{\mathbf{w}}^{\perp}\mathbf{Y} + \alpha Y_{k+1}]}}\right) = N\left(\frac{-a - \alpha b + \bar{\mathbf{w}}^{\perp}\boldsymbol{\mu} + \alpha\mu_{k+1}}{\sqrt{\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}(1 + 2\alpha x_{k+1} + \alpha^2\mathfrak{B}_{k+1,k+1}/(\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}))}}\right).
$$

Now, we take

$$
b = \frac{1}{2}(x_{k+1} + 1)\log x \quad \Rightarrow \quad a = \log\bigl(x + \mathrm{e}^{b}\bigr) = \log x + \log\bigl(1 + x^{-(1/2)(1-x_{k+1})}\bigr).
$$

Using these substitutions, we obtain

$$
\begin{aligned}
&\frac{-a - \alpha b + \bar{\mathbf{w}}^{\perp}\boldsymbol{\mu} + \alpha\mu_{k+1}}{\sqrt{\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}(1 + 2\alpha x_{k+1} + \alpha^2\mathfrak{B}_{k+1,k+1}/(\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}))}} \\
&\quad = \frac{-\log x(1 + (\alpha/2)(1 + x_{k+1})) - \log(1 + x^{-(1/2)(1-x_{k+1})}) + \bar{\mathbf{w}}^{\perp}\boldsymbol{\mu} + \alpha\mu_{k+1}}{\sqrt{\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}(1 + 2\alpha x_{k+1} + \alpha^2\mathfrak{B}_{k+1,k+1}/(\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}))}} \\
&\quad \le -\frac{\log x}{\sqrt{\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}}}\frac{1 + (\alpha/2)(1 + x_{k+1})}{\sqrt{1 + 2\alpha x_{k+1} + \alpha^2\mathfrak{B}_{k+1,k+1}/(\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}})}} + C_{k+1},
\end{aligned}
$$

where C_{k+1} is a constant independent of x. Next, reasoning as in the proof of Theorem 1, we see that there exist α small enough and $\varepsilon_{k+1} > 0$ such that for all $x > 1$,

$$
\frac{-a - \alpha b + \mathbb{E}[\bar{\mathbf{w}}^{\perp}\mathbf{Y} + \alpha Y_{k+1}]}{\sqrt{\operatorname{Var}[\bar{\mathbf{w}}^{\perp}\mathbf{Y} + \alpha Y_{k+1}]}} \le -\frac{\log x}{\sqrt{\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}}}(1 + \varepsilon_{k+1}) + C_{k+1}.
$$

The rest of the proof is similar to that of Theorem 1 modulo some trivial changes. □

Proof of Theorem 2. It is clear that the sets $\mathcal{P}_4$, $\mathcal{P}_3$, and $\mathcal{P}_2$ defined by (38), (39), and (40), respectively, are not empty.

Upper estimate. Fix a positive function φ such that $\varphi(x) \to 0$ as $x \to \infty$. Then we have

$$\begin{aligned}\mathbb{P}\big[X^{(m)} \geq x\big] \leq & \sum_{1\leq p\leq m} \mathbb{P}\big[X_p \geq \big(1-\varphi(x)\big)(X_{m+1}+\cdots+X_n+x)\big] \\ & + \sum_{1\leq p,q\leq m,\, p\neq q} \mathbb{P}\bigg[X_p \geq \frac{\varphi(x)}{m-1}(X_{m+1}+\cdots+X_n+x), \\ & \qquad X_q \geq \frac{\varphi(x)}{m-1}(X_{m+1}+\cdots+X_n+x)\bigg].\end{aligned} \tag{55}$$

Formula (55) can be established as follows. Let E_1, E_2, and F_i with $1 \leq i \leq m$, be random variables. Then it is not hard to prove that the following set theoretical inclusion holds:

$$\begin{aligned}\big\{F_1+\cdots+F_m \geq E_1+(m-1)E_2\big\} \subset & \bigcup_{p=1}^{m}\{F_p \geq E_1\} \\ & \cup \bigg[\bigcup_{1\leq p,q\leq m,\, p\neq q}\{F_p \geq E_2, F_q \geq E_2\}\bigg].\end{aligned} \tag{56}$$

Next, using (56) with

$$F_p = X_p, \qquad E_1 = \big(1-\varphi(x)\big)(X_{m+1}+\cdots+X_n+x),$$

and

$$E_2 = \frac{\varphi(x)}{m-1}(X_{m+1}+\cdots+X_n+x),$$

and taking into account the countable subadditivity of $\mathbb{P}$, we obtain (55).

To estimate the terms in the first sum in formula (55), we introduce the following probability measure:

$$\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} = \frac{e^{\log(1-\varphi(x))\mathbf{e}^p\mathfrak{B}^{-1}\mathbf{Y}}}{\mathbb{E}[e^{\log(1-\varphi(x))\mathbf{e}^p\mathfrak{B}^{-1}\mathbf{Y}}]}, \tag{57}$$

where $\mathbf{Y} = (Y_1, \ldots, Y_n)$, and $\mathbf{e}^p$ is the vector with pth component equal to 1 and all the other components equal to zero. Note that the measure $\widetilde{\mathbb{P}}$ depends on p. However, we omit the parameter p in the symbol $\widetilde{\mathbb{P}}$ for the sake of simplicity. It is not hard to see that under the probability $\widetilde{\mathbb{P}}$, we have $\mathbf{Y} \sim N(\boldsymbol{\mu} + \log(1-\varphi(x))\mathbf{e}^p, \mathfrak{B})$. In other words, the law of the random vector

$$\big(Y_p - \log\big(1-\phi(x)\big), Y_{m+1}, \ldots, Y_n\big)$$

under $\widetilde{\mathbb{P}}$ coincides with the law of the random vector $(Y_p, Y_{m+1}, \ldots, Y_n)$ under $\mathbb{P}$, which means that

$$\widetilde{\mathbb{P}}\big[X_p \geq \big(1-\varphi(x)\big)(X_{m+1}+\cdots+X_n+x)\big] = \mathbb{P}[X_p \geq X_{m+1}+\cdots+X_n+x]. \tag{58}$$

It follows from (57) that

$$\begin{aligned}\mathbb{P}\big[X_p \ge \big(1-\varphi(x)\big)(X_{m+1}+\cdots+X_n+x)\big]\\ &= \widetilde{\mathbb{E}}\bigg[\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\widetilde{\mathbb{P}}}\mathbf{1}_{\{X_p\ge(1-\varphi(x))(X_{m+1}+\cdots+X_n+x)\}}\bigg] \\ &= \mathbb{E}\big[\mathrm{e}^{\log(1-\varphi(x))\mathbf{e}^p\mathfrak{B}^{-1}\mathbf{Y}}\big]\widetilde{\mathbb{E}}\big[\mathrm{e}^{-\log(1-\varphi(x))\mathbf{e}^p\mathfrak{B}^{-1}\mathbf{Y}}\mathbf{1}_{\{X_p\ge(1-\varphi(x))(X_{m+1}+\cdots+X_n+x)\}}\big].\end{aligned} \tag{59}$$

Next, let r and q be positive numbers satisfying $\frac{1}{r}+\frac{1}{q}=1$. Then, using (59), (58) and Hölder's inequality, we obtain

$$\begin{aligned}&\mathbb{P}\big[X_p \ge \big(1-\varphi(x)\big)(X_{m+1}+\cdots+X_n+x)\big]\\ &\quad\le \mathbb{E}\big[\mathrm{e}^{\log(1-\varphi(x))\mathbf{e}^p\mathfrak{B}^{-1}\mathbf{Y}}\big]\widetilde{\mathbb{E}}\big[\mathrm{e}^{-r\log(1-\varphi(x))\mathbf{e}^p\mathfrak{B}^{-1}\mathbf{Y}}\big]^{1/r}\\ &\qquad\times\widetilde{\mathbb{P}}\big[X_p\ge\big(1-\varphi(x)\big)(X_{m+1}+\cdots+X_n+x)\big]^{1/q}\\ &\quad= \mathbb{E}\big[\mathrm{e}^{\log(1-\varphi(x))\mathbf{e}^p\mathfrak{B}^{-1}\mathbf{Y}}\big]^{1-1/r}\mathbb{E}\big[\mathrm{e}^{-(r-1)\log(1-\varphi(x))\mathbf{e}^p\mathfrak{B}^{-1}\mathbf{Y}}\big]^{1/r}\\ &\qquad\times\mathbb{P}[X_p\ge X_{m+1}+\cdots+X_n+x]^{1/q}\\ &\quad= \mathrm{e}^{(r-1)/2\log^2(1-\varphi(x))a_{pp}}\mathbb{P}\big[X_p\ge(X_{m+1}+\cdots+X_n+x)\big]^{1/q}.\end{aligned}$$

Next, set

$$\varphi(x)=\frac{1}{\log^2 x},\qquad r=r(x)=\log^3 x,\quad\text{and}\quad \frac{1}{q(x)}=1-\frac{1}{\log^3 x}. \tag{60}$$

Then we have

$$\begin{aligned}\exp\bigg\{\frac{r(x)-1}{2}\log^2\big(1-\varphi(x)\big)a_{pp}\bigg\} &= \exp\bigg\{\frac{\log^3(x)-1}{2}\log^2\big(1-\log^{-2}(x)\big)a_{pp}\bigg\}\\ &= 1+\mathrm{O}\bigg(\frac{1}{\log x}\bigg),\end{aligned}$$

and hence, by Theorem 3,

$$\begin{aligned}&\mathbb{P}\big[X_p\ge\big(1-\varphi(x)\big)(X_{m+1}+\cdots+X_n+x)\big]\\ &\quad\le \delta_{1,p}(\log x)^{\delta_{2,p}}x^{\delta_{3,p}}\exp\bigg\{-\frac{\log^2 x}{2\delta_{4,p}}\bigg\}\big(1+\mathrm{O}\big((\log x)^{-1}\big)\big)\end{aligned}$$

as $x\to\infty$.

It remains to estimate the terms in the second sum in (55). For any two integers p and q with $1\le p,q\le m$ and $p\ne q$, let $\Delta_{m,n}^{p,q}$ be the set of weights $\mathbf{w}\in\mathbb{R}^n$ with $w_i=0$ for $i=1,\dots,m$, $i\ne p$, $i\ne q$; $w_p\ge 0$, $w_q\ge 0$; $w_i\le 0$ for $i=m+1,\dots,n$; and $\sum w_i=1$. Recall that

$$\Delta_{m,n}^p=\big\{\mathbf{w}\in\Delta_{m,n}^{p,q}\colon\ w_q=0\big\}.$$

By Jensen's inequality, for any $\mathbf{w} \in \Delta_{m,n}^{p,q}$,

$$\begin{aligned}
&\mathbb{P}\Big[X_p \geq \frac{\varphi(x)}{m-1}(X_{m+1}+\cdots+X_n+x), X_q \geq \frac{\varphi(x)}{m-1}(X_{m+1}+\cdots+X_n+x)\Big] \\
&\leq \mathbb{P}\Big[\sum_{i=1}^n w_i \log X_i \geq (w_p+w_q)\log\frac{\varphi(x)}{m-1}+\log x\Big] \\
&\leq N\Big(-\frac{(w_p+w_q)\log\varphi(x)/(m-1)+\log x-\sum_{i=1}^n w_i\mu_i}{\sqrt{\mathbf{w}^{\perp}\mathfrak{B}\mathbf{w}}}\Big) \\
&= \exp\Big\{-\frac{\log^2 x}{2\mathbf{w}^{\perp}\mathfrak{B}\mathbf{w}}+\mathrm{O}(\log x\cdot\log\log x)\Big\}
\end{aligned} \tag{61}$$

as $x \to \infty$.

Since the matrix $\mathfrak{B}$ is invertible and positive definite, the mapping $\mathbf{w} \mapsto \mathbf{w}^{\perp}\mathfrak{B}\mathbf{w}$ is strictly convex. This implies that

$$\min_{\mathbf{w}\in\Delta_{m,n}^{p,q}} \mathbf{w}^{\perp}\mathfrak{B}\mathbf{w} < \max\Big(\min_{\mathbf{w}\in\Delta_{m,n}^{p}} \mathbf{w}^{\perp}\mathfrak{B}\mathbf{w}, \min_{\mathbf{w}\in\Delta_{m,n}^{q}} \mathbf{w}^{\perp}\mathfrak{B}\mathbf{w}\Big).$$

Since

$$\delta_{4,p} = \min_{\mathbf{w}\in\Delta_{m,n}^{p}} \mathbf{w}^{\perp}\mathfrak{B}\mathbf{w},$$

we conclude from the estimate in (61) that the terms in the second sum in formula (55) provide a negligible contribution to the asymptotics, so that

$$\begin{aligned}
\mathbb{P}\big[X^{(m)} \geq x\big] &\leq \sum_{p=1}^m \delta_{1,p}(\log x)^{\delta_{2,p}} x^{\delta_{3,p}} \exp\Big\{-\frac{\log^2 x}{2\delta_{4,p}}\Big\}\big(1+\mathrm{O}\big((\log x)^{-1}\big)\big) \\
&= \delta_1(\log x)^{\delta_2} x^{\delta_3} \exp\Big\{-\frac{\log^2 x}{\delta_4}\Big\}\big(1+\mathrm{O}\big((\log x)^{-1/2}\big)\big)
\end{aligned}$$

as $x \to \infty$.

Lower estimate. Let F_p, $0 \leq p \leq m$, be random variables. Then for every such p, the following inclusion is valid:

$$\{F_p \geq F_0\} \subset \bigcup_{q\neq p}\{F_p \geq F_0, F_q \geq F_0\} \cup \{F_p \geq F_0, F_q < F_0 \text{ for all } q \neq p\}. \tag{62}$$

In addition, the sets $\{F_p \geq F_0, F_q < F_0 \text{ for all } q \neq p\}$, $1 \leq p \leq m$, are disjoint. Now, setting $F_p = X_p$, $1 \leq p \leq m$, and $F_0 = X_{m+1}+\cdots+X_n+x$ in (62), we can easily derive the following lower bound for the probability of our interest:

$$\mathbb{P}\big[X^{(m)} \geq x\big] \geq \sum_{p=1}^{m} \mathbb{P}[X_p \geq X_{m+1} + \cdots + X_n + x]$$
$$- \sum_{1\leq p,q,\leq m, p\neq q} \mathbb{P}[X_p \geq X_{m+1} + \cdots + X_n + x, X_q \geq X_{m+1} + \cdots + X_n + x].$$

Similarly to the first part of the proof, we can now show that the terms in the second line make a negligible contribution to the limit. It follows that

$$\mathbb{P}\big[X^{(m)} \geq x\big] \geq \delta_1 (\log x)^{\delta_2} x^{\delta_3} \exp\left\{-\frac{\log^2 x}{2\delta_4}\right\}\big(1 + \mathrm{O}\big((\log x)^{-1/2}\big)\big)$$

as $x \to \infty$. □

4. Numerics

4.1. Implementation and domain of validity of the asymptotic formulas

Formulas (13) and (11), as well as formula (15) have a vertical asymptote at $x = 1$, which means that these formulas are not valid for $x \geq 1$ and in practice have very poor accuracy unless x is much smaller than one, which may correspond to very small probabilities. To partially alleviate this difficulty, we suggest to use for numerical computations the following natural modification of formula (13), where the asymptote is shifted towards the center of the distribution:

$$\begin{aligned}\mathbb{P}[X \leq x] = \widetilde{C}\left(\log\frac{1}{x} + \bar{\mathbf{w}}^{\perp}\boldsymbol{\mu} + \mathcal{E}(\bar{\mathbf{w}})\right)^{-(1+\bar{n})/2} \exp\left\{-\frac{(\log x - \bar{\mathbf{w}}^{\perp}\boldsymbol{\mu} - \mathcal{E}(\bar{\mathbf{w}}))^2}{2\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}}\right\} \\ \times \big(1 + \mathrm{O}\big(|\log x|^{-1}\big)\big) \qquad \text{as } x \to 0.\end{aligned} \tag{63}$$

This formula clearly has the same asymptotics as (13), however its domain of validity is larger and numerical experiments show that it has a better performance. The expression for the distribution is well defined for all x such that

$$x \leq x^* = \mathrm{e}^{\bar{\mathbf{w}}^{\perp}\boldsymbol{\mu} + \mathcal{E}(\bar{\mathbf{w}})}.$$

For example, assume that $Y_1, \ldots, Y_n$ are identically distributed with variance σ^2 and constant correlation ρ. Table 1 gives the values of the probability $\mathbb{P}[X \leq x^*]$ for different values of σ, ρ and n. Formulas (15), (41) and (43) can be modified in a similar manner. Note that the formulas for the conditional law of Corollaries 1, 3 and 4 do not contain a vertical asymptote due to the cancellation of the logarithmic singularities.

Table 1. Location of the vertical asymptote in formula (63) for various values of model parameters

σ	0.3	1.0	0.3	1.0	0.3	1.0	0.3	1.0
ρ	0.5	0.5	0.0	0.0	0.5	0.5	0.0	0.0
n	5	5	5	5	20	20	20	20
$\mathbb{P}[X \le x^*]$	0.47	0.40	0.40	0.24	0.46	0.38	0.27	0.041

4.2. Using the asymptotic formulas directly

To illustrate the performance of the asymptotic formulas of Theorems 1 and 2 numerically, we have taken a 4×4 covariance matrix with the following entries: $b_{ij} = \sigma_i \sigma_j \rho$ (constant correlation) where $\sigma = \{2, 2.3, 3, 3\}$. The distribution functions

$$\mathbb{P}[X \le x] = \mathbb{P}\left[\mathrm{e}^{Y_1} + \mathrm{e}^{Y_2} + \mathrm{e}^{Y_3} + \mathrm{e}^{Y_4} \le x\right]$$

and

$$\mathbb{P}\left[X^{(2)} \ge x\right] = \mathbb{P}\left[\mathrm{e}^{Y_1} + \mathrm{e}^{Y_2} - \mathrm{e}^{Y_3} - \mathrm{e}^{Y_4} \ge x\right]$$

have been computed first, using the asymptotic formulas given in Theorems 1 and 2, with the modification suggested in formula (63). The corresponding asymptotic approximations will be denoted below by $F_{\mathrm{a}}(x)$ and $F_{\mathrm{a}}^{(2)}(x)$, respectively. Then we evaluated Monte Carlo estimates $F_{\mathrm{mc}}(x)$ and $F_{\mathrm{mc}}^{(2)}(x)$ of these quantities (the Monte Carlo algorithm is described in detail later in this section). To evaluate the quality of the asymptotic approximation, we plot the ratios $\frac{F_{\mathrm{mc}}(x)}{F_{\mathrm{a}}(x)}$ and $\frac{F_{\mathrm{mc}}^{(2)}(x)}{F_{\mathrm{a}}^{(2)}(x)}$ for a wide range of values of x. These ratios, plotted as functions of $\log x$, are shown in Figure 1 for two values of the correlation coefficient ρ.

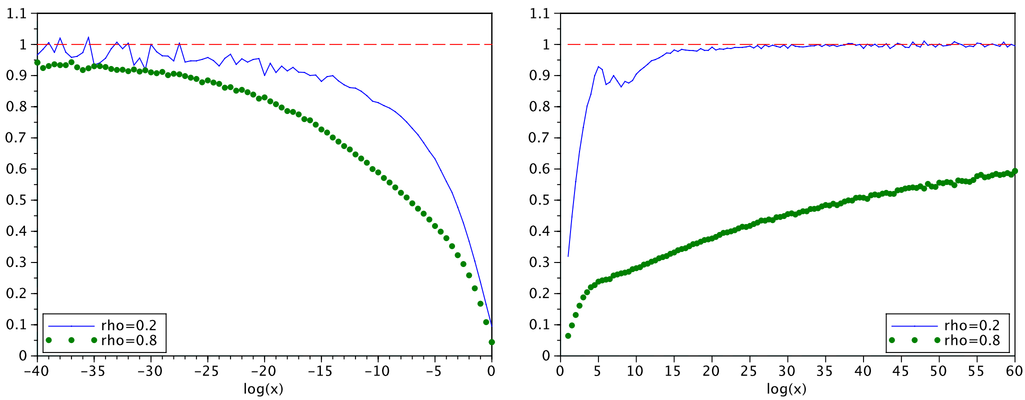

Figure 1. Ratios of the Monte Carlo estimate of the distribution function (survival function) to the estimate obtained using the asymptotic formulas. Left: $\mathbb{P}[X \le x]$. Right: $\mathbb{P}[X^{(2)} \ge x]$. The fluctuations in the curves are due to the Monte Carlo error.

In the evaluation of the asymptotic formula for $\mathbb{P}[X \le x]$, one needs to solve the quadratic programming problem formulated in (6). For the first value, $\rho = 0.2$, the solution to this problem is $\bar{\mathbf{w}} \approx \{0.44\ 0.30\ 0.13\ 0.13\}$. Thus, here we are in the setting of the "special case", where the asymptotics is obtained directly by Laplace's method (see Lemma 1). For the second value, $\rho = 0.8$, the solution is $\bar{\mathbf{w}} \approx \{0.83\ 0.17\ 0\ 0\}$, so only the first two components make a contribution to the asymptotics.

In the evaluation of the asymptotic formula for $\mathbb{P}[X^{(2)} \ge x]$, one needs to solve the problem in (36) twice, for $p = 1$ and $p = 2$, and compare the resulting minimum values. Here, for $\rho = 0.2$, the solutions are $\bar{\mathbf{w}}_1 = \bar{\mathbf{w}}_2 = \{1\ 0\ 0\}$, and $p = 2$ gives a larger minimum value, so that the asymptotic behavior of the distribution function is determined by the second component of the vector Y only. For $\rho = 0.8$, the solutions are $\bar{\mathbf{w}}_1 \approx \{1.32\ -0.16\ -0.16\}$ and $\bar{\mathbf{w}}_2 \approx \{1.1\ -0.05\ -0.05\}$, and once again, the minimum value is greater for $p = 2$. Therefore, in this case the asymptotic behavior is determined by the second, third and fourth components of $\mathbf{Y}$.

Analyzing Figure 1, one can make the following observations, which turn out to be rather generic:

- As expected, the ratio of the distribution functions converges to one, but this convergence is very slow. This observation is consistent with the logarithmic error bounds in Theorems 1 and 2.
- Although the ratio of the estimates converges to one very slowly, this ratio is never very far from one (compared to the value of the probability itself), which means that the asymptotic formula gives the right order of magnitude for a wide range of probabilities. For instance, for $\rho = 0.8$, the values of x, shown in the left graph, correspond to the range of probabilities from $\sim 5 \times 10^{-93}$ for $\log x = -40$ to 0.2 for $\log x = 0$.

4.3. Efficient Monte Carlo estimation of tail probabilities

As we have already seen, due to the slow convergence, the asymptotic formulas in Theorems 1 and 2 typically provide only order-of-magnitude approximations of the distribution function of the sum/difference of log-normal random variables. When a more precise estimate is needed, and the dimension n is large, one can use a Monte Carlo estimator. In such a case, as we will next explain, the asymptotic formulas can be utilized to construct very efficient variance reduction procedures. To save space, we will only discuss the case of distribution functions. Similar ideas can be used to reduce the variance of Monte Carlo estimates of densities, conditional expectations or other quantities of interest.

Left tail of X

For the distribution function $F(x) = \mathbb{P}[X \le x]$, the standard estimate is the following:

$$\widehat{F}_N(x) = \frac{1}{N}\sum_{k=1}^{N} \mathbf{1}_{\{\sum_{i=1}^n \exp\{Y_i^{(k)}\} \le x\}}, \tag{64}$$

where $\mathbf{Y}^{(1)}, \ldots, \mathbf{Y}^{(N)}$ are i.i.d. vectors with the law $N(\boldsymbol{\mu}, \mathfrak{B})$. However, this estimate is not a suitable approximation of the tail of the distribution function. Indeed, the variance of $\widehat{F}_N(x)$ is given by

$$\operatorname{Var}\widehat{F}_N(x)=\frac{F(x)-F^2(x)}{N}\sim\frac{F(x)}{N},\qquad x\to 0,$$

and the relative error, that is,

$$\frac{\sqrt{\operatorname{Var}\widehat{F}_N(x)}}{F(x)}\sim\frac{1}{\sqrt{NF(x)}},$$

explodes very quickly as $x\to 0$ (it behaves like $\mathrm{e}^{c\log^2 x}$ for some constant c). The usual way to reduce variance in the Gaussian context is via importance sampling. The idea is to rewrite the formula for F as follows:

$$F(x)=\mathbb{E}\left[\exp\left\{-\boldsymbol{\Lambda}^\perp\mathfrak{B}^{-1}(\mathbf{Y}-\boldsymbol{\mu})-\frac{1}{2}\boldsymbol{\Lambda}^\perp\mathfrak{B}^{-1}\boldsymbol{\Lambda}\right\}\mathbf{1}_{\{\sum_{i=1}^n\exp\{Y_i+\Lambda_i\}\le x\}}\right],$$

where $\boldsymbol{\Lambda}\in\mathbb{R}^n$ is a vector that will be chosen later. Note that if $\boldsymbol{\Lambda}=0$, then the standard estimate is recovered. The goal is to find a nonzero $\boldsymbol{\Lambda}$ such that the corresponding estimate

$$\widehat{F}_N^{\boldsymbol{\Lambda}}(x)=\frac{1}{N}\sum_{k=1}^N\exp\left\{-\boldsymbol{\Lambda}^\perp\mathfrak{B}^{-1}\big(Y^{(k)}-\boldsymbol{\mu}\big)-\frac{1}{2}\boldsymbol{\Lambda}^\perp\mathfrak{B}^{-1}\boldsymbol{\Lambda}\right\}\mathbf{1}_{\{\sum_{i=1}^n\exp\{Y_i^{(k)}+\Lambda_i\}\le x\}}\tag{65}$$

has variance smaller than that of the standard estimate.

Simple computations show that the variance of $\widehat{F}_N^{\boldsymbol{\Lambda}}(x)$ is given by

$$\begin{aligned}\operatorname{Var}\widehat{F}_N^{\boldsymbol{\Lambda}}(x)&=\frac{1}{N}\operatorname{Var}\left[\exp\left\{-\boldsymbol{\Lambda}^\perp\mathfrak{B}^{-1}(\mathbf{Y}-\boldsymbol{\mu})-\frac{1}{2}\boldsymbol{\Lambda}^\perp\mathfrak{B}^{-1}\boldsymbol{\Lambda}\right\}\mathbf{1}_{\{\sum_{i=1}^n\exp\{Y_i+\Lambda_i\}\le x\}}\right]\\&=\frac{1}{N}\big\{\mathbb{E}\big[\exp\big\{-2\boldsymbol{\Lambda}^\perp\mathfrak{B}^{-1}(\mathbf{Y}-\boldsymbol{\mu})-\boldsymbol{\Lambda}^\perp\mathfrak{B}^{-1}\boldsymbol{\Lambda}\big\}\mathbf{1}_{\{\sum_{i=1}^n\exp\{Y_i+\Lambda_i\}\le x\}}\big]-F^2(x)\big\}\\&=\frac{1}{N}\left\{\exp\left\{\frac{1}{2}\boldsymbol{\Lambda}^\perp\mathfrak{B}^{-1}\boldsymbol{\Lambda}\right\}\mathbb{E}\big[\exp\big\{-\boldsymbol{\Lambda}^\perp\mathfrak{B}^{-1}(\mathbf{Y}-\boldsymbol{\mu})\big\}\mathbf{1}_{\{\sum_{i=1}^n\exp\{Y_i\}\le x\}}\big]-F^2(x)\right\}\\&=\frac{1}{N}\left\{\exp\big\{\boldsymbol{\Lambda}^\perp\mathfrak{B}^{-1}\boldsymbol{\Lambda}\big\}\mathbb{P}\left[\sum_{i=1}^n\exp\{Y_i-\Lambda_i\}\le x\right]-F^2(x)\right\}.\end{aligned}$$

Let

$$V(\boldsymbol{\Lambda},x)=\exp\big\{\boldsymbol{\Lambda}^\perp\mathfrak{B}^{-1}\boldsymbol{\Lambda}\big\}\mathbb{P}\left[\sum_{i=1}^n\mathrm{e}^{Y_i-\Lambda_i}\le x\right].$$

Since $F(x)$ does not depend on $\boldsymbol{\Lambda}$, the optimal variance reduction is obtained by minimizing $V(\boldsymbol{\Lambda},x)$ as a function of $\boldsymbol{\Lambda}$. Our idea is to obtain an explicit estimate by replacing the probability in the previous expression by an asymptotically equivalent expression given in Theorem 1. In other words, we compute an approximation to the optimal $\mathbf{\Lambda}$ by minimizing

$$\widetilde{V}(\mathbf{\Lambda}, x) = \mathbf{\Lambda}^{\perp}\mathfrak{B}^{-1}\mathbf{\Lambda} - \frac{1}{2}\sum_{i,j=1}^{\bar{n}} \bar{a}_{ij}\left(\log\frac{\bar{A}_1 + \cdots + \bar{A}_{\bar{n}}}{\bar{A}_i} + \bar{\mu}_i - \log x - \Lambda_i\right)$$
$$\times\left(\log\frac{\bar{A}_1 + \cdots + \bar{A}_{\bar{n}}}{\bar{A}_j} + \bar{\mu}_j - \log x - \Lambda_j\right).$$

To obtain the expression above, we have omitted all the factors in the formula in Theorem 1, which do not depend on $\mathbf{\Lambda}$, and have also taken the logarithm of the resulting expression. The optimal value $\mathbf{\Lambda}^*$ of $\mathbf{\Lambda}$ can be found by solving the following system of equations:

$$\frac{\partial\widetilde{V}}{\partial\Lambda_i}(\mathbf{\Lambda}^*) = 0, \qquad 1 \le i \le n.$$

This system can be rewritten as follows:

$$2\sum_{j=1}^{n} a_{ij}\Lambda_j^* + \sum_{j=1}^{\bar{n}} \bar{a}_{ij}\left(\log\frac{\bar{A}_1 + \cdots + \bar{A}_{\bar{n}}}{\bar{A}_j} + \bar{\mu}_j - \log x - \Lambda_j\right) = 0, \qquad 1 \le i \le \bar{n},$$
$$2\sum_{j=1}^{n} a_{ij}\Lambda_j^* = 0, \qquad \bar{n} < i \le n.$$

Applying the matrix $\mathfrak{B}$ to the previous system, we obtain

$$2\Lambda_k^* + \sum_{i,j=1}^{\bar{n}} b_{ki}\bar{a}_{ij}\left(\log\frac{\bar{A}_1 + \cdots + \bar{A}_{\bar{n}}}{\bar{A}_j} + \bar{\mu}_j - \log x - \Lambda_j^*\right) = 0, \tag{66}$$

for all $k = 1, \ldots, n$. When $k \le \bar{n}$, the formula in (66) simplifies to

$$\Lambda_k^* + \log\frac{\bar{A}_1 + \cdots + \bar{A}_{\bar{n}}}{\bar{A}_k} + \bar{\mu}_k - \log x = 0.$$

Substituting this into (66), we see that for all k, the optimal value Λ_k^* is given by

$$\Lambda_k^* = \sum_{i,j=1}^{\bar{n}} b_{ki}\bar{a}_{ij}\left(\log x - \log\frac{\bar{A}_1 + \cdots + \bar{A}_{\bar{n}}}{\bar{A}_j} - \bar{\mu}_j\right). \tag{67}$$

Note that since the optimal vector $\mathbf{\Lambda}^*$ depends on x, we cannot apply Theorem 1 directly to characterize the asymptotic behavior of the function $V(\mathbf{\Lambda}^*, x)$ as $x \to 0$. Nevertheless, this function can be estimated from above by using Jensen's inequality as follows:

$$
\begin{aligned}
V\left(\boldsymbol{\Lambda}^*, x\right) &\le \mathrm{e}^{\boldsymbol{\Lambda}^{*\perp}\mathfrak{B}^{-1}\boldsymbol{\Lambda}^*}\mathbb{P}\left[\sum_{i=1}^{n}\bar{w}_i\left(Y_i - \Lambda_i^* - \log\bar{w}_i\right) \le \log x\right] \\
&= \mathrm{e}^{\boldsymbol{\Lambda}^{*\perp}\mathfrak{B}^{-1}\boldsymbol{\Lambda}^*}N\left(\frac{\bar{\mathbf{w}}^{\perp}\boldsymbol{\Lambda}^* + \log x + \sum_{i=1}^{n}\bar{w}_i(\log\bar{w}_i - \mu_i)}{\sqrt{\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}}}\right),
\end{aligned}
$$

where $\bar{\mathbf{w}}$ is the solution of (6) and N is the standard normal distribution function. Substituting the expression in (67) for $\boldsymbol{\Lambda}^*$ and using (9), we obtain

$$
\bar{\mathbf{w}}^{\perp}\boldsymbol{\Lambda}^* = \log x + \sum_{i=1}^{n}\bar{w}_i(\log\bar{w}_i - \mu_i)
$$

and

$$
\begin{aligned}
\boldsymbol{\Lambda}^{*\perp}\mathfrak{B}^{-1}\boldsymbol{\Lambda}^* &= \sum_{i,j=1}^{\bar{n}}\bar{a}_{ij}(\log x + \log\bar{w}_i - \bar{\mu}_i)(\log x + \log\bar{w}_j - \bar{\mu}_j) \\
&= \frac{1}{\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}}\left\{\log x + \sum_{i=1}^{n}\bar{w}_i(\log\bar{w}_i - \mu_i)\right\}^2 \\
&\quad + \sum_{i,j=1}^{\bar{n}}\bar{a}_{ij}(\log\bar{w}_i - \bar{\mu}_i)(\log\bar{w}_j - \bar{\mu}_j) - \frac{1}{\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}}\left\{\sum_{i=1}^{n}\bar{w}_i(\log\bar{w}_i - \mu_i)\right\}^2.
\end{aligned}
$$

Therefore, as $x \to 0$,

$$
V\left(\boldsymbol{\Lambda}^*, x\right) \lesssim C\frac{\exp\{-(1/(\bar{\mathbf{w}}^{\perp}\mathfrak{B}\bar{\mathbf{w}}))\{\log x + \sum_{i=1}^{n}\bar{w}_i(\log\bar{w}_i - \mu_i)\}^2\}}{\log 1/x},
$$

where the constant C is independent of x. Comparing the previous estimate with the asymptotics of $F(x)$ (see formula (13)), we see that, for a different constant C,

$$
V\left(\boldsymbol{\Lambda}^*, x\right) \lesssim CF^2(x)\left(\log\frac{1}{x}\right)^{\bar{n}} \tag{68}
$$

as $x \to 0$. This means in particular that our estimator is *logarithmically efficient* in the sense of Asmussen and Glynn [5], Section VI.1.

To test the performance of the proposed variance reduction algorithm, we have computed the Monte Carlo estimates with and without variance reduction for different levels x, using the same numerical values of the parameters as above. Table 2 shows the relative error of the estimate (65), where the value of $\boldsymbol{\Lambda}^*$ is given by (67) as well as the ratio of the *standard deviation* of the estimate (64) to that of the estimate (65). The relative errors appear quite stable and the reduction factors

Table 2. Relative errors of the variance reduction estimate (65) with the value of $\mathbf{\Lambda}^*$ given by (67), and the factors by which the *standard deviation* of the estimate is reduced with the variance reduction algorithm

$\rho = 0.2$				$\rho = 0.8$			
x	$\mathbb{P}[X \le x]$	rel. error	red. factor	x	$\mathbb{P}[X \le x]$	rel. error	red. factor
0.006738	0.0000027	0.43%	152.8	0.0002035	0.0000012	0.27%	269
0.01831	0.0000424	0.37%	38.07	0.0009119	0.0000331	0.26%	69.08
0.04979	0.0004639	0.32%	14.48	0.004089	0.0005282	0.25%	16.07
0.1353	0.003457	0.28%	6.188	0.01832	0.005085	0.25%	5.312
0.3679	0.01798	0.24%	3.152	0.08209	0.02998	0.26%	2.256
1	0.06603	0.20%	1.845	0.3679	0.1141	0.27%	1.078

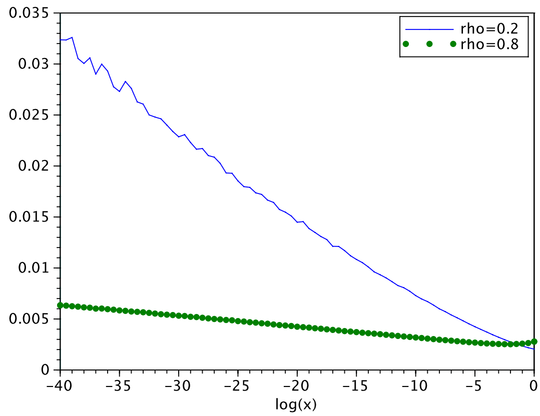

Figure 2. Relative error of the variance reduction estimate (65) with the value of $\mathbf{\Lambda}^*$ given by (67).

are greater than one for all values of x and in general quite spectacular, ranging from 4–5 for not so small probabilities of order of 1% to hundreds for probabilities of order of 10^{-6}.

Figure 2 shows the relative error of the estimate (65) with the value of $\mathbf{\Lambda}^*$ given by (67) (the standard deviation divided by the value of the estimate, computed over 10^6 trajectories). As shown by the theoretical analysis of the variance, the relative error grows only logarithmically in x, which means that even for very small probabilities (such as 10^{-100}), our estimator requires a reasonable number of trajectories to obtain adequate precision.

Additional tests

To evaluate the robustness of our variance reduction method with respect to the choice of the parameters of the model and the number of variables n, we have performed additional tests, assuming this time that $Y_1, \dots, Y_n$ are identically distributed with law $N(0, \sigma^2)$ and have a constant correlation ρ. Table 3 shows the *standard deviation* reduction factors for various values of σ, ρ

Table 3. *Standard deviation* reduction factors for the additional tests. The probability $\mathbb{P}[X \le x]$ is approximately equal to 10^{-3} for all tests

σ	0.3	1.0	0.3	1.0	0.3	1.0	0.3	1.0
ρ	0.5	0.5	0.0	0.0	0.5	0.5	0.0	0.0
n	5	5	5	5	20	20	20	20
x	2.5	0.55	3.4	1.6	10.5	2.7	16.9	14.3
red. factor	15.7	14.7	14.0	10.1	15.2	14.2	11.4	4.8

and n. For each test, the value of x was selected so that the probability $\mathbb{P}[X \le x]$ is approximately equal to 10^{-3} (belongs to the interval $(0.9 \times 10^{-3}, 1.1 \times 10^{-3})$). We see that in all tests but one, the standard deviation is reduced by a factor greater than 10, which means that, for equal precision, the computation would be accelerated by a factor greater than 100.

Right tail of a log-normal difference

In this case, the standard estimate of the survival function has the form

$$\widehat{F}_N(x) = \frac{1}{N}\sum_{k=1}^{N} \mathbf{1}_{\{\sum_{i=1}^m \exp\{Y_i^{(k)}\} - \sum_{i=m+1}^n \exp\{Y_i^{(k)}\} \ge x\}}, \tag{69}$$

and the alternative estimate which may potentially reduce variance is as follows:

$$\begin{aligned}\widehat{F}_N^{\mathbf{\Lambda}}(x) = \frac{1}{N}\sum_{k=1}^{N} \exp\left\{-\mathbf{\Lambda}^{\perp}\mathfrak{B}^{-1}\left(\mathbf{Y}^{(k)} - \boldsymbol{\mu}\right) - \frac{1}{2}\mathbf{\Lambda}^{\perp}\mathfrak{B}^{-1}\mathbf{\Lambda}\right\} \\ \times \mathbf{1}_{\{\sum_{i=1}^m \exp\{Y_i^{(k)}+\Lambda_i\} - \sum_{i=m+1}^n \exp\{Y_i^{(k)}+\Lambda_i\} \ge x\}}.\end{aligned} \tag{70}$$

To find the optimal value of $\mathbf{\Lambda}$, we need to minimize

$$\exp\{\mathbf{\Lambda}^{\perp}\mathfrak{B}^{-1}\mathbf{\Lambda}\}\mathbb{P}\left[\sum_{i=1}^{m}\exp\{Y_i^{(k)} - \Lambda_i\} - \sum_{i=m+1}^{n}\exp\{Y_i^{(k)} - \Lambda_i\} \ge x\right],$$

and once again, the main idea is to minimize the asymptotic approximation to this function, given in Theorem 2. Assuming for simplicity that the set $\mathcal{P}_4$ defined in (38) is a singleton, $\mathcal{P}_4 = \{p\}$, the problem reduces to that of minimizing the following function:

$$\begin{aligned}\widetilde{V}(\mathbf{\Lambda}, x) = \mathbf{\Lambda}^{\perp}\mathfrak{B}^{-1}\mathbf{\Lambda} - \frac{1}{2}\sum_{i,j=1}^{\bar{n}^{(p)}} \bar{a}_{ij}^{(p)}\left(\log\frac{\bar{A}_1^{(p)} + \cdots + \bar{A}_{\bar{n}^{(p)}}^{(p)}}{|\bar{A}_i^{(p)}|} + \bar{\mu}_i^{(p)} - \log x - \Lambda_{\bar{k}^{(p)}(i)}\right) \\ \times\left(\log\frac{\bar{A}_1^{(p)} + \cdots + \bar{A}_{\bar{n}^{(p)}}^{(p)}}{|\bar{A}_j^{(p)}|} + \bar{\mu}_j^{(p)} - \log x - \Lambda_{\bar{k}^{(p)}(j)}\right).\end{aligned}$$

Table 4. Relative errors of the variance reduction estimate (70) with the value of $\mathbf{\Lambda}^*$ given by (71), and those of the plain Monte Carlo estimate (69), whenever available

$\rho = 0.2$				$\rho = 0.8$			
x	$\mathbb{P}[X^{(2)} \geq x]$	rel. error	rel. error, plain MC	x	$\mathbb{P}[X^{(2)} \geq x]$	rel. error	rel. error, plain MC
e	0.2672	0.153%	0.166%	e	0.1121	0.74%	0.281%
e^5	0.01564	2.37%	00.792%	e^5	2.134×10^{-3}	4.26%	2.17%
e^{10}	6.771×10^{-6}	57.1%	29.5%	e^{10}	3.759×10^{-7}	1.15%	376%
e^{15}	3.459×10^{-11}	0.274%	–	e^{15}	9.765×10^{-13}	0.44%	–
e^{20}	1.724×10^{-18}	0.318%	–	e^{20}	2.654×10^{-20}	0.502%	–
e^{25}	8.050×10^{-28}	0.358%	–	e^{25}	6.872×10^{-30}	0.561%	–

Next, reasoning as in the proof of (67), we see that the optimal value $\mathbf{\Lambda}^*$ of $\mathbf{\Lambda}$ is given by

$$\Lambda_k^* = \sum_{i,j=1}^{\bar{n}^{(p)}} b_{k,\bar{k}^{(p)}(i)} \bar{a}_{ij}^{(p)} \left(\log x - \log \frac{\bar{A}_1^{(p)} + \cdots + \bar{A}_{\bar{n}^{(p)}}^{(p)}}{|\bar{A}_j^{(p)}|} - \bar{\mu}_j^{(p)} \right). \tag{71}$$

However, here the computation remains only heuristic, since there is no simple upper bound for the variance of the estimator with the optimal $\mathbf{\Lambda}^*$.

Numerical tests (see Table 4) are much less conclusive than those for the left tail presented in the previous paragraph. For moderate values of x, the algorithm does not lead to any variance reduction and may even increase variance. However, very far in the tail, when the probability in question is so small that it cannot be computed with the conventional Monte Carlo estimator in reasonable time, the variance reduction estimator becomes very efficient. We conclude that for the case of log-normal differences, our variance reduction algorithm can be potentially very useful for the simulation of extremely rare events (with probability smaller than 10^{-6}), but further research and further improvements to the algorithm are necessary before it can be used in the context of not-so-rare events, such as those with probability of 10^{-2}–10^{-3} arising, for example, in the Value at Risk calculations in financial risk management.

5. Risk management in the multidimensional Black–Scholes model

The tail estimates obtained in this paper can be applied to risk management problems in the context of the n-dimensional Black–Scholes model. Suppose that the dynamics of the asset price vector $\mathbf{S}_t = (S^1, \dots, S^n)$ is described by the following n-dimensional stochastic process:

$$\log S_t = \log \mathbf{S}_0 + \boldsymbol{\theta} t - \frac{\operatorname{diag}(\mathfrak{B})t}{2} + \mathfrak{B}^{1/2}\mathbf{W}_t, \tag{72}$$

where $\mathbf{W}$ is an n-dimensional standard Brownian motion, $\mathfrak{B}$ is the covariance matrix, $\boldsymbol{\theta}$ is the drift vector, and $\operatorname{diag}(\mathfrak{B})$ stands for the main diagonal of $\mathfrak{B}$.

Consider a portfolio containing the assets S^i, $1 \le i \le n$ with weights $\xi_1, \dots, \xi_n$, and the price process X defined by

$$X_t = \sum_{i=1}^{n} \xi_i S_t^i, \qquad t \ge 0. \tag{73}$$

The initial condition for the process X is given by $X_0 = \sum_{i=1}^{n} \xi_i S_0^i$. The process X can be alternatively expressed as $X_t = \sum_{i=1}^{n} \operatorname{sgn} \xi_i \exp\{Y_t^i\}$, where

$$Y_t^i = \log S_0^i + \log|\xi_i| + \theta_i t - \frac{b_{ii} t}{2} + \sum_{j=1}^{n} \gamma_{ij} W_t^j, \qquad 1 \le i \le n. \tag{74}$$

In (74), the symbols γ_{ij} stand for the elements of the matrix $\mathfrak{B}^{1/2}$. We also set

$$\mu_{i,t} = \log S_0^i + \log|\xi_i| + \theta_i t - \frac{b_{ii} t}{2}, \qquad 1 \le i \le n. \tag{75}$$

In the sequel, t will be fixed, and the asymptotic formulas obtained in Sections 2 and 3 will be applied to the random variable X_t defined in (73). The Gaussian data associated with the case described above are given by the following: the mean vector is $\vee\mu = (\mu_{1,t}, \dots, \mu_{n,t})$ and the covariance matrix is $t\mathfrak{B}$.

For the purposes of risk management, it is important to solve two classes of problems in relation with the portfolio X:

- Quantify the behavior of one portfolio in specific adverse scenarios of market evolution, which are typically defined in terms of another portfolio (the benchmark). This can be done using our characterization of the asymptotic behavior of a Gaussian vector conditionally on the sum or difference of exponentials of its components (Corollaries 3 and 4). We address this issue in detail in the following paragraph.
- Evaluate various risk measures for the portfolio X, such as the probability of loss of a given magnitude or the Value at Risk (the quantile function). The probability of loss may be approximated using the asymptotic formulas of Section 2 (for portfolios with only positive weights) or Section 3 (for portfolios with both positive and negative weights). The asymptotic behavior of the Value at Risk when the confidence level tends to one is characterized in Section 5.2.

5.1. Behavior of log-normal portfolios under adverse scenarios

Suppose that an investor holds a portfolio containing assets $S^1, \dots, S^n$ with weights $v_1, \dots, v_n$. The value of such a portfolio is given by

$$V_t = \sum_{i=1}^{n} v_i S_t^i. \tag{76}$$

The 1996 Market Risk Amendment to Basel I [8] as well as Basel II and Basel III Capital Accords require banks and investment firms to conduct stress tests to determine their ability to respond to adverse market events. These adverse scenarios are typically defined in terms of the performance of a certain benchmark and correspond to a stylized version of certain crisis events observed in the past. We will next describe some examples of plausible stress scenarios and explain how the corresponding benchmark process X can be defined.

- *Equity market fall of a certain magnitude*. This is the most common stress scenario. The benchmark process $\{X\}_{t\geq 0}$ under such a scenario is the normalized market index, having the initial value 1, and the adverse event is $\{X_t = x\}$ for some $t > 0$ and x which is supposed small. The weights ξ_i are then positive and equal to the normalized market capitalizations of the stocks.
- *A certain difference in performance between the equity markets of two geographical areas or two sectors*. For instance, one may assume that the American markets outperform the European ones, or that small capitalization shares outperform large capitalization ones. Let $X^a_t = \sum_{i=1}^m \xi_i S^i_t$ be the market index of the first area, where $\xi_1, \dots, \xi_m$ are the positive market capitalization weights of the stocks $S^1, \dots, S^m$, and $X^b_t = \sum_{i=m+1}^n \xi_i S^i_t$ be the market index of the second area. The stress scenario is the event

$$\left\{\frac{X^a_t}{X^a_0} - \frac{X^b_t}{X^b_0} = x\right\}.$$

  This can be dealt with in our framework by taking

$$X_t = \sum_{i=1}^m \frac{\xi_i}{X^a_0} S^i_t - \sum_{i=m+1}^n \frac{\xi_i}{X^b_0} S^i_t$$

  with the stress scenario $\{X_t = x\}$. Here the value of x is large.
- *A certain difference in performance between two benchmarks*. The investor may be interested, for example, in the event when her portfolio severely underperforms the market. This is similar to the case considered above, except that the two benchmarks may contain the same stocks. Let the two benchmarks be given by $X^a_t = \sum_{i=1}^n \xi^a_i S^i_t$ and $X^b_t = \sum_{i=1}^n \xi^b_i S^i_t$. We are once again interested in the stress scenario $\{\frac{X^a_t}{X^a_0} - \frac{X^b_t}{X^b_0} = x\}$. This is equivalent to taking

$$X_t = \sum_{i=1}^n \left\{\frac{\xi^a_i}{X^a_0} - \frac{\xi^b_i}{X^b_0}\right\} S^i_t$$

  and using the stress scenario $\{X_t = x\}$ with x large.

Our next goal is to characterize the asymptotic behavior of various conditional expected values for the portfolio with the price process given by (76) under the stress scenarios described above. This can be done for the individual stocks or for the entire portfolio. In the former case, we approximate the conditional probabilities of the form

$$e_i(t,x) = \mathbb{E}\left[S_t^i \,\middle|\, \sum_{k=1}^n \xi_k S_t^k = x\right], \tag{77}$$

while in the latter case we deal with the following conditional probabilities:

$$\mathbb{E}[V_t | X_t = x] = \sum_{i=1}^n v_i e_i(t,x). \tag{78}$$

The quantities $e_i(t,S)$ can be estimated using formulas (16) and (42) for the conditional Laplace transform, since

$$e_i(t,x) = \frac{1}{\xi_i}\mathbb{E}\left[\exp\{Y_t^i\} \,\middle|\, \sum_{k=1}^n \exp\{Y_t^k\} = x\right] \tag{79}$$

for all $1 \le i \le n$. The following results are thus direct consequences of Corollaries 1 and 4.

Theorem 4. *Suppose that the weights $\xi_1,\dots,\xi_n$ are positive and that assumption $(\mathcal{A})$ holds for the covariance matrix $\mathfrak{B}$. Then the following are true*:

1. *If $1 \le i \le \bar{n}$, then as $x \to 0$,*

$$\begin{aligned} e_i(t,x) &= \frac{x}{\xi_i}\frac{\bar{A}_i}{\bar{A}_1 + \cdots + \bar{A}_{\bar{n}}}\left(1 + \mathrm{O}\left(\left(\log\frac{1}{x}\right)^{-1}\right)\right) \\ &= x\frac{\bar{w}_i}{\xi_i}\left(1 + \mathrm{O}\left(\left(\log\frac{1}{x}\right)^{-1}\right)\right). \end{aligned}$$

2. *If $\bar{n} < i \le n$, then as $x \to 0$,*

$$\begin{aligned} e_i(t,x) = x^{\sum_{j=1}^{\bar{n}} \bar{A}_j b_{ij}} S_0^i \exp&\left\{\theta_i t - \sum_{p,q=1}^{\bar{n}} b_{pi}\bar{a}_{pq}\left(\log\frac{\bar{A}_1 + \cdots + \bar{A}_{\bar{n}}}{\bar{A}_q} + \mu_{q,t}\right)\right\} \\ &\times \exp\left\{-\frac{t}{2}\sum_{p,q=1}^{\bar{n}} \bar{a}_{pq} b_{pi} b_{qi}\right\}\left(1 + \mathrm{O}\left(\left(\log\frac{1}{x}\right)^{-1}\right)\right). \end{aligned}$$

Remark 8. Since $\sum_{j=1}^{\bar{n}} \bar{A}_j b_{ij} > 1$ for $i \notin \bar{I}$ (see Remark 3), it follows from Theorem 4 that the assets in the market index can be classified into two categories, depending on the behavior of their conditional expectation under the conditional law.

- "Safe assets", whose conditional expectations decay proportionally to the value x of the market index. Those are exactly the assets, which enter the Markowitz minimal variance portfolio (solution of problem (6)) with strictly positive weights.

- "Dangerous assets", whose conditional expectations decay faster than the index.

The next assertion concerns the second and the third typical stress scenarios described above.

Theorem 5. *Suppose that for $m \le n$ the weights $\xi_1, \dots, \xi_m$ are positive and $\xi_{m+1}, \dots, \xi_n$ are negative, that assumption $(\mathcal{A}_1^i)$ holds for matrix $\mathfrak{B}$ with every $i = 1, \dots, m$, and that the set $\mathcal{P}_4$ defined in (38) is a singleton, $\mathcal{P}_4 = \{p\}$. Then the following are true.*

1. *If $i \in I^{(p)}$, then as $x \to +\infty$,*

$$e_i(t,x) = \frac{x}{\xi_i} \frac{|\bar{A}_i^{(p)}|}{\sum_{j=1}^{\bar{n}^{(p)}} \bar{A}_j^{(p)}} \times \left(1 + \mathrm{O}\left(\left(\log\frac{1}{x}\right)^{-1}\right)\right)$$
$$= x\frac{|\bar{w}_i|}{\xi_i} \times \left(1 + \mathrm{O}\left(\left(\log\frac{1}{x}\right)^{-1}\right)\right).$$

2. *If $i \notin I^{(p)}$, then as $x \to +\infty$,*

$$e_i(t,x) = S_0^i x^{\sum_{j=1}^{\bar{n}^{(p)}} \bar{A}_j^{(p)} b_{\bar{k}^{(p)}(j),i}}$$
$$\times \exp\left\{\theta_i t - \sum_{j,k=1}^{\bar{n}^{(p)}} \bar{a}_{jk}^{(p)} b_{\bar{k}^{(p)}(j),i} \left(\log\frac{\sum_{l=1}^{\bar{n}^{(p)}} \bar{A}_l^{(p)}}{|\bar{A}_k^{(p)}|} + \mu_{\bar{k}^{(p)}(k),t}\right)\right\}$$
$$\times \exp\left\{-\frac{t}{2}\sum_{j,k=1}^{\bar{n}^{(p)}} \bar{a}_{jk}^{(p)} b_{\bar{k}^{(p)}(j),i} b_{\bar{k}^{(p)}(k),i}\right\}\left(1 + \mathrm{O}\left(\left(\log\frac{1}{x}\right)^{-1}\right)\right).$$

Remark 9. It follows from assumption $(\mathcal{A}_1^p)$ and Remark 7 that for $i \notin I^{(p)}$,

$$\sum_{j=1}^{\bar{n}^{(p)}} \bar{A}_j^{(p)} b_{\bar{k}^{(p)}(j),i} < 1.$$

Therefore, the assets in the benchmark can once again be classified into the following two categories:

- Those assets, whose conditional expectations, given the stress scenario, grow proportionally to x. This category includes exactly one asset among $S_1, \dots, S_m$, that one with the highest relative asymptotic variance with respect to $S_{m+1}, \dots, S_n$. It may or may not include some assets among $S_{m+1}, \dots, S_n$.
- Those assets, whose conditional expectations, given the stress scenario, grow slower than x.

In other words, the fact that the portfolio $S^1 + \dots + S^m$ strongly outperforms the portfolio $S^{m+1} + \dots + S^n$ can be attributed asymptotically to a very strong performance of a single stock among $S^1, \dots, S^m$, which may be partially offset by the performance of some stocks from the second group.

5.2. Log-normal portfolios and Value at Risk

Our goal in this subsection is to find a sharp asymptotic formula for the Value at Risk (VaR_α) of the portfolio described in Section 5. The price X_t at time t for this portfolio is defined by (73). We study the case where the confidence level α tends to one, and restrict ourselves to the portfolios with only positive weights. The case of portfolios with both positive and negative weights can be handled similarly.

For a portfolio, the value at risk VaR_α, $0<\alpha<1$, over the time period $t>0$ is defined as the smallest number k such that the probability of a loss greater than k over the time interval t is equal to α. It is not hard to see that

$$\mathrm{VaR}_\alpha = \inf\bigl\{k\colon\ \mathbb{P}(X_t \le X_0 - k) = 1-\alpha\bigr\}.$$

The next theorem provides an asymptotic formula for the function $\alpha \mapsto \mathrm{VaR}_\alpha$ as the confidence level α tends to one.

Theorem 6. *Suppose assumption* $(\mathcal{A})$ *holds for the covariance matrix* $\mathfrak{B}$. *Then the following asymptotic formula is valid*:

$$\begin{aligned}\mathrm{VaR}_\alpha = X_0 - \exp\Biggl\{&-\sqrt{\frac{2t}{\bar{A}_1+\cdots+\bar{A}_{\bar{n}}}\log\frac{1}{1-\alpha}} \\ &+ \frac{\sum_{k=1}^{\bar{n}}\bar{A}_k(\log(\bar{A}_1+\cdots+\bar{A}_{\bar{n}})/\bar{A}_k+\mu_{k,t})}{\bar{A}_1+\cdots+\bar{A}_{\bar{n}}}\Biggr\} \\ \times\Biggl(1+\mathrm{O}\Biggl(&\frac{\log\log 1/(1-\alpha)}{\sqrt{\log 1/(1-\alpha)}}\Biggr)\Biggr)\end{aligned} \tag{80}$$

as $\alpha \to 1$.

Proof. Let us fix $t>0$, and denote by F_t^{-1} the generalized inverse function of the function $F_t(x)=\mathbb{P}(X_t\le x)$. Then we have

$$\mathrm{VaR}_\alpha = X_0 - F_t^{-1}(1-\alpha). \tag{81}$$

Therefore, in order to characterize the asymptotic behavior of the function $\alpha\mapsto\mathrm{VaR}_\alpha$ as $\alpha\to1$, it suffices to find an asymptotic formula for the function $y\mapsto F_t^{-1}(y)$ as $y\to0$.

We will first study the asymptotics near zero of the inverse function F^{-1} of any function F, having the following form:

$$F(x) = c_1\left(\frac{1}{x}\right)^{c_2}\left(\log\frac{1}{x}\right)^{c_3}\exp\left\{-c_4\log^2\frac{1}{x}\right\}\left(1+\mathrm{O}\left(\left(\log\frac{1}{x}\right)^{-1}\right)\right) \tag{82}$$

as $x\to0$. It is assumed in (82) that the constants satisfy the following conditions: $c_1>0$, $c_2\in\mathbb{R}$, $c_3\in\mathbb{R}$, and $c_4>0$. We also assume the continuity and the invertibility of the function F near zero.

Lemma 3. *Under the previous restrictions, the following asymptotic formula holds as* $y \to 0$:

$$F^{-1}(y) = \exp\{-\sqrt{\phi(y)}\}\left(1 + \mathrm{O}\left(\left(\log\frac{1}{y}\right)^{-1}\right)\right), \tag{83}$$

where

$$\phi(y) = \frac{1}{c_4}\log\frac{1}{y} + c_2 c_4^{-3/2}\sqrt{\log\frac{1}{y}} + \frac{c_3}{2c_4}\log\log\frac{1}{y}\left(\frac{c_2^2}{2c_4^2} + \frac{c_3}{2c_4}\log\frac{1}{c_4} + \frac{1}{c_4}\log c_1\right) + \frac{c_2 c_3}{4}c_4^{-3/2}\frac{\log\log 1/y}{\sqrt{\log 1/y}}.$$

Using the Taylor formula for the function $u \mapsto \sqrt{1+u}$ with two terms, we obtain a simpler formula from (83).

Corollary 6. *The following asymptotic formula holds*:

$$F^{-1}(y) = \exp\left\{-\sqrt{\frac{1}{c_4}\log\frac{1}{y}} - \frac{c_2}{2c_4}\right\}\left(1 + \mathrm{O}\left(\frac{\log\log 1/y}{\sqrt{\log 1/y}}\right)\right) \tag{84}$$

as $y \to 0$.

Note that formula (84) uses only the constants c_2 and c_4.

Proof of Lemma 3. Let $y > 0$, and let $F(u_y) = y$. Then $F^{-1}(y) = u_y$. Next, using (82), we obtain

$$\log\frac{1}{y} = c_4\log^2\frac{1}{u_y} - c_2\log\frac{1}{u_y} - c_3\log\log\frac{1}{u_y} - \log c_1 + \mathrm{O}\left(\left(\log\frac{1}{u_y}\right)^{-1}\right) \tag{85}$$

as $y \to 0$. The previous formula implies the following two-sided estimate:

$$a_1\sqrt{\log\frac{1}{y}} \le \log\frac{1}{u_y} \le a_2\sqrt{\log\frac{1}{y}}, \qquad 0 < y < y_0,$$

for some constants $a_1 > 0$ and $a_2 > 0$.

Put $z_y = \log^2\frac{1}{u_y}$. Then formula (85) gives

$$z_y = \frac{1}{c_4}\log\frac{1}{y} + \frac{c_2}{c_4}\sqrt{z_y} + \frac{c_3}{2c_4}\log z_y + \frac{1}{c_4}\log c_1 + \mathrm{O}\left(\left(\log\frac{1}{y}\right)^{-1/2}\right) \tag{86}$$

as $y \to 0$.

Our next goal is to use iterations in formula (86). We will replace any occurrence of z_y on the right-hand side of (86) by the whole expression on the right-hand side of (86). The following simple formulas will be needed in the sequel: $\log(1+s) = \mathrm{O}(s)$ and $\sqrt{1+s} = 1 + \frac{1}{2}s + \mathrm{O}(s^2)$ as $s \to 0$. Let us put

$$h = \frac{c_3}{2c_4}\log z_y + \frac{1}{4}\log c_1 + \mathrm{O}\left(\left(\log\frac{1}{y}\right)^{-1/2}\right), \tag{87}$$

where the O-term is the same as in formula (86). We have

$$\begin{aligned} \log z_y &= \log\left(\frac{1}{c_4}\log\frac{1}{y} + \frac{c_2}{c_4}\sqrt{z_y} + h\right) \\ &= \log\frac{1}{c_4} + \log\log\frac{1}{y} + \mathrm{O}\left(\left(\log\frac{1}{y}\right)^{-1/2}\right) \end{aligned} \tag{88}$$

as $y \to 0$. Moreover,

$$\begin{aligned} \sqrt{z_y} &= \frac{1}{\sqrt{c_4}}\sqrt{\log\frac{1}{y}}\sqrt{1 + \frac{c_2\sqrt{z_y}}{\log 1/y} + \frac{c_4 h}{\log 1/y}} \\ &= \frac{1}{\sqrt{c_4}}\sqrt{\log\frac{1}{y}} + \frac{c_2}{2\sqrt{c_4}}\frac{\sqrt{z_y}}{\sqrt{\log 1/y}} + \frac{\sqrt{c_4}h}{2\sqrt{\log 1/y}} + \mathrm{O}\left(\left(\log\frac{1}{y}\right)^{-1/2}\right) \end{aligned}$$

as $y \to 0$. Next, we iterate again, and obtain

$$\begin{aligned} \sqrt{z_y} = \frac{1}{\sqrt{c_4}}\sqrt{\log\frac{1}{y}} & \\ + \frac{c_2}{2\sqrt{c_4}\sqrt{\log 1/y}} & \\ \times \Bigg[\frac{1}{\sqrt{c_4}}\sqrt{\log\frac{1}{y}} &+ \frac{c_2\sqrt{z_y}}{2\sqrt{c_4}\sqrt{\log 1/y}} + \frac{\sqrt{c_4}h}{2\sqrt{\log 1/y}} + \mathrm{O}\left(\left(\log\frac{1}{y}\right)^{-1/2}\right)\Bigg] \\ + \frac{\sqrt{c_4}\hat{h}}{2\sqrt{\log 1/y}} &+ \mathrm{O}\left(\left(\log\frac{1}{y}\right)^{-1/2}\right) \end{aligned} \tag{89}$$

as $y \to 0$. In (89), the symbol $\hat{h}$ stands for the result of substituting the expression for $\log z_y$ given in (88) into formula (87). It is not hard to see that

$$\frac{\sqrt{c_4}\hat{h}}{2\sqrt{\log 1/y}} = \frac{c_3}{4\sqrt{c_4}}\frac{\log\log 1/y}{\sqrt{\log 1/y}} + \mathrm{O}\left(\left(\log\frac{1}{y}\right)^{-1/2}\right) \tag{90}$$

as $y \to 0$. Now, taking into account (89) and (90), we get

$$\sqrt{z_y} = \frac{1}{\sqrt{c_4}}\sqrt{\log\frac{1}{y}} + \frac{c_2}{2c_4} + \frac{c_3}{4c_4}\frac{\log\log 1/y}{\sqrt{\log 1/y}} + \mathrm{O}\left(\left(\log\frac{1}{y}\right)^{-1/2}\right) \tag{91}$$

as $y \to 0$. Finally, we can estimate z_y. Using (86), (88), and (91), we see that

$$
\begin{aligned}
z_y &= \frac{1}{c_4}\log\frac{1}{y} + c_2 c_4^{-3/2}\sqrt{\log\frac{1}{y}} + \frac{c_3}{2c_4}\log\log\frac{1}{y}\left(\frac{c_2^2}{2c_4^2} + \frac{c_3}{2c_4}\log\frac{1}{c_4} + \frac{1}{c_4}\log c_1\right) \\
&\quad + \frac{c_2 c_3}{4} c_4^{-3/2}\frac{\log\log 1/y}{\sqrt{\log 1/y}} + \mathrm{O}\left(\left(\log\frac{1}{y}\right)^{-1/2}\right) \qquad (92) \\
&= \phi(y) + \mathrm{O}\left(\left(\log\frac{1}{y}\right)^{-1/2}\right)
\end{aligned}
$$

as $y \to 0$.

We will next find an asymptotic formula for the function F^{-1}. We will use the formula

$$F^{-1}(y) = u_y = \exp\{-\sqrt{z_y}\} \qquad (93)$$

and the following simple lemma.

Lemma 4. *Let $z_y = \phi(y) + \mathrm{O}(\psi(y))$ as $y \to 0$, where the functions ϕ and ψ are positive and such that $\phi(y) \to \infty$ and $\frac{\psi(y)}{\sqrt{\phi(y)}} \to 0$ as $y \to 0$. Then*

$$u_y = \exp\{-\sqrt{\phi(y)}\}\left(1 + \mathrm{O}\left(\frac{\psi(y)}{\sqrt{\phi(y)}}\right)\right)$$

as $y \to 0$.

Proof of Lemma 4. We have

$$
\begin{aligned}
-\sqrt{z_y} &= -\sqrt{\phi(y) + \mathrm{O}(\psi(y))} = -\sqrt{\phi(y)}\sqrt{1 + \mathrm{O}\left(\frac{\psi(y)}{\phi(y)}\right)} \\
&= -\sqrt{\phi(y)}\left(1 + \mathrm{O}\left(\frac{\psi(y)}{\phi(y)}\right)\right),
\end{aligned}
$$

and hence,

$$
\begin{aligned}
u_y &= \exp\{-\sqrt{z_y}\} = \exp\{-\sqrt{\phi(y)}\}\exp\left\{\mathrm{O}\left(\frac{\psi(y)}{\sqrt{\phi(y)}}\right)\right\} \\
&= \exp\{-\sqrt{\phi(y)}\}\left(1 + \mathrm{O}\left(\frac{\psi(y)}{\sqrt{\phi(y)}}\right)\right).
\end{aligned}
$$

□

Now, taking into account (93), (92), and Lemma 4 with the function ϕ such as in (92) and the function ψ given by $\psi(y) = (\log\frac{1}{y})^{-1/2}$, we establish Lemma 3. □

We are now ready to complete the proof of Theorem 6. Applying Theorem 1 to the random variable X_t given by (73), we see that condition (82) holds. Note that

$$c_2 = -\frac{1}{t}\sum_{k=1}^{\bar{n}} \bar{A}_k\left(\log\frac{\bar{A}_1+\cdots+\bar{A}_{\bar{n}}}{\bar{A}_k} + \bar{\mu}_{m,t}\right) \quad \text{and} \quad c_4 = \frac{\bar{A}_1+\cdots+\bar{A}_{\bar{n}}}{2t}.$$

Now, it is easy to see that (81) and Corollary 6 imply formula (80). □

Acknowledgements

We would like to thank the two anonymous referees and the editor in chief Eric Moulines, whose insightful remarks lead to a substantial improvement of the paper.

References

[1] Aitchison, J. and Brown, J.A.C. (1957). *The Lognormal Distribution, with Special Reference to Its Uses in Economics*. Cambridge: Cambridge Univ. Press. MR0084966

[2] Albrecher, H., Asmussen, S. and Kortschak, D. (2012). Tail asymptotics for dependent subexponential variables. *Sib. Math. J.* **53** 965–983.

[3] Albrecher, H., Asmussen, S. and Kortschak, D. (2006). Tail asymptotics for the sum of two heavy-tailed dependent risks. *Extremes* **9** 107–130. MR2329800

[4] Asmussen, S., Blanchet, J., Juneja, S. and Rojas-Nandayapa, L. (2011). Efficient simulation of tail probabilities of sums of correlated lognormals. *Ann. Oper. Res.* **189** 5–23. MR2833609

[5] Asmussen, S. and Glynn, P.W. (2007). *Stochastic Simulation: Algorithms and Analysis. Stochastic Modelling and Applied Probability* **57**. New York: Springer. MR2331321

[6] Asmussen, S. and Rojas-Nandayapa, L. (2008). Asymptotics of sums of lognormal random variables with Gaussian copula. *Statist. Probab. Lett.* **78** 2709–2714. MR2465111

[7] Avellaneda, M., Boyer-Olson, D., Busca, J. and Friz, P. (2003). Application of large deviation methods to the pricing of index options in finance. *C. R. Math. Acad. Sci. Paris* **336** 263–266. MR1968270

[8] Bank for International Settlements (1996). Amendment to the capital accord to incorporate market risks.

[9] Barouch, E. and Kaufman, G.M. (1976). On sums of log-normal random variables. Working paper, no. 831-76. M.I.T. Alfred P. Sloan School of Management.

[10] Barouch, E., Kaufman, G.M. and Glasser, M.L. (1986). On sums of lognormal random variables. *Stud. Appl. Math.* **75** 37–55. MR0850438

[11] Bayer, C., Friz, P. and Laurence, P. On the probability density function of baskets. Preprint. Available at arXiv:1306.2793.

[12] Bayer, C. and Laurence, P. (2014). Asymptotics beats Monte-Carlo: The case of correlated local vol baskets. *Comm. Pure Appl. Math.* **67** 1618–1657.

[13] Bleistein, N. and Handelsman, R.A. (1995). *Asymptotic Expansions of Integrals*. New York: Holt, Rinehart and Winston.

[14] Carmona, R. and Durrleman, V. (2003). Pricing and hedging spread options. *SIAM Rev.* **45** 627–685. MR2047349

[15] Crow, E.L. and Shimizu, K., eds. (1988). *Lognormal Distributions: Theory and Applications. Statistics: Textbooks and Monographs* **88**. New York: Dekker. MR0939191

[16] Dufresne, D. (2004). The log-normal approximation in financial and other computations. *Adv. in Appl. Probab.* **36** 747–773. MR2079912

[17] Embrechts, P. and Puccetti, G. (2010). Bounds for the sum of dependent risks having overlapping marginals. *J. Multivariate Anal.* **101** 177–190. MR2557627

[18] Foss, S., Korshunov, D. and Zachary, S. (2011). *An Introduction to Heavy-Tailed and Subexponential Distributions*. *Springer Series in Operations Research and Financial Engineering*. New York: Springer. MR2810144

[19] Foss, S. and Richards, A. (2010). On sums of conditionally independent subexponential random variables. *Math. Oper. Res.* **35** 102–119. MR2676758

[20] Gao, X., Xu, H. and Ye, D. (2009). Asymptotic behavior of tail density for sum of correlated lognormal variables. *Int. J. Math. Math. Sci.* Art. ID 630857, 28. MR2533549

[21] Geluk, J. and Tang, Q. (2009). Asymptotic tail probabilities of sums of dependent subexponential random variables. *J. Theoret. Probab.* **22** 871–882. MR2558656

[22] Johnson, N.L. and Kotz, S. (1970). *Distributions in Statistics*. *Continuous Univariate Distributions*. 1. Boston, MA: Houghton Mifflin Co. MR0270475

[23] Kahane, J.-P. (1985). Sur le chaos multiplicatif. *Ann. Sci. Math. Québec*. **9** 105–150. MR0829798

[24] Ko, B. and Tang, Q. (2008). Sums of dependent nonnegative random variables with subexponential tails. *J. Appl. Probab.* **45** 85–94. MR2409312

[25] Kortschak, D. and Albrecher, H. (2009). Asymptotic results for the sum of dependent non-identically distributed random variables. *Methodol. Comput. Appl. Probab.* **11** 279–306. MR2511246

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

[27] Lo, C.F. (2012). The sum and difference of two lognormal random variables. *J. Appl. Math.* **2012** 838397.

[28] Rhodes, R. and Vargas, V. Gaussian multiplicative chaos and applications: A review. Preprint. Available at arXiv:1305.6221.

[29] Rojas-Nandayapa, L. (2008). Risk probabilities: Asymptotics and simulation. Ph.D. dissertation, Department of Mathematical Sciences, Faculty of Sciences, University of Aarhus.

[30] Senaratne, D. and Tellambura, C. (2009). Numerical computation of the log-normal sum distribution. In *GLOBECOM* 2009 1–6.

[31] Szyszkowicz, S.S. and Yanikomeroglu, H. (2007). On the tails of the distribution of the sum of lognormals. In *Proceedings of IEEE International Conference on Communications* (*ICC'*07).

[32] Tellambura, C. (2008). Bounds on the distribution of a sum of correlated log-normal random variables and their application. *IEEE Trans. Commun.* **56** 1241–1248.

[33] Tellambura, C. and Senaratne, D. (2010). Accurate computation of the MGF of the log-normal distribution and its application to sum of log-normals. *IEEE Trans. Commun.* **58** 1568–1577.

[34] Vanduffel, S., Chen, X., Dhaene, J., Goovaerts, M., Henrard, L. and Kaas, R. (2008). Optimal approximations for risk measures of sums of lognormals based on conditional expectations. *J. Comput. Appl. Math.* **221** 202–218. MR2458764

[35] Yuen, K.C. and Yin, C. (2012). Asymptotic results for tail probabilities of sums of dependent and heavy-tailed random variables. *Chin. Ann. Math. Ser. B* **33** 557–568. MR2996531

Received October 2013 and revised May 2014