EFFICIENT TAIL ESTIMATION FOR SUMS OF CORRELATED LOGNORMALS

Jose Blanchet

Industrial Engineering and Operations Research
Columbia University

Sandeep Juneja

School of Technology and Computer Science
Tata Institute of Fundamental Research

Leonardo Rojas-Nandayapa
Department of Mathematical Sciences

ABSTRACT

Our focus is on efficient estimation of tail probabilities of sums of correlated lognormals. This problem is motivated by the tail analysis of portfolios of assets driven by correlated Black-Scholes models. We propose three different procedures that can be rigorously shown to be asymptotically optimal as the tail probability of interest decreases to zero. The first algorithm is based on importance sampling and is as easy to implement as crude Monte Carlo. The second algorithm is based on an elegant conditional Monte Carlo strategy which involves polar coordinates and the third one is an importance sampling algorithm that can be shown to be strongly efficient.

1 INTRODUCTION

We are concerned with the problem of efficient estimation of tail probabilities of sums of random variables that are correlated and possess heavy tails. This situation might arise, for instance, when computing the probability of large losses or high returns on a portfolio of correlated asset prices. A basic model in the financial literature is the so-called Black-Scholes model, in which stock prices follow a lognormal distribution. Although the Black-Scholes model is mostly used as a benchmark for pricing, in risk management, where model complexity can impact computational burden enormously, the Black-Scholes model assumptions are popular (cf. McNeil, Frey, and Embrechts 2005). Motivated by these types of problems, our focus is on efficient tail estimation of sums of correlated lognormals. More precisely, let $\mathbf{Y} = (Y_1, Y_2, ..., Y_d)^T$ be a d-dimensional vector distributed jointly Gaussian with mean $\mu = (\mu_1, ..., \mu_d)^T$ and covariance matrix Σ (we say that $\mathbf{Y} \sim N(\mu, \Sigma)$). Finally, define $X_i = \exp(Y_i)$ and $S_d = X_1 + ... + X_d$. We are interested in the efficient estimation of $\alpha(b) = P(S_d > b)$ as $b \nearrow \infty$.

Recall that an *unbiased* estimator Z_b for $\alpha(b)$ is said to be *weakly efficient* or *asymptotically optimal* if $\log E Z_b^2 / \log \alpha(b) \to 2$ as $b \nearrow \infty$. Moreover, the estimator is *strongly efficient or is said to have bounded relative error* if $\sup_{b \geq 0} E Z_b^2 / \alpha(b)^2 < \infty$. These notions are standard in rare event simulation (cf. Asmussen and Glynn 2008, Bucklew 2004, Juneja and Shahabuddin 2006).

Most of the literature on rare-event simulation for heavy-tailed systems has focused on random walk-type models (cf. Asmussen and Kroese 2006, Juneja and Shahabuddin 2002, Dupuis, Leder, and Wang 2007, Blanchet and Glynn 2008, Blanchet, Glynn, and Liu 2007). In contrast, we consider a rare-event simulation problem that involves the sum of highly dependent increment distributions. The dependence structure makes the available rare-event simulation algorithms for tails of sums of iid heavy-tailed increments difficult to apply in our current setting because they rely heavily on the iid assumption.

As we mentioned before that our current setting relates to applications in finance, in the context of tail probabilities of assets driven by correlated Black-Scholes models. In this context, a popular approach that is often suggested is approximating the prices by a *t*-distributed model. Such approximation is motivated by means of a Taylor expansion which is often called a Delta approximation, if it involves the first derivative only or Delta-Gamma approximation, if the first and second derivatives are considered Glasserman 2000. The use of *t*-distributions is appealing in these settings in order to capture the heavy-tailed behavior which is present in the original lognormal model (which is approximated by means of the Delta-Gamma development). Efficient rare-event simulation procedures are then designed for the Delta-Gamma approximation with *t*-distributed factors or quadratic forms of Gaussian factors Glasserman 2000, Glasserman, Heidelberger, and Shahabuddin 1998. The simulation estimators that we propose and analyze here avoid the need for a Delta-Gamma approximation by working directly with the lognormal factors in an efficient way. So, we do not incur in bias errors that are inherent to the use of the Delta-Gamma approximation and, at the same time, efficiency of the estimators is preserved.

Our contributions are as follows. We analyze and propose three estimators for $\alpha(b)$. The first estimator is based on an importance sampling strategy that scales the covariance matrix by a factor that grows at a suitable speed as $b \nearrow \infty$. Since the sampler involves a simple scaling, the estimator is straightforward to implement and it can be shown to be asymptotically optimal as $b \nearrow \infty$. The second estimator is based on a conditional Monte Carlo strategy which is applied after parameterizing the underlying Gaussian factors in polar coordinates. The strategy is to condition on the angular component and integrate the radial contribution in closed form. Finally, the third of our estimator takes advantage of the fact that the largest of the increments dominates the large deviations behavior of the sums of correlated lognormals. The strategy is decompose the tail event of interest in two contributions, a dominant piece corresponding to the tail of the maximum and a remaining contribution. The dominant contribution is analyzed by means of a strongly efficient algorithm for the maximum of multivariate Gaussians and the remaining contribution is independently handled using the importance sampling strategy utilized in the design of the first estimator. The third estimator can be shown to be strongly efficient.

The rest of the paper is organized as follows. Basic large deviations results for sums of correlated lognormals are briefly discussed in Section 2. The description of our first estimator is given in Section 3, whereas our conditional Monte Carlo estimator is analyzed in Section 4. Our third estimator is discussed in Section 5. Finally, numerical examples are given in our last section, namely, Section 6.

2 LARGE DEVIATIONS FOR SUMS OF CORRELATED LOGNORMALS

It is well known Embrechts, Klüppelberg, and Mikosch 1997 that the lognormal distribution belongs to the class of subexponential distributions and therefore, if the X_i's are iid, then

$$\alpha(b) \sim P\left(\max_{1 \le j \le d} X_j > b\right) \sim dP(X_j > b) \qquad (1)$$

as $b \nearrow \infty$. The previous approximation in particular implies that, asymptotically as $b \nearrow \infty$, the event $\{S_d > b\}$ occurs due to the contribution of a large increment (the maximum) which must exceed b.

Our focus, however, is on general correlation structure among the underlying Gaussian factors. Assuming only that all pairwise correlations between the Y_i's are strictly less than 1, we obtain Asmussen and Rojas-Nandayapa 2008 that basically the same intuition behind (1) holds true. More

precisely, we have that

$$\alpha(b) \sim P\left(\max_{1 \le j \le d} X_j > b\right) \sim \sum_{j=1}^{d} P(X_j > b) \qquad (2)$$

as $b \nearrow \infty$. The previous asymptotic result is a consequence of a basic feature about multivariate Gaussians, namely, the asymptotic independence at the tails. That is, if Z_1 and Z_2 are correlated standard Gaussians with $cov(Z_1, Z_2) < 1$, then $P(Z_1 > b | Z_2 > b) \to 0$ as $b \nearrow \infty$.

Another elementary, yet important, piece of information that we must recall in order to proceed with the description and the motivation of our algorithms is the following elementary asymptotic result for Gaussian random variables known as Mill's ratio Resnick . If $Y_i \sim N\left(\mu_i, \sigma_i^2\right)$ then as $y \nearrow \infty$

$$P(Y_i > y) \sim \frac{\sigma_i}{(2\pi)^{1/2}(y - \mu_i)} \exp\left(-\frac{(y - \mu_i)^2}{2\sigma_i^2}\right). \qquad (3)$$

Note that, in particular, approximation (3) indicates that the tail of the a Gaussian (or equivalently the tail of a lognormal) is heavier for larger values of the variance parameter σ_i^2.

In addition to the asymptotic approximations (3) and (2) the discussion above tells us useful information that we shall leverage off in the design of efficient simulation estimators. First, the tail of the maximum drives the tail of the sum, even in the case of correlated lognormals (we will use this piece in Section 5) and, second, the tail of the sum is driven (at least in logarithmic scale) by those Gaussian components that have the largest variance parameter, regardless of the correlation structure as long as no redundant components are present (the estimators in Sections 3 and 4 are based on this observation).

3 ASYMPTOTICALLY OPTIMAL IMPORTANCE SAMPLING VIA VARIANCE SCALING

A principle that is popular in financial risk analysis is that high variance or volatility is associated to high risk. Of course, one has to be careful when applying this principle in light of what is meant by risk. Typically, the notion of risk is associated to tail behavior and, in general, variance has little to do with tail behavior. However, as we saw in Section 2, in the case of Gaussian random variables, the variance controls the tail behavior of the underlying factors.

Using the previous principle a natural importance sampling strategy that one might consider for computing $\alpha(b)$ is one that induces high variances. This motivates considering as importance sampler a distribution a $N(\mu, \Sigma/(1-\theta))$ for some $0 < \theta < 1$; in other words, relative to the nominal (original) probability distribution, we just inflate the covariance matrix by the factor $1/(1-\theta)$. This importance

sampling distribution is denoted by $P_\theta(\cdot)$ and we shall use the notation $E_\theta(\cdot)$ for the associated expectation operator.

The estimator induced by this simple strategy is

$$Z_1(b) = \frac{I(S_d > b)\exp\left(-\mathbf{Y}^T\Sigma^{-1}\mathbf{Y}/2\right)/\det(\Sigma)^{1/2}}{\exp\left(-\mathbf{Y}^T\Sigma^{-1}\mathbf{Y}(1-\theta)/2\right)\det(\Sigma/(1-\theta))^{1/2}}$$

$$= I(S_d > b)\frac{\exp\left(-\theta\mathbf{Y}^T\Sigma^{-1}\mathbf{Y}/2\right)}{(1-\theta)^{d/2}}.$$

The next lemma summarizes a useful representation for the second moment of $Z_1(b)$ under the importance sampling distribution. However, in order to state such representation we introduce another family of probability measures (in addition to the P_θ's), which we shall denote by $(Q_\theta : 0 \leq \theta \leq 1)$. We use $Q_\theta(\cdot)$ to denote a probability measure under which $\mathbf{Y}$ is $N(\mu, \Sigma/(1+\theta))$.

Proposition 1

$$E_\theta Z_1^2(b) = (1-\theta^2)^{-d/2}Q_\theta(S_d > b) \qquad (4)$$

Proof.

$E_\theta Z_1^2(b)$

$$= \int \frac{I(e^{y_1} + ... + e^{y_n} > b)\exp\left(-2\theta(\mathbf{y}-\mu)^T\Sigma^{-1}(\mathbf{y}-\mu)/2\right)}{(1-\theta)^d}$$

$$\times \frac{\exp\left(-(1-\theta)(\mathbf{y}-\mu)^T\Sigma^{-1}(\mathbf{y}-\mu)/2\right)}{(2\pi)^{d/2}\det(\Sigma/(1-\theta))}dy_1...dy_n$$

$$= \int I(e^{y_1} + ... + e^{y_n} > b)$$

$$\times \frac{\exp\left(-(1+\theta)(\mathbf{y}-\mu)^T\Sigma^{-1}(\mathbf{y}-\mu)/2\right)(1+\theta)^{-d/2}}{(1-\theta)^d(1-\theta)^{-d/2}(2\pi)^{d/2}\det(\Sigma/(1+\theta))}$$

$$= (1-\theta^2)^{-d/2}Q_\theta(S_n > b).$$

∎

As an immediate consequence of the previous result we obtain that the estimator $Z_1(\theta)$ is logarithmic efficient if one chooses $\theta(b) \to 1$ at an appropriate speed. In particular, the asymptotic behavior of $Q_\theta(\cdot)$ (as given by (3)), suggests selecting θ close to unity.

Theorem 1 *Suppose that* $0 \leq 1 - \theta(b) = o\left(1/(\log(b))^2\right)$, *then*

$$E_{\theta(b)}Z_1^2(b) \sim o\left(\log(b)^d\right)2^{-d/2}\alpha(b)^2 \qquad (5)$$

as $b \nearrow \infty$ *and in particular,* $Z_1(b)$ *is logarithmically efficient.*

Proof. [Sketch of the Proof]The proof proceeds by taking advantage of (4) directly together with approximations such as (2) and (3). The only missing details in the argument, which is given in Blanchet, Juneja, and Rojas-Nandayapa 2008, involve covering asymptotics for (4) when the covariance matrix is changing according to $\theta(b)$. ∎

One can choose $\theta(b)$ in many ways which are consistent with the condition that $1 - \theta(b) = o\left(1/(\log(b))^2\right)$ as $b \to \infty$. One of them involves finding $\theta(b)$ that minimizes the asymptotic expressions for the second moment of the estimator given by (4). A simpler approach is to find the unique positive root $\theta(b)$ (which exists for b large enough) to the equation $E_{\theta(b)}S_d = b$. This root-finding procedure does not contribute significantly to the computational cost of the algorithm because is done just once and is straightforward to see that the condition $1 - \theta(b) = o\left(1/\log(b)^2\right)$ is satisfied as $b \nearrow \infty$. We implemented the numerical examples that we will show in Section 6 using the root finding approach.

Although the estimator $Z_1(\theta)$ possesses two very convenient features, namely, is very easy to implement and is asymptotically optimal, it also has the disadvantage that the premultiplying factor in the asymptotic variance expression (5) might grow substantially. For instance, in the case of the root finding selection, such factor is of order $O\left(\log(b)^{d/2}\right)$. So, for moderate values of b and d, the variance performance of the estimator might degrade in a significant way. To cope with this problem one can introduce additional variance reduction techniques. In the next two sections we introduce a couple of alternatives that help address this problem, one can obviously think about combining the strategies that we shall discuss in future sections.

4 ASYMPTOTICALLY OPTIMAL CONDITIONAL MONTE CARLO

The estimator discussed in the previous section showed a successful (in the sense of achieving asymptotic optimality) importance sampling strategy based on a simple parametric family of distributions that places all the effort of the sampler uniformly on the variances. In other words, we induce the rare event by multiplying the whole covariance matrix by the same factor, namely, $1/(1-\theta)$.

Now, recall that the vector $\mathbf{Y}$ can be represented as $\mathbf{Y} = \mu + C\mathbf{Z}$, where $\mathbf{Z} = (Z_1, ..., Z_d)^T$ is a vector of iid standard Gaussians and $CC^T = \Sigma$. In turn, we can represent $\mathbf{Z}$ in polar coordinates, as $\mathbf{Z} = R\phi$, where R is the radial distance, ϕ is a point selected uniformly over the surface of the unit ball in d dimensions and ϕ and R are independent. It then follows that we can interpret the importance sampling strategy in the previous section as a change in scale applied to R only. In turn, this suggests that the randomness of R contributes substantially to the variance of the estimator and this motivates our search for a conditional Monte Carlo

algorithm that integrates in closed form the contribution of the radial distance to the evaluation of $\alpha(b)$.

More precisely, let $C_{i,\cdot}$ denote the i-th column of the matrix C and write $W_i = C_{i,\cdot}\phi$, so that $X_i = \exp(\mu_i + RW_i)$. Our conditional Monte Carlo estimator takes the form

$$Z_2(b) = P(S_d > b \,|\, \phi) = P(S_d > b \,|\, W_1, ..., W_d).$$

In order to generate $Z_2(b)$, we need to generate copies of ϕ and, given a realization $W_i = w_i$, we need to evaluate $P(R \in A_b)$, for A_b of the form

$$A_b = \{r \geq 0 : \exp(\mu_1 + rw_1) + ... + \exp(\mu_d + rw_d) > b\}.$$

Generating copies of ϕ is easily done by computing $\mathbf{Z}/\|\mathbf{Z}\|_2$. We also know that $R^2 \sim \chi_d^2$ (chi-squared with d degrees of freedom) or, equivalently, R^2 is distributed gamma with mean d and variance $2d$. So we just need to find parameterize the set A_b in an appropriate way. Fortunately, this can be easily done in terms of a root finding procedure as we now describe. Define

$$f(r) = \exp(\mu_1 + rw_1) + ... + \exp(\mu_d + rw_d)$$

and note that $f(\cdot)$ is the sum of exponentials and therefore is strictly convex. As a consequence, A_b can be characterized in terms of the *non-negative roots of the equation $f(r) = b$* as follows: a) No solution exists on $(0, \infty)$ and $f(0) \geq b$ in which case $A_b = (0, \infty)$, b) No solution exists on $(0, \infty)$ and $f(0) < b$ in which case $A_b = \oslash$, c) there is only one solution on $(0, \infty)$, say $r_+(b) > 0$ and $f(0) \leq b$, in which case $A_b = (r_+(b), \infty)$, d) there is only one solution on $(0, \infty)$ and $f(0) > b$, in which case $A_b = (0, r_+(b))$ and e) there are two solutions $0 \leq r_-(b) < r_+(b)$, in which case $A_b = (0, r_-(b)) \cup (r_+(b), \infty)$. The following lemma, which follows directly from the fact that $f(\cdot)$ is the sum of exponential functions, is useful to find the precise form of A_b given particular parameters μ_i and w_i for $1 \leq i \leq d$.

Lemma 1 *i) If $\min_{1 \leq i \leq d} w_i > 0$ then there is at most one solution to the equation $f(r) = b$ on $(-\infty, \infty)$. Moreover, a non-negative solution exists if and only if $f(0) > b$; if no solution exists, then $A_b = \oslash$ whereas if a solution $r_+(b) > 0$ exists then $A_b = (0, r_+(b))$.*

ii) If $\min_{1 \leq i \leq d} w_i > 0$ then there is at most one solution to the equation $f(r) = b$ on $(-\infty, \infty)$. Moreover, a non-negative solution exists if and only if $f(0) < b$; if no solution exists, then $A_b = (0, \infty)$ whereas if a solution $r_+(b) > 0$ exists then $A_b = (r_+(b), \infty)$.

iii) If $\min_{1 \leq i \leq d} w_i < 0 < \max_{1 \leq i \leq d} w_i$ then the equation $f'(r^*) = 0$ has exactly one solution on $(-\infty, \infty)$ and $f(r^*) = \min_{r \in \mathbb{R}} f(r)$.

Parts i) and ii) from the previous lemma can be applied to construct the set A_b in some of the cases described by

a) to d) above (we just need to apply Newton-Raphson's method in order to find the required root when it exists). The situation described by iii) is helpful to construct the set A_b when there is more than one root to the equation $f(r) = b$. In such situation we first run Newton-Raphson's method to find r^* such that $f'(r^*) = b$. If $f(r^*) \geq b$ then $A_b = (0, \infty)$, otherwise, there are exactly two roots of the equation $f(r) = b$, which can be easily located by applying Newton-Raphson's method with initial conditions $r^* - 1$ and $r^* + 1$ respectively. The form of A_b can then be easily obtained by looking at the sign of the roots and following the line of reasoning indicated in a) or e) above.

The analysis of the total computational cost required to implement the estimator $Z_2(b)$ requires a careful study of the Newton-Raphson procedures indicated in the previous paragraph. A detailed discussion of the issues involved is given in Blanchet, Juneja, and Rojas-Nandayapa 2008.

We close the section with the following result, which summarizes the efficiency properties of our conditional Monte Carlo estimator.

Theorem 2 *Assuming that the pairwise correlations of the Y_i's are strictly less than unity, then*

$$\frac{\log EZ_2^2(b)}{\log \alpha(b)} \to 2$$

as $b \nearrow \infty$.

Proof. [Sketch of the Proof] Consider the problem $\max_{\|\phi\|_2 = 1} C_{i,\cdot}\phi$. By Cauchy-Schwarz inequality, the (unique) solution to this problem is easily seen to satisfy $\phi_i^* = C_{i,\cdot}/\|C_{i,\cdot}\|_2$. Since the pairwise correlations between the Y_i's are strictly less than unity, then is not difficult to see that $\phi_i^* \neq \phi_j^*$ for $i \neq j$. Now, let $\mathcal{T} = \{i : \sigma_i = \max_{1 \leq j \leq d} \sigma_j\}$ and select $i^* \in \mathcal{T}$ such that $\mu_{i^*} \geq \mu_j$ for all $j \in \mathcal{T}$. Since $\sigma_{i^*} = \|C_{i^*,\cdot}\|_2$ we have (using the fact that $\phi_i^* \neq \phi_j^*$ for $i \neq j$) that for each $\delta > 0$, there exists a deterministic constant $k(\delta) > 0$ such that for all $r \geq 0$

$$f(r) \leq \exp(\mu_{i^*} + r\sigma_{i^*}(1 + \delta)) + k.$$

Therefore, elementary calculus and approximation (2) allows to conclude

$$Z_2(b) \leq P(\exp(\mu_{i^*} + R\sigma_{i^*}/(1 + \delta)) > b - k)$$
$$= O\left(\alpha(b)^{1/(1+\delta)} \log(b)^d\right)$$

as $b \nearrow \infty$. Since $\delta > 0$ is arbitrary, the statement of the theorem follows. ∎

5 STRONGLY EFFICIENT IMPORTANCE SAMPLING ESTIMATOR

In Section 2 we noted that large values of S_d are basically due to the contribution of a single large jump (the maximum) and in Section 3 we constructed a weakly efficient estimator using an importance sampler based on the fact that, roughly speaking (i.e. in logarithmic sense) and according to (3) and (2), the variances dictate the tail behavior S_d. The idea in this section is to combine these two intuitive observations in order to produce a strongly efficient importance sampling estimator. First, note that

$$\alpha(b) = \alpha_1(b) + \alpha_2(b),$$

where

$$\alpha_1(b) = P\left(\max_{1 \leq i \leq d} X_i > b\right),$$

$$\alpha_2(b) = P\left(S_d > b, \max_{1 \leq i \leq d} X_i \leq b\right).$$

In view of (2) we must have that $\alpha_2(b) = o(\alpha_1(b))$ as $b \nearrow \infty$, so the most important contribution comes from the term $\alpha_1(b)$. The strategy is to design independent and unbiased estimators, say $Z_{3,1}(b)$ and $Z_{3,2}(b)$, for the terms $\alpha_1(b)$ and $\alpha_2(b)$ respectively. The gain comes if $Z_{3,1}(b)$ is strongly efficient for $\alpha_1(b)$ even if $Z_{3,2}(b)$ has a coefficient of variation of order $O(\alpha(b)/\alpha_2(b))$ as $b \nearrow \infty$. In other words, $Z_{3,2}(b)$ may not be strongly efficient for $\alpha_2(b)$, but its coefficient of variation could grow slowly enough so that the combined estimator $Z_3(b) = Z_{3,1}(b) + Z_{3,2}(b)$ for $\alpha(b)$ is strongly efficient.

For $Z_{3,2}(b)$ we propose to use (recall the notation introduced in Section 3) P_θ as our importance sampling distribution (i.e. $\mathbf{Y}$ has distribution $N(\mu, \Sigma/(1+\theta))$) and set

$$Z_{3,2}(b) = I\left(S_d > b, \max_{1 \leq i \leq d} X_i \leq b\right) \frac{\exp\left(-\theta \mathbf{Y}^T \Sigma^{-1} \mathbf{Y}/2\right)}{(1-\theta)^{d/2}}.$$

Just as in Section 3, we conclude that

$$E_\theta Z_{3,2}(b)^2 = \left(1 - \theta^2\right)^{-d/2} Q_\theta\left(S_d > b, \max_{1 \leq i \leq d} X_i \leq b\right).$$

A key insight behind the efficiency analysis of $Z_{3,2}(b)$ is that $\alpha_2(b)/\alpha_1(b)$ goes to zero at least at a polynomial rate in b, which dominates any function with growth $O(\log(b)^a)$ for $a > 0$ (which, in turn, is the type of behavior that arises when selecting $E_{\theta(b)} S_d = b$ as suggested in Section 3.

Finally, we turn our attention to $Z_{3,1}(b)$, which involves computing $\alpha_1(b) = P\left(\max_{1 \leq j \leq d} Y_j > \log(b)\right)$. We shall use $f_j(y_j)$ to denote the marginal density of Y_j evaluated at

$y_j \in \mathbb{R}$ and y_{-j} to denote the vector $(y_1, .., y_{j-1}, y_{j+1}, ..., y_d)$. The expression $f(y_{-j}|y_j)$ is used to denote the conditional density of $\mathbf{Y}_{-j} = (Y_1, .., Y_{j-1}, Y_{j+1}, ..., Y_d)^T$ evaluated at y_{-j} given $Y_j = y_j$. The density of the vector $\mathbf{Y}$ evaluated at y is denoted by $f(y)$. Note that for all j we have that $f(y) = f_j(y_j) f(y_{-j}|y_j)$. We consider as importance sampling density $g(\cdot)$ defined via

$$g(y) = \sum_{j=1}^{d} p_j(b) f_j(y_j) f(y_{-j}|y_j) \frac{I(y_j > \log(b))}{P(Y_j > \log(b))},$$

where

$$p_j(b) = P(Y_j > \log(b)) / \left(\sum_{i=1}^{d} P(Y_i > \log(b))\right).$$

We shall use the notation $Var_g(\cdot)$ to denote the variance operator under the probability measure induced by $g(\cdot)$.

In other words, we first select the j-th index with probability proportional to $P(Y_j > \log(b))$. Then, given that j^* has been selected we sample Y_{j^*} given that $Y_{j^*} > \log(b)$. Finally, we sample the rest of the components under the nominal distribution given that $Y_{j^*} = y_{j^*}$ (i.e. we use the law $f(\cdot|y_{j^*})$). The corresponding estimator is

$$Z_{3,1}(b) = f(\mathbf{Y})/g(\mathbf{Y})$$
$$= \frac{\sum_{i=1}^{d} P(Y_i > \log(b))}{\sum_{j=1}^{d} I(Y_j > \log(b))} \leq \sum_{i=1}^{d} P(Y_i > \log(b)).$$

This sampler is proposed and studied in Adler, Blanchet, and Liu 2008. It is shown that if the pairwise correlations among the Y_i's are strictly less than unity then the coefficient of variation of the estimator $Z_{3,1}(b)$ converges to zero as $b \nearrow \infty$, a property that is even stronger than bounded relative error.

Combining the discussion indicated in previous paragraphs about the estimator $Z_{3,1}(b)$ and $Z_{3,2}(b)$ we arrive at the following result, which summarizes the performance of the estimator $Z_3(b) = Z_{3,1}(b) + Z_{3,2}(b)$.

Theorem 3 *Suppose that* $0 \leq 1 - \theta(b) = o\left(1/(\log(b))^2\right)$. *Then, the unbiased estimator* $Z_3(b)$ *is strongly efficient in the sense that*

$$\sup_{b \geq 0} \frac{Var(Z_3(b))}{\alpha(b)^2}$$
$$= \sup_{b \geq 0} \left(\frac{Var_g(Z_{3,1}(b))}{\alpha(b)^2} + \frac{Var_{\theta(b)}(Z_{3,2}(b))}{\alpha(b)^2}\right) < \infty$$

6   NUMERICAL EXPERIMENTS

Table 1: $\mathbb{P}(S_{10} > 50)$

Method	Estimator	CV
AK	0.002588204	0.450933837
IS	0.002536418	7.668013424
CM	0.002619456	2.990221532
ISS	0.002574308	6.469782624

Table 2: $\mathbb{P}(S_{10} > 75)$

Method	Estimator	CV
AK	0.000250592	0.362576756
IS	0.000233741	12.834976956
CM	0.000239655	4.958485430
ISS	0.000200905	9.356164179

Table 3: $\mathbb{P}(S_{10} > 100)$

Method	Estimator	CV
AK	0.000048978	0.263833033
IS	0.000068366	31.283224854
CM	0.000048858	7.368722011
ISS	0.000066627	34.425338337

Table 4: $\mathbb{P}(S_{10} > 12500)$

Method	Estimator	CV
IS	0.000615578	19.997044444
CM	0.000575373	10.528366085
ISS	0.000548409	0.847424920

Table 5: $\mathbb{P}(S_{10} > 25000)$

Method	Estimator	CV
IS	0.000213384	22.000276130
CM	0.000257832	12.333104669
ISS	0.000241408	0.936462139

Table 6: $\mathbb{P}(S_{10} > 37500)$

Method	Estimator	CV
IS	0.000254008	26.417180649
CM	0.000141805	18.442528446
ISS	0.000148114	1.904853911

We implemented the three estimators described above in three examples corresponding to low, medium and high correlations. In particular, Example 1 assumes that the Gaussian factors involved are iid; in Example 2 we use a common correlation coefficient equal to .4 and Example 3 involves a common correlation coefficient equal to .9. In the construction of the tables we use the following abbreviations: IS denotes the importance sampling strategy based on variance scaling discussed in Section 3, CM corresponds to the conditional Monte Carlo method discussed in Section 4, ISS relates to the strongly efficient importance sampling strategy described in Section 5. Finally, CV denotes the empirical coefficient of variation of the estimator (i.e. the empirical standard deviation divided by the empirical mean). In Example 1 we have also implemented the estimator of Asmussen and Kroese 2006 for comparison, denoted by AK, which shows excellent performance. We used 1,000 replications for each of the estimators shown in all the examples.

We note, however, that even in the iid case, the performance of the ISS estimator is a good competitor against the AK estimator. Overall, not surprisingly given that ISS has bounded relative error, its performance is the strongest among the three estimators discussed. The CM estimator seems to achieve higher variance reduction than IS, however, the rootfinding procedure involved in the implementation of CM makes somewhat slower than IS. Still, CM might be a good alternative in terms of robustness, given that it always achieves variance reduction (compared to IS which, in some problem instances and for relatively small values of the tail parameter, might not achieve variance reduction).

Example 1   In Tables 1–3 we use 10 lognormal r.v.'s iid with parameters $\mu = 0$ and $\sigma^2 = 1$.

Example 2   In Tables 4–7 we use 10 lognormal with

$$\mu_i = -i \qquad \sigma_i^2 = i \qquad i = 1,\ldots,n \qquad (6)$$

and common correlation $\rho = .4$.

Example 3   In Tables 8–11 we use 10 lognormal with

$$\mu_i = -i \qquad \sigma_i^2 = i \qquad i = 1,\ldots,n \qquad (7)$$

and common correlation $\rho = .9$.

ACKNOWLEDGEMENTS

Table 7: $\mathbb{P}(S_{10} > 50000)$

Method	Estimator	CV
IS	0.000143007	27.407531643
CM	0.000101505	15.435680709
ISS	0.000101904	0.681132772

Table 8: $\mathbb{P}(S_{10} > 12500)$

Method	Estimator	CV
IS	0.000777656	18.153949459
CM	0.000611142	11.076718679
ISS	0.000418335	15.884860953

Table 9: $\mathbb{P}(S_{10} > 25000)$

Method	Estimator	CV
IS	0.000233472	19.799092654
CM	0.000315217	12.805770365
ISS	0.000168077	7.429588684

Table 10: $\mathbb{P}(S_{10} > 37500)$

Method	Estimator	CV
IS	0.000161555	25.666874186
CM	0.000177439	15.047519196
ISS	0.000119861	16.046135126

Table 11: $\mathbb{P}(S_{10} > 50000)$

Method	Estimator	CV
IS	0.000110665	28.735442408
CM	0.000114764	15.797186276
ISS	0.000069387	7.862441931

REFERENCES

Adler, R., J. Blanchet, and J. C. Liu. 2008. Fast simulation of gaussian random fields. In *Winter Simulation Conference, to appear*.

Asmussen, S., and P. Glynn. 2008. *Stochastic simulation: Algorithms and analysis*. New York: Springer-Verlag.

Asmussen, S., and D. Kroese. 2006. Improved algorithms for rare event simulation with heavy tails. *Advances in Applied Probability* 38:545–558.

Asmussen, S., and L. Rojas-Nandayapa. 2008. Asymptotics of sums of lognormal random variables with gaussian copula. *Stat. Probab. Lett.* 78.

Blanchet, J., and P. Glynn. 2008. Efficient rare-event simulation for the maximum of a heavy-tailed random walk. *Ann. of Appl. Probab.* 18:1351–1378.

Blanchet, J., P. W. Glynn, and J. Liu. 2007. Fluid heuristics, Lyapunov bounds and efficient importance sampling for a heavy-tailed G/G/1 queue. *QUESTA* 57:99–113.

Blanchet, J., S. Juneja, and L. Rojas-Nandayapa. 2008. Efficient rare-event simulation for portfolios of black-scholes assets. Technical report, Columbia University, New York, NY.

Bucklew, J. 2004. *Introduction to rare-event simulation*. New York: Springer-Verlag.

Dupuis, P., K. Leder, and H. Wang. 2007. Importance sampling for sums of random variables with regularly varying tails. *ACM TOMACS* 17:1–21.

Embrechts, P., C. Klüppelberg, and T. Mikosch. 1997. *Modelling extremal events for insurance and finance*. New York: Springer-Verlag.

Glasserman, P. 2000. *Monte carlo methods in finance*. New York: Springer.

Glasserman, P., P. Heidelberger, and P. Shahabuddin. 1998. Gaussian importance sampling and stratification: Computational issues. In *Proceedings of the 1998 Winter Simulation Conference*, ed. D. J. Meideros, E. F. Watson, J. S. Carson, and M. S. Manivannan, 685–693. Piscataway, New Jersey: IEEE Press.

Juneja, S., and P. Shahabuddin. 2002. Simulating heavy-tailed processes using delayed hazard rate twisting. *ACM TOMACS* 12:94–118.

Juneja, S., and P. Shahabuddin. 2006. Rare event simulation techniques: An introduction and recent advances. In *Simulation*, ed. S. G. Henderson and B. L. Nelson, Handbooks in Operations Research and Management Science. Amsterdam, The Netherlands: Elsevier.

McNeil, A., R. Frey, and P. Embrechts. 2005. *Quantitative risk management: Concepts, techniques and tools*. New Jersey: Princeton University Press.

Resnick, S. I. *Adventures in stochastic processes*. Boston: Birkhauser.

AUTHOR BIOGRAPHIES

JOSE BLANCHET is Assistant Professor in the Department of Industrial Engineering and Operations Research at Columbia University. Jose holds a M.Sc. in Engineering-Economic Systems and Operations Research and a Ph.D. in Management Science and Engineering, both from Stanford University. He also holds two B.Sc. degrees: one in

Actuarial Science and another one in Applied Mathematics from ITAM (México). José worked for two years as an analyst in Protego Financial Advisors, a leading investment bank in Mexico. He has research interests in applied probability and stochastic simulation. He serves or has served in the editorial board of journals such as the *Applied Probability Journals* and *TOMACS*. His email address is `<jose.blanchet@columbia.edu>`.

SANDEEP JUNEJA received his Ph.D. in Operations Research from Stanford University in 1994. He is currently an Associate Professor in the School of Technology and Computer Science at the Tata Institute of Fundamental Research in India. His research interests include applied probability, computational finance, simulation and rare-event analysis. He serves on the editorial boards of Management Science and ACM TOMACS. His email address is `<juneja@tifr.res.in>` and his web page is www.tcs.tifr.res.in/∼sandeepj.

LEONARDO ROJAS-NANDAYAPA is currently a PhD student at the Department of Mathematical Sciences, Aarhus University. He holds a BSc in Applied Mathematics from ITAM (México) and a MSc in Statistics and Probability from IIMAS-UNAM (México). His main research interests are in Applied Probability, in particular Simulation Methodology and Insurance and Financial Mathematics. His email address is `<leorojas@imf.au.dk>`.