Andreas SCHIESTL

Modeling credit risk: A comparison of firm value and intensity based models

DIPLOMARBEIT

zur Erlangung des akademischen Grades eines
Diplom-Ingenieurs

Diplomstudium Operations Research, Statistik, Finanz- und Versicherungsmathematik

Graz University of Technology

Technische Universität Graz

Betreuer/in:
Ao.Univ.-Prof. Dipl.-Ing. Dr.techn. Eranda Dragoti-Cela
Institut für Optimierung und Diskrete Mathematik

Dipl.-Ing. Dr.techn. Philipp Mayer
Institut für Analysis und Computational Number Theory

Graz, im Mai 2010

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz, am ....

(Unterschrift)

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitely marked all material which has been quotes either literally or by content from the used sources.

....

date (signature)

Vorwort

Da der Markt für die unterschiedlichsten Kreditinstrumente in den letzten Jahren kontinuierlich gewachsen ist, spielt die Bestimmung des Ausfallrisikos bzw. des Kreditrisikos eine immer tragendere Rolle. Vor allem ausgeklügelte quantitative Methoden sollen hierbei helfen, dieses finanzielle Risiko richtig zu bemessen. Schlussendlich haben uns auch die Begebenheiten am Finanzmarkt im Jahre 2007 nochmals deutlich vor Augen geführt, wie wichtig die Einbindung des Kreditrisikos ins Risikomanagement ist.

Preface

Since the market of credit instruments has been growing continuously over the last years, quantitative modeling of default risk, or more generally credit risk, is getting more and more important. Sophisticated quantitative methodologies help to measure and manage this financial risk. Finally the financial crisis of 2007 showed, that credit risk plays an essential role in risk management.

Acknowledgement

Firstly I would like to thank my advisors at University, Ao.Univ.-Prof. Dipl.-Ing. Dr.techn. Eranda Dragoti-Cela and Dipl.-Ing. Dr.techn. Philipp Mayer, for their expert support and providing information at any time. Special thanks is due to Dipl.-Ing. Dr.techn. Philipp Mayer, who corrected my diploma thesis several times and helped me a lot to improve this text.

Furthermore I owe thanks to the Raiffeisen Landesbank Steiermark for giving me the opportunity to write this diploma thesis, especially to Claire Hemery for her support particularly in practical issues.

I would also like to thank my girlfriend Martina, who always supported and encouraged me while writing this diploma thesis. She stood by me all the time, despite the big effort of time I had to put in this work.

Last but not least I would like to use this occasion to thank my parents for all their support during my studies and for giving me the opportunity to achieve my aims. Thanks for all the patience and support by words and deeds in every situation of my life.

Contents

Abstract

The main subject of this diploma thesis is the estimation of the default probability of a single firm. A variety of models have been derived to model this default risk in a quantitative way, but basically two main approaches have proved to be particularly useful and well suited: so called structural models on the one hand and intensity models on the other. This thesis will focus on the most popular representatives of these approaches and compare them regarding feasibility and their ability to match real market data.

In the first chapter of my diploma thesis we will first point out the most important definitions and some preliminary considerations concerning default risk. Afterwards we will derive the technical setup of the different approaches, namely the Itô integral, the Black-Scholes formula, first passage times and the Poisson process and variants thereof.

In Chapter 2 we will focus on structural models, more precisely the Merton model and the Black and Cox model. These models are also known as firm value models and especially the Merton model used to be and still is quite popular in practice, mainly because of its simplicity. An illustration of the Merton model is given in Figure 1.

In Chapter 3 we will focus on the second approach, the intensity models. We will start with a simple deterministic model and then we will introduce stochastic default intensities on the one hand, and stochastic interest rates on the other hand. Finally we will discuss the SSRD model, which allows for dependence between interest rate and default intensities. An illustration of the intensity setup is given by Figure 2.

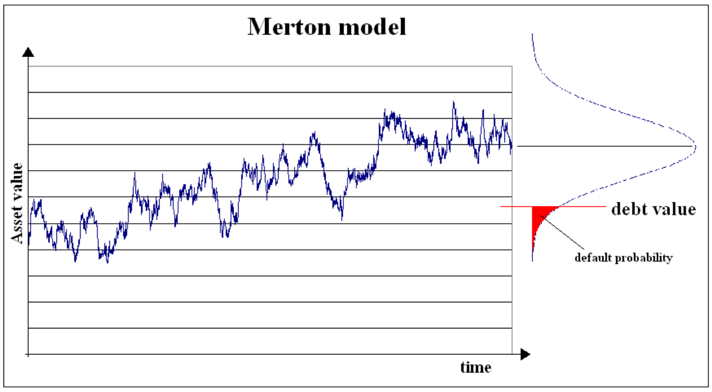

Figure 1: Merton model

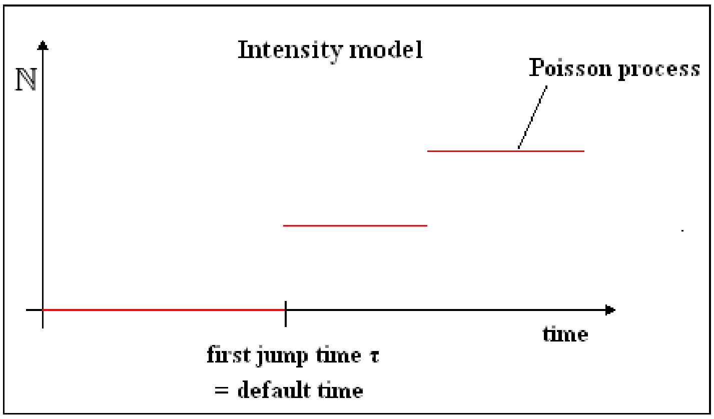

Figure 2: Intensity model

As already mentioned, in my diploma thesis I want to study the practical feasibility of the different models, therefore the last section will focus on the application of the discussed models. We will discuss two case studies with real market data and compare the obtained results.

Chapter 1

Introduction

1.1 Preliminary considerations

First we will settle the most important definitions and basic information about credit risk and the instruments discussed later on. For further information I refer to [2], on which the framework of this section is based.
One main object of this diploma thesis will be to calculate the price of a so-called defaultable corporate bond. But before we can go into details, the event of default and the main instrument, the corporate bond, need to be defined.
Since the corporate bond is a special type of bond, we will start with the definition of the most simple version of a bond, the default-free zero coupon bond.

Definition 1.1.1.: Zero coupon bond

The zero coupon bond pays the face value 1 at maturity. Since there are no periodic payments, the price of the zero coupon bond is its discounted face value. In the sequel the price of a zero coupon bond will be denoted by $\mathrm{B}(t, T)$.

Definition 1.1.2.: Coupon bond

The coupon bond pays at predefined dates $T_1, T_2, ..., T_n$ a predefined payment, the coupon, and at maturity the face value 1, and a coupon. Its price is the overall discounted value of the payments. Therefore a coupon bond can be seen as a sum of zero coupon bonds.

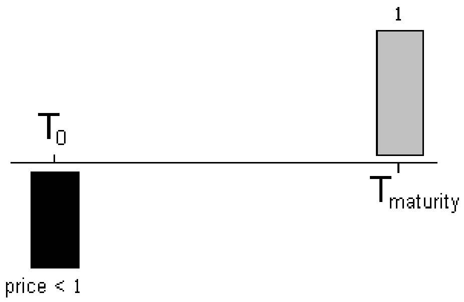

Figure 1.1: Cashflow of a zero coupon bond

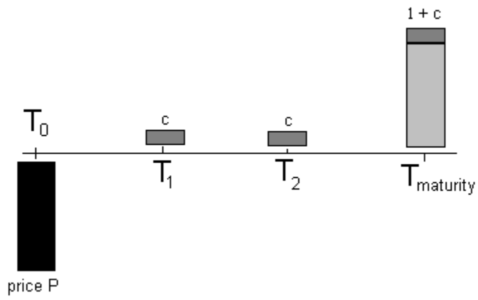

Figure 1.2: Cashflow of a coupon bond

Definition 1.1.3.: Corporate bond

A corporate bond is a bond issued by a corporate. By issuing the bond, the corporate commits to make specific payments on future dates. In return the corporate charges a fee.

Definition 1.1.4.: Default

A counterparty defaults, when it can not fulfill a contractual commitment to meet its obligations.

Given this definition, default risk is the financial risk, that future payments will not be paid in full amount. Credit risk is the financial risk that occurs due to any change in credit quality. For example downgrading of a firms credit rating, which causes a loss in the value of the bond, issued by this firm. So default risk is a special instance of credit risk.

Remark: When the issuer of a corporate bond defaults, the bondholder will not receive the promised payments (in full amount). Therefore, in this debt instrument, default risk has to be taken into account.

Definition 1.1.5.: Default time

The default time τ of a corporate is the first point in time, the corporate can not fulfill its obligations.

Definition 1.1.6.: Recovery rules

The recovery rules specify the timing and the amount of the payments, the corporate has to/is able to pay in case of default before or at maturity.

Remark: Normally the recovery scheme is assumed to be a fixed fraction δ of the future payments, called the recovery rate.

Given these first definitions one can already specify the payoff $\mathrm{D}^{\delta}(T,T)$ at time T of a zero coupon corporate bond, assuming that the face value equals 1 and the recovery rate δ is paid at maturity T in case of default:

$$\mathrm{D}^{\delta}(T,T) = \mathbb{I}_{\{\tau > T\}} + \delta \mathbb{I}_{\{\tau \leq T\}}, \tag{1.1}$$

where T is the maturity of the zero coupon bond and τ is the default time.
In my further explanations $\mathrm{D}(t,T)$ will denote the price of a defaultable bond at time t with maturity T.

With given (riskless) zero coupon bond prices $\mathrm{B}(t,T)$, the price of the zero coupon corporate bond is given by

$$\begin{aligned} \mathrm{D}^{\delta}(0,T) &= \mathbb{E}\Big(\mathrm{B}(0,T)\mathbb{I}_{\{\tau > T\}} + \delta\,\mathrm{B}(0,T)\mathbb{I}_{\{\tau \leq T\}}\Big) \\ &= \mathrm{B}(0,T)\mathbb{P}(\tau > T) + \delta\,\mathrm{B}(0,T)\mathbb{P}(\tau \leq T) \\ &= \mathrm{B}(0,T) - \mathrm{LGD}\,\mathbb{P}(\tau \leq T) \end{aligned}$$

with LGD (=Loss Given Default) $= (1-\delta)\,\mathrm{B}(0,T)$.

Another credit instrument, we will need in the sequel, is the Credit Default Swap, or short CDS. This contract can be seen as a default insurance.

Definition 1.1.7.: Credit default swap

In this agreement periodic payments or an upfront fee from the buyer of the CDS is exchanged for a promised payment (from the seller), if a pre-specified credit event occurs.

Usually, the defined credit event is default, since then the CDS contract can be used to eliminate the risk, that promised payments are not paid. But, of course, other definitions are possible, for example a change in the firms credit rating.

Remark: Normally there are three counterparties engaged in the CDS. The seller of the CDS , let us call it “A”, the buyer of the CDS, “B”, and the credit event is usually a default of a third party, “C”. The CDS is the protection of “B”, when “C” defaults and can not meet its obligations towards “B”. So “B” avoids his financial risk, when “C” defaults, by paying a contractual fixed fee to “A”. For better understanding, the cashflows are given in figure 1.3.

Typically the nominal of the CDS is set accordingly to the nominal of the underlying bond and the insured sum is the expected LGD.

Definition 1.1.8.: CDS - spread

The spread of a CDS is the amount/fee, in percent of the nominal, that the buyer has to pay to the seller at predefined dates $T_1, T_2, ...$ until maturity or default of “C”.

Since CDS-spreads are quoted on the market, they reflect the markets opinion on the default probability of the firm “C”. Hence the market CDS-spreads can be used to calibrate credit risk models.

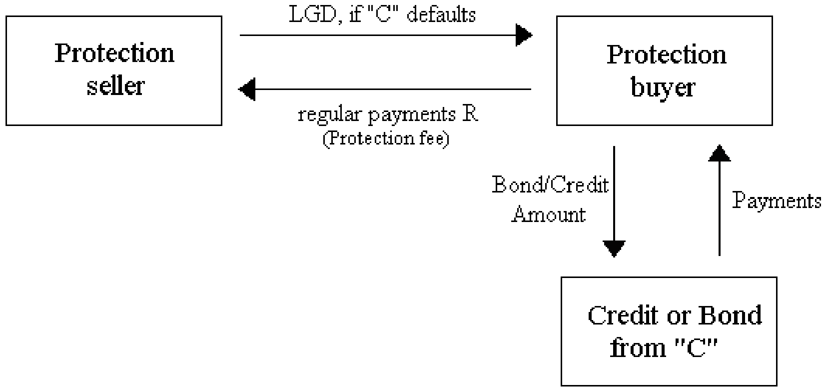

Figure 1.3: CDS - cashflow with associated bond

1.2 Overview of the main approaches of modeling credit risk

In this section I will give a short review of the main quantitative approaches, that can be used to model the default risk. It is important to point out, that most of the models can also be used to model credit risk and not only default risk, nevertheless I will concentrate on the default event.
In principle there are two main groups of models. Structural models and reduced form or intensity models.
In structural models the value of a firm and its capital structure is modeled by some stochastic process and the default event is triggered, when this process falls below some specific barrier. These models are quite popular, since there is a direct link between the firm value and the default event. Due to this link one can establish a connection between the observable market data, as e.g. the stock price of a firm, and the credit risk model. Furthermore the input parameters have an economic interpretation. My explanations concerning structural models are mostly based on [2], [3] and [19]. For further readings see for instance [1], [4] and [18].
As mentioned above, the second approach are reduced form models. In these models the capital structure is not modeled at all. Instead the credit event is described directly without using the firms value: default is triggered by an exogenously given jump process. We can distinguish between approaches that model the default time (= intensity models) and those modeling the migration between rating classes (=credit migration models). For intensity models we will follow in large parts [6] and [11], respectively. Also [1], [18] are a good account on intensity models and credit risk in general.
Of course not all proposed models can be classified in this way. For example, recently some hybrid models (see e.g. [1], [13]) were introduced, that try to combine the advantages of both approaches. However in this diploma thesis the focus will lie on structural models, or more precisely, on the Merton model and the Black and Cox model on the one hand, and on intensity models on the other hand.

1.3 Technical setup

In this section I review some of the main building blocks for the following models. For structural models the most important tool is the Brownian motion and some related concepts, as the Itô-formula or first passage times, which will be considered in some detail.
On the other hand, for the intensity models, the major ingredient is a Poisson process, either with constant, time varying or stochastic intensity. We will discuss this kind of processes at the end of this chapter.

1.3.1 Brownian motion

Let $\mathrm{N}(\mu, \sigma^2)$ denote the normal distribution with expected value μ and variance σ^2. Then the Brownian motion is defined by:

Definition 1.3.1.: Brownian motion

A Brownian motion $(B_t)_{t \geq 0}$ is a continuous stochastic process defined on a probability space $(\Omega, \mathbb{A}, \mathbb{P})$, with the following properties:

- $B_0 = 0$ almost surely.
- All paths of B are almost surely continuous.
- B has independent increments with distribution $B_t - B_s \sim N(\mu, \sigma^2(t-s))$ for $0 \leq s \leq t$.

Remark: A path of the Brownian motion B is given by the function $\omega \to B_t(\omega)$, for a fixed $\omega \in \Omega$.

A Brownian motion with $\mu = 0$ and $\sigma = 1$ is called standard Brownian motion or Wiener process and is often denoted by $(W_t)_{t \geq 0}$.

Definition 1.3.2.: Geometric Brownian motion

A stochastic process $(S_t)_{t\geq 0}$ is called a geometric Brownian motion, if

$$S_t = S_0 \cdot \exp((\mu - \frac{\sigma^2}{2})t + \sigma W_t).$$

Definition 1.3.3.: Filtration

A filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ is a nondecreasing family of sub - σ - algebras.

Remark: In this chapter we will use the so called natural filtration of a Brownian motion $\mathbb{F} = \sigma(\{B_s | s \leq t\} \cup \mathcal{N})$, with $\mathcal{N}$ being the empty sets. It reflects the information process up to time t.

Definition 1.3.4.: Adapted process

A stochastic process X is called adapted to $\mathbb{F}$, if the random variable X_t is $\mathcal{F}_t$-measurable for every t.

Remark: We say, B_t is a Brownian motion with respect to a Filtration $\mathbb{F}$, if B_t is adapted to $\mathbb{F}$ and the increments $B_t - B_s$ are independent of $\mathcal{F}_s$ for all $0 \leq s \leq t$.

Definition 1.3.5.: Martingale / local martingale

Let $(\Omega, \mathbb{A}, \mathbb{P})$ be a probability space and $\mathbb{F}$ a filtration on this probability space. A $\mathbb{F}$ - adapted family of random values $(M_t)_{t\geq 0}$, is called:

- Martingale, if $\mathbb{E}(|M_t|) < \infty$ and $\mathbb{E}(M_t | \mathcal{F}_s) = M_s$ for all $s \leq t$.
- Local martingale, if there exist an series of increasing, $\mathcal{F}_t$ - measurable stopping times $\tau_k : \Omega \to [0, \infty)$, such that the stopped process $M_t^{\tau_k} = M_{\min(t,\tau_k)}$ is a martingale.

Remark: When W_t is a standard Brownian motion with respect to $\mathbb{F}$, then W_t is a martingale. Furthermore, the processes $M_t = W_t^2 - t$ and $Y_t = \exp(\theta W_t + \frac{\theta^2}{2} t)$ are martingales.

Definition 1.3.6.: Quadratic Variation

Let X_t be a stochastic process. If there exists a random variable $[X, X]_t$ for every t, such that

$$\sum_{k=1}^{n} (X_{t_k} - X_{t_{k-1}})^2 \xrightarrow{\mathbb{P}} [X, X]_t$$

when the fineness of the decomposition $0 = t_0 < t_1 < ... < t_n = t$ goes to zero, then $[X, X]_t$ is a stochastic process and it is called the quadratic variation of X at time t.

For the quadratic variation of a standard Brownian motion we have the following theorem.

Theorem 1.3.7: ([16])

The quadratic variation of a standard Brownian motion is given as $[W, W]_t = t$.

Corollary 1.3.8: ([16])

The paths of a Brownian motion have infinite variation.

1.3.2 The Itô Formula

For the application of the Brownian motion to financial modeling we need a concept for an integral with respect to the Brownian motion. As the paths of the Brownian motion are very rough and in particular not of finite variation, the integral can not be defined in a pathwise manner using for example the Lebesgue-Stieltjes integral. Itô [14] solved this problem and it turns out, that the right way to define an integral with respect to the Brownian motion is the following.

Definition 1.3.9.: Itô Integral defined by Riemann-Stieltjes-sums

Let $(Y_t)_{t\geq 0}$ be a quadratic integrable ($\int_0^\infty |Y_s|^2 \, ds < \infty$), measurable and adapted process, $(B_t)_{t\geq 0}$ a Brownian motion and $0 = t_1 < t_2 < \cdots < t_n = t$ a partition of $[0, t]$, then the Itô-integral is defined as:

$$\int_0^t Y_s \, dBs = \lim \sum_{i=1}^{n} Y_{t_{i-1}} (B_{t_i} - B_{t_{i-1}})$$

when the fineness of the decomposition $0 = t_0 < t_1 < ... < t_n = t$ goes to zero. This definition is well defined, which means that the limes does not depend on the decomposition, as long as its fineness goes to zero.

Remark: $X_t = Y_t \quad \mathbb{P} - a.s.$ means, that $\mathbb{P}(X_t \neq Y_t) = 0 \quad \forall t$. We will use this notation in the sequel.

Definition 1.3.10.: Itô process

A real-valued, adapted process X is called Itô-process, if it can be represented by:

$$X_t = X_0 + \int_0^t K_s\, ds + \int_0^t H_s\, dBs, \tag{1.2}$$

which is often denoted by

$$dX_t = K_t dt + H_t dBt,$$

with

- X_0 is a $\mathcal{F}_0$-measurable random variable
- $K = (K_t)_{0 \leq t \leq T}$ is adapted and $\int_0^T |K_s|\, ds < \infty$ $\mathbb{P}$ - a.s.
- $H = (H_t)_{0 \leq t \leq T}$ is measurable, adapted and $\int_0^T {H_s}^2\, ds < \infty$ $\mathbb{P}$ - a.s.

Remark:

- This representation is well-defined. This means if $X_t = X_0' + \int_0^t K_s'\, ds + \int_0^t H_s'\, dBs$, then $X_0 = X_0'$, $K = K'$ and $H = H'$, $\lambda \times \mathbb{P}$-almost surely, with λ being the Lebesgue measure.
- If X is a local martingale $\Rightarrow K = 0$ $\lambda \times \mathbb{P}$-almost surely.

Definition 1.3.11.: Integral with respect to an Itô process

Let X be an Itô process with representation (1.2) and Y an adapted process. Then we define:

$$\int_0^t Y_s\, dX_s := \int_0^t Y_s K_s\, ds + \int_0^t Y_s H_s\, dB_s$$

whenever the right-hand side is defined.

Theorem 1.3.12: If Y has continuous paths and $\int_0^t Y_s \, dX_s$ exists, then

$$\sum_{i=1}^{n} Y_{t_{i-1}}(X_{t_i} - X_{t_{i-1}}) \xrightarrow{\mathbb{P}} \int_0^t Y_s \, dX_s$$

and

$$\sum_{j=1}^{n} Y_{t_{j-1}}(X_{t_j} - X_{t_{j-1}})^2 \xrightarrow{\mathbb{P}} \int_0^t Y_s {H_s}^2 \, ds$$

The sums on the left are called “Riemann-Stieltjes sums”.

Proof: For a proof see for instance [16].

Remark: This theorem in particular shows that the definition of the integral with respect to the Brownian motion (Definition 1.3.9) is consistent with the definition of integrals with respect to Itô-Processes. Furthermore, we obtain that, if X is an Itô process with representation (1.2), then $[X, X]_t = \int_0^t {H_s}^2 \, ds$.

Theorem 1.3.13: Itô-formula

If f: $\mathbb{R} \to \mathbb{R}$ is two times continuously differentiable and X is an Itô process, then the process $f(X) = (f(X_t))_{0 \leq t \leq T}$ is an Itô process and

$$f(X_t) = f(X_0) + \int_0^t f'(X_s) \, dX_s + \frac{1}{2} \int_0^t f''(X_s) \, d[X, X]_s. \tag{1.3}$$

Remark: (1.3) is often written in the differential notation

$$df(X_t) = f'(X_t) dX_t + \frac{1}{2} f''(X_t) d[X, X]_t. \tag{1.4}$$

This explains, why the Itô formula is often called the “chain-rule” of stochastic calculus.

For the proof we first need to define a stopping time.

Definition 1.3.14.: Stopping time

Let $(\Omega, \mathbb{A}, \mathbb{P})$ be a probability space and $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ a filtration defined on $(\Omega, \mathbb{A}, \mathbb{P})$. Then the random time τ is a stopping time, if $\{\tau \leq t\} \in \mathcal{F}_t \; \forall t$.

Proof:(Sketch)
We introduce a stopping time T_n, that is defined as follows:

$$T_n = \begin{cases} 0, & \text{if } |X_0| \geq n, \\ \inf(t \geq 0 : |X_t| \geq n), & \text{if } |X_0| < n, \\ \infty, & \text{if } |X_0| < n \text{ and } \{t \geq 0 : |X_t| \geq n\} = \emptyset \end{cases}$$

The resulting sequence is then nondecreasing with $\lim T_n = \infty$. So we can prove the Itô theorem for the stopped processes $X_{t \wedge T_n}$, and we get the desired result for $n \to \infty$. We may assume, therefore, that $X_0(\omega)$ and the random function $X_t(\omega)$ is bounded on $[0, \infty)$ by some constant C. Then f, f' and f'' are also bounded.

Let us have a look on the Taylor decomposition of $f(X)$.
For a given decomposition Π, $0 = t_0 < t_1 < ... < t_n = t$, of $[0, t]$ we get:

$$f(X_{t_j}) - f(X_{t_{j-1}}) = f'(X_{t_{j-1}})(X_{t_j} - X_{t_{j-1}}) + \frac{1}{2} f''(\eta_j)(X_{t_j} - X_{t_{j-1}})^2$$

with $\eta_j(\omega) = X_{t_{j-1}}(\omega) + \theta_k(\omega)(X_{t_j}(\omega) - X_{t_{j-1}}(\omega))$ for some appropriate $\theta_k(\omega)$ satisfying $0 < \theta_k(\omega) < 1$ and $\omega \in \Omega$.

Summing up over $1 \leq j \leq n$ we obtain

$$\begin{aligned} f(X_{t_n}) - f(X_0) = & \underbrace{\sum_{j=1}^{n} f'(X_{t_{j-1}})(X_{t_j} - X_{t_{j-1}})}_{\xrightarrow{\mathbb{P}} \int_0^t f'(X_s)\, dX_s} \\ & + \frac{1}{2} \sum_{j=1}^{n} f''(\eta_j)(X_{t_{j-1}} - X_{t_j})^2. \end{aligned}$$

For the first term, we can use Theorem 1.3.12. For the second one, we want to show, that

$$\frac{1}{2} \sum_{j=1}^{n} (f''(X_{t_{j-1}}) - f''(\eta_j))(X_{t_{j-1}} - X_{t_j})^2 \xrightarrow{\mathbb{P}} 0.$$

Let us define $[X,X]_{t,\Pi}$ being the quadratic variation over the decomposition Π. This means, that

$$[X,X]_{t,\Pi} = \sum_{t_k \in \Pi} (X_{t_k} - X_{t_{k-1}})^2.$$

Then we obtain, that

$$\frac{1}{2}\sum_{j=1}^{n}(f''(X_{t_{j-1}}) - f''(\eta_j))(X_{t_{j-1}} - X_{t_j})^2 \leq [X,X]_{t,\Pi} \max_{\{1\leq j\leq n\}} |f''(X_{t_{j-1}}) - f''(\eta_j)|,$$

Since $[X,X]_{t,\Pi} = \sum_{j=1}^{n}(X_{t_j} - X_{t_{j-1}})^2 \xrightarrow{\mathbb{P}} \int_0^t {H_s}^2\, ds$ and $\int_0^t {H_s}^2\, ds < \infty$ by assumption, there exists a constant C, so that we obtain applying the Cauchy-Schwartz inequality

$$\mathbb{E}\Bigg([X,X]_{t,\Pi} \max_{\{1\leq j\leq n\}} |f''(X_{t_{j-1}}) - f''(\eta_j)|\Bigg) \leq \sqrt{C}\sqrt{\mathbb{E}\Bigg(\max_{\{1\leq j\leq n\}} |f''(X_{t_{j-1}}) - f''(\eta_j)|\Bigg)^2}.$$

Now using that X is continuous and the bounded convergence theorem, this term converges to zero.
Another application of Theorem 1.3.12 then yields

$$\frac{1}{2}\sum_{j=1}^{n} f''(X_{t_{j-1}})(X_{t_{j-1}} - X_{t_j})^2 \xrightarrow{\mathbb{P}} \frac{1}{2}\int_0^t f''(X_s)\, d[X,X]_s.$$

□

1.3.3 Black - Scholes formula

The Black-Scholes model is an option pricing model and was invented by Fischer Black and Myron Scholes in 1973 [5]. We will see later on, that the stock price process in the Black-Scholes model follows a geometric Brownian motion, therefore we will show, that a geometric Brownian motion is the solution of the stochastic differential equations

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

We have defined a geometric Brownian motion by:

$$X_t = X_0 \cdot \exp((\mu - \frac{\sigma^2}{2})t + \sigma B_t)$$

Note that $X_t = X_0 \exp(\xi_t)$ with $\xi_t = (\mu - \frac{\sigma^2}{2})t + \sigma B_t$.
Now applying Itô's formula yields

$$\begin{aligned} dX_t &= X_0 e^{\xi_t} d\xi_t + \frac{1}{2} X_0 e^{\xi_t} (d\xi_t)^2 \\ &= X_t((\mu - \frac{\sigma^2}{2})dt + \sigma dB_t) + \frac{1}{2} X_t \underbrace{((\mu - \frac{\sigma^2}{2})dt + \sigma dB_t)^2}_{=\sigma^2 dt} \\ &= X_t(\mu dt + \sigma dB_t). \end{aligned}$$

Definition 1.3.15.: Black Scholes asset price model

In the Black Scholes model the market consists of two liquid assets, namely a (riskless) zero coupon bond and a stock. The respective price processes are given as

- $S_t^0 = e^{rt}$ for the bond, where r is the riskless interest rate r.
- S_t^1 for the stock where

$$dS_t^1 = S_t^1(\mu dt + \sigma dW_t).$$

In the following we will denote the discounted price of the stock by $\tilde{S_t^1} = \frac{S_t^1}{S_t^0}$.

In the Black-Scholes Model, we assume the market for these two assets to be frictionless, i.e.:

- trading in continuous time is possible
- all assets are infinitely divisible
- unrestricted borrowing and lending at the same interest rate r is possible
- there are no transaction costs and no restriction on short-selling
- no bankruptcy or reorganization costs in case of default have to be paid

Definition 1.3.16.: Trading strategy

A trading strategy is an adapted two dimensional process $X = (X_t)_{0 \leq t \leq T} = (X_{t,0}, X_{t,1})_{0 \leq t \leq T}$, where $X_{t,0}$ is the number of bonds and $X_{t,1}$ is the number

of shares of the stock held at time t. So X_t represents the Black-Scholes portfolio at time t and its value is given by $V_t(X) = X_{t,0}S_t^0 + X_{t,1}S_t^1$.
A trading strategy is called self financing, if $dV_t(X) = X_{t,0}dS_t^0 + X_{t,1}dS_t^1$.

Definition 1.3.17.: Arbitrage free markets

An arbitrage opportunity is a self-financing trading strategy X with $V_0(X) = 0$ and $\mathbb{P}(V_t(X) \geq 0) = 1$, such that $\mathbb{P}(V_t(X) > 0) > 0$.
If a financial market does not provide arbitrage opportunities, we call it arbitrage free.

Definition 1.3.18.: Equivalent measures

Two measures $\mathbb{P}, \mathbb{Q}$ on $(\Omega, \mathbb{A})$ are called equivalent, if $\forall A \in \mathbb{A}$: $\mathbb{P}(A) = 0 \Leftrightarrow \mathbb{Q}(A) = 0$.

Theorem 1.3.19: Theorem of Girsanov

Assume B_t to be a Brownian motion on $(\Omega, \mathbb{A}, \mathbb{P})$. For a measurable, adapted process H_t, with $\int_0^T H_s^2\, ds < \infty$ $\mathbb{P}$ - a.s., we consider the stochastic process $\hat{B}_t$, that is defined by:

$$\hat{B}_t = B_t + \int_0^t H_s\, ds.$$

Furthermore we assume, that the exponential process

$$Z_t = \exp\Big(-\int_0^t H_s\, dB_s - \frac{1}{2}\int_0^t H_s^2\, ds\Big),$$

that is always a local martingale, is even a martingale. Then $\hat{B}$ is a Brownian motion on $(\Omega, \mathbb{A}, \mathbb{P}^*)$, where $\mathbb{P}^*$ is equivalent to $\mathbb{P}$ and defined by the Radon-Nikodym-derivative $\frac{d\mathbb{P}^*}{d\mathbb{P}} = Z_T$.

Definition 1.3.20.: Equivalent martingale measure/risk neutral probability measure

A probability measure $\mathbb{P}^*$ is called equivalent martingale measure or risk neutral probability measure on $(\Omega, \mathbb{A}, \mathbb{P})$, if $\mathbb{P}^*$ is equivalent to $\mathbb{P}$ and the discounted price process $\tilde{S}_t^1$ is a $\mathbb{P}^*$-local martingale.

The following theorem indicates, why we are looking for such an equivalent martingale measure.

Theorem 1.3.21: First fundamental theorem of asset pricing

The no arbitrage assumption is equivalent to the existence of an equivalent probability measure $\mathbb{P}^*$, such that the discounted price process $\tilde{S}_t^1$ is a $\mathbb{P}^*$-martingale.

Proof: For the proof see for instance [9].

A Claim C_T is a $\mathcal{F}_T$-measurable random variable, specifying a payoff at time T. According to the first fundamental theorem of asset pricing its price is given by

$$\tilde{V}_t(C_T) = \mathbb{E}_{\mathbb{P}^*}(\tilde{C}_T|\mathbb{F}_t).$$

Since we assume an arbitrage free market, there exists an equivalent martingale measure $\mathbb{P}^*$, such that the discounted price process $\tilde{S}_t^1$ is a martingale.
As an Itô process $\tilde{S}_t^1$ can only be a martingale, if it has the form: $\tilde{S}_t^1 = \int_0^t \tilde{H}_s\, dW_s$, with W being a standard Brownian motion. It follows that $d\tilde{S}_t^1 = \tilde{H}_t dW_t$ has to hold.

Let us now have a look on the differential equation of the discounted stock price $\tilde{S}_t^1$:

$$\begin{aligned} d\tilde{S}_t^1 = d((S_t^0)^{-1}S_t^1) &= (S_t^0)^{-1}dS_t^1 + S_t^1 d(S_t^0)^{-1} + \underbrace{d(S_t^0)^{-1}dS_t^1}_{=0} \\ &= \tilde{S}_t^1(\mu dt + \sigma dB_t) - r\tilde{S}_t^1 dt \\ &= \tilde{S}_t^1((\mu - r)dt + \sigma dB_t) \\ &= \sigma\tilde{S}_t^1(\frac{\mu - r}{\sigma}dt + dB_t) \end{aligned}$$

By setting $W_t = B_t + \frac{\mu - r}{\sigma}t \Leftrightarrow dW_t = dB_t + \frac{\mu - r}{\sigma}dt$, we get

$$d\tilde{S}_t^1 = \sigma\tilde{S}_t^1 dW_t.$$

Now we need Girsanov's theorem: If we change the probability measure in a way, that W is a Brownian motion under $\mathbb{P}^*$, then $\tilde{S}_t^1$ is a local martingale.

Hence by setting

$$\frac{d\mathbb{P}^*}{d\mathbb{P}} = Z_T,$$

where

$$Z_t = \exp(-\int_0^t H_s \, dB_s - \frac{1}{2}\int_0^t H_s^2 \, ds) = \underbrace{\exp(-\frac{\mu - r}{\sigma} B_t - \frac{1}{2}\frac{\mu - r}{\sigma}^2 t)}_{= \text{martingale}}$$

we obtain an equivalent measure $\mathbb{P}^*$ under which $\tilde{S}_t^1$ is a local martingale. Hence $\mathbb{P}^*$ is an equivalent martingale measure.

As mentioned before, the stochastic differential equation for $\tilde{S}_t^1$ has an explicit solution:

$$\tilde{S}_t^1 = \tilde{S}_0^1 \exp(\sigma W_t - \frac{\sigma^2}{2} t).$$

This solution is called the exponential martingale of the Brownian motion W_t.

Let us now derive the famous Black - Scholes formula. The payoff C_T of a call option is given by $C_T = f(S_T^1) = (S_T^1 - K)_+$. The discounted value then equals:

$$\begin{aligned} \mathbb{E}_{\mathbb{P}^*}(\tilde{C}_T | \mathbb{F}_t) &= \mathbb{E}_{\mathbb{P}^*}(f(S_T^1) e^{-r(T-t)} | \mathbb{F}_t) \\ &= e^{-r(T-t)} \mathbb{E}_{\mathbb{P}^*}(f(\frac{S_T^1}{S_t^1} S_t^1) | \mathbb{F}_t)). \end{aligned}$$

Note that

$$\frac{S_T^1}{S_t^1} = \exp((r - \frac{\sigma^2}{2})(T - t) + \sigma(W_T - W_t))$$

is independent of $\mathcal{F}_t$.

Combining the last two expressions, we obtain:

$$\mathbb{E}_{\mathbb{P}^*}(\tilde{C}_T | \mathbb{F}_t) = e^{-r(T-t)} \mathbb{E}_{\mathbb{P}^*}\Big(f\big(S_t^1 \exp((r - \frac{\sigma^2}{2})(T - t) + \sigma(W_T - W_t))\big) | \mathbb{F}_t \Big)$$

where $f(x) = (x - K)_+$.

$$V_t = e^{-r(T-t)} \int_{\mathbb{R}} (S_t^1 \cdot \exp((r - \frac{\sigma^2}{2})(T - t) + \sigma\sqrt{T - t} u) - K)_+ \cdot e^{-\frac{u^2}{2}} \, du$$

And by straight-forward calculations one obtains:

$$V_t = S_t^1 \Phi(d_1) - K e^{-r(T-t)} \Phi(d_2)$$

with

$$d_1 = \frac{\ln(\frac{S_t^1}{K}) + (r + \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}$$
$$d_2 = d_1 - \sigma\sqrt{T-t}.$$

This is exactly the Black Scholes formula for the price of a call option on a stock in a dividend free market.

Including dividends in the Black Scholes formula:

The original Black and Scholes model does not allow for dividend payments, but it can easily be modified to do so. This was first done by Merton [19] in 1973. He assumed that the price of the stock fulfills the following differential equation:

$$dS_t^1 = S_t^1((\mu - \kappa)dt + \sigma dW_t)$$

where κ is the so called dividend yield.
Using the same calculus as above, we get the Black Scholes formula for a call on a dividend paying stock as follows:

$$V_0 = S_0^1 e^{-\kappa(T-t)} \Phi(d_1) - K e^{-r(T-t)} \Phi(d_2)$$

with

$$d_1 = \frac{\ln(\frac{S_0^1}{K}) - (r - \kappa + \frac{\sigma^2}{2})(T-t)}{\sigma^2\sqrt{T-t}}$$
$$d_2 = \frac{\ln(\frac{S_0^1}{K}) - (r - \kappa - \frac{\sigma^2}{2})(T-t)}{\sigma^2\sqrt{T-t}}.$$

1.3.4 First passage times

In this section, we will present some characteristics concerning Brownian motions and first passage times. For further information concerning first passage times see for instance [20], [16], [22] and [23].
We are interested in the joint distribution of the first passage time τ of a process Y_t, with $Y_t = Y_0 + \nu t + \sigma W_t$, with $Y_0, \nu \in \mathbb{R}, \sigma > 0$ and W_t being a standard Brownian motion. Formally τ is defined by:

$$\tau = \inf\{t \geq 0 : Y_t < 0\}. \tag{1.5}$$

Let W be a standard Brownian motion, then we define:

$$M_t^W = \max_{u \in [0,t]} W_u \ , \ m_t^W = \min_{u \in [0,t]} W_u.$$

Since W has infinite variation, it follows immediately that for every $t > 0$

$$\mathbb{P}(M_t^W > 0) = 1 \qquad \text{and} \qquad \mathbb{P}(m_t^W < 0) = 1.$$

Theorem 1.3.22: Strong Markov property of the Brownian motion

Let $(B_t)_{t \geq 0}$ be a Brownian motion with respect to $\mathbb{F}$ and T a stopping time with $\mathbb{P}(T < \infty) = 1$. Then the process $B^{(T)} = (B_t^{(T)})_{t \geq 0}$ with

$$B_t^{(T)} = B_{t+T} - B_T$$

is a Brownian motion and $B^{(T)}$ is independent from $\mathcal{F}_T$.

Proof: See [16].

The above theorem can be used to derive the so called "reflection principle".

Lemma 1.3.23: Reflection principle
For every $t > 0, y \geq 0$ and $x \leq y$ the following formula holds:

$$\mathbb{P}(W_t \leq x, M_t^W \geq y) = \mathbb{P}(W_t \geq 2y - x) = \mathbb{P}(W_t \leq x - 2y)$$

Proof: We define the stopping time $T_y = \inf\{t \geq 0 | W_t \geq y\}$.

According to the strong Markov property, $W_{T_y+t} - y$ is a Brownian motion.
We start with:

$$\begin{aligned}\mathbb{P}(W_t \leq x, M_t^W \geq y) &= \mathbb{P}(W_t \leq x | M_t^W \geq y) \cdot \mathbb{P}(M_t^W \geq y) \\ &= \mathbb{P}(W_{T_y+(t-T_y)} - y \leq x - y | T_y < t) \cdot \mathbb{P}(M_t^W \geq y).\end{aligned}$$

Now we can use the strong Markov property:

$$\begin{aligned}&= \mathbb{P}(W_{T_y+(t-T_y)} - y \geq y - x | T_y < t) \cdot \mathbb{P}(M_t \geq y) \\ &= \mathbb{P}(W_t \geq y - x + y | T_y < t) \cdot \mathbb{P}(M_t \geq y) \\ &= \mathbb{P}(W_t \geq 2y - x, T_y < t) \\ &= \mathbb{P}(W_t \geq 2y - x, M_t^W > y).\end{aligned}$$

If $W_t \geq 2y - x$ and $x \leq y$ it follows that $W_t \geq y$, and therefore it is ensured that $M_t^W > y$. Knowing this we get

$$\mathbb{P}(W_t \geq 2y - x, M_t^W > y) = \mathbb{P}(W_t \geq 2y - x).$$

□

Now, we will have a look on the more general process $X_t = \nu t + \sigma W_t$, with $\sigma > 0$, $\nu \in \mathbb{R}$ and W_t being a standard Brownian motion under $\mathbb{P}$. From Girsanov's theorem we know, that X is a Brownian motion under an equivalent probability measure and thus we get:

$$\mathbb{P}(M_t^X > 0) = 1 \qquad \text{and} \qquad \mathbb{P}(m_t^X < 0) = 1$$

for every $t > 0$.

Lemma 1.3.24: Let $X_t = \nu t + \sigma W_t$. Then, for every $t > 0, y \geq 0$ and $x \leq y$ the following formula holds:

$$\mathbb{P}(X_t \leq x, M_t^X \geq y) = e^{2\nu y \sigma^{-2}} \mathbb{P}(X_t \geq 2y - x + 2\nu t). \tag{1.6}$$

Proof: Let us set

$$I := \mathbb{P}(X_t \leq x, M_t^X \geq y) = \mathbb{P}(X_t^\sigma \leq \frac{x}{\sigma}, M_t^{X^\sigma} \geq \frac{y}{\sigma}),$$

with $X_t^\sigma = W_t + \frac{\nu t}{\sigma}$. Thus, without loss of generality we may assume $\sigma = 1$.
By applying Girsanov's theorem, we get the equivalent probability measure defined by:

$$\frac{d\hat{\mathbb{P}}}{d\mathbb{P}} = e^{-\nu W_T - \nu^2 \frac{T}{2}} \, , \, \mathbb{P} - a.s.$$

or inverted:

$$\frac{d\mathbb{P}}{d\hat{\mathbb{P}}} = e^{\nu W_T + \nu^2 \frac{T}{2}} = e^{\nu X_T - \nu^2 \frac{T}{2}} \, , \, \hat{\mathbb{P}} - a.s.$$

We get:

$$I = \mathbb{E}_{\hat{\mathbb{P}}}(e^{\nu X_T - \nu^2 \frac{T}{2}} \mathbb{I}_{\{X_t \le x, M_t^X \ge y\}}).$$

Since X follows a Brownian motion under $\hat{\mathbb{P}}$, an application of the reflection principle gives:

$$\begin{aligned} I &= \mathbb{E}_{\hat{\mathbb{P}}}\big(e^{\nu(2y - X_T) - \nu^2 \frac{T}{2}} \mathbb{I}_{\{2y - X_t \le x, M_t^X \ge y\}}\big) \\ &= \mathbb{E}_{\hat{\mathbb{P}}}\big(e^{\nu(2y - X_T) - \nu^2 \frac{T}{2}} \mathbb{I}_{\{X_t \ge 2y - x\}}\big) \\ &= e^{2\nu y} \mathbb{E}_{\hat{\mathbb{P}}}\big(e^{-\nu X_T - \nu^2 \frac{T}{2}} \mathbb{I}_{\{X_t \ge 2y - x\}}\big). \end{aligned}$$

Since $2y - x \ge y$, we can define another equivalent probability measure $\tilde{\mathbb{P}}$ by setting

$$\frac{d\tilde{\mathbb{P}}}{d\hat{\mathbb{P}}} = e^{-\nu X_T - \nu^2 \frac{T}{2}} \, , \, \hat{\mathbb{P}} - a.s.$$

It follows:

$$I = e^{2\nu y} \mathbb{E}_{\hat{\mathbb{P}}}(e^{-\nu X_T - \nu^2 \frac{T}{2}} \mathbb{I}_{\{X_t \ge 2y - x\}}) = e^{2\nu y} \tilde{\mathbb{P}}(X_t \ge 2y - x)$$

and $\tilde{W}_t = X_t + \nu t$ follows a standard Brownian motion under $\tilde{\mathbb{P}}$. Since $X_t = \tilde{W}_t - \nu t$, we obtain

$$I = e^{2\nu y} \tilde{\mathbb{P}}(\tilde{W}_t + \nu t \ge 2y - x + 2\nu t).$$

Since $\tilde{W}_t$ is a standard Brownian motion under $\tilde{\mathbb{P}}$ and W_t under $\mathbb{P}$, it follows that

$$\tilde{\mathbb{P}}(\tilde{W}_t + \nu t \ge 2y - x + 2\nu t) = \mathbb{P}(\underbrace{W_t + \nu t}_{= X_t} \ge 2y - x + 2\nu t),$$

and we finally obtain

$$I = e^{2\nu y} \mathbb{P}(X_t \ge 2y - x + 2\nu t).$$

□

Corollary 1.3.25: For every $x, y \in \mathbb{R}$ such that $y \geq 0$ and $x \leq y$, we have

$$\mathbb{P}(X_t \leq x, M_t^X \geq y) = e^{2\nu y \sigma^{-2}} N(\frac{x - 2y - \nu t}{\sigma\sqrt{T}})$$

and

$$\mathbb{P}(X_t \leq x, M_t^X \leq y) = N(\frac{x - \nu t}{\sigma\sqrt{T}}) - e^{2\nu y \sigma^{-2}} N(\frac{x - 2y - \nu t}{\sigma\sqrt{T}}).$$

Proof: The first equation can be obtained by

$$\mathbb{P}(X_t \geq 2y - x + 2\nu t) = \mathbb{P}(-\sigma W_t \leq x - 2y - \nu t) = N(\frac{x - 2y - \nu t}{\sigma\sqrt{T}}),$$

since $-\sigma W_t$ is normally distributed with mean zero and variance $\sigma^2 t$. The second formula follows from

$$\mathbb{P}(X_t \leq x) = \mathbb{P}(X_t \leq x, M_t^X \leq y) + \mathbb{P}(X_t \leq x, M_t^X \geq y).$$

□

Corollary 1.3.26: For every $y \geq 0$ we have:

$$\mathbb{P}(M_t^X \leq y) = N(\frac{y - \nu t}{\sigma\sqrt{T}}) - e^{2\nu y \sigma^{-2}} N(\frac{-y - \nu t}{\sigma\sqrt{T}}).$$

The same calculations can be used for the minimum value, since the connection is the following:

$$\mathbb{P}(\max_{u \in [0,t]}(\sigma W_u - \nu u) \geq -y) = \mathbb{P}(\min_{u \in [0,t]}(-\sigma W_u + \nu u) \leq y) = \mathbb{P}(\min_{u \in [0,t]} X_u \leq y).$$

Therefore we get the following corollaries:

Corollary 1.3.27: For every $x, y \in \mathbb{R}$ such that $y \leq 0$ and $y \leq x$, we have:

$$\mathbb{P}(X_t \geq x, m_t^X \geq y) = N(\frac{-x + \nu t}{\sigma\sqrt{T}}) - e^{2\nu y \sigma^{-2}} N(\frac{2y - x + \nu t}{\sigma\sqrt{T}}).$$

Corollary 1.3.28: For every $y \leq 0$ we have:

$$\mathbb{P}(m_t^X \geq y) = N(\frac{-y + \nu t}{\sigma\sqrt{T}}) - e^{2\nu y \sigma^{-2}} N(\frac{y + \nu t}{\sigma\sqrt{T}}).$$

Using these results, we obtain the joint probability of $Y_t = Y_0 + X_t$ and its first passage time τ.
It is obvious, that

$$\mathbb{P}(\tau \geq s) = \mathbb{P}(\inf_{t \in [0,s]} Y_t \geq 0). \tag{1.7}$$

Corollary 1.3.29: For every $s > 0$ and $y \geq 0$ we have:

$$\mathbb{P}(Y_s \geq y, \tau \geq s) = N(\frac{-y + Y_0 + \nu s}{\sigma\sqrt{s}}) - e^{-2\nu y \sigma^{-2} Y_0} N(\frac{-y - Y_0 + \nu s}{\sigma\sqrt{s}}).$$

Proof: Follows directly of Corollary 1.3.27.

1.3.5 Poisson processes

In this section we will discuss the stochastic foundations of intensity models, namely the Poisson process and variants thereof. We will deal with the time-homogeneous Poisson process, the time-inhomogeneous Poisson process and the Cox process. For further information I refer to [6] and [21].

Time-homogeneous Poisson process:

Definition 1.3.30.: Time-homogeneous Poisson process

A process $N = (N_t)_{t \geq 0}$ is called Poisson process with intensity λ, when:

- $N_0 = 0$
- N is an integer-valued, right continuous increasing process
- N has stationary and independent increments
- N follows in any interval of length t a Poisson distribution with mean λt, i.e.:
$$\mathbb{P}(N_{(t+s)} - N_s = k) = e^{-\lambda t}(\lambda t)^k / k!.$$

Characteristics of the time homogeneous Poisson process:

Let $\tau^1, \tau^2, \ldots$ be the jump times of N. Given the above definition, it follows that
$$\mathbb{P}(N_t = 0) = \mathbb{P}(\tau^1 > t) = e^{-\lambda t}.$$
This means, that τ^1 is exponential distributed with mean $1/\lambda$.
Furthermore we have
$$\lim_{t \to 0} \mathbb{P}(N_t \geq 2)/t = 0$$
$$\lim_{t \to 0} \mathbb{P}(N_t = 1)/t = \lambda.$$
Note that $\lambda = \mathbb{E}(N_t)/t = Var(N_t)/t$.

We are interested in the first jump time τ^1. We already know that τ^1 is a random variable and it is exponential distributed with mean $1/\lambda$.

It follows, that the infinitesimal probability of the first jump at time t is given by

$$\mathbb{P}(\tau \in [t, t+dt) | \tau \geq t) = \frac{\mathbb{P}(\tau \in [t, t+dt))}{\mathbb{P}(\tau > t)} = \frac{e^{-\lambda t} - e^{-\lambda(t+dt)}}{e^{-\lambda t}} \approx \lambda dt.$$

Time-inhomogeneous Poisson process:

A time inhomogeneous Poisson process M_t is again a integer-valued increasing, right continuous process with $M_0 = 0$, but in contrast to the time homogeneous Poisson process with time-varying intensity λ_t.
We define

$$\Lambda(t) := \int_0^t \lambda_u \, du,$$

where λ_t is called intensity or hazard rate and $\Lambda(t)$ the cumulative intensity, cumulative hazard rate or hazard function.
Then the time-inhomogeneous Poisson process M_t with intensity λ_t is given by $M_t = N_{\Lambda(t)}$, where N_t is a time-homogeneous Poisson Process with intensity $\lambda = 1$.
With time-varying λ_t, the increments of M_t are no longer identically distributed, albeit still independent.

Again, our interest lies on the first jump time of M, denoted by τ. With this definition N jumps the first time at $\Lambda(\tau)$. Since N is a standard Poisson process, we know that the first jump time is an exponential distributed random variable with mean 1 and hence

$$\Lambda(\tau) =: \xi \sim \text{Exp}(1).$$

By inverting this formula we get

$$\tau = \Lambda^{-1}(\xi).$$

Having not jumped before t, the probability of a jump in the time period $[t, t+dt]$ is therefore

$$\mathbb{P}(\tau \in [t, t+dt] | \tau > t, \mathbb{F}_t) \approx \lambda_t dt,$$

where the factor λ_t is strictly positive.

Cox process:

The Cox process C_t is again an integer-valued, increasing, right continuous process, with $C_0 = 0$ and with stochastic intensity λ_t, where λ_t is a right continuous and $\mathcal{F}_t$ - adapted process . We define

$$\Lambda(T) = \int_0^T \lambda_t \, dt,$$

with $\lambda_t > 0$.
Then the Cox process C_t with intensity λ_t is defined as $C_t = N_{\Lambda(T)}$, where N_t is again a time-homogeneous Poisson Process with intensity $\lambda = 1$.

Also with stochastic intensity, we can use the link to the standard Poisson process N to describe the first jump time τ of C. More precisely we have

$$\Lambda(\tau) =: \xi \sim \text{Exp}(1)$$

and

$$\tau = \Lambda^{-1}(\xi).$$

The infinitesimal probability of the first jump is again given by

$$\mathbb{P}(\tau \in [t, t + dt] | \tau > t, \mathbb{F}_t) = \lambda_t dt.$$

Chapter 2

Structural models

2.1 Introduction

In this chapter we will focus on structural models, more precisely on the Merton and the Black and Cox model. For further information on structural models I refer to [2], on which the framework of this chapter is based.
The first structural model was introduced by Merton [19] in 1974 and is nowadays known as the Merton model. The models setup is as follows: the firm value V_t is assumed to follow a geometric Brownian motion and the liabilities are given by one zero coupon bond. At the bond's maturity T, the firm has to be able to pay out the bond's face value L. If the firms value is below L a default occurs!
The big disadvantage of Merton's model is, that default can only occur at maturity T, which is widely acknowledged to be unrealistic. For this reason so called first passage time models were introduced, e.g. the Black and Cox model, to which we will come back later.
There are plenty of other structural models. They may be classified by the following components (cf. [2]):

- the dynamics of the asset value of the firm
- the structure of the firms liabilities
- the default event (barrier)
- the recovery rule
- other economic quantities

2.2 Merton's approach

Merton's model has the following setup. The market has to be frictionless and the short term interest rate r is constant and deterministic, which implies that the price of a zero coupon bond with maturity T at time t equals

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

Let the firm's equity value be denoted by $E(V_t)$ (or short E_t), the firm's debt by $D(V_t)$, and the firm value at time t by $V_t = E(V_t) + D(V_t)$. The firm value process V_t is assumed to follow a geometric Brownian motion under the equivalent martingale measure $\mathbb{P}^*$:

$$dV_t = V_t((r-\kappa)dt + \sigma_V dW_t^*)$$

where κ is the non-negative payout ratio, which reflects the outflow of capital of the firm (for example dividends), W_t^* is a standard Brownian motion and σ_V is the constant volatility of V_t.

The credit event (in this case the default) is triggered, when the firm value process $V_T \leq L$.

Remark: Here one big problem of Merton's model can be observed. Both, firm value and its volatility are not observable in general. But there is a way to obtain these quantities in terms of the observable equity value and its volatility. We will come back to this topic later.

The payment to the bondholders at time T in this model is

$$X(T) = L \cdot \mathbb{I}_{\{V_T \geq L\}} + V_T \cdot \mathbb{I}_{\{V_T < L\}} = \min(V_T, L) = L - (L - V_T)^+. \qquad (2.1)$$

with $x^+ = \max(x, 0)$.

Hence the payment can be decomposed in a fixed amount L and a put option on the firm value with maturity T and strike L. This put is also called "put-to-default".

So the value of the firm's debt is:

$$\mathrm{D}(t,T) = D(V_t) = L \cdot \mathrm{B}(t,T) - P_t, \tag{2.2}$$

where P_t is the price of the put-to-default.

With (2.1) we get

$$E(V_T) = V_T - D(V_T) = V_T - \min(V_T, L) = (V_T - L)^+ \tag{2.3}$$

Again we see that this is the payoff of a European Call-Option, and with the call-put-parity we get:

$$C_t - P_t = V_t - L\,\mathrm{B}(t,T). \tag{2.4}$$

Now using the Black-Scholes-formula leads us to the following result:

Proposition 2.2.1: The price of a defaultable zero coupon bond is

$$\mathrm{D}(t,T) = V_t \cdot e^{-\kappa(T-t)} N(-d_1(V_t, T-t)) + L\,\mathrm{B}(t,T) N(d_2(V_t, T-t))$$

where κ is again the non-negative payout ratio, N is the standard Gaussian cumulative distribution function and

$$\begin{aligned} d_1(V_t, T-t) &= \frac{\ln(V_t/L) + (r - \kappa + \frac{1}{2}\sigma_V^2)(T-t)}{\sigma_V \sqrt{T-t}} \\ d_2(V_t, T-t) &= \frac{\ln(V_t/L) + (r - \kappa - \frac{1}{2}\sigma_V^2)(T-t)}{\sigma_V \sqrt{T-t}}. \end{aligned}$$

Proof: The Black-Scholes price of a put option with strike L on a dividend-paying stock equals:

$$P_t = L\,\mathrm{B}(t,T) N(-d_2(V_t, T-t)) - V_t \cdot e^{-\kappa(T-t)} N(-d_1(V_t, T-t)).$$

From (2.2) we get:

$$\begin{aligned} \mathrm{D}(t,T) &= L\,\mathrm{B}(t,T) - P_t \\ &= L\,\mathrm{B}(t,T)(1 - N(-d_2(V_t, T-t))) + V_t \cdot e^{-\kappa(T-t)} N(-d_1(V_t, T-t)). \end{aligned}$$

Since $N(-x) = 1 - N(x)$, we get the final result

$$\mathrm{D}(t,T) = V_t \cdot e^{-\kappa(T-t)} N(-d_1(V_t, T-t)) + L\,\mathrm{B}(t,T) N(d_2(V_t, T-t)).$$

$\square$

Distance to default

The probability of default in this model is closely linked to the distance-to-default, which is defined as:

$$DD_t = \frac{\ln(V_t/L) + (r - \kappa - \frac{1}{2}\sigma_V^2)(T-t)}{\sigma_V \sqrt{T-t}}.$$

The default probability then equals $\mathbb{P}(V_T \leq L) = N(-DD_t)$.

Remark: The distance to default is the number of standard deviations, that the firms value is above or below the default barrier.

Credit spread

We define the yield of the non-defaultable bond and the defaultable bond via

$$Y(t,T) = -\frac{\ln B(t,T)}{T-t} \quad , \quad Y^d(t,T) = -\frac{\ln D(t,T)}{T-t}.$$

The credit spread is then given by

$$S(t,T) = Y^d(t,T) - Y(t,T).$$

In the Merton model we assume a riskless interest rate r, so it follows that $Y(t,T) = r$. Using Proposition 2.2.1 we find:

$$S(t,T) = -\frac{ln(\frac{V_t}{B(t,T)} e^{-\kappa(T-t)} N(-d_1(V_t, T-t)) + L \cdot N(d_2(V_t, T-t)))}{T-t}.$$

Estimation of the firm value

As already mentioned, we have to deal with the non-observability of the asset value and its volatility. But we can bypass this problem by using some helpful characteristics of the observable equity value.
Following Jones et al. (1984) the estimation of the firm value is based on two equations. The first one can be derived from Merton's formula:

$$E_t = V_t N(d_1(V_t, T-t)) - Le^{-r(T-t)} N(d_2(V_t, T-t)). \tag{2.5}$$

Using Itô's lemma, one can check, that the dynamics of E_t under $\mathbb{P}^*$ are:

$$dE_t = rE_t dt + V_t N(d_1(V_t, T-t))\sigma_V dW_t^*.$$

Therefore the volatility of the firm's equity has the following representation:

$$\sigma_E = \frac{V_t}{E_t} N(d_1(V_t, T-t))\sigma_V. \tag{2.6}$$

Given now the observable input parameters E, σ_E, r and L, we are able to numerically solve equation (2.5) for $t = 0$ and (2.6) iteratively and we get the firm value V_0 and its volatility σ_V.

2.2.1 Advantages and drawbacks of the Merton's model:

Since the Merton model was the first credit risk model, there are a few things, that are too simplifying in this approach. But, nevertheless, it was/is quite popular because of its simplicity and its clear references to the input parameters.

The main drawbacks of the Merton model are:

- default can only occur at maturity T
- the non-observability of the asset value, although it can be bypassed
- the constant parameters, for example the constant short rate r
- for small maturities the produced credit spreads are close to zero, which is not consistent with empirical studies

Remark: Another limitation that could come to mind might be the stochastic process. Assuming that the firms value follows a geometric Brownian motion, induces lognormal distributed dynamics for V_t. However, [6] found that "this assumption is quite robust and, according to KMV's empirical studies, actual data confirm quite well for this hypothesis."

2.3 Merton model with stochastic interest rate

In this section I want to point out, how stochastic parameters can be taken into account in the Merton model.
Therefore we consider the short rate r to follow the Vasicek model under the spot martingale measure $\mathbb{P}^*$, i.e.:

$$dr_t = (\alpha - \beta r_t)dt + \sigma_r d\tilde{W}_t,$$

with $\tilde{W}_t$ being a standard Brownian motion. For further details on the Vasicek model and other short rate models, see for instance [6].

The value process V_t is here

$$dV_t = V_t(r_t dt + \sigma_V dW_t^*),$$

with W_t^* being again a standard Brownian motion.
Again, this equation holds under the spot martingale measure $\mathbb{P}^*$. Furthermore we assume the Brownian motions $\tilde{W}$ and W^* to be correlated with constant correlation $\rho_{V,r}$. With $b(t,T)$ being the volatility of a default free zero coupon bond in the Vasicek model, i.e.:

$$b(t,T) = \sigma_r \beta^{-1}(1 - e^{-\beta(T-t)}), \forall t \in [0,\mathrm{T}].$$

we define $\sigma^2(t,T)$ by

$$\sigma^2(t,T) = \int_t^T (\sigma_V^2 - 2\rho_{V,r}\sigma_V b(u,T) + b^2(u,T))du.$$

As shown by Jamshidian (1989) [15], the put-to-default can be obtained in this setup by:

$$P_t = L\,\mathrm{B}(t,T)N(-d_2(V_t,t,T)) - V_t N(-d_1(V_t,t,T)),$$

where

$$d_1(V_t,t,T) = \frac{\ln(V_t/\,\mathrm{B}(t,T)) - \ln(L) + \frac{1}{2}\sigma^2(t,T)}{\sigma(t,T)}.$$

$$d_2(V_t,t,T) = \frac{\ln(V_t/\,\mathrm{B}(t,T)) - \ln(L) - \frac{1}{2}\sigma^2(t,T)}{\sigma(t,T)}.$$

Since we know that $D(t,T) = LB(t,T) - P_t$, we get

$$\begin{aligned} D(t,T) &= LB(t,T)(1 - N(-h_2(V_t,t,T))) + V_t N(-h_1(V_t,t,T)) \\ &= LB(t,T)(N(h_2(V_t,t,T))) + V_t N(-h_1(V_t,t,T)). \end{aligned}$$

2.4 Black and Cox model

We have seen in the last chapter, that the main restriction of the Merton model is, that default can only occur at maturity T. To avoid this unrealistic feature Black and Cox [3] introduced a first passage time model, where default is assumed to occur the first time the firm value V_t falls below an pre-specified time-dependent barrier L_t.
Let be τ the default time, then we have:

$$\tau = \inf\{t \geq 0 : V_t < L_t\}. \tag{2.7}$$

This setup allows a lot more modeling choices. For example, the default barrier can be assumed to be stochastic itself and the recovery payoff may be specified in various ways.

2.4.1 The model

As mentioned before Black and Cox's approach is based on Merton's model. However, it is generalized in numerous ways: not only premature default is allowed, also more specific features of debt contracts, like safety covenants, debt subordination etc., can be included. Assuming continuous payout at a rate κ (the dividend yield) the value process is given by:

$$dV_t = V_t((r - \kappa)dt + \sigma_V dW_t^*). \tag{2.8}$$

The other parameters are set as in Merton's model.

Next we focus on the safety covenants. The safety covenants define, when the bondholders have the right to force the firm into bankruptcy or reorganization to protect their capital. More precisely, as soon as the firm value process falls below the safety covenants, the bondholders take over control of the firm and thus default occurs. In Black and Cox' approach the safety covenants are modeled by a deterministic barrier $l(t)$. One of the most popular choices is to set $l(t) = Ke^{-\gamma(T-t)}$, where γ is the continuous payoff the bondholders receive for their capital. Besides undershooting the safety covenants there is a second possibility

of default. Namely, if at maturity T $V_T < L$.
Thus setting

$$L_t = \begin{cases} l(t), & \text{for } t < T, \\ L, & \text{for } t = T, \end{cases}$$

default occurs, when $V_t < L_t$ for the first time, or mathematically, default time τ is set as:

$$\tau := \inf\{t \in [0,T] : V_t < L_t\}. \tag{2.9}$$

Economically it makes sense to assume that $\gamma > r$, since the bondholders will only provide their capital, if the expected return is higher than the riskless interest rate r. Furthermore we demand that

$$L_t \leq Le^{-\gamma(T-t)} \tag{2.10}$$

i.e.: $K \leq L$.

This condition makes sense, since the bondholders will not force the firm into bankruptcy, when the firm value is higher than the future payment L. When $K < L$, the bondholders do not force the firm into bankruptcy immediately when the firm value is below the (with γ) discounted value of the debt L.

The next step we will make, is to develop a formula for the price of a defaultable bond. Since the Black and Cox model allows more specifications, we will also include a recovery rule. So let us set β_2 the recovery rate, when the bondholders force the firm into bankruptcy, which means $\tau < T$, and β_1 the recovery rate at maturity T, when $V_T < L$.

To get the formula, we distinguish different possibilities of default. So when no default occurs, the value of the bond at time t is $Le^{-r(T-t)}$.
When default occurs at maturity T, which means that $V_t > L_t$ for all $t \in [t,T)$, but $V_T < L$, the value of the payoff at time t is $\beta_1 V_T e^{-r(T-t)}$. Last but not least default can occur at $\tau < T$, in which case we have a payoff of $K\beta_2 e^{-\gamma(T-\tau)} e^{-r(t-\tau)}$.

By combining the different possibilities we get

$$\begin{aligned} D(t,T) &= \mathbb{E}_{\mathbb{P}^*}(Le^{-r(T-t)}\mathbb{I}_{\{\tau\geq T, V_T\geq L\}}|\mathbb{F}_t) \\ &+ \mathbb{E}_{\mathbb{P}^*}(\beta_1 V_T e^{-r(T-t)}\mathbb{I}_{\{\tau\geq T, V_T< L\}}|\mathbb{F}_t) \\ &+ \mathbb{E}_{\mathbb{P}^*}(K\beta_2 e^{-\gamma(T-\tau)}e^{-r(t-\tau)}\mathbb{I}_{\{t<\tau<T\}}|\mathbb{F}_t). \end{aligned}$$

Using now the results on the first passage time of a Brownian motion one can in fact evaluate the expectations above.
For notional convenience we will use the notation $\nu = r - \kappa - \frac{1}{2}\sigma_V$, $\tilde{\nu} = \nu - \gamma = r - \kappa - \gamma - \frac{1}{2}\sigma_V$ and $\tilde{a} = \tilde{\nu}\sigma_V^{-2}$ in the sequel.

Proposition 2.4.1: Assume that $\tilde{\nu}^2 + 2\sigma_V^2(r-\gamma) > 0$. Then the price of a defaultable bond at time t is given by

$$\begin{aligned} D(t,T) &= LB(t,T)(N(h_1(V_t,T-t)) - R_t^{2\tilde{a}}N(h_2(V_t,T-t))) \\ &+ \beta_1 V_t e^{-\kappa(T-t)}(N(h_3(V_t,T-t)) - N(h_4(V_t,T-t))) \\ &+ \beta_1 V_t e^{-\kappa(T-t)}R_t^{2\tilde{a}+2}(N(h_5(V_t,T-t)) - N(h_6(V_t,T-t))) \\ &+ \beta_2 V_t(R_t^{\theta+\zeta}N(h_7(V_t,T-t)) + R_t^{\theta-\zeta}N(h_8(V_t,T-t))), \end{aligned}$$

where

$$R_t = l(t)/V_t \qquad \theta = \tilde{a}+1 \qquad \zeta = \sigma_V^{-2}\sqrt{\tilde{\nu}^2 + 2\sigma^2(r-\gamma)}$$

and

$$\begin{aligned} h_1(V_t,T-t) &= \frac{\ln(V_t/L) + \nu(T-t)}{\sigma\sqrt{T-t}} \\ h_2(V_t,T-t) &= \frac{\ln(l^2(t)) - \ln(LV_t) + \nu(T-t)}{\sigma\sqrt{T-t}} \\ h_3(V_t,T-t) &= \frac{\ln(L/V_t) - (\nu+\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \\ h_4(V_t,T-t) &= \frac{\ln(K/V_t) - (\nu+\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \\ h_5(V_t,T-t) &= \frac{\ln(l^2(t)) - \ln(LV_t) + (\nu+\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \\ h_6(V_t,T-t) &= \frac{\ln(l^2(t)) - \ln(KV_t) + (\nu+\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \\ h_7(V_t,T-t) &= \frac{\ln(l(t)/V_t) + \zeta\sigma^2(T-t)}{\sigma\sqrt{T-t}} \\ h_8(V_t,T-t) &= \frac{\ln(l(t)/V_t) - \zeta\sigma^2(T-t)}{\sigma\sqrt{T-t}}. \end{aligned}$$

Before we can prove the proposition, we need an elementary lemma:

Lemma 2.4.2: For any $a \in \mathbb{R}$, $b > 0$ and any $y > 0$, we have

$$\int_0^y x\, dN\Big(\frac{\ln x + a}{b}\Big) = e^{\frac{1}{2}b^2 - a} N\Big(\frac{\ln y + a - b^2}{b}\Big) \tag{2.11}$$

and

$$\int_0^y x\, dN\Big(\frac{-\ln x + a}{b}\Big) = e^{\frac{1}{2}b^2 + a} N\Big(\frac{-\ln y + a + b^2}{b}\Big). \tag{2.12}$$

Let $a, b, c \in \mathbb{R}$ satisfy $b < 0$ and $c^2 > a$. Then for every $y > 0$

$$\int_0^y e^{ax}\, dN\Big(\frac{b - cx}{\sqrt{x}}\Big) = \frac{d + c}{2d} g(y) + \frac{d - c}{2d} h(y), \tag{2.13}$$

where $d = \sqrt{c^2 - 2a}$ and

$$g(y) := e^{b(c-d)} N\Big(\frac{b - dy}{\sqrt{y}}\Big)\ ,\ h(y) := e^{b(c+d)} N\Big(\frac{b + dy}{\sqrt{y}}\Big).$$

Proof

The proof of (2.11) and (2.12) is quite standard, so we will focus on (2.13). Denote

$$f(y) = \int_0^y e^{ax}\, dN\Big(\frac{b - cx}{\sqrt{x}}\Big) = \int_0^y e^{ax} n\Big(\frac{b - cx}{\sqrt{x}}\Big)\Big(- \frac{b}{2x^{\frac{3}{2}}} - \frac{c}{2\sqrt{x}}\Big)\, dx$$

with n being the probability density function of a standard normal distribution. Now we will have a look on the right side of (2.13):

$$\begin{aligned} g'(x) &= e^{b(c - \sqrt{c^2 - 2a})} n\Big(\frac{b - \sqrt{c^2 - 2a}x}{\sqrt{x}}\Big)\Big(- \frac{b}{2x^{\frac{3}{2}}} - \frac{\sqrt{c^2 - 2a}}{2\sqrt{x}}\Big) \\ &= e^{ax} n\Big(\frac{b - cx}{\sqrt{x}}\Big)\Big(- \frac{b}{2x^{\frac{3}{2}}} - \frac{d}{2\sqrt{x}}\Big) \end{aligned}$$

and

$$\begin{aligned} h'(x) &= e^{b(c + \sqrt{c^2 - 2a})} n\Big(\frac{b + \sqrt{c^2 - 2a}x}{\sqrt{x}}\Big)\Big(- \frac{b}{2x^{\frac{3}{2}}} + \frac{\sqrt{c^2 - 2a}}{2\sqrt{x}}\Big) \\ &= e^{ax} n\Big(\frac{b - cx}{\sqrt{x}}\Big)\Big(- \frac{b}{2x^{\frac{3}{2}}} + \frac{d}{2\sqrt{x}}\Big). \end{aligned}$$

Therefore we get:

$$g'(x) + h'(x) = -e^{ax}\frac{b}{x^{\frac{3}{2}}}n\Big(\frac{b - cx}{\sqrt{x}}\Big)$$

and

$$g'(x) - h'(x) = -e^{ax}\frac{d}{x^{\frac{1}{2}}}n\Big(\frac{b - cx}{\sqrt{x}}\Big).$$

Hence we can represent f as follows:

$$f(x) = \frac{1}{2}\int_0^y (g'(x) + h'(x) + \frac{c}{d}(g'(x) - h'(x)))\,dx.$$

Since $\lim_{y\to 0+} g(y) = \lim_{y\to 0+} h(y) = 0$, we have for every $y > 0$:

$$f(y) = \frac{1}{2}(g(y) + h(y)) + \frac{c}{2d}(g(y) - h(y)).$$

□

Proof of Proposition 2.4.1: Since most of the calculations needed for the proof are are standard but lengthly, we will not go into detail and only sketch the proof here. For full details see [16].

We need to find the following conditional expectations:

$$D_1(t,T) = L\,\mathrm{B}(t,T)\mathbb{P}^*\{V_T \geq L, \tau \geq T|\mathbb{F}_t\},$$
$$D_2(t,T) = \beta_1\,\mathrm{B}(t,T)\mathbb{E}_{\mathbb{P}^*}(V_T\mathbb{I}_{\{V_T<L,\tau\geq T\}}|\mathbb{F}_t),$$
$$D_3(t,T) = K\beta_2 V_t e^{-\gamma T}\mathbb{E}_{\mathbb{P}^*}(e^{(\gamma-r)\tau}\mathbb{I}_{\{t<\tau<T\}}|\mathbb{F}_t).$$

For notional simplicity, we set $t = 0$. Let us start with $D_1(0,T)$:
Using Corollary 1.3.29 we find

$$\mathbb{P}^*\{V_T \geq L, \tau \geq T\} = N\Big(\frac{\ln\frac{V_0}{L} + \nu T}{\sigma\sqrt{T}}\Big) - R_0^{2\tilde{a}}N\Big(\frac{\ln\frac{l^2(0)}{LV_0} + \nu T}{\sigma\sqrt{T}}\Big)$$

with $R_0 = \frac{l(0)}{V_0}$ and hence

$$D_1(0,T) = LB(0,T)(N(h_1(V_0,T)) - R_0^{2\tilde{a}}N(h_2(V_0,T))).$$

For $D_2(0,T)$, which is the part, where default is triggered at time T, we consider:

$$\frac{D_2(0,T)}{\beta_1\,\mathrm{B}(0,T)} = \mathbb{E}_{\mathbb{P}^*}(V_T \mathbb{I}_{\{V_T<L,\tau\geq T\}}), = \int_K^L x\, d\mathbb{P}^*\{V_T < x, \tau \geq T\}.$$

Using again Corollary 1.3.29 we get:

$$d\mathbb{P}^*\{V_T < x, \tau \geq T\} = dN\Big(\frac{\ln\frac{x}{V_0} - \nu T}{\sigma\sqrt{T}}\Big) + R_0^{2\tilde{a}} dN\Big(\frac{\ln\frac{l^2(0)}{xV_0} + \nu T}{\sigma\sqrt{T}}\Big).$$

Denoting

$$K_1(0) = \int_K^L x\, dN\Big(\frac{\ln x - \ln V_0 - \nu T}{\sigma\sqrt{T}}\Big)$$

and

$$K_2(0) = \int_K^L x\, dN\Big(\frac{-2\ln l(0) - \ln x - \ln V_0 + \nu T}{\sigma\sqrt{T}}\Big),$$

using Lemma 2.4.2, we have:

$$K_1(0) = V_0 e^{(r-\kappa)T}\Big(N\Big(\frac{\ln\frac{L}{V_0} - \hat{\nu}T}{\sigma\sqrt{T}}\Big) - N\Big(\frac{\ln\frac{K}{V_0} - \hat{\nu}T}{\sigma\sqrt{T}}\Big)\Big),$$

and

$$K_2(0) = V_0 R_0^2 e^{(r-\kappa)T}\Big(N\Big(\frac{\ln\frac{l^2(0)}{LV_0} + \hat{\nu}T}{\sigma\sqrt{T}}\Big) - N\Big(\frac{\ln\frac{l^2(0)}{KV_0} + \hat{\nu}T}{\sigma\sqrt{T}}\Big)\Big),$$

where $\hat{\nu} = \nu + \sigma^2 = r - k + \frac{1}{2}\sigma^2$. Since $D_2(0,T) = \beta_1\,\mathrm{B}(0,T)(K_1(0) + R_0^{\tilde{a}} K_2(0))$, we find:

$$\begin{aligned} D_2(t,T) &= \beta_1 V_0 e^{-\kappa T}(N(h_3(V_0,T)) - N(h_4(V_0,T))) \\ &\quad + \beta_1 V_0 e^{-\kappa T} R_0^{2\tilde{a}+2}(N(h_5(V_0,T)) - N(h_6(V_0,T))). \end{aligned}$$

Let us finally consider $D_3(0,T)$. We observe:

$$l(0)\mathbb{E}_{\mathbb{P}^*}(e^{(\gamma-r)\tau}\mathbb{I}_{\{\tau<T\}}) = l(0)\int_0^T e^{(\gamma-r)s}\, d\mathbb{P}^*\{\tau \leq s\},$$

where

$$\mathbb{P}^*\{\tau \leq s\} = N\Big(\frac{\ln\frac{l(0)}{V_0} - \tilde{\nu}s}{\sigma\sqrt{s}}\Big) + \Big(\frac{l(0)}{V_0}\Big)^{2\tilde{a}} N\Big(\frac{\ln\frac{l(0)}{V_0} + \tilde{\nu}s}{\sigma\sqrt{s}}\Big).$$

Note that $l(0) < V_0$ and thus $\ln \frac{l(0)}{V_0} < 0$. Now using (2.13) we obtain:

$$
\begin{aligned}
l(0) \int_0^T e^{(\gamma - r)s}\, dN\Big(\frac{\ln \frac{l(0)}{V_0} - \tilde{\nu} s}{\sigma \sqrt{s}}\Big) &= \frac{V_0(\tilde{a} + \zeta)}{2\zeta} R_0^{\theta - \zeta} N(h_8(V_0, T)) \\
&- \frac{V_0(\tilde{a} - \zeta)}{2\zeta} R_0^{\theta + \zeta} N(h_7(V_0, T))
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{l(0)^{2\tilde{a}+1}}{V_0^{2\tilde{a}}} \int_0^T e^{(\gamma - r)s}\, dN\Big(\frac{\ln \frac{l(0)}{V_0} + \tilde{\nu} s}{\sigma \sqrt{s}}\Big) &= \frac{V_0(\tilde{a} + \zeta)}{2\zeta} R_0^{\theta + \zeta} N(h_7(V_0, T)) \\
&- \frac{V_0(\tilde{a} - \zeta)}{2\zeta} R_0^{\theta - \zeta} N(h_8(V_0, T)).
\end{aligned}
$$

All together we obtain:

$$
D_3(0, T) = \beta_2 V_0 (R_0^{\theta + \zeta} N(h_7(V_0, T)) + R_0^{\theta - \zeta} N(h_8(V_0, T))).
$$

□

2.4.2 Further characteristics

Distance to default:

The distance to default is similar to the one in Merton's model, the only difference is the time-dependent barrier, that has to be taken into account.
So the formula is changed into

$$
DD_t = \frac{\ln(V_t / l(t)) + (r - \kappa - \frac{1}{2}\sigma_V^2)(T - t)}{\sigma_V \sqrt{T - t}}. \tag{2.14}
$$

The distance to default represents again the number or standard deviations, the firm value is above the default barrier.

Estimation of the firm value:

Since the Black and Cox model needs the same input parameters as the Merton model, in particular also the firm value and its volatility, we have to employ the same procedure as in the Merton model. The estimation of firm value and volatility has to be done in all firm value models, which was one of the motivations to develop alternative risk models.

2.5 Black and Cox model with stochastic interest rate

As in the Merton model, we will also incorporate interest rate risk in the Black and Cox model. Furthermore we allow the parameters κ and σ_V to be time-dependent. To maintain the analytical tractability of the model despite this extension, we have to assume that the barrier L_t can be represented as $l(t) = K\,\mathrm{B}(t,T)f(t)$ for some constant K and some function $f : [0,T] \to \mathbb{R}_+$. Furthermore the volatility of the forward firm's value $F_V(t,T) := V_t/\,\mathrm{B}(t,T)$ has to be a deterministic function.

Having the second condition in mind we will use the Gaussian Heath-Jarrow-Morton setup for the stochastic interest rates, which means assuming deterministic bond price volatilities.
Therefore, the dynamics of the firm value and the default free zero coupon bond under the martingale measure $\mathbb{P}^*$ are given by:

$$dV_t = V_t((r_t - \kappa(t))dt + \sigma_V(t)dW_t^*),$$

and

$$d\,\mathrm{B}(t,T) = \mathrm{B}(t,T)(r_t dt + b(t,T)dW_t^*),$$

where W^* is a d-dimensional standard Brownian motion under $\mathbb{P}^*$, $\kappa : [0,T] \to \mathbb{R}$ is deterministic and $\sigma_V, b : [0,T] \to \mathbb{R}^d$ are bounded deterministic functions.

As typically done in the Heath-Jarrow-Morton setup, we use the zero coupon bond as numeraire and the forward martingale measure $\mathbb{P}_T$. For further information regarding the HJM model or the change of numeraire technique, see [6].
As already mentioned above, the forward value of the firm is given by:

$$F_V(t,T) = V_t/\,\mathrm{B}(t,T).$$

Its dynamics satisfy the following equation under the forward martingale measure $\mathbb{P}^T$:

$$dF_V(t,T) = -\kappa(t)F_V(t,T)dt + F_V(t,T)(\sigma_V - b(t,T))dW_t^T,$$

with

$$W_t^T = W_t^* - \int_0^t b(u,T)\,du.$$

For any $t \leq T$ we define:

$$L_t = \begin{cases} K\,\mathrm{B}(t,T)e^{\int_t^T \kappa(u)\,du}, & \text{for } t < T, \\ L, & \text{for } t = T, \end{cases}$$

and

$$\kappa(t,T) = \int_t^T \kappa(u)\,du \ , \ \sigma(t,T) = \sqrt{\int_t^T ||\sigma_V(u) - b(u,T)||^2\,du}$$

where $||\cdot||$ is the Euclidean norm in $\mathbb{R}^d$.

For motional issues we set:

$$\eta_+(t,T) = \kappa(t,T) + \frac{1}{2}\sigma^2(t,T) \ , \ \eta_-(t,T) = \kappa(t,T) - \frac{1}{2}\sigma^2(t,T).$$

Proposition 2.5.1: Let the barrier process L_t be given as above. For any $t < T$, the forward price $F_D(t,T) = \mathrm{D}(t,T)/\,\mathrm{B}(t,T)$ of a defaultable bond equals:

$$\begin{aligned} F_D(t,T) = {} & L(N(\hat{h}_1(F_V(t,T),t,T)) - \frac{F_V(t,T)}{K}e^{-\kappa(t,T)}N(\hat{h}_2(F_V(t,T),t,T))) \\ & + \beta_1 F_V(t,T)e^{-\kappa(t,T)}(N(\hat{h}_3(F_V(t,T),t,T)) - N(\hat{h}_4(F_V(t,T),t,T))) \\ & + \beta_1 K(N(\hat{h}_5(F_V(t,T),t,T)) - N(\hat{h}_6(F_V(t,T),t,T))) \\ & + \beta_2 K J_1(F_V(t,T),t,T) + \beta_2 F_V(t,T)e^{-\kappa(t,T)}J_2(F_V(t,T),t,T), \end{aligned}$$

where

$$\begin{aligned} \hat{h}_1(F_V(t,T),t,T) &= \frac{\ln(F_V(t,T)/L) - \eta_+(t,T)}{\sigma(t,T)} \\ \hat{h}_2(F_V(t,T),t,T) &= \frac{2\ln(K) - \ln(LF_V(t,T)) + \eta_-(t,T)}{\sigma(t,T)} \\ \hat{h}_3(F_V(t,T),t,T) &= \frac{\ln(L/F_V(t,T)) + \eta_-(t,T)}{\sigma(t,T)} \\ \hat{h}_4(F_V(t,T),t,T) &= \frac{\ln(K/F_V(t,T)) + \eta_-(t,T)}{\sigma(t,T)} \\ \hat{h}_5(F_V(t,T),t,T) &= \frac{2\ln(K) - \ln(LF_V(t,T)) + \eta_+(t,T)}{\sigma(t,T)} \\ \hat{h}_6(F_V(t,T),t,T) &= \frac{\ln(K/F_V(t,T)) + \eta_+(t,T)}{\sigma(t,T)} \end{aligned}$$

$$J_1(F_t,t,T) = \int_t^T e^{\kappa(u,T)}\, dN(\frac{\ln(K/F_V(t,T)) + \eta_+(t,T)}{\sigma(t,u)})$$
$$J_2(F_t,t,T) = \int_t^T e^{\kappa(u,T)}\, dN(\frac{\ln(K/F_V(t,T)) + \eta_-(t,T)}{\sigma(t,u)}).$$

2.6 Summary

One big advantage of structural models is, that the parameters in the models are linked to economic fundamentals. Therefore the influence of the variables on the result are quite clear and comprehensible. For example, by increasing the firm's volatility, it is quite clear, that the probability of default will rise too.

One of the main drawbacks is, as mentioned above, the non-observability of some input parameters, although there are ways to bypass this problem. But even when we assume, that the determination is correct, it is still questionable, if the equity market is a reliable source of information on the credit quality of a company.
Another drawback is the so called "predictability of default". Since default is triggered by the firm value and a predefined barrier, we know at time t the distance to default, and if this distance of default is large, the firm value would have to fall very fast to trigger the default event. However, in the considered models the firm value is continuous and hence, a sharp fall is implausible, implying that the short-term probability of default is close to zero. With the same considerations, the recovery rate can also be seen as a predictable variable, since when the firm defaults, the recovery rate is the result of the remaining firm value.

These drawbacks have led to further extensions of the Black and Cox model. For example, Duffie and Lando (2001), Giesecke (2005) and Jarrow and Protter (2004) assume, that the investors only infer a distribution function of the firm value or the barrier. Another approach, for example by Zhou (2001), is to allow for jumps of the firm value process. With these extensions, sudden default is possible and therefore the predictability of default is eliminated.
For further information on structural models see [1], [2], [4] and [18].

Chapter 3

Intensity Models

3.1 Introduction

In this chapter we will discuss the second approach for modeling default risk, namely intensity models. The framework of this chapter is based on [6].
These models are totally different from the structural models we have considered in Chapter 2, since default is not triggered by market information. Instead intensity models (or reduced form models) describe default by means of an exogenously given jump process. More precisely, the default time τ is the first jump time of a stochastic process, in the most simple setting the Poisson process. So the default time is closely linked to the intensity λ. As pointed out above, default has an exogenous component, that is independent of all the default free market data. This means that monitoring the default free market does not give complete information about the default process and there is no direct link to the economic parameters.
Intensity models are typically fitted to credit default spreads and a big advantage of intensity models is, that they are in general easier to calibrate to credit default swap or corporate bond data.

As mentioned above default is triggered by the first jump of a Poisson/Cox process. According to Section 1.3.5 the survival probability $\mathbb{P}(\tau > t)$ is given by

$$\begin{aligned}\mathbb{P}(\tau \geq t) &= \mathbb{P}(\Lambda(\tau) \geq \Lambda(t)) \\ &= \mathbb{P}(\xi \geq \int_0^t \lambda_u \, du) \\ &= \mathbb{E}(\mathbb{P}(\xi \geq \int_0^t \lambda_u \, du | \mathbb{F}_t^\lambda)) \\ &= \mathbb{E}(e^{-\int_0^t \lambda_u \, du}).\end{aligned}$$

In case of deterministic intensities λ_t or even constant intensity λ, this formula simplifies to

$$\mathbb{P}(\tau > t) = e^{-\int_0^t \lambda_u \, du},$$

respectively

$$\mathbb{P}(\tau > t) = e^{-\lambda t}.$$

Notice, that the survival probabilities $\mathbb{P}(\tau > t)$ have the same structure as discount factors. In particular the formulas above are equivalent to the price of a zero coupon bond with stochastic short rate r_t

$$B(0,t) = \mathbb{E}(e^{-\int_0^t r_u \, du}),$$

deterministic short rate r_t

$$B(0,t) = e^{-\int_0^t r(u) \, du}$$

or constant short rate r

$$B(0,t) = e^{-rt}.$$

Due to the close connection of intensity models to short rate models, it is natural to use short rate models to model the hazard rate λ_t.

3.2 A deterministic intensity model

In this first example we will see, how one can fit an intensity model to given credit spreads.
As a first example, we will assume that the hazard rate λ is deterministic and piecewise constant, i.e. $\lambda_t = \lambda_i$ for t $\in [K_i, K_{i+1}]$, where K_i, for $i = 1, \dots, n$, are the maturities of the CDS contracts observable on the market. This means, that the cumulative hazard function is

$$\Lambda(t) = \int_0^t \lambda_u \, du = \sum_{i=1}^{k(t)} (K_{i+1} - K_i)\lambda_i + (t - K_{k(t)-1})\lambda_{k(t)} \qquad (3.1)$$

where $k(t)$ is the index of the first K_i following t. If $t = K_j$ for some $j = 1, \dots, n$ then

$$\Lambda_j = \Lambda(K_j) = \sum_{i=1}^{j} (K_i - K_{i-1})\lambda_i. \qquad (3.2)$$

Let us shortly recapitulate the cash flows in a CDS contract. The buyer of the CDS gets a predefined value, when a third party defaults, usually the LGD.
In return for this insurance she pays the seller of the CDS a constant rate R at predefined dates $T_1, ...T_n$.

The sellers expected payout can thus be written as:

$$\begin{aligned} LGD \cdot \mathbb{E}(\text{Disc}(0,\tau)\mathbb{I}_{\{T_0<\tau<T_n\}}|\mathbb{F}) &= LGD \int_{T_0}^{T_n} \mathbb{E}(\text{Disc}(0,u))\mathbb{P}(\tau \in [u, u+du]) \, du \\ &= LGD \int_{T_0}^{T_n} B(0,u)\lambda_u e^{-\int_0^u \lambda_s \, ds} \, du \\ &= LGD \sum_{i=1}^{n} \lambda_i \int_{T_i-1}^{T_i} e^{-\Lambda_{i-1} - \lambda_{u-T_{i-1}}} B(0,u) \, du \end{aligned}$$

where $\text{Disc}(0,T)$ is the discount factor from T to now.

Remark: It is important to point out, that this derivation only holds, if we assume independence of the intensity rate λ and the non-defaultable zero coupon bond price, respectively the short rate r.

The buyers expected payout is similar, but she receives a fixed rate R at predefined dates $T_1, ...T_n$ or until default. In case of default, she receives the rate R from the last payment date T_i until default at time τ. This payment is called the "accrued interest".
So the cash flow can be written as:

$$\begin{aligned}
&\mathbb{E}\Big(\sum_{i=1}^{n} R \cdot \mathbb{I}_{\{T_i<\tau\}}|\mathbb{F}\Big) + R \cdot \mathbb{E}\Big(\mathrm{Disc}(0,\tau)(\tau - T_{k(\tau)-1})\mathbb{I}_{\{T_0<\tau<T_n\}}|\mathbb{F}\Big)\\
=&\sum_{i=1}^{n} R\alpha_i \mathbb{E}\Big(\mathrm{Disc}(0,T_i)|\mathbb{F}\Big)\mathbb{P}(T_i<\tau) + R\int_{T_0}^{T_n} \mathbb{E}\Big(\mathrm{Disc}(0,u)(u-T_{k(u)-1})\Big)\mathbb{P}(\tau\in[u,u+du])\,du\\
=&R\sum_{i=1}^{n} B(0,T_i)\alpha_i e^{-\Lambda(T_i)} + R\int_{T_0}^{T_n} B(0,u)(u-T_{k(u)-1})\lambda(u)e^{-\int_0^u \lambda(s)\,ds}\,du\\
=&R\sum_{i=1}^{n} B(0,T_i)\alpha_i e^{-\Lambda(T_i)} + R\sum_{i=1}^{n}\lambda_i\int_{T_i-1}^{T_i} e^{-\Lambda_{i-1}-\lambda(u-T_{i-1})}B(0,u)(u-T_{i-1})\,du
\end{aligned}$$

where α_i is the time span from T_{i-1} to T_i.

All together we get a formula for the CDS price:

$$\begin{aligned}
CDS_{T_0,T_n}(0,R,LGD,\Lambda(\cdot)) = {}& R\sum_{i=1}^{n}\lambda_i\int_{T_i-1}^{T_i} e^{-\Lambda_{i-1}-\lambda_i(u-T_{i-1})}B(0,u)(u-T_{i-1})\,du\\
&+R\sum_{i=1}^{n} B(0,T_i)\alpha_i e^{-\Lambda(T_i)}\\
&-LGD\sum_{i=1}^{n}\lambda_i\int_{T_i-1}^{T_i} e^{-\Lambda_{i-1}-\lambda(u-T_{i-1})}B(0,u)\,du.
\end{aligned} \tag{3.3}$$

In the next section we will see, how we can use this result, to derive the implied hazard rate from the actual CDS-prices.

3.2.1 Numeric example

Let us have a look at a numeric example. To use the intensity model, derived in the previous section, we need the CDS spreads for some time steps K_i and the appropriate interest rate curve. So let us assume a company with the following spreads:

Maturity T_n	CDS Spread in BP
1Y	100
3Y	110
5Y	115
7Y	118
10Y	120

Table 3.1: CDS-Spreads

As common in market, the CDS spreads are given in Basis Points (BP), which is 1/100 of a percent. Furthermore we assume annual payments and a constant recovery rate, which is set to 40%. The interest rate, used in our calculations, is given in Figure 3.1.

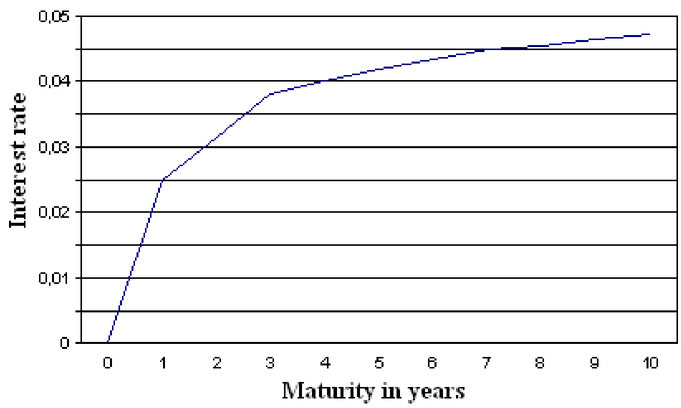

Figure 3.1: Interest rate curve

Now we can use formula (3.3) to solve the equations $CDS_{K_0,K_i}(0, R_i, LGD, \Lambda(\cdot)) = 0$ with the given spread R_i and $K_1, ...K_n$ = 1,3,5,7,10 for $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ and λ_5. The results for piecewise constant intensities are shown in Figure 3.2. In addition we have also solved the equations for piecewise linear intensities, these results are given in Figure 3.3.

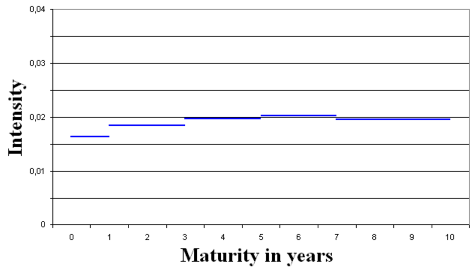

Figure 3.2: Piecewise constant intensity

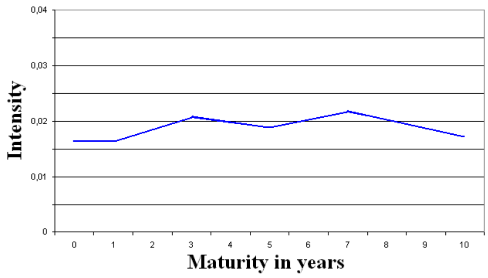

Figure 3.3: Piecewise linear intensity

For this standard company the results are quite satisfying. But let us now consider a more stressed company, for which the CDS-spreads are given in Table 3.2. This example is taken from [6], it is real market data and coming from Parmalat, December 2003:

Maturity T_n	CDS Spread in BP
1Y	5050
3Y	2100
5Y	1500
7Y	1250
10Y	1100

Table 3.2: CDS-Spreads of a Parmalat 2003

Given the spreads of Table 3.2 and a recovery rate of 15%, we obtain the results, shown in Figure 3.4 and 3.5.

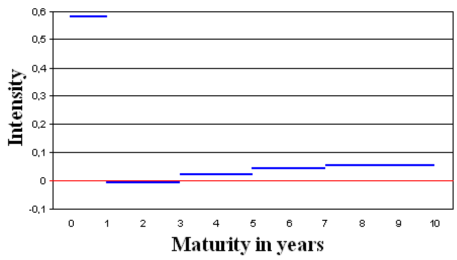

Figure 3.4: Piecewise constant intensity

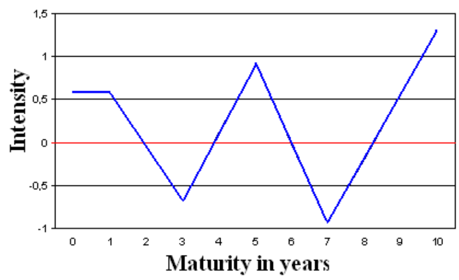

Figure 3.5: Piecewise linear intensity

As we can see, the calibration results in negative hazard rates, which is not consistent with the models assumptions.
By comparing the two choices of intensities, we see, that the model with piecewise constant intensity is more robust against negative intensities. One could have expected this result, since the obtained constant intensity is an average intensity over its time span, so extreme values are compensated. By introducing linear intensities, this behavior is lost.
We have seen, that in very distressed situations even the constant intensity model results in negative intensities. But according to [6], this only happens in a pathological situation and therefore it is a clear signal that one should pay attention to this firm.

3.3 A stochastic intensity model

For deterministic intensities, we have derived a valuation formula for the price of a CDS, or respectively a numerical method to bootstrap the hazard rates for given CDS spreads R_i. But if we want to allow the model for stochastic volatilities, this leads to a Cox process.
We already noted, that the survival probability is closely linked to the prize of a zero coupon bond and, therefore, that is reasonable to use a short rate model for the stochastic intensity. In the sequel we will use a particular model, namely the CIR++ model, which is introduced in the following section.

3.3.1 The CIR model

Before we can discuss the CIR++ model, we first need to introduce the well known CIR(Cox-Ingersoll-Ross)-model [7]. Since the CIR model is a short rate model, we will derive the model for the short rate r, but keeping in mind, that we will use it to model the default intensities λ.
The dynamics of the short rate r are given by:

$$dr(t) = k(\theta - r(t))dt + \sigma\sqrt{r(t)}dW_t,$$

with $r(0) = r_0$ and r_0, k, θ and σ positive constants and $2k\theta > \sigma^2$. This condition assures, that r remains positive.

Here r follows for all $t > 0$ a noncentral χ^2-distribution., which means that the density function of r is given by:

$$p_{r(t)}(x) = p_{\chi^2(v,\xi_t)/c_t}(x) = c_t p_{\chi^2(v,\xi_t)}(c_t x),$$

with

$$\begin{aligned} c_t &= \frac{4k}{\sigma^2(1-\exp(-kt))} \\ v &= \frac{4k\theta}{\sigma^2} \\ \xi_t &= c_t r_0 \exp(-kt) \end{aligned}$$

The density of the (noncentral) χ-squared distribution is given by:

$$p_{\chi^2(v,\xi)}(z) = \sum_{i=0}^{\infty} \frac{e^{-\xi/2}(\xi/2)^i}{i!} p_{\chi^2(v+2i)}(z),$$

with $p_{\chi^2(v)}(z)$ being the density of a central χ-squared distribution.

Modeling the intensity λ by a short rate model, the default probability is given by the price of a zero coupon bond.
In the CIR model, the price of a zero coupon bond with maturity T at time t is

$$\mathrm{B}(t,T) = \zeta_1(t,T)e^{-\zeta_2(t,T)r(t)}, \tag{3.4}$$

where

$$\zeta_1(t,T) = \Big(\frac{2h\exp(k+h)(T-t)/2}{2h+(k+h)(\exp((T-t)h)-1)}\Big)^{2k\theta/\sigma^2}, \tag{3.5}$$

$$\zeta_2(t,T) = \frac{2(\exp((T-t)h)-1)}{2h+(k+h)(\exp((T-t)h)-1)}, \tag{3.6}$$

$$h = \sqrt{k^2+2\sigma^2}.$$

3.3.2 The CIR++ model

The CIR++ model is a generalization of the CIR model to enable exact fitting on market data. Therefore the model is modified the following way:

$$dx(t) = k(\theta - x(t))dt + \sigma\sqrt{x(t)}dW_t, \text{ with } x_0 = 0 \tag{3.7}$$

$$r(t) = x(t) + \phi(t),$$

where $\phi(t)$ is a non-negative function and x_0, k, θ and σ are positive constants such that $2k\theta > \sigma^2$.

The advantage of this model is, that the analytic tractability of the initial model is maintained, while by introducing the function $\phi(t)$ any observed term structure can be fitted exactly.

For simplicity we define the parameter vector $\alpha = (k, \theta, \sigma)$. When we assume exact fitting of the initial discount factors given by the markets zero coupon prices $B^{mkt}(0, T)$, we have:

$$\phi(t) = \phi^{CIR}(t, \alpha),$$

where

$$\begin{aligned}\phi^{CIR}(t, \alpha) &= f^{mkt}(0, t) - f^{CIR}(0, t; \alpha),\\ f^{CIR}(0, t; \alpha) &= \frac{2k\theta(exp(th) - 1)}{2h + (k + h)(exp(th) - 1)} + x_0 \frac{4h^2 exp(th)}{[2h + (k + h)(exp(th) - 1)]^2}\end{aligned}$$

with $h = \sqrt{k^2 + 2\sigma^2}$, $f^{mkt}(0, t)$ being the instantaneous interest rate with maturity t available on the market and $f^{CIR}(0, t; \alpha)$ the interest rate given by the CIR model.

The price of a zero coupon bond with maturity T in this model is given by

$$\mathrm{B}(t, T) = \tilde{\zeta}_1(t, T) e^{-\zeta_2(t,T) r(t)}, \tag{3.8}$$

with

$$\tilde{\zeta}_1(t, T) = \frac{B^{mkt}(0, T)\zeta_1(0, t) \exp(-\zeta_2(0, t)x_0)}{B^{mkt}(0, t)\zeta_1(0, T) \exp(-\zeta_2(0, T)x_0)} \zeta_1(t, T) e^{\zeta_2(t,T)\phi^{CIR}(t,\alpha)},$$

with $\zeta_1(t, T)$ and $\zeta_2(t, T)$ be given by (3.5) and (3.6).

For further information on the CIR/CIR++ or interest rate models in general I refer to [6].

3.3.3 Setup of the model

Modeling the default intensity λ by the CIR++ model we define

$$\lambda_t = x_t^\alpha + \psi(t, \alpha), \text{ with } t \geq 0,$$

where ψ is a deterministic function, depending on the parameter vector $\alpha = (\kappa, \theta, \sigma, x_0^\alpha)$, with $\kappa, \theta, \sigma, x_0^\alpha$ being positive deterministic constants. We are free to set the starting parameter x_0^α as long as:

$$\psi(0, \alpha) = \lambda_0^{mkt} - x_0^\alpha.$$

The process x is defined by (3.7) and to ensure the positivity of λ, we set the condition:

$$2\kappa\theta > \sigma^2.$$

Calibration:

In Section 3.2, we have derived the CDS price for piecewise constant intensities. A more general formula is given by

$$\begin{aligned} CDS_{a,b}(0, R, LGD, \Lambda(\cdot)) = R[&- \int_{T_a}^{T_b} B(0,t)(t - T_{k(t)-1})\, d\mathbb{P}(\tau \geq t)] \\ &+ \sum_{i=a+1}^{b} B(0, T_i)\alpha_i \mathbb{P}(\tau \geq T_i) \\ &+ LGD \int_{T_a}^{T_b} B(0,t)\, d\mathbb{P}(\tau \geq t), \end{aligned} \tag{3.9}$$

where α_i is the time span between T_{i-1} and T_i and τ is the default time, respectively the first jump of the underlying Cox process.

Brigo and Mercurio [6] suggest to calibrate the model the following way: First, according to Section 3.2, solve the $CDS_{T_0,T_i}(0, R_i, LGD, \Lambda(\cdot)) = 0$ with the given spread R_i and T_i for λ_i. Therefore we are free to use linear or constant intensities. These intensities are now seen as the market intensities λ^{mkt}. Next, we use the following derivation to calibrate the model.

The survival probabilities in the CIR++ model equals:

$$\mathbb{P}(\tau > t)_{model} = \mathbb{E}(e^{-\Lambda(t)}) = \mathbb{E}(\exp(-\Psi(t, \alpha) - X^\alpha(t)),$$

with

$$\Psi(t,\alpha) = \int_0^t \psi(s,\alpha)ds \ , \ X^\alpha(t) = \int_0^t x_s^\alpha ds.$$

So we have to make sure that:

$$\mathbb{E}(\exp(-\Psi(t,\alpha) - X^\alpha(t)) = \mathbb{P}(\tau > t)_{mkt} = e^{-\Lambda^{mkt}(t)}.$$

To obtain exact fitting of the initial market default intensities, we have

$$\Psi(t,\alpha) = \Lambda^{mkt}(t) + \ln(\mathbb{E}(e^{-X^\alpha(t)})) = \Lambda^{mkt}(t) + \ln(B^{CIR}(0,t,x_0,\alpha)).$$

On the one hand, we need to find α subject to a non-negative ψ, or in other words non-negative and increasing Ψ, on the other hand we want the contain the departure of λ from its time-homogeneous component x_t^α. Therefore we find α by $\min \int_0^T \psi(s,\alpha)^2 ds \quad \text{subject to} \quad \psi(s,\alpha) \geq 0 \forall s.$

If we still assume independence of the short rate and the default intensities, we obtain the CDS-price at time 0 by

$$\begin{aligned} CDS_{T_0,T_n}(0,R,LGD,\Lambda(\cdot)) = {} & R\int_{T_a}^{T_b} B(0,u)(T_{k(u)-1} - u)\lambda(u)e^{-\Lambda(u)}du \\ & + R\sum_{i=a+1}^{b} B(0,T_i)\alpha_i e^{-\Lambda(T_i)} \\ & - LGD\int_{T_a}^{T_b} B(0,u)\lambda(u)e^{-\Lambda(u)}du. \end{aligned}$$

By construction, this price will exactly fit the market price.

3.4 The SSRD model

3.4.1 Dependence between interest rate and default probability

Let us have a look on the price of a defaultable bond:

$$\begin{aligned}
\mathrm{D}(0,T) &= \mathbb{E}(\mathrm{Disc}(0,T)\mathbb{I}_{\{\tau>T\}}) = \mathbb{E}(\mathrm{Disc}(0,T)\mathbb{I}_{\{\Lambda(\tau)>\Lambda(T)\}}) \\
&= \mathbb{E}(\mathbb{E}(\mathrm{Disc}(0,T)\mathbb{I}_{\{\xi>\Lambda(T)\}}|\mathcal{F}_t)) = \mathbb{E}(\mathrm{Disc}(0,T)\mathbb{E}(\mathbb{I}_{\{\xi>\Lambda(T)\}}|\mathcal{F}_t)) \\
&= \mathbb{E}(\mathrm{Disc}(0,T)\,\mathrm{Exp}(-\Lambda(T))) = \mathbb{E}(e^{-\int_0^T (r_u+\lambda_u)\,du})
\end{aligned}$$

with $\mathrm{Disc}(0,\cdot)$ being the discount factor for time T.

It is obvious, that for independent short rate r and intensity λ, the formula can be simplified as:

$$\mathrm{D}(0,T) = B(0,T)\mathbb{P}(\tau > T). \tag{3.10}$$

Up to now, we only have discussed intensity models, where independence of r and the probability of default is assumed. We will now introduce the SSRD (=shifted square root diffusion) model which allows to take dependence into account.

We assume, that the short rate r follows a CIR++ model, while default is triggered by a Cox process.

As [2] have shown, these two processes are independent, if they are defined on the same probability space. Therefore the short rate r and the default time τ are independent as long as λ_t and r_t are independent.

So dependence can only be incorporated in this setup by correlating the interest rate r and the intensities λ_t.

3.4.2 Setup of the model

Short rate model:

We assume the short rate r to follow the CIR++ model. We have

$$dx_t^\alpha = k(\theta - x_t^\alpha)dt + \sigma\sqrt{x_t^\alpha}dW_t$$
$$r_t = x_t^\alpha + \phi(t,\alpha), \text{ with } t \geq 0,$$

with starting condition

$$\phi(0,\alpha) = r_0^{mkt} - x_0^\alpha.$$

As we have already pointed out in Section 3.3.2, the closed form formula for the the price of a zero coupon bond with maturity T is:

$$B(t,T) = \frac{B^{mkt}(0,T)\zeta_1(0,t;\alpha)\exp(-\zeta_2(0,t;\alpha)x_0)}{B^{mkt}(0,t)\zeta_1(0,T;\alpha)\exp(-\zeta_2(0,T;\alpha)x_0)} B^{CIR}(t,T,r_t - \phi^{CIR}(t,\alpha);\alpha),$$

where

$$B^{CIR}(t,T,x_t;\alpha) = \mathbb{E}_t(e^{-\int_t^T x^\alpha(u)\,du}) = \zeta_1(t,T;\alpha)\exp(-\zeta_2(t,T;\alpha)x_t)$$

is the bond price formula of the CIR model.

Intensity model:

We also assume the intensities to follow the CIR++ model. Therefore the intensities λ are given by

$$\lambda_t = y_t^\beta + \psi(t,\beta), \text{ with } t \geq 0,$$

where ψ is a deterministic function, depending on the parameter vector $\beta = (\kappa,\mu,\nu,y_0^\beta)$, with κ,μ,ν,y_0^β being positive deterministic constants.
As above, we are free to select the value of y_0^β as long as:

$$\psi(0,\beta) = \lambda_0^{mkt} - y_0.$$

The process y is defined by:

$$dy_t^\beta = \kappa(\mu - y_t^\beta)dt + \nu\sqrt{y_t^\beta}dZ_t,$$

with Z being a standard Brownian motion and to ensure the positivity of λ, we set the condition:

$$2\kappa\mu > \nu^2.$$

Since our motivation was, to introduce dependence between the interest rate and the intensities, we assume the Brownian motions W and Z to have a constant correlation ρ.

3.4.3 Calibration of the SSRD model

Recall that the survival probability in the intensity model is given by:

$$\mathbb{P}(\tau \geq t) = \mathbb{P}(\Lambda(\tau) \geq \Lambda(t)) = \mathbb{E}(e^{-\int_0^t \lambda_u \, du}).$$

To calibrate the model, we need a CDS-formula for dependent r and λ. A general formula is given by:

$$\begin{aligned} CDS_{a,b}(t, R, \text{LGD}) = \mathbb{I}_{\{\tau > t\}} \mathbb{E}\Big(& \text{Disc}(t,\tau)(\tau - T_{k(\tau)-1}) R \mathbb{I}_{\{T_a < \tau < T_b\}} \\ & + \sum_{i=a+1}^{b} \text{Disc}(t, T_i)\alpha_i R \mathbb{I}_{\{\tau > T_i\}} - \mathbb{I}_{\{T_a < \tau < T_b\}} \text{Disc}(t,\tau)\, \text{LGD}\, | \mathcal{G}_t \Big) \end{aligned}$$

where $\mathbb{G} = (\mathcal{G}_t)_{t\geq 0}$, $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$, $\mathcal{G}_t = \mathcal{F}_t \vee \sigma(\tau < u, u \leq t)$ and $\mathbb{F}_t$ is the information on the default free market up to time t.

To calibrate the model to actual CDS-rates, we would have to solve the following equations for α, β and ρ:

$$CDS_{0,b}^{SSRD}(0, R_{0,b}^{mkt}, \alpha, \beta, \phi(\cdot, \alpha), \psi(\cdot, \beta), \rho) = 0.$$

For a consistent calibration of the model, we would have to include also interest rate derivatives, like zero coupon bonds, swaps or caps, and fit the parameters of the interest rate model and the intensity model at the same time. This method, however, would be very time intense and hence a different approach based on some approximation might be a good alternative. Therefore we need the following proposition.

Proposition 3.4.1: ([2]) Filtration switching formula
Let $\mathcal{F}_t \subseteq \mathcal{G}_t$, then for any $\mathbb{G}$-measurable random variable Y and $t \in \mathbb{R}$ the following formula holds.

$$\mathbb{E}(\mathbb{I}_{\{\tau>T\}}Y|\mathcal{G}_t) = \frac{\mathbb{I}_{\{\tau>t\}}}{\mathbb{P}(\tau > t|\mathcal{F}_t)}\mathbb{E}(\mathbb{I}_{\{\tau>T\}}Y|\mathcal{F}_t) \text{ , for } t < T.$$

Let us now have a closer look on the first term of the CDS-formula.

$$\begin{aligned}
&\mathbb{I}_{\{\tau>t\}}\mathbb{E}\Bigg(\text{Disc}(t,\tau)(\tau - T_{k(\tau)-1})R\, \mathbb{I}_{\{T_a<\tau<T_b\}}|\mathcal{G}_t\Bigg) \\
&= \frac{\mathbb{I}_{\{\tau>t\}}}{\mathbb{P}(\tau > t|\mathcal{F}_t)}\mathbb{E}\Bigg(\text{Disc}(t,\tau)(\tau - T_{k(\tau)-1})R\, \mathbb{I}_{\{T_a<\tau<T_b\}}|\mathcal{F}_t\Bigg) \\
&= \frac{\mathbb{I}_{\{\tau>t\}}}{\exp(-\int_0^t \lambda_s\, ds)}\mathbb{E}\Bigg(\int_t^\infty \text{Disc}(t,s)(s - T_{k(s)-1})R\, \mathbb{I}_{\{T_a<s<T_b\}}\mathbb{I}_{\{\tau\in[s,s+ds]\}}|\mathcal{F}_t\Bigg) \\
&= \frac{\mathbb{I}_{\{\tau>t\}}}{\exp(-\int_0^t \lambda_s\, ds)}\mathbb{E}\Bigg(\mathbb{E}\Bigg(\int_{T_a}^{T_b} \text{Disc}(t,s)(s - T_{k(s)-1})R\, \mathbb{I}_{\{\tau\in[s,s+ds]\}}|\mathcal{F}_{T_b}\Bigg)|\mathcal{F}_t\Bigg) \\
&= \frac{\mathbb{I}_{\{\tau>t\}}}{\exp(-\int_0^t \lambda_s\, ds)}\mathbb{E}\Bigg(\int_{T_a}^{T_b} \text{Disc}(t,s)(s - T_{k(s)-1})R\, \mathbb{P}\Bigg(\tau \in [s, s+ds]|\mathcal{F}_{T_b}\Bigg)|\mathcal{F}_t\Bigg) \\
&= \frac{\mathbb{I}_{\{\tau>t\}}}{\exp(-\int_0^t \lambda_s\, ds)}\mathbb{E}\Bigg(\int_{T_a}^{T_b} \text{Disc}(t,s)(s - T_{k(s)-1})R\, \exp(-\int_0^s \lambda_u\, du)\lambda_s\, ds|\mathcal{F}_t\Bigg) \\
&= \mathbb{I}_{\{\tau>t\}}\mathbb{E}\Bigg(\int_{T_a}^{T_b} \text{Disc}(t,s)(s - T_{k(s)-1})R\, \exp(-\int_t^s \lambda_u\, du)\lambda_s\, ds|\mathcal{F}_t\Bigg).
\end{aligned}$$

Similar calculations can be carried out for the second and third term and we obtain the following general CDS-formula, that holds for dependent r and λ.

$$CDS_{a,b}(t,R,\text{LGD}) = \mathbb{I}_{\{\tau>t\}}\Big(R\sum_{i=a+1}^{b}\alpha_i\mathbb{E}(\exp(-\int_t^{T_i}(r_u+\lambda_u)\,du)|\mathbb{F}_t) \quad (3.11)$$
$$+ R\int_{T_a}^{T_b}\mathbb{E}(\exp(-\int_t^{u}(r_s+\lambda_s)\,ds)\lambda_u|\mathbb{F}_t)(u-T_{k(u)-1})\,du$$
$$- \text{LGD}\int_{T_a}^{T_b}\mathbb{E}(\exp(-\int_t^{u}(r_s+\lambda_s)\,ds)\lambda_u|\mathbb{F}_t)du\Big).$$

To bypass the joint calibration of the short rate and the intensity model, we will assume $\rho = 0$. Let us have a look on the following expectations:

$$\mathbb{E}(\exp(-\int_t^{T_i}(r_u+\lambda_u)\,du)|\mathbb{F}_t) = \mathbb{E}(\exp(-\int_t^{T_i}r_u\,du)|\mathbb{F}_t)\mathbb{E}(\exp(-\int_t^{T_i}\lambda_u\,du)|\mathbb{F}_t) \quad (3.12)$$
$$= \exp(\Psi(t,\beta)-\Psi(T_i,\beta))B^{CIR}(t,T_i,y_t,\beta)$$
$$\times \exp(\Phi(t,\alpha)-\Phi(T_i,\alpha))B^{CIR}(t,T_i,x_t,\alpha)$$

where B^{CIR} is the zero coupon bond price in the CIR model given by formula (3.4).

$$\mathbb{E}(\exp(-\int_t^{u}(r_s+\lambda_s)\,ds)\lambda_u|\mathbb{F}_t) = \mathbb{E}(\exp(-\int_t^{u}r_s\,ds)|\mathbb{F}_t)\mathbb{E}(\exp(-\int_t^{u}\lambda_s\,ds)\lambda_u|\mathbb{F}_t)$$
$$= \mathbb{E}(\exp(-\int_t^{u}r_s\,ds)|\mathbb{F}_t)(-\frac{d}{du}\mathbb{E}(\exp(-\int_t^{u}\lambda_s\,ds)|\mathbb{F}_t)) \quad (3.13)$$
$$= -\exp(\Phi(t,\alpha)-\Phi(u,\alpha))B^{CIR}(t,u,x_t,\alpha)$$
$$\times \frac{d}{du}\exp(\Psi(t,\beta)-\Psi(u,\beta))B^{CIR}(t,u,y_t,\beta).$$

Since we have obtained Φ and Ψ as the difference of x_t^α and y_t^β to the market data, (3.12) and (3.13) can be reduced to

$$B^{mkt}(t,T_i)(e^{-(\Lambda(T_i)-\Lambda(t))}) \quad \text{and} \quad B^{mkt}(t,u)\lambda^{mkt}(u)(e^{-(\Lambda(u)-\Lambda(t))}).$$

Using these results and by setting $t = 0$, we can simplify formula (3.11) to

$$\begin{aligned} CDS_{a,b}(t, R, \text{LGD}) = R &\int_{T_a}^{T_b} \text{B}(0,u)(T_{k(u)-1} - u)\lambda(u)e^{-(\Lambda(u))}\, du \\ &+ R \sum_{i=a+1}^{b} \text{B}(0,T_i)\alpha_i e^{-\Lambda(T_i)} \\ &- \text{LGD} \int_{T_a}^{T_b} \text{B}(0,u)\lambda(u)e^{-(\Lambda(u))}\, du. \end{aligned} \tag{3.14}$$

So for calibration of the SSRD model we have to take the following steps. First, we separately calibrate the short rate model to actual cap prices and the intensity model to CDS spreads. Then we have determined all input parameters except ρ, which is then set on a specific value, based on historical data or market view.

It is important to point out, that for an exact calibration we would have to simultaneously calibrate the interest rate model and the intensity model. But as shown in [6], for pricing credit default swaps with the SSRD model, the correlation ρ is negligible. Therefore the presented workaround is feasible and will fit to market data.
The obtained model is able to replicate the actual CDS- and cap prices, but also keeps the feature of dependence between the interest rate r and the intensity λ, which is needed for more complicated products.

3.4.4 Simulation in the SSRD model

When we have $\rho = 0$, we have shown that the two independent CIR-processes have non-central χ^2 distributed marginals and most prices are given in closed form. When dependence is introduced, it is not possible, to characterize the joint distribution of r and λ and therefore numerical methods have to be used.
Typically this is made by applying a discretization scheme for the stochastic differential equations and then to simulate the Gaussian increments corresponding to the joint Brownian motion.
We will now discuss a few methods, that can be used in the SSRD model with $\rho \neq 0$.

We will discuss three different approaches, namely the Euler scheme, the implicit Euler scheme and the Gaussian dependence mapping.

Euler scheme in different approaches:

The Euler scheme is the easiest discretization scheme. We split the interval $[0, T]$ into $0 = t_0 < t_1 < \cdots < t_n = T$. Next, the Brownian motion Z (of the intensity process) can be represented by means of a third Brownian motion V, i.e. $Z_t = \rho W_t + \sqrt{1-\rho^2} V_t$, with V being independent of W.
Given this setup, the increments of W and Z in the interval $[t_i, t_{i+1}]$ can be simulated by the increments of the independent Brownian motions W and V.
We obtain:

$$\begin{aligned} x^\alpha_{t_{i+1}} &= x^\alpha_{t_i} + k(\theta - x^\alpha_{t_i})(t_{i+1} - t_i) + \sigma\sqrt{x^\alpha_{t_i}}(W_{t_{i+1}} - W_{t_i}) \\ y^\beta_{t_{i+1}} &= y^\beta_{t_i} + \kappa(\mu - y^\beta_{t_i})(t_{i+1} - t_i) + \nu\sqrt{y^\beta_{t_i}}(Z_{t_{i+1}} - Z_{t_i}). \end{aligned}$$

The drawback of this simplifying approach is, that we do not ensure positivity and since $x^\alpha_{t_{i+1}}$ contains the square root of $x^\alpha_{t_i}$, the process is not well defined in general.
There are a few approaches, to bypass the problem. One possibility is proposed by [8]. $x^\alpha_{t_{i+1}}$ is defined by

$$x^\alpha_{t_{i+1}} = x^\alpha_{t_i} + k(\theta - x^\alpha_{t_i})(t_{i+1} - t_i) + \sigma\sqrt{x^\alpha_{t_i}\mathbb{I}_{\{x^\alpha_{t_i} > 0\}}}(W_{t_{i+1}} - W_{t_i}).$$

Here, positivity is not ensured, but since we only take the square root in case of positivity, the process is well defined.
Another approach, according to [10], is to take the absolute value of $x^\alpha_{t_{i+1}}$:

$$x^\alpha_{t_{i+1}} = |x^\alpha_{t_i} + k(\theta - x^\alpha_{t_i})(t_{i+1} - t_i) + \sigma\sqrt{x^\alpha_{t_i}}(W_{t_{i+1}} - W_{t_i})|.$$

Implicit Euler scheme:

For CIR processes it is possible to obtain ad-hoc schemes.
Let us therefore have a look on the diffusion process of the square root of x_t^α.

$$d\sqrt{x_t^\alpha} = \frac{k\theta - \frac{\sigma^2}{4}}{2\sqrt{x_t^\alpha}}dt - \frac{k}{2}\sqrt{x_t^\alpha}dt + \frac{\sigma}{2}dW_t \tag{3.15}$$

It follows, that $d\sqrt{x_t^\alpha}dW_t = \sigma dt/2$. We will use this result in the next calculations.

When $\max(t_{i+1} - t_i, 0 \le i \le n) \to 0$, we can transform x_t^α as follows

$$\begin{aligned}
x_t^\alpha &= x_0^\alpha + \int_0^t k(\theta - x_s^\alpha)\, ds + \sigma \int_0^t \sqrt{x_s^\alpha}\, dW_s \\
&= \lim\Big[x_0^\alpha + \sum_{i;t_i<t} k(\theta - x_{t_i}^\alpha)(t_{i+1} - t_i) + \sigma \sum_{i;t_i<t} \sqrt{x_{t_i}^\alpha}(W_{t_{i+1}} - W_{t_i})\Big] \\
&= \lim\Big[x_0^\alpha + \sum_{i;t_i<t} k(\theta - x_{t_{i+1}}^\alpha)(t_{i+1} - t_i) + \sigma \sum_{i;t_i<t} \sqrt{x_{t_{i+1}}^\alpha}(W_{t_{i+1}} - W_{t_i}) \\
&- \sigma \sum_{i;t_i<t} (\sqrt{x_{t_{i+1}}^\alpha} - \sqrt{x_{t_i}^\alpha})(W_{t_{i+1}} - W_{t_i})\Big] \\
&= \lim\Big[x_0^\alpha + \sum_{i;t_i<t} (k\theta - kx_{t_{i+1}}^\alpha)(t_{i+1} - t_i) + \sigma \sum_{i;t_i<t} \sqrt{x_{t_{i+1}}^\alpha}(W_{t_{i+1}} - W_{t_i})\Big] \\
&- \lim \sum_{i;t_i<t} \frac{\sigma^2}{2}(t_{i+1} - t_i) \\
&= \lim\Big[x_0^\alpha + \sum_{i;t_i<t} (k\theta - \frac{\sigma^2}{2} - kx_{t_{i+1}}^\alpha)(t_{i+1} - t_i) + \sigma \sum_{i;t_i<t} \sqrt{x_{t_{i+1}}^\alpha}(W_{t_{i+1}} - W_{t_i})\Big].
\end{aligned}$$

Therefore we introduce the following implicit scheme:

$$x_{t_{i+1}}^\alpha = x_{t_i}^\alpha + (k\theta - \frac{\sigma^2}{2} - kx_{t_{i+1}}^\alpha)(t_{i+1} - t_i) + \sigma\sqrt{x_{t_{i+1}}^\alpha}(W_{t_{i+1}} - W_{t_i}).$$

Then $\sqrt{x_{t_{i+1}}^\alpha}$ is the unique positive square root of the polynomial

$$P(X) = (1 + k(t_{i+1} - t_i))X^2 - \sigma(W_{t_{i+1}} - W_{t_i})X - (x_{t_i}^\alpha + (k\theta - \frac{\sigma^2}{2})(t_{i+1} - t_i))$$

and hence is given by

$$x_{t_{i+1}}^\alpha = \left(\frac{\sigma(W_{t_{i+1}} - W_{t_i}) + \sqrt{\Delta_{t_i}}}{2(1 + k(t_{i+1} - t_i))}\right)^2$$

with

$$\Delta_{t_i} = \sigma^2(W_{t_{i+1}} - W_{t_i})^2 + 4(x^\alpha_{t_i} + (k\theta - \frac{\sigma^2}{2})(t_{i+1} - t_i))(1 + k(t_{i+1} - t_i))).$$

A second approach starts with equation (3.15). Setting up the implicit scheme, we get again a second-degree equation, given by

$$(1 + \frac{k}{2}(t_{i+1} - t_i))x^\alpha_{t_{i+1}} - (\frac{\sigma}{2}(W_{t_{i+1}} - W_{t_i}) + \sqrt{x^\alpha_{t_i}})\sqrt{x^\alpha_{t_{i+1}}} + \frac{k\theta - \frac{\sigma^2}{4}}{2}(t_{i+1} - t_i) = 0.$$

By solving the equation and taking again the positive square root, we obtain

$$x^\alpha_{t_{i+1}} = \left(\frac{\frac{\sigma}{2}(W_{t_{i+1}} - W_{t_i}) + \sqrt{x^\alpha_{t_i}} + \sqrt{\tilde{\Delta}_{t_i}}}{2(1 + \frac{k}{2}(t_{i+1} - t_i))}\right)^2,$$

with

$$\tilde{\Delta_{t_i}} = (\frac{\sigma}{2}(W_{t_{i+1}} - W_{t_i}) + \sqrt{x^\alpha_{t_i}})^2 + 2(1 + \frac{k}{2}(t_{i+1} - t_i))(k\theta - \frac{\sigma^2}{4})(t_{i+1} - t_i).$$

The big problem of the Monte Carlo algorithm is, that we need a large number of simulations to achieve a good precision. For CDS-simulations, this is even more complicated, since the variance of the CDS price obtained by Monte Carlo simulation is quite large in comparison to the obtained value.
One method to reduce the number of simulations is to introduce a threshold barrier B, so that $\mathbb{P}(\Lambda(T) < B) \simeq 1$. This means, that we do not have to simulate values of ξ, that are larger than B, since in this case no default occurs and we already know the value of the CDS, which is $R_{0,b}(0) \sum_{i=1}^{b} \mathrm{B}(0, T_i)\alpha_i$.
For $\Lambda(\tau) \sim$ exponential distributed with parameter 1 we have

$$\mathbb{E}(CDS) = \mathbb{E}(CDS|\Lambda(\tau) < B)(1 - e^{-B}) + \mathbb{E}(CDS|\Lambda(\tau) \geq B)e^{-B}.$$

This will lead to an increase of our efficiency by $(1 - e^{-B})$, nevertheless, the number of simulations, or in other words the amount of time needed, remains quite large.

Gaussian dependence mapping:

Due to the drawbacks of Monte Carlo simulations for pricing CDS, Brigo and Mercurio [6] mentioned a different approach, that does not need Monte Carlo, called Gaussian dependence mapping. As we have seen in Section 3.4.3, the CDS formula can be displayed the following way, for $t = 0$ and $T_a = 0$:

$$\begin{aligned} CDS_{0,b}(0, R, \text{LGD}) = R \sum_{i=1}^{b} \alpha_i \mathbb{E}\Big(\exp(-\int_0^{T_i} (r_u + \lambda_u)\, du)\Big) \\ + R \int_0^{T_b} \mathbb{E}\Big(\exp(-\int_0^{u} (r_s + \lambda_s)\, ds)\lambda_u\Big)(u - T_{k(u)-1})\, du \\ - LGD \int_0^{T_b} \mathbb{E}\Big(\exp(-\int_0^{u} (r_s + \lambda_s)\, ds)\lambda_u\Big) du. \end{aligned} \tag{3.16}$$

The problem is in the case of $\rho \neq 0$, that we have no explicit formula for $\mathbb{E}(\exp(-\int_0^T (x_s^\alpha + y_s^\beta)\, ds))$.

According to [6] "the idea Gaussian dependence mapping approach is now, to map the two-dimensional CIR dynamics in an analogous tractable two-dimensional Gaussian dynamics, that preserves as much as possible of the original CIR structure."
By taking a Vasicek process for x and y we can use the following lemma:

Lemma 3.4.1: Let $x_t^{\alpha,V}$ and $y_t^{\beta,V}$ be two Vasicek processes as follows:

$$\begin{aligned} dy_t^{\beta,V} &= \kappa(\mu - y_t^{\beta,V})dt + \nu dZ_t \\ dx_t^{\alpha,V} &= k(\theta - x_t^{\alpha,V})dt + \sigma dW_t \end{aligned}$$

with $dZ_t dW_t = \rho dt$. Then $A = \int_0^T (x_s^{\alpha,V} + y_s^{\beta,V})\, ds$ and $B = y_T^{\beta,V}$ are Gaussian random variables with respective mean and variance given as follows:

$$\begin{aligned} m_A &= (\mu + \theta)T - [(\theta - x_0)g(k, T) + (\mu - y_0)g(\kappa, T)] \\ m_B &= \mu - (\mu - y_0)e^{-\kappa T}, \\ \sigma_A^2 &= \frac{\nu^2}{\kappa}(T - 2g(\kappa, T) + g(2\kappa, T)) + \frac{\sigma^2}{k}(T - 2g(k, T) + g(2k, T)), \\ &+ \frac{2\rho\nu\sigma}{\kappa k}(T - g(\kappa, T) - g(k, T) + g(k + \kappa, T)) \\ \sigma_B^2 &= \nu^2 g(2\kappa, T). \end{aligned}$$

The correlation is given as:

$$\tilde{\rho} = \frac{1}{\sigma_A \sigma_B}\Big[\frac{\nu^2}{\kappa}\Big(g(\kappa,T) - g(2\kappa,T)\Big) + \frac{\rho\nu\sigma}{k}\Big(g(\kappa,T) - g(\kappa+k,T)\Big)\Big]$$

where $g(k,T) = \frac{1-e^{-kT}}{k}$.

With the above formula and taking y^V the degenerated case with $\mu = \kappa = y_0 = 1$ and $\nu = 0$, we can calculate:

$$\mathbb{E}\Big[\exp\Big(-\int_0^T x_s^{\alpha,V}\,ds\Big)\Big] = \exp(-\theta t + (\theta - x_0)g(k,T) + \frac{1}{2}\frac{\sigma}{k}^2(t - 2g(k,T) + g(2k,T))),$$

which is exactly the bond price formula in the Vasicek model.

Then the expectation of the CIR process is matched with the expectation of the Vasicek process, such that

$$\mathbb{E}\Bigg(\exp(-\int_0^T x_s^{\alpha_T,V}\,ds)\Bigg) = \mathbb{E}\Bigg(\exp(-\int_0^T x_s^{\alpha}\,ds)\Bigg)$$
$$\mathbb{E}\Bigg(\exp(-\int_0^T y_s^{\beta_T,V}\,ds)\Bigg) = \mathbb{E}\Bigg(\exp(-\int_0^T y_s^{\beta}\,ds)\Bigg).$$

These equations are all analytically known.

Therefore we approximate the expectations needed in equation (3.16) by

$$\mathbb{E}\Bigg(\exp(-\int_0^T (x_s^{\alpha} + y_s^{\beta})\,ds)\Bigg) \approx \mathbb{E}\Bigg(\exp(-\int_0^T (x_s^{\alpha_T,V} + y_s^{\beta_T,V})\,ds)\Bigg)$$
$$\mathbb{E}\Bigg(\exp(-\int_0^T (x_s^{\alpha} + y_s^{\beta})\,ds)y_T^{\beta}\Bigg) \approx \mathbb{E}\Bigg(\exp(-\int_0^T (x_s^{\alpha_T,V} + y_s^{\beta_T,V})\,ds)y_T^{\beta_T,V}\Bigg) + \Delta,$$

with

$$\Delta = \mathbb{E}\Bigg(\exp(-\int_0^T x_s^{\alpha}\,ds)\Bigg)\mathbb{E}\Bigg(\exp(-\int_0^T y_s^{\beta}\,ds)y_T^{\beta}\Bigg)$$
$$- \mathbb{E}\Bigg(\exp(-\int_0^T x_s^{\alpha_T,V}\,ds)\Bigg)\mathbb{E}\Bigg(\exp(-\int_0^T y_s^{\beta_T,V}\,ds)y_T^{\beta_T,V}\Bigg).$$

3.5 Summary

As already mentioned above, the intensity approach does not use economic fundamentals, it focuses directly on the default event. This direct approach leads to advantages and drawbacks shortly discussed below.

One big advantage of intensity models in comparison with structural models is the better ability to match actual CDS-spreads. This fact is closely linked to the observability of market data required by the model, since intensity models only need the actual zero coupon curve (default-free market information) and CDS-spreads (default information) for calibration. Furthermore intensity models are based on a quite general approach, since in principle, we are free to use any short rate model of intensities.
It is important to mention, that λ does not only depend on the credit quality of the firm. λ is linked directly to the CDS-price, which is influenced by many factors, that cannot be separated, since the CDS-data is mapped directly to the intensities. For example, increasing CDS-rates do not have to be based on worse credit quality of this single firm, but may also depend on other influence factors, like insufficient liquidity on the CDS market, macroeconomic reasons, et cetera. Due to this inseparability of influence factors it is not possible to simulate for example macroeconomic changes and their impact on the default probability.
Another drawback of intensity models is that for most of the rate models leading to positive intensities, we do not get explicit formulas for the prices of CDS-rates or for corporate bonds. So we have to use Monte Carlo simulations or other approximation methods to obtain satisfying results. Here, variance reduction methods and Quasi Monte Carlo methods have to be used and we refer to [17], [8] and [10] for further information.

Chapter 4

Application of the models

In my diploma thesis I want to discuss also the practical feasibility of the models and how good they perform in matching real market data. Therefore we will discuss two real market examples of the 30^{th} of December 2009 and compare the results obtained by the different models.
This diploma thesis arose in the context of my work at the Raiffeisen Landesbank Steiermark. Therefore our field of interest lies in financial bonds and, as we will see later, banks are a special field of application, especially for structural models.
For the first example, Credit Suisse, we will calculate the default probabilities using the different models and compare the results. We will focus on the models with deterministic interest rate, since the application of the more complex models would go beyond the scope of this diploma thesis. To double check the results, we will take the obtained default probabilities to value two bonds, issued by Credit Suisse, and compare them with the market price. This procedure should give information about the reliability of the results, since the default probabilities should not differ significantly whether they are obtained by CDS- or bond data.
The second example, Banca Intesa, is a bank with similar market situation to Credit Suisse, since they have the same credit rating and very similar CDS spreads. As we will see later, the situation is though little different due to the different balance sheet data. Nevertheless the markets default probabilities should not differ significantly to the ones of Credit Suisse due to the similar CDS spreads. We will see later, how structural models can handle this situation.

Remark: All data for the following examples are taken from Bloomberg Terminal.

Consider the default free interest rate curve of the 30^{th} of December 2009 shown below:

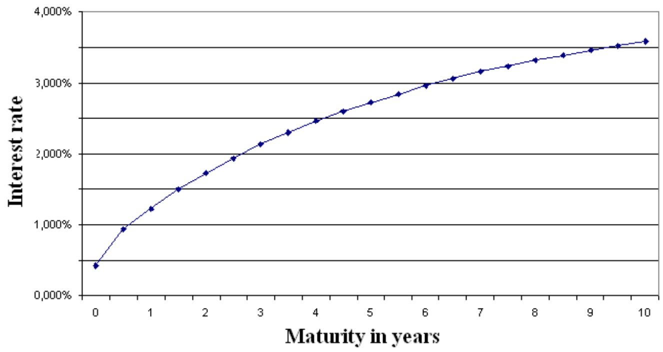

Figure 4.1: Zero coupon curve 30^{th} of December 2009

Remark: Here, we are already confronted with the first difficulty, because we have to choose the risk free interest rate. According to [12], interest swap rates are a good approximation for default free interest rates, since default risk does not play a crucial role in Swap contracts. This approach is also widely used in practice, therefore the above interest curve is stripped from the swap rate curve of the 30^{th} of December 2009.

4.1 Credit Suisse

Before we can start with the different models, we have to quote first the given data. The CDS-spreads of Credit Suisse on the 30^{th} of December 2009 are given in Table 4.1.

We will first start with intensity models, in particularly with the simple model of piecewise constant/linear intensities. These intensities are also the starting basis of the CIR++ model.
For the intensity approach we only need as input parameters the actual CDS rates and the interest rate curve. Using formula (3.3), we get the results, shown in Figure 4.2.

Maturity T_n	CDS Spread in BP
1Y	33
2Y	38
3Y	44
4Y	55
5Y	60
7Y	68
10Y	72

Table 4.1: CDS-Spreads

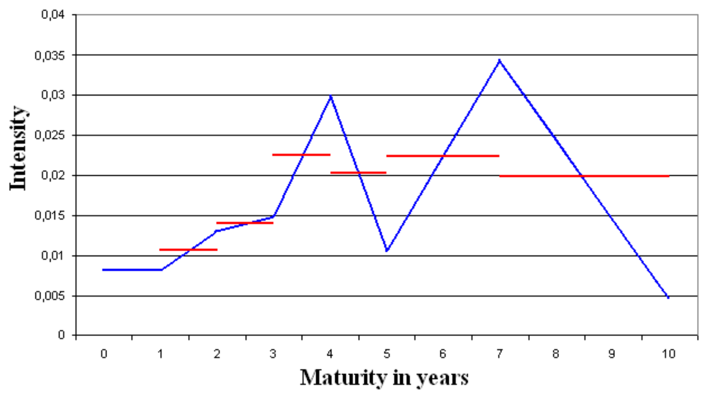

Figure 4.2: piecewise constant/linear default intensities

By taking these intensities as the "market intensities", we can fit the CIR++ model to this data. By using the least squares method, we obtain the results plotted in Figure 4.3. The parameters of the fitted CIR++ model are given by:

$$\kappa = 0,065939 \qquad \mu = 0,00001 \qquad \sigma = 0,00036315 \qquad y_0 = 0,00819139$$

We can compute the default probability by using the Bond price formula of the CIR++ model. Since we have fitted the CIR++ model to the piecewise linear default intensities, we will obtain the same probabilities for the CIR++ model and the piecewise constant intensity model. The default probabilities are given in Table 4.2.

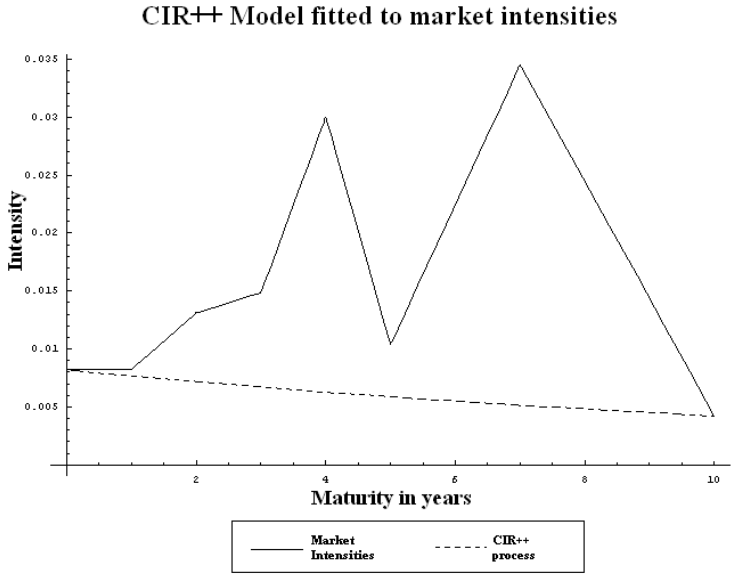

Figure 4.3: CIR++ intensity model

Maturity T_n	Constant Intensities	Piecewise intensities / CIR++
1Y	0,816 %	0,816 %
2Y	1,867 %	1,867%
3Y	3,228 %	3,231 %
5Y	7,272 %	7,275 %
7Y	11,306 %	11,345 %
10Y	16,442 %	16,363 %

Table 4.2: Default probabilities for Credit Suisse, 30^{th} of December 2009

Next we compare the results with the ones obtained by structural models. For structural models we need the equity value E_0, its volatility σ_E and the debt value D_0, which are normally observable on the market. The equity value is calculated by

$$E_0 = \text{Number of shares} * \text{Value of shares}$$

and σ_E can be estimated by the logarithmic changes of the equity value. The debt value for a common corporate is given by the nominal of the issued bonds, which is observable on market. When it comes to balance sheet data of banks, the segmentation of liabilities and equity is different, as illustrated in figure 4.4.

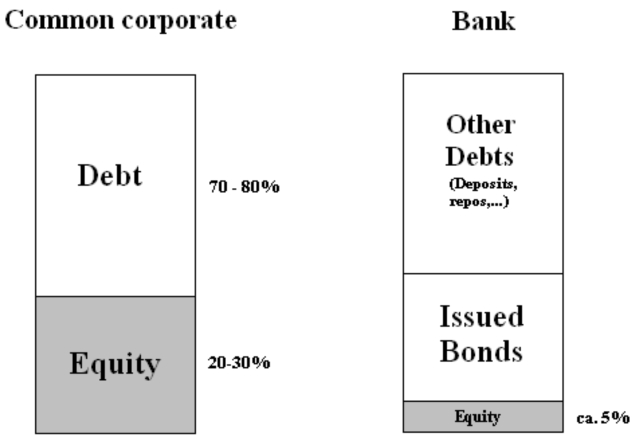

Figure 4.4: Liabilities and equity of banks in comparison to corporates

Nevertheless we will first use the equity value E_0 and the total liabilities to calculate the default probabilities. I will not focus on the estimation of the total liabilities, since we will see, that this approach will not lead to satisfying results. I have taken the total liabilities of the quarterly financial statement for my calculations.

$$E_0 = 35.819 \qquad \sigma_E = 30,245\% \qquad D_0 = 660.329$$

in TEUR. The balance sheet data of Credit Suisse is given in CHF, therefore I used the currency rate of the 30^{th} of December 2009, which equals 1,4888.

Remark: σ_E is estimated by the logarithmic changes of the equity value in CHF, since it should not contain the FX volatility.

Furthermore I assume the debt value D_t to be given by

$$D_t = D_0 \cdot e^{\gamma t}$$

where γ is according to the Merton and Black and Cox model the interest rate the investors get for their capital. A good estimator for γ is the riskless interest rate r plus the CDS spread according to the maturity, since in principle, the investor can invest in the bond and buy a CDS as insurance in case of default.

With these input data, we calculate the starting firm value V_0 and its volatility according to the method mentioned in Section 2.2:

$$V_0 = 696.149 \qquad \sigma_V = 1,5562\%$$

The 10 year default probability then equals:

$$\mathbb{P}(\tau < 10) = 64,6\%$$

It is obvious that this probability is overestimated and in total contrariety to the market's CDS spreads.

As a second approach we will assume the "other debts" as constant. As illustrated in figure 4.5, the model is only sensitive on the distance to default, therefore a constant factor in D_t and V_t do not affect the result.

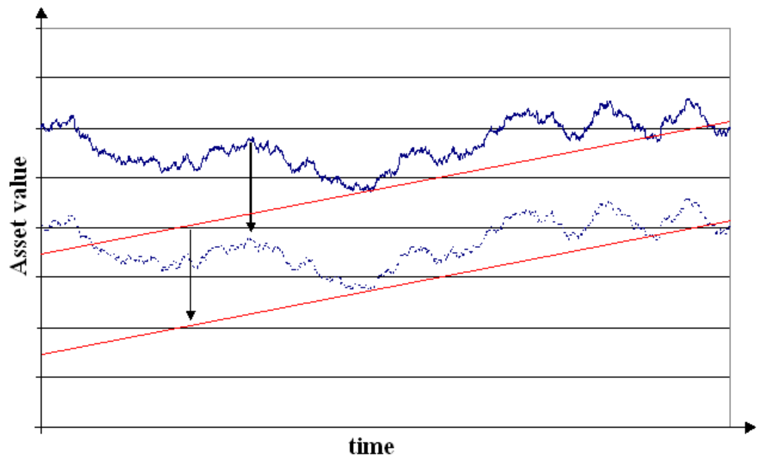

Figure 4.5: CIR++ intensity model

With this assumption the model is equivalent to the structural models for common corporates, since we may estimate the "changing" total asset value as equity value E_0 and debt value, respectively the nominal of the issued bonds, D_0. On the one hand this is quite a strong hypothesis, but on the other hand it can be seen as a way to translate the bank data to the corporate setup of the Merton and the Black and Cox model.

Remark: For the debt structure we definitively have to take into account the nominal of the issued bonds, since the nominal and the interest has to be paid in fully amount until/at maturity to avoid default. As we will see later, the obtained default probabilities seem to be quite consistent with the ones obtained by the intensity models, but even overestimated for maturities over 5 years. Therefore, including the other liabilities would lead to overestimation of the default probabilities for longer maturities.

Using this approach, we get the following results:

$$E_0 = 35.819 \qquad \sigma_E = 30,245\% \qquad D_0 = 93.386 \qquad D_t = D_0 \cdot e^{\gamma t}$$

We calculate again the starting firm value V_0 and its volatility according to the method mentioned in Section 2.2:

$$V_0 = 129.205 \qquad \sigma_V = 8,385\%$$

As we can see, the debt-to-equity ratio is now 2,6%, which is in the usual range. Using the Merton Model and the Black and Cox Model we obtain the results given in Table 4.3.

Maturity T_n	Merton	Black and Cox
1Y	0,007 %	0,01 %
2Y	0,45 %	0,54 %
3Y	1,91 %	2,16 %
5Y	6,97 %	7,52 %
7Y	12,75 %	13,40 %
10Y	20,59 %	21,28 %

Table 4.3: Default probabilities for Credit Suisse, 30^{th} of December 2009

The results obtained by this approach seem more in accordance with the market CDS rates. The Black and Cox model produces higher default probabilities than the Merton's model. We have expected this kind of result, since the Black and Cox model has a similar setup, but allows also default before maturity T.
Here we can see one characteristic of structural models, which was already

mentioned in the conclusion of Chapter 2. The short term default probabilities produced by the Merton and also the Black and Cox model are approximately zero. Therefore, by pricing for example a CDS or a corporate bond with a short duration, the calculated prices do not match the market prices.

For testing purposes, we will calculate the price of two coupon bonds, issued by Credit Suisse, given the default probabilities of the different models. Since we use deterministic interest rates, we can use formula 3.10 to calculate the present value of the bonds.
An annual 4,625% coupon bond with maturity 07.06.2010 had a bid price of 104,0179 and an ask price of 104,2019. So the middle price, we want to estimate by our models is 104,1099. This example is very simple, since only one payment is outstanding. Using the different models, we get the following results:

Model	Price	YtM
Merton	104,2461	0,834 %
Black and Cox	104,2460	0,835 %
Market price	104,1099	1,136 %
Constant intensities	104,099	1,160 %
CIR++ / linear intensities	104,099	1,160 %

Table 4.4: Coupon bond Price

The price of the intensity models is the same, since we assume in the first interval $[0, T_1]$ a constant intensity. As we can see in Table 4.4, in case of short term contracts the results produced by structural models and intensity models are quite different. As a measure of the difference one may use the corresponding yield-to-maturity given also in Table 4.4. Notice that taking no default risk into account would cause a price of 104,252. So for contracts with short duration structural models cannot be used in practice and intensity models clearly have to be preferred.

The second bond, is an annual 5,125 coupon bond with maturity 18.09.2017. Its bid price was 106,943, the ask price 107,436, which results in a middle price of 107,189.

Model	Price	YtM
Merton	106,302	4,35 %
Black and Cox	105,904	4,41 %
Market price	107,190	4,22 %
Constant intensities	106,996	4,25 %
CIR++ / linear intensities	106,799	4,28 %

Table 4.5: Coupon bond Price

As we can see from Table 4.5, also here the intensity models match the market prices better than the structural models.

The Merton and the Black and Cox model reflect the drawbacks mentioned in Section 2, especially the underestimation of the short term default probabilities. Nevertheless I have to mention, that the small number of input factors is on the one hand the big advantage of intensity models, but on the other hand also its drawback. Given the interest rate, intensity models use a bijective mapping of CDS-rates to intensities, therefore intensity models will match market prices, but a proceeding analysis of the influence factors is not in the scope of the model.

4.2 Banca Intesa

The second example we will discuss is Banca Intesa. Banca Intesa is an Italian bank with the same credit rating as Credit Suisse. As shown in Table 4.6 also the CDS spreads are in the same range, therefore the results obtained by the different models should not differ significantly in comparison to Credit Suisse. As we will see later, the debt-to-equity value is quite different, therefore we want to test how structural models can handle this situation.

Maturity T_n	CDS Spread in BP
1Y	33
2Y	38
3Y	42
4Y	54
5Y	59
7Y	65
10Y	70

Table 4.6: CDS-Spreads

Using intensity models, the probability of default is only influenced by the interest rate and the actual CDS-spreads. Therefore the obtained results, given in Table 4.7, are very similar to the ones of Credit Suisse.

Maturity T_n	Constant Intensities	Piecewise intensities / CIR++
1Y	0,829 %	0,831 %
2Y	1,843 %	1,843%
3Y	3,115 %	3,118 %
5Y	7,195 %	7,196 %
7Y	10,767 %	10,803 %
10Y	16,166 %	16,124 %

Table 4.7: Default probabilities

For structural models, we estimate again the firm value by the equity value E_0, its volatility σ_E and the debt value D_0 and D_t. Using the same approach as in section 4.1, we have the following input data:

$$E_0 = 34.868 \qquad \sigma_E = 28,416\% \qquad D_0 = 185.243 \qquad D_t = 185.243 \cdot e^{\gamma t},$$

with γ being again the riskless interest rate plus the CDS spread in basis points. To obtain the total asset value and its volatility we use again the approach described in Section 2.2.

$$V_0 = 220.111 \qquad \sigma_V = 4,5015\%$$

As we can see we have here a debt-to-equity ratio of 4,9, whereas the debt-to-equity ratio of Credit Suisse was 2,6. For the default probabilities of the Merton and Black and Cox Model this leads to the following results:

Maturity T_n	Merton	Black and Cox
1Y	0,009 %	0,010 %
2Y	0,053 %	0,56%
3Y	2,27 %	2,33 %
5Y	8,53%	8,62 %
7Y	15,71 %	15,78 %
10Y	25,83 %	25,88%

Table 4.8: Default probabilities

By comparing the default probabilities for Banca Intesa and Credit Suisse, the long term default probabilities differ clearly, despite the similar initial situation.

The reason for this difference may be based on the technical setup of structural models in general. Structural models have to deal with more hypothesis in the modeling and the calibration, therefore there is a bigger margin of error. The different balance sheet data of banks even worsens this problem.
Concluding we can say, that using structural models requires a lot of know-how and a well-founded analysis of the market situation, including balance sheet data, industry sector,et cetera.
Due to the observability and the simplicity of the input data, in this special application, the direct approach of intensity models works more reliable.

4.3 Conclusion

In my opinion, the feasibility of the different models and the special demands of banks as counterparties, make intensity models more applicable than structural models.
First, the non-observability of data and therefore the estimation of the required input parameters make structural models hard to calibrate and automatize, and second, these models do not really match actual market data well.
Exactly these issues are the advantages of the intensity models, since all needed input parameters are observable in real time and the obtained probabilities of default match the markets opinion.
As a drawback of the intensity models we have to mention, that we need numerical methods to solve the set of equations $CDS_{T_0,T_i}(0, R_i, LGD, \Gamma(\cdot)) = 0$. Nevertheless we only have to solve the system of equations assuming independence, the computational time requirements should remain in tolerable limits.

Within intensity models, the choice of the model to use, is clearly dependent on the field of application. If the field of interest is the default probabilities, the simplified method with piecewise constant or linear intensities could be sufficient, whereas if the dependence on the interest rate has to be taken into account, then the SSRD model can better fit the requirements.

Last but not least we have to mention, that these models are only one tool among others. This means that there are of course other influence factors on CDS rates or corporate bond prices than the default probability that are not taken into account by these models, like liquidity or the debt maturity structure. But nevertheless these models are a good support to manage default risk in a quantitative way.

Bibliography

[1] Ammann M. (2004): *Credit Risk Valuation: Methods, Models, and Applications*. Springer Finance

[2] Bielecki T., Rutkowski M. (2002): *Credit Risk: Modeling, Valuation and Hedging*. Springer Finance

[3] Black F., Cox J.C. (1976): *Valuing corporate securities: Some effects of bond indenture provisions*. In The Journal of Finance 31, pp. 351-367.

[4] Bluhm C., Overbeck L., Wagner C. (2002): *An Introduction to Credit Risk Modeling*. Chapman & Hall/CRC Financial Mathematics Series.

[5] Black F., Scholes M. (1973): *The pricing of options and corporate liabilities*. In Journal of political economy 81, pp. 637-654..

[6] Brigo D., Mercurio F. (2006): *Interest Rate Models - Theory and Practice*. 2nd Edition, Springer Finance

[7] Cox J. C., Ingersoll J. E. JR., Ross S. A. (1985): *A Theory of the Term Structure of Interest Rates*. In Econometrica, Vol. 53, No. 2, pp. 385-407.

[8] Deelstra G., Delbean F. (1998): *Convergence of Discretized Stochastic (interest rate) Processes with Stochastic Drift Term*. In Appl. Stochastic Models Data Anal. 14, 77-84 (1998).

[9] Delbaen, Freddy ; Schachermayer, Walter (1993): *A general version of the fundamental theorem of asset pricing*. In Mathematische Annalen, Volume 300, pp. 463-520.

[10] Diop A.: *Sur la discrétisation et le comportement à petit bruit d'EDS multidimensionnelles dont les coefficients sont à dérievées simguliere*. Ph. D. Thesis, INRIA.

[11] Duffie D.,Singleton K.J. (2003: *Credit Risk: Pricing, Measurement, and Management*. Princeton Series in Finance

[12] Hull J., Predescu M., White A. (2004): *Bond Prices, Default Probabilities and Risk Premiums*. In Journal of Credit Risk, Vol. 1, No. 2, (2005), pp. 53-60..

[13] Hurd T., Yi C. (2008): *In Search of Hybrid Models for Credit Risk: from Leland-Toft to Carr-Linetsky*. available at http://www.math.mcmaster.ca/ yichuang/MyPapers/HurdYi08b.pdf

[14] Ito K., McKean H.P. (1965): *Diffusion Processes and Their Sapmle Paths*. Springer

[15] Jamshidian F. (1989): *An exact bond option pricing formula*. In The Journal of Finance 44, pp. 205-209.

[16] Karatzas I., Shreve S. E. (1998): *Brownian Motion and Stochastic Calculus*. 2nd Edition, Springer

[17] Lord R., Koekkoek R., van Dijk D.J.C. (2006): *A comparison of biased simulation schemes for stochastic volatility models*. In Tinbergen Institute Discussion Papers: 06-046/4. Tinbergen Institute

[18] Meissner G. (2004): *Credit Derivatives: Application, Pricing, and Risk Management*. Blackwell Publishers

[19] Merton R.C. (1974): *On the pricing of corporate debt: The risk structure of interest rates*. In The Journal of Finance 29, pp. 449-470.

[20] Musiela M., Rutkowski M. (1997): *Martingale methods in financial modelling*. Springer

[21] Ross S. (1995): *Stochastic Processes*. Second edition, Wiley

[22] Shreve S. E. (2004): *Stochastic Calculus for Finance I: The Binomial Asset Pricing Model*. Springer Finance

[23] Shreve S. E. (2004): *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer Finance