Financial Engineering of the Stochastic Correlation in Credit Risk Models

by

Hamidreza Arian

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy,
Graduate Department of Mathematics
University of Toronto

Abstract

Financial Engineering of the Stochastic Correlation
in Credit Risk Models

Hamidreza Arian
Doctor of Philosophy
Graduate Department of Mathematics
University of Toronto
Hamidreza Arian
(2010)

The main objective of this thesis is to implement stochastic correlation into the existing structural credit risk models. There are two stochastic models suggested for the covariance matrix of the assets' prices. In our first model, to induce the stochasticity into the structure of the correlation, we assume that the eigenvectors of the covariance matrix are constant but the eigenvalues are driven by independent Cox-Ingersoll-Ross processes. To price equity options on this framework we first transform the calculations from the pricing domain to the frequency domain. Then we derive a closed formula for the Fourier transform of the Green's function of the pricing PDE. Finally we use the method of images to find the price of the equity options. The same method is used to find closed formulas for marginal probabilities of defaults and CDS prices. In our second model, the covariance of the assets follows a Wishart process, which is an extension of the CIR model to dimensions greater than one. The popularity of the Heston model, which uses the CIR process to model the stochastic volatility, could be a promising point for using Wishart process to model stochastic correlation. We give closed form solutions for equity options, marginal probabilities of defaults, and some other major financial derivatives. For the calculation of our pricing formulas we make a bridge between two recent trends in pricing theory; from one side, pricing of barrier options by Lipton (2001) and Sepp (2006) and from other side the development of Wishart processes by Bru (1991), Gourieroux (2005) and Fonseca *et al.* (2007a, 2006, 2007b). After obtaining the mathematical results above, we then estimate the parameters of the two models we have developed by an evolutionary algorithm. We prove a theorem which guarantees the convergence of the evolutionary algorithm to the set of optimizing parameters. After estimating the parameters of the two stochastic correlation models, we conduct a comparative analysis of our stochastic correlation models. We give an approximation formula for the joint and marginal probabilities of default for General Motors and Ford. For the *marginal* probabilities of default, a closed formula is given and for the *joint* probabilities of default an approximation formula is

suggested. To show the convergence properties of this approximation method, we perform the Monte Carlo simulation in two forms: a full and a partial Monte Carlo simulation. At the end, we compare the marginal and joint probabilities with full and partial Monte Carlo simulations.

To Mehry, Reza and Ismael

Acknowledgments

I would like to thank my thesis advisors Prof. Luis Seco and Prof. Marcos Escobar for their limitless support, guidance and helpful discussions throughout my doctoral studies. I hope that the end of my doctoral program will become the beginning of a lifelong friendship with Luis and Marcos.

I sincerely thank Prof. Matt Davison, Prof. Robert Jerrard, Prof. Robert McCann, Prof. Adrian Nachman and Prof. Mary Pugh for their helpful comments on my thesis.

I would like to thank Ida Bulat, the angel of our department, for all her patience, kindness and help during these years.

I would also like to thank Prof. Sebastian Ferrando from Ryerson University for providing me with their Applied Mathematics Lab. I thank Khurrum Jawaid, the assistant of the Applied Mathematics Lab at Ryerson University for all his help. I also thank Ioannis Anapolitanos from University of Toronto's Mathematics department for his helpful comments and discussions on the PDE theory related to my work.

I would also like to thank Prof. Sebastian Jamingual and his group in the statistics department of the University of Toronto. I have learned a lot in their biweekly mathematical finance seminars in the Fields institute.

Many people have influenced my interests and knowledge of science in the past whom I will never forget. Among them, I would like to thank Prof. Hossein Hamedani, Prof. Esmaeil Mahdavi Harsini, Prof. Hassan Nojoumi, Prof. Behzad Rouhani, Prof. Hossein Seyyed Mousavi, Prof. Mehrdad Shahshahani, Prof. Shiva Zamani and Prof. Bijan Z. Zangeneh.

I would like to thank my parents, Mehry and Reza to whom this thesis is dedicated.

I thank my lovely sisters Zoya and Miriam and all my other family members, especially my grandparents, for all the limitless love they give me. Without them, the whole world means nothing to me.

List of Publications

- *H. Arian, M. Escobar and L. Seco.*(2008). The CreditGrades Model with Stochastic Covariance Wishart Process , *submitted*
- *H. Arian, M. Escobar and L. Seco.*(2008). Pricing and Calibration of a Stochastic Correlation CreditGrades Model , *submitted*
- *H. Arian, M. Escobar and L. Seco.*(2008). Derivative Pricing with Stochastic Switching Correlations , *RiskLab Working Paper*
- *H. Arian, M. Escobar and L. Seco.*(2008). A Comparative Analysis of two Stochastic Correlation Credit Risk Models, *RiskLab Working Paper*

Contents

Chapter 1

Introduction

1.1 Literature Overview

In 1974, Merton presented a model which assumed equity as an option on company assets and the debt as the strike price of the option. In the event that the asset price lies below the debt value at maturity, default occurs. Merton's model has several drawbacks. First, the default can only happen at maturity. Black and Cox (1976) were the first to give an extension in which the default could happen prior to maturity time. Second, Merton's model doesn't give a straightforward connection with equity markets, while equity options are widely traded in the market and they bear a very rich set of information about the company's credit. In fact, to the best of our knowledge Hull *et al.* (2004) is the first to calibrate the parameters of the Merton's model with equity options by considering equity options as compound options on the firm's asset. They use results from Geske (1979) to identify equity options as compound options on the assets of the company. Their work is very interesting, especially because most of the dynamic information about a company can be obtained from the options on the firm's equity. This is a very strong advantage when it comes to calibrating the model's parameters. When volatility/correlation are stochastic Geske's results are no longer analytically tractable. Since Hull *et al.* (2004) assumes a deterministic volatility/correlation in their model, we think their model fails to give a closed form formula for equity options in the presence of stochastic volatility/correlation. One of the popular versions of Merton's structural credit risk model is the CreditGrades model. The CreditGrades model was jointly developed by CreditMetrics, JP Morgan, Goldman Sachs and Deutsche Bank. The original version of the CreditGrades model assumes that volatility is deterministic. In this thesis, we extend the CreditGrades model, by means of stochastic covariance Wishart processes focusing on the role of stochastic correlation.

The first mathematical result for valuation of options in the presence of stochastic volatility was given by Heston (1993). In that article closed-form formulas for the characteristic function and the probability distribution of the joint logprice-volatility process $(\ln A_t, \lambda_t)$ were obtained. The results in that paper also give a closed-form solution for the valuation problem of a vast range of popular financial derivatives on A_t. Implementation of the model has been done with direct integration, Fast Fourier Transform (FFT) and fractional FFT methods assuming the asset price is observable (see Kilin (2007)). Even though there are many results on both stochastic volatility as well as on the modeling of financial instruments with Levy processes, there is not much work done regarding stochastic correlation. Market data clearly shows that the correlation between financial instruments is not constant and changes stochastically over time. Modeling stochastic correlation has difficulties from the analytical as well as the estimation point of view. We think this is the main reason that very little has been done so far regarding stochastic modeling of the joint behavior of financial assets. One of the attempts to fix this gap started with a paper about Wishart processes by Bru (1991), which followed by a series of papers by Gourieroux and Sufana (2004). The Wishart process is a positive-definite symmetric random matrix process which satisfies a certain stochastic differential equation. The Wishart process is a natural extension of the Cox-Ingersoll-Ross process. Several authors have recently brought the finance community's attention to the Wishart process and showed that the Wishart process is a good candidate for modeling multi-name assets' covariance. Risk is usually measured using the covariance matrix. Therefore Wishart processes can be seen as a tool to model dynamic behavior of multivariate risk. Bru (1991) proposes the Wishart process as a generalized squared Bessel process for dimensions greater than one. She then studied existence and uniqueness characteristics, additivity properties, first hitting time of the smallest eigenvalue and the distributions of the Wishart process. Gourieroux *et al.* (2004) then use Bru (1991) and suggest the Wishart process as an extension of the CIR process. Via the Laplace transform and the distribution of the Wishart process, Gourieroux prices derivatives with a Wishart stochastic covariance matrix. This approach can be used to model risk in the structural credit risk framework but has some drawbacks in the estimation side of the problem. The Wishart process models the covariance process in a way that the marginal and joint parameters are mixed; this invalidates the method of calibrating marginal and joint parameters separately, which is the usual one employed by most of the existing estimation techniques in the literature. Gourieroux and Sufana (2004) gives a discrete version of the Wishart process known as the Wishart Autoregressive (WAR) process. The WAR process is specifically useful for simulation purposes for Wishart processes with integer degrees of freedom. Gourieroux shows that both the Wishart process and the WAR process preserve the analytical tractability property of the CIR process. In fact, these two models are both affine models which yield closed formulas for most popular financial derivatives. Gourieroux' results have been continued in a series of papers by Fonseca *et al.* (2006, 2007b,a). Fonseca *et al.* (2006) use a multi-factor Heston model to show the flexibility of the Wishart process to capture volatility smile and skew. One of the main advantages of this paper is the introduction of a correlation structure between stock's noise and the volatility's noise. Moreover Fonseca focuses on the role of stochastic correlation brought by the Wishart process. He gives closed formulas for the stochastic correlation between several factors of the model. Fonseca extends his approach in Fonseca *et al.* (2007b) to model the multivariate risk with the Wishart process. In this paper several risky assets are considered and the pricing problem for one dimensional vanilla options and multidimensional geometric basket options on the assets are solved. The results show the consistency of the Wishart stochastic covariance model with the smile and skew effects observed in the market. Even though we know the analytical properties of the Wishart process, little has been done on the calibration of the parameters of the Wishart process. Fonseca *et al.* (2007a) estimates the Wishart Stochastic Correlation Model on the stock indexes SP500, FTSE, DAX and CAC40 under the historical measure. Because of the different nature of our problem, we use a different method to calibrate the parameters of the model under the risk neutral measure.

The Dynamic Conditional Correlation Model introduced by Engle (2002) is a new class of multivariate GARCH models which is considered as one of the generalizations of the Bollerslev's Constant Conditional Correlation Model. The Dynamic Conditional Correlation Model assumes the return of the assets follows a normal distribution $r_t|\mathcal{F}_t \sim \mathcal{N}(0, H_t)$, where $H_t = D_t R_t D_t$ is the covariance matrix, $D_t = diag(\sqrt{h_t^i})$ is the diagonal matrix of standard deviations and R_t is the correlation matrix. h_t^i's follow a univariate GARCH dynamics and the dynamics of the correlation R_t are given by

$$\begin{cases} Q_t &= (1-\alpha-\beta)\bar{Q} + \alpha\epsilon_{t-1}\epsilon_{t-1}^{'} + \beta Q_{t-1}, \\ R_t &= Q_t^{*-1} Q_t Q_t^{*-1}, \end{cases} \tag{1.1}$$

with $Q_t^* = diag(\sqrt{q_{ij}})$ and $\epsilon_t \sim \mathcal{N}(0, R_t)$. The Dynamic Conditional Correlation Model is usually calibrated in two steps. In the first step, the volatilities are calibrated and in the second step, the correlation parameters are estimated. This can not be done in the case of Wishart process, because the joint and marginal information are shared among the same parameters of the model. Nelson (1990) presents a discrete time model that converges to a continuous time stochastic volatility model and uses that to estimate the parameters. Fonseca *et al.* (2007a) approximate the Wishart process by a sequence of discrete time processes satisfying a stochastic difference equation which converge in probability to the Wishart process. This approach, as pointed out by Nelson (1990) and Fonseca *et al.*

(2007a) does not guarantee consistent estimators of the continuous time model, which is a striking disadvantage for this method of calibration.

Hamida and Cont (2004) used a genetic algorithm to estimate the volatility surface of a local volatility model by option prices. Genetic or evolutionary algorithms, together with the simulated annealing variant developed by Ingber, are probabilistic search methods whose main idea is to minimize an objective function based on its performance on a search space. At the beginning of the algorithm an initial population is chosen from the search space which undergoes a cycle of mutation-crossover-selection procedures to approach the areas in the search space with population individuals that minimize the objective function. A special application of the evolutionary algorithm is to match the theoretical prices of a suggested model with their corresponding market prices, provided that either explicit formulas or fast algorithms exist for the prices of the derivatives under study.

1.2 Contributions

The main objective of this thesis is the study and development of continuous multidimensional processes applicable in credit risk analysis with emphasis on stochastic correlation. The results obtained are as follows:

1. Credit Risk Analysis - Principal Component Model

In Arian *et al.* (2008d) , we develop a new model for credit risk based on stochastic correlation. In this model, we assume that the eigenvectors of the covariance matrix are constant but the eigenvalues are driven by independent Cox-Ingersoll-Ross processes. To price equity options in this framework we first transform the calculations from the pricing domain to the frequency domain. Then we derive a closed formula for the Fourier transform of the Green's function of the pricing PDE. Finally we use the method of images to find the price of the equity options. The same method is used to find closed formulas for marginal probabilities of defaults and CDS prices. Our method is inspired by Lipton's work (Lipton (2001)) on pricing barrier options which translate to equity options when considering the default assumption. Next, we proceed with model calibration. The calibration is done via an evolutionary algorithm method which aims to minimize an error function based on a probabilistic method which searches a candidate space. Even though the evolutionary algorithm is criticized for its low speed of convergence, our calibration algorithm yields fast results based on our prior assumptions about the parameters and the conditions we make on the search space.

2. Credit Risk Analysis - Stochastic Correlation Wishart Model

In Arian *et al.* (2008b), we present a structural credit risk model which considers stochastic correlation between the assets of the companies. The covariance of the assets follows a Wishart process, which is an extension of the CIR model to dimensions greater than one. The Wishart process is an affine symmetric positive definite process. Bru (1991), Gourieroux (2005) and Fonseca *et al.* (2007a, 2006, 2007b) have brought the finance community's attention to this process as a natural extension of the Heston's stochastic volatility model, which has been a very successful univariate model for option pricing and reconstruction of volatility smiles and skews. The popularity of the Heston model could be a promising point for using Wishart process to model stochastic correlation. Here, we take a next step and use what has been discovered so far about the Wishart process in credit risk modeling. We give closed form solutions for equity options, marginal probabilities of default, and some other major financial derivatives. For calculation of our pricing formulas we build a bridge between two recent trends in pricing theory; from one side, pricing of barrier options by Lipton (2001) and Sepp (2006) and from other side the development of Wishart processes by Bru (1991), Gourieroux (2005) and Fonseca *et al.* (2007a, 2006, 2007b). In the second part, we estimate the parameters of the model by an evolutionary algorithm. We prove a theorem which guarantees the convergence of the evolutionary algorithm to the set of optimizing parameters. The parameters are found by searching a compact bounded set $E \subset \mathbb{R}^n$. An initial population is selected from the parameter space E and then the individuals go under a mutation-selection procedure to search for the best parameters which minimize the error function.

3. Ill-posed Inverse Problems

Building on the analytical results in Arian *et al.* (2008d) and Arian *et al.* (2008b), we move to the calibration problem of those models. Since the volatility and correlation are modeled as a single covariance process, the solution obtained uses evolutionary algorithms instead of the usual approach to estimate marginals and dependence structures separately; this is the content of Arian *et al.* (2008d) and Arian *et al.* (2008b), and we should mention that our approach for the calibration of the model is inspired by Hamida and Cont (2004). In Arian *et al.* (2008d,b), we have presented the evolutionary algorithm as a probabilistic tool to calibrate the model and have also obtained theoretical results regarding the convergence of the algorithm. These results are also used in Arian *et al.* (2008a) to price more complex derivatives on two stochastically correlated assets, as described below.

4. Pricing Multi-name Derivatives with Stochastic Correlation

In Arian *et al.* (2008a), we have developed a method to price multi-name derivatives on the equities of two firms based on the models developed in Arian *et al.* (2008d,b). This problem is strongly related to the reflection principle for stochastically correlated Brownian motions which is still an open problem. We were unable to solve the problem in closed form. Instead, and building on existing methods to price Bermudan options, we have developed approximation formulas for the joint probabilities of default as well as for the default correlation.

1.3 Future Research

Valuation of multi-name derivatives in the presence of stochastic correlation is still an open problem, which we plan to study with the reflection principle for two Brownian motions with stochastic correlation, which is an interesting area all in itself.

In Arian *et al.* (2008d,b) we have developed two stochastic correlation structural credit risk models. We plan to pursue the modeling of stochastic correlation in other areas of credit risk, such as intensity models or factor models.

Chapter 2

Preliminaries

Stochastic analysis lies at the heart of financial engineering. In this chapter we present the mathematical tools and their classic financial applications which are later used in this thesis. For the main part of this chapter we use Øksendal (2003) and Karatzas and Shreve (1991, 1998).

2.1 Probability Theory

The modern probability theory uses the key ideas from measure theory to define a probability space. The triple $(\Omega, \mathcal{F}, \mathcal{Q})$ is called a probability space if each of its components $\Omega, \mathcal{F}$ and $\mathcal{Q}$ have the following properties. We call Ω the sample space. A σ-algebra $\mathcal{F}$ on Ω is a family of subsets of Ω such that it contains the empty set and is closed under complements and countable unions. In other words $\mathcal{F}$ should have the following properties

- $\varnothing \in \mathcal{F}$,
- if $F \in \mathcal{F}$ then $F^C \in \mathcal{F}$,
- if $A_1, A_2, \ldots \in \mathcal{F}$ then $\bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{F}$.

The pair $(\Omega, \mathcal{F})$ is called a measurable space. A probability measure $\mathcal{Q}$ on a measurable space $(\Omega, \mathcal{F})$ is a function $\mathcal{Q} : \mathcal{F} \longrightarrow [0, 1]$ such that

- $\mathcal{Q}(\varnothing) = 0, \mathcal{Q}(\Omega) = 1$,
- If $A_1, A_2, \ldots \in \mathcal{F}$ are disjoint, then $\mathcal{Q}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathcal{Q}(A_i)$.

The triple $(\Omega, \mathcal{F}, \mathcal{Q})$ is called a complete probability space, if $\mathcal{F}$ contains all the $\mathcal{Q}-$null sets. A filtration refers to the set of information available up to a specific point in time. A filtered probability space $(\Omega, \mathcal{F}, \mathcal{Q}, \mathbb{F})$ is a probability space $(\Omega, \mathcal{F}, \mathcal{Q})$ equipped with a filtration $\mathbb{F}$. A filtration $\mathbb{F}$ is a family $(\mathcal{F}_t)_{t\geq 0}$ such that $\mathcal{F}_t \subset \mathcal{F}$ and $\mathcal{F}_s \subset \mathcal{F}_t$ for all $0 \leq s < t < \infty$ and $\mathbb{F}$ is right-continuous, i.e. $\mathcal{F}_t = \mathcal{F}_{t+}$. If $\mathcal{F}_0$ contains all subsets of the $\mathcal{Q}-$null sets of $\mathcal{F}$, we say that $(\Omega, \mathcal{F}, \mathcal{Q}, \mathbb{F})$ is a complete filtered probability space.

Financial engineering uses the theory of stochastic processes to model financial instruments such as the stock prices, volatility of the stock prices, interest rates and foreign exchange rates. A stochastic processes is an indexed set of random variables. In the continuous case the index is $\mathbb{R}_+$ and in the discrete case the index is $\mathbb{N}$. Most of the stochastic processes used are Markov. A stochastic Process X_t is called Markov if for any bounded measurable function g and any $t \geq s$

$$E\left[g(X_t)|X_u \quad \text{for } 0 \leq u \leq s\right] = E\left[g(X_t)|X_s\right].$$

A random variable is defined as a measurable function on a probability space. On the other hand, a stochastic process is defined on a filtered probability space. A stochastic process is a family of random variables $X = (X_t)_{t\geq 0} = (X(t))_{t\geq 0}$ defined on the filtered probability space $(\Omega, \mathcal{F}, \mathcal{Q}, \mathbb{F})$. Therefore for any given point in time t, we have a random variable

$$\omega \longmapsto X_t(\omega),$$

and for any $\omega \in \Omega$ we have the trajectory of X_t corresponding to ω

$$t \longmapsto X_t(\omega).$$

For any random variable X, there exists a smallest σ-algebra which contains all the pullback sets of the Borel σ-algebra. Similarly, for any stochastic process X the natural filtration $\mathbb{F}(X)$ is defined as

$$\mathcal{F}_t = \mathcal{F}(X_s : 0 \leq s \leq t), \tag{2.1}$$

with $\mathcal{F}(X_s : 0 \leq s \leq t)$ being the smallest σ-algebra which contains all sets

$$X_s^{-1}(B) = \left\{\omega \in \Omega : X_s(\omega) \in B, \text{ for } B \in \mathcal{B} \text{ and } 0 \leq s \leq t\right\}.$$

For any $\mathcal{F}$-measurable random variable X and any sub-σ-field $\mathcal{G} \subset \mathcal{F}$, the conditional expectation of X with respect to $\mathcal{G}$ is defined as a $\mathcal{G}$-measurable random variable $Y :=$

$E[X|\mathcal{G}]$ such that

$$E[XZ] = E[YZ] \quad \text{for any } \mathcal{F} - \text{measurable random variable } Z.$$

This is easy to check that if $\mathcal{H}$ is another σ-algebra such that $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$, then for any $\mathcal{F}$-measurable random variable X,

$$E[E[X|\mathcal{G}]|\mathcal{H}] = E[X|\mathcal{G}].$$

This is called the law of iterated expectations.

Brownian motion is one of the most important stochastic processes and a key element in the modern financial engineering. Let $(\Omega, \mathcal{F}, \mathcal{Q}, \mathbb{F})$ be the underlying filtered probability space. The stochastic process $W = (W_t)_{t\geq 0} = (W(t))_{t\geq 0}$ is called a $\mathcal{Q}$-Brownian motion if it has the following properties

- $W(0) = 0$,
- for all $0 \leq s_1 \leq t_1 \leq s_2 \leq t_2$, $W(t_1) - W(s_1)$ is independent of $W(t_2) - W(s_2)$,
- for any $u \geq 0$ the distribution of $W(t+u) - W(t)$ depends on u only,
- for any $u, t \geq 0$, $W(t+u) - W(t) \sim N(0, u)$.

Brownian motion is an example of a martingale. A stochastic process $X = (X(t))_{t\geq 0}$ is a martingale if it has a finite first moment and

$$E_Q(X_t|\mathcal{F}_s) = X_s \quad \mathcal{Q} - \text{a.s. for all} \quad 0 \leq s \leq t.$$

2.2 Stochastic Calculus

Stochastic calculus deals with new probabilistic definitions and methods for integration and differentiation. In this section we review the fundamental concepts of Itô calculus which has wide applications in financial engineering. First we define the stochastic Itô integral which will later equip us to define an Itô process. Suppose W_t is an m-dimensional Brownian motion and $\mu(t)$ and $\sigma(t)$ are m-dimensional measurable stochastic processes such that

$$\int_0^t |\mu(s)|\, ds < \infty \quad , \quad \int_0^t \sigma_j^2(s) ds < \infty,$$

for all $t \geq 0$ and $j = 1, \ldots, m$. Then the Itô integral of the process $\sigma(t)$ with respect to the Brownian motion W_t is defined as

$$\int_0^t \sigma(s)\, dW_s = \lim_{n\to\infty} \sum_{j=0}^{\lfloor 2^n t - 1 \rfloor} \sigma(j/2^n).[W_{(j+1)/2^n} - W_{j/2^n}],$$

where the limit is in $\mathcal{L}^2(\Omega)$. A stochastic process $X(t)$ defined as below is called an Itô process

$$\begin{aligned} X(t) &= X(0) + \int_0^t \mu(s)ds + \int_0^t \sigma(s)dW(s) \qquad (2.2) \\ &= X(0) + \int_0^t \mu(s)ds + \sum_{j=1}^{m} \int_0^t \sigma_j(s)dW_j(s). \end{aligned}$$

In differential form, we write 2.2 as

$$\begin{aligned} dX(t) &= \mu(t)dt + \sigma(t)dW(t) \\ &= \mu(t)dt + \sum_{j=1}^{m} \sigma_j(t)dW_j(t). \end{aligned}$$

Example 1. *(Geometric Brownian Motion) One of the most commonly used Itô processes is the Geometric Brownian motion. A GBM follows the stochastic differential equation*

$$dX(t) = \mu X(t)dt + \sigma X(t)dW(t),$$

where μ and σ are constants. The Geometric Brownian motion has the closed form

$$X(t) = X(0)e^{\left\{(\mu - \frac{1}{2}\sigma^2)t + \sigma dW(t)\right\}},$$

and the expectation and the variance given by

$$\begin{aligned} E(X_t) &= X_0 e^{\mu t}, \\ Var(X_t) &= X_0^2 e^{2\mu t}(e^{\sigma^2 t} - 1). \end{aligned}$$

Black and Scholes (1973) model the dynamics of the stock by a Geometric Brownian motion.

Theorem 1. *(Martingale Representation Theorem) If $\{M_t\}_{t\geq 0}$ is a martingale with respect to the filtration generated by the Brownian motion W_t, then there exists an $\mathcal{F}_t$-adapted random variable σ_t such that:*

$$M_t = M_0 + \int_0^t \sigma_s dW_s,$$

$\mathcal{Q}$-almost surely for every $t \geq 0$.

Proof. See Øksendal (2003), p. 51-52. □

Now consider two Itô processes $X_1(t)$ and $X_2(t)$

$$\begin{aligned} dX_i(t) &= \mu_i(t)dt + \sigma_i(t)dW(t) \\ &= \mu_i(t)dt + \sum_{j=1}^{m} \sigma_{ij}(t)dW_j(t). \end{aligned}$$

We define the quadratic covariance of X_1 and X_2 as

$$\langle X_1, X_2 \rangle_t = \sum_{j=1}^{m} \int_0^t \sigma_{1j}(s) \cdot \sigma_{2j}(s)ds.$$

In the special case if $X_1 = X_2 = X$, the quadratic covariance $\langle X, X \rangle_t$ is called the quadratic variation of the stochastic process X

$$\langle X, X \rangle_t = \sum_{j=1}^{m} \int_0^t \sigma_j^2(s)ds = \int_0^t \|\sigma(x)\|^2 ds.$$

In the following theorem we give the very useful Itô's formula. This is an essential tool to derive the so called Black-Scholes partial differential equation and has other important applications almost anywhere in the modern quantitative finance.

Theorem 2. *(Itô's Lemma) Let $W = (W(t))_{t\geq 0}$ be a m-dimensional Brownian motion, $m \in \mathbb{N}$, and $X = (X(t))_{t\geq 0}$ be an Itô process with*

$$dX(t) = \mu(t)dt + \sigma(t)dW(t) = \mu(t)dt + \sum_{j=1}^{m} \sigma_j(t)dW_j(t). \tag{2.3}$$

Furthermore, let $G : \mathbb{R} \times [0, \infty) \longrightarrow \mathbb{R}$ be twice continuously differentiable in the first variable, with the derivatives denoted by G_x and G_{xx}, and once continuously differentiable *in the second, with the derivative denoted by G_t. Then we have for all $t \in [0, \infty)$*

$$\begin{aligned} G(X(t),t) &= G(X(0),0) + \int_0^t G_t(X(s),s)ds \\ &+ \int_0^t G_x(X(s),s)dX(s) \\ &+ \frac{1}{2}\int_0^t G_{xx}(X(s),s)d\langle X\rangle(s), \end{aligned}$$

or

$$\begin{aligned} dG(X(t),t) &= (G_t(X(t),t) + G_x(X(t),t)\mu(t) \\ &+ \frac{1}{2}G_{xx}(X(t),t)\,\|\sigma(t)\|^2)dt \\ &+ G_x(X(t),t)\sigma(t)dW(t). \end{aligned}$$

Proof. See Karatzas and Shreve (1991), p. 150-153. □

The existence and uniqueness question is an important mathematical question which rises regarding any differential equation including SDE's. In the case of the Geometric Brownian motion, the solution exists because there can be obtained in closed form. Unfortunately, this is not always the case. For instance, the Cox-Ingersoll-Ross (CIR) Process defined by

$$dX_t = \alpha(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t. \tag{2.4}$$

for constant α, θ and σ. This stochastic differential equation has several applications in finance

- Cox, Ingersoll, Ross applied this process in interest rates modelling.
- Heston modelled stock's volatility with the CIR process.

In contrast to the Geometric Brownian motion, the existence of a CIR process is not obvious. A famous theorem by Yamada and Watanabe (Theorem 1) guarantees the existence of the solution of the stochastic differential equation defining the CIR process. Even though there is no closed form representation for the CIR process, its probability transition density and characteristic function entail closed form formulas. This is one of the most important advantages of the models using this process and gives closed form solutions for the price of a vast range of derivatives in those models. In what follows we give several classic results regarding the existence and uniqueness of the solutions of stochastic differential equations. In general $X(t)$ is a strong solution for the stochastic differential equation

$$\begin{aligned} dX(t) &= \mu(X(t), t)dt + \sigma(X(t), t)dW(t), \\ X(0) &= x, \end{aligned}$$

if for all $t \geq 0$

$$X(t) = x + \int_0^t \mu(X(s), s)ds + \int_0^t \sigma(X(s), s)dW(s) \quad Q - \text{a.s.}$$

The following theorem is the first standard result for the existence and uniqueness problem of the stochastic differential equations.

Theorem 3. *(Existence and Uniqueness) Let μ and σ of the stochastic differential equation be continuous functions such that for all $t \geq 0, x, y \in \mathbb{R}^n$ and for some constant $K > 0$ the following conditions hold:*

- *(Lipschitz condition)*

$$\|\mu(x,t) - \mu(y,t)\| + \|\sigma(x,t) - \sigma(y,t)\| \leq K \|x - y\|,$$

- *(Growth condition)*

$$\|\mu(x,t)\|^2 + \|\sigma(x,t)\|^2 \leq K^2(1 + \|x\|^2),$$

then there exists a unique, continuous strong solution $X = (X(t))_{t \leq 0}$ of the SDE and a constant C, depending only on K and $T > 0$, such that

$$E_{\mathcal{Q}}[\|X(t)\|^2] \leq C(1 + \|x\|^2)e^{Ct} \quad \textit{for all} \quad t \in [0, T]. \tag{2.5}$$

Moreover,

$$E_{\mathcal{Q}}[\sup_{0 \leq t \leq T} \|X(t)\|^2] < \infty.$$

Proof. See Øksendal (2003), p. 69-71. □

The theorem 3 is not strong enough to cover the existence of the CIR process, because the square root function is not Lipschitz. In a one-dimensional case, Yamada and Watanabe have relaxed the Lipschitz condition for the diffusion coefficient

Proposition 1. *(Yamada and Watanabe) Let us suppose that the coefficients of the one-dimensional equation* $(n = m = 1)$

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \tag{2.6}$$

satisfy the conditions

$$\begin{aligned} \|\mu(x,t) - \mu(y,t)\| &\leq K\|x-y\|, \\ \|\sigma(x,t) - \sigma(y,t)\| &\leq h\|x-y\|, \end{aligned}$$

for every $0 \leq t < \infty$ *and* $x \in \mathbb{R}$, $y \in \mathbb{R}$, *where* K *is a positive constant and* $h : [0, \infty) \longmapsto [0, \infty)$ *is a strictly increasing function with* $h(0) = 0$ *and*

$$\int_{0,\epsilon} h^{-2}(u)du = \infty; \quad \forall \epsilon > 0. \tag{2.7}$$

Then the strong uniqueness holds for the equation (2.6).

Proof. See Karatzas and Shreve (1991), p. 291-292. □

Another example of a stochastic differential equation is the Ornstein-Uhlenbeck Process

$$dX_t = -\alpha X_t dt + \sigma dW_t.$$

The expectation and variance of the Ornstein-Uhlenbeck Process is given by

$$\begin{aligned} m(t) &= E(X_t) = m(0)e^{-\alpha t}, \\ V(t) &= Var(X_t) = \frac{\sigma^2}{2\alpha} + (V(0) - \frac{\sigma^2}{2\alpha})e^{-2\alpha t}. \end{aligned}$$

An interesting observation is the fact that the CIR process is the square of the Ornstein-Uhlenbeck process. This simple fact will be used later to extend the Heston's model to dimensions greater than one. Another extended form of the Ornstein-Uhlenbeck process is

$$dX_t = \alpha(\phi - X_t)dt + \sigma dW_t. \tag{2.8}$$

Girsanov's theorem is a very interesting mathematical result which at the first glance seems very unlikely to give an applicable result in an empirical field like finance. Later we will see that this theorem will give us an insight into a new measure which exists in the market and is related to the valuation of options. First we explain the concept of the Radon-Nikodyn derivative which is essential in the structure of the Girsanov's theorem. If $\mathcal{P}$ and $\mathcal{Q}$ are two probability measures with respect to a σ-algebra $\mathcal{F}$, we say P is absolutely continuous with respect to $\mathcal{Q}$ if there exists a measurable function f such that

- $I|f| < \infty$,
- $\int g(x)d\mathcal{P}(x) = \int g(x)f(x)d\mathcal{Q}(x)$ for any integrable function g .

The function f is called the Radon-Nikodyn Derivative of $\mathcal{P}$ with respect to $\mathcal{Q}$ and is denoted by $\frac{d\mathcal{P}}{d\mathcal{Q}}$. In addition to the Radon-Nykodyn derivative, the following lemma is essential in the statement of the Girsanov's theorem

Lemma 1. *(Novikov Condition) Let $\gamma = (\gamma(t))_{t>0}$ be a m-dimensional progressively measurable stochastic process, $m \in \mathbb{N}$, with*

$$\int_0^t \gamma_j^2(s)ds < \infty \;\; Q - a.s. \;\; \textit{for all}\; t \geq 0, j = 1, \ldots, m,$$

and let the stochastic process $L(\gamma) = (L(\gamma,t))_{t\geq 0} = (L(\gamma(t),t))_{t\geq 0}$ for all $t \geq 0$ be defined by

$$L(\gamma, t) = e^{-\int_0^t \gamma(s)' dW(s) - \frac{1}{2}\int_0^t \|\gamma(s)\|^2 ds}.$$

Then $L(\gamma)$ is a continuous $\mathcal{Q}-$martingale if

$$E_Q[e^{\frac{1}{2}\int_0^T \|\gamma(s)\|^2 ds}] < \infty.$$

Proof. See Karatzas and Shreve (1991), p. 198-199. □

Now with the Radon-Nikodym notation, one can define the measure $\tilde{\mathcal{Q}} = \mathcal{Q}_{L(\gamma,T)}$ by $d\tilde{\mathcal{Q}} = L(\gamma, T).d\mathcal{Q}$, or in the integral form

$$\tilde{\mathcal{Q}}(A) = \int_A L(\gamma, T)d\mathcal{Q} \quad \text{for all} \quad A \in \mathcal{F}_T,$$

which is a probability measure because $L(\gamma, T)$ is a $\mathcal{Q}$-martingale. In the following we provide the Girsanov theorem, which shows, how a $\tilde{\mathcal{Q}}-$Brownian motion $\tilde{W} = (\tilde{W}(t))_{t\in[0,T]}$ starting with a $\mathcal{Q}-$Brownian motion $W = (W(t))_{t\geq 0}$ can be constructed.

Theorem 4. *(Girsanov Theorem - first version)*

Let $(\Omega, F, \mathcal{P})$ be a probability space and F be the standard filtration $\{F_t : 0<t<T\}$ of a Brownian motion B. Let X be an Ito process in R^N,

$$X_t = x + \int_0^t \mu_s ds + \int_0^t \sigma_s dB_s \quad 0 \leq t \leq T.$$

Suppose $v = (v^1, \ldots, v^N)$ is a vector of processes in L^1 such that there exists some θ, satisfying

$$E\left[\exp\left(\frac{1}{2}\int_0^T \theta_s \cdot \theta_s ds\right)\right] < \infty,$$

with

$$\sigma_t \theta_t = \mu_t - \nu_t \quad 0 \leq t \leq T.$$

Then there exists a probability measure $\mathcal{Q}$ equivalent to $\mathcal{P}$ such that

$$\overline{B_t} = B_t + \int_0^t \theta_s ds \quad 0 \leq t \leq T,$$

defines a standard Brownian motion in R^d on $(\Omega, \mathcal{F}, \mathcal{Q})$ that has the same standard filtration F. The process X is also an Ito process with respect to $(\Omega, \mathcal{F}, \mathcal{Q})$ and

$$X_t = x + \int_0^t v_s ds + \int_0^t \sigma_s d\overline{B}_s \quad 0 \leq t \leq T.$$

Proof. See Øksendal (2003), p. 163-164. □

Theorem 5. *(Girsanov Theorem - second version) Let θ be an n-dimensional predictable process $(\theta_1(t), \ldots, \theta_n(t))$ and $\phi(t)$ a predictable process with:*

$$\int_0^t ||\theta(u)||^2 du < \infty \qquad \int_0^t |\phi(u)|\lambda(u) du < \infty,$$

for finite t. Define the process L by $L(0) = 1$ and

$$\frac{dL(t)}{L(t-)} = \sum_{i=1}^{n} \theta_i(t) dW^i(t) - (\phi(t) - 1) dM(t).$$

Assume $E[L] = 1$ for finite L. Then there is a probability measure $\mathcal{Q}$ equivalent to $\mathcal{P}$ with $d\mathcal{Q} = Ld\mathcal{P}$ such that:

$$dW(t) - \theta(t)dt = dW^*(t),$$

defines W^* as a $\mathcal{Q}$-Brownian motion.

Proof. See Øksendal (2003), p. 163-164. □

2.3 Ordinary and Partial Differential Equations

Affine models are the most powerful and efficient models in the finance literature. These models give closed form formulas for pricing problem of a vast range of derivatives in these models. In the process of finding the final price, as we will see later in the third and fourth chapter, we need to solve a partial differential equation with a certain boundary condition dependent on the specific derivative under study. We will then convert the PDE problem to the problem of solving a system of ordinary differential equations. Therefore, in this section, we will take a closer look at the basic theorems regarding the existence and uniqueness of the solutions of the ordinary and partial differential equations and their relations to Affine models. The first theorem below gives the standard existene conditions for a system of first order ODE's

Theorem 6. *(First Order ODE's)*

Let

$$\begin{aligned} x_1' &= F_1(t, x_1, ..., x_n), \\ &\cdots \\ x_n' &= F_n(t, x_1, ..., x_n), \end{aligned} \tag{2.9}$$

be a system of ODE. Let each of the functions F_i and the partial derivatives

$$\frac{\partial F_1}{\partial x_1}, ..., \frac{\partial F_1}{\partial x_n}, ... \frac{\partial F_n}{\partial x_n},$$

be continuous in a region R of $(t, x_1, \ldots, x_n)$ space defined by

$$\alpha < t < \beta, \alpha_1 < x_1 < \beta_1, ..., \alpha_n < x_n < \beta_n,$$

and let the point

$$(t^0, x_1^0, ..., x_n^0),$$

be in R. Then there is an interval $|t - t_0| < h$ in which there exists a unique solution of the system of ODE that also satisfies the initial conditions.

Proof. see Boyce and DiPrima (2001). □

Later on in this thesis we will encounter a linear case of the above theorem. As a special case, If F_i are linear on x_j,

$$\begin{aligned} x_1^{'} &= p_{11}(t) \cdot x_1 + ... + p_{1n}(t) \cdot x_n + g_1(t), \\ &\quad ... \\ x_n^{'} &= p_{n1}(t) \cdot x_1 + ... + p_{nn}(t) \cdot x_n + g_n(t), \end{aligned} \tag{2.10}$$

and p_{ij}, g_i are continuous on an open interval I, $\alpha < t < \beta$. Then there exists a unique solution of the system that also satisfies the initial conditions, the solution exists throughout the interval I. In particular if we have the asymmetric matrix Riccati differential equation

$$X' = A \cdot X + X \cdot B + X \cdot P \cdot X + Q,$$

where A, B, P, Q are continuous matrix-valued functions with real entries on $[t_0, t_1]$. This is a particular case of a non-linear system of ODE, therefore we know that as long as A, B, P, Q are differentiable there is a unique solution on an interior interval.

Theorem 7. *(First Order PDE's)*

Let

$$\begin{aligned} a(x, u) \cdot \nabla u(x) + b(x, u) &= 0, \\ u(x_1, \ldots, x_{n-1}, 0) &= \phi(x_1, \ldots, x_{n-1}). \end{aligned}$$

We make the following assumptions:

- $a(x, u) = (a_1(x, u), \ldots, a_n(x, u))$ *where a_i is a differentiable function on R^{n+1}.*

- *$b(x,u)$ is differentiable on R^{n+1}.*
- *ϕ is differentiable on R^{n-1}.*
- *$a_n(x_1,\dots,x_{n-1},0,\phi) \neq 0$ for every $(x_1,\dots,x_{n-1})$ in R^{n-1}.*

Then the PDE has a unique solution u defined on a neighborhood of $R^{n-1}\times 0$.

Proof. see Boyce and DiPrima (2001). □

In our derivative valuation, we will use special types of ordinary differential equations called Riccati equations. Riccati equations arise commonly in the analysis of affine models. The one dimensional Riccati equation follows the dynamics

$$\frac{da(h)}{dh} = b[a(h)-c_o][a(h)-c_1].$$

To solve this equation, we first rewrite it as

$$da(h)\left[\frac{1}{a(h)-c_o} - \frac{1}{a(h)-c_1}\right] = b(c_o-c_1)dh.$$

After integration

$$\frac{a(h)-c_o}{a(h)-c_1} = \frac{a(0)-c_o}{a(0)-c_1}\exp[b(c_o-c_1)h],$$

the solution can be written as

$$a(h) = c_1 + \frac{[a(0)-c_1](c_o-c_1)}{a(0)-c_1-[a(0)-c_o]\exp[b(c_o-c_1)h]}.$$

Solving the multidimensional Riccati equation is not as straightforward as its univariate case. In fact, the general multidimensional Riccati differential equation does not have a closed form solution. The multidimensional Riccati equation we deal with in the framework of the Wishart process satisfies the matrix Riccati equation

$$\frac{dX(h)}{dh} = A'X(h) + X(h)A + 2X(h)\Lambda X(h) + C_1, \tag{2.11}$$

with initial condition $X(0)=C_0$. The coefficients Λ, C_1, C_0 are symmetric matrices, Λ is positive definite, and A is a square matrix. To solve this equation we use the following two lemmas

Lemma 2.3.1. *If X^* solves the following system*

$$A'X^* = X^*A + 2X^*\Lambda X^* + C_1 = 0.$$

Then $Z(h) = X(h) - X^*$ satisfies

$$\frac{dZ(h)}{dh} = A^{*'}Z(h) + Z(h)A^* + 2Z(h)\Lambda Z(h),$$

with initial condition $Z(0) = C_0^*$, where $A^* = A + 2\Lambda X^*$ and $C_0^* = C_0 - X^*$.

Proof. In 2.11 substitute $X(h)$ by $Z(h) + X^*$. □

Lemma 2.3.2. *If $Z(h)$ satisfies*

$$\frac{dZ(h)}{dh} = A^{*'}Z(h) + Z(h)A^* + 2Z(h)\Lambda Z(h),$$

with $Z(0) = C_0^*$ then

$$Z(h) = \exp(A^*h)' \left[C_0^{*-1} + 2\int_0^h \exp(A^*u)\Lambda \exp(A^*u)' du \right]^{-1} \exp(A^*h).$$

Proof. Define $\Lambda(h)$ by

$$Z(h) = \exp(A^*h)'\Lambda(h)\exp(A^*h).$$

Differentiating from both sides of the above equation

$$\begin{aligned}
\frac{dZ(h)}{dh} &= A^{*'}\exp(A^*h)'\Lambda(h)\exp(A^*h) \\
&\quad + \exp(A^*h)'\Lambda(h)\exp(A^*h) + \exp(A^*h)'\frac{d\Lambda(h)}{dh}\exp(A^*h) \\
&= A^{*'}Z(h) + Z(h)A^* + \exp(A^*h)'\frac{d\Lambda(h)}{dh}\exp(A^*h),
\end{aligned}$$

implies

$$\exp(A^*h)'\frac{d\Lambda(h)}{dh}\exp(A^*h) = 2\exp(A^*h)'\Lambda(h)\exp(A^*h)\Lambda\exp(A^*h)'\Lambda(h)\exp(A^*h),$$

and the result follows.

□

By lemmas 2.3.1 and 2.3.2 the solution of Riccati equation 2.11 is given by

$$
\begin{aligned}
X(h) = X^* + [&\exp(A + 2\Lambda X^*)h]' \\
&\left\{ (C_0 - X^*)^{-1} + 2\int_0^h [\exp(A + 2\Lambda X^*)u]\Lambda[\exp(A + 2\Lambda X^*)u]' du \right\}^{-1} \\
&[\exp(A + 2\Lambda X^*)h],
\end{aligned}
$$

where X^* satisfies

$$A'X^* + X^*A + 2X^*\Lambda X^* + C_1 = 0.$$

2.4 Derivative Pricing

One of the most important results of the modern financial engineering is to give formulas for the price of a contract with uncertain payoffs in future. These contract are known as contingent claims. In a more specific language, if the underlying process is a vector $S = (S_1, S_2, \ldots, S_n)$ of assets with generated information $\mathcal{F}_t$ up to time t, a contingent claim at maturity T is any $\mathcal{F}_T$-measurable random variable. In the most simple case when there is a function Φ such that the payoff is $\Phi(S_T)$, we call the contingent claim, simple. In the following we introduce some of the most popular contingent claims and describe their properties in brief.

1. Forward Contract: A forward contract is an agreement to buy or cell an underlying asset at a specific maturity date for a specific price. There is no price to enter a forward contract and the payoff in the case of a call contract is $S_T - F_{0,T}$, where S is the spot price and $F_{0,T}$ is the forward price seen at the beginning of the contract. By the risk neutral valuation, the value at t of a forward contract started at 0 is

$$E_t^Q[\exp^{-\int_t^T r(s)ds} \cdot (S_T - F_{0,T})].$$

The forward contracts are traded over the counter.

2. Forward Rate Agreement : In a forward rate agreement, two parties agree that a certain interest rate apply to a specified principal during a period of time in future.

3. Futures Contract: A futures contract is an agreement to purchase or sell a given asset at a future date at a preset time. The main difference between a futures contract and a forward contract is in that for a futures contract, the profits and losses are settled on a

daily cash flow basis and therefore the credit risk is not significant. The payoff at maturity is given by $S_T - F_{0,T}$, where $F_{0,T}$ is the futures price seen at the beginning of the contract and therefore the price of the futures contract started at 0 is given $E_t^Q[(S_T - F_{0,T})]$.

4. Swap : Swaps are financial instruments which are used to exchange cash flows at fixed points in future. To enter into a swap no cash is needed. One type of a swap is the commodity swap in which an exchange of a fixed cash flow with a floating cash flow occurs. The floating cash flow depends on the price of a certain commodity.

5. Option : An option is a contract to sell or buy a tradable asset prior to at a specific maturity date. Options are used to obtain protection against movements of assets by creating ceilings or floors for prices of the assets. The buyer of the option should pay the price of the option in advance to be given the permission to exercise the option prior or at the maturity date. The buyer is exposed to the default risk of the the seller of the contract. There are different options traded in the market:

- European Option: European puts (or calls) gives the holder the right to buy (or sell) an asset for a specific strike price K at a particular maturity date T. For a European call the payoff at maturity is $\max(S_T - K, 0)$ and for a European put the payoff at maturity is $\max(K - S_T, 0)$. European options could also be written on future contracts. In this case if F_{T,T_1} is the future price at T for maturity T_1, the payoff at maturity for a call option is $\max(F_{T,T_1} - K, 0)$.
- American Option: Give the holder the right to exercise the option at any time prior to the maturity date.
- Asian Option: The payoff depends on the historical average of the underlying asset.
- Barrier Option: The payoff depends on the event that the assets price process hits a certain barrier prior to maturity.
- Lookback Options: The payoff depends on the max or min of the underlying asset over a specific period of time.
- Compound Option: Is an option on another option.
- Rainbow Option: Is an option on the max or min of several assets.
- Exchange Option: Gives the holder the right to exchange the underlying asset for another tradable asset. If the first asset is S_1 and the second asset is S_2, the payoff at maturity is $\max(S_1(T) - S_2(T), 0)$.
- Basket Option: is written on a portfolio.

- Shout Option: The holder can at any time reset the strike price to the current level of the asset's price.
- Volatility Option: is written on the realized historical volatility.
- Conditional Option: Gives the holder the right to buy or sell an asset with a certain strike price conditioned on the occurrence of a certain event for another tradable or non-tradable asset. In this content, there are Conditional European, Conditional American, Conditional Asian, Conditional Barrier and etc. The only difference with the above mentioned cases is that the payoff depends on another variable instead of the underlying asset itself. The new variable may not be a tradable asset and is considered to be correlated to the underlying asset.

Now we explain the basic financial concepts and applications implemented by the theorems in section 2.2. For any tradable asset with the dynamics

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dB_t,$$

the martingale measure $\mathcal{Q}$ is the measure generated by the Girsanov's theorem which gives the process X_t a new dynamics

$$dX_t = r(t)dt + \sigma(X_t, t)dB_t,$$

where $r(t)$ is the risk free interest rate. Under this measure, the discounted price process $\frac{X(t)}{B(t)}$ is a martingale. An arbitrage opportunity is an adapted and square integrable portfolio process Π_t such that produces guaranteed profit. In other words, if the value of the portfolio is denoted by $V(\Pi_t)$, then there is a $T > 0$ such that $V_0(\Pi) = 0$ and $P[V_T(\Pi) > 0] > 0$. It can be proved that there are no arbitrate opportunities if and only if there is a martingale measure $\mathcal{Q}$ equivalent to the real world measure $\mathcal{P}$. The following theorem is the formal statement of the Risk-Neutral Valuation Formula

Theorem 8. *(Risk-Neutral Valuation Formula) Assume that the price processes for the underlying assets is modeled by geometric Brownian motions, the market is complete and X is a given contingent claim. Let B_0 be a risk free asset with a price process given by $dB_0(t) = B_0(t)r(t)dt,\ B_0(0) = 1$, where $r(t)$ is the risk free instantaneous interest rate. The arbitrage price process of X is given by the risk-neutral valuation formula*

$$V_x(t) = B_0(t)\mathbb{E}_{\mathcal{Q}}[\frac{X}{B(T)}|\mathcal{F}_t] = \mathbb{E}_{\mathcal{Q}}[Xe^{-\int_t^T r(u)du}|\mathcal{F}_t],$$

where $\mathcal{Q}$ is the unique equivalent martingale measure.

Proof. See Øksendal (2003) □

Assuming the interest rate to be constant, the price of the contingent claim is given by

$$V_x(t) = e^{-r(T-t)}\mathbb{E}_{\tilde{Q}}[X|\mathcal{F}_t].$$

If the contingent claim X is simple, i.e. $X = \Phi(S(T))$ with a sufficiently smooth function Φ, then the price process is given by $V_x(t) = G(S(t), t)$, where G solves the Black-Scholes partial differential equation

$$\begin{aligned} G_t(s,t) + rsG_s(s,t) + \frac{1}{2}\sigma^2 s^2 G_{ss}(s,t) - rG(s,t) &= 0, \\ G(s,T) &= \Phi(s). \end{aligned}$$

The PDE can be obtained by using the Feynman-Kac formula which links stochastic differential equations to partial differential equations

Theorem 9. *(Feynman-Kac Formula) For given $T > 0$ suppose $f(x,t)$ has continuous second and first derivative respect to x and t, respectively, and satisfies the partial differential equation*

$$\begin{gathered} f_x(x,t)\mu(x,t) + f_t(x,t) + \frac{1}{2}tr[\sigma(x,t)\sigma(x,t)^T f_{xx}(x,t)] - r(x,t)f(x,t) + h(x,t) = 0, \\ (x,t) \in R^N \times [0,T), \end{gathered}$$

with the boundary condition $f(x,T) = g(x)$. Then $f(x,t)$ can be represented by

$$\begin{aligned} f(x,t) &= E_t\left[\int_t^T \left[\left(\exp\{-\int_t^s r(X_u,u)du\}\right) h(X_s,s)\right] ds\right] \\ &+ \left[\left(\exp\{-\int_t^T r(X_u,u)du\}\right) g(X_T)\right], \end{aligned}$$

where X is assumed to solve the stochastic differential equation

$$\begin{aligned} dX_s &= \mu(X_s,s)ds + \sigma(X_s,s)dB_s, \\ X(t) &= x. \end{aligned}$$

Proof. See Øksendal (2003), p. 143-144. □

Consider n tradable or non-tradable assets

$$dX_{t,i} = m_i(X_t, t)dt + s_i(X_t, t)dB_{t,i},$$

with securities on X_j with prices $F_{t,j}, j = 1, ..., n$ following the dynamics

$$\frac{dF_{t,j}}{F_{t,j}} = \mu_j(X_t, t)dt + \sum_{i=1}^{n} \sigma_{i,j}(X_t, t)dB_{t,i}.$$

Then $\mu - r = \sum\limits_{i=1}^{n} \lambda_i \sigma_i$, where λ_i is the market price of risk for the asset X_i. If the asset X_i is tradable then $\dfrac{m_i - r}{s_i} = \lambda_i$ and in the risk neutral world X_i follows the dynamics

$$dX_{t,i} = \left(m_i(X_t, t) - \lambda_i \cdot s_i(X_t, t)\right) dt + s_i(X_t, t)dB^{Q}_{t,i}.$$

Market price of risk is a key element in the definition of the risk-neutral world and other similar imaginary environments. If in the real world, a security F has the observable dynamics under the $\mathcal{P}$-measure

$$\frac{dF_t}{F_t} = \mu dt + \sigma dB_t,$$

then in the risk neutral world, it follows the imaginary dynamics under the $\mathcal{Q}$-measure

$$\frac{dF_t}{F_t} = (\mu - \lambda\sigma)dt + \sigma dB^{Q}_t,$$

where as before λ is the market price of risk related to the interest rate by $\mu - r = \lambda\sigma$. By taking other values of λ we can define other imaginary worlds like the risk neutral world. The risk neutral world could also be defined based on a numeraire. Consider the bank account $dN = rNdt$ as the numeraire. The process $\frac{F}{N}$ satisfies the real world dynamics

$$d\left(\frac{F}{N}\right) = (\mu - r)\left(\frac{F}{N}\right) dt + \sigma \left(\frac{F}{N}\right) dB_t,$$

and therefore by the Ito's lemma, it satisfies the following risk neutral equation

$$d\left(\frac{F}{N}\right) = \sigma \left(\frac{F}{N}\right) dB^{Q}_t.$$

Therefore choosing a risk free asset for the numeraire leads to the well known risk neutral world. Alternatively, if the process N_t satisfies the dynamics

$$\frac{dN}{N} = rdt + \sigma_N dB_t^Q,$$

then pricing the derivatives based on this numeraire leads to a new financial world. Under the previous $\mathcal{Q}$-measure, the process $\frac{F}{N}$ satisfies

$$d\left(\frac{F}{N}\right) = \sigma_N^2 \left(\frac{F}{G}\right) dt + (\sigma + \sigma_N)\left(\frac{F}{N}\right) dB_t^Q,$$

which is not a martingale. In order to price derivative in a no arbitrage content, the process $\frac{F}{N}$ should be a martingale. The new world which has this feature is called the forward risk neutral world with respect to N, and its correspondent measure is called the FQN-measure. Under the FQN-measure, $\frac{F}{N}$ satisfies

$$d\left(\frac{F}{N}\right) = (\sigma + \sigma_N)\left(\frac{F}{N}\right) dB_t^{FQN}.$$

The price of the contingent claims can be derived with respect to the risk neutral measure or the FQN-measure. If a contingent claim matures at time T and has the payoff $\Phi(T)$ at maturity, its risk neutral price is defined as

$$Y(t) = E_t^Q\left[\frac{Y(T)}{e^{\int_t^T r(s)ds}}|t\right],$$

and its forward risk neutral price with respect to the numeraire N is defined as

$$E_t^{FQN}\left[\frac{\Phi(T)}{N(T)}|t\right].$$

Now consider two numeraires N_1 and N_2. In a forward risk neutral world with respect to N_i, a traded security F follows

$$df = \left(r + \sigma_F \cdot \sigma_{N_i}\right) Fdt + \sigma_F f dB_t^{FQN_i}.$$

Therefore when moving from the first forward risk neutral world to the second one, the growth rate of the price of any traded security F increases by $\sigma_F \cdot \left(\sigma_{N_1} - \sigma_{N_2}\right)$.

The sensitivities of the contingent claims with respect to the current stock price, the volatility, the time to maturity and the interest rate are called the Greeks and are used for hedging purposes. If V is the value of the contingent claim, S the value of the underlying, T the time to maturity, and r the risk-free interest rate, the Greeks are defined as $\Delta = \frac{\partial V}{\partial S}, \nu = \frac{\partial V}{\partial \sigma}, \Theta = \frac{\partial V}{\partial T}, \rho = \frac{\partial V}{\partial r}, \Gamma = \frac{\partial^2 V}{\partial S^2},$.

Chapter 3

Principal Component Analysis in Credit Risk Modeling

We develop a new model for credit risk based on stochastic correlation. To induce the stochasticity into the structure of the correlation, we assume that the eigenvectors of the covariance matrix are constant but the eigenvalues are driven by independent Cox-Ingersoll-Ross processes. To price equity options on this framework we first transform the calculations from the pricing domain to the frequency domain. Then we derive a closed formula for the Fourier transform of the Green's function of the pricing PDE. Finally we use the method of images to find the price of the equity options. Same method is used to find closed formulas for marginal probabilities of defaults and CDS prices. Our method is inspired by Lipton's work (Lipton (2001)) on pricing the barrier options which translates to equity options when considering the default assumption. Next, we proceed with the calibration of the model. The calibration is done via an evolutionary algorithm method which aims to minimize an error function based on a probabilistic method. Even though the evolutionary algorithm is criticized for its low speed of convergence, our calibration algorithm yields fast results based on our prior assumptions about the parameters and the conditions we make on the search space.

3.1 Introduction

In 1974, Merton presented a model which assumed equity as an option on the asset and the debt as the strike price of the option. In the event that the asset lies below the debt at maturity, default occurs. Merton's model has several drawbacks. First, the default can only happen at maturity. Black and Cox (1976) were the first to give an extension in which the default could happen prior to maturity time. To the best of our knowledge Hull *et al.* (2004) is the first to calibrate the parameters of the Merton's model with equity options by considering equity options as compound options on the firm's asset. They use results from Geske (1979) to identify equity options as compound options on the asset of the company. Their work is very interesting, specially because most of the dynamic information about a company can be obtained from the options on the firm's equity. This is a very strong advantage when it comes to calibrating the model's parameters. When volatility is stochastic Geske's results are not analytically tractable anymore. Since Hull *et al.* (2004) assumes a deterministic volatility in their model, we think their model fails to give a closed form formula for equity options in the presence of stochastic volatility-correlation. One of the popular versions of the Merton's structural credit risk model is the CreditGrades model. The CreditGrades model was jointly developed by CreditMetrics, JP Morgan, Goldman Sachs and Deutsche Bank. The original version of the CreditGrades model assumes that volatility is deterministic. We extend the CreditGrades model, by means of stochastic covariance principal component model focusing on the role of stochastic correlation.

The Black and Scholes's model is the most well-known approach for modeling financial assets. In this model, asset prices are driven by Gaussian random processes with independent increments and variances proportional to time. To study the behavior of a certain asset in this model, knowing the mean and the standard variation of the underlying Gaussian variable is enough. The first advantage of the Black and Scholes model is its tractability which provides closed form solutions for prices of derivatives on the underlying asset. Moreover, the model has a very small parameter set and therefore the calibration of the model is less involved. On the other hand, the Black and Scholes's model is not flexible enough to capture certain features of the market. Two famous examples are volatility smiles and skews. As a result of the demand from market, other models have been presented to improve Black and Scholes (1973), like Hull and White (1987), Stein and Stein (1991) and Heston (1993). Among these models, Heston's model is the most well-known mathematical solution for the inefficiencies of the Black-Scholes Model, which uses CIR processes to monitor the dynamics of the volatilities of the assets. Heston's model is analytically tractable but its calibration is more difficult, firstly because it has more parameters than the constant volatility model, and secondly because some of the parameters related to the non-observable volatility process are difficult to estimate. Despite difficulties in calibrating the Heston's model, it is able to reproduce volatility smile and skew.

Almost all existing models in the finance literature assume the correlation between assets, equities, interest rates and etc. are deterministic. But market observations show that cor-

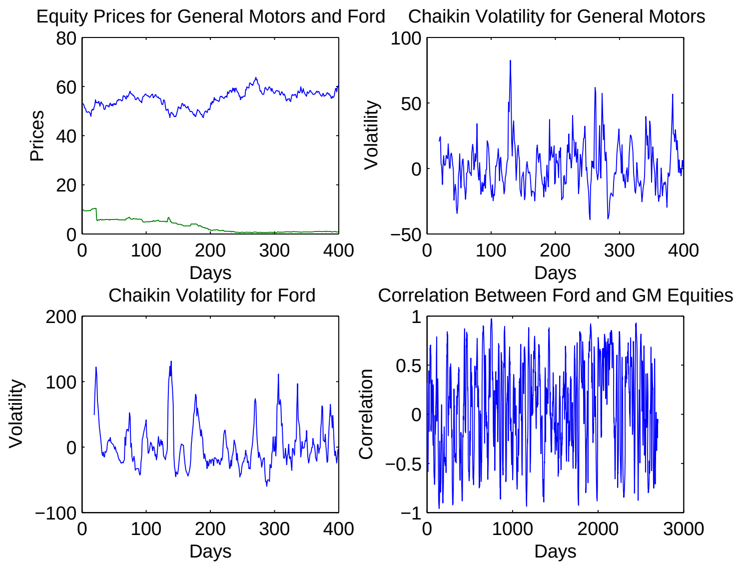

Figure 3.1: Historical Volatility and Correlation of General Motors and Ford

relations change over time and are stochastic. One of the main difficulties for modeling correlation is to construct a process which remains in the interval $[-1, 1]$, has nice analytical properties and is easy to implement. Consider two Brownian motions $B_1(t)$ and $B_2(t)$ with correlation ρ. We write this in a mathematical form as $dB_1(t).dB_2(t) = \rho dt$. Figure 3.1 shows the historical volatility of GM and Ford's equity prices as well as their correlations. It can be seen from the figure that not only correlation is not constant, but it also has stochastic behavior. Interestingly enough, it has even more fluctuations and stochasticity than the volatility itself. This is just one of the observations from the market which arises the need to deal with correlation risk. Based on financial data, we expect correlation to have some standard properties, and this will direct our intuition to pick our model. First, the process should remain in the interval $[-1, 1]$. Moreover, we expect correlation to have a mean reversion behavior. In addition to the above properties, we expect the correlation to have negligible probabilities in its boundaries. Before we present our model, we briefly go over the present stochastic correlation models presented so far in the finance literature.

- Inspired by the standard stochastic volatility models starting from Heston's paper (Heston (1993)), Emmerich (2006) proposes the following dynamics for modeling correlation

$$d\rho = \kappa(\theta - \rho)dt + \alpha\sqrt{1 - \rho^2}dW_t.$$

Then one can separately consider volatility as a stochastic process as in Heston

(1993) and study a stochastic volatility-correlation model. There are two problems with this idea: First, the model violates analytical tractability since it's not affine and second, it is not easily extendable to dimensions higher than two.

- In Escobar *et al.* (2007, 2006) and Arian *et al.* (2008d), the applications of the principal component model in credit risk is studied. The main idea is to identify the covariance matrix by its eigenvectors and eigenvalues. In the above papers, the authors have assumed that the eigenvectors of the covariance matrix are constant but the eigenvalues follow a CIR process. This implies a stochastic structure for the correlation between the assets. Escobar *et al.* (2007, 2006) price the collateralized debt obligations under Merton's model using a tree approach and the principal component model. Arian *et al.* (2008d) extends the CreditGrades model using the method of images and the principal component model.

- On the other hand, there has been a series of works by Bru (1991), Gourieroux (2005) and Fonseca *et al.* (2007a, 2006, 2007b) which consider the stochasticity of volatility and correlation into one stochastic covariance matrix process. Wishart Covariance Models are affine and bear closed form formulas for some popular financial derivatives. But since they model volatility and correlation as one covariance process, it is more challenging to calibrate the marginal and joint parameters, because they are mixed in the model. We will return to this point in the section 3.5.2 where we deal with implementation of our model.

Hamida and Cont (2004) have used a genetic algorithm to estimate the volatility surface of a local volatility model by option prices. Genetic or evolutionary algorithms are probabilistic search methods which are quite popular in engineering and physics. Lester Ingber developed a variation of the genetic algorithm called adapted simulated annealing. The main idea in the evolutionary algorithm is to minimize an objective function based on its performance on a search space. At the beginning of the algorithm an initial population is chosen from the search space which undergoes a cycle of mutation-crossover-selection procedure to approach the areas in the search space with population individuals that minimize the objective function. A special application of the evolutionary algorithm is to match the theoretical prices of a suggested model with their corresponding market prices, provided that there is a formula for the price of the derivatives under study. Based on our knowledge Hamida and Cont (2004) are the first researchers who have used the evolutionary algorithm to estimate a financial model.

This chapter is organized as follows: In section 3.2 we introduce the Cox-Ingersoll-Ross process and some of its properties. This process will be used to model the eigenvalue of the covariance matrix. We explain our extension of the CreditGrades model in section 3.3. In section 3.4 we give closed formulas for prices of equity options and probabilities of default. We end this chapter by suggesting a probabilistic method to estimate the model's parameters based on the evolutionary algorithm in section 3.5.

3.2 The Cox-Ingersoll-Ross Process

The principal component model we suggest uses the Cox-Ingersoll-Ross (CIR) process as the source of stochasticity in the covariance matrix driving the assets' prices. The CIR process is a positive one dimensional stochastic process which can represent stochastic volatility, interest rate and some other financial concepts. In this section, we first introduce the CIR process by its dynamics, study its distribution and conditional Laplace transformation and connect the CIR process with the Ornstein-Uhlenbeck process.

By its definition, the Cox−Ingersoll−Ross (CIR) process λ_t satisfies the stochastic differential equation

$$\begin{aligned} d\lambda(t) &= \kappa(\theta - \lambda(t))dt + \sigma\sqrt{\lambda(t)}dZ_t, \\ \lambda(0) &= \lambda_0, \end{aligned} \tag{3.1}$$

where Z_t is a standard Brownian motion and the parameters κ, θ, σ and λ_0 are positive real numbers. Because of the term $(\theta - \lambda(t))$ and the positivity of κ, the process $\lambda(t)$ has mean reverting properties. The existence and uniqueness of the solutions of the stochastic differential equation above comes from a theorem by Yamada and Watanabe (Theorem 1). The intuitive reasoning for the positivity of the CIR process comes from the observation that when λ_t hits the boundary 0, the diffusion term vanishes and the drift term is equal to $\kappa\theta dt$ which is positive and rejects the process toward the positive direction of the real axis. There is also an alternative way to prove the existence of the CIR processes of certain type based on the fact that the sum of squares of N independent Ornstein-Uhlenbeck processes is a CIR process. To see this, assume that we have N independent Ornstein-Uhlenbeck processes

$$d\omega_t^{(i)} = a\omega_t^{(i)}dt + \sigma_\omega dW_t^{(i)} \text{ for } i = 1, 2, \ldots, N.$$

Then Ito's lemma applied to

$$\lambda(t) = \sum_{i=1}^{N} (\omega_t^{(i)})^2,$$

shows that $\lambda(t)$ is a CIR process with parameters $\kappa = -2a, \sigma = 2\sigma_\omega$ and $\theta = -\frac{N\sigma_\omega^2}{2a}$. Therefore any CIR process with parameters κ, θ, σ such that $\frac{4\kappa\theta}{\sigma^2}$ is an integer could be written as the square summation of $N = \frac{4\kappa\theta}{\sigma^2}$ independent Ornstein-Uhlenbeck processes.

One of the advantages of the CIR process is that its drift and diffusion term are affine functions of this process. More specifically

$$\begin{aligned} E(d\lambda_t|\lambda_t) &= \kappa(\theta - \lambda(t))dt, \\ V(d\lambda_t|\lambda_t) &= \sigma_t^2 \lambda_t dt. \end{aligned}$$

This results in closed form solution for the Laplace transform of the CIR process and therefore gives closed form solutions for a vast range of financial derivatives in the models which have their stochasticity based on the CIR process.

From now on, we use the notation $\mathbb{E}_t(.) = \mathbb{E}(.|\lambda_t)$. The conditional Laplace transform of the CIR process is given by

$$\psi_{t,h}(u) = \mathbb{E}(exp(-u\lambda_{t+h})|\lambda_t).$$

Knowing the Laplace transform of the CIR process, one can find its probability distribution by the means of an inverse transform. It turns out that the distribution function of the CIR process does not have a nice closed form like its Laplace transform. The following proposition gives a closed form solution for the Laplace transform of the CIR process

Proposition 2. *The conditional Laplace transform of the CIR process λ_t defined by* (3.1) *is*

$$\begin{aligned} \psi_{t,h}(u) &= \mathbb{E}_t(\exp(-u\lambda_{t+h})) \\ &= \exp\left(-[A(h,u)\lambda_t + B(h,u)]\right), \end{aligned}$$

where the functions $A(h,u)$ and $B(h,u)$ satisfy the system of ordinary differential equations

$$\begin{aligned}
\frac{dA(h,u)}{dh} &= -\kappa A(h,u) - \frac{\sigma^2}{2}A^2(h,u), \\
\frac{dB(h,u)}{dh} &= \kappa\theta A(h,u),
\end{aligned}$$

with the initial conditions

$$\begin{cases} A(0,u) = u, \\ B(0,u) = 0, \end{cases}$$

which is a standard Ricatti differential equation with the solutions

$$\begin{aligned}
A(h,u) &= \frac{u\exp(-\kappa h)}{1+\frac{\sigma^2 u}{2\kappa}(1-\exp(-\kappa h))}, \\
B(h,u) &= \frac{2\kappa\theta}{\sigma^2}\log(1+\frac{\sigma^2 u}{2\kappa}(1-\exp(-\kappa h))).
\end{aligned}$$

Proof. See Gourieroux (2005). □

This is worthwhile to emphasize that the Laplace transform of the process λ_t is exponentially affine which is another property of the affine processes. By the representation of the function $A(h,u)$, we have

$$\lim_{h\to\infty} A(h,u) = 0,$$

provided that κ is strictly positive. In this case, the process λ_t is stationary and its probability distribution at horizon h approaches its stationary distribution with the Laplace transform given by

$$\begin{aligned}
\mathbb{E}_t(exp(-u\lambda_\infty)) &= \lim_{h\to\infty} \exp(-[A(h,u)\lambda_t + B(h,u)]) \\
&= -\frac{2\kappa\theta}{\sigma^2}\log(1+\frac{\sigma^2 u}{2\kappa}) \\
&= (1+\frac{\sigma^2 u}{2\kappa})^{-\frac{2\kappa\theta}{\sigma^2}}.
\end{aligned}$$

Knowing the Laplace transform of the CIR process, one can find its probability distribution function as follows

Proposition 3. *The conditional probability density function of the CIR process λ_t is given by*

$$f(\lambda_{t+h}|\lambda_t) = \exp\left(-\frac{(\lambda_{t+h}-\rho(h)\lambda_t)}{\zeta(h)}\right)\frac{\lambda_{t+h}^{\nu-1}}{\zeta(h)^{\nu}}\sum_{z=0}^{\infty}\left(\frac{1}{z!\Gamma(\nu+z)}(\frac{\lambda_{t+h}\rho(h)\lambda_t}{\zeta(h)})^z\right), \quad (3.2)$$

where the parameters ζ, ν and ρ are given by

$$\begin{aligned}\zeta &= \frac{\sigma^2}{2\kappa}(1-e^{-\kappa h}),\\ \nu &= \frac{2\kappa\theta}{\sigma^2},\\ \rho(h) &= e^{-\kappa h}.\end{aligned}$$

Proof. See Gourieroux (2005). □

The integrated CIR process Λ_u is defined as

$$\Lambda_u = \int_t^u \lambda_s\, ds.$$

The Laplace transform of this process is defined like the CIR process. Finding the Laplace transform of the integrated CIR process is important in financial applications because this process appears in some of the calculations related to pricing of derivatives. In analogy to proposition 2, we have the following proposition regarding the Laplace transform of the integrated CIR process.

Proposition 4. *The conditional Laplace transform of the integrated CIR process Λ_t is*

$$\begin{aligned}\psi^*_{t,h}(u) &= \mathbb{E}_t(\exp(-u\Lambda_{t+h}))\\ &= \mathbb{E}_t(\exp(-u\int_t^{t+h}\lambda_s\, ds))\\ &= \exp\left(-[A^*(h,u)\lambda_t + B^*(h,u)]\right),\end{aligned}$$

where the functions $A^*(h,u)$ and $B^*(h,u)$ are given by the formulas

$$\begin{aligned}A^*(h,u) &= \frac{2u}{\alpha(u)+\kappa} - \frac{4u\alpha(u)}{(\alpha(u)+\kappa)[(\alpha(u)+\kappa)exp(\alpha(u)h)+\alpha(u)-\kappa]},\\ B^*(h,u) &= -\frac{\kappa\theta}{\sigma^2}(\alpha(u)+\kappa)h + \frac{2\kappa\theta}{\sigma^2}\log\left[(\alpha(u)+\kappa)\exp(\alpha(u)h)+\alpha(u)-\kappa\right] - \frac{2\kappa\theta}{\sigma^2}\log\left(2\alpha(u)\right),\end{aligned}$$

with $\alpha(u) = \sqrt{\kappa^2 + 2u\sigma^2}$.

Proof. See Gourieroux (2005). □

The CIR process has vast applications in finance. As two important examples we can name the term structure model for interest rates and the Heston's stochastic volatility model. In all the models in which the CIR process is involved, the closed form formulas for the Laplace transform of the CIR process as well as the integrated CIR process induce closed form solutions for prices of derivatives. In the next section we will use the CreditGrades model with the CIR-eigenvalue covariance matrix, and derive closed form solutions for price of equity options and marginal probabilities of default. We will finish by estimating the parameters of the model.

3.3 The Model

We assume the assets and equities are defined on an underlying probability space $(\Omega, \mathcal{F}, \mathbb{Q})$ where $\{\mathcal{F}_t\}_{t\geq 0}$ is the filter generated by the random resources of the model up to time t and $\mathbb{Q}$ is the risk-neutral measure. We assume that assets are driven by the Brownian motion W_t, the CIR processes running the covariance matrix of the assets are driven by the Brownian motion Z_t and two Brownian motions W_t and Z_t are uncorrelated. In our study, we restrict ourselves to two reference firms, but the same method works for n reference companies. We fix the following notations for the rest of the chapter

- $A_i(t)$: i^{th} firm's asset price per share ,
- $S_i(t)$: i^{th} firm's equity price per share ,
- $B_i(t)$: i^{th} firm's debt per share ,
- R_i: i^{th} firm's recovery rate.

We assume that $B_i(t)$ follows the dynamics $dB_t = (r(t) - d_i(t))B_t dt$, where $r(t)$ is the risk free interest rate and $d_i(t)$ is the dividend yield for the i^{th} firm. We therefore have

$$B(t) = B(0)\exp(\int_0^t r(s) - d_i(s)\, ds).$$

If the value of the company falls below the recovery part of the debt, then the company defaults. Therefore the hitting barrier for the asset is defined as the recovery part of the debt $D_i(t) = R_i B_i(t)$. We give the default time as

$$\eta_i = \inf\{0 \le t | A_i(t) \le D_i(t)\}.$$

We define the equity's value prior to default by $S_i(t) = A_i(t) - D_i(t)$. After the default, equity's value is zero. Therefore zero is an absorbing state for $S_i(t)$ and the default time of the company in terms of its equity would be $\eta_i = \inf\{0 < t | S_i(t) \le 0\}$. This forms a connection between our option pricing theory and Lipton (2001) concerning barrier down-and-out options. Absorbing probabilities, which here in this content are probabilities of default, are available only for a limited class of stochastic processes. We show that our stochastic correlation principal component model is one of them.

We assume that the i^{th} firm's value $A_i(t)$ is driven by the dynamics

$$dA_i(t) = diag(A_i(t))[(r(t) - d_i(t))\mathbb{I}dt + \sqrt{\Sigma_t}dW_t],$$

where $W_t \in M_{m\times 1}$, $A_t \in M_{n\times 1}$, $\Sigma_t = ED_tE'$ with

$$D_t = diag(\lambda_i)_{i=1}^m \ ; \ E = (\alpha_{ij})_{n\times m}.$$

Each λ_i follows a CIR process of the type

$$d\lambda_i(t) = \kappa_i(\theta_i - \lambda_i)dt + \sigma_i\sqrt{\lambda_t}dZ_t^i.$$

In the two assets case, the above dynamics follows

$$\begin{aligned} dA_1(t) &= A_1(t)\left((r(t) - d_1(t))dt + \alpha_{11}\sqrt{\lambda_1(t)}dW_1(t) + \alpha_{12}\sqrt{\lambda_2(t)}dW_2(t)\right), \\ dA_2(t) &= A_2(t)\left((r(t) - d_2(t))dt + \alpha_{21}\sqrt{\lambda_1(t)}dW_1(t) + \alpha_{22}\sqrt{\lambda_2(t)}dW_2(t)\right), \end{aligned} \tag{3.3}$$

where the eigenvalues of the covariance process follow

$$\begin{aligned} d\lambda_1(t) &= \kappa_1(\theta_1 - \lambda_1)dt + \sigma_1\sqrt{\lambda_1(t)}dZ_t^1, \\ d\lambda_2(t) &= \kappa_2(\theta_2 - \lambda_2)dt + \sigma_2\sqrt{\lambda_2(t)}dZ_t^2. \end{aligned}$$

Assuming ζ as the angle that the first eigenvector makes with the real axis, the eigenvector matrix E is given by

$$E = \begin{pmatrix} \cos(\zeta) & -\sin(\zeta) \\ \sin(\zeta) & \sin(\zeta) \end{pmatrix}.$$

We assume that assets are driven by the Brownian motion W_t, the covariance matrix of the assets is driven by the Brownian motion Z_t and two Brownian motions W_t and Z_t are uncorrelated. The reason we make the independence assumption between stock and its volatility is that closed form formulas for the value of double-barrier options and equity options are not available when the asset and its volatility are correlated as pointed out by Lipton (2001) and Sepp (2006).

The infinitesimal generator of the joint process (S, Σ), $\mathcal{A}_{(S,\Sigma)}$, appears in the pricing PDE. Here we find a formula for this operator to use it for our pricing purposes in the next section. Since $S_i(t) = A_i(t) - D_i(t)$, the equity satisfies the stochastic differential equation

$$dS_t = [(r_t - d_t)S_t]dt + [S_t + D_t]\Sigma_t^{\frac{1}{2}} dW_t.$$

$\mathcal{A}_{(S,\Sigma)}$ can be divided into three terms related to the stock's operator, the covariance operator and their joint operator

$$\mathcal{A}_{(S,\Sigma)} = \mathcal{A}_S + \mathcal{A}_\Sigma + \mathcal{A}_{<S,\Sigma>}.$$

Since dZ_t and dW_t are independent, the last term is zero. From the dynamics of the equity, we know that

$$\begin{aligned} \mathcal{A}_S(W) &= (r(t) - d_1(t))S_1 W_{S_1} + \frac{1}{2}(S_1 + D_1(t))^2 (\sum_{j=1}^m \alpha_{1j}^2 \lambda_j) W_{S_1 S_1} \\ &= (r(t) - d_2(t))S_2 W_{S_2} + \frac{1}{2}(S_2 + D_2(t))^2 (\sum_{j=1}^m \alpha_{2j}^2 \lambda_j) W_{S_2 S_2}, \end{aligned} \tag{3.4}$$

and from the classical results regarding the infinitesimal generator of the CIR process

$$\mathcal{A}_\Sigma(W) = \sum_{i=1}^m [\kappa_i(\theta_i - \lambda_i) W_{\lambda_i} + \frac{1}{2}\sigma_i^2 \lambda_i W_{\lambda_i \lambda_i}].$$

Therefore if W is a derivative on the first underlying asset only, we have

$$\mathcal{A}_{(S,\Sigma)}(W) = (r(t)-d_1(t))S_1W_{S_1}+\frac{1}{2}(S_1+D_1(t))^2(\sum_{j=1}^{m}\alpha_{1j}^2\lambda_j)W_{S_1S_1}+\sum_{i=1}^{m}[\kappa_i(\theta_i-\lambda_i)W_{\lambda_i}+\frac{1}{2}\sigma_i^2\lambda_iW_{\lambda_i\lambda_i}]. \tag{3.5}$$

In the next section, we derive closed formulas for the price of equity options and marginal probabilities of default.

3.4 Derivative Pricing; Analytical Results

In the two factor model, the first asset follows

$$\begin{aligned} dA_1(t) &= A_1(t)\left((r(t)-d_1(t))dt+\alpha_{11}\sqrt{\lambda_1(t)}dW_1(t)+\alpha_{12}\sqrt{\lambda_2(t)}dW_2(t)\right), \\ d\lambda_1(t) &= \kappa_1(\theta_1-\lambda_1)dt+\sigma_1\sqrt{\lambda_1(t)}dZ_t^1, \\ d\lambda_2(t) &= \kappa_2(\theta_2-\lambda_2)dt+\sigma_2\sqrt{\lambda_2(t)}dZ_t^2. \end{aligned}$$

Calculating equity option prices is essential to calibrate the stochastic correlation CreditGrades model since this model uses the information available from equity options to estimate model parameters. Later, we will use the evolutionary algorithm method to match the theoretical results of our extended CreditGrades model with market data. One of the advantages of the CreditGrades model compared to Merton's model, is the straightforward link it makes with the equity option markets. The price of the equity option can be calculated by discounting the payoff function at maturity. The only subtle point here in pricing these options lies in the specific dynamics of the equity itself and the possibility of default for the company. In the Black-Scholes model, the stock follows geometric Brownian motion which is a strictly positive process with a log-normal distribution and never hits zero. In the CreditGrades model, equity is modeled as a process satisfying a shifted log-normal distribution which hits the state zero when the company defaults. Because of the absorbing property of the state zero for the equity process, there is a resemblance between pricing the equity options and the pricing of the down-and-out options. By considering the barrier condition for equity, the payoff of an equity call option is given by $\left(S_T\mathbb{I}_{\{\eta>T\}}-K\right)^+$. Therefore the price of an equity call option can be written as

$$\begin{aligned} V_{call}(t,\Sigma,S,K) &= E^{Q}_{(t,\Sigma,S)}\left(\exp(-\int_t^T r(s)\,ds)\,(S_T\mathbb{I}_{\eta>T}-K)^+\right) \\ &= E^{Q}_{(t,\Sigma,S)}\left(\exp(-\int_t^T r(s)\,ds)\,(S_T-K)^+\,\mathbb{I}_{\eta>T}\right). \end{aligned} \tag{3.6}$$

Similarly, the payoff of an equity put option is $(K - S_T\mathbb{I}_{\eta>T})^+$. Therefore the price of an equity put option is given by

$$\begin{aligned} V_{put}(t,\Sigma,S,K) &= E^{Q}_{(t,\Sigma,S)}\left(\exp(-\int_t^T r(s)\,ds)\,(K-S_T\mathbb{I}_{\eta>T})^+\right) \\ &= E^{Q}_{(t,\Sigma,S)}\left(\exp(-\int_t^T r(s)\,ds)\,\left((K-S_T)^+\mathbb{I}_{\eta>T}+K\mathbb{I}_{\eta\leq T}\right)\right). \end{aligned} \tag{3.7}$$

Equations (3.6) and (3.7) give the put-call parity formula for equity options:

$$V_{call}(t,\Sigma,S,K) - V_{put}(t,\Sigma,S,K) = V_{call}(t,\Sigma,S,0) - K\exp(-\int_t^T r(s)\,ds). \tag{3.8}$$

The following proposition gives a closed form solution for the price of an equity call option on the first asset. Proposition 5 and equation (3.8) give the price of an equity put option. This result is an essential tool to calibrate the model in the next section.

Proposition 5. *The price of a call option on $S_1(t)$ with maturity date T and strike price K is given by*

$$W(t,S_j) = (D_1(T)+K)\exp(-\int_t^T r(s)\,ds)Z(\tau,y),$$

with

$$
\begin{aligned}
Z(\tau, y) &= e^y - e^b - \frac{e^{\frac{1}{2}y}}{\pi} \int_0^{+\infty} \frac{e^{A(\tau,k)+\sum_{i=1}^m B_i(\tau,k)\lambda_i}(\cos(yk) - \cos((y-2b)k))}{k^2+\frac{1}{4}} \, ds, \\
B_i(\tau, k) &= -\alpha_{1i}^2(k^2+\frac{1}{4})\frac{1-e^{-\zeta_i\tau}}{\psi_-^{(i)} + \psi_+^{(i)}e^{-\zeta_i\tau}}, \\
A(\tau, k) &= \sum_{i=1}^m -\frac{\kappa_i\theta_i}{\sigma_i^2}\left(\tau\psi_+^{(i)} + 2\ln(\frac{\psi_-^{(i)} + \psi_+^{(i)}e^{-\tau\zeta_i}}{2\zeta_i})\right), \\
y &= \ln(\frac{S+D(t)}{D(T)+K}) + \int_t^T (r(s)-d(s))\, ds, \\
b &= \ln(\frac{D(t)}{D(T)+K}) + \int_t^T (r(s)-d(s))\, ds, \\
\psi_\pm^{(i)} &= \mp\kappa_i + \zeta_i, \\
\zeta_i &= \sqrt{\kappa_i^2 + \alpha_{1i}^2\sigma_i^2(k^2+\frac{1}{4})}.
\end{aligned}
$$

Proof. By risk neutral valuation, W satisfies

$$W_t + \mathcal{A}_{(\lambda,S)} - rW = 0,$$

where $\mathcal{A}_{(\lambda,S)}$ is the infinitesimal generator of the SDE driving the equity. By substitution

$$W_t + (r(t)-d_1(t))S_1W_{S_1} + \frac{1}{2}(S_1+D_1(t))^2(\sum_{j=1}^m \alpha_{1j}^2\lambda_j)W_{S_1S_1} + \sum_{i=1}^m[\kappa_i(\theta_i-\lambda_i)W_{\lambda_i} + \frac{1}{2}\sigma_i^2\lambda_iW_{\lambda_i\lambda_i}] - rW = 0.$$

Henceforth, wherever it doesn't cause confusion, for simplicity we drop the index i. We change the variables to:

$$
\begin{aligned}
x = \ln(\frac{S+D(t)}{D(t)}) \quad , \quad \alpha = \ln\frac{(D(T)+K}{D(T))}, \\
G(t,x) &= \exp(\int_t^T r(s)\, ds)\frac{W(t,S)}{D(T)},
\end{aligned}
$$

which gives

$$G_t + \frac{1}{2}(\sum_{j=1}^m \alpha_{1j}^2\lambda_j)(G_{xx} - G_x) + \sum_{i=1}^m[\kappa_i(\theta_i-\lambda_i)G_{\lambda_i} + \frac{1}{2}\sigma_i^2\lambda_iG_{\lambda_i\lambda_i}] = 0.$$

We perform the second change of variables as

$$
\begin{aligned}
y = x - a \quad &, \quad \tau = T - t, \\
F(\tau, y) &= e^{-a} G(t, x),
\end{aligned}
$$

which transforms the PDE to

$$
-F_\tau + \frac{1}{2}(\sum_{j=1}^{m} \alpha_{1j}^2 \lambda_j)(F_{yy} - F_y) + \sum_{i=1}^{m} [\kappa_i(\theta_i - \lambda_i) F_{\lambda_i} + \frac{1}{2}\sigma_i^2 \lambda_i F_{\lambda_i \lambda_i}] = 0.
$$

Finally we perform the third change of variables

$$
U(\tau, y) = e^{-\frac{y}{2}} F(\tau, y),
$$

to get

$$
-U_t + \frac{1}{2}(\sum_{j=1}^{m} \alpha_{1j}^2 \lambda_j)(U_{yy}) + \sum_{i=1}^{m} [\kappa_i(\theta_i - \lambda_i) U_{\lambda_i} + \frac{1}{2}\sigma_i^2 \lambda_i U_{\lambda_i \lambda_i}] - \frac{1}{8}(\sum_{j=1}^{m} \alpha_{1j}^2 \lambda_j) U = 0.
$$

We claim that the Fourier transform of the Green's function for the above PDE is of the form

$$
q(\tau, \lambda, Y) = \int_{-\infty}^{+\infty} e^{ikY + A(\tau,k) + \sum_{j=1}^{m} B_j(\tau,k)\lambda_j} \, dk, \tag{3.9}
$$

with

$$
\begin{aligned}
B_i(\tau, k) &= -\alpha_{1i}^2 (k^2 + \frac{1}{4}) \frac{1 - e^{-\zeta_i \tau}}{\psi_-^{(i)} + \psi_+^{(i)} e^{-\zeta_i \tau}}, \\
A(\tau, k) &= \sum_{i=1}^{m} -\frac{\kappa_i \theta_i}{\sigma_i^2} \left(\tau \psi_+^{(i)} + 2 \ln (\frac{\psi_-^{(i)} + \psi_+^{(i)} e^{-\tau \zeta_i}}{2\zeta_i}) \right), \\
\psi_\pm^{(i)} &= \mp \kappa_i + \zeta_i, \\
\zeta_i &= \sqrt{\kappa_i^2 + \alpha_{1i}^2 \sigma_i^2 (k^2 + \frac{1}{4})}.
\end{aligned}
$$

We know that $q(\tau, \lambda, Y)$ satisfies the corresponding PDE. Plugging $q(\tau, \lambda, Y)$ into the PDE, one gets the following Ricatti ODE's for $A(\tau, k)$ and $B_j(\tau, k)$'s

$$-(A_\tau+\sum_{j=1^m} B_\tau^{(j)}\lambda_j)+\frac{1}{2}(\sum_{j=1}^m \alpha_{1j}^2\lambda_j)(-k^2)+\sum_{i=1}^m [\kappa_i(\theta_i-\lambda_i)B_i+\frac{1}{2}\sigma_i^2\lambda_i B_i]-\frac{1}{8}(\sum_{j=1}^m \alpha_{1j}^2\lambda_j)=0.$$

The above ODE's functions, A and B_i, are independent of the variables λ_i. Therefore assuming $\lambda_j = 0$ for $j = 1, 2, \ldots, m$, we get

$$A_\tau - \sum_{i=1}^m \kappa_i\theta_i B_i = 0. \tag{3.10}$$

Now assume $\lambda_j = 0$ for $j \in \{1, 2, \ldots, m\} - \{u\}$, to get

$$(B_u)_\tau - \frac{1}{2}\sigma_u^2 B_u^2 + \kappa_u B_u + \frac{1}{2}\alpha_{1u}^2(k^2+\frac{1}{4}) = 0, \tag{3.11}$$

for $u = 1, 2, \ldots, m$. Now we have $m+1$ ODE's from equations (3.10) and (3.11) and $m+1$ unknown functions with initial conditions

$$A(0,k) = 0 \qquad , B_u(0,k) = 0 \quad \text{for } u = 1, 2, \ldots, m.$$

The system of ODE's (3.11) are standard Ricatti equations which after a change of variable

$$-\frac{1}{2}\sigma_u^2 B_u = \frac{(C_u)_\tau}{C_u},$$

turn to the system of second order linear ODE's

$$(C^{(u)})_{\tau\tau} + \kappa_u C_\tau^{(u)} - \frac{1}{4}\alpha_{1u}^2\sigma_u^2(k^2+\frac{1}{4})C^{(u)} = 0,$$

with initial conditions $C_\tau^{(u)}(0,k) = 0$ and $C^{(u)}(0,k) = 1$. The solution to the above ordinary differential equation with initial conditions is given by

$$C^{(u)} = c_+ e^{\Lambda_+\tau} + c_- e^{\Lambda_-\tau},$$

where the coefficients Λ_+ and Λ_- satisfy

$$\Lambda_\pm^2 + \kappa_u\Lambda_\pm - \frac{1}{4}\alpha_{1u}^2\sigma_u^2(k^2+\frac{1}{4}) = 0,$$

with solutions

$$\begin{aligned}\Lambda_{\pm} &= \frac{-\kappa_i \pm \sqrt{\kappa_i^2 + \alpha_{1i}^2\sigma_i^2(k^2+\frac{1}{4})}}{2} \\ &= \frac{-\kappa_i}{2} \pm \frac{\zeta_i}{2}.\end{aligned}$$

After matching the coefficients c_+ and c_- from the initial conditions, one has the function $C^{(u)}$ as

$$\begin{aligned}C^{(u)}(\tau,k) &= (\frac{\kappa_i+\zeta_i}{2\zeta_i})e^{(\frac{-\kappa_i}{2}+\frac{\zeta_i}{2})\tau} + (\frac{-\kappa_i+\zeta_i}{2\zeta_i})e^{(\frac{-\kappa_i}{2}-\frac{\zeta_i}{2})\tau} \\ &= \frac{\psi_-^{(i)}e^{\frac{\psi_+^{(i)}}{2}\tau} + \psi_+^{(i)}e^{-\frac{\psi_-^{(i)}}{2}\tau}}{2\zeta_i} \\ &= e^{\frac{\psi_+^{(i)}}{2}\tau}(\frac{\psi_-^{(i)}+\psi_+^{(i)}e^{-\zeta_i\tau}}{2\zeta_i}),\end{aligned}$$

which finally gives the function $B_{(u)}(\tau,k)$ as

$$\begin{aligned}B_{(u)}(\tau,k) &= \frac{1}{-\frac{1}{2}\sigma_u^2}.\frac{(C_u)_\tau}{C_u} \\ &= -\alpha_{1i}^2(k^2+\frac{1}{4})\frac{1-e^{-\zeta_i\tau}}{\psi_-^{(i)}+\psi_+^{(i)}e^{-\zeta_i\tau}}.\end{aligned}$$

The representation for the function $A(\tau,k)$ comes from equation (3.10)

$$\begin{aligned}A(\tau,k) &= \sum_{i=1}^{m} -\frac{2\kappa_i\theta_i}{\sigma_i^2}\ln(C_\tau^{(u)}) \\ &= \sum_{i=1}^{m} -\frac{\kappa_i\theta_i}{\sigma_i^2}\left(\psi_+^{(i)}\tau + 2\ln(\frac{\psi_-^{(i)}+\psi_+^{(i)}e^{-\zeta_i\tau}}{2\zeta_i})\right).\end{aligned}$$

Note that $q(\tau,\lambda_1,\lambda_2,Y)$ has a structure that is invariant with respect to the change of variables

$$y \to -y \quad \text{and} \quad k \to -k.$$

Therefore the Fourier transform absorbed at $x = b$ is

$$q^{(b)}(\tau,\lambda,y,y') = q(\tau,\lambda,y'-y) - q(\tau,\lambda,y'+y-2b).$$

The above expression is all that one needs to use with Duhamel's formula

$$
\begin{aligned}
U(\tau,y) &= \frac{1}{2\pi}\int_0^{+\infty}\left(e^{\frac{y'}{2}}-e^{-\frac{y'}{2}}\right)q^{(b)}(\tau,\lambda,y,y')\,dy' \\
&= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_0^{+\infty}\left(e^{\frac{y'}{2}}-e^{-\frac{y'}{2}}\right)e^{iky'}(e^{-iky}-e^{ik(y-2b)})e^{A(\tau,k)+\sum_{j=1}^m B_j(\tau,k)\lambda_j}\,dy'dk \\
&= \frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{A(\tau,k)+\sum_{j=1}^m B_j(\tau,k)\lambda_j}\left(e^{-iky}-e^{ik(y-2b)}\right)\int_0^{+\infty}\left(e^{(\frac{1}{2}+ik)y'}-e^{(-\frac{1}{2}+ik)y'}\right)dy'dk \\
&= \frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{A(\tau,k)+\sum_{j=1}^m B_j(\tau,k)\lambda_j}\left(e^{-iky}-e^{ik(y-2b)}\right)\left(-\frac{1}{\frac{1}{2}+ik}+2\pi\delta(k-\frac{i}{2})+\frac{1}{-\frac{1}{2}+ik}\right)dk \\
&= e^{\frac{y}{2}}-e^{-\frac{1}{2}(y-2b)}-\frac{1}{2\pi}\int_{-\infty}^{+\infty}\frac{e^{A(\tau,k)+\sum_{j=1}^m B_j(\tau,k)\lambda_j}\left(e^{iky}-e^{ik(y-2b)}\right)}{k^2+\frac{1}{4}}\,dk.
\end{aligned}
$$

Since $Z(\tau,y)=e^{\frac{y}{2}}U(\tau,y)$,

$$
Z(\tau,y)=e^y-e^b-\frac{e^{\frac{y}{2}}}{2\pi}\int_{-\infty}^{+\infty}\frac{e^{A(\tau,k)+\sum_{j=1}^m B_j(\tau,k)\lambda_j}(\cos(yk)-\cos((y-2b)k))}{k^2+\frac{1}{4}}\,dk.
$$

□

Similar techniques can be used to find the marginal probabilities of default. Suppose $P(t,T,S_i)$ is the survival probability for the i^{th} company

$$
P(t,T,S_i)=\mathbb{Q}_{(t,\Sigma,S_i)}(S_i(\tau)>0|t<\tau\le T).
$$

Using the Feynman-Kac formula, $P(t,T,S_i)$ satisfies the partial differential equation $P_t+\mathcal{A}_{(S,\Sigma)}P=0$ with boundary conditions $P(t,T,0)=0$ and $P(T,T,S)=1$. We have the following proposition for the survival probabilities

Proposition 6. *The survival probability for the i^{th} firm is given by*

$$
P(t,T,S_i)=\frac{2e^{\frac{y}{2}}}{\pi}\int_0^{+\infty}\frac{e^{A(\tau,k)+\sum_{j=1}^m B_j(\tau,k)\lambda_j}k\sin(ky)}{k^2+\frac{1}{4}}\,dk.
$$

Proof. Substituting for the infinitesimal generator from equation (3.5), $P(t,T,S_i)$ solves

$$
\begin{cases}
P_t+(r(t)-d_1(t))S_1P_{S_1}+\frac{1}{2}(S_1+D_1(t))^2(\sum_{j=1}^m\alpha_{1j}^2\lambda_j)P_{S_1S_1} \\
+\sum_{i=1}^m[\kappa_i(\theta_i-\lambda_i)P_{\lambda_i}+\frac{1}{2}\sigma_i^2\lambda_iP_{\lambda_i\lambda_i}]=0 \\
P(t,T,0)=0,P(T,T,S)=1.
\end{cases}, \tag{3.12}
$$

Using the change of variables $y_i = \ln\left(\frac{S_1(t)+D_1(t)}{D_1(t)}\right)$, $\tau = T - t$ and $P(t,T,S_i) = e^{\frac{y}{2}} U(\tau, y_i)$, the PDE (3.12) transforms to

$$\begin{cases} -U_t + \frac{1}{2}(\sum_{j=1}^m \alpha_{1j}^2 \lambda_j)(U_{yy}) + \sum_{i=1}^m [\kappa_i(\theta_i - \lambda_i)U_{\lambda_i} + \frac{1}{2}\sigma_i^2 \lambda_i U_{\lambda_i \lambda_i}] - \frac{1}{8}(\sum_{j=1}^m \alpha_{1j}^2 \lambda_j) U = 0, \\ U(\tau, 0) = e^{-\frac{y_i}{2}}, U(0, y_i) = 0. \end{cases} \tag{3.13}$$

In the proof of proposition 5 we showed that the Fourier transform of the Green's function for the above PDE is of the form

$$q(\tau, \lambda, Y) = \int_{-\infty}^{+\infty} e^{ikY + A(\tau,k) + \sum_{j=1}^m B_j(\tau,k)\lambda_j} \, dk, \tag{3.14}$$

where the functions $A(\tau, k)$ and $B_j(\tau, k)$ are given by

$$\begin{aligned} A(\tau,k) &= \sum_{i=1}^m -\frac{\kappa_i \theta_i}{\sigma_i^2} \left(\tau \psi_+^{(i)} + 2 \ln\left(\frac{\psi_-^{(i)} + \psi_+^{(i)} e^{-\tau \zeta_i}}{2\zeta_i}\right) \right), \\ B_i(\tau,k) &= -\alpha_{1i}^2 (k^2 + \frac{1}{4}) \frac{1 - e^{-\zeta_i \tau}}{\psi_-^{(i)} + \psi_+^{(i)} e^{-\zeta_i \tau}}. \end{aligned}$$

To find a bounded solution reflected at $x = 0$, we use the method of images to write the absorbed aggregated Green's function as

$$q^{(0)}(\tau, \Sigma, y, y') = q(\tau, \Sigma, y' - y) - q(\tau, \Sigma, y' + y).$$

Now by Duhamel's formula

$$\begin{aligned} U(\tau, y) &= \frac{1}{2\pi} \int_0^{+\infty} e^{-\frac{y'}{2}} q^{(0)}(\tau, \Sigma, y, y') \, dy' \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_0^{+\infty} e^{-\frac{y'}{2}} e^{iky'} (e^{-iky} - e^{iky}) e^{A(\tau,k) + \sum_{j=1}^m B_j(\tau,k)\lambda_j} \, dy' dk \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{A(\tau,k) + \sum_{j=1}^m B_j(\tau,k)\lambda_j} (e^{-iky} - e^{iky}) \int_0^{+\infty} e^{(-\frac{1}{2} + ik)y'} \, dy' dk \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{A(\tau,k) + \sum_{j=1}^m B_j(\tau,k)\lambda_j} (e^{-iky} - e^{iky}) \left(\frac{1}{-\frac{1}{2} + ik}\right) dk \\ &= \frac{2}{\pi} \int_0^{+\infty} \frac{e^{A(\tau,k) + \sum_{j=1}^m B_j(\tau,k)\lambda_j} k \sin(ky)}{k^2 + \frac{1}{4}} \, dk. \end{aligned}$$

Therefore the survival probability is given by

$$P(t,T,S_i) = e^{\frac{y}{2}}U(\tau,y_i) = \frac{2e^{\frac{y}{2}}}{\pi}\int_0^{+\infty} \frac{e^{A(\tau,k)+\sum_{j=1}^m B_j(\tau,k)\lambda_j} k\sin(ky)}{k^2+\frac{1}{4}}\,dk.$$

□

Knowing the probability of the default, one can find the CDS spread for the underlying company. Assume that the CDS spread is denoted by S, the periodic payments occur at $0 = T_0 < T_1 < \ldots < T_N = T$, the notional is N, the time of default is denoted by τ and the recovery rate is the constant R. The *fixed leg* of the CDS is the value at time $t = 0$ of the cash flow corresponding to the payments the buyer makes. With the above notation we have

$$\begin{aligned}
\text{Fixed Leg} &= E^{\mathbb{Q}}(\sum_{i=0}^{N} e^{-\int_t^{T_i} r(s)\,ds} SN(T_i - T_{i-1})\mathbb{I}_{\tau\geq T_i}) \\
&= SN\sum_{i=0}^{N} e^{-\int_t^{T_i} r(s)\,ds}(T_i - T_{i-1})Q(t,T_i).
\end{aligned} \tag{3.15}$$

On the other hand the *floating leg*, which is the value of the protection cash flow at $t = 0$, is

$$\begin{aligned}
\text{Floating Leg} &= E^{\mathbb{Q}}(\sum_{i=0}^{N} e^{-\int_t^{T_i} r(s)\,ds}(1-R)\mathbb{I}_{T_{i-1}<\tau\leq T_i}) \\
&= (1-R)\sum_{i=0}^{N} e^{-\int_t^{T_i} r(s)\,ds}(Q(t,T_{i-1}) - Q(t,T_i)).
\end{aligned} \tag{3.16}$$

The CDS spread S is chosen such that the contract has a fair value at $t = 0$. By setting the fixed leg equal to the floating leg, the equations (4.27) and (4.28) imply

$$S = \frac{(1-R)\sum_{i=0}^{N} e^{-\int_t^{T_i} r(s)\,ds}(Q(t,T_{i-1}) - Q(t,T_i))}{\sum_{i=0}^{N} e^{-\int_t^{T_i} r(s)\,ds}Q(t,T_i)}.$$

3.5 Model Calibration with Evolutionary Algorithms

In this section, we present the calibration of the model as an optimization problem. Our approach for the calibration of our model is inspired by Hamida and Cont (2004).

We use evolutionary algorithms as a probabilistic tool to do this task and verify the theoretical results regarding the convergence of the algorithm. By calibrating our credit risk model, we tackle an ill-posed inverse problem. The solution to the minimization problem is not unique and in some cases, a global solution might not exist. If the model is not flexible enough, it may not capture the market's behavior. On the other hand, if the model is too complex, it might fail to give a unique solution for the optimization problem to find the parameters, making the calibration an ill-posed problem. The results of our model calibration method will be later used to price more complex derivatives on two stochastically correlated assets in chapter 5. There are numerous research papers on calibration of structural credit risk models which assume the volatility of the asset is constant. For instance, Hull *et al.* (2004) considers constant volatility for the assets and uses equity options to implement the model and to find the probabilities of default. When assuming the volatility of the asset to be constant (as in Black and Scholes), implementing the Merton model is easy and can be done by two approaches discussed in Hull *et al.* (2004). Assuming the volatility to be drawn by a CIR process makes the implementation more difficult. In this section we tackle this problem and give estimation results for options on General Motors' and Ford's equities.

3.5.1 Evolutionary Algorithms

Evolutionary algorithms give solutions for non-convex optimization problems. The parameters are found by searching through a compact bounded set $E \subset \mathbb{R}^n$. An initial population is selected from the parameter space E and then the individuals go under a mutation-selection procedure to search for the best parameters which minimize the error function. Searching process is done through mutation and during selection better optimizers are being recognized. Our special application of this method of calibration is to implement the CreditGrades model for credit risk.

During mutation, individuals go through random transformations which help the algorithm search the regions of the space. Then the candidate parameters are selected based on their performance according to the error function and an annealing parameter β_n. By letting $\beta_n \to +\infty$ we make sure that the algorithm eventually converges to the global minima regions. There's usually one middle transformation between mutation and selection called crossover. Crossover is not needed to guarantee the convergence of the algorithm, but it helps the algorithm to find the solution faster. In this stage, every pair of individuals is used to get another parameter. Crossover helps the algorithm to search the space better. Now suppose we are in the n^{th} step of the algorithm with the population $\theta_1^n, \theta_2^n, \ldots, \theta_N^n$. In the selection part of the algorithm, the individual θ_j^n is selected to remain in the population for the next step with probability $\exp(-\beta_n G(\theta_j))$ and if it's not selected, the individual θ_j^n is selected instead with probability

$$\frac{\exp(-\beta_n G(\theta_j^n))}{\sum_{k=1}^N (\exp(-\beta_n G(\theta_k^n)))}.$$

Note that some of the parameters might appear more than once after this step. We assume that $\beta_n \to \infty$. This forces the population to approach the regions where the error function is minimized. Based on the condition above, the selection kernel is given by

$$S_t^n(x, dy) = e^{-\beta_n G(x)}\delta_x(dy) + (1 - e^{-\beta_n G(x)})\frac{e^{-\beta_n G(x)}\mu(dy)}{\int \mu(dz)e^{-\beta_n G(x)}}.$$

After setting up the ingredients of the algorithm, one should verify the convergence behavior of the individuals undergoing the transitions. For this purpose, we force some mixing conditions on the mutation kernel $M(x, dy)$, for instance a truncated gaussian kernel on the search space E bears this condition. On the other hand, we choose the selection pressure such that it moves the Markov chain to the global minima of the error function. We consider a very large sample ($N \sim \infty$) and instead of θ_n^i's, work with

$$\mu_n^N = \frac{1}{N}\sum_{i=1}^N \delta_{\theta_n^i} \in M_1(E),$$

where $M_1(E)$ is the space of probability measures on E. By Moral and Miclo (2003), the flow $(\mu_n^N)_{n\geq 0}$ converges weakly to $(\mu_n)_{n\geq 0}$ when $N \to \infty$, where $(\mu_n)_{n\geq 0}$ is given by the recursive equation

$$\mu_{n+1} = \mu_n M_n S_{\mu_n}^n,$$

with the selection kernel given by

$$S_t^n(x, dy) = e^{-\beta_n G(x)}\delta_x(dy) + (1 - e^{-\beta_n G(x)})\frac{e^{-\beta_n G(x)}\mu(dy)}{\int \mu(dz)e^{-\beta_n G(x)}}. \tag{3.17}$$

It can be proved as in Moral and Miclo (2003) that if M is mixing and β_n is increasing, the population concentrates on semi-minimum areas of the error function G.

Proposition 7. *If $\beta_n = n^\alpha$ where $\alpha \in (0, 1)$ and $M_n = M$ has the mixing property*

$$\exists\epsilon, \text{ such that } \forall x, y \in E \quad M(x, .) \geq \epsilon M(y, .),$$

then

$$\forall \delta, \mu_n(G(\theta) \geq G^* + \delta) \xrightarrow{n\uparrow\infty} 0,$$

where

$$\begin{aligned} G^* &= \inf\{\int_E G(\theta)\mu(d\theta); \mu \in M_1(E) \text{ with } I(\mu) < \infty\}, \\ I(\mu) &= \inf_K \int \mu(dx) H(K(x,.)|M(x,.)), \end{aligned}$$

and the infimum for $I(\mu)$ is taken over all Markov kernels K with stationary distribution μ.

Proof. See Moral and Miclo (2003). □

Remark 1. *The value G^* is not always equal to the minimum of G over the search space. Later on,we will show that for our model, G^* coincides with G (see proposition 25 part c).*

3.5.2 Model Calibration

Assume the model has a parameter set θ. Take $E \subset \mathbb{R}^m$ to be a compact set in which our search algorithm to find θ takes place. For every parameter θ, there is a risk-neutral measure Q^θ which gives the price of the derivatives on the underlying asset

$$C_i^\theta(t,T) = E_t^{Q^\theta}(e^{-\int_t^T r(s)\,ds}\Phi(T)),$$

where $E_t^{Q^\theta}$ is the risk-neutral expectation corresponding to the parameter θ and $\Phi(T)$ is the pay-off of the derivative at time T. To find a parameter θ such that the theoretical prices of all the instruments match their corresponding market prices is not feasible because of the noisy observations of the market prices hidden in the bid-ask spread and the possible model mis-specifications. Hence we define an error function to match the theoretical prices with the market prices. By *calibrating* the model's parameters, we aim to solve the inverse problem of optimizing the error function under stochastic search scenarios for the parameter θ. The definition of the error function depends on the problem in hand. We will give our definition of the error function below. By considering the error function G, one can see how well the parameter set θ is doing with respect to the market data. The less the error $G(\theta)$ is, the closer theoretical prices are to market prices and the better the parameter θ. Now we specify the components of our evolutionary algorithms as follows : parameters of the model, calculation of the optimizer function, initial Distribution, mutation kernel, crossover outputs, selection criteria and convergence of the algorithm.

Parameters of the Model

Let's first make a quick review of the model. The assets and the equities follow the dynamics

$$\begin{aligned} dA_i(t) &= diag(A_i(t))[(r(t) - d_i(t))\mathbb{I}dt + \sqrt{\Sigma_t}dW_t], \\ dS_t &= [(r_t - d_t)S_t]dt + [S_t + D_t]\Sigma_t^{\frac{1}{2}}dW_t. \end{aligned}$$

By our assumption, the covariance matrix can be decomposed into $\Sigma_t = ED_tE'$ where $D_t = diag(\lambda_i)_{i=1}^m$ is a diagonal matrix in which $\{\lambda_i\}_{i=1}^m$ are the eigenvalues of the covariance matrix and they follow the CIR process

$$d\lambda_i(t) = \kappa_i(\theta_i - \lambda_i)dt + \sigma_i\sqrt{\lambda_t}dZ_t^i.$$

The matrix $E = (\alpha_{ij})_{n\times m}$ is an orthonormal matrix such that its columns are eigenvectors of the covariance matrix. The set of the parameters for the first and the second eigenvalues are

$$\begin{array}{cccc} \kappa_1, & \theta_1, & \sigma_1, & \lambda_1, \\ \kappa_2, & \theta_2, & \sigma_2, & \lambda_2, \end{array}$$

where λ_1 and λ_2 are the initial eigenvalues. And the eigenvectors are

$$E = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}.$$

The matrix E can be identified with the angle that the first eigenvalue makes with the $x-$axis, ζ . Therefore the parameters set can be considered as $\theta = (\kappa_1, \theta_1, \sigma_1, \lambda_1, \kappa_2, \theta_2, \sigma_2, \lambda_2, \zeta)$ with nine elements. We put some conditions on the parameters to optimize the convergence time. The bounds for the parameters are as

$$0 < \kappa_i < 20 \quad , \quad 0 < \theta_i < 0.2 \quad , \quad 0 < \sigma_i < 10 \quad , \quad 0 < \lambda_i < 0.2.$$

The upper boundaries for the parameters θ_i and λ_i were obtained by looking at the at-the-money volatilities of the equity options prices. The bounds for κ_i and σ_i were chosen by the literature reports on the calibration of the Heston model. Moreover, we assume that $\kappa_i.\theta_i - \sigma_i^2 > 0$ to ensure the stability of the eigenvalues.

Calculation of the Optimizer Function

The stochastic covariance matrix is

$$\Sigma_t = ED_tE' = \begin{pmatrix} \alpha_{11}^2\lambda_1 + \alpha_{12}^2\lambda_2 & \alpha_{11}\alpha_{21}\lambda_1 + \alpha_{12}\alpha_{22}\lambda_2 \\ \alpha_{11}\alpha_{21}\lambda_1 + \alpha_{12}\alpha_{22}\lambda_2 & \alpha_{21}^2\lambda_1 + \alpha_{22}^2\lambda_2 \end{pmatrix}.$$

As it can be seen from the above matrix expression, the marginal and joint parameters of the model are shared among the processes λ_1 and λ_2. Therefore, unlike many other calibration methods regarding multidimensional models, it is difficult to classify the parameters into marginal and joint parameters and then estimate them separately. For this reason, we define the objective function in a way that captures both marginal and joint behavior of the companies' equities. We define the objective function as

$$\begin{aligned} G(\theta) &= \sum_{i=1}^{I} \omega_i^2 |C_i^\theta(t,T_i) - C_i^*(t,T_i)|^2 \\ &+ \sum_{i=1}^{I} \omega_i \nu_i |C_i^\theta(t,T_i).D_i^\theta(t,T_i) - C_i^*(t,T_i).D_i^*(t,T_i)| \\ &+ \sum_{i=1}^{I} \nu_i^2 |D_i^\theta(t,T_i) - D_i^*(t,T_i)|^2, \end{aligned} \tag{3.18}$$

where I is the number of options available from each company, $C_i^\theta(t,T)$ and $D_i^\theta(t,T)$ are the theoretical prices for the first and the second company, respectively, and $C_i^*(t,T)$ and $D_i^*(t,T)$ are the corresponding market prices. As it can be seen, the objective function has three parts. The first and the third line of the objective function match the marginal parameters and the second line matches the joint parameters. Another important fact is that there is always a bid-ask spread for prices of financial instruments. Therefore it is impractical to expect $G(\theta) = 0$. Because of this fact we define a degree of accuracy as

$$\delta = \sum_{i=1}^{I} \omega_i^2 |C_i^{bid^2} - C_i^{ask^2}| + \sum_{i=1}^{I} \omega_i\nu_i |C_i^{bid} D_i^{bid} - C_i^{ask} D_i^{ask}| + \sum_{i=1}^{I} \nu_i^2 |D_i^{bid^2} - D_i^{ask^2}|,$$

and are interested in finding values of θ such that $G(\theta) < \delta$. As was said before, the market price of a derivative is assumed to be the average of the bid-ask prices. For calculation of the error function we assume

$$\omega_i = \frac{1}{|C_i^{bid^2} - C_i^{ask^2}|} \quad , \quad \nu_i = \frac{1}{|D_j^{bid^2} - D_j^{ask^2}|}.$$

Table 3.1: Parameters of the first and second eigenvalues

Covariance Matrix	Mean-Reversion Speed	Asymptotic Eigenvalue	Eigenvalue Volatility	Initial Eigenvalue	Eigenvector Matrix	
1^{st} Eigenvalue	κ_1	θ_1	σ_1	λ_1	α_{11}	α_{12}
2^{nd} Eigenvalue	κ_2	θ_2	σ_2	λ_2	α_{21}	α_{22}

Initial Distribution

We assume the initial probability μ_0 is the truncated Gaussian measure with mean θ_0 and identity covariance matrix on the set E

$$\mu_0 \sim a\mathbb{I}_E \mathcal{N}(\theta_0, \mathbb{I}_m).$$

a is a normalization constant and $\theta_0 = (\kappa_{1,0}, \theta_{1,0}, \sigma_{1,0}, \lambda_{1,0}, \kappa_{2,0}, \theta_{2,0}, \sigma_{2,0}, \lambda_{2,0}, \zeta_0)$. We set the initial parameters $\theta_{i,0}$ and $\lambda_{i,0}$ as the one year at-the-money implied volatility derived from the market prices of the i^{th} company. We draw N samples $X_0 = \{\theta_1, \theta_2, \ldots, \theta_N\}$ from E independently with distribution μ_0 as our initial population in the first step of the algorithm.

Mutation Kernel

We use the following mutation kernel for our purpose:

$$\begin{aligned} M(x, dy) &= M(x,x)\delta_x(dy) + \mathbb{I}_E \frac{\exp\left(\frac{-1}{2}(y-x)' A^{-1}(y-x)\right)}{\sqrt{(2\pi)^n |A|}} dy, \\ M(x,x) &= (1 - N(x, \mathbb{I})[E]). \end{aligned}$$

$M(x,x)$ is the probability that mutation leaves the point x unchanged. To check that the mutation kernel $M(x, dy)$ is mixing, one can use the above equations to write

$$M(x, \mathcal{G}) = M(x, x)\mu(\mathcal{G} \cap E) + N(x, \mathbb{I})[\mathcal{G} \cap E],$$

where $\mu(.)$ denotes the Lebesgue's measure. Since all the Borel sets can be approximated from below by closed sets, assuming a closed set $\mathcal{G}$, it can be seen that

$$\begin{aligned} \inf_{x,y|M(y,y)\neq 0} \frac{M(x,x)}{M(y,y)} &> 0, \\ \inf_{x,y,\mathcal{G}|\mu(\mathcal{G})\neq 0} \frac{N(x,\mathbb{I})[\mathcal{G} \cap E]}{N(y,\mathbb{I})[\mathcal{G} \cap E]} &> 0, \end{aligned}$$

which implies that the mutation kernel $M(x, dy)$ is mixing by setting ϵ in the definition of the mixing property as

$$\epsilon = \min \left(\inf_{x,y|M(y,y)\neq 0} \frac{M(x,x)}{M(y,y)}, \inf_{x,y,\mathcal{G}|\mu(\mathcal{G})\neq 0} \frac{N(x,\mathbb{I})[\mathcal{G} \cap E]}{N(y,\mathbb{I})[\mathcal{G} \cap E]} \right).$$

The mixing property of the mutation kernel assures us that the algorithm searches all the space for the possible optimizer parameters and it does not get stuck in a local minimum area.

Crossover Outputs

Crossover is not necessary for the convergence of the algorithm but can reduce the convergence time. Considering two parameter sets θ_1 and θ_2, crossover produces a third parameter set $\bar{\theta}$ such that

$$\bar{\theta} = \frac{G(\theta_2)}{G(\theta_1) + G(\theta_2)}\theta_1 + \frac{G(\theta_1)}{G(\theta_1) + G(\theta_2)}\theta_2.$$

Therefore the parameter with the less corresponding error value will be closer to the newly generated parameter set $\bar{\theta}$. Because the search space E is convex, it is closed under the crossover operation and therefore $\bar{\theta}$ will be in E provided that $\theta_1, \theta_2 \in E$.

Selection Criteria

For our purpose we use the following selection kernel

$$S_t^n(x, dy) = e^{-\beta_n G(x)}\delta_x(dy) + (1 - e^{-\beta_n G(x)})\frac{e^{-\beta_n G(x)}\mu(dy)}{\int \mu(dz)e^{-\beta_n G(x)}}.$$

The selection kernel is constructed in such a way to guarantee that the population of the parameters concentrate on the minimum of the error function by allowing β_n approach to infinity. According to the selection kernel at the n^{th} step, among the population parameters $\theta_1^n, \theta_2^n, \ldots, \theta_N^n$, the individual θ_j^n will be selected to remain in the pool with probability $\exp(-\beta_n G(\theta_j))$. We increase the selection pressure by assuming $\beta = n^\alpha$ for some $\alpha \in (0, 1)$.

Convergence of the Algorithm

Now we prove the following proposition which guarantees the convergence of our algorithm to the global minima of G:

Proposition 8. *With the above assumptions and definitions of the error function G and the distribution of the population μ_n, we have the following results*

a) The function G is continuous, bounded and of bounded oscillations.

b) The truncated Gaussian kernel $M(x, dy)$ is mixing.

c) $G^* = \inf_K \{\int_E G(\theta)\mu(d\theta); \mu \in M_1(E)$ with $I(\mu) < \infty\} = \inf_E G$, where the infimum is taken over all Markov Kernels $K(x, dy)$ with stationary measure μ.

d) For all $\delta > 0$

$$\mu_n(G(\theta) \geq \inf_E G + \delta) \xrightarrow{n\uparrow\infty} 0.$$

Proof. The proof is similar to the proof of Proposition 2 in Hamida and Cont (2004), some modifications are needed for the special case of stochastic covariance.

a) The prices of the derivatives we use for the algorithm are continuous functions of the parameters. This can be seen either from the results regarding the continuity of the solutions of parabolic partial differential equations with respect to their coefficients and initial values (see Ladyzhenskaya *et al.* (1968)), or simply from the closed form solutions we derive for the price of derivatives. Besides, by arbitrage arguments and because the call options can not be more expensive than the underlying stock, it turns out that the error function is bounded on the compact set E and therefore it is of bounded oscillations.

b) It is obvious.

c) As we showed before $M(x, x) > 0$ for all $x \in E$. Therefore by Moral, G^* coincides with the essential minimum of the error function with respect to π, the invariance measure of the irreducible kernel M. Then it can be proved that π is absolutely continuous with respect to the Lebesgue's measure and therefore

$$G^* = \inf_E G.$$

d) Comes from a), b), c) and proposition 1 in Hamida and Cont (2004).

□

Empirical Results

General Motors is one of the biggest news makers in the credit market which holds significant pension liabilities. In May 2005, General Motors' stock fell more than 20% and its CDS spread widened significantly. In addition S&P downgraded General Motors from investment grade, BBB, to sub-investment grade, BB. One of the difficulties in estimating the credit risk of General Motors with structural credit risk models is in the level of its liabilities. General Motors is the biggest corporate bond issuer in North America which issues %80 of its bonds through its financial services subsidiary, General Motors Acceptance Corporation. General Motors Acceptance Corporation acts like a bank and most of its issued bonds are secured. Therefore, only 20% of General Motors total debts contribute to its implied debt per share. Stamicar and Finger (2005) state that for the above reasons, the model debt per share for General Motors should be 20% of its overall debt value. Ford is the second largest automobile producer in the world. Using structural models for the credit risk analysis of Ford is challenging for the same reason as in the above argument for General Motors. It is very difficult to estimate the true level of the liabilities of the company. Similarly to General Motors, majority of the financial liabilities of the Ford company is related to its financial subsidiary, Ford Motors Credit Company. In May 2005, S&P downgraded Ford from investment grade BBB to sub-investment grade BB. During this period the correlation between the CDS spreads of the General Motors and Ford jumped to a high value but it quickly returned to its normal mean. Stamicar and Finger (2005) relate the credit rating decline of the Ford Motors to several factors including Ford's declining market share, high health care costs, slide in sales of sports utility vehicles, and growing competition between the auto-makers. The period May-March 2005 is an indication of the nonconstant correlation between the two big auto-maker companies and implies that during crisis periods default correlation increases.

In this section we perform some numerical tests on our model using real data from General Motors and Ford. We use the data from 120 options on General motors' equity and 64 options on Ford's equity. The data are given in tables A.1 and A.4 in the appendix A. First we perform the algorithm on the General Motor's equity options only to compare our model with the CreditGrades model with Heston stochastic volatility proposed in

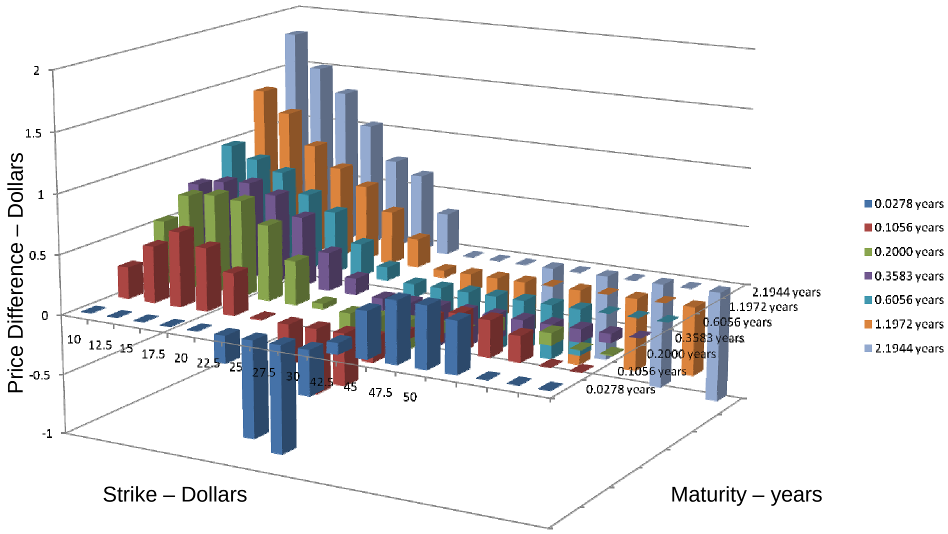

Figure 3.2: The difference between market prices and stochastic correlation model's results for General Motors's equity options in U.S. dollars as a function of strike price and maturity date. The maturity dates range from $T = 0.0278$ years to $T = 2.1944$ years. The strike prices range from $K = 10$ dollars to $K = 50$ dollars. At the time of analysis, The value of General Motors' equity is $S_0 = \$25.86$

Sepp (2006). For this reason, in the error function (3.18), we let the Ford's weights be $\nu_i = 0$

$$G(\theta) \quad = \quad \sum_{i=1}^{I} \omega_i^2 |C_i^{\theta}(t, T_i) - C_i^*(t, T_i)|^2, \tag{3.19}$$

so that the error term accounts for the General Motors equity market information only. After running the evolutionary algorithm, we get the optimizing parameter vector as

$$\theta = (12.2334, 0.0755, 2.0904, 0.0869, 11.2558, 0.0490, 6.6987, 0.1269, 0.0060).$$

An interesting observation is that in the absence of the second and third terms in the error function (3.18), the parameter $\zeta = 0.0060 \simeq 0$. Remember that the parameter ζ

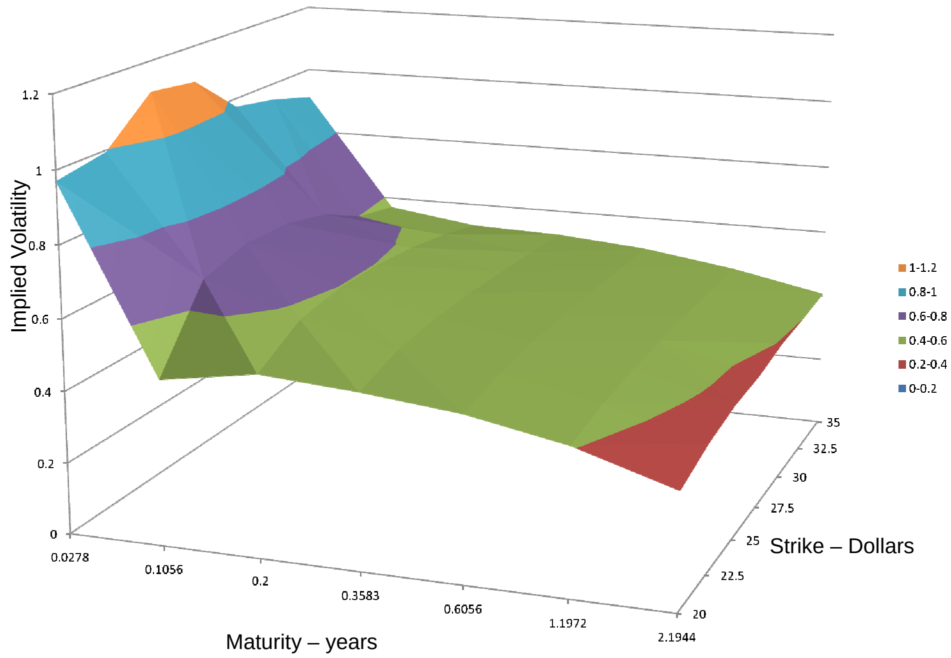

Figure 3.3: Implied Volatility surface of the stochastic correlation model

is the angle that the first eigenvector makes with the x-axis. Therefore the eigenvector matrix is given by

$$E = \begin{pmatrix} \cos(\zeta) & -\sin(\zeta) \\ \sin(\zeta) & \sin(\zeta) \end{pmatrix} = \begin{pmatrix} 0.9999 & -0.0060 \\ 0.0060 & 0.9999 \end{pmatrix} \simeq \mathbb{I}_2,$$

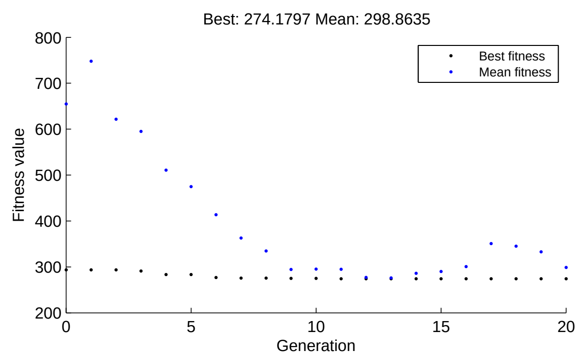

Figure 3.4: Convergence of the evolutionary algorithm for parameters of the principal component model for General Motors and Ford's equity options

which implies that the eigenvector matrix is very close to the identity matrix. This implies that when calibrating the marginal factors, the calibrated two factor stochastic correlation model is close to the one factor Heston stochastic volatility model introduced in Sepp (2006). The error terms regarding the option prices are shown in figure 3.2. In Sepp (2006), figures 7.1, 7.2 and 7.3 show the corresponding error terms for the three models studied in the paper, as a function of strike and maturity. Our model shows improvements with respect to the regular diffusion model, the stochastic variance model and the double-exponential jump diffusion model. The stochastic correlation model gives a slightly better fit to the market data compared to the single factor stochastic volatility model introduced in Sepp (2006). We have an error term about 5% less than the regular diffusion model and 2% less than the stochastic variance model and the double-exponential jump diffusion model. The figure 3.3 shows the implied volatility surface of our model with respect to strike and maturity. As it can be seen the stochastic correlation model is capable of reproducing the volatility smile and skews.

Now consider the complete error function given in equation (3.18)

$$
\begin{aligned}
G(\theta) &= \sum_{i=1}^{I} \omega_i^2 |C_i^\theta(t,T_i) - C_i^*(t,T_i)|^2 \\
&+ \sum_{i=1}^{I} \omega_i \nu_i |C_i^\theta(t,T_i).D_i^\theta(t,T_i) - C_i^*(t,T_i).D_i^*(t,T_i)| \\
&+ \sum_{i=1}^{I} \nu_i^2 |D_i^\theta(t,T_i) - D_i^*(t,T_i)|^2.
\end{aligned}
\tag{3.20}
$$

This function includes the error terms regarding the General Motors and Ford's equity prices altogether. The first line represents the error term with respect to General Motors equity options, the second term represents the joint behavior of the price differences, and the third term represents the error term with respect to Ford Motors equity options. After running the evolutionary algorithm with the above error function, the convergence of the algorithm is shown in figure 3.4. We get the solution of the algorithm as

$$\theta = (3.1041, 0.0144, 6.6155, 0.0293, 8.5067, 0.1600, 5.6187, 0.0760, -0.9091).$$

This time the parameter ζ which represents the angle that the first eigenvector makes with the x-axis is not close to zero and therefore the eigenvector matrix is far from identity and is given by

$$E = \begin{pmatrix} \alpha_{11} = 0.614 & \alpha_{12} = 0.789 \\ \alpha_{21} = -0.789 & \alpha_{22} = 0.614 \end{pmatrix}.$$

Table 3.2: Parameters of the first and the second eigenvalues

Covariance Matrix	Mean-Reversion Speed	Asymptotic Eigenvalue	Eigenvalue Volatility	Initial Eigenvalue	Eigenvector Matrix	
1^{st} EV	$\kappa_1 = 3.1049$	$\theta_1 = 0.0141$	$\sigma_1 = 6.6152$	$\lambda_1 = 0.0291$	$\alpha_{11} = 0.6142$	$\alpha_{12} = 0.7891$
2^{nd} EV	$\kappa_2 = 8.5065$	$\theta_2 = 0.1607$	$\sigma_2 = 5.6188$	$\lambda_2 = 0.7660$	$\alpha_{21} = -0.789$	$\alpha_{22} = 0.6144$

Table 3.2 shows the final results of the estimation algorithm. Because the eigenvector matrix is not close to identity, the effects of both CIR eigenvalue processes are more involved in the dynamics of the assets of both companies. We should also mention that since the CIR processes are not additive (the sum of two CIR processes is not a CIR process), the marginal volatility structure of each company's asset is different from the single factor CreditGrades model with Heston volatility structure.

3.6 Conclusion

We give a stochastic correlation structural credit risk model. The eigenvalues of the covariance matrix are driven by CIR processes and the eigenvectors are constant. In this framework, we priced equity options by using a combination of Fourier transform and method of images. We extend this methodology to calculate the probabilities of default for the companies. We estimate the parameters of the model by evolutionary algorithms. For the purpose of the calibration, we considered two companies : General Motors and Ford Motors. We used data from 200 options on the equities of the two companies. For future research, we are interested to extend the empirical results we gained here to credit portfolios with more firms. Studies show that usually two or three eigenvalues of the covariance matrix are enough to capture the majority of the information regarding the covariance matrix. In that case, we would be using only two or three CIR processes with their corresponding eigenvectors. This gives a very small set of parameters which makes the calibration problem less involved. As another future research perspective, we are also interested in implementing the stochasticity in correlation using other tools. For this purpose, one can either use another credit risk model or another dynamics for the covariance matrix.

Chapter 4

CreditGrades Model with Stochastic Covariance Wishart Process

We present a structural credit risk model which considers stochastic correlation between the assets of the companies. The covariance of the assets follows a Wishart process, which is an extension of the CIR model to dimensions greater than one. Wishart process is an affine symmetric positive definite process. Bru (1991), Gourieroux *et al.* (2004) and Fonseca *et al.* (2007a, 2006, 2007b) brought the finance community's attention to this process as a natural extension of Heston's stochastic volatility model, which has been a very successful univariate model for option pricing and reconstruction of volatility smiles and skews. The popularity of the Heston model could be a promising point for using Wishart process to model stochastic correlation. We will use the results discovered so far about the Wishart process in credit risk modeling. We give closed form solutions for equity options, marginal probabilities of defaults, and some other major financial derivatives. For calculation of our pricing formulas we make a bridge between two recent trends in pricing theory; from one side, pricing of barrier options by Lipton (2001) and Sepp (2006) and from other side the development of Wishart process by Bru (1991), Gourieroux *et al.* (2004) and Fonseca *et al.* (2007a, 2006, 2007b). In the second part of the paper, we estimate the parameters of the model by evolutionary algorithms. We prove a theorem which guarantees the convergence of the evolutionary algorithm to the set of optimizing parameters. The parameters are found by searching through a compact bounded set $E \subset \mathbb{R}^n$. An initial population is selected from the parameter space E and then the individuals go under a mutation-selection procedure to search for the best parameters which minimize the error function.

4.1 Introduction

The first mathematical result for valuation of options in the presence of stochastic volatility was given by Heston (1993). In that article closed-form formulas for the characteristic function and the probability distribution of the joint logprice-volatility process $(\ln A_t, \lambda_t)$ were obtained. The results in that paper also give a closed-form solution for the valuation problem of a vast range of popular financial derivatives on A_t. Implementation of the model has been done with direct integration, FFT and fractional FFT methods assuming the asset is observable (see Kilin (2007)). Even though there are numerous research papers on stochastic volatility and modeling the financial instruments with Levy processes, there is not much work done regarding stochastic correlation. Market data implies that the correlation between financial instruments is not constant and changes stochastically over time. Modeling stochastic correlation has difficulties from the analytical as well as the estimation point of view. We think this is the main reason that very little has been done so far regarding stochastic modeling of the joint behavior of financial assets. One of the attempts to fix this gap began with a paper on Wishart processes by Bru (1991), which followed by a series of papers by Gourieroux *et al.* (2004). The Wishart process is a positive-definite symmetric matrix which satisfies a certain stochastic differential equation. The Wishart process is a natural extension of the Cox-Ingersoll-Ross process. Several authors have recently brought the finance community's attention to the Wishart process and showed that the Wishart process is a good candidate for modeling the covariance of assets. Risk is usually measured by the covariance matrix. Therefore Wishart process can be seen as a tool to model dynamic behavior of multivariate risk. Bru (1991) proposes the Wishart process as a generalized squared Bessel process for dimensions greater than one. She then verifies existence and uniqueness, additivity property, first hitting time of the smallest eigenvalue and the distributions of the Wishart process. Gourieroux *et al.* (2004) then uses the mathematical results from Bru and suggest the Wishart process as an extension of the CIR process. Via the Laplace transform and the distribution of the Wishart process, Gourieroux prices derivatives with a Wishart stochastic covariance matrix. This approach can be used to model risk in the structural credit risk framework but has some drawbacks in the estimation side of the problem. The Wishart process models the covariance process in a way that the marginal and joint parameters are mixed and therefore this is very difficult to calibrate marginal and joint parameters separately, which is the usual method by most of the existing estimation techniques in the literature. Gourieroux *et al.* (2004) give a discrete version of the Wishart process known as the Wishart Autoregressive (WAR) process. The WAR process is specifically useful for simulation purposes for Wishart processes with integer degrees of freedom. Gourieroux show that both Wishart process and WAR process preserve the analytical tractability property of the CIR process. In fact these two models are both affine models which yield closed formulas for most popular financial derivative in the model. Gourieroux results have been continued in a series of papers by Fonseca *et al.* (2006, 2007b,a). The results of Fonseca *et al.* (2006) uses a multi-factor Heston model to show the flexibility of the Wishart process to capture volatility smile and skew. One of the main advantages of this paper is the introduction of a correlation structure between stock's noise and the volatility's noise. Moreover, Fonseca focuses on the role of stochastic correlation brought by the Wishart process. He gives closed formulas for the stochastic correlation between several factors of the model. Fonseca extends his approach in Fonseca *et al.* (2007b) to model the multivariate risk by the Wishart process. In this paper several risky assets are considered and the pricing problem for one dimensional vanilla options and multidimensional geometric basket options on the assets are solved. The results show the consistency of the Wishart stochastic covariance model with the smile and skew effects observed in the market. Even though we know the analytical properties of the Wishart process, little has been done on the calibration of the parameters of the Wishart process. Fonseca *et al.* (2007a) estimates the Wishart Stochastic Correlation Model by using the stock indexes SP500, FTSE, DAX and CAC40 under the historical measure. Because of the different nature of our problem, we use a different method to calibrate the parameters of the model under the risk neutral measure.

Merton's model Merton (1974) is the first structural credit risk model proposed which considers the company's equity as an option on the firm's asset. There has been numerous extensions for the Merton's model in the literature including incorporating early defaults, stochastic interest rates, stochastic default barriers and jumps in the asset's price process. One of the drawbacks of the Merton's model is that it provides no connection betweem credit risk and the equity markets. This connection has been the main objective in introducing the CreditGrades model. The CreditGrades model was jointly developed by CreditMetrics, JP Morgan, Goldman Sachs and Deutsche Bank. The original version of the CreditGrades model assumes that volatility is deterministic. We extend the CreditGrades model, using stochastic covariance Wishart process focusing on the role of stochastic correlation. The performance of a company is usually monitored by observing its equity's volatility or the CDS spread. CreditGrades model can be considered as a down-and-out barrier credit risk model. This means that default is triggered if the value of the asset reaches a certain level identified by the recovery part of the debt. Stamicar and Finger (2005) has extended the CreditGrades model to price equity options by introducing the equity as a shifted log-normal process. Sepp (2006) has extended Stamicar's idea by embedding the Heston's volatility into the model and pricing equity derivative. His calibrated model captures the market data better than the constant volatility CreditGrades model used by Stamicar. Both Stamicar and Sepp models are univariate credit risk models. We extend the CreditGrades model by use of the Wishart processes to dimensions greater than one implementing stochastic correlation into the dynamics of the assets. We give closed formulas for equity derivatives based on our Stochastic Correlation model. Finally we use an evolutionary algorithm to calibrate the parameters of the model. Fonseca *et al.* (2007b, 2006) introduces a correlation structure between dZ_t and dW_t and gives a closed formula for the characteristic function of the joint price-volatility process resulting in pricing non-exotic derivatives like European call-put options. In the framework of CreditGrades or the first passage time model, products like equity call options are evaluated with the assumption that the company might default prior to the maturity of the option which affects the payoff at maturity. This implies the need to be able to value barrier options. However, absorbing probabilities are not available where there is a non-zero correlation between the assets and their volatilities even in the univariate case (see Lipton (2001) and Sepp (2006)). Therefore we follow Gourieroux *et al.* (2004) and assume no correlation structure between the assets and their volatilities.

The Dynamic Conditional Correlation Model introduced by Engle (2002) is a new class of multivariate GARCH model which is considered as one of the generalizations of the Bollerslev's Constant Conditional Correlation Model. The Dynamic Conditional Correlation Model assumes the return of the assets follows a normal distribution $r_t|\mathcal{F}_t \sim \mathcal{N}(0, H_t)$, where $H_t = D_t R_t D_t$ is the covariance matrix, $D_t = diag(\sqrt{h_t^i})$ is the diagonal matrix of standard deviations and R_t is the correlation matrix. h_t^i's follow a univariate GARCH dynamics and the dynamics of the correlation R_t is given by

$$\begin{cases} Q_t &= (1-\alpha-\beta)\bar{Q} + \alpha\epsilon_{t-1}\epsilon_{t-1}^{'} + \beta Q_{t-1}, \\ R_t &= Q_t^{*-1} Q_t Q_t^{*-1}, \end{cases} \tag{4.1}$$

with $Q_t^* = diag(\sqrt{q_{ij}})$ and $\epsilon_t \sim \mathcal{N}(0, R_t)$. The Dynamic Conditional Correlation Model is usually calibrated in two steps. In the first step, the volatilities are calibrated and in the second step, the correlation parameters are estimated. This can not be done in the case of Wishart process, because the joint and marginal information are shared among the same parameters of the model. Nelson (1990) presents a discrete time model that converges to a continuous time stochastic volatility model and uses that to estimate the parameters. Fonseca *et al.* (2007a) approximate the Wishart process by a sequence of discrete time processes satisfying a stochastic difference equation which converge in probability to the Wishart process. This approach, as pointed out by Nelson (1990) and Fonseca *et al.* (2007a) does not guarantee consistent estimators of the continuous time model.

There are numerous approaches for estimating the parameters of a model with hidden parameters. One of the approaches recently suggested is the evolutionary algorithm which is a search process using mutation-crossover-selection cycles. The Heston Model has four parameters κ, θ, σ and v_0 and each parameter has its own economic interpretation. Even though Wishart process is the extension of the CIR process, it does not give a clear interpretation of its parameters. This is the main difficulty in calibrating the model.

In section 4.2 we explain the market behavior of the correlation process and we introduce a stochastic differential equation suitable for capturing these properties. In section 4.3 we suggest Wishart process as a candidate to model the covariance matrix of the assets' prices. We then talk about the Wishart autoregressive process which is the discrete version of the Wishart process in section 4.4. We introduce our model for the assets' prices in section 4.5 and solve the pricing problem for some derivatives on the equities in section 4.6. We end this chapter by suggesting an evolutionary algorithm method to calibrate the parameters of the model in section 4.7.

4.2 Modeling Correlation as a Stochastic Process

It is not difficult to construct a pair of Brownian motions with a given correlation structure. Assume that V_t, W_t and K_t are independent Brownian Motions. Define ρ_t as

$$d\rho_t = a(\rho_t)dt + b(\rho_t)dK_t,$$

and define the Brownian motion Z_t as:

$$Z_t = \int_0^t \rho_t dV_t + \int_0^t \sqrt{1-\rho_t^2}dW_t.$$

The fact that Z_t is a Brownian Motion is obvious. One can also check that

$$E(Z_t V_t) = \int_0^t \rho_t du.$$

This is an obvious generalization of the constant correlation case.

There are two approaches presented in Emmerich (2006) for modeling the correlation as a stochastic process. The first one is by using a suitable transformation which maps

$$X_t : (-\infty, +\infty) \to (-1, +1).$$

A suitable candidate would be $X_t = \frac{2}{\pi}Arctan(\alpha(W_t + \gamma))$. This approach is not intuitive and is hard to calibrate. Another approach which seems more interesting is to model correlation as the solution of a nonlinear stochastic differential equation

$$d\rho_t = \kappa(\theta - \rho_t)dt + \sqrt{1-\rho_t^2}dK_t. \tag{4.2}$$

Emmerich (2006) has made efforts to force conditions on the parameters of (2.1) such that the process ρ_t fits the best with the data gathered from the market. There are two important observations which are cited in the literature and observed from market data. One is that the mass in the boundaries 1 and +1 are almost zero. And the second one is that we expect the correlation to stay around its mean in the long run.

We now examine the analytical properties of the boundaries of (4.2). Suppose l and r are the boundaries of a general time-homogenous diffusion process

$$dX_t = a(X_t)dt + b(X_t)dW_t.$$

Then define the following functions based on the parameters a and b

$$\begin{aligned}
s(v) = \quad & \exp\left(-\int_{v_0}^{v}\frac{2a(u)}{b^2(u)}du\right), \\
S(x) = \int_{x_0}^{x} s(v)dvdS(x) \quad , \quad & S[c,d] = \int_c^d s(v)dv, \\
S(l,x] = lim_{u\to l^+}S[u,x] \quad , \quad & m(x) = \frac{1}{b^2(x)s(x)}, \\
M[c,d] = \int_c^d m(x)dx \quad , \quad & \Sigma(l) = \int_l^x M(l,x]dx.
\end{aligned}$$

We say that the left boundary l is *attractive* if there exists an x such that

$$S(l,x] < \infty,$$

and *attainable* if

$$\Sigma(l) < \infty.$$

For the stochastic differential equation (4.2), it has been shown in Emmerich (2006) that the left boundary -1 is attractive and attainable if $\frac{\kappa}{\alpha^2}(1+\theta) < 1$, and right boundary $+1$ is attractive and attainable if $\frac{\kappa}{\alpha^2}(1-\theta) < 1$. Therefore for the correlation to be unattainable near the boundaries -1 and $+1$ we must have

$$\kappa \geq \frac{\alpha^2}{1 \pm \theta},$$

which is the expectation we have from the behavior of the correlation near the boundaries.

The transition probability of (4.2) is given by the Fokker-Planck equation. Emmerich (2006) has calculated the stationary transition probability of (4.2) and shows that it matches the market behavior well. As a general well-known result, the transition probability of the process

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t,$$

should satisfy the PDE

$$\frac{\partial}{\partial t}(p(t,x)) + \frac{\partial}{\partial x}(a(t,x)p(t,x)) + \frac{\partial^2}{\partial x^2}(b^2(t,x)p(t,x)) = 0.$$

Now we find the transition density of the solution of the stochastic equation proposed for the correlation

$$d\rho_t = \kappa(\theta - \rho_t)dt + \alpha\sqrt{1-\rho_t^2}dK_t. \tag{4.3}$$

It can be proved that the transition probability for the above mean reverting process is convergent

$$p(x) = \lim_{t \to \infty} p(t,x).$$

For simplicity, suppose that $\alpha = 1$. In the simple case $\theta = 0$, $p(t,x)$ satisfies the *stationary* Fokker-Planck equation:

$$(1-\kappa)p(x) + x(2-\kappa)p'(x) - \frac{1}{2}(1-x^2)p''(x) = 0,$$

with two conditions

$$\int_{-1}^{1} p(x)dx = 1,$$

$$\int_{-1}^{1} xp(x)dx = 0.$$

The first condition is to ensure that $p(x)$ is actually a probability measure and the second condition is the mean perseverance assumption of the limit probability $p(x)$. The solution under the assumption $\theta = 0$ is given by the proposition below

Proposition 9. *The stationary transition density for the equation* (4.3) *with* $\theta = 0$ *is given by*

$$p(x) = c(1 - x^2)^{(\kappa - 1)}.$$

c is a constant given by $c = \frac{1}{2.F(\frac{1}{2}, 1-\kappa, \frac{3}{2}, 1)}$ *where F is the hyper-geometric function*

$$F(a, b, c, y) = 1 + \frac{ab}{c}x + \frac{a(a+1)b(b+1)}{2!c(c+1)}x^2 + \dots.$$

Proof. See Emmerich (2006). □

Assuming $\theta \neq 0$, $p(x)$ is the solution of the ODE

$$(1 - \kappa)p(x) + x(2 - \kappa)p'(x) + \kappa\theta p'(x) - \frac{1}{2}(1 - x^2)p''(x) = 0.$$

with two conditions

$$\int_{-1}^{1} p(x)dx = 1,$$
$$\int_{-1}^{1} xp(x)dx = 0.$$

The solution of the above ODE is given by the following proposition

Proposition 10. *The stationary transition density for the equation* (4.3) *with* $\theta \neq 0$ *is given by*

$$p(x) = c\left(\frac{1-x}{1+x}\right)^{\kappa\theta}(1 - x^2)^{\kappa - 1},$$

where c is again a constant found by the condition $\int_{-1}^{1} p(x)dx = 1$.

Proof. See Emmerich (2006). □

Considering above statements about the density function $p(x)$, we can see that as we expected:

- p is concentrated on $[-1, 1]$.
- p is symmetric with respect to θ. This implies that $p^{\theta}(x) = p^{-\theta}(-x)$.
- If $\theta = 0$ the global maximum is attained at $x = 0$.
- The probability measure approaches zero around the boundaries.

Therefore, Emmerich's model (4.2) seems to be a suitable one for modeling the correlation. In what follows, we will show that even though the model (4.2) can capture the correlation implied from the market well, it is not suitable to model correlation between assets. One of the main drawbacks of Emmerich's approach is its analytical non-tractability. In the next two sections, we introduce Wishart process, and its discrete version, Wishart Autoregressive process, as a suitable candidate for modeling the joint stochastic volatility and stochastic correlation. Wishart process and Wishart Autoregressive process both entail affine analytical properties which allow us to derive close formulas for the prices of a vast range of financial derivatives.

4.3 The Wishart Process

4.3.1 Dynamics of the Stochastic Volatility Matrix: Existence, Uniqueness and Analytic Properties

In this section we model the covariance between assets as a stochastic Wishart process. Our model allows correlation and volatility be stochastic. Heston (1993) chooses the CIR process[1] to model volatility of the asset. One of the nice properties of his model is its analytical tractability which makes the calibration of the hidden parameters of the model less involved. Gourieroux *et al.* (2004) has introduced Wishart process as a direct multivariate extension of CIR process. The standard Heston model is known to fit market data better than the Black and Scholes model in capturing the volatility smile and skew. But this model still seems to lack enough flexibility to capture the market behavior. Fonseca *et al.* (2006) shows that a multi-factor Heston model captures the volatility smiles and skews even better, thanks to structural complexity of Wishart process. By considering the stochastic covariance Wishart process, we have more flexibility and degree of freedom in the marginal, while analytic tractability is preserved when extending CIR process to Wishart process. We first present Wishart process of integer degree of freedom and

[1]CIR Processes are of the forms $d\lambda_t = \kappa(\theta - \lambda_t)dt + \alpha{\lambda_t}^{\zeta}dW_t$ when $\zeta = \frac{1}{2}$. The $\frac{3}{2}$ model is similar to the Heston model, but assumes that the randomness of the variance process varies with $\zeta = \frac{3}{2}$

derive their matrix stochastic differential equation which later on will give a natural representation of Wishart process with fractional degree of freedom. A Wishart process with integer degree of freedom K is a sum of K independent n-dimensional Ornstein-Uhlenbeck process. We give the formal definition below

Definition 1. *Consider $\{U^{(k)}\}_{k=1}^{K}$ as an independent set of Ornstein-Uhlenbeck processes*

$$dU_t^{(k)} = AU_t^{(k)}dt + QdW_t^{(k)},$$

where $A \in M_n$ and $Q \in GL(n)$. Then a Wishart process of degree K is defined as

$$\Sigma_t = \sum_{k=1}^{K} U_t^{(k)} U_t^{(k)'},$$

where $U_t^{(k)'}$ is the transpose of the vector $U_t^{(k)}$.

Ito's lemma can be used to find a diffusion SDE for the process Σ_t

$$\begin{aligned} d\Sigma_t &= \sum_{k=1}^{K} \left(dU_t^{(k)}.U_t^{(k)'} + U_t^{(k)}.U_t^{(k)'} + dU_t^{(k)}.dU_t^{(k)'} \right) \\ &= (KQ'Q + A\Sigma_t + \Sigma_t A')dt + \sum_{k=1}^{K} \left(QdW_t^{(k)} U_t^{(k)'} + U_t^{(k)} dW_t^{(k)'} Q' \right). \end{aligned}$$

As it can be seen, the drift term of the SDE above contains Σ_t, but the diffusion part contains the terms $U_t^{(k)}$ and $U_t^{(k)'}$ separately. Bru (1991) and Gourieroux *et al.* (2004) show that Σ_t also satisfies the following matrix SDE

$$d\Sigma_t = (KQ'Q + A\Sigma_t + \Sigma_t A')dt + QdW_t^{(k)}\Sigma_t^{\frac{1}{2}} + \Sigma_t^{\frac{1}{2}} dW_t^{(k)'} Q',$$

where W_t is an $n \times n$ standard Brownian motion matrix. Assuming $n = 1$ in the above stochastic differential equation, one has

$$d\Sigma_t = (Kq^2 + m\Sigma_t)dt + 2q\sqrt{\Sigma_t}dW_t,$$

which is a CIR process. This shows that Wishart process is a multidimensional extension of the CIR process. There is an impact of joint dynamics on idiosyncratic behaviors, which is implemented in the Wishart Covariance Model. This effect can be clearly seen in the equation above when considered element-wise. Before Gourieroux *et al.* (2004) and

Fonseca *et al.* (2006, 2007b,a), this effect had not been documented in the multidimensional dynamic models. From the above representation of Σ_t, it can be seen that Σ_t has an affine structure. More precisely, straightforward calculation of the drift and volatility of Σ_t implies

$$\begin{cases} E(d\Sigma_t) = (KQ'Q + A\Sigma_t + \Sigma_t A')dt, \\ Cov(\alpha' d\Sigma_t \beta, \delta' d\Sigma_t \omega) = (4\alpha'\Sigma_t\beta\delta'Q'Q\omega)dt, \end{cases}$$

for α, β, δ and $\omega \in \mathbb{R}^n$, which implies the drift and quadratic variations of Σ_t are affine functions of Σ_t. For more detailed properties of the Wishart process we refer the reader to Bru (1991) and Gourieroux *et al.* (2004).

The new SDE above also defines a Wishart process with fractional degree of freedom. K is called the Gindikin coefficient of the Wishart process. If K is a positive real number, we have the extended definition of the Wishart process as below

Definition 2. *A Wishart process with degree of freedom β is a strong solution of the stochastic differential equation*

$$d\Sigma_t = (\Omega\Omega' + A\Sigma_t + \Sigma_t A')dt + QdW_t^{(k)}\Sigma_t^{\frac{1}{2}} + \Sigma_t^{\frac{1}{2}}dW_t^{(k)'}Q', \tag{4.4}$$

where $\Omega\Omega' = \beta Q'Q$.

To satisfy the condition $\Omega\Omega' = \beta Q'Q$, Without loss of generality one may assume $\Omega = \sqrt{\beta}Q'$. β is called the Gindikin's coefficient and Q is the volatility of the volatility matrix. The existence of the solution of the stochastic differential equation above has been established by a series of theorems in Bru (1991). We first present the first existence theorem which deals with a simplified stochastic differential equation and its eigenvalues

Theorem 10. *If $(B_t)_{t\geq 0}$ is a $p \times p$ Brownian motion, then for every $p \times p$ symmetric matrix $s_0 = (s_{ij}(0)) \in \zeta_p^+$ with distinct eigenvalues labeled*

$$\lambda_1(0) > \ldots > \lambda_p(0) \geq 0.$$

The stochastic differential equation

$$dS_t = \sqrt{S_t}dB_t + dB_t'\sqrt{S_t} + \alpha\mathbb{I}dt,$$

has (1) a unique solution in ζ_p^+ (in the sense of probability law) if $\alpha \in (p-1, p+1)$ and (2) a unique strong solution in ζ_p^+ if $\alpha \geq p+1$. The eigenvalues of such a solution never *collide almost surely, for all* $t \geq 0$, $\lambda_1(t) > \ldots > \lambda_p(t) \geq 0$ *with* $\lambda_p(t) > 0$ *if* $\alpha \geq p+1$ *and satisfy the stochastic differential system*

$$d\lambda_i = 2\sqrt{\lambda_i}d\nu_i + \left(\alpha + \sum_{k\neq i} \frac{\lambda_i + \lambda_k}{\lambda_i - \lambda_k}\right) dt \qquad 1 \leq i \leq p,$$

where $\nu_1(t), \ldots, \nu_p(t)$ *are independent Brownian motions.*

Proof. See Bru (1991) □

The proof is based on the results from Ikeda and Watanabe (1981). Bru then gives a second existence theorem with more general drift and diffusion structure for the stochastic differential equation driving the Wishart process

Theorem 11. *If* $(B_t)_{t\geq 0}$ *is a* $p \times p$ *Brownian motion,* $\alpha \geq p+1$, a *is in the group* $GL(p)$, $b \in \zeta_p^-$, s_0 *is in* ζ_p^+ *and has all its eigenvalues distinct, and* τ *is the collision time, then on* $[0, \tau)$ *the stochastic differential equation*

$$dS_t = \sqrt{S_t}dB_t\sqrt{a'a} + \sqrt{a'a}dB_t'\sqrt{S_t} + (bS_t + S_t b)dt + na'adt \qquad S_0 = s_0,$$

has a unique solution if b *and* $\sqrt{a'a}$ *commute.*

Proof. See Bru (1991) □

From the structure of the SDE, it is obvious that Σ_t is symmetric. Assuming $\Omega\Omega' = \beta Q'Q$ and $\beta > n-1$, one can also show that Σ_t is positive definite(see Bru (1991)). Moreover, the Wishart process because of its drift term has mean reverting properties. If the asymptotic limit of the volatility is Σ_∞, then it satisfies

$$-\Omega\Omega' = M\Sigma_\infty + \Sigma_\infty M'.$$

Wishart process defined by Definition 2, with $\beta > n-1$, M negative definite and $Q \in GL(n)$, is an affine symmetric positive definite process. Bru (1991), Gourieroux *et al.* (2004) and Fonseca *et al.* (2007a, 2006, 2007b) have brought the finance community's attention to this process as a natural extension of the Heston's stochastic volatility model, which has been a very successful univariate model for option pricing and reconstruction of volatility smiles and skews. Popularity of the Heston model could be a promising point for using Wishart process to model stochastic correlation. What we will do in what follows is to take the next step and use what has been discovered so far about the Wishart process in credit risk modeling.

4.3.2 Parameter Interpretation when n = 2

If ρ_t is the implied correlation between the assets, it satisfies the following SDE

$$d\rho_t = (A_t\rho_t^2 + B_t\rho_t + C_t)dt + \sqrt{1-\rho_t^2}D_t dN_t, \tag{4.5}$$

where A_t, B_t and C_t are functions of Σ_{11}, Σ_{22} and parameters of the Wishart process, D_t is a stochastic term and a function of Σ_t and the parameters and N_t is a noise term. The expression for A_t, B_t, C_t and D_t is given below.

One useful observation is to compare (4.5) with equation (4.2) given by Emmerich (2006) as a stochastic dynamics for the correlation process. When ρ_t is low and close to zero, the quadratic term is negligible and the mean reversion behavior of (4.5) would be close to the mean reversion behavior of (4.2). In what follows we will show that $B_t < 0$, which is intuitively anticipated because of the mean reversion property of the correlation process.

In the two dimensional case the Wishart process has the matrix form

$$\Sigma(t) = \begin{pmatrix} \Sigma_{11}(t) & \Sigma_{12}(t) \\ \Sigma_{21}(t) & \Sigma_{22}(t) \end{pmatrix},$$

where $\sqrt{\Sigma_{11}(t)}$ and $\sqrt{\Sigma_{22}(t)}$ are volatilities of the first and the second assets respectively and the correlation between the two assets is given by

$$\rho_t = \frac{\Sigma_{12}(t)}{\sqrt{\Sigma_{11}(t)\Sigma_{22}(t)}}.$$

Here we expand the Wishart's stochastic differential equation in the case of $n = 2$. Let's denote the symmetric square root of the covariance matrix by $\sigma(t)$

$$\begin{pmatrix} \sigma_{11}(t) & \sigma_{12}(t) \\ \sigma_{12}(t) & \sigma_{22}(t) \end{pmatrix} = \begin{pmatrix} \Sigma_{11}(t) & \Sigma_{12}(t) \\ \Sigma_{12}(t) & \Sigma_{22}(t) \end{pmatrix}^{\frac{1}{2}}.$$

Note that the matrix $\sigma(t)$ is symmetric as well as $\Sigma(t)$. Moreover

$$\Sigma(t) = \sigma(t)^2 = \begin{pmatrix} \sigma_{11}^2(t) + \sigma_{12}^2(t) & \sigma_{11}(t)\sigma_{12}(t) + \sigma_{12}(t)\sigma_{22}(t) \\ \sigma_{11}(t)\sigma_{12}(t) + \sigma_{12}(t)\sigma_{22}(t) & \sigma_{22}^2(t) + \sigma_{12}^2(t) \end{pmatrix}.$$

This gives the following relation between Σ_{ij}'s and σ_{ij}'s

$$\begin{aligned}
\Sigma_{11}(t) &= \sigma_{11}^2(t) + \sigma_{12}^2(t),\\
\Sigma_{12}(t) &= \sigma_{11}(t)\sigma_{12}(t) + \sigma_{12}(t)\sigma_{22}(t),\\
\Sigma_{22}(t) &= \sigma_{22}^2(t) + \sigma_{12}^2(t).
\end{aligned}$$

Equation (4.4) yields

$$\begin{aligned}
d\begin{pmatrix} \Sigma_{11}(t) & \Sigma_{12}(t) \\ \Sigma_{12}(t) & \Sigma_{22}(t) \end{pmatrix} &= \left(\beta \begin{pmatrix} Q_{11} & Q_{21} \\ Q_{12} & Q_{22} \end{pmatrix} \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}\right.\\
&+ \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} \Sigma_{11}(t) & \Sigma_{12}(t) \\ \Sigma_{12}(t) & \Sigma_{22}(t) \end{pmatrix}\\
&+ \left.\begin{pmatrix} \Sigma_{11}(t) & \Sigma_{12}(t) \\ \Sigma_{12}(t) & \Sigma_{22}(t) \end{pmatrix} \begin{pmatrix} M_{11} & M_{21} \\ M_{12} & M_{22} \end{pmatrix}\right) dt\\
&+ \begin{pmatrix} \sigma_{11}(t) & \sigma_{12}(t) \\ \sigma_{12}(t) & \sigma_{22}(t) \end{pmatrix} \begin{pmatrix} dW_{11}(t) & dW_{12}(t) \\ dW_{21}(t) & dW_{22}(t) \end{pmatrix} \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}\\
&+ \begin{pmatrix} Q_{11} & Q_{21} \\ Q_{12} & Q_{22} \end{pmatrix} \begin{pmatrix} dW_{11}(t) & dW_{12}(t) \\ dW_{21}(t) & dW_{22}(t) \end{pmatrix} \begin{pmatrix} \sigma_{11}(t) & \sigma_{12}(t) \\ \sigma_{12}(t) & \sigma_{22}(t) \end{pmatrix}.
\end{aligned}$$

The matrix SDE (4.4) yields the following equations for Σ_{11}, Σ_{12} and Σ_{22}

$$\begin{aligned}
d\Sigma_{11}(t) &= (\beta(Q_{11}^2 + Q_{21}^2) + 2M_{11}\Sigma_{11} + 2M_{12}\Sigma_{12})dt\\
&+ 2\sigma_{11}(Q_{11}dW_{11}(t) + Q_{21}dW_{12}(t)) + 2\sigma_{12}(Q_{11}dW_{21}(t) + Q_{21}dW_{22}(t)),\\
d\Sigma_{12}(t) &= (\beta(Q_{11}Q_{12} + Q_{21}Q_{22}) + M_{11}\Sigma_{12} + M_{12}\Sigma_{22} + M_{21}\Sigma_{11} + M_{22}\Sigma_{12})dt\\
&+ \sigma_{11}(Q_{12}dW_{11}(t) + Q_{22}dW_{12}(t)) + \sigma_{12}(Q_{12}dW_{21}(t) + Q_{22}dW_{22}(t))\\
&+ \sigma_{12}(Q_{11}dW_{11}(t) + Q_{21}dW_{12}(t)) + \sigma_{22}(Q_{11}dW_{21}(t) + Q_{21}dW_{22}(t)),\\
d\Sigma_{22}(t) &= (\beta(Q_{12}^2 + Q_{22}^2) + 2M_{21}\Sigma_{12} + 2M_{22}\Sigma_{22})dt\\
&+ 2\sigma_{12}(Q_{12}dW_{11}(t) + Q_{22}dW_{12}(t)) + 2\sigma_{22}(Q_{12}dW_{21}(t) + Q_{22}dW_{22}(t)).
\end{aligned}$$

Considering the equations for Σ_{ij}'s, the quadratic variations are

$$\begin{cases} d\Sigma_{11}(t).d\Sigma_{11}(t) &= 4\Sigma_{11}(t)(Q_{11}^2 + Q_{21}^2)dt, \\ d\Sigma_{22}(t).d\Sigma_{22}(t) &= 4\Sigma_{22}(t)(Q_{12}^2 + Q_{22}^2)dt, \\ d\Sigma_{11}(t).d\Sigma_{22}(t) &= 4\Sigma_{12}(t)(Q_{11}Q_{12} + Q_{21}Q_{22})dt, \\ d\Sigma_{12}(t).d\Sigma_{12}(t) &= (\Sigma_{11}(t)(Q_{12}^2 + Q_{22}^2) + 2\Sigma_{12}(t)(Q_{11}Q_{12} + Q_{21}Q_{22}) + \Sigma_{22}(t)(Q_{11}^2 + Q_{21}^2))dt, \\ d\Sigma_{11}(t).d\Sigma_{12}(t) &= (2\Sigma_{11}(t)(Q_{11}Q_{12} + Q_{21}Q_{22}) + 2\Sigma_{12}(t)(Q_{11}^2 + Q_{21}^2))dt \\ d\Sigma_{22}(t).d\Sigma_{12}(t) &= (2\Sigma_{22}(t)(Q_{11}Q_{12} + Q_{21}Q_{22}) + 2\Sigma_{12}(t)(Q_{12}^2 + Q_{22}^2))dt. \end{cases}$$

Taking derivatives on both sides of $\rho_t = \frac{\Sigma_{12}(t)}{\sqrt{\Sigma_{11}(t)\Sigma_{22}(t)}}$, one obtains

$$\begin{aligned} d\rho_t &= \frac{d\Sigma_{12}(t)}{\sqrt{\Sigma_{11}(t)\Sigma_{22}(t)}} + \Sigma_{12}(t)d(\frac{1}{\sqrt{\Sigma_{11}(t)\Sigma_{22}(t)}}) + d(\Sigma_{12}(t)).d(\frac{1}{\sqrt{\Sigma_{11}(t)\Sigma_{22}(t)}}) \\ &= \frac{1}{\sqrt{\Sigma_{11}(t)\Sigma_{22}(t)}}(\beta Q_{11}Q_{12} + \beta Q_{21}Q_{22} + M_{21}\Sigma_{11}(t) + M_{12}\Sigma_{22}(t) + M_{11}\Sigma_{12}(t) + M_{22}\Sigma_{12}(t))dt \\ &+ \Sigma_{12}(t)\left(-\frac{1}{2\sqrt{\Sigma_{22}(t)\Sigma_{11}^3(t)}}(\beta Q_{11}^2 + \beta Q_{21}^2 + 2M_{11}\Sigma_{11}(t) + 2M_{12}\Sigma_{12}(t))\right. \\ &- \frac{1}{2\sqrt{\Sigma_{11}(t)\Sigma_{22}^3(t)}}(\beta Q_{12}^2 + \beta Q_{22}^2 + 2M_{21}\Sigma_{12}(t) + 2M_{22}\Sigma_{22}(t)) \\ &+ \frac{3}{8\sqrt{\Sigma_{11}(t)\Sigma_{22}(t)}\Sigma_{11}^2(t)}d\Sigma_{11}(t).d\Sigma_{11}(t) + \frac{3}{8\sqrt{\Sigma_{11}(t)\Sigma_{22}(t)}\Sigma_{22}^2(t)}d\Sigma_{22}(t).d\Sigma_{22}(t) \\ &+ \left.\frac{1}{4\sqrt{(\Sigma_{11}(t)\Sigma_{22}(t))^3}}d\Sigma_{11}(t).d\Sigma_{22}(t)\right) dt - \frac{1}{2\sqrt{\Sigma_{22}(t)\Sigma_{11}^3(t)}}d\Sigma_{11}(t).d\Sigma_{12}(t) \\ &- \frac{1}{2\sqrt{\Sigma_{11}(t)\Sigma_{22}^3(t)}}d\Sigma_{22}(t).d\Sigma_{12}(t) + \text{ Diffusion Term }. \end{aligned} \tag{4.6}$$

The diffusion term of the stochastic correlation process can be found similarly by direct calculation. Differentiating the equation $\rho_t^2 = \frac{\Sigma_{12}^2(t)}{\Sigma_{11}(t)\Sigma_{22}(t)}$ yields

$$2\rho_t d\rho_t = (\dots)dt + \frac{2\Sigma_{12}(t)}{\Sigma_{11}(t)\Sigma_{22}(t)}d\Sigma_{12}(t) + \frac{\Sigma_{12}^2(t)}{\Sigma_{22}(t)}d(\frac{1}{\Sigma_{11}(t)}) + \frac{\Sigma_{12}^2(t)}{\Sigma_{11}(t)}d(\frac{1}{\Sigma_{22}(t)}),$$

or equivalently

$$d\rho_t = (\dots)dt + \frac{1}{\sqrt{\Sigma_{11}(t)\Sigma_{22}(t)}}\left(d\Sigma_{12}(t) - \frac{\Sigma_{12}(t)}{2\Sigma_{11}(t)}d\Sigma_{11}(t) - \frac{\Sigma_{12}(t)}{2\Sigma_{22}(t)}d\Sigma_{22}(t)\right).$$

Using the quadratic variations (4.6) one can obtain the diffusion part in (4.5) by writing $d\rho_t.d\rho_t$ as

$$\begin{aligned}
d\rho_t.d\rho_t &= \frac{1}{\Sigma_{11}(t)\Sigma_{22}(t)}\Big(\Sigma_{11}(t)(Q_{12}^2+Q_{22}^2)+2\Sigma_{12}(t)(Q_{11}Q_{12}+Q_{21}Q_{22})+\Sigma_{22}(t)(Q_{11}^2+Q_{21}^2)\\
&+ \Sigma_{12}^2(t)((\frac{Q_{11}^2+Q_{21}^2}{\Sigma_{11}(t)})+\frac{2\Sigma_{12}(t)(Q_{11}Q_{12}+Q_{21}Q_{22})}{\Sigma_{11}(t)\Sigma_{22}(t)}+(\frac{Q_{22}^2+Q_{21}^2}{\Sigma_{22}(t)}))\\
&- \frac{2\Sigma_{12}(t)}{\Sigma_{11}(t)}(\Sigma_{11}(t)(Q_{11}Q_{12}+Q_{21}Q_{22})+\Sigma_{12}(t)(Q_{11}^2+Q_{21}^2))\\
&- \frac{2\Sigma_{12}(t)}{\Sigma_{22}(t)}(\Sigma_{22}(t)(Q_{11}Q_{12}+Q_{21}Q_{22})+\Sigma_{12}(t)(Q_{12}^2+Q_{22}^2))\Big)dt,
\end{aligned}$$

and therefore the quadratic variation of the correlation process is given by

$$d\rho_t.d\rho_t = (1-\rho_t^2)(\frac{Q_{12}^2+Q_{22}^2}{\Sigma_{22}(t)}+\frac{Q_{11}^2+Q_{21}^2}{\Sigma_{11}(t)}-\frac{2\rho_t(Q_{11}Q_{12}+Q_{21}Q_{22})}{\sqrt{\Sigma_{11}(t)\Sigma_{22}(t)}})dt. \tag{4.7}$$

Equations (4.6) and (4.7) imply that the correlation between two assets ρ_t follows the dynamics

$$d\rho_t = (A_t\rho_t^2+B_t\rho_t+C_t)dt+\sqrt{1-\rho_t^2}D_t dW_t,$$

where the coefficients A_t, B_t and C_t are given by

$$\begin{aligned}
A_t &= \frac{Q_{11}Q_{12}+Q_{21}Q_{22}}{\sqrt{\Sigma_{11}(t)\Sigma_{22}(t)}}-\sqrt{\frac{\Sigma_{22}(t)}{\Sigma_{11}(t)}}M_{12}-\sqrt{\frac{\Sigma_{11}(t)}{\Sigma_{22}(t)}}M_{21},\\
B_t &= -\frac{\Omega_{11}^2+\Omega_{12}^2}{2\Sigma_{11}(t)}-\frac{\Omega_{21}^2+\Omega_{22}^2}{2\Sigma_{22}(t)}+\frac{Q_{11}^2+Q_{21}^2}{2\Sigma_{11}(t)}+\frac{Q_{12}^2+Q_{22}^2}{2\Sigma_{22}(t)},\\
C_t &= \frac{\Omega_{11}\Omega_{21}+\Omega_{12}\Omega_{22}-2Q_{11}Q_{12}-2Q_{21}Q_{22}}{\sqrt{\Sigma_{11}(t)\Sigma_{22}(t)}}+\sqrt{\frac{\Sigma_{22}(t)}{\Sigma_{11}(t)}}M_{12}+\sqrt{\frac{\Sigma_{11}(t)}{\Sigma_{22}(t)}}M_{21}.
\end{aligned}$$

Note that $\Omega\Omega' = \beta Q'Q$ implies $\Omega_{11}^2+\Omega_{12}^2 = \beta(Q_{11}^2+Q_{21}^2)$ and $\Omega_{21}^2+\Omega_{22}^2 = \beta(Q_{12}^2+Q_{22}^2)$. Since $n=2$, we have $\beta > 2-1 = 1$. This gives

$$\begin{aligned}
B_t &= -\frac{\Omega_{11}^2+\Omega_{12}^2}{2\Sigma_{11}(t)}-\frac{\Omega_{21}^2+\Omega_{22}^2}{2\Sigma_{22}(t)}+\frac{Q_{11}^2+Q_{21}^2}{2\Sigma_{11}(t)}+\frac{Q_{12}^2+Q_{22}^2}{2\Sigma_{22}(t)}\\
&= (1-\beta)(\frac{Q_{12}^2+Q_{22}^2}{2\Sigma_{22}(t)}) \leq 0,
\end{aligned}$$

which implies the negativity of B_t. This guarantees that the correlation has a mean reverting property which is in consistency with market observations.

4.4 The Wishart Autoregressive Process

The Wishart Autoregressive process is a stochastic positive definite matrix process which is suitable for modeling covariance matrix of some financial instruments including stocks, interest rates and exchange rates. In this section we give definitions of the WAR process, derive analytical results regarding the characteristic function and the distribution of this process and simulate the bivariate WAR process.

4.4.1 Definition

The simplest form of a Wishart Autoregressive process is constructed from the Gaussian $VAR(1)$ process. A Gaussian vector autoregressive process $\{x_t\}_{t\in\mathbb{N}}$ is defined by

$$x_{t+1} = Mx_t + \epsilon_{t+1} \ ; \ \epsilon_{t+1} \sim \mathcal{N}(0, \Sigma), \tag{4.8}$$

where ϵ_t 's are independent. From the above definition, it is obvious that $\{x_t\}_{t\in\mathbb{N}}$ is Markov. Moreover, $\{x_t\}_{t\in\mathbb{N}}$ is stationary if all the eigenvalues of the matrix M have modulus less than one. We define the $WAR(1)$ process Ω_t by

$$\Omega_t = \sum_{k=1}^{K} x_t^k (x_t^k)',$$

where $\{x_t^{(k)}\}_{k=1}^{K}$ are independent copies of the $AR(1)$ process defined by (4.8). One should note that the matrix Ω_t in above has rank K and therefore it is called the WAR(1) process of rank K. It is obvious that Ω_t is symmetric and positive definite. When $M = 0$, Ω_t is called the Wishart white noise.

4.4.2 Analytical Properties

To obtain analytical results for models which use Ω_t as a factor, one needs to find the conditional distribution function of Ω_t. This distribution function is hard to work with and has a complicated form. Therefore as a standard trick in affine models, we transform the computation from the pricing domain to the frequency domain by considering the Laplace transform of the $WAR(1)$ process. The Laplace transform of an n-dimensional random variable X is defined by

$$\psi_X(\lambda) = \mathbb{E}(exp(\lambda.X)),$$

for $\lambda \in \mathbb{R}^n$. In the case of a symmetric matrix process Ω_t, it makes sense to define the Laplace transform as the expectation of the exponential of the linear combination of Ω_t^{ij}. But if $\Gamma = (\alpha_{ij})$ is a symmetric matrix, because of the symmetry of Ω_t

$$Tr(\Gamma\Omega_t) = \sum_{i,j} \alpha_{ij}\Omega_t^{ij},$$

which is a linear combination of Ω_t^{ij} 's. Conversely any linear combination of Ω_t^{ij} 's can be rewritten as $Tr(\Gamma\Omega_t)$ for some symmetric matrix Γ, This justifies the notation of the proposition below, which gives a closed form solution for the characteristic function of Ω_t

Proposition 11. *Ω_t is a Markov Process. Moreover, the conditional Laplace transform of Ω_t is of the form*

$$\begin{aligned}
\psi_t(\Gamma) &= E_t(\exp Tr(\Gamma\Omega_{t+1})) \\
&= E(\exp \sum_{k=1}^{K} x'_{t+1}\Gamma x_{t+1}|x_t) \\
&= \frac{\exp Tr(M'\Gamma(I-2\Sigma\gamma)^{-1}M\Omega_t)}{(det(I-2\Sigma\Gamma))^{\frac{K}{2}}} .
\end{aligned}$$

Proof. See Gourieroux and Sufana (2004). □

Once can check that Ω_t is non-singular if $K > n-1$. The process Ω_t is denoted by $W_n(K, M, \Sigma)$ where M is called the autoregressive parameter and Σ is the innovation variance. As we mentioned before, the density function of the $WAR(1)$ process , $f(\Omega_{t+1}|\Omega_t)$, has a complicated form. By taking the inverse transformation of $\psi_t(\Gamma)$, one can find the conditional transition density function of the $WAR(1)$ process

Proposition 12. *The transition PDF of the Wishart Autoregressive Process at horizon h is given by*

$$\begin{aligned}
f(Y_{t+h}|Y_t) &= \frac{1}{2^{\frac{Kn}{2}}\Gamma_n(\frac{K}{2})(\det\Sigma(h))^{-\frac{K}{2}}}(\det Y_{t+h})^{\frac{K-n-1}{2}} \\
&\quad \exp(\frac{-1}{2}Tr\left(\Sigma(h)^{-1}(Y_{t+h}+M(h)Y_tM(h)')\right))\, {}_0F_1\,(\frac{K}{2},\frac{1}{4}M(h)Y_tM(h)'Y_{t+h}),
\end{aligned} \tag{4.9}$$

where

$$\Gamma_n(\frac{K}{2}) = \int_{A>>0} exp(Tr(-A))(\det A)^{\frac{K-n-1}{2}},$$

is the multidimensional gamma function and ${}_0F_1$ *is the hypergeometric function which has a series expansion*

$${}_0F_1\ (a; X) = \sum_{k=0}^{+\infty} \sum_{\kappa} (a)_\kappa \frac{C_\kappa(X)}{k!},$$

where $C_\kappa(X)$ *is the zonal polynomial of* X *corresponding to* κ*, the argument* X *is a symmetric matrix,* $\sum_\kappa$ *denotes summation over all partitions* $\kappa = (k_1, k_2, \ldots, k_m)$*,* $k_1 \geq k_2 \geq \ldots \geq k_m$ *of* k *and the generalized hypergeometric coefficient* $(a)_\kappa$ *is given by*

$$(a)_\kappa = \prod_{i=1}^{m} (a - \frac{1}{2}(i-1))_{k_i},$$

with $(a)_k = a(a+1)(a+2)\ldots(a+k-1)$ *and* $(a)_0 = 1$*.*

Proof. See Muirhead (2005), section 10.3 . □

Even though the Laplace transform includes all the probabilistic information about the process Ω_t, it is also beneficial to have the first and second order moments of Ω_t

Proposition 13. *For the conditional moments of the WAR(1) process, we have*

$$E(\Omega_{t+1}|\Omega_t) = M\Omega_t M' + K\Sigma,$$

and

$$\begin{aligned} Cov_t(\gamma'\Omega_t\alpha, \delta'\Omega_t\beta) &= \gamma' M\Omega_t M'\delta\alpha' M\Omega_t\beta + \gamma' M\Omega_t M'\beta\alpha'\Sigma\delta + \alpha' M\Omega_t M'\delta\gamma'\Sigma\beta \\ &\quad + \alpha' M\Omega_t M'\beta\gamma'\Sigma\delta + K[\gamma'\Sigma\beta\delta'\Sigma\delta + \alpha'\Sigma\beta\gamma'\Sigma\delta]. \end{aligned}$$

Proof. See Gourieroux and Sufana (2004). □

As a result of proposition 13, Ω_t is a weak AR(1) process:

$$\Omega_{t+1} = M\Omega_t M' + K\Sigma + \eta_{t+1},$$

where η_t has mean zero. One of the advantages of the family of the Wishart process is its invariance under portfolio allocation. If the returns of n assets are given by the vector r_t with covariance matrix Ω_t, the returns of n portfolios $P_1, P_2, \dots, P_n$ defined by $\bar{r}_t = Pr_t$ has the covariance $P'\Omega_t P$. The proposition below shows that if $\Omega_t \in WAR(1)$ then $P'\Omega_t P \in WAR(1)$ as well.

Proposition 14. *If Ω_t is a $W_n(K, M, \Sigma)$ process and P is an invertible matrix, then $P'\Omega_t P$ is a $W_n(K, P'M(P')^{-1}, P'\Sigma P)$ process.*

Proof. See Gourieroux and Sufana (2004). □

As the first result of proposition 14, if $\Omega_t \sim W_n(K, M, \Sigma)$ then one can write Ω_t as $\Omega_t = \Sigma^{\frac{1}{2}}\Omega_t^*\Sigma^{\frac{1}{2}}$, where $\Omega_t^* \sim W_n(K, \Sigma^{\frac{1}{2}}M\Sigma^{\frac{-1}{2}}, \mathbb{I})$. As the second result, if M is diagonalizable (i.e. it can be written as $M = EDE^{-1}$ for some diagonal matrix D), then $\hat{\Omega}_t = E^{-1}\Omega_t(E^{-1})' \in W_n(K, D, E^{-1}\Sigma(E^{-1})')$. The above proposition shows one of the nice properties of the Wishart process which has mathematical and economical interpretations. Mathematically, this means that Wishart process is invariant under linear invertible transformations. From economical point of view the above proposition implies the invariance of the Wishart process under portfolio allocations. If the returns of n risky assets are given by the stochastic vector r_t with covariance matrix Σ_t, and n portfolios of these risky assets are associated with the columns of the matrix α, the returns of the portfolios are given by $r_t(\alpha) = \alpha' r_t$. This implies that the covariance of the portfolio is $V_t[r_{t+1}(\alpha)] = \alpha' Y_t \alpha$. Therefore the above proposition implies that if the covariance matrix of the risky assets follows a Wishart process, the covariance matrix of any portfolio linearly created from those assets follows a Wishart process as well. This is the content of the following proposition

Proposition 15. *If a portfolio has weights α', then the covariance matrix of the portfolio is $\alpha'\Omega_t\alpha$ with the following first and second moments*

$$\begin{aligned} E_t(\alpha'\Omega_{t+1}\alpha) &= \alpha' M\Omega_{t+1}M'\alpha + \alpha' K\Sigma\alpha, \\ V_t(\alpha'\Omega_{t+1}\alpha) &= \alpha' M\Omega_{t+1}M'\alpha\alpha'\Sigma\alpha + 2K(\alpha'\Sigma\alpha)^2. \end{aligned}$$

Proof. For proof see Gourieroux and Sufana (2004). □

Knowing the expression of the Laplace transform of the $WAR(1)$ process, we extend the definition of the WAR process to Wishart Autoregressive process with integer autoregressive order. This is the most general definition of the WAR process we deal with in this content.

Definition 3. *Ω_t is a $WAR(p)$ process if the characteristic function of Ω_t is of the form*

$$\begin{aligned}\psi_t(\Gamma) &= E_t(\exp Tr(\Gamma\Omega_{t+1})) \\ &= \frac{\exp Tr(\Gamma(I-2\Sigma\gamma)^{-1}\sum_{j=1}^{p} M_j\Omega_{t-j+1}M_j^T)}{(det(I-2\Sigma\Gamma))^{\frac{K}{2}}},\end{aligned}$$

where $M_i \in M_{n\times n}$ are the matrix autoregressive coefficients. We denote the $WAR(p)$ process above by $W_n(K; M_1, M_2, \ldots, M_p, \Sigma)$.

$WAR(p)$ does not bear the classical representation of the $WAR(1)$ process as the summation of the square of $AR(p)$ processes. But similarly to $WAR(1)$ process, one can check that

$$\begin{aligned}E_t(\Omega_{t+1}) &= \sum_{j=1}^{p} M_j\Omega_{t-j+1}M_j' + K\Sigma, \\ V_t(\alpha'\Omega_{t+1}\alpha) &= 4\alpha'(\sum_{j=1}^{p} M_j\Omega_{t-j+1}M_j')\alpha\alpha'\Sigma\alpha + 2K(\alpha'\Sigma\alpha)^2,\end{aligned}$$

and therefore Ω_t is a weak $AR(p)$ process

$$\Omega_{t+1} = \sum_{j=1}^{p} M_j\Omega_{t-j+1}M_j' + K\Sigma + \eta_{t+1},$$

where η_{t+1} has mean zero.

By writing x_{t+h} inductively as a function of x_t, one can study the long term behavior of $\{x_t\}_{t=0}^n$ and $\{\Omega_t\}_{t=0}^n$. In fact, if x_t follows (4.8), then x_{t+h} follows

$$x_{t+h} = M^h x_t + \epsilon_{t+1}^{(h)} \; ; \; \epsilon_{t+1} \sim \mathcal{N}(0, \Sigma(h)),$$

with $\Sigma(h) = \sum_{j=0}^{h-1} M^j\Sigma(M^j)'$. One can rewrite Ω_{t+h} as

$$\Omega_{t+h} = \sum_{k=1}^{K} x_{t+h}^k (x_{t+h}^k)',$$

and conclude that Ω_{t+h} is a Wishart $W_n(K, M^h, \Sigma(h))$ process. This implies that

$$E_t(\Omega_{t+h}|\Omega_t) = M^h\Omega_{t-j+1}M^h + K\Sigma(h).$$

If all the eigenvalues of M have modulus less than 1, Ω_t is strictly stationary with the stationary distribution $W(K, 0, \Sigma_\infty)$ where Σ_∞ is given by the proposition below

Proposition 16. *The conditional $WAR(1)$ process at horizon h is a Wishart process. More precisely*

$$\Omega_{t+h}|\Omega_t \sim W_n(K, M^h, \Sigma(h)).$$

Moreover, if the eigenvalues of M are strictly less than 1 in modulus, $(\Omega_t)_{t>0}$ is stationary with marginal distribution

$$W_n(K, 0, \Sigma_\infty),$$

where Σ_∞ solves

$$\Sigma_\infty = M\Sigma_\infty M' + K\Sigma.$$

Proof. See Gourieroux and Sufana (2004). □

The Wishart Autoregressive process is a discrete version of the continuous Wishart process if proper conditions are satisfied by its parameters. If there exists a matrix A such that $M = e^A$, then Ω_t is the time discretization of a continuous Wishart process. Moreover, if K is a positive integer, Ω_t can be written as a sum of the squares of K independent Ornstein-Uhlenbeck processes. Consider K independent Ornstein-Uhlenbeck processes

$$dX_t^{(k)} = AX_t^{(k)}dt + QdW_t \qquad k = 1, 2, \ldots, K.$$

where $A \in M_{n\times n}$, $Q \in GL(n)$ and W_t is a standard Brownian motion. Then X_t is the continuous time version of a Gaussian $AR(1)$ process

$$x_{t+1} = Mx_t + \epsilon_{t+1} \; ; \; \epsilon_{t+1} \sim \mathcal{N}(0, \Sigma), \tag{4.10}$$

where $M = \exp(A)$ and $\Sigma = \int_0^1 e^{sA}QQ'(e^{sA})'\,ds$. Now define the continuous time stochastic process Ω_t as

$$\Omega_t = \sum_{k=1}^{K} X_t^{(k)}(X_t^{(k)})'.$$

By Ito's lemma

$$d\Omega_t = (KQ'Q + A\Omega_t + \Omega_t A') + \sum_{l=1}^{n} (X_t^{(k)}Q' + Q(X_t^{(k)})')dW_t^l,$$

which has n Brownian motions involved. This SDE has a disadvantage that in its diffusion term, $X_t^{(k)}$ and $(X_t^{(k)})'$ appear separately. It can be shown (see Gourieroux and Sufana (2004)) that Ω_t also satisfies

$$d\Sigma_t = (\Omega\Omega' + A\Sigma_t + \Sigma_t A')dt + QdW_t^{(k)}\Sigma_t^{\frac{1}{2}} + \Sigma_t^{\frac{1}{2}}dW_t^{(k)'}Q',$$

where W_t is an $n \times n$ matrix Brownian motion. This SDE therefore has n^2 random sources which obviously gives more information than $\{\Omega_s\}_{0<s<t}$ itself. More precisely, because of its symmetric properties, $\{\Omega_s\}_{0<s<t}$ has $\frac{n(n+1)}{2}$ independent random sources. Therefore the σ-algebra generated by $\{\Omega_s\}_{0<s<t}$ is strictly included in the σ-algebra generated by the n^2 Brownian motions $\{W_s\}_{0<s<t}$ for all $t > 0$. This is worthwhile to remark that in the case $n = 1$, one would get the CIR process for Ω_t which is the continuous version of the autoregressive gamma process.

One of the problems of the Wishart autoregressive process is the high number of its parameters. Gourieroux and Sufana (2004) proposes a general-to-specific approach based on the rank analysis of the autoregressive matrix M which reduces the number of the parameters significantly. The next two propositions give different factor representations for the autoregressive matrix M. First, and for the sake of simplicity, assume that Rank$(M) = 1$. Therefore there exist matrices β and α such that $M = \beta\alpha'$. In this case, there are some nice properties for the characteristic function of Ω_t. If the parameter M can be written as $\beta\alpha'$ for $\beta, \alpha \in \mathbb{R}^n$, then $\{\alpha' x_t^k\}$ satisfies

$$\alpha' x_{t+1}^k = \alpha' x_t^k + \alpha' \epsilon_{t+1}^h,$$

where $\alpha'\epsilon_{t+1}^k \sim \mathcal{N}(0, \alpha'\Sigma\ \alpha)$. This motivates the following proposition

Proposition 17. *i) If $M = \beta\alpha'$, where α and $\beta \in \mathbb{R}$, then the distribution of Ω_{t+1} given Ω_t depends on $\alpha'\Omega_t\alpha$ only. More precisely*

$$\psi_t(\Gamma) = \frac{\exp\left(Tr((\beta'\Gamma(I - 2\Sigma\gamma)^{-1}\beta)\alpha'\Omega_t\alpha)\right)}{(det(I - 2\Sigma\Gamma))^{\frac{K}{2}}}.$$

Moreover, the factor $\alpha'\Omega_t\alpha$ is a $W_1(K, \alpha'\beta, \alpha'\Sigma\alpha)$ process.

ii)Suppose $M = \beta\alpha'$, where $\beta, \alpha \in M_{n\times p}$ have full rank. Then the distribution of Ω_{t+1} conditioned on Ω_t depends on $\alpha'\Omega_t\alpha$ only, which is a $W_p(K, \alpha'\beta, \alpha'\Sigma\alpha)$ process. Moreover if the columns of $C \in M_{n\times p}$ are perpendicular to β $(C.\beta = 0)$

$$C'\Omega_t C \sim W_p(K, 0, C'\Sigma C),$$

hence $C'\Omega_t C$ is a so-called white noise process.

Proof. See Gourieroux and Sufana (2004). □

Similarly to above, if $C \in M_{n\times n}$ is invertible, then

$$C'\Omega_t C \sim W_p(K, C'M(C')^{-1}, C'\Sigma C).$$

Particularly, by choosing C' in some special form and assuming $M = \beta\alpha'$ one can derive additional specific results. For instance if

$$\text{Columns of C } \perp \beta,$$

then $C'M = C'\beta\alpha = 0$ and therefore

$$C'\Omega_t C \sim W_p(K, 0, C'\Sigma C).$$

The following proposition is an extension of the statement above.

Proposition 18. *Suppose $\Omega_t \sim W_n(K, M, \Sigma)$ and $a \in M_{n\times p}$ is full rank. Then $a'\Omega_t a$ is Markov iff there exists a $Q \in M_{p\times p}$ such that*

$$a'M = Qa'.$$

Moreover

$$a'\Omega_t a \sim W_n(K, Q, a'\Sigma a).$$

Proof. See Gourieroux and Sufana (2004). □

Wishart Autoregressive process can be used to construct WAR-in-mean process in the same way that Wishart process construct the stochastic covariance model. We have used Wisahrt process in section 4.5 to give the model below for dynamics of returns

$$d\ln A_t = (\mu_i + Tr(D_i\Omega_t))dt + \Sigma_t^{\frac{1}{2}} dW_t^S,$$

where Σ_t is the continuous version of the WAR process. We can define a WAR-in-mean process similarly as the return vector r_t of n risky assets such that the conditional distribution of r_{t+1} with respect to $\mathcal{F}_t$ is $\mathcal{N}(\mu_i + Tr(D_i\Omega_t), \Omega_t)$ when μ_i and D_i are parameters of the model. In other words r_t satisfies

$$r_{t+1} = (\mu_i + Tr(D_i\Omega_t))_i + \epsilon_{t+1} \ ; \ \epsilon_{t+1} \sim \mathcal{N}(0, \Omega_t).$$

This gives the discrete time return process associated with WAR process. As a result, the discrete WAR-in-mean process is a discretization of a continuous return process with Wishart covariance process if the conditions above are satisfied. Our discrete time assumption is that $r_{t+1}|r_t, (\Omega_t)$ has the drift $m_t = (\mu_i + Tr(D_i\Omega_t))_i$ and the volatility matrix Ω_t. For instance, when $n = 2$

$$\begin{aligned} r^1_{t+1} &= b_1 + d_{11}\Omega_{11} + 2d_{12}\Omega_{12} + d_{22}\Omega_{22} + \epsilon^1_{t+1}, \\ r^2_{t+1} &= b_2 + \bar{d}_{11}\Omega_{11} + 2\bar{d}_{12}\Omega_{12} + \bar{d}_{22}\Omega_{22} + \epsilon^2_{t+1}, \end{aligned}$$

with

$$V_t(\epsilon^1_{t+1}, \epsilon^2_{t+1}) = \Omega_t.$$

The return $Tr(D_i\Omega_t)$ is called the risk premium. To check the positivity of $Tr(D_i\Omega_t)$ when D_i is positive definite, suppose D is positive definite. Then it can be written as $D = \sum_{k=1}^n d_k d_k^{'}$ where $d_k \in \mathbb{R}^n$. Therefore

$$\begin{aligned} Tr(D\Omega_t) &= Tr(\sum_{k=1}^n d_k d_k^{'} \Omega_t) \\ &= \sum_{k=1}^n Tr(d_k d_k^{'} \Omega_t) \\ &= \sum_{k=1}^n Tr(d_k \Omega_t d_k^{'}) \geq 0, \end{aligned}$$

which guarantees the risk premium $Tr(D_i\Omega_t)$ to be positive. Note that since Ω_t is positive definite $d_k\Omega_t d_k^{'} \geq 0$. Moreover, one can define an order $>>$ on the positive definite matrices such that

$$A >> B \Longleftrightarrow A - B \text{ is positive definite .}$$

If $\Omega_t^{(1)}$ and $\Omega_t^{(2)}$ are two Wishart processes such that $\Omega_t^{(2)} >> \Omega_t^{(1)}$ then $Tr(D(\Omega_t^{(2)} - \Omega_t^{(1)})) \geq 0$ or $Tr(D\Omega_t^{(2)}) \geq Tr(D\Omega_t^{(1)})$. Therefore $Tr(D\Omega_t)$ preserves the positive definite order. This is referred to as the monotonicity property of the risk premium.

There are two situations for the calibration of the Wishart process: When the volatility matrix is observed or when the volatility matrix is not observed. Gourieroux and Sufana (2004) calibrate the Wishart autoregressive process when Ω_t is observable. In this case, it can be proved that if the volatility matrix Ω_t is observed, K and Σ are identifiable but M is identifiable up to its sign. Moreover, Σ is first-order identifiable up to a scale factor and M is first-order identifiable up to its sign, but K is second-order identifiable.

4.4.3 Wishart Autoregressive process when n = 2

In this section, we simulate a 2×2-WAR(1) process. In this case, if the Wishart Autoregressive process is given by

$$\Omega(t) = \begin{pmatrix} \Omega_{11}(t) & \Omega_{12}(t) \\ \Omega_{12}(t) & \Omega_{22}(t) \end{pmatrix}.$$

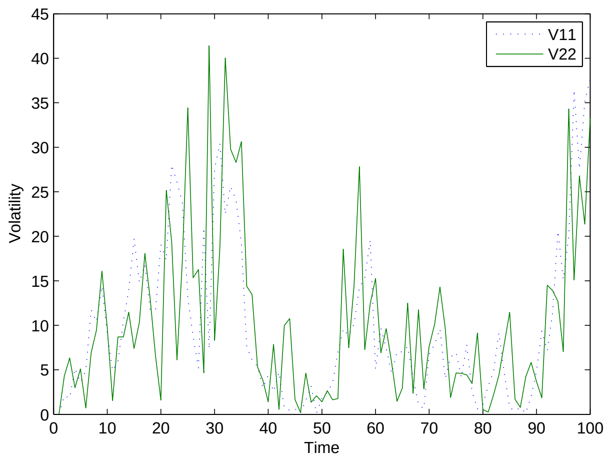

Figure 4.1: Simulated WAR(1) Stochastic Volatilities

$\Omega_{11}(t)$ and $\Omega_{22}(t)$ denote the volatility and the correlation is $\rho(t) = \frac{\Omega_{12}(t)}{\sqrt{\Omega_{11}(t).\Omega_{22}(t)}}$. Assume that the parameters of the WAR(1) model are given by

$$M = \begin{pmatrix} 0.5 & 0 \\ 0.9 & 0 \end{pmatrix}, \Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

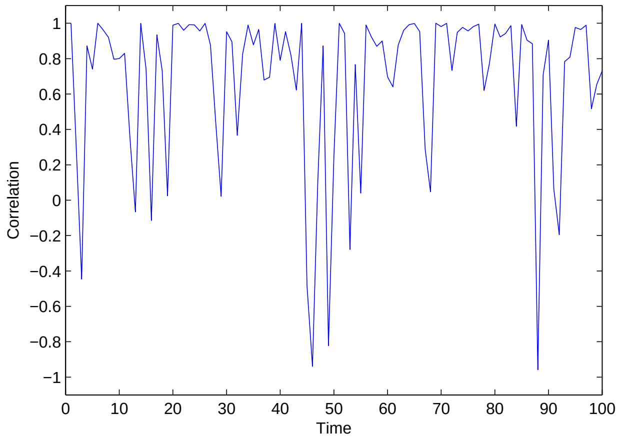

Figure 4.2: Simulated WAR(1) Stochastic Correlation

Here the Gindikin coefficient is $K = 2$. We use 100 time steps for the purpose of our simulation. The figure 4.1 shows the generated volatility for the first and the second asset by the WAR(1) process. One of the interesting observations about the WAR(1) process is that it is able to reproduce the volatility clustering phenomena, that are the periods of high/low volatilities. The figure 4.2 shows the correlation generated by the WAR(1) process. As it can be seen from the figure 4.2, the correlation has a fluctuating stochastic behavior. Figure 4.3 shows the maximum and minimum eigenvalues of the simulated Wishart process.

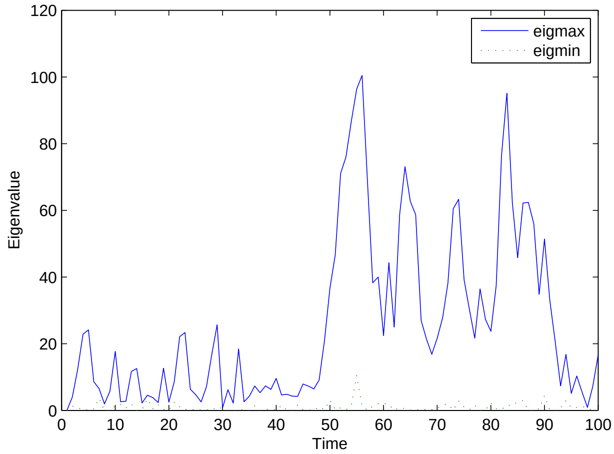

Figure 4.3: Maximum and Minimum Eigenvalues of the Simulated Wishart Process

4.5 The Model

In this section, we introduce our stochastic correlation CreditGrades model. CreditGrades model is an extension of the Merton model jointly developed by CreditMetrics, JP Morgan, Goldman Sachs and Deutsche Bank. The original version of the CreditGrades model assumes that volatility is deterministic. We extend CreditGrades model, by means of stochastic covariance Wishart process focusing on the role of stochastic correlation. For numerical implementation, we restrict ourselves to two reference firms, but the analytical results are valid for arbitrary number of firms.

4.5.1 The Dynamics of the Assets

The assets are defined on a probability space $(\Omega, \mathcal{F}, \mathbb{Q})$ where $\{\mathcal{F}_t\}_{t\geq 0}$ is the information up to time t and $\mathbb{Q}$ is the risk-neutral measure equivalent to the real-world measure $\mathbb{P}$. Let's assume that the i^{th} firm's asset price per share is given by $A_i(t)$. Here we review the results regarding the dynamics of the assets with stochastic covariance Wishart process [2]. As before, assume the assets' prices follow the multivariate real-world model

$$
\begin{aligned}
d\ln A_t &= (\mu_i + (Tr(D_i\Sigma_t)))dt + \Sigma_t^{\frac{1}{2}} dW_t^A, \\
d\Sigma_t &= (\Omega\Omega' + M\Sigma_t + \Sigma_t M')dt + QdW_t^{\sigma}\Sigma_t^{\frac{1}{2}} + \Sigma_t^{\frac{1}{2}} dW_\sigma Q'.
\end{aligned}
$$

Here the log-price process has the drift term $E_t(d\ln A_t) = (\mu_i + (Tr(D_i\Sigma_t)))dt$ and the quadratic variation $V_t(d\ln A_t) = \Sigma_t dt$. Moreover, we assume that the Brownian motions driving the assets and the Brownian motions driving the Wishart process are uncorrelated. The vector μ is constant and D_i is a symmetric positive definite matrix. Assuming D_i to be a symmetric positive definite matrix, $Tr(D_i\Sigma_t) > 0$ accounts for risk premium. For the transition distribution of A_{t+h} given A_t and (Σ_t) we have

$$
\ln A_{t+h}|A_t, (\Sigma_t) \sim \mathcal{N}\left(\ln A_t + \int_t^{t+h} \mu + Tr(D_i\Sigma_u)\, du, \int_t^{t+h} \Sigma_u\, du\right),
$$

and the unconditional probability function can be found by Integration over the distribution function of $\int_t^{t+h} \Sigma_u\, du$. Considering $a'\Sigma_t a$ as a one-dimensional process, its drift and quadratic variation is given by

[2] We do not need the analytical results in this subsection to price the options we consider. This is because we are interested in pricing equity options, which are options on equity and the equity itself is a derivative on the asset. However, we include the analytical results regarding the characteristic function of the joint asset and covariance matrix for completeness.

$$
\begin{aligned}
E_t\left(d(a'\Sigma_t a)\right) &= (a'\Omega\Omega^T a + a'M\Sigma_t a + a'\Sigma_t M'a)dt, \\
V_t\left(d(a'\Sigma_t a)\right) &= 4(a'\Sigma_t a)(a'Q'Qa)dt.
\end{aligned}
$$

Note that when Σ_t hits the boundary of the symmetric definite matrices, there exists an a such that $a\Sigma_t a = 0$, $\Sigma_t^{\frac{1}{2}} a = 0$ and $a\Sigma_t^{\frac{1}{2}} = 0$. Therefore from above $V_t\left(d(a'\Sigma_t a)\right) = 0$ and $E_t\left(d(a'\Sigma_t a)\right) > 0$ which implies a *deterministic* reflection towards the positive direction of the symmetric matrices. This intuitive and not so accurate statement justifies the fact that the process Σ_t introduced by the dynamics above lives in the space of symmetric positive definite matrices. The following proposition is the main result regarding the characteristic function and the real-world distribution function of the joint process $(\ln A_t, \Sigma_t)$

Proposition 19. *For the characteristic function of the joint log-price and volatility we have*

$$
\begin{aligned}
\phi_{t,h}(\psi_0, \psi_1, \psi_2, C_1, C_2) &= E_t\Bigg(exp\int_t^{t+h}(\psi_0 \ln A_u + \psi_1)du + \psi_2^{'} \ln S_{t+h} \\
&\quad + \int_t^{t+h} Tr(C_1\Sigma_u)du + Tr(C_2\Sigma_{t+h})\Bigg) \\
&= \exp\left(A(h)' \ln S_t + Tr(B(h)\Sigma_t) + c(h)\right),
\end{aligned}
$$

where $A(h), B(h)$ and $c(h)$ satisfy the system of Ricatti ODEs

$$
\begin{aligned}
\frac{dA(h)}{dh} &= \psi_0, \\
\frac{dB(h)}{dh} &= B(h)M + M'B(h) + 2B(h)Q'QB(h) + \frac{1}{2}A(h)A(h)' + \sum_{i=1}^{n} A_i(h)D_i + C_1, \\
\frac{dc(h)}{dh} &= A(h)'\mu + Tr(B(h)\Omega\Omega') + \psi_1,
\end{aligned}
$$

with initial conditions $A(0) = \psi_2$, $B(0) = C_2$ and $c(0) = 0$.

Proof. See Gourieroux *et al.* (2004). □

The system of differential equations in the above proposition can be solved in closed forms. From the first equation, the function $A(h)$ can be easily found as $A(h) = \psi_0 h + \psi_2$. Other coefficients can be found recursively. The function $B(h)$ satisfies the well-known Ricatti equation which has been well studied leading to closed solutions (see Abou-Kandil *et al.* (2003) for the general theory and Grasselli and Tebaldi (2004) for applications in Affine term structure models). For the simple case $\psi_0 = 0$, We have

$$\begin{aligned} B(h) &= B^* + (\exp{(M + 2Q'QB^*)h})' \left((C_2 - B^*)^{-1}\right. \\ &+ \left. 2\int_0^h (\exp{(M + 2Q'QB^*)u})Q'Q(\exp{(M + 2Q'QB^*)u})'\, du\right)^{-1} (\exp{(M + 2Q'QB^*)h}), \end{aligned}$$

where B^* satisfies

$$M'B^* + B^*M + 2B^*Q'QB^* + \frac{1}{2}\psi_2\psi_2' + \sum_{i=1}^{n} \psi_2^{(i)} D_i + C_1 = 0.$$

From the last differential equation in proposition 19, $c(h)$ can be found as

$$c(h) = (\psi_2'\mu + \psi_1)h + Tr[\Omega\Omega' \int_0^h B(u)\, du].$$

Similar calculations can lead to characteristic function of the joint log-price and volatility process under the risk-neutral measure. For the sake of simplicity and without loss of generality, for the rest of this section we assume the risk free rate is zero. As a consequence of the Girsanov's theorem, The Radon-Nikodym derivative between the real-world and the risk-neutral probabilities is given by

$$m(t, t+h) = exp\left(\int_t^{t+h} [\alpha' d\ln A_u + Tr((\Gamma)_u d\Sigma_u)] + \int_t^{t+h} \alpha_{0u} + \Gamma_{0u}\, du\right),$$

where $\alpha_u, \Gamma_u, \alpha_{0u}$ and Γ_{0u} are coefficients dependant on the parameters of the model. The drift and the diffusion coefficients for the process $(\ln A_t, \Sigma_t)$ under the risk-neutral probability can be found by the expression above. Assume E_t^Q be the conditional expectation with respect to the risk neutral probability, then the drift and the diffusion of the processes $(\ln A_t, \Sigma_t)$ under the risk-neutral probability is given by

$$\begin{aligned} E_t^Q(d\ln A_t) &= -\frac{1}{2}(\sigma_{ii})dt, \\ E_t^Q(d\Sigma_t) &= E_t(d\Sigma_t) + 2(\Sigma_t\Gamma_t Q'Q + Q'Q\Gamma_t'\Sigma_t)dt. \end{aligned}$$

Note that in the drift expression above the risk-free interest rate is hidden since we have assumed a zero risk free interest rate. Proposition below from Gourieroux *et al.* (2004) derives the characteristic function of the joint log-price and volatility process under the risk-neutral measure

Proposition 20. *For the characteristic function of the joint log-price and volatility under the risk-neutral probability measure, we have*

$$\begin{aligned}\phi_{t,h}(\psi_0,\psi_1,\psi_2,C_1,C_2) &= E_t^Q\left(\exp\int_t^{t+h}(\psi_0\ln A_u+\psi_1)du+\psi_2\ln A_{t+h}\right.\\ &\quad\left.+\int_t^{t+h}Tr(C_1\Sigma_u)du+Tr(C_2\Sigma_{t+h})\right)\\ &= \exp\left(A^*(h)\ln A_t+Tr(B^*(h)\Sigma_t)+c^*(h)\right),\end{aligned}$$

where $A^*(h), B^*(h)$ and $c^*(h)$ satisfy the system of Ricatti ODEs

$$\begin{aligned}\frac{dA^*(h)}{dh} &= \psi_0,\\ \frac{dB^*(h)}{dh} &= B^*(h)M^*+M^{*'}B^*(h)+2B^*(h)Q'QB^*(h)+\frac{1}{2}A^*(h)A^*(h)'-\frac{1}{2}diag(A_i^*(h))+C_1,\\ \frac{dc(h)}{dh} &= Tr(B^*(h)\Omega\Omega')+\psi_1,\end{aligned}$$

with initial conditions $A^*(0)=\psi_2$, $B^*(0)=C_2$ and $c^*(0)=0$ and $M^*=M+2Q'Q\Gamma'$.

Proof. See Gourieroux *et al.* (2004). □

The coefficients $A^*(h), B^*(h)$ and $c^*(h)$ can be found similarly to their real-world counterparts. For $A^*(h)$, the first equation yields $A^*(h)=\psi_0 h+\psi_2$. Assuming $\psi_0=0$, We have

$$\begin{aligned}B^*(h) &= B^*+(\exp(M^*+2Q'QB^*)h)'\left((C_2-B^*)^{-1}\right.\\ &+ \left.2\int_0^h(\exp(M^*+2Q'QB^*)u)Q'Q(\exp(M^*+2Q'QB^*)u)'\,du\right)^{-1}(\exp(M^*+2Q'QB^*)h),\end{aligned}$$

where B^* satisfies

$$M^{*\prime}B^*+B^*M^*+2B^*Q'QB^*+\frac{1}{2}\psi_2\psi_2^{'}-\frac{1}{2}diag(\psi_2^{(i)})+C_1=0,$$

and finally for $c^*(h)$

$$c^*(h)=\psi_1 h+Tr[\Omega\Omega'\int_0^h B^*(u)\,du].$$

Up to now we have described the dynamics of the assets' prices with stochastic covariance structure come from the Wishart process. The asset's price for a firm is not directly observed from the market. This leads us to structural credit risk models which introduce the equity as a form of a derivative on the asset of the company. Then the role of the model is in connecting the equity market to the default event. For example Gourieroux *et al.* (2004) propose the following dynamics for the assets and liabilities in the Merton's model

$$
\begin{aligned}
\begin{pmatrix} d\ln A_i(t) \\ d\ln L_i(t) \end{pmatrix} &= \begin{pmatrix} \mu_A + Tr(D_{i,A}\Sigma_i(t) + Tr(D_A\Sigma(t))) \\ \mu_L + Tr(D_{i,L}\Sigma_i(t) + Tr(D_L\Sigma(t))) \end{pmatrix} dt + \Sigma_i^{\frac{1}{2}} dW_t^i + \Sigma^{\frac{1}{2}} dW_t^i, \\
d\Sigma_t &= (\Omega\Omega^T + M\Sigma_t + \Sigma_t M')dt + Q dW_t^\sigma \Sigma_t^{\frac{1}{2}} + \Sigma_t^{\frac{1}{2}} dW_\sigma Q', \\
d\Sigma_t &= (\Omega\Omega^T + M\Sigma_t + \Sigma_t M')dt + Q dW_t^\sigma \Sigma_t^{\frac{1}{2}} + \Sigma_t^{\frac{1}{2}} dW_\sigma Q'.
\end{aligned}
$$

Then the equity is defined as a call option on the asset with liability as the strike price. This is a direct extension of the Merton model with multivariate stochastic volatility. In that case the price of a bond, an equity and a CDS would be

$$
\begin{aligned}
B_i(t, t+h) &= E_t^Q \left(\frac{A_i}{L_i} \mathbb{I}_{\{A_i < L_i\}} + \mathbb{I}_{\{A_i > L_i\}} \right), \\
A_i(t) &= E_t^Q (A_i - L_i)^+, \\
CDS_i(t) &= P_t^* (A_i < L_i),
\end{aligned}
$$

which have closed form formulas based on the closed formulas for the conditional Laplace transform of the joint price−covariance process. We prefer not to use this model for several reasons. First, we found the CreditGrades model more popular in financial markets because of its ability to link the structural framework to equity derivatives. On the other hand, the model proposed by Gourieroux has one common Wishart process and one distinct Wishart processes for each asset. Therefore the number of parameters for the model is high and the calibration is extremely ill-posed because of the high degree of freedom imposed by the number of parameters inside the model. We found that in the case of two companies, assuming only one wishart process driving the covariance matrix gives a fairly flexible model to capture market's behavior, while at the same time provides a less number of parameters.

In the next section, we introduce the equity process based on the CreditGrades' perspective rather than Merton's perspective taking advantage of the flexibility of the CreditGrades model. This will enrich the credit risk modeling with possibility of early default and also a straightforward link between credit risk and the equity market. We will derive a formula for the infinitesimal generator of the joint equity-covariance process below. This operator will play an important role in the partial differential equation of the equity option's price.

4.5.2 The Stochastic Correlation CreditGrades Model

Now that we have identified the dynamics of the assets, we explain the mechanism of the CreditGrades model. As before, we assume that the i^{th} firm's value $A_i(t)$ is driven by the dynamics

$$\begin{cases} dA_i(t) = diag(A_i(t))[(r(t) - d_i(t))\mathbb{I}dt + \sqrt{\Sigma_t}dW_t], \\ d\Sigma_t = (\Omega\Omega' + M\Sigma_t + \Sigma_t M')dt + \sqrt{\Sigma_t}dZ_t Q + Q'dZ_t'\sqrt{\Sigma_t}, \end{cases} \tag{4.11}$$

where $\Omega\Omega' = \beta Q'Q$ for some $\beta > n-1$ and M is a negative definite matrix. We assume that assets are driven by the Brownian motion W_t, the covariance matrix of the assets which follows a Wishart process is driven by the Brownian motion Z_t and two Brownian motions W_t and Z_t are uncorrelated. Even though we make this assumption, in the literature there are reports regarding correlation between stock's noise and its volatility varying from -0.27 to -0.62 (see Rockinger and Semenova (2005) and Chacko and Viceira (2003)). The reason we make this assumption is that closed form formulas for the value of double-barrier options and equity options are not available when the asset and its volatility are correlated as pointed out by Lipton (2001) and Sepp (2006). We assume that $S_i(t)$ is the i^{th} firm's equity price per share, $B_i(t)$ is the i^{th} firm's debt per share and R_i is the i^{th} firm's recovery rate. $B_i(t)$ has a deterministic growth rate $r(t) - d_i(t)$, where $r(t)$ is the risk free interest rate and $d_i(t)$ is the dividend yield for the i^{th} firm. The recovery part of the debt, $D_i(t) = R_i B_i(t)$, is the default barrier for the asset. Therefore the default time based on CreditGrades model is given by

$$\eta_i = \inf\{0 \leq t | A_i(t) \leq D_i(t)\}.$$

In the framework of CreditGrades model, the equity's value is given by

$$S_i(t) = \begin{cases} A_i(t) - D_i(t) & , t < \eta_i, \\ 0 & , \eta_i \leq t. \end{cases}$$

In terms of the equity, the default time can be written as $\eta_i = \inf\{0 < t | S_i(t) \leq 0\}$. Zero is an absorbing state for the equity process which makes the pricing of the equity option similar to pricing of down-and-out options studied by Lipton (2001). By using $S_i(t) = A_i(t) - D_i(t)$, the dynamics of $D_i(t)$ and equation (4.11), the equity follows a shifted log-normal SDE[3]

$$\begin{cases} dS_i(t) = diag(S_i(t).(r(t) - d_i(t)))\mathbb{I}dt + diag(S_i(t) + D_i(t))\sqrt{\Sigma_t}dW_t, \\ d\Sigma_t = (\Omega\Omega' + M\Sigma_t + \Sigma_t M')dt + \sqrt{\Sigma_t}dZ_tQ + Q'dZ_t'\sqrt{\Sigma_t}. \end{cases} \tag{4.12}$$

Note that the solution of the dynamics above can reach negative values but not before the stopping time η_i. We force sufficient conditions on the wishart process to make Σ_t mean reverting. For our purposes, we assume M is negative definite and $\Omega\Omega' = \beta Q'Q$ for some $\beta > n - 1$. Moreover, without loss of generality, we assume $\Omega = \sqrt{\beta}Q'$. We first derive the infinitesimal generator of the joint process (S, Σ). This operator will appear in the pricing PDE for equity options and the probabilities of default in the next section

Proposition 21. *The infinitesimal generator of the joint process (S, Σ) is given by*

$$\begin{aligned} \mathcal{A}_{(S,\Sigma)} &= [(r(t) - d(t))S] \bigtriangledown_S + \frac{1}{2}[(S + D)\bigtriangledown_S]\Sigma[(S + D)\bigtriangledown_S]' \\ &+ Tr[(\Omega\Omega' + M\Sigma_t + \Sigma_t M')D + 2\Sigma DQ'QD], \end{aligned}$$

where $D = \left(\frac{\partial}{\partial\Sigma_{ij}}\right)_{ij}$, Tr is the trace of a matrix, and we've used the notation $[(r-d)S] = Vec\left((r - d_i)S_i\right)$ and $[(S + D)\bigtriangledown_S] = Vec((S_i + D_i)\frac{\partial}{\partial S_i})$

Proof. Since $S_i(t) = A_i(t) - D_i(t)$

$$dS_t = [(r_t - d_t)S_t]dt + [S_t + D_t]\Sigma_t^{\frac{1}{2}}dW_t.$$

$\mathcal{A}_{(S,\Sigma)}$ can be divided to

$$\mathcal{A}_{(S,\Sigma)} = \mathcal{A}_S + \mathcal{A}_\Sigma + \mathcal{A}_{<S,\Sigma>}.$$

Since dZ_t and dW_t are independent, the last term is zero. By Bru (1991)

$$\mathcal{A}_\Sigma = Tr[(\Omega\Omega' + M\Sigma_t + \Sigma_t M')D + 2\Sigma DQ'QD].$$

[3]Note that $S_i(t)$ with the above dynamics is allowed to gain negative values but not prior to the stopping time η_i. Even though it might seem unreasonable to allow $S_i(t)$ have negative values, this doesn't affect any of the pricing formulas since whenever the process $S_i(t)$ is involved, it is followed by the truncating factor $\mathbb{I}_{\{\eta_i > \tau\}}$ (as in equations (4.13) and (4.25) for the payoffs of call and put options.)

To find $\mathcal{A}_S$, by the dynamics of S_t (4.12)

$$\begin{aligned}
\mathcal{A}_S &= \sum_{i=1}^{n}\left((r-d_i)S_i\frac{\partial}{\partial S_i}\right)+\frac{1}{2}\sum_{i,j=1}^{n}\left((S_i+D_i)(S_j+D_j)\Sigma_{ij}\frac{\partial^2}{\partial S_i\partial S_j}\right)\\
&= [(r(t)-d(t))S]\,\bigtriangledown_S+\frac{1}{2}[(S+D)\bigtriangledown_S]\Sigma[(S+D)\bigtriangledown_S]'.
\end{aligned}$$

□

4.6 Derivative Pricing; Analytical Results

In this section, we tackle the *pricing problem* of our credit risk model. We will use the fourier transform and method of images to solve the pricing problem for European calls and puts on the equity. This will give us the privilege to use equity options along with the the evolutionary algorithm to estimate the parameters of the model. The parameters can then be used to calculate the probabilities of default for each individual firm.

4.6.1 Equity Call Options

The price of a European Call option on the equity is calculated by discounting the risk-neutral expectation of the payoff at maturity. Since $\left(S_T\mathbb{I}_{\{\eta>T\}}-K\right)^+=(S_T-K)^+\,\mathbb{I}_{\{\eta>T\}}$ the price of the call option could be rewritten as

$$\begin{aligned}
V_{call}(t,\Sigma,S,K) &= E^Q_{(t,\Sigma,S)}\left(\exp(-\int_t^T r(s)\,ds)\,(S_T\mathbb{I}_{\eta>T}-K)^+\right)\\
&= E^Q_{(t,\Sigma,S)}\left(\exp(-\int_t^T r(s)\,ds)\,(S_T-K)^+\,\mathbb{I}_{\eta>T}\right).
\end{aligned} \tag{4.13}$$

The price of a single name derivative on one of the equities satisfies the partial differential equation $W_t+\mathcal{A}_{(\Sigma,S)}-rW=0$. Specially, the price of an equity call option is given by the PDE

$$\begin{cases}
W_t+\frac{1}{2}\Sigma_{ii}(S_i+D_i(t))^2W_{S_iS_i}+(r(t)-d_i(t))S_iW_{S_i}+\mathcal{A}_\Sigma W-rW=0,\\
W(t,0)=0, W(T,S)=(S-K)^+,
\end{cases} \tag{4.14}$$

where $\mathcal{A}_{(\Sigma,S)}$ is the infinitesimal generator of the joint process (S,Σ) given by the proposition 21. We first change the variables by $x_i = \ln(\frac{S_i+D_i(t)}{D_i(t)})$, $a_i = \ln(\frac{D_i(T)+K}{D_i(T)})$ and $G(t,x_i) = \exp(\int_t^T r(s)\,ds)\frac{W(t,S_i)}{D_i(T)}$ to transform the PDE (4.14) to

$$\begin{cases} G_t + \frac{1}{2}\Sigma_{ii}(G_{xx} - G_x) + \mathcal{A}_\Sigma G = 0, \\ G(t,0) = 0, W(T,x) = (e^x - e^a)^+. \end{cases} \tag{4.15}$$

To use the method of images, we need to eliminate the drift term first, hence we change the variables by $y_i = x_i - a_i$, $\tau = T - t$ and $U(\tau, y_i) = e^{-a}e^{-\frac{y_i}{2}}G(t,x_i)$. Then PDE (4.15) transforms to

$$\begin{cases} -U_\tau + \frac{1}{2}\Sigma_{ii}U_{y_iy_i} + \mathcal{A}_\Sigma U - \frac{1}{8}\Sigma_{ii}U = 0, \\ U(\tau,0) = 0, U(0,y) = (e^{\frac{y}{2}} - e^{-\frac{y}{2}})^+. \end{cases} \tag{4.16}$$

The PDE (4.16) is our reference PDE to solve the pricing problem for equity options on $S_i(t)$. We have the following proposition for the Fourier transform of the Green's function of PDE (4.16)

Proposition 22. *The Fourier transform of the Green's function of PDE* (4.16) *is given by*

$$q_j(\tau,\Sigma,Y) = \int_{-\infty}^{+\infty} e^{ikY_j + A(\tau,k) + Tr(B(\tau,k)\Sigma)}\,dk, \tag{4.17}$$

where

$$\begin{aligned} B(\tau,k) &= (\Lambda_{22}(\tau,k))^{-1}(\Lambda_{21}(\tau,k)), \\ A(\tau,k) &= Tr(\Omega\Omega'\int_0^\tau B(u,k)\,du), \end{aligned} \tag{4.18}$$

with

$$\Lambda = \begin{pmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{pmatrix} = \exp\tau \begin{pmatrix} M & -2Q'Q \\ -\frac{1}{2}(k^2+\frac{1}{4})\mathbb{I} & -M' \end{pmatrix}.$$

Proof. Define $\mathcal{X} = e^{ikY_j + A(\tau,k) + Tr(B(\tau,k)\Sigma)}$, Then one can check

$$
\begin{aligned}
q_\tau &= \int_{-\infty}^{+\infty} \mathcal{X}\,(A_\tau + Tr(B_\tau\Sigma))dk,\\
q_{yy} &= \int_{-\infty}^{+\infty} \mathcal{X}\left(-k^2\right)dk,\\
\mathcal{A}_\Sigma q &= \int_{-\infty}^{+\infty} \mathcal{X} Tr[(\Omega\Omega' + M\Sigma + \Sigma M')D + 2\Sigma D Q' Q D]\,dk.
\end{aligned}
$$

Substituting into (4.16) yields

$$
-[A_\tau + Tr(B_\tau\Sigma)] + \frac{1}{2}\Sigma_{jj}(-k^2) + Tr[(\Omega\Omega' + M\Sigma + \Sigma M')B + 2\Sigma B Q' Q B] - \frac{1}{8}\Sigma_{jj} = 0. \quad (4.19)
$$

Note that the functions satisfying the ODE above (i.e. $A(\tau, k)$ and $B(\tau, k)$) do not depend on the variable Σ. So we set $\Sigma = 0$ to get

$$
A_\tau - Tr(\Omega\Omega' B) = 0, \quad (4.20)
$$

and then by substituting (4.20) into (4.19)

$$
-[Tr(B_\tau\Sigma)] + \frac{1}{2}\Sigma_{jj}(-k^2) + Tr[(M\Sigma + \Sigma M')B + 2\Sigma B Q' Q B] - \frac{1}{8}\Sigma_{jj} = 0.
$$

To solve the above ODE, we do the following trick : rearrange the equation as

$$
\begin{cases}
Tr(H\Sigma) = & \frac{1}{2}\Sigma_{jj}(k^2 + \frac{1}{4}),\\
H(\tau, k) = & -B_\tau + BM + M'B + 2BQ'QB.
\end{cases}
$$

Therefore $H\Sigma$ satisfies

$$
\sum_{l,m=1}^{n} (H_{lm}\Sigma_{lm}) = \frac{1}{2}\Sigma_{jj}(k^2 + \frac{1}{4}).
$$

Since the function H_{ij} is independent of Σ, assuming Σ to be a zero matrix except for the $(l, m)^{th}$ entry yields

$$
H_{lm} = \begin{cases}
\frac{1}{2}(k^2 + \frac{1}{4}) & \text{if } l = m.\\
0 & \text{otherwise .}
\end{cases}
$$

Therefore

$$-B_\tau + BM + M'B + 2BQ'QB - \frac{1}{2}(k^2 + \frac{1}{4})\mathbb{I} = 0, \tag{4.21}$$

where $\mathbb{I}$ is the identity matrix. This matrix Ricatti equation has been studied in the literature (see Abou-Kandil *et al.* (2003)) and in Affine term structure models (see Grasselli and Tebaldi (2004)). As a standard method to solve this ODE, we change the variables as

$$B(\tau, k) = C(\tau, k)^{-1} D(\tau, k).$$

After substitution to (4.21), one gets

$$\begin{cases} D_\tau = -\frac{1}{2}(k^2 + \frac{1}{4})C + DM, \\ C_\tau = -CM' - 2DQ'Q, \end{cases}$$

or in the matrix form $X_\tau = XE$, where

$$X = \begin{pmatrix} D & C \end{pmatrix} \text{ and } E = \begin{pmatrix} M & -2Q'Q \\ -\frac{1}{2}(k^2 + \frac{1}{4})\mathbb{I} & -M' \end{pmatrix},$$

which has the solution

$$X(\tau, k) = X_0 e^{\tau E} = \begin{pmatrix} D_0 & C_0 \end{pmatrix} \begin{pmatrix} \Lambda_{11}(\tau, k) & \Lambda_{12}(\tau, k) \\ \Lambda_{21}(\tau, k) & \Lambda_{22}(\tau, k) \end{pmatrix},$$

where $\Lambda(\tau, k) = e^{\tau E} \in GL(2n)$. Therefore

$$\begin{cases} D(\tau, k) = D_0\Lambda_{11} + C_0\Lambda_{21}, \\ C(\tau, k) = D_0\Lambda_{12} + C_0\Lambda_{22}. \end{cases}$$

Now note that $B(\tau, k) = C(\tau, k)^{-1} D(\tau, k)$. Since $B(0, k) = 0$, we should have $D(0, k) = 0$ which implies $D_0 = 0$. Therefore

$$\begin{cases} D(\tau, k) = C_0\Lambda_{21}(\tau, k), \\ C(\tau, k) = C_0\Lambda_{22}(\tau, k). \end{cases}$$

And finally

$$B(\tau, k) = (\Lambda_{22}(\tau, k))^{-1}(\Lambda_{21}(\tau, k)).$$

$A(\tau, k)$ can be found by integration from (4.20)

$$A(\tau, k) = Tr(\Omega\Omega' \int_0^\tau B(\tau, k)\, du).$$

□

Now that we have found the fourier transform of the Green's function of the pricing PDE (4.16), we solve the pricing problem for an equity call option by the method of images.

Proposition 23. *The price of a call option on $S_j(t)$ with maturity date T and strike price K is given by*

$$W(t, S_j) = (D(T) + K)\exp(-\int_t^T r(s)\, ds)Z(\tau, y),$$

with

$$\begin{aligned} y &= \ln(\frac{S + D(t)}{D(T) + K}) + \int_t^T (r(s) - d(s))\, ds, \\ b &= \ln(\frac{D(t)}{D(T) + K}) + \int_t^T (r(s) - d(s))\, ds, \end{aligned}$$

and the function Z is defined by

$$Z(\tau, y) = e^y - e^b - \frac{e^{\frac{1}{2}y}}{\pi}\int_0^{+\infty} \frac{e^{A(\tau,k)+Tr(B(\tau,k)\Sigma)}(cos(yk) - cos((y - 2b)k))}{k^2 + \frac{1}{4}}\, ds, \tag{4.22}$$

with $A(\tau, k)$ and $B(\tau, k)$ given by

$$B(\tau, k) = (\Lambda_{22}(\tau, k))^{-1}(\Lambda_{21}(\tau, k)), \tag{4.23}$$

$$A(\tau, k) = Tr(\Omega\Omega' \int_0^\tau B(u, k)\, du). \tag{4.24}$$

Proof. The previous proposition gives the fourier transform of the Green's function of the pricing PDE. Now note that $q(\tau, \Sigma, Y)$ is invariant with respect to the change of variables

$y \to -y$ and $k \to -k$, therefore $q(\tau, \Sigma, Y)$ is an even function with respect to Y. This implies that the fourier transform of the Green's function absorbed at $x = b$ is

$$q^{(b)}(\tau, \Sigma, y, y') = q(\tau, \Sigma, y' - y) - q(\tau, \Sigma, y' + y - 2b).$$

By Duhamel's formula

$$\begin{aligned}
U(\tau, y) &= \frac{1}{2\pi}\int_0^{+\infty}\left(e^{\frac{y'}{2}} - e^{-\frac{y'}{2}}\right) q^{(b)}(\tau, \Sigma, y, y')\, dy' \\
&= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_0^{+\infty}\left(e^{\frac{y'}{2}} - e^{-\frac{y'}{2}}\right) e^{iky'}(e^{-iky} - e^{ik(y-2b)})e^{A(\tau,k)+Tr(B(\tau,k)\Sigma)}\, dy'dk \\
&= \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{A(\tau,k)+Tr(B(\tau,k)\Sigma)}(e^{-iky} - e^{ik(y-2b)})\int_0^{+\infty}\left(e^{(\frac{1}{2}+ik)y'} - e^{(-\frac{1}{2}+ik)y'}\right) dy'dk \\
&= \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{A(\tau,k)+Tr(B(\tau,k)\Sigma)}(e^{-iky} - e^{ik(y-2b)})\left(-\frac{1}{\frac{1}{2}+ik} + 2\pi\delta(k - \frac{i}{2}) + \frac{1}{-\frac{1}{2}+ik}\right) dk \\
&= e^{\frac{y}{2}} - e^{-\frac{1}{2}(y-2b)} - \frac{1}{2\pi}\int_{-\infty}^{+\infty}\frac{e^{A(\tau,k)+Tr(B(\tau,k)\Sigma)}(e^{iky} - e^{ik(y-2b)})}{k^2 + \frac{1}{4}}\, dk.
\end{aligned}$$

Remembering the consequent changes of variables

$$\begin{aligned}
G(t, x_i) &= e^{\int_t^T r(s)ds}\frac{W(t, S_i)}{D_i(T)}, \\
U(\tau, y_i) &= e^{-a}e^{-\frac{y_i}{2}}G(t, x_i),
\end{aligned}$$

one can conclude that $Z(\tau, y) = e^{\frac{y}{2}}U(\tau, y)$ and finally

$$Z(\tau, y) = e^y - e^b - \frac{e^{\frac{y}{2}}}{2\pi}\int_{-\infty}^{+\infty}\frac{e^{A(\tau,k)+Tr(B(\tau,k)\Sigma)}(\cos(yk) - \cos((y-2b)k))}{k^2 + \frac{1}{4}}\, dk.$$

□

For large values of k, the integrand in (4.22) is exponentially decreasing which makes it easy to evaluate the integral numerically.

Remark 2. *As we have mentioned before, our result covers Sepp (2006) as a special case. If in the dynamics of the asset* (4.11)*, we assume* $n = 1$ *and for the parameters we let* $M = -\frac{\kappa}{2}$, $Q = \frac{\sigma}{2}$ *and* $\Omega\Omega' = \kappa\theta$*, Propositions 22 and 23 yield* $B(\tau, k) = \frac{\Lambda_{12}}{\Lambda_{22}}$*. Now to find* Λ_{12} *and* Λ_{22}*, by proposition 22 we have*

$$E = \begin{pmatrix} -\frac{\kappa}{2} & -\frac{1}{2}(k^2 + \frac{1}{4}) \\ -\frac{\sigma^2}{2} & -\frac{\kappa}{2} \end{pmatrix}.$$

$\Lambda = e^{\tau E}$ is a 2×2 matrix with

$$\begin{aligned}\Lambda_{12}(\tau, k) &= \frac{(\kappa^2 - \zeta^2)(-e^{\frac{\tau\zeta}{2}} + e^{\frac{-\tau\zeta}{2}})}{-2\sigma^2\zeta}, \\ \Lambda_{22}(\tau, k) &= \frac{\sigma^2(-(\kappa+\zeta)e^{\frac{\tau\zeta}{2}} + (\kappa-\zeta)e^{\frac{-\tau\zeta}{2}})}{-2\sigma^2\zeta},\end{aligned}$$

where $\zeta = \sqrt{\kappa^2 + \sigma^2(k^2 + \frac{1}{4})}$. Therefore (4.23) implies

$$\begin{aligned}B(\tau, k) &= \frac{\Lambda_{12}(\tau, k)}{\Lambda_{12}(\tau, k)} \\ &= (k^2 + \frac{1}{4})\frac{1 - e^{-\tau\zeta}}{(\kappa+\zeta) + (-\kappa+\zeta)e^{-\tau\zeta}},\end{aligned}$$

The function $A(\tau, k)$ can be found by integration from (4.24). This gives the price of equity call option in the presence of Heston stochastic volatility (as in Sepp (2006) equations (3.5)-(3.7)).

4.6.2 Put-Call Parity

The price of a European put option on the equity is calculated by discounting the risk-neutral expectation of the payoff at maturity. Similarly to payoff of the call option, one can check that $(K - S_T\mathbb{I}_{\eta>T})^+ = (K - S_T)^+ \mathbb{I}_{\{\eta>T\}} + K\mathbb{I}_{\{\eta\leq T\}}$. Therefore the price of the put option could be rewritten as

$$\begin{aligned}V_{put}(t, \Sigma, S, K) &= E^Q_{(t,\Sigma,S)}\left(\exp(-\int_t^T r(s)\,ds)\,(K - S_T\mathbb{I}_{\eta>T})^+\right) \\ &= E^Q_{(t,\Sigma,S)}\left(\exp(-\int_t^T r(s)\,ds)\left((K - S_T)^+\mathbb{I}_{\eta>T} + K\mathbb{I}_{\eta\leq T}\right)\right).\end{aligned} \tag{4.25}$$

Equations (4.13) and (4.25) give the put-call parity for the equity options

$$\begin{aligned}V_{call}(t, \Sigma, S, K) - V_{put}(t, \Sigma, S, K) &= E^Q_{(t,\Sigma,S)}\left(\exp(-\int_t^T r(s)\,ds)\,((S_T - K)\mathbb{I}_{\eta>T} - K\mathbb{I}_{\eta\leq T})\right) \\ &= E^Q_{(t,\Sigma,S)}\left(\exp(-\int_t^T r(s)\,ds)\,((S_T - K + K)\mathbb{I}_{\eta>T} - K)\right) \\ &= V_{call}(t, \Sigma, S, 0) - K\exp(-\int_t^T r(s)\,ds).\end{aligned}$$

4.6.3 Survival Probabilities

Suppose $P(t,T,S_i)$ is the survival probability for the i^{th} company

$$P(t,T,S_i) = \mathbb{Q}_{(t,\Sigma,S_i)}(S_i(\tau) > 0 | t < \tau \leq T),$$

then using Feynman-Kac formula, $P(t,T,S_i)$ satisfies the partial differential equation $P_t + \mathcal{A}_{(S,\Sigma)}P = 0$.

Proposition 24. *The survival probability for the i^{th} firm is given by*

$$P(t,T,S_i) = \frac{2e^{\frac{y}{2}}}{\pi} \int_0^{+\infty} \frac{e^{A(\tau,k)+Tr(B(\tau,k)\Sigma)} k \sin{(ky)}}{k^2 + \frac{1}{4}}\, dk,$$

with functions $A(\tau,k)$ and $B(\tau,k)$ as in equations (4.18)*.*

Proof. Proposition 21 give the specific PDE for survival probability

$$\begin{cases} P_t + \frac{1}{2}\Sigma_{ii}(S_i + D_i(t))^2 P_{S_i S_i} + (r(t) - d_i(t))S_i P_{S_i} + \mathcal{A}_\Sigma P = 0, \\ P(t,T,0) = 0, P(T,T,S) = 1. \end{cases} \tag{4.26}$$

Using the change of variables $y_i = \ln{(\frac{S_i(t)+D_i(t)}{D_i(t)})}$, $\tau = T - t$ and $P(t,T,S_i) = e^{\frac{y}{2}} U(\tau, y_i)$, the PDE (4.26) transforms to

$$\begin{cases} -U_\tau + \frac{1}{2}\Sigma_{ii} U_{y_i y_i} + \mathcal{A}_\Sigma U - \frac{1}{8}\Sigma_{ii} U = 0, \\ P(\tau, 0) = e^{-\frac{y_i}{2}}, P(0, y_i) = 0. \end{cases}$$

This PDE is the same as (4.16). In Proposition 22, we have proved that the aggregated Green's function for this PDE is of the form (4.17). To find a bounded solution reflected at $x = 0$, we use the method of images to write the absorbed aggregated Green's function as

$$q^{(0)}(\tau, \Sigma, y, y') = q(\tau, \Sigma, y' - y) - q(\tau, \Sigma, y' + y).$$

Now by Duhamel's formula

$$
\begin{aligned}
U(\tau, y) &= \frac{1}{2\pi} \int_0^{+\infty} e^{-\frac{y'}{2}} q^{(0)}(\tau, \Sigma, y, y')\, dy' \\
&= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_0^{+\infty} e^{-\frac{y'}{2}} e^{iky'} (e^{-iky} - e^{iky}) e^{A(\tau,k)+Tr(B(\tau,k)\Sigma)}\, dy' dk \\
&= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{A(\tau,k)+Tr(B(\tau,k)\Sigma)} (e^{-iky} - e^{iky}) \int_0^{+\infty} e^{(-\frac{1}{2}+ik)y'}\, dy' dk \\
&= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{A(\tau,k)+Tr(B(\tau,k)\Sigma)} (e^{-iky} - e^{iky}) (\frac{1}{-\frac{1}{2}+ik})\, dk \\
&= \frac{2}{\pi} \int_0^{+\infty} \frac{e^{A(\tau,k)+Tr(B(\tau,k)\Sigma)} k \sin(ky)}{k^2 + \frac{1}{4}}\, dk.
\end{aligned}
$$

Remembering the change of variable $P(t, T, S_i) = e^{\frac{y}{2}} U(\tau, y_i)$, one can find the survival probability from the above formula for $U(\tau, y_i)$ as

$$
P(t, T, S_i) = \frac{2e^{\frac{y}{2}}}{\pi} \int_0^{+\infty} \frac{e^{A(\tau,k)+Tr(B(\tau,k)\Sigma)} k \sin(ky)}{k^2 + \frac{1}{4}}\, dk.
$$

□

4.6.4 Pricing Credit Default Swaps

Credit default swaps are one of the most popular credit derivatives traded in the market. A CDS provides protection against the default of a firm, known as reference entity. The buyer of the contract pays periodic payments, called CDS spreads, until the default time or maturity date. In return, the seller of the CDS provides the buyer with the unrecovered part of the notional if default occurs. The valuation problem of a CDS is then to give the CDS spread a value such that the contract begins with a zero value. This means that the value of the floating leg and the fixed leg should coincide when the contract is written. Assume that the CDS spread is denoted by S, the periodic payments occur at $0 = T_0 < T_1 < \ldots < T_N = T$, the notional is N, the time of default is denoted by τ and the recovery rate is the constant R. The *fixed leg* of the CDS is the value at time $t = 0$ of the cash flow corresponding to the payments the buyer makes. With the above notation we have

$$\begin{aligned}
\text{Fixed Leg} &= E^{\mathbb{Q}}\Big(\sum_{i=0}^{N} e^{-\int_t^{T_i} r(s)\,ds} SN(T_i - T_{i-1})\mathbb{I}_{\tau \geq T_i}\Big) \\
&= SN\sum_{i=0}^{N} e^{-\int_t^{T_i} r(s)\,ds}(T_i - T_{i-1})Q(t, T_i).
\end{aligned} \tag{4.27}$$

On the other hand, the *floating leg*, which is the value of the protection cash flow at $t = 0$, is

$$\begin{aligned}
\text{Floating Leg} &= E^{\mathbb{Q}}\Big(\sum_{i=0}^{N} e^{-\int_t^{T_i} r(s)\,ds}(1 - R)\mathbb{I}_{T_{i-1} < \tau \leq T_i}\Big) \\
&= (1 - R)\sum_{i=0}^{N} e^{-\int_t^{T_i} r(s)\,ds}(Q(t, T_{i-1}) - Q(t, T_i)).
\end{aligned} \tag{4.28}$$

The CDS spread S is chosen such that the contract has a fair value at $t = 0$. By setting the fixed leg equal to the floating leg, the equations (4.27) and (4.28) imply

$$S = \frac{(1 - R)\sum_{i=0}^{N} e^{-\int_t^{T_i} r(s)\,ds}(Q(t, T_{i-1}) - Q(t, T_i))}{\sum_{i=0}^{N} e^{-\int_t^{T_i} r(s)\,ds}Q(t, T_i)}.$$

4.7 Model Calibration with Evolutionary Algorithms

So far we have developed a credit risk model which has the Wishart process as the driving covariance matrix process. The Wishart process is a new stochastic correlation process in finance community and the features and properties of this process are still under investigation. The majority of the research done on the Wishart process is related to analytical properties of this process and little has been done related to the calibration problem of the Wishart process. Based on our knowledge, the only calibration paper regarding the Wishart process is Fonseca *et al.* (2007a). Using the characteristic function of the Wishart process, Fonseca estimates the Wishart Stochastic Correlation Model from the stock indexes data from SP500, FTSE, DAX and CAC40. In this section, we develop a calibration method to estimate the parameters of the extended CreditGrades model introduced in section 4.5. We use a probabilistic approach which searches a candidate space and finds parameters which match the market prices with the theoretical prices based on our option pricing formula. The solutions of the calibration algorithm are the sets of parameters which minimize the error function associated to the difference between market data and our theoretical results. We examine our results on options on General Motors' and Ford's equities.

4.7.1 Evolutionary Algorithms

In this section we briefly overview the evolutionary algorithm and explain its basic features. For a detailed discussion of the evolutionary algorithm we refer the reader to Hamida and Cont (2004) and Arian *et al.* (2008d). First we explain the nature of the random cycles happening in the algorithm and then we elaborate on the convergence properties of the algorithm.

The evolutionary algorithm has three main steps : mutation, crossover and selection. At each step, the set of population goes through a sequence of random flows which search the space to find optimizer regions and selects better parameters according to criteria derived from the error function. The algorithm starts with choosing N independent initial candidates, $\theta_1, \theta_2, \ldots, \theta_N$, with an initial distribution μ_0 as the initial population in the first step of the algorithm. Then the population set undergo a mutation - crossover - selection procedure to approach the set of optimizer parameter sets. Mutation is a search process in the space E to find better choices for the parameter θ. Search in the parameter space is done via a transition kernel $M(x, dy)$. After mutation, the population set goes under the crossover process, in which new parameters are produced from the existing ones. The way that the new parameters are selected depends on the problem in hand. We should also mention that crossover is not necessary to guarantee the convergence of the algorithm, but if chosen correctly, it can make the convergence faster by searching the space more effectively. The last part of the evolutionary algorithm cycle is the selection procedure. The performance of each parameter is examined in this part. For choosing the parameters, priority is given to the ones with better performance. The performance of each parameter is determined by an objective function and also an annealing parameter β_n. By letting β_n approach infinity, the evolutionary algorithm concentrates on the minimum areas of the objective function without being captured in a local minimum area.

So far we have explained the structure of the evolutionary algorithm. We should also make sure that the evolutionary algorithm concentrates the population on the minimum areas of the error function. For this purpose, we prove a convergence theorem based on the results from Moral and Miclo (2003) which ensure that the algorithm returns the set of parameters that match the market data with our theoretical results up to the bid-ask spread. In the literature, there are enormous results regarding the convergence of evolutionary algorithms. They usually prove the convergence without using the crossover step. In order to make sure that the algorithm does not get stuck in a local minimum, the normal mixing requirement on the mutation kernel is sufficient to make the algorithm search the whole space. On the other hand, a careful choice of the selection criteria is important to force the population set to the extremum areas of the objective function. For a more detailed discussion on the convergence of the evolutionary algorithm we refer the reader to Hamida and Cont (2004), Moral and Miclo (2003) and Arian *et al.* (2008d).

4.7.2 Model Calibration

We show the parameter set of the model by a vector θ which belongs to a compact set E. For any parameter set θ, the risk-neutral price of a derivative on the underlying asset is given by

$$C_i^{\theta}(t,T) = E_t^{Q^{\theta}}(e^{-\int_t^T r(s)\,ds}\Phi(T)),$$

where $\Phi(T)$ is the pay-off at maturity and $E_t^{Q^{\theta}}$ is the risk-neutral expectation with respect to the parameter set θ. Because of the model mis-specification and noisy observations of the market data, it is not possible to find a parameter set θ_0 such that the price given by the equation above matches all the quoted market prices for derivatives. The standard method to estimate the parameters is then to define a distance or error function for any given parameter $\theta \in E$ with respect to the market data. A proper parameter would minimize the error function in the search space where the evolutionary algorithm takes place. In what follows, we first determine the parameters of the model. Then we give a definition for the objective function which evaluates the distance each parameter has with respect to the optimum parameter. Then we give a distribution for our initial population and give the mutation and selection kernels of the evolutionary algorithm. Finally we prove a convergence theorem which guarantees that the algorithm in long run concentrates on the minimum of the error function. We finally apply the evolutionary algorithm to empirical data from options on General Motors' and Ford's equities.

Based on our assumption, the firms' asset values follow the dynamics

$$\begin{aligned} dA_i(t) &= diag(A_i(t))[(r(t)-d_i(t))\mathbb{I}dt + \sqrt{\Sigma_t}dW_t], \\ d\Sigma_t &= (\beta Q'Q + M\Sigma_t + \Sigma_t M')dt + \sqrt{\Sigma_t}dZ_tQ + Q'dZ_t'\sqrt{\Sigma_t}, \end{aligned}$$

where the parameters of the model are the Gindikin's coefficient $\beta > n-1 = 1$, negative definite autoregressive matrix M and the Wishart's volatility matrix $Q \in GL(n)$. In the two dimensional case, the parameters are shown in table 4.2

Table 4.1: Parameters of the Wishart Process

Wishart Covariance	Gindikin's Coefficient	Autoregressive Matrix		Wishart's Volatility Matrix		Initial Covariance	
Parameters	β	m_{11}	m_{12}	q_{11}	q_{12}	Σ_{11}	Σ_{12}
		m_{21}	m_{22}	q_{21}	q_{22}	Σ_{21}	Σ_{22}

To guarantee the positive definiteness of the Wishart process, we assume $\beta > 1$. Since the matrix M is negative definite, we have $m_{11} < 0, m_{22} < 0$ and $(m_{12} + m_{21})^2 < 4.m_{11}.m_{22}$. Moreover, for the positive definite matrix Σ we should have $\sigma_{11} > 0, \sigma_{22} > 0$ and $(\sigma_{12} + \sigma_{21})^2 < 4.\sigma_{11}.\sigma_{22}$. Finally, the covariance matrix of the Wishart process $Q \in GL(n)$ is invertible.

The objective function matching the theoretical prices with market data is given by

$$\begin{aligned} G(\theta) &= \sum_{i=1}^{I} \omega_i^2 |C_i^\theta(t,T_i) - C_i^*(t,T_i)|^2 \\ &+ \sum_{i=1}^{I} \omega_i \nu_i |C_i^\theta(t,T_i).D_i^\theta(t,T_i) - C_i^*(t,T_i).D_i^*(t,T_i)| \\ &+ \sum_{i=1}^{I} \nu_i^2 |D_i^\theta(t,T_i) - D_i^*(t,T_i)|^2. \end{aligned} \tag{4.29}$$

I is the number of options, $C_i^\theta(t,T)$ and $D_i^\theta(t,T)$ are the theoretical prices and $C_i^*(t,T)$ and $D_i^*(t,T)$ are the market prices for the derivatives. The reason that we define the objective function with equation (4.29) is the fact that the joint and marginal parameters of the extended CreditGrades model are mixed. Therefore the standard methods to estimate the joint and marginal parameters separately does not work. The first and third lines of the equation above calculate the error related to marginal information of the first and the second firms, respectively, and the second line calculates the joint information. Based on the equation above, we define a degree of accuracy

$$\begin{aligned}\delta &= \sum_{i=1}^{I} \omega_i^2 |C_i^{bid^2} - C_i^{ask^2}| \\ &+ \sum_{i=1}^{I} \omega_i \nu_i |C_i^{bid} D_i^{bid} - C_i^{ask} D_i^{ask}| \\ &+ \sum_{i=1}^{I} \nu_i^2 |D_i^{bid^2} - D_i^{ask^2}|.\end{aligned}$$

A values of θ such that $G(\theta) < \delta$ would be considered as an answer for the calibration problem. The weights in the equation (4.29) are given by

$$\omega_i = \frac{1}{|C_i^{bid^2} - C_i^{ask^2}|} \quad , \quad \nu_i = \frac{1}{|D_j^{bid^2} - D_j^{ask^2}|}.$$

In the first step of the algorithm, we draw N samples $X_0 = \{\theta_1, \theta_2, \ldots, \theta_N\}$ from E independently according to a truncated normal distribution with mean θ_0 on the set E

$$\mu_0 \sim a\mathbb{I}_E \mathcal{N}(\theta_0, \mathbb{I}_m).$$

a is chosen such that the above measure is in fact a probability measure.

We use the following mutation kernel for the algorithm

$$\begin{aligned}M(x, dy) &= M(x,x)\delta_x(dy) + \mathbb{I}_E \frac{\exp\left(\frac{-1}{2}(y-x)' A^{-1}(y-x)\right)}{\sqrt{(2\pi)^n |A|}} dy, \\ M(x,x) &= (1 - N(x, \mathbb{I})[E]).\end{aligned}$$

Mixing property of the mutation kernel is one of the fundamental features of the algorithm which is necessary to guarantee the convergence of the algorithm. For a proof that the mutation kernel defined by (4.30) is mixing see Arian *et al.* (2008d).

During crossover, two parameters mix and produce another parameter. Suppose that two input parameters are θ_1 and θ_2. We define the output parameter $\bar{\theta}$ as

$$\bar{\theta} = \frac{G(\theta_2)}{G(\theta_1) + G(\theta_2)} \theta_1 + \frac{G(\theta_1)}{G(\theta_1) + G(\theta_2)} \theta_2.$$

In the selection part of the algorithm, the probability that the individual θ_j^n is selected is $\exp(-\beta_n G(\theta_j))$. Otherwise, another individual θ_j^n is selected with probability

$$\frac{\exp(-\beta_n G(\theta_j^n))}{\sum_{k=1}^{N}(\exp(-\beta_n G(\theta_k^n)))}.$$

This implies that the selection kernel is given by

$$S_t^n(x,dy) = e^{-\beta_n G(x)}\delta_x(dy) + (1-e^{-\beta_n G(x)})\frac{e^{-\beta_n G(x)}\mu(dy)}{\int \mu(dz)e^{-\beta_n G(x)}}.$$

By setting the annealing parameter β_n such that $\beta_n \to \infty$, we make sure the population approaches the regions where the error function is minimized.

The following proposition guarantees the convergence of the evolutionary algorithm to the global minima of error function (4.29). Assume the population is large ($N \sim \infty$) and define the probability measures μ_n^N by

$$\mu_n^N = \frac{1}{N}\sum_{i=1}^{N}\delta_{\theta_n^i} \in M_1(E).$$

Moral and Miclo (2003) show that if $N \to \infty$

$$(\mu_n^N)_{n\geq 0} \xrightarrow{weakly} (\mu_n)_{n\geq 0},$$

where $(\mu_n)_{n\geq 0}$ is given by the recursive equation $\mu_{n+1} = \mu_n M_n S_{\mu_n}^n$, with the selection kernel

$$S_t^n(x,dy) = e^{-\beta_n G(x)}\delta_x(dy) + (1-e^{-\beta_n G(x)})\frac{e^{-\beta_n G(x)}\mu(dy)}{\int \mu(dz)e^{-\beta_n G(x)}}.$$

It can be proved as in Moral and Miclo (2003) that if M is mixing and β_n is increasing, the population concentrates on semi-minimum areas of the error function G. Based on this result we have the following proposition

Proposition 25. *With the above assumptions and definitions of the error function G and the distribution of the population μ_n, we have the following results*

a) The function G is continuous, bounded and of bounded oscillations.

b) The truncated Gaussian kernel $M(x,dy)$ is mixing.

c) $G^* = \inf_K \{\int_E G(\theta)\mu(d\theta); \mu \in M_1(E)$ with $I(\mu) < \infty\} = \inf_E G$, where the infimum is taken over all Markov Kernels $K(x,dy)$ with stationary measure μ.

d) For all $\delta > 0$

$$\mu_n(G(\theta) \geq \inf_E G + \delta) \xrightarrow{n\uparrow\infty} 0.$$

Proof. See Arian *et al.* (2008d) □

For our empirical test, we choose two of the biggest automobile manufacturers in the world, General Motors and Ford. Implementing a structural model for the credit risk analysis of any of the two companies is difficult because the true level of liabilities for them is not clear from their balance sheets. General Motors and Ford have corresponding financial subsidiaries named General Motors Acceptance Corporation and Ford Motors Credit Company, respectively. Most of the bonds issued by these two financial institutions are secured and therefore the true level of liabilities for each company is only a percentage of the amount reflected by their balance sheet. Stamicar and Finger (2005) conduct a detailed study on both companies specially during year 2005 and argue that for the above reason, the model debt per share should be considered as $20-25\%$ of the overall debt per share. We consider this assumption for our empirical study below. Besides of the level of liabilities for each company, Stamicar points out that in May 2005, S&P downgraded both companies from grade BBB to BB. During this period, the correlation between CDS spread of both companies increased rapidly and then it came back to its normal mean. This is in consistency with our expectation of the change in correlation during crisis periods.

Using data from Bloomberg for values of 64 call options on General Motors' equities and 64 call options on Ford's equities, we calibrate the parameters of our stochastic correlation CreditGrades model. The data we use for empirical study is given in tables A.1 and A.4 in the appendix A. Table 4.2 shows the estimated parameters from General Motor's and Ford's data.

Table 4.2: Parameters of the Wishart Process

Wishart Covariance	Gindikin's Coefficient	Autoregressive Matrix		Wishart's Volatility Matrix		Initial Covariance	
Parameters	$\beta = 4.447$	$m_{11} = -185.015$	$m_{12} = 0.011$	$q_{11} = 0.025$	$q_{12} = 0$	$\Sigma_{11} = 0.025$	$\Sigma_{12} = 0$
		$m_{21} = 0.011$	$m_{22} = -52.53$	$q_{21} = 0$	$q_{22} = 0.017$	$\Sigma_{21} = 0$	$\Sigma_{22} = 0.017$

The error function we used for the purpose of our calibration is

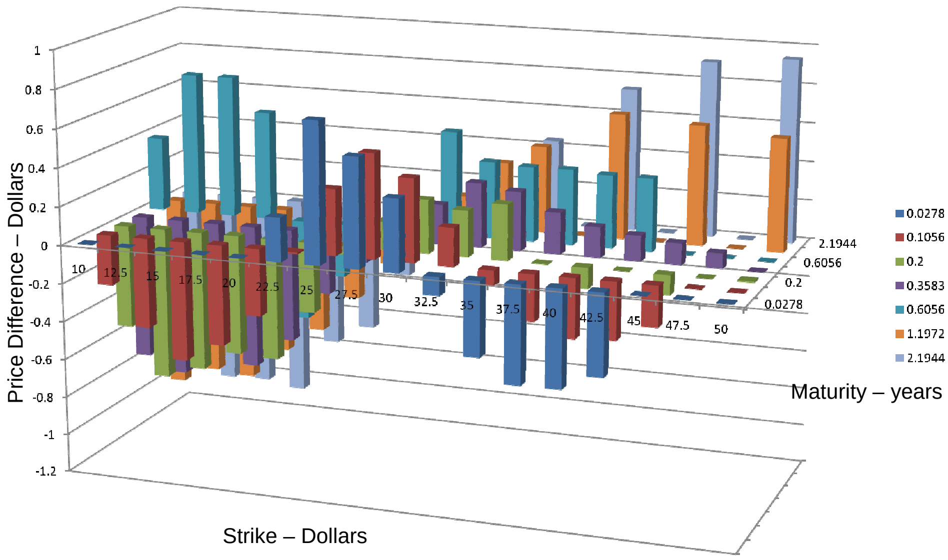

Figure 4.4: The difference between market prices and stochastic correlation model's results for General Motors's equity options in U.S. dollars as a function of strike price and maturity date. The maturity dates range from $T = 0.0278$ years to $T = 2.1944$ years. The strike prices range from $K = 10$ dollars to $K = 50$ dollars. At the time of analysis, the value of General Motors' equity is $S_0 = \$25.86$

$$
\begin{aligned}
G(\theta) \quad = \quad & \sum_{i=1}^{I} \omega_i^2 |C_i^\theta(t, T_i) - C_i^*(t, T_i)|^2 \\
+ \quad & \sum_{i=1}^{I} \omega_i \nu_i |C_i^\theta(t, T_i).D_i^\theta(t, T_i) - C_i^*(t, T_i).D_i^*(t, T_i)| \\
+ \quad & \sum_{i=1}^{I} \nu_i^2 |D_i^\theta(t, T_i) - D_i^*(t, T_i)|^2 .
\end{aligned}
\tag{4.30}
$$

This implies a price difference for each maturity-strike node for General Motors and Ford. The price difference generated for General Motors is shown in figure 4.4. In Sepp (2006), figures 7.1, 7.2 and 7.3 show the corresponding error terms for the three models studied in the paper, as a function of strike and maturity. Our model shows improvements with respect to the regular diffusion model, the stochastic variance model and the double-exponential jump diffusion model. The stochastic correlation model gives a slightly better fit to the market data compared to the single factor stochastic volatility model introduced in Sepp (2006). We have an error term about 5% less than the regular diffusion model and 2% less than the stochastic variance model and the double-exponential jump diffusion model. The implied volatility surface corresponding to the parameters for the General Motors equity options is shown in figure 4.5. As it can be seen, the stochastic correlation CreditGrades model is capable of reproducing volatility smile and skew.

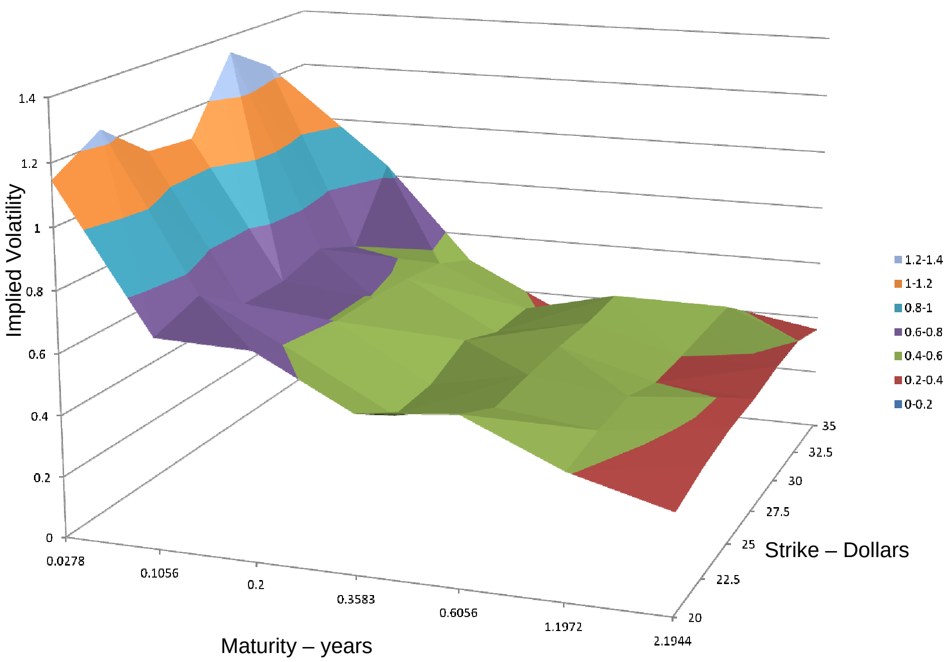

Figure 4.5: Implied Volatility surface of the stochastic correlation model

4.8 Conclusion

We presented a structural credit risk model which considers stochastic correlation between the assets of the companies. The stochasticity of the volatility and correlation come from a Wishart process which drives the covariance matrix of the assets. To model credit risk, we use the so called CreditGrades model. Using the affine properties of the joint log-price and volatility process, we solved the pricing problem of the equity options. We used our analytical techniques to derive closed form solution for equity options, probabilities of defaults and prices of CDSs issued by the companies. After solving the pricing problem for our stochastic correlation credit risk model, we used our theoretical results for the equity option prices to calibrate the parameters of the model. The calibration problem is done via a probabilistic approach called the evolutionary algorithm. The calibration problem we deal with is an ill-posed inverse problem. We implemented the algorithm's components to fit the needs of the particular problem here. We showed that by putting certain conditions on the components of the algorithm, the parameters population converges to the minimum of the error function which matches the market prices with our suggested theoretical prices. The results we acquire are close to market observations and our model is capable of reflecting the volatility smile and skews.

Chapter 5

Implementation of the Default Correlations

We conduct a comparative analysis of two stochastic correlation models studied in Arian *et al.* (2008d,b) and the standard CreditGrades model with constant covariance structure. Getting a closed form solution for the joint probability of default is difficult because of the nature of the mathematical problem involved. This problem involves a treatment of the infinitesimal operator of the assets' prices and their covariance. Zhou (2001) finds a closed form formula for the joint probability of default for a model with constant covariance. To find closed form solutions for the marginal probabilities, the standard independence assumption between the asset and its volatility is considered as in Lipton (2001). In Arian *et al.* (2008d,b), we have extended the CreditGrades model with two stochastic correlation structures and solved the pricing and the calibration problems. In this paper, we use the results of the calibration algorithms from Arian *et al.* (2008d,b) and give an approximation formula for the joint and marginal one year probabilities of default for General Motors and Ford. For the *marginal* probabilities of default, a closed formula is given in Arian *et al.* (2008d,b). For the *joint* probabilities of default an approximation formula has been given in Arian *et al.* (2008a). To show the convergence properties of this approximation method, we perform the Monte Carlo simulation in two forms; a full and a partial monte carlo simulation. At the end, we compare the marginal and joint probabilities with full and partial monte carlo simulations for General Motors and Ford.

5.1 Introduction

Credit Derivatives have been traded in large scales in the last several years and they have played an important role in the recent financial crisis. Some of credit derivatives, like CDS's, are dependent on the credit structure of only one credit entity while some other forms of the credit derivatives like Collateralized Debt Obligations (CDOs) depend on the credit valuation of more than one company. The second type of credit derivatives imply the need for calculation of the credit correlation between all firms within the credit portfolio. Based on our knowledge, all of the credit risk models assume constant correlation among the underlying firms. Analytical representation of the correlated default is a very difficult mathematical problem which has been worked out in Rebholz (1994) and Zhou (2001) in the case of constant correlation in two dimensions. Zhou (2001) suggests an analytical method to calculate correlated default in the first-passage-time framework developed by Black and Cox (1976). The first-passage-time model relaxes the highly restrictive assumption that default can only occur on a pre-specified time in future. Besides of the constant correlation assumption in that paper, their mathematical method has this drawback that can not be extended to dimensions more than two. On the other hand, Patras (2005) uses the reflection principle which can hardly be implemented for stochastic correlation models developed in the financial literature. Here we explore the theoretical and numerical results of the dynamics of stochastic dependence structure of credit risk models developed in Arian *et al.* (2008d) and Arian *et al.* (2008b) which have extended one of the most popular structural credit risk models in the market, the CreditGrades model, to accommodate stochastic correlation.

The structure of this chapter is as follows: In section 5.2 we discuss the works done related to calculation of constant correlated defaults. In section 5.3 we develop a method to evaluate stochastic correlated default probabilities with an insight to value multi-name credit derivatives. Finally, section 5.3 is devoted to implementation of the joint default probability for the constant correlation case and the models developed in Arian Arian *et al.* (2008d,b).

5.2 Constant Default Correlation

Implementation of correlated defaults is difficult because the correlation is not observable and most of the exchange traded credit derivatives in the market do not contain joint default information. In the case of constant correlation, joint probability of default for two obligers has been derived by Zhou (2001), He *et al.* (1998) and Rebholz (1994). In a first-passage-time framework, The joint probability of default for two obligers which are constantly correlated has been given by Zhou (2001). Prior to this paper, the probability of default had been studied only for a single firm in first-passage-time models. The first drawback of results obtained by Zhou is that the correlation is assumed constant. Secondly, the approach is not extendable for a study on more than two firms. Assume two firms with default indicators $D_i(t)$ for $i = 1, 2$. If the i^{th} firm default by time t, $D_i(t) = 1$ and if it doesn't $D_i(t) = 0$. Therefore we have

$$D_i(t) = \begin{cases} 1 & \text{if firm } i \text{ defaults by } t, \\ 0 & \text{otherwise.} \end{cases}$$

In the case that two companies have independent default indicators, $P(D_1(t) = 1 \text{ and } D_2(t) = 1) = P(D_1(t) = 1).P(D_2(t) = 1)$ which implies an independent default correlation. The general definition of the default correlation is given below.

Definition 4. *For two companies with default indicators $D_1(t)$ and $D_2(t)$, the default correlation is defined as*

$$Corr[D_1(t), D_2(t)] \equiv \frac{E[D_1(t).D_2(t)] - E[D_1(t)].E[D_2(t)]}{\sqrt{Var\,[D_1(t)].\ Var\,[D_2(t)]}}. \tag{5.1}$$

The default correlation between companies A and B up to time t is defined as

$$\begin{aligned}\rho_D(t) &= Corr[D_A(t), D_B(t)] \\ &= \frac{P[D_A(t).D_B(t)] - P[D_A(t)].P[D_B(t)]}{[P(D_A(t))(1 - P(D_A(t)))]^{\frac{1}{2}}.[P(D_B(t))(1 - P(D_B(t)))]^{\frac{1}{2}}} \\ &= \frac{P[D_A(t)] + P[D_B(t)] - P[D_A(t) + D_B(t) + D_B(t)] - P[D_A(t)].P[D_B(t)]}{[P(D_A(t))(1 - P(D_A(t)))]^{\frac{1}{2}}.[P(D_B(t))(1 - P(D_B(t)))]^{\frac{1}{2}}},\end{aligned}$$

where $P[D_A(t)]$ and $P[D_B(t)]$ are the probabilities that the company A and B default, respectively, and $P[D_A(t) + D_B(t)]$ is the probability that at least one of the firms default. One can show that the joint probability of default can be written in terms of the default correlation and marginal probabilities of default

$$E[D_1(t).D_2(t)] = E[D_1(t)].E[D_2(t)] + Corr[D_1(t), D_2(t)].\sqrt{\text{Var }[D_1(t)].\text{ Var }[D_2(t)]}.$$

The formulation of the default correlation in the framework of the first-passage-time model is related to hitting time of a two dimensional diffusion process to a certain boundary determined by liabilities of the companies. If the asset of the i^{th} firm is denoted by V_i, then the constant correlation first-passage-time model assumes that

$$\begin{pmatrix} d\ln(V_1) \\ d\ln(V_2) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} dt + \Omega \begin{pmatrix} dz_1 \\ dz_2 \end{pmatrix}, \tag{5.2}$$

where μ_i is the drift term for the i^{th} firm and the matrix $\Omega\Omega'$ represents the covariance matrix of the assets dynamics and is given by

$$\Omega.\Omega' = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

where all the coefficients are constant and ρ is the *constant* correlation between the assets prices. Each firm defaults if the value of its asset falls below the recovery part of the overall liability of the company. This quantity which is denoted here by $C_i(t)$, follows a simple dynamic

$$C_i(t) = K_i e^{\lambda_i t},$$

based on the Black and Cox (1976) assumption. In the equation above, λ_i represents the growth rate of the i^{th} firm's overall liability. In this case, the default time is defined as the stopping time

$$\tau_i = \min_{t \geq 0}\{t | e^{-\lambda_i t} V_i(t) \leq K_i\}.$$

The marginal default probability in the above framework is given by

$$P[D_i(t) = 1] = N(-\frac{Z_i}{\sqrt{t}} - \frac{\mu_i - \lambda_i}{\sigma_i}\sqrt{t}) + e^{\frac{2(\lambda_i - \mu_i)Z_i}{\sigma_i}} N(-\frac{Z_i}{\sqrt{t}} + \frac{\mu_i - \lambda_i}{\sigma_i}\sqrt{t}).$$

Even though calculating the marginal default probability given by the above equation is not challenging, finding the joint probability of default including the probability that both companies default $P[D_1(t) = 1 \text{ and } D_2(t) = 1]$, and the probability that at least one firm defaults $P[D_1(t) = 1 \text{ or } D_2(t) = 1]$, is a difficult mathematical problem. The difficulty of this problem is the main reason we know little about stochastic correlated defaults. Here we intend to give a method Rebholz (1994) has offered to calculate the joint probability of default with constant correlation in the simple case where the growth rate of the assets coincides with the growth rate of the liability(i.e. $\lambda_i = \mu_i$). Defining $\tau = \min\{\tau_1, \tau_2\}$, we have

$$P[D_1 = 1 \text{ or } D_2 = 1] = P[\tau_1 \le t \text{ or } \tau_2 \le t] = P[\tau \le t].$$

Now let's transform asset prices by

$$X_i(t) = -\ln[e^{-\lambda_i t} V_i(t)/V_i(0)]. \tag{5.3}$$

The process $(X_1(t), X_2(t))$ is a two dimensional Brownian motion and follows

$$\begin{pmatrix} dX_1 \\ dX_2 \end{pmatrix} = -\Omega \begin{pmatrix} dz_1 \\ dz_2 \end{pmatrix}, \tag{5.4}$$

with initial conditions $X_1(0) = 0$ and $X_2(0) = 0$. Defining $b_i = -\ln[K_i/V_i(0)]$, the problem of finding the joint probability of default is equivalent to the problem of calculating the probability that the process $(X_1(t), X_2(t))$ hits the barrier (b_1, b_2). Let's denote the probability density function of the process $(X_1(t), X_2(t))$ prior to stopping time τ by $f(x_1, x_2, t)$. Then we have

$$\begin{aligned} F(y_1, y_2, t) &= P[X_1(t) < y_i \text{ and } X_2(t) < y_2 | X_1(s) < b_1 \text{ and } X_2(s) < b_2, \text{ for } 0 < s < t] \\ &= \int_{-\infty}^{y_1} \int_{-\infty}^{y_2} f(x_1, x_2, t) dx_1 dx_2. \end{aligned}$$

In this case $F(b_1, b_2, t) = P[\tau > 1] = 1 - P[\tau \le t]$ refers to the possibility that hitting has not occurred prior to time t. Now the problem is to find an expression for the density function $f(x_1, x_2, t)$. The density function satisfies the Kolmogorov's forward equation

$$\frac{\sigma_1^2}{2}\frac{\partial^2 f}{\partial x_1^2} + \rho\sigma_1\sigma_2 \frac{\partial^2 f}{\partial x_1 \partial x_2} + \frac{\sigma_2^2}{2}\frac{\partial^2 f}{\partial x_2^2} = \frac{\partial f}{\partial t} \quad (x_1 < b_1, x_2 < b_2), \tag{5.5}$$

with the boundary conditions

$$\begin{aligned} & f(-\infty, x_2, t) = f(x_1, -\infty, t) = 0, \\ & f(x_1, x_2, 0) = \delta(x_1)\delta(x_2), \\ & \int_{-\infty}^{b_1} \int_{-\infty}^{b_2} f(x_1, x_2, t) dx_1 dx_2 \le 1 \qquad \text{for } t > 0, \\ & f(b_1, x_2, t) = f(x_1, b_2, t) = 0. \end{aligned} \tag{5.6}$$

To solve the PDE (5.5) with boundary conditions (5.6), we first change the variables by

$$
\begin{aligned}
u_1 &= \frac{x_1}{\sigma_1}, \\
u_2 &= \frac{1}{\sqrt{1-\rho^2}}(\frac{x_2}{\sigma_2} - \rho\frac{x_1}{\sigma_1}).
\end{aligned}
$$

This eliminates the joint partial differential term and transform the partial differential equation (5.5) to

$$
\frac{1}{2}\frac{\partial^2 f}{\partial u_1^2} + \frac{1}{2}\frac{\partial^2 f}{\partial u_2^2} = \frac{\partial f}{\partial t}. \tag{5.7}
$$

With this transformation, the hitting barrier becomes

$$
\begin{aligned}
u_1 &= \frac{b_1}{\sigma_1}, \\
u_2 &= \frac{1}{\sqrt{1-\rho^2}}(\frac{b_2}{\sigma_2} - \rho\frac{u_1}{\sigma_1}).
\end{aligned}
$$

By the series of transformations,

$$
\begin{aligned}
v_1 &= u_1 - \frac{b_1}{\sigma_1}, \\
v_2 &= u_2 - \frac{1}{\sqrt{1-\rho^2}}(\frac{b_2}{\sigma_2} - \rho\frac{u_1}{\sigma_1}),
\end{aligned}
$$

and

$$
\begin{aligned}
w_1 &= -(\sqrt{1-\rho^2})v_1 + \rho v_2, \\
w_2 &= -\rho v_1 - (\sqrt{1-\rho^2})v_2,
\end{aligned}
$$

the PDE (5.7) remains unchanged and the hitting barrier entails the simpler form

$$
\begin{aligned}
w_1 &= 0, \\
w_2 &= -(\frac{\rho}{\sqrt{1-\rho^2}})w_1.
\end{aligned}
$$

The above transformations yield the following relation between initial variables x_1, x_2 and final variables w_1, w_2

$$x_1 = b_1 - \sigma_1[(\sqrt{1-\rho^2})w_1 + \rho w_2], \tag{5.8}$$

$$x_2 = b_2 - \sigma_2 w_2. \tag{5.9}$$

By change of variables $w_1 = r\cos(\theta)$ and $w_2 = r\sin(\theta)$, the PDE (5.7) changes to

$$\frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2} + \frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} = 2\frac{\partial f}{\partial t}, \tag{5.10}$$

with the boundary conditions

$$f(r,0,t) = f(r,\alpha,t) = f(\infty,\theta,t) = 0, \tag{5.11}$$
$$f(r,\theta,0) = \delta(r-r_0)\delta(\theta-\theta_0),$$
$$\int_{\theta=0}^{\alpha}\int_{r=0}^{\infty} J.f(r,\theta,t)drd\theta \leq 1, t > 0,$$

where

$$x_1(r,\theta) = b_1 - \sigma_1[(\sqrt{1-\rho^2})r\cos(\theta) + \rho r\sin(\theta)], \tag{5.12}$$

$$x_2(r,\theta) = b_2 - \sigma_2 r\sin(\theta), \tag{5.13}$$

$$\alpha = \begin{cases} \tan^{-1}(-\frac{\sqrt{1-\rho^2}}{\rho}) & if \rho < 0 \\ \pi + \tan^{-1}(-\frac{\sqrt{1-\rho^2}}{\rho}) & otherwise, \end{cases} \tag{5.14}$$

$$J = r\sigma_1\sigma_2\sqrt{1-\rho^2}. \tag{5.15}$$

By Zhou (2001), the solution of the above PDE is given by

$$f = \frac{2r}{J.\alpha.t} e^{-\frac{r^2+r_0^2}{2t}} \sum_{n=1}^{\infty} \sin(\frac{n\pi\theta}{\alpha}) \sin(\frac{n\pi\theta_0}{\alpha}) I_{\frac{n\pi}{\alpha}}(\frac{rr_0}{t}). \tag{5.16}$$

This gives the following theorem

Theorem 12. *The solutions of the PDE* (5.5) *with boundary conditions* (5.6) *is given by*

$$f(x_1,x_2,t) = \frac{2}{\sigma_1\sigma_2\sqrt{1-\rho^2}\alpha t} e^{-\frac{r^2+r_0^2}{2t}} \times \sum_{n=1}^{\infty} \sin(\frac{n\pi\theta}{\alpha}) \sin(\frac{n\pi\theta_0}{\alpha}) I_{\frac{n\pi}{\alpha}}(\frac{rr_0}{t}), \tag{5.17}$$

where

$$x_1 = b_1 - \sigma_1[(\sqrt{1-\rho^2})r\cos(\theta) + \rho r\sin(\theta)],$$
$$x_2 = b_2 - \sigma_2 r\sin(\theta).$$

Now we just need to plug in the probability density function to find $F(b_1, b_2, t)$

$$\begin{aligned} F(b_1, b_,t) &= \int_{-\infty}^{b_1}\int_{-\infty}^{b_2} f(x_1, x_2, t)dx_1dx_2 \\ &= \int_{\theta=0}^{\infty}\int_{r=0}^{\infty} J.f(r,\theta,t)d\theta dr \\ &= \frac{2}{\alpha.t}e^{-\frac{r_0^2}{2t}}\sum_{n=1}^{\infty}\sin(\frac{n\pi\theta_0}{\alpha})\int_{\theta=0}^{\alpha}\sin(\frac{n\pi\theta}{\alpha})d\theta\int_{r=0}^{\infty} re^{-\frac{r^2}{2t}}I_{\frac{n\pi}{\alpha}}(\frac{rr_0}{t})dr. \end{aligned}$$

Using the well-known identities

$$\begin{aligned} \int_{\theta=0}^{\alpha}\frac{n\pi}{\alpha}\sin(\frac{n\pi\theta}{\alpha})d\theta &= 1-(-1)^n, \\ \int_0^{\infty} re^{-c_1r^2}I_v(c_2r)dr &= \frac{c_2}{8c_1}\sqrt{\frac{c_2^2}{8c_1}}e^{\frac{c_2^2}{8c_1}}[I_{\frac{1}{2}(v+1)}(\frac{c_2^2}{8c_1}) + I_{\frac{1}{2}(v-1)}(\frac{c_2^2}{8c_1})]. \end{aligned}$$

The probability that at least one firm defaults prior to time t reads

$$P[D_1(t)=1 \text{ or } D_2(t)=1] = 1-\frac{2r_0}{\sqrt{2\pi t}}.e^{-\frac{r_0^2}{4t}}.\sum_{n=1,3,\ldots}\frac{1}{n}.\sin(\frac{n\pi\theta_0}{\alpha}).[I_{\frac{1}{2}(\frac{n\pi}{\alpha}+1)}(\frac{r_0^2}{4t})+I_{\frac{1}{2}(\frac{n\pi}{\alpha}-1)}(\frac{r_0^2}{4t})], \tag{5.18}$$

where I_ν is the modified Bessel function of order ν and the coefficients of the above expression are given by

$$\alpha = \begin{cases} \tan^{-1}(-\frac{\sqrt{1-\rho^2}}{\rho}) & if\ \rho < 0 \\ \pi + \tan^{-1}(-\frac{\sqrt{1-\rho^2}}{\rho}) & otherwise. \end{cases}$$
$$\theta_0 = \begin{cases} \tan^{-1}(-\frac{Z_2\sqrt{1-\rho^2}}{Z_1-\rho Z_2}) & if\ (.) > 0 \\ \pi + \tan^{-1}(-\frac{Z_2\sqrt{1-\rho^2}}{Z_1-\rho Z_2}) & otherwise. \end{cases}$$
$$r_0 = Z_2/\sin(\theta_0).$$

This is worthwhile to mention that in this simple case, the marginal probability of default is given by $P[D_i(t) = 1] = 2.N(-\frac{\ln(V_{i,0}/K_i)}{\sigma_i\sqrt{t}})$. In the more general case where $\lambda_i \neq \mu_i$, the joint probability of default is given by the proposition below

Proposition 26. *In the framework of the constant correlation first-passage-time model, the probability that at least one firm defaults is given by*

$$P[D_1(t) = 1 \ or \ D_2(t) = 1] = 1 - \frac{2}{\alpha t} e^{a_1 x_1 + a_2 x_2 + a_t t}. \sum_{n=1}^{\infty} \sin(\frac{n\pi\theta_0}{\alpha}).e^{-\frac{r_0^2}{2t}} \int_0^{\alpha} \sin(\frac{n\pi\theta}{\alpha}) g_n(\theta) d\theta,$$

where the function g_n is given by

$$g_n(\theta) = \int_0^{\infty} r.e^{-\frac{r^2}{2t}}.e^{d_1 r \sin(\theta-\alpha) - d_2 r \cos(\theta-\alpha)}.I_{\frac{n\pi}{\alpha}}(\frac{r r_0}{t}) dr,$$

and the coefficients of the above representation are given by

$$\begin{aligned}
\alpha &= \begin{cases} \tan^{-1}(-\frac{\sqrt{1-\rho^2}}{\rho}) & if \ \rho < 0 \\ \pi + \tan^{-1}(-\frac{\sqrt{1-\rho^2}}{\rho}) & otherwise. \end{cases} \\
\theta_0 &= \begin{cases} \tan^{-1}(-\frac{Z_2\sqrt{1-\rho^2}}{Z_1 - \rho Z_2}) & if \ (.) > 0 \\ \pi + \tan^{-1}(-\frac{Z_2\sqrt{1-\rho^2}}{Z_1 - \rho Z_2}) & otherwise. \end{cases} \\
r_0 &= Z_2 / \sin(\theta_0). \\
a_1 &= \frac{(\lambda_1 - \mu_1)\sigma_2 - (\lambda_2 - \mu_2)\rho\sigma_1}{(1-\rho^2)\sigma_1^2\sigma_2}, \\
a_2 &= \frac{(\lambda_2 - \mu_2)\sigma_1 - (\lambda_1 \mu_1)\rho\sigma_2}{(1-\rho^2)\sigma_2^2\sigma_1}, \\
a_t &= \frac{a_1^2\sigma_1^2}{2} + \rho a_1 a_2 \sigma_1 \sigma_2 + \frac{a_2^2\sigma_2^2}{2} - a_1(\lambda_1 - \mu_1) - a_2(\lambda_2 - \mu_2), \\
d_1 &= a_1\sigma_1 + \rho a_2 \sigma_2, \\
d_2 &= a_2 \sigma_2 \sqrt{1-\rho^2}.
\end{aligned}$$

Proof. See Rebholz (1994) □

Remark 3. *(Comparison with the Merton model)*

Merton model assumes that the default can only happen at a single point in time. To see the impact this assumption has on the probabilities of default, we look at the marginal probability of default for the i^{th} firm assuming that $\lambda_i = \mu_i$ under the Merton's model

$$P[D_i(t)=1] = P[V_i(t) \leq C_i(t)] = N(-\frac{Z_i}{\sqrt{t}}), \tag{5.19}$$

whereas the marginal probabilities of default in the Black-Cox framework is given by

$$P[D_i(t)=1] = 2.N(-\frac{\ln(V_{i,0}/K_i)}{\sigma_i\sqrt{t}}).$$

This is interesting to note that the marginal probability of default in the Marton's framework is half the marginal probability of default in the Black-Cox framework. This justifies Jones' argument in Jones et al. *(1984) that Merton model underestimates the probability of default. On the other hand, Merton model has a simple structure and the correlated default can be easily found by integrating with respect to the probability density of a two dimensional Gaussian variable as*

$$\begin{aligned} P[D_1(t)=1 \text{ and } D_2(t)=1] &= P[V_1(t) \leq C_1(t) \text{ and } V_2(t) \leq C_2(t)] \\ &= \int_{-\infty}^{Z_1/\sqrt{t}} \int_{-\infty}^{Z_2/\sqrt{t}} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp[\frac{x_1^2 - 2\rho x_1 x_2 + x_2^2}{2(1-\rho^2)}] dx_1.dx_2. \end{aligned}$$

5.3 Stochastic Default Correlation

The characteristic function of a stochastic process completely defines its probability distribution and its probability density function. If $Y_t = \ln A_t$, the characteristic function of Y_t is defined as

$$\varphi_{Y_t}(\lambda) = E_t(e^{i\lambda Y_t}) = \int_{\mathbb{R}} e^{i\lambda y} f_{Y_t}(y)\, dy.$$

Similarly, the multivariate characteristic function of a random variable Y_t which takes its values in $\mathbb{R}^n$ is defined for the vector $\lambda \in \mathbb{R}^n$

$$\varphi_{Y_t}(\lambda) = E_t(e^{i\lambda.Y_t}) = \int_{\mathbb{R}^n} e^{i\lambda.y} f_{Y_t}(y)\, dy.$$

The *characteristic function* of a stochastic process is the complex conjugate of its *Fourier transform*

$$\varphi_{Y_t}(\lambda) = \int_{\mathbb{R}^n} e^{i\lambda.y} f_{Y_t}(y)\, dy = \overline{\int_{\mathbb{R}^n} e^{-i\lambda.y} f_{Y_t}(y)\, dy} = \overline{\psi_{Y_t}(\lambda)},$$

where $\psi_{Y_t}(t)$ is the fourier transform of the stochastic process Y_t. In both of the cases above, f_{Y_t} can be recovered either from its characteristic function or its fourier transform

$$\begin{aligned} f_{Y_t}(y) &= \frac{1}{2\pi}\int_{\mathbb{R}} e^{-i\lambda y}\varphi_{Y_t}(y)\,dy \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} e^{-i\lambda y}\overline{\psi_{Y_t}(y)}\,dy. \end{aligned}$$

Characteristic function and fourier tranform have strong applications in option pricing theory which result in a so-called convolution based method which switches the computation from asset pricing domain to the frequency domain. By Feynman-Kac theorem, the value of an option with payoff $\Phi(S)$ at time T is given by

$$V(t,S_0) = \mathbb{E}^Q_{t,S_0}\left(e^{-\int_t^T r(s)\,ds}\Phi(S(T)\right).$$

By the change of variable $S = e^x$, one can rewrite the above equation as

$$V(t,x) = e^{-\int_t^T r(s)\,ds}\int \Phi(e^{x_T})f(x_T|x,\Sigma)\,dx_T,$$

where $f(x_T|x,\Sigma)$ is the density function of the log-price $x_t = \ln S_t$. The density function $f(x_T)$ is very difficult to obtain. However, its characteristic function and fourier transform have nice analytical formulas.

Theorem 13 (Feynman-Kac Representation)**.** *Assume Φ is a non-negative function and $V(t,x,y) \in C^{1,2,2}$ is a solution for the following PDE with bounded derivatives*

$$\begin{cases} \frac{\partial V}{\partial t} + \frac{1}{2}\sigma_1^2x^2\frac{\partial^2 V}{\partial x^2} + \rho\sigma_1\sigma_2 xy\frac{\partial^2 V}{\partial xy} + \frac{1}{2}\sigma_2^2y^2\frac{\partial^2 V}{\partial y^2} + \mu_1 x\frac{\partial V}{\partial x} + \mu_2 y\frac{\partial V}{\partial y} - rV = 0, \\ V(T,x,y) = \Phi(x,y). \end{cases}$$

Then $V(t,x,y)$ has the diffusion representation

$$V(t,x,y) = \mathbb{E}_{t,x,y}\left(e^{-\int_t^T r(s)\,ds}\Phi(S(T))\right).$$

Proof. See Karatzas and Shreve (1998) $\square$

For various choices of the terminal function ϕ we get different multi-dimensional derivatives. For instance, if $\Phi(x,y) = (x-y-K)^+$, $V(t,x,y)$ is the value of a spread option, if $\Phi(x,y) = (x-y)^+$, $V(t,x,y)$ is the value of an exchange option, and if

$\Phi(x, y) = (x - K_1)^+(y - K_2)^+$, $V(t, x, y)$ is the value of a correlation option. All of the above derivatives's values satisfy the PDE

$$W_t + \mathcal{A}_{(\Sigma,S)} - rW = 0,$$

but with different initial conditions. In fact, it is well-known that for the price of correlation options and exchange options, there are closed form formulas in the presence of constant volatility but there are no closed form formulas for the price of spread options available.

5.3.1 Joint Probability of Survival/Default

The joint probability of default in the presence of constant volatility and correlation has been studied in Rebholz (1994); He *et al.* (1998) and Patras (2005). These papers show how complicated the problem could be even when the correlation is deterministic. In this section, we give an asymptotic technique to find probabilities of default. We give approximations which converge to the actual joint probabilities of default. Let's define

$$\Lambda(\tau, \Sigma, S) = \mathbb{Q}_{(\tau,\Sigma,S)}(S(t + \tau) >> 0), \tag{5.20}$$

where $\mathbb{Q}$ is the risk-neutral measure, $S = (S_1, S_2)$ and we use the notation

$$\{S >> 0\} \iff \{S_1 > 0 \text{ and } S_2 > 0\}.$$

Here we emphasize that since our model is time-homogenous, Λ is time-homogenous as well (i.e. it does not depend on the initial time t). Denote by $P[t, T]$ the joint survival probability of the firms

$$P[t, T] = \mathbb{Q}_{(\tau,\Sigma,S)}(S(u) >> 0 \text{ for all } u \in [t, T]). \tag{5.21}$$

We approximate $P[t, T]$ by dividing the interval $[t, T]$ into n pieces

$$t = T_0 < T_1 < \ldots < T_n = T.$$

Define A_j to be the event that $\{S(T_j) >> 0\}$. We give a closed form formula for $P_{(t,\Sigma,S)}(A_1, \ldots, A_n)$ and we prove the convergence $P_{(t,\Sigma,S)}(A_1, \ldots, A_n) \xrightarrow{n \to +\infty} P[t, T]$.

Proposition 27. *With the above conditions and definitions*

$$P_{(t,\Sigma,S)}(A_1,\dots,A_n) = \pi_1\pi_2\dots\pi_n,$$

where

$$\pi_j = \int_{\Sigma(T_{j-1})}\int_{S(T_{j-1})>>0} \Lambda(T_j - T_{j-1}, \Sigma(T_{j-1}), S(T_{j-1}))\, dP_{(t,\Sigma,S)}(dS_{j-1}\Sigma_{i-1}),$$

where Λ *is given by equation* (5.20).

Proof. The events $\{A_j\}_1^n$ are not independent, but if we define

$$B_j = A_j|(A_i)_{i=1}^{j-1}, (\Sigma_i)_{i=1}^{j-1}.$$

Then one can check that

$$\begin{aligned} P_{(t,\Sigma,S)}(A_1,\dots,A_n) &= P(A_1).\frac{P(A_1,A_2)}{P(A_1)}\dots\frac{P(A_1,A_2,\dots A_n)}{P(A_1,A_2,\dots A_{n-1})} \\ &= P(B_1)P(B_2)\dots P(B_n). \end{aligned}$$

The following line completes the proof

$$\begin{aligned} \pi_j &= P_{(t,\Sigma,S)}(B_j) = P_{(t,\Sigma,S)}(A_j|(A_i)_{i=1}^{j-1}, (\Sigma_i)_{i=1}^{j-1}) = P_{(t,\Sigma,S)}(A_j|A_{j-1}, \Sigma_{j-1}) \\ &= \int_{\Sigma(T_{j-1})}\int_{S(T_{j-1})>>0} \Lambda(T_j - T_{j-1}, \Sigma(T_{j-1}), S(T_{j-1}))\, dP_{(t,\Sigma,S)}(dS_{j-1}\Sigma_{i-1}), \end{aligned}$$

Where P is the probability distribution of (S,Σ) □

Proposition 28. *The events* $\{A_1, A_2, \dots, A_n\}$ *converge in probability to the continuous survival event. In other words*

$$P_{(t,\Sigma,S)}(A_1,\dots,A_n) \xrightarrow{n\to+\infty} P(t,T).$$

Proof. Assume the subintervals $\mathcal{I}_n = \{T_0^n < T_1^n < \dots < T_n^n\}$ such that $T_j = t + \frac{j}{n}(T-t)$. Let's define

$$\begin{aligned} C_N &= \{\omega \in \Omega | \exists N \text{ and } \exists j \leq N \text{ such that } S_{T_j^n} << 0\}, \\ C &= \{\omega \in \Omega | \exists u \in [t,T] \text{ such that } S_u << 0\}. \end{aligned}$$

Since the trajectories of S_t are continuous, for almost any $\omega \in C$, there is an N such that $\omega \in C_N$. Therefore, $C_N \uparrow C$. Since $P(C) < \infty$, we have $P(C_N) \uparrow P(C)$ or

$$P_{(t,\Sigma,S)}(A_1, \ldots, A_n) \xrightarrow{n\to+\infty} P(t,T).$$

□

5.4 Financial Applications

In this section, we find the default correlation of three models based on constant correlation, Wishart correlation and principal component correlation by Monte Carlo simulation. The monte carlo simulation is done in the CreditGrades framework which has been studied in Stamicar and Finger (2005), Sepp (2006) and Arian *et al.* (2008c,d).

5.4.1 Correlated Default for the Standard CreditGrade model with Constant Covariance Matrix

The formulation of the default correlation in the framework of the first-passage-time model is related to the hitting time of a two dimensional diffusion process to a certain boundary related to liabilities of the companies. If the asset of the i^{th} firm is denoted by A_i, then the constant correlation CreditGrades model assumes that

$$\begin{pmatrix} d\ln(V_1) \\ d\ln(V_2) \end{pmatrix} = \begin{pmatrix} r_1(t) - d_1(t) \\ r_2(t) - d_2(t) \end{pmatrix} dt + \begin{pmatrix} \sqrt{1-\rho^2}\sigma_1 & \rho\sigma_1 \\ 0 & \sigma_2 \end{pmatrix} \begin{pmatrix} dz_1 \\ dz_2 \end{pmatrix}, \tag{5.22}$$

where $\Omega\Omega'$ represents the covariance matrix of the assets dynamics and is given by

$$\Omega.\Omega' = \begin{pmatrix} \sqrt{1-\rho^2}\sigma_1 & \rho\sigma_1 \\ 0 & \sigma_2 \end{pmatrix} \begin{pmatrix} \sqrt{1-\rho^2}\sigma_1 & 0 \\ \rho\sigma_1 & \sigma_2 \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

where all the coefficients, including the correlation ρ, are constant. For each firm, the liability per share follows the dynamics $dB_t = (r(t) - d_i(t))B_t dt$, where $r(t)$ is the risk free interest rate and $d_i(t)$ is the dividend yield for i^{th} firm. Therefore the liability per share satisfies

$$B(t) = B(0)\exp(\int_0^t r(s) - d_i(s)\, ds).$$

Assuming the recovery rate for the i^{th} firm as R_i, recovery part of the debt per share is given by $D_i(t) = R_i B_i(t)$. If the value of company's asset falls below recovery part of the debt, the company defaults. Therefore, default time for the company is defined as the stopping time

$$\eta_i = \inf\{0 \le t | A_i(t) \le D_i(t)\}.$$

Similarly, we define the corresponding stopping times

$$\begin{aligned} \underline{\eta} &= \inf\{0 \le t | A_1(t) \le D_1(t) \text{ and } A_2(t) \le D_2(t)\}, \\ \overline{\eta} &= \inf\{0 \le t | A_1(t) \le D_1(t) \text{ or } A_2(t) \le D_2(t)\}, \end{aligned}$$

Table 5.1: Parameters of the Constant Correlation Model

Volatility of the First Asset	Volatility of the Second Asset	Cross-Assets Correlation
0.3197	0.2781	0.3325

as the first time that both companies default and the first time that at least one of the companies default, respectively. By this setting, for each company i, we define equity's value prior to default by $S_i(t) = A_i(t) - D_i(t)$. After the default time η_i, equity's value is zero. This makes zero an absorbing state for the asset's price process. We use the evolutionary algorithm to calibrate the parameters of the constant correlation Credit-Grades model to market prices from General Motors and Ford's equity options prices on November 8^{th}, 2005 as in Arian *et al.* (2008d,b). The results of the calibration algorithm is shown in table 5.1.

We use these parameters to perform a monte carlo simulation on the joint processes of the General Motors and Ford Motors based on the constant correlation model (5.22). Two sample trajectories are shown in figure 5.1 for General Motors and Ford Motors asset prices and the recovery parts of their corresponding liabilities. In the monte carlo study, whenever the sample trajectory of the asset price hits the recovery part of the liability, default occurs. Then the probability of default can be calculated as the number of the sample trajectories which yield default divided by the total number of the trajectories.

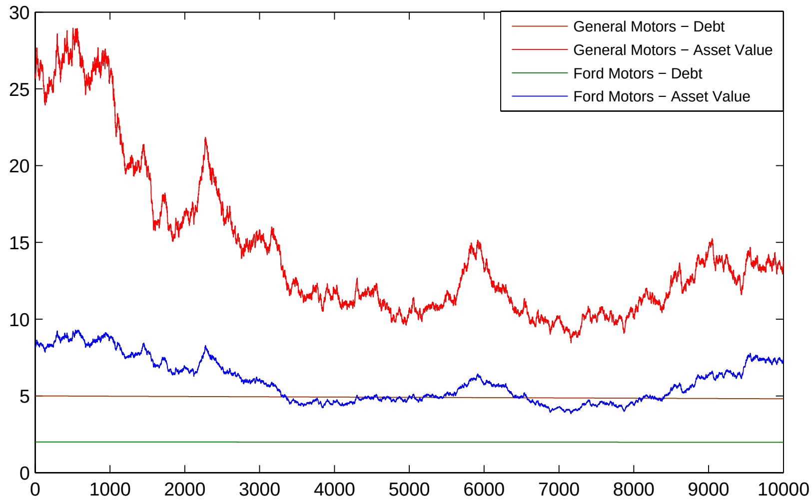

Figure 5.1: Simulated price of General Motors' and Ford's assets and the recovery part of their liabilities with constant correlation model

The same technique can be used to find the correlation default. The results are shown in table 5.4.

5.4.2 Correlated Default for the Extended CreditGrades model with Stochastic Covariance Wishart Process

In this section, we implement the Wishart process into the CreditGrades framework as in Arian *et al.* (2008d). Assume that the assets' prices for two companies follow the dynamics

$$\begin{pmatrix} d\ln(V_1) \\ d\ln(V_2) \end{pmatrix} = \begin{pmatrix} r_1(t) - d_1(t) \\ r_2(t) - d_2(t) \end{pmatrix} dt + \sqrt{\Sigma_t} \begin{pmatrix} dz_1 \\ dz_2 \end{pmatrix}, \tag{5.23}$$

where the covariance matrix is a two dimensional Wishart process which satisfies the dynamics

$$d\begin{pmatrix}\Sigma_{11}(t) & \Sigma_{12}(t)\\ \Sigma_{12}(t) & \Sigma_{22}(t)\end{pmatrix} = \left(\beta\begin{pmatrix}Q_{11}(t) & Q_{21}(t)\\ Q_{12}(t) & Q_{22}(t)\end{pmatrix}\begin{pmatrix}Q_{11}(t) & Q_{12}(t)\\ Q_{21}(t) & Q_{22}(t)\end{pmatrix}\right.$$
$$+\begin{pmatrix}M_{11}(t) & M_{12}(t)\\ M_{21}(t) & M_{22}(t)\end{pmatrix}\begin{pmatrix}\Sigma_{11}(t) & \Sigma_{12}(t)\\ \Sigma_{12}(t) & \Sigma_{22}(t)\end{pmatrix}$$
$$\left.+\begin{pmatrix}\Sigma_{11}(t) & \Sigma_{12}(t)\\ \Sigma_{12}(t) & \Sigma_{22}(t)\end{pmatrix}\begin{pmatrix}M_{11}(t) & M_{21}(t)\\ M_{12}(t) & M_{22}(t)\end{pmatrix}\right)dt$$
$$+\begin{pmatrix}\sigma_{11}(t) & \sigma_{12}(t)\\ \sigma_{12}(t) & \sigma_{22}(t)\end{pmatrix}\begin{pmatrix}dW_{11}(t) & dW_{12}(t)\\ dW_{21}(t) & dW_{22}(t)\end{pmatrix}\begin{pmatrix}Q_{11}(t) & Q_{12}(t)\\ Q_{21}(t) & Q_{22}(t)\end{pmatrix}$$
$$+\begin{pmatrix}Q_{11}(t) & Q_{21}(t)\\ Q_{12}(t) & Q_{22}(t)\end{pmatrix}\begin{pmatrix}dW_{11}(t) & dW_{12}(t)\\ dW_{21}(t) & dW_{22}(t)\end{pmatrix}\begin{pmatrix}\sigma_{11}(t) & \sigma_{12}(t)\\ \sigma_{12}(t) & \sigma_{22}(t)\end{pmatrix}.$$

Table 5.2: Parameters of the Wishart Process

Wishart Covariance	Gindikin's Coefficient	Autoregressive Matrix		Wishart's Volatility Matrix		Initial Covariance	
Parameters	$\beta = 4.447$	$m_{11} = -185.015$	$m_{12} = 0$	$q_{11} = 0.025$	$q_{12} = 0$	$\Sigma_{11} = 0.025$	$\Sigma_{12} = 0$
		$m_{21} = 0$	$m_{22} = -52.53$	$q_{21} = 0$	$q_{22} = 0.017$	$\Sigma_{21} = 0$	$\Sigma_{22} = 0.017$

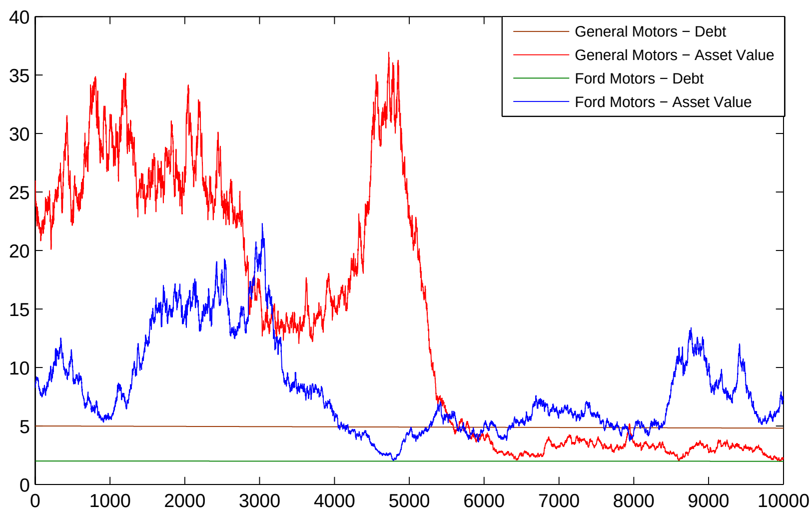

Figure 5.2: Simulated price of General Motors' and Ford's assets and the recovery part of their liabilities with stochastic covariance Wishart model

As before, the liability of the company is denoted by $B_i(t)$, the recovery rate is R_i and the dividend yield is d_i. We have calibrated the CreditGrades model with Wishart process in Arian *et al.* (2008b) by equity option data from General Motors and Ford. The results of the calibration is given in table 5.2. Using these sets of parameters we perform a Monte

Carlo simulation to reproduce the asset prices and evaluate probabilities of default when each trajectory hits the recovery part of the debt. Two simulated paths are shown in figure 5.2 for General Motors and Ford. The results of the Monte Carlo simulation is given in table 5.4.

5.4.3 Correlated Default for the Extended CreditGrades model with Principal Component Model

As a second extension of the CreditGrades model, in Arian *et al.* (2008d), we assume the covariance matrix of the assets has constant eigenvectors but stochastic eigenvalues following a CIR process. More precisely and in a mathematical framework if the values of assets are denoted by A_i, we have

$$\begin{aligned} dA_1(t) &= A_1(t)\left((r(t)-d_1(t))dt + \alpha_{11}\sqrt{\lambda_1(t)}dW_1(t) + \alpha_{12}\sqrt{\lambda_2(t)}dW_2(t)\right), \\ dA_2(t) &= A_2(t)\left((r(t)-d_2(t))dt + \alpha_{21}\sqrt{\lambda_1(t)}dW_1(t) + \alpha_{22}\sqrt{\lambda_2(t)}dW_2(t)\right), \end{aligned} \tag{5.24}$$

where eigenvalues of the covariance process follow

$$\begin{aligned} d\lambda_1(t) &= \kappa_1(\theta_1-\lambda_1)dt + \sigma_1\sqrt{\lambda_1(t)}dZ_t^1, \\ d\lambda_2(t) &= \kappa_2(\theta_2-\lambda_2)dt + \sigma_2\sqrt{\lambda_2(t)}dZ_t^2. \end{aligned}$$

Table 5.3: Parameters of the first and the second eigenvalues

Covariance Matrix	Mean-Reversion Speed	Asymptotic Eigenvalue	Eigenvalue Volatility	Initial Eigenvalue	Eigenvector Matrix	
1^{st} EV	$\kappa_1 = 3.1049$	$\theta_1 = 0.0141$	$\sigma_1 = 6.6152$	$\lambda_1 = 0.0291$	$\alpha_{11} = 0.6142$	$\alpha_{12} = 0.7891$
2^{nd} EV	$\kappa_2 = 8.5065$	$\theta_2 = 0.1607$	$\sigma_2 = 5.6188$	$\lambda_2 = 0.7660$	$\alpha_{21} = -0.789$	$\alpha_{22} = 0.6144$

This model has been calibrated via evolutionary algorithms in Arian *et al.* (2008d) using data from equity option prices of General Motors and Ford on November 8^{th}, 2005. The results of the calibration is given in table 5.3. Using parameter outputs from evolutionary algorithms, we perform a monte carlo method to simulate the assets prices. One sample path for General Motors and Ford is shown in figure 5.3. The results for the probabilities of defaults are shown in table 5.4. To verify the convergence properties of proposition 27 which approximates the joint probabilities of default, we perform another independent

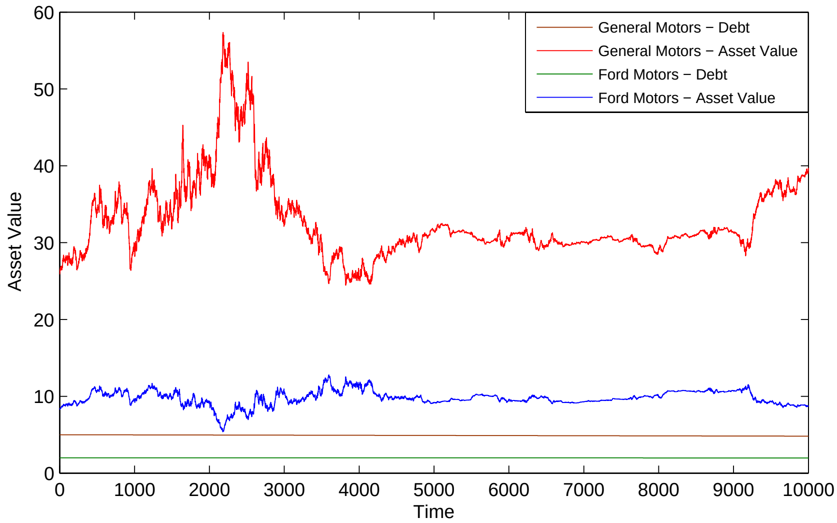

Figure 5.3: Simulated price of General Motors' and Ford's assets and the recovery part of their liabilities with stochastic covariance principal component model

partial monte carlo simulation which concentrates on only discrete points of time, meaning that the default could only occur on those pre-determined dates. Using 12 point to simulate the proposition we get table 5.4. The table shows that both marginal and joint probabilities for partial monte carlo simulations are very close to results obtained by the full monte carlo simulation.

Table 5.4: One year probabilities of default for general motors and ford with full Monte Carlo and partial Monte Carlo simulations

Cross-Asset Correlation	Probability of Default for General Motors	Probability of Default For Ford	Joint Default Probability
Constant Correlation - Full Monte Carlo	0.0077	0.0056	0.0012
Wishart Correlation - Full Monte Carlo	0.0091	0.0052	0.0035
Principal Component Correlation - Full Monte Carlo	0.0108	0.0063	0.0040
Constant Correlation - Partial Monte Carlo	0.0065	0.0054	0.0011
Wishart Correlation - Partial Monte Carlo	0.0080	0.0047	0.0030
Principal Component Correlation - Partial Monte Carlo	0.0101	0.0059	0.0035

5.5 Conclusion

We compared two *stochastic correlation* credit risk models with a *constant correlation* credit risk model. The main objective was to calculate marginal and joint probabilities of default. In Arian *et al.* (2008d,b), we estimate two stochastic correlation extensions of the CreditGrades model. We used these calibration results to estimate the marginal and joint probabilities of default in the CreditGrades framework with constant covariance, principal component covariance and Wishart covariance for General Motors and Ford. We performed the Monte Carlo simulation in full and partial modes and compared the simulated results for the marginal and joint probabilities of defaults. The full Monte Carlo simulation considers full path of the assets' value. The company defaults if its asset hits the recovery part of the debt at any time up to maturity. Obviously, the probability associated to the second event is lower than the probability associated to the first event. We show sample simulation trajectories for the constant covariance, principal component covariance and the Wishart covariance models.

Appendix A

Market Data

Table A.1: Ford Equity Option Prices - November 2005

Spot	Strike	Term	Mid Price	Bid	Offer
8.26	2.5	0.0278	5.75	5.7	5.8
8.26	5	0.0278	3.25	3.2	3.3
8.26	7.5	0.0278	0.8	0.75	0.85
8.26	10	0.0278	0.025	0	0.05
8.26	12.5	0.0278	0.025	0	0.05
8.26	15	0.0278	0.025	0	0.05
8.26	17.5	0.0278	0.025	0	0.05
8.26	20	0.0278	0.025	0	0.05
8.26	22.5	0.0278	0.025	0	0.05
8.26	2.5	0.1056	5.75	5.7	5.8
8.26	5	0.1056	3.3	3.2	3.4
8.26	7.5	0.1056	1	0.95	1.05
8.26	10	0.1056	0.075	0.05	0.1
8.26	12.5	0.1056	0.025	0	0.05
8.26	15	0.1056	0.025	0	0.05
8.26	17.5	0.1056	0.025	0	0.05
8.26	20	0.1056	0.025	0	0.05
8.26	22.5	0.1056	0.025	0	0.05
8.26	25	0.1056	0.025	0	0.05
8.26	30	0.1056	0.025	0	0.05
8.26	2.5	0.2	5.8	5.7	5.9
8.26	5	0.2	3.35	3.3	3.4

Table A.2: Ford Equity Option Prices - November 2005 (continued)

8.26	7.5	0.2	1.15	1.1	1.2
8.26	10	0.2	0.15	0.1	0.2
8.26	12.5	0.2	0.025	0	0.05
8.26	15	0.2	0.025	0	0.05
8.26	17.5	0.2	0.025	0	0.05
8.26	20	0.2	0.025	0	0.05
8.26	22.5	0.2	0.025	0	0.05
8.26	25	0.2	0.025	0	0.05
8.26	2.5	0.3583	5.8	5.7	5.9
8.26	5	0.3583	3.35	3.3	3.4
8.26	7.5	0.3583	1.325	1.3	1.35
8.26	10	0.3583	0.325	0.3	0.35
8.26	12.5	0.3583	0.075	0.05	0.1
8.26	15	0.3583	0.025	0	0.05
8.26	17.5	0.3583	0.025	0	0.05
8.26	20	0.3583	0.025	0	0.05
8.26	22.5	0.3583	0.025	0	0.05
8.26	25	0.3583	0.025	0	0.05
8.26	30	0.3583	0.025	0	0.05
8.26	2.5	0.6056	5.8	5.7	5.9
8.26	5	0.6056	3.4	3.3	3.5
8.26	7.5	0.6056	1.5	1.45	1.55
8.26	10	0.6056	0.475	0.45	0.5
8.26	12.5	0.6056	0.15	0.1	0.2
8.26	15	0.6056	0.05	0	0.1
8.26	17.5	0.6056	0.025	0	0.05
8.26	20	0.6056	0.025	0	0.05
8.26	22.5	0.6056	0.025	0	0.05
8.26	2.5	1.1972	5.8	5.7	5.9
8.26	5	1.1972	3.6	3.5	3.7
8.26	7.5	1.1972	1.875	1.85	1.9
8.26	10	1.1972	0.825	0.8	0.85
8.26	12.5	1.1972	0.275	0.25	0.3
8.26	15	1.1972	0.15	0.1	0.2
8.26	17.5	1.1972	0.075	0.05	0.1
8.26	20	1.1972	0.05	0	0.1

Table A.3: Ford Equity Option Prices - November 2005 (continued)

8.26	2.5	2.1944	5.75	5.6	5.9
8.26	5	2.1944	3.85	3.7	4
8.26	7.5	2.1944	2.45	2.4	2.5
8.26	10	2.1944	1.4	1.35	1.45
8.26	12.5	2.1944	0.8	0.75	0.85
8.26	15	2.1944	0.425	0.4	0.45

Table A.4: General Motors Equity Option Prices - November 2005

Spot	Strike	Term	Mid Price	Bid	Offer
25.86	22.5	0.0278	3.6	3.5	3.7
25.86	25	0.0278	1.55	1.5	1.6
25.86	27.5	0.0278	0.35	0.3	0.4
25.86	30	0.0278	0.075	0.05	0.1
25.86	32.5	0.0278	0.025	0	0.05
25.86	35	0.0278	0.025	0	0.05
25.86	37.5	0.0278	0.025	0	0.05
25.86	40	0.0278	0.025	0	0.05
25.86	42.5	0.0278	0.025	0	0.05
25.86	5	0.1056	20.9	20.8	21
25.86	7.5	0.1056	18.4	18.3	18.5
25.86	10	0.1056	15.9	15.8	16
25.86	12.5	0.1056	13.45	13.3	13.6
25.86	15	0.1056	11.05	10.9	11.2
25.86	17.5	0.1056	8.6	8.5	8.7
25.86	20	0.1056	6.35	6.2	6.5
25.86	22.5	0.1056	4.25	4.1	4.4
25.86	25	0.1056	2.475	2.4	2.55
25.86	27.5	0.1056	1.175	1.1	1.25
25.86	30	0.1056	0.475	0.4	0.55
25.86	32.5	0.1056	0.15	0.1	0.2
25.86	35	0.1056	0.075	0.05	0.1
25.86	37.5	0.1056	0.05	0	0.1
25.86	40	0.1056	0.05	0	0.1
25.86	42.5	0.1056	0.05	0	0.1
25.86	45	0.1056	0.025	0	0.05
25.86	2.5	0.2	23.4	23.3	23.5
25.86	5	0.2	20.9	20.8	21
25.86	7.5	0.2	18.45	18.3	18.6
25.86	10	0.2	16.05	15.9	16.2
25.86	12.5	0.2	13.75	13.6	13.9
25.86	15	0.2	11.35	11.2	11.5
25.86	17.5	0.2	9.1	9	9.2
25.86	20	0.2	6.95	6.8	7.1

Table A.5: General Motors Equity Option Prices - November 2005 (continued)

25.86	22.5	0.2	5	4.9	5.1
25.86	25	0.2	3.3	3.2	3.4
25.86	27.5	0.2	2.025	1.95	2.1
25.86	30	0.2	1.075	1	1.15
25.86	32.5	0.2	0.525	0.45	0.6
25.86	35	0.2	0.275	0.25	0.3
25.86	37.5	0.2	0.125	0.1	0.15
25.86	40	0.2	0.1	0.05	0.15
25.86	45	0.2	0.05	0	0.1
25.86	50	0.2	0.025	0	0.05
25.86	55	0.2	0.05	0	0.1
25.86	60	0.2	0.025	0	0.05
25.86	65	0.2	0.025	0	0.05
25.86	70	0.2	0.025	0	0.05
25.86	5	0.3583	20.9	20.8	21
25.86	7.5	0.3583	18.45	18.3	18.6
25.86	10	0.3583	16.1	16	16.2
25.86	12.5	0.3583	13.75	13.6	13.9
25.86	15	0.3583	11.5	11.4	11.6
25.86	17.5	0.3583	9.35	9.2	9.5
25.86	20	0.3583	7.35	7.2	7.5
25.86	22.5	0.3583	5.5	5.4	5.6
25.86	25	0.3583	4	3.9	4.1
25.86	27.5	0.3583	2.7	2.6	2.8
25.86	30	0.3583	1.775	1.75	1.8
25.86	32.5	0.3583	1.025	0.95	1.1
25.86	35	0.3583	0.6	0.55	0.65
25.86	37.5	0.3583	0.375	0.3	0.45
25.86	40	0.3583	0.225	0.15	0.3
25.86	42.5	0.3583	0.125	0.05	0.2
25.86	45	0.3583	0.075	0	0.15
25.86	47.5	0.3583	0.075	0	0.15
25.86	5	0.6056	20.9	20.8	21
25.86	7.5	0.6056	18.5	18.4	18.6
25.86	10	0.6056	16.15	16	16.3
25.86	12.5	0.6056	13.85	13.7	14

Table A.6: General Motors Equity Option Prices - November 2005 (continued)

25.86	15	0.6056	11.7	11.6	11.8
25.86	17.5	0.6056	9.65	9.5	9.8
25.86	20	0.6056	7.85	7.7	8
25.86	22.5	0.6056	6.15	6.1	6.2
25.86	25	0.6056	4.75	4.7	4.8
25.86	27.5	0.6056	3.5	3.4	3.6
25.86	30	0.6056	2.525	2.45	2.6
25.86	32.5	0.6056	1.725	1.7	1.75
25.86	35	0.6056	1.15	1.1	1.2
25.86	37.5	0.6056	0.8	0.75	0.85
25.86	40	0.6056	0.475	0.4	0.55
25.86	42.5	0.6056	0.325	0.25	0.4
25.86	2.5	1.1972	23.35	23.2	23.5
25.86	5	1.1972	20.95	20.8	21.1
25.86	7.5	1.1972	18.45	18.3	18.6
25.86	10	1.1972	16.25	16.1	16.4
25.86	12.5	1.1972	14.2	14	14.4
25.86	15	1.1972	12.2	12.1	12.3
25.86	17.5	1.1972	10.45	10.3	10.6
25.86	20	1.1972	8.9	8.7	9.1
25.86	22.5	1.1972	7.45	7.3	7.6
25.86	25	1.1972	6.15	6	6.3
25.86	27.5	1.1972	4.9	4.7	5.1
25.86	30	1.1972	4	3.9	4.1
25.86	32.5	1.1972	3.125	2.95	3.3
25.86	35	1.1972	2.45	2.35	2.55
25.86	40	1.1972	1.375	1.35	1.4
25.86	45	1.1972	0.8	0.75	0.85
25.86	50	1.1972	0.425	0.35	0.5
25.86	55	1.1972	0.25	0.15	0.35
25.86	60	1.1972	0.125	0.05	0.2
25.86	65	1.1972	0.075	0	0.15
25.86	70	1.1972	0.05	0	0.1
25.86	2.5	2.1944	23.4	23.2	23.6
25.86	5	2.1944	20.9	20.7	21.1
25.86	7.5	2.1944	18.4	18.2	18.6
25.86	10	2.1944	16.2	16	16.4

Table A.7: General Motors Equity Option Prices - November 2005 (continued)

25.86	12.5	2.1944	14.4	14.1	14.7
25.86	15	2.1944	12.8	12.5	13.1
25.86	17.5	2.1944	11.25	11	11.5
25.86	20	2.1944	9.8	9.5	10.1
25.86	22.5	2.1944	8.65	8.5	8.8
25.86	25	2.1944	7.4	7.1	7.7
25.86	30	2.1944	5.5	5.3	5.7
25.86	35	2.1944	3.8	3.7	3.9
25.86	40	2.1944	2.55	2.35	2.75
25.86	45	2.1944	1.65	1.5	1.8
25.86	50	2.1944	1.05	0.9	1.2
25.86	55	2.1944	0.65	0.55	0.75
25.86	60	2.1944	0.375	0.25	0.5

List of Tables

List of Figures

Bibliography

Abou-Kandil, H., Freiling, G., Ionescu, V., and Jank, G. (2003). *Matrix Riccati Equations in Control and Systems Theory*. Springer.

Arian, H., Escobar, M., and Seco, L. (2008a). A comparative analysis of two stochastic correlation credit risk models. *RiskLab Working Paper*.

Arian, H., Escobar, M., and Seco, L. (2008b). The creditgrades model with stochastic covariance wishart process. *RiskLab Working Paper*.

Arian, H., Escobar, M., and Seco, L. (2008c). Derivative pricing with stochastic switching correlations. *RiskLab Working Paper*.

Arian, H., Escobar, M., and Seco, L. (2008d). Pricing and calibration of a stochastic correlation credit risk model. *RiskLab Working Paper*.

Black, F. and Cox, J. (1976). Valuing corporate securities: Some effects of bond indenture provisions. *Journal of Finance*, **31**(2), 351–67.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*.

Boyce, W. and DiPrima, R. (2001). *Elementary Differential equations and boundary value problems*. Wiley and Sons.

Bru, M.-F. (1991). Wishart processes. *Journal of Theoretical Probability*, **4**(4), 725–751.

Carmona, R. and Durrleman, V. (2003). Pricing and hedging of spread options. *SIAM Review*.

Chacko, G. and Viceira, L. (2003). Spectral gmm estimation of continuous-time processes. *Journal of Econometrics*, **116**, 259–292.

Deelstra, G. and Delbaen, F. (2000). Existence of solutions of stochastic differential equations related to the bessel process. *Forthcoming*.

Emmerich, C. (2006). Modelling correlation and a stochastic process. *Working Paper.*

Escobar, M., Gotz, B., Seco, L., and Zagst, R. (2006). Pricing of spread options on stochastically correlated underlyings. *Working paper.*

Escobar, M., Gotz, B., Seco, L., and Zagst, R. (2007). Pricing of a c.d.o. on stochastically correlated underlyings.

Fonseca, J. D., Grasselli, M., and Tebaldi, C. (2006). A multifactor volatility heston model. *Quantitative Finance. (formerly RR-31)*, **8**, 591–604.

Fonseca, J. D., Grasselli, M., and Ielpo, F. (2007a). Estimating the wishart affine stochastic correlation model using the empirical characteristic function. *Working paper ESILV, RR-35.*

Fonseca, J. D., Grasselli, M., and Tebaldi, C. (2007b). Option pricing when correlations are stochastic: an analytical framework. *Review of Derivatives Research (Working paper ESILV, RR-33)*, **10**(2).

Geske, R. (1979). The valuation of compound options. *Journal of Financial Economics*, **7**(1), 63–81.

Gourieroux, C. (2005). Wishart auturegressive model for stochastic risk. *Working Paper.*

Gourieroux, C. and Sufana, R. (2004). The wishart autoregressive process of multivariate stochastic volatility. *Working Paper.*

Gourieroux, C., Jasiak, J., and Sufana, R. (2004). Derivative pricing with multivariate stochastic volatility: Application to credit risk. *Working Paper.*

Grasselli, M. and Tebaldi, C. (2004). Solvable affine term structure models. *Mathematical Finance*, **18**(1), 135–153.

Hamida, S. B. and Cont, R. (2004). Recovering volatility from option prices by evolutionary optimization. *Rapport Interne CMAP No 534, Ecole Polytechnique*, **8**(4).

He, H., Keirstead, W., and Rebholz, J. (1998). Double lookbacks. *Mathematical Finance.*

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility, with applications to bond and currency options. *Review of Financial Studies*, **6**, 327–343.

Hull, J. (2002). *Options, Futures, and Other Derivatives.* Prenctice Hall.

Hull, J. and White, A. (1987). The pricing of options on assets with stochastic volatilities. *Journal of Finance*, **42**(2), 281–300.

Hull, J. and White, A. (2004). Valuation of a c.d.o. and nth to default c.d.s. without monte carlo simulation. *Journal of Derivatives*, **12**(2), 8–23.

Hull, J., Nelken, I., and White, A. (2004). Merton's model, credit risk and volatility skews. *Journal of Credit Risk*, **1**(1), 1–27.

Hull, J., Predescu, M., and White, A. (2005). The valuation of correlation-dependent credit derivatives using a structural model. *Working Paper*.

Ikeda, N. and Watanabe, S. (1981). *Stochastic Differential Equations and Diffusion Processes*. Amsterdam, North-Holland.

Jones, E., Mason, S., and Rosenfield, E. (1984). Contingent claim analysis of corporate capital structures:an empirical investigation. *Journal of Finance*, **39**(3), 611–625.

Karatzas, I. and Shreve, S. (1991). *Brownian Motion and Stochastic Calculus*. Springer, Berlin.

Karatzas, I. and Shreve, S. (1998). *Methods of Mathematical Finance*. Springer-Verlag, New York.

Kilin, F. (2007). Accelerating the calibration of stochastic volatility models.

Ladyzhenskaya, O., Solonnikov, O., and Uralceva, N. (1968). Linear and quasilinear equations of parabolic type. *English Translation: American Mathematical Society*.

Lipton, A. (2001). *Mathematical methods for foreign exchange: A Financial Engineers Approach*. World Scientific.

Merton, R. C. (1974). On the pricing of corporate debt: the risk structure of interest rates. *Journal of Finance*.

Moral, P. D. and Miclo, L. (2003). Annealed feynman-kac models. *Communications in Mathematical Physics*, **235**(2), 191–214.

Muirhead, R. (2005). *Aspects of Multivariate Statistical Theory*. Wiley Series in Probability and Statistics.

Nelson, D. B. (1990). Arch models as diffusion approximations. *Journal of Econometrics*, **45**, 7–38.

Øksendal, B. (2003). *Stochastic Differential Equations*. Springer-Verlag.

Patras, F. (2005). A reflection principle for correlated defaults. *Stochastic Processes and their Applications*, **116**(4), 690–698.

Rebholz, J. (1994). Planar diffusions with applications to mathematical finance. *PhD Thesis, University of California-Berekley.*

Rockinger, M. and Semenova, M. (2005). Estimation of jump-diffusion process via empirical characteristic function. *FAME Research Paper Series rp150, International Center for Financial Asset Management and Engineering.*

Sepp, A. (2006). Extended creditgrades model with stochastic volatility and jumps. *Wilmott Magazine*, pages 50–62.

Stamicar, R. and Finger, C. (2005). Incorporating equity derivatives into the creditgrades model. *RiskMetrics Group.*

Stein, E. M. and Stein, J. (1991). Stock price distributions with stochastic volatility: an analytic approach. *The Review of Financial Studies*, **4**(4), 727–752.

Tretiakova, I. (2005). The stochastic nature of default correlation. *MSc thesis, University of Toronto.*

Zhou, C. (2001). An analysis of default correlations and multiple defaults. *The Review of Financial Studies*, **14**(2), 555–576.