Proceedings of the 2004 Winter Simulation Conference
R. G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters, eds.

A UNIFIED APPROACH FOR FINITE-DIMENSIONAL, RARE-EVENT MONTE CARLO SIMULATION

Zhi Huang

Fixed Income Research
Lehman Brothers
New York, NY 10019, U.S.A.

Perwez Shahabuddin

Department of IEOR
Columbia University
New York, NY 10027, U.S.A.

ABSTRACT

We consider the problem of estimating the small probability that a function of a finite number of random variables exceeds a large threshold. Each input random variable may be light-tailed or heavy-tailed. Such problems arise in financial engineering and other areas of operations research. Specific problems in this class have been considered earlier in the literature, using different methods that depend on the special properties of the particular problem. Using the Laplace principle (in a restricted finite-dimensional setting), this paper presents a unified approach for deriving the log-asymptotics, and developing provably efficient fast simulation techniques using the importance sampling framework of hazard rate twisting.

1 INTRODUCTION

We consider the problem of estimating the *small* probability that a general function of a fixed number of random variables exceeds a large threshold. The input random variables to the function may be light-tailed or heavy-tailed and are assumed independent; the latter may not necessarily be a limitation due to the existence of copula methods that transform dependent random variables into independent ones. Recently, several specific cases of such problems have been considered in the literature, mainly in the area of financial engineering (e.g., Glasserman, Heidelgerger and Shahabuddin 2002; Huang and Shahabuddin 2003), but also in other areas (e.g., stochastic PERT networks have been considered in Juneja, Karandikar and Shahabuddin 2004 and stochastic shortest path problem in Kroese and Rubinstein 2004). In most of these problems, either the function is very complex or the number of input variables is very large making the analytical and/or numerical computation of this probability very difficult. This coupled with the fact that the event of interest is rare, makes naive simulation also very time consuming. While some papers use adaptive importance sampling changes of measure with no formal log-asymptotics and simulation efficiency results (e.g., Kroese and Rubinstein 2004; however, the adaptive approach uses less system specific properties), others derive log-asymptotics and develop provably efficient (in a certain sense), importance sampling based fast simulation methods using the special properties of the problem at hand.

In this paper we present a unified approach for deriving the log-asymptotics for such probabilities as the threshold goes to infinity (the probability of interest becomes smaller). We then develop a provably efficient importance sampling change of measure under the general framework of hazard rate twisting (Juneja and Shahabuddin 2002). Unlike exponential twisting, hazard rate twisting can be applied to both light-tailed and heavy-tailed random variables. As an illustration, we apply the approach to value-at-risk problems that arise in financial engineering. With this new approach, one can now attempt to develop asymptotics and provably efficient fast simulation techniques directly for the loss probability, rather than its quadratic approximation (only the latter seems to have been done in the literature so far). Hazard rate twisting of input random variables has also been suggested as one of the approaches in Kroese and Rubinstein 2004; however the parameters of the twisting are obtained in an adaptive manner and no formal efficiency results are proved.

Rare event asymptotics and importance sampling based fast simulation have been extensively studied in the light-tailed setting in areas like queueing, reliability, insurance (see, e.g., Bucklew 1990, Heidelberger 1995, Juneja and Shahabuddin 2004 for expositions and surveys). Most of these works deal with estimating probabilities concerning infinite sums and stopped sums of sequences of random variables. Hence these problems are "infinite-dimensional", which in general are much more difficult than the finite-dimensional problems that we consider. However, we consider general functions of random variables rather than simple sums, and this introduces a complexity that is not present in much of the light-tailed, rare-event simulation literature.

In the heavy-tailed setting, even though rare event asymptotics has been studied extensively (see, e.g., Embrechts, Kluppelberg and Mikosch 1997), the techniques used are very different from those used for light-tailed asymptotics. This is because the manner in which rare events occur in the two settings are very different. Provably efficient fast simulation techniques in this area have remained elusive, since unlike the light-tailed setting, it is not easy to transfer the information from asymptotics into importance sampling changes of measure. Till date, in the "infinite-dimensional" setting, provably efficient simulation techniques exist only for estimating the probability that a sum of a geometric number of i.i.d heavy-tailed random variable exceeds a large threshold (Asmussen and Binswanger 1997, Asmussen, Binswanger and Hojgaard 2000, and Juneja and Shahabuddin 2002). This has applications only in some very simple models in queueing and insurance. Some partial success has been obtained for the case of random walks, for some specific heavy-tailed distributions (Boots and Shahabuddin 2000). Some work has also been done regarding heavy-tailed simulation for specific problems in the domain of this paper (see references mentioned in the first paragraph) using innovative approaches that are targeted to the problem at hand. Each such case seems to require a new approach.

This work suggests that at least for the "finite-dimensional" problems that we consider, it is possible to develop a unified theory for both light-tailed and heavy-tailed asymptotics (and fast simulation). It also shows how classical results (like the Laplace principle), that have so far been mainly been used in the domain of light-tailed asymptotic theory, can also be used in the heavy-tailed setting (even though in this restricted finite dimensional setting). Furthermore, it shows that just as exponential twisting played a central role in light-tailed importance sampling, hazard rate twisting seems to play a central role when we consider light-tailed and heavy-tailed simulations under a unified framework.

A formal statement of the problem is given in Section 2. In Section 3 we give specific examples of this problem from the literature. In Section 4 we first develop a large-deviation principle (in the restricted finite-dimensional setting) for a non-negative random vector that may include both light and heavy-tailed random variables. We then introduce a normalization that transforms the problem of estimating the tail probability into a form where one can make use of this large deviation principle, thus arriving at a log-asymptotics for the tail probability. In Section 5 we use a version of the Laplace principle to show that hazard rate twisting, with an appropriate selection of twisting parameter, is asymptotically logarithmically efficient, i.e., the simulation using this change of measure remains efficient as the probability tends to zero. The hazard rate twisting in this case turns out to be equivalent to twisting with the rate function of the large-deviations principle for the random vector. An alternate equivalent formulation of the asymptotics and importance sampling is presented in Section 6. In Section 7 we show how to extend the methodology to the case of input random variables taking values on the real line. Finally, in Section 8 we apply the methodology to one of the examples of Section 3. We should caution the reader that due to space limitations, this paper makes many simplifications and states some key results without proof (the algorithms are explained fully in this paper). For a detailed exposition the reader is referred to Huang and Shahabuddin (2004).

2 THE PROBLEM

Let $X = (X_1, \ldots, X_m)$ be a vector of independent random variables. For simplicity in presentation we will assume that each X_i is a continuous random variable with support $\mathcal{S}_i = (0, \infty)$ or $\mathcal{S}_i = [0, \infty)$ (e.g., gamma or exponential or Pareto). X_i's with support $(-\infty, 0)$ or $(-\infty, 0]$ can easily be accomodated by defining a random variable that is the negative of X_i. Similarly, random variables with support (c, ∞), for some non-zero constant c, can be accomodated by defining another random variable that subtracts c from X_i. Later in Section 8, we show how to extend this study to the case where some or all of the X_i's have support $(-\infty, \infty)$ (e.g., normal). Also, the X_i's may be light-tailed or heavy-tailed to the right (see later for definitions); similarly for left tails, if any.

We consider the problem of estimating $\alpha(y) = P(Y > y)$, where $Y = h(X)$, h is a continuous function taking values in the set of real numbers (thus Y inherits the same regularity properties as those of the X_i's), and y is a large. Let x to denote the vector $(x_1, \ldots, x_m)$. The continuity condition also implies that the only way $h(x) \to \infty$ is by $||x|| \to \infty$. Hence the only way the event $\{Y \geq y\}$ can happen for large y, is by one or more of the X_i's either becoming either very large or very small. We will first derive log-asymptotics for this quantity as $y \to \infty$, and then give a fast simulation method for estimating this quantity that is asymptotically logarithmically efficient (see, e.g., Heidelberger 1995; also defined later).

Let the cumulative distribution function (cdf) of X_i be F_i, and let $\bar{F}_i = 1 - F_i$. Define the hazard function as $\Lambda_{X_i}(x_i) := -\ln(\bar{F}_i(x_i))$. Note that given our assumptions, $\Lambda_{X_i}(x_i)$ is strictly increasing in x_i over the range where the pdf of X_i is positive. For any generic random variable W we will let $\Lambda_W(w)$ denote its hazard function, and let $M_W(\theta)$ denote its moment generating function. A generic random variable W is defined to be light-tailed to the right if $M_W(\theta) < \infty$ for some $\theta > 0$; otherwise it is defined to be heavy-tailed to the right (similar definitions apply for the left tails). For any two functions, say $g_1(w)$ and $g_2(w)$, $g_1 \sim g_2$ denotes that $\lim_{w \to \infty} g_1(w)/g_2(w) = 1$.

3 EXAMPLES

We now present several examples from the literature where specific cases of this problem have been considered.

Example 1 (*Heavy-tailed simulation*): This is the problem of estimating $P(Y > y)$ where $Y = \sum_{i=1}^{m} X_i$ and X_i's belong to a large class of heavy-tailed distributions called subexponential distributions (see, e.g., Embrechts, Kluppelberg, Mikosch 1997, for a rigorous definition). This is the first heavy-tailed simulation problem to be considered in the rare event simulation literature. Different, provably efficient methods for the fast simulation have been suggested in Asmussen and Binswanger (1997), Asmussen, Binswanger and Hojgaard (2000), Juneja and Shahabuddin (2002).

Example 2 (*Light and Heavy-tailed Porfolio Value-at-Risk*): Consider a portfolio that consists of stocks, options and other instruments. Let $S(t) = (S_1(t), \ldots, S_m(t))$ be the vector of risk factors (e.g., stock prices, prices of commodities, exchange rates) on which the instruments are based. Let $V(S(s), s)$ be the the value of the portfolio at time s. Note that in addition to the risk factors, the value of the portfolio also depends on the time s, since the value of some instruments may be time-sensitive (e.g., options with a expiration date). Let t be the current time and let $L = V(S(t), t) - V(S(t + \Delta), t + \Delta)$ be the loss over the future interval $(t, t + \Delta)$. Usually Δ is small (e.g., 1 day or 2 weeks) and hence it is assumed that the constitution of the portfolio is unchanged over this interval. The problem is to estimate $P(L > x)$ for large x. Importance sampling changes of measure for estimating this quantity may also be used for estimating the value-at-risk, i.e., estimating x such the $P(L > x) = p$ for some given p, where p is small (e.g., 0.01 or 0.005); see, e.g., Glasserman, Heidelberger and Shahabuddin 2000.

Since L is usually a complicated function of the $S_i(t)'s$, it is customary to investigate a "quadratic approximation" to L. If $\Delta S = S(t+\Delta) - S(t)$, then a quadratic approximation is of the form

$$L \approx a_0 + a^T \Delta S + \Delta S^T A \Delta S \equiv a_0 + Y,$$

where a_0 is a constant, a is a vector, and A is a matrix. There are various ways of arriving at a quadratic approximation including doing a Taylor series expansion of L in terms of ΔS, or doing a regression on historical data.

Glasserman, Heidelberger and Shahabuddin (2000) suggested the approach of first developing a provably efficient importance sampling change of measure (on the risk factors) for estimating $P(Y > x - a_0)$, for large x, and then using the same change of measure for estimating $P(L > x)$. Since $L \approx Y + a_0$, the importance sampling change of measure is likely to be also efficient for the latter. The problem then is to develop provably efficient changes of measure for estimating $P(Y > y)$ where $y = x - a_0$ is large.

Note that the components of ΔS may be correlated and this makes Y a function of dependent random variables. However ΔS is usually chosen from the family of "elliptic distributions" where one can use Cholesky factorization to express ΔS as a function of independent random variables. Various distributions for ΔS have been considered in the literature in order to model varying degrees of tail heaviness of financial data. For example, the case when ΔS is polynomial tailed is modelled by $\Delta S =^d t(0, \Sigma)$, where $t(0, \Sigma)$ is a particular multivariate t distribution. This multivariate t has the same distribution as $\sqrt{V} N(0, \Sigma)$ where $V = \nu / \chi^2_\nu$ (χ^2_ν is the chisquare random variable with ν degrees of freedeom). Using this fact one may show that

$$Y =^d V \sum_{i=1}^{m} \lambda_i Z_i^2 + \sqrt{V} \sum_{i=1}^{m} b_i Z_i \equiv h(Z, V),$$

where λ_i and b_i are constants. Asymptotics and provably efficient fast simulation techniques for the estimation of $P(Y + a_0 > x)$ were developed in Glasserman, Heidelberer and Shahabuddin (2002).

However, there is no guarantee that a change of measure that is proved to be efficient for the quadratic approximation will also be efficient for case of the actual loss. As we will see later, using the approach in this paper one can attempt to investigate asymptotics and provably efficient fast simulation directly for $P(L > x)$.

Example 3 (*PERT Networks with Light-tailed and Heavy-tailed Acivity Durations*): Let the directed graph $G = (V, A)$ be the network representation of a project where A is the set of directed edges representing activities and V is the set of nodes. The directed graph represents the fact that any activity originating from a node cannot be started until all the activities that feed into the node (if any) have been completed. There is one node called the "project starting node" and another node called the "project completion node".

Let the set of edges (activities) be denoted by $\{1, 2, \ldots, m\}$. Let the duration of activity $i \in A$ be given by a non-negative random variable X_i. The X_i's are assumed independent and may be heavy-tailed or light-tailed. If we let X_i correspond to the length of the directed edge i, then the project completion time T is simply the maximum path length from the project starting node to the project completion node (see, e.g, Adalakha and Kulkarni 1989). Hence T is a function of $(X_1, \ldots, X_m)$. The problem is to estimate $P(T > y)$ for large y. Log-asymptotics and fast simulation for this problem has been considered in Juneja, Karandikar and Shahabuddin (2004).

4 LOG ASYMPTOTICS OF THE PROBABILITY

4.1 Large Deviation Principle for Random Vectors

We first specialize the definition of a family of measures satisfying a large deviations principle (LDP) (see, e.g., Dembo and Zeitouni 1998, Pg 5) to the specific type of random variables that we have.

Definition 4.1 *Consider a family of continuous random variables $X_{i,\epsilon}$ (indexed by ϵ) that have support $\mathcal{S}_i$ that is independent of ϵ. Then $X_{i,\epsilon}$ satisfies a LDP with rate function $I_i(x)$ iff for any $\Gamma_i \subset \mathcal{S}_i$*

$$\lim_{\epsilon\to 0} \epsilon \ln P(X_{i,\epsilon} \in \Gamma_i) = -\inf_{x_i \in \Gamma_i} I_i(x_i).$$

There are certain standard constructions of families of such random variables, starting with "seed" random variables that are light-tailed. For example, if X_i is an exponential random variable with mean $1/\lambda$, then $X_{i,\epsilon} := \epsilon X_i$ satisfies LDP with rate function $I_i(x_i) = \lambda x_i$, for $x_i \geq 0$. Similarly if X_i is a normal random variable with mean 0 and variance σ^2, then $X_{i,\epsilon} = \sqrt{\epsilon} X_i$ satisfies a LDP with rate function $I_i(x_i) = x_i^2 \sigma^2/2$. The lemma below gives a general method for constructing a family of random variables that satisfies the LDP, starting from a "seed" random variable that may be light-tailed or heavy-tailed.

Lemma 4.2 *Let X_i be a continuous random variable with support $[0, \infty)$ (or $(0, \infty)$) and hazard function $\Lambda_{X_i}(x_i)$. Let $\Lambda_i(x_i)$ be any strictly increasing function of x_i such that $\Lambda_i \sim \Lambda_{X_i}$. Then $X_{i,\epsilon} = \Lambda_i^{-1}(\epsilon \Lambda_i(X_i))$ satisfies the LDP with rate function $I_i(x_i) = \Lambda_i(x_i)$. Also, $I_i(x_i)$ is a "good" rate function in the sense that the level sets $\{x_i : I_i(x_i) \leq a\}$ are compact for each a.*

For example, if X_i is Pareto with scale parameter 1 and shape parameter α (i.e., $\bar{F}_{X_i}(x_i) = 1/(1+x_i)^\alpha$ for $x_i \geq 0$) and $\Lambda_{X_i}(x_i) = \alpha \ln(1 + x_i)$), then $\lambda_{X_i}^{-1}(\epsilon \Lambda_{X_i}(x_i)) = (1 + x_i^\epsilon) - 1$. Hence the family of random variables $(1 + X_i^\epsilon) - 1$ satisfies a LDP with $I(x_i) = \Lambda_{X_i}(x_i)$.

We now adapt an existing result (see, e.g., Dembo and Zeitouni 1998, Pg 129) to come up with a LDP for random vectors of a finite number of independent random variables, given LDPs for the individual random variables.

Lemma 4.3 *Let $X_{i,\epsilon}$'s be independent for each ϵ, and assume that for each i, the family of of probablity measures corresponding to $X_{i,\epsilon}$ is "exponentially tight". If $X_{i,\epsilon}$ satisfies LDP with rate function $I_i(x_i)$, then $X_\epsilon = (X_{1,\epsilon}, \ldots, X_{m,\epsilon})$ satisfies a LDP with rate function $I(x) = \sum_{i=1}^m I_i(x_i)$, where $I(x)$ is a good rate function.*

4.2 Log-asymptotics of the Tail Probability

For the random vector X in our problem, define $X_{i,\epsilon}$ as in Lemma 4.2, and let $X_\epsilon = (X_{1,\epsilon}, \ldots, X_{m,\epsilon})$. Exponential tightness (see, e.g., Dembo and Zeitouni (1998), Pg 8, for a definition) corresponding to $X_{i,\epsilon}$ can easily be verified. Then Lemma 4.3 implies that for any $\Gamma \subset \mathcal{S}$, where $\mathcal{S} := \mathcal{S}_1 \times \mathcal{S}_2 \times \ldots \times \mathcal{S}_m$,

$$\lim_{\epsilon\to 0} \epsilon \ln P(X_\epsilon \in \Gamma) = -\inf_{x\in\Gamma} \sum_{i=1}^m \Lambda_i(x_i) \tag{1}$$

Recall that we are interested in investigating the log-asymptotics of $P(h(X)/y > 1)$ as $y \to \infty$. Note that $\{x : h(x)/y > 1\}$ defines a region in $\mathcal{S}$, and thus one hopes to use the formulation given by (1). However, the difficulty in using this formulation is that the quantity, ϵ, with respect to which we are taking the limit forms part of the probablity measure of the random variables. In contrast, in our problem the limit is with respect to a quantity, y, that is *exterior* to the probability measure. We now give a method for reconciling this.

First, we make use of the fact that under fairly general conditions, if the family of sets $\Gamma_\epsilon \subset \mathcal{S}$ is such that $\Gamma_\epsilon \to \Gamma$ for $\Gamma \in \mathcal{S}$ (in a sense that will be made precise later), then a principle similar to the LDP holds for $P(X_\epsilon \in \Gamma_\epsilon)$, i.e.,

$$\lim_{\epsilon\to 0} \epsilon \ln P(X_\epsilon \in \Gamma_\epsilon) = -\inf_{x\in\Gamma} \sum_{i=1}^m \Lambda_i(x_i). \tag{2}$$

Second, we express $h(X)$ as a function of a random-vector X_ϵ that satisfies the LDP, and then link the ϵ in the LDP to the y in our problem. Note that one may express $X_i = g_{i,\epsilon}^{-1}(g_{i,\epsilon}(X_i)) = g_{i,\epsilon}^{-1}(X_{i,\epsilon})$, where $g_{i,\epsilon}(x) := \Lambda_i^{-1}(\epsilon \Lambda_i(x))$. Substituting $X_i = g_{i,\epsilon}^{-1}(X_{i,\epsilon})$ in $h(X_1, \ldots, X_m)$ we get that

$$h(X_1, \ldots, X_m) = h_\epsilon(X_{1,\epsilon}, \ldots, X_{m,\epsilon}),$$

where $h_\epsilon(x_1, \ldots, x_m) = h(g_{1,\epsilon}^{-1}(x_1), \ldots, g_{m,\epsilon}^{-1}(x_m))$.

In order to link the y and the ϵ that appear in $P(h_\epsilon(X_\epsilon)/y \geq 1))$ (that is the same as $P(h(X)/y \geq 1)$), we express $\epsilon = 1/q(y)$ (or equivalently, $y = 1/r(\epsilon)$ where $r(\epsilon) = 1/q^{-1}(1/\epsilon)$) for some appropriately chosen function $q(y)$ that satisfies the following condition.

Condition 4.4

1. *$q(y)$ is increasing in y and $q(y) \to \infty$ as $y \to \infty$.*
2. *$h_0(x) := \lim_{y\to\infty} h_{1/q(y)}(x)/y$ is such that $\{x : h_0(x) > 1\}$ is non-empty and its closure does not include 0 (the $h_0(x)$ may have values ∞ and $-\infty$).*

Define $\Gamma_\epsilon := \{x : r(\epsilon) h_\epsilon(x) > 1\}$ (that is the equivalent to the set $\{x : h_{1/q(y)}(x)/y > 1\}$) and $\Gamma := \{x : h_0(x) > 1\}$. We use the notation $\Gamma_\epsilon \to \Gamma$ (as $\epsilon \to 0$) in the sense suggested by the second part of Condition 4.4. Using (2) we have

$$\lim_{y\to\infty} \frac{1}{q(y)} \ln P(h(X) > y)$$
$$= \lim_{\epsilon\to 0} \epsilon \ln P(r(\epsilon)h_\epsilon(X_\epsilon) > 1) = -\inf_{x\in\Gamma} \sum_{i=1}^{m} \Lambda_i(x_i).$$

4.3 Algorithm for Log-Asymptotics

To summarize the above discussion we present the algorithm for determining the log asymptotics.

1. Find a function $q(y)$ that satisfies Condition 4.4.
2. Solve the optimization problem $I_{opt} = \inf_{x\in\Gamma} \sum_{i=1}^{m} \Lambda_i(x_i)$, where $\Gamma := \{x : h_0(x) > 1\}$

Then $\lim_{y\to\infty} \ln P(h(X) > y)/q(y) = -I_{opt}$.

This asymptotic gives useful information in the sense that $P(h(X) > y) = e^{-I_{opt}q(y)(1+o(1))}$. Also, as we will see later, the solution to the optimization problem gives useful information about the manner in which the rare-event occurs in the system. Note that the crucial part in the above procedure is identifying a $q(y)$ that satisfies Condition 4.4.

5 FAST SIMULATION

5.1 Preliminaries: Importance Sampling and Hazard Rate Twisting

For y large, the event $\{Y > y\}$ may be rare, and we use importance sampling to simulate for $P(Y > y)$ more efficiently. In particular, if $\tilde{f}_i(x)$ is a new probability density function for X_i, with the same support as X_i, then we may express

$$P(Y > y) = E(I(Y > y)) = \tilde{E}(I(Y > y)l(X)) \quad (3)$$

where $l(x) = \prod_{i=1}^{m}(f_i(x_i)/\tilde{f}_i(x_i))$, and the $\tilde{E}(\cdot)$ indicates that X_i's have the (new) pdfs $\tilde{f}_i$'s. The quantity within the expectation on the RHS (right-hand side) of (3) forms an unbiased, "importance sampling" estimator of $P(Y > y)$.

The attempt is to find $\tilde{f}_i$'s so that the variance of this new estimator is as low as possible. More specifically, we want to $\tilde{E}(I(Y > y)l^2(X))$ to be the least possible. The change of measure $(\tilde{f}_1, \ldots, \tilde{f}_n)$ is called "asymptotically logarithmically efficient" or "ALE" iff

$$\liminf_{y\to\infty} \frac{\ln \tilde{E}(I(Y > y)l^2(X))}{2\ln P(Y > y)} \geq 1. \quad (4)$$

This means that the exponential rate of decrease of the second moment is twice the exponential rate of decrease of the probability one is trying to estimate. Non-negativity of the variance implies that this is the fastest possible rate for any unbiased estimator. This is the reason why ALE is also referred to as "asymptotic optimality". Note that for standard simulation, $\ln \tilde{E}(I(Y > y)l(X)))/(2\ln P(Y > y)) \sim 1/2$.

We will use the change of measure called hazard rate twisting that was introduced in Juneja and Shahabuddin (2002). Again let $\Lambda_i(x) \sim \Lambda_{X_i}(x)$. Then the hazard rate twisted density with amount θ, $0 < \theta < 1$, is given by

$$f_{X_i,\theta}(x) = \frac{f_{X_i}(x)e^{\theta\Lambda_i(x)}}{M_{\Lambda_i(X_i)}(\theta)}. \quad (5)$$

Here $M_{\Lambda_i(X_i)}(\theta) \equiv \int_0^\infty f_{X_i}(x)e^{\theta\Lambda_i(x)}dx$ is the normalization constant that is needed to make $f_{X_i,\theta}(x)$ a pdf. Note that unlike exponential twisting that may not be defined for heavy-tailed random variables (since the moment generating function may be infinite for all $\theta > 0$), hazard rate twisting is defined for all $0 < \theta < 1$. This is because $\Lambda_i(X_i)$ is random variable that has an exponential tail with rate 1 (see Huang and Shahabuddin 2003) and hence the normalization constant $M_{\Lambda_i(X_i)}(\theta)$ is defined for all $0 < \theta < 1$.

Note that if we use $\Lambda_i(x) = \Lambda_{X_i}(x)$ in (5) then $M_{\Lambda_i(X_i)}(\theta) = 1/(1-\theta)$. This is because $\Lambda_{X_i}(X_i)$ is an exponential random variable with rate 1 (see Huang and Shahabuddin 2003). Originally, this was called hazard rate twisting in Juneja and Shahabuddin (2002), and using a general $\Lambda_i \sim \Lambda_{X_i}$ was called "asymptotic" hazard rate twisting. In the interest of brevity we have used "hazard rate twisting" to denote both of them.

When we apply hazard rate twisting to each X_i, the likelihood ratio is

$$l(X) = \prod_{i=1}^{m} \frac{f_{X_i}(x)}{f_{X_i,\theta}(x)} = (\prod_{i=1}^{m} M_{\Lambda_i(X_i)}(\theta))e^{-\theta\sum_{i=1}^{m}\Lambda_i(X_i)}.$$

Hence the second moment of the (single sample) importance sampling estimator is

$$\tilde{E}(I(h(X) > y)l^2(X)) = E(I(h(X) > y)l(X))$$
$$= (\prod_{i=1}^{m} M_{\Lambda_i(X_i)}(\theta))E\left(I(h(X) > y)e^{-\theta\sum_{i=1}^{m}\Lambda_i(X_i)}\right). \quad (6)$$

To investigate ALE properties, we now develop a log-asymptotics for this second moment as $y \to \infty$.

5.2 Log-asymptotics for the Second Moment

To enable the above we use a modified version of Varadhan's Integral Lemma (see, e.g., Dembo and Zeitouni 1998, Pg 137) that was proved in Glasserman, Heidelberger and

Shahabuddin (1999). We first adapt this lemma to our simplified setting.

Lemma 5.1 *Let X_ϵ be a family of continuous random vectors with support $\mathcal{S}$ for all ϵ. Let $\phi(x)$ be a continuous, negative-valued function with domain $\mathcal{S}$ and let X_ϵ satisfy the LDP with good rate function $I(x)$. Then for all $\Gamma \subset \mathcal{S}$,*

$$\lim_{\epsilon\to 0} \epsilon \ln E\left(e^{\frac{\phi(X_\epsilon)}{\epsilon}} I(X_\epsilon \in \Gamma)\right) = \sup_{x\in\Gamma}(\phi(x) - I(x)). \quad (7)$$

Similar to (2), under some fairly general conditions, this Lemma extends to the case where we replace Γ by Γ_ϵ in the left hand side of (7), where Γ_ϵ is such that $\Gamma_\epsilon \to \Gamma$ in the sense described previously.

To apply this lemma we first need to express the last term of (6) in a suitable form. In particular, defining the sets Γ and Γ_ϵ as before, and using the transformations defined earlier we have that $\tilde{E}(I(h(X) > y)l^2(X))$ may be expressed as

$$[\prod_{i=1}^{m} M_{\Lambda_i(X_i)}(\theta)] E(I(X_\epsilon \in \Gamma_\epsilon) e^{-\frac{\theta}{\epsilon}\sum_{i=1}^m \Lambda_i(X_{i,\epsilon})}).$$

Note that we made use of the convenient fact that

$$\begin{aligned}\sum_{i=1}^m \Lambda_i(X_i) &= \sum_{i=1}^m \Lambda_i(g_{i,\epsilon}^{-1}(g_{i,\epsilon}(X_i))) \\ &= \sum_{i=1}^m \Lambda_i(g_{i,\epsilon}^{-1}(X_{i,\epsilon})) = \sum_{i=1}^m \Lambda_i(X_{i,\epsilon})/\epsilon,\end{aligned}$$

since $g_{i,\epsilon}^{-1}(x) = \Lambda_i^{-1}(\Lambda_i(x)/\epsilon)$.

In our case, we have already shown that X_ϵ, where $X_{i,\epsilon} = g_{i,\epsilon}(X_i)$, satisfies the LDP with good rate function $I(x) = \sum_{i=1}^m \Lambda_i(x_i)$. First, lets try using a θ, $0 < \theta < 1$, that is independent of ϵ (or equivalently, independent of y). Applying the modification of Lemma 5.1 mentioned before (i.e., with Γ replaced by Γ_ϵ on the LHS of (7)), we have that

$$\epsilon \ln\left([\prod_{i=1}^m M_{\Lambda_i(X_i)}(\theta)] E(I(X_\epsilon \in \Gamma_\epsilon) e^{-\frac{\theta}{\epsilon}\sum_{i=1}^m \Lambda_i(X_{i,\epsilon})})\right) \to -(1+\theta)\inf_{x\in\Gamma}\sum_{i=1}^m \Lambda_i(x_i),$$

as $\epsilon \to 0$. Or equivalently,

$$\lim_{y\to\infty} \frac{1}{q(y)} \ln\left(\tilde{E}(I(h(X) > y)l^2(X))\right) = -(1+\theta)\inf_{x\in\Gamma}\sum_{i=1}^m \Lambda_i(x_i).$$

Using (3) and (8), we see that the limit in (4) is $(1+\theta)/2$ which is slightly less than desired since $0 < \theta < 1$.

To achieve ALE one may choose $\theta = \theta_\epsilon = 1 - b\epsilon$, where b is some positive constant, so that $\theta_\epsilon \to 1$ as $\epsilon \to 0$. Now that θ varies with ϵ, we will need the following assumption:

Assumption 5.2 $\lim_{\epsilon\to 0} \epsilon \ln[M_{\Lambda_i(X_i)}(\theta_\epsilon)] = 0$.

This assumption is true quite generally; sufficient conditions for this to hold may be found in Juneja, Karandikar and Shahabuddin (2004). It can be checked that it holds for the case $\Lambda_i(\cdot) = \Lambda_{X_i}(\cdot)$ for which (as mentioned before) $M_{\Lambda_i(X_i)}(\theta_\epsilon) = 1/(1-\theta_\epsilon) = 1/(b\epsilon)$.

Note that $E(I(X_\epsilon \in \Gamma_\epsilon) e^{-\frac{\theta}{\epsilon}\sum_{i=1}^m \Lambda_i(X_{i,\epsilon})})$ is bounded below by $E(I(X_\epsilon \in \Gamma_\epsilon) e^{-\frac{1}{\epsilon}\sum_{i=1}^m \Lambda_i(X_{i,\epsilon})})$, whose logarithmic limit is $-\inf_{x\in\Gamma} 2\sum_{i=1}^m \Lambda_i(x_i)$. Also for any $\delta > 0$, there exists ϵ_0, such that for all $\epsilon \geq \epsilon_0$, $E(I(X_\epsilon \in \Gamma_\epsilon) e^{-\frac{\theta}{\epsilon}\sum_{i=1}^m \Lambda_i(X_{i,\epsilon})})$ is bounded above by $E(I(X_\epsilon \in \Gamma_\epsilon) e^{-\frac{(1-\delta)}{\epsilon}\sum_{i=1}^m \Lambda_i(X_{i,\epsilon})})$. Note that the logarithmic limit for the latter quantity is $-(2-\delta)\inf_{x\in\Gamma}\sum_{i=1}^m \Lambda_i(x_i)$. Since this is true for all $\delta > 0$, using Assumption 5.2 we obtain

$$\frac{1}{q(y)} \ln\left(\tilde{E}(I(h(X) > y)l^2(X))\right) \to -2\inf_{x\in\Gamma}\sum_{i=1}^m \Lambda_i(x_i),$$

as $y \to \infty$. This gives ALE.

5.3 Importance Sampling Algorithm

Preprocessing: Determine the function $q(\cdot)$ satisfying Condition 4.4. Set $\theta = 1 - b/q(y)$ where b is some positive constant.

Sampling:

- For $i = 1, \ldots, m$, generate X_i from $f_{X_i,\theta}$, as defined by (5).
- Compute the output sample

$$\left(\prod_{i=1}^m M_{\Lambda_i(X_i)}(\theta)\right) I(h(X) > y) e^{-\theta\sum_{i=1}^m \Lambda_i(X_i)}. \quad (8)$$

The complete algorithm will involve generating n output samples and computing the sample mean.

6 AN ALTERNATE FORMULATION

As mentioned in Huang and Shahabuddin (2003) (see also Kroese and Rubinstein 2004), hazard rate twisting on X_i by amount θ, as given by (5), is equivalent to "exponentially twisting" $V_i := \Lambda_i(X_i)$ by amount θ, $0 < \theta < 1$. Exponentially twisting V_i by amount θ means that the new pdf of V_i is given by

$$f_{V_i,\theta}(v_i) = \frac{f_{V_i}(v_i)e^{\theta v_i}}{M_{V_i}(\theta)}.$$

Note that $\Lambda_i(X_i)$ is exponentially tailed with rate 1 (see Huang and Shahabuddin 2003), and hence exponential twisting by amount θ, $0 < \theta < 1$, is permissible.

Hence another equivalent approach for estimating $P(h(X_1, \ldots, X_m) > y)$ would be to first express $P(h(X_1, \ldots, X_m) > y)$ as $P(h(\Lambda_1^{-1}(V_1), \ldots, \Lambda_m^{-1}(V_m)) > y)$ and then use exponential twisting on the V_i's by amount θ. Indeed, the output sample obtained using this method has exactly the same distribution as the quantity in (8). This approach has also been suggested as one of the approaches in Kroese and Rubinstein (2004). However, as mentioned before, they used an adaptive approach for the determination of the best θ and no log-asymptotics and simulation efficiency results were provided.

One could also use this alternate formulation (i.e, of expresssing the probability of interest as $P(h(\Lambda_1^{-1}(V_1), \ldots, \Lambda_m^{-1}(V_m)) > y)$) to derive the log-asymptotics, making use of the fact that ϵV_i satisfies a LDP with rate function $I(v_i) = v_i$. First one will need to derive the analogue of the function $q(y)$ (above), that turns out to be the same as $q(y)$. Next, one would need to solve the optimization problem $I_{opt} = \inf_{\{v: h_0(\Lambda_1^{-1}(v_1),\ldots,\Lambda_m^{-1}(v_m))>1\}} \sum_{i=1}^m v_i$. By making the monotonic substitution $v_i = \Lambda_i(x_i)$, we see that this problem is equivalent to the optimization problem in Section 4.3. Hence the two approaches not only yield the same asymptotics (as was to be expected) but are also algorithmically equivalent.

7 RANDOM VARIABLES ON $(-\infty, \infty)$

We now give a brief description of the case when some or all of the X_i's have support $(-\infty, \infty)$. We had omitted this case earlier for the sake of simplicity.

7.1 LDP and Asymptotics

Consider a X_i for which the support is $(-\infty, \infty)$. In this case it is easy to see that $\Lambda_i^{-1}(\epsilon\Lambda_i(X_i))$ need not necessarily satisfy a LDP (e.g., use the double exponential distribution and use the set $(-\infty, a]$ for some constant $a < 0$).

Without loss of generality, we will assume that $F_i(0) = 0.5$ (otherwise we can simply translate the X_i's by a constant amount to make sure that this holds). In addition to $\Lambda_{X_i}(x_i)$, define $\bar{\Lambda}_{X_i}(x_i) = -\ln F(x_i)$. Note that $\bar{\Lambda}_{X_i}(x_i)$ is always positive, is strictly decreasing in x_i, $\bar{\Lambda}_{X_i}(0) = \Lambda_{X_i}(0)$, and $\bar{\Lambda}_{X_i}(x_i) \to \infty$ as $x_i \to -\infty$. Similar to before, let Λ_i be a continuous, strictly increasing function over $[0, \infty)$, such $\Lambda_i(x_i) \sim \Lambda_{X_i}(x_i)$. Let $\bar{\Lambda}_i$ be a continuous, strictly decreasing function over $(-\infty, 0)$ such that $\bar{\Lambda}_i(x_i)/\bar{\Lambda}_{X_i}(x_i) \to 1$ as $x_i \to -\infty$, and $\bar{\Lambda}_i(0) = \Lambda_i(0)$. Let

$$\begin{aligned} \tilde{\Lambda}_i(x_i) &= \Lambda_i(x_i) \quad \text{for} \quad x_i \geq 0 \\ & \bar{\Lambda}_i(x_i) \quad \text{for} \quad x_i < 0. \end{aligned}$$

Note that $\tilde{\Lambda}_i(x_i)$ is a continuous function, and it tends to infinity as $|x_i| \to \infty$.

As before, define $g_{i,\epsilon}(x_i) = \Lambda_i^{-1}(\epsilon\Lambda_i(x_i))$ for $x_i \geq 0$, and define $\bar{g}_{i,\epsilon}(x_i) := \bar{\Lambda}_i^{-1}(\epsilon\bar{\Lambda}_i(x_i))$ for $x_i < 0$. Also, define

$$\begin{aligned} \tilde{g}_{i,\epsilon}(x_i) &= g_{i,\epsilon}(x_i) \quad \text{for} \quad x_i \geq 0 \\ & \bar{g}_{i,\epsilon}(x_i) \quad \text{for} \quad x_i < 0. \end{aligned}$$

Then $X_{i,\epsilon} := \tilde{g}_{i,\epsilon}(X_i)$ satisfies a LDP with rate function $\tilde{\Lambda}_i(x_i)$.

For example, take the case of $X_i =^d N(0,1)$. It is well known the $\Lambda_{X_i}(x_i) \sim x_i^2/2$, and hence we may use $\Lambda_i(x_i) = x_i^2/2$ for $x_i > 0$. By symmetry, $\lim_{x_i \to -\infty} \bar{\Lambda}_{X_i}(x_i)/(x_i^2/2) = 1$ and hence we may use $\bar{\Lambda}_i(x_i) = x_i^2/2$ for $x_i < 0$. One can then work out that $X_{i,\epsilon} = \sqrt{\epsilon}X_i$.

In this case the transformation $\tilde{g}_{i,\epsilon}(\cdot)$ is no longer monotonic, and hence $\tilde{g}_{i,\epsilon}^{-1}(\cdot)$ for use in the h function is no longer defined. However, since $X_{i,\epsilon} > 0$ if and only if $X_i > 0$, we can express

$$\begin{aligned} X_i &= g_{i,\epsilon}^{-1}(\tilde{g}_{i,\epsilon}(X_i))I(X_i \geq 0) + \bar{g}_{i,\epsilon}^{-1}(\tilde{g}_{i,\epsilon}(X_i))I(X_i < 0) \\ &= g_{i,\epsilon}^{-1}(X_{i,\epsilon})I(X_{i,\epsilon} \geq 0) + \bar{g}_{i,\epsilon}^{-1}(X_{i,\epsilon})I(X_{i,\epsilon} < 0) \end{aligned}$$

We then follow the same procedure as before.

7.2 Double-Tailed Hazard Rate Twisting

The simplest way to extend hazard rate twisting to $(-\infty, \infty)$ would be to allow the x in (5) to lie in $(-\infty, \infty)$. However this increases only the weight of the positive tail (or only the weight of the negative tail if one transforms X_i to $-X_i$ before commencing). But in several cases, the way $h(x)$ becomes large is by a x_i becoming very large or very small (e.g., $h(x_1) = x_1^2$). To be sure that we are not missing out any important region, we need to increase the weight of both the tails of X_i, if X_i has support $(-\infty, \infty)$. The relevant extension to hazard rate twisting is suggested directly from the previous sub-section, i.e., use $\tilde{\Lambda}_i(x)$ instead of $\Lambda_i(x)$ in (5). This simple modification again yields ALE.

8 ILLUSTRATION OF ALGORITHMS

We now apply the algorithms we developed above to some of the financial engineering problems described in Example 2 of Section 3.

Example 2 (Contd.): *(a) Quadratic Approximation.* First we consider asymptotics and fast simulation for $P(Y > y)$ for the case when $\Delta S = t(0, \Sigma)$, and the portfolio is delta-hedged (i.e., the b_i's are zero). Since $V = \frac{\nu}{\chi_\nu^2}$, one can easily check that $\bar{F}_V(v) \sim c/v^{\nu/2}$. Hence $\Lambda_V(v) \sim -\ln(c/v^{\nu/2})$ (note that $g_1(w) \sim g_2(w)$ does *not* imply $g(g_1(w)) \sim g(g_2(w))$ for any general functions g_1, g_2 and g; however the result holds in this case). Hence we may use $\Lambda_{m+1}(v) = (\nu/2)\ln(1+v)$ (we use the notation $\Lambda_{m+1}(v)$, since V is the $(m+1)st$ input to h). Also, $\sqrt{\epsilon}Z_i$ satisfies a LDP with rate function $z_i^2/2$. Hence if we let $z := (z_1, \dots z_m)$, then

$$h_\epsilon(z, v) = [(1+v)^{1/\epsilon} - 1]\sum_{i=1}^{m} \frac{\lambda_i z_i^2}{2\epsilon},$$

or equivalently

$$\frac{1}{y} h_{1/q(y)}(z, v) = \frac{[e^{q(y)\ln(1+v)} - 1]}{y} q(y) \sum_{i=1}^{m} \frac{\lambda_i z_i^2}{2}.$$

For any fixed (z, v), since $[e^{q(y)\ln(1+v)} - 1]q(y)/y$ is the term that varies with y, we select $q(y)$ so that this term converges to different values when $\ln(1+v) > 0$ and $\ln(1+v) < 0$. This suggests using $q(y) = \ln y$. When one does this, one finds that

$$\begin{aligned} h_0(z, v) &= \infty \quad \text{if} \quad v > e - 1 \\ & 0 \quad \text{if} \quad v \le e - 1 \end{aligned}$$

Hence $\Gamma = \{(z, v) : v > e - 1\}$. Carrying out the optimization, we get that $I_{opt} = \nu/2$. In this case the unique optimal solution is $v^* = e - 1$, and $z_i^* = 0$ for each i. This conveys that the most likely way the event $\{h(Z_1, \dots, Z_m, V) > y)\}$ happens for large y is by V becoming large and Z_i's being in their usual range.

The asymptotic is consistent with results in Glasserman, Heidelberger and Shahabuddin (2002). However, the change of measure is different since the latter uses more system specific information.

(b)*Actual Loss Function.* Next we consider asymptotics and fast simulation for $P(L > y)$, for portfolios consisting of stocks, and standard European calls and puts. For simplicity, we only consider a two instruments, two risk factors case, where the risk factors are dependent; one can then see that the basic methodology could easily be extended to a general portfolio consisting of shares, calls and puts. We assume the changes in risk factors, $\Delta S = (\Delta S_1, \Delta S_2)$, to be normally distributed (as assumed, for example, in Glasserman, Heidelberger and Shahabuddin 2000) with means 0, variances σ_1^2 and σ_2^2, respectively, and correlation ρ. In this case one can express $\Delta S_1 = \sigma_1 Z_1$ and $\Delta S_2 = \sigma_2(\rho Z_1 + \sqrt{1-\rho^2} Z_2)$, where Z_1 and Z_2 are independent $N(0, 1)$'s. Assume that the portfolio consists of shorting one standard call on each of the risk factors. Then one can easily derive the expression for L as some function $h(Z_1, Z_2)$ using the Black-Scholes formula (with some rounding off to take care of the negative values of the risk factors which may occur with small probability, since we are using normal instead of lognormal distribution). Since $\sqrt{\epsilon}Z_i$ satisfies a LDP with rate function $z_i^2/2$, we have that $h_\epsilon(z_1, z_2) = h(z_1/\sqrt{\epsilon}, z_2/\sqrt{\epsilon})$. One can then easily see that $\lim_{\epsilon\to 0}\sqrt{\epsilon}h_\epsilon(z_1, z_2) = h_0(z_1, z_2)$, where

$$\begin{aligned} h_0(z_1, z_2) &= \sigma_1 z_1 I(z_1 > 0) \\ &\quad + \sigma_2(\rho z_1 + \sqrt{1-\rho^2} z_2) I(\rho z_1 + \sqrt{1-\rho^2} z_2 > 0). \end{aligned}$$

Hence one may use $y = 1/\sqrt{\epsilon}$ or equivalently, $\epsilon = 1/y^2$. Minimizing $z_1^2/2 + z_2^2/2$ over $h_0(z_1, z_2) \ge 1$ we get that $z_1^* = (\sigma_1 + \sigma_2\rho)/A$ and $z_2^* = \sigma_2\sqrt{1-\rho^2}/A$, where $A = (\sigma_1 + \sigma_2\rho)^2 + (\sigma_2\sqrt{1-\rho^2})^2$. The optimal value $I_{opt} = 1/A$. Hence

$$\lim_{y\to\infty} \frac{1}{y^2} \ln P(L > y) = 1/A.$$

For a general portfolio, even though one may easily be able to formulate the optimization problem, it may be hard to solve it. However, note that for the fast simulation we need *not* solve the optimization problem, since all we need is the $q(y)$ for use in the θ. As long as we have identified such a $q(y)$ (in this case $q(y) = y^2$), it should be enough for the fast simulation.

ACKNOWLEDGMENTS

This work was partially supported by the National Science Foundation (U.S.A.) Grant DMI 03-00044. The work of the first author was performed while she was at Columbia University.

REFERENCES

Adalakha, V. G., and V. G. Kulkarni. 1989. A classified bibliography of research on stochastic PERT networks: 1966-1987. *INFOR* 27(3) 272-296.

Asmussen, S., and K. Binswanger. 1997. Simulation of ruin probabilities for subexponential claims. *ASTIN Bulletin* 27(2): 297-318.

Asmussen, S., K. Binswanger and B. Hojgaard. 2000. Rare event simulation for heavy-tailed distributions. *Bernoulli* 6(2): 303-322.

Boots, N.K., and P. Shahabuddin. 2000. Simulating GI/GI/1 queues and insurance risk processes with subexponential distributions. In *Proceedings of the 2000 Winter Simulation Conference*, ed. J.A. Joines, R.R. Barton, K. Kang, and P.A. Fishwick, 656-665, IEEE Press (extended version available).

Bucklew, J.A. 1990. *Large Deviations Techniques in Decision, Simulation, and Estimation*. John Wiley & Sons, Inc.

Dembo, A., and O. Zeitouni. 1998. *Large Deviations Techniques and Applications.* Second Edition, Springer, New York.

Embrechts, P., C. Klüppelberg and T. Mikosch. 1997. *Modelling Extremal Events*. Springer-Verlag, Heidelberg.

Glasserman, P., P. Heidelberger, and P. Shahabuddin. 1999. Asymptotically optimal importance sampling and stratification for pricing path-dependent options. *Mathematical Finance* 9(2): 117-152.

Glasserman, P., P. Heidelberger, and P. Shahabuddin. 2000. Variance reduction techniques for estimating value-at-risk. *Management Science* 46(10): 1349-1364.

Glasserman, P., P. Heidelberger, and P. Shahabuddin. 2002. Portfolio value-at-risk with heavy-tailed risk factors. *Mathematical Finance* 12(3): 239-269.

Heidelberger, P. 1995. Fast simulation of rare events in queueing and reliability models. *ACM Transactions on Modeling and Computer Simulation* 6: 43-85.

Huang, Z., and P. Shahabuddin. 2003. Rare-event, heavy-tailed simulations using hazard function transformations, with applications to value-at-risk. In *Proceedings of the 2003 Winter Simulation Conference*, ed. S. Chick, P.J. Sanchez, D. Ferrin, and D.J. Morrice, 276-284, IEEE Press (extended version available).

Huang, Z., and P. Shahabuddin. 2004. A unified approach for finite-dimensional, rare-event asymptotics and simulation. Technical Report, Dept. of IEOR, Columbia University, NY 10027, U.S.A..

Juneja, S., and P. Shahabuddin. 2002. Simulating heavy-tailed processes using delayed hazard rate twisting. *ACM Transactions on Modeling and Computer Simulation* 12: 94-118.

Juneja, S., and P. Shahabuddin. 2004. Rare-Event Simulation Techniques: An Introduction and Recent Advances. Technical Report, Dept. of IEOR, Columbia University, NY 10027, U.S.A..

Juneja, S., Karandikar, R.L., and P. Shahabuddin. 2004 Tail asymptotes and fast simulation of delay probabilities in stochastic PERT networks. Technical Report, Dept. of IEOR, Columbia University, NY 10027, U.S.A..

Kroese, D., and R.Y. Rubinstein. 2004. The transform likelihood ratio method for rare-event simulation with heavy-tails. *Queueing Systems* 46(3-4): 317-351.

AUTHOR BIOGRAPHIES

ZHI HUANG works at Lehman Brothers. She received her B.S. in Computer Science (2000) from University of Science and Technology in China, and a M.S (2001) and Ph.D. (2004) in Operations Research from Columbia University. Her research interests include financial modeling using analytical techniques and simulation. Her e-mail address is <zh52@columbia.edu>.

PERWEZ SHAHABUDDIN is a professor in the Industrial Engineering and Operations Research Department at Columbia University. From 1990 to 1995, he was a Research Staff Member at IBM T.J. Watson Research Center. He received a B.Tech. in Mechanical Eng. (1984) from the Indian Institute of Technology, Delhi, India, and a M.S. in Statistics (1987) and a Ph.D. in Operations Research (1990) from Stanford University. He serves as the Department Editor for Stochastic Models and Simulation at *Management Science*. His e-mail address is <perwez@ieor.columbia.edu>.