Interacting Particle Systems for the Computation of CDO Tranche Spreads with Rare Defaults

Douglas Vestal · **René Carmona** ·
Jean-Pierre Fouque

Abstract We propose an Interacting Particle System method to accurately calculate the distribution of the losses in a highly dimensional portfolio by using a selection and mutation algorithm. We demonstrate the efficiency of this method for computing the probability of rare defaults on a toy model for which we have explicit formulas. This method has the advantage of accurately computing small probabilities without requiring the user to compute a change of measure as in the Importance Sampling method. This method will be useful for computing the senior tranche spreads in *Collateralized Debt Obligations* (CDOs).

Keywords Interacting Particle Systems · Rare Defaults · Monte Carlo Methods · Credit Derivatives · Variance Reduction

Mathematics Subject Classification (2000) 60H35 · 91B70

1 Introduction

The past few years have seen an explosion in the credit risk market. At the same time, the field of credit risk and credit derivatives research has substantially increased. As the credit derivative products have grown in complexity, so has the need for fast and accurate numerical methods to accurately price such derivatives.

D. Vestal · J.P. Fouque
Department of Statistics and Applied Probability, University of California, Santa Barbara, CA 93106-3110, USA
E-mail: vestal@pstat.ucsb.edu, fouque@pstat.ucsb.edu

R. Carmona
Department of Operations Research & Financial Engineering, Princeton University, E-Quad, Princeton, NJ 08544, USA
E-mail: rcarmona@princeton.edu

In this paper, we consider the pricing of CDOs under the first passage model. A CDO is a credit derivative that pools together many different firms (125 is a typical contract) and sells exposure to different default levels of the portfolio; the so-called tranches. This segmentation of the risk in a portfolio enables the buyer to purchase only the tranches that they deem appropriate for their hedging positions. Since the tranche levels are fairly standardized, there are also new products called *bespoke* CDOs that sell customized default levels.

The main difficulty in pricing CDOs is the high-dimensional nature of the problem. To accurately price the CDO, the distribution of the joint default for many names is needed. Even if the distribution of joint defaults is found explicitly, there is no guarantee that the expectations that are then needed to compute tranche spreads can be found analytically. Therefore, the user has to rely on numerical schemes to compute the CDO prices. Due to the high-dimensional nature of the problem, PDE based methods are ruled out and Monte Carlo (MC) methods are heavily relied upon. While MC methods are easy to implement in high-dimensional problems, they do suffer from slow convergence. In addition, due to the rare nature of joint defaults, the computational problem is exacerbated since many MC simulations are needed to observe the joint default of many names. Therefore, variance reduction and efficiency become very important for these MC methods.

The main variance reduction technique used is importance sampling. There have been many successful applications of importance sampling towards credit risk [1,12]. However, most authors have concentrated on multifactor Gaussian copula models or the reduced-form model for credit risk. We have not found any reference in the literature that applies importance sampling to the first passage model. In addition, the difficulty with implementing importance sampling is computing the change of measure under which one simulates the random variables. More information about importance sampling and the difficulty required can be found in [11] and the references therein.

In this paper we are concerned with first passage models where it is impractical, if not impossible, to compute the importance sampling change of measure explicitly. Examples we have in mind are first passage models with stochastic volatility (see for instance [10]), and/or regime switching models with many factors.

In this situation, our solution is to use an Interacting Particle System (IPS) technique that can briefly be loosely described as applying importance sampling techniques in path space. That is, we never compute a change of measure explicitly but rather decide on a judicious choice of a potential function and a selection/mutation algorithm under the original measure such that the measure in path space converges to the desired "twisted" measure. This "twisted" measure will be exactly the change of measure that one would use in an importance sampling scheme. The advantage is that the random processes are always simulated under the original measure, but through the selection/mutation algorithm converge to the desired importance sampling measure. The IPS techniques and theoretical support we use to design our algorithm can be found

in the book [5] where Pierre Del Moral developed the theory of IPS, and in [6] which provides applications as well as a toy model very similar to the one used in this paper in the one-dimensional case.

In this paper, we are interested in the intersection of two events; a complicated model that doesn't lend itself to importance sampling, and the computation of rare probabilities under such a model. The rest of the paper is organized as follows. Section 2 discusses the first passage model that we will be using as a toy model. Section 3 gives an overview of Feynman-Kac measures and the associated IPS interpretation. Section 4 provides the algorithm we propose and outlines its implementation on our toy model. Section 5 discusses the numerical results of using IPS on our toy model and the comparison with traditional MC.

2 Problem Formulation

In contrast to intensity-based approaches to credit risk where default is given by an exonously defined process, default in the firm value model has a very nice economical appeal. The firm value approach, or structural model, models the total value of the firm that has issued the defaultable bonds. Typically, the value of the firm includes the value of equity (shares) and the debt (bonds) of the firm [15]. There are two main approaches to modeling default in the firm value approach: one is that default can only happen at the maturity of the bond, and the second is that default can occur any time before maturity. The latter is referred to as the *first-passage model* and is the one we consider in this paper.

2.1 Review of the First Passage Model

We follow both [14,2] and assume that the value of the firm, S_t, follows geometric Brownian motion. We also assume that interest rates are constant. Under the risk-neutral probability measure $\mathbb{P}$ we have,

$$dS(t) = rS(t)dt + \sigma S(t)dW(t), \tag{2.1}$$

where r is the risk-free interest rate, and σ is constant volatility. At any time $t \leq T$, the price of the unit-notional nondefaultable bond is $\Gamma(t,T) = e^{-r(T-t)}$.

We also assume that at time 0, the firm issued non-coupon corporate bonds expiring at time T. The price of the defaultable bond at time $t \leq T$ is denoted by $\bar{\Gamma}(t,T)$.

In [14] default was assumed to only occur at the expiration date T. Furthermore, default at time T is triggered if the value of the firm was below some default threshold B; that is if $S_T \leq B$. Therefore, assuming zero recovery, the price of the defaultable bond satisfies,

$$\bar{\Gamma}(t,T) = \mathbb{E}\left[e^{-r(T-t)}\mathbf{1}_{S_T > B}|S_t\right]$$
$$= \Gamma(t,T)\mathbb{P}(S_T > B|S_t)$$
$$= \Gamma(t,T)N(d_2),$$

where $N(\cdot)$ is the standard cumulative normal distribution function and,

$$d_2 = \frac{\ln\left(\frac{S_t}{B}\right) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}.$$

The next model, the Black-Cox model, is also known as the first passage approach. Developed in [2], default can occur anytime before the expiration of the bond and the barrier level, $B(t)$, is some deterministic function of time. In [2], they assume that the default barrier is given by the function $B(t) = Ke^{\eta t}$ (exponentially increasing barrier) with $K > 0$ and $\eta \geq 0$. The default time τ is defined by

$$\tau = \inf\{t : S_t \leq B(t)\}.$$

That is, default happens the first time the value of the firm passes below the default barrier. This has the very appealing economical intuition that default occurs when the value of the firm falls below its debt level, thereby not allowing the firm to pay off its debt.

Assuming zero recovery, we can price the defaultable bond by pricing a barrier option. Therefore,

$$\bar{\Gamma}(t,T) = \mathbf{1}_{\tau > t}\mathbb{E}\left[e^{-r(T-t)}\mathbf{1}_{\tau > T}|S_t\right]$$
$$= \mathbf{1}_{\tau > t}\Gamma(t,T)\mathbb{P}(\tau > T|S_t)$$
$$= \mathbf{1}_{\tau > t}\Gamma(t,T)\left(N(d_2^+) - \left(\frac{S_t}{B(t)}\right)^p N(d_2^-)\right),$$

where

$$d_2^{\pm} = \frac{\pm\ln\left(\frac{S_t}{B(t)}\right) + (r - \eta - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}},$$

$$p = 1 - \frac{2(r - \eta)}{\sigma^2}.$$

In addition, we denote the probability of default for the firm between time t and time T by $P(t,T)$. Hence,

$$P(t,T) = 1 - \left(N(d_2^+) - \left(\frac{S_t}{B(t)}\right)^p N(d_2^-)\right). \tag{2.2}$$

The yield spread of the defaultable bond is defined as

$$Y(t,T) = -\frac{1}{T-t}\ln\left(\frac{\bar{\Gamma}(t,T)}{\Gamma(t,T)}\right).$$

We remark that in the first passage model above, and all firm value models, the yield spread for very short maturities goes to 0 in contrast to the empirical evidence found in [8]. However, by incorporating a fast mean-reverting stochastic volatility σ_t, the authors in [9] were able to raise the yield spread for short maturities.

2.2 Multiname Model

For our purposes, we consider a CDO written on N firms under the first passage model. That is, we assume that the firm values for the N names have the following dynamics,

$$dS_i(t) = rS_i(t)dt + \sigma_i S_i(t)dW_i(t), \quad i = 1, \ldots, N \qquad (2.3)$$

where r is the risk-free interest rate, σ_i is constant volatility, and the driving innovations $dW_i(t)$ are infinitesimal increments of Wiener processes W_i with correlation

$$d\langle W_i, W_j\rangle_t = \rho_{ij}dt.$$

Each firm i is also assumed to have a deterministic boundary process $B_i(t)$ and default for firm i is given by

$$\tau_i = \inf\left\{t : S_i(t) \leq B_i(t)\right\}. \qquad (2.4)$$

We define the loss function as

$$L(T) = \sum_{i=1}^{N} \mathbf{1}_{\{\tau_i \leq T\}}. \qquad (2.5)$$

That is, $L(T)$ counts the number of firms among the N firms that have defaulted before time T. We remark that in the independent homogeneous portfolio case, the distribution of $L(T)$ is $Binomial(N, P(0, T))$ where $P(0, T)$ is defined in (2.2).

It is well known (see for instance [7]) that the spread on a single tranche can be computed from the knowledge of expectations of the form,

$$\mathbb{E}\{(L(T_i) - K_j)^+\},$$

where T_i is any of the coupon payment dates, and K_j is proportional to the attachment points of the tranche.

The most interesting and challenging computational problem is when all of the names in the portfolio are correlated. In [18], the distribution of losses for $N = 2$ is found by finding the distribution of the hitting times of a pair of correlated Brownian motions. However, the distribution is given in terms of modified Bessel functions and a tractable general result for $N > 2$ is not available. Since the distribution of $L(T)$ is not known in the dependent case, for $N > 2$, Monte Carlo methods are generally used to calculate the spread

on the tranches. Since N is typically very large (125 names is a standard contract), PDE based methods are ruled out and one has to use Monte Carlo.

Instead of computing the spread on the tranches numerically, our goal is to calculate the probability mass function for $L(T)$, that is calculate

$$\mathbb{P}(L(T) = i) = p_i(T), \quad i = 0, \dots, N. \tag{2.6}$$

In this manner, we will then be able to calculate *all* expectations that are a function of $L(T)$, not just the spreads. In addition, as the reader will see, our method is dynamically consistent in time so that we can actually calculate, for all coupon dates $T_j \leq T$,

$$\mathbb{P}(L(T_j) = i) = p_i(T_j), \quad i = 0, \dots, N,$$

with one Monte Carlo run. This is in contrast to a lot of importance sampling techniques where the change of measure has a dependence on the final time through the Girsanov transformation thereby requiring a different MC run for each coupon date.

3 Feynman-Kac Path Measures and IPS

Feynman-Kac path measures, and their subsequent interacting particle system interpretation, are closely related to the stochastic filtering techniques used in mathematical finance. In this paper, we adapt an original interacting particle system developed in [6] to the computation of the probability mass function (2.6). In [6], the authors develop a general interacting particle system method for calculating the probabilities of rare events. For the sake of completeness, we provide a fairly thorough overview of IPS as developed in [6] and the main results that will be the foundation of this paper. Briefly, the method can be described as formulating a Markov process and conducting a mutation-selection algorithm so that the chain is "forced" into the rare event regime.

We suppose that we have a continuous time non-homogeneous Markov chain, $(\tilde{X}_t)_{t \in [0,T]}$. However, for the variance analysis and to foreshadow the method ahead, we only consider the chain $(X_p)_{0 \leq p \leq n} = (\tilde{X}_{pT/n})_{0 \leq p \leq n}$, where n is fixed. The chain X_n takes values in some measurable state space $(E_n, \mathcal{E}_n)$ with Markov transitions $K_n(x_{n-1}, dx_n)$. We denote by Y_n the historical process of X_n, that is

$$Y_n \stackrel{\text{def.}}{=} (X_0, \dots, X_n) \in F_n \stackrel{\text{def.}}{=} (E_0 \times \cdots \times E_n).$$

Let $M_n(y_{n-1}, dy_n)$ be the Markov transitions associated with the chain Y_n. Let $B_b(E)$ denote the space of bounded, measurable functions with the uniform norm on some measurable space $(E, \mathcal{E})$. Then, given any $f_n \in B_b(F_n)$, and the pair of potentials/transitions (G_n, M_n), we have the following Feynman-Kac measure defined by

$$\gamma_n(f_n) = \mathbb{E}\left[f(Y_n) \prod_{1 \leq k < n} G_k(Y_k) \right]. \tag{3.1}$$

We denote by $\eta_n(\cdot)$ the normalized measure defined as

$$\eta_n(f_n) = \frac{\mathbb{E}\left[f(Y_n)\prod_{1\leq k<n}G_k(Y_k)\right]}{\mathbb{E}\left[\prod_{1\leq k<n}G_k(Y_k)\right]} = \gamma_n(f_n)/\gamma_n(1). \qquad (3.2)$$

In addition, in [6] they assume that the potential functions are chosen such that

$$\sup_{(y_n,\bar{y}_n)\in F_n^2} G_n(y_n)/G_n(\bar{y}_n) < \infty.$$

However, the authors note that this condition can be relaxed by considering traditional cut-off techniques, among other techniques (see [6,5] for more details).

A very important observation is that

$$\gamma_{n+1}(1) = \gamma_n(G_n) = \eta_n(G_n)\gamma_n(1) = \prod_{p=1}^{n}\eta_p(G_p).$$

Therefore, given any bounded measurable function f_n, we have

$$\gamma_n(f_n) = \eta_n(f_n)\prod_{1\leq p<n}\eta_p(G_p).$$

The above relationship is crucial because it allows us to relate the un-normalized measure in terms of only the normalized "twisted" measures. In our study, we will also make use of the distribution flow (γ_n^-,η_n^-) defined exactly the same way as (γ_n,η_n) except we replace G_p by its inverse

$$G_p^- = 1/G_p\,.$$

Then, using the definition for γ_n and η_n it is easy to see that $\mathbb{E}[f_n(Y_n)]$ admits the following representation,

$$\mathbb{E}\left[f_n(Y_n)\right] = \mathbb{E}\left[f_n(Y_n)\prod_{1\leq p<n}G_p^-(Y_p)\times\prod_{1\leq p<n}G_p(Y_p)\right]$$

$$= \gamma_n\left(f_n\prod_{1\leq p<n}G_p^-\right)$$

$$= \eta_n\left(f_n\prod_{1\leq p<n}G_p^-\right)\prod_{1\leq p<n}\eta_p(G_p).$$

Finally, it can be checked that the measures $(\eta_n)_{n\geq 1}$ satisfy the nonlinear recursive equation

$$\eta_n = \Phi_n(\eta_{n-1}) \overset{\text{def.}}{=} \int_{F_{n-1}}\eta_{n-1}(dy_{n-1})G_{n-1}(y_{n-1})M_n(y_{n-1},\cdot)/\eta_{n-1}(G_{n-1}),$$

starting from $\eta_1 = M_1(x_0,\cdot)$.

3.1 IPS Interpretation and General Algorithm

The above definitions and results lend themselves to a very natural interacting path-particle interpretation. We denote the Markov chain taking values in the product space F_n^M with transformation Φ_n by $\xi_n = (\xi_n^i)_{1 \leq i \leq M}$, for each time $n \geq 1$. One constructs a numerical algorithm so that each path-particle

$$\xi_n^i = (\xi_{0,n}^i, \xi_{1,n}^i, \ldots, \xi_{n,n}^i) \in F_n = (E_0 \times \cdots \times E_n),$$

is sampled almost according to the twisted measure η_n.

We start with an initial configuration $\xi_1 = (\xi_1^i)_{1 \leq i \leq M}$ that consists of M independent and identically distributed random variables with distribution,

$$\eta_1(d(y_0, y_1)) = M_1(x_0, d(y_0, y_1)) = \delta_{x_0}(dy_0) K_1(y_0, dy_1),$$

i.e., $\xi_1^i \stackrel{\text{def.}}{=} (\xi_{0,1}^i, \xi_{1,1}^i) = (x_0, \xi_{1,1}^i) \in F_1 = (E_0 \times E_1)$. Then, the elementary transitions $\xi_{n-1} \to \xi_n$ from F_{n-1}^M into F_n^M are defined by

$$\mathbb{P}(\xi_n \in d(y_n^1, \ldots, y_n^M) | \xi_{n-1}) = \prod_{j=1}^M \Phi_n(m(\xi_{n-1}))(dy_n^i), \qquad (3.3)$$

where $m(\xi_{n-1}) \stackrel{\text{def.}}{=} \frac{1}{M} \sum_{i=1}^M \delta_{\xi_{n-1}^i}$, and $d(y_n^1, \ldots, y_n^M)$ is an infinitesimal neighborhood of the point $(y_n^1, \ldots, y_n^M) \in F_n^m$. From the definition of Φ_n, one can see that (3.3) is the overlapping of a simple selection and mutation transition,

$$\xi_{n-1} \in F_{n-1}^M \stackrel{\text{selection}}{\longrightarrow} \hat{\xi}_{n-1} \in F_{n-1}^M \stackrel{\text{mutation}}{\longrightarrow} \xi_n \in F_n^M.$$

The selection stage is performed by choosing randomly and independently M path-particles

$$\hat{\xi}_{n-1}^i = (\hat{\xi}_{0,n-1}^i, \hat{\xi}_{1,n-1}^i, \ldots, \hat{\xi}_{n-1,n-1}^i) \in F_{n-1},$$

according to the Boltzmann-Gibbs particle measure

$$\sum_{j=1}^M \frac{G_{n-1}(\xi_{0,n-1}^j, \ldots, \xi_{n-1,n-1}^j)}{\sum_{i=1}^M G_{n-1}(\xi_{0,n-1}^i, \ldots, \xi_{n-1,n-1}^i)} \delta_{(\xi_{0,n-1}^j, \ldots, \xi_{n-1,n-1}^j)}.$$

Then, for the mutation stage, each selected path-particle $\hat{\xi}_{n-1}^i$ is extended by

$$\xi_n^i = ((\xi_{0,n}^i, \ldots, \xi_{n-1,n}^i), \xi_{n,n}^i)$$
$$= ((\hat{\xi}_{0,n}^i, \ldots, \hat{\xi}_{n-1,n}^i), \xi_{n,n}^i) \in F_n = F_{n-1} \times E_n,$$

where $\xi_{n,n}^i$ is a random variable with distribution $K_n(\hat{\xi}_{n-1,n-1}^i, \cdot)$. In other words, the transition is made by applying the original kernel K_n. All of the

mutations are performed independently. We just quote the results from [5,6] in stating the weak convergence result:

$$\eta_n^M \stackrel{\text{def.}}{=} \frac{1}{M} \sum_{i=1}^{M} \delta_{(\xi_{0,n}^i, \xi_{1,n}^i, \ldots, \xi_{n,n}^i)} \stackrel{N \to \infty}{\longrightarrow} \eta_n.$$

Furthermore, there are several propagation of chaos estimates that ensure that $(\xi_{0,n}^i, \xi_{1,n}^i, \ldots, \xi_{n,n}^i)$ are asymptotically independent and identically distributed with distribution η_n [5]. Therefore, we can form the particle approximation γ_n^M defined as

$$\gamma_n^M(f_n) = \eta_n^M(f_n) \prod_{1 \leq p < n} \eta_p^M(G_p).$$

Lemma 1 ([6]). *γ_n^M is an unbiased estimator for γ_n, in the sense that for any $p \geq 1$ and $f_n \in B_b(F_n)$ with $\|f_n\| \leq 1$, we have*

$$\mathbb{E}(\gamma_n^M(f_n)) = \gamma_n(f_n),$$

and in addition

$$\sup_{M \geq 1} \sqrt{M} \mathbb{E}[|\gamma_n^M(f_n) - \gamma_n(f_n)|^p]^{1/p} \leq c_p(n),$$

for some constant $c_p(n) < \infty$ whose value does not depend on the function f_n.

Proof Refer to [6].

4 Pricing CDOs using IPS

In this section, we present our adaptation of the interacting particle system approach to computing rare event probabilities in credit risk with a structural based approach by applying it to the following model.

Our Markov process is the $3 \times N$ dimensional process $(\tilde{X}_t)_{t \in [0,T]}$ defined as:

$$\tilde{X}_t = \Big(S_1(t), \min_{u \leq t} S_1(u), \mathbf{1}_{\tau_1 \leq t}, S_2(t), \min_{u \leq t} S_2(u), \mathbf{1}_{\tau_2 \leq t}, \cdots$$
$$\cdots, S_N(t), \min_{u \leq t} S_N(t), \mathbf{1}_{\tau_N \leq t} \Big),$$

where the dynamics of $S_i(t)$ are given in equation (2.3). We assume a constant barrier B_i for each firm $1 \leq i \leq N$. While it is redundant to also include $\mathbf{1}_{\tau_i \leq t}$ in the above expression since we also know $\min_{u \leq t} S_i(u)$, we keep track of it because it will tell us the default *time* of the firm when we implement the algorithm numerically. We divide the time interval $[0,T]$ into n equal intervals $[t_{i-1}, t_i]$, $i = 1, 2, \ldots, n$. These are the times we stop and perform the selection step. We introduce the chain $(X_p)_{0 \leq p \leq n} = (\tilde{X}_{pT/n})_{0 \leq p \leq n}$ and the whole history of the chain is denoted by $Y_p = (X_0, \ldots, X_p)$.

Since it is not possible to sample directly from the distribution of $(X_p)_{0 \leq p \leq n}$ for $N > 2$, we will have to apply an Euler scheme during the mutation stage; we let $\triangle t$ denote the sufficiently small time step used. In general $\triangle t$ will be chosen so that $\triangle t << T/n$.

Our general strategy is to find a potential function that increases the likelihood of default among the firms. In the IPS algorithm, given a particular choice of weight function $G(\cdot)$, particles with low scores are replaced by particles with high scores. Therefore, we would like to select a potential function $G(\cdot)$ that places a higher score on firms which have reduced their distance to default during the previous mutation step. Since the rare event in this case is that the minimum of the firm value falls below a certain level, we would like to put more emphasis on particles whose firm minimums are decreasing during a mutation step. Therefore, we fix some parameter $\alpha < 0$, and define the potential function,

$$G^\alpha(Y_p) = \exp[\alpha(V(X_p) - V(X_{p-1}))], \qquad (4.1)$$

where

$$V(X_p) = \sum_{i=1}^{N} \log(\min_{u \leq t_p} S_i(u)).$$

The choice of $\alpha < 0$ may seem peculiar initially, but it is chosen to be negative because the potential function $G^\alpha(Y_p)$ can be written in the form,

$$
\begin{aligned}
G^\alpha(Y_p) &= \exp[\alpha(V(X_p) - V(X_{p-1}))] \\
&= \exp\left[\sum_{i=1}^{N} \alpha \log\left(\min_{u \leq t_p}(S_i(u)/\min_{u \leq t_{p-1}} S_i(u))\right)\right],
\end{aligned}
$$

where,

$$\log\left(\min_{u \leq t_p}(S_i(u)/\min_{u \leq t_{p-1}} S_i(u))\right) \leq 0.$$

Therefore, to place more weight on the firms whose minimum has decreased, we must multiply by $\alpha < 0$. In addition, by choosing the weight function as we did, if the minimum value did not decrease during the last mutation, then that firm has a small relative contribution to the total empirical measure. Therefore, we will be putting more weight onto path-particles whose minimum has decreased the most between two mutation times.

In addition, there are several computational advantages for choosing the weight function above. Chiefly among them are:

1. Our choice of weight function, while not unique in this regard, will only require us to keep track of (X_{p-1}, X_p) instead of the full history $Y_p = (X_0, X_1, \cdots, X_p)$ thereby minimizing the increased dimensionality of using an IPS scheme.

2. In addition, our weight function has the added advantage of having the property that $\prod_{1 \leq k < p} G(Y_k) = exp[\alpha(V(X_{p-1}) - V(X_0))]$ thereby ensuring that the Feynman-Kac measures defined in equations (3.1) and (3.2) are simpler to analyze.

4.1 Detailed IPS Algorithm

Our algorithm is built with the weight function defined in equation (4.1).

Initialization. We start with M identical copies, $\hat{X}_0^{(i)}$, $1 \leq i \leq M$, of the initial condition X_0. That is,

$$\hat{X}_0^{(i)} = (S_1(0), S_1(0), 0, S_2(0), S_2(0), 0, \cdots, S_N(0), S_N(0), 0), \quad 1 \leq i \leq M.$$

We also have a set of "parents", $\hat{W}_0^{(i)}$, defined by $\hat{W}_0^{(i)} = \hat{X}_0^{(i)}$. We denote $V_0 \stackrel{\text{def.}}{=} V(\hat{W}_0^{(i)})$. This forms a set of M particles $(\hat{W}_0^{(i)}, \hat{X}_0^{(i)})$, $1 \leq i \leq M$.

Now suppose that at time p, we have the set of M particles $(\hat{W}_p^{(i)}, \hat{X}_p^{(i)})$, $1 \leq i \leq M$.

Selection Stage

We first compute the normalizing constant,

$$\hat{\eta}_p^M = \frac{1}{M} \sum_{i=1}^M \exp\left[\alpha\left(V(\hat{X}_p^{(i)}) - V(\hat{W}_p^{(i)})\right)\right]. \tag{4.2}$$

Then, we choose independently M particles according to the empirical distribution,

$$\eta_p^M(d\check{W}, d\check{X}) = \frac{1}{M\hat{\eta}_p^M} \sum_{i=1}^M \exp\left[\alpha\left(V(\hat{X}_p^{(i)}) - V(\hat{W}_p^{(i)})\right)\right] \times \delta_{(\hat{W}_p^{(i)}, \hat{X}_p^{(i)})}(d\check{W}, d\check{X}). \tag{4.3}$$

The particles that are selected are denoted $(\check{W}_p^{(i)}, \check{X}_p^{(i)})$.

Mutation Stage

For each of the selected particles, $(\check{W}_p^{(i)}, \check{X}_p^{(i)})$, we apply an Euler scheme from time t_p to time t_{p+1} with step size $\triangle t$ for each $\check{X}_p^{(i)}$ so that $\check{X}_p^{(i)}$ becomes $\hat{X}_{p+1}^{(i)}$. We then set $\hat{W}_{p+1}^{(i)} = \check{X}_p^{(i)}$ It should be noted, that each of the particles are evolved independently and that the true dynamics (given in equation (2.3)) of X_p is applied rather than some other measure. It is this fact that separates IPS from IS (Importance Sampling).

Then let,

$$f(\hat{X}_n^{(i)}) = \sum_{j=1}^N \mathbf{1}_{\{\min_{u \leq T} S_j^{(i)}(u) \leq B_j\}}$$

denote the number of firms that have defaulted by time T for the ith particle. Then, the estimator for $P(L(T) = k) = p_k(T)$ is given by

$$P_k^M(T) = \left[\frac{1}{M} \sum_{i=1}^M \mathbf{1}_{f(\hat{X}_n^{(i)})=k} \exp(-\alpha(V(\hat{W}_n^{(i)}) - V_0))\right] \times \left[\prod_{p=0}^{n-1} \hat{\eta}_p^M\right]. \tag{4.4}$$

This estimator is unbiased in the sense that $\mathbb{E}[P_k^M(T)] = p_k(T)$. The unbiasedness follows directly from Lemma 1.

4.2 Single-Name Case: Variance Analysis

We analyze the variance of the estimator in equation (4.4) for the single name case. Therefore, we take $N = 1$, with constant barrier B and we are interested in computing, using IPS, the probability of default before maturity T. That is, we compute

$$P^B(0, T) = \mathbb{P}(\min_{u \leq T} S(u) \leq B) = \mathbb{E}[\mathbf{1}_{\min_{u \leq T} S(u) \leq B}].$$

Of course, we have an explicit formula for $P^B(0, T)$ given by (2.2), and this case will be precisely our toy model used to compare the variance for IPS and pure MC. In the more general case, where N is large and the names are correlated, we will provide an empirical comparison. It should be noted that we are only interested in values of B that make the above event rare.

We remark that it is a standard result that the variance associated with the traditional Monte Carlo method for computing $P^B(0, T)$ is $P^B(0, T)(1 - P^B(0, T))$. We also remark that for a single name the Markov chain $(X_p)_{0 \leq p \leq n}$ defined in Section 4 simplifies to

$$X_p = (S(t_p), \min_{u \leq t_p} S(u), \mathbf{1}_{\tau \leq t_p}).$$

Then, following the setup described in Section 3, we see that the rare event probability $P^B(0, T)$ has the following Feynman-Kac representation:

$$P^B(0, T) = \gamma_n(L_n^{(B)}(1)),$$

where $L_n^{(B)}(1)$ is given by the weighted indicator function defined for any path $y_n = (x_0, \ldots, x_n) \in F_n$ by

$$L_n^{(B)}(1)(y_n) = L_n^{(B)}(1)(x_0, \ldots, x_n) = \mathbf{1}_{\{\min_{u \leq T} S(u) \leq B\}} \prod_{1 \leq p < n} G_p^-(x_0, \ldots, x_p)$$

$$= \mathbf{1}_{\{\min_{u \leq T} S(u) \leq B\}} e^{-\alpha(V(x_{n-1}) - V(x_0))}$$

$$= \mathbf{1}_{\{\min_{u \leq T} S(u) \leq B\}} e^{-\alpha(\log(\min_{u \leq t_{n-1}} S(u)/S_0))}.$$

Also, notice that $||L_n^{(B)}(1)(y_n)|| \leq 1$ since $\log(\min_{u \leq t_{n-1}} S(u)/S_0) \leq 0$ and $-\alpha > 0$ by assumption.

Next, following the IPS selection-mutation algorithm outlined in Section 4.1, we form the estimator

$$P_M^B(0, T) = \gamma_n^M(L_n^{(B)}(1)) = \eta_n^M(L_n^{(B)}(1)) \prod_{1 \leq p < n} \eta_p^M(G_p). \tag{4.5}$$

By Lemma 1, $P_M^B(0, T)$ is an unbiased consistent estimator of $P^B(0, T)$. While many estimators are unbiased, the key to determining the efficiency of our estimator is to look at its variance. As such, we have the following central limit theorem for our estimator.

Theorem 1. *The estimator $P_B^M(0,T)$ given in equation (4.5) is unbiased, and it satisfies the central limit theorem*

$$\sqrt{M}\mathbb{E}[P_M^B(0,T) - P^B(0,T)] \overset{M\to\infty}{\longrightarrow} N(0, \sigma_n^B(\alpha)^2),$$

with the asymptotic variance

$$\sigma_n^B(\alpha)^2$$
$$= \sum_{p=1}^{n}\left[\mathbb{E}\left\{e^{\alpha\log(\min_{u\leq t_{p-1}} S(u))}\right\} \times \mathbb{E}\left\{P_{B,p,n}^2 e^{-\alpha\log(\min_{u\leq t_{p-1}} S(u))}\right\} - P^B(0,T)^2\right],$$

$$(4.6)$$

where $P_{B,p,n}$ is the collection of functions defined by

$$P_{B,p,n}(x) = \mathbb{E}\left\{\mathbf{1}_{\min_{t\leq T} S(t)\leq B}|X_p = x\right\},$$

and $P^B(0,T)$ is given by (2.2).

Proof The proof follows directly by applying Theorem 2.3 in [6] with the weight function that we have defined in (4.1).

In the constant volatility single-name case, the asymptotic variance $\sigma_n^B(\alpha)^2$ can be obtained explicitly in terms of double and triple integrals with respect to explicit densities. This will be used in our comparison of variances for IPS and pure MC in Section 5.1. The details of these explicit formulas are given in the Appendix A.

As shown numerically in the next section the variance for IPS is of order p^2 with $p = P^B(0,T)$ (small in the regime of interest), in contrast to being of order p for the direct MC simulation. This is indeed a very significant variance reduction in the regime p small, as already observed in [6], in a different context.

5 Numerical Results

In this section we investigate numerically the results of implementing the IPS procedure for estimating the probability mass function of the loss function for single names and multinames.

5.1 Single-Name Case

For the single-name case, we compute the probability of default for different values of the barrier using IPS and traditional Monte Carlo. In addition, for each method, we implemented the continuity correction for the barrier level described in [4] to account for the fact that we are using a discrete approximation to the continuous barrier for both IPS and MC. For the different values

of the barrier we use, we can calculate the exact probability of default from equation (2.2).

The following are the parameters we used for both IPS and MC.

r	σ	S_0	$\triangle t$	T	n (# of mutations in IPS)	M
.06	.25	80	.001	1	20	20000

The number of simulations M is the same for IPS and MC, and from an empirical investigation, we chose $\alpha = -18.5$ in the IPS method. The results are shown in Figure 5.1.
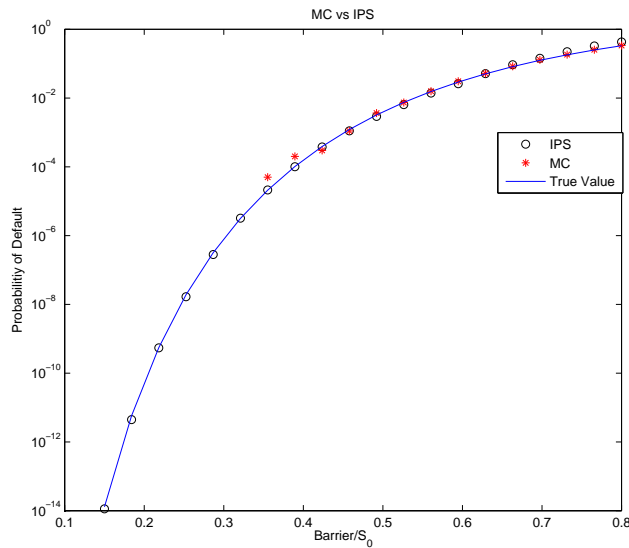
Fig. 5.1 Default probabilities for different barrier levels for IPS and MC

Indeed probabilities of order 10^{-14} will be irrelevant in the context of default probabilities but the user can see that IPS is capturing the rare events probabilities for the single name case whereas traditional Monte Carlo is not able to capture these values below 10^{-4}.

In Figure 5.2 we show how the variance decreases with the barrier level, and therefore with the default probability, for MC and IPS. In the IPS case the variance is obtained empirically and using the integral formulas derived in the Appendix. We deduce that the variance for IPS decreases as p^2 (p is the default probability), as opposed to p in the case of MC simulation.

Each MC and IPS simulation gives an estimate of the probability of default (whose theoretical value does not depend on the method) as well as an estimate of the standard deviation of the estimator (whose theoretical value does depend on the method). Therefore, it is instructive from a practical point of
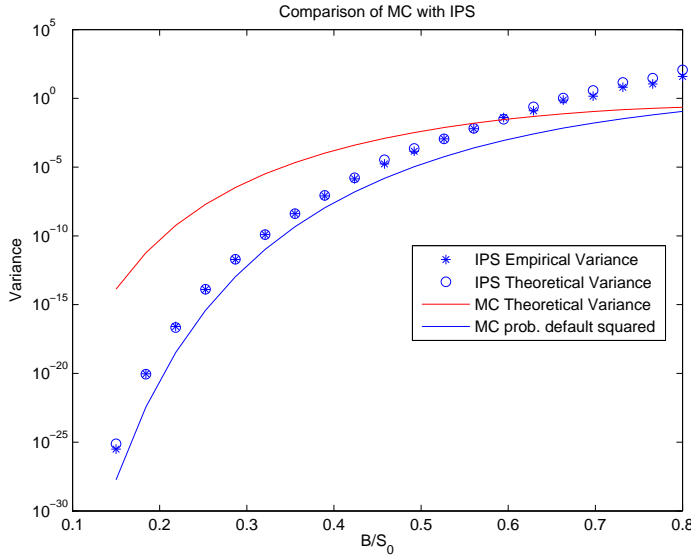
Fig. 5.2 Variances for different barrier levels for IPS and MC

view to compare the two methods by comparing the empirical ratios of their standard deviation to the probability of default for each method. If $p(B)$ is the probability of default for a certain barrier level B, then the standard deviation, $p_2(B)$, for traditional MC is given by,

$$p_2^{\mathrm{MC}}(B) = \sqrt{p(B)} \times \sqrt{(1 - p(B))},$$

and the theoretical ratio for MC is given by

$$\frac{p_2^{\mathrm{MC}}(B)}{p(B)} = \frac{\sqrt{(1 - p(B))}}{\sqrt{p(B)}},$$

which can computed using (2.2).

For IPS, the corresponding ratio is

$$\frac{p_2^{\mathrm{IPS}}(B)}{p(B)} = \frac{\sigma_n^B(\alpha)}{p(B)},$$

where $\sigma_n^B(\alpha)$ is given in Theorem 1. It is computed using the formula given in Corollary 1 in the Appendix.

In Figure 5.3 one sees that there are specific regimes where it is more efficient to use IPS as opposed to traditional MC for certain values of the barrier level (below $.6 \times S_0$). This is to be expected since IPS is well suited to rare event probabilities whereas MC is not.
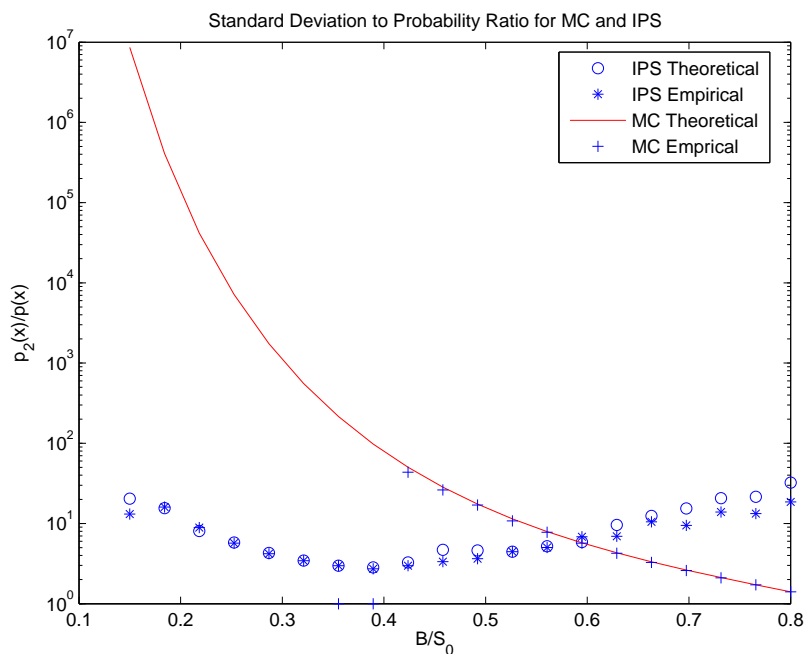
Fig. 5.3 Standard deviation-to-probability ratio for MC and IPS

5.2 Multiname Case

For the multiname case, we tested using 25 firms ($N = 25$). In addition, we took all of the firms to be homogeneous, meaning that they have the same parameters, starting value, and default barrier in (2.3). The following are the parameters that we used:

r	σ_i	$S_i(0)$	B_i	$\triangle t$	T	n	M
.06	.3	90	36	.001	1	20	10000

In addition, we took the correlation between the driving Brownian motions to be $\rho_{ij} = .4$ for $i \neq j$. The above parameters give, in the independent case, a probability of default of .0018, a realistic default probability for highly rated firms. For the IPS simulation we used $\alpha = -18.5/25$ so that it is consistent with the value used in the single-name case.

The following picture illustrates the difference between using MC and IPS to estimate the Loss probability mass function. That is, we calculate numerically $P(L(T) = k)$ where $L(T)$ is the number of firms that have defaulted before time T as defined in (2.5). In Figure 5.4 we plot the pmf on a log scale more adapted to our range of values. One can see that the IPS method is picking up more of the tail events for the pmf of the loss function and the distinction between IPS and MC becomes clear. in this regime, MC in only

good for estimating the pmf for $K = 0, 1, 2, 3, 4$, but IPS is good for estimating the pmf for $K = 0, 1, \ldots, 10$. Considering the fact that most contracts are written only on the first 40% of losses and $10 = .4 \times 25$, we see that IPS does a good job of describing the rare, but economically (in the sense of contracts) significant events.
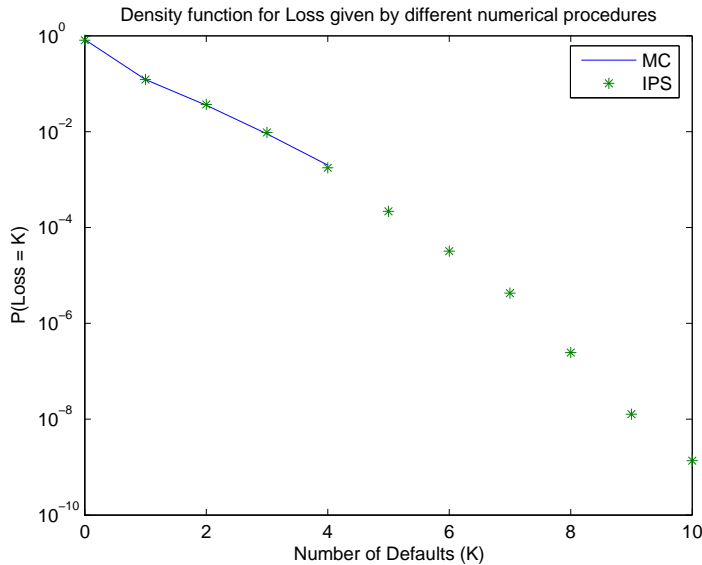

Fig. 5.4 Probability mass function of the Loss shown in log-scale, with $N = 25$.

6 Conclusion

In this paper, we adapted an original IPS approach to the computation of rare probabilities [6] to the field of credit risk under the first passage model. We showed that with our choice of weight function, IPS performs better than traditional MC methods for computing the probability of tail events in credit risk under a simple toy model. We also derived an explicit formula (up to double and triple integrals) for the asymptotic variance of the IPS estimator in the single-name case. In addition, we also showed that there are specific regimes where IPS is better suited to credit risk than traditional MC methods. For practical purposes our algorithm can be adapted to more complicated models, for instance, with stochastic volatility [10].

Appendix

A Formulas for the Variance in the Single-Name Constant Volatility Case

The asymptotic variance $\sigma_n^B(\alpha)^2$ of the IPS estimator is given by (4.6) in Theorem 1. As a Corollary, we deduce in this Appendix the explicit formulas used in Section 5.1.

Corollary 1. *Given the choice of weight function in (4.1) with $\alpha < 0$, and constant barrier B, we have that*

$$
\sigma_n^B(\alpha)^2
$$

$$
= \sum_{p=1}^{n} \left[f(\alpha, t_{p-1}) e^{-\frac{1}{2}\theta^2 t_p} \left(\int_{-\infty}^{(\frac{1}{\sigma})\ln(B/S_0)} \int_{y}^{0} \int_{y}^{\infty} e^{-\alpha\sigma x + \theta z} \Psi_{(t_{p-1},t_p)}(x,y,z)\,dz\,dx\,dy \right.\right.
$$

$$
+ \int_{-\infty}^{(\frac{1}{\sigma})\ln(B/S_0)} \int_{x}^{\infty} e^{-\alpha\sigma x + \theta z} \Upsilon_{(t_{p-1},t_p)}(x,z)\,dz\,dx
$$

$$
+ \int_{(\frac{1}{\sigma})\ln(B/S_0)}^{0} \int_{y}^{0} \int_{y}^{\infty} h(t_p,z)^2 e^{-\alpha\sigma x + \theta z} \Psi_{(t_{p-1},t_p)}(x,y,z)\,dz\,dx\,dy
$$

$$
\left.+ \int_{(\frac{1}{\sigma})\ln(B/S_0)}^{0} \int_{x}^{\infty} h(t_p,z)^2 e^{-\alpha\sigma x + \theta z} \Upsilon_{(t_{p-1},t_p)}(x,z)\,dz\,dx \right)
$$

$$
\left. - \left(1 - N(d_2^+(0,0)) + \left(\frac{S_0}{B}\right)^{1-\frac{2r}{\sigma^2}} N(d_2^-(0,0)) \right)^2 \right],
\tag{A.1}
$$

where,

$$
f(\alpha, t_{p-1}) = \frac{\alpha\sigma + \theta}{\alpha\sigma + 2\theta} e^{\alpha\sigma(\alpha\sigma + 2\theta)t_{p-1}/2} Erfc\left(\frac{(\alpha\sigma + \theta)\sqrt{t_{p-1}}}{\sqrt{2}} \right) + \frac{\theta}{\alpha\sigma + 2\theta} Erfc\left(\frac{-\theta\sqrt{t_{p-1}}}{\sqrt{2}} \right),
$$

$$
Erfc(x) = \frac{2}{\sqrt{\pi}} \int_{x}^{\infty} e^{-v^2}\,dv,
$$

$$
\theta = \frac{r - \frac{1}{2}\sigma^2}{\sigma},
$$

$$
\Psi_{(t_{p-1},t_p)}(x,y,z)
$$

$$
= \frac{2\sqrt{2}e^{-(2x-2y+z)^2/2t_p}}{\sqrt{\pi t_p t_{p-1}^2 (t_p - t_{p-1})^2}} \times \left[\frac{\sigma_{(t_{p-1},t_p)}(\mu_{(t_{p-1},t_p)}^{(1)}(x,y,z) - x + z - 2y)}{\sqrt{2\pi}} e^{-(x - \mu_{(t_{p-1},t_p)}^{(1)}(x,y,y))^2/2\sigma_{(t_{p-1},t_p)}^2} \right.
$$

$$
+ \left(\sigma_{(t_{p-1},t_p)}^2 + (\mu_{(t_{p-1},t_p)}^{(1)}(x,y,z))^2 \right.
$$

$$
\left.\left. + \mu_{(t_{p-1},t_p)}^{(1)}(x,y,z)(z - 2y - 2x) - 2x(z - 2y) \right) \left(1 - \Phi\left(\frac{x - \mu_{(t_{p-1},t_p)}^{(1)}(x,y,z)}{\sigma_{(t_{p-1},t_p)}} \right) \right) \right],
$$

$$\Upsilon_{(t_{p-1},t_p)}(x,z)$$

$$= \sqrt{\frac{2}{\pi t_p t_{p-1}{}^2}} \times \left[e^{-(2x-z)^2/2t_p} \left(\frac{\sigma_{(t_{p-1},t_p)}}{\sqrt{2\pi}} e^{-(x-\mu^{(2)}_{(t_{p-1},t_p)}(x,z))^2/2\sigma^2_{(t_{p-1},t_p)}} \right. \right.$$

$$\left. + (\mu^{(2)}_{(t_{p-1},t_p)}(x,z) - 2x)\left(1 - \Phi\left(\frac{x - \mu^{(2)}_{(t_{p-1},t_p)}(x,z)}{\sigma_{(t_{p-1},t_p)}}\right)\right) \right)$$

$$- e^{-z^2/2t_p}\left(\frac{\sigma_{(t_{p-1},t_p)}}{\sqrt{2\pi}} e^{-(x-\mu^{(3)}_{(t_{p-1},t_p)}(x,z))^2/2\sigma^2_{(t_{p-1},t_p)}} \right.$$

$$\left. \left. + (\mu^{(3)}_{(t_{p-1},t_p)}(x,z) - 2x)\left(1 - \Phi\left(\frac{x - \mu^{(3)}_{(t_{p-1},t_p)}(x,z)}{\sigma_{(t_{p-1},t_p)}}\right)\right) \right) \right],$$

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy,$$

$$\mu^{(1)}_{(t_{p-1},t_p)}(x,y,z) = \frac{2x(t_p - t_{p-1}) + (2y - z)t_{p-1}}{t_p},$$

$$\mu^{(2)}_{(t_{p-1},t_p)}(x,z) = \frac{2x(t_p - t_{p-1}) + zt_{p-1}}{t_p},$$

$$\mu^{(3)}_{(t_{p-1},t_p)}(x,z) = \frac{2xt_p - zt_{p-1}}{t_p},$$

$$\sigma^2_{(t_{p-1},t_p)} = \frac{t_{p-1}}{t_p}(t_p - t_{p-1}),$$

$$h(t_p,z) = \left(1 - N(d_2^+(t_p,z)) + \left(\frac{S_0 e^{\sigma z}}{B}\right)^{1-\frac{2r}{\sigma^2}} N(d_2^-(t_p,z)) \right),$$

$$d_2^{\pm}(t_p,z) = \frac{\pm(\ln(S_0) + \sigma z - \ln(B)) + (r - \frac{1}{2}\sigma^2)(T - t_p)}{\sigma\sqrt{T - t_p}}.$$

Proof As can be seen in (4.6), we need to compute the joint distribution of

$$\left(\min_{u \leq t_{p-1}} S(u), \min_{u \leq t_p} S(u), S(t_p) \right).$$

First, recall that since $S(u) = S_0 e^{(r - \frac{1}{2}\sigma^2)u + \sigma W_u}$, $\sigma > 0$ by assumption, and log is an increasing function we have:

$$\mathbb{E}\left\{ e^{\alpha \log(\min_{u \leq t_{p-1}} S(u))} \right\} = \mathbb{E}\left\{ e^{\alpha \log\left(\min_{u \leq t_{p-1}} S_0 e^{(r - \frac{1}{2}\sigma^2)u + \sigma W_u}\right)} \right\}$$

$$= S_0^{\alpha} \mathbb{E}\left\{ e^{\alpha(\min_{u \leq t_{p-1}}(r - \frac{1}{2})u + \sigma W_u)} \right\}$$

$$= S_0^{\alpha} \mathbb{E}\left\{ e^{\alpha\sigma(\min_{u \leq t_{p-1}} \widehat{W}_u)} \right\},$$

where $\widehat{W}_u = \theta u + W_u$ is a Brownian motion with deterministic drift $\theta = \left(\frac{r - \frac{1}{2}\sigma^2}{\sigma}\right)$. Therefore, the above computation simplifies to computing the moment generating function of the running minimum of Brownian motion with drift. This formula is well known (see for

instance [3]) and so we have,

$$\mathbb{E}\left\{ e^{\alpha \log(\min_{u \le t_{p-1}} S(u))} \right\}$$

$$= S_0^\alpha \left(\frac{\alpha\sigma + \theta}{\alpha\sigma + 2\theta} e^{\alpha\sigma(\alpha\sigma+2\theta)t_{p-1}/2} \mathrm{Erfc}\left(\frac{(\alpha\sigma+\theta)\sqrt{t_{p-1}}}{\sqrt{2}} \right) + \frac{\theta}{\alpha\sigma+2\theta}\mathrm{Erfc}\left(\frac{-\theta\sqrt{t_{p-1}}}{\sqrt{2}} \right) \right)$$

$$:= S_0^\alpha f(\alpha, t_{p-1}), \tag{A.2}$$

where

$$\mathrm{Erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-v^2}\, dv.$$

Now, we compute $\mathbb{E}\left\{ P_{B,p,n}^2 e^{-\alpha \log(\min_{u \le t_{p-1}} S(u))} \right\}$. The general goal will be to write everything in terms of expectations of functionals of $\widehat{W}_u$. First, we note that

$$P_{B,p,n}(x) = \mathbb{E}\left\{ \mathbf{1}_{\min_{u \le T} S(u) \le B} | X_p = x \right\}$$

$$= \mathbb{E}\left\{ \mathbf{1}_{\min_{u \le T} S(u) \le B} | X_p = (\min_{u \le t_p} S(u), S(t_p), \mathbf{1}_{\tau \le t_p}) \right\}$$

$$= \mathbf{1}_{\min_{u \le t_p} S(u) \le B} + \mathbf{1}_{\min_{u \le t_p} S(u) > B} \mathbb{P}(\min_{t_p \le u \le T} S(u) \le B | S(t_p))$$

$$= \mathbf{1}_{\min_{u \le t_p} S(u) \le B} + \mathbf{1}_{\min_{u \le t_p} S(u) > B} \left(1 - N(d_2^+) + \left(\frac{S(t_p)}{B} \right)^{1-\frac{2r}{\sigma^2}} N(d_2^-) \right),$$

where

$$d_2^\pm = \frac{\pm \ln\left(\frac{S(t_p)}{B} \right) + (r - \frac{1}{2}\sigma^2)(T - t_p)}{\sigma\sqrt{T - t_p}}.$$

In the formula for $d_2^\pm$ we will find it useful to substitute the formula $S(t_p) = S_0 e^{\sigma \widehat{W}_{t_p}}$ and write the dependence on t_p and $\widehat{W}_{t_p}$ explicitly as,

$$d_2^\pm(t_p, \widehat{W}_{t_p}) = \frac{\pm \ln\left(\frac{S_0 e^{\sigma \widehat{W}_{t_p}}}{B} \right) + (r - \frac{1}{2}\sigma^2)(T - t_p)}{\sigma\sqrt{T - t_p}}$$

$$= \frac{\pm \left(\ln(S_0) + \sigma\widehat{W}_{t_p} - \ln(B) \right) + (r - \frac{1}{2}\sigma^2)(T - t_p)}{\sigma\sqrt{T - t_p}},$$

In addition, we also substitute $S(u) = S_0 e^{\sigma \widehat{W}_u}$ into the expression for $P_{B,p,n}$ and rearrange to get,

$$P_{B,p,n}(x) = \mathbf{1}_{\min_{u \le t_p} \widehat{W}_u \le (\frac{1}{\sigma}) \ln(B/S_0)}$$

$$+ \mathbf{1}_{\min_{u \le t_p} \widehat{W}_u > (\frac{1}{\sigma}) \ln(B/S_0)} \left(1 - N(d_2^+(t_p, \widehat{W}_{t_p})) + \left(\frac{S_0 e^{\sigma \widehat{W}_{t_p}}}{B} \right)^{1-\frac{2r}{\sigma^2}} N(d_2^-(t_p, \widehat{W}_{t_p})) \right)$$

Hence,

$$P_{B,p,n}(x)^2 = \mathbf{1}_{\min_{u \le t_p} \widehat{W}_u \le (\frac{1}{\sigma}) \ln(B/S_0)} + \mathbf{1}_{\min_{u \le t_p} \widehat{W}_u > (\frac{1}{\sigma}) \ln(B/S_0)} h(t_p, \widehat{W}_{t_p})^2,$$

where

$$h(t_p, \widehat{W}_{t_p}) = \left(1 - N(d_2^+(t_p, \widehat{W}_{t_p})) + \left(\frac{S_0 e^{\sigma \widehat{W}_{t_p}}}{B} \right)^{1-\frac{2r}{\sigma^2}} N(d_2^-(t_p, \widehat{W}_{t_p})) \right).$$

Hence, plugging in the expression for $P_{B,p,n}^2$ into $\mathbb{E}\left\{P_{B,p,n}^2 e^{-\alpha \log(\min_{u\le t_{p-1}} S(u))}\right\}$ we have

$$\mathbb{E}\left\{P_{B,p,n}^2 e^{-\alpha \log(\min_{u\le t_{p-1}} S(u))}\right\}$$

$$= S_0^{-\alpha}\mathbb{E}\left\{\mathbf{1}_{\min_{u\le t_p}\widehat{W}_u\le(\frac{1}{\sigma})\ln(B/S_0)} e^{-\alpha\sigma \min_{u\le t_{p-1}}\widehat{W}_u}\right\}$$

$$+ S_0^{-\alpha}\mathbb{E}\left\{\mathbf{1}_{\min_{u\le t_p}\widehat{W}_u>(\frac{1}{\sigma})\ln(B/S_0)} h(t_p,\widehat{W}_{t_p})^2 e^{-\alpha\sigma \min_{u\le t_{p-1}}\widehat{W}_u}\right\}, \qquad \text{(A.3)}$$

where the expectation above is taken with respect to the measure $\mathbb{P}$ for which $\widehat{W}_u$ is a Brownian motion with drift. Recall that under $\mathbb{P}$,

$$d\widehat{W}_t = \theta dt + dW_t$$

where W_t is a $\mathbb{P}$ standard Brownian motion and $\theta = \left(\frac{r-\frac{1}{2}\sigma^2}{\sigma}\right)$. Using Girsanov's theorem (see [13] or [16]), $\widehat{W}_t$ is a standard Brownian motion under $\widehat{\mathbb{P}}$ and the Radon-Nikodym density, $Z(t)$, is given by

$$Z(t) = \exp\left\{-\int_0^t \theta dW_u - \frac{1}{2}\int_0^t \theta^2 du\right\}$$

$$= \exp\left\{-\int_0^t \theta(d\widehat{W}_u - \theta u) - \frac{1}{2}\theta^2 t\right\}$$

$$= \exp\left\{-\theta\widehat{W}_t + \frac{1}{2}\theta^2 t\right\}.$$

Therefore, we rewrite (A.3) as an expectation under $\widehat{\mathbb{P}}$ to get

$$\mathbb{E}\left\{P_{B,p,n}^2 e^{-\alpha \log(\min_{u\le t_{p-1}} S(u))}\right\}$$

$$= S_0^{-\alpha}\widehat{\mathbb{E}}\left\{\mathbf{1}_{\min_{u\le t_p}\widehat{W}_u\le(\frac{1}{\sigma})\ln(B/S_0)} e^{-\alpha\sigma \min_{u\le t_{p-1}}\widehat{W}_u} Z(t_p)^{-1}\right\}$$

$$+ S_0^{-\alpha}\widehat{\mathbb{E}}\left\{\mathbf{1}_{\min_{u\le t_p}\widehat{W}_u>(\frac{1}{\sigma})\ln(B/S_0)} h(t_p,\widehat{W}_{t_p})^2 e^{-\alpha\sigma \min_{u\le t_{p-1}}\widehat{W}_u} Z(t_p)^{-1}\right\}$$

$$= S_0^{-\alpha}\widehat{\mathbb{E}}\left\{\mathbf{1}_{\min_{u\le t_p}\widehat{W}_u\le(\frac{1}{\sigma})\ln(B/S_0)} e^{-\alpha\sigma \min_{u\le t_{p-1}}\widehat{W}_u} e^{\theta\widehat{W}_{t_p}-\frac{1}{2}\theta^2 t_p}\right\}$$

$$+ S_0^{-\alpha}\widehat{\mathbb{E}}\left\{\mathbf{1}_{\min_{u\le t_p}\widehat{W}_u>(\frac{1}{\sigma})\ln(B/S_0)} h(t_p,\widehat{W}_{t_p})^2 e^{-\alpha\sigma \min_{u\le t_{p-1}}\widehat{W}_u} e^{\theta\widehat{W}_{t_p}-\frac{1}{2}\theta^2 t_p}\right\}$$

$$= S_0^{-\alpha} e^{-\frac{1}{2}\theta^2 t_p}\left(\int\int\int \mathbf{1}_{y\le(\frac{1}{\sigma})\ln(B/S_0)} e^{-\alpha\sigma x+\theta z}\Gamma_{(t_{p-1},t_p)}(dx,dy,dz)\right.$$

$$+ \left.\int\int\int \mathbf{1}_{y>\frac{1}{\sigma}\ln(B/S_0)} h(t_p,z)^2 e^{-\alpha\sigma x+\theta z}\Gamma_{(t_{p-1},t_p)}(dx,dy,dz)\right), \qquad \text{(A.4)}$$

where,

$$\Gamma_{(t_{p-1},t_p)}(dx,dy,dz) = \widehat{\mathbb{P}}(\min_{u\le t_{p-1}}\widehat{W}_u\in dx, \min_{u\le t_p}\widehat{W}_u\in dy, \widehat{W}_{t_p}\in dz), \qquad \text{(A.5)}$$

and as stated before, $\widehat{W}_u$ is a standard Brownian motion under $\widehat{\mathbb{P}}$.

The formula for $\Gamma_{(t_{p-1},t_p)}(dx,dy,dz)$ is given by

$$\Gamma_{(t_{p-1},t_p)}(dx,dy,dz) = \Psi_{(t_{p-1},t_p)}(x,y,z)\mathbf{1}_{y<x,y\leq z,x\leq 0}dzdydx$$
$$+\Upsilon_{(t_{p-1},t_p)}(x,z)\mathbf{1}_{x\leq z,x\leq 0}\delta_x(dy)dzdx,$$

where the functions $\Psi_{(t_{p-1},t_p)}$ and $\Upsilon_{(t_{p-1},t_p)}$ are given in Corollary 1.

The derivation of these formulas is obtained by:

1. Introducing the value of $\widehat{W}_u$ at the intermediate time t_{p-1} in (A.5).
2. Using the Markov property at present time t_{p-1}.
3. Using the classical joint distribution of a Brownian motion and its running minimum.
4. Re-integrating with respect to $\widehat{W}_{t_{p-1}}$.

The details of this derivation are in [17]. Substituting the formulas for $\Psi_{(t_{p-1},t_p)}$ and $\Upsilon_{(t_{p-1},t_p)}$ into (A.4) ends the proof of Corollary 1.

References

1. A. BASSAMBOO AND S. JAIN, *Efficient importance sampling for reduced form models in credit risk*, in WSC '06: Proceedings of the 38th conference on Winter simulation, Winter Simulation Conference, 2006, pp. 741–748.
2. F. BLACK AND J. C. COX, *Valuing corporate securities: Some effects of bond indenture provisions*, Journal of Finance, 31 (1976), pp. 351–67.
3. A. BORODIN AND P. SALMINEN, *Handbook of Brownian motion: Facts and formulae*, Basel: Birkauser, 2002.
4. M. BROADIE, P. GLASSERMAN, AND S. KOU, *A continuity correction for discrete barrier options*, Mathematical Finance, 7 (1997), pp. 325–349.
5. P. DEL MORAL, *Feynman-Kac Formulae: Genealogical and Interacting Particle Systems with Applications*, Springer-Verlag, 2004.
6. P. DEL MORAL AND J. GARNIER, *Genealogical particle analysis of rare events*, Annals of Applied Probability, 15 (2005), pp. 2496–2534.
7. A. ELIZALDE, *Credit risk models iv: Understanding and pricing cdos*, working papers, CEMFI, Apr. 2006.
8. Y. EOM, J. HELWEGE, AND J. Z. HUANG, *Structural models of corporate bond pricing: An empirical analysis*, Review of Financial Studies, 17 (2004), pp. 499–544.
9. J.-P. FOUQUE, R. SIRCAR, AND K. SOLNA, *Stochastic volatility effects on defaultable bonds*, Applied Mathematical Finance, 13 (2006), pp. 215–244.
10. J.-P. FOUQUE, B. WIGNALL, AND X. ZHOU, *Modeling correlated defaults: First passage model under stochastic volatility*, Journal of Computational Finance, 11 (2008).
11. P. GLASSERMAN, *Monte Carlo Methods in Financial Engineering*, Springer, 2003.
12. P. GLASSERMAN AND J. LI, *Importance sampling for portfolio credit risk*, Management Science, 51 (2005), pp. 1643–1656.
13. I. KARATAS AND S. E. SHREVE, *Brownian Motion and Stochastic Calculus*, Springer, 1991.
14. R. C. MERTON, *On the pricing of corporate debt: the risk structure of interest rates*, Journal of Finance, 29 (1974), pp. 449–470.
15. P. J. SCHONBUCHER, *Credit Derivative Pricing Models: Models, Pricing and Implementation*, Wiley Finance Series, 2003.
16. S. E. SHREVE, *Stochastic Calculus for Finance II: Continuous-Time Models*, Springer, 2003.
17. D. VESTAL, *Interacting particle systems for pricing credit derivatives*, PhD Dissertation, University of California Santa Barbara, (2008).
18. C. ZHOU, *An analysis of default correlations and multiple defaults*, Review of Financial Studies, 14 (2001), pp. 555–76.