

COMPUTATION OF MULTIVARIATE BARRIER CROSSING PROBABILITY AND ITS APPLICATIONS IN CREDIT RISK MODELS

Joonghee Huh* and Adam Kolkiewicz†

ABSTRACT

In this paper we consider computational methods of finding exit probabilities for a class of multivariate diffusion processes. Although there is an abundance of results for one-dimensional diffusion processes, for multivariate processes one has to rely on approximations or simulation methods. We adopt a Large Deviations approach to approximate barrier crossing probabilities of a multivariate Brownian Bridge. We use this approach in conjunction with simulation methods to develop an efficient method of obtaining barrier crossing probabilities of a multivariate Brownian motion. Using numerical examples, we demonstrate that our method works better than other existing methods. We mainly focus on a three-dimensional process, but our framework can be extended to higher dimensions. We present two applications of the proposed method in credit risk modeling. First, we show that we can efficiently estimate the default probabilities of several correlated credit risky entities. Second, we use this method to efficiently price a credit default swap (CDS) with several correlated reference entities. In a conventional approach one normally adopts an arbitrary copula to capture dependency among counterparties. The method we propose allows us to incorporate the instantaneous variance-covariance structure of the underlying process into the CDS prices.

1. INTRODUCTION AND MOTIVATION

In this paper we develop an efficient computational technique of approximating exit probabilities for a class of multivariate correlated diffusion processes. Our objective is to construct a practical method to deal with dependent default events in the context of structural credit risk models.

The presented method is based on the assumption that the underlying process follows a multivariate correlated Brownian motion process. Currently no analytical results for exit probabilities for this process exist, and typically we have to resort to simulation methods that use some form of a discrete-time approximation of the process. Suppose that τ denotes the first time a process $\{\mathbf{X}(t) \in \mathbb{R}^n, t \in [0, T]\}$ crosses a prespecified, possibly time-dependent, deterministic barrier $\mathbf{b}(t) \in \mathbb{R}^n, t \in [0, T]$, where $\mathbf{X}(t) = [X_1(t), \dots, X_n(t)]'$, $\mathbf{b}(t) = [b_1(t), \dots, b_n(t)]'$, and \mathbf{A}' denotes the transpose of a vector or matrix \mathbf{A} . We assume $X_i(0) > b_i(0)$, for all $i = 1, \dots, n$. We can represent this exit time as $\tau = \min\{\tau_1, \dots, \tau_n\}$, where $\tau_i = \inf\{t \in [0, T]: X_i(t) \leq b_i(t)\}$, $i = 1, \dots, n$. In a naive simulation approach, τ can be approximated by the exit time of a discretized process, $\hat{\tau} = \min\{\hat{\tau}_1, \dots, \hat{\tau}_n\}$, where $\hat{\tau}_i = \min\{k\varepsilon : X_i(k\varepsilon) \leq b_i(k\varepsilon), k = 1, 2, \dots, M\}$ and $\varepsilon = T/M$. Thus crossings can occur only at a set of M discrete time points. However, in this method the estimate $\hat{\tau}$ will be biased in the sense $E[\hat{\tau}] > \tau$, which follows from the observation that for all trajectories of the process we have $\hat{\tau} \geq \tau$,

* Joonghee Huh, PhD, FRM, is an investment Vice President at Prudential Financial, 100 Mulberry St. Gateway Center 2, 5th floor, Newark, NJ 07102-5096, joonghee.huh@prudential.com.

† Adam Kolkiewicz, PhD, is an Associate Professor in the Department of Statistics and Actuarial Science, University of Waterloo, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada, wakolkie@uwaterloo.ca.

and the equality occurs only with probability zero. Formally the inequality $\hat{\tau} \geq \tau$ is a consequence of the fact that in the definition of τ the infimum is taken over a larger set of possible exit times than in the definition of $\hat{\tau}$. A more intuitive explanation can be based on the observation that even if we know that the values $X(j\varepsilon)$ and $X((j+1)\varepsilon)$ are above the barrier for some $j \in \{1, \dots, M-1\}$, there is still the possibility that the process may breach the barrier at some time between $j\varepsilon$ and $(j+1)\varepsilon$. Ignoring this possibility leads to a slow convergence of $\hat{\tau}$ to τ as the number of time steps increases to infinity.¹ In practice, this is manifested by a non-negligible approximation error for finite values of M . A formal investigation of the rate of convergence of the discretization error for multidimensional diffusion processes is carried out by Gobet (2000). A comprehensive exposition of numerical methods for stochastic differential equations is provided by Kloeden and Platen (1992).

To refine a discrete-time approximation of the exit probability, we need to determine the probability that the process $\{X(t)\}$ crosses the barrier between discrete simulation times. This problem involves computation of the exit probability of a Brownian Bridge process, which describes the dynamics of a Brownian motion when its end values are known. Currently analytical representations of the exit probability for a Brownian Bridge exist only if the dimension of the process does not exceed two. For higher dimensions we have to rely on approximations or simulations. In this context asymptotic results based on the theory of Large Deviations have proven to be particularly useful (Baldi 1995; Baldi, Caramellino, and Iovino 1999; Baldi and Caramellino 2000). The existing results, however, are not sufficient for credit risk models, because they usually deal with a standard Brownian motion. In the paper we demonstrate that a direct application of the result by Baldi to a general multivariate Brownian motion does not capture the covariance structure of the process. This finding was the motivation for the new method that we present in the paper. Our approach still uses the Large Deviations Theory, but it captures the correlation structure of the process. The Large Deviations Theory has many applications in finance and actuarial science, and Boyle, Feng, and Tian (2008) and Pham (2007) provide excellent overviews. Recent work by Boyle et al. (2008) uses the Large Deviations Theory in perturbation of parameters to price options in incomplete markets.

In conjunction with the proposed method for a multivariate Brownian Bridge, we suggest a time-efficient simulation algorithm for computing exit probabilities of a multivariate Brownian motion process. In typical applications in finance, the method requires only a small number of subintervals, and the algorithm allows us to compute the probability that in a portfolio of defaultable instruments at least one obligor defaults. It can also be extended to deal with the problem of finding the distribution of the number of defaults in a portfolio. Such information is required, for example, to price credit instruments such as first n to default. In view of the scope of possible applications and its efficiency, the proposed method compares quite favorably with other approximations, such as the one proposed by Shevchenko (2003). We provide more information about the latter one in Section 3.3. The developed method provides also the basis for extensions to other models, such as general diffusion processes or diffusion processes with jumps. We make additional comments about this in Section 5.

The presented technique for computing exit probabilities can be a useful tool in credit risk modeling, because often the default event is represented as an exit time of a certain underlying process. The recent credit crunch in financial markets demonstrates the critical importance of proper credit risk quantification and management. In light of large exposures to the credit market, credit risk is also significant to insurance companies, which has recently been discussed by Li (2006). The proposed method is also applicable to pricing of other financial instruments, such as portfolios of barrier options or equity default swaps. The latter are structured similarly to credit default swaps except that payouts occur when an equity default event happens, which occurs whenever a given share price drops below a fixed fraction of the spot value at the inception of the contract (see, e.g., Medova and Smith 2006).

In Section 2 we present two methods of computing the exit probability of a multivariate correlated Brownian Bridge process. Both approaches use the Large Deviations Theory. Although the first method

¹ For a formal explanation of the sense $\hat{\tau}$ converges to τ , see, for example, Billingsley (1999).

can be considered an extension of the results presented by Baldi (1995), the second uses the Large Deviations Theory in a novel way. In Section 3 we present an algorithm to calculate the exit probability of a multivariate Brownian motion process. There we show how the results from Section 2 can be incorporated into a simulation framework. We also discuss other methods for approximating exit probabilities. In Section 4 we present two applications of the method to credit risk models. In particular, we compute the probability of at least one firm defaulting and perform the valuation of a credit default swap in which a payment is triggered by the first default event of several underlying entities. Finally, Section 5 concludes with suggestions for further research. Technical results are presented in appendices.

2. EXIT PROBABILITY FOR BROWNIAN BRIDGE

Suppose that $\{X(t), t \in [t_0, t_1]\}$ is a one-dimensional Brownian motion process with zero drift and a constant diffusion coefficient σ . Assuming that the end values of the process are known, this conditioned process is known as a Brownian Bridge. Let $p_b(x, y)$ denote the probability that the conditioned process breaches a constant barrier b between times t_0 and t_1 , given that $X(t_0) = x$ and $X(t_1) = y$. It is well known that this probability has an analytical representation of the following form (see, e.g., Section 4.3 of Karatzas and Shreve 1991):

$$p_b(x, y) = \exp \left\{ -\frac{2(b-x)(b-y)}{\sigma^2(t_1-t_0)} \right\}. \quad (2.1)$$

For general multivariate Brownian motions such analytical results do not exist. When the process is two-dimensional, the probability that at least one of the barriers is breached can be semi-analytically represented in terms of an infinite sum of Bessel functions, as shown by He, Keirstead, and Rebholz (1998). For processes of dimensions three or higher, however, one has to rely on simulation techniques or some asymptotic approximations.

In the paper we consider asymptotic techniques based on the Large Deviations Theory. The proposed methods can be applied directly in cases when the asymptotic approximations provide accurate results, or they can be embedded into Monte Carlo simulations. In this section we present two methods of computing the exit probability for a Brownian Bridge process based on the Large Deviations Theory.

2.1 Large Deviations Method

Suppose that for a given $\varepsilon > 0$ the process $\{X(t), t \in [s, s + \varepsilon]\}$ satisfies

$$dX(t) = \Sigma dW(t), \quad (2.2)$$

where $\Sigma\Sigma'$ is an $n \times n$ variance-covariance matrix of the process, which we assume to be nonsingular, and $\{W(t)\}$ is a standard n -dimensional Brownian motion with zero drift and unit variance. We assume also that the initial and terminal values of the process, $X(s) = \mathbf{x}$ and $X(s + \varepsilon) = \mathbf{y}$, are known and belong to a given set $D \subset \mathbb{R}^n$. We are interested in finding the probability that the process $\{X(t)\}$ will exit D in the time interval $(s, s + \varepsilon)$:

$$\Pr\{X(t) \notin D \text{ for some } t \in (s, s + \varepsilon) | X(s) = \mathbf{x}, X(s + \varepsilon) = \mathbf{y}\}.$$

In the paper we assume that D is a convex open set whose boundary is defined by barriers represented by a collection of $(n - 1)$ -dimensional hyperplanes in the \mathbb{R}^n space. We will denote them by $B_i(t)$, $i = 1, \dots, k$, where k is at least equal to n . When $n = 2$, the barriers consist of a set of lines, and for $n = 1$, a barrier is equivalent to a point in \mathbb{R}^1 . This specification of D covers all our practical applications.

We should note that our assumption that the drift term in (2.2) is zero is not a restrictive one, because it is possible to show that this parameter is irrelevant when the end values of the process are

given. Intuitively, this is a consequence of the fact that the process has to follow a direction set by the end point. Formally, this can be justified by using properties of a pinned process, as discussed, for example, in Karlin and Taylor (1981, Chapter 15, Section 9).

It is known that by a suitable transformation of the time scale, the process $\{X(t)\}$ has the same law as the solution of either

$$dX^\varepsilon(t) = \sqrt{\varepsilon}\Sigma dW(t), \quad \text{for } t \in [0, 1], \quad (2.3)$$

with $X^\varepsilon(0) = \mathbf{x}$ and $X^\varepsilon(1) = \mathbf{y}$, or

$$dX^\varepsilon(t) = -\frac{X^\varepsilon(t) - \mathbf{y}}{1 - t} dt + \sqrt{\varepsilon}\Sigma dW(t), \quad \text{for } t \in [0, 1], \quad (2.4)$$

with $X^\varepsilon(0) = \mathbf{x}$ (see, e.g., Karlin and Taylor 1981). In the drift term in (2.4), we divide a vector by a scalar, which should be understood as a componentwise operation. Similarly, the multiplication of the matrix Σ by the scalar $\sqrt{\varepsilon}$ is also componentwise. The drift term represents the pull-back effect to the end point $X^\varepsilon(1) = \mathbf{y}$, which becomes stronger as $t \rightarrow 1$. These facts imply that finding the exit probability in the initial formulation of the problem is equivalent to solving the same problem for a process defined by (2.4).

Observe that under our assumption that $\mathbf{x}, \mathbf{y} \in D$, the exit probability will converge to 0 as $\varepsilon \rightarrow 0$, which follows from the fact that the solution of the deterministic ODE

$$d\mathbf{x}(t) = -\frac{\mathbf{x}(t) - \mathbf{y}}{1 - t} dt,$$

with $\mathbf{x}(0) = \mathbf{x}$, is a straight line that joins \mathbf{x} and \mathbf{y} . This observation justifies applications of asymptotic methods based on the Large Deviations Theory. Below we show that in the one-dimensional case this theory recovers exactly the analytical result (2.1). For a multivariate Brownian Bridge, asymptotic approximations for the case of independent Brownian motions have been developed by Baldi (1995). One of the methods studied in this paper uses some of the results obtained by Baldi, and therefore we briefly present them below. In Section 2.2 an alternative approach based on the Large Deviations Theory is proposed. A concise overview of the Large Deviations Theory is provided in Appendix A.

Following Baldi (1995), we first assume that the instantaneous variance-covariance matrix $\Sigma\Sigma'$ for the process $\{X^\varepsilon(t)\}$ in (2.4) is equal to the identity matrix \mathbf{I} . By $\|\cdot\|$ we denote the Euclidean norm, that is, for a vector $\mathbf{z} = [z_1, \dots, z_n]'$ we have $\|\mathbf{z}\| = \sqrt{\langle \mathbf{z}, \mathbf{z} \rangle}$, where the brackets $\langle \cdot, \cdot \rangle$ denote a scalar product that assigns to a pair of n -dimensional vectors \mathbf{v} and \mathbf{z} the number $\langle \mathbf{v}, \mathbf{z} \rangle = \mathbf{v}'\mathbf{z}$. We also define the following set of functions:

$$\Gamma_{\mathbf{x}}(D) = \{\gamma \in \Gamma^n : \gamma(0) = \mathbf{x} \text{ and } \gamma(t) \in D^C \text{ for some } t \in (0, 1)\}, \quad (2.5)$$

where D^C is the complement of D , and Γ^n denotes the set of all absolutely continuous functions on $[0, 1]$ with values in \mathbb{R}^n . For the exit time $\tau = \inf\{t \in (0, 1) : X^\varepsilon(t) \notin D\}$, where we use the convention that the infimum of an empty set is equal to ∞ , we want to find $\Pr_{X^\varepsilon}\{\tau < 1\}$. By specializing the general result presented in Appendix A (eqs. A.7–A.11) to the specific diffusion process (2.4), we obtain the following representation of the exit probability as $\varepsilon \rightarrow 0$:

$$\Pr_{X^\varepsilon}\{\tau < 1\} = \exp\left(-\frac{u(\mathbf{x}, \mathbf{y}, D)}{\varepsilon}\right) (1 + o(\varepsilon^m)) \quad (2.6)$$

for every $m > 0$, where

$$u(\mathbf{x}, \mathbf{y}, D) = \inf_{\gamma \in \Gamma_{\mathbf{x}}(D)} J_{\mathbf{y}}(\gamma), \quad (2.7)$$

and the action functional J_y is defined as

$$J_y(\gamma) = \frac{1}{2} \int_0^1 \left\| \frac{d\gamma(t)}{dt} + \frac{\gamma(t) - y}{1-t} \right\|^2 dt. \tag{2.8}$$

Here the little- o notation $o(\varepsilon^m)$ represents a function that converges to zero at a faster rate than ε^m , as $\varepsilon \rightarrow 0$. Intuitively the action functional (2.8) is a function that assigns a certain value to each function from $\Gamma_x(D)$.

For some values of x and y , the infimum in (2.7) can be attained at more than one point. In such cases the right-hand side of (2.6) must be modified by introducing a multiplier. For the class of sets D defined earlier, however, such situations will occur only with probability zero, and hence in the context of Monte Carlo simulations they can be ignored. For example, when D is an intersection of two half-spaces in \mathbb{R}^2 , the set of x and y for which the minimizer is not unique is equal to a straight line, which follows easily from the equivalent formulation of the optimization problem given by (2.10). Therefore, from now on we assume that x and y are such that the infimum in (2.7) is attained at a single point.

Representation (2.6) suggests the following Large Deviations approximation of the exit probability:

$$\Pr_{X^\varepsilon} \{ \tau < 1 \} \sim \Pr_{X^\varepsilon}^{LD} \{ \tau < 1 \} := \exp \left(-\frac{u(x, y, D)}{\varepsilon} \right). \tag{2.9}$$

In Sections 3 and 4 we use (2.9) for simulated values of x and y , which suggests that we need an efficient way of calculating the function u . Using the calculus of variations, Baldi (1995) has demonstrated that the infimum in (2.7) can be obtained by solving the following minimization problem:

$$u(x, y, D) = \inf_{\phi \in \partial D} \frac{1}{2} \{ (\|x - \phi\| + \|y - \phi\|)^2 - \|x - y\|^2 \}, \tag{2.10}$$

where ∂D denotes the boundary of D . This is a finite-dimensional optimization problem, which can be solved using standard procedures such as the MATLAB function “fmincon.” The minimizer of (2.10) allows for a simple geometric interpretation; it defines the shortest path from x to y that touches the boundary ∂D .

To illustrate the method, suppose that $\{X(t), t \in [0, 1]\}$ is a one-dimensional Brownian motion with zero drift and a constant diffusion coefficient σ . We are interested in finding the probability that the process breaches a barrier b between times 0 and 1, given that $X(0) = x, X(1) = y$, and $x, y > b$. In this case we can first rescale to problem by dividing $X(t)$ and the barrier by σ . Let us denote the transformed barrier and the initial and end points by b^*, x^* , and y^* , respectively. Then representation (2.10) gives us

$$u(x^*, y^*, (b^*, \infty)) = \frac{1}{2} [((x^* - b^*) + (y^* - b^*))^2 - (x^* - y^*)^2] = 2(b^* - x^*)(b^* - y^*),$$

which in the original variables is equal to $2(b - x)(b - y)/\sigma^2$. When we substitute this value for u in (2.9) and take $\varepsilon = 1$, we recover the exact analytical formula for the exit probability (2.1).

We now extend the approximation (2.9) to the case of a multivariate Brownian Bridge process with a general constant variance-covariance matrix. Because we have assumed that $\Sigma\Sigma'$ is nonsingular, it follows that we can define a matrix $K = (\Sigma')^{-1}$, which is nonsingular and has the property $(\Sigma\Sigma')^{-1} = KK'$. Using K' , we can now transform D to obtain

$$K'D := \{K'z: \text{all } z \in D\}. \tag{2.11}$$

By specializing the general expression for the action functional (A.9) to the Brownian Bridge process (2.4) with the variance-covariance matrix $\Sigma\Sigma'$, we can obtain the action functional for this process

$$J_y^K(\gamma) = \frac{1}{2} \int_0^1 \left\langle \mathbf{K}\mathbf{K}' \left(\frac{d\gamma}{dt}(t) + \frac{\gamma(t) - \mathbf{y}}{1-t} \right), \frac{d\gamma}{dt}(t) + \frac{\gamma(t) - \mathbf{y}}{1-t} \right\rangle dt. \quad (2.12)$$

Using some basic properties of the scalar product in Euclidian spaces, the right-hand side of (2.12) can be rewritten as

$$\frac{1}{2} \int_0^1 \left\| \frac{d\eta}{dt}(t) + \frac{\eta(t) - \mathbf{w}}{1-t} \right\|^2 dt, \quad (2.13)$$

where $\eta(t) = \mathbf{K}'\gamma(t)$ and $\mathbf{w} = \mathbf{K}'\mathbf{y}$. From this observation and the fact that \mathbf{K}' is invertible, it follows that $J_y^K((\mathbf{K}')^{-1}\eta(t)) = J_{\mathbf{K}'\mathbf{y}}(\eta(t))$, for each $\eta(t) \in \Gamma_{\mathbf{K}'\mathbf{x}}(\mathbf{K}'D)$. Therefore the minimization of J_y^K in (2.12) over $\Gamma_{\mathbf{x}}(D)$ is equivalent to the minimization of $J_{\mathbf{K}'\mathbf{y}}$ over $\Gamma_{\mathbf{K}'\mathbf{x}}(\mathbf{K}'D)$. This implies that for the Brownian Bridge process (2.4) with the variance-covariance matrix $\Sigma\Sigma'$, the Large Deviations approximation of the exit probability is given by (2.9), where now u is calculated as

$$u(\mathbf{K}'\mathbf{x}, \mathbf{K}'\mathbf{y}, \mathbf{K}'D). \quad (2.14)$$

Thus we can obtain u from the minimization problem (2.10) solved for the transformed values of \mathbf{x} , \mathbf{y} , and D .

This result provides a method of approximating the exit probability for a general multivariate Brownian Bridge process. It turns out, however, that this Large Deviations approximation does not depend on the correlation structure of the process. To explain this, let us focus on a two-dimensional process and a set D equal to $D_{b_1, b_2} = (b_1, \infty) \times (b_2, \infty)$, where b_1 and b_2 are such that $X_1(0) > b_1$, $X_2(0) > b_2$, $X_1(1) > b_1$, and $X_2(1) > b_2$. Let $E_1(E_2)$ denote the event that the first component (the second component) crosses the level $b_1(b_2)$, that is,

$$E_i = \{X_i(t) \leq b_i \text{ for some } t, 0 \leq t \leq 1\}, \quad i = 1, 2. \quad (2.15)$$

Then for the Large Deviations approximation (2.9) of the exit probability from D_{b_1, b_2} we have the following result, proof of which is presented in Appendix B.

Proposition 2.1

Suppose that $\{X^\varepsilon(t)\}$ follows the Brownian Bridge process (2.4) with a nonsingular variance-covariance matrix $\Sigma\Sigma'$. Then the Large Deviations approximation (2.9) of the exit probability from $D_{b_1, b_2} = (b_1, \infty) \times (b_2, \infty)$, given by

$$\Pr_{X^\varepsilon}^{LD}\{\tau < 1\} = \exp\left(-\frac{u(\mathbf{K}'\mathbf{x}, \mathbf{K}'\mathbf{y}, \mathbf{K}'D_{b_1, b_2})}{\varepsilon}\right),$$

satisfies

$$\Pr_{X^\varepsilon}^{LD}\{\tau < 1\} = \max\{\Pr(E_1), \Pr(E_2)\},$$

where E_i , $i = 1, 2$, are defined in (2.15).

Proposition 2.1 can be generalized to higher dimensions. In this case the Large Deviations approximation of $\Pr(E_1 \cup E_2 \cup \dots \cup E_n)$ will be the maximum of $\Pr(E_1), \Pr(E_2), \dots, \Pr(E_n)$. This result diminishes the usefulness of this Large Deviations approximation, because it neglects the covariance structure of the process. Consequently the method can give quite inaccurate approximations, which has been confirmed by our numerical examples presented in Section 2.3. Throughout the paper we will refer to this method as the *LD1* method. In the next section we propose another way of utilizing the Large Deviations Theory.

2.2 The Proposed Method

For ease of exposition we focus on a two-dimensional process, but later we generalize the results to higher dimensions. Suppose that $\mathbf{X}(t) = [X_1(t), X_2(t)]'$ follows the stochastic differential equation (2.4)

with Σ' denoting a 2×2 positive definite matrix. To simplify the presentation, we assume that we have two constant barriers b_1 and b_2 such that $X_1(0) > b_1, X_2(0) > b_2, X_1(1) > b_1,$ and $X_2(1) > b_2.$ The method can be easily modified in cases when the barriers are not parallel to the axis or when the initial points are below them.

For the two events E_1 and E_2 defined in (2.15), the exit probability from D_{b_1, b_2} is given by $\Pr(E_1 \cup E_2).$ In this section we propose a method of approximating this probability by applying the Large Deviations Theory to the probability $\Pr(E_1 \cap E_2)$ rather than directly to $\Pr(E_1 \cup E_2).$ As we discuss in Section 2.1, the latter approach does not capture the correlation structure of the process, and the proposed method is designed to remedy this situation. Clearly the method is feasible by the elementary inclusion-exclusion principle

$$\Pr(E_1 \cup E_2) = \Pr(E_1) + \Pr(E_2) - \Pr(E_1 \cap E_2) \tag{2.16}$$

and the fact that from (2.1) we can easily compute the marginal barrier crossing probabilities, $\Pr(E_1)$ and $\Pr(E_2).$ We should observe that $\Pr(E_1 \cap E_2)$ can no longer be interpreted as the exit probability. However, as $\varepsilon \rightarrow 0,$ the sequence of processes $\{X^\varepsilon(t)\}$ in (2.4) satisfies the Large Deviations Principle with the action functional being the same as for the problem of finding $\Pr(E_1 \cup E_2).$

Let $\Gamma_x^*(D_{b_1}, D_{b_2})$ be the set of absolutely continuous functions with values in \mathbb{R}^2 such that both of their components breach corresponding barriers between time 0 and time 1:

$$\Gamma_x^*(D_{b_1}, D_{b_2}) = \{\gamma \in \Gamma^2 : \gamma(0) = \mathbf{x} \text{ and } \gamma(t_1) \in D_{b_1}^C, \gamma(t_2) \in D_{b_2}^C \text{ for some } t_1, t_2 \in (0,1)\},$$

where $D_{b_i} = \{\mathbf{x} = [x_1, x_2]' \in \mathbb{R}^2 : x_i > b_i\}, i = 1, 2.$ The key to the proposed method is the observation that the probability $\Pr(E_1 \cap E_2)$ can also be approximated using the Large Deviations method, which follows from Remark A.1 in Appendix A. In the case when $\Sigma' = \mathbf{I},$ the function u in (2.9) can be calculated from the formula

$$u^*(\mathbf{x}, \mathbf{y}, D_{b_1}, D_{b_2}) = \inf_{\gamma \in \Gamma_x^*(D_{b_1}, D_{b_2})} J_\gamma(\gamma). \tag{2.17}$$

For a general variance-covariance matrix $\Sigma',$ we can apply a transformation similar to the one presented in Section 2.1, giving us u in (2.9) of the form $u^*(\mathbf{K}'\mathbf{x}, \mathbf{K}'\mathbf{y}, F_1, F_2),$ where $F_i = \mathbf{K}'D_{b_i}, i = 1, 2.$

Evaluation of the infimum that appears in the representation (2.17) can be reduced to a convex optimization problem over a finite-dimensional domain. Some calculations, which we present in Appendix C, show that $u^*(\mathbf{K}'\mathbf{x}, \mathbf{K}'\mathbf{y}, F_1, F_2)$ can be equivalently represented as

$$u^*(\mathbf{K}'\mathbf{x}, \mathbf{K}'\mathbf{y}, F_1, F_2) = \min_{\eta \in S_2} \inf_{\substack{\phi_1, \phi_2 \\ \phi_i \in \partial F_i}} \frac{1}{2} \{(\|\mathbf{K}'\mathbf{x} - \phi_{\eta(1)}\| + \|\phi_{\eta(1)} - \phi_{\eta(2)}\| + \|\mathbf{K}'\mathbf{y} - \phi_{\eta(2)}\|)^2 - \|\mathbf{K}'\mathbf{x} - \mathbf{K}'\mathbf{y}\|^2\}, \tag{2.18}$$

where the minimum is taken over the set S_2 that includes all permutations of $\{1, 2\}$ and $\eta(i)$ denotes the i th coordinate of a permutation η from $S_2.$ Once the value of u^* is known, the probability $\Pr(E_1 \cap E_2)$ is obtained from expression (2.9). A geometric interpretation of the problem (2.18) is also possible; in this case the minimizer corresponds to the shortest path that connects $\mathbf{K}'\mathbf{x}$ and $\mathbf{K}'\mathbf{y}$ and touches the boundaries ∂F_1 and $\partial F_2.$

The above approach can be generalized to the case when the dimension of the process is greater than two. Then to compute $\Pr(E_1 \cap \dots \cap E_n)$ we can use (2.9) with u equal to

$$\min_{\eta \in S_n} \inf_{\substack{\phi_1, \phi_2, \dots, \phi_n \\ \phi_i \in \partial F_i, i=1, \dots, n}} \frac{1}{2} \{ (\|K'x - \phi_{\eta(1)}\| + \|\phi_{\eta(1)} - \phi_{\eta(2)}\| + \|\phi_{\eta(2)} - \phi_{\eta(3)}\| + \dots + \|K'y - \phi_{\eta(n)}\|)^2 - \|K'x - K'y\|^2 \}, \quad (2.19)$$

where S_n represents the set of all permutations of $\{1, \dots, n\}$. Hereafter, we will refer to this approach as the *LD2* method.

2.3 Implementation and Numerical Examples

2.3.1 Implementation of the Proposed Method

For a multidimensional Brownian Bridge process, we need to solve the optimization problem specified in (2.19). Because this is a convex optimization problem, for each permutation η in (2.19) we can employ standard optimization tools, such as the procedure “fmincon” in Matlab. Alternatively, we can use an approximating algorithm that is presented in Appendix D. We use the first method for computing exit probabilities of a Brownian Bridge, and the second method for the Brownian motion process discussed in the following sections, because in this case the optimization problem needs to be solved many times.

When we are interested in finding an exit probability for three-dimensional processes, we compute $\Pr(E_1 \cup E_2 \cup E_3)$ by evaluating the right-hand side of the following expression:

$$\Pr(E_1 \cup E_2 \cup E_3) = \sum_{i=1}^3 \Pr(E_i) - \sum_{i < j} \Pr(E_i \cap E_j) + \Pr(E_1 \cap E_2 \cap E_3).$$

In our method the first-order terms are calculated analytically by (2.1). The second-order intersection terms can be evaluated using the representation by He, Keirstead, and Rebholz (1998). Alternatively, these terms can be calculated efficiently using the *LD2* method, because the optimization problem (2.18) in two dimensions can be solved analytically using a geometric argument (Huh 2007). We briefly outline this approach. To simplify the explanation, here we denote Kx as \tilde{x} and Ky as \tilde{y} . Then, the optimization problem (2.18) can be reduced to

$$\inf_{\phi_1 \in l_1, \phi_2 \in l_2} (\|\tilde{x} - \phi_1\| + \|\phi_1 - \phi_2\| + \|\tilde{y} - \phi_2\|), \quad (2.20)$$

where l_1 and l_2 are the two lines that represent boundaries of two half-planes whose intersection forms the region to which the points \tilde{x} and \tilde{y} belong. By $\mathbf{b} = [b_1, b_2]'$ we denote the intersection point of l_1 and l_2 . Suppose that we want to calculate the shortest path from \tilde{x} to \tilde{y} that reaches l_1 first, and then l_2 . Let us denote by $\text{ref}_{l_1}(\tilde{x})$ the reflection of \tilde{x} around the line l_1 .² We define similarly $\text{ref}_{l_2}(\tilde{y})$. The coordinates of these reflected points can be easily obtained analytically. Let l be the line joining these points. It is easy to verify that if l crosses both l_1 and l_2 , then the shortest path from \tilde{x} to \tilde{y} has the same length as the line segment joining $\text{ref}_{l_1}(\tilde{x})$ and $\text{ref}_{l_2}(\tilde{y})$:

$$\inf_{\phi_1 \in l_1, \phi_2 \in l_2} (\|\tilde{x} - \phi_1\| + \|\phi_2 - \phi_1\| + \|\tilde{y} - \phi_2\|) = \|\text{ref}_{l_1}(\tilde{x}) - \text{ref}_{l_2}(\tilde{y})\|.$$

² A formal definition of the reflection is presented in Appendix B.

On the other hand, if l does not cross l_1 and l_2 , then

$$\begin{aligned} & \inf_{\phi_1 \in l_1, \phi_2 \in l_2} (\|\tilde{\mathbf{x}} - \phi_1\| + \|\phi_2 - \phi_1\| + \|\tilde{\mathbf{y}} - \phi_2\|) \\ &= \begin{cases} \inf_{\phi \in l_1} (\|\tilde{\mathbf{x}} - \phi\| + \|\tilde{\mathbf{y}} - \phi\|), & \text{if for optimal } \phi^*, \text{ path } \tilde{\mathbf{x}} \rightarrow \phi^* \rightarrow \tilde{\mathbf{y}} \text{ crosses } l_2 \\ \inf_{\phi \in l_2} (\|\tilde{\mathbf{x}} - \phi\| + \|\tilde{\mathbf{y}} - \phi\|), & \text{if for optimal } \phi^*, \text{ path } \tilde{\mathbf{x}} \rightarrow \phi^* \rightarrow \tilde{\mathbf{y}} \text{ crosses } l_1 \\ \|\tilde{\mathbf{x}} - \mathbf{b}\| + \|\tilde{\mathbf{y}} - \mathbf{b}\|, & \text{otherwise} \end{cases} \\ &= \begin{cases} \|\text{ref}_{l_1}(\tilde{\mathbf{x}}) - \tilde{\mathbf{y}}\|, & \text{if shortest path to } l_1 \text{ also crosses } l_2 \\ \|\tilde{\mathbf{x}} - \text{ref}_{l_2}(\tilde{\mathbf{y}})\|, & \text{if shortest path to } l_2 \text{ also crosses } l_1 \\ \|\tilde{\mathbf{x}} - \mathbf{b}\| + \|\tilde{\mathbf{y}} - \mathbf{b}\|, & \text{otherwise.} \end{cases} \end{aligned}$$

In the above equation the first two cases correspond to the situation when the shortest path from $\tilde{\mathbf{x}}$ to $\tilde{\mathbf{y}}$ that touches one line happens to cross the other line. It is easy to verify that this particular path would also be the shortest path from $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ that touches both lines. If such a situation does not occur, the shortest path is the one that touches the point \mathbf{b} .

By considering in the same way the alternative order of reaching l_1 and l_2 , we can solve (2.18) without using numerical methods.

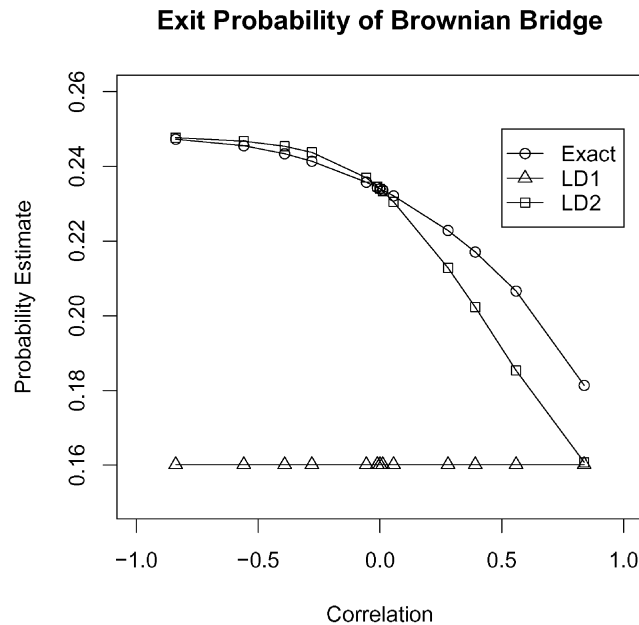
The evaluation of $\Pr(E_1 \cap E_2 \cap E_3)$ takes more time, because we need to solve the optimization problem specified in (2.19). In many situations $\Pr(E_1 \cap E_2 \cap E_3)$ will be very small compared with $\Pr(E_i \cap E_j)$, $i, j = 1, 2, 3$, and hence by taking this probability equal to zero will not increase significantly the overall error. To implement such a technique, we need to have some criteria to determine whether this probability is small enough to be ignored. A possible approach is as follows. If $\min\{\Pr(E_1), \Pr(E_2), \Pr(E_3)\}$ and $\min\{\Pr(E_1 \cap E_2), \Pr(E_2 \cap E_3), \Pr(E_1 \cap E_3)\}$ are smaller than prespecified threshold values, say, α_1 and α_2 , then we let $\Pr(E_1 \cap E_2 \cap E_3) = 0$. Otherwise, we compute $\Pr(E_1 \cap E_2 \cap E_3)$ using the method described in the previous section. The numerical examples that we present in Section 4.1.2 provide some insight into the problem of proper selection of α_1 and α_2 .

2.3.2 Numerical Examples

Suppose that a two-dimensional Brownian motion has fixed initial and end values of $\mathbf{x} = [0.784, 0.778]'$ and $\mathbf{y} = [0.788, 0.781]'$, respectively, and that the corresponding barriers are $b_1 = b_2 = 1$. The variances of the first and the second components of the two-dimensional Brownian motion are equal to 0.1 and 0.08, respectively. For a given time horizon of $\varepsilon = 0.5$, we have estimated the probability that at least one barrier is breached. For different values of the correlation coefficient between the two components of the process, we have obtained estimates using the *LD1* and *LD2* methods, described in Sections 2.1 and 2.2, respectively. As a benchmark, we have also calculated the exit probabilities using the first 20 terms of an infinite series representation given by He, Keirstead, and Rebholz (1998). We have found that this number of terms is sufficient to provide accurate approximations up to the third decimal place. We refer to these values as ‘‘Exact.’’

Figure 1, which depicts all these estimates, suggests that the proposed method gives good results for different values of the correlation coefficient ρ , especially for negative or small positive correlations. This phenomenon can be explained by the fact that for negative values of ρ it is less likely that both coordinates of the process will touch the corresponding barriers, implying that the exit probability $\Pr(E_1 \cup E_2)$ is well approximated by $\Pr(E_1) + \Pr(E_2)$. Estimates of the probability $\Pr(E_1 \cup E_2)$ based on the *LD1* method are constant for all correlation values. We have obtained similar results under different settings of the initial value, end value, and time horizon.

Figure 1
Estimates of Exit Probabilities Using the Two Methods



3. EXIT PROBABILITY FOR A MULTIVARIATE BROWNIAN MOTION

We now consider a correlated multivariate Brownian motion and suggest an algorithm for computing exit probabilities that takes into account the crossings at the end points and at intermediate times. In particular, we focus on computing the probability that any of the components crosses its respective barrier during the given time horizon.

3.1 General Algorithm

The algorithm presented in this section is in principle applicable to a Brownian motion of any dimension. However, as the dimension increases, the efficiency of the method will very likely deteriorate because of the additional time needed to solve the optimization problem in (2.19). This does not mean that the method cannot be applied to a high-dimensional Brownian motion, because the accuracy of the method depends on the number of terms we use in the inclusion-exclusion formula. In some problems it may be possible to determine that terms of the form $P(E_{j_1} \cap \cdots \cap E_{j_i})$ for i greater than certain number, say, i_0 , contribute very little to the exit probability. This will mean that only terms of order up to i_0 need to be evaluated.

To simplify the exposition, we focus on a three-dimensional stochastic process $\mathbf{X}(t)$ that follows a Brownian motion of the form

$$d\mathbf{X}(t) = \mathbf{U}dt + \Sigma d\mathbf{W}(t), \quad t \in [0, T], \quad (3.1)$$

with the barrier levels $\mathbf{B} = [b_1, b_2, b_3]'$, where \mathbf{U} is a three-dimensional drift vector. We first generate values of the underlying process at discrete time intervals for a specified time horizon. If the process breaches the barrier at any of the discrete points, then the simulated path is considered to have crossed the barrier. Otherwise, for each subinterval and corresponding initial and terminal values of the process, we need to approximate the probability of the process crossing the barrier using the techniques described in the previous section.

First, we describe the algorithm when the number of subintervals, M , is equal to one. Then we present the general version.

For $M = 1$, the algorithm with N runs of simulation and $T = 1$ is as follows:

1. Initialize $Default_Cnt = 0$, and take $\mathbf{X}(0) = [x_1(0), x_2(0), x_3(0)]'$, which is the initial value of \mathbf{X} , such that $x_1(0) > b_1$, $x_2(0) > b_2$, and $x_3(0) > b_3$.
2. Generate a value $\mathbf{X}(1) = [x_1(1), x_2(1), x_3(1)]'$ according to the Brownian motion process (3.1).
3. If for some i we have $x_i(1) \leq b_i$, then at least one component of the process has breached the barrier. In this case increment $Default_Cnt = Default_Cnt + 1$.
4. If $x_1(1) > b_1$, $x_2(1) > b_2$, and $x_3(1) > b_3$, then none of the simulated values of the components have crossed the barrier at the discrete point. Denote by CP the probability that at least one component of \mathbf{X} crosses its corresponding component of the barrier \mathbf{B} with fixed initial and terminal values $\mathbf{X}(0)$ and $\mathbf{X}(1)$. We increment $Default_Cnt$ by CP , that is, $Default_Cnt = Default_Cnt + CP$. We discuss computation of CP later in this section.
5. Repeat Steps 2–4 N times.
6. As an estimate of the exit probability,³ take $Default_Cnt/N$.

If we have more than one subinterval (i.e., $M > 1$), Steps 2, 3, and 4 of the above procedure should be changed to Steps 2', 3', and 4':

- 2'. Generate values $\mathbf{X}(i\varepsilon) = [x_1(i\varepsilon), x_2(i\varepsilon), x_3(i\varepsilon)]'$, $i = 1, \dots, M$, at each discrete time step according to the dynamics (3.1), where $\varepsilon = T/M$.
- 3'. If there is $i \in \{1, \dots, M\}$ such that either $X_1(i\varepsilon) \leq b_1$, $X_2(i\varepsilon) \leq b_2$ or $X_3(i\varepsilon) \leq b_3$, then the given simulated path is considered to have breached the barrier, and we increment $Default_Cnt = Default_Cnt + 1$.
- 4'. If $\forall i \in \{1, \dots, M\}$, $x_1(i\varepsilon) > b_1$, $x_2(i\varepsilon) > b_2$, and $x_3(i\varepsilon) > b_3$, then the simulated path has not crossed the barrier at the discrete points. Denote by CP_i , $i = 1, \dots, M$, the exit probability of the process \mathbf{X} on the subinterval $((i - 1)\varepsilon, i\varepsilon)$ with fixed initial and terminal values $\mathbf{X}((i - 1)\varepsilon)$ and $\mathbf{X}(i\varepsilon)$, respectively. The exit probability of $\{\mathbf{X}(t), 0 \leq t \leq T\}$ is calculated as $1 - \prod_{i=1}^M (1 - CP_i)$, and we increment $Default_Cnt$ by $1 - \prod_{i=1}^M (1 - CP_i)$, that is, $Default_Cnt = Default_Cnt + (1 - \prod_{i=1}^M (1 - CP_i))$.

In the description of the above algorithm, CP or CP_i 's are the exit probabilities of the multidimensional Brownian Bridge process. Depending on how these probabilities are calculated, we have the following three specifications of this simulation framework:

- *Crude Monte Carlo Method (CMC)*: CP and CP_i 's are assumed to be 0.
- *Large Deviations Method (LDMC1)*: CP and CP_i 's are computed using the method *LD1* described in Section 2.1
- *Proposed Method (LDMC2)*: CP and CP_i 's are computed using the method *LD2* described in Section 2.2.

3.2 Efficient Version of the Algorithm

In this section we propose a modification of the method *LDMC2*. This approach is motivated by the fact that *LDMC2* is computationally costly, because it involves solving an optimization problem at each time interval for each simulation path. To reduce the computational burden of the optimization, we have devised an algorithm that is much faster and still produces accurate results. We will refer to this method as a Fast Large Deviations Monte Carlo (*FLDMC*) method.

³ Intuitively this formula can be explained as follows. If all values of CP were equal to zero, then, by the Law of Large Numbers, $Default_Cnt/N$ would converge to the probability that the end point of the process is below the barrier. After including CP into $Default_Cnt$, as in Step 4, the fraction $Default_Cnt/N$ approximates the sum of the exit probability because of the end point crossing the barrier and the probability that the process has crossed the barrier at an intermediate time given that the process was above the barrier at the end point. Formally this can be justified by conditioning on the end value of the process.

The method exploits the fact that we have an analytical representation of the probability that at least one barrier is breached during the given time horizon when the components of the process are uncorrelated. The algorithm is a two-stage procedure. In the first stage, we run a small number of simulations to determine the ratio between the probability of barrier crossing under the given correlation structure and that under zero correlation. In the second stage, where we run the full set of simulations, we compute the barrier crossing probability under zero correlation and then use the estimated ratio to adjust for the correlation effect.

3.2.1 Step 1: Preliminary Runs

We perform the following procedures:

1. We generate a small number, for example, $N = 200$, of values of the multivariate correlated Brownian motion at the end of the first time interval.
2. For the j th simulation run, we know the initial and the terminal values, $\mathbf{X}^{(j)}(0) = \mathbf{x}_0$ and $\mathbf{X}^{(j)}(\epsilon) = \mathbf{y}^{(j)}$, of the process. If $\mathbf{y}^{(j)}$ does not breach the barrier, we compute the exit probability CP^{ρ} that at least one component breaches its respective barrier given that the correlation coefficient is ρ . For this we use the method *LD2* from Section 2.2. In addition, using the analytical representation we compute CP^0 , which is the exit probability under zero correlation. Then we find $Ratio^{(j)} = CP^{\rho}/CP^0$.
3. After generating N paths, let N' be the number of simulations where the end values did not breach the barrier. In this stage we need to obtain the average ratio of the exit probability under the correlation structure ρ and under zero correlation. We set $RatioSum$ to be the sum of $Ratio^{(j)}$'s over the paths for which $\mathbf{y}^{(j)}$ did not breach the barrier. We can then obtain the desired average ratio by $Ratio = RatioSum/N'$. This ratio is used in the second stage.

3.2.2 Step 2: Full Simulation

Now we perform a full set of simulations with a large number of runs, that is, $N = 80,000$. The general simulation framework still remains the same as that described in Section 3.1, but now we modify the way we calculate CP and CP_i 's. These numbers are obtained by a method similar to *LD2* in Section 2.2; however, now the term $\Pr(E_1 \cup E_2 \cup E_3)$ is calculated differently.

For a given subinterval we have an initial value and a terminal value for each simulated path. When the instantaneous correlation is 0, the probability that at least one component breaches its respective barrier can be calculated explicitly. For the three-dimensional process, we denote this probability as $\Pr^0(E_1 \cup E_2 \cup E_3)$. It can be expressed as

$$\Pr^0(E_1 \cup E_2 \cup E_3) = \sum_{i=1}^3 \Pr(E_i) - \sum_{i < j} \Pr(E_i)\Pr(E_j) + \Pr(E_1)\Pr(E_2)\Pr(E_3). \quad (3.2)$$

Because the marginal barrier crossing probability is given by formula (2.1), the computation of (3.2) is quite fast.

The quantity that we ultimately need to compute is $\Pr^{\rho}(E_1 \cup E_2 \cup E_3)$, which is the exit probability under the specified correlation structure ρ . This can be obtained by adjusting $\Pr^0(E_1 \cup E_2 \cup E_3)$ with the factor $Ratio$ from Step 1 as follows:

$$\Pr^{\rho}(E_1 \cup E_2 \cup E_3) = \Pr^0(E_1 \cup E_2 \cup E_3) \times Ratio.$$

Once we obtain $\Pr^{\rho}(E_1 \cup E_2 \cup E_3)$, the remaining parts follow in the same way as in *LDMC2*. When $M > 1$, the factor $Ratio$ is calculated using the first subinterval as described in Section 3.2.1. Then the same value of $Ratio$ is applied to approximate values of the barrier crossing probabilities of other subintervals, namely, CP_i , $i = 1, \dots, M$.

This algorithm renders accurate results if the the ratio between the barrier crossing probability under the given correlation and that under zero correlation is more or less constant with respect to different time steps and different locations of initial and terminal values relative to the barrier levels. Although

at this time we are not able to offer any formal analysis of the accuracy of the method, our numerical experiments affirm that this method greatly improves time efficiency while retaining accuracy of the estimates.

3.3 Comparison with an Existing Method

The work of Shevchenko (2003) proposes an efficient simulation method to find multivariate crossing probability. The author simulates trajectories of the underlying process at discrete time intervals and then computes the barrier crossing probability of a Brownian Bridge between discrete points using the Fréchet bounds. These bounds have been established for multivariate distributions defined by means of a copula function (e.g., Nelsen 1999). This method can also be used to approximate barrier crossing probabilities for a continuous-time process using a small number of discrete intervals. It can be shown, however, that the resulting method is equivalent to the method *LDMCI*. In addition, our numerical studies indicate that for a given number of subintervals *FLDMC* works at the same level of computational efficiency as Shevchenko's method, giving at the same time more accurate results.

We should also mention that the method presented by Shevchenko is well suited only for computing the probability that at least one component breaches its respective barrier. It cannot be easily extended to problems where the distribution of the number of components breaching barriers is required. On the other hand, *FLDMC* can be relatively easily adapted to find such distributions.

4. APPLICATIONS IN CREDIT RISK MODELS

Credit risk is the possibility of a financial loss due to a default event of a counterparty or migration of counterparty's credit rating. A broad array of applications require pricing models that are tractable and at the same time flexible enough to reproduce important real features of credit events. Models that attempt to describe default processes can be divided into two main categories: structural models and reduced-form models. Initiated by Merton (1974), the first category of models uses the evolution of firms' structural variables, such as asset and debt values, to determine the time of default. In Merton's model a firm defaults if at the time of servicing the debt its assets are below the debt. Black and Cox (1976) extends the Merton model and provides a more realistic description of default by allowing it to occur at any time between the inception of the contract and its maturity. It can be described as a first-passage-time model, because it specifies the default time as the first time the firm's asset value hits a lower barrier. The barrier can be determined exogenously, as in Black and Cox (1976) and Longstaff and Schwartz (1995), and then it acts as a safety covenant to protect the bondholders. Alternatively, it can be determined endogenously, and then it corresponds to the stockholders' attempt to maximize the value of the firm (Leland 1994; Leland and Toft 1996).

In contrast to structural models, in reduced-form models (also known as intensity-based models) the time of default is not determined directly by the value of the firm but is the first jump of an exogenously given jump process whose hazard rate is allowed to depend on economic factors. Parameters that enter the hazard rate must be inferred from market data, such as spreads between defaultable or nondefaultable instruments. As a result, these models differ fundamentally from structural models in the degree of predictability, because for the structural models the time of default is predictable given knowledge of the current value of the firm. Because of their mathematical tractability and flexibility to reproduce realistic implied short-term credit spreads, reduced-form models have become popular among academics and practitioners. Empirical evidence concerning these models, however, is rather limited. Also, under these models the underlying mechanism that triggers defaults is not modeled explicitly. For an overview of different credit risk models, the interested reader is referred to McDonald (2006), Lando (2004), or Schönbucher (2003).

An appealing feature of structural default models is that they provide an explicit link between default of a firm and the asset value of the firm. However, they cannot explain the observed spreads between prices of defaultable and nondefaultable instruments over a short time period after the inception. This can be explained by the fact that exit times for continuous stochastic processes are predictable, which

implies that in structural models there is always a gradual degradation in credit quality and defaults cannot come as a surprise. As a result, these models underestimate default probabilities. To address this problem, hybrids of structural and reduced form models have been proposed. For example, Zhou (2001) and Chen and Panjer (2003) model the firm's value as a diffusion with jumps process, which leads to a structural model where a default can occur either as a result of jump to default or as the first time a process breaches a barrier in a continuous manner. Another first-passage approach has been recently proposed by Douady and Jeanblanc (2002), who use credit ratings as a mechanism that triggers default. Assuming that a bond issuer is assigned a continuous in time rating that follows a jump-diffusion process, they show that most other credit models can be seen either as particular cases or as limit cases of the proposed model. They also demonstrate that this approach is tractable, in terms of pricing, hedging, and calibrations, and flexible enough to reproduce empirically consistent short-term credit spreads. Credit derivatives that can be priced using this model include credit default swaps, Brady bond options, and convertible bonds, for one issuer, and basket protection, tranche insurance, and first n to default, for several issuers.

In managing a portfolio of credit risks, such as CDO tranches, it is crucial to model properly the joint default behavior of several counterparties. This aspect of credit models has attracted much attention among practitioners and academics. The current market standard is represented by copula models, as discussed in Li (1999, 2000), with the dependency between default times modeled through one of the standard copulas, like Gaussian, Student's t , Clayton, or Marshall-Olkin. A shortcoming of copula models is that they do not correspond to any financial intuition and describe the dependency of default events only statistically. Therefore, in such approaches it is difficult to link the dependency with observable economic factors. In this context the first-passage models have the strong advantage that they do not require the specification of joint distributions of the pairwise, triplet-wise, etc., defaults, because in practice these values are very difficult to obtain from either historical or current market data. For structural models based on diffusion processes, only specifications of pairwise correlations are necessary. When the underlying follows a jump-diffusion process, however, one also needs to model a dependency structure of the jumps.

Practical implementations of structural or rating-based models rely on the availability of efficient methods of estimating exit time probabilities. As explained in the Introduction, except for very few special cases, these probabilities do not have analytic representations, especially for models based on multivariate stochastic processes. Hence models for portfolios of credit risks will typically require applications of numerical methods, such as a Monte Carlo simulation technique or a finite-difference method. The latter, however, typically become ineffective for problems where the dimension of the underlying process is higher than three.

In this section we present two applications of the methods we have developed in Section 3. The first example computes, under the real-world measure, the probability that at least one entity in a portfolio defaults, and the second example obtains the fair premium of a credit default swap under a risk-neutral probability measure.

4.1 COMPUTATION OF THE PROBABILITY OF DEFAULT

4.1.1 Formulation of the Problem

We consider a portfolio consisting of three credit risky bonds issued by three distinct counterparties. Our objective is to determine the probability that at least one counterparty defaults within a horizon of one year. For this we adopt the structural approach, according to which the total value of assets of a firm is represented by a stochastic process, and the obligor is considered to have defaulted if this process crosses a certain threshold. The asset value of the firm, $\mathbf{S} = [S_1(t), S_2(t), S_3(t)]'$, is assumed to follow a three-dimensional geometric Brownian motion

$$d\mathbf{S}(t) = \text{diag}[\boldsymbol{\gamma}]\mathbf{S}(t) dt + \text{diag}[\mathbf{S}(t)]\boldsymbol{\Sigma} d\mathbf{W}(t), \quad (4.1)$$

where $\{W(t)\}$ is a three-dimensional standard Brownian motion, Σ is a 3×3 matrix corresponding to a square root of an instantaneous variance-covariance matrix, for example, its Cholesky decomposition, and $\gamma = [\gamma_1, \gamma_2, \gamma_3]'$ is a drift vector. For a generic vector $x = [x_1, \dots, x_n]'$, by $\text{diag}[x]$ we denote an $n \times n$ matrix with diagonal elements given by x_1, \dots, x_n and all other elements equal to zero. For a vector of default levels $C = [c_1, c_2, c_3]'$, the i th obligor is considered to default if S_i crosses c_i .

If the process (4.1) is estimated using historical data of some proxies of the assets value of the firm, then this dynamic would be specified under the real-world probability measure. In such a case the exit probability that we calculate below can be used for forecasting future default behavior and for risk management (see, e.g., Schönbucher 2003, Chapter 9, or Duffie and Singleton 2003).

The log-transformation of (4.1) and the application of Itô's lemma result in the following stochastic differential equation:

$$d \ln S(t) = U dt + \Sigma dW(t), \tag{4.2}$$

where $\ln S(t) = [\ln S_1, \ln S_2, \ln S_3]'$, $U = [\mu_1, \mu_2, \mu_3]'$, $\mu_i = \gamma_i - \sigma_i^2/2$, and σ_i^2 is the i th diagonal element of $\Sigma\Sigma'$. Hence, by setting $X = \ln S$ and $B = [b_1, b_2, b_3]'$ = $[\ln c_1, \ln c_2, \ln c_3]'$, we find that $X(t)$ solves the equation

$$dX(t) = U dt + \Sigma dW(t), \tag{4.3}$$

and the threshold level is given by $B = [b_1, b_2, b_3]'$. In this context we apply the simulation approach described in Section 3.1.

For the three-dimensional process (4.3), we have the parameterizations of the drift and variance terms as given in Table 1.

We can rewrite the barrier crossing probability of a one-dimensional Brownian process (2.1) as

$$p_\beta(x, y) = \exp \left\{ -\frac{2d_0d_1}{\sigma^2(t_1 - t_0)} \right\},$$

where d_0 and d_1 are distances to the barrier from the initial value and the terminal value, respectively. We notice that the effects of variance, time horizon, and distances to a default barrier are all confounded. Hence, for the numerical illustration, we change only one of these variables to analyze the accuracy of our results in various situations. In our study we vary the level of default threshold denoted by B , and we fix the time horizon $T = 1$. For illustration purposes, we report the results under one setting of the default threshold level, $B = [4.51, 4.48, 4.47]'$, as other levels of default threshold provide us with similar results.

In our studies the correlation ρ between each pair of the components was allowed to take values from the set $\{-0.2, 0, 0.2, 0.8\}$. We focus more on high positive values of ρ because in such cases, as suggested by our empirical studies presented in Section 2.3.2, the *LD2* method gives less accurate results. We also considered different values of the number of subintervals, M , the number of simulation runs, N , and the levels of the thresholds α_1 and α_2 , which, as described in Section 2.3.1, determine whether to calculate the third-order term. For different combinations of these values, we have examined the effectiveness of different methods discussed in Section 3, namely, *CMC*, *LDMC1*, *LDMC2*, and *FLDMC*.

Table 1
Parameters of Brownian Motions for Underlying Processes, Initial Value = 4.54

Component	μ	σ^2
1	0.07	0.02
2	0.0325	0.035
3	0.03	0.015

4.1.2 Results and Findings

First, we compare the performance of *CMC* to that of *LDMC2*. Figures 2 and 3 show estimates from *CMC* and *LDMC2* as we vary the threshold values of α_1 and α_2 . These threshold values are 0.1, 0.03 for Case 1, 0.8, 0.6 for Case 2, and 1, 1 for Case 3, respectively. In Case 3 the third-order intersection term is always assumed to be zero. In Figures 2 and 3 the two left-hand graphs display the estimates using the *LDMC2* method with various thresholds levels α_1 and α_2 . The two right-hand graphs show the results obtained using *CMC*. In both cases we plot the values of the estimates as a function of the number of subintervals. We do not know the exact value of default probability except for the case where $\rho = 0$. However, it is reasonable to assume that the true value is equal to the level at which the curve resulted from *LDMC2* flattens and stabilizes at sufficiently large values of M . Also, we know that the estimate from *CMC* should converge to the true value as $M \rightarrow \infty$. When $\rho = 0$, the true default probability in this setting can be calculated explicitly and is 0.9736. In the context of Case 1 with the *LDMC2* method we obtain the values 0.9728 and 0.9733 for $M = 1$ and 2, respectively. To achieve the

Figure 2
Comparison of Estimates from *LDMC2* and *CMC* with Various Values of α_1 and α_2 (Part 1)

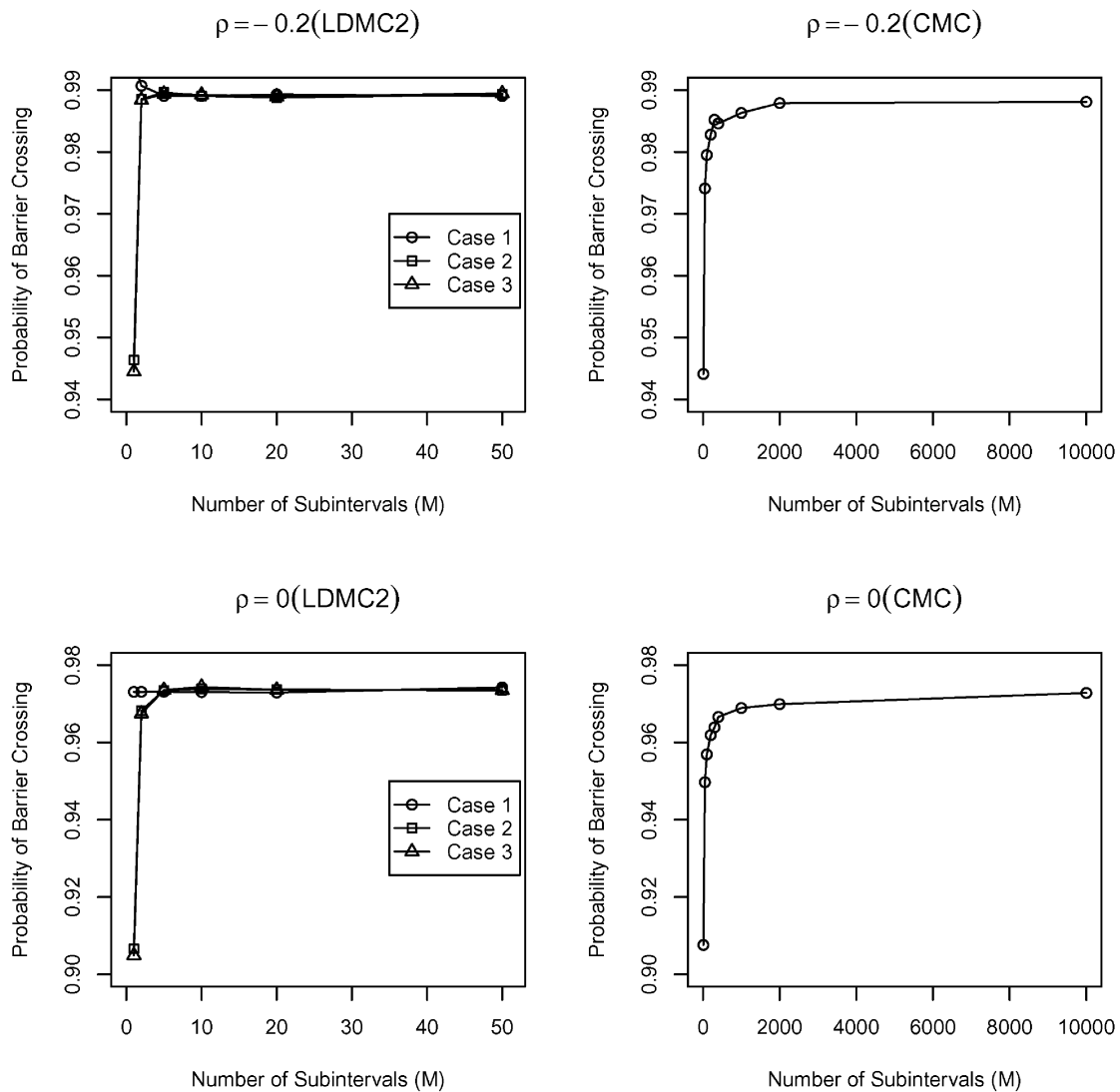
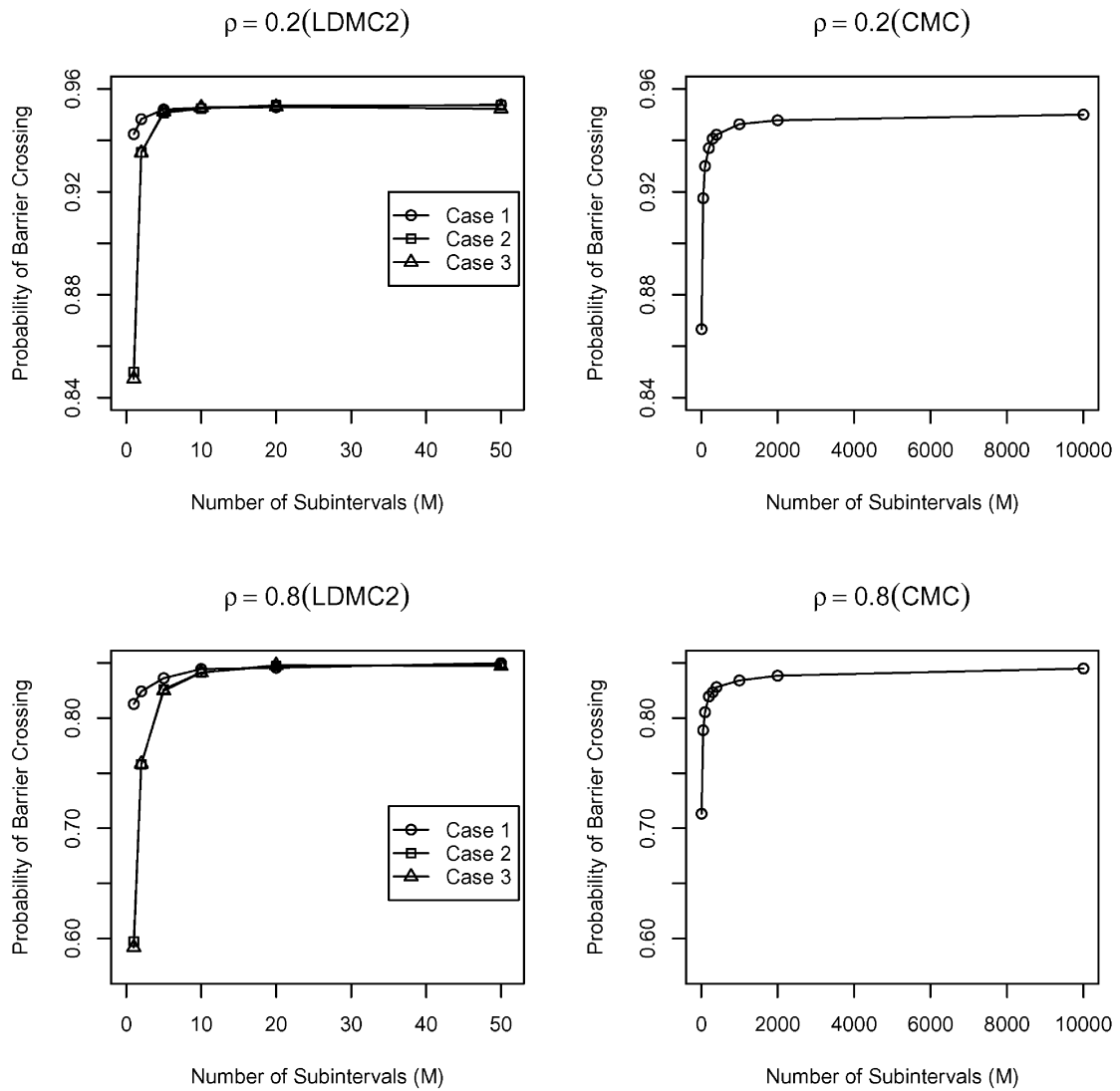


Figure 3
Comparison of Estimates from LDMC2 and CMC with Various Values of α_1 and α_2 (Part 2)



same level of accuracy using the *CMC* method we need more than $M = 10,000$ subintervals. Even when M is as large as 10,000, we do not completely remove the downward bias.

Let us now consider the effects of different levels of the threshold values α_1 and α_2 . We have found that for all the correlation structures we have examined, the *LDMC2* estimates are closer to the true value when values of α_1 and α_2 are low. For such threshold values the estimates also converge faster as the number of time intervals M increases. On the other hand, for small α 's the computational time increases, because then the algorithm evaluates the third-order term more frequently. This observation holds true even in situations in which we have different levels of default thresholds. However, we have also noticed that as the default thresholds are moved further from their initial values of the underlying processes, the third-order term plays a less significant role, and consequently the effect of different values of α_1 and α_2 diminishes.

Figures 2 and 3 also provide insight into the performance of *LDMC2* under different values of the correlation coefficient. When $\rho = -0.2$ or 0, the method approximates well the true value already for

$M = 1$ or $M = 2$. As ρ increases and becomes positive, we need slightly larger values of M . This outcome is consistent with our earlier findings in Section 2.3.2.

Figure 4 compares the convergence of the estimates to the true value as a function of M , using *LDMC1*, *LDMC2*, and *CMC*. We can observe that *LDMC2* achieves the fastest convergence under all correlation structures. Similar observations can be made when the location of the default threshold relative to the initial value of the Brownian motion varies.

Table 2 examines the amount of time required to attain accurate results. For each simulation method specified in each row, the column M exhibits the number of subintervals needed, and the column *Time* displays the amount of time needed to evaluate the default probability using the corresponding number of subintervals given under the column M .

As discussed above, *LDMC2* results in the fastest convergence with the smallest number of subintervals than any other method, and its computational time efficiency is better than *CMC*. However, *LDMC2* is not necessarily computationally more efficient than *LDMC1*. This is attributed to the fact that *LDMC2* involves solving an optimization problem, whereas *LDMC1* requires only the operation of

Figure 4
Comparison of Estimates of *LDMC1*, *LDMC2* (Case 1), and *CMC*

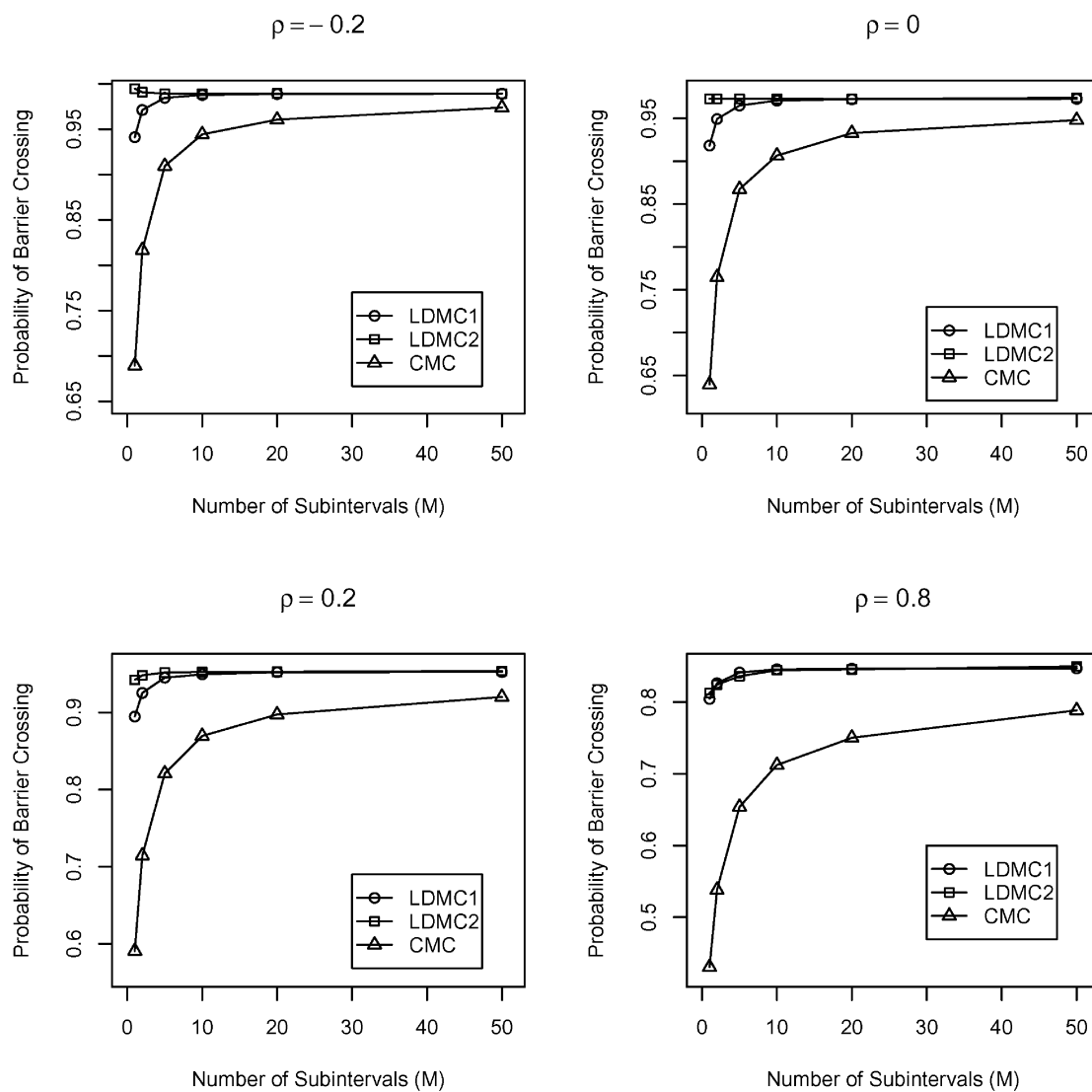


Table 2
Number of M Needed to Obtain Accurate Results with Different Simulation Methods

Simulation Type	$\rho = -0.2$		$\rho = 0$		$\rho = 0.2$		$\rho = 0.8$	
	M	Time (sec)	M	Time (sec)	M	Time (sec)	M	Time (sec)
LDMC2: Case 1	1	193.3	1	220.9	2	216.5	10	517.0
Case 2	2	59.9	5	98.6	5	149.2	10	507.2
Case 3	2	61.1	5	98.5	5	149.2	10	507.9
CMC	10,000	1,375.4	10,000	1,347	10,000	1,383.3	10,000	1,379.7
LDMC1	5	4.6	20	11.7	5	7.4	10	19.9

taking a maximum value among three closed-form solutions of marginal default probabilities. There are still other potential ways of making *LDMC2* more time efficient, such as creating a more efficient optimization routine or applying low-discrepancy sequences for generating random numbers. However, these are beyond the scope of this paper.

We have seen that *LDMC2* performs best in terms of the convergence rate as a function of M , and that *LDMC1* is the best in terms of computational efficiency for a given M . By performing additional numerical experiments, we have found that *FLDMC*, described in Section 3.2, is as good as *LDMC2* in convergence as a function of M , and as good as *LDMC1* in time efficiency for a given M . Hence, our study suggests that *FLDMC* is the best method to apply in this situation.

Figure 5 presents the estimates of *LDMC1*, *LDMC2*, and *FLDMC* under different values of the correlation coefficient as we vary the number of subintervals. From the graph we can see that the estimates from *FLDMC* are almost the same as those from *LDMC2*. In terms of convergence rate as a function of the number of subintervals, *FLDMC* outperforms *LDMC1*, except in the case of high correlation, that is, $\rho = 0.8$.

Figure 6 shows that *FLDMC* takes a similar amount of time as *LDMC1* for a given number of subintervals. From these observations we can conclude that *FLDMC* is the method of choice because it is time efficient and exhibits the fastest convergence rate as a function of M . Even at high correlation $\rho = 0.8$, this method does not perform much worse than *LDMC1*.

4.2 Pricing Credit Default Swaps

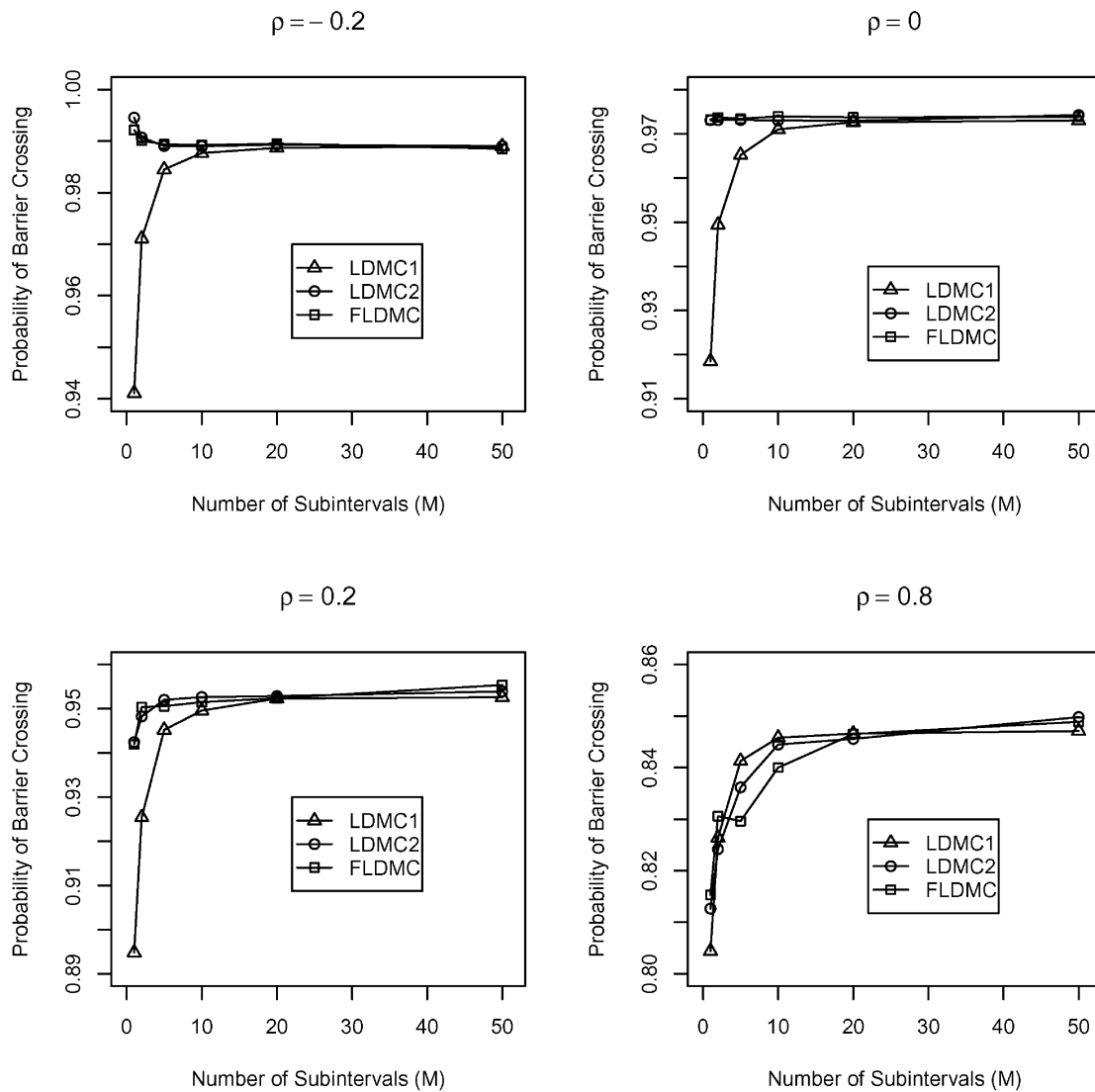
The second application of the proposed simulation techniques deals with the pricing of a credit basket derivative. In particular, we examine a credit default swap (CDS) with several reference entities, namely, three risky bonds, under the assumption of no default by the CDS issuer. In this example we consider the first-to-default swap, where the payment is triggered by the first default during a specified time horizon of T years. The protection buyer pays quarterly a certain amount of fixed premium, s , to a protection seller. In the credit event of the first default among the reference entities, the protection buyer receives a constant rebate R from the protection seller.

4.2.1 Determination of Swap Rate

In Hull and White (2001) a fair premium s is calculated by equating the expected present value of all the future cash flows of premium payments to the expected present value of the payout in the case of the credit event under a risk-neutral probability. The authors give the following simplified formula for the premium:

$$s = \frac{\int_0^T R \cdot \theta(t) \cdot v(t) dt}{\int_0^T \theta(t) \cdot u(t) dt + \pi u(T)}, \quad (4.4)$$

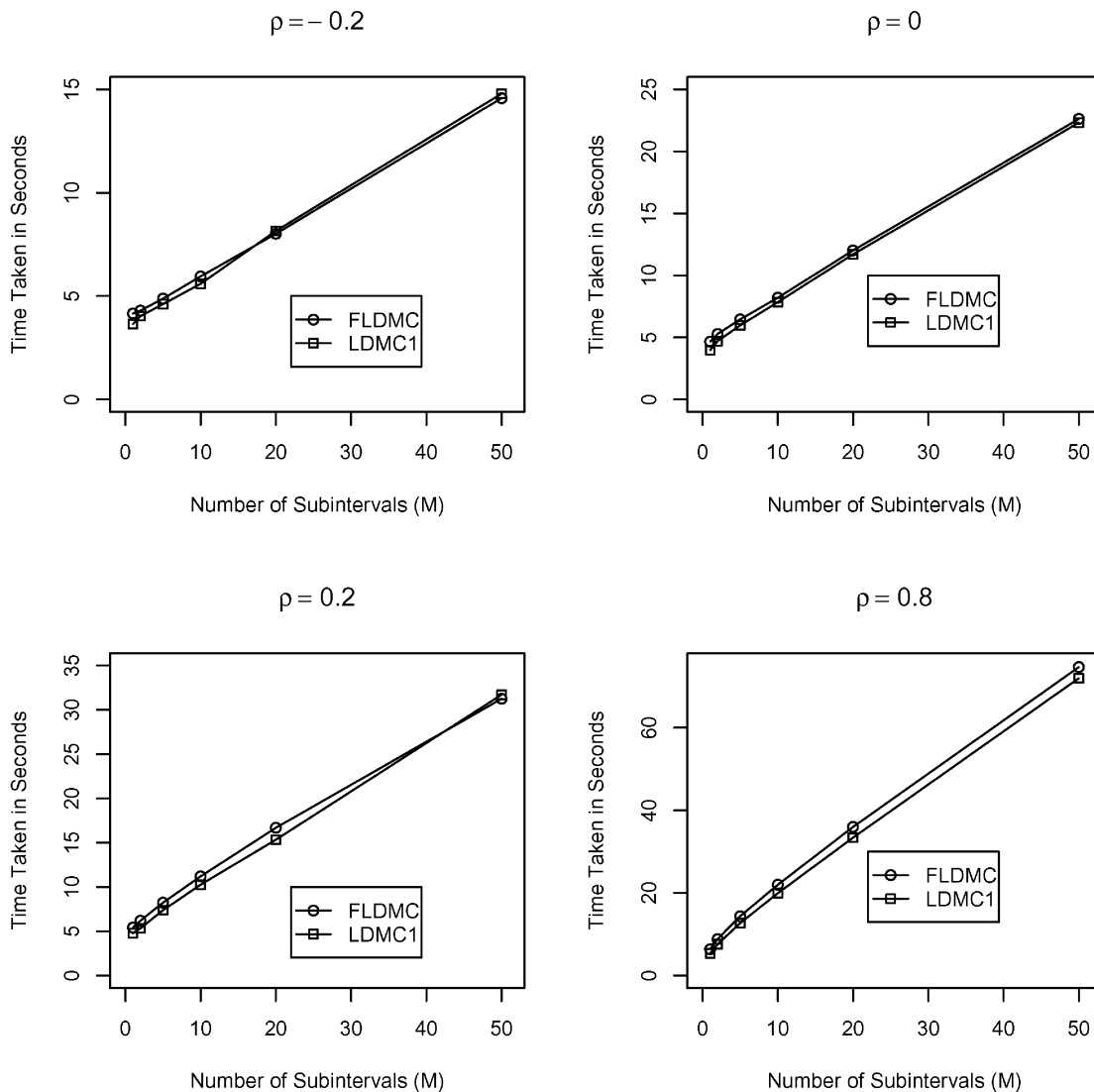
Figure 5
Comparison of Convergence Rates of LDMC1, LDMC2, and FLDMC



where $\theta(t) \cdot \Delta t$ is the risk-neutral probability of default by a reference entity between times t and $t + \Delta t$ with no earlier default, $u(t)$ is the present value of quarterly payments at the rate of \$1 per year on the CDS payment dates between time 0 and time t , and $v(t)$ is the present value of \$1 received at time t . Also, π is the probability of no default by a reference entity during the life of the credit default swap. The determination of s involves evaluating the integrals in the numerator and the denominator of (4.4).

As in Section 4.1, we assume that the i th reference entity defaults if the i th coordinate of the process (4.1) crosses a certain threshold. Now, however, we will interpret $X_j(t)$ as the creditworthiness of the j th entity at time t and assume that (4.1) is specified under a risk-neutral measure inferred from market information, such as spreads between some basic defaultable and nondefaultable instruments (Hull and White 2001; Douady and Jeanblanc 2002). As described by Hull and White (2001), such a calibrated model is consistent with the risk-neutral default probabilities and allows us to price more complicated credit instruments in a consistent manner. We should note that because of our interpretation, the

Figure 6
Comparison of Time Efficiency for *FLDMC* and *LDMC1*



process (4.1) does not have to represent a tradeable security, and hence the parameter γ is not necessarily equal to the risk-free rate of return.

Although the calibrated default barriers may be time dependent, the proposed method can still be applied. For this, we can first approximate the barriers by a piecewise linear functions and then use, for each subinterval, the method proposed in Section 2. For simplicity of exposition, however, below we outline an algorithm that evaluates the integrals in (4.4) under the assumptions that barriers are constant.

In the case when the components of the portfolio are assumed independent, we can generate the default time of each component by using the distribution of the exit time for a Brownian motion. The minimum of default times of these components will give a simulated value of the exact first-to-default time. Hence, the evaluation of the integrals in (4.4) is straightforward.

Now we formulate a procedure for evaluating the integrals in (4.4) when the underlying components are dependent.

1. With the assumption of a correlated multivariate Brownian motion, we simulate trajectories of the underlying process at the endpoints corresponding to the quarterly CDS premium payment dates.
2. For each subinterval we determine the occurrence of a credit event (first-to-default). If the terminal value of any component at the j th subinterval has crossed the barrier, then we know for certain that the credit event has already occurred. If not, then we compute the risk-neutral probability of occurrence of the credit event, that is, the probability that any one component of the process may have crossed the barrier in this subinterval. This probability can be estimated using *LDMC1* or *LDMC2* from Section 2.1, or *FLDMC* from Section 3.2. Once we obtain this probability, actual occurrence of the credit event is determined by a Bernoulli trial.
3. If we determine that a credit event has occurred in the j th subinterval, we can estimate the default time τ by generating additional trajectories of a Brownian Bridge process over a finer partition of this subinterval. The number of steps will depend on the required accuracy of the generated first-to-default time. Because in Step 2 we obtain Large Deviations estimates of default probabilities, the following approximation of the default time can also be considered. Suppose that in the reference basket the p th entity has the highest marginal probability of default. Let X_0^p and X_1^p be the values of the p th component of the underlying process at times $(j-1)\varepsilon$ and $j\varepsilon$, respectively, and ε is the size of the subinterval. According to the Large Deviations Theory, the event of the barrier crossing occurs via the most likely path of the underlying process. Hence, as the entity that defaults, we accept the one that has the highest probability. The most likely path is the length-minimizing path that joins the points X_0^p to X_1^p and touches the default threshold, B_p , of the p th component. The corresponding elapsed time τ_p at which this minimizing path touches its barrier B_p is given by

$$\tau_p = \frac{|X_0^p - B_p|}{|X_0^p - B_p| + |X_1^p - B_p|}.$$

This expression can be obtained by solving a problem of calculus of variation, where the minimizing path is composed of line segments driven at uniform speed (Baldi 1995). Intuitively, this means that the elapsed time is given by the proportion of the length of the line segment joining X_0^p and B_p to the overall length of the minimizing path. Using this expression, we can estimate the default time by $\tau = (j-1)\varepsilon + \tau_p\varepsilon$.

This procedure does not simulate default times from the exact distribution but is less computationally demanding than the previous one and for small time steps should lead to reasonably accurate approximations of the integrals in (4.4). Our simulation results confirm this intuition.

4. After finding an estimate of the default time for a given trajectory, we can obtain an estimate of the integrals in (4.4). By repeating Steps 1–3, we can approximate these integrals using corresponding sample means. The probability of no default π is estimated by the proportion of trajectories that have no occurrences of the credit event during the time horizon T .

4.2.2 Numerical Examples

Here we apply the methods *FLDMC*, *LDMC2*, and *CMC* to price a first-to-default credit default swap. In particular, the reference entity consists of three credit risky bonds issued by three distinct counterparties. The asset values of these counterparties are modeled by a multivariate correlated Brownian motion process whose parameters of the marginal processes are specified in Table 3.

For different values of the instantaneous correlation ρ between these components, we calculate the fair premium s of the credit default swap as described in Section 4.2, with $T = 3$ and $R = 100$. Because the maturity of the CDS is three years with quarterly premium payments, we take subinterval points corresponding to premium payment dates, resulting in the number of subintervals to be 12 ($M = 12$) for the simulation methods *LDMC2* and *FLDMC*. We perform 80,000 runs of simulations ($N = 80,000$). The thresholds for α_1 and α_2 , which are used in the computation of barrier crossing probability within a subinterval, are set to be 0.1 and 0.03, respectively. In Step 1 of the *FLDMC* method, we have used 200 runs for computing the ratio. In the method *CMC* we take many subintervals and assume that the probability of default between the endpoints of subintervals is 0.

Table 3
**Parameterizations of Brownian Motion for
 Underlying Process, Initial Value = 4.54**

Firm	μ	σ^2	Default Threshold
1	0.07	0.02	4.30
2	0.0325	0.035	4.27
3	0.03	0.015	4.28

Table 4
Simulation Result for Premium Value s of Credit Default Swap

Simulation Type	$\rho = -0.2$		$\rho = 0$		$\rho = 0.2$		$\rho = 0.8$	
	s	Time (sec)	s	Time (sec)	s	Time (sec)	s	Time (sec)
FLDMC	8.47 (0.01)	68	7.74 (0.02)	69	7.01 (0.01)	71	4.78 (0.02)	78
LDMC2	8.47 (0.01)	1,128	7.70 (0.01)	1,188	6.99 (0.02)	1,379	4.66 (0.02)	1,657
CMC: $M = 1,200$	8.19	516	7.46	529	6.74	525	4.57	549
$M = 8,400$	8.38	3,563	7.55	3,606	6.89	3,646	4.68	3,816
$M = 12,000$	8.43	5,057	7.71	5,128	6.91	5,194	4.67	5,432
$M = 24,000$	8.45 (0.02)	9,307	7.71 (0.03)	9,444	6.96 (0.03)	9,564	4.71 (0.02)	9,962

Table 4 summarizes the results obtained by different simulation methods. The column s shows the fair premium s , and the column *Time* shows the amount of time taken for simulation. The values in parentheses are standard errors. The estimates of *FLDMC* and *LDMC2* are very similar in value, but their computational times differ greatly. This observation reaffirms our earlier statement that *FLDMC* is to a great extent more time efficient than *LDMC2*, without losing accuracy. As a way of verifying correctness of our estimates, we used the method *CMC* with very fine subintervals. The estimates from *CMC* should converge to the true values as M increases to infinity. We observe that *FLDMC* and *LDMC2* estimates are close to these asymptotic values, confirming that our proposed simulation algorithms provide accurate approximations of the CDS prices.

5. EXTENSIONS

The simulation methodology discussed in this paper can be readily used in applications involving computation of barrier crossing probabilities of a multivariate correlated Brownian motion. Other possible applications include modeling collateralized debt obligations with several reference entities and pricing barrier-style options whose underlying is a sum of several stocks.

In this paper we assume the underlying process to be a multivariate Brownian motion process with linear barriers. By using the standard Euler's scheme approximation, the presented results can be applied to more general diffusion processes (Gobet 2000). However, as indicated by Baldi and Caracciolo (2000), an extension of Large Deviations methods to multidimensional pinned diffusion processes is more delicate, and more research is needed. The proposed method may also be applied to diffusion processes with jumps, for example, of the type proposed by Zhou (1997) in the context of structural credit risk models. For this we can divide the given time horizon into a finite number of subintervals corresponding to simulated jump times of the process. Because in each interval the process is a diffusion, we can apply the developed method. For processes with a low intensity of jumps, it is reasonable to expect that the accuracy of the method will not change significantly. For more general Lévy processes more rigorous studies are necessary, and this will be the subject of future research.

6. ACKNOWLEDGMENTS

Adam Kolkiewicz acknowledges the financial support provided by the National Science and Engineering Research Council of Canada. The authors also thank Paolo Baldi, Elias S. W. Shiu, and an anonymous referee for detailed comments that significantly improved the presentation.

APPENDICES

A. OVERVIEW OF THE LARGE DEVIATIONS THEORY

Here we provide results from the Large Deviations Theory that are used in the paper. For a comprehensive exposition of the theory, see Dembo and Zeitouni (1993) or Shwartz and Weiss (1995).

In a general formulation the Large Deviations Principle (LDP) characterizes the limiting behavior, as $\varepsilon \rightarrow 0$, of a family of probability measures $\{\mu_\varepsilon\}$ on $(\mathcal{X}, \mathcal{B})$, where \mathcal{X} is a topological space and \mathcal{B} is a σ -algebra on \mathcal{X} . It can be stated as follows.

DEFINITION A.1

We say that a sequence of measures $\{\mu_\varepsilon\}$ satisfies the Large Deviations Principle with a rate function J if for all $\Gamma \in \mathcal{B}$,

$$- \inf_{x \in \Gamma^\circ} J(x) \leq \liminf_{\varepsilon \rightarrow 0} \varepsilon \log \mu_\varepsilon(\Gamma), \quad (\text{A.1})$$

$$- \inf_{x \in \bar{\Gamma}} J(x) \geq \limsup_{\varepsilon \rightarrow 0} \varepsilon \log \mu_\varepsilon(\Gamma), \quad (\text{A.2})$$

where Γ° and $\bar{\Gamma}$ denote the interior and the closure of Γ , respectively.

If the Large Deviations Principle is applicable to a given problem, then quite often it provides reasonably accurate approximations of the probability of a given event occurring. In the paper we use the following simple consequence of the definition.

REMARK A.1

Suppose that $\{\mu_\varepsilon\}$ satisfies the Large Deviations Principle with a rate function J , and $\Gamma_1, \Gamma_2, \dots, \Gamma_N \in \mathcal{B}$. Then, for $\Gamma = \cap \Gamma_i$, we have $\Gamma \in \mathcal{B}$ and

$$- \inf_{x \in \Gamma^\circ} J(x) \leq \liminf_{\varepsilon \rightarrow 0} \varepsilon \log \mu_\varepsilon(\Gamma), \quad (\text{A.3})$$

$$- \inf_{x \in \bar{\Gamma}} J(x) \geq \limsup_{\varepsilon \rightarrow 0} \varepsilon \log \mu_\varepsilon(\Gamma). \quad (\text{A.4})$$

We now briefly explain how the Large Deviations Theory can be applied to the problem of diffusion exit from a given domain. For a detailed treatment on the subject, see Freidlin and Wentzell (1998) or Dembo and Zeitouni (1993). Suppose that we have an n -dimensional stochastic process that follows a stochastic differential equation of the form

$$dX^\varepsilon(t) = b(X^\varepsilon(t), t) dt + \sqrt{\varepsilon} \Sigma(X^\varepsilon(t), t) dW(t), \quad (\text{A.5})$$

with $X^\varepsilon(0) = \mathbf{x}$, where $\{W(t), t \geq s\}$ is an n -dimensional standard Brownian motion and ε is a small positive number. Let D be an open set in \mathbb{R}^n such that $\mathbf{x} \in D$, and $\tau = \inf\{t > 0 : X^\varepsilon(t) \notin D\}$. For $T > 0$ we are interested in finding the exit probability $\Pr_{X^\varepsilon}\{\tau < T\}$. Because typically it is difficult to find the exact value of $\Pr_{X^\varepsilon}\{\tau < T\}$, we may consider using the Large Deviations Principle to approximate it when ε is small. In this problem $\{\mu_\varepsilon\}$ in Definition A.1 is the measure \Pr_{X^ε} induced by the process $\{X^\varepsilon\}$, and the event of interest corresponds to the collection of all continuous trajectories that exit set D before time T . Then, under certain regularity conditions on the coefficients of the process $\{X^\varepsilon\}$, the Large Deviations approach leads to the following representation:

$$\Pr_{X^\varepsilon} \{ \tau \leq T \} \sim \exp \left(-\frac{u(\mathbf{x}, D)}{\varepsilon} \right), \tag{A.6}$$

where

$$u(\mathbf{x}, D) = \inf_{\gamma \in \Gamma_x(D)} J(\gamma), \tag{A.7}$$

$$\Gamma_x(D) = \{ \gamma \in \Gamma^n : \gamma(0) = \mathbf{x}, \gamma(t) \in D^c \text{ for some } t \in (0, T) \}. \tag{A.8}$$

The asymptotic in (A.6) is understood in the sense that the ratio of the logarithms of the two terms tends to 1 as $\varepsilon \rightarrow 0$. $J(\gamma)$ in equation (A.7) corresponds to the rate function in Definition (A.1) and in the context of the above problem is often referred to as the action functional (Freidlin and Wentzell 1998).⁴ It is defined as follows:

$$J(\gamma) = \frac{1}{2} \int_0^T \langle A^{-1}(\gamma(t), t) \left(\frac{d\gamma}{dt}(t) - b(\gamma(t), t) \right), \frac{d\gamma}{dt}(t) - b(\gamma(t), t) \rangle dt, \tag{A.9}$$

$$A(\gamma(t), t) = \Sigma(\gamma(t), t) \Sigma'(\gamma(t), t). \tag{A.10}$$

For the Brownian Bridge process (2.4), and \mathbf{x} and D that satisfy certain regularity conditions, more precise approximations are possible. By applying an expansion of the exit probability derived by Fleming and James (1992), Baldi (1995) has obtained the following representation:

$$\Pr_{X^\varepsilon} \{ \tau \leq T \} = \exp \left(-\frac{u(\mathbf{x}, D)}{\varepsilon} \right) \exp(-\varpi) (1 + \psi_1 \varepsilon + \dots + \psi_m \varepsilon^m + o(\varepsilon^m)), \quad \text{as } \varepsilon \rightarrow 0,$$

where ϖ and $\psi_k, k = 1, \dots, m$ are functions that solve a system of partial differential equations with boundary conditions depending on the boundary of the region D . By using an argument similar to the one presented by Baldi (1995, Example 4.5), it can be shown that for the class of sets D defined in Section 2.1 and for almost all $(\mathbf{x}, \mathbf{y}) \in D$, the coefficients ϖ and ψ_k vanish, leading to the approximation

$$\Pr_{X^\varepsilon} \{ \tau \leq T \} = \exp \left(-\frac{u(\mathbf{x}, D)}{\varepsilon} \right) (1 + o(\varepsilon^m)), \quad \text{as } \varepsilon \rightarrow 0, \tag{A.11}$$

for every $m > 0$.

If the minimizing path γ in (A.7) is unique, then $\Pr_{X^\varepsilon} \{ \tau \leq T \}$ is asymptotically the same as $\Pr_{X^\varepsilon} \{ \tau \leq T, X^\varepsilon(t) \in B_\delta(\gamma(t)) \text{ for } 0 \leq t \leq T \}$, where $B_\delta(\gamma(t))$ denotes the neighborhood (a tube) of radius δ of $\gamma(t)$. This implies that the barrier hitting probability is asymptotically determined only by the portion of the barrier that this minimizing path goes through.

B. PROOF OF PROPOSITION 2.1

Let us define two lines $l_i = \{ [b_1, b_2]' + \alpha \mathbf{e}_i : \alpha \in \mathbb{R}^1 \}, i = 1, 2$, where $\{ \mathbf{e}_1, \mathbf{e}_2 \}$ is the standard basis for \mathbb{R}^2 , that is, $[\mathbf{e}_1, \mathbf{e}_2] = \mathbf{I}$, where \mathbf{I} is a 2×2 identity matrix. Then for the half-lines $l_i^+ = \{ [b_1, b_2]' + \alpha \mathbf{e}_i : \alpha \geq 0 \}, i = 1, 2$, the set $l_1^+ \cup l_2^+$ represents the boundary of D_{b_1, b_2} . By linearity of the transformation by the matrix \mathbf{K}' , it follows that the image of D_{b_1, b_2} , given by $\mathbf{K}'D_{b_1, b_2}$, is a convex set whose boundary is $\mathbf{K}'l_1^+ \cup \mathbf{K}'l_2^+$.

In the first part of the proof we show that

$$u(\mathbf{K}'\mathbf{x}, \mathbf{K}'\mathbf{y}, \mathbf{K}'D_{b_1, b_2}) = 2 \min \{ d(\mathbf{K}'\mathbf{x}, \mathbf{K}'l_1) d(\mathbf{K}'\mathbf{y}, \mathbf{K}'l_1), d(\mathbf{K}'\mathbf{x}, \mathbf{K}'l_2) d(\mathbf{K}'\mathbf{y}, \mathbf{K}'l_2) \}, \tag{B.1}$$

where $d(\mathbf{z}, l)$ denotes the distance between a point \mathbf{z} and a line l . In the second part, we establish a connection between the right-hand side of (B.1) and the exit probability of a one-dimensional Brownian Bridge.

⁴ Dembo and Zeitouni (1993) use the name "good rate function."

For $r \geq 0$, let us define $O(r) = \{z \in \mathbb{R}^2 : (\|K'x - z\| + \|K'y - z\|)^2 - \|K'x - K'y\|^2 \leq r\}$, which represents the set of all points whose total distance to $K'x$ and $K'y$ is less than or equal to $\|K'x - K'y\|^2 + r$. Because $O(0)$ is equal to the line segment that joins the points $K'x$ and $K'y$, the following number is well defined and finite:

$$r^* = \sup\{r : O(r) \cap (K'l_1^+ \cup K'l_2^+) = \emptyset\},$$

where \emptyset represents an empty set. It is easy to verify that for each $r > 0$, the set $O(r)$ is convex and has a tangent line at each boundary point. These properties and the definition of r^* imply that $O(r^*) \subseteq K'D_{b_1, b_2}$, and at each point $z^* \in O(r^*) \cap (K'l_1^+ \cup K'l_2^+)$ the tangent line is equal to either $K'l_1^+$ or $K'l_2^+$. The latter property implies that the vertex $K'(b_1, b_2)'$ does not belong to $O(r^*) \cap (K'l_1^+ \cup K'l_2^+)$. From the definition of r^* and (2.10), it also follows that $r^* = 2u(K'x, K'y, K'D_{b_1, b_2})$.

Suppose that $z^* \in O(r^*) \cap (K'l_1^+ \cup K'l_2^+)$, and that the tangent line to $O(r^*)$ at z^* is the first line, $K'l_1$. Then, by the definition of r^* we have

$$r^* = \inf_{\phi \in \partial K'D_{b_1, b_2}} \{(\|K'x - \phi\| + \|K'y - \phi\|)^2 - \|K'x - K'y\|^2\} \tag{B.2}$$

$$= \inf_{\phi \in K'l_1} \{(\|K'x - \phi\| + \|K'y - \phi\|)^2 - \|K'x - K'y\|^2\}, \tag{B.3}$$

and the infimum is attained at $\phi = z^*$. Equation (B.3) follows from the fact that the vertex $K'(b_1, b_2)'$ does not belong to $O(r) \cap (K'l_1^+ \cup K'l_2^+)$, because it implies that z^* must belong to the interior of l_1^+ .

To show that (B.1) follows from (B.3), we need some additional notation. For a point z in \mathbb{R}^2 and a line l , we denote by $P_l(z)$ the projection of z onto l . By some basic properties of a projection, we have $d(z, l) = \|P_l(z) - z\|$. For a point z we also define its reflection around the line l as the point $ref_l(z)$ for which the line segment that joins z and $ref_l(z)$ is orthogonal to l and $d(ref_l(z), l) = d(z, l)$.

We first observe that z^* at which the infimum in (B.3) is attained can be characterized as the point on $K'l_1$ at which the line segment that joins $ref_{K'l_1}(K'x)$ and $K'y$ crosses $K'l_1$. This follows from the fact that the shortest distance between two points corresponds to a linear segment. Then, by applying the Pythagorean Formula twice, we obtain

$$\begin{aligned} r^* &= (\|K'x - z^*\| + \|K'y - z^*\|)^2 - \|K'x - K'y\|^2 = \|K'y - ref_{K'l_1}(K'x)\|^2 - \|K'x - K'y\|^2 \\ &= \|P_{K'l_1}(K'x) - P_{K'l_1}(K'y)\|^2 + (\|P_{K'l_1}(K'x) - K'x\| + \|P_{K'l_1}(K'y) - K'y\|)^2 - \|K'x - K'y\|^2 \\ &= (\|P_{K'l_1}(K'x) - K'x\| + \|P_{K'l_1}(K'y) - K'y\|)^2 - (\|P_{K'l_1}(K'x) - K'x\| - \|P_{K'l_1}(K'y) - K'y\|)^2 \\ &= 4\|P_{K'l_1}(K'x) - K'x\| \cdot \|P_{K'l_1}(K'y) - K'y\|, \end{aligned}$$

which together with the relation $r^* = 2u(K'x, K'y, K'D_{b_1, b_2})$ proves (B.1).

Now we will show that (B.1) implies the result stated in Proposition 2.1. Because the linear transformation by K' is a one-to-one mapping, the probability that the i th coordinate of the process $\{X^\epsilon(t)\}$ reaches the line l_i , $i = 1, 2$, is the same as the probability that the transformed process $\{K'X^\epsilon(t)\}$ reaches the line $K'l_i$. Suppose that we want to find the probability that $\{K'X^\epsilon(t)\}$ reaches $K'l_1$. For this let us introduce new orthogonal coordinates such that the first one is parallel to l_1 . Let $\{c_1, c_2\}$ be the orthonormal basis corresponding to these coordinates. Then the i th column in $C := [c_1, c_2]$ gives the coordinates of c_i in the standard basis. Let us denote the process $\{K'X^\epsilon(t)\}$ in these new coordinates by $\{Y(t)\} = \{[Y_1(t), Y_2(t)]'\}$. From elementary linear algebra we have the following relation: $\{Y(t)\} = \{C^{-1}K'X^\epsilon(t)\}$. Because the vectors $\{c_1, c_2\}$ are orthonormal, we have $C^{-1} = C'$. From this it follows that $\{K'X^\epsilon(t)\}$ in these new coordinates is a two-dimensional Brownian Bridge with the covariance matrix given by ϵI . Therefore, by (2.1) the probability that $\{K'X^\epsilon(t)\}$ reaches the line $K'l_1$ is given by

$$\exp\{-2d(K'x, K'l_1) d(K'y, K'l_1)/\epsilon\}.$$

A similar result can be obtained for the second line l_2 . This, together with (2.9) and (B.1), proves the result.

C. SIMPLER REPRESENTATION OF THE OPTIMIZATION PROBLEM IN SECTION 2.2

Here we show that u^* defined as

$$u^* = u^*(\mathbf{Kx}, \mathbf{Ky}, F_1, F_2) = \inf_{\eta \in \Gamma_{\mathbf{Kx}(F_1, F_2)}^*} J_{\mathbf{Ky}}(\eta)$$

can be obtained by solving a finite dimensional convex optimization problem. For notational convenience, we represent \mathbf{Kx} as $\tilde{\mathbf{x}}$ and \mathbf{Ky} as $\tilde{\mathbf{y}}$. Let $t_i, i = 1, 2$, denote the argument at which the minimizing path η reaches the boundary of F_i , and $\phi_i = \eta(t_i)$.

We will consider three cases depending on the relative position of t_1 and t_2 . Assume first that $t_1 < t_2$. We want to minimize $J_{\tilde{\mathbf{y}}}(\eta)$ defined in (2.8) with respect to η . By selecting $\eta(t)$ on the interval $(t_2, 1]$ in such a way that

$$\frac{d\eta}{dt}(t) = -\frac{\eta(t) - \tilde{\mathbf{y}}}{1 - t}, \quad \eta(t_2) = \phi_2,$$

we obtain the integrand $\|d\eta/dt(t) + (\eta(t) - \tilde{\mathbf{y}})/(1 - t)\|^2$ to be zero on $(t_2, 1]$. This leads to the following representation:

$$u^* = \inf_{\eta \in \Gamma_{\tilde{\mathbf{x}}(F_1, F_2)}^*} \frac{1}{2} \left\{ \int_0^{t_1} \left\| \frac{d\eta}{dt}(t) + \frac{\eta(t) - \tilde{\mathbf{y}}}{1 - t} \right\|^2 dt + \int_{t_1}^{t_2} \left\| \frac{d\eta}{dt}(t) + \frac{\eta(t) - \tilde{\mathbf{y}}}{1 - t} \right\|^2 dt \right\} \tag{C.1}$$

$$= \inf_{t_1, t_2, \phi_1, \phi_2} \left\{ \frac{1}{2} \left[\frac{\|\tilde{\mathbf{x}} - \phi_1\|^2}{t_1} + \frac{\|\tilde{\mathbf{y}} - \phi_1\|^2}{1 - t_1} - \|\tilde{\mathbf{x}} - \tilde{\mathbf{y}}\|^2 \right] + \frac{1}{2} \left[\frac{\|\phi_1 - \phi_2\|^2}{t_2 - t_1} + \frac{\|\tilde{\mathbf{y}} - \phi_2\|^2}{1 - t_2} - \frac{\|\phi_1 - \tilde{\mathbf{y}}\|^2}{1 - t_1} \right] \right\} \tag{C.2}$$

$$= \inf_{t_1, t_2, \phi_1, \phi_2} \frac{1}{2} \left\{ \frac{\|\tilde{\mathbf{x}} - \phi_1\|^2}{t_1} + \frac{\|\phi_1 - \phi_2\|^2}{t_2 - t_1} + \frac{\|\tilde{\mathbf{y}} - \phi_2\|^2}{1 - t_2} - \|\tilde{\mathbf{x}} - \tilde{\mathbf{y}}\|^2 \right\}. \tag{C.3}$$

The minimization problem in (C.1) is a problem of calculus of variations with fixed end points, for which it is known that the infimum is attained by a function composed of line segments joining the end points (Baldi 1995). This justifies the expression in (C.2). By minimizing (C.3) with respect to t_1 and t_2 , we obtain the optimal values:

$$t_1^* = \frac{\|\tilde{\mathbf{x}} - \phi_1\|}{\|\tilde{\mathbf{x}} - \phi_1\| + \|\phi_1 - \phi_2\| + \|\tilde{\mathbf{y}} - \phi_2\|}, \quad t_2^* = \frac{\|\phi_1 - \phi_2\| + \|\tilde{\mathbf{x}} - \phi_1\|}{\|\tilde{\mathbf{x}} - \phi_1\| + \|\phi_1 - \phi_2\| + \|\tilde{\mathbf{y}} - \phi_2\|}.$$

By substituting t_1^* and t_2^* into (C.3), we arrive at the following convex minimization problem:

$$u^* = \inf_{\phi_1, \phi_2} \frac{1}{2} \{ (\|\tilde{\mathbf{x}} - \phi_1\| + \|\phi_1 - \phi_2\| + \|\tilde{\mathbf{y}} - \phi_2\|)^2 - \|\tilde{\mathbf{x}} - \tilde{\mathbf{y}}\|^2 \}. \tag{C.4}$$

A similar approach leads to the representation

$$u^* = \inf_{\phi_1, \phi_2} \frac{1}{2} \{ (\|\tilde{\mathbf{x}} - \phi_2\| + \|\phi_2 - \phi_1\| + \|\tilde{\mathbf{y}} - \phi_1\|)^2 - \|\tilde{\mathbf{x}} - \tilde{\mathbf{y}}\|^2 \}, \tag{C.5}$$

when $t_1 \geq t_2$. The value of u^* that is used in the Large Deviations approximation is equal to the minimum of (C.4) and (C.5).

D. AN ALTERNATIVE ALGORITHM FOR OPTIMIZATION

We describe an algorithm that provides an approximation of the optimal solution for the problem (2.19). Let ∂F_i denote the boundary of $F_i, i = 1, \dots, n$, and ϕ_0 and ϕ_{n+1} denote the initial value \mathbf{Kx}

and the terminal value Ky , respectively. After initializing values of $\phi_1, \phi_2, \dots, \phi_n$, we perform the following procedure:

1. For each i , where $i = 1, \dots, n$, we fix the values of ϕ_{i-1} and ϕ_{i+1} and then find $\phi_i^* \in \partial F_i$ such that the length of the path $\phi_{i-1} \rightarrow \phi_i^* \rightarrow \phi_{i+1}$ is minimized. We update $\phi_i = \phi_i^*$.
2. We repeat Step 1 until certain convergence criteria are reached, or the number of iterations reach a prespecified value.

This approach performs a piecewise optimization, and hence it does not guarantee yielding the global optima. However, we have found in our simulation studies that most of the time it gives values that are quite close to the global optima. In terms of time efficiency, this algorithm is much faster than solving the convex optimization problem using standard libraries, such as the MATLAB function “fmincon.”

REFERENCES

- BALDI, P. 1995. Exact Asymptotics for the Probability of Exit from a Domain and Application to Simulation. *Annals of Probability* 23(4): 1644–70.
- BALDI, P., AND L. CARAMELLINO. 2000. Asymptotics of Hitting Probabilities for General One-Dimensional Pinned Diffusions. *Annals of Applied Probability* 12: 1071–95.
- BALDI, P., L. CARAMELLINO, AND M. G. IOVINO. 1999. Pricing General Barrier Options: A Numerical Approach Using Sharp Large Deviations. *Mathematical Finance* 9(4): 293–321.
- BILLINGSLEY, P. 1999. *Convergence of Probability Measures*. 2nd edition. New York: Wiley.
- BLACK, F., AND J. C. COX. 1976. Valuing Corporate Securities: Some Effects of Bond Indenture Provisions. *Journal of Finance* 31: 351–67.
- BOYLE, P. P., S. FENG, AND W. TIAN. 2008. Large Deviation Techniques and Financial Applications. In *Handbooks in Operations Research and Management Science: Financial Engineering*, edited by J. Birge and V. Linetsky. Amsterdam: Elsevier.
- BOYLE, P. P., S. FENG, W. TIAN, AND T. WANG. 2008. Robust Stochastic Discount Factors. *Review of Financial Studies*, forthcoming.
- CHEN, C. J., AND H. PANJER. 2003. Unifying Discrete Structural Models and Reduced-Form Models in Credit Risk Using a Jump-Diffusion Process. *Insurance: Mathematics and Economics* 33: 357–80.
- DEMBO, A., AND O. ZEITOUNI. 1993. *Large Deviations Techniques and Applications*. New York: Springer.
- DOUADY, R., AND M. JEANBLANC. 2002. A Rating Based Model for Credit Derivatives. *European Investment Review* 1: 17–29.
- DUFFIE, D., AND K. J. SINGLETON. 2003. *Credit Risk: Pricing, Measurement, and Management*. Princeton: Princeton University Press.
- FLEMING, W. H., AND M. R. JAMES. 1992. Asymptotic Series and Exit Time Probabilities. *Annals of Probability* 20(3): 1369–84.
- FREIDLIN, M. I., AND A. D. WENTZELL. 1998. *Random Perturbations of Dynamical Systems*. 2nd edition. New York: Springer.
- GOBET, E. 2000. Weak Approximation of Killed Diffusion Using Euler Schemes. *Stochastic Processes and Their Applications* 87: 167–97.
- HE, H., W. P. KEIRSTEAD, AND J. REBHOLZ. 1998. Double Lookbacks. *Mathematical Finance* 8(3): 201–28.
- HUH, J. 2007. Computation of Multivariate Barrier Crossing Probability, and Its Applications in Finance. Ph.D. thesis, University of Waterloo.
- HULL, J., AND A. WHITE. 2001. Valuing Credit Default Swaps II: Modeling Default Correlations. *Journal of Derivatives* 8(3): 12–31.
- KARATZAS, I., AND S. E. SHREVE. 1991. *Brownian Motion and Stochastic Calculus* 2nd edition. New York: Springer.
- KARLIN, S., AND H. M. TAYLOR. 1981. *A Second Course in Stochastic Process*. New York: Academic Press.
- KLOEDEN, P., AND E. PLATEN. 1992. *Numerical Solution of Stochastic Differential Equations*. New York: Springer.
- LANDO, D. 2004. *Credit Risk Models: Theory and Applications*. Princeton: Princeton University Press.
- LELAND, H. E. 1994. Corporate Debt Value, Bond Covenants and Optimal Capital Structure. *Journal of Finance* 49: 1213–52.
- LELAND, H. E., AND K. B. TOFT. 1996. Optimal Capital Structure, Endogenous Bankruptcy and the Term Structure of Credit Spreads. *Journal of Finance* 51: 987–1019.
- LI, D. X. 1999. The Valuation of Basket Credit Derivatives. *CreditMetrics Monitor* (April): 34–50.
- . 2000. On Default Correlation: A Copula Approach. *Journal of Fixed Income* 9(4): 43–50.
- . 2006. It’s All about Credit. *North American Actuarial Journal* 10(1): iii–iv.
- LONGSTAFF, F. A., AND E. S. SCHWARTZ. 1995. A Simple Approach to Credit Risk of Large Corporate Bond and Loan Portfolio. *Journal of Finance* 50(3): 789–819.
- MCDONALD, R. L. 2006. *Derivatives Markets*. 2nd edition. Boston: Addison-Wesley.
- MEDOVA, E., AND R. SMITH. 2006. A Structural Approach to EDS Pricing. *Risk* 19(4): 84–88.
- MERTON, R. C. 1974. On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journal of Finance* 29: 449–70.
- NELSEN, R. B. 1999. *An Introduction to Copulas*. New York: Springer.

- PHAM, H. 2007. Some Applications and Methods of Large Deviations in Finance and Insurance. In *Paris-Princeton Lectures on Mathematical Finance 2004*, edited by Rene A. Carmona et al., pp. 191–244. Berlin: Springer.
- SCHÖNBUCHER, P. J. 2003. *Credit Derivatives Pricing Models: Models, Pricing and Implementation*. Chichester: Wiley.
- SHEVCHENKO, P. V. 2003. Addressing the Bias in Monte Carlo Pricing for Multi-asset Options with Multiple Barriers through Discrete Sampling. *Journal of Computational Finance* 6(3): 1–20.
- SHWARTZ, A., AND A. WEISS. 1995. *Large Deviations for Performance Analysis*. London: Chapman & Hall.
- ZHOU, C. 1997. *A Jump-Diffusion Approach to Modelling Credit Risk and Valuing Defaultable Securities*. Washington, DC: Federal Reserve Board.
- . 2001. An Analysis of Default Correlations and Multiple Defaults. *Review of Financial Studies* 14(2): 555–76.

Discussions on this paper can be submitted until January 1, 2009. The authors reserve the right to reply to any discussion. Please see the Submission Guidelines for Authors on the inside back cover for instructions on the submission of discussions.