Efficient simulation of tail probabilities for sums of log-elliptical risks

Dominik Kortschak[1] and Enkelejd Hashorva[2]

October 18, 2012

Abstract: In the framework of dependent risks it is a crucial task for risk management purposes to quantify the probability that the aggregated risk exceeds some large value u. Motivated by Asmussen et al. (2011) in this paper we introduce a modified Asmussen-Kroese estimator for simulation of the rare event that the aggregated risk exceeds u. We show that in the framework of log-Gaussian risks our novel estimator has the best possible performance i.e., it has asymptotically vanishing relative error. For the more general class of log-elliptical risks with marginal distributions in the Gumbel max-domain of attraction we propose a modified Rojas-Nandayapa estimator of the rare events of interest, which for specific importance sampling densities has a good logarithmic performance. Our numerical results presented in this paper demonstrate the excellent performance of our novel Asmussen-Kroese algorithm.

Key words and phrases: Asmussen-Kroese estimator; Rojas-Nandayapa estimator; log-elliptical distribution; log-Gaussian distribution; asymptotically vanishing relative error.

1 Introduction

Efficient simulation of the tails of aggregated dependent risks has been the topic of many recent research papers, culminating in the contribution Asmussen et al. (2011). The fact that risks – here a synonym for random variables – are considered to be dependent, poses considerable difficulties in understanding the tail behavior of the aggregated risk. Nevertheless in diverse applications from finance and insurance (Goovaerts et al. (2005), Valdez et al. (2009), Asmussen et al. (2011)), risk management (Vanduffel et al. (2008), Mitra and Resncik (2009)), wireless communications (Pratesi et al. (2006), Tellambura (2008)) a few to be mentioned here, correlated log-Gaussian (log-normal) risks appear naturally.

In this paper we will allow that the parameters of the log-normal distribution depend on u. Therefore let $\boldsymbol{N} = (N_1, \ldots, N_d)^\top$ be a vector of d independent standard Gaussian random variables, and let $A_u, u > 0$ be a lower non-singular triangular matrix. For $\Sigma_u = A_u(A_u)^\top$ assume that Σ_u is a correlation matrix, i.e.,

$$\sigma_{11}(u) = \cdots = \sigma_{dd}(u) = 1, \quad \sigma_{ij}(u) \in [-1, 1], \quad i \neq j, \quad u > 0.$$

[1] Department of Actuarial Science, Faculty of Business and Economics, University of Lausanne, Bâtiment Extranef, UNIL-Dorigny, 1015 Lausanne, Switzerland, email: Dominik.Kortschak@unil.ch

[2] Department of Actuarial Science, Faculty of Business and Economics, University of Lausanne, Bâtiment Extranef, UNIL-Dorigny, 1015 Lausanne, Switzerland, email: Enkelejd.Hashorva@unil.ch

Set for $u > 0$

$$(Y_1(u), \dots, Y_d(u))^\top = A_u \boldsymbol{N} \tag{1.1}$$

and define

$$S(u) = \sum_{i=1}^{d} X_i(u), \quad \text{with} \quad X_i(u) = \lambda_i e^{\beta_i \gamma_u Y_i(u)}, \quad i \le d,$$

where $\lambda_i, \beta_i, \gamma_u, u > 0$ are given positive constants. In this paper we are interested in the numerical estimation of

$$\alpha(u) = \mathbb{P}\left(S(u) > u\right).$$

For $d = 2$ and both A_u, γ_u being constant with respect to u, the asymptotic expansion

$$\alpha(u) \sim \mathbb{P}(X_1(u) > u) + \mathbb{P}(X_2(u) > u), \quad u \to \infty \tag{1.2}$$

has been first derived in Asmussen and Rojas-Nandayapa (2008) (see for a heuristical derivation Albrecher et al. (2006)). Similar asymptotic results to (1.2) for general γ_u, A_u have been derived in Asmussen et al. (2011) and Hashorva (2011). In our notation $f(u) \sim g(u)$ means that $\lim_{u\to\infty} f(u)/g(u) = 1$ for f, g two given function.

In the light of known numerical examples (see e.g. Mitra and Resnik (2009)) the asymptotic expansion of $\alpha(u)$ given in (1.2) is too crude to be useful in practice. Hence one seeks for numerical solutions for $\alpha(u)$. A widely used numerical method for this kind of problems is Monte Carlo simulation. Since $\alpha(u) \to 0$ as $u \to \infty$ we are in the classical situation of rare event simulations. By definition, see e.g., Asmussen and Glynn (2007), an unbiased estimator $Z(u)$ of $\alpha(u)$ (i.e., a family of random variables satisfying $\mathbb{E}\left\{Z(u)\right\} = \alpha(u)$) is said to be (asymptotically as $u \to \infty$) logarithmically efficient if

$$\lim_{u\to\infty} \frac{\log \mathbb{E}\left\{Z(u)^2\right\}}{\log \alpha(u)} = 2.$$

A concept that goes beyond that is introduced in Junea (2007), namely $Z(u)$ has asymptotically vanishing relative error if further

$$\lim_{u\to\infty} \mathbb{E}\left\{ \left(\frac{Z(u)}{\alpha(u)} \right)^2 \right\} = 1. \tag{1.3}$$

Such estimator of $\alpha(u)$ reaches the best possible asymptotic performance.

It is well-known (see e.g., Cambanis et al. (1981)) that the d-dimensional standard Gaussian random vector $\boldsymbol{N}$ has the stochastic representation

$$\boldsymbol{N} \stackrel{d}{=} R\boldsymbol{U},$$

with $R > 0$ such that R^2 is chi-square distributed with d degrees of freedom being further independent of the random vector $\boldsymbol{U}$ which is uniformly distributed on the unit sphere of $\mathbb{R}^d$ (hereafter $\boldsymbol{U}$ will be reserved only for such random vectors). If we drop the distributional assumption on R, supposing only that it has some distribution function F, then $\boldsymbol{Y}(u), u > 0$ with stochastic representation

$$\boldsymbol{Y}(u) \stackrel{d}{=} \exp(A_u R \boldsymbol{U})$$

is a log-elliptical random vector; see Cambanis et al. (1981) for the basic distributional properties of elliptical random vectors. The framework of multivariate log-elliptical risks is useful in finance and insurance models (see e.g., Hamada and Valdez (2008), Valdez et al. (2009)). A key advantage when working with elliptical and log-elliptical risks is that in our model there is no distributional restriction on each individual risk; we impose only asymptotic constrains which are satisfied by a large class of possible marginal distributions. Rojas-Nandayapa (2008) provided an estimator that also works for this class of distributions.

Organisation of the paper: In the following we review some key results from the literature. Section 3 gives details of our novel Asmussen-Kroese estimator of $\alpha(u)$ which has excellent performance for log-Gaussian risks. In Section 4 we shall introduce the modified Rojas-Nandayapa estimator which can be utilised for log-elliptical risks. The numerical illustrations presented in Section 5 show the excellent performance of our modified Asmussen-Kroese estimator. The proofs of all results are relegated to Section 6, which is followed by an Appendix.

2 Details for known estimators

When $X_1(u), \ldots, X_d(u)$ are independent random variables with common distribution function F an estimator $Z_{AK}(u)$ of $\alpha(u)$ (referred to as Asmussen-Kroese estimator) is introduced in Asmussen and Kroese (2005). Namely, we have (set $\overline{F} = 1 - F$)

$$Z_{AK}(u) = d \cdot \overline{F}\left(\max\left(u + X_d(u) - S(u), \max_{1 \le i < d} X_i(u)\right)\right),$$

which is motivated by the following decomposition

$$\alpha(u) = \sum_{j=1}^{d} \Psi_j(u), \quad \text{with} \quad \Psi_j(u) = \mathbb{P}\left(S(u) > u, X_j(u) = M(u)\right) \quad \text{and} \quad M(u) = \max_{1 \le i \le d} X_j(u). \tag{2.4}$$

Accounting for the dependence of the risks, in the setup of log-Gaussian risks, Asmussen et al. (2011) introduces three different estimators of $\alpha(u)$. The first one denoted by $Z_{IS}(u)$ is an importance sampling estimator where the importance sampling distribution is log-Gaussian but the matrix Σ is multiplied by some constant γ_u, which is deduced from an asymptotic argument. Related to this estimator is $Z_{IS-CE}(u)$ where again the importance sampling distribution is log-Gaussian but this time also the mean vector can be different. The parameters are then chosen with the cross entropy method.

The third estimator of $\alpha(u)$ introduced in the aforementioned paper has a vanishing relative error. Write $\mathbb{P}(S(u) > u)$ as

$$\mathbb{P}\left(S(u) > u, \max_{i \le d} X_i(u) > u\right) + \mathbb{P}\left(S(u) > u, \max_{i \le d} X_i(u) \le u\right) := \alpha_1(u) + \alpha_2(u).$$

For the first term $\alpha_1(u)$ an importance sampling estimator that has vanishing relative error is suggested therein, whereas for the second term $\alpha_2(u)$ an importance sampling estimator equivalent to $Z_{IS}(u)$ respectively $Z_{IS-CE}(u)$ is employed. The sum of these estimators is denoted by $Z_{ISVE}(u)$ and $Z_{ISVE-CE}(u)$, respectively.

The more general case of log-elliptical risks is addressed in Rojas-Nandayapa (2008). The main idea of Rojas-Nandayapa estimator of $\alpha(u)$ is that for a log-elliptical random vector we have $S(u) = h(R, A_u, \boldsymbol{U})$ for some function h, where R and $\boldsymbol{U}$ are independent. Thus conditioning on $\boldsymbol{U}$ yields

$$\alpha(u) = \mathbb{P}\left(S(u) > u\right) = \mathbb{P}\left(h(R, A_u, \boldsymbol{U}) > u\right) = \mathbb{E}\left\{\mathbb{P}\left(h(R, A_u, \boldsymbol{U}) > u\right) | \boldsymbol{U})\right\}, \quad u > 0.$$

Denote in the following by $\boldsymbol{u}$ a simulated value (outcome) of $\boldsymbol{U}$. Since Σ_u is assumed to be positive definite, for any fixed u, the equation $h(R, A_u, \boldsymbol{u}) = u$ solved for $r > 0$ has at most two solutions denoted by $\psi_L(u, \boldsymbol{u})$ and $\psi_U(u, \boldsymbol{u})$.

For a given outcome $\boldsymbol{u}$ the function h can be $S1$) strictly decreasing, $S2$) decreasing or increasing, and $S3$) strictly increasing. Both properties $S1, S2, S3$ are examined in Rojas-Nandayapa (2008), p. 62. We define $\psi_L(u, \boldsymbol{u}), \psi_U(u, \boldsymbol{u})$ as therein, for instance if $S2$ holds, then there exist at most two different solutions satisfying

$$\lim_{u \to \infty} \psi_L(u, \boldsymbol{u}) = -\infty, \text{ and } \lim_{u \to \infty} \psi_U(u, \boldsymbol{u}) = \infty.$$

The Rojas-Nandayapa estimator of $\alpha(u)$ is defined as

$$Z_R(u) \quad = \quad \mathbb{P}\left(R < \psi_L(u, \boldsymbol{U})\right) \mathbb{I}_{\{\psi_L(u, \boldsymbol{U}) > 0\}} + \mathbb{P}\left(R > \psi_U(u, \boldsymbol{U})\right), \quad u > 0. \tag{2.5}$$

Summarising, the algorithm proposed in Rojas-Nandayapa (2008) consists of the following steps:

A. Simulate the random vector $\boldsymbol{U}$ which is uniformly distributed on the unit sphere of $\mathbb{R}^d$.

B. Calculate $\psi_L(u, \boldsymbol{U}), \psi_U(u, \boldsymbol{U})$.

C. Return $Z_R(u)$ as in (2.5).

As shown in the aforementioned paper $Z_R(u)$ is unbiased and logarithmically (asymptotic) efficient under certain restrictions on the random radius R.

3 A novel Asmussen-Kroese estimator

One reason that Asmussen-Kroese estimator has a good asymptotic behavior in the independent case is that heuristically when the sum is large then one element is large and all the others behave in a normal way. In this section we want to present a new modification of Asmussen-Kroese estimator that is better suited for log-Gaussian risks. In this paper, for the efficient estimation of the tail probability $\alpha(u) = \mathbb{P}(S(u) > u)$ for u large we use the decomposition (2.4). We shall consider the estimation, for each index $j \le d$, of the partial max-sum probability $\Psi_j(u)$ defined in (2.4). In order to compensate for the role of different components being maximal, (corresponding to different indexes j) we shall utilise a stratification idea. Specifically, when $\mathbb{P}(\mathcal{I} = i) = \frac{\mathbb{P}(X_i(u)>u)}{\sum_{j=1}^d \mathbb{P}(X_j(u)>u)}, i \le d$ and $\mathcal{I}$ is a random variable, then

$$\alpha(u) = \mathbb{P}(S(u) > u) = \left(\sum_{i=1}^d \mathbb{P}(X_i(u) > u)\right) \sum_{i=1}^d \mathbb{P}(\mathcal{I} = i) \frac{\Psi_i(u)}{\mathbb{P}(X_i(u) > u)},$$

which leads to our novel modified Asmussen-Kroese estimator of $\alpha(u)$

$$Z_{MAK}(u) = \left(\sum_{i=1}^d \mathbb{P}(X_i(u) > u)\right) \sum_{i=1}^d \mathbb{I}_{\{\mathcal{I}=i\}} \frac{Z_i(u)}{\mathbb{P}(X_i(u) > u)}, \tag{3.6}$$

where $\mathbb{I}_{\{\cdot\}}$ is the indicator random variable and $Z_i(u)$ is our modified Asmussen-Kroese estimator of $\Psi_i(u)$. In view of Lemma A.3 in Appendix, it is enough to show that Z_i is an efficient estimator for $\Psi_i(u)$.

Our novel Asmussen-Kroese estimator of the partial max-sum probability $\Psi_j(u)$ is constructed by modifying the classical Asmussen-Kroese estimator (see e.g., [5]). We will assume that A_u is chosen such that $Y_j = N_j$. Essentially, instead of conditioning on $X_i(u), i \le d, i \ne j$ like for Asmussen-Kroese estimator we condition on $N_i, i \le d, i \ne j$, which leads to the following estimator (set $M(u) = \max_{i \le d} X_i(u)$)

$$\begin{aligned} &Z_j(u) = \mathbb{P}\left(S(u) > u, X_j(u) = M(u) \middle| \boldsymbol{N}_{-j}\right) \\ &= \mathbb{P}\left(\sum_{i=1}^d \lambda_i e^{\beta_i \gamma_u a_{ij}(u) N_j + \sum_{k \ne j} \beta_i \gamma_u a_{ik}(u) N_k} > u, \lambda_j e^{\beta_j \gamma_u a_{jj}(u) N_j} = \max_{i \le d} \lambda_i e^{\beta_i \gamma_u a_{ij}(u) N_j + \sum_{k \ne j} \beta_i \gamma_u a_{ik}(u) N_k} \middle| \boldsymbol{N}_{-j}\right), \end{aligned}$$

where $a_{ij}(u)$ is the ijth entry of the matrix A_u and $\boldsymbol{N}_{-j} = (N_1, \ldots, N_{j-1}, N_{j+1}, \ldots, N_d)$. Throughout in the sequel $\gamma_u, u > 0$ are constants satisfying $\lim_{u \to \infty} \gamma_u = \gamma \in (0, \infty)$ and β_i, λ_i are positive constants. For $e(x), x \in \mathbb{R}$ some function (to be specialized later) we define

$$e_i^*(u) \quad = \quad \beta_i \gamma_u u e\left(\left(\frac{u}{\lambda_i}\right)^{\frac{1}{\beta_i \gamma_u}}\right)\left(\frac{u}{\lambda_i}\right)^{-\frac{1}{\beta_i \gamma_u}}. \tag{3.7}$$

The main result of this section is the next theorem which establishes the asymptotic properties of $Z_{MAK}(u)$.

Theorem 3.1. *Define $J = \{j : \beta_j = \max_{i \le d} \beta_i\}$ and set $e(x) = x^{-1}\log(x), x > 0$. If further for all $c > 0$ and $\epsilon > 0$ there exists $u_0 > 0$ such that for all $u > u_0$ and $i \ne j \in J$*

$$\sigma_{ij}(u) + c\sqrt{\frac{1 - \sigma_{ij}(u)^2}{\log(u)}} \quad \le \quad \frac{\beta_j}{\beta_i} \frac{\log(\epsilon e_i^*(u))}{\log(u)}, \tag{3.8}$$

then the modified Asmussen-Kroese estimator $Z_{MAK}(u)$ of $\alpha(u)$ has asymptotically vanishing relative error.

Remark 3.2. *a) Condition* (3.8) *is forced only when* $\liminf_{u\to\infty} \sigma_{ij}(u) = 1$*, since when* $\sigma_{ij}(u) \le \rho < 1$ *for all* u *large* (3.8) *is satisfied for any* $c > 0$.

b) In order to evaluate $Z_j(u)$ *we have to modify the matrix* A_u *in such a way that* $Y_j = N_j$ *which means that for every* j *we have to compute a Cholesky factorization of a matrix. Further we need to determine* x *satisfying the equation* $\sum_{i=1}^d c_i e^{d_i x} = u$*. As shown in Rojas-Nandayapa (2008) such an* x *can be quite efficiently found by Newton's method.*

c) The recent paper Kortschak (2011) derives second-order asymptotic results for dependent risks with regularly varying tails. Similar results for our framework where risks have distributions in the Gumbel MDA (and therefore have no regularly varying tails), will be derived in a forthcoming manuscript.

4 The modified Rojas-Nandayapa estimator

An key result of this section is Theorem 4.1 below, which motivates a modification of the algorithm of Rojas-Nandayapa (2008). Our novel modified Rojas-Nandayapa estimator introduced in (4.10) is logarithmically efficient, and moreover behaves asymptotically significantly better than the original one. Specifically, our algorithm is constructed under the following modifications:

i) As for Asmussen-Kroese estimator we condition on the element which is the maximum.

ii) We use importance sampling on $\boldsymbol{\Theta} := A_u \boldsymbol{U}$.

iii) We employ the same stratification method as in Eq. (3.6).

We note in passing that Rojas-Nandayapa (2008) considers only the case that A_u is constant in u.

For a given index j assume that A_u is chosen in such a way that $\Theta_j = (A_u \boldsymbol{U})_j = U_j$. We will only change the distribution of Θ_j which possesses the probability density function (pdf)

$$f(\theta) \quad = \quad \frac{\Gamma(d/2)}{\sqrt{\pi}\Gamma((d-1)/2)}(1-\theta^2)^{\frac{d-3}{2}}, \quad \theta \in (-1,1), \tag{4.9}$$

where $\Gamma(\cdot)$ is the Euler gamma function. We write f_{IS} for the corresponding pdf of Θ_j under the importance sampling measure. We then use the estimator

$$\hat{Z}_j(u) = \mathbb{P}\Big(S(u) > u, X_j(u) = M(u) \Big| \boldsymbol{\Theta}\Big) \frac{f(\Theta_j)}{f_{IS}(\Theta_j)}$$

to estimate $\Psi_j(u)$ and

$$Z_{RN}(u) = \left(\sum_{i=1}^d \mathbb{P}\left(X_i(u) > u\right) \right) \sum_{i=1}^d \mathbb{I}_{\{\mathcal{I}=i\}} \frac{\hat{Z}_i(u)}{\mathbb{P}\left(X_i > u\right)} \tag{4.10}$$

as an estimator for $\alpha(u) = \mathbb{P}(S(u) > u)$. As in Section 3 we only have to show that the estimators $\hat{Z}_j(u)$ are asymptotically efficient.

For our investigations we shall assume that the distribution function F of R with infinite upper endpoint, belongs to the Gumbel MDA with some positive scaling function ν, i.e.,

$$\lim_{u\to\infty} \frac{1 - F(u + x\nu(u))}{1 - F(u)} \quad = \quad \exp(-x), \quad \forall x \in \mathbb{R}, \tag{4.11}$$

which we abbreviated hereafter as $F \in GMDA(\nu)$ or $R \in GMDA(\nu)$. We suppose in the following that

$$\lim_{u\to\infty} \nu(u) = 0, \quad \text{and} \quad \lim_{u\to\infty} u\nu(\log(u)) = \infty. \tag{4.12}$$

Theorem 4.1. *Suppose that* (4.12) *holds. If further, for j with $\beta_j = \max_{1\le i\le d} \beta_i$ condition* (3.8) *is satisfied for any $i \neq j, i \le d$, then we have*

$$\mathbb{P}(S(u) > u) \quad \sim \quad \sum_{i=1}^{d} \mathbb{P}(X_i(u) > u). \tag{4.13}$$

Remark 4.2. *a) The sum in* (4.13) *can be reduced to the sum over the indices i such that $\beta_i = \max_{1\le j\le d} \beta_j$ and $\lambda_i = \max_{j:\beta_j=\beta_i} \lambda_j$.*

b) The scaling function $\nu(\cdot)$ is asymptotically equivalent to the mean excess function $\mathbb{E}\{(R - x)|R > x\}$.

If $\beta_i = \beta_j$ and $\lim_{u\to\infty} \log(e^*_j(u))/\log(u) = 1$, then (3.8) *is for example fulfilled when $(\sigma_{ij}(u) < 1)$*

$$\limsup_{u\to\infty} \frac{-\log\left(\frac{e^*_j(u)}{u}\right)}{(1 - \sigma_{ij}(u))\log(u)} < 1.$$

Note that $e^*_j(\cdot)$ above is defined in (3.7) *where $e(u) = u\nu(\log(u))$.*

We shall consider in the following importance sampling pdf f_{IS} given by

$$f_{IS}(a, b, x) = 2^{-(a+b-1)} \frac{\Gamma(a + b)}{\Gamma(a)\Gamma(b)} (1 + x)^{a-1}(1 - x)^{b-1}, \quad x \in [-1, 1], \tag{4.14}$$

with a, b positive constants. For our numerical results we choose the constant a to be large, say equal to 10. Next set

$$beta = \max_{i\le d} \beta_i, \quad \lambda = \max_{i:\beta_i=\beta} \lambda_i, \quad d_m = |\{i : \beta_i = \beta, \lambda_i = \lambda\}|. \tag{4.15}$$

Whenever the index i is such that $\beta_i = \beta$ and $\lambda_i = \lambda$ we define $e^*(u) := e^*_i(u)$.

Theorem 4.3. *Let the assumptions of Theorem 4.1 be fulfilled. Further assume that the function $e(u) = u\nu(\log u)$ is of bounded variation i.e., for all $c > 0$*

$$\limsup_{u\to\infty} \frac{e(cu)}{e(u)} < \infty. \tag{4.16}$$

If the importance sampling pdf f_{IS} has parameters $a > 0$ and $b = \beta(u) = \log(u)/\log(u/e^*(u))$, then

$$\frac{\mathbb{E}\left\{\hat{Z}^2_{RN}(u)\right\}}{\mathbb{P}(S(u) > u)^2} \sim \frac{e\Gamma(d/2)}{2\sqrt{\pi}\Gamma((d - 1)/2)} \log\left(\frac{u\log(u)}{e^*(u)}\right). \tag{4.17}$$

Remark 4.4. *a) In the log-Gaussian case it follows that the standard error* $\sqrt{\frac{\mathbb{E}\left\{\hat{Z}^2_{RN}(u)\right\}}{\mathbb{P}(S(u)>u)^2}}$ *is of order* $\sqrt{\log(\log(u))}$ *and hence it remains small even for relatively large values u.*

b) For the original Rojas-Nandayapa estimator $Z_R(u)$ *defined in (3.8) we obtain under the same conditions as in Theorem 4.3 that (recall* d_m *is defined in* (4.15)*)*

$$\frac{\mathbb{E}\left\{Z^2_R(u)\right\}}{\mathbb{P}\left(S(u)>u\right)^2} \gtrsim \frac{1}{d_m}\frac{2\sqrt{\pi}}{\Gamma(d/2)}\left(\frac{u\log(u)}{e^*(u)}\right)^{\frac{d-1}{2}}.$$

In the log-Gaussian case it follows that the standard error is of order $\log(u)^{\frac{d-1}{4}}$ *and hence significantly bigger than for the modified estimator.*

5 Numerical examples

In this section we present some examples on rare-event estimation. In order to compare our results, we refer to the examples of Asmussen et al. (2011). Specifically, we consider the case of a multivariate log-Gaussian distribution with $d = 10$,

$$\mu_i = i - 10, \quad \sigma^2_{ii} = i, \quad i \le d$$

and

$$\rho_{ij} \in \{0, 0.4, 0.9\}, \quad u \in \{20000, 40000, 500000\}.$$

In order to obtain reliable estimates for the variance we performed 10^7 simulations for each proposed estimator. Beside the standard error ($\sqrt{\mathbb{V}\text{ar}\{Z(u)\}}$) and the coefficient of variation $\sqrt{\mathbb{V}\text{ar}Z(u)}/\mathbb{E}\left\{Z(u)\right\}$ we also provide the needed time for the evaluation (for $5 * 10^5$ simulations since this is the number of simulations used in Asmussen et al. (2011), computations where carried out in R [23]) and the Efficiency defined by

$$\frac{Var\{\text{CMC-estimator}\} \times \text{Computation-time}\{\text{CMC-estimator}\}}{Var\{\text{Estimator}\} \times \text{Computation-time}\{\text{Estimator}\}}.$$

We compare our estimators to the Crude Monte Carlo estimator $Z_{CMC} = \mathbb{I}_{\{S(u)>u\}}$ and the importance sampling estimators defined in the aforementioned paper (compare Section 2).

$\underline{\rho = 0}$: In this case the by far best estimator is the novel modified Asmussen-Kroese estimator (MAK) that corresponds in this case to the classical Asmussen-Kroese estimator; in Table 2 for example it outperforms the other estimators by a factor of 100. Further the performance of the modified Rojas-Nandayapa estimator lies between the one of the IS respectively IS-CE and ISVE respectively ISVE-CE.

Comparing our implementation with that of Asmussen et al. (2011) we see that our implementation of the estimators IS-CE, ISVE and ISVE-CE is considerably slower. However, for other values of ρ our implementation of ISVE and ISVE-CE is more efficient. Perhaps in Asmussen et al. (2011) a different implementation for the independent case was used. For IS-CE we do not have a plausible explanation why our estimator is slower, but it only shows that the used implementation of an estimator can be important for the comparison with other estimators. If we compare the standard errors we see that there can be considerable differences. Here one should note that the standard error is only estimated and hence can only be estimated with a certain amount of uncertainty. Since we used considerably more simulations than in Asmussen et al. (2011) we will assume that our results are more accurate.

In order to get an idea for the uncertainty involved, one can consider Table 3 and the results for estimator IS-CE. We see that although the reported standard error is small the error of the estimation is relatively large, which suggest that the distribution of IS-CE is rather skewed. Therefore, one should mistrust the standard error for this particular estimator.

Since the comparison of our findings with those in Asmussen et al. (2011) for the other values of ρ is similar as for $\rho = 0$ we will concentrate next on our numerical results.

$\rho = 0.4$: In this case we see that our modified Rojas-Nandayapa (RN) estimator has standard error that is comparable to the one of ISVE-CE which is the best of the estimators in Asmussen et. al. (2011). However the RN estimator suffers from a long computation time and hence in practice the ISVE-CE estimator is still preferable. On the other hand we see that our MAK estimator is by far the best in terms of standard variation as well as in terms of efficiency. We have a speed up to a factor 8 for $u = 20000$ to a factor of 33 for $u = 500000$.

$\rho = 0.9$: We observe that all estimators decrease there performance. Our RN estimator has standard error that is comparable to the one of ISVE-CE which is the best estimator in Asmussen et. al. (2011). As explained above, RN estimator suffers from a long computation time. Similarly, our MAK estimator is by far the best in terms of standard variation as well as in terms of efficiency; we have a speed up of a factor 2 for $u = 20000$ to a factor of 5 for $u = 500000$.

Summarizing, our numerical findings show that the novel MAK estimator proposed in this paper is by far the best from the considered ones. Since the efficiency of MAK in the above examples is at least a factor 2 better than for the other estimators, which means that the evaluation time of $\alpha(u)$ for a given precision is at most half as long as for the other estimators, our estimator shows clear advantages for practical applications.

Method	Estimation	Standard error	Variation coeff.	Time	Efficiency
RN	0.00102	0.000914	0.892	67.6	34.1
MAK	0.00102	$2.81e-05$	0.0275	40.1	60700
IS	0.00102	0.0166	16.2	2.82	2.49
IS-CE	0.00344	7.41	2150	13	$2.7e-06$
ISVE	0.00102	0.000472	0.461	14	620
ISVE-CE	0.00102	0.00024	0.235	14.5	2300
CMC	0.00104	0.0322	31	1.86	1

Table 1: $\rho = 0,\ u = 20000$

Method	Estimation	Standard error	Variation coeff.	Time	Efficiency
RN	0.000463	0.000415	0.897	66	75.7
MAK	0.000463	$9.11e-06$	0.0197	39.5	264000
IS	0.000465	0.00963	20.7	2.79	3.33
IS-CE	0.000415	0.013	31.4	12.9	0.396
ISVE	0.000463	0.000197	0.425	14	1590
ISVE-CE	0.000463	0.00023	0.496	14.6	1120
CMC	0.000464	0.0215	46.4	1.86	1

Table 2: $\rho = 0,\ u = 40000$

Method	Estimation	Standard error	Variation coeff.	Time	Efficiency
RN	$1.79e-05$	$1.82e-05$	1.01	63	1560
MAK	$1.8e-05$	$7.93e-08$	0.00442	39.5	$1.31e+08$
IS	$1.83e-05$	0.000879	48.1	2.69	15.6
IS-CE	$1.67e-05$	$3.62e-05$	2.17	12.6	1960
ISVE	$1.79e-05$	$1.57e-06$	0.0872	14	946000
ISVE-CE	$1.79e-05$	$2.27e-07$	0.0127	14.7	42800000
CMC	$1.75e-05$	0.00418	239	1.85	1

Table 3: $\rho = 0,\ u = 500000$

Method	Estimation	Standard error	Variation coeff.	Time	Efficiency
RN	0.00105	0.000954	0.908	70.2	32.2
MAK	0.00105	0.000146	0.139	46.1	2090
IS	0.00105	0.0168	16.1	2.92	2.48
IS-CE	0.00104	0.0236	22.8	12.8	0.288
ISVE	0.00105	0.00271	2.58	14.1	19.9
ISVE-CE	0.00105	0.00073	0.695	14.6	265
CMC	0.00105	0.0324	30.8	1.96	1

Table 4: $\rho = 0.4$, $u = 20000$

Method	Estimation	Standard error	Variation coeff.	Time	Efficiency
RN	0.000473	0.000428	0.906	69.9	70.4
MAK	0.000473	$5.66e-05$	0.12	45.4	6220
IS	0.000468	0.00956	20.4	2.87	3.44
IS-CE	0.000479	0.0569	119	12.9	0.0217
ISVE	0.000472	0.00123	2.6	14	42.9
ISVE-CE	0.000472	0.000402	0.85	14.5	385
CMC	0.000471	0.0217	46.1	1.92	1

Table 5: $\rho = 0.4$, $u = 40000$

Method	Estimation	Standard error	Variation coeff.	Time	Efficiency
RN	$1.81e-05$	$1.83e-05$	1.01	68.2	1540
MAK	$1.81e-05$	$1.15e-06$	0.0637	42.6	623000
IS	$1.8e-05$	0.000861	47.9	2.78	17.1
IS-CE	$1.76e-05$	0.00107	61.1	13	2.35
ISVE	$1.81e-05$	$3.54e-05$	1.96	14.1	2000
ISVE-CE	$1.81e-05$	$1.14e-05$	0.629	14.8	18500
CMC	$1.82e-05$	0.00427	234	1.94	1

Table 6: $\rho = 0.4$, $u = 500000$

Method	Estimation	Standard error	Variation coeff.	Time	Efficiency
RN	0.00113	0.0012	1.06	84.6	18.1
MAK	0.00113	0.000493	0.437	58.3	155
IS	0.00113	0.0188	16.6	2.89	2.15
IS-CE	0.00113	0.0022	1.95	13.1	34.9
ISVE	0.00113	0.0096	8.49	14.1	1.69
ISVE-CE	0.00113	0.00155	1.38	15	60.7
CMC	0.00112	0.0335	29.8	1.96	1

Table 7: $\rho = 0.9,\ u = 20000$

Method	Estimation	Standard error	Variation coeff.	Time	Efficiency
RN	0.000519	0.000542	1.04	83.8	41.1
MAK	0.000519	0.000215	0.414	57.6	381
IS	0.000515	0.0108	21.1	2.85	3.02
IS-CE	0.000519	0.00105	2.02	13	71
ISVE	0.000519	0.00545	10.5	14.1	2.41
ISVE-CE	0.000519	0.000706	1.36	14.9	136
CMC	0.000519	0.0228	43.9	1.95	1

Table 8: $\rho = 0.9,\ u = 40000$

Method	Estimation	Standard error	Variation coeff.	Time	Efficiency
RN	$2.08e-05$	$2.29e-05$	1.11	81.8	1010
MAK	$2.08e-05$	$7.22e-06$	0.348	54.8	15200
IS	$1.95e-05$	0.000915	46.8	2.78	18.7
IS-CE	$2.08e-05$	$4.91e-05$	2.37	12.7	1420
ISVE	$2.1e-05$	0.000476	22.7	14	13.7
ISVE-CE	$2.08e-05$	$3.19e-05$	1.53	14.7	2900
CMC	$2.28e-05$	0.00477	209	1.91	1

Table 9: $\rho = 0.9,\ u = 500000$

6 Proofs

We prove next a lemma which is of independent interest, and then continue with the proofs of the main results.

Lemma 6.1. *Let $\boldsymbol{U}$ be uniformly distributed on the unit sphere of $\mathbb{R}^d, d \geq 2$ and $\Sigma = AA^\top$ be a correlation matrix ($\sigma_{ii} = 1$ and $-1 \leq \sigma_{ij} = \sigma_{ji} \leq 1$) with A a lower triangular non-singular matrix. Then the components of the random vector $\boldsymbol{\theta} = A\boldsymbol{U}$ satisfy for any $i \leq d$*

$$|\Theta_i - \sigma_{i1}\Theta_1| = |\Theta_i - \sigma_{i1}U_1| \leq \sqrt{1-\sigma_{i1}^2}\sqrt{1-\Theta_1^2} = \sqrt{1-\sigma_{i1}^2}\sqrt{1-U_1^2}. \tag{6.18}$$

Proof. By the assumptions $\sum_{j=1}^d a_{ij}^2 = \sigma_{ii} = 1, i \leq d$ and

$$\Theta_1 = U_1, \quad \Theta_i - \sigma_{i1}U_1 = \sum_{j=2}^d a_{ji}U_j.$$

Since $\sum_{j=2}^d a_{ij}^2 = 1 - \sigma_{i1}^2$ and $\sum_{j=2}^d U_j^2 = 1 - U_1^2$ the claim follows by Cauchy-Schwarz inequality. □

Corollary 6.2. *Under Assumption (3.8) we have that for every j with $\beta_j = \beta_1$ and every $\epsilon > 0$ there exist c, u_0 positive such that*

$$\theta_i \leq \theta_j \frac{\beta_j}{\beta_i} \frac{\log(\epsilon e_j^*(u))}{\log(u)} \tag{6.19}$$

holds for all $\theta_j > 1 - c/\log(u)$.

Proof. Condition (3.8) and (6.18) imply

$$\theta_i \leq \theta_j \left(\sigma_{ij}(u) + \sqrt{1-\sigma_{ij}(u)^2}\sqrt{\frac{1}{\theta_j^2} - 1}\right) \leq \theta_j \frac{\beta_j}{\beta_i} \frac{\log(\epsilon e_j^*(u))}{\log(u)},$$

and hence the claim follows. □

Proof of Theorem 3.1 In view of Lemma A.3 (in Appendix) we have to analyze the estimators $Z_j(u)$. For simplicity we assume that $j = 1$. Next, suppose that $\beta_1 = \max_{1 \leq i \leq d} \beta_i$. We have to show that in this case

$$\lim_{u\to\infty} \frac{\mathbb{E}\left\{Z_1(u)^2\right\}}{\mathbb{E}\left\{Z_1(u)\right\}^2} \leq 1. \tag{6.20}$$

For a constant c such that

$$c > \sqrt{4\frac{\log(d)}{(\beta_1\gamma_u)^2}}$$

we split the mean into two cases: $\max_{2 \leq i \leq d} N_i > c\sqrt{\log(u)}$ and $\max_{2 \leq i \leq d} N_i \leq c\sqrt{\log(u)}$. For the first case note that

$$\mathbb{E}\left\{Z_1(u)^2 \mathbb{I}_{\left\{\max_{2\leq i\leq d} N_i > c\sqrt{\log(u)}\right\}}\right\} \leq \sum_{i=2}^d \mathbb{E}\left\{Z_1(u)^2 \mathbb{I}_{\left\{N_i > c\sqrt{\log(u)}\right\}}\right\}$$

$$
\begin{aligned}
&\leq \sum_{i=2}^{d} \mathbb{E}\left\{\mathbb{P}\left(\lambda_1 e^{\beta_1\gamma_u N_1} > \frac{u}{d}\right)^2 \mathbb{I}_{\left\{N_i > c\sqrt{\log(u)}\right\}}\right\} \\
&= d\, \mathbb{P}\left(\lambda_1 e^{\beta_1\gamma_u N_1} > \frac{u}{d}\right)^2 \mathbb{P}\left(N_1 > c\sqrt{\log(u)}\right) \\
&\approx \exp\left(-\left(\frac{\log\left(\frac{u}{d\lambda_1}\right)}{\beta_1\gamma_u}\right)^2 - \frac{c^2\log(u)}{2}\right) \\
&\approx \exp\left(-\left(\frac{\log\left(\frac{u}{\lambda_1}\right)}{\beta_1\gamma_u}\right)^2 - \left(\frac{c^2}{2} - 2\frac{\log(d)}{(\beta_1\gamma_u)^2}\right)\log(u)\right) \\
&= o\left(\mathbb{P}\left(\lambda_1 e^{\beta_1\gamma_u N_1} > u\right)^2\right),
\end{aligned}
$$

where $\approx$ is a logarithmic asymptotic and also the last equality holds on this logarithmic scale.

For the second case we have that (with $c_1 > 0$ is a suitable constant)

$$
\begin{aligned}
&\mathbb{E}\left\{Z_1(u)^2 \mathbb{I}_{\left\{\max_{2\leq i\leq d} N_i \leq c\sqrt{\log(u)}\right\}}\right\} \\
&\leq \mathbb{E}\left\{\mathbb{P}\left(\lambda_1 e^{\beta_1\gamma_u N_1} + \sum_{i=2}^{d}\lambda_i e^{\beta_i\gamma_u a_{i1}N_1 + \sum_{j=2}^{i}\beta_i\gamma_u a_{ij}N_j} > u, \lambda_1 e^{\beta_1\gamma_u N_1} > \frac{u}{d}\right)^2 \mathbb{I}_{\left\{\max_{2\leq i\leq d} N_i \leq c\sqrt{\log(u)}\right\}}\right\} \\
&\leq \mathbb{E}\left\{\mathbb{P}\left(\lambda_1 e^{\beta_1\gamma_u N_1} + \sum_{i=2}^{d}\lambda_i e^{\beta_i\gamma_u a_{1i}N_1 + \sum_{j=2}^{i}\beta_i\gamma_u\sqrt{1-a_{1i}^2}c\sqrt{\log(u)}} > u, \lambda_1 e^{\beta_1\gamma_u N_1} > \frac{u}{d}\right)^2\right\} \\
&\leq \mathbb{E}\left\{\mathbb{P}\left(\lambda_1 e^{\beta_1\gamma_u N_1} + \sum_{i=2}^{d}\lambda_i e^{\beta_i\gamma_u a_{1i}N_1 + d\beta_i\gamma_u\sqrt{1-a_{1i}^2}c\sqrt{\log(u)}} > u, \lambda_1 e^{\beta_1\gamma_u N_1} > \frac{u}{d}\right)^2\right\} \\
&= \mathbb{P}\left(\lambda_1 e^{\beta_1\gamma_u N_1} + \sum_{i=2}^{d}\lambda_i e^{\beta_i\gamma_u N_1\left(a_{1i} + dc\sqrt{1-a_{1i}^2}\sqrt{\frac{\log(u)}{N_1^2}}\right)} > u, \lambda_1 e^{\beta_1\gamma_u N_1} > \frac{u}{d}\right)^2 \\
&\leq \mathbb{P}\left(\lambda_1 e^{\beta_1\gamma_u N_1} + \sum_{i=2}^{d}\lambda_i e^{\beta_i\gamma_u N_1\left(a_{1i} + dc\sqrt{1-a_{1i}^2}\sqrt{\frac{(\beta_1\gamma_u)^2\log(u)}{\log(u/(d\lambda_1))^2}}\right)} > u\right)^2 \\
&\lesssim \mathbb{P}\left(\lambda_1 e^{\beta_1\gamma_u N_1} + \sum_{i=2}^{d}\lambda_i e^{\beta_i\gamma_u N_1\left(a_{1i} + c_1\sqrt{\frac{1-a_{1i}^2}{\log(u)}}\right)} > u\right)^2 \\
&\leq \mathbb{P}\left(\lambda_1 e^{\beta_1\gamma_u N_1} + \sum_{i=2}^{d}\lambda_i e^{\beta_1\gamma_u N_1 \frac{\log(\epsilon e_1^*(u))}{\log(u)}} > u\right)^2,
\end{aligned}
$$

where we write a_{ij} instead of $a_{ij}(u)$. Next, we can find another constant c_2 such that for every ϵ there exists a $u_\epsilon > 1$

$$
c_2 > \sup_{u>u_\epsilon}\sum_{i=2}^{d}\lambda_i e^{\left(\log\left(1-c_2\epsilon\frac{e_1^*(u)}{u}\right)-\log(\lambda_1)\right)\frac{\log(\epsilon e_1^*(u))}{\log(u)}}.
$$

If we set

$$
L_1 = \frac{\log\left(\frac{u-c_2\epsilon e_1^*(u)}{\lambda_1}\right)}{\beta_1\gamma_u}
$$

then for $u > u_\epsilon$

$$\lambda_1 e^{\beta_1\gamma_u L_1} + \sum_{i=2}^{d} \lambda_i e^{\beta_1\gamma_u L_1 \frac{\log(\epsilon e_1^*(u))}{\log(u)}}$$
$$= u - c_2\epsilon e_1^*(u) + \epsilon e_1^*(u) \sum_{i=2}^{d} \lambda_i e^{\left(\log\left(1-c_2\epsilon\frac{e_1^*(u)}{u}\right)-\log(\lambda_1)\right)\frac{\log(\epsilon e_1^*(u))}{\log(u)}} \le u.$$

Hence (6.20) follows from

$$\mathbb{E}\left\{[Z_1(u)]^2 \mathbb{I}_{\left\{\max_{2\le i\le d} N_i \le c\sqrt{\log(u)}\right\}}\right\} \lesssim \mathbb{P}\left(\lambda_1 e^{\beta_1\gamma_u N_1} > u - c_2\epsilon e_j^*(u)\right)^2$$

and letting $\epsilon \to 0$. On the other hand if $\beta_1 < \max_{i\le d}\beta_i$, then for all u large

$$Z_1(u) \le \mathbb{P}\left(\lambda_1 e^{\beta_1\gamma_u N_1} > \frac{u}{d}\right)$$

implying

$$\limsup_{u\to\infty} \frac{\left(\mathbb{E}\left\{Z_1(u)^2\right\}\right)}{\Psi_1(u)\mathbb{P}\left(S(u) > u\right)} = 0,$$

and hence the claim follows from Lemma A.3. □

Proof of Theorem 4.1 Define $\mathbf{\Theta}_u := A_u\boldsymbol{U}$ and write Θ_i for its ith component. Note that by the assumption Θ_i has distribution function not depending on u. Condition (4.12) implies $\exp(R) \in GMDA(e)$ with $e(u) = u\nu(\log u)$. By the Davis-Resnick tail property (see e.g., Hashorva (2012) or Hashorva (2013)) for any $c > 1$ and $\mu > 0$

$$\lim_{u\to\infty} \frac{\mathbb{P}\left(R > \log(cu)\right)}{\left(\frac{e(u)}{u}\right)^{\mu}\mathbb{P}\left(R > \log(u)\right)} = 0. \tag{6.21}$$

In order to show the proof we use the next equality, that holds for all $u > d\lambda_j$

$$\begin{aligned}\mathbb{P}\left(S(u) > u\right) &= \sum_{j=1}^{d} \mathbb{P}\left(S(u) > u, X_j(u) = \max_{k\le d} X_k(u)\right)\\ &= \sum_{j=1}^{d}\int_0^1 \mathbb{P}\left(\sum_{i=1}^{d}\lambda_i e^{R\Theta_i\beta_i\gamma_u} > u, \lambda_j e^{R\Theta_j\beta_j\gamma_u} = \max_{k\le d}\lambda_k e^{R\Theta_k\beta_k\gamma_u}\middle|\Theta_j = \theta\right) f(\theta)d\theta,\end{aligned}$$

where without loss of generality we will assume that depending on j an A_u is chosen such that $\Theta_j = U_j$ and the pdf f is given by (4.9). For a fixed j, we split the integral above into two parts determined through $a(u) = 1 - 2\log(d)/\log(u), u > d\lambda_j$. Then we have that

$$\int_0^{a(u)} \mathbb{P}\left(\sum_{i=1}^{d}\lambda_i e^{R\Theta_i\beta_i\gamma_u} > u, \lambda_j e^{R\Theta_j\beta_j\gamma_u} > \max_{k\ne j}\lambda_k e^{R\Theta_k\beta_k\gamma_u}\middle|\Theta_j = \theta\right) f(\theta)d\theta$$
$$\le \int_0^{a(u)} \mathbb{P}\left(d\lambda_j e^{R\theta\beta_j\gamma_u} > u\right) f(\theta)d\theta$$
$$= o\left(\mathbb{P}\left(X_j(u) > u\right)\right).$$

The last equality follows as a combination of (6.21) and Lemmas A.1, A.2 in Appendix.

Further for any $\epsilon \in (0,1)$ and $u > u_0$ (u_0 from condition 3.8) we obtain by Corollary 6.2

$$
\begin{aligned}
&\int_{a(u)}^{1} \mathbb{P}\left(e^{R\Theta_i\beta_i\gamma_u} > u \middle| \Theta_j = \theta\right) f(\theta) d\theta \\
&\lesssim \int_{a(u)}^{1} \mathbb{P}\left(\sum_{i=1}^{d} \lambda_i e^{R\Theta_i\beta_i\gamma_u} > u, \lambda_j e^{R\theta\beta_j\gamma_u} > \max_{k\neq j} \lambda_k e^{R\Theta_k\beta_k\gamma_u} | \Theta_j = \theta\right) f(\theta) d\theta \\
&\leq \int_{a(u)}^{1} \mathbb{P}\left(\lambda_j e^{R\theta\beta_j\gamma_u} + \sum_{i\neq j} \lambda_i e^{R\theta\beta_j\gamma_u \frac{\log(\epsilon e_j^*(u))}{\log(u)}} > u\right) f(\theta) d\theta.
\end{aligned}
$$

Next, we choose c such that for all $\tau < 1$ there exists a u_τ such that

$$
c > \sup_{u > u_\tau} \sum_{i\neq j} \lambda_i e^{\left(\log\left(1 - \frac{c\tau e_j^*(u)}{u}\right) - \log(\lambda_j)\right) \frac{\log(\tau e_j^*(u))}{\log(u)}}.
$$

Both constants u_τ and c exist since we assume that $\lim_{u\to\infty} e(u) = \infty$. Hence for $u > u_\tau$

$$
\int_{a(u)}^{1} \mathbb{P}\left(\lambda_j e^{R\theta\beta_j\gamma_u} + \sum_{i\neq j} \lambda_i e^{R\theta\beta_j\gamma_u \frac{\log(\epsilon e_j^*(u))}{\log(u)}} > u\right) f(\theta) d\theta \quad \leq \quad \int_{a(u)}^{1} \mathbb{P}\left(\lambda_j e^{R\theta\beta_j\gamma_u} > u - c\epsilon e_j^*(u)\right) f(\theta) d\theta.
$$

Assuming that $\beta_j = \max_{i\leq d}\beta_i$, Lemma A.1 implies thus

$$
\Psi_j(u) = \mathbb{P}\left(S(u) > u, X_j(u) > \max_{k\neq j} X_k(u)\right) \sim \mathbb{P}\left(X_j(u) > u\right).
$$

If $\beta_j < \max_{i\leq d}\beta_i$ and k is such that $\beta_k = \max_{i\leq d}\beta_i$, then for every $c_1 > 1$

$$
\begin{aligned}
\Psi_j(u) &\leq \mathbb{P}\left(X_k(u) > u/d\right) \\
&= \mathbb{P}\left(X_k(u) > \left(\frac{u}{d\lambda_k}\right)^{\frac{\beta_k}{\beta_j}}\right) \lesssim \mathbb{P}\left(X_k(u) > c_1 u\right),
\end{aligned}
$$

hence the claim follows. □

Proof of Theorem 4.3 Again we have to analyze the estimator $\hat{Z}_j(u)$. Denote by $f_{-j}(\boldsymbol{\theta}_{-j}|\theta)$ the conditional density of $\boldsymbol{\Theta}_{-j} := (\Theta_1, \ldots, \Theta_{j-1}, \Theta_{j+1}, \ldots, \Theta_d)$ given $\Theta_j = \theta$. The second moment of the estimator is given by

$$
\int_{-1}^{1} \int \mathbb{P}\left(S(u) > u, X_j(u) = \max_{k\leq d} X_k(u) \middle| \boldsymbol{\Theta} = \boldsymbol{\theta}\right)^2 f_{-j}(\boldsymbol{\theta}_{-j}|\theta)\, d\boldsymbol{\theta}_{-j} \frac{f(\theta)}{f_{IS}(\theta)} f(\theta)\, d\theta. \tag{6.22}
$$

We assume next that $\beta_j = \max_{1\leq i\leq d}\beta_j$. As in the proof of Theorem 4.1 we split the integral into parts, where Θ_j is between $a(u)$ and 1 respectively -1 and $a(u)$. By the same method as in the proof of Theorem 4.1 for some $c > 0$ and all $\epsilon \in (0,1)$ we obtain

$$
\begin{aligned}
&\int_{a(u)}^{1} \mathbb{P}\left(\lambda_j e^{R\Theta_j\beta_j\gamma_u} > u \middle| \Theta_j = \theta\right)^2 \frac{f(\theta)}{f_{IS}(\theta)} f(\theta)\, d\theta \\
&\lesssim \int_{a(u)}^{1} \int \mathbb{P}\left(S(u) > u, X_j(u) > \max_{k\neq j} X_k(u) \middle| \boldsymbol{\Theta} = \boldsymbol{\theta}\right)^2 f_{-j}(\boldsymbol{\theta}_{-j}|\theta) d\boldsymbol{\theta}_{-j} \frac{f(\theta)}{f_{IS}(\theta)} f(\theta)\, d\theta
\end{aligned}
$$

$$\lesssim \int_{a(u)}^{1} \mathbb{P}\left(\lambda_j e^{R\Theta_j\beta_j\gamma_u} > u - c\epsilon e_j^*(u) \middle| \Theta_j = \theta\right)^2 \frac{f(\theta)}{f_{IS}(\theta)} f(\theta)\, d\theta.$$

As in the proof of Lemma A.1 we substitute $\theta = \frac{\log(u)}{\log(u)+\log(1+xe_j^*(u)/u)}$. Set next $\beta(u) := \log(u)/\log(u/e^*(u))$ and note that uniformly for $u \to \infty$

$$\begin{aligned}
&\frac{f\left(\frac{\log(u)}{\log(u)+\log(1+xe_j^*(u)/u)}\right)}{B(a,\beta(u)) f_{IS}\left(\frac{\log(u)}{\log(u)+\log(1+xe_j^*(u)/u)}\right)} \\
&\sim 2^{\frac{d-3}{2}+\beta(u)} \frac{\Gamma(d/2)}{\sqrt{\pi}\Gamma((d-1)/2)} \left(\frac{\log(1+xe_j^*(u)/u)}{\log(u)+\log(1+xe_j^*(u)/u)}\right)^{\frac{d-3}{2}-\beta(u)+1} \\
&\leq e 2^{\frac{d-3}{2}+\beta(u)} \frac{\Gamma(d/2)}{\sqrt{\pi}\Gamma((d-1)/2)} \left(\frac{xe_j^*(u)}{u\log(u)}\right)^{\frac{d-1}{2}} x^{-\beta(u)}
\end{aligned}$$

and

$$\frac{f\left(\frac{\log(u)}{\log(u)+\log(1+xe_j^*(u)/u)}\right)}{f_{IS}\left(\frac{\log(u)}{\log(u)+\log(1+xe_j^*(u)/u)}\right)} \sim e 2^{\frac{d-3}{2}} \frac{B(a,\beta(u))\Gamma(d/2)}{\sqrt{\pi}\Gamma((d-1)/2)} \left(\frac{e_j^*(u)}{u\log(u)}\right)^{\frac{d-1}{2}} x^{\frac{d-1}{2}},$$

where $B(a,\beta(u)) = \Gamma(a)\Gamma(\beta(u))/\Gamma(a+\beta(u))$. Consequently, as in the proof of Lemma A.1

$$\begin{aligned}
&\int_{a(u)}^{1} \mathbb{P}\left(\lambda_j e^{R\Theta_j\beta_j\gamma_u} > u - c\epsilon e_j^*(u) \middle| \Theta_j = \theta\right)^2 \frac{f(\theta)}{f_{IS}(\theta)} f(\theta) d\theta \\
&\sim (1+O(\epsilon)) \left(\frac{2^{\frac{d-3}{2}}\Gamma(d/2)}{\sqrt{\pi}\Gamma((d-1)/2)} \left(\frac{e_j^*(u)}{u\log(u)}\right)^{\frac{d-1}{2}} \mathbb{P}\left(\lambda_j e^{R\beta_j\gamma_u} > u\right)\right)^2 \\
&\quad \times eB(a,\beta(u)) \int_0^\infty e^{-2x} x^{\frac{d-3}{2}+\frac{d-1}{2}} dx \\
&\sim -\frac{e\Gamma(d/2)}{2\sqrt{\pi}\Gamma((d-1)/2)} \log\left(\frac{e_j^*(u)}{u\log(u)}\right) \mathbb{P}\left(X_j(u) > u\right)^2.
\end{aligned} \tag{6.23}$$

Since $B(a,\beta) \sim 1/\beta$ as $\beta \to 0$, analogously to the proof of Lemma A.2 we have

$$\begin{aligned}
&\int_{-1}^{a(u)} \int \mathbb{P}\left(S(u) > u, X_j(u) > \max_{k\neq j} X_k(u) \middle| \boldsymbol{\Theta} = \boldsymbol{\theta}\right)^2 f_{-j}(\boldsymbol{\theta}_{-j}|\theta) d\boldsymbol{\theta}_{-j} \frac{f(\theta)}{f_{IS}(\theta)} f(\theta) d\theta \\
&\leq \int_{-1}^{a(u)} \mathbb{P}\left(X_j(u) > u/d \middle| \Theta_j = \theta\right)^2 \frac{f(\theta)}{f_{IS}(\theta)} f(\theta) d\theta \\
&\sim \int_{0}^{a(u)} \mathbb{P}\left(X_j(u) > u/d \middle| \Theta_j = \theta\right)^2 \frac{f(\theta)}{f_{IS}(\theta)} f(\theta) d\theta \\
&= \log\left(\frac{u\log(u)}{e_j^*(u)}\right) o\left(\mathbb{P}(X_j(u) > u)^2\right).
\end{aligned}$$

Next assume that $\beta_j < \max_{1\leq i\leq d} \beta_i$. The second moment of the estimator is (asymptotically) given by (6.22). As in the proof of Theorem 4.1 for every $c > 1$ we obtain

$$\begin{aligned}
&\int_{-1}^{1} \int \mathbb{P}\left(S(u) > u, X_j(u) > \max_{k\neq j} X_k(u) \middle| \boldsymbol{\Theta} = \boldsymbol{\theta}\right)^2 f_{-j}(\boldsymbol{\theta}_{-j}|\theta) d\boldsymbol{\theta}_{-j} \frac{f(\theta)}{f_{IS}(\theta)} f(\theta) d\theta \\
&\lesssim \int_{0}^{1} \mathbb{P}\left(dX_j(u) > u \middle| \Theta_j = \theta\right)^2 \frac{f(\theta)}{f_{IS}(\theta)} f(\theta) d\theta.
\end{aligned}$$

We can proceed as in the proof of Theorem 4.3 to get that

$$\mathbb{E}\left\{[\hat{Z}_j(u)]^2\right\} = o\left(\log\left(\frac{u\log(u)}{e_j^*(u)}\right)\left(\frac{e\left(\left(\frac{u}{d\lambda_j}\right)^{\frac{1}{\beta_j\gamma_u}}\right)}{e\left(\left(\frac{u}{\lambda_j}\right)^{\frac{1}{\beta_j\gamma_u}}\right)}\right)^{\frac{d-1}{2}}\right)\mathbb{P}\left(X_j(u) > u/d\right)^2$$

and hence the claim follows by condition (4.16) and Lemma A.3. □

A Appendix

In the sequel we consider some positive random variable R such that its distribution function F has an infinite upper endpoint. We have the following representation for $F \in GMDA(\nu)$, see e.g., Resnick (1987)

$$1 - F(u) \quad = \quad c(u)\exp\Big(-\int_{x_0}^{u}\frac{g(t)}{\nu(t)}\,dt\Big), \tag{A.24}$$

with x_0 some constant and c, g two positive measurable functions such that $\lim_{u\to\infty} c(u) = \lim_{u\to\infty} g(u) = 1$. Further we assume that $e(u) = u\nu(\log(u))$ is a scaling function of $\exp(R)$, i.e., $\exp(R) \in GMDA(e)$. This holds in particular when $\lim_{u\to\infty}\nu(u) = 0$. We define $e^*(u)$ by (3.7) for some λ, β positive, i.e.,

$$e^*(u) = \beta\gamma_u u e\left(\left(\frac{u}{\lambda}\right)^{\frac{1}{\beta\gamma_u}}\right)\left(\frac{u}{\lambda}\right)^{-\frac{1}{\beta\gamma_u}},$$

with γ_u such that $\lim_{u\to\infty}\gamma_u = \gamma \in (0,\infty)$. We proceed with two lemmas and then conclude this section with two results, the first shows an unbiased estimator for sums of certain probabilities, whereas the second provides an upper bound on the linear combination of the components of uniformly distributed random vectors on the unit sphere of $\mathbb{R}^d$.

Lemma A.1. *Let R be a positive random variable, and let f be the pdf given by* (4.9)*. If* $\exp(R) \in GMDA(e)$*, then for any $\beta, \lambda, m, \varepsilon$ positive and some $k > 0$*

$$\int_{a(u)}^{1}\mathbb{P}\left(\lambda e^{R\theta\beta\gamma_u} > u - \epsilon e^*(u)\right)^m f(\theta)d\theta = m^{\frac{d-1}{2}}(1+O(\epsilon))\frac{2^{\frac{d-3}{2}}\Gamma(d/2)}{\sqrt{\pi}}\left(\frac{e^*(u)}{u\log(u)}\right)^{\frac{d-1}{2}}\mathbb{P}\left(\lambda e^{R\beta\gamma_u} > u\right)^m, \tag{A.25}$$

with $a(u) \le 1 - k/\log(u)$ such that $\lim_{u\to\infty} a(u) = 1$, and γ_u some positive constants such that $\lim_{u\to\infty}\gamma_u = \gamma \in (0,\infty)$.

Proof. The assumption that $\exp(R) \in GMDA(e)$ implies

$$\xi(u) := e^*(u)/u \to 0, \quad u \to \infty. \tag{A.26}$$

Next, set $b(x,u) = \log(1 + x\xi(u))$ and $c = \frac{2^{\frac{d-3}{2}}\Gamma(d/2)}{\sqrt{\pi}\Gamma((d-1)/2)}$. We have

$$\int_{a(u)}^{1}\mathbb{P}\left(\lambda e^{R\theta\beta\gamma_u} > u - \epsilon e^*(u)\right)^m f(\theta)d\theta$$

$$
\begin{aligned}
&\sim c\int_{a(u)}^{1} \mathbb{P}\left(\lambda e^{R\theta\beta\gamma_u} > u-\epsilon e^*(u)\right)^m (1-\theta)^{\frac{d-3}{2}} d\theta \\
&= c\int_0^{\frac{(u-\epsilon e^*(u))^{1/a(u)-1}-1}{\xi(u)}} \frac{\xi(u)\log(u-\epsilon e^*(u))}{(x\xi(u)+1)\left(\log(u-\epsilon e^*(u))+b(x,u)\right)^2} \\
&\qquad \times \mathbb{P}\left(\lambda e^{R\beta\gamma_u} > (u-\epsilon e^*(u))^{1+\frac{\log(1+x\xi(u))}{\log(u-\epsilon e^*(u))}}\right)^m \left(\frac{b(u,x)}{\log(u-\epsilon e^*(u))+b(u,x)}\right)^{\frac{d-3}{2}} dx \\
&\sim c\left(\frac{\xi(u)}{\log(u)}\right)^{\frac{d-1}{2}} \int_0^{\frac{(u-\epsilon e^*(u))^{1/a(u)-1}-1}{\xi(u)}} \frac{1}{1+x\xi(u)}\left(1+\frac{b(u,x)}{\log(u-\epsilon e^*(u))}\right)^{-\frac{d+1}{2}} \\
&\qquad \times \mathbb{P}\left(\lambda e^{R\beta\gamma_u} > (u-\epsilon e^*(u))\,(1+x\xi(u))\right)^m \left(\frac{b(u,x)}{\xi(u)}\right)^{\frac{d-3}{2}} dx.
\end{aligned}
$$

It follows that $R \in GMDA(\nu)$, where $\nu(\log(u)) = e(u)/u$, hence Eq. (6.31) of Hashorva (2009) implies for any $\varepsilon > 0$ and some η_1, η_2 positive constants

$$
\frac{\mathbb{P}(R > u + x\nu(u))}{\mathbb{P}(R > u)} \le \eta_1 (1+\eta_2 x)^{-1/\varepsilon}. \tag{A.27}
$$

Consequently, by the dominated convergence theorem

$$
\begin{aligned}
&\int_0^{\frac{u^{1/a(u)-1}-1}{\xi(u)}} \frac{\xi(u)\log(u)}{(x\xi(u)+1)\left(\log(u)+b(u,x)\right)^2} \\
&\qquad \times \mathbb{P}\left(\lambda e^{R\beta\gamma_u} > u + x(1+O(\epsilon))e^*(u)\right)^m \left(\frac{b(u,x)}{\log(u)+b(u,x)}\right)^{\frac{d-3}{2}} dx \\
&\sim \left(\frac{\xi(u)}{\log(u)}\right)^{\frac{d-1}{2}} \mathbb{P}\left(\lambda e^{R\beta\gamma_u} > u\right)^m \int_0^\infty e^{-mx(1+O(\epsilon))} x^{\frac{d-3}{2}} dx \\
&\sim m^{\frac{d-1}{2}} (1+O(\epsilon)) \frac{2^{\frac{d-3}{2}}\Gamma(d/2)}{\sqrt{\pi}} \left(\frac{\xi(u)}{\log(u)}\right)^{\frac{d-1}{2}} \mathbb{P}\left(\lambda e^{R\beta\gamma_u} > u\right)^m,
\end{aligned}
$$

and thus the proof is complete. ☐

Lemma A.2. *Under the assumptions of Lemma A.1, for any $\beta, \lambda, \varepsilon$ positive and some $k > d$*

$$
\int_0^{1-\log(k)/\log(u)} \mathbb{P}\left(\lambda e^{R\theta\beta\gamma_u} > u/d\right) f(\theta)d\theta \ll \frac{2\Gamma(d/2)}{\sqrt{\pi}} \left(\frac{e^*(u)}{u\log(u)}\right)^{\frac{d-1}{2}} \mathbb{P}\left(\lambda e^{R\beta\gamma_u} > u\right), \tag{A.28}
$$

where for two functions $h_1(u) \ll h_2(u)$ means $h_1(u) = o(h_2(u))$.

Proof. Choose $b(u) \le a(u) := 1 - log(k)/\log(u)$ with $\lim_{u\to\infty} b(u) = 1$, such that

$$
\mathbb{P}\left(\lambda e^{Rb(u)\beta\gamma_u} > u/d\right) = o\left(\int_0^1 \mathbb{P}\left(\lambda e^{R\theta\beta\gamma_u} > u\right) f(\theta)d\theta\right).
$$

Set $\xi(u) = e^*(u)/u, b(x,u) = \log(1+x\xi(u))$. By substituting

$$
\theta = \frac{\log(u/k)}{\log(u+xe^*(u))}
$$

and for some $c > 0$, we have that

$$\begin{aligned}\int_{b(u)}^{a(u)} \mathbb{P}\left(\lambda e^{R\theta\beta\gamma_u} > u/d\right) f(\theta) d\theta \quad &\sim \quad c \int_0^{\frac{(u/k)^{1/b(u)}-u}{e^*(u)}} \frac{e^*(u)\log(u/k)}{(u+xe^*(u))\log(u+xe^*(u))^2} \\ &\times \quad \mathbb{P}\left(\lambda e^{R\beta\gamma_u} > (u/d)^{\frac{\log(u+xe^*(u))}{\log(u/k)}}\right)\left(\frac{b(u,x)+\log(k)}{\log(u)+b(u,x)}\right)^{\frac{d-3}{2}} dx \\ &\lesssim \quad c\log(u)^{-\frac{d-1}{2}} \int_0^{\frac{(u/k)^{1/b(u)}-u}{e^*(u)}} \frac{e^*(u)}{(u+xe^*(u))} \\ &\times \quad \mathbb{P}\left(\lambda e^{R\beta\gamma_u} > \frac{k}{d}\left(u+xe^*(u)\right)\right)\left(\frac{b(u,x)+\log(k)}{1+b(u,x)/\log(u)}\right)^{\frac{d-3}{2}} dx.\end{aligned}$$

Next, (6.21) implies

$$\mathbb{P}\left(\lambda e^{R\beta\gamma_u} > \frac{k}{d}\left(u+xe^*(u)\right)\right) \lesssim \left(\frac{e\left(\left(\frac{1}{\lambda}(u+xe^*(u))\right)^{\frac{1}{\beta\gamma_u}}\right)}{\left(\frac{1}{\lambda}(u+xe^*(u))\right)^{\frac{1}{\beta\gamma_u}}}\right)^{\frac{d-1}{2}} \mathbb{P}\left(\lambda e^{R\beta\gamma_u} > u+xe^*(u)\right)$$

From the representation theorem for self-neglecting functions (cf. Bingham et al. (1987)) it follows that for every $\delta > 0$ and u large enough and $x > 0$ we have

$$e(u+xe(u))/e(u) \le (1+\delta)(1+\delta x).$$

Together with (A.27) it follows that for every $\epsilon > 0$ there exist η_1 and η_2 such that

$$\left(\frac{e\left(\left(\frac{1}{\lambda}(u+xe^*(u))\right)^{\frac{1}{\beta\gamma_u}}\right)}{\left(\frac{1}{\lambda}(u+xe^*(u))\right)^{\frac{1}{\beta\gamma_u}}}\right)^{\frac{d-1}{2}} \mathbb{P}\left(\lambda e^{R\beta\gamma_u} > u+xe^*(u)\right) \lesssim \eta_1(1+\eta_2 x)^{-1/\epsilon}\left(\frac{e^*(u)}{u}\right)^{\frac{d-1}{2}} \mathbb{P}\left(\lambda e^{R\beta\gamma_u} > u\right)$$

holds uniformly for $x > 0$, and hence the proof follows with similar arguments as that of Lemma A.1. □

Lemma A.3. *Assume that that $A_i, i \le d$ are events and Z_i is an unbiased estimator for $\mathbb{P}(A_i)$ for $i \le d$. Let $\mathcal{I}$ be an integer valued random variable with*

$$\mathbb{P}(\mathcal{I} = i) = \frac{z_i}{z}, \quad z := \sum_{j=1}^d z_j,$$

with $z_i > 0$ some positive constants. Then $Z := z\sum_{k=1}^d \mathbb{I}_{\{\mathcal{I}=k\}} \frac{Z_k}{z_k}$ is an unbiased estimator for $\sum_{i=1}^d \mathbb{P}(A_i)$ with

$$\mathbb{E}\left\{Z^2\right\} = z\sum_{i=1}^d \frac{\mathbb{E}\left\{Z_i^2\right\}}{z_i}.$$

Proof. The proof follows by straightforward calculations. □

Acknowledgments. The authors would like to thank the referees for their careful reading and helpful comments. Dominik Kortschak has been supported by the Swiss National Science Foundation Project 200021-124635/1. Both authors also acknowledge partial support from Swiss National Science Foundation Project 200021-1401633/1.

References

[1] Albrecher, H., Asmussen, S., and Kortschak, D. (2006) Tail asymptotics for the sum of two heavy tailed dependent risks. *Extremes*,**9**,2, 107–130.

[2] Albrecher, H., and Kortschak, D. (2009) Asymptotic results for the sum of dependent non-identically distributed random variables *Meth. Comp. Appl. Probab.* **11**,3, 279–306.

[3] Asmussen, S., Blanchet, J., Juneja, S., Rojas-Nandayapa, L. (2011) Efficient simulation of tail probabilities of sums of correlated lognormals. *Ann. Oper. Res.* **189**,1, 5–23.

[4] Asmussen, S., and Kroese, D.P. (2006) Improved algorithms for rare event simulation with heavy tails. *Adv. Appl. Probab.*, **38**,2,545–558.

[5] Asmussen, S., and Rojas-Nandayapa, L. (2008) Sums of dependent log-Gaussian random variables with Gaussian copula. *Stat. Prob. Lett.* **78**, 2709–2714.

[6] Asmussen, S., and Glynn, P.W. (2007) *Stochastic simulation. Algorithms and analysis.* Berlin, Springer.

[7] Bingham, N.N., Goldie, C.M., and Teugels, J. (1987) *Regular Variation.* Cambridge University Press.

[8] Cambanis, S., Huang, S., and Simons, G. (1981) On the theory of elliptically contoured distributions. *J. Multivariate Analysis,* **11**,3, 368–385.

[9] Davis, R.A, and Resnick, S.I. (1988) Extremes of moving averages of random variables from the domain of attraction of the double exponential distribution. *Stoch. Proc. Appl.* **30**, 41–68.

[10] Hamada, M., and Valdez, E. (2008) CAPM and option pricing with elliptically contoured distributions. *The Journal of Risk & Insurance*, **75**,2, 387-409.

[11] Geluk, J., and Tang, Q. (2009) Asymptotic tail probabilities of sums of dependent subexponential random variables. *J. Theoretical Probability* **22**,4, 871–882.

[12] Goovaerts, M., Kaas, R., Laeven, R., Tang, Q., and Vernic, R. (2005) The tail probability of discounted sums of Pareto-like losses in insurance. *Scandinavian Actuarial Journal*, **6**, 446–461.

[13] Hashorva, E. (2009) Exact tail asymptotics of Dirichlet distributions. (*arXiv:0904.0144v1*).

[14] Hashorva, E. (2011) Aggregation and conditional limits of risks in log-elliptical models. Preprint.

[15] Hashorva, E. (2012) Exact tail asymptotics in bivariate scale mixture models. *Extremes*, **15**, 109–128.

[16] Hashorva, E. (2013) Minima and maxima of elliptical triangular arrays and spherical processes. *Beroulli*, in press.

[17] Juneja, S. (2007) Estimating tail probabilities of heavy-tailed distributions with asymptotically zero relative error. *QUESTA*, **57**, 115–127.

[18] Kortschak, D. (2011) Second order tail asymptotics for the sum of dependent, tail-independent regularly varying risks. *Extremes*, **15**, 353–388.

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

[20] Pratesi, M., Santiccu, F., and Graziosi, F. (2006) Generalized moment matching for the linear combination of lognormal RVs: application to outage analysis in wireless systems. *IEEE Trans. Wireless Communications*, **5**, 1122-1132.

[21] Mitra, A., and Resncik, S.I. (2009) Aggregation of rapidly varying risks and asymptotic independence. *Adv. Appl. Probab.* **41**, 797–828.

[22] Rojas-Nandayapa, L. (2008) Risk probabilities: Asymptotics and simulation. *PhD Thesis, Aarhus University, Denmark.*

[23] R Development Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2009. ISBN 3-900051-07-0.

[24] Resnick, S.I. (1987) *Extreme Values, Regular Variation and Point Processes. Springer,* New York.

[25] Tellambura, S. (2008) Bounds on the distribution of a sum of correlated lognormal random variables and their application. *IEEE Trans. Communications,* **56**, 1241-1248.

[26] Valdez, E.A., Dhaene, J., Maj, M., and Vanduffel, S. (2009) Bounds and approximations for sums of dependent log-elliptical random variables. *Insurance: Mathematics and Economics*, **44**,3, 385–397.

[27] Vanduffel, S., Chena, C.X., Dhaene, J., Goovaerts, M., Henrard, L., Kaas, R. (2008) Optimal approximations for risk measures of sums of lognormals based on conditional expectations. *J. Comp. Appl. Math.* **221**, 202-218.