Non-linear filtering applied to a new model with jumps in credit risk.

by

Maja Alvilde Bratlien Larsen

Thesis

for the degree of

Master of science

(Master i Modellering og Dataanalyse)

Faculty of Mathematics and Natural Sciences
University of Oslo

November 2015

Det matematisk- naturvitenskapelige fakultet
Universitetet i Oslo

Abstract

Lending money has been one of the basic activities of banks for centuries. However, credit evaluation and pricing of loans are still not well understood, since the assessment of the impact of credit risk on prices in bond markets, which is one of the most challenging types of financial risk, is in general difficult and subject to the complex interplay of factors as e.g. recovery risk and market risk. Roughly speaking, credit risk describes the exposure of losses due to changes of the vanity of borrowers as e.g. the issuer of a corporate bond. The severe global crisis of 2008, which was significantly caused by the sudden occurrence of illiquidity of credit markets, has shown the urgent need for a better understanding this sort of risk. In this thesis we will present a new quantitative model which is developed to control these kind of risk. We will estimate the parameters of the new model by using non-linear filtering techniques. Based on these estimations future stock prices will be computed.

This thesis consists of 7 chapters, where chapter 1 is an introduction to the mathematical notation and definitions which provide us with a foundation throughout this thesis. In chapter 2 an overview of Lévy process is given. Chapter 3 and 4 defines financial derivatives and discuss the challenges of the modelling of credit risk and basic approaches to such risk. We introduce the most common credit models, focusing on the Merton model which will be the reference model in a later chapter. In chapter 5 the theory of non-linear filtering will be given. In Chapter 6 we introduce a new model and we fit this model to empirical data. We emphasize that we focus on the simulations in this context. Concluding comments will be given in the last section. Chapter 7 suggests possible extensions to this thesis.

R code is given in the Appendix.

Acknowledgement

First and foremost I would like to thank my supervisor, Frank Norbert Proske, who has provided me with an interesting topic. His enthusiasm and advice during the last year have been very valuable to me.

I want to thank my fellow students at room B802 and particularly Lars and Tor Martin for motivating as well as academic conversations during these last years. Your company has been priceless! I would like to thank my family, especially my father Tor Espen for his support and my sister Kjerstin for proofreading.

Last, but not least, I would like to thank friends and my boyfriend Eivind, for their love and encouragement.

Contents

1 Basic mathematical tools **3**
- 1.0.1 Measure theory . . . 3
- 1.0.2 Stochastic processes and martingales . . . 5
- 1.0.3 Brownian motion, Itô integration and Itôs formula . . . 6
- 1.0.4 Change of measure and numeraire . . . 9

2 Lévy Processes **13**
- 2.0.5 Theory about Lévy processes . . . 14
- 2.0.6 Increasing Lévy processes . . . 18
- 2.0.7 Stable distributions and processes . . . 19
- 2.0.8 Lévy processes as Markov processes and martingales . 21

3 Credit risk and credit derivatives **23**
- 3.1 What is credit risk? . . . 23
- 3.2 Credit derivatives . . . 24
- 3.3 Bonds and Zero-Coupon Bonds . . . 24
  - 3.3.1 Corporate bonds . . . 25
- 3.4 Credit Default Swap(CDS) . . . 27
  - 3.4.1 Pricing CDS . . . 27
- 3.5 Portfolio credit derivatives . . . 28
  - 3.5.1 Index (Portfolio) CDS . . . 28

4 Modelling Credit Risk **31**
- 4.1 The Merton model . . . 32
  - 4.1.1 Finding the value of the firm with Merton . . . 35
  - 4.1.2 Maximum likelihood estimation . . . 37
  - 4.1.3 Merton's jump diffusion model . . . 38
- 4.2 First-passage modelling . . . 39
- 4.3 Challenges of the modelling of credit risk . . . 39

5 Non-linear filtering theory **41**

5.1 What is non-linear filtering? 41
5.1.1 Filtering problem 42
5.1.2 Non-linear filtering techniques and theory 43

6 Non-linear filtering applied to our new model 47
6.1 Our new model 47
6.2 Estimation of the parameters of the signal process 48
6.2.1 Simulation approach 49
6.2.2 Numerical results 51
6.2.3 Choosing λ to be different 53
6.3 Simulations of the price process 55
6.3.1 Comparing the model with the classical Merton model 58
6.4 Conclusion 59

7 Extensions 61
7.1 The signal process 61
7.2 The observation process 63
7.3 Intensity based credit models 63

8 Appendix 65
8.0.1 Appendix: chapter 2 65
8.0.2 Appendix: chapter 5 66
8.0.3 Appendix: chapter 6 66
8.1 R-code: Chapter 1 69
8.2 R-code: Chapter 2 70
8.3 R-code: Chapter 3 72
8.4 R-code: Chapter 4 73
8.5 R-code: Chapter 6 76

Chapter 1

Basic mathematical tools

In this chapter we will introduce the mathematical framework and notation which we will use throughout the thesis. This gives us a summary of the basic concepts concerning stochastic analysis. We will relate the concepts to applications in finance, which will give us a *toolbox* for the theory discussed in this thesis.

This chapter is based on [Sch03], [Ben04], [CT04], [App09], [Øks95] and the lecture notes of [Kie08].

1.0.1 Measure theory

We start with some measure theory. Let Ω represent the sample space. This is a set $\Omega \neq \emptyset$, [1] representing the collection of all possible outcomes of a random experiment.

We are often interested in finding the probability of an event to occur in an experiment, and for this purpose we need the following definitions:

[1] The notation $\emptyset$ represents the empty set.

Definition 1.1. ***σ-algebra***
A family $\mathcal{F}$ of subsets of Ω is called a σ-algebra (on Ω), if

(i) $\emptyset \in \mathcal{F}$

(ii) for $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$ *where* $A^c := \Omega - A$

(iii) $A_1, A_2, \ldots \in \mathcal{F} \Rightarrow \cup_{i \geq 1} A_i \in \mathcal{F}$.

Definition 1.2. ***Probability measure***
A function

$$\mathbb{P} : \mathcal{F} \to [0, 1]$$

is called probability measure on $(\Omega, \mathcal{F})$*, if*

(i) $\mathbb{P}(\emptyset) = 0, \quad \mathbb{P}(\Omega) = 1$

(ii) $A_1, \cdots, A_n, \cdots, \in \mathcal{F}$ *with* $A_i \cap A_j = \emptyset, i \neq j$ *(disjoint).*
This implies that $\mathbb{P}(\cup_{i \geq 1} A_i) = \sum_{i \geq 1} \mathbb{P}(A_i)$.

Elements of $\mathcal{F}$ are called events and and the triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called the *probability space.*

We say that the probability space if *complete* is $N \in \mathcal{F}$ is a null set, i.e. $\mathbb{P}(N) = 0$, then subsets A of N are null sets too.

Example 1.3. Lebesgue(-Borel) measure
The Lebesgue(-Borel) measure on [0,1] is an example of a probability measure. Set Ω = [0,1], $\mathcal{F} = B(\mathbb{R}) \cap [0, 1] = \{A \cap [0, 1] : A \in B(\mathbb{R})\}$.
It can be shown that there exist a unique probability measure

$$\lambda : \mathcal{F} \to [0, 1]$$
$$\text{s.t.}$$
$$\lambda([a, b)) = b - a \text{ (length of the interval [a,b))}.$$

Characteristic function

In chapter 2 we will look at characteristic functions of Lévy processes, which motivates us to give the following definition:

Definition 1.4. ***Characteristic function***
If X is a random variable with cumulative distribution function F, then its characteristic function ϕ_X is defined as

$$\phi_X(t) = \mathbf{E}[e^{itX}] = \int_{-\infty}^{\infty} e^{itx} F(dx), \quad t \in \mathbb{R}, i = \sqrt{-1}.$$

This implies that the characteristic function always exits since $|e^{itX}| = 1, 0 \leq t \leq T$.

Characteristic functions of random variables characterize the distribution. This means that if two random variables have the same characteristic function, it implies that they also have the same distribution.

1.0.2 Stochastic processes and martingales

Definition 1.5. ***Stochastic process***
A Stochastic process $X_t, 0 \leq t \leq T$ is a family of random variables parametrized by time t. That is, for each given time $t \in [0, T]$, X_t is a random variable.

Example 1.6. Stock prices
Let S_t denote the price of a stock at time $t, 0 \leq t \leq T$. For each $t \in [0, T]$, S_t is modelled by a random variable and hence the process $S_t, 0 \leq t \leq T$ is a stochastic process.

Definition 1.7. ***Filtration, $\mathcal{F}_t$.***
Let $\mathcal{F}_t, 0 \leq t \leq T$ be a family of σ-algebras on $(\Omega, \mathcal{F}, \mathbb{P})$ such that

$$\mathcal{F}_{t_1} \subset \mathcal{F}_{t_2}$$

for all $0 \leq t_1 \leq t_2 \leq T$.
Then $\mathcal{F}_t, 0 \leq t \leq T$ is called a filtration on $(\Omega, \mathcal{F}, \mathbb{P})$.

The σ-algebra $\mathcal{F}_t$ can be interpreted as a collection of information up to time t. More information is available when time passes by and hence the chances of determining the events we are looking for are more certain.

A process $X_t, 0 \leq t \leq T$ is said to be a $\mathcal{F}_t$-adapted process if the value at time t is revealed by the information $\mathcal{F}_t$.

Definition 1.8. ***$\mathcal{F}_t$-adapted***
A stochastic process $X_t, 0 \leq t \leq T$ is said to be $\mathcal{F}_t$-adapted if, for each $t, 0 \leq t \leq T$, the value of X_t is revealed at time t: the random variable X_t is $\mathcal{F}_t$-measurable.

Definition 1.9. ***Stopping time, optional time*** *A random variable* $\tau : \Omega \to [0, \infty]$ *is*

1. *a stopping time if the set* $\{\tau \leq t\} \in \mathcal{F}_t, \forall t$.
2. *a optional time if* $\{\tau < t\} \in \mathcal{F}_t, \forall t$.

Definition 1.10. ***σ-algebra at a stopping time, Martingale***

1 *For all cádlág processes, we define the* ***stopping time σ-algebra*** $\mathcal{F}_\tau$ *as*

$$\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap \{\tau \leq t\} \in \mathcal{F}_t, \forall t \geq 0\}.$$

2 *X is a* ***(sub-/super-) martingale*** *(with respect to* $\mathbb{F}$ *and* $\mathbb{P}$*) if*

$$\mathbf{E}[X_t|\mathcal{F}_s] \begin{cases} \leq X_s(super - martingale) \\ = X_s(martingale) \\ \geq X_s(sub - martingale) \end{cases}$$

a.s. for all $0 \leq s \leq t$*, provided* $E[|X_t|] < \infty$ *for all t.*

In other words, when looking at a (super-)martingale we look at the present value to predict the future value.

A cádlág process $X_t, 0 \leq t \leq T$, that is a process with right-continuous paths and existing left limits is called a local martingale if there exists an increasing sequence of stopping times T_n with $T_n \to \infty$ for $n \to \infty$ a.e. such that $X_{t \wedge T_n} \mathbf{1}_{T_n > 0}$ is an uniformly integrable martingale. This means that any martingale is a local martingale, but not necessarily vice versa. A local martingale is a martingale up to some stopping time T_n.

Definition 1.11. ***Equivalent martingale measure***
A measure $\mathbb{Q} \sim \mathbb{P}$ *such that the normalized process* $S_i(t) = \frac{S_i(t)}{S_0(t)}, 0 \leq t \leq T, 1 \leq i \leq N$*, is a (local) martingale measure w.r.t.* $\mathbb{Q}$ *is called an equivalent (local) martingale measure.*

If there exists an equivalent (local) martingale measure, then the market has no arbitrage. In other words the possibility of earning money from a zero investment without taking any risk does no exist.

1.0.3 Brownian motion, Itô integration and Itôs formula

Definition 1.12. ***Brownian motion***
$B = \{B_t\}_{t \geq 0}$ *is called a Brownian motion on* $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ *if*

- $B_0 = 0$ *a.e.*
- *B has independent increments.*
  For $0 \leq t_1 < \ldots < t_n$ *gives* $B_{t_1}, B_{t_2} - B_{t_1}, ..., B_{t_n} - B_{t_{n-1}}$ *are independent.*
- *B has stationary and Gaussian distributed increments.*
  For $t_1 \leq t_2, B_{t_2} - B_{t_1}$ *has the same distribution as* $B_{t_2-t_1}$*, with* $B_{t_2-t_1} \sim \mathcal{N}(0, t_2 - t_1)$*.*

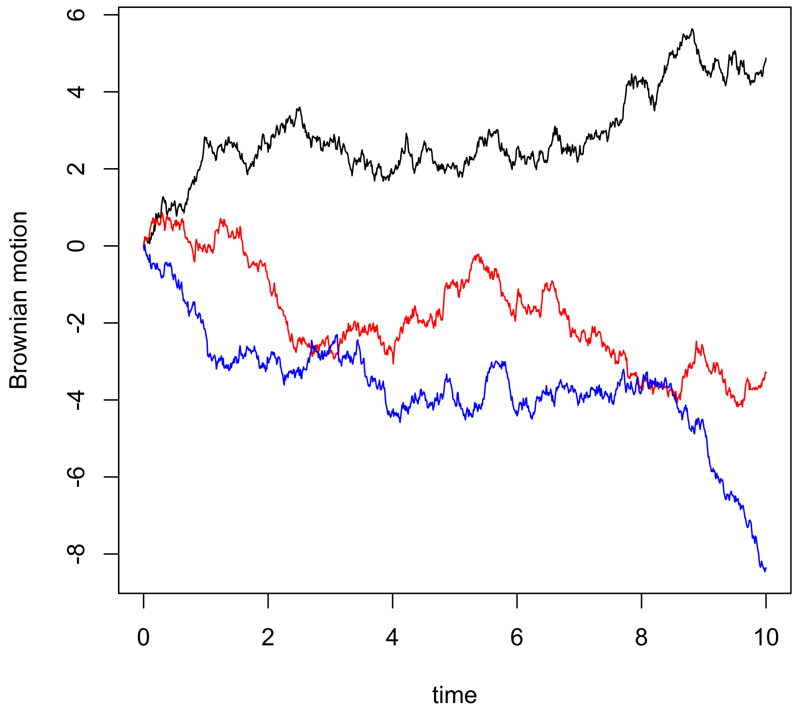


Figure 1.1: Three sample paths of Brownian motion.

The Brownian motion turns out to be a Markov process. This is because of its independence and stationary property.

From one day to another, the price of a stock or the credibility of a firm either stay the same, or move up or down by jumps. Brownian motion is a continuous stochastic process, which means that it does not capture this scenario. In reality, this economical behaviour of jumps causes unpredictable results when only taking into account continuous movements in the case of

Brownian motion. For this purpose we will later in this thesis look at Lévy processes.

Definition 1.13. ***Itô Processes***
Let B_t be Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$. A Itô Process is a stochastic process $X_t, 0 \leq t \leq T$ on $(\Omega, \mathcal{F}, \mathbb{P})$ of the form:

$$X_t = X_0 + \int_0^t u(s,\omega)ds + \int_0^t v(s,\omega)dB_s, 0 \leq t \leq T, \tag{1.1}$$

where u and v satisfies:

$$\mathbb{P}[\int_0^t u(s,\omega)^2 ds < \infty \quad \text{for all } t \geq 0] = 1 \tag{1.2}$$

$$\mathbb{P}[\int_0^t |u(s,\omega)| ds < \infty \quad \text{for all } t \geq 0] = 1. \tag{1.3}$$

We usually write equation (1.1) on the shorter differential form:

$$dX_t = udt + vdB_t, 0 \leq t \leq T. \tag{1.4}$$

Definition 1.14. ***Itô integrable***
Let $B_s, 0 \leq s \leq t$ be a Brownian motion with respect to a usual filtration $\mathcal{F}_s, 0 \leq s \leq t$. A stochastic process $X_s, 0 \leq s \leq t$ is called Itô integrable on the interval $[0,t]$ if:

- *$X_s, 0 \leq s \leq t$ is adapted.*
- $\int_0^t \mathrm{E}[X_s^2]ds < \infty$.

The Itô integral for e.g. bounded continuous adapted X is defined as the random variable

$\int_0^t X(s,\omega)dB(s,\omega) = \lim_{n\to\infty} \sum_{i=1}^{n-1} X(s_i,\omega)(B(s_{i+1},\omega) - B(s_i,\omega))$,

where the limit is in the sense of variance.

Theorem 1.15. ***Expectation and variance of the Itô integral***
The expectation and variance of the Itô integral are

$$\mathbf{E}[\int_0^t X_s dB_s] = 0, \quad \mathbf{Var}[\int_0^t X_s dB_s] = \int_0^t \mathbf{E}[X_s^2]ds, 0 \leq t \leq T. \tag{1.5}$$

The relation for the variance is known as the Itô isometry.

Definition 1.16. ***Itôs formula*** *(Short hand version):*

$$df(t,X_t) = \frac{\partial f(t,X_t)}{\partial t}dt + \frac{\partial f(t,X_t)}{\partial x}dX_t + \frac{1}{2}\frac{\partial^2 f(t,X_t)}{\partial x^2}(dX_t)^2, 0 \leq t \leq T. \tag{1.6}$$

1.0.4 Change of measure and numeraire

Radon-Nikodym theorem

The Radon-Nikodym theorem states that the change of probability measure $\mathbb{Q}$ to another measure $\mathbb{P} << \mathbb{Q}$ is uniquely characterized by the corresponding Radon-Nikodym density L with expectation $\mathbf{E}_{\mathbb{Q}}[L] = 1, L \geq 0$. In other words, for all measurable X the expected values under the new probability measure $\mathbb{P}$ is given by:

$$\mathbf{E}_{\mathbb{P}}[X] = \mathbf{E}_{\mathbb{Q}}[LX]. \tag{1.7}$$

This relation is, in literature, usually denoted by:

$$\frac{d\mathbb{P}}{d\mathbb{Q}} = L. \tag{1.8}$$

Hence an interpretation of the Radon-Nikodym density is a likelihood ratio between the two probability measures.

When we change the measure, we change the probability but the random variables remain unchanged. A Brownian motion under the probability measure $\mathbb{Q}$ is not necessarily a Brownian motion under the new probability measure $\mathbb{P}$. Girsanov's theorem determines which processes are Brownian motion under $\mathbb{P}$.

Theorem 1.17. ***The Girsanov theorem***
Let $Y_t \in \mathbb{R}^n$ be a Itô process of the form

$$dY_t = a(t,\omega)dt + dB_t, \quad 0 \leq t \leq T, Y_0 = 0,$$

where $T \leq \infty$ is a given constant and B_t is a n-dimensional Brownian-motion. Put

$$M_t = \exp(-\int_0^t a(s,\omega)dB_s - \frac{1}{2}\int_0^t a^2(s,\omega)ds), \quad 0 \leq t \leq T. \tag{1.9}$$

Assume that M_t is martingale w.r.t. $\mathcal{F}_t^{(n)}$ and $\mathbb{Q}$, that is $a(t,\omega), 0 \leq t \leq T$ satisfies Novikov's condition

$$\mathbf{E}[\exp(\frac{1}{2}\int_0^T a^2(s,\omega)ds)] < \infty, \tag{1.10}$$

where $\mathbf{E} = \mathbf{E}_{\mathbb{Q}}$ is the expectation w.r.t. $\mathbb{Q}$. Define the measure $\mathbb{P}$ on $\mathcal{F}_T^{(n)}$ by

$$d\mathbb{P}(\omega) = M_T(\omega)d\mathbb{Q}(\omega). \tag{1.11}$$

Then $\mathbb{P}$ is a probability measure on $\mathcal{F}_T^{(n)}$ and Y_t is a n-dimensional Brownian motion w.r.t. $\mathbb{P}$, for $0 \leq t \leq T$.

Proof. See [Øks95]. □

In other words, Girsanov's theorem tells us that if we change the drift coefficient of a given Itô process, then the law of the process will not change. By applying Girsanov's theorem one can move from the original measure $\mathbb{Q}$ to an equivalent measure $\mathbb{P}$, e.g. pricing assets as stocks in an arbitrage free market.

Example 1.18. Geometric Brownian motion
Geometric Brownian motion is a dynamical model which describes the price S_t of a underlying stock at times $t, 0 \leq t \leq T$.
The model is given by the following:

$$S_t = S_0 \exp(\mu t + \sigma B_t), \quad 0 \leq t \leq T, S_0 = x \tag{1.12}$$

where μ is the drift, σ represents the volatility and $B_t, 0 \leq t \leq T$ is Brownian motion.

The dynamics of the Geometric Brownian motion is given by:

$$dS_t = \alpha S_t dt + \sigma S_t dB_t. \tag{1.13}$$

Here $\alpha = (\mu + \frac{1}{2}\sigma^2)$. [2]

We define the dynamics:

$$dW_t = dB_t + \tfrac{\alpha - r}{\sigma}dt.$$

[2] for further calculation see Appendix.

By substituting the expression in the dynamics of equation (1.13) which gives:

$$
\begin{aligned}
dS_t &= \alpha S_t dt + \sigma S_t dB_t \\
&= r S_t dt + (\alpha - r) S_t dt + \sigma S_t dB_t \\
&= r S_t dt + \sigma S_t (dB_t + \frac{\alpha - r}{\sigma} dt) \\
&= r S_t dt + \sigma S_t dW_t.
\end{aligned}
\tag{1.14}
$$

By looking at the expectation:

$$
\begin{aligned}
\mathbf{E}[W_t] &= \mathbf{E}[\lambda t + B_t] \\
&= \mathbf{E}[\lambda t] + \mathbf{E}[B_t] \\
&= \frac{\alpha - r}{\sigma} t \\
&\neq 0
\end{aligned}
$$

from definition (1.12) we know that W_t is not a Brownian motion under the probability $\mathbb{Q}$.

If we now set $a(t, \omega) = \frac{\alpha - r}{\sigma}$ in theorem 1.17 gives:

$$W_t = B_t + \tfrac{\alpha - r}{\sigma} t.$$

Hence we have Brownian motion under the probability $\mathbb{P}$.

Chapter 2

Lévy Processes

In finance we are often interested in modelling the dynamics of the underlying asset. A popular model used for this purpose is the Black-Scholes model which describes diffusion.

However, observed asset returns certain empirical properties which are not captured by the Black-Scholes model. When we look at how the price behaves over time we see jumps. This has led to a development of a large number of jump diffusion models and a widely studied class is the exponential Lévy process. Lévy processes have become an extremely popular and important tool in mathematical finance. This is so because it describes the financial market in a more accurate and realistic way than models based on continuous processes. In the real world we observe price processes with sudden changes, which are captured by jumps and this is something price analysts have to take into consideration.

The aim of this chapter is to provide an overview of Lévy processes and their application to mathematical finance. Most of the material is borrowed from [CT04]. We begin with the definition of Lévy processes and some basic concepts. We will then introduce non-negative Lévy processes, namely subordinators and α-stable distributions.

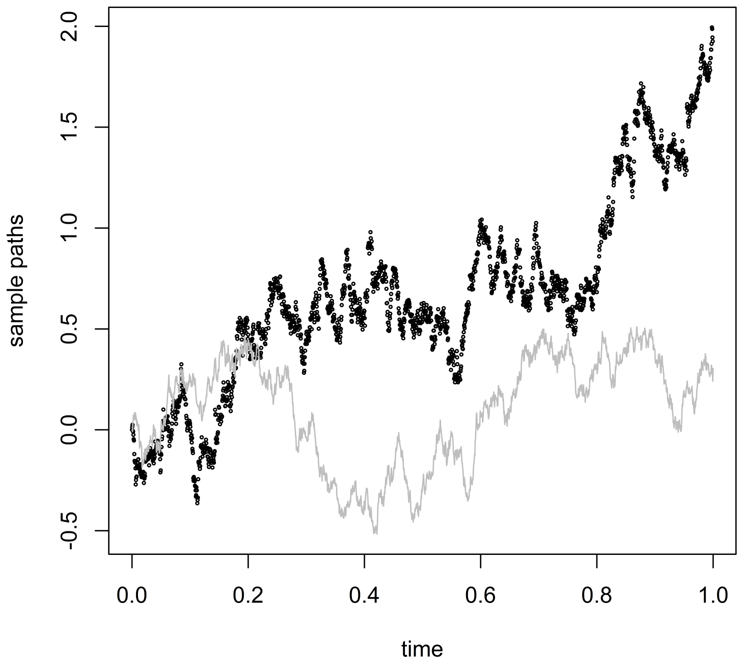


Figure 2.1: Sample path of Brownian motion (grey) where $\alpha = 2$ and a Lévy process (black) where $\alpha = 1.9$.

2.0.5 Theory about Lévy processes

Let us begin with the definition of Lévy processes:

Definition 2.1. ***Lévy Process***
A cádlág [1] *stochastic process* $X_t, 0 \leq t \leq T$ *on* $(\Omega, \mathcal{F}, \mathbb{P})$ *with values in* $\mathbb{R}^d$ *such that* $X_0 = 0$ *is called a Lévy process if it possesses the following properties:*

1. *Independent increments: for every increasing sequence of times* $t_0, \cdots, t_n$*, the random variables* $X_{t_0}, X_{t_1} - X_{t_0}, \cdots, X_{t_n} - X_{t_{n-1}}$ *are independent.*

2. *Stationary increments: the law of* $X_{t+h} - X_t$ *does not depend on* t*.*

[1] A cádlág has the property of being right continuous and has left limits. In some literature the term RCLL (right continuous left limits)is often used.

3. *Stochastic continuity:* $\forall \epsilon > 0, \lim_{h \to 0} \mathbb{P}(|X_{t+h} - X_t| \geq \epsilon) = 0$.

The third condition entails that for a given time $t, 0 \leq t \leq T$ the probability of observing a jump is zero. In other words, discontinuity can only occur at random times.

Definition 2.2. ***Infinite divisibility***
A probability distribution F on $\mathbb{R}^d$ is said to be infinitely divisible if for any integer $n \geq 2$, there exist n i.i.d. random variables $X_1, \cdots, X_n$ such that the sum $X_1 + \cdots + X_n$ has distribution F.

There is a strong interplay between infinite divisible distributions and Lévy processes. In fact, for every $t \geq 0$ a Lévy process X_t has an infinitely divisible distribution. And if F is an infinitely divisible distribution, then it exists a Lévy process such that the distribution of X_1 is given by F.

Example 2.3. Normally distributed variables
A simple example is where $X_1, \cdots, X_n$ are independent, identically normally distributed with mean $\frac{\mu}{n}$ and variance $\frac{\sigma^2}{n}$. Then $Y = \sum_{i=1}^{n} X_i$ is also normally distributed, but with mean μ and variance σ^2. In other words the distribution is the same but the parameters are modified.

Definition 2.4. ***Characteristic function***
The characteristic function of the $\mathbb{R}^d$-valued random variable X is the function $\Phi_X : \mathbb{R}^d \to \mathbb{R}$ defined by

$$\forall z \in \mathbb{R}^d, \Phi_X(z) = \mathbf{E}[\exp(iz.X)] = \int_{\mathbb{R}^d} e^{iz.x} d\mu_X(x). \tag{2.1}$$

The characteristic function of a Lévy process, $X_t, 0 \leq t \leq T$ is given by

$$\mathbf{E}[e^{iz.X_t}] = e^{t\psi(z)}, \quad z \in \mathbb{R}^d, \tag{2.2}$$

where the continuous function $\psi : \mathbb{R}^d \mapsto \mathbb{R}$ is called the characteristic exponent of X.

Since $X_t, 0 \leq t \leq T$ is a Lévy process we know that it has a infinitely divisible distribution. This gives $\Psi = \Psi_{X_1}$ and by linearity we have $\Psi_{X_1} = t\Psi_{X_1} = t\Psi$. This entails that if we know the distribution of X_1, we can say something about the whole process.

Definition 2.5. ***Compound Poisson Process***
A compounded Poisson process with intensity λ>0 and jump size distribution f is a stochastic process X_t defined as

$$X_t = \sum_{i=1}^{N_t} Y_i, \tag{2.3}$$

where jump sizes Y_i are i.i.d. with distribution f and (N_t) is a Poisson process with intensity λ, independent from $(Y_i)_{i\geq 1}$.

Proposition 2.6. ***Characteristic function of a compound Poisson process***
Let $X_t, 0 \leq t \leq T$ be a compounded Poisson process on $\mathbb{R}^d$. Its characteristic function has the following representation:

$$\mathbf{E}[\exp(iu.X_t)] = \exp(t\lambda \int_{\mathbb{R}^d} (e^{iu.x} - 1) f(dx)), \tag{2.4}$$

where λ denotes the jump intensity and f the jump size distribution.

Definition 2.7. ***Lévy measure***
Let $X_t, 0 \leq t \leq T$ be a Lévy process on $\mathbb{R}^d$. The measure ν on $\mathbb{R}^d$ defined by:

$$\nu(A) = E[\#\{t \in [0,1] : \Delta X_t \neq 0, \Delta X_t \in A\}], A \in \mathcal{B}(\mathbb{R}^d) \tag{2.5}$$

is called the Lévy measure of $X : \nu(A)$.

As we can see, the Lévy measure describes the expected number of jumps per unit of time.

Proposition 2.8. ***Lévy-Itô decomposition***
Let $X_t, 0 \leq t \leq T$ be a Lévy process on $\mathbb{R}^d$ and ν its Lévy measure.

1. *ν is a Radon measure on $\mathbb{R}^d \backslash 0$ and verifies:*

$$\int_{|x|\leq 1} |x|^2 \nu(dx) < \infty \qquad \int_{|x|\geq 1} \nu(dx) < \infty. \tag{2.6}$$

2. *The jump measure of X, denoted by J_X, is a Poisson random measure on $[0, \infty[\times\mathbb{R}^d$ with intensity measure $\nu(dx)dt$.*

3. *There exist a vector γ and a d-dimensional Brownian motion $B_t, 0 \leq t \leq T$ with covariance matrix A such that:*

$$X_t = \gamma t + B_t + X_t^l + \lim_{\epsilon \downarrow 0} \tilde{X}_t^\epsilon, \tag{2.7}$$

where:

$$X_t^l = \int_{|x| \geq 1, s \in [0,t]} x J_X(ds \times dx) \tag{2.8}$$

and:

$$\tilde{X}_t^\epsilon = \int_{\epsilon \leq |x| < 1, s \in [0,t]} x J_X(ds \times dx) - \nu(dx) ds \tag{2.9}$$

$$= \int_{\epsilon \leq |x| < 1, s \in [0,t]} x \tilde{J}_X(ds \times dx). \tag{2.10}$$

The so-called Lévy-Itô decomposition (2.7) entails that all Lévy processes can be decomposed into three parts $X = X^1 + X^2 + X^3$, where $X^1 = \gamma t + B_t$ is a Brownian motion with drift, $X^2 = X_t^l$ is a compounded Poisson process with jumps with size bigger or equal to one, and $X^3 = \tilde{X}_t^\epsilon$ is a compensated sum of jumps smaller than one.

The reason why we can not set $\epsilon = 0$ immediately in X^3 is because it may have infinitely many small jumps. And by letting ϵ go to zero we avoid this problem.

The characteristic triplet (A, ν, γ) of the Lévy process characterizes the distribution through its characteristic function.

Theorem 2.9. ***Lévy-Khinchin representation***
Let $X_t, 0 \leq t \leq T$ be a Lévy process on $\mathbb{R}^d$ with characteristic triplet $(\mathcal{A}, \nu, \gamma)$. Then

$$E[e^{iz.X_t}] = e^{t\psi(z)}, z \in \mathbb{R}^d \tag{2.11}$$

with $\psi(z) = -\frac{1}{2} z.Az + i\gamma.z + \int_{\mathbb{R}^d} (e^{iz.x} - 1 - iz.x\mathbf{1}_{|x| \leq 1})\nu(dx)$.

From the Lévy-Itô decomposition we know that the Lévy process can be decomposed into three parts $X = X^1 + X^2 + X^3$, where $X^i, i = 1, 2, 3$ are independent processes. This gives us the characteristic exponent of a Lévy process given by equation (2.11). The Lévy-Khinchin representation combined with the Itô decomposition actually tells us that the small jumps are independent of the big jumps.

2.0.6 Increasing Lévy processes

Subordinators are Lévy processes with increasing increments, in other words they do not have any negative jumps. This means that μ is defined in $(0, \infty)$ and has no mass in $(-\infty, 0)$. Mathematically, a Lévy process is a subordinator if and only if one of the equivalent conditions of following proposition is satisfied.

Proposition 2.10. ***Subordinator***
Let $X_t, 0 \leq t \leq T$ be a Lévy process on $\mathbb{R}$. The following conditions are equivalent:

i. $X_t \geq 0$ *a.s. for some* $t > 0$.

ii. $X_t \geq 0$ *a.s. for every* $t > 0$.

iii. *Sample paths of* $X_t, 0 \leq t \leq T$ *are almost surely non-decreasing:* $t \geq s \Rightarrow X_t \geq X_s$ *a.s.*

iv. *The characteristic triplet of* $X_t, 0 \leq t \leq T$ *satisfies* $A = 0$, $\nu((-\infty, 0]) = 0$, $\int_0^\infty (x \wedge 1)\nu(dx) < \infty$ *and* $b \geq 0$. *That is,* $X_t, 0 \leq t \leq T$ *has no diffusion component, only positive jumps of finite variation.*

A subordinator is often used as a building block, as time changes, to built other Lévy processes. This property is especially convenient when we construct Lévy-based models in finance.

Theorem 2.11. ***Subordination of a Lévy process***
Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $X_t, 0 \leq t \leq T$ be a Lévy process on $\mathbb{R}^d$ with characteristic exponent $\Psi(u)$ and triplet (A, ν, γ) and let $S_t, 0 \leq t \leq T$ be a subordinator with Laplace exponent $l(u)$ and triplet $(0, \rho, b)$. Then the process $Y_t, 0 \leq t \leq T$ defined for each $\omega \in \Omega$ by $Y(t, \omega) = X(S(t, \omega), \omega)$ is a Lévy process. Its characteristic function is

$$\mathbf{E}[e^{iuY_t}] = e^{tl(\Psi(u))}. \tag{2.12}$$

I.e., the characteristic exponent of Y is obtained by composition of the Laplace exponent of S with the characteristic exponent of X. The triplet (A^Y, ν^Y, γ^Y) of Y is given by:

$$A^Y = bA \tag{2.13}$$

$$\nu^Y(B) = b\nu(B) + \int_0^\infty p_s^X(B)\rho(ds), \quad \forall B \in \mathcal{B}(\mathbb{R}^d), \tag{2.14}$$

$$\gamma^Y = b\gamma + \int_0^\infty \rho(ds) \int_{|x|\leq 1} x p_s^X(dx), \tag{2.15}$$

where p_t^X is the probability distribution of $X_t, 0 \leq t \leq T$.

Proof. The proof is carried out in [CT04]. □

If $S_t, 0 \leq t \leq T$ is a subordinator, its trajectories are increasing:

$$S_t \geq 0, \quad \forall t \geq 0$$

and hence we can use it as a *time change* of other Lévy processes.

2.0.7 Stable distributions and processes

Self-similarity is a remarkable property of Brownian motion. This entails that: if B is Brownian motion on $\mathbb{R}$ then

$$(\frac{B_{at}}{\sqrt{a}}) \stackrel{d}{=} (B_t), \quad 0 \leq t \leq T. \tag{2.16}$$

More generally, a Lévy process has the property of being self-similar if:

$$\forall a > 0, \quad \exists b(a) < 0 : \quad \left(\tfrac{X_{at}}{b(a)}\right) \stackrel{d}{=} (X_t), \quad 0 \leq t \leq T.$$

The characteristic function of X_t is given by $\Phi_{X_t}(z) = e^{-t\Psi(z)}$, which leads to the following definition:

Definition 2.12. ***Stable distribution***
A Random variable $X \in \mathbb{R}^d$ is said to have stable distribution if for every $a > 0$ there exists $b(a)$ and $c(a) \in \mathbb{R}^d$ such that

$$\Phi_X(z)^a = \Phi_X(zb(a))e^{ic.z}, \quad \forall z \in \mathbb{R}^d. \tag{2.17}$$

It is said to have a strictly stable distribution if

$$\Phi_X(z)^a = \Phi_X(zb(a)), \quad \forall z \in \mathbb{R}^d. \tag{2.18}$$

For every stable distribution there exists a constant $\alpha \in (0, 2]$, called the index of stability, such that $b(a) = a^{1/\alpha}$ in equation (2.17). A stable distribution with index α is also referred to as α-**stable distribution.**

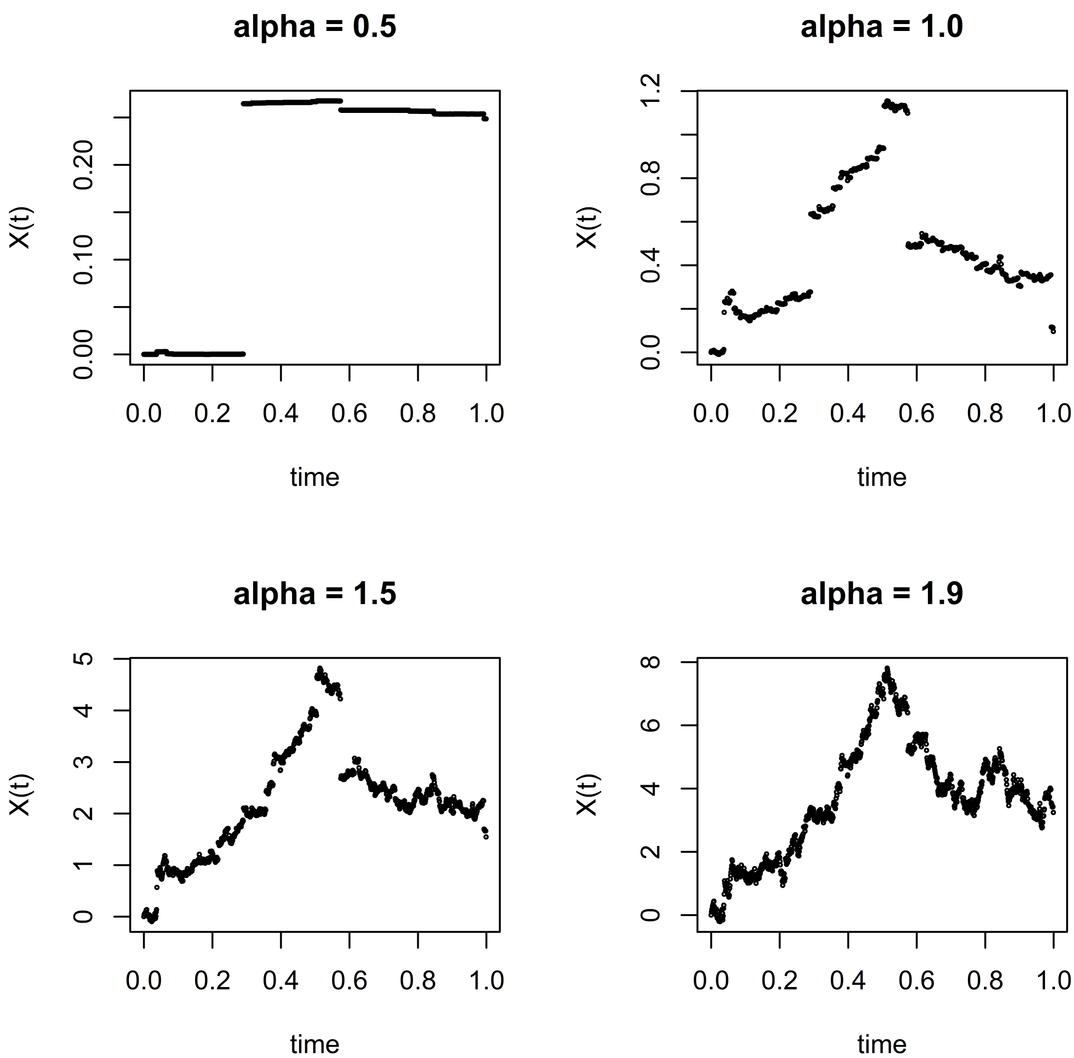


Figure 2.2: α-stable processes with α-values equal to 0.5,1.0,1.5 and 1.9

An example of an α-stable distribution is where we set $\alpha = 2$. In this case we have a Gaussian distribution. In fact, the Gaussian distribution is the only 2-stable distribution, proposition 2.13. As we can see by Figure 2.2 the sample paths begin to look like the trajectory of a Brownian motion as α increases.

Proposition 2.13. ***Stable distributions and Lévy processes***
A distribution on $\mathbb{R}^d$ is α-stable with $0 < \alpha < 2$ if and only if it is infinitely divisible with characteristic triplet $(0, \nu, \gamma)$ and there exists a finite measure λ on S, a unit sphere of $\mathbb{R}^d$, such that

$$\nu(B) = \int_S \lambda(d\xi) \int_0^\infty \mathbf{1}_B(r\xi) \frac{dr}{r^{1+\alpha}}. \tag{2.19}$$

A distribution on R^d is α-stable with $\alpha = 2$ if and only if it is Gaussian.

2.0.8 Lévy processes as Markov processes and martingales

Lévy processes have the Markov property because of its independent increments. We know that, for time $s, 0 \leq s \leq t$ the Lévy process satisfies

$$X_{s+t} - X_s \stackrel{d}{=} X_t.$$

Its transition kernel is given by:

$$P_{s,t}(x, B) = P(X_t \in B | X_s) \quad \forall B \in \mathcal{B}.$$

If we now consider a stopping time τ, $0 \leq \tau \leq t$ then the process $Y_t = X_{\tau+t} - X_\tau$ is again a Lévy process, independent from the filtration $\mathcal{F}_\tau = A \subset \mathcal{F} : \tau \cap A \in \mathcal{F}_t, t \geq 0$ and with the same distribution as $X_t, 0 \leq t \leq T$. This implies that the Lévy process has the strong Markov property.

Lévy processes have independent increments which entails that we can construct different martingales.

Proposition 2.14. ***Lévy processes as Martingales***
Let $X_t, 0 \leq t \leq T$ be a real-valued process with independent increments. Then

1. $\left(\frac{e^{iuX_t}}{\mathbf{E}[e^{iuX_t}]}\right), t \geq 0$ *is a martingale* $\forall u \in \mathbb{R}$
2. *If for some* $u \in \mathbb{R}, \mathbf{E}[e^{uX_t}] < \infty, \forall t \geq 0$ *then* $\left(\frac{e^{iuX_t}}{\mathbf{E}[e^{iuX_t}]}\right)$ *is a martingale.*
3. *If* $\mathbf{E}[X_t] < \infty, \forall t \geq 0$ *then* $M_t = X_t - \mathbf{E}[X_t]$ *is a martingale (and also a process with independent increments).*
4. *If* $\mathbf{Var}[X_t] < \infty, \forall t \geq 0$ *then* $(M_t)^2 - \mathbf{E}[(M_t)^2]$ *is a martingale, where* M *is the martingale defined above.*

If X_t is a Lévy process, for all of the processes of this proposition to be martingales it suffices that the corresponding moments be finite for one value of t.

In finance we are often interested in whether the Lévy process itself or its exponential is a martingale. We can verify this by checking the following proposition:

Proposition 2.15. ***Martingale condition for Lévy process***
Let $X_t, 0 \leq t \leq T$ be a Lévy process on $\mathbb{R}$ with characteristic triplet (A, ν, γ).

1. *X_t is a martingale if and only if*

$$\int_{|x|\geq 1} |x|\nu(dx) < \infty$$

and

$$\gamma + \int_{|x|\geq 1} x\nu(dx) = 0$$

2. *e^{X_t} is a martingale if and only if $\int_{|x|\geq 1} e^x \nu(dx) < \infty$ and*

$$\frac{A}{2} + \gamma + \int_{-\infty}^{\infty} (e^x - 1 - x\mathbf{1}_{|x|\leq 1})\nu(dx) = 0$$

Chapter 3

Credit risk and credit derivatives

In this chapter we will define credit risk and discuss pricing methods of a traditional credit derivative. This will give us a brief introduction to the financial notation, which we will use throughout this thesis.

The material of this chapter are mainly borrowed from [WHD99], [Sch03], [BR02] and [Kie08].

3.1 What is credit risk?

[Sch03], defines credit risk as

Definition 3.1. ***Credit risk***
The risk that an obligor does not honour his payment obligations.

In other words, credit risk [1] describes the risk that an obligor does not manage to pay off interest or a principle of a loan. E.g. a bank lending out money with the risk of not getting it back by the time of maturity of the contract. Intuitively with longer time to maturity comes greater risk.

There are commercial rating agencies ranking creditworthiness of companies. A firm's credit rating is a measure of the firm's probability to default. Example of such agencies are Moody's Investors Service and Standard & Poor's Corporation. In Moody's, the gradations of creditworthiness are indicated

[1] The terms "credit risk" and "default risk" has the same meaning, unless otherwise stated.

by rating symbols from Aaa to C (lowest to highest risk)[2].

However, many major financial institutions use their own internal rating systems.

3.2 Credit derivatives

Definition 3.2. ***Credit derivatives***

- *A credit derivative is a derivative security that is primarily used to transfer, hedge or manage credit risk.*
- *A credit risk derivative is a derivative security whose payoff is materially affected by credit risk.*

Secondary (or derived) products where values and pay-off are channelled through contract clauses set up in advanced, are called derivatives in finance. A credit derivative is a contract between two or more parties which allows the participants to manage their exposure to credit risk. Options, swaps and forward contracts are examples of such financial assets. These agreements make it possible to trade credit risk. Credit derivatives are often traded *over-the-counter*(OTC), this means that the trading is done directly between the parties, without any supervision.

A popular type of credit derivative is the Credit Default Swap (CDS) which will be explained later on in this chapter.

3.3 Bonds and Zero-Coupon Bonds

Bonds are investment in debt and help government and private companies to raise capital. The borrower makes fixed payments to the investor at certain times $0 = t_0 \leq t_1 \leq \ldots \leq t_n = T$, where T is the time of maturity of the contract. The last payment at time of maturity is usually larger than the others and is known as the *face value* or *face* of the bond. The time to maturity of a bond varies from a year to a century (or sometimes longer). A special case of a bond is the zero-coupon bond (ZCB), where there are no (coupon) payments in addition to the face value paid at maturity

[2]for more information see `https://www.moodys.com/ratings-process/Ratings-Definitions/002002`.

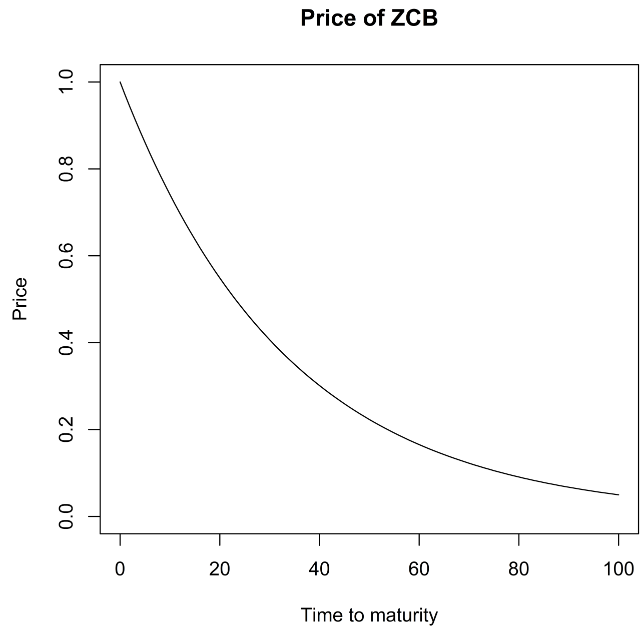


Figure 3.1: Price of a ZCB with fixed interest rate, $r = 0.03$.

Definition 3.3. ***Zero Coupon Bond (ZCB)*** *A T-maturity ZCB (pure discount bond) is a contract that guarantees its holder the payment of one unit of currency at time T, with no intermediate payments. The contract value at time $t < T$ is denoted by $P(t,T)$. Clearly $P(T,T) = 1$ for all T.*

The borrower of the ZCB gains/loses on the difference between the payment at time zero and the amount they receive at maturity. A monetary unit today is not worth the same tomorrow and to relate the different time values of currencies, we simply compare the ZCB prices with different maturity times T. E.g. we want to answer: *how much do we need to pay today to get a dollar back in 10 years?*

3.3.1 Corporate bonds

A corporate bond is an investment in debt security issued by a corporation. In other words, the investors are lending money to the company issuing the bond, e.g. the company promises today (at time $t = 0$) a payment of one unit at the time of maturity of the contract ($= T$). The investors do not own equity by the company and hence do not receive any dividends declared and paid in the company. The investors receive interest and principal of the bond, regardless of how profitable the company becomes. By issuing the bonds a corporation commits itself to make specified payments to the bondholders at some future date (fixed maturity), and the corporation charges a fee for this commitment.

The backing for the bond, is the payment ability of the company. There is a chance the company may fail to pay back the debt and default of payment may occur at a random time $\tau, 0 \leq \tau \leq T$. This default risk makes the creditworthiness of the company. Hence corporate bonds are more risky than government bonds, which are considered to be risk-free, commonly referred to as treasury bonds.

There will always be a probability that the company defaults. As a result of this the bondholders are exposed to risk and hence the implied interest rates of corporate bonds are usually higher than for treasury bonds. Measures of the excess return on a corporate bond on an equivalent treasury bond is referred to as the *credit spread*[3].

The price of a zero-coupon corporate bond at time $t, 0 \leq t \leq T$ is given by:

$$p(t,T)^d = \mathbf{E}[\underbrace{e^{-r(T-t)}\mathbf{1}_{\{\tau>T\}}}_{\text{no default}} + \underbrace{e^{-r(T-t)}R\mathbf{1}_{\{\tau\leq T\}}}_{\text{default}}|\mathcal{F}_t],$$

under some pricing measure, where R is the recovery rate and recovery payments are done at time of default, $\tau < T$. In the financial jargon, it is common to use the generic term *loss given default (LGD)* to describe the loss of value in case of default. When the face value of a zero-coupon corporate bond is one, LDG equals $(1 - R)$.

At time zero, the price of a zero-coupon corporate bond is given by:

$$\begin{aligned} p^d(0,T) &= \mathbf{E}[\underbrace{e^{-rT}\mathbf{1}_{\{\tau>T\}}}_{\text{no default}} + \underbrace{e^{-rT}R\mathbf{1}_{\{\tau\leq T\}}}_{\text{default}}|\mathcal{F}_t] \\ &= e^{-rT}\mathbb{P}(\tau > T) + \mathbf{E}[e^{-rt}R\mathbf{1}_{\{0\leq\tau\leq T\}}] \\ &= e^{-rT}\mathbb{P}(\tau > T) + Re^{-rT}\int_0^T d\mathbb{P}(\tau \leq t), \end{aligned}$$

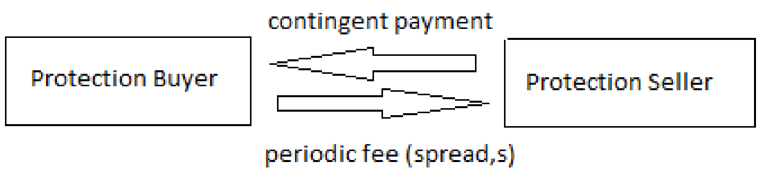
where we need a model which specifies the probability of default, $\mathbb{P}(\tau \leq t)$.

[3]The main goal of many credit models is to determine the credit spread.

3.4 Credit Default Swap(CDS)

Definition 3.4. ***Credit Default Swap (CDS).***
Exchange of a periodic payment against a one-off contingent payment if some credit event occurs on a reference asset.

Credit default swaps, also known as default insurances, are basic protection contracts, which have become quite popular in the last few years. The contract ensures protection against default.



In these agreements, periodic fixed payments from the protection buyer are exchanged for the promise of some specified payment from the protection seller to be made only if a particular, pre-specified credit event occurs (typically a default at time τ). If a credit event occurs during the life time of the default swap, $t \leq \tau$ and $0 \leq t \leq T$, the seller pays the buyer an amount to cover the loss, and the contract then terminates. If no credit event has occurred prior to maturity of the contract, $T > \tau$, both sides end their obligation to each other.

3.4.1 Pricing CDS

We look at the payment schedule $0 < t_1 < \ldots < t_n = T$. The expected discounted cash flows, with a deterministic interest rate r can then be represented as:

$$ED_{buyer} = \mathbf{E}[\sum_{i=1}^{n} e^{-rt_i} k \mathbf{1}_{\tau > t_i}]$$
$$EDS_{seller} = \mathbf{E}[e^{-r\tau}(1-R)\mathbf{1}_{\tau \leq T}].$$

To obtain a fair spread we set $ED_{buyer} = EDS_{seller}$ and solve the equation for k. This gives:

$$k = \frac{\mathbf{E}[e^{-r\tau}(1-R)\mathbf{1}_{\{\tau \leq T\}}]}{\sum_{i=1}^{n} e^{-rt_i}\mathbf{E}[\mathbf{1}_{\{\tau > t_i\}}]}$$

and by modelling the default probabilities the premium k can be evaluated.

3.5 Portfolio credit derivatives

Derivatives on credit portfolios are products with a payment stream which depends on credit-risky assets. The value of such a portfolio depends on the individual default probabilities of the assets in the portfolio, and the dependence structure within the portfolio (such as macro-economic variables, industry sector and geographic location).

3.5.1 Index (Portfolio) CDS

Assuming a portfolio consist of l assets, and the nominal of each asset is denoted by N_i, the portfolio has the face value: $N = N_1 + \ldots + N_l$.

The portfolio loss process is given by:

$$L_t = \sum_{i=1}^{l} (1 - R_i)\mathbf{1}_{\tau_i \leq t} \tag{3.1}$$

for $t \in [0, T]$. Given the loss of the portfolio we can calculate the remaining notional of the portfolio as the initial value minus the loss:

$$N_t = N_0 - L_t. \tag{3.2}$$

The expected discounted cash flows for the protection seller and the protection buyer, with a deterministic interest rate r becomes:

$$\begin{aligned} ED_{buyer} &= \mathbf{E}[\sum_{i=1}^{n} e^{-rt_i} k \Delta t_i N_{t_i}] \\ EDS_{seller} &= \mathbf{E}[\sum_{i=1}^{n} e^{-rt_i}(L_{t_i} - L_{t_{i-1}})], \end{aligned}$$

where $\Delta t_i = t_i - t_{i-1}$, for $i = 1, \ldots, n$.

To obtain a fair spread we set $ED_{buyer} = EDS_{seller}$ and solve the equation for k.

Chapter 4

Modelling Credit Risk

Models to describe default processes for defaultable financial instruments are primarily divided into two models: *structural* and *reduced form*.

Structural modelling (also referred to as the *firm value modelling*).
By using the structural modelling approach we consider the credit risk that is specific to a particular firm. That is, the credit event is moved by the firm's value relative to some threshold τ.

Reduced form modelling (also referred to as the *intensity based modelling*).
By using the reduced form modelling approach we do not consider the relation between default and the value of the firm. In contrast to the structural approach, the default is defined as the first jump of an exogenously given jump process. E.g. the default time $\tau, 0 \leq \tau \leq T$ is the first jump of typically a Poisson process.

In this chapter we will focus on the structural modelling approach, especially the Merton model (with extension), since we will later on compare a new model with this particular model.

The theory of this chapter is mainly borrowed from [Mer75], [Ben04],[BM01] and [BR02].

4.1 The Merton model

The Merton model is an application of the Black & Scholes option pricing model to the firm's debt. The Merton model considers a company to default if it does not have the ability to pay back its debt by the time of maturity T. That is, at the maturity data T default is defined when the value of the liabilities exceeds the value of the assets in the balance sheet. In other words, if the obligations of the firm is less than the liabilities we define it as a *default*.

The Merton model makes the following assumptions [Mer75]:

1 Frictionless market. There are no transaction costs, bankruptcy costs or taxes. Assets are divisible and trading takes place continuously in time with no restrictions on short selling of all assets. Borrowing and lending is possible at the same, constant interest r.

2 There are sufficient investors in the market place with comparable wealth levels, such that each investor can buy as much of an asset he wants at the market price. And the stock pays no dividends or other distributions during the life of the option.

3 The risk-free interest rate r is constant and known with certainty. This means that the discount factor is given by:

$$B(t,T) = e^{-r(T-t)}$$

4 The option is European, which means that it only can be exercised at the time of maturity T.

5 The evolution of the firm's value V_t follows the dynamics:

$$dV_t = (\mu_V - \gamma)V_t dt + \sigma_V V_t dB_t \tag{4.1}$$

where μ_V is the expected return on the firm's assets per unit time, $\gamma > 0$ is the payout of the firm per unit time, which means that if $\gamma < 0$ then there is an inflow of capital. σ_V is the volatility (constant) of the firm's assets per unit time, and B_t is a Brownian motion.

We use Itô's formula on equation (4.1):

Introduce the function:

$$f(t,x) = V_0 \exp((\mu_V - \gamma - \frac{1}{2}\sigma_V^2)t + \sigma_V x)$$

and calculate:

$$\begin{aligned}
\frac{\partial f(t,x)}{\partial t} &= (\mu_V - \gamma - \frac{1}{2}\sigma_V^2) f(t,x) \\
\frac{\partial f(t,x)}{\partial x} &= \sigma_V f(t,x) \\
\frac{\partial^2 f(t,x)}{\partial x^2} &= \sigma_V^2 f(t,x).
\end{aligned}$$

We find from Itô's formula with $X_t = B_t$ that:

$$df(t,B_t) = (\mu_V - \gamma - \frac{1}{2}\sigma_V^2) f(t,B_t)dt + \sigma_V f(t,B_t)dB_t + \frac{1}{2}\sigma_V^2 f(t,B_t)(dB_t)^2$$

Since $(dB_t)^2 = dt$ we are left with the dynamics of equation (4.1).

From these calculations we find that the firms value at time $t, 0 \leq t \leq T$ can be written as:

$$V_t = V_0 \exp((\mu_V - \gamma - \frac{1}{2}\sigma_V)t + \sigma_V B_t), \tag{4.2}$$

where V_0 is the value of the firm today.

Figure 4.1 and 4.2 exhibit simulated paths of the firm's value given by equation (4.2). As we can see from Figure 4.2 the sample paths increase in the case where $\gamma < 0$. As we know this is when there is an inflow of capital (payout in Figure 4.1 with $\gamma > 0$).

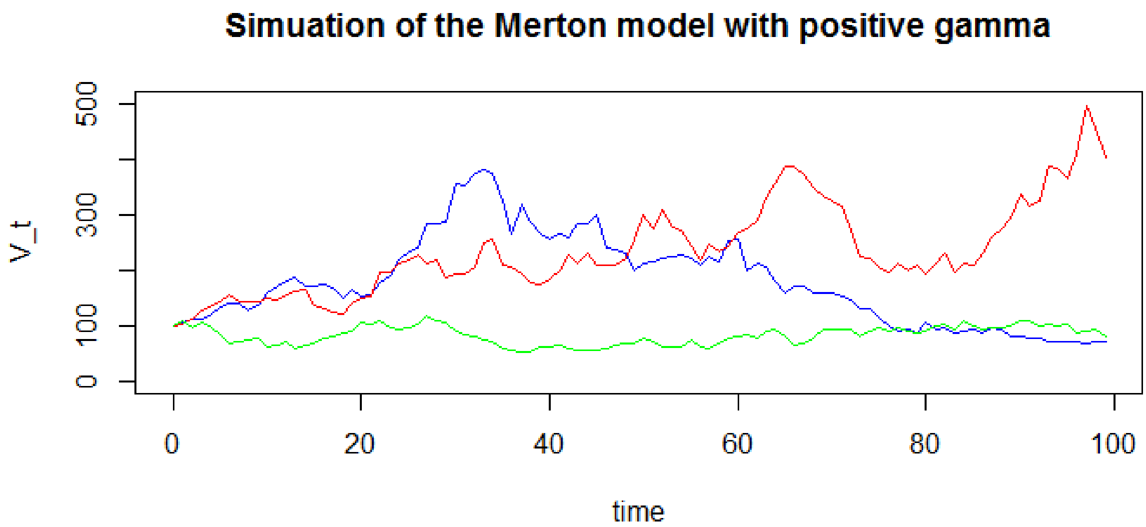


Figure 4.1: Three sample paths of the Merton model with parameters $\mu_V = 0.02, \gamma = 0.01$ and $\sigma_V = 0.09$

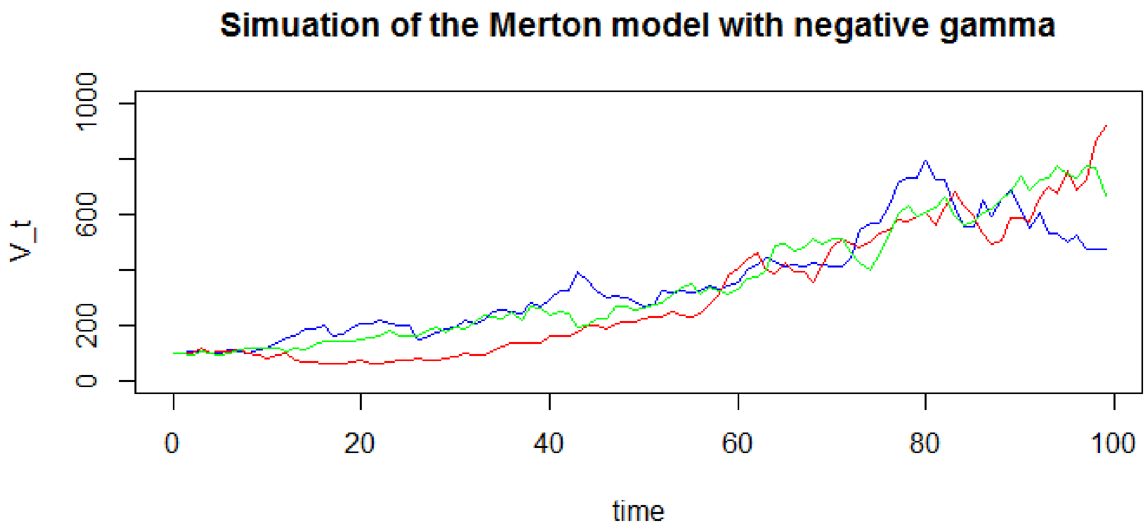


Figure 4.2: Three sample paths of the Merton model with parameters $\mu_V = 0.02, \gamma = -0.01$ and $\sigma_V = 0.09$

4.1.1 Finding the value of the firm with Merton

Under the assumptions of the Merton model, the firm's value can be found by the following equation:

$$V_t = E_t + D_t, \tag{4.3}$$

where E_t is the notation for the firm's equity and D_t denotes the firm's debt at time t, $0 \leq t \leq T$.

Let us now consider the debt D_t to be a defaultable ZCB with face value D.

This means that:

$$\text{Firm's value} < \text{Value of } D$$

implies that the firm defaults. And similarly

$$\text{Firm's value} > \text{Value of } D$$

implies that the assets of the firm exceed the liabilities and hence there are no default.

The replicating portfolio at maturity T for the payoff to the bondholder can be expressed as:

Debt

$$\begin{aligned} D_T &= \min(D, V_T) \\ &= D - \underbrace{\max(D - V_T, 0)}_{\text{Put option}}, \end{aligned}$$

where the put option represents the loss given default (LGD). And as we can see the bond can be hedged by buying a put.

Similarly:
Equity

$$\begin{aligned} E_T &= V_T - \min(V_T, D) \\ &= \underbrace{\max(V_T - D, 0)}_{\text{Call option}}. \end{aligned}$$

The value of the debt for all $t, 0 \leq t \leq T$ is given by:

$$D_t = e^{-r(T-t)}D - P_t, \tag{4.4}$$

where P_t denotes the price of the put option at time t and r is the deterministic interest rate.

By the put-call parity[1] we have that:

$$\begin{aligned} E_t &= V_t - D_t \\ &= V_t - De^{-r(T-t)} + P_t \\ &= C_t, \end{aligned}$$

where C_t denotes the price of the call option.

Theorem 4.1. ***Black & Scholes Option pricing formula**, [Ben04]*
The price of a call option with strike D and exercise time T is

$$C_t = V_t\Phi(d_1) - De^{-(\mu_V-\gamma)(T-t)}\Phi(d_2), \tag{4.5}$$

where

$$\begin{aligned} d_1 &= d_2 + \sigma_V\sqrt{T-t} \\ d_2 &= \tfrac{ln(\frac{V_t}{D})+(\mu_V-\gamma-\sigma_V^2/2)(T-t)}{\sigma_V\sqrt{T-t}} \end{aligned}$$

and

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

Since we are needing the Black & Scholes formula this also means that we will work under the risk neutral measure $\mathbb{Q}$ and hence by solving the risk neutral expected discount payoff, it is simply straight forward by applying Theorem 4.1, we find that the equity value at time $t, 0 \leq t \leq T$ is given by:

$$\begin{aligned} E_t &= C_t \\ &= V_t\Phi(d_1) - De^{-(\mu_V-\gamma)(T-t)}\Phi(d_2), \end{aligned}$$

where $\Phi(d_2)$ denotes the probability of exercising the call option. In other words the probability of no default.

[1] See Appendix

4.1.2 Maximum likelihood estimation

We know that the firm's value at time $t, 0 \leq t \leq T$ is given by:

$$V_t = V_0 \exp(\underbrace{(\mu_V - \gamma - \frac{1}{2}\sigma_V)}_{=\alpha} t + \sigma_V B_t). \tag{4.6}$$

We now set $t_i - t_{i-1} = \Delta t$ and look at the logarithmic transformation of V_t:

$$X_{t_i} = \log(\frac{V_{t_i}}{V_{t_{i-1}}}) = \alpha \Delta t + \sigma_V (B_{t_i} - B_{t_{i-1}})$$

for $i = 1, 2, \ldots$

From the definition of Brownian motion we know that the increments $B_{t_i} - B_{t_{i-1}}$ for $i = 1, 2, \ldots$, are independent and normally distributed random variables with zero expectation and variance Δt. This means that

$$X_{t_i} \sim \mathcal{N}(\alpha \Delta t, \sigma_V^2 \Delta t).$$

Further, the X_{t_i}'s are i.i.d. and by using the maximum likelihood technique we can estimate α and σ_V.

Having N logarithmic transformations of V_t, the maximum likelihood estimators of α and σ_V^2 are given by

$$\hat{\alpha} = \frac{1}{N\Delta t} \sum_{i=1}^{N} x_i, \tag{4.7}$$

$$\hat{\sigma}_V^2 = \frac{1}{N\Delta t} \sum i = 1^N (x_i - \Delta\hat{\alpha})^2. \tag{4.8}$$

Remark: We can choose which time scale we may prefer e.g. days, months or years.

4.1.3 Merton's jump diffusion model

As we have mentioned earlier in the chapter of Lévy processes, the market may have sudden changes and this will affect the firm's value. Zhou (1996) extended the Merton approach by modelling the firms value process V_t as a geometric jump-diffusion process [BR02]. By including a jump component to equation (4.1):

$$dV_t = (\mu_V - \gamma)V_t dt + \sigma_V V_t dB_t + V_t dY_t, \tag{4.9}$$

where Y_t denotes a compound Poisson process.

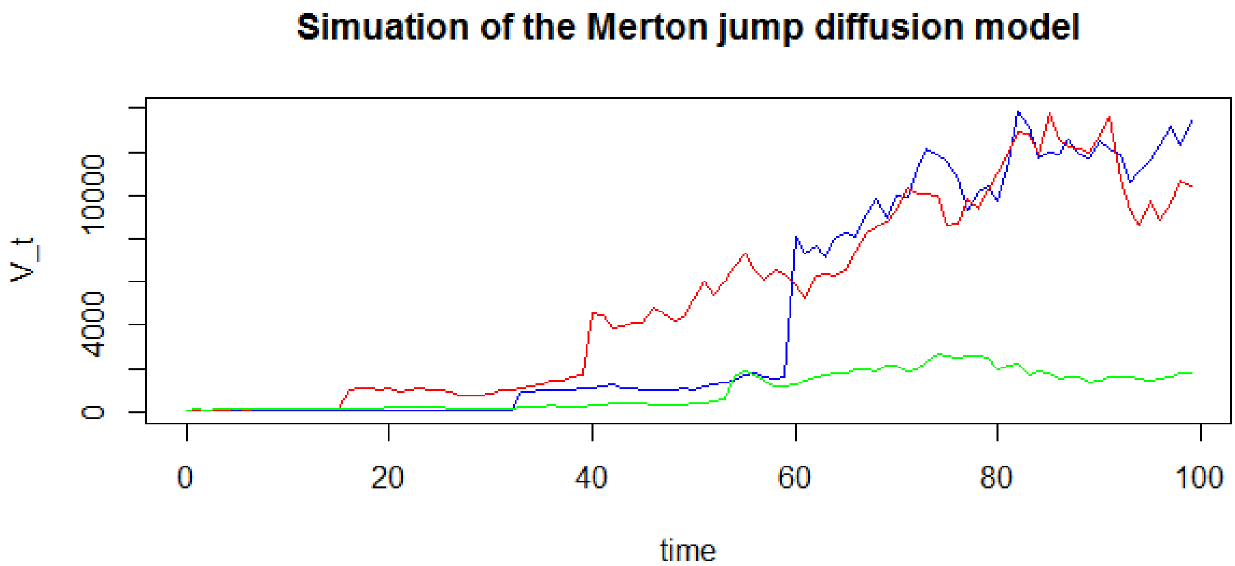


Figure 4.3: Three sample paths of the jump diffusion Merton model (also called the Zhou model) with parameters $\mu_V = 0.02, \gamma = 0.01$, $\sigma_V = 0.09$ and a compounded Poisson process $Y_t = \sum_{i=1}^{N(\lambda)} X_i$ with exponential distributed X_i.

From Figure 4.3 we see the jumps are causing sudden changes in the trajectories compared to Figure 4.1 and Figure 4.2.

4.2 First-passage modelling

As we have seen in this chapter the Merton model only concludes that a firm has (not) defaulted by the time of maturity, T. This is not a realistic assumption since the firm can default at any time. In response to this, within the structural framework, is to model the default as the first passage time that the firm's asset value falls below a certain threshold, d.[2]
The first-passage time is then modelled by:

$$\tau = \inf\{t > 0 : V_t \leq d\}. \tag{4.10}$$

4.3 Challenges of the modelling of credit risk

We have in this chapter focused on the structural modelling approach, represented by the Merton model.

By using the Merton model we can directly use the Black & Scholes option pricing formula. This is clearly an advantage but as we stated earlier in this chapter the model requires many and some unrealistic assumptions to be fulfilled, e.g. the firm has a single issue of zero-coupon debt. Hence the question becomes *does the purpose of the model disappears by requiring all these assumptions?*

Another challenge of the Merton model is the restrictions of the default time. That is, a default can only be defined at the time of maturity of the debt and hence a default can not occur at an earlier stage during the period. However, this can be captured by extended models such as the first-passage time. On the other hand, by using the Merton approach we avoid to determine a default in the case where the firm's value falls to a minimum level before maturity but manages to recover and meet the payment of the debt by maturity.

[2]This idea was introduced by Black and Cox (1976), for more information of such models we refer to the book [BM01],page 702-704.

Chapter 5

Non-linear filtering theory

Estimating unobservable variables from empirical data is a familiar problem in mathematical finance. An important issue in pricing and risk analysis is to estimate the dynamics of the underlying assets. Non-linear filtering has been studied in literature since the 1960's and by introducing some of these techniques, we can extract information from the observed process and estimate an unobserved process.

This chapter will give a summary of the basic concepts concerning non-linear filtering theory, which give us the mathematics we need for solving our problem in the following chapter.

The material in this chapter are mainly borrowed from [MBP03], [BDP15] and [MPMB09].

5.1 What is non-linear filtering?

In non-linear filtering theory we consider a partially observable process $(X, Y) = (X_t, Y_t)_{0 \leq t \leq T} \in \mathbb{R}^2$ defined on a probability space $(\Omega, \mathcal{F}_t, \mathbb{P})$.

$Y_t, 0 \leq t \leq T$ is the observed process which we extract information from into the process $X_t, 0 \leq t \leq T$ that we want to estimate. This means that the unobserved process is partially observed by the observable process.

1. The process $Y_t, 0 \leq t \leq T$ where we have the information is called *the observation process.*

2. The process $X_t, 0 \leq t \leq T$ that we estimate by extracting information

from the observation process $Y_t, 0 \leq t \leq T$ is referred to as *the signal process.*

5.1.1 Filtering problem

Let the parametrization process $X_t, 0 \leq t \leq T$ be the signal process, in the non-linear filtering problem. The process follows the dynamics given by the stochastic differential equation(SDE):

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t^X \quad , 0 \leq t \leq T, \tag{5.1}$$

where b, σ are Borel functions and $B_t^X, 0 \leq t \leq T$ is a Brownian motion.

Since this is a filtering problem it means that we can extract some information about $X_t, 0 \leq t \leq T$ out of the observation process $Y_t, 0 \leq t \leq T$, which is described by the equation:

$$dY_t = h(t, X_t)dt + \sigma dB_t^Y + \int_{\mathbb{R}_0} \zeta N_\lambda(dt, d\zeta) \quad , 0 \leq t \leq T, \tag{5.2}$$

where h is the (non-linear) observation function, σ is a constant, $B_t^Y, 0 \leq t \leq T$ is a Brownian motion process and N_λ is an integer-valued random measure with the predictable compensator:

$$\hat{\mu}(dt, d\xi, \omega) = \lambda(t, X_t, \xi)dt\nu(d\xi) \tag{5.3}$$

for a Lévy measure ν and a function λ[1].

As we can see from equation (5.2) the dynamics of the observation process consists of an information drift dependent on the signal process. The two other components are some Gaussian noise plus a pure jump part, whose jump intensity depends on the signal. The pure jump part is independent of the Brownian motion part.

The aim of this thesis is to obtain a least square estimate of $f(X_t)$ given the observations up to time $t, 0 \leq t \leq T$. In other words, evaluate the optimal filter, given by the following conditional expectation:

$$\mathbf{E}_{\mathbb{P}}[f(X_t)|\mathcal{F}_t^Y], \tag{5.4}$$

[1] For further restriction of the λ-function we refer to the book [Sch03].

where $\mathbf{E}_{\mathbb{P}}$ denotes the expectation w.r.t. $\mathbb{P}$ and f is a suitable real-valued, Borel measurable function and $\mathcal{F}_t^Y$ is a σ-algebra generated by the observations $Y_s, 0 \leq s \leq t \leq T$.

This estimate depends, in general, non-linearly on the observations and is known as *the non-linear filter.*

In order to have a strong solution to the system (5.1) and (5.2), we require that the coefficients b, σ, h and λ fulfill a linear growth and Lipschitz condition, that is:

$$||b(x)|| + ||\sigma(x)|| + ||h(t,x)|| + \int_{\mathbb{R}_0} |\lambda(t,x,\zeta)|\nu(d\zeta) \leq C(1 + ||x||) \tag{5.5}$$

and

$$\begin{aligned} &||b(x) - b(y)|| + ||\sigma(x) - \sigma(y)|| + ||h(t,x) - h(t,y)|| \\ &+ \int_{\mathbb{R}_0} |\lambda(t,x,\zeta) - \lambda(t,y,\zeta)|\nu(d\zeta) \\ &\leq C||x - y|| \end{aligned} \tag{5.6}$$

for all x, y, t and a constant $C < \infty$, where $|| \cdot ||$ stands for a vector or matrix norm.

5.1.2 Non-linear filtering techniques and theory

To solve the non-linear filtering problem we need some techniques as well as theory.

From the paper [BDP15] we consider the density process:

$$\begin{aligned} \Lambda_t = \exp\Big(&\int_0^t h(s,X_s)dB_s - \frac{1}{2}\int_0^t h^2(s,X_s)ds - \int_0^t \int_{\mathbb{R}_0} \log(\lambda(s,X_s,z)N(ds,dz)) \\ &- \int_0^t \int_{\mathbb{R}_0} (1 - \lambda(s,X_s,z))ds\nu(dz)\Big) \quad ,0 \leq t \leq T \end{aligned}$$

and assume that

$$\mathbf{E}_{\mathbb{P}}[\Lambda_T] = 1$$

as well as

$$\int_{\mathbb{R}_0} |z|\nu(dz) < \infty.$$

Lemma 5.1. *Define*

$$d\mathbb{Q} = \Lambda_t d\mathbb{P}.$$

Then $\mathbb{Q}$ is a probability measure, and under $\mathbb{Q}$ we have that:

$$Y_t = B_t + L_t,$$

where $B_t = B_t^Y - \int_0^t (-h(s, X_s))ds$, $0 \le t \le T$ is a Brownian motion part, $L_t = \int_0^t \int_{\mathbb{R}_0} \zeta N(ds, d\zeta), 0 \le t \le T$ is a pure jump Lévy process with respect to the Poisson random measure N with compensator $ds\nu(d\xi)$. Here we assume that $\sigma = 1$ in equation (5.2).

Further, the process B, L and X are independent under $\mathbb{Q}$ and Λ_t is a martingale.

The inverse Radon-Nikodym derivative is denoted by:

$$\frac{d\mathbb{P}}{d\mathbb{Q}} = (\Lambda_t)^{-1} = M_t, 0 \le t \le T,$$

where

$$\begin{aligned} M_t &:= \Lambda_t^{-1} \\ &= \exp\Big(\int_0^t h(s, X_s)dB_s - \frac{1}{2}\int_0^t ||h(s, X_s)||^2 ds\Big) \\ &+ \int_0^t \int_{\mathbb{R}_0^m} \log \lambda(s, X_s, \zeta)N(ds, d\zeta) \\ &+ \int_0^t \int_{\mathbb{R}_0^m} (1 - \lambda(s, X_s, \zeta))ds\nu(d\zeta), \quad 0 \le t \le T. \end{aligned} \tag{5.7}$$

Then by [BDP15] and Girsanov's theorem the observation process $Y_t, 0 \le t \le T$ becomes a Lévy process being independent of the signal process $X_t, 0 \le t \le T$ under the new probability measure $\mathbb{Q}$.

In other words, the system of (5.1) and (5.2) has the following representation under $\mathbb{Q}$:

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t^X, 0 \le t \le T \tag{5.8}$$
$$dY_t = dB_t + dL_t, 0 \le t \le T. \tag{5.9}$$

For any $\mathcal{F}_T^X$-integrable function f, the Kallianpur-Striebel formula (which is a consequence of Bayes formula for conditional expectation) is given by:

$$\mathbf{E}_{\mathbb{P}}[f(X_t)|\mathcal{F}_t^Y] = \frac{\mathbf{E}_{\mathbb{Q}}[M_t f(X_t)|\mathcal{F}_t^Y]}{\mathbf{E}_{\mathbb{Q}}[M_t|\mathcal{F}_t^Y]} \tag{5.10}$$

where $M_t, 0 \le t \le T$ is given by (5.7).

In fact, in the case of Lipschitz continuous coefficients[2] b, σ, h and λ we have the following proposition:

Proposition 5.2. *Assume that the functions b, σ, h and λ are bounded and satisfy conditions (5.5) and (5.6). Let $X_t^i, 0 \le t \le T, i \ge 1$ be a sequence of i.i.d. copies of the solution of the signal process $X_t, 0 \le t \le T$:*

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t^X$$

on our probability space, being independent of the observation process $Y_t, 0 \le t \le T$. Denote by $M_t, 0 \le t \le T$ the stochastic exponential:

$$\begin{aligned} M_t &= (\Lambda_t)^{-1} \\ &= \exp\Big(\int_0^t h(s, X_s)dB_s - \frac{1}{2}\int_0^t h^2(s, X_s)ds \\ &+ \int_0^t \int_{\mathbb{R}_0} \log(\lambda(s, X_s, \zeta)N(ds, d\zeta)) \\ &+ \int_0^t \int_{\mathbb{R}_0} (1 - \lambda(s, X_s, \zeta))ds\nu(d\zeta)\Big) \end{aligned} \tag{5.11}$$

based on $X_t, 0 \le t \le T$. Let f be a bounded, continuous function. Then:

$$Z^l(f) = \frac{1}{l}\sum_{i=1}^{l} M_t^i f(X_t^i) \xrightarrow{l \to \infty} \mathbf{E}_{\mathbb{Q}}[M_t f(X_t)|\mathcal{F}_t^Y] \quad a.e. \tag{5.12}$$

[2]for definition of Lipschitz continuous coefficients see Appendix.

for all t*. Moreover, for all* t *there exists a constant* C *such that:*

$$\mathbf{E}_{\mathbb{Q}}[(Z^l(f) - \mathbf{E}_{\mathbb{Q}}[M_t f(X_t)|\mathcal{F}_t^Y])^2] \leq \frac{1}{l} C \parallel f \parallel^2 \tag{5.13}$$

for all $l \geq 1$*.*

Proof. For a proof see [BDP15]. □

Chapter 6

Non-linear filtering applied to our new model

As we have seen earlier in this thesis, the firms value can be modelled by using the Merton model. In this chapter we introduce a new model, which can be considered as a generalization of the Merton model. The new model is simulated by using non-linear filtering techniques and empirical market data.

In the following sections we will solve a non-linear filtering problem numerically. In other words, we will estimate unobservable variables from observed data.

When the new simulation approach is implemented we will compare our new model with the Merton model by using US market data.

In the last section a summary of the results and a conclusion is given.

6.1 Our new model

The signal process in the new model is given by:

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t^X \quad , 0 \leq t \leq T, \tag{6.1}$$

where $B_t^X, 0 \leq t \leq T$ is Brownian motion and in our model we will set the functions b and σ to be

$$b(\tilde{x}) = \xi_2(\xi_3 - x)$$

$$\sigma(\tilde{x}) = \xi_4.$$

As we can see the signal process, $X_t, 0 \leq t \leq T$, is now described by the Vasicek model, where ξ_2 is the speed of reversion, ξ_3 is the long term mean level and ξ_4 is the instantaneous volatility.

The observation process, which is our new model for the observed price process S_t, takes the form:

$$\begin{aligned} Y_t &= \log(S_t) - \log(S_0) \\ &= at + \sigma B_t + \int_0^t \int_{\mathbb{R}_0^4} \zeta N_\lambda(dt, d\zeta) \quad, 0 \leq t \leq T, \end{aligned} \tag{6.2}$$

where S_t is the stock price at time $t, 0 \leq t \leq T$, $B_t, 0 \leq t \leq T$ is Brownian motion and $N_\lambda(ds, d\zeta)$ is the jump measure with compensator of the form

$$\hat{\mu}(dt, d\zeta) = \lambda(t, X_t, \zeta) dt \nu(d\zeta).$$

Further we assume that $a = \sigma = 0$, hence the observation process is given by a pure jump process.

6.2 Estimation of the parameters of the signal process

Let us consider a λ-function with respect to the compensator given by

$$\lambda(t, X_t, \zeta) = \frac{\mathcal{N}_\zeta(X_t^{(1)}, X_t^{(4)})}{\mathcal{N}_\zeta(0,1)}, 0 \leq t \leq T, \tag{6.3}$$

where $\mathcal{N}_\zeta(\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{1}{2}(\frac{\zeta-\mu}{\sigma})^2)$, is the Gaussian density with mean μ and standard deviation σ and $X_t, 0 \leq t \leq T$ is a multidimensional signal process given by (6.5).

Our main goal is to determine the conditional expectation

$$\mathbf{E}_{\mathbb{P}}[f(X_t) | \mathcal{F}_t^Y] \tag{6.4}$$

where f is a measurable function and $\mathcal{F}_t^Y$ is the σ-algebra, generated by $Y_s, 0 \leq s \leq t$.

To simulate the observation process we need to determine a prior distribution for i.i.d. ξ_i and we choose this to be $\xi_i \sim \mathcal{N}(0,1) \quad \forall i$.

The corresponding filter problem results in the following four-dimensional signal process:

$$dX_t = \begin{cases} dX_t^{(1)} = X_t^{(2)}(X_t^{(3)} - X_t^{(1)})dt + X_t^{(4)}dB_t^X \\ dX_t^{(2)} = 0 \\ dX_t^{(3)} = 0 \\ dX_t^{(4)} = 0 \qquad , 0 \le t \le T, \end{cases} \tag{6.5}$$

and the observation process is given by equation (6.2).

The start values $X_{t_0}^{(1)} = \xi_1, X_{t_0}^{(2)} = \xi_2, X_{t_0}^{(3)} = \xi_3, X_{t_0}^{(4)} = \xi_4$ are drawn from a normal distribution with mean zero and variance equal to one.

Further we assume that the Lévy-measure, ν, is given by:

$$\nu(A) = \int_A \mathcal{N}_z(0,1)dz.$$

That is, the Lévy-measure is represented by the integral of a Gaussian density.

The aim is to determine the estimates of the signal process, $X_t, 0 \le t \le T$. That is, to find $\hat{\xi_2}$, $\hat{\xi_3}$ and $\hat{\xi_4}$.

The inverse of the Radon-Nikodym density $\Lambda_t, 0 \le t \le T$ under the change of measure $\mathbb{Q}$ is given by:

$$\begin{aligned} M_t = \exp\Big(& \underbrace{\int_0^t h(s,X_s)dB_s}_{I_1} - \underbrace{\frac{1}{2}\int_0^t h^2(s,X_s)ds}_{I_2} \\ & + \underbrace{\int_0^t \int_{\mathbb{R}_0} \log(\lambda(s,X_s,\zeta)N(ds,d\zeta))}_{I_3} \\ & + \underbrace{\int_0^t \int_{\mathbb{R}_0} (1-\lambda(s,X_s,\zeta))ds\nu(d\zeta)}_{I_4} \Big), 0 \le t \le T. \end{aligned} \tag{6.6}$$

6.2.1 Simulation approach

To estimate the signal process $X_t, 0 \le t \le T$ we start by considering a discrete-time analogue of (6.5):

$$X_{t_{i+1}} = \begin{cases} X^{(1)}_{t_{i+1}} = X^{(1)}_{t_i} + X^{(2)}_{t_i}(X^{(3)}_{t_i} - X^{(1)}_{t_i})(t_{i+1} - t_i) + X^{(4)}_{t_i}(B^X_{t_{i+1}} - B^X_{t_i}) \\ X^{(2)}_{t_i} = \xi_2 \\ X^{(3)}_{t_i} = \xi_3 \\ X^{(4)}_{t_i} = \xi_4, \end{cases} \tag{6.7}$$

where

$$\begin{aligned} \Delta t &= t_{i+1} - t_i, \quad 0 = t_0 \leq t_1 \leq \ldots \leq t_n = T(= \text{maturity}) \\ X_i &= X_{t_i} \\ (B_{t_{i+1}} - B_{t_i}) &= \sqrt{t_{i+1} - t_i}\eta_i, \quad \eta_i \sim \mathcal{N}(0,1). \end{aligned}$$

Further define

$$\begin{aligned} \Delta Y_t &= Y_{t_{i+1}} - Y_{t_0} \\ &= \log(S_{t_{i+1}}) - \log(S_{t_0}) \\ &= \log(\frac{S_{t_{i+1}}}{S_{t_0}}). \end{aligned}$$

We have the observation process given by $Y_t = \log(S_t) - \log(S_0), 0 \leq t \leq T$ and we say a jump occur if:

$$|Y_s| > \bar{Y} = \tfrac{1}{N}\textstyle\sum_{i=1}^{N} |Y_i|,$$

where N is the number of observations in the dataset.

We want to simulate $M_t, 0 \leq t \leq T$ and for simplifying the calculation we set the function $h = 1$ in equation (6.6) , which gives for $0 \leq t \leq T$:

1. $I_1 = B_t$
2. $I_2 = \frac{1}{2}t$
3. $I_3 = \sum_{0<s\leq t} \log(\lambda(s, X_s, \Delta Y_s))\mathbf{1}_{|\Delta Y_s| > \bar{Y}}$

4.

$$
\begin{aligned}
I_4 &= \int_0^t \int_{\mathbb{R}_0} (1-\lambda(s,X_s,\zeta))\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\zeta^2}d\zeta ds \\
&= \int_0^t \int_{\mathbb{R}_0} \big(1-\frac{\mathcal{N}_\zeta(X_s^{(1)},X_s^{(4)})}{\mathcal{N}_\zeta(0,1)}\big)\mathcal{N}_\zeta(0,1)d\zeta ds \\
&= \int_0^t \underbrace{\int_{\mathbb{R}_0} (\mathcal{N}_\zeta(0,1)-\mathcal{N}_\zeta(X_s^{(1)},X_s^{(4)}))d\zeta}_{=0,\text{ since we intergrate densities}}\, ds \\
&= 0.
\end{aligned}
$$

We can extend Proposition 5.2 to hold for λ given by (6.3)[1]. Hence we have by the strong law of large numbers [2] and Proposition 5.2, the following result:

$$
\frac{1}{n}\sum_{i=1}^{n} M_t^i f(X_t^{(i)}) \to E_{\mathbb{Q}}[M_t f(X_t)|\mathcal{F}_t^Y] \tag{6.8}
$$

as $n \to \infty$.

Further we calculate the conditional expectation, $\mathbf{E}_{\mathbb{P}}[f(X_t)|\mathcal{F}_t^Y]$, by applying the Kallianpur-Striebel formula:

$$
E_{\mathbb{P}}[f(X_t)|\mathcal{F}_t^Y] = \frac{E_{\mathbb{Q}}[M_t f(X_t)|\mathcal{F}_t^Y]}{E_{\mathbb{Q}}[M_t|\mathcal{F}_t^Y]}, 0 \le t \le T. \tag{6.9}
$$

6.2.2 Numerical results

We use daily US Market data of the closing price for General Electric Company from January 2006 to December 2010, taken from Yahoo Finance [3]. By using these data , plotted in Figure 6.1, we will try to estimate the parameters of the signal process.

As we can see by Figure 6.2 the estimates converge, and take the following

[1] The same proof given by [MPMB09] can be used.

[2] Definition is given in Appendix.

[3] `http://finance.yahoo.com/q/hp?s=GE&a=00&b=1&c=2006&d=11&e=31&f=2010&g=d&z=66&y=0`

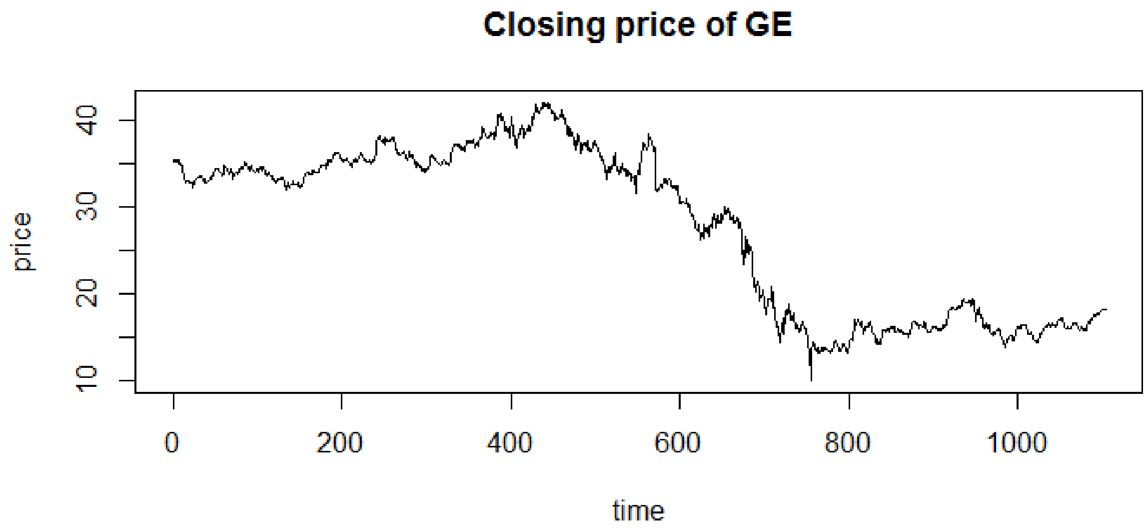


Figure 6.1: General Electric Company closing prices from January 2006 to December 2010.

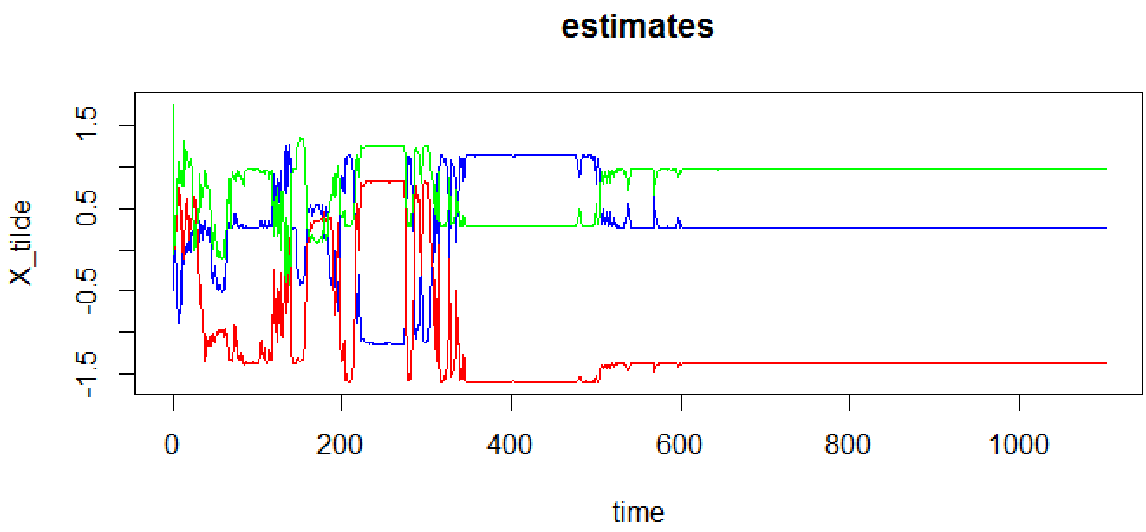


Figure 6.2: Plot of the estimates of the signal process (with General Electrics Company data), where the blue graph represents ξ_2, the red is ξ_3 and the green ξ_4.

values:

$$\xi_2 \to \hat{\xi}_2 = 0.262893$$
$$\xi_3 \to \hat{\xi}_3 = -1.376566$$
$$\xi_4 \to \hat{\xi}_4 = 0.9725912.$$

By using λ as given in equation (6.3) the compensator is given by:

$$\begin{aligned}\hat{\mu}(dt, d\zeta) &= \lambda(t, X_t, \zeta)dt\nu/(d\zeta) \\ &= \frac{\mathcal{N}_\zeta(X_t^{(1)}, X_t^{(4)})}{\mathcal{N}_\zeta(0,1)} dt\mathcal{N}_\zeta(0,1)d\zeta \\ &= \mathcal{N}_\zeta(X_t^{(1)}, X_t^{(4)})dtd\zeta.\end{aligned} \tag{6.10}$$

Unfortunately we are not able to find any literature that captures a λ given by equation (6.3) and a compensator given by equation (6.10), which provides us with an efficient tool to simulate our new model (6.2).

6.2.3 Choosing λ to be different

Using λ as is equation (6.3) we get some complications concerning further simulations. We now look at another λ-function, which will hopefully enable us to simulate our new model.

We choose

$$\lambda(t, X_t) = (\| X_t \| + 1) \tag{6.11}$$
$$= (\sqrt{|X_t^1|^2 + |X_t^2|^2 + |X_t^3|^2 + |X_t^4|^2} + 1). \tag{6.12}$$

We now have a λ-function which does not depend on the spatial variable coming from the jumps, and by letting λ be of this form we avoid the complications we get in equation (6.10).

The following expression determines the times when jumps occur

$$\mu(t) = \int_0^t \lambda(s, X_s)ds, \quad 0 \le t \le T. \tag{6.13}$$

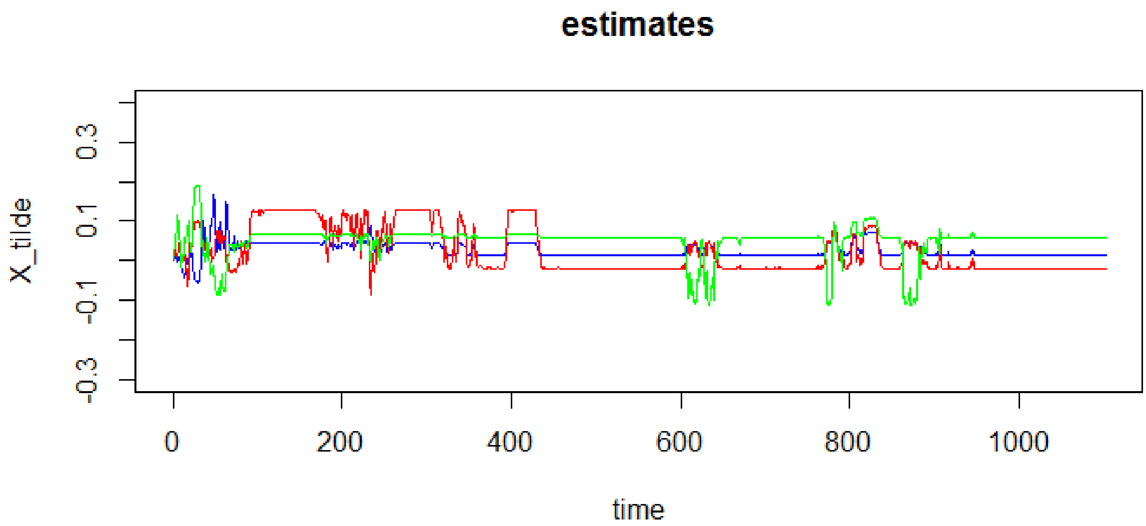


Figure 6.3: Plot of the estimates of the signal process (with General Electrics Company data and λ as is equation (6.11)), where the blue graph represents ξ_2, the red is ξ_3 and the green ξ_4

As we can see by Figure 6.3 the estimates stabilize, and takes the following values:

$$\xi_2 \to \hat{\xi_2} = 0.01440555$$
$$\xi_3 \to \hat{\xi_3} = -0.01900694$$
$$\xi_4 \to \hat{\xi_4} = 0.05796884.$$

We want to predict future closing prices of General Electrics Company. By equation (5.2) with $h = \sigma = 0$ we have the price process given by:

$$S_t = S_0 \exp(Y_t), 0 \leq t \leq T, \tag{6.14}$$

where the observation process is given by a pure jump Lévy process, $Y_t = \int_0^t \int_{\mathbb{R}_0} \zeta N_\lambda(ds, d\zeta), 0 \leq t \leq T$.

6.3 Simulations of the price process

We simulate the price process by conducting the following steps:

Step 1: **Simulation of the Vasicek model.**
Take the values $\hat{\xi}_2$,$\hat{\xi}_3$ and $\hat{\xi}_4$ into the signal process (6.5). Simulate ten different paths of the signal process , see Figure 6.4, where the red path illustrates the average. We now have ten different paths, $(X_t^{(1)}, X_t^{(2)}, \ldots, X_t^{(10)}), 0 \leq t \leq T$.

Step 2: **Simulation of the jump times.**
Take the ten paths of the signal process into the integral given by (6.13).

Step 3: **The modified Cramer Lundberg model.**[4]
Look at a compound Poisson process with

$$Z_s = \sum_{i=1}^{N(s)} X_i,$$

where $N(s)$ is a standard Poisson process with parameter $\lambda = 1$ and the jump sizes X_i are i.i.d. with $X_i \sim \mathcal{N}(0,1)$ for $i = 1, \ldots, 10$.

This gives us the following ten paths: $Z_s^{(1)}, Z_s^{(2)}, \ldots, Z_s^{(10)}$ where $0 \leq s \leq \tilde{T}$,

$$\tilde{T} = \max_{1 \leq i \leq 10, 0 \leq s \leq T} \mu^{(i)}(s),$$

where

$$\mu^{(i)}(t) = \int_0^t \lambda(s, X_s^{(i)}) ds.$$

Step 4: **The final simulation**
Set:

$$Y_t^{(1)} = Z^{(1)}_{\mu_t^{(1)}}, \ldots, Y_t^{(10)} = Z^{(10)}_{\mu_t^{(10)}}.$$

We now have ten paths of the observation process $Y_t, 0 \leq t \leq T$. And we can finally simulate paths of the closing price $S_t = S_0 \exp(Y_t), 0 \leq t \leq T$.

[4]The jump size can be negative, hence we are looking at a modified Cramer Lundberg model.

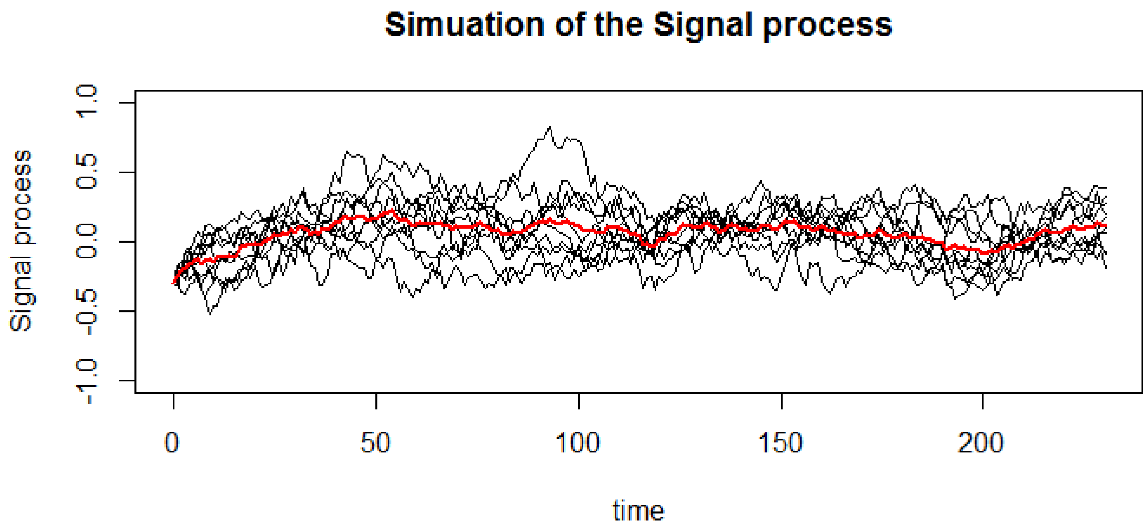


Figure 6.4: Ten paths of the signal process $X_t, 0 \leq t \leq T$, with $\hat{\xi}_2 = 0.01440555, \hat{\xi}_3 = -0.01900694$ and $\hat{\xi}_4 = 0.05796884$, the red path is the average.

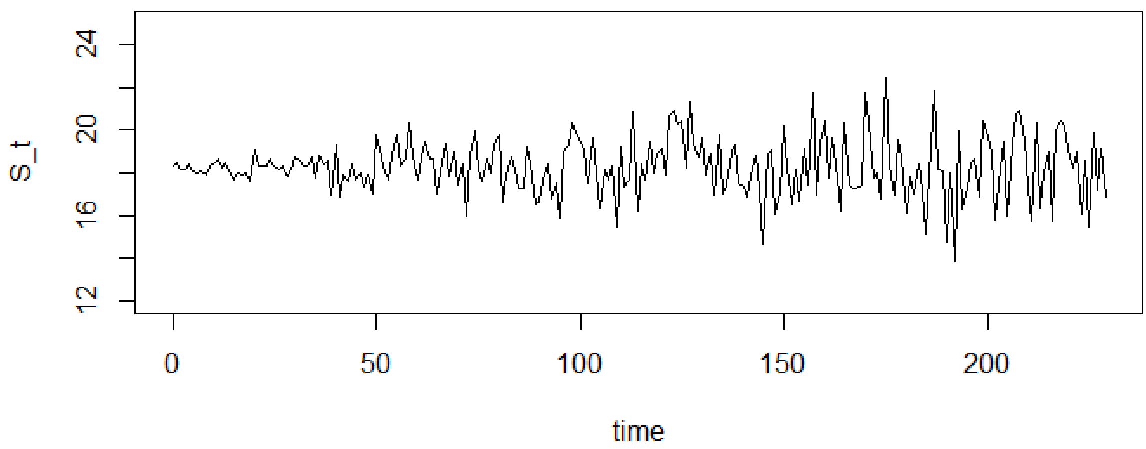


Figure 6.5: A single path of the $S_t^{(1)} = S_0 \exp(Y_t^{(1)}), 0 \leq t \leq T$.

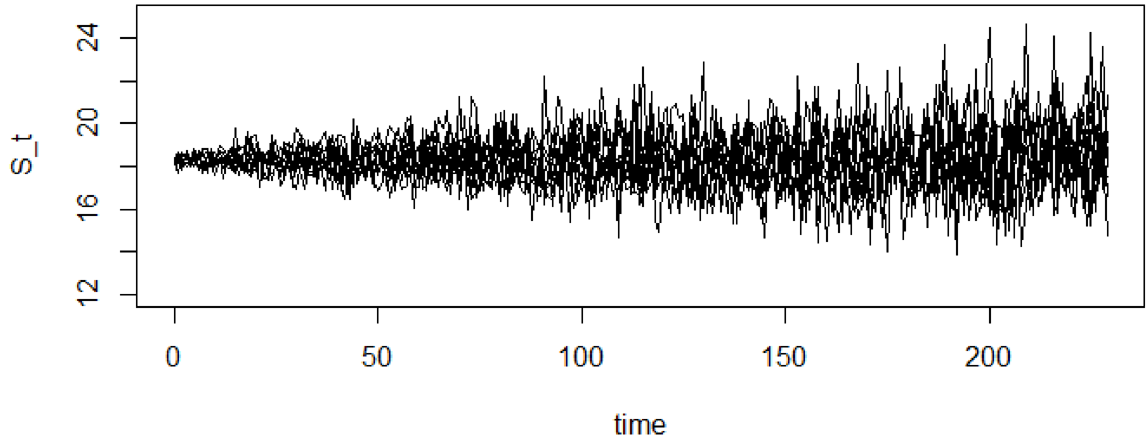


Figure 6.6: Ten paths of $S_t^{(i)} = S_0 \exp(Y_t^{(i)}), 0 \leq t \leq T$, for $i = 1, \ldots, 10$.

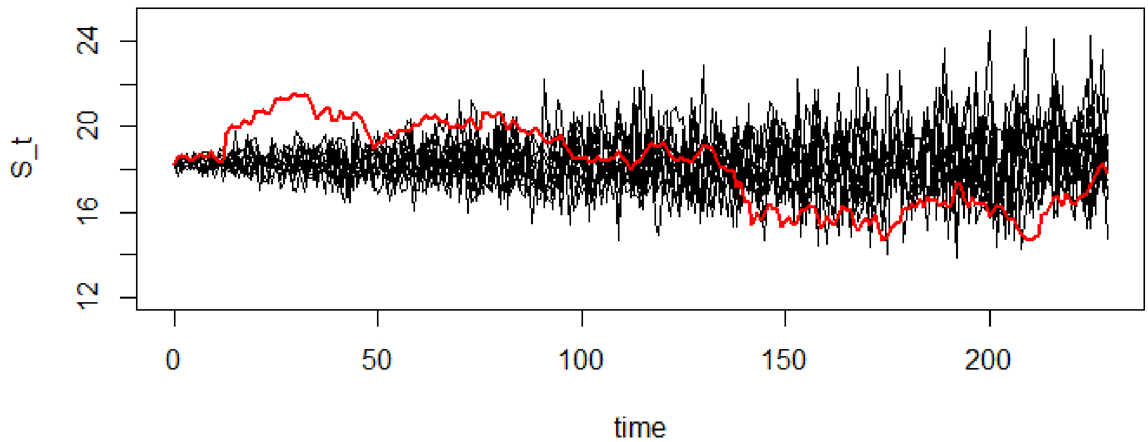


Figure 6.7: Ten paths of our estimated price process versus the observed price process of General Electrics Company from January 2011 to December 2011 (red path).

Intuitively, the uncertainty of prediction of the stock price comes greater with time. Figure 6.8 shows that this is also the case when trying to estimate the closing prices of General Electrics Company with the new model.

6.3.1 Comparing the model with the classical Merton model

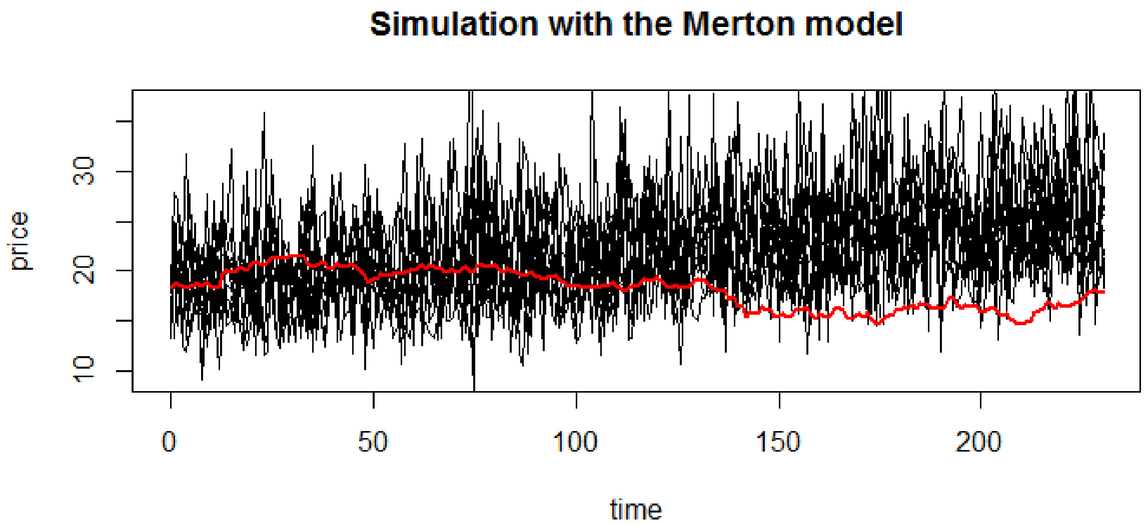


Figure 6.8: Ten simulation paths of the Merton model with calculated estimates from MLE. The red path shows the actual closing price of General Electrics Company from January 2011 to December 2011.

We have simulated the Merton model, presented in chapter 5, and the new model based on the same data set. As we can see from Figure 6.9 the predictions of the closing prices for General Electrics Company based on the Merton model is more volatile and this increases with time.

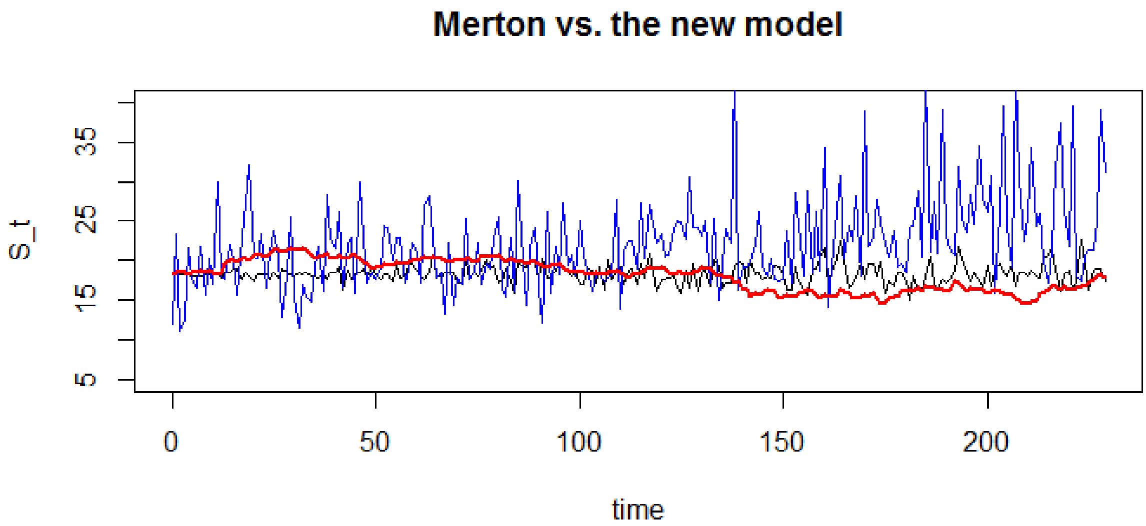


Figure 6.9: Simulation of the Merton model versus the new model. The black path is the new model, the blue path represents the Merton model and the red path is the actual closing price of General Electrics Company from January 2011 to December 2011.

6.4 Conclusion

In this thesis we have discussed the challenges of the modelling of credit risk and basic approaches concerning modelling of such risk.

We have introduced a new model and estimated parameters of this model by using non-linear filtering techniques and a new simulation method. Based on this estimation we manage to simulate closing prices of stocks of US market data. We then compared the new model with a classical credit risk model, the Merton model.

As we have seen by the result, our new model is adaptable and makes a good fit to the empirical market data. The model is quite complex, and with this comes great flexibility in the sense that we can adapt underlying densities and parameters within the model to fit the dataset. On the other hand, with the complexity comes the chance of overfitting.

However, at this point we can not state that our new model will give a closer match to the statistical properties of observed market data than the classical Merton model. As a conclusion we can state that we have manage to simulate a new model which fits market data in a good way.

Chapter 7

Extensions

In this chapter we will share ideas of further constructions and extensions of the new model presented in chapter 6. By extending the model we will try to make it more realistic in the sense that it can capture more of the economical information in the market.

7.1 The signal process

We assumed that the signal process was given by the Vasicek model, that is:

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t^X, 0 \leq t \leq T \tag{7.1}$$

where B_t^X is Brownian motion and where b and σ is given by

$$b(\tilde{x}) = \xi_2(\xi_3 - x)$$

$$\sigma(\tilde{x}) = \xi_4.$$

The Vasicek model is easy to work with and it is mean reverting. One could have used another Ornstein-Uhlenbeck process as e.g. the CIR model.

Hence, the functions b and σ would have been given by:

$$b(\tilde{x}) = \xi_2(\xi_3 - x)$$

$$\sigma(\tilde{x}) = \sqrt{x}\xi_4.$$

Regime switching

When looking at economic time series we often see dramatic breaks caused by events like a financial crisis or changes in government policy. In our model this phenomena may be captured by a *Regime switching mean-reversion model* where the function b in the signal process (7.1) is given by:

$$b(\tilde{x}) = \begin{cases} a(b_1 - x) & x \geq \tau \\ a(b_2 - x) & \text{else,} \end{cases} \tag{7.2}$$

for mean reversion coefficient $a \geq 0$ and the long-run average levels $b_1, b_2 \geq 0$ depending on a critical threshold $\tau > 0$.

The same algorithm and recursion as in chapter 6 can be used for this purpose.

N-dimensional signal process

An idea is to extend the signal process $X_t, 0 \leq t \leq T$ to a n-dimension process with $n > 4$, as we have estimated the signal process with 4-dimensions. In this way we can capture more of the economical impacts.

However, with more parameters to estimate the chance of overfitting comes greater, we might experience unstable parameter estimates and the simulations get more complicated.

Research of the convergence in the signal process

The simulations of the signal process $X_t, 0 \leq t \leq T$ showed convergence of ξ_2, ξ_3 and ξ_4. To explain this result we have to look further into the mathematics of the non-linear filtering theory [MPMB09]. This is beyond the scope of this thesis, but for further studies of the model this would have been interesting to address.

7.2 The observation process

In our new model given by (6.2) the observation process is given by a pure jump Lévy process. We could extend the model by looking at a jump diffusion process by adding a drift term and Brownian motion:

$$Y_t = \log(\frac{S_t}{S_0}) = \int_0^t h(s, X_s)ds + \sigma B_t^Y + \int_0^t \int_{\mathbb{R}_0} \zeta N_\lambda(ds, d\zeta), 0 \leq t \leq T \quad (7.3)$$

where B_t^Y is Brownian motion, h a function and $\sigma \neq 0$. This may give the model a better fit to the dataset we are looking at.

7.3 Intensity based credit models

We have in this thesis focused on the structural modelling approach. As mentioned there is also an approach called *Intensity based modelling.* We could extend this thesis by considering an alternative to the structural framework.

In intensity-based credit risk models the default time $\tau, 0 \leq \tau < T$ of a defaultable bond can be modelled by a Hazard rate process $\gamma_t, 0 \leq t \leq T$ of the $(\mathcal{F}_t)$-intensity of τ given by

$$P(\tau > t|\mathcal{F}_t) = \exp(-\textstyle\int_0^t \gamma_u du).$$

The dynamics of $\gamma_t, 0 \leq t \leq T$ can be described by e.g. the Vasicek model, such that:

$$d\gamma_t = k(\theta - \gamma_t)dt + \sigma dL_t, \quad (7.4)$$

where k, θ, σ are constants and $L_t, 0 \leq t \leq T$ is a square integrable Lévy martingale with Lévy measure ν.

In order to estimate "time-dependent Lévy measures" of more realistic models for $\gamma_t, 0 \leq t \leq T$ one could assume that the compensator μ of the jump measure of $L_t, 0 \leq t \leq T$ can be parametrized by a compensator of the form (as in the structural approach):

$$\hat{\mu}(dt, d\zeta) = \lambda(t, X_t, \zeta)dt\nu(d\zeta).$$

We may also replace the Lévy process $L_t, 0 \leq t \leq T$ in equation (7.4) by a process L_t^λ given by:

$$L_t^\lambda = B_t + \int_0^t \int_{\mathbb{R}_0} \zeta(N_\lambda(ds, d\zeta) - \hat{\mu}(dt, d\zeta)), \quad 0 \leq t \leq T, \tag{7.5}$$

where N_λ is an integer-valued random measure with predictable compensator $\hat{\mu}$ and $B_t, 0 \leq t \leq T$ Brownian motion.

Chapter 8

Appendix

8.0.1 Appendix: chapter 2

Itôs formula applied of the Geometric Brownian motion
Let:

$$S_t = S_0 \exp(\mu t + \sigma B_t). \tag{8.1}$$

set:

$$f(t,x) = S_0 \exp(\mu t + \sigma x)$$

Having:
$\frac{\partial f(t,x)}{\partial t} = \mu f(t,x)$, $\frac{\partial f(t,x)}{\partial x} = \sigma f(t,x)$ and $\frac{\partial^2 f(t,x)}{\partial x^2} = \sigma^2 f(t,x)$.
By Itôs formula we get:

$$df(t,x) = \mu f(t,x)dt + \sigma f(t,x)dx + \tfrac{1}{2} f(t,x)(dx)^2$$

By letting $f(t,x) = f(t,B_t) = S_t$ and having that $(dB_t)^2 = dt$:

$$dS_t = (\mu + \frac{1}{2}\sigma^2)S_t dt + \sigma S_t dB_t \tag{8.2}$$

This gives:

$$S_t = S_0 + \int_0^t (\mu + \frac{1}{2}\sigma^2)S_u du + \int_0^t \sigma S_u dB_u. \tag{8.3}$$

8.0.2 Appendix: chapter 5

Put-Call Parity[Ben04]
The put-call parity is given by:

$$P_t^c - P_t^p = S_t - Ke^{-r(T-t)} \tag{8.4}$$

where P_t^c is the price of a call option and P_t^p is the price of a put option.

The prices of a call and put is, respectively,

$$\begin{aligned} P_t^c &= e^{-r(T-t)}\mathbf{E}_{\mathbb{Q}}[\max(0, S_T - K)] \\ P_t^p &= e^{-r(T-t)}\mathbf{E}_{\mathbb{Q}}[\max(0, K - S_T)]. \end{aligned}$$

Using the fact that:

$$\max(x - K, 0) = (x - K) + \max(K - x, 0)$$

we find:

$$\begin{aligned} P_t^c &= e^{-r(T-t)}\mathbf{E}_{\mathbb{Q}}[\max(0, S_T - K)|\mathcal{F}_t] \\ &= e^{-r(T-t)}\mathbf{E}_{\mathbb{Q}}[(S_T - K)|\mathcal{F}_t] + \mathbf{E}_{\mathbb{Q}}[\max(0, K - S_T)|\mathcal{F}_t] \\ &= e^{rt}\mathbf{E}_{\mathbb{Q}}[e^{-rT}S_T|\mathcal{F}_t] - e^{-r(T-t)}K + P_t^p \\ &= e^{rt}e^{-rt}S_t - e^{-r(T-t)}K + P_t^p \\ &= S_t - e^{-r(T-t)}K + P_t^p. \end{aligned}$$

Where the third equality follows from the martingale property of $e^{-rt}S_t$ w.r.t. $\mathbb{Q}$.

8.0.3 Appendix: chapter 6

Vasicek model

The dynamics of the stochastic signal process X_t is expressed by the Vasicek model and hence given by the stochastic differential equation (SDE):

$$dX_t = k(\theta - X_t)dt + \sigma dB_t \qquad , X_0 = x_0,$$

where θ is the long term mean, k is the speed of reversion, σ is the volatility and B_t is Brownian motion.

Remark: observe that in this case X_t can be negative.

The equation in (8.0.3) can be solved by using It$\hat{'}$os formula with

$f(t,x) = xe^{kt}$, where $f(t,\gamma_t) = \gamma_t e^{kt}$.

We obtain:

$$\begin{aligned} d(f(t,X_t) &= kX_t e^{kt}dt + e^{kt}dX_t \\ &= kX_t e^{kt}dt + e^{kt}(k(\theta - X_t)dt + \sigma dB_t) \\ &= kX_t e^{kt}dt + e^{kt}(k(\theta - X_t)dt + e^{kt}\sigma dB_t) \\ &= e^{kt}k\theta dt + e^{kt}\sigma dB_t. \end{aligned} \tag{8.5}$$

Integrating from 0 to t we get:

$$e^{kt}X_t = X_0 + k\theta \int_0^t e^{ks}ds + +\sigma \int_0^t e^{ks}dB_s. \tag{8.6}$$

Solving for X_t we obtain:

$$\begin{aligned} X_t &= e^{-kt}X_0 + k\theta \int_0^t e^{-k(t-s)}ds + \sigma \int_0^t e^{-k(t-s)}dB_s \\ &= e^{-kt}X_0 + e^{-kt}\theta k\Big[\frac{1}{k}e^{ks}\Big]_{s=0}^t + \sigma \int_0^t e^{-k(t-s)}dB_s \\ &= e^{-kt}X_0 + \theta(1 - e^{-kt}) + \sigma \int_0^t e^{-k(t-s)}dB_s. \end{aligned} \tag{8.7}$$

The mean of $X_t, 0 \leq t \leq T$ is given by:

$$\begin{aligned} \mathbb{E}[X_t|\mathcal{F}_0] &= \mathbb{E}[\underbrace{e^{-kt}X_0 + \theta(1 - e^{-kt})}_{deterministic} + \underbrace{\sigma \int_0^t e^{-k(t-s)}dB_s}_{Stochastic}|\mathcal{F}_0] \\ &= e^{-kt}X_0 + \theta(1 - e^{-kt}) + \mathbb{E}[\sigma \int_0^t e^{-k(t-s)}dB_s|\mathcal{F}_0] \end{aligned}$$

Having that: $\sigma \int_0^t e^{-k(t-s)} dB_s \sim \mathcal{N}(0, \left(\sigma \int_0^t e^{-k(t-s)} dB_s\right)^2)$, and using the fact that: $(dB_s)^2 = ds$ gives:

$\sigma \int_0^t e^{-k(t-s)} dB_s \sim \mathcal{N}(0, \frac{\sigma^2}{2k}(1-e^{-2kt}))$

The distribution of X_t is then given by:

$$X_t \sim \mathcal{N}(X_0 e^{-kt} + \theta(1-e^{-kt}), \tfrac{\sigma^2}{2k}(1-e^{-2kt})).$$

We look at at the case where t goes to infinity and get the stationary distribution:

$$\lim_{t \to \infty} X_t \sim \mathcal{N}(\theta, \frac{\sigma^2}{2k}). \tag{8.8}$$

As one can see the limited distribution is constant and independent of the time, t. Also notice that the process is mean reverting, where θ is the long term mean level.

Definition 8.1. ***Lipschitz continuity*** *[Lin14].*
A function $f : X \to Y$ between metric spaces is said to be Lipschitz continuous with Lipschitz constant K if $d_Y(f(x), f(y)) \leq K d_X(x, y)$. Where $d_X(x, y)$ (similarly $d_Y(f(x), f(y))$)denotes the metric on X, that is the distance between two points x and y in X.

8.1 R-code: Chapter 1

Listing 8.1: Simulation of Brownian motion

```
#Brownian motion
setwd("C:/Users/maja/Desktop/MO2015")

png("BM.png", width=15, height=15, units='cm', res=1500)

T = 10#maturity
n = 100 #steps from today to maturity
dt = 1/n #step
t = seq(0,T,dt)

dBt_1 = rnorm(length(t),0,1)*sqrt(dt)
dBt_2 = rnorm(length(t),0,1)*sqrt(dt)
dBt_3 = rnorm(length(t),0,1)*sqrt(dt)
Bt_1 = rep(0,length(t))
Bt_2 = rep(0,length(t))
Bt_3 = rep(0,length(t))

for (i in 1:(length(t)-1))
{
  Bt_1[i+1] =  Bt_1[i] + dBt_1[i]
  Bt_2[i+1] =  Bt_2[i] + dBt_2[i]
  Bt_3[i+1] =  Bt_3[i] + dBt_3[i]
}

plot(t,Bt_1 ,type = 'l', xlab = 'time', ylab = 'Brownian
   motion',xlim =c(0,max(t)) ,ylim=c(min(Bt_1,Bt_2,Bt_3),
   max(Bt_1,Bt_2,Bt_3)) )
lines(t,Bt_2, type = 'l', col = 'red')
lines(t,Bt_3, type = 'l', col = 'blue')
dev.off()
```

8.2 R-code: Chapter 2

Listing 8.2: Sample paths of Brownian motion and Lévy processes.

```
#Simulating Brownian motion and a Levy process
setwd("C:/Users/maja/Desktop/MO2015")

png("alphastabilevsbm.png", width=15, height=15, units='
   cm', res=1500)

n = 2000
dt = 1/n
t = seq(0,1,dt)

#Brownian motion:
dBt = rnorm(length(t),0,1)*sqrt(dt)
Bt = rep(0,length(t))

for (i in 1:(length(t)-1))
{
  Bt[i+1] =  Bt[i] + dBt[i]
}



#alpha-stable distributions
lambda = 0.7
#n independent ,random variables uniformly distributed on
    (-pi/2,pi/2)
gamma = runif(n,-pi/2,pi/2)

#n independent standard exponential random variables,
   with parameter lambda
W = rexp(n,lambda)

#set alpha-value
alpha = 1.9

#computing delta X:

delta_X = c()
delta_X[1] = 0
for (i in 2:n)
{
  delta_X[i] = ((t[i]-t[i-1])^(1/alpha))*(sin(alpha*gamma
      [i])/(cos(gamma[i])^(1/alpha)))*(cos((1-alpha)*gamma
      [i])/W[i])^((1-alpha)/alpha)
}

X= c()
X[1] = delta_X[1]
for (k in 2:n)
{
  X[k] = X[k-1] + delta_X[k]
}

plot(t[1:n], X,cex = 0.3,ylim = c(min(X,Bt),max(X,Bt)),
   xlab='time', ylab='sample paths')
lines(t,Bt ,type = 'l', col='grey')
```

```
dev.off()
```

Listing 8.3: α-stable processes.

```
#alpha-stable distributions
setwd("C:/Users/maja/Desktop/MO2015")

png("alphastabileproc.png", width=15, height=15, units='
    cm', res=1500)

n = 1000
dt = 1/n
lambda = 0.1
t = seq(0,1,dt)
#n independent ,random variables uniformly distributed on
     (-pi/2,pi/2)
gamma = runif(n,-pi/2,pi/2)

#n independent standard exponential random variables,
    with parameter lambda
W = rexp(n,lambda)

#set alpha-value
alpha = c(0.5,1,1.5,1.9)
delta_X = matrix(0,length(alpha),n)
X= matrix(0,length(alpha),n)
for  (j in 1:length(alpha))
{
#computing delta X:

for (i in 2:n)
{
  delta_X[j,i] = ((t[i]-t[i-1])^(1/alpha[j]))*(sin(alpha[
      j]*gamma[i])/(cos(gamma[i])^(1/alpha[j])))*(cos((1-
      alpha[j])*gamma[i])/W[i])^((1-alpha[j])/alpha[j])
}

X[j,1] = delta_X[j,1]
for (k in 2:n)
{
  X[j,k] = X[j,k-1] + delta_X[j,k]
}

}


par(mfrow=c(2,2))
plot(t[1:n], X[1,],cex = 0.3, ylab = 'X(t)', xlab='time',
    main='alpha = 0.5')
plot(t[1:n], X[2,],cex = 0.3, ylab = 'X(t)', xlab='time',
    main='alpha = 1.0')
plot(t[1:n], X[3,],cex = 0.3, ylab = 'X(t)', xlab='time',
    main='alpha = 1.5')
plot(t[1:n], X[4,],cex = 0.3, ylab = 'X(t)', xlab='time',
    main='alpha = 1.9')


dev.off()
```

8.3 R-code: Chapter 3

Listing 8.4: Illustration of the price of a ZCB.

```
#ZCB
setwd("C:/Users/maja/Desktop/MO2015")
png("ZCB.png", width=15, height=15, units='cm', res=1500)

#constant interest rate
r_constant = 0.03
time = seq(0,100)
price = exp(-r_constant*time)
plot(time,price,type = 'l',ylim = c(0,1),xlim = c(0,100),
    main = 'Price of ZCB',xlab='Time to maturity', ylab='
    Price')

dev.off()
```

8.4 R-code: Chapter 4

Listing 8.5: Simulation of the Merton model

```
#The Merton model.
setwd("C:/Users/maja/Desktop/MO2015")
png("merton.png", width=15, height=15, units='cm', res
   =1500)

mu_V = 0.02
gamma = 0.01
sigma_V= 0.09
dt = 1
#Time steps:
m = 100
#Simulations:
n = 1000
Vm = matrix(0,n,m+1)
t = seq(0,m-1,1)



for (j in 1:n)
{
  dBt = rnorm(length(t),0,1)*sqrt(dt)
  Bt = rep(0,length(t))
  V = c()
  V[1] = 100
  for (i in 1:m)
  {
    Bt[i+1] =  Bt[i] + dBt[i]
    dV = (mu_V - gamma)*V[i]*dt + sigma_V*V[i]*(Bt[i+1] -
        Bt[i])
    V[i+1] = V[i] + dV
    Vm[j,1] = V[1]
    Vm[j,i+1] = V[i+1]
  }

}
#Average:
av = c()
for (k in 1:m)
{
  av[k] = 1/(length(Vm[,k]))*sum(Vm[,k])
}

Vm = Vm[,-(m+1)]

plot(t,Vm[1,], type = 'l', col = 'blue', main = '
   Simuation of the Merton model with positive gamma',
   xlab = 'time',ylab = 'V_t',  ylim = c(0,500) )
lines(t,Vm[2,], type = 'l',col = 'red')
lines(t,Vm[3,], type = 'l', col = 'green')

dev.off()

#Using the same code with gamma = -0.01 in "Merton model
   with negative gamma".
```

Listing 8.6: Simulation of the Merton jump diffusion model

```
#Merton jump diffusion model.
setwd("C:/Users/maja/Desktop/MO2015")

png("mertonjumpdiff.png", width=15, height=15, units='cm'
    , res=1500)

#Set parameters:
mu_V = 0.02
gamma = 0.01
sigma_V= 0.09
dt = 1
#Time steps:
m = 100
#Simulations:
n = 100
Vm = matrix(0,n,m+1)
t = seq(0,m-1,1)

#Parameters in the jump component:
lambda_T = 0.2*m
mu = 0.3

for (j in 1:n)
{
  dBt = rnorm(m,0,1)*sqrt(dt)
  Bt = rep(0,m)
  Yt = rep(0,m)
  V = c()
  V[1] = 100
  N = c()
  xi = rep(0,m)
  dYt = rep(0,m)

    for (k in 1:m)
    {
    N[k] = rpois(1,lambda_T)
    xi[k] = rexp(length(N[k]),mu)
    if (N[k]>30)
    {
      dYt[k] = sum(xi[k])
    }
    }


  for (i in 1:m)
    {
    Yt[i+1] = Yt[i] + dYt[i]
    Bt[i+1] =  Bt[i] + dBt[i]
    dV = (mu_V - gamma)*V[i]*dt + sigma_V*V[i]*(Bt[i+1] -
        Bt[i]) + V[i]*(Yt[i+1]-Yt[i])
    V[i+1] = V[i] + dV
    Vm[j,1] = V[1]
    Vm[j,i+1] = V[i+1]
  }

}

Vm = Vm[,-(m+1)]

plot(t,Vm[1,], type = 'l', col = 'blue', main = '
   Simuation of the Merton jump diffusion model', xlab =
   'time',ylab = 'V_t', ylim = c(min(Vm[1,],Vm[2,],Vm
```

```
    [3,]),max(Vm[1,],Vm[2,],Vm[3,])))
lines(t,Vm[2,], type = 'l',col = 'red')
lines(t,Vm[3,], type = 'l', col = 'green')

dev.off()
```

8.5 R-code: Chapter 6

Listing 8.7: Simulation with the first λ

```
#simulation with the model with General Electrics Company
    data.
#Data from yahoo-finance:
data_GM = read.csv("C:/Users/maja/Desktop/MO2015/datasett
   /GE General Electric2006-2010.csv", sep=";")
Y = data_GM[,2] #Closing price.
#Simulations:
n= 100
#Timesteps:
T = length(Y) #Maturity
dt = 1 #steps from today to maturity
t = seq(0,(T-1),dt) #interval
Delta_Y = c()
Delta_Y[1] = Y[1]
for (i in 1:(length(Y)-1))
{
  Delta_Y[i+1] = log(Y[i+1])-log(Y[1])
}

#Brownian motion
Bt = matrix(0,length(t),n)
for (j in 1:n)
{
  dBt = rnorm(length(t),0,1)*sqrt(dt)
  for (i in 1:(length(t)-1))
  {
    Bt[i+1,j] =  Bt[i,j] + dBt[i]
  }
}

#Draw random numbers from Normal distr.(a priori distr.)
   for xi-values:
xi = matrix(0,4,n)
for (j in 1:4)
{
  xi[j,] =rnorm(n,0,1)
}

#X-Matix, creating an array:
X = array(0,dim = c(4,length(t),n))

#Start values:
for (j in 1:n)
{
  X[1,1,j] = xi[1,j]
  X[2,1,j] = xi[2,j]
  X[3,1,j] = xi[3,j]
  X[4,1,j] = xi[4,j]
}

for (j in 1:n)
{
  for (i in 1:(length(t)-1))
  {
    X[1,i+1,j] = X[1,i,j] + xi[2,j]*(xi[3,j]-X[1,i,j])*dt
        + xi[4,j]*rnorm(1,0,1)*sqrt(dt)
    X[2,i+1,j] = xi[2,j]
```

```
    X[3,i+1,j] = xi[3,j]
    X[4,i+1,j] = xi[4,j]
  }
}

#lambda function
epsilon = 1
lambda = function(t,x,y,z)
{
  epsilon*(((1/sqrt(2*pi*(y^2)))*exp(-(1/2)*((z-x)/y)^2))
     /((1/sqrt(2*pi*1))*exp(-(1/2)*(z)^2)))
}

#jump function
normz = function(z)
{
  ((1/sqrt(2*pi))*exp(-(1/2)*(z)^2))
}

#Projection:
f = function(x)
{
  x
}

#I_3:
jump = abs(mean(Delta_Y))
#Make matix:
I3 = matrix(0,length(t),n)
for (j in 1:n)
{
  for (i in 1:length(t))
    {
    if (abs(Delta_Y[i]) > jump)
    {
      I3[i,j] = log(lambda(i,X[1,i,j],X[4,i,j],Delta_Y[i
         ]))
    }
    else
    {
      I3[i,j] = 0
    }
  }
}

#Trapez method.
I_3 = matrix(0,length(t)-1,n)
for (j in 1:n)
{
  for (i in 1:(length(t)-1))
  {
    I_3[i,j] = (((i+1)-i)*((I3[i+1,j] + I3[i,j])/2))
  }
}

I_t3 = matrix(0,length(t)-1,n)
for (j in 1:n)
{
  for (i in 1:(length(t)-2))
  {
    I_t3[1,j] = I_3[1,j]
    I_t3[i+1,j] = I_t3[i,j] + I_3[i+1,j]
```

```
  }
}

#I_4=0
I_1  = Bt
I_2 = matrix(0,length(t),n)
for (i in 1:length(t))
{
  I_2[i,] = (1/2)*t[i]
}

#M_t:
M_t = matrix(0,length(t)-1,n)
for (j in 1:n)
{
  for (i in 1:(length(t)-1))
  {
    M_t[i,j] = exp(I_1[i+1,j] - I_2[i+1,j] +I_3[i,j])
  }
}

#Sum over diff. times:
expect_Mtf = matrix(0,4,length(t)-1)
for (j in 1:4)
{
  for (i in 1:(length(t)-1))
  {
    expect_Mtf[j,i] = (1/length(M_t[i,]))*sum(M_t[i,]*f(X
        [j,i,]))
  }
}

expect_Mt = matrix(0,4,length(t)-1)
for (j in 1:4)
{
  for (i in 1:(length(t)-1))
  {
    expect_Mt[j,i] = (1/length(M_t[i,]))*sum(M_t[i,])
  }
}

X_tilde = expect_Mtf/expect_Mt
time = seq(0,(T-2),dt)
#plot
plot(time,X_tilde[2,], type = 'l', col = 'blue', xlab='
   time', ylab='X_tilde',main= 'estimates', ylim = c(min(
   X_tilde[2,],X_tilde[3,],X_tilde[4,] ),max(X_tilde[2,],
   X_tilde[3,],X_tilde[4,])) )
lines(time,X_tilde[3,], type = 'l',col = 'red')
lines(time,X_tilde[4,], type = 'l', col = 'green')

#xi-values(last):
#xi_2=0.262893
#xi_3=-1.376566
#xi_4=0.9725912
```

Listing 8.8: Simulation with a different λ

```
#simulation with the model with General Electrics company
    data.
#Data from yahoo-finance:
data_GM = read.csv("C:/Users/maja/Desktop/MO2015/datasett
   /GE General Electric2006-2010.csv", sep=";")
Y = data_GM[,2] #Closing price.
#Simulations:
n= 100
#Timesteps:
T = length(Y) #Maturity
dt = 1 #steps from today to maturity
t = seq(0,(T-1),dt) #interval
#Plot of the closing price:
plot(t, Y, type = 'l', main = 'Closing price of GE', ylab
    = 'price', xlab = 'time')
Delta_Y = c()
Delta_Y[1] = Y[1]
for (i in 1:length(Y))
{
  Delta_Y[i] = log(Y[i])-log(Y[1])
}

#Plot of log-return:
plot(t, Delta_Y, type = 'l', main = 'log-return', ylab =
   'log-return', xlab = 'time')

#Brownian motion
Bt = matrix(0,length(t),n)
for (j in 1:n)
{
  dBt = rnorm(length(t),0,1)*sqrt(dt)
  for (i in 1:(length(t)-1))
  {
    Bt[i+1,j] =  Bt[i,j] + dBt[i]
  }
}

#Draw random numbers from Normal distr.(a priori distr.)
   for xi-values:
xi = matrix(0,4,n)
for (j in 1:4)
{
  xi[j,] =rnorm(n,0,0.1)
}

#X-Matix, creating an array:
X = array(0,dim = c(4,length(t),n))
#Startvalues;
for (j in 1:n)
{
  X[1,1,j] = xi[1,j]
  X[2,1,j] = xi[2,j]
  X[3,1,j] = xi[3,j]
  X[4,1,j] = xi[4,j]
}
for (j in 1:n)
{
  for (i in 1:(length(t)-1))
  {
    X[1,i+1,j] = X[1,i,j] + xi[2,j]*(xi[3,j]-X[1,i,j])*dt
```

```
        + xi[4,j]*rnorm(1,0,1)*sqrt(dt)
    X[2,i+1,j] = xi[2,j]
    X[3,i+1,j] = xi[3,j]
    X[4,i+1,j] = xi[4,j]
  }
}

#lambda function
lambda = function(t,x_1,x_2,x_3,x_4)
{
  (sqrt(x_1^2 + x_2^2 + x_3^2 +x_4^2) + 1)
}

#jump function
normz = function(z)
{
  ((1/sqrt(2*pi))*exp(-(1/2)*(z)^2))
}

#Projection:
f = function(x)
{
  x
}

#I_3:
jump = abs(mean(Delta_Y))

#Make matix:
I3 = matrix(0,length(t),n)
for (j in 1:n)
{
  for (i in 1:length(t))
  {
    if (abs(Delta_Y[i]) > jump)
    {
      I3[i,j] = log(lambda(i,X[1,i,j],X[2,i,j],X[3,i,j],X
          [4,i,j]))
    }
    else
    {
      I3[i,j] = 0
    }
  }
}

I4 = matrix(0,length(t),n)
for (j in 1:n)
{
  for (i in 1:length(t))
  {
    I4[i,j] = (1-lambda(i,X[1,i,j],X[2,i,j],X[3,i,j],X[4,
        i,j]))
  }
}

#Trapez method.
I_3 = matrix(0,length(t)-1,n)
I_4 = matrix(0,length(t)-1,n)
for (j in 1:n)
{
  for (i in 1:(length(t)-1))
  {
```

```
    I_3[i,j] = (((i+1)-i)*((I3[i+1,j] + I3[i,j])/2))
    I_4[i,j] = (((i+1)-i)*((I4[i+1,j] + I4[i,j])/2))
  }
}
I_t3 = matrix(0,length(t)-1,n)
I_t4 = matrix(0,length(t)-1,n)
for (j in 1:n)
{
  for (i in 1:(length(t)-2))
  {
    I_t3[1,j] = I_3[1,j]
    I_t3[i+1,j] = I_t3[i,j] + I_3[i+1,j]
    I_t4[1,j] = I_4[1,j]
    I_t4[i+1,j] = I_t4[i,j] + I_4[i+1,j]
  }
}

I_1  = Bt
I_2 = matrix(0,length(t),n)
for (i in 1:length(t))
{
  I_2[i,] = (1/2)*t[i]
}

#M_t:
M_t = matrix(0,length(t)-1,n)
for (j in 1:n)
{
  for (i in 1:(length(t)-1))
  {
    M_t[i,j] = exp(I_1[i+1,j] - I_2[i+1,j] + I_3[i,j] + I
       _4[i,j])
  }
}

#Sum over diff. times:
expect_Mtf = matrix(0,4,length(t)-1)
for (j in 1:4)
{
  for (i in 1:(length(t)-1))
  {
    expect_Mtf[j,i] = (1/length(M_t[i,]))*sum(M_t[i,]*f(X
       [j,i,]))
  }
}

expect_Mt = matrix(0,4,length(t)-1)
for (j in 1:4)
{
  for (i in 1:(length(t)-1))
  {
    expect_Mt[j,i] = (1/length(M_t[i,]))*sum(M_t[i,])
  }
}

X_tilde = expect_Mtf/expect_Mt
time = seq(0,(T-2),dt)
#plot:
plot(time,X_tilde[2,], type = 'l', col = 'blue', xlab='
   time', ylab='X_tilde',main= 'estimates', ylim = c
   (-0.3,0.4))
```

```
lines(time,X_tilde[3,], type = 'l',col = 'red')
lines(time,X_tilde[4,], type = 'l', col = 'green')

#xi-values (converges):
xi_2 = X_tilde[2,T-1] #=0.01440555
xi_3 = X_tilde[3,T-1] #=-0.01900694
xi_4 = X_tilde[4,T-1] #=0.05796884
#Data from January 2011 to December 2011:
data_GM2011 = read.csv("C:/Users/maja/Desktop/M02015/
   datasett/GE General Electric2011.csv", sep=";")
obs_2011 = data_GM2011[,2]
T = length(obs_2011)
#Startvalue:
xi_1 =rnorm(1,0,0.1)
#timestep:
m=T
#paths:
n = 10
xm = matrix(0,n,T+1)
#Signal process
for (j in 1:n)
{
  dBt = rnorm(length(t),0,1)*sqrt(dt)
  Bt = rep(0,length(t))
  x = c()
  x[1] = xi_1
  for (i in 1:T)
  {
    Bt[i+1] =  Bt[i] + dBt[i]
    dx = xi_2*(xi_3-x[i])*dt + xi_4*(Bt[i+1] - Bt[i])
    x[i+1] = x[i] + dx
    xm[j,1] = x[1]
    xm[j,i+1] = x[i+1]
  }

}
#average:
av = c()
for (k in 1:T)
{
  av[k] = 1/(length(xm[,k]))*sum(xm[,k])
}

xm = xm[,-(T+1)]

#plot signal process
time_1 =seq(0,T-1,1)
plot(time_1,xm[1,], type = 'l',xlab = 'time', ylab = '
   Signal process', main = 'Simuation of the Signal
   process',ylim = c(-1,1))
for (i in 2:10)
{
  lines(time_1,xm[i,], type = 'l')
}
#and the average:
lines(time_1,av, type = 'l', col = 'red',  lwd= 2)

mu = matrix(0,n,T)
for(j in 1:T)
{
  for (i in 1:n)
```

```
    {
      mu[i,j] = lambda(j,xm[i,j],xi_2,xi_3,xi_4)
    }
}

#Trapez method.
mu_1 = matrix(0,n,T-1)
for (j in 1:(T-1))
{
  for (i in 1:n)
  {
    mu_1[i,j] = (((i+1)-i)*((mu[i,j+1] + mu[i,j])/2))
  }
}
mu1 = matrix(0,n,T-1)
for (j in 1:(T-2))
{
  for (i in 1:n)
  {
    mu1[i,1] = mu_1[i,1]
    mu1[i,j+1] = mu1[i,j] + mu_1[i,j+1]
  }
}
#Step 3.:
L = matrix(0,n,T-1)
for (i in 1:n)
{
  for (j in 1:(T-1))
  {
    L[i,j] = sum(runif(rpois(1,mu1[i,j]),-0.01,0.01))
  }

}
S_0 = data_GM[length(data_GM[,2]),2]
time = seq(0,T-2,1)
dataGM2011 = data_GM2011[,2]
dataGM2011 = dataGM2011[-(length(dataGM2011))]
#Plot:
plot(time,S_0*exp(L[1,]), type = 'l', ylab = 'S_t', ylim
    = c(12, 25) )
for (i in 2:10)
{
  lines(time,S_0*exp(L[i,]), type = 'l')
}
lines(time, dataGM2011, type = 'l', col = 'red', lwd = 2)

#To compare the models:
plot(time,S_0*exp(L[1,]), type = 'l', ylab = 'S_t', ylim
    = c(5, 40), main='Merton vs. the new model' )
lines(time,Vt[,1],type = 'l', col = 'blue')
lines(time, dataGM2011, type = 'l', col = 'red', lwd = 2)
```

Listing 8.9: Simulation with the Merton model

```
#GE data from Jan 2006 to Dec 2010
data_GM = read.csv("C:/Users/maja/Desktop/MO2015/datasett
   /GE General Electric2006-2010.csv", sep=";")
Y = data_GM[,2] #Closing price.
#Timesteps:
N = length(Y) #Maturity
dt = 1 #steps from today to maturity
Delta_Y = c()
Delta_Y[1] = Y[1]
#log-return
for (i in 1:(length(Y)-1))
{
  Delta_Y[i+1] = log(Y[i+1]/Y[i])
}
N= length(Delta_Y)
delta_t = 24
#MLE estimates
alpha_hat = (1/(delta_t*N))*sum(Delta_Y)
var_hat = c()
for (i in 1:(N-1))
{
  var_hat[i] = (Delta_Y[i]-delta_t*alpha_hat)^2
}

sigma_hat = sqrt(1/(delta_t*N-1)*sum(var_hat))

#GE data from Jan 2011 to Dec 2011
data_GM2011 = read.csv("C:/Users/maja/Desktop/MO2015/
   datasett/GE General Electric2011.csv", sep=";")
obs_2011 = data_GM2011[,2]
T = length(obs_2011)
Vt = c()
dBt = rnorm(T,0,1)*sqrt(dt)
Bt = rep(0,T)
Vt[1] =data_GM[length(data_GM[,2]),2]
V_0 = 18.29 #last value of december 2010
Vt[1] = V_0
Xt = matrix(0,T-1,10)
for (j in 1:10)
{
for (i in 1:(T-1))
{
  Xt[i,j] = alpha_hat*i + sigma_hat*rnorm(1,0,1)
}
}

Vt = V_0*exp(Xt)
t = seq(0,(T-1),1) #time interval
plot(t,Vt[,1], type='l', ylab = 'price', xlab = 'time',
   main = 'Simulation with the Merton model' )
for (j in 2:10)
{
  lines(t,Vt[,j], type = 'l')
}
lines(t,obs_2011, type = 'l', col = 'red', lwd=2)
```

Bibliography

[App09] D. Applebaum. *Lévy Processes and Stochastic Calculus.* Cambridge University Press, 2009.

[BDP15] E. Bølviken, S. Duedahl, and F. Proske. *Modeling and estimation of stochastic transition rates in life insurance with regime switching based on generalized Cox processes. Preprint UiO.* 2015.

[Ben04] F. E. Benth. *Option theory with stochastic analysis.* Springer-Verlag, Berlin, 2004.

[BM01] D. Brigo and F. Mercurio. *Interest rate models-theory and practice.* Springer-Verlag, Berlin, 2001.

[BR02] T.R. Bielecki and M. Rutkowski. *Credit Risk: Modeling, Valuation and Hedging.* Springer, 2002.

[CT04] R. Cont and P. Tankov. *Financial modelling with jump processes.* Chapman & Hall/CRC, Boca Raton, FL, 2004.

[Kie08] R. Kiesel. Credit modelling and deivatives, cma intensive course, september 15-18,2008.

[Lin14] T. Lindstrøm. Lecture notes from the course: Mat2400 at uio, mathematical analysis, 2014.

[MBP03] T. Meyer-Brandis and F. Proske. *Explicit Solution and a Nonlinear Filtering Problem for Lévy Processes with Applications to Finance.* 2003.

[Mer75] R.C. Merton. Option pricing when underlying stock returns are discontinuous, journal of financial economics, July, 1975.

[MPMB09] V. Mandrekar, F. Proske, and T. Meyer-Brandis. *A bayes formula for non-linear fitering with Gaussian and Cox noise.* 2009.

[Øks95] B. Øksendal. *Stochastic Differential Equations, fourth edition.* Springer, 1995.

[Sch03] P. Schönbucher. *Credit Derivatives Pricing Models.* Wiley Finance, 2003.

[WHD99] P. Willmott, S. Howison, and J. Dewynne. *The Mathematics of Financial Derivatives.* Cambridge University press, 1999.