On the Efficient Simulation of the Left-Tail of the Sum of Correlated Log-normal Variates

Mohamed-Slim Alouini, Nadhir Ben Rached, Abla Kammoun and Raul Tempone

Abstract. The sum of Log-normal variates is encountered in many challenging applications such as performance analysis of wireless communication systems and financial engineering. Several approximation methods have been reported in the literature. However, these methods are not accurate in the tail regions. These regions are of primordial interest as small probability values have to be evaluated with high precision. Variance reduction techniques are known to yield accurate, yet efficient, estimates of small probability values. Most of the existing approaches have focused on estimating the right-tail of the sum of Log-normal random variables (RVs). Here, we instead consider the left-tail of the sum of correlated Log-normal variates with Gaussian copula, under a mild assumption on the covariance matrix. We propose an estimator combining an existing mean-shifting importance sampling approach with a control variate technique. This estimator has an asymptotically vanishing relative error, which represents a major finding in the context of the left-tail simulation of the sum of Log-normal RVs. Finally, we perform simulations to evaluate the performances of the proposed estimator in comparison with existing ones.

Keywords. Sum of correlated Log-normal, small probability values, variance reduction techniques, left-tail of the sum of correlated Log-normal variates, importance sampling, control variate, asymptotically vanishing relative error.

1 Introduction

The Log-normal distribution is encountered in many applications such as financial engineering [13] and performance analysis of wireless communications systems in which it was shown to be a good fit for realistic propagation channels [15, 20, 24]. Therefore, investigating the distribution of sums of Log-normal random variables (RVs) is of primordial practical interest. In fact, the pricing of Asian or basket options is closely related to the distribution of the sum of Log-normal variates [13]. Another application using the distribution of sums of Log-normal variates is the evaluation of the value at risk, defined as $(1-\alpha)$ quantile of the loss distribution, for a sufficiently small value of $\alpha$ [4]. In wireless communication systems, the distribution of sums of Log-normal RVs is of major practical interest in the problem of evaluating the outage probability at the output of receivers using
diversity techniques [9].

The distribution of the sums of Log-normal variates is not known in a closed form. This has led researchers to propose various approximation techniques [7, 11, 12, 14, 23]. However, as illustrated in [7], these approximations are not always reliable especially in the tail regions, which is what we are interested in.

Variance reduction Monte Carlo (MC) methods constitute good alternatives to efficiently estimate tail probabilities of the distribution of sums of Log-normal RVs. Numerous studies on the distribution of sums of Log-normal variates can be found in the literature. As previously mentioned, most of them focus on the right-tail region instead of the left-tail one. Over the years, several efficient variance reduction techniques have been developed to estimate the right-tail of the distribution of sums of independent Log-normal RVs [5, 8, 18]. In the correlated setting as well, efficient simulation approaches of the right-tail of sums of correlated Log-normal variates have been proposed [1, 10, 21].

The estimation of the distribution of the left-tail of sums of Log-normal RVs has received less interest than that of the right-tail region. Estimating the left-tail of the distribution of sums of Log-normal variates is motivated by the problem of evaluating outage probabilities in the performance analysis of wireless communication systems operating over a Log-normal fading environment. It is only recently that researchers have shifted their attention to the left-tail region [4, 9, 17]. In fact, the authors in [4] have used the well-known importance sampling (IS) technique (i.e. exponential twisting technique [22]) and have proved that the proposed estimator achieves the asymptotic optimality property under the independent and identically distributed (i.i.d) assumption. However, given the fact that the application of the exponential twisting technique requires knowledge of the moment generating function (MGF), which is out-of-reach for the Log-normal distribution, the authors in [4] have instead considered an estimator of the MGF [3]. In another study, two unified IS approaches have been introduced in [9] using the well-known hazard rate twisting technique [8, 18]. In particular, for the Log-normal setting, the first estimator was shown to achieve the asymptotic optimality criterion under the assumption of independent and not necessarily identically distributed sums of Log-normal RVs. The asymptotic optimality property holds again using the second IS scheme under the i.i.d assumption. Comparisons between the estimator presented in [4] and the two latter estimators have been performed in [9]. The sole work the authors are aware of it that deals with the efficient simulation of the left-tail of the sum of correlated Log-normal variates is in [17]. In fact, an IS estimator has been proposed by shifting the mean of the corresponding multivariate normal distribution. This estimator was shown to achieve the asymptotic optimality property under a mild assumption.

In the present paper, we aim to further improve the mean-shifting IS estimator
More precisely, we consider the left-tail simulation of sums of correlated Log-normal variates, i.e. the probability that a sum of correlated Log-normal variates is less than a sufficiently small threshold, and we provide an improved estimator of the IS scheme proposed in [17]. Our methodology is based on the results presented in [17] to describe the asymptotic behavior and considers the case where the left-tail of the distribution of the sum is determined by only one dominant component. In this setting, we improve the IS estimator of [17] by combining it with a control variate type of variance reduction technique. The introduced control variate is a function of the dominant component that characterizes the tail behavior of the left-tail of the sum. The main result of our study is that the improved estimator using the proposed control variate technique has the asymptotically vanishing relative error property which is the most desired property in the context of rare event simulations. This result represents a valuable contribution to the problem of estimating the left-tail of sums of Log-normal RVs since as it was mentioned above the existing estimators in the corresponding literature were only proved to achieve a weaker property; namely the asymptotic optimality criterion [4, 9, 17]. Simulation results show that our proposed approach yields a substantial amount of variance reduction compared to the mean-shifting IS approach of [17], a reduction which increases as we decrease the probability of interest, i.e. as we decrease the threshold value to zero.

The rest of the paper is organized as follows. In section II, we describe the problem setting and we review the concepts of IS and control variate. Section III is devoted to presenting the IS scheme of [17]. The main idea of the paper is provided in Section IV where we show how to improve the IS approach of [17] through the use of control variate. In the same section, the main result proving the asymptotically vanishing relative error property is provided. Finally, some selected simulation results are provided in Section V to assess the performance of the proposed approach which achieves a substantial amount of computational gain over the mean shifting IS of [17].

\section{Problem Setting}

We consider a random vector $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_N)^t$ with a $N$-dimensional multivariate normal distribution, with a mean vector $\mu$ and a covariance matrix $\Sigma$. We assume that $\Sigma$ is positive definite. The probability density function (PDF) of the random vector $\mathbf{Y}$ is given by:

$$f(\mathbf{y}) = \frac{\exp \left( -\frac{1}{2} (\mathbf{y} - \mu)^t \Sigma^{-1} (\mathbf{y} - \mu) \right)}{\sqrt{(2\pi)^N |\Sigma|}},$$

(2.1)
where $|\Sigma|$ is the determinant of the matrix $\Sigma$ and $y^t$ is the transpose of the vector $y$. Similarly to [17], we define the following quantities: the elements of $\Sigma$ and its inverse $\Sigma^{-1}$ will be denoted by $\Sigma_{ij}$ and $\Sigma_{ij}^{-1}$, respectively. Moreover, we define $A_k = \sum_{j=1}^{N} \Sigma_{kj}^{-1}$. Let $\bar{w}$ be the unique solution of

$$
\min_{\Delta_N} w^t \Sigma w
$$

with $\Delta_N = \{w \in \mathbb{R}^N \text{ such that } w_i \geq 0, i = 1, 2, \cdots, N, \text{ and } \sum_{i=1}^{N} w_i = 1\}$. Let us denote by $\bar{n}$ the cardinal of the set $\{1, 2, \cdots, N, \text{ such that } \bar{w}_i \neq 0\}$ and $\bar{I} = \{\bar{k}(1), \cdots, \bar{k}(\bar{n})\}$ the corresponding indexes. Then, we denote by $\bar{\mu} \in \mathbb{R}^{\bar{n}}$ the vector with entries $\bar{\mu}_i = \mu_{\bar{k}(i)}$ and $\bar{\Sigma}$ the $\bar{n} \times \bar{n}$ matrix with elements $\bar{\Sigma}_{ij} = \Sigma_{\bar{k}(i)\bar{k}(j)}$. Moreover, we denote by $\bar{\Sigma}^{-1}$ the inverse of $\bar{\Sigma}$ with elements $\bar{\Sigma}_{ij}^{-1}$. Finally, the row sums of $\bar{\Sigma}^{-1}$ is denoted by $\bar{A}_k = \sum_{j=1}^{\bar{n}} \bar{\Sigma}_{kj}^{-1}$. As it was mentioned in [17], we assume without loss of generality that $\bar{I} = \{1, 2, \cdots, \bar{n}\}$.

Let us consider the Log-normal random vector $X = (X_1, X_2, \cdots, X_N)^t$ so that $Y_i = \log(X_i), i = 1, 2, \cdots, N$. Our goal is to efficiently estimate the left-tail of the sum of $X_i$’s, i.e. the probability that the sum of correlated Log-normal RVs with Gaussian copula falls below a sufficiently small threshold:

$$
\alpha(\gamma_{th}) = P_f \left( \sum_{i=1}^{N} X_i \leq \gamma_{th} \right).
$$

(2.3)

As mentioned previously, the above quantity is of paramount practical interest in the performance evaluation of wireless communication systems as it corresponds to the probability that the communication system is in outage, i.e. the probability that the system fails to operate correctly. The simplest method to estimate $\alpha(\gamma_{th})$ is to use the naive MC estimator. However, it is well-known that the naive estimator is computationally expensive when estimating rare events, i.e. regions of sufficiently small values of $\alpha(\gamma_{th})$ which is the focus of our study.

To construct computationally efficient estimators, variance reduction techniques represent good alternatives [19]. When used appropriately, variance reduction techniques are known to yield accurate estimates with less computational effort than with the naive MC sampler. Here, we consider two instances of variance reduction techniques, IS and control variate.

### 2.1 Review of IS

IS has been extensively used for the efficient simulation of rare events [1, 4, 8, 18]. The main idea of IS is to introduce an IS distribution $g(\cdot)$ under which sampling
is performed, instead of the original PDF $f(\cdot)$. The probability of interest $\alpha(\gamma_{th})$ can then be re-written as follows:

$$\alpha(\gamma_{th}) = \mathbb{E}_f \left[ 1_{\left( \sum_{i=1}^{N} \exp(Y_i) \leq \gamma_{th} \right)} \right]$$

$$= \mathbb{E}_g \left[ 1_{\left( \sum_{i=1}^{N} \exp(Y_i) \leq \gamma_{th} \right)} L \left( Y_1, Y_2, \ldots, Y_N \right) \right], \quad \text{(2.4)}$$

where $1_{(\cdot)}$ represents the indicator function, and $\mathbb{E}_f[\cdot]$ and $\mathbb{E}_g[\cdot]$ are the expectation operators under the PDFs $f(\cdot)$ and $g(\cdot)$, respectively. $L(\cdot)$ is the likelihood ratio which is defined as:

$$L \left( Y_1, Y_2, \ldots, Y_N \right) = \frac{f \left( Y_1, Y_2, \ldots, Y_N \right)}{g \left( Y_1, Y_2, \ldots, Y_N \right)}.$$ \quad \text{(2.5)}

Following the above probability change of measure, the IS estimator is defined as:

$$\hat{\alpha}_{IS}(\gamma_{th}) = \frac{1}{M} \sum_{k=1}^{M} 1_{\left( \sum_{i=1}^{N} \exp(Y_i^{(k)}) \leq \gamma_{th} \right)} L \left( Y_1^{(k)}, Y_2^{(k)}, \ldots, Y_N^{(k)} \right), \quad \text{(2.6)}$$

where $\{(Y_1^{(k)}, Y_2^{(k)}, \ldots, Y_N^{(k)})\}_{k=1}^{M}$ are independent realizations of the random vector $Y$ under the PDF $g(\cdot)$. The crucial step when using IS is the choice of the IS distribution $g(\cdot)$ that results in a variance reduction. In fact, a good IS distribution encourages the sampling of important realizations, that is samples that belong to the rare set $\{\sum_{i=1}^{N} \exp(Y_i) \leq \gamma_{th}\}$, and also tries to maintain the likelihood ratio $L(\cdot)$ constant in the rare region.

### 2.2 Review of Control Variate

Control variate is also a variance reduction technique that can, if adequately used, yield a substantial amount of variance reduction. This method can be combined with IS in order to further reduce the variance. Let us suppose we want to estimate $\alpha = \mathbb{E}_g \left[ T_{\gamma_{th}}(Y) \right]$ where $T_{\gamma_{th}}(Y) = 1_{\left( \sum_{i=1}^{N} \exp(Y_i) \leq \gamma_{th} \right)} L \left( Y_1, \ldots, Y_N \right)$ is an IS estimator and let $Z_{\gamma_{th}}(Y)$ be a control variate with a known expected value $P(\gamma_{th})$. The idea of control variate technique is to consider the following estimator of $\alpha(\gamma_{th})$

$$T_{\gamma_{th}}'(Y) = T_{\gamma_{th}}(Y) + \beta \left( Z_{\gamma_{th}}(Y) - P(\gamma_{th}) \right), \quad \text{(2.7)}$$

where $\beta \in \mathbb{R}$. Obviously, the above estimator is an unbiased estimator of $\alpha(\gamma_{th})$. The variance of $T_{\gamma_{th}}'(Y)$ is given by the expression:

$$\text{var}_g \left[ T_{\gamma_{th}}'(Y) \right] = \text{var}_g \left[ T_{\gamma_{th}}(Y) \right] + 2 \beta \text{cov}_g \left[ T_{\gamma_{th}}(Y), Z_{\gamma_{th}}(Y) \right]$$

$$+ \beta^2 \text{var}_g \left[ Z_{\gamma_{th}}(Y) \right]. \quad \text{(2.8)}$$
Hence, the optimal value of $\beta$ is the value that minimizes the variance of $T'_{\gamma_{th}}(Y)$ and obtained, through a simple computation, by:

$$
\beta^* = -\frac{\text{cov}_g [T_{\gamma_{th}}(Y), Z_{\gamma_{th}}(Y)]}{\text{var}_g [Z_{\gamma_{th}}(Y)]}.
$$

(2.9)

By plugging the optimal value $\beta^*$ into (2.8), we easily find that the variance of $T'_{\gamma_{th}}(Y)$ is given by

$$
\text{var}_g [T'_{\gamma_{th}}(Y)] = \left(1 - \rho_{T_{\gamma_{th}}(Y), Z_{\gamma_{th}}(Y)}^2\right) \text{var}_g [T_{\gamma_{th}}(Y)],
$$

(2.10)

where $\rho_{T_{\gamma_{th}}(Y), Z_{\gamma_{th}}(Y)}$ is the correlation coefficient between the two RVs $T_{\gamma_{th}}(Y)$ and $Z_{\gamma_{th}}(Y)$. From the above result, we conclude that, in order to further reduce the variance of $T_{\gamma_{th}}(Y)$, the RV $Z_{\gamma_{th}}(Y)$ has to be selected so that it is highly correlated with the RV $T_{\gamma_{th}}(Y)$. The estimator of $\alpha(\gamma_{th})$ following the control variate technique with any value of $\beta$ is as follows

$$
\hat{\alpha}_{IS-CV}(\gamma_{th}) = \frac{1}{M} \sum_{k=1}^{M} \left( T_{\gamma_{th}}(Y^{(k)}) + \beta \left( Z_{\gamma_{th}}(Y^{(k)}) - P(\gamma_{th}) \right) \right).
$$

(2.11)

One should note that the optimal value $\beta^*$ is generally unknown and has to be estimated via sample covariance and sample variance estimators. However, it is worth mentioning that estimating $\beta^*$ with the same simulated data as those used in getting the estimate, introduces some dependence in the above estimator. However, it was shown in [16] that working with the estimated value of $\beta^*$ yields the same estimator’s performances as by working with $\beta^*$ for large number of samples.

2.3 Performance Metrics

In practice, many criteria have been used to measure the efficiency of an unbiased estimator such as the asymptotically vanishing relative error, the bounded relative error, and the asymptotic optimality [2]. For any estimator $T'_{\gamma_{th}}(Y)$ of $\alpha(\gamma_{th})$ with $Y$ is distributed according to the PDF $g(\cdot)$, we have from the non-negativity of the variance of $T'_{\gamma_{th}}(Y)$

$$
\mathbb{E}_g [T'_{\gamma_{th}}^2(Y)] \geq \alpha^2(\gamma_{th}).
$$

(2.12)

Using the fact that $\log(\alpha(\gamma_{th})) < 0$, we get:

$$
\frac{\log \left( \mathbb{E}_g [T'_{\gamma_{th}}^2(Y)] \right)}{\log(\alpha(\gamma_{th}))} \leq 2.
$$

(2.13)
We say that the estimator $T_{\gamma_{th}}' (Y)$ achieves the asymptotic optimality property if:

$$\lim_{\gamma_{th} \to 0} \frac{\log \left( \mathbb{E}_g \left[ T_{\gamma_{th}}'^2 (Y) \right] \right)}{\log (\alpha(\gamma_{th}))} = 2. \quad (2.14)$$

An equivalent definition of the asymptotic optimality criterion is: $\forall \epsilon > 0$, we have:

$$\lim_{\gamma_{th} \to 0} \frac{\text{var}_g \left[ T_{\gamma_{th}}' (Y) \right]}{\alpha^2 - \epsilon(\gamma_{th})} = 0. \quad (2.15)$$

Two interesting interpretations can be deduced from the asymptotic optimality property. First, when $\alpha(\gamma_{th})^2 \to 0$ with an exponential rate, the second moment of $T_{\gamma_{th}}' (Y)$ converges to zero with the same exponential rate. This is the best exponential rate that the second moment may converge with. Second, when the asymptotic optimality property holds, the number of simulation runs $M$ required to meet a fixed accuracy requirement satisfies $M = o(\alpha(\gamma_{th})^{-\epsilon})$ for all $\epsilon > 0$. Such a result ensures that an estimator with the asymptotic optimality criterion will certainly yield a substantial amount of variance reduction compared to a naive MC simulation which requires a number of runs of the order of $\alpha(\gamma_{th})^{-1}$ to meet the same accuracy requirement.

A stronger criterion than the asymptotic optimality is the bounded relative error. In fact, this property holds when

$$\lim_{\gamma_{th} \to 0} \sup \frac{\text{var}_g \left[ T_{\gamma_{th}}' (Y) \right]}{\alpha^2(\gamma_{th})} < +\infty. \quad (2.16)$$

In this case, the number of simulation runs needed to meet a fixed accuracy requirement remains bounded, regardless of how small $\alpha(\gamma_{th})$ is.

An even stronger criterion is the asymptotically vanishing relative error which applies when

$$\lim_{\gamma_{th} \to 0} \sup \frac{\text{var}_g \left[ T_{\gamma_{th}}' (Y) \right]}{\alpha^2(\gamma_{th})} = 0. \quad (2.17)$$

This means that the number of samples is getting smaller as we decrease the probability of interest while ensuring a fixed accuracy requirement.
3 Mean-Shifting Approach

In this section, we review the mean-shifting IS scheme proposed in [17]. The main idea is to consider an IS distribution resulting from shifting the mean of the multivariate normal distribution. More precisely, the IS PDF is chosen to be a multivariate normal with a mean vector \( \mu + \Lambda \) and a covariance matrix \( \Sigma \) with \( \Lambda \in \mathbb{R}^N \):

\[
g(y) = \frac{\exp \left( -\frac{1}{2} (y - \mu - \Lambda)^t \Sigma^{-1} (y - \mu - \Lambda) \right)}{\sqrt{(2\pi)^N |\Sigma|}}. \tag{3.1}
\]

The likelihood ratio following this IS scheme is given as follows

\[
L(Y_1, Y_2, \cdots, Y_N) = \exp \left( -\Lambda^t \Sigma^{-1} (Y - \mu) + \frac{1}{2} \Lambda^t \Sigma^{-1} \Lambda \right). \tag{3.2}
\]

Hence, the probability of interest \( \alpha(\gamma_{th}) \) is re-written as

\[
\alpha(\gamma_{th}) = \mathbb{E}_g \left[ \exp \left( -\Lambda^t \Sigma^{-1} (Y - \mu) + \frac{1}{2} \Lambda^t \Sigma^{-1} \Lambda \right) 1(\sum_{i=1}^N \exp(Y_i) \leq \gamma_{th}) \right] = \mathbb{E}_f \left[ \exp \left( -\Lambda^t \Sigma^{-1} (Y - \mu) - \frac{1}{2} \Lambda^t \Sigma^{-1} \Lambda \right) 1(\sum_{i=1}^N \exp(Y_i + \Lambda_i) \leq \gamma_{th}) \right]. \tag{3.3}
\]

By simple computation, it was shown in [17] that the second moment of \( T_{\gamma_{th}}(Y) \) is given by

\[
\mathbb{E}_g \left[ T_{\gamma_{th}}^2(Y) \right] = \exp \left( \Lambda^t \Sigma^{-1} \Lambda \right) P_f \left( \sum_{i=1}^N \exp(Y_i - \Lambda_i) \leq \gamma_{th} \right). \tag{3.4}
\]

The remaining step is to determine the value of \( \Lambda \) that guarantees a variance reduction, compared to the naive MC estimator. Obviously, the optimal value is obtained by minimizing the second moment of the RV \( T_{\gamma_{th}}(Y) \) with respect to \( \Lambda \). However, the second moment in (3.4) is not known explicitly. The authors in [17] have then proposed to find a value of \( \Lambda \) that minimizes an asymptotic equivalent given by replacing the probability in (3.4) with an equivalent expression [17]. The value of \( \Lambda \) proposed by the authors in [17] is:

\[
\Lambda_k^* = \sum_{i,j=1}^{n} \Sigma_{ki} \bar{\Sigma}_{ij}^{-1} \left( \log(x) - \log \left( \frac{\bar{A}_1 + \cdots + \bar{A}_n}{\bar{A}_j} \right) - \bar{\mu}_j \right). \tag{3.5}
\]
By using this value of $\Lambda$, they show that the second moment of the IS estimator satisfies:

$$
\mathbb{E}_g [T_{\gamma_{th}}^2 (Y)] \leq C \alpha^2(\gamma_{th}) \left( \log \left( \frac{1}{\gamma_{th}} \right) \right) \bar{n},
$$

(3.6)

where $C$ is a constant that does not depend on $\gamma_{th}$. This means that the asymptotic optimality property (2.14) holds. This result is based on the asymptotic behavior of $\alpha(\gamma_{th})$ as $\gamma_{th} \to 0$, see Theorem 1 of [17].

### 4 Improved Algorithm Using Control Variate

The results presented in this section constitute the core of our work. Our objective is to combine the IS scheme (presented in Section 3) with a control variate technique to further achieve a variance reduction. We assume the following:

**Assumption A**: there exist an index $i \in \{1, 2, \cdots, N\}$ such that $\sqrt{\Sigma_{ii}} < \rho_{ij} \sqrt{\Sigma_{jj}}$ for all $j \neq i$, where $\rho_{ij} = \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}\Sigma_{jj}}}$ denotes the correlation coefficient between $Y_i$ and $Y_j$.

With no loss of generality, we may suppose that $i = 1$. Under the above assumption, it was previously shown in [17] that the asymptotic behavior of the left-tail of the sum is dominated by only one component that corresponds to the index $i = 1$ in assumption A, i.e.:

$$
\alpha(\gamma_{th}) \sim \frac{\sqrt{\Sigma_{11}}}{\gamma_{th} \to 0} \left( \frac{1}{\gamma_{th}} \right) \exp \left( - \frac{(\log (\gamma_{th}) - \mu_1)^2}{2\Sigma_{11}} \right).
$$

(4.1)

Moreover, under assumption A, $\bar{n} = 1$ and the solution to (2.2) is given by:

$$
\bar{w}_1 = 1, \text{ and } \bar{w}_j = 0 \text{ for all } j \neq 1.
$$

(4.2)

Under assumption A, the value of $\Lambda^*$ in (3.5) can be simplified to:

$$
\Lambda^*_1 = \log (\gamma_{th}) - \mu_1, \text{ and } \Lambda^*_k = \frac{\Sigma_{k1}}{\Sigma_{11}} (\log (\gamma_{th}) - \mu_1) \text{ for all } k \neq 1.
$$

(4.3)

Now, we present how the control variate technique may be combined with the previous IS scheme in order to achieve a further variance reduction. The control variable $Z_{\gamma_{th}} (Y)$ is selected as follows:

$$
Z_{\gamma_{th}} (Y) = 1_{(\exp(Y_i) \leq \gamma_{th})} L (Y_1, Y_2, \cdots, Y_N),
$$

(4.4)
where the random vector $\mathbf{Y}$ is distributed according to $g(\cdot)$. The expected value $P(\gamma_{th})$ of the control variable $Z_{\gamma_{th}}(\mathbf{Y})$ is given through straightforward computation:

$$
P(\gamma_{th}) = \mathbb{E}_g [Z_{\gamma_{th}}(\mathbf{Y})] = \Phi \left( \frac{\log(\gamma_{th}) - \mu_1}{\sqrt{\Sigma_{11}}} \right), \quad (4.5)$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. The main idea of the above choice of the control variable is that, with the values of $\Lambda^*$ in (4.3) and under the PDF $g(\cdot)$, $X_1$ still represents the dominant component in the sense that it determines the asymptotic behavior of the left-tail of the sum. Hence, it is likely that each realization that belongs to the set $\{X_1 \leq \gamma_{th}\}$ will also belong to the set $\{\sum_{i=1}^N X_i \leq \gamma_{th}\}$ for a sufficiently small threshold value $\gamma_{th}$.

The covariance between $T_{\gamma_{th}}(\mathbf{Y})$ and $Z_{\gamma_{th}}(\mathbf{Y})$ which is useful in the computation of the optimal value $\beta^*$ in (2.9) is given by

$$
\text{cov}_g [T_{\gamma_{th}}(\mathbf{Y}), Z_{\gamma_{th}}(\mathbf{Y})] = \mathbb{E}_g \left[ 1(\sum_{i=1}^N \exp(Y_i) \leq \gamma_{th}) \mathcal{L}(Y_1, Y_2, \ldots, Y_N) \right] - \alpha(\gamma_{th}) P(\gamma_{th})
$$

$$
= \mathbb{E}_g \left[ 1(\sum_{i=1}^N \exp(Y_i) \leq \gamma_{th}) \mathcal{L}(Y_1, Y_2, \ldots, Y_N) \right] - \alpha(\gamma_{th}) P(\gamma_{th})
$$

$$
= \text{var}_g [T_{\gamma_{th}}(\mathbf{Y})] + \alpha^2(\gamma_{th}) - \alpha(\gamma_{th}) P(\gamma_{th}). \quad (4.6)
$$

The variance of $Z_{\gamma_{th}}(\mathbf{Y})$ is given in a closed-form expression. Using a simple computation, we obtain:

$$
\mathbb{E}_g [Z_{\gamma_{th}}^2(\mathbf{Y})] = \mathbb{E}_f \left[ 1(\exp(Y_1) \leq \gamma_{th}) \mathcal{L}(Y_1, Y_2, \ldots, Y_N) \right]
$$

$$
= \exp \left( \frac{1}{2} \Lambda^* \Sigma^{-1} \Lambda^* + \Lambda^* \Sigma^{-1} \mu \right) \int_{\exp(y_1) \leq \gamma_{th}} \exp(-\Lambda^* \Sigma^{-1} y) f(y) dy \cdots dy_N
$$

$$
= \exp \left( \Lambda^* \Sigma^{-1} \Lambda^* \right) P_f (\exp(Y_1 - \Lambda_1^*) \leq \gamma_{th}). \quad (4.7)
$$

Now, using the value of $\Lambda_1^*$ in (4.3), we get:

$$
\mathbb{E}_g [Z_{\gamma_{th}}^2(\mathbf{Y})] = \exp \left( \Lambda^* \Sigma^{-1} \Lambda^* \right) \Phi \left( \frac{2 \log(\gamma_{th}) - \mu_1}{\sqrt{\Sigma_{11}}} \right). \quad (4.8)
$$

The following lemma is very useful to study the performance of the estimator $T'_{\gamma_{th}}(\mathbf{Y})$. 

Lemma 1. There exists a constant $C_1$ such that
\[
\frac{\mathbb{E}_g \left[ Z_{\gamma_{th}}^2 (Y) \right] - \mathbb{E}_g \left[ T_{\gamma_{th}}^2 (Y) \right]}{\mathbb{E}_g \left[ Z_{\gamma_{th}}^2 (Y) \right]} \leq C_1 \sqrt{\log \left( \frac{1}{\gamma_{th}} \right)} \sqrt{P_1 - P_2}. \tag{4.9}
\]
with $P_1 = \mathbb{E}_g \left[ \mathbf{1}_{(\exp(Y_1) \leq \gamma_{th})} \right] = \frac{1}{2}$ and $P_2 = \mathbb{E}_g \left[ \mathbf{1}_{\left( \sum_{i=1}^N \exp(Y_i) \leq \gamma_{th} \right)} \right]$.

Proof. Via Cauchy Schwartz inequality, it follows that
\[
\mathbb{E}_g \left[ Z_{\gamma_{th}}^2 (Y) \right] - \mathbb{E}_g \left[ T_{\gamma_{th}}^2 (Y) \right] = \mathbb{E}_g \left[ L^2 \left( \mathbf{1}_{(\exp(Y_1) \leq \gamma_{th})} - \mathbf{1}_{\left( \sum_{i=1}^N \exp(Y_i) \leq \gamma_{th} \right)} \right) \right] \leq \sqrt{\mathbb{E} \left[ L^4 \mathbf{1}_{(\exp(Y_1) \leq \gamma_{th})} \right]} \sqrt{P_1 - P_2}. \tag{4.10}
\]
Now, let us compute $\mathbb{E} \left[ L^4 \mathbf{1}_{(\exp(Y_1) \leq \gamma_{th})} \right]$. Using a similar computation as (4.7), we get:
\[
\mathbb{E}_g \left[ L^4 \mathbf{1}_{(\exp(Y_1) \leq \gamma_{th})} \right] = \mathbb{E}_f \left[ L^3 \mathbf{1}_{(\exp(Y_1) \leq \gamma_{th})} \right] = \exp \left( \frac{3}{2} \Lambda^* \Sigma^{-1} \Lambda^* + 3 \Lambda^* \Sigma^{-1} \mu \right) \int_{\{\exp(y_1) \leq \gamma_{th}\}} \exp \left( -3 \Lambda^* \Sigma^{-1} y \right) f(y) dy_1 \cdots dy_N
\]
\[
= \exp \left( \frac{3}{2} \Lambda^* \Sigma^{-1} \Lambda^* + 3 \Lambda^* \Sigma^{-1} \mu - \frac{1}{2} \mu^T \Sigma^{-1} \mu + \frac{1}{2} (\mu - 3 \Lambda^*)^T \Sigma^{-1} (\mu - 3 \Lambda^*) \right)
\times \int_{\{\exp(y_1) \leq \gamma_{th}\}} f(y + 3 \Lambda^*) dy_1 \cdots dy_N
\]
\[
= \exp \left( 6 \Lambda^* \Sigma^{-1} \Lambda^* \right) \Phi \left( \frac{4 \left( \log (\gamma_{th}) - \mu_1 \right)}{\sqrt{\Sigma_{11}}} \right). \tag{4.11}
\]
We then use the asymptotic behavior of $\Phi(\cdot)$ [1]
\[
\Phi(x) \sim \frac{1}{\sqrt{2\pi(-x)}} \exp \left( -\frac{1}{2} x^2 \right) \text{ as } x \to -\infty. \tag{4.12}
\]
By combining (4.8), (4.10), and (4.11), we get:
\[
\frac{\mathbb{E}_g \left[ Z_{\gamma_{th}}^2 (Y) \right] - \mathbb{E}_g \left[ T_{\gamma_{th}}^2 (Y) \right]}{\mathbb{E}_g \left[ Z_{\gamma_{th}}^2 (Y) \right]} \leq \frac{\exp \left( 3 \Lambda^* \Sigma^{-1} \Lambda^* \right) \sqrt{\Phi \left( \frac{4 \left( \log (\gamma_{th}) - \mu_1 \right)}{\sqrt{\Sigma_{11}}} \right)} \sqrt{P_1 - P_2}}{\exp \left( \Lambda^* \Sigma^{-1} \Lambda^* \right) \Phi \left( \frac{2 \left( \log (\gamma_{th}) - \mu_1 \right)}{\Sigma_{11}} \right)}. \tag{4.13}
\]
From the expression of $\mathbf{\Lambda}^*$ in (4.3), we observe that $\mathbf{\Lambda}^t \Sigma^{-1} \mathbf{\Lambda}^* = \frac{(\log(\gamma_{th}) - \mu_1)^2}{\Sigma_{11}}$. Subsequently, we obtain:

$$
\mathbb{E}_g \left[ Z_{\gamma_{th}}^2 (\mathbf{Y}) \right] - \mathbb{E}_g \left[ T_{\gamma_{th}}^2 (\mathbf{Y}) \right] \leq C_1 \sqrt{P_1 - P_2} \sqrt{\log \left( \frac{1}{\gamma_{th}} \right)} 
\times \exp \left( \frac{2(\log(\gamma_{th}) - \mu_1)^2}{\Sigma_{11}} \right) \exp \left( -\frac{4}{\Sigma_{11}} (\log(\gamma_{th}) - \mu_1)^2 \right) \exp \left( -\frac{2}{\Sigma_{11}} (\log(\gamma_{th}) - \mu_1)^2 \right). \tag{4.14}
$$

Thus, the proof is concluded.

In the next lemma, we study the asymptotic behavior of $P_1 - P_2$.

**Lemma 2.** For $i \in \{2, \cdots, N\}$, let us define $a_i = \frac{\Sigma_{ij}}{\Sigma_{11}} - 1$ and $c_i = \exp \left( \mu_i - \frac{\Sigma_{ji}}{\Sigma_{11}} \mu_1 \right)$ with $a_i > 0$ from Assumption A. Let $c_{i_{0}} \gamma_{th}^{a_{i_{0}}} = \max_i c_i \gamma_{th}^{a_i}$, then, for a sufficiently small $\gamma_{th}$, there a constant $C_2$ such that

$$
P_1 - P_2 \leq C_2 \gamma_{th}^{a_{i_{0}}}. \tag{4.15}
$$

**Proof.** Let us first re-write $P_1 - P_2$ as follows:

$$
P_1 - P_2 = P_g \left( \exp(Y_1) \leq \gamma_{th}, \sum_{i=1}^{N} \exp(Y_i) \geq \gamma_{th} \right). \tag{4.16}
$$

Using the value of $\mathbf{\Lambda}^*$ in (4.3), the above expression can be expressed as:

$$
P_1 - P_2 = P \left( \exp(Y_1) \leq 1, \exp(Y_1) + \sum_{i=2}^{N} c_i \gamma_{th}^{a_i} \exp(Y_i) \geq 1 \right), \tag{4.17}
$$

with $P(\cdot)$ is the probability measure under which $\mathbf{Y}$ (in this case) is a multivariate Gaussian vector with zero mean and covariance matrix $\Sigma$. If $\lambda$ is the minimum eigenvalue of $\Sigma^{-1}$, then the above expression is upper-bounded by

$$
P_1 - P_2 \leq d_1 \tilde{P} \left( \exp(Y_1) \leq 1, \exp(Y_1) + c_{i_{0}} \gamma_{th}^{a_{i_{0}}} \sum_{i=2}^{N} \exp(Y_i) \geq 1 \right), \tag{4.18}
$$

with $d_1 = \frac{1}{\sqrt{|\Sigma| \lambda^N}}$ and $\tilde{P}(\cdot)$ the probability measure under which $\mathbf{Y}$ is an independent Gaussian random vector with zero mean and covariance matrix $I_N/\lambda$.
where $I_N$ denotes the identity matrix of order $N$. One should note that, for a sufficiently small threshold, the index $i_0$ is independent of $\gamma_{th}$. On the other hand, the probability on the right-hand side can be written as:

$$
P \left( \exp(Y_1) \leq 1, \exp(Y_1) + c_{i_0} \gamma_{th}^{a_{i_0}} \sum_{i=2}^{N} \exp(Y_i) \geq 1 \right)
$$

$$
= \tilde{P} \left( \exp(Y_1) \leq 1, \exp(Y_1) + c_{i_0} \gamma_{th}^{a_{i_0}} \sum_{i=2}^{N} \exp(Y_i) \geq 1, c_{i_0} \gamma_{th}^{a_{i_0}} \sum_{i=2}^{N} \exp(Y_i) \leq 1 \right)
$$

$$
+ \tilde{P} \left( \exp(Y_1) \leq 1, \exp(Y_1) + c_{i_0} \gamma_{th}^{a_{i_0}} \sum_{i=2}^{N} \exp(Y_i) \geq 1, c_{i_0} \gamma_{th}^{a_{i_0}} \sum_{i=2}^{N} \exp(Y_i) \geq 1 \right)
$$

$$
= I_1(\gamma_{th}) + I_2(\gamma_{th}). \quad (4.19)
$$

Let $\tilde{f}(\cdot)$ be the univariate normal PDF with mean zero and variance $1/\lambda$. Then, the quantity $I_1(\gamma_{th})$ is expressed as follows:

$$
I_1(\gamma_{th})
$$

$$
= \int_{\{c_{i_0} \gamma_{th}^{a_{i_0}} \sum_{i=2}^{N} \exp(y_i) \leq 1\}} \tilde{P} \left( \exp(Y_1) \leq 1, \exp(Y_1) \geq 1 - c_{i_0} \gamma_{th}^{a_{i_0}} \sum_{i=2}^{N} \exp(y_i) \right)
$$

$$
\times \tilde{f}(y_2) \cdots \tilde{f}(y_N)
$$

$$
= \int_{\{c_{i_0} \gamma_{th}^{a_{i_0}} \sum_{i=2}^{N} \exp(y_i) \leq 1\}} \left[ \frac{1}{2} - \Phi \left( \sqrt{\lambda} \log \left( 1 - c_{i_0} \gamma_{th}^{a_{i_0}} \sum_{i=2}^{N} \exp(y_i) \right) \right) \right]
$$

$$
\times \tilde{f}(y_2) \cdots \tilde{f}(y_N)
$$

$$
= \int_{\{c_{i_0} \gamma_{th}^{a_{i_0}} \sum_{i=2}^{N} \exp(y_i) \leq \frac{1}{2}\}} \left[ \frac{1}{2} - \Phi \left( \sqrt{\lambda} \log \left( 1 - c_{i_0} \gamma_{th}^{a_{i_0}} \sum_{i=2}^{N} \exp(y_i) \right) \right) \right]
$$

$$
\times \tilde{f}(y_2) \cdots \tilde{f}(y_N)
$$

$$
+ \int_{\frac{1}{2} \leq c_{i_0} \gamma_{th}^{a_{i_0}} \sum_{i=2}^{N} \exp(y_i) \leq 1} \left[ \frac{1}{2} - \Phi \left( \sqrt{\lambda} \log \left( 1 - c_{i_0} \gamma_{th}^{a_{i_0}} \sum_{i=2}^{N} \exp(y_i) \right) \right) \right]
$$

$$
\times \tilde{f}(y_2) \cdots \tilde{f}(y_N)
$$

$$
\leq \int_{\{c_{i_0} \gamma_{th}^{a_{i_0}} \sum_{i=2}^{N} \exp(y_i) \leq \frac{1}{2}\}} \left[ \frac{1}{2} - \Phi \left( \sqrt{\lambda} \log \left( 1 - c_{i_0} \gamma_{th}^{a_{i_0}} \sum_{i=2}^{N} \exp(y_i) \right) \right) \right]
$$

$$
\times \tilde{f}(y_2) \cdots \tilde{f}(y_N)
$$

$$
+ \tilde{P} \left( \sum_{i=2}^{N} \exp(Y_i) \geq \frac{1}{2c_{i_0} \gamma_{th}^{a_{i_0}}} \right). \quad (4.20)
$$
If we let \( Z = \sum_{i=2}^{N} \exp(Y_i) \) and \( \tilde{f}_Z(\cdot) \) its corresponding PDF, we have:

\[
I_1(\gamma_{th}) \leq \int_{\{c_{io}^{-\gamma_{th}} z \leq \frac{1}{2}\}} \left[ \frac{1}{2} - \Phi \left( \sqrt{\lambda} \log \left( 1 - c_{io}^{-\gamma_{th}} z \right) \right) \right] \tilde{f}_Z(z) \, dz
\]

\[
+ \hat{P} \left( Z \geq \frac{1}{2 c_{io}^{-\gamma_{th}}} \right).
\]

Let \( g_{\gamma_{th}}(z) = \Phi \left( \sqrt{\lambda} \log \left( 1 - c_{io}^{-\gamma_{th}} z \right) \right) \). Via a simple computation, we prove that, for all \( z \leq \frac{1}{2 c_{io}^{-\gamma_{th}}} \)

\[
\left| g'_{\gamma_{th}}(z) \right| \leq \sqrt{\frac{2 \lambda}{\pi}} c_{io}^{\alpha_{io}}.
\]

Therefore, we can write:

\[
I_1(\gamma_{th}) \leq \sqrt{\frac{2 \lambda}{\pi}} c_{io}^{\alpha_{io}} \mathbb{E}[f_z[Z]] + \hat{P} \left( \sum_{i=2}^{N} \exp(Y_i) \geq \frac{1}{2 c_{io}^{-\gamma_{th}}} \right).
\]

On the other hand, we have

\[
I_2(\gamma_{th})
\]

\[
= \hat{P} \left( \exp(Y_1) \leq 1, \exp(Y_1) + c_{io}^{\alpha_{io}} \sum_{i=2}^{N} \exp(Y_i) \geq 1, c_{io}^{\alpha_{io}} \sum_{i=2}^{N} \exp(Y_i) \geq 1 \right)
\]

\[
= \hat{P} \left( \exp(Y_1) \leq 1, c_{io}^{\alpha_{io}} \sum_{i=2}^{N} \exp(Y_i) \geq 1 \right)
\]

\[
= \frac{1}{2} \hat{P} \left( \sum_{i=2}^{N} \exp(Y_i) \geq \frac{1}{c_{io}^{\alpha_{io}}} \right).
\]

Therefore, by combining (4.23) and (4.24), we get

\[
P_1 - P_2 \leq d_1 \left( \sqrt{\frac{2 \lambda}{\pi}} c_{io}^{\alpha_{io}} \mathbb{E}[f[Z]] + \frac{1}{2} \hat{P} \left( \sum_{i=2}^{N} \exp(Y_i) \geq \frac{1}{c_{io}^{\alpha_{io}}} \right) \right)
\]

\[
+ \hat{P} \left( \sum_{i=2}^{N} \exp(Y_i) \geq \frac{1}{2 c_{io}^{\alpha_{io}}} \right).
\]

On the other hand, the asymptotic behavior of the right-tail of the sum of Log-normal variates is given by \([6]\)

\[
\hat{P} \left( \sum_{i=2}^{N} \exp(Y_i) \geq \frac{1}{c_{io}^{\alpha_{io}}} \right) \sim \frac{c_1}{-\log(\gamma_{th})} \exp \left( -\frac{\lambda(\log(c_{io}^{\alpha_{io}}))^2}{2} \right) \text{ as } \gamma_{th} \to 0.
\]

On the other hand, the asymptotic behavior of the right-tail of the sum of Log-normal variates is given by \([6]\)

\[
\hat{P} \left( \sum_{i=2}^{N} \exp(Y_i) \geq \frac{1}{c_{io}^{\alpha_{io}}} \right) \sim \frac{c_1}{-\log(\gamma_{th})} \exp \left( -\frac{\lambda(\log(c_{io}^{\alpha_{io}}))^2}{2} \right) \text{ as } \gamma_{th} \to 0.
\]

Hence, the proof is concluded.
The results in Lemma 1 and Lemma 2 are instrumental in our study of the goodness of the proposed estimator $T'_{\gamma_{th}}(Y)$. The next proposition provides interesting results on the correlation coefficient as well as the optimal value $\beta^*$.

**Proposition 1.** The correlation coefficient $\rho_{T'_{\gamma_{th}}(Y),Z_{\gamma_{th}}(Y)}$ between $T'_{\gamma_{th}}(Y)$ and $Z_{\gamma_{th}}(Y)$ goes to 1 while the optimal value $\beta^*$ approaches $-1$ as $\gamma_{th} \to 0$.

**Proof.** Let us first prove that $\beta^*$ approaches $-1$ as $\gamma_{th}$ goes to zero. In fact, from the expression of $\beta^*$ in (2.9), we have

$$\beta^* = -\frac{\mathbb{E}_g[T^2_{\gamma_{th}}(Y)] - \alpha(\gamma_{th})P(\gamma_{th})}{\text{var}_g[Z_{\gamma_{th}}(Y)]}.$$  \hspace{1cm} (4.27)

When taking into account the results of Lemma 1 and Lemma 2, we have $\mathbb{E}_g[Z^2_{\gamma_{th}}(Y)] \sim \mathbb{E}_g[T^2_{\gamma_{th}}(Y)]$ as $\gamma_{th} \to 0$. Moreover, the results in (4.1) and (4.8) imply that $\alpha^2/\mathbb{E}_g[Z^2_{\gamma_{th}}(Y)] \to 0$ as $\gamma_{th}$ goes to zero. In particular, this shows that $\text{var}_g[Z_{\gamma_{th}}(Y)] \sim \mathbb{E}_g[T^2_{\gamma_{th}}(Y)]$ as $\gamma_{th} \to 0$. By combining these results and the fact that $\alpha(\gamma_{th}) \sim P(\gamma_{th})$, we conclude that $\beta^*$ approaches $-1$ as $\gamma_{th} \to 0$.

Now, let us prove that $\rho_{T_{\gamma_{th}}(Y),Z_{\gamma_{th}}(Y)}$ approaches 1 as $\gamma_{th} \to 0$. The expression of $\rho_{T_{\gamma_{th}}(Y),Z_{\gamma_{th}}(Y)}$ is given by

$$\rho_{T_{\gamma_{th}}(Y),Z_{\gamma_{th}}(Y)} = \frac{\mathbb{E}_g[T^2_{\gamma_{th}}(Y)] - \alpha(\gamma_{th})P(\gamma_{th})}{\sqrt{\text{var}_g[T_{\gamma_{th}}(Y)]}\text{var}_g[Z_{\gamma_{th}}(Y)]}. \hspace{1cm} (4.28)$$

Given that $\text{var}_g[Z_{\gamma_{th}}(Y)] \sim \text{var}_g[T_{\gamma_{th}}(Y)]$ as $\gamma_{th} \to 0$ (this follows from Lemma 1 and 2 and the fact that $\alpha^2/\mathbb{E}_g[Z^2_{\gamma_{th}}(Y)] \to 0$ and $\alpha(\gamma_{th}) \sim P(\gamma_{th})$ as $\gamma_{th}$ goes to 0), it follows that

$$\rho_{T_{\gamma_{th}}(Y),Z_{\gamma_{th}}(Y)} \sim \frac{\mathbb{E}_g[T^2_{\gamma_{th}}(Y)] - \alpha(\gamma_{th})P(\gamma_{th})}{\text{var}_g[Z_{\gamma_{th}}(Y)]}. \hspace{1cm} (4.29)$$

Hence, following the same arguments as in the first part of the proof, we conclude that $\rho_{T_{\gamma_{th}}(Y),Z_{\gamma_{th}}(Y)}$ approaches 1 as $\gamma_{th}$ goes to zero. \hfill \Box

The main conclusions that can be drawn from the above results are the following: Since $\rho_{T_{\gamma_{th}}(Y),Z_{\gamma_{th}}(Y)}$ approaches 1, the improved estimator is guaranteed to achieve a variance reduction, with respect to the mean shifting IS scheme, that increases as we decrease the threshold value. The second interesting point is that we may select $\beta$ to be equal to $-1$ instead of working with the optimal unknown value $\beta^*$. In fact, in addition to retrieving the same estimator’s performances as with $\beta = \beta^*$, we avoid, when working with $\beta = -1$, the approximation of $\beta^*$.
using simulated data as this might cause problematic statistical issues, especially when the number of samples is not large enough [16].

The next theorem exhibits the main result of our work. In fact, we prove that the estimator $T'_{\gamma th}(Y)$ has an asymptotically vanishing relative error.

**Theorem 1.** The estimator $T'_{\gamma th}(Y)$ has the asymptotically vanishing relative error property with $\beta = -1$, that is

$$\limsup_{\gamma th \to 0} \frac{\var_g [T'_{\gamma th}(Y)]}{\alpha^2(\gamma th)} = 0.$$  

**Proof.** By replacing $\beta = -1$ into (2.8), it follows that

$$\var_g [T'_{\gamma th}(Y)] = \var_g [T_{\gamma th}(Y)] - 2 \cov_g [T_{\gamma th}(Y), Z_{\gamma th}(Y)] + \var_g [Z_{\gamma th}(Y)]$$

$$= \mathbb{E}_g [Z_{\gamma th}^2(Y)] - \mathbb{E}_g [T_{\gamma th}^2(Y)] + 2 \alpha(\gamma th) P(\gamma th) - \alpha^2(\gamma th) - P^2(\gamma th).$$

Hence, by dividing the above expression by $\alpha^2(\gamma th)$ and by using Lemma 1 and Lemma 2, we get

$$\frac{\var_g [T'_{\gamma th}(Y)]}{\alpha^2(\gamma th)} = \frac{\mathbb{E}_g [Z_{\gamma th}^2(Y)] - \mathbb{E}_g [T_{\gamma th}^2(Y)]}{\alpha^2(\gamma th)}$$

$$+ \frac{2 \alpha(\gamma th) P(\gamma th) - \alpha^2(\gamma th) - P^2(\gamma th)}{\alpha^2(\gamma th)}$$

$$\leq \frac{C_3 \exp \left( \frac{3}{\Sigma_{11}} (\log(\gamma th) - \mu_1)^2 \right)}{\alpha^2(\gamma th)} \sqrt{\Phi \left( \frac{4(\log(\gamma th) - \mu_1)}{\sqrt{\Sigma_{11}}} \right)} \gamma_{th}^{\alpha_{th}/2} + o(1)$$

$$\leq C_4 \gamma_{th}^{\alpha_{th}/2} \left( \log \left( \frac{1}{\gamma th} \right) \right)^{3/2} + o(1).$$

Hence, the proof is concluded. □

## 5 Simulation Results

In this section, we present a selection of simulation results that were performed to validate our work and to investigate the amount of variance reduction achieved by
using the control variate technique compared to the IS scheme of [17]. The problem parameters are given as follows: the covariance matrix of the 4-dimensional Gaussian vector is given by

\[
\Sigma = \begin{bmatrix}
1 & 2 & 2 & 2 \\
2 & 5 & 4 & 4 \\
2 & 4 & 4.5 & 4 \\
2 & 4 & 4 & 4.5
\end{bmatrix}
\]

and the mean vector is \( \mu = (4, 4, 4, 4)^t \). It is important to note that, with the above choice of the covariance matrix, assumption A is satisfied and hence our choice of the control variable in (4.4) is justified. We now define some performance metrics that serve to compare the proposed estimator with that of [17]. We define the squared coefficient of variation of the estimator \( T_{\gamma_{th}}(Y) \) as the ratio of its variance to its squared mean

\[
CV(T_{\gamma_{th}}(Y)) = \frac{\text{var}_g[T_{\gamma_{th}}(Y)]}{\alpha^2(\gamma_{th})}.
\]

Similarly, we define the squared coefficient of variation of the estimator \( T_{\gamma_{th}}(Y) \) as

\[
CV(T_{\gamma_{th}}(Y)) = \frac{\text{var}_g[T_{\gamma_{th}}(Y)]}{\alpha^2(\gamma_{th})}.
\]

The squared coefficient of variation, also referred as squared relative error in some references, is an interesting performance metric that serves to indicate the number of samples needed to achieve a fixed accuracy requirement. In fact, from the central limit theorem, the number of samples should be proportional to the squared coefficient of variation in order to maintain the width of the confidence interval constant.

We also define the amount of variance reduction between the mean-shifting IS approach of [17] and our proposed approach as:

\[
\xi = \frac{\text{var}_g[T_{\gamma_{th}}(Y)]}{\text{var}_g[T'_{\gamma_{th}}(Y)]}.
\]

In a first experiment, we aim to validate that the correlation between \( T_{\gamma_{th}}(Y) \) and \( Z_{\gamma_{th}}(Y) \) approaches 1 as \( \gamma_{th} \) decreases to 0. To this end, we plot in Fig. 1 the correlation coefficient \( \rho_{T_{\gamma_{th}}(Y), Z_{\gamma_{th}}(Y)} \) as a function of \( \gamma_{th} \). This figure clearly shows that the correlation coefficient \( \rho_{T_{\gamma_{th}}(Y), Z_{\gamma_{th}}(Y)} \) approaches 1 as the event
of interest becomes rarer and rarer, i.e. as we decrease the threshold value $\gamma_{th}$. This result supports the performance of the proposed control variate technique in ensuring a variance reduction with respect to the mean shifting approach of [17]. Interestingly, this reduction is increasing as we decrease the value of $\gamma_{th}$.

In a second experiment, we investigate the value of $\beta^*$ as a function of the $\gamma_{th}$. We plot in Fig. 2 the approximate value of $\beta^*$ as a function of $\gamma_{th}$. From this figure, we easily observe that $\beta^*$ approaches $-1$ as $\gamma_{th}$ decreases, which validates the result in Proposition 1. Thus, this result suggests that instead of working with the optimal unknown value $\beta^*$, we should work with a fixed value of $\beta$ equal to $-1$.

In Fig. 3, we plot the estimated value of $\alpha(\gamma_{th})$ as a function of $\gamma_{th}$ using the mean-shifting IS estimator of [17] as well as our proposed estimator with $\beta = \beta^*$ and $\beta = -1$. We should note that we use the version of [16] where $\beta^*$ is estimated using the same simulated data as for the estimation. While this could result in a dependence across the replicants, it was noted in [16] that when $M$ is sufficiently large, the dependence can be ignored and we can retrieve the same performance when using the approximate value of $\beta^*$ instead of the optimal unknown value. In Fig. 3, we observe that the three curves coincide perfectly and thus the three approaches yield accurate estimates of $\alpha(\gamma_{th})$.

We now focus on studying the efficiency of the three approaches. To this end, we plot in Fig. 4 the squared coefficient of variation given by the mean-shifting approach and the two versions of the control variate techniques, i.e. when $\beta$ is equal to the approximate value of $\beta^*$ and when $\beta$ is equal to $-1$. We point out from this figure that the squared coefficient of variation corresponding to the
mean-shifting IS scheme is increasing as we decrease the threshold values. This is expected from our analysis, since the mean-shifting approach is only asymptotically optimal. More precisely, from (4.1) and (4.7), the quantity $CV \left( T_{\gamma_{th}} (Y) \right)$ is equivalent to a constant time $\log \left( \frac{1}{\gamma_{th}} \right)$. Hence, in order to meet a fixed accuracy requirement, i.e. maintain the width of the confidence interval fixed, the number of simulation runs required by the mean-shifting approach should be of the order of $\log \left( \frac{1}{\gamma_{th}} \right)$.

On the other hand, the squared coefficient of variation of the control variate
estimator, using the two values of $\beta$, is decreasing as we decrease the threshold values. This observation is expected as our proposed estimator $T'_{\gamma_{th}}(Y)$ with $\beta = -1$ has the asymptotically vanishing relative error property as it was proven in Theorem 1. Hence, the number of samples needed by the proposed control vari-
Finally, we aim to quantify the amount of variance reduction $\xi$ achieved by both versions of the proposed control variate estimator in comparison with the mean-shifting one. To do so, we plot in Fig. 5 the value of $\xi$ as a function of $\gamma_{th}$. Interestingly, we observe that, for both versions of the control variate approaches, the amount of variance reduction is increasing as decrease the probability of interest. Moreover, as we decrease $\gamma_{th}$, the value of $\xi$ with $\beta = -1$ approaches the optimal value corresponding to $\beta = \beta^*$, which is $\frac{1}{1 - \rho_{T_{\gamma_{th}}(Y), Z_{\gamma_{th}}(Y)}}$.

6 Conclusion

In this paper, we tackled the important issue of estimating the probability that a sum of correlated Log-normal variates with Gaussian copula falls below a certain threshold. We developed a variance reduction technique based on the combination of a control variate approach with a previously developed mean-shifting importance sampling approach. Under a mild assumption on the covariance matrix of the corresponding multivariate Gaussian random vector, we demonstrated that our proposed estimator has the asymptotically vanishing relative error property. Simulations were performed in order to validate our theoretical results and to quantify the performance of our proposed estimator, with and without the use of the control variate technique. Our work is an important step forward in the field of left-tail estimation of sum of Log-normal random variables. Indeed, previous studies have been limited to the weaker property of the estimator, namely the asymptotic optimality criterion.

Bibliography


Received ???.

**Author information**

Mohamed-Slim Alouini, King Abdullah University of Science and Technology (KAUST), Saudi Arabia.
E-mail: slim.alouini@kaust.edu.sa

Nadir Ben Rached, King Abdullah University of Science and Technology (KAUST), Saudi Arabia.
E-mail: nadhir.benrached@kaust.edu.sa

Abla Kammoun, King Abdullah University of Science and Technology (KAUST), Saudi Arabia.
E-mail: abla.kammoun@kaust.edu.sa

Raul Tempone, King Abdullah University of Science and Technology (KAUST), Saudi Arabia.
E-mail: raul.tempone@kaust.edu.sa