Exact and Efficient Simulation of Correlated Defaults

Kay Giesecke
Management Science & Engineering
Stanford University
giesecke@stanford.edu
www.stanford.edu/~giesecke

Joint work with H. Takada, H. Kakavand, and M. Mousavi
Corporate defaults cluster

Joint work with F. Longstaff, S. Schaefer and I. Strebulaev
Correlated default risk

Important applications

- Risk management of credit portfolios
  - Prediction of correlated defaults and losses
  - Portfolio risk measures: VaR etc.

- Optimization of credit portfolios

- Risk analysis, valuation, and hedging of portfolio credit derivatives
  - Collateralized debt obligations (CDOs)
Default timing

• Consider a portfolio of $n$ defaultable assets
  – Default stopping times $\tau^i$ relative to $(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{F}$
  – Default indicators $N^i_t = I(\tau^i \leq t)$
  – Vector of default indicators $N = (N^1, \ldots, N^n)$

• The **portfolio default process** $1_n \cdot N$ counts defaults
  – At the center of many applications
Bottom-up model of default timing

- Name \( i \) defaults at intensity \( \lambda^i \)
  - A martingale is given by \( N^i - \int_0^\cdot (1 - N_s^i) \lambda_s^i \, ds \)
  - \( \lambda^i \) represents the conditional default rate: for small \( \Delta > 0 \)
    \[
    \lambda^i_i \Delta \approx \mathbb{P}(i \text{ defaults during } (t, t + \Delta] \mid \mathcal{F}_t)
    \]
- The vector process \( \lambda = (\lambda^1, \ldots, \lambda^n) \) is the modeling primitive
  - Component processes are correlated: diffusion, common or correlated or feedback jumps
  - Large literature
Model computation

- We require \( \mathbb{E}(f(N_T)) \) for \( T > 0 \) and real-valued \( f \) on \( \{0, 1\}^n \)
  - \( \mathbb{P}(1_n \cdot N_T = k) \)
  - \( \mathbb{P}(\tau_i > t) \) for constituents \( i \)

- Semi-analytical transform techniques
  - Limited to (one-) factor doubly-stochastic intensity models

- Monte Carlo simulation
  - Larger class of intensity models
  - Treatment of more complex instruments such as cash CDOs
Simulation by time-scaling

• Widely used
  – \( \tau^i \) has the same distribution as \( \inf \{ t : \int_0^t \lambda^i_s ds = \text{Exp}(1) \} \)
  – In practice: approximate \( \lambda^i \) on discrete-time grid, integrate, and record the hitting time of the integrated process

• Potential problems
  – Discretization may introduce bias
    ∗ Magnitude?
    ∗ Computational effort
    ∗ Allocation of resources
  – Can be computationally burdensome (often \( n \geq 100 \))
Time-scaling vs. exact methods

Distribution of $1_{100} \cdot N_2$
Exact and efficient simulation

- Our approach has two parts
  1. Construct a time-inhomogeneous, continuous-time Markov chain $M \in \{0, 1\}^n$ with the property that $M_t = N_t$ in law
  2. Estimate $\mathbb{E}(f(N_T)) = \mathbb{E}(f(M_T))$ by simulating $M$
    - **Exact**: avoids intensity discretization
    - **Efficient**: adaptive variance reduction scheme

- Powerful simulation engine applicable to many intensity models in the literature
Multivariate Markovian projection

Proposition

- Let $M$ be a Markov chain that takes values in $\{0, 1\}^n$, starts at $0_n$, has no joint transitions in any of its components and whose $i$th component has transition rate $h^i(\cdot, M)$ where
  \[
  h^i(t, B) = \mathbb{E}(\lambda^i_t I(\tau^i > t) \mid N_t = B), \quad B \in \{0, 1\}^n
  \]

Then $M_t = N_t$ in distribution:
  \[
  \mathbb{P}(M_t = B) = \mathbb{P}(N_t = B), \quad B \in \{0, 1\}^n
  \]

Markov counting process

- \( M \) is a Markov point process in its own filtration \( \mathbb{G} \)

- The Markov counting process \( 1_n \cdot M \) has \( \mathbb{G} \)-intensity

\[
1_n \cdot h(t, M_t) = \sum_{k=0}^{n-1} H(t, k) \mathbb{I}(T_k \leq t < T_{k+1})
\]

where \( h = (h^1, \ldots, h^n) \), and \( (T_k) \) is the strictly increasing sequence of event times of \( 1_n \cdot M \), and

\[
H(t, k) = 1_n \cdot h(t, M_{T_k}), \quad t \geq T_k
\]

- Compare: original portfolio default process \( 1_n \cdot N \) has \( \mathbb{F} \)-intensity

\[
\sum_{i=1}^{n} \lambda^i_t I(\tau^i > t)
\]
Markov counting process

- The $G$ inter-arrival intensities $H(t, k)$ of the Markov counting process $1_n \cdot M$ are deterministic

- Exact simulation of arrival times of $1_n \cdot M$
  - Time-scaling method based on $H(t, k)$
  - Equivalently, inverse method based on
    
    $$
    \mathbb{P}(T_{k+1} - T_k > s \mid \mathcal{G}_{T_k}) = \exp\left(\int_{T_k}^{T_k + s} H(t, k) dt\right)
    $$
    
    - Sequential acceptance/rejection based on $H(t, k)$

- Exact simulation of the component $I_k \in \{1, 2, \ldots, n\}$ of $M$ in which the $k$th transition took place:

  $$
  \mathbb{P}(I_k = i \mid \mathcal{G}_{T_k}) = \frac{h^i(T_k, M_{T_{k-1}})}{H(T_k, k - 1)}
  $$
Variance reduction

• Interested in $\mathbb{P}(1_n \cdot N_T = k)$ for large $k$
  – Need to force mimicking chain $M$ into rare-event regime

• Selection/mutation scheme
  – Evolve $R$ copies $(V^r_p)$ of $M$ over grid $p = 0, 1, \ldots, m$ under $\mathbb{P}$
  – At each $p$, select $R$ particles by sampling with replacement

$$
\mathbb{P}($particle \ r \ selected) = \frac{1}{R \eta_p} \exp [\delta 1_n \cdot (V^r_p - V^r_{p-1})]
$$

where $\eta_p = \frac{1}{R} \sum_{r=1}^{R} \exp[\delta 1_n \cdot (V^r_p - V^r_{p-1})]$ and $\delta > 0$

– Final estimator of $\mathbb{P}(1_n \cdot N_T = k)$ corrects for selections

$$
\frac{\eta_0 \cdots \eta_{m-1}}{R} \sum_{r=1}^{R} I(1_n \cdot V^r_m = k) \exp (-\delta 1_n \cdot V^r_m)
$$
Selection/mutation scheme

• The selection mechanism adaptively forces the mimicking Markov chain $M$ into the rare-event regime
  – Del Moral & Garnier (2005, AAP)
  – Carmona & Crepey (2009, IJTAF)
  – Carmona, Fouque & Vestal (2009, FS)
  – Twisting of Feynman-Kac path measures
  – Well-suited to deal with different model specifications

• Mutations are generated under the reference measure $\mathbb{P}$ via the exact A/R scheme

• Estimators are unbiased

• Choice of $R$, $m$ and $\delta$
Calculating the projection

- Need \( h^i(t, B) = \mathbb{E}(\lambda^i_t I(\tau^i > t) \mid N_t = B) \) for given \((\lambda^1, \ldots, \lambda^n)\)
- We show how to calculate \( h^i(t, B) \) for a range of
  - Multi-factor doubly-stochastic models \( \lambda^i_t = X^i_t + \alpha^i \cdot Y_t \)
  - Multi-factor frailty models \( \lambda^i_t = X^i_t + \mathbb{E}(\alpha^i \cdot Y_t \mid F_t) \)
  - Self-exciting models \( \lambda^i_t = X^i_t + c^i(t, N_t) \)

in terms of the transform

\[
\phi(t, u, z, Z) = \mathbb{E} \left[ \exp \left( -u \int_0^t Z_s ds - zZ_t \right) \right], \quad Z \in \{X^i, Y\}
\]

- This extends the reach of our exact method to most models in the literature, and beyond
Numerical results

Self-exciting intensity model for $n = 100$

- Suppose the intensities $\lambda^i = X^i + \sum_{j \neq i}^{n} \beta^{ij} N^j$
  - Feedback specification can be varied
  - Analytical solutions not known

- Suppose the idiosyncratic factor follows the Feller diffusion

$$dX^i_t = \kappa_i(\theta_i - X^i_t)dt + \sigma_i \sqrt{X^i_t}dW^i_t$$

where $(W^1, \ldots, W^n)$ is a standard Brownian motion

- Parameters selected randomly (relatively high credit quality)
Numerical results

Projection for self-exciting intensity model

- The projected intensity is given by

\[ h^i(t, B) = \mathbb{E}(\lambda_t^i I(\tau^i > t) \mid N_t = B) \]
\[ = (1 - B^i) \left\{ - \frac{\partial_z \phi(t, 1, z, X^i)|_{z=0}}{\phi(t, 1, 0, X^i)} + \sum_{j \neq i}^n \beta^{ij} B^j \right\} \]

- The transform \( \phi(t, u, z, X^i) \) is in closed form, and so is \( h^i(t, B) \)
  - Can add compound Poisson jumps without reducing tractability
  - General affine jump diffusion dynamics
Numerical results

Simulation results for $\mathbb{E}((C_1 - 3)^+) \text{ where } C_1 = N_1 \cdot 1_n$

<table>
<thead>
<tr>
<th>Method</th>
<th>Trials</th>
<th>Steps</th>
<th>Bias</th>
<th>SE</th>
<th>RMSE</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exact</td>
<td>5,000</td>
<td>N/A</td>
<td>0</td>
<td>0.0239</td>
<td>0.0239</td>
<td>0.10 min</td>
</tr>
<tr>
<td></td>
<td>7,500</td>
<td>N/A</td>
<td>0</td>
<td>0.0193</td>
<td>0.0193</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>10,000</td>
<td>N/A</td>
<td>0</td>
<td>0.0165</td>
<td>0.0165</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>50,000</td>
<td>N/A</td>
<td>0</td>
<td>0.0073</td>
<td>0.0073</td>
<td>1.69</td>
</tr>
<tr>
<td></td>
<td>100,000</td>
<td>N/A</td>
<td>0</td>
<td>0.0052</td>
<td>0.0052</td>
<td>5.51</td>
</tr>
<tr>
<td></td>
<td>1,000,000</td>
<td>N/A</td>
<td>0</td>
<td>0.0016</td>
<td>0.0016</td>
<td>463.78</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time Scaling</th>
<th>Trials</th>
<th>Steps</th>
<th>Bias</th>
<th>SE</th>
<th>RMSE</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>5,000</td>
<td>71</td>
<td>0.0735</td>
<td>0.0246</td>
<td>0.0775</td>
<td>1842.15</td>
<td></td>
</tr>
<tr>
<td>7,500</td>
<td>87</td>
<td>0.0697</td>
<td>0.0199</td>
<td>0.0725</td>
<td>2628.33</td>
<td></td>
</tr>
<tr>
<td>10,000</td>
<td>100</td>
<td>0.0174</td>
<td>0.0171</td>
<td>0.0244</td>
<td>3255.12</td>
<td></td>
</tr>
</tbody>
</table>
Numerical results

Convergence of RMS errors

![Graph showing convergence of RMS errors over total simulation time. The graph compares exact and time-scaling results, with the exact results showing a steady decrease in RMSE as the total simulation time increases.]
Numerical results

Variance reduction for $\mathbb{P}(C_1 = k)$, $R = 10,000$ particles, $m = 4$

<table>
<thead>
<tr>
<th>$k$</th>
<th>$\delta$</th>
<th>Particles</th>
<th>$P(C_1 = k)$</th>
<th>Trials</th>
<th>$P(C_1 = k)$</th>
<th>VarRatio</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>0.8</td>
<td>10,000</td>
<td>0.00162340</td>
<td>17,742</td>
<td>0.00220</td>
<td>10.14</td>
</tr>
<tr>
<td>13</td>
<td>0.85</td>
<td>10,000</td>
<td>0.00068818</td>
<td>18,065</td>
<td>0.00066</td>
<td>12.33</td>
</tr>
<tr>
<td>14</td>
<td>0.85</td>
<td>10,000</td>
<td>0.00029433</td>
<td>18,387</td>
<td>0.00027</td>
<td>63.88</td>
</tr>
<tr>
<td>15</td>
<td>1.05</td>
<td>10,000</td>
<td>0.00016310</td>
<td>19,032</td>
<td>0.00011</td>
<td>121.80</td>
</tr>
<tr>
<td>16</td>
<td>1.05</td>
<td>10,000</td>
<td>0.00006790</td>
<td>19,032</td>
<td>0.00005</td>
<td>236.92</td>
</tr>
<tr>
<td>17</td>
<td>1.15</td>
<td>10,000</td>
<td>0.00002597</td>
<td>19,355</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>1.15</td>
<td>10,000</td>
<td>0.00000970</td>
<td>19,355</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>1.15</td>
<td>10,000</td>
<td>0.00000500</td>
<td>19,355</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>1.15</td>
<td>10,000</td>
<td>0.00000203</td>
<td>19,355</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>1.15</td>
<td>10,000</td>
<td>0.00000106</td>
<td>19,355</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>1.3</td>
<td>10,000</td>
<td>0.00000039</td>
<td>19,677</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
Numerical results

Probabilities $\mathbb{P}(C_1 = k)$, $R = 10,000$ particles, $m = 4$ selections
Numerical results

Variance reduction for $\mathbb{P}(C_1 = k)$, $R = 1,000$ particles, $m = 4$

<table>
<thead>
<tr>
<th>$k$</th>
<th>$\delta$</th>
<th>Particles</th>
<th>$P(C_1 = k)$</th>
<th>Trials</th>
<th>$P(C_1 = k)$</th>
<th>VarRatio</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>0.8</td>
<td>1,000</td>
<td>0.00156206</td>
<td>1,600</td>
<td>0.00125</td>
<td>14.16</td>
</tr>
<tr>
<td>13</td>
<td>0.85</td>
<td>1,000</td>
<td>0.00073476</td>
<td>1,600</td>
<td>0.00125</td>
<td>32.62</td>
</tr>
<tr>
<td>14</td>
<td>0.85</td>
<td>1,000</td>
<td>0.00024347</td>
<td>1,600</td>
<td>0.00188</td>
<td>565.20</td>
</tr>
<tr>
<td>15</td>
<td>1.05</td>
<td>1,000</td>
<td>0.00009063</td>
<td>1,726</td>
<td>0.00058</td>
<td>1562.75</td>
</tr>
<tr>
<td>16</td>
<td>1.05</td>
<td>1,000</td>
<td>0.00009381</td>
<td>1,759</td>
<td>0.00057</td>
<td>2951.99</td>
</tr>
<tr>
<td>17</td>
<td>1.15</td>
<td>1,000</td>
<td>0.00006339</td>
<td>1,790</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>1.15</td>
<td>1,000</td>
<td>0.00002132</td>
<td>1,823</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>1.15</td>
<td>1,000</td>
<td>0.00000972</td>
<td>1,887</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>1.15</td>
<td>1,000</td>
<td>0.00000040</td>
<td>1,887</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>1.15</td>
<td>1,000</td>
<td>0.00000078</td>
<td>1,918</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>1.3</td>
<td>1,000</td>
<td>0.00000028</td>
<td>1,983</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
Numerical results

Probabilities $\mathbb{P}(C_1 = k)$, $R = 1,000$ particles, $m = 4$ selections
Numerical results

Variance ratios for $\mathbb{P}(C_1 = k)$, varying $R$, $m = 4$ selections

<table>
<thead>
<tr>
<th>Number of Defaults</th>
<th>Variance Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection/Mutation, R=10,000 Particles</td>
<td></td>
</tr>
<tr>
<td>Selection/Mutation, R=1,000 Particles</td>
<td></td>
</tr>
</tbody>
</table>

Kay Giesecke
Numerical results

Variance ratios for $\mathbb{P}(C_1 = k)$, $R = 1,000$ particles, varying $m$
Numerical results

Probabilities $\mathbb{P}(C_1 = k), \ R = 1,000$ particles, varying $m$
Conclusions

• Exact and efficient simulation engine for portfolio credit risk
  – Based on multivariate Markovian projection
  – Variance reduction via selection/mutation scheme

• Broadly applicable
  – Multi-factor doubly-stochastic models
  – Multi-factor frailty models
  – Self-exciting models

• Full portfolio and single-name functionality
Conclusions

• Our results address a gap in the literature on intensity-based models of portfolio credit risk
  – Bassamboo & Jain (2006, WSC)

• Our results complement the simulation methods developed for copula-based models of portfolio credit risk
  – Bassamboo, Juneja & Zeevi (2008, OR)
  – Chen & Glasserman (2008, OR)
  – Glasserman & Li (2005, MS)

• Our results are relevant in several other application areas, including reliability, insurance, queuing
References


