Efficient Importance Sampling for the First Passage Time Problem

Chuan-Hsiang Han *

Yen-An Chen †

Abstract

Motivated from the first passage model for default in credit risk, this paper extends the two-dimensional problem studied by Zhou (2001) to any finite dimension by means of Monte Carlo simulation. We derive a variational problem to characterize the importance sampling scheme for estimating joint default probabilities. The large deviation principle is applied to prove that the proposed importance sampling is asymptotic optimal.

Keywords: first passage time problem, importance sampling, variational problem, asymptotic optimality, large deviation principle

1 Introduction

Estimation of the joint default probability under a structural-form model emerged pretty early in the presence of stochastic financial theory. In models of Black and Scholes [2] and Merton [10], a firm’s default time can only happen at expiration of its debt when an issuer’s asset value is less than the debt value. Black and Scholes modeled the asset value process by a geometric Brownian motion, then Merton incorporated an additional compound

*Corresponding Author. Department of Quantitative Finance, National Tsing Hua University, Hsinchu, Taiwan, 30013, ROC. Email: chhan@mx.nthu.edu.tw. Work supported by MOST 104-2115-M-007-009.

†Department of Mathematics, National Taiwan University, Taipei, Taiwan, 10617, ROC. Email: r03246004@ntu.edu.tw.
Poisson jump term. Black and Cox [3] generalized these models by allowing that default can occur at any time before the expiration of debt. They considered a first passage time problem for the geometric Brownian motion in one dimension. The joint default for two-dimensional geometric Brownian motions was treated by Zhou [11]. A comprehensive technical review with financial applications can be found in [4].

In this paper, we focus on generalizing the joint default from a two-dimensional first passage time problem studied in [11] to any finite dimensions through the importance sampling analyzed by the large deviation principle. A high-dimensional setup of the first passage time problem under correlated Brownian motions is the following. We assume that each firm value process \( S_it \), \( 1 \leq i \leq n \), is governed by the following dynamics,

\[
dS_{it} = \alpha_i(t)dt + \sum_{j=1}^{n} \beta_{ij}(t)dW_{jt}, \tag{1}
\]

where \( \alpha_i(t) \) denotes the drift term and \( \beta_{ij}(t) \) denotes the volatility of its associated Brownian motion \( W_{jt} \). All Brownian motions \( W_{j}, j = 1, \cdots, n \) are mutually independent. Each firm also has a constant default barrier \( B_i \) for \( 1 \leq i \leq n \), and its default happens at the first time when the asset value \( S_{it} \) falls below its barrier level \( B_i \). Therefore, default time \( \tau_i \) of the \( i \)-th firm is defined by

\[
\tau_i = \inf\{t \geq 0 : S_{it} \leq B_i\}. \tag{2}
\]

Assume that \( \mathcal{F}_t \) is the natural filtration generated from all \( S_{it}, i = 1, \cdots, n \), under a probability space \( (\Omega, \mathcal{F}, P) \). The joint default probability with a terminal time \( T \) is defined by

\[
DP = E \{ \prod_{i=1}^{n} I(\tau_i \leq T) | \mathcal{F}_0 \}. \tag{3}
\]

Let the notation \( ' \) (prime) denote the transpose of a vector. When the debt level \( B = (B_1, \cdots, B_n)' \) of a company is much smaller than its initial asset value \( S_0 = (S_{10}, \cdots, S_{n0})' \), the joint default event is rare.

In general, there is no closed-form solution for such joint default probability (3) in high dimension, so one has to rely on numerical methods. Monte Carlo simulation is employed to overcome the curse of dimensionality caused by deterministic approaches such as numerical partial differential equation, the binomial tree method, etc. Moreover, it is crucial to reduce the variance of Monte Carlo estimators for convergence improvement. Carmona, Fouque, and Vestal [5] studied a first passage time problem and estimate the loss density function for a credit portfolio under a stochastic volatility model. They used interacting particle systems for variance reduction. Al-
ternatively, we propose an efficient Monte Carlo method which incorporates importance sampling for variance reduction in order to accurately estimate joint default probabilities. Numerical results are quite remarkable and it is easy to implement. In particular, the Black-Cox model in high dimension is treated.

There are many ways to analyze the variance of an importance sampling estimator. In a strong sense, one can minimize the variance, say over a parametrized space, by solving relevant optimization problems. See for example [1], in which authors proposed an adaptive scheme, namely Robbins Monro algorithms, to solve for a class of importance sampling estimators. They utilized constant change of drift in high dimension and solve nontrivial optimization problems. See also [7] for using a singular and regular perturbation method to approximate the optimal change of measure.

On the other hand, minimizing asymptotically the decay rate of variance, instead of the variance itself, of an importance sampling estimator is considered in a weak sense. We adopt this asymptotic variance analysis by means of the large deviation principle. We prove that the proposed importance sampling estimator for the high-dimensional first passage time problem has a zero variance rate. That is, our proposed importance sampling scheme is asymptotically optimal. Numerical results of our intensive experiments indicate the proposed importance sampling is robust and efficient.

The organization of this paper is as follows: Section 2 provides the derivation of an innovative importance sampling for the probability estimation of the first passage time model in high dimension. In Section 3 we prove that the proposed importance sampling method is asymptotically optimal (or called efficient) by an application of Freidlin-Wentzell Theorem. In Section 4, numerical results are presented. Then we conclude in Section 5.

2 Importance Sampling for High-Dimensional First Passage Time Problem

In this section, we propose an importance sampling scheme in order to improve the convergence of Monte Carlo simulation. The basic Monte Carlo simulation approximates the joint default probability defined in (3) by the following estimator

\[ DP \approx \frac{1}{N} \sum_{i=1}^{n} I(\tau^{(k)}_i \leq T), \]  

(4)
where $\tau^{(k)}_i$ denotes the $k$-th i.i.d. sample of the $i$-th default time defined in (2) and $N$ denotes the total number of simulation. By Girsanov Theorem, one can construct an equivalent probability measure $\tilde{P}$ defined by the following Radon-Nikodym derivative
\[
\frac{dP}{d\tilde{P}} = Q_T(h.) = \exp\left( \int_0^T h(s, S_s)d\tilde{W}_s - \frac{1}{2} \int_0^T \|h(s, S_s)\|^2 ds \right),
\]
where we denote $S_s = (S^1_s, \ldots, S^n_s)$ the vector of state variables (asset value processes) and $\tilde{W}_s = (\tilde{W}^1_s, \ldots, \tilde{W}^n_s)$ the vector of standard Brownian motions. The function $h(s, S_s)$ is assumed to satisfy Novikov’s condition such that $\tilde{W}_t = W_t + \int_0^t h(s, S_s)ds$ is a vector of Brownian motions under $\tilde{P}$. Thus, the joint default probability defined in (3) becomes
\[
DP = \tilde{E}\{\Pi_{i=1}^n I(\tau_i \leq T)Q_T(h)|\mathcal{F}_0\}. \tag{6}
\]

Based on the dynamics (1), we illustrate two case studies for the construction of the drift change vector function $h$.

**Case 1: Correlated Brownian motions.** For $1 \leq i,j \leq n$, $\alpha_i(t) = 0$, $\beta_{ij}(t) = \nu_{ij} \geq 0$, and $\nu = (\nu_{ij})_{1 \leq i,j \leq n}$, the drift change is
\[
h(t) = \nu^{-1} \check{f}^*(t), \tag{7}
\]
where $a = \nu \cdot \nu^T$, $\check{B} = \sqrt{\nu}B$, $\frac{1}{\sqrt{\nu}} = \max_{i \in 1\ldots n} \{|B_i|\}$, and
\[
\check{f}^*(t) := \{f(t) \in H_1 : \inf_{0 \leq t \leq T} f(t) < \check{B}\} \frac{1}{2} \int_0^T \check{f}(t) a^{-1} \check{f}(t) dt, \tag{8}
\]
where $H_1$ is a continuous function space defined on $[0,T]$ with the initial value being 0.

**Case 2: Geometric Brownian motions.** For $1 \leq i,j \leq n$, $\alpha_i(t) = S_i\mu_i$, $\beta_{ij}(t) = S_i\nu_{ij}$, the drift change is
\[
h = \frac{1}{\sqrt{\nu}} \nu^{-1} \check{f}^*(t), \tag{9}
\]
where $a = \nu \cdot \nu^T$, $\check{B} = \sqrt{\nu}(\ln \frac{B_1}{S_10}, \ldots, \ln \frac{B_n}{S_n0})$, $\frac{1}{\sqrt{\nu}} = \max_{i \in 1\ldots n} \{\ln \frac{B_i}{S_i0}\}$, and
\[
\check{f}^*(t) := \{f(t) \in H_1 : \inf_{0 \leq t \leq T} f(t) < \check{B}\} \frac{1}{2} \int_0^T \check{f}(t) a^{-1} \check{f}(t) dt. \tag{10}
\]
3 Asymptotic Variance Analysis by Large Deviation Principle

We provide theoretical verification to prove that two importance sampling methods developed above are asymptotic optimal. Our proofs are based on the Freidlin-Wentzell theorem \(6\) in large deviation theory in order to approximate the default probability and the second moment of the importance sampling estimator defined in (6). A scaling factor \(\varepsilon > 0\) is considered to create a rare event so that the current asset value \(S_0\) is much larger than its debt value \(B\), i.e. \(0 < B < S_0\) in the vector convention. Our asymptotic results show that the second moment approximation is the square of the first moment (or default probability) approximation. Therefore, the optimality of variance reduction in an asymptotic sense is obtained so that the proposed importance sampling scheme is efficient.

Theorem 1. (Efficient Importance Sampling for Correlated Brownian Motions) Let \(S_t = (S_{1t}, S_{2t}, \ldots, S_{nt})'\) denote the system of \(n\) assets value following the process \(dS_t = \nu dW_t\) with the standard Brownian motion \(W_t = (W_{1t}, W_{2t}, \ldots, W_{nt})'\), the volatility matrix \(\nu = \{\nu_{i,j}\}_{i,j=1\ldots n}\), the initial value \(S_0 = (S_{10}, S_{20}, \ldots, S_{n0})'\), and the default boundary \(B = (B_1, B_2, \ldots, B_n)'\) denote the default boundary. We introduce a scale \(\varepsilon\) defined from \(\frac{1}{\sqrt{\varepsilon}} = \max_{i=1\ldots n}\{\|B_i\|\}\) and \(B = \sqrt{\varepsilon}B\). The default probability estimated by the importance sampling is defined by

\[
P_1^\varepsilon = E \left\{ \prod_{i=1}^n I \left( \inf_{0 \leq t \leq T} S_{it} \leq B_i \right) \right\}
\]

\[
= E \left\{ \prod_{i=1}^n I \left( \inf_{0 \leq t \leq T} S_{it} \leq B_i \right) Q_T(h) \right\},
\]

where the Radon-Nykodym derivative \(Q(h)\) is defined in (5). The second moment of this estimator is denoted by

\[
P_2^\varepsilon = E \left\{ \prod_{i=1}^n I \left( \inf_{0 \leq t \leq T} S_{it} \leq B_i \right) Q_T^2(h) \right\},
\]

by the choice \(h = \nu^{-1}f^*(t)\), where \(a = \nu \cdot \nu^T\) and

\[
f^*(t) := \arg \inf_{f(t) \in H_1} \left\{ \inf_{0 \leq t \leq T} f(t) < B \right\} \left\{ \frac{1}{2} \int_0^T f(t)^T a^{-1} f(t) dt \right\}.
\]
When \( \varepsilon \) is small enough, we obtain a zero variance rate, i.e.,

\[
\lim_{\varepsilon \to 0} \varepsilon^2 \ln \left( \frac{P_2^\varepsilon(h)}{(P_1^\varepsilon)^2} \right) = 0,
\]

so that the importance sampling scheme is asymptotically optimal (efficient).

**Proof:** Recall that the default probability is defined by

\[
P_1 = P(\inf_{0 \leq t \leq T} S_t < B) := P(\bigcap_{i=1}^{n} \{\inf_{0 \leq t \leq T} S_{t_i} < B_i\})
\]

\[
= P(\inf_{0 \leq t \leq T} \sqrt{\varepsilon} \nu W_t < \bar{B})
\]

By an application of Freidlin-Wentzell Theorem [6],

\[
\lim_{\varepsilon \to 0} \varepsilon \log P_1 = -\inf_{\{f(t) \in H_1; \inf_{0 \leq t \leq T} f < B\}} \frac{1}{2} \int_0^T \hat{f}(t)^T a^{-1} \hat{f}(t) dt := -I(f^*(t))
\]

(15)

where \( a = \nu \cdot \nu^T \) and \( f^*(t) := \arg \inf_{\{f(t) \in H_1; \inf_{0 \leq t \leq T} f < B\}} \frac{1}{2} \int_0^T \hat{f}(t)^T a^{-1} \hat{f}(t) dt \).

The second moment becomes

\[
\tilde{E} \left\{ I(\inf_{0 \leq t \leq T} S_t < B) e^{2 \frac{1}{2} \int_0^T \hat{dW}_t} \right\} e^{-\frac{1}{2} \int_0^T \|h\|^2 dt}
\]

\[
= \hat{E} \left\{ I(\inf_{0 \leq t \leq T} S_t < B) e^{2 \frac{1}{2} \int_0^T \hat{dW}_t} \frac{d\tilde{P}}{d\tilde{P}} \right\} e^{-\frac{1}{2} \int_0^T \|h\|^2 dt}
\]

(16)

\[
= \hat{E} \left\{ I(\inf_{0 \leq t \leq T} S_t < B) \right\} e^{\frac{1}{2} \int_0^T \|h\|^2 dt}
\]

\[
= \hat{E} \left\{ I(\inf_{0 \leq t \leq T} S_t - \nu h t < B) \right\} e^{\frac{1}{2} \int_0^T \|h\|^2 dt}
\]

where the measure change \( d\tilde{P}/d\tilde{P} \) is defined by \( Q_T(2h) \) for the second line.

Substituting the scaling of \( \varepsilon \) and \( \bar{B} \) again, the last equation in (16)

\[
E \left\{ I(\inf_{0 \leq t \leq T} \sqrt{\varepsilon} \nu W_t - f^*(t) t < \bar{B}) \right\} e^{\frac{1}{2} \int_0^T \|h\|^2 dt}
\]

(17)

Applying Freidlin-Wentzell Theorem again, we obtain the decay rate of the
second moment
\[
\lim_{\varepsilon \to 0} \varepsilon \ln P^\varepsilon_2(h) = -\frac{1}{2} \int_0^T (2f^*(t))^T a^{-1}(2f^*(t)) dt + \int_0^T (f^*(t))^T \nu^{-1} \nu^{-1} f^*(t)) dt
\]
(18)
\[
= 2 \lim_{\varepsilon \to 0} \varepsilon \ln P^\varepsilon_1.
\]
\[
= -2I(f^*(t))
\]

The last two equalities are obtained from (14). We confirmed that the variance rate of this importance sampling is asymptotically zero so that this scheme is efficient.

**Theorem 2.** (Efficient Importance Sampling for Geometric Brownian Motions) Let \( S_t = (S_{1t}, S_{2t}, \ldots, S_{nt}) \) denote the system of \( n \) assets value following the log-normal process
\[
dS_t = S_t (\mu dt + \nu dW_t)
\]
with the drift rate \( \mu = (\mu_1, \mu_2, \ldots, \mu_n) \), the Brownian motion \( W_t = (W_{1t}, W_{2t}, \ldots, W_{nt}) \), the volatility matrix \( \nu = \{\nu_{i,j}\}_{i,j=1}^n \), the initial value \( S_0 = (S_{10}, S_{20}, \ldots, S_{n0}) \), and the default boundary \( B = (B_1, B_2, \ldots, B_n) \) denote the default boundary. We introduce a scale \( \varepsilon \) defined from \( \frac{1}{\sqrt{\varepsilon}} = \max \{|\ln B_{i}/S_{i0}|\} \) and \( \bar{B} = \sqrt{\varepsilon}(\ln B_{1}/S_{10}, \ldots, \ln B_{n}/S_{n0}) \). The default probability estimated by the importance sampling is defined by
\[
P^\varepsilon_1 = E\left\{\prod_{i=1}^n 1\left(\inf_{0\leq t \leq T} S_{it} \leq B_i \right) \right\}
\]
(19)
\[
= E\left\{\prod_{i=1}^n 1\left(\inf_{0\leq t \leq T} S_{it} \leq B_i \right) Q_T(h) \right\},
\]
where the Radon-Nykodym derivative \( Q(h) \) is defined in (5). The second moment of this estimator is denoted by
\[
P^\varepsilon_2 = E\left\{\prod_{i=1}^n 1\left(\inf_{0\leq t \leq T} S_{it} \leq B_i \right) Q_T^2(h) \right\},
\]
(20)
by the choice \( h = \frac{1}{\sqrt{\varepsilon}} \nu^{-1} f^*(t) \), where \( a = \nu \cdot \nu^T \) and
\[
f^*(t) := \arg\inf_{\{f(t) \in H_1 : \inf_{0\leq t \leq T} f(t) < B\}} \frac{1}{2} \int_0^T f(t)^T a^{-1} f(t) dt.
\]
(21)

When \( \varepsilon \) is small enough, we obtain a zero variance rate, i.e.,
\[
\lim_{\varepsilon \to 0} \varepsilon^2 \ln \left(\frac{P^\varepsilon_2(h)}{(P^\varepsilon_1)^2} \right) = 0,
\]

7
so that the importance sampling scheme is asymptotically optimal (efficient).

Before the proof, we note that $\nu^T \cdot \nu$ is the correlation structure of the system $W_t$. And we denote

$$\nu = \begin{pmatrix}
\nu_{1,1} & \nu_{1,2} & \nu_{1,3} & \cdots & \nu_{1,n} \\
\nu_{2,1} & \nu_{2,2} & \nu_{2,3} & \cdots & \nu_{2,n} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\nu_{n,1} & \nu_{n,2} & \nu_{n,3} & \cdots & \nu_{n,n}
\end{pmatrix} := \begin{pmatrix}
\tilde{\nu}_1 \\
\tilde{\nu}_2 \\
\vdots \\
\tilde{\nu}_n
\end{pmatrix}$$

(22)

for convenience.

**Proof:** Recall that the default probability is defined by

$$P_1 = P\left( \inf_{0 \leq t \leq T} S_t < B \right) := P\left( \bigcap_{i=1}^{n} \{ \inf_{0 \leq t \leq T} S_{it} < B_{i1} \} \right)$$

$$= P\left( \inf_{0 \leq t \leq T} \begin{pmatrix}
\mu_1 - \frac{1}{2} \langle \tilde{\nu}_1, \tilde{\nu}_1 \rangle \\
\mu_2 - \frac{1}{2} \langle \tilde{\nu}_2, \tilde{\nu}_2 \rangle \\
\vdots \\
\mu_n - \frac{1}{2} \langle \tilde{\nu}_n, \tilde{\nu}_n \rangle
\end{pmatrix} t + \nu \begin{pmatrix}
W_{1t} \\
W_{2t} \\
\vdots \\
W_{nt}
\end{pmatrix} < \ln \begin{pmatrix}
\frac{B_1}{S_{10}} \\
\frac{B_2}{S_{20}} \\
\vdots \\
\frac{B_n}{S_{n0}}
\end{pmatrix} \right)$$

$$= P\left( \inf_{0 \leq t \leq T} \sqrt{\varepsilon} \begin{pmatrix}
\mu_1 - \frac{1}{2} \langle \tilde{\nu}_1, \tilde{\nu}_1 \rangle \\
\mu_2 - \frac{1}{2} \langle \tilde{\nu}_2, \tilde{\nu}_2 \rangle \\
\vdots \\
\mu_n - \frac{1}{2} \langle \tilde{\nu}_n, \tilde{\nu}_n \rangle
\end{pmatrix} t + \sqrt{\varepsilon} \nu \begin{pmatrix}
W_{1t} \\
W_{2t} \\
\vdots \\
W_{nt}
\end{pmatrix} < \tilde{B} \right)$$

(23)

where we have used the strictly monotonicity of the logarithmic transformation and we introduce a scaling $\sqrt{\varepsilon} = \max_{i=1,\ldots,n} | \ln \frac{B_i}{S_{i0}} |$, also define $\tilde{B}$ as $\tilde{B} = \sqrt{\varepsilon} \ln(\frac{B_1}{S_{10}}, \ldots, \frac{B_n}{S_{n0}})$. For small parameter $\varepsilon$, the default probability will be small in financial intuition because the debt to asset value, $B_i/S_{i0}$, is small. By an application of Freidlin-Wentzell Theorem, it is easy to prove that the rate function of (23) is

$$\lim_{\varepsilon \to 0} \varepsilon \log P_1 = - \inf_{\{f(t) \in H_1: \inf_{0 \leq t \leq T} f < B\}} \int_0^T \frac{1}{2} \dot{f}(t)^T a^{-1} \dot{f}(t) dt := -I(f^*(t))$$

(24)
where \( a = \nu \cdot \nu^T \) and \( f^*(t) := \arg \inf_{\{f(t) \in H_1: \inf \int_{0 \leq t \leq T} f(t) < \bar{B} \}} \frac{1}{2} \int_0^T \dot{f}(t)^T a^{-1} \dot{f}(t) dt \).

The second moment becomes

\[
\hat{E} \left\{ I \left( \inf_{0 \leq t \leq T} S_t < \bar{B} \right) e^{2 \int_0^T h dW_t - \int_0^T [h]^2 dt} \right\} \\
= \hat{E} \left\{ I \left( \inf_{0 \leq t \leq T} S_t < \bar{B} \right) e^{2 \int_0^T h d\hat{W}_t} \frac{d\hat{P}}{d\hat{P}} \right\} e^{-\int_0^T [h]^2 dt} \\
= \hat{E} \left\{ I \left( \inf_{0 \leq t \leq T} S_t < \bar{B} \right) e^{\int_0^T [h]^2 dt} \right\}
\]

(25)

where the measure change \( d\hat{P} / d\hat{P} \) is defined by \( Q_T(2h) \) for the second line. Recall that under the measure change defined by \( \hat{P} \), the price dynamics \( S_t^j \) becomes

\[
S_0^j \exp \left\{ \left( \mu_j - \frac{1}{2} (\bar{\nu}_j, \bar{\nu}_j) - (\bar{\nu}_j, h) \right) t + (\bar{\nu}_j, \hat{W}_t) \right\}
\]

(26)

Thus, the second moment (25) becomes

\[
\hat{E} \left\{ I \inf_{0 \leq t \leq T} \sqrt{\nu} \left( \begin{array}{c}
\mu_1 - \frac{1}{2} (\bar{\nu}_1, \bar{\nu}_1) \\
\mu_2 - \frac{1}{2} (\bar{\nu}_2, \bar{\nu}_2) \\
\vdots \\
\mu_n - \frac{1}{2} (\bar{\nu}_n, \bar{\nu}_n)
\end{array} \right) - \nu \left( \begin{array}{c}
h_1 \\
h_2 \\
\vdots \\
h_n
\end{array} \right) t + \sqrt{\nu} \left( \begin{array}{c}
\hat{W}_1^j \\
\hat{W}_2^j \\
\vdots \\
\hat{W}_n^j
\end{array} \right) < \bar{B} \right\} e^{\int_0^T [h]^2 dt}.
\]

(27)

We incorporated the same scaling to rescale our problem for the last line above. With \( h = \frac{1}{\sqrt{\nu}} \nu^{-1} \dot{f}^* \), we can apply Freidlin-Wentzell Theorem again,

\[
\lim_{\epsilon \to 0} \epsilon \ln P_{1}^\epsilon(h)
\]

\[
= - \frac{1}{2} \int_0^T (2\dot{f}^*)^T a^{-1} (2\dot{f}^*) dt + \int_0^T (\dot{f}^* \nu^{-1} T \nu^{-1} \dot{f}^*) dt
\]

(28)

\[
= 2 \lim_{\epsilon \to 0} \epsilon \ln P_{1}^\epsilon.
\]

We confirmed that the variance rate of this importance sampling is asymptotically zero so that this scheme is efficient. \( \square \)
4 Numerical Results

Three numerical experiments are conducted in the case of geometric Brownian motions for dimensions one, three, and many others. These results indicate that the proposed importance sampling scheme is quite efficient and robust to several market model scenarios.

When implementing the importance sampling scheme proposed in the previous section, it is essential to solve a variational problem in order to obtain the efficient importance sampling estimator. Here we introduce a procedure to simplify the variational problem into quadratic programming with linear constraint. The idea is to consider the minimizer of rate functional to be straight lines passing through the origin. See the use of this linear approximation in [9]. That is, consider $f(t) = (f_1(t), f_2(t), \cdots, f_n(t))'$ for $k_1, \cdots, k_n \in \mathbb{R}$. Since every $f_i(t)$ needs to touch the barrier $B_i$ before time $T$, the constraints are simplified to $k_i \leq \frac{B_i}{T}$.

Hence, we can approximate the optimum by solving

$$\inf_{k_i \leq \frac{B_i}{T}} \frac{1}{2} k^T a^{-1} k T,$$

where $k = (k_1, \cdots, k_n)'$. Note that the solution of this quadratic programming can be obtained easily but it does not satisfy the original variation problem. However, through various experiments below, this simple approximation remains feasible for the use of importance sampling.

The single name default probability has an exact solution:

$$1 - \mathcal{N}(d_2^+) + \mathcal{N}(d_2^-) \left( \frac{S_0}{B} \right)^{1-2\mu/\sigma^2}$$

with $d_2^\pm = \frac{\pm \ln(S_0/B) + (\mu \sigma^2/2)T}{\sigma \sqrt{T}}$ [4]. This result is obtained from the distribution of the running minimum of Brownian motion. Table 1 records several results obtained by basic Monte Carlo (BMC), exact solution, and importance sampling (IS). The number of simulation is $10^4$ and an Euler discretization for (1) is used by taking time step size $T/400$, where $T$ is one year. Other parameters are $S_0 = 100, \mu = 0.05$ and $\sigma = 0.4$.

There is no closed-form solution for the joint default probability of three names except for the case of zero correlation. In Table 2, we consider 5 instances of a 3-dimensional geometric Brownian motion. All of them are of the same initial parameters, except for the correlation structure.

$N = 10^5$, $dt = 10^{-4}$, $S_0 = (100, 100, 200)$, $\mu = (0.07, 0.05, 0.05)$, $\sigma = (0.4, 0.4, 0.3)$, $B = (30, 30, 120)$
\[
\rho_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \rho_2 = \begin{pmatrix} 1 & 0.3 & 0.3 \\ 0.3 & 1 & 0.3 \\ 0.3 & 0.3 & 1 \end{pmatrix}, \quad \rho_3 = \begin{pmatrix} 1 & -0.3 & -0.3 \\ -0.3 & 1 & -0.3 \\ -0.3 & -0.3 & 1 \end{pmatrix},
\]
\[
\rho_4 = \begin{pmatrix} 1 & 0.5 & 0.3 \\ 0.5 & 1 & -0.2 \\ 0.3 & -0.2 & 1 \end{pmatrix}, \quad \rho_5 = \begin{pmatrix} 1 & -0.1 & -0.7 \\ -0.1 & 1 & 0.2 \\ -0.7 & 0.2 & 1 \end{pmatrix}.
\]

Table 3 records multi-name joint default probabilities by basic Monte Carlo (BMC), and importance sampling (IS) under high-dimensional Black-Cox model. Let \( n \) denote the dimension, the total number of firms. The number of simulation is \( 3 \times 10^4 \) and an Euler discretization for (1) is used by taking time step size \( T \)/100, where \( T \) is one year. Other parameters are \( S_0 = 100, \mu = 0.05, \sigma = 0.3, \rho = 0.3, \) and \( B = 50 \).

Table 1: Comparison of single-name default probabilities by basic Monte Carlo (BMC), exact solution, and importance sampling (IS) with different thresholds \( B \).

<table>
<thead>
<tr>
<th>( B )</th>
<th>exact</th>
<th>CMC</th>
<th>s.e.</th>
<th>IS</th>
<th>s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>0.0945</td>
<td>0.0886</td>
<td>0.0028</td>
<td>0.0890</td>
<td>0.0016</td>
</tr>
<tr>
<td>20</td>
<td>7.73e-5</td>
<td>-</td>
<td>-</td>
<td>7.16e-5</td>
<td>2.32e-6</td>
</tr>
<tr>
<td>1</td>
<td>1.33e-30</td>
<td>-</td>
<td>-</td>
<td>1.81e-30</td>
<td>3.44e-31</td>
</tr>
</tbody>
</table>

Table 2: Comparisons of three-name joint default probabilities by basic Monte Carlo (BMC), exact solution, and importance sampling (IS) with different correlation structures \( \rho \).

<table>
<thead>
<tr>
<th>( \rho )</th>
<th>exact</th>
<th>CMC</th>
<th>s.e.</th>
<th>IS</th>
<th>s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho_1 )</td>
<td>7.93e-07</td>
<td>0</td>
<td>0</td>
<td>7.96e-07</td>
<td>4.34e-08</td>
</tr>
<tr>
<td>( \rho_2 )</td>
<td>-2e-05</td>
<td>1.41e-05</td>
<td>4.13e-05</td>
<td>4.13e-05</td>
<td>1.06e-06</td>
</tr>
<tr>
<td>( \rho_3 )</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>5.88e-13</td>
<td>4.37e-14</td>
</tr>
<tr>
<td>( \rho_4 )</td>
<td>-6e-05</td>
<td>2.45e-05</td>
<td>2.34e-05</td>
<td>2.34e-05</td>
<td>1.65e-06</td>
</tr>
<tr>
<td>( \rho_5 )</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>3.24e-11</td>
<td>5.12e-12</td>
</tr>
</tbody>
</table>
Table 3: Comparison of single-name default probabilities by basic Monte Carlo (BMC), exact solution, and importance sampling (IS) with various dimension size \( n \).

<table>
<thead>
<tr>
<th>( n )</th>
<th>CMC s.e.</th>
<th>IS s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.1e-03</td>
<td>3.31e-04</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>20</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>25</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>30</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>35</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>40</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>45</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>50</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

One can combine these importance sampling schemes with other techniques to solve for empirical problems under complicated models. For example, Han [8] measures the systemic risk under a class of stochastic volatility/correlation models. A combination of the homogenization of stochastic volatility matrix models with the efficient importance sampling proposed in this paper becomes a feasible computational method to treat empirical problems such as SRISK.

5 Conclusion

Estimation of joint default probabilities in a first passage time model is tackled by importance sampling. We extend the two-dimensional problem studied by Zhou (2001) to any finite dimension and prove that our proposed importance sampling method is asymptotically optimal by means of the large deviation principle. A variational problem is characterized to construct the optimal importance sampling.
References


