Rare Event Simulation using Interacting Particle Systems for Rare Credit Portfolio Losses

Abhishek Shah
ans556@nyu.edu
Computer Science, NYU

Prithvi Krishna Gattamaneni
pkg238@nyu.edu
Computer Science, NYU

Raghav Singhal
rs4070@nyu.edu
Mathematics, NYU

Srivas Venkatesh
sv1358@nyu.edu
Computer Science, NYU

Abstract

With the recent explosion of the credit market, it is imperative to understand the risk profiles of these large credit portfolios. This helps segment and structure the portfolio based on the risk and price it appropriately. However this is a very high dimensional problem and suffers from the typical issues of such problems. Moreover the joint probability of multiple defaults occurring is very low and hence estimating such probabilities is harder too. In this project we explore Interacting Particle Systems methodology to sample these rare events and estimate the default probabilities from these simulations. We also explore the use of a few different models for the credit portfolios.

1 Motivation

We assume that we have a Markov chain $S = (S_n)_{0 \leq n \leq T}$. At each time step n, we have d correlated assets or sources. $S_n = (S_n^1, ..., S_n^d) \in E$. We wish to understand the nature of probabilities of rare events which are of the form $L(T) > K$, where L_T is the number of defaults at the maturity time T, can be expressed as below.

$$\{L(T) \geq K\} = \{V_T(S_T) \geq K\} = \{V_T(S_0, ..., S_T) \geq K\}$$

$V(T)$ can be thought of as a real positive function that is a risk measure of these rare events. Calculating $\mathbb{P}(V_T(S_T) \geq K)$ through the standard Monte Carlo procedure is not feasible because of the difficulty faced in ensuring that we generate a large number of samples to realize the rare event. One partial remedy is to provide a tight upper bound based on large deviation ideas. For any $\lambda \geq 0$ we have :

$$\mathbb{P}(V_T(S_T) \geq K) = \mathbb{E}\left(\mathbf{1}_{\{V_T(S_T) \geq K\}} e^{\lambda V_T(S_T)} e^{-\lambda V_T(S_T)}\right) \leq e^{-\lambda K} \mathbb{E}\left(\mathbf{1}_{\{V_T(S_T) \geq K\}} e^{\lambda V_T(S_T)}\right) \tag{1}$$

If we denote $\mathbb{E}^\lambda$ the expectation under the probability $\mathbb{P}^{(\lambda)}$ defined by

$$d\mathbb{P}^{(\lambda)} \propto e^{\lambda V_T(S_T)} d\mathbb{P}$$

we get

$$\mathbb{E}\left(\mathbf{1}_{\{V_T(S_T) \geq K\}} e^{\lambda V_T(S_T)}\right) = \mathbb{E}^{(\lambda)}\left(\mathbf{1}_{\{V_T(S_T) \geq K\}}\right) \mathbb{E}\left(e^{\lambda V_T(S_T)}\right)$$

and we know that

$$\mathbb{E}^{(\lambda)}\left(\mathbf{1}_{\{V_T(S_T) \geq K\}}\right) \leq 1$$

from which we get

$$\mathbb{P}\left(V_T(S_T) \geq K\right) \leq e^{-\sup_{\lambda \geq 0}(\lambda K - \Lambda(\lambda))} \tag{2}$$

where $\Lambda(\lambda)$ is the Fenchel Transformation which is $\log(\mathbb{E}(\lambda V_T(S_T)))$.
Examining the above, we observe that the desired probability can be approximated by searching for a proper λ. This large deviation type approach is widely used, but in the above form, it requires extensive calculations in order to obtain a reasonable approximation of the desired probability.

Del Moral and Garnier provide in Del Moral and Garnier [2005] a zero bias estimate with interacting particle systems. The approach is to construct a genealogical tree based model instead of the above large deviation type inequality. Using the same change of measure from $\mathbb{P}$ to $\mathbb{P}^{(\lambda)}$ we can say that the target probability

$$\mathbb{P}(V_T(S_T) \geq K) = \mathbb{E}\left(\mathbf{1}_{\{V_T(S_T)\geq K\}} e^{\lambda V_T(S_T)} e^{-\lambda V_T(S_T)}\right)$$

can be rewritten as

$$\mathbb{E}^{(\lambda)}\left(\mathbf{1}_{\{V_T(S_T)\geq K\}} e^{-\lambda V_T(S_T)}\right) \mathbb{E}\left(e^{\lambda V_T(S_T)}\right) = \mathbb{E}^{(\lambda)}(f_T(S_T))\mathbb{E}(e^{\lambda V_T(S_T)})$$

where $f_t(S_T) := \mathbf{1}_{\{V_T(S_T)\geq K\}} e^{-\lambda V_T(S_T)}$. With the convention that $V_0 = 0$, we get the following decomposition

$$e^{\lambda V_T(S_T)} \equiv \prod_{p=1}^{T} e^{\lambda(V_p(S_p) - V_{p-1}(S_{p-1}))}$$

By using the notation

$$\mathcal{X}_k = (S_k, S_{k+1})$$

for $0 \leq k < T$, the above produce can be rewritten as

$$\prod_{p=1}^{T} G_{p-1}(\mathcal{X}_{p-1})$$

where

$$G_{p-1}(\mathcal{X}_{p-1}) := e^{\lambda(V_p(S_p) - V_{p-1}(S_{p-1}))}$$

Using the notation that $F_T(\mathcal{X}_T) = f_T(S_T)$ we get

$$\mathbb{E}^{(\lambda)}(f_T(S_T)) = \frac{\mathbb{E}(F_T(\mathcal{X}_T)\prod_{p=1}^{T} G_p(\mathcal{X}_p))}{\mathbb{E}(\prod_{p=1}^{T} G_p(\mathcal{X}_p))} := \eta_T(F_T) \tag{3}$$

These quantities above can be approximated efficiently by interacting particle systems.

2 Theory

2.1 Interacting Particle Systems

Particle methods depend upon the existence of a background Markov chain which is denoted by $X = X_n n \geq 0$. This chain is not assumed to be time homogeneous. The random element X_n takes values in some measurable state space (E_n, ε_n) that can change with n. $K_n(x_{n-1}, dx_n)$ is the Markov transition kernels. $B_b(E)$ is the space of bounded measurable functions of the measurable space (E, ε)

2.1.1 Feynman-Kac Path Expectations

Let Y_n be the history of X_n.

$$Y_n := (X_0, ..., X_n) \in F_n := (E_0 \times ... \times E_n), n \geq 0$$

$Y_{n n\geq 0}$ is itself a Markov chain and we denote by $M_n(y_{n-1}, dy_n)$ its transition kernel. For each $n \geq 0$, we choose a multiplicative potential function G_n defined on F_n and we define the Feynman-Kac expectations by

$$\gamma_n(f_n) = \mathbb{E}\left[f(Y_n)\prod_{1\leq p<n} G_k(Y_p)\right] \tag{4}$$

We define $\eta_n(.)$ the corresponding normalized measure defined as

$$\eta_n(f_n) = \frac{\mathbb{E}\left[f_n(Y_n)\prod_{1\leq k<n} G_k(Y_k)\right]}{\mathbb{E}\left[\prod_{1\leq k<n} G_k(Y_k)\right]} = \gamma_n(f_n)/\gamma_n(1). \tag{5}$$

We also observe that

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

Hence given any bounded measurable function f_n on F_n we have

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

Using the notation $G_p^- = 1/G_p$ and the above definitions of γ_n and η_n we see that

$$\begin{aligned}
\mathbb{E}[f_n(Y_n)] &= \mathbb{E}\left[f_n(Y_n)\prod_{1\leq p<n} G_p^-(Y_p) \times \prod_{1\leq p<n} G_p(Y_p)\right] &(6)\\
&= \gamma_n\left(f_n\prod_{1\leq p<n} G_p^-\right) &(7)\\
&= \eta_n\left(f_n\prod_{1\leq p<n} G_p^-\right)\prod_{1\leq p<n}\eta_p(G_p) &(8)
\end{aligned}$$

We can also check by inspection that the measures $(\eta_n)_{n\geq 1}$ satisfy the nonlinear recursive equation.

$$\eta_n = \Phi_n(\eta_{n-1}) := \int_{F_{n-1}} \eta_{n-1}(dy_{n-1})\frac{G_{n-1}(y_{n-1})}{\eta_{n-1}(G_{n-1})}M_n(y_{n-1}, .)$$

starting from $\eta_1 = M_1(x_0, .)$. This dynamical equation on the space of measures is known as Stettners equation in filtering theory. We state it to justify the selection/mutation decomposition of each step of the particle algorithm introduced below.

2.1.2 IPS Interpretation and Monte Carlo Algorithm

We introduce a natural interacting path-particle system. For a given integer M, using the transformation Φ_n, we construct a Markov chain $\{\xi_n\}_{n\geq 0}$ whose state $\xi_n = (\xi_n^j)_{1\leq j\leq M}$ at time n can be interpreted as a set of M Monte Carlo samples of path-particles

$$\xi_n^j = (\xi_{0,n}^j, \xi_{1,n}^j, ..., \xi_{n,n}^j) \in F_n = (E_0 \times ... \times E_n)$$

The transition mechanism of this Markov chain can be described as follows. We start with an initial configuration $\xi_1 = (\xi_1^j)_{1\leq j\leq M}$ 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 $\xi_1^j := (\xi_{0,1}^j, \xi_{1,1}^j) = (x_0, \xi_{1,1}^j) \in F_1 = (E_0 \times E_1)$ where the $\xi_{1,1}^j$ are drawn independently of each other from the distribution $K_1(x_0,.)$. Then the one-step transition taking $\xi_{n-1} \in F_{n-1}^M$ into $\xi \in F_n^M$ is given by a random draw from the distribution

$$\mathbb{P}\big(\xi_n \in d(y_n^1, ..., y_n^M)|\xi_{n-1}\big) = \prod_{j=1}^{M} \Phi(m(\xi_{n-1}))(dy_n^j) \tag{9}$$

where the notation $m(\xi_{n-1})$ is used for the empirical distribution of the ξ_{n-1}^j, i.e

$$m(\xi_{n-1}) := \frac{1}{M}\sum_{j=1}^{M} \delta_{\xi_{n-1}^j}$$

From the definition of Φ_n, we see that the above is the superposition of a section procedure followed by a mutation given by the transition of the original Markov chain. More precisely,

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

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

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

according to the Boltzmann-Gibbs measure

$$\sum_{j=1}^{M} \frac{G_{n-1}(\xi_{0,n-1}^j, ..., \xi_{n-1,n-1}^j)}{\sum_{k=1}^{M} G_{n-1}(\xi_{0,n-1}^k, ..., \xi_{n-1,n-1}^k)} \delta_{(\xi_{0,n-1}^j, ..., \xi_{n-1,n-1}^j)} \tag{10}$$

and for the mutation stage, each selected part-particle $\hat{\xi}_{n-1}^j$ is extended as follows.

$$\begin{aligned} \xi_n^j &= (\hat{\xi}_{n-1}^j, \xi_{n,n}^j) \\ &= ((\hat{\xi}_{0,n-1}^j, ..., \hat{\xi}_{n-1,n-1}^j), \xi_{n,n}^j) \in F_n = F_{n-1} \times E_n \end{aligned}$$

where $\xi_{n,n}^j$ is a random variable with distribution $K_n(\hat{\xi}_{n-1,n-1}, .)$. In other words, the transition step is a mere extension of the path particle with an element drawn at random using the transition kernel K_n of the original Markov chain. All of the mutations are performed independently. But most importantly all of these mutations are happening with the original transition distribution of the chain. This is a contrast with importance sampling where the Monte Carlo transitions are from twisted transition distributions obtained from a Girsanov-like change of measure. So from a practical approach, a black-box providing random samples from the original chain transition distribution is enough for the implementation of the IPS algorithm.

Another result states that for each fixed time n, the empirical historical path measure

$$\eta_n^M := m(\xi_n) = \frac{1}{M}\sum_{j=1}^{M} \delta_{\xi_{0,n}^j, \xi_{1,n}^j, ..., \xi_{n,n}^j}$$

converges in distribution as $M \to \infty$, toward the normalized Feynman-Kac measure η_n. Moreover there are several propagation of chaos estimates that ensure that $(\xi_{0,n}^j, \xi_{1,n}^j, ..., \xi_{n,n}^j)$ are asymptotically independent and identically distributed with distribution η_n. This justifies for each measurable function $\tilde{f}_n$ on F_n the choice of

$$\gamma_n^M(\tilde{f}_n) = \eta_n^M(\tilde{f}_n) \prod_{1 \le p < n} \eta_p^M(G_p) \tag{11}$$

for a particle approximation of the expectation $\gamma_n(\tilde{f}_n)$

2.2 Merton's Model with Stochastic Volatility

This is a model for structural credit risk based on Merton's model with a stochastic volatility term. We then apply IPS system on this model to estimate the rare default probabilities as described in Carmona et al. [2009]. The details of the model and application of IPS to this model is mentioned below.

2.2.1 Credit Portfolio Model

Given a portfolio of credit instruments related to N firms, where each underlying asset evolves according to the following SDE:

$$dS_i(t) = rS_i(t)dt + \sigma_i\sigma(t)S_i(t)dW_i(t) \tag{12}$$

where r is the risk-free interest rate, σ_i is a non-random volatility factor, and the correlation structure of the driving Wiener processes W_i is given by:

$$d\langle W_i, W_j\rangle_t = \rho_{ij}dt \tag{13}$$

and $\sigma(t)$ evolves according to another stochastic differential equation:

$$d\sigma(t) = \kappa(\bar{\sigma} - \sigma(t))dt + \gamma\sqrt{\sigma(t)}dW(t) \tag{14}$$

where $\kappa, \bar{\sigma}, \gamma$ are constants and the Wiener Process satisfies $\forall i = 1, 2....., N$:

$$d\langle W_i, W\rangle_t = \rho_\sigma dt \tag{15}$$

Now, for each asset we take a deterministic barrier, $B_i(t)$, or in other words a threshold , so that if the asset price falls under that barrier price at any time, the firm defaults. We then define a stopping time τ_i:

$$\tau_i = inf\left\{t \geq 0 : S_i(t) \leq B_i(t)\right\} \tag{16}$$

We now define the Portfolio Loss Function $L(t)$ as the number of defaults till a given time t:

$$L(t) = \sum_{i=0}^{n} \mathbf{1}_{\{\tau_i \leq t\}} \tag{17}$$

Since the spreads of CDO tranches are derived from the knowledge of a finite number of expectations of the form:

$$\mathbf{E}[(L(T) - K)^+] \tag{18}$$

where T is the coupon payment date and K is an acceptable number of defaults, beyond which we start accumulating losses. So to evaluate such expectation, we estimate the probabilities of default. For that purpose, we evaluate $\forall k = 0, 1, ..., N$

$$\mathbb{P}(L(T) = k) = \mathbf{p}_k(T) \tag{19}$$

2.2.2 Discretization of the Model

For the implementation of our algorithm and for computational efficiency we select two time step, $\Delta t = \frac{1}{20}$ which is used to perform the selection step, and $\delta t = 10^{-3}$ which will be used in the Euler Step.

The Markov Chain that we simulate is given as (Note that X_n is $2N + 1$ dimensional):

$$X_n = \left(\sigma\left(n\Delta t\right), \left(S_i\left(n\Delta t\right)\right)_{1\leq i\leq N}, \min_{0\leq m\leq n}\left(\left(S_i\left(m\Delta t\right)\right)\right)\right) \tag{20}$$

A remarkable and distinctive feature of the Interacting Particle System approach is that the evolution dynamics of the underlying process is preserved, that is the Markov Chain X_n follows the same evolution dynamics as the continuous Model.

We also take a constant Barrier, $B_i = 36$, and define the stopping time τ_i as :

$$\tau_i = \min\{n \geq 0 : S_i(n\Delta t) \leq B_i\} \tag{21}$$

Now, we define the potential function such that we assign more weight to portfolios with lower values, that is we assign more weight to rare events so the likelihood of defaults increases. The potential is a function of X_p and another parameter $\alpha > 0$:

$$G_p(Y_p) = \exp[-\alpha(V(X_p) - V(X_{p-1}))] \tag{22}$$

where $V(X_p) = \sum_{i=1}^{N} \log(min_{0\leq m\leq p} S_i(m\Delta t))$ so the potential can be written as:

$$G_p(Y_p) = \exp\left[-\alpha \sum_{i=1}^{N} \log \frac{min_{0\leq m\leq p} S_i(m\Delta t)}{min_{0\leq m\leq p-1} S_i(m\Delta t)}\right] \tag{23}$$

where $Y_p = (X_0, X_1, ..., X_p)$ but we only need the last two values not the earlier values. Notice that different values of α will give different Loss Distributions $\mathbb{P}(L(T) = k)$ for all k, because in the selection step those portfolios with lower number of defaults are assigned a lower weight leading to different number of defaults in each portfolio. The choice of the potential function and the parameter α lead to enough number of sample paths with large number of defaults even if the number of samples is significantly lower than what would be required by a plain Monte Carlo simpler. We however follow an idea mentioned in Carmona and CRÉPEY [2009], where the best α is selected for each k.

2.2.3 Single Asset Constant Volatility Analysis

The simplest case of the above model is a case of single asset and no stochastic volatility. As we know the exact solution for Geometric Brownian Motion, we can obtain the hitting time distribution analytically. This can then be used to compare against estimated values. The solution for the Geometric Brownian Motion is as follows.

$$S_t = S_0 \exp\left(\left(r - \frac{\sigma^2}{2}\right)t + \sigma W_t\right) \tag{24}$$

which implies that

$$\begin{aligned}\mathbb{P}[\tau_B \leq T] = \mathbb{P}[\min_{t\leq T} S_t \leq B] &= \mathbb{P}[\min_{t\leq T}(r - \frac{\sigma^2}{2})t + \sigma W_t \leq \log\frac{B}{S_0}] \\ &= \mathbb{P}[\min_{t\leq T}(r - \frac{\sigma^2}{2})t + \sigma\sqrt{t}N(0,1) \leq \log\frac{B}{S_0}]\end{aligned} \tag{25}$$

Now, using Girsonov's Theorem, we can get an analytic expression for the distribution above:

$$\mathbb{P}[\tau_B \leq T] = 1 - \left(\Phi(d^+) - \left(\frac{S_0}{B}\right)^p \Phi(d^-)\right) \tag{26}$$

where :

$$d^+ = \frac{\log\frac{S_0}{B} + (r - \frac{\sigma^2}{2})t}{\sigma\sqrt{T}}$$

$$d^- = \frac{-\log\frac{S_0}{B} + (r - \frac{\sigma^2}{2})t}{\sigma\sqrt{T}}$$

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

and Φ is the cumulative distribution function for a standard normal.

2.3 Local Intensity Model

Local Intensity model belongs to family of Markovian intensity models of credit risk. This model is applied using IPS framework as mentioned in Carmona and CRÉPEY [2009].

2.3.1 Local Intensity Model as Markov Point Process

Assume there are n names in the credit risk portfolio. The calculation of cumulative loss $L = L_t; t \geq 0$ in the credit portfolio is modeled as a Markov point process as mentioned in **?** This Markov point process is characterized by the local intensity function $\lambda(t, L_t)$ given by a deterministic function $\lambda(t,i)_{t\geq 0, i\geq 0}$ satisfying $\lambda(t,i) = 0$ for $i \geq n$. We need to satisfy the last condition to ensure that process L is stopped when n^{th} level as n that is the maximum number of defaults that can occur because there are n names in the portfolio. We assume that $L_0 = 0$ and L is considered as a pure birth process. Thus, the probability of a jump in the infinitesimal time interval $(t, t + dt)$ is given by $\lambda(t, L_t)dt$.
The infinitesimal generator $\mathcal{G}_t$ of the process is given by the $(n + 1) \times (n + 1)$ matrices:

$$\mathcal{G}_t = \begin{bmatrix} -\lambda(t,0) & \lambda(t,0) & 0 & 0 & 0 \\ 0 & -\lambda(t,1) & \lambda(t,1) & 0 & 0 \\ & & \cdots & & \\ 0 & 0 & 0 & -\lambda(t,n-1) & \lambda(t,n-1) \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Now lets consider $p_i(t) = \mathbb{P}L_t = i$ and $p(t) = [p_i(t)]_{i=0\ldots n}$ and it satisfies Kolmogorov equation:

$$(\partial_t - (G)_t^*)\, p = 0 \text{ on } (0, +\infty), p(0) = \delta_0, \tag{27}$$

which is in fact a system of ordinary differential equations where $\mathcal{G}_t^*$ denotes the transpose of the $\mathcal{G}$ matrix as it is a real matrix. Now Kolmogorov equation can be expressed in terms of possibly infinitely many matrix multiplications. And hence the solution to the Kolmogorov equation is given by a matrix exponential of the form

$$p(t) = exp(t\mathcal{G}^*)\delta_0, \quad t \geq 0 \tag{28}$$

We also set $\tilde{t_0} = 0$. Given a fixed maturity T, say $T = 5yr$, we define $t_i = min(\tilde{t_i}, T)$ for each $i = 0, \ldots, n$. Consequently, $t_i < T$ and $L_{t_i} = i$ if and only if there are at least i jumps of L before the maturity T on a given trajectory.

2.3.2 Modeling Markov Jumps

The sample paths of the loss process L are piecewise constant. Let $\tilde{t_i}$ denote the ith ordered jump time of L, for each $i = 1, \ldots, n$, (or $\tilde{t_i} = +\infty$, in case there are less than i jumps on a given trajectory). For the implementation of the algorithm the time steps are chosen randomly from exponential distribution according to a given formula:

$$\Delta t \sim \frac{1}{1 - \frac{L_t}{n}} \times \exp(-x) \tag{29}$$

$$\tilde{t_i} = t_{i-1} + \Delta t \tag{30}$$

$t_i = min(\tilde{t_i}, T)$ for each $i = 0, \ldots, n$ where T is maturity time. Consequently, $t_i < T$ and $L_{t_i} = i$ if and only if there are at least i jumps of L before the maturity T on a given trajectory. The Markov chain $X_{i0\leq i\leq n}$ is defined by:

$$X_i = (t_i, L_{t_i}), 0 \leq i \leq n. \tag{31}$$

We use this notation of Markov chain through out the paper for the discussion of Local Intensity Models. We define the potential function for Local Intensity Model that is used in the Selection Process of IPS so as to assign more weight to portfolios with lower values. This will help in assigning more weights to rarer events so that likelihood of default increases. The potential is a function of X_i and α. However, in this case, X_i is just used to determine if more weights should be put on the jumps that will increase the defaults. The potential function is defined as follows:

$$\omega^{\alpha}(X) = \begin{cases} \exp(\alpha), & \text{if } t < T \\ 1, & \text{otherwise} \end{cases} \tag{32}$$

The values of α parameter will play a crucial role in determining the important samples that will help in increasing the number of defaults. The α that gave the best results was chosen by the given formula:

$$\alpha(\ell) = argmax_{\alpha} \#\{j; 0 \leq j \leq m, L_T(\omega_{\alpha}^j, \alpha) = \ell\} \tag{33}$$

3 Algorithm

3.1 IPS Algorithm for Modified Merton's Model

We let $\Delta t = \frac{T}{n}$ and divide the Time interval $[0, T]$ in to equal intervals. We denote the chain $X_p = \tilde{X}_{\frac{pT}{n}}$, where $\tilde{X}$ evolves according to the continuous time dynamics,and denote the whole history of the chain as $Y_p = (X_0, X_1, ..., X_p)$. We use the potential function defined in the equation (22). And we select a smaller time step to calculate the Euler Step, and for our experiment we have chosen $\delta t = 10^{-3}$.

Also due to the form of the potential function, we do not have to track the entire history of the particle, only its current value X_p and that of its "parent" X_{p-1}, which is denoted as $\hat{W}_p$ in the following description.

3.1.1 Initialization

We take M particles, where each particle represents a complete portfolio, with identical initial values. So $\forall j \in \{, 1, .., M\}$,

$$\hat{X}_0^j = \Big(\sigma(0), (S_1(0), \cdots, S_N(0))_{1 \le i \le N}, (S_1(0), \cdots, S_N(0))\Big) \tag{34}$$

And we define the initial parent $\hat{W}_0^j = \hat{X}_0^j$.

3.1.2 Selection Stage

Suppose at time p we have a set of M particles, $(\hat{W}_p^j, \hat{X}_p^j)$, with $1 \le j \le M$. We then compute a normalization constant $\hat{\eta}_p^M$ as:

$$\hat{\eta}_p^M = \frac{1}{M} \sum_{j=1}^{M} \exp\Big[\alpha\Big(V(\hat{X}_p^{(j)})\Big) - V(\hat{W}_p^{(j)})\Big] \tag{35}$$

Then we choose M independent samples using the following distribution:

$$\eta_p^M(dW, dX) = \frac{1}{M\hat{\eta}_p^M} \sum_{j=1}^{M} \exp\Big[\alpha\Big(V(\hat{X}_p^{(j)})\Big) - V(\hat{W}_p^{(j)})\Big] \times \delta_{(\hat{W}_p^j, \hat{X}_p^j)}(dW, dX) \tag{36}$$

The particles selected are then denoted as $(\breve{W}_p^j, \breve{X}_p^j)$.

3.1.3 Mutation Stage

This stage sets the IPS apart from other importance sampling methods, as we use the exact dynamics of the model to sample points. We chose the Euler-Maruyama method to solve for the Asset Prices and the Stochastic Volatility with the time step δt mentioned above.

For each particle $(\breve{W}_p^j, \breve{X}_p^j)$, we evolve it using the Euler-Maruyama scheme from t_p to t_{p+1}, so $\breve{X}_p^j$ becomes $\hat{X}_{p+1}^j$. Note that, each particle, that is a portfolio, is evolved independently.

3.1.4 Termination Stage

At the Maturity Time, we compute the number of losses in each particle, that is a portfolio, for all M particles by computing the function f_n defined as follows:

$$f(X_n^{(j)}) = \sum_{i=1}^{N} \mathbf{1}_{\{X_n^{(j)}(N+1+i) \le B_i\}} \tag{37}$$

where the last N components of X_n are the minimums of the asset values. The estimates $\hat{p}_k^M(T)$ for the number of defaults, $\mathbf{p}_k(T) = \mathbb{P}(L(T) = k)$, is defined as:

$$\hat{p}_k^M(T) = \left[\frac{1}{M} \sum_{j=1}^{M} \mathbf{1}_{\{f(\hat{X}_n^{(j)})=k\}} \exp\left[\alpha \left(V(\hat{W}^{(j)}) - V(\hat{X}_0) \right) \right] \right] \times \left[\prod_{p=0}^{n-1} \hat{\eta}_p^M \right] \tag{38}$$

As explained in the theory section, the above estimator is an unbiased estimator.

3.2 IPS algorithm for Markovian Intensity Models

We let $\Delta t = \frac{T}{n}$ and divide the Time interval $[0, T]$ in to equal intervals. We denote the chain $X_p = \tilde{X}_{\frac{pT}{n}}$, where $\tilde{X}$ evolves according to the continuous time dynamics,and denote the whole history of the chain as $Y_p = (X_0, X_1, ..., X_p)$. We use the potential function defined in the equation (22). And we select a smaller time step to calculate the Euler Step, and for our experiment we have chosen $\delta t = 10^{-3}$.

Also due to the form of the potential function, we do not have to track the entire history of the particle, only its current value X_p and that of its "parent" X_{p-1}, which is denoted as $\hat{W}_p$ in the following description.

3.2.1 Initialization

We take M particles, where each particle represents a complete portfolio, with identical initial values. So $\forall j \in \{, 1, .., M\}$, And we define the initial parent $\hat{W}_0^j = \hat{X}_0^j$.

$$\hat{X}_0^{(j)} = \left(t_0^{(j)}, L_{t_0}^{(j)} \right), \quad 1 \leq j \leq M$$

$t_0^{(j)} = 0$ and $L_{t_0}^{(j)} = 0 \; \forall 1 \leq j \leq M$ considering L as a pure birth process.

3.2.2 Selection Stage

Sample independent M particles with repetition from input M paths according to empirical distribution under the given Gibbs measure. $\left(t_{i-1}^{(j)}, L_{t_{i-1}}^{(j)} \right)$ becomes $\left(\hat{t}_{i-1}^{(j)}, \hat{L}_{t_{i-1}}^{(j)} \right)$.

$$\begin{aligned}
\eta_i^M(dX) &= \frac{1}{M\hat{\eta}_i^M} \sum_{j=1}^{M} \left(\omega^\alpha(X_i^{(j)}) \right) \times \delta_{X_i^{(j)}}(dX) \\
\text{Where} \quad \hat{\eta}_i^M &= \mathbb{E}_i^m \omega^\alpha(X) = \frac{1}{M} \sum_{j=1}^{M} \left(\omega^\alpha(X_i^{(j)}) \right) \\
\omega^\alpha(X) &= \begin{cases} \exp(\alpha), \text{if } t < T \\ 1, \text{otherwise} \end{cases}
\end{aligned}$$

Then we choose M independent samples using the following distribution:

$$\eta_p^M(dW, dX) = \frac{1}{M\hat{\eta}_p^M} \sum_{j=1}^{M} \exp\left[\alpha \left(V(\hat{X}_p^{(j)}) \right) - V(\hat{W}_p^{(j)}) \right] \times \delta_{(\hat{W}_p^j, \hat{X}_p^j)}(dW, dX) \tag{39}$$

The particles selected, are then denoted as $(\breve{W}_p^j, \breve{X}_p^j)$.

3.2.3 Mutation Stage

Evolve paths from $\left(\hat{t}_{i-1}^{(j)}, \hat{L}_{t_{i-1}}^{(j)}\right)$ to $\left(t_i^{(j)}, L_{t_i}^{(j)}\right)$. The function λ is used to take each step until t evolves to maturity time T.

$$t_i^{(j)} = \min\left(\hat{t}_{i-1}^{(j)} + \lambda\left(\hat{t}_{i-1}^{(j)}, \hat{L}_{t_{i-1}}^{(j)}\right), T\right)$$
$$\lambda\left(\hat{X}\right) \sim \frac{1}{1 - \frac{L_t}{n}} \times \exp(-x)$$
$$L_{t_i}^{(j)} = \hat{L}_{t_{i-1}}^{(j)} + 1 \text{ if } t_n^{(j)} \neq T$$

3.2.4 Termination Stage

Terminate after running n loops of mutation and selection. Here, $n =$ maximum number of defaults that can happen. Estimate default probability $p_k(T) = \mathbb{P}\left(L\left(T\right) = k\right)$ using formula below.

$$\tilde{p}_\ell^m(T, \alpha) = \mathbb{E}_n^m \delta_\ell(\mathcal{L}(X)) exp(-\alpha \mathcal{L}(X)) \prod_{i=0}^{n-1} \mathbb{E}_i^m \omega^\alpha(X)$$

4 Results

4.1 IPS algorithm for Local Intensity Models

For this particular setup, the 3 experiments were performed. The parameters were set to the given values:

Parameter	Value
n	125
α	0.4 - 4.0
α-step	0.4
M	5000,20000,100000

In figure (a) of 1, 2 and 3, the relationship between α values and the number of defaults is shown. It can be observed that there is large difference in the number of defaults with the increase in α. Infact, with the change of α, there is significant change in the number of defaults. Hence, choice of α plays a crucial role in simulating rare defaults. For example, the results of m=5000, 1153 portfolios were able to simulate 40 defaults.

In figure (b) of 1, 2 and 3, the relationship between number of defaults and the log probabilities are shown. Also, changing the values of m drastically, doesn't change the log probabilities which explains the efficiency of the algorithm. However, it can be also observed that for different number of samples, the probabilities obtained for the higher number of defaults are significantly low. The number of defaults improve from normal monte carlo case when the value of $\alpha = 0$.

4.2 IPS Algorithm for Modified Merton's Model

The following parameters were used across experiments and observations were made on these:

Parameter	Value
General Parameters	
N	125
M	20 - 2000
α	0.05 - 0.5 (100 values in between)
T	$1, 2, 3, 4, 5$
Stochastic Volatility Parameters (Equation 14)	
$\sigma(0)$	0.4
κ	3.5
$\bar{\sigma}$	0.4
γ	0.7
ρ_σ	-0.06
Asset Parameters (Equation 12)	
$S_i(0)$	90
r	0.06
ρ	0.1
B_i	36
N_{sel}	20
δt	10^{-3}

For the sake of simplistic simulations we take the starting point, barrier level, correlation structure of all assets to be the same.

Figure 4 shows the probability distribution of the number of loss L(T) for $T = 1, 2, 3, 4$. We see that we have managed to obtain the probability for large number of defaults with only a few number of samples ($M = 200/2000$). Also as we would expect the probability of larger defaults increases as the simulated maturity period increases.

In Figure 5, we see the effect of number of samples taken. The point to observe here is that given enough values of α, increasing M from 200 to 2000 does not drastically affect the generated probabilities. This speaks to the efficiency of the algorithm in selecting appropriate paths to simulate rare events.

Another point of note is that we run 100 IPS runs with values of α equally spaced between 0.05 to 0.5 and choose the best α for a given range of defaults as suggested in Carmona and CRÉPEY [2009]. The reason for doing so is because for a given value of α, the IPS algorithm is able to estimate $\mathbb{P}(L(T) = k)$ without bias only for some values of k. This is a well studied fact as the IPS algorithm tends to concentrate on a small set of paths. This effect is clearly seen in Figure 6.

4.2.1 Single Asset Constant Volatility Case

We proceed with the analysis of a single asset constant volatility as it can be compared against known solution of a default over a maturity period T. For this we used a constant volatility of 0.25 and $\alpha = 18.5$.

We then use the analytically derived default probabilities from Section 2.2.3 to compare against our estimated probabilities to show how good these estimates are.

As shown in Figure 7, we can see that the IPS algorithm accurately estimates the default probabilities for a good range of barrier$/S_0$ value. It even manages to trace default probabilities of the order of 10^{-14} which is quite remarkable. This indicates that the IPS algorithm is quite effective in sampling these rare events and is doing so quite accurately. In the same figure we can see that the MC algorithm is unable to estimate low probability defaults just as expected.

5 Conclusions

In the two models of credit defaults that we applied the IPS algorithm to, we see the effectiveness of the IPS algorithm in sampling rare events. As seen clearly in our results the selection steps select paths which lead to more samples of the rare event. Using these resampled paths we are able to compute an estimate of the probability of rare events as described in Section **??**.

Also as indicated in Carmona et al. [2009] as well as Section 2.2.2, we evolve similar to the underlying process. This makes the implementation and understanding of the model easier.

Thus we believe that IPS technique is quite useful in sampling rare events in very high dimensional problems.

References

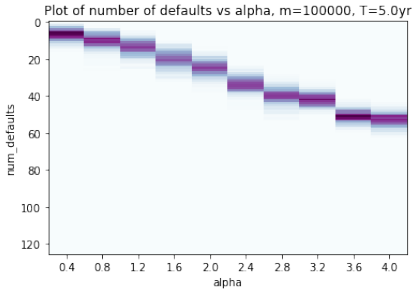
Our code, 2017. URL https://github.com/alemc2/rare-credit-portfolio.

R. Carmona and S. CRÉPEY. Importance sampling and interacting particle systems for the estimation of markovian credit portfolios loss distribution. *Forthcoming in International Journal of Theoretical and Applied Finance*, 2009.

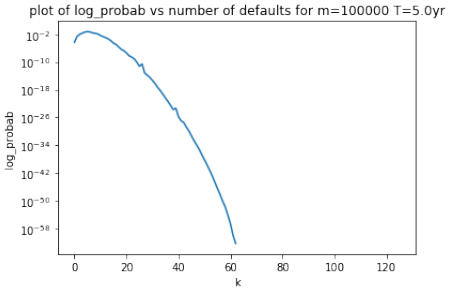
R. Carmona, J.-P. Fouque, and D. Vestal. Interacting particle systems for the computation of rare credit portfolio losses. *Finance and Stochastics*, 13(4):613–633, 2009. ISSN 1432-1122. doi: 10.1007/s00780-009-0098-8. URL http://dx.doi.org/10.1007/s00780-009-0098-8.

P. Del Moral and J. Garnier. Genealogical particle analysis of rare events. *Ann. Appl. Probab.*, 15 (4):2496–2534, 11 2005. doi: 10.1214/105051605000000566. URL http://dx.doi.org/10.1214/105051605000000566.

List of Figures

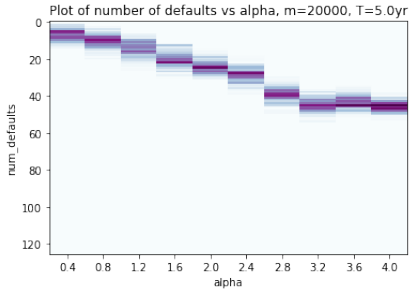


(a) number of defaults vs α

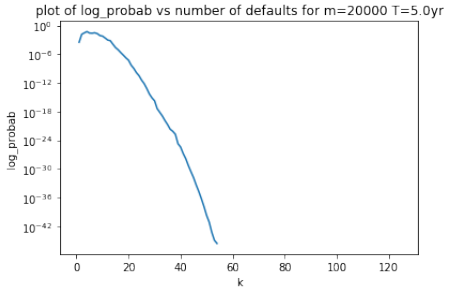


(b) Log probabilities vs number of defaults

Figure 1: Plot of runs of Local Intensity Model with $m = 10^5, T = 5.0yr$

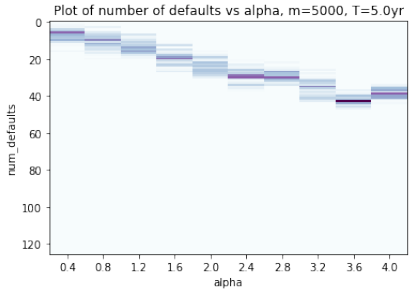


(a) number of defaults vs α

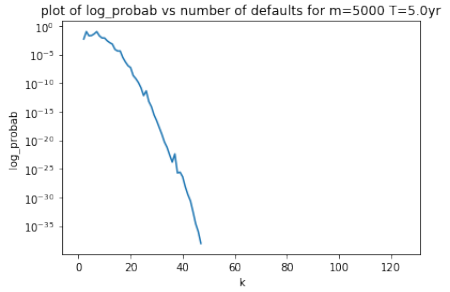


(b) Log probabilities vs number of defaults

Figure 2: Plot of runs of Local Intensity Model with $m = 20000, T = 5.0yr$



(a) number of defaults vs α



(b) Log probabilities vs number of defaults

Figure 3: Plot of runs of Local Intensity Model with $m = 5000, T = 5.0yr$

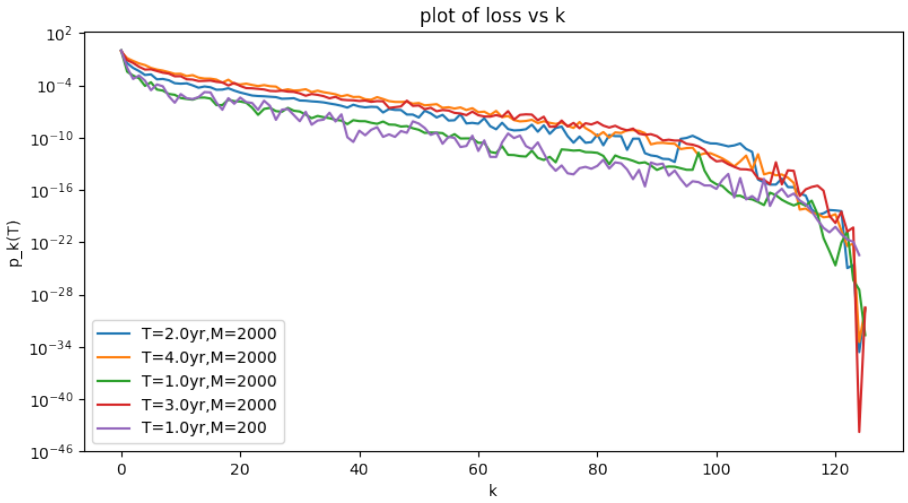


Figure 4: Plot of default probabilities for each level and maturity times based on IPS

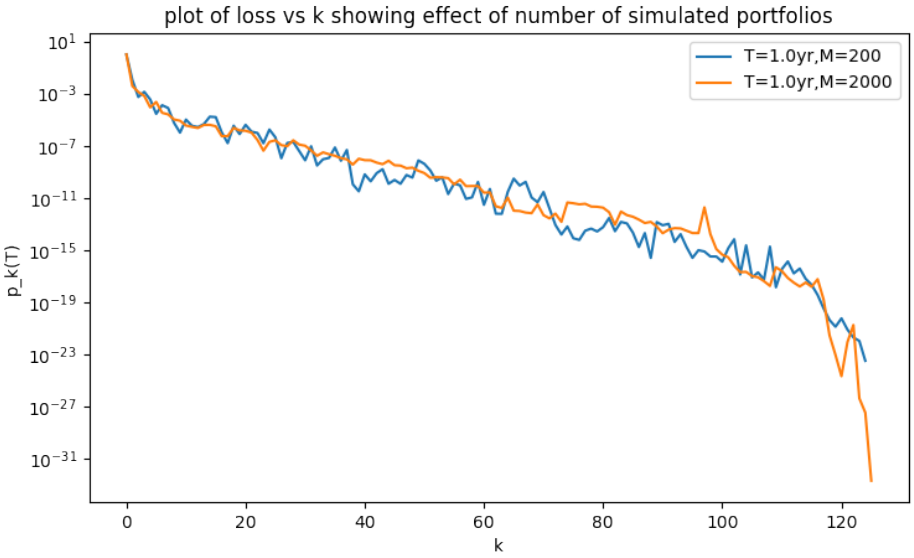


Figure 5: Plot comparing default probabilities for maturity $T = 1$ with different number of portfolios

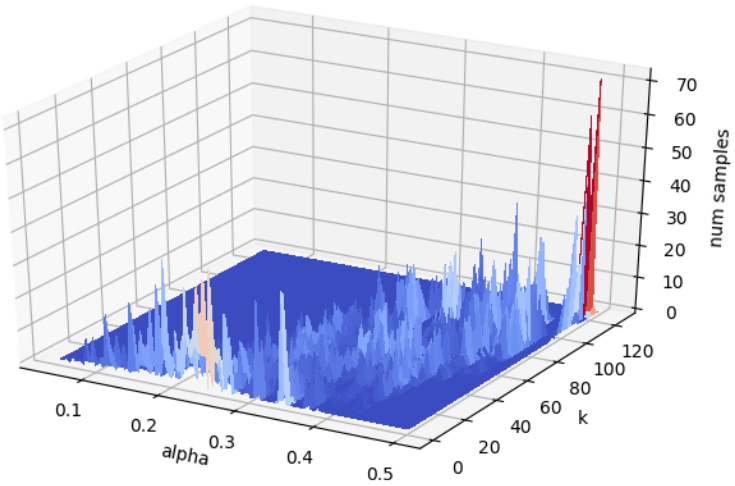


Figure 6: Plot of number of defaults simulated for α, k combination

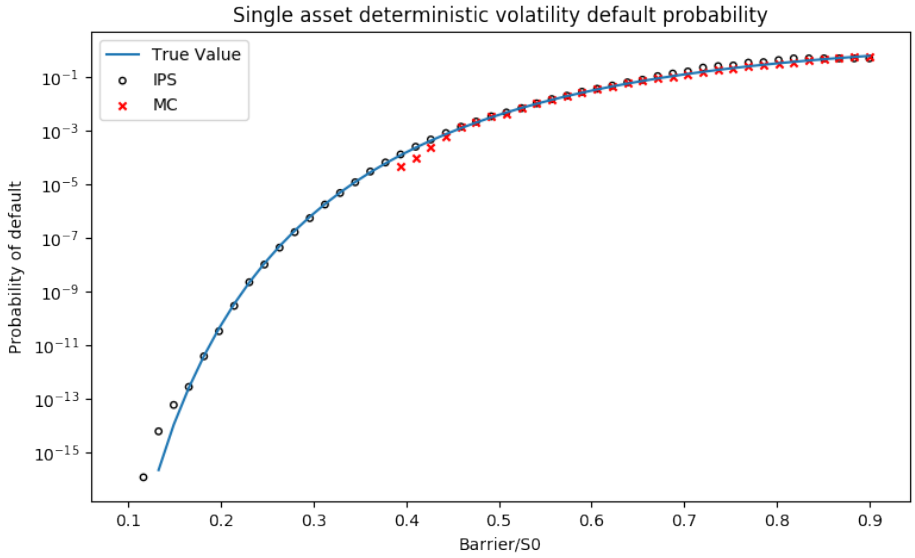


Figure 7: Default probabilities of different barrier levels for single asset deterministic volatility case