Title: Pricing of a CDO on stochastically correlated underlyings.

Authors: M. Escobar, B. Gotz, L. Seco, R. Zagst.

2007: # 1
Pricing of a CDO on stochastically correlated underlyings.

Authors: Marcos Escobar $^1$, Barbara Götz $^2$, Luis Seco $^3$, Rudi Zagst $^4$

Abstract

In this report, we propose a method to price Collateralized debt obligations (CDO) within Merton’s structural model on underlyings with a stochastic mean-reverting covariance dependence. There are two key elements in our development, first we reduce dimensionality and complexity using principal component analysis on the assets’ covariance matrix. Second, we approximate this continuous multidimensional structure using a tree method. Trinomial-tree models can be developed for both the principal components and the eigenvalues assuming the eigenvectors constant over time and the eigenvalues stochastic. Our method allows us to compute the joint default probabilities for $k$ defaults of stochastically correlated underlyings and the value of CDOs, without having lost much accuracy. Furthermore we provide a method to estimate the parameters to fit the model.

Key words: stochastic covariance matrix, CDO, trinomial-trees, principal component analysis.

JEL classification: G13, C63
Contents

1 Valuation of a $k^*$ to Default Swap on constant correlated underlyings 1
   1.1 Introduction .................................................. 1
   1.2 Merton Model ................................................. 2
   1.3 Outline of the Method ....................................... 3
   1.4 Tree Building for the Principal Components ............... 4
   1.5 Results ..................................................... 7

2 Pricing of a CDO on stochastically correlated underlyings 13
   2.1 Introduction ................................................ 13
   2.2 Outline of the Method ...................................... 15
      2.2.1 Reduction of the complexity: Principal Component Analysis .... 15
      2.2.2 Tree Building ......................................... 17
   2.3 Underlying Processes ...................................... 17
   2.4 Trinomial-Tree Approximation for the Transform of the Brownian motion .... 18
   2.5 Trinomial-Tree for the Eigenvalue Matrix .................. 19
   2.6 Combined Tree for Principal Components and Eigenvalues ........... 21
      2.6.1 Computation of the Means ............................ 22
   2.7 Parameter Estimation ....................................... 24
      2.7.1 Stochastic volatility models as hidden Markov models ........ 24
      2.7.2 Application for the parameter estimation of the eigenvalue processes 26
   2.8 Example .................................................... 28
   2.9 Summary and Conclusion ................................... 30
Chapter 1

Valuation of a $k^*$ to Default Swap on constant correlated underlyings

1.1 Introduction

In the following we want to price a $k^*$ to Default Swap on $m$ constantly correlated underlyings. The grantor pays the secured party a fixed sum in the case of $k^*$ defaults in the portfolio. The model we propose here is based on the assumptions of Merton [43]. We consider $m$ different companies and define $V_i(t)$ to be the value of the assets of a company $i$, $i \in \{1, \ldots, m\}$ at time $t$. The system of processes is defined on a filtered probability space $(\Omega, \mathcal{F}, \tilde{Q}, \mathcal{F})$ where $\mathcal{F}_0$ contains all subsets of the $(\tilde{Q}-)$ null sets of $\mathcal{F}$ and $\mathcal{F}$ is right-continuous. As we assume the market to be arbitrage-free the processes are defined under the risk neutral measure $\tilde{Q}$.

In order to value credit derivatives on $m$ values with a tree model we have to reduce complexity and dimensionality. Thus, we will apply Principal Component Analysis on the Brownian motions. We will build a tree for the valuation of the $k$ highest principal components at maturity $T$ and express the results in terms of the original data to find the probabilities of the joint defaults of the companies.

In Section 1.2 we introduce the Merton model, in Section 1.3 we outline the method. To value the principal components we build a tree in Section 1.4. The results are described in Section 1.5.
1.2 Merton Model

The Merton Model is counted among the firm value credit models. In contrast to intensity and spread-based models for default risk those models take a more fundamental approach to valuing defaultable debt by linking the prices of stock and all debt issued by a particular company.

Firm value models assume that there is a fundamental process $V_i$ that describes the total value of assets of a company $i$, $i \in \{1, \ldots, m\}$. This process for $V_i$ moves stochastically. In this context default can be modeled in two ways. The approaches of Black and Scholes [8] and Merton [43] use $V_i(T)$ to pay off debt at maturity $T$ of the contract, i.e. if $V_i$ is insufficient to repay the debt, the company defaults. This means that a default event cannot occur during the life time of the contract. In a more advanced approach of Black and Cox [7] default is triggered if $V_i$ falls below a barrier $K_i$ during the runtime of the contract.

In both models the dynamics $V_i$ are assumed to follow a geometric Brownian motion

$$dV_i = rV_i dt + \sigma_i V_i dW_i, \ i \in \{1, \ldots, m\},$$

(1.1)

where $r$ is the constant risk-free rate of return and $\sigma_i$ the volatility parameter, and $W_i$ is a Brownian motion. At maturity both, equity $S_i$ and debt $L_i$ with face value $K_i$, can be viewed as derivatives on the value $V_i$ of the firm’s assets. The payoffs at maturity time $T$ can be described as the following:

$$L_i(V_i, T) = \min(K_i, V_i(T))$$

(1.2)

$$S_i(V_i, T) = \max(V_i(T) - K_i, 0)$$

(1.3)

The payoff of the debt $L_i$ is the minimum of the face value $K_i$ and the firm value $V_i(T)$, i.e. when the company value $V_i(T)$ falls below the face value $K_i$, the default of the company $i$ is triggered. In this case the value of the bond $L_i(V_i, T)$ is equal to $V_i(T)$ (see Merton [43]). The payoff of the shares is identical to the payoff of a European call option on the firm value with strike price $K_i$. It must hold that $S_i + L_i = V_i$ in each time period $t$, $t \leq T$, to exclude arbitrage opportunities. The valuation of the debt is based on the idea that a hedge portfolio exists which allows us to hedge the bond with the share and vice versa because there is only one underlying source of uncertainty. Setting up such a portfolio $\Pi_i$ and applying Itô’s lemma to it we get

$$d\Pi_i = dL_i + \Delta_i dS_i$$

$$= \left( \frac{\partial L_i}{\partial t} + \frac{1}{2} \frac{\partial^2 L_i}{\partial V_i^2} \Delta_i \frac{\partial S_i}{\partial t} + \Delta_i \frac{\partial S_i}{\partial V_i} \right) dt + \left( \frac{\partial L_i}{\partial V_i} + \Delta_i \frac{\partial S_i}{\partial V_i} \right) dV_i, \ i \in \{1, \ldots, m\}.$$
Choosing $\Delta_i = -\frac{\partial L_i}{\partial V_i} - \frac{\partial S_i}{\partial V_i}$ allows us to eliminate the stochastic $dV_i$ and yields the following pricing partial differential equation assuming that the portfolio earns the risk-free rate of interest $r$ foreclosing any arbitrage opportunity (for a proof see Hull [28], p. 504-505):

$$\frac{\partial}{\partial t} \Pi_i + \frac{1}{2} \sigma_i^2 V_i^2 \frac{\partial^2}{\partial V_i^2} \Pi_i + r V_i \frac{\partial}{\partial V_i} \Pi_i - r \Pi_i = 0 \quad (1.4)$$

Its value is given by the Black-Scholes formula for a European call option (see Schönbucher [51]).

### 1.3 Outline of the Method

In the Merton model the companies’ values are assumed to follow a log normal distribution $V_i(t) = V_i(0)e^{(r-\frac{\sigma_i^2}{2})t + \tilde{W}_i(t)}$, $i \in \{1, \ldots, m\}$ (see 1.1), where $\tilde{W}_i(t) = \sigma_i W_i(t)$ is a Brownian motion, i.e. normally distributed with expectation 0 and variance $\sigma_i^2 t$. We start in $t = 0$ and define $\tilde{W}(0)$ as a $m$-dimensional row vector, where $m$ is the number of companies in the portfolio. For the singular Brownian motions we assume $\text{Cov}[d\tilde{W}_i, d\tilde{W}_j] = \rho_{ij} \sigma_i \sigma_j t$, where $i, j \in \{1, \ldots, m\}$ and $i \neq j$. Now Principal Component Analysis is applied to $\tilde{W}(0)$. For a detailed illustration of Principal Component Analysis see Section ??.

Principal Component Analysis (see Alexander [1], [2] and Joliffe [37]) will allow us to find $m$ uncorrelated variables at each time $t$, called the principal components of $\tilde{W}(t)$, the vector of the components $\tilde{W}_i(t)$, $i \in \{1, \ldots, m\}$. Each principal component is a simple linear combination of the original returns as we can see below. Moreover, it is possible to state how much of the original variation in the data is explained by each principal component, which are ordered according to the amount of variation they explain. Let the covariance matrix $\Sigma$ of $\tilde{W}(t)$ be

$$\Sigma = ADA'$$

where $A$ is a $(m \times m)$ orthogonal matrix with columns represented by the eigenvectors of $\Sigma$ ordered by the amount of variation they explain and $D$ is a $(m \times m)$ diagonal matrix with the respective eigenvalues on its diagonal, i.e. $D = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_m)$.

The principal components of $\tilde{W}(0)$ are given by

$$B(0) = \tilde{W}(0)A$$

Hence, a linear transformation of the original risk factor returns has been made in such a way that transformed risk factors are orthogonal, i.e. they have zero instantaneous correlation. The new risk factors are ordered by the amount of the variation they explain. As $A$ is a orthogonal matrix the relationship (1.6) is equivalent to $\tilde{W} = BA'$, i.e. $\tilde{W}_i(t) = \sigma_i W_i(t)$.
\(a_1 B_1(t) + a_2 B_2(t) + \ldots + a_m B_m(t)\). Most of the time the first 2-3 eigenvectors are sufficient to describe more than 90% of the variation in the system (see Alexander [1] and [2]), which allows us to reduce the complexity and dimensionality of our system by removing the higher eigenvectors. Thus, we assume that two eigenvectors are sufficient to describe the process of \(\tilde{W}\). The effect of using additional eigenvalues will be analyzed in Section 1.5. In the following the method is explained using two eigenvectors. Thus, we set

\[
B^*(t) = \tilde{W}(t) A^*, \quad B^* = (B_1(t), B_2(t)),
\]

where \(A^*\) includes the first 2 columns of \(A\), i.e. it is a \(m \times 2\) orthogonal matrix. Therefore, the \(B_k(t), k = 1, 2\), have zero correlation and \(B_k(t) \sim N(0, D)\). In terms of the original data \(\Sigma(t), \tilde{W}_i(t)\) and \(V_i(t)\), the transformation (1.7) is equivalent to

\[
\Sigma = A^* DA^{*\prime} + \Sigma_\epsilon \quad \text{(1.8)}
\]

\[
\tilde{W}_i(t) = a_{i1} B_1(t) + a_{i2} B_2(t) + \epsilon_i \quad \text{(1.9)}
\]

\[
\ln V_i(t) = \mu_i(t) + \tilde{W}_i(t) + \ln V_i(0) \quad \text{(1.10)}
\]

where \(\mu_i(t)\) is the mean of the log returns of the company value \(V_i\), \(i \in \{1, \ldots, m\}\), in \(t\). \(\epsilon\) describes the noise caused by only including two eigenvectors and \(\Sigma_\epsilon\) its variation in \(t\). In the following, we will ignore \(\epsilon_i\) and \(\Sigma_\epsilon\).

For \(B_k, k = 1, 2\) this procedure implies a normal distribution, \(N(0, \lambda_k^2 t)\), uncorrelated principal components \(B_1\) and \(B_2\) and the following diffusion

\[
\frac{dB_k}{\Delta t} = \sqrt{\lambda_k} dZ_k \quad \text{(1.11)}
\]

\[
E[dZ_1 dZ_2] = 0 \quad \text{(1.12)}
\]

where \(\lambda_k\) is the \(k\)th eigenvalue and \(dZ_k\) a Wiener process.

The implied \(B_k(T)\) can then be retransformed to \(\tilde{W}(T) = B_k(T) A^{*\prime}\), where \(T\) denotes the values at maturity.

## 1.4 Tree Building for the Principal Components

The two-dimensional process (1.11) for \(dB_k\) is approximated using a binomial-tree for each dimension \(k, k = \{1, 2\}\), which can be built independently as their Brownian motions are uncorrelated. For the discretization the lifetime of the derivative is divided in \(n = \frac{T}{\Delta t}\) equal time steps, where \(\Delta t\) is the length of one such time steps. We denote the nodes of one such tree by \((j, t)\), where \(t\) indicates the number of time steps passed since \(t = 0\), i.e. the root of the tree and \(j\) is the number of upwards movements \(\Delta B_k\) and \(t - j\) the number
of downwards movements $-\Delta B_k$, i.e. the value $B_k(t) = B_k(0) + j \cdot \Delta B_k + (t - j) \cdot (-\Delta B_k)$. It is assumed that the underlying can jump to two different values one time step ahead. The principal component can increase to $\Delta B_k$ after one time step or decrease to $-\Delta B_k$:

$$B_k(t + 1) = B_k(t) = \begin{cases} 
\Delta B_k & \text{with probability } p_k \\
-\Delta B_k & \text{with probability } 1 - p_k 
\end{cases} \quad (1.13)$$

Such a tree is illustrated in Figure 1.1.

![Structure of a recombining tree for $B_k$](image)

Figure 1.1: Structure of a recombining tree for $B_k$

The convergence of this discrete two-point distribution to the continuous process is ensured by matching the first two moments. Note that $E[B_k(0)] = 0$.

$$E(\Delta B_k) = \Delta B_k p_k + (1 - p_k)(-\Delta B_k) = 0 \quad (1.14)$$

$$Var(\Delta B_k) = E(\Delta B_k^2) - E(\Delta B_k)^2 = p_k \Delta B_k^2 + (1 - p_k)\Delta B_k^2 = \lambda_k \Delta t \quad (1.15)$$

These equations are solved for
\[ p_k = 0.5 = 1 - p_k \]
\[ \Delta B_k = \sqrt{\lambda_k \Delta t} \]  
\[ (1.16) \]

\textit{Proof.} Set

\[ p_k \equiv 0.5 \]  
\[ (1.17) \]

Substituting (1.17) in (1.15) we get:

\[ \Delta B_k^2 = \lambda_k \Delta t \]
\[ \Delta B_k = \sqrt{\lambda_k \Delta t} \]  
\[ (1.18) \]

\( B_k(T) \) can then be retransformed to \( \tilde{W}(T) = B_k(T) A^* \), where \( T \) denotes the values at maturity. For the retransformation the values for \( B_{1j}(T), B_{2i}(T) \) with the probabilities 
\[ 0.5^n \cdot \binom{n}{j}, 0.5^n \cdot \binom{n}{i} \]  
resulting from the two independent trees with \( n + 1 \) branches in \( T \) each have to be combined to

\[ \tilde{W}(T) = a_1 \cdot B_{1j}(T) + a_2 \cdot B_{2i}(T), \ j, i \in \{0, \ldots, n\} \]  
\[ (1.19) \]

with probability

\[ p = 0.5^{2n} \cdot \binom{n}{j} \cdot \binom{n}{i}, \]  
\[ (1.20) \]

\[ (1.21) \]

where \( n \) is the number of time steps computed in the tree and \( i, j \in \{1 \ldots n\} \). \( \tilde{W}(T) \) is then a \( ((n + 1)^2 \times m) \)-matrix, where \( m \) is the number of companies in the portfolio, i.e. there are in total \( (n + 1)^2 \) different scenarios for each \( \tilde{W}_i \) in \( T \). \( \tilde{V}_i(T) \) can now be expressed in terms of \( \ln V_i(T) \):

\[ \ln V_i(T) = (r - \frac{\sigma_i^2}{2}) \cdot T + \ln V_i(0) + \tilde{W}_i(T), \]

It is assumed that the level of the default threshold is provided for each company in the portfolio. By subtracting the threshold from \( V_i(T) \) it can be determined if a company \( i \) defaults in a specific scenario. Starting from there, the value of different shaped credit derivatives can be determined. The value of the derivative is calculated as the discounted sum of the values of the derivative in each scenario weighted by the respective probability

\[ p = 0.5^{2n} \cdot \binom{n}{j} \cdot \binom{n}{i}. \]
1.5 Results

In the following we analyze the results from our valuation. The influence of the number of time steps and the proportion of the variance explained by the chosen dimensionality are examined in more detail.

All the calculations are run for the following basic scenario:

\[
\begin{align*}
\text{Maturity} &= 3 \text{ years}, \quad \Delta t = \frac{1}{n}, \quad r = 0.04, \quad m = 10 \text{ companies}, \\
K_i &= 0.75 \cdot V_i(0)
\end{align*}
\]

We used the covariance matrix in Table 1.1 for the computation of the principal components. The eigenvectors of our correlation matrix explain the following percentages of the volatilities illustrated in Table 1.2. The results of the calculations oscillate and converge. They oscillate the more, the less time steps and eigenvectors are used for the calculation (see Figures 1.5, 1.6 and 1.7). For a fixed number of eigenvalues the calculation of the value of the derivative converges the more time steps are used (see Figures 1.2, 1.3 and 1.4). To see this effect better we subtract the value of the derivative computed with 100 time steps from all others for two, three and four eigenvectors respectively to get an estimate for the error. See Figures 1.5, 1.6 and 1.7. Also note that the interval of the oscillations of the error get smaller and smaller for each additional eigenvalue used for the computations. The difference in calculation precision due to varying the dimensionality is illustrated in Figure (1.8). One can see that the difference in calculation precision is the lower the more of the volatility can be explained by the dimensions used.
Table 1.1: Covariance matrix

<table>
<thead>
<tr>
<th></th>
<th>0.002</th>
<th>2000.0</th>
<th>0.001</th>
<th>2000.0</th>
<th>900.0</th>
<th>900.0</th>
<th>0.003</th>
<th>2000.0</th>
<th>0.007</th>
<th>2000.0</th>
<th>0.005</th>
<th>2000.0</th>
<th>0.006</th>
<th>2000.0</th>
<th>0.007</th>
<th>2000.0</th>
<th>0.007</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.002</td>
<td></td>
<td></td>
<td>0.002</td>
<td></td>
<td>0.001</td>
<td></td>
<td>0.003</td>
<td></td>
<td>0.003</td>
<td></td>
<td>0.005</td>
<td></td>
<td>0.006</td>
<td></td>
<td>0.007</td>
<td></td>
<td>0.007</td>
</tr>
<tr>
<td>2000.0</td>
<td>-0.002</td>
<td>0.015</td>
<td>0.015</td>
<td>0.017</td>
<td>0.012</td>
<td>0.017</td>
<td>0.027</td>
<td>0.18</td>
<td>0.14</td>
<td>0.014</td>
<td>0.2</td>
<td>0.027</td>
<td>0.027</td>
<td>0.007</td>
<td>0.027</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>0.001</td>
<td>-0.008</td>
<td>0.009</td>
<td>0.016</td>
<td>0.019</td>
<td>0.024</td>
<td>0.029</td>
<td>0.034</td>
<td>0.18</td>
<td>0.14</td>
<td>0.014</td>
<td>0.2</td>
<td>0.027</td>
<td>0.027</td>
<td>0.007</td>
<td>0.027</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>0.003</td>
<td>-0.008</td>
<td>0.009</td>
<td>0.016</td>
<td>0.019</td>
<td>0.024</td>
<td>0.029</td>
<td>0.034</td>
<td>0.18</td>
<td>0.14</td>
<td>0.014</td>
<td>0.2</td>
<td>0.027</td>
<td>0.027</td>
<td>0.007</td>
<td>0.027</td>
<td>0.007</td>
<td></td>
</tr>
</tbody>
</table>
1 Valuation of a $k^*$ to Default Swap on constant correlated underlyings

Table 1.2: Volatility explained by the Eigenvectors

<table>
<thead>
<tr>
<th>Number of Eigenvectors</th>
<th>Percentage of the Volatility explained</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>66.045%</td>
</tr>
<tr>
<td>2</td>
<td>79.24%</td>
</tr>
<tr>
<td>3</td>
<td>85.824%</td>
</tr>
<tr>
<td>4</td>
<td>89.6124%</td>
</tr>
<tr>
<td>5</td>
<td>92.9287%</td>
</tr>
<tr>
<td>6</td>
<td>95.1642%</td>
</tr>
<tr>
<td>7</td>
<td>96.9177%</td>
</tr>
<tr>
<td>8</td>
<td>98.472%</td>
</tr>
<tr>
<td>9</td>
<td>99.4742%</td>
</tr>
<tr>
<td>10</td>
<td>99.99929%</td>
</tr>
</tbody>
</table>

Figure 1.2: Effect of Varying the Number of Time Steps with $k = 2$

Figure 1.3: Effect of Varying the Number of Time Steps with $k = 3$
Figure 1.4: Effect of Varying the Number of Time Steps with $k = 4$

Figure 1.5: Effect of Varying the Number of Time Steps on the Estimated Error for $k = 2$

Figure 1.6: Effect of Varying the Number of Time Steps on the Estimated Error for $k = 3$
Figure 1.7: Effect of Varying the Number of Time Steps on the Estimated Error for $k = 4$
Figure 1.8: Effect of Varying the Number of Time Steps and the Number of Eigenvalues
Chapter 2

Pricing of a CDO on stochastically correlated underlyings

2.1 Introduction

The British Banking Association estimates in its Credit Derivatives Report 2006 [5] that by the end of 2006 the size of the global credit derivatives market will be $20 trillion. Two of the most popular examples of credit derivatives are credit default swaps (CDS) and collateralised debt obligations (CDO). These credit derivative products strongly depend on the joint behavior of the underlying companies, i.e. the covariances. These covariances seem to change stochastically. The recent popularity of discrete stochastic correlation models like the CGARCH proposed by Engle [21] supports this idea. On the other hand, some of the most popular procedures to price credit derivatives are non-stochastic in their correlation structures, e.g. the Linear Factor Model (LFM) (see Hull and White [33]), the Factor-Copula Model (FCM) (see Laurent and Gregory [41]) or the Intensity-Factor Models (see Duffie and Garleanu [19], Hurd and Kuznetsov [35]). There is, nevertheless, one recent attempt, the Structural Model (see Hull and White [34]), where this feature is added within a factor model framework.

The dependence structure implied by these models and the simplicity of the pricing scheme are the key elements for a good pricing performance. In this chapter, we suggest a method which allows us to value CDOs on assets with a stochastic mean-reverting covariance matrix. To reduce dimensionality and complexity we apply Principal Component Analysis on the assets’ instantaneous covariance matrix. Other publications applying principal component analysis are in support of the fact that two to three eigenvalues are sufficient to describe most of the variation in the portfolio (see Alexander [1], [2]). Thus, we
use the first two eigenvalues and eigenvectors to compute the principal components of the Brownian motions of the assets. For these eigenvalues and principal components we suggest a system of processes and develop a trinomial-tree model to approximate the system. In this model the respective trees for the principal components and their volatilities, the eigenvalues, can be combined easily as we assume that their processes are independent. Thus, the joint default probabilities can be found by simply multiplying the marginal probabilities. We provide a path-independent method to approximate the means of the singular assets, which allows us to compute company values based on the values of the principal components and their respective probability. Using these asset values at maturity, the scenarios in which companies default can be calculated. As their probability is known the values of the tranches of the CDO can be computed. Furthermore, we provide a method to fit our model to market data.

In Section 2.2 we outline our method and model the processes in Section 2.3. We derive a trinomial-tree model for the principal components in Section 2.4 and for the eigenvalues in Section 2.4. The combined tree system is described in Section 2.6. Section 2.7 introduces the parameter estimation method and Section 2.8 prices a CDO for an example. We conclude in Section 2.9.
2.2 Outline of the Method

We consider \( m \) different companies and define \( V_i(t) \) to be the value of the assets of a company \( i, \ i \in \{1, \ldots, m\} \) at time \( t \). The system of processes is defined on a filtered probability space \((\Omega, \mathcal{F}, \tilde{Q}, \mathbb{F})\) where \( \mathcal{F}_0 \) contains all subsets of the \((\tilde{Q}–)\) null sets of \( \mathcal{F} \) and \( \mathbb{F} \) is right-continuous. As we assume the market to be arbitrage-free the processes are defined under the risk neutral measure \( \tilde{Q} \). The dynamics of \( V_i \) are assumed to follow a Geometric Brownian motion

\[
dV_i = rV_i dt + \sigma_i V_i dW_i, \ i \in \{1, \ldots, m\},
\]

where \( r \) is the constant risk-free rate of return and \( \sigma_i \) the volatility parameter, and \( W_i \) is a Brownian motion. Each company is assumed to be funded by equity \( S_i \) as well as one bond with face value \( K_i \) and present value \( L_i \). The approach of Merton [43] uses \( V_i \) to pay off debt at maturity of the contract. If \( V_i \) is insufficient to repay the debt, the company defaults. Thus, at maturity both, equity \( S_i \) and the bond \( L_i \) can be viewed as derivatives on the value \( V_i \) of the firm’s assets, and the payoffs at maturity time \( T \) can be described as follows:

\[
L_i(V_i, T) = \min(K_i, V_i(T)) \\
S_i(V_i, T) = \max(V_i(T) - K_i, 0)
\]

The payoff of the bond at maturity time \( T \) is the minimum of the face value \( K_i \) of the bond and the firm value \( V_i(T) \), i.e. when the company value \( V_i(T) \) falls below the face value \( K_i \), the default of the company \( i \) is triggered. In this case the value of the bond \( L_i \) is equal to \( V_i \) (see Merton 1974). The payoff of the shares is identical to the payoff of a European call option on the firm value with strike price \( K_i \). As the companies’ values are assumed to follow a log normal distribution, \( V_i(t) = V_i(0)e^{(r - \frac{\sigma_i^2}{2})t + \tilde{W}_i(t)} \), where \( \tilde{W}_i(t) = \sigma_i W_i(t) \) is a Brownian motion, i.e. normally distributed with expectation 0 and variance \( \sigma_i^2 t \). In our framework we assume that the covariance matrix follows a mean-reverting process. To handle this complexity and to reduce dimensionality implied by the \( m \) companies in the portfolio we will apply Principal Component Analysis on \( \tilde{W} = (\tilde{W}_1, \ldots, \tilde{W}_m) \).

2.2.1 Reduction of the complexity: Principal Component Analysis

Principal Component Analysis (see Alexander [1], [2] and Joliffe [37]) will allow to find \( m \) uncorrelated variables at each time \( t \), called the principal components of \( \tilde{W}(t) \). Each principal component is a simple linear combination of the original returns as we can see...
below. Moreover, it is possible to state how much of the original variation in the data is explained by each principal component, which are ordered according to the amount of variation they explain. Now let the stochastic covariance matrix \( \Sigma(t) \) of \( \tilde{W}(t) \) be

\[
\Sigma(t) = AD(t)A',
\]

(2.1)

where \( A \) is a \((m \times m)\) orthogonal (time independent) matrix with columns represented by the eigenvectors of \( \Sigma(t) \) ordered by the amount of variation they explain and \( D(t) \) is a (time dependent) \((m \times m)\) diagonal matrix with the respective eigenvalues on its diagonal, which are stochastic, i.e. \( D(t) = diag(\lambda_1(t), \lambda_2(t), \ldots, \lambda_m(t)) \). The principal components of \( \tilde{W} \) are given by

\[
B = \tilde{W}A
\]

(2.2)

Hence, a linear transformation of the original risk factor returns has been made in such a way that transformed risk factors are orthogonal, i.e. they have zero instantaneous correlation. The new risk factors are ordered by the amount of the variation they explain. As \( A \) is a orthogonal matrix the relationship (2.2) is equivalent to \( \tilde{W} = BA' \), i.e. \( \tilde{W}_i(t) = a_{i1}B_1(t)+a_{i2}B_2(t)+\ldots+a_{im}B_m(t) \). Most of the time the first 2-3 eigenvectors are sufficient to describe more than 90% of the variation in the system (see Alexander [1], [2]), which allows us to reduce the complexity and dimensionality of our system by removing the higher eigenvectors. Thus, we set

\[
B^*(t) = \tilde{W}(t)A^*, \quad B^* = (B_1(t), B_2(t))
\]

(2.3)

where \( A^* \) includes the first 2 columns of \( A \), i.e. it is a \( m \times 2 \) orthogonal matrix. Therefore, the \( B_k(t) \), the \( k \)th column of \( B(t) \), have zero correlation and \( B_k(t) \sim N(0, D(t)) \). As we assume the covariance matrix of the companies in the portfolio to be stochastic, the covariance matrix of \( B^*(t) \), \( D(t) = diag(\lambda_1(t), \lambda_2(t)) \), is also stochastic because we perform a linear transformation. In terms of the original data \( \Sigma(t) \), \( \tilde{W}_i(t) \) and \( V_i(t) \), the transformation (2.3) is equivalent to

\[
\Sigma(t) = A^*D(t)A'^* + \Sigma_\epsilon(t)
\]

(2.4)

\[
\tilde{W}_i(t) = a_{i1}B_1(t)+a_{i2}B_2(t)+\epsilon_i
\]

(2.5)

\[
\ln V_i(t) = \mu_i(t) + \tilde{W}_i(t) + \ln V_i(0),
\]

(2.6)

where \( \mu_i(t) \), the mean of the log returns of the company \( i, \ i \in \{1, \ldots, m\} \) in \( t \), is a function of the volatilities of the singular time steps. \( \epsilon_i \) describes the noise caused by only including two eigenvectors and \( \Sigma_\epsilon \) its variation. In the following we will ignore \( \epsilon_i \) and \( \Sigma_\epsilon \).
2.2.2 Tree Building

Starting in $t = 0$ with

$$B^*(0) = \tilde{W}(0)A^*, \quad (2.7)$$

we are going to build a combined tree model for the stochastic volatility and the underlying $B^*$ to get values and probabilities for $B^*$ and $D$ at maturity $T$. As seen above, these results can be expressed in terms of the original variables $\tilde{W}_i(T)$ and $V_i(T)$, $i \in \{1, \ldots, m\}$:

$$\Sigma(T) = A^* D(T) A^T \quad (2.8)$$

$$\tilde{W}_i(T) = a_{i1} B_1(T) + a_{i2} B_2(T) + \epsilon_i \quad (2.9)$$

$$\ln V_i(T) = \mu_i(T) + \tilde{W}_i(T) + \ln V_i(0), \quad (2.10)$$

These data allow us to calculate how many companies default, assuming we know the face values $K_i$ of the bonds and the respective probabilities, i.e. all the information needed to compute the value of the tranches of a CDO, without having lost much accuracy.

2.3 Underlying Processes

We suggest the following model for the underlying system of the principal components and eigenvectors $k, j$, where $k, j \in \{1, 2\}$:

$$dB_k = \sqrt{\lambda_k} dZ_k \quad (2.11)$$

$$d\lambda_k = d_k (b_k - \lambda_k) dt + c_k \sqrt{\lambda_k} dQ_k \quad (2.12)$$

$$E(dZ_k dQ_k) = 0 \quad (2.13)$$

$$E(dZ_k dZ_j) = 0 \quad \forall k \neq j \quad (2.14)$$

$$E(dQ_k dQ_j) = 0, \quad (2.15)$$

where $d_k, b_k$ and $c_k$ are fixed constants, $\lambda_k$ is the kth eigenvalue and $dZ_k$ and $dQ_k$ are independent Wiener processes\(^1\).

\(^1\)The processes for $B_k(t)$ and $d\lambda_k$ are known to exist. The original data $\tilde{W}_i(t)$, $S_i(t)$, $i \in \{1, \ldots, m\}$ and $\Sigma(t)$ can be expressed approximately as a linear transformation of $B_k(t)$, $D_k(t)$ and $A$, that is

$$\tilde{W}_i = a_{i1} B_1 + a_{i2} B_2 \quad (2.16)$$

$$\ln V_i = \ln V_0 + \mu_i + a_{i1} B_1 + a_{i2} B_2 \quad (2.17)$$

$$\sigma_i^2 = a_{i1}^2 \lambda_1 + a_{i2}^2 \lambda_2, \quad (2.18)$$

where $a_{i1}, a_{i2} \in A^*$. Therefore the processes for $W_i$, $V_i$ and $\Sigma(t)$ exist.
2.4 Trinomial-Tree Approximation for the Transform of the Brownian motion

The process (2.11) for $dB_k$, $k \in \{1, 2\}$, is approximated using a trinomial-tree because in this framework the spacing parameter can be set independent from the volatility of the process. In the following, the nodes are denoted by $(j, t)$, where $j$ is the number of upwards movements and $t$ indicates the number of time steps passed since $t = 0$, i.e. the root of the tree. Let $B_k(t) = B_k(0) + j \Delta B_k$, then

$$
\Delta B_k(t) := B_k(t + 1) - B_k(t) = \begin{cases} 
\Delta B_k & \text{with probability } p_{j,j+1,t} \\
0 & \text{with probability } p_{j,j,t} \\
-\Delta B_k & \text{with probability } p_{j,j-1,t}
\end{cases}
$$

(2.19)

The convergence of this discrete three-point distribution to the continuous process is ensured by matching the first two moments. Note that $E(B_k(0)) = 0$. Again, let $B_k(t) = B_k(0) + j \Delta B_k = j \Delta B_k$. Then,

$$
E(\Delta B_k(t)) = p_{j,j+1,t} \Delta B_k - p_{j,j-1,t} \Delta B_k = 0 \quad (2.20)
$$

$$
Var(\Delta B_k(t)) = E(\Delta B_k^2(t)) = p_{j,j+1,t} \Delta B_k^2 + p_{j,j-1,t} \Delta B_k^2 = \lambda_k (j \Delta B_k, (t-1) \Delta t) \Delta t \quad (2.21)
$$

$$
p_{j,j+1,t} + p_{j,j,t} + p_{j,j-1,t} = 1 \quad (2.22)
$$

From Equation (2.20) follows

$$
p_{j,j+1,t} = p_{j,j-1,t} \quad (2.24)
$$

Substituting Equation (2.24) in Equation (2.22) we get

$$
\Delta B_k^2(2p_{j,j+1,t}) = \lambda_k (j \Delta B_k, (t-1) \Delta t) \Delta t \quad (2.25)
$$

Solving this for $p_{j,j+1,t}$ we obtain

$$
p_{j,j+1,t} = \frac{1}{2} \frac{\lambda_k (j \Delta B_k, (t-1) \Delta t) \Delta t}{\Delta B_k^2} \quad (2.26)
$$

and

$$
p_{j,j-1,t} = \frac{1}{2} \frac{\lambda_k (j \Delta B_k, (t-1) \Delta t) \Delta t}{\Delta B_k^2} \quad (2.27)
$$

$$
p_{j,j,t} = 1 - \frac{\lambda_k (j \Delta B_k, (t-1) \Delta t) \Delta t}{\Delta B_k^2} \quad (2.28)
$$

To guarantee positive probabilities $p_{j,j+1,t}, p_{j,j,t}$ and $p_{j,j-1,t}$ we choose appropriate spacing and timing parameters. From equations (2.26), (2.27) and (2.28) it can be shown that
the probabilities are positive if
\[ \lambda_k(j \Delta B_k, (t - 1) \Delta t) < \frac{\Delta B_k^2}{\Delta t}. \]  
(2.29)
is satisfied (see Section ??). Increasing the number of time steps, Hull and White [31] and Brennan and Schwartz [11] find it desirable to keep the ratio \( \frac{\Delta B_k^2}{\Delta t} \) constant to ensure convergence. In order to develop a recombining tree with a spacing parameter not depending on the volatility, i.e. the eigenvalue, the eigenvalue has to be restricted. However, in order not to restrain the model too much we use the following relationship between the variances of the original data and the eigenvalues
\[ \sum_{i=1}^{m} \text{Var}(V_i(t)) = \text{trace}(\Sigma(t)) = \text{trace}(\Sigma(t)AA') \]
\[ = \text{trace}(A'\Sigma(t)A) = \text{trace}(D(t)) = \sum_{k=1}^{m} \lambda_k(t) \]
(see Joliffe [37]). Thus, we set
\[ \Delta B_k^2 \equiv \gamma \Delta t \sum_{k=1}^{m} \lambda_k(0) \]  
(2.30)
and inequality (2.29) transforms to
\[ \lambda_k(j \Delta B_k, (t - 1) \Delta t) < \gamma \sum_{k=1}^{m} \lambda_k(0). \]  
(2.31)
A sufficient choice of \( \gamma \) is possible as all \( \lambda_k(t) \) are bounded in the trinomial tree. The tree building and the combination of the trees are shown for the first two principal components. Table (2.1) illustrates the possible movements of \( B_1 \) and \( B_2 \) after two time steps. The probabilities of each joint combination are obtained by simply multiplying the probabilities of the marginals as the marginals are uncorrelated. These marginal probabilities are, however, influenced by the volatilities \( \sqrt{\lambda_1(t-1)} \) and \( \sqrt{\lambda_2(t-1)} \) respectively.

### 2.5 Trinomial-Tree for the Eigenvalue Matrix

The process (2.12) is implemented using a trinomial-tree suggested by Hull and White [31]. This method requires a constant volatility parameter for the process, which can be obtained with the following transformation of the process:
\[ \phi_k = \sqrt{\lambda_k}, \; k \in \{1, 2\} \]  
(2.32)
Table 2.1: Possible Movements of $B_1(t)$ and $B_2(t)$ after two time steps

<table>
<thead>
<tr>
<th>$2\Delta B_1(t)$</th>
<th>$2\Delta B_1(t)$</th>
<th>$2\Delta B_1(t)$</th>
<th>$2\Delta B_1(t)$</th>
<th>$2\Delta B_1(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2\Delta B_2(t)$</td>
<td>$\Delta B_2(t)$</td>
<td>$0$</td>
<td>$-\Delta B_2(t)$</td>
<td>$-2\Delta B_2(t)$</td>
</tr>
<tr>
<td>$\Delta B_1(t)$</td>
<td>$\Delta B_1(t)$</td>
<td>$\Delta B_1(t)$</td>
<td>$\Delta B_1(t)$</td>
<td>$\Delta B_1(t)$</td>
</tr>
<tr>
<td>$2\Delta B_2(t)$</td>
<td>$\Delta B_2(t)$</td>
<td>$0$</td>
<td>$-\Delta B_2(t)$</td>
<td>$-2\Delta B_2(t)$</td>
</tr>
<tr>
<td>$\Delta B_1(t)$</td>
<td>$-\Delta B_1(t)$</td>
<td>$-\Delta B_1(t)$</td>
<td>$-\Delta B_1(t)$</td>
<td>$-\Delta B_1(t)$</td>
</tr>
<tr>
<td>$2\Delta B_2(t)$</td>
<td>$\Delta B_2(t)$</td>
<td>$0$</td>
<td>$-\Delta B_2(t)$</td>
<td>$-2\Delta B_2(t)$</td>
</tr>
<tr>
<td>$-2\Delta B_1(t)$</td>
<td>$-2\Delta B_1(t)$</td>
<td>$-2\Delta B_1(t)$</td>
<td>$-2\Delta B_1(t)$</td>
<td>$-2\Delta B_1(t)$</td>
</tr>
<tr>
<td>$2\Delta B_2(t)$</td>
<td>$\Delta B_2(t)$</td>
<td>$0$</td>
<td>$-\Delta B_2(t)$</td>
<td>$-2\Delta B_2(t)$</td>
</tr>
</tbody>
</table>

Using Itô’s Lemma, this leads to

$$d\phi_k = d\sqrt{\lambda_k} = \left(\frac{1}{2\sqrt{\lambda_k}}(b_k d_k - d_k \lambda_k) - \frac{1}{8}\lambda_k^{-\frac{3}{2}}c_k^2 \lambda_k\right)dt + \frac{c_k}{2}dQ_k$$

$$= (\frac{b_k d_k}{2\phi_k} - \frac{d_k}{2}\phi_k - \frac{1}{8}c_k^2 \frac{1}{\phi_k})dt + \frac{c_k}{2}dQ_k$$

$$= (\frac{4b_k d_k - c_k^2}{8}\frac{1}{\phi_k} - \frac{d_k}{2}\phi_k)dt + \frac{c_k}{2}dQ_k$$

$$= q_k dt + \nu_k dQ_k,$$

where $\nu_k = \frac{q_k}{2}$, $q_k = \frac{\alpha_{1,k}}{\phi_k} - \alpha_{2,k}\phi_k$ and $\alpha_{1,k} = \frac{4b_k d_k - c_k^2}{8}$, $\alpha_{2,k} = \frac{d_k}{2}$.

Again the lifetime of the derivative is divided in $n = \frac{T}{\Delta t}$ equal time steps, where $\Delta t$ is the length of one such time step. After each time step the tree branches out. The nodes are denoted by $(l, t)$, where $l$ is the number of upwards movements of the eigenvalue, i.e. the value $\phi_k(l, t) = \phi_k(0) + l\Delta \lambda$, and $t$ indicates the number of time steps passed since $t = 0$. As before, we assume that $\lambda_k$ can increase, move aside, or decrease. However, as $q$ is unbounded, the model we used before may not converge. Thus, this method was modified by Hull and White (1990). Beside the possibility that $\Delta \phi_k$ is equal to $\Delta \phi_k(l, t), 0$ and $-\Delta \phi_k(l, t)$ two other branching methods are relevant for this implementation. The three methods are illustrated in Figures 2.1-2.3, where $\kappa$ is the move increment. To ensure
convergence the probabilities are set to (see Section ??):

\[
\begin{align*}
    p_{l,\kappa - 1} &= \nu_k^2 \frac{\Delta t}{(\Delta \phi_k)^2} + \frac{\eta_k^2}{2(\Delta \phi_k)^2} - \frac{\eta_k}{2\Delta \phi_k}, \\
    p_{l,\kappa} &= 1 - \nu_k^2 \frac{\Delta t}{(\Delta \phi_k)^2} - \frac{\eta_k^2}{(\Delta \phi_k)^2}, \\
    p_{l,\kappa + 1} &= \nu_k^2 \frac{\Delta t}{2(\Delta \phi_k)^2} + \frac{\eta_k^2}{2(\Delta \phi_k)^2} + \frac{\eta_k}{2\Delta \phi_k},
\end{align*}
\]  
(2.33 - 2.35)

where \( \eta = q_k(l,t)\Delta t + (l - \kappa)\Delta \phi_k \).

When \( \Delta \phi_k \) is set to \( \nu_k \sqrt{3\Delta t} \) the following dynamic rules for the choice of \( \kappa \) can be implemented to ensure positive probabilities:

\[
\kappa = \begin{cases} 
    l + 1 & \text{if } \frac{q_k(l,t)\Delta t}{\Delta \phi} \geq \sqrt{\frac{2}{3}} \\
    l & \text{if } -\sqrt{\frac{2}{3}} < \frac{q_k(l,t)\Delta t}{\Delta \phi} < \sqrt{\frac{2}{3}} \\
    l - 1 & \text{if } \frac{q_k(l,t)\Delta t}{\Delta \phi} \leq -\sqrt{\frac{2}{3}}
\end{cases}
\]  
(2.36)

These dynamic rules of choice for \( \kappa \) imply minimum and maximum values for \( \phi_k(t) \):

\[
-\sqrt{\frac{2}{3}} \leq \frac{q_k(l,t)\Delta t}{\Delta \phi} \leq \sqrt{\frac{2}{3}} \iff \phi_{k,\min} = -\beta + \sqrt{\beta^2 + 4\alpha_1 \alpha_2} \leq \phi_k(t) \leq \phi_{k,\max} = \phi_{k,\max} = \phi_{k,\max}
\]  
(2.37)

where \( \beta = \sqrt{\frac{2}{3}} \frac{\Delta \phi_k}{\Delta t}, \alpha_1 = \frac{4db - c^2}{8}, \alpha_2 = \frac{1}{2}d \).

The branching method is changed to \( \kappa = l - 1 \) at a node \((\psi,t)\), where \( \psi \) is the largest integer that \( \phi_k(\psi,t) = \phi_k(0) + \psi \Delta \phi \leq \phi_{k,\max} \) and to \( \kappa = l + 1 \) at a node \((\zeta,t)\), where \( \zeta \) is the smallest integer that \( \phi_k(\zeta,t) = \phi_k(0) + \zeta \Delta \phi \geq \phi_{k,\min} \).

### 2.6 Combined Tree for Principal Components and Eigenvalues

In order to obtain spacing parameters, which are independent from the volatility in the trees for the principal components, we had to restrict the volatility (see (2.31)) generated
by the trees for $\sqrt{\lambda_k}$, $k \in \{1, 2\}$. Thus, additionally to the restrictions on the volatility imposed by (2.37) it has to hold that

$$\phi_k(j \Delta B_k, (t-1)\Delta t) < \sqrt{\gamma \sum_{k=1}^{m} \lambda_k(0)} \quad (2.38)$$

Hence, whichever of the restrictions (2.37) and (2.38) is reached earlier in a time step $t$ it causes the tree for $\phi_k$ to change the branching method in this node $(s,t)$. $s$ is $\min(\psi, \vartheta)$, where $\vartheta$ is the largest integer that $\phi_k(\vartheta, t) = \phi_k(0) + \vartheta \Delta \phi \leq \sqrt{\gamma \sum_{k=1}^{m} \lambda_k(0)}$ in (2.38) and $\psi$ is the largest integer that $\phi_k(\psi, t) = \phi_k(0) + \psi \Delta \phi \leq \phi_{k,\text{max}}$ in (2.37). As we assume the Brownian motions of the principal components of the assets and of the eigenvalues to be independent, the joint probabilities can be obtained by simply multiplying the marginal probabilities. The probabilities for the movements of the principal components ((2.26) to (2.28)) also apply in the case of stochastic volatilities.

The nodes in the combined tree are denoted by $(j, w, l, v, t)$, where $j$ and $w$ indicate the number of up or down moves of the first and the second principal component respectively. $l$ and $v$ specify the level of the two eigenvalues. The trees of the eigenvalues and the principal components are arranged in series, i.e. the eigenvalues in $t$ influence the probabilities of the principal components to move up or down in $t+1$ (in the tree the principal components increase or decrease $n$ times, the eigenvalues $n-1$ times until $T$). A particular node in the combined tree branches in $3^4$ different nodes in the next time step.

### 2.6.1 Computation of the Means

In $T$, $\mu_i(T)$, $i \in \{1, \ldots, m\}$, has to be added to the retransformed Brownian motions $\tilde{W}_i(T)$. $\mu_i(T)$ is actually the sum of the means of the singular time steps and can be obtained by combining the two trees of the eigenvalues to a tree for the variances of the underlyings with nodes $(l, v, t)$ to $(3)^{n-1}$ nodes in $T$, assuming $n$ time steps:

$$\mu_i(l, v, t|l, v, t-1; \ldots; l, v, t = 0) = \sum_{t=1}^{n} (r - \frac{\sigma_i(l, v, t)^2}{2}) \Delta t, \quad \forall \text{ combinations } l, v,$$

where $\sigma_i(l, v, t)^2$ is computed by

$$\lambda_1(l, t)a_{1i}^2 + \lambda_2(v, t)a_{2i}^2 \quad \forall \text{ combinations } l, v. \quad (2.39)$$

However, the determination of the exact means would result in a path-dependent pricing as $\mu_i(T)$ depends on the path in the combined volatility tree. This results in the computation of $3^{2(n-1)}$ different paths, where $n$ is the number of time steps computed in the tree. The
distribution of the values of the means in $T$ is illustrated for an example in section 2.8. We will compute the mean at each node $(l, v, t)$ in the variance tree by weighting the means of the preceding nodes with the conditional probabilities to reach the respective node $(l, v, t)$. In this way each node of the combined variance tree in time step $n$ gives rise to one mean value $\mu_i(T)$, i.e. there are $((n-1) \cdot 2 + 1)^2$ different mean values in $T$. This is illustrated in Figure 2.4 for $n = 2$. In Section 2.8 we compare the distribution of the values resulting from this method with the path-dependent computation method. This comparison shows that the distributions are similar.

These mean values allow us to express the principal components and eigenvalues in $T$ in terms of $\ln V_i$ and $V_i$, $i \in \{1, \ldots, m\}$:

$$\Sigma(T) \approx A^* D(T) A'$$

$$\ln V_i(T) \approx \ln V_i(0) + \mu_i(T) + \tilde{W}_i(T)$$

$$V_i(T) \approx V_i(0) \exp (\mu_i(T) + \tilde{W}_i(T))$$

By subtracting the face values $D_i$ from $V_i$ it can be determined if and how many companies default in each scenario. The pay-off function of each tranche of the CDO can then be

![Figure 2.4: Tree for the Mean Values for $n = 2$ and two eigenvalues](image_url)
weighted with the joint probabilities of the respective scenario \((i, j, l, v, T)\) and discounted, which results in the value of the tranche in \(t = 0\).

### 2.7 Parameter Estimation

The estimation procedure we propose to estimate the parameters of our model is inspired by the work of Genon-Catalot et al. [25] for stochastic volatility models in the case of a fixed sampling interval \(\Delta t\).

#### 2.7.1 Stochastic volatility models as hidden Markov models

Genon-Catalot et al. [25] show that a stochastic volatility model can be viewed as a hidden Markov model.

**Definition 1. (Hidden Markov model)**(Genon-Catalot et al. [25] Definition 3.1)
A stochastic process \((S(i), i \geq 1)\), with state space \((S, B(S))\), is a hidden Markov model if the following hold:

- **(Hidden chain)** We are given (but do not observe) a strictly stationary Markov chain \(U(1), U(2), \ldots, U(z), \ldots\) with state space \((U, B(U))\).
- For all \(z\), given \((U(1), U(2), \ldots, U(z))\), the \(S(i), i = 1, \ldots, z\), are conditionally independent, and the conditional distribution of \(S(i)\) depends only on \(U(i)\).
- The conditional distribution of \(S(i)\) given \(U(i) = u\) does not depend on \(i\).

**Proposition 1. (Genon et al. [25] Proposition 3.1)**
The process \((S(i), i \geq 1)\) is strictly stationary. If the hidden Markov chain \((U(i), i \geq 1)\) is ergodic, then \((S(i), i \geq 1)\) is also ergodic.

For a proof see Genon et al. [25].

Now consider a Markov process which is defined by a stochastic equation

\[
    dX = \mu(X)dt + \sigma(X)dZ, \quad X_0 = \eta,
\]

(2.43)
where \( W \) is a standard Brownian motion and \( \eta \) is a real random variable defined on \( \Omega \) and independent of \( Z \). Genon et al. [25] make the standard assumptions on functions \( \mu(x) \) and \( \sigma(x) \), ensuring that the solution (see Assumption 1) of (2.43) is a positive recurrent diffusion on an interval \( (l, r) \), \(-\infty \leq l < r \leq \infty\), (see Assumption 2) and a strictly stationary ergodic time-reversible process (see Assumption 3).

**Assumption 1.**

The functions \( \mu(x) \) and \( \sigma(x) \) are defined on \( (l, r) \), and satisfy

\[
\mu(x) \in C^1(l, r), \sigma^2(x) \in C^2(l, r), \sigma(x) > 0 \quad \forall x \in (l, r),
\]

and

\[
\exists K > 0, \forall x \in (l, r), |\mu(x)| \leq K(1 + |x|) \text{ and } \sigma^2(x) \leq K(1 + x^2).
\]

For \( x_0 \in (l, r) \), define the scale and speed densities of diffusion \( (X_t) \),

\[
s(x) = \exp(-2 \int_{x_0}^{x} \frac{\mu(u)}{\sigma^2(u)} du), \quad m(x) = \frac{1}{\sigma^2(x)s(x)}.
\]

**Assumption 2.**

\[
\int_{l}^{\infty} s(x)dx = \infty, \int_{r}^{-\infty} s(x)dx = \infty, \int_{l}^{r} m(x)dx = M < \infty.
\]

Let us define the stationary density

\[
\pi^*(x) = \frac{1}{M}m(x)1_{x \in (l, r)}.
\]

**Assumption 3.**

The initial random variable \( \eta \) has distribution \( \pi^*(dx) = \pi^*(x)dx \).

Now let \( (Y, V)_{t \geq 0} \) be a two-dimensional diffusion process given by

\[
dY = \sigma dZ, Y(0) = 0,
\]

\[
V = \sigma^2, dV = d(b - V)dt + c\sqrt{V}dQ, V(0) = \eta,
\]

where \( \eta \) is a random variable, independent of \( (Z, Q) \), and assume that Assumptions 1 to 3 hold for \( (Y, V)_{t \geq 0} \). Then, the diffusion \( V \) is strictly stationary, ergodic and time-reversible.
(see Proposition 1 and Genon et al. [25]). For positive $\Delta t$ Genon-Catalot et al. [25] define, for $i \geq 1$,

$$S(i) = \frac{1}{\sqrt{\Delta t}} \int_{(i-1)\Delta t}^{i\Delta t} \sigma_s dZ_s$$

(2.51)

and

$$U(i) = (V(i), V(\Delta t)), \quad V(i) = \frac{1}{\Delta t} \int_{(i-1)\Delta t}^{i\Delta t} V_s ds,$$  

(2.52)

where $V(\Delta t)$ are values of $V$ at fixed intervals $\Delta t$.

**Proposition 2.** (Genon-Catalot et al. [25] Proposition 4.1)

Assume that the above hidden diffusion $V$ satisfies the Assumptions 1 to 3 and that $E(V(0)^2)$ is finite. Then, $E(\overline{V}(1)) = E(\overline{V}(0)) = \beta$, and

$$E(\overline{V}(1)^2) = b^2 + \frac{c^2 b^2 (d\Delta t - 1 + e^{-d\Delta t})}{d^2(\Delta t)^2},$$

(2.53)

$$E(\overline{V}(1)\overline{V}(2)) = b^2 + \frac{c^2 b^2 (1 - e^{-d\Delta t})^2}{2d^2(\Delta t)^2}.$$  

(2.54)

For a proof see Genon-Catalot et al. [25].

The following functions of the observations

$$\hat{m}_1 = \frac{1}{n} \sum_{i=1}^{n} Z(i)^2, \quad \hat{m}_2 = \frac{1}{3n} \sum_{i=1}^{n} Z(i)^4, \quad \hat{m}_{12} = \frac{1}{n} \sum_{i=1}^{n-1} Z(i)^2 Z(i+1)^2,$$

(2.55)

where $n$ is the number of observations, are consistent estimators of $b^2, E(\overline{V}(1)^2)$ and $E(\overline{V}(1)\overline{V}(2))$ respectively. For a proof see Genon-Catalot et al. [25].

### 2.7.2 Application for the parameter estimation of the eigenvalue processes

In our multidimensional model an orthogonal transformation of the returns leads to a set of univariate series which can be estimated using standard univariate methods.

For our estimation, we propose a two steps estimation procedure. First, the constant orthogonal matrix is found by singular decomposition on the stationary covariance structure to compute $A^*$, then each principal component is modeled as independent stochastic volatility model (SV) of the kind of Cox-Ingersoll and Ross, also known by the work of
Heston (1993). For the estimation of the singular SV models we choose the procedure of Genon-Catalot et al. [25].

In this thesis we assume that the conditions to use the estimation method of Catalot-Genon et al. [25] are satisfied. The proof is beyond the scope of the thesis.

We consider a diffusion process \((B(t), D(t))\), where we regard \(B(t)\), the matrix of the principal components, as observable at \(n\) discrete times with regular sampling interval \(\Delta t\) and denote the observations by \(B(\Delta t)\). \(D(t)\), the matrix of the eigenvalues, is assumed to be ergodic and rules the diffusion of \((B(\Delta t))\). The observations can be viewed as a hidden Markov model. For \(\Delta t\) positive, we define, for \(i \geq 1\),

\[
B(i) = \frac{1}{\sqrt{\Delta t}} \int_{(i-1)\Delta t}^{i\Delta t} \sqrt{\lambda(s)} dZ
\]

and

\[
D(i) = \frac{1}{\Delta t} \int_{(i-1)\Delta t}^{i\Delta t} D(s) d s.
\]

We subtract in each sampling interval the mean \(\bar{V}_i\) from all data series \(V_i\) to obtain \(m\) data sets whose mean is zero. For these \(m\) sets we compute the first two principal components to obtain two-dimensional \(B(i)\).

Thus, for our model (2.11 - 2.15) we can apply that

\[
E(D(1)) = E(D(0)) = b,
\]

\[
E(D(1)^2) = b^2 + \frac{c^2}{2d} \frac{2d(1+e^{-d\Delta t})}{d^2(\Delta t)^2},
\]

\[
E(D(1)D(2)) = b^2 + \frac{c^2}{2d} \frac{(1-e^{-d\Delta t})^2}{d^2(\Delta t)^2}.
\]

Consistent estimators of \(b, E(D(1)^2)\) and \(E(D(1)D(2))\) are the following functions of the observations

\[
\hat{m}_1 = \frac{1}{n} \sum_{i=1}^{n} Z(i)^2, \quad \hat{m}_2 = \frac{1}{3n} \sum_{i=1}^{n} Z(i)^4, \quad \hat{m}_{12} = \frac{1}{n} \sum_{i=1}^{n-1} Z(i)^2 Z(i+1)^2.
\]

which allow us to estimate the parameters \(d, b, c\) for the diffusions of the first two eigenvalues.

The literature on estimation methods for continuous-time multidimensional stochastic volatility processes is almost non-existent. This method, like any other moment-based procedure, has several difficulties for small and medium sample sizes as well as for some ranges of the parameters (see Sørensen [55] for a survey). We have, however, already tested the method on simulations, i.e. we tried to retrieve the parameters of the simulations using our parameter estimation method. This showed that the results for the mean-reversion level and the mean-reversion speed were quite close.
2.8 Example

In the following we value the tranches of a CDO. The underlying security is a portfolio of coupon paying corporate bonds on 10 firms. We assume the face values $K_i$, $i \in \{1, \ldots, 10\}$ of these bonds to be 1 and assume that the process of the company values show the same characteristics as the respective stock price processes. Thus, in order to estimate the parameters of the eigenvalue process we use the stock returns of ten companies (MMM.N, ABT.N, AA.N, AXP.N, BUD.N, AVP.N, BAX.N, BDK.N, BNI.N, BMY.N) from 1983 to October 2006. We use the parameter estimation method described above for the mean-reversion level and the mean-reversion speed. As the method has still to be improved, considering the volatilities of the eigenvalue, we assume plausible values for the latter. Thus, we value the tranches for the following scenario:

| Basic Scenario: $r = 0.05$, Maturity = 1 year, $\Delta t = \frac{1}{n}$, $n = 3$, $b_1 = 0.2713032$, $b_2 = 0.126441$, $d_1 = d_2 = 0.5$, $c_1 = 0.14$, $c_2 = 0.12$, $K_i = 0.75 \cdot V_i(0)$, Tranche 1 = from 0 to 1 default, Tranche 2 = from 2 to 3 defaults, Tranche 3 = from 3 to 4 defaults, Tranche 4 = from 4 to 5 defaults, Tranche 5 = 5 defaults |

We value the tranches using the path-dependent and the non path-dependent mean value computation in order to compare the distributions we get for the mean values in $T$ and the values of the tranches. The results for the values of the tranches of these two methods are close (see Tables 2.2 and 2.3).

Table 2.2: Values of the Tranches using the path dependent method

<table>
<thead>
<tr>
<th>Tranche</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1140942079</td>
</tr>
<tr>
<td>2</td>
<td>0.0745104126</td>
</tr>
<tr>
<td>3</td>
<td>0.0727544611</td>
</tr>
<tr>
<td>4</td>
<td>0.0543120038</td>
</tr>
<tr>
<td>5</td>
<td>0.0377722636</td>
</tr>
</tbody>
</table>

Furthermore, we analyze the distribution of the means of the two methods. Using the Mann-Whitney Test we cannot reject the $H_0$ hypothesis that the two distributions are alike. The two distributions are shown in the appendix and analyzed in Tables 2.4 and 2.5.

However, the values of the tranches differ more when we assume higher volatilities for the process of the eigenvalues. In the following we have calculated the value for the volatilities $c_1=0.25$ and $c_2=0.13$ (see Tables 2.6 and 2.7).
Table 2.3: Values of the Tranches using the non path-dependent method

<table>
<thead>
<tr>
<th>Tranche</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1149043351</td>
</tr>
<tr>
<td>2</td>
<td>0.0745104126</td>
</tr>
<tr>
<td>3</td>
<td>0.0727544611</td>
</tr>
<tr>
<td>4</td>
<td>0.0543120038</td>
</tr>
<tr>
<td>5</td>
<td>0.0376654103</td>
</tr>
</tbody>
</table>

Table 2.4: Moments of the Mean Values

<table>
<thead>
<tr>
<th></th>
<th>path-dependent Method</th>
<th>Non path-dependent Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.026241208</td>
<td>0.026125472</td>
</tr>
<tr>
<td>Variance</td>
<td>4.92252E−06</td>
<td>5.82982E−06</td>
</tr>
<tr>
<td>Skewness</td>
<td>−0.242396164</td>
<td>−0.195523145</td>
</tr>
<tr>
<td>Minimum</td>
<td>0.018953</td>
<td>0.018953</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.031259</td>
<td>0.031259</td>
</tr>
</tbody>
</table>

The deviation in value of the first tranche can be explained by the fact that the differences in skewness between the distributions of these two methods become higher for higher volatilities of the eigenvalue process although the span of both distributions, i.e. minimum and maximum values, remains the same for both distributions (see Tables 2.4 and 2.8). This imprecision can be handled by using more time steps for the computation of the tranches when the processes of the eigenvalues are quite volatile. The two distributions are shown in the appendix and analyzed in Tables 2.8 and 2.9.
Table 2.5: Classification of the Mean Values

<table>
<thead>
<tr>
<th>Interval</th>
<th>From</th>
<th>To</th>
<th>path-dependent Method</th>
<th>Non path-dependent Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.018</td>
<td>0.019</td>
<td>0.00050859</td>
<td>0.00050859</td>
</tr>
<tr>
<td>2</td>
<td>0.019</td>
<td>0.02</td>
<td>0.00631195</td>
<td>0.00631195</td>
</tr>
<tr>
<td>3</td>
<td>0.02</td>
<td>0.021</td>
<td>1.93E – 05</td>
<td>1.93E – 05</td>
</tr>
<tr>
<td>4</td>
<td>0.021</td>
<td>0.022</td>
<td>0.0555033</td>
<td>0.0665629</td>
</tr>
<tr>
<td>5</td>
<td>0.022</td>
<td>0.023</td>
<td>0.00628108</td>
<td>0.06341643</td>
</tr>
<tr>
<td>6</td>
<td>0.023</td>
<td>0.024</td>
<td>0.1024631</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0.024</td>
<td>0.025</td>
<td>0.00262061</td>
<td>0.037027</td>
</tr>
<tr>
<td>8</td>
<td>0.025</td>
<td>0.026</td>
<td>0.407302</td>
<td>0.4609296</td>
</tr>
<tr>
<td>9</td>
<td>0.026</td>
<td>0.027</td>
<td>0.2555679</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>0.027</td>
<td>0.028</td>
<td>0.18859835</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>0.028</td>
<td>0.029</td>
<td>0.050709</td>
<td>0.305454</td>
</tr>
<tr>
<td>12</td>
<td>0.029</td>
<td>0.03</td>
<td>0.11467721</td>
<td>0.02033294</td>
</tr>
<tr>
<td>13</td>
<td>0.03</td>
<td>0.031</td>
<td>0.0384517</td>
<td>0.0369733</td>
</tr>
<tr>
<td>14</td>
<td>0.031</td>
<td>0.032</td>
<td>9.83E – 04</td>
<td>0.00246123</td>
</tr>
</tbody>
</table>

2.9 Summary and Conclusion

We have developed and implemented a tree model to price tranches of a CDO on a portfolio of stochastically correlated underlyings. Dimensionality and complexity can be reduced by using only the first few eigenvectors and eigenvalues of the underlyings instead of the actual values. We have built trees for the principal components and the stochastic eigenvalues, which allows us to compute the underlying probabilities of default in this portfolio. This model relaxes the constant correlation assumption in the existing literature. Furthermore, we provide a method to match the parameters of the stochastic processes of the principal components as well as the eigenvalues to market data and show the parameter estimation and the valuation of the tranches for an example.
Table 2.6: Values of the Tranches using the path-dependent method with higher volatilities

<table>
<thead>
<tr>
<th>Tranche</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1291563212</td>
</tr>
<tr>
<td>2</td>
<td>0.0745361782</td>
</tr>
<tr>
<td>3</td>
<td>0.0727397773</td>
</tr>
<tr>
<td>4</td>
<td>0.0544242608</td>
</tr>
<tr>
<td>4</td>
<td>0.0392558059</td>
</tr>
</tbody>
</table>

Table 2.7: Values of the Tranches using the non path-dependent method with higher volatilities

<table>
<thead>
<tr>
<th>Tranche</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1423252046</td>
</tr>
<tr>
<td>2</td>
<td>0.0745268594</td>
</tr>
<tr>
<td>3</td>
<td>0.0727548047</td>
</tr>
<tr>
<td>4</td>
<td>0.0544113424</td>
</tr>
<tr>
<td>5</td>
<td>0.0406599125</td>
</tr>
</tbody>
</table>

Table 2.8: Moments of the Mean Values

<table>
<thead>
<tr>
<th></th>
<th>path-dependent Method</th>
<th>Non path-dependent Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.026235437</td>
<td>0.026026984</td>
</tr>
<tr>
<td>Variance</td>
<td>1.82868E−05</td>
<td>1.54905E−05</td>
</tr>
<tr>
<td>Skewness</td>
<td>−0.432561011</td>
<td>−0.330387539</td>
</tr>
<tr>
<td>Minimum</td>
<td>0.013258</td>
<td>0.013258</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.034434</td>
<td>0.034434</td>
</tr>
</tbody>
</table>
Table 2.9: Classification

<table>
<thead>
<tr>
<th>Interval</th>
<th>From</th>
<th>To</th>
<th>path-dependent Method</th>
<th>Non path-dependent Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.013</td>
<td>0.014</td>
<td>0.00480658</td>
<td>0.00438662</td>
</tr>
<tr>
<td>2</td>
<td>0.014</td>
<td>0.015</td>
<td>0.004299903</td>
<td>0.004719743</td>
</tr>
<tr>
<td>3</td>
<td>0.015</td>
<td>0.016</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.016</td>
<td>0.017</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.017</td>
<td>0.018</td>
<td>0.0207461</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.018</td>
<td>0.02</td>
<td>0.05232637</td>
<td>0.15160554</td>
</tr>
<tr>
<td>8</td>
<td>0.02</td>
<td>0.021</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.021</td>
<td>0.022</td>
<td>0.0612924</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.0224</td>
<td>0.023</td>
<td>0.05483224</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.0234</td>
<td>0.024</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0.024</td>
<td>0.026</td>
<td>0.417319</td>
<td>0.5115732</td>
</tr>
<tr>
<td>14</td>
<td>0.026</td>
<td>0.027</td>
<td>0.0126069</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.027</td>
<td>0.028</td>
<td>0.0124291</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>0.0284</td>
<td>0.029</td>
<td>0.1732454</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>0.029</td>
<td>0.036</td>
<td>0.00560917</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>0.03</td>
<td>0.0314</td>
<td>0.009681</td>
<td>0.142688</td>
</tr>
<tr>
<td>19</td>
<td>0.031</td>
<td>0.032</td>
<td>0.1349391</td>
<td>0.15352308</td>
</tr>
<tr>
<td>20</td>
<td>0.032</td>
<td>0.033</td>
<td>0.00436899</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>0.033</td>
<td>0.034</td>
<td>0.0279369</td>
<td>0.0151704</td>
</tr>
<tr>
<td>22</td>
<td>0.034</td>
<td>0.035</td>
<td>0.003556733</td>
<td>0.016323433</td>
</tr>
</tbody>
</table>
Appendix A

General Trinomial Tree Probabilities

In the trinomial tree the probabilities of \( y(l, t) \) moving to \( y(\kappa - 1, t) \), \( y(\kappa, t) \) and \( y(\kappa + 1, t) \) are chosen to match the first and second moments of the three point-jump process of the change in \( y(l, t) \) to the continuous distribution. Thus, the following equations must be satisfied:

\[
p_{l,\kappa - 1}(\kappa - 1 - l)\Delta y + p_{l,\kappa}(\kappa - l)\Delta y + p_{l,\kappa + 1}(\kappa + 1 - l)\Delta y = \mu \Delta t \tag{A.1}
\]

\[
p_{l,\kappa - 1}(\kappa - 1 - l)^2\Delta y^2 + p_{l,\kappa}(\kappa - l)^2\Delta y^2 + p_{l,\kappa + 1}(\kappa + 1 - l)^2\Delta y^2 - (\mu \Delta t)^2 = c^2 \Delta t \tag{A.2}
\]

\[
p_{l,\kappa - 1} + p_{l,\kappa} + p_{l,\kappa + 1} = 1 \tag{A.3}
\]

where

\[\mu = a(b - y).\]

It follows from (A.3)

\[p_{l,\kappa} = 1 - p_{l,\kappa - 1} - p_{l,\kappa + 1}. \tag{A.4}\]

Substituting Equation (A.4) in (A.1) and reformulating it we get \((\mu(l, t) := \mu)\)

\[-p_{l,\kappa - 1}\Delta y + (\kappa - l)\Delta y + p_{l,\kappa + 1}\Delta y = \mu \Delta t, \tag{A.5}\]

which is equivalent to

\[p_{l,\kappa + 1} = \frac{\mu \Delta t}{\Delta y} - (\kappa - l) + p_{l,\kappa - 1} \tag{A.6}\]

Substituting Equations (A.4) and (A.6) in (A.2) we have

\[
p_{l,\kappa - 1}(\kappa - 1 - l)^2(\Delta y)^2 + \\
(1 - p_{l,\kappa - 1} - \frac{\mu \Delta t}{\Delta y} + (\kappa - l) - p_{l,\kappa - 1})(\kappa - l)^2(\Delta y)^2 + \\
+\left(\frac{\mu \Delta t}{\Delta y} - (\kappa - l) + p_{l,\kappa - 1}\right)(\kappa + 1 - l)^2(\Delta y)^2 = \mu^2(\Delta t)^2 + c^2 \Delta t,
\]
which is equivalent to

\[ p_{l,\kappa - 1} = \frac{c^2 \Delta t}{2(\Delta y)^2} + \frac{\mu^2 (\Delta t)^2 + 2(l - \kappa) \Delta y \mu \Delta t + (l - \kappa)^2 (\Delta y)^2}{2(\Delta y)^2} - \frac{\eta}{2} \Delta y \text{ (A.7)} \]

With \( \eta = \mu \Delta t + (l - \kappa) \Delta y \)

\[ p_{l,\kappa - 1} = \frac{c^2 \Delta t}{2(\Delta y)^2} + \frac{\eta^2}{2(\Delta y)^2} - \frac{\eta}{2} \Delta y. \]

Substituting Equation (A.7) in (A.6) we get

\[ p_{l,\kappa + 1} = \frac{c^2 \Delta t}{2(\Delta y)^2} + \frac{\eta^2}{2(\Delta y)^2} + \frac{\eta}{2} \Delta y \text{ (A.8)} \]

Substituting Equations (A.7) and (A.8) in (A.4) we obtain

\[ p_{l,\kappa} = 1 - \frac{c^2 \Delta t}{(\Delta y)^2} - \frac{\eta^2}{(\Delta y)^2} \text{ (A.9)} \]
Appendix B

Specific Choice of Tree Probabilities

As Hull and White do not provide the proof for this dynamic rule in their paper the restrictions are shown in the following. If \( \kappa = l \) the probabilities are

\[
p_{l,l+1} = \frac{c^2 \Delta t}{2(\Delta y)^2} + \frac{\mu^2(\Delta t)^2}{2(\Delta y)^2} + \frac{\mu \Delta t}{2\Delta y} \tag{B.1}
\]

\[
p_{l,l} = 1 - \frac{e^2 \Delta t}{(\Delta y)^2} - \frac{\mu^2(\Delta t)^2}{(\Delta y)^2} \tag{B.2}
\]

\[
p_{l,l-1} = \frac{c^2 \Delta t}{2(\Delta y)^2} + \frac{\mu^2(\Delta t)^2}{2(\Delta y)^2} - \frac{\mu \Delta t}{2\Delta y} \tag{B.3}
\]

To obtain positive probabilities, which are smaller than 1 we have to ensure that Equations (B.1) to (B.3) are positive:

\[
p_{l,l+1} = \frac{c^2 \Delta t}{2(\Delta y)^2} + \frac{\mu^2(\Delta t)^2}{2(\Delta y)^2} + \frac{\mu \Delta t}{2\Delta y} \geq 0 \iff \frac{\mu^2(\Delta t)^2 + \mu \Delta t \Delta y}{c^2 \Delta t} \geq -1
\]

Substituting \( c^2 \Delta t = \frac{1}{3} (\Delta y)^2 \) results in:

\[
\frac{\mu^2(\Delta t)^2 + \mu \Delta t \Delta y}{(\Delta y)^2} \geq -\frac{1}{3} \iff \left( \frac{\mu \Delta t}{\Delta y} + \frac{1}{2} \right)^2 + \frac{1}{12} \geq 0,
\]

which does not impose any constraints on the parameters.

\[
p_{l,l} = 1 - \frac{c^2 \Delta t}{(\Delta y)^2} - \frac{\mu^2(\Delta t)^2}{(\Delta y)^2} \geq 0 \iff \frac{(\Delta y)^2 - \mu^2(\Delta t)^2}{c^2 \Delta t} \geq 1
\]

Substituting \( c^2 \Delta t = \frac{1}{3} (\Delta y)^2 \) results in:

\[
\frac{(\Delta y)^2 - \mu^2(\Delta t)^2}{(\Delta y)^2 \Delta t} \geq \frac{1}{3} \iff -\sqrt{\frac{2}{3}} \leq \frac{\mu \Delta t}{\Delta y} \leq \sqrt{\frac{2}{3}}
\]

\[
p_{l,l-1} = \frac{c^2 \Delta t}{2(\Delta y)^2} + \frac{\mu^2(\Delta t)^2}{2(\Delta y)^2} - \frac{\mu \Delta t}{2\Delta y} \geq 0 \iff \frac{\mu \Delta t \Delta y - \mu^2(\Delta t)^2}{\sigma^2 \Delta t} \leq 1
\]

Substituting \( c^2 \Delta t = \frac{1}{3} (\Delta y)^2 \) results in:

\[
\frac{\mu \Delta t \Delta y - \mu^2(\Delta t)^2}{\sigma^2 \Delta t} \leq \frac{1}{3} \iff -(\frac{\mu \Delta t}{\Delta y} - \frac{1}{2})^2 - \frac{1}{12} \leq 0,
\]
which does not impose any constraints on the parameters. Thus,
\[ -\sqrt{\frac{2}{3}} \leq \frac{\mu \Delta t}{\Delta y} \leq \sqrt{\frac{2}{3}} \].

If \( \kappa = l + 1 \)

\[ p_{l,t+2} = \frac{c^2 \Delta t}{2(\Delta y)^2} + \frac{\mu^2(\Delta t)^2}{2(\Delta y)^2} - \frac{\mu \Delta t}{2(\Delta y)} \geq 0 \iff \frac{\mu \Delta t \Delta y - \mu^2(\Delta t)^2}{c^2 \Delta t} \leq 1 \]

Substituting \( c^2 \Delta t = \frac{1}{3}(\Delta y)^2 \) results in:
\[ \frac{\mu \Delta t \Delta y - \mu^2(\Delta t)^2}{(\Delta y)^2} \leq \frac{1}{3} \iff -\left( \frac{\mu \Delta t}{\Delta y} - \frac{1}{2} \right)^2 - \frac{1}{12} \leq 0, \]

which does not impose any constraints on the parameters.

\[ p_{l,t+1} = -\frac{c^2 \Delta t}{2(\Delta y)^2} + \frac{2\mu(l,l)\Delta t}{2(\Delta y)^2} - \frac{\mu^2(\Delta t)^2}{4(\Delta y)^2} \geq 0 \iff \frac{2\mu(l,l)\Delta t \Delta y - \mu^2(\Delta t)^2}{c^2 \Delta t} \geq 1 \]

Substituting \( c^2 \Delta t = \frac{1}{3}(\Delta y)^2 \) results in:
\[ \frac{2\mu(l,l)\Delta t \Delta y - \mu^2(\Delta t)^2}{(\Delta y)^2} \geq \frac{1}{3} \iff 1 - \sqrt{\frac{2}{3}} \leq \frac{\mu \Delta t}{\Delta y} \leq 1 + \sqrt{\frac{2}{3}} \]

\[ p_{l,t} = 1 + \frac{c^2 \Delta t}{2(\Delta y)^2} + \frac{\mu^2(\Delta t)^2}{2(\Delta y)^2} - \frac{3\mu \Delta t}{2(\Delta y)} \geq 0 \iff -\frac{2(\Delta y)^2 - \mu^2(\Delta t)^2 + 3\mu(l,l)\Delta t \Delta y}{2(\Delta y)^2} \leq 1 \]

Substituting \( c^2 \Delta t = \frac{1}{3}(\Delta y)^2 \) results in:
\[ -\frac{2(\Delta y)^2 - \mu^2(\Delta t)^2 + 3\mu(l,l)\Delta t \Delta y}{(\Delta y)^2} \leq \frac{1}{3} \]

As \( \frac{3\mu \Delta t \Delta y - \mu^2(\Delta t)^2}{(\Delta y)^2} \leq 1 \)

\[ -\frac{2(\Delta y)^2 + 2\mu(l,l)^2(\Delta t)^2}{(\Delta y)^2} \leq -\frac{2}{3} \iff -\sqrt{\frac{2}{3}} \leq \frac{\mu \Delta t}{\Delta y} \leq \sqrt{\frac{2}{3}} \]

Therefore for \( \kappa = l + 1 \)

\[ 1 - \sqrt{\frac{2}{3}} \leq \frac{\mu \Delta t}{\Delta y} \leq \sqrt{\frac{2}{3}} \].

If \( \kappa = l - 1 \)

\[ p_{l,t-1} = \frac{c^2 \Delta t}{2(\Delta y)^2} - \frac{2\mu(l,l)\Delta t}{2(\Delta y)^2} - \frac{\mu^2(\Delta t)^2}{2(\Delta y)^2} \geq 0 \iff \frac{2\mu(l,l)\Delta t \Delta y + \mu^2(\Delta t)^2}{c^2 \Delta t} \leq -1 \]

Substituting \( c^2 \Delta t = \frac{1}{3}(\Delta y)^2 \) results in:
\[ \frac{2\mu(l,l)\Delta t \Delta y + \mu^2(\Delta t)^2}{(\Delta y)^2} \leq -\frac{1}{3} \iff -1 - \sqrt{\frac{2}{3}} \leq \frac{\mu \Delta t}{\Delta y} \leq -1 + \sqrt{\frac{2}{3}} \]

\[ p_{l,t-2} = \frac{c^2 \Delta t}{2(\Delta y)^2} + \frac{\mu^2(\Delta t)^2}{2(\Delta y)^2} + \frac{\mu \Delta t}{2(\Delta y)} \geq 0 \iff \frac{\mu \Delta t \Delta y + \mu^2(\Delta t)^2}{c^2 \Delta t} \geq -1 \]
Substituting $c^2 \Delta t = \frac{1}{3} (\Delta y)^2$ results in:

$$\frac{\mu \Delta t \Delta y + \mu^2 \Delta t}{(\Delta y)^2} \geq -\frac{1}{3} \iff (\frac{\mu \Delta t}{\Delta y} + \frac{1}{2})^2 + \frac{1}{12} \geq 0,$$

which does not impose any constraints on the parameters.

$$p_{l,t} = 1 + \frac{c^2 \Delta t}{2(\Delta y)^2} + \frac{\mu^2(\Delta t)^2}{2(\Delta y)^2} + \frac{3 \mu \Delta t}{(\Delta y)^2} \geq 0 \iff \frac{-2(\Delta y)^2 - \mu^2(\Delta t)^2 - 3\mu(l,t)\Delta t \Delta y}{c^2 \Delta t} \leq 1$$

Substituting $c^2 \Delta t = \frac{1}{3} (\Delta y)^2$ results in:

$$\frac{-2(\Delta y)^2 - \mu^2(\Delta t)^2 - 3\mu(l,t)\Delta t \Delta y}{(\Delta y)^2 \Delta t} \leq \frac{1}{3}$$

As $3\frac{-\mu \Delta t \Delta y - \mu^2(\Delta t)^2}{(\Delta y)^2 \Delta t} \leq 1$

$$\frac{(-2\Delta y)^2 + 2\mu(l,t)^2(\Delta t)^2}{(\Delta y)^2 \Delta t} \leq -\frac{2}{3} \iff -\sqrt{2} < \frac{\mu \Delta t}{\Delta y} \leq \sqrt{2}$$

Hence, for $\kappa = l - 1$

$$-\sqrt{\frac{2}{3}} \leq \frac{\mu \Delta t}{\Delta y} \leq -1 + \sqrt{\frac{2}{3}}$$

Considering all three branching methods the following dynamic rules for the choice of the parameter $\kappa$ can be derived:

$$\kappa = \begin{cases} 
  l + 1 & \text{if } \frac{\mu \Delta t}{\Delta y} \geq \sqrt{\frac{2}{3}} \\
  l & \text{if } -\sqrt{\frac{2}{3}} < \frac{\mu \Delta t}{\Delta y} < \sqrt{\frac{2}{3}} \\
  l - 1 & \text{if } \frac{\mu \Delta t}{\Delta y} \leq -\sqrt{\frac{2}{3}} 
\end{cases} \quad (B.4)$$
Appendix C

Results of the Pricing of an Example

C.1 Mean Values

Figure C.1: Mean Values using the path-dependent method
C. Results of the Pricing of an Example

C.2 Mean Values using higher volatilities

Figure C.2: Mean Values using the non path-dependent method

Figure C.3: Mean Values using the path-dependent method
Figure C.4: Mean Values using the non path-dependent method
Bibliography


