Credit Risk Modeling with Random Fields

Inaugural-Dissertation

zur Erlangung des Doktorgrades
an den Naturwissenschaftlichen Fachbereichen
(Mathematik)

der Justus-Liebig-Universität Gießen

vorgelegt von
Thorsten Schmidt

Gießen 2003

D - 26

Dekan: Prof. Dr. Albrecht Beutelspacher

Gutachter: Prof. Dr. Winfried Stute (Gießen)
Prof. Dr. Ludger Overbeck (Gießen)

Datum der Disputation: 11.09.2003

Preface

The demand for investments with higher returns in areas other than the stock market has increased enormously due to the stock market crash in the last two years. In exchange for an attractive yield the investors take a credit risk, and as a result methodologies for pricing and hedging credit derivatives as well as for risk management of credit risky assets became very important. The efforts of the Basel Committee is just one of many examples which substantiate this.

In the last years the credit markets developed at a tremendous speed while at the same time the number of corporate defaults increased dramatically. It is therefore not surprising that the demand for credit derivatives is growing rapidly.

In view of this, the goal of this work is twofold. In the first part, a survey of the credit risk literature is given, which offers a quick introduction into the area and presents the mathematical methods in a unifying way. Second, we propose two new models of credit risk, focusing on different needs. The first model generalizes existing models using random fields in Hilbert spaces. The second model uses Gaussian random fields leading to explicit formulas for a number of derivatives, for which we propose two calibration procedures.

This work is organized as follows. In Chapter 1, a survey of the credit risk literature is given. This includes structural models, hazard rate models, methods incorporating credit ratings, models for baskets of credit risky bonds, hybrid models, market models and commercial models. In the last section we illustrate several credit derivatives. Generally the mathematical framework for the models is provided and some models are discussed in greater detail. Additionally, an explicit formula for the default intensity in the imperfect information model of Duffie and Lando (2001) is derived.

Chapters 2 and 3 focus on credit risk modeling using stochastic differential equations (SDEs) in infinite dimensions. Although known in interest rate theory, the application of these methods is new to credit risk. Chapter 2 contains an introduction to SDEs in Hilbert spaces providing an Itô formula which is adequate for our purposes. In Chapter 3 a Heath-Jarrow-Morton formulation of credit risk in infinite dimensions is given. The work of Duffie and Singleton (1999) and Bielecki and Rutkowski (2000) was enhanced with alternative recovery models and extended to infinite dimensions. These new models comprise most of the known credit risk models and still offer frameworks which are tractable. Recent research in Özkan and Schmidt (2003) extends this further to Lévy processes in infinite dimensions.

In Chapter 4, a credit risk model is presented which uses Gaussian random fields and transfers the framework of Kennedy (1994) to credit risk. In contrast to the functional analytic approach in the previous two chapters, the methods used in this section concentrate on deriving formulas for pricing and hedging. Explicit expressions for the prices of several credit default options are obtained and an example for hedging credit derivatives is presented.

Based on these pricing formulas, two calibration methodologies are provided. The first calibration procedure fits the model to prices of derivatives using a least squares approach. As the data for derivatives like credit default swaptions is still scarce, the second approach takes this into account and in addition uses historical data. This new approach allows to calibrate perfectly to market prices and is applicable using only a small amount of credit derivatives data.

I am most grateful to my supervisor, Prof. Dr. Winfried Stute, for his vital support. His fascinating lectures and his way of inspiring mathematics were a highly valuable encouragement. Always having time for fruitful discussions is just one example of his continual support throughout the making of this thesis. I also warmly thank my friends and colleagues from the “Stochastik-AG”.

Special thanks go to Sue, Charlie and Oli for spending hours and hours reading cryptic notes. I wish, especially, to thank my family for their education which encouraged the search for answers and helping me whenever I needed them.

Finally, I thank my dearest Kirsten as she brightens my life with her love.

Contents

Chapter 1

Credit Risk - A Survey

1.1 Introduction

The first regulations of lending and interest were mentioned in Hammurabi's Code of Laws. Hammurabi was a famous Babylonian king, who lived circa 1800 BC. The most remarkable source for his legal code is a stone slab discovered in 1901 which is preserved in the Louvre, Paris. Other cuneiform tables record a number of textbook-like interest rate problems. For example, the cuneiform table "VAT 8528" poses the following problem[1]:

> "If I lent one mina of silver at the rate of 12 shekels (1/60 of a mina) per year, and I received in repayment, one talent (60 minas) and 4 minas. For how long was the money lent?"

As long as lending is subject to a person's employment, there is risk of losing part of the loan, which in modern financial language would be called *credit risk*. A common definition of credit risk is the following:

> "Credit risk refers to the possibility that a contractual counterpart may not be able to meet his obligations so that the lender faces a financial loss."

The financial object, which is subject to credit risk, is a so-called *bond*. In today's financial markets there is a vast variety of bonds traded, from Treasuries issued by different countries or states to bonds issued by corporates. Generally speaking, a bond is a certificate confirming that its owner, the creditor, has lent a certain amount of money to a specified issuer. The lent sum is called the *principal* or *face value* of the bond and has to be repaid at a fixed date, called *maturity* of the bond. Additionally the bond offers a fixed rate of interest and appears as an example of a fixed-income instrument.

Even if the creditor has no kind of ownership rights, it is important to note that in the event of business liquidation, bond holders have priority over shareholders in terms of ability to reclaim capital.

[1] See Neugebauer (1969). Further historical information on interest rates in history may be found in chapter two of James and Webber (2000).

The risk of the bond holder to lose a certain portion of his investment is the above mentioned credit risk. Accordingly, the creditworthiness of the issuer is an important kind of information. Agencies like Moody's and Standard & Poor's classify the creditworthiness of the issuers by the so-called *rating*. As a consequence, market participants demand higher yields for lower rated bonds as a compensation for the taken risk. The excess return of the corporate bond over a Treasury bond, i.e., a bond which is assumed to be free of credit risk, is called the *credit spread*; see Bielecki and Rutkowski (2002).

A *default* occurs if the issuer is not able to meet his obligations. The precise definition of a default is complicated, because it is itself negotiable; see Tavakoli (1998). Certainly, an amount of money is lost, and the post-default value of the bond, which is called *recovery*, significantly differs from the pre-default value. For this reason, spread-widening risk or changes in credit quality are also implied when talking about credit risk.

The occurrence of credit risk raises the demand for possibilities to manage them. This is when *credit derivatives* come into play. They enable protection against different types of credit risk to the effect that certain risk profiles are achieved. For example, credit derivatives can be used, if an investor wants to hedge himself against a credit risk, but not against interest risk. As both are entangled in a bond, credit derivatives provide the tailor-made possibility to trade this specific risk.

It is important to distinguish between *reference risk* and *counterparty risk*. The former refers to a contract of two default free parties, where the contract relates to the credit risk of some reference entity. If, on the other hand, over-the-counter derivatives are traded, which are in contrast to exchange-traded contracts not backed by a clearinghouse or an exchange, then each party faces the default risk of its counterparty.

We introduce several classes of models of credit risk, which serve different needs. Some try to determine the magnitude of credit risk in a certain product while others are more suitable for the management of whole portfolios or for pricing derivatives.

Structural models date back to the Nobel Prize paper of Merton (1974). They make a specific assumption about the capital structure of a company, which leads to a precise specification when obligations cannot be fulfilled. Therefore, the probability of a default can be determined and further calculations done. A commercial implementation of this model is presented in Section 1.8.1.

Conversely, *hazard rate models* focus on modeling the time, at which the default event occurs, while the capital structure of the company is not modeled at all. The default event is specified in terms of an exogenous jump process, which itself might depend on interest rates, credit ratings, firms assets or others. Often also called reduced-form or intensity based models, they were first mentioned in Pye (1974). An important class of hazard rate models incorporate *credit ratings*, readily available information on the creditworthiness of the bonds issuer.

So-called *hybrid models* try to combine these ideas and incorporate both hazard rates and the capital structure of the company. From this perspective these interesting models are relatively new in the financial literature and a lot of research is going on in this field.

In the section on *basket models* we present two methods of modeling a portfolio of credit risky securities. Basket models are mainly used to value credit derivatives with a first-to-default feature.

Market models represent the transfer of a very successful class of interest rate models to credit risk. They mainly cover the fact, that yields (or bonds, respectively) in the market are available with respect to a finite number (less than 20) of maturity times, and not for any maturity as assumed by most other models.

Quite different are the *commercial models* which represent readily available software packages. These models show the implementation of several methods handling credit risk and applications to large portfolios.

Finally we present certain *credit derivatives* in a precise specification. These include credit default swaps and swaptions, credit default options, credit spread options and options with a first-to default feature, and provide the basis for deriving prices in different models.

1.2 Structural Models

The first class of models tries to measure the credit risk of a corporate bond by relating the firm value of the issuing company to its liabilities. If the firm value at maturity T is below a certain level, the company is not able to pay back the full amount of money, so that a default event occurs.

1.2.1 Merton (1974)

In his landmark paper Merton (1974) applied the framework of Black and Scholes (1973) to the pricing of a corporate bond. A corporate bond promises the repayment F at maturity T. Since the issuing company might not be able to pay the full amount of money back, the payoff is subject to default risk.

Let V_t denote the firm's value at time t. If, at time T, the firm's value V_T is below F, the company is not able to make the promised repayment so that a default event occurs. In Merton's model it is assumed that there are no bankruptcy costs and that the bond holder receives the remaining V_T, thus facing a financial loss.

If we consider the payoff of the corporate bond in this model, we see that it is equal to F in the case of no default ($V_T \geq F$) and V_T otherwise, i.e.,

$$1_{\{V_T > F\}} F + 1_{\{V_T \leq F\}} V_T = F - (F - V_T)^+.$$

If we split the single liability into smaller bonds with face value 1, then we can replicate the payoff of this bond by a portfolio of a riskless bond $B(t,T)$ with face value 1 (long) and $1/F$ puts with strike F (short).

Consequently the price of the corporate bond at time t, which we denote by $\bar{B}(t,T)$, equals the price of the replicating portfolio:

$$\begin{aligned}\bar{B}(t,T) &= B(t,T) - 1/F \cdot P(F,V_t,t,T,\sigma_V) \\ &= e^{-r(T-t)} - \frac{1}{F}\left(Fe^{-r(T-t)}\Phi(-d_2) - V_t\Phi(-d_1)\right) \\ &= e^{-r(T-t)}\Phi(d_2) + \frac{V_t}{F}\Phi(-d_1), \end{aligned} \tag{1.1}$$

where $\Phi(\cdot)$ is the cumulative distribution function of a standard normal random variable.

Furthermore, $P(F,V_t,t,T,\sigma_V)$ denotes the price of a European put on the underlying V with strike F, evaluated at time t, when maturity is T and the volatility of the underlying is σ_V. This price is calculated using the Black and Scholes option pricing formula. The constants d_1 and d_2 are

$$\begin{aligned} d_1 &= \frac{\ln \frac{V_t}{Fe^{-r(T-t)}} + \frac{1}{2}\sigma^2(T-t)}{\sigma\sqrt{T-t}} \\ d_2 &= d_1 - \sigma\sqrt{T-t}. \end{aligned}$$

If the current firm value V_t is far above F the put is worth almost nothing and the price of the corporate bond equals the price of the riskless bond. If, otherwise, V_t approaches F the put becomes more valuable and the price of the corporate bond reduces significantly. This is the premium the buyer receives as a compensation for the credit risk included in the contract. Price reduction implies a higher yield for the bond. The excess yield over the risk-free rate is directly connected to the creditworthiness of the bond and is called the *credit spread*. In this model the credit spread at time t equals

$$\begin{aligned} s(t,T) &= -\frac{1}{T-t}\ln\left[\bar{B}(t,T)e^{r(T-t)}\right] \\ &= -\frac{1}{T-t}\ln\left(\Phi(d_2) + \frac{V_t}{F\cdot e^{-r(T-t)}}\Phi(-d_1)\right), \end{aligned}$$

see Figure 1.1.

The question of *hedging* the corporate bond is easily solved in this context, as hedging formulas for the put are readily available. To replicate the bond the hedger has to trade the risk-free bond and the firm's share simultaneously[2]. This reveals the fact that in Merton's model the corporate bond is a derivative on the risk-free bond and the firm's share.

We face the following problems within this model:

- The credit spreads for short maturity are close to zero if the firm value is far above F. This is in contrast to observations in the credit markets, where these *short maturity spreads* are not negligible because even close to maturity the bond holder

[2]The hedge consists primarily of hedging $\frac{1}{F}$ put and is a straightforward consequence of the Black-Scholes Delta-Hedging.

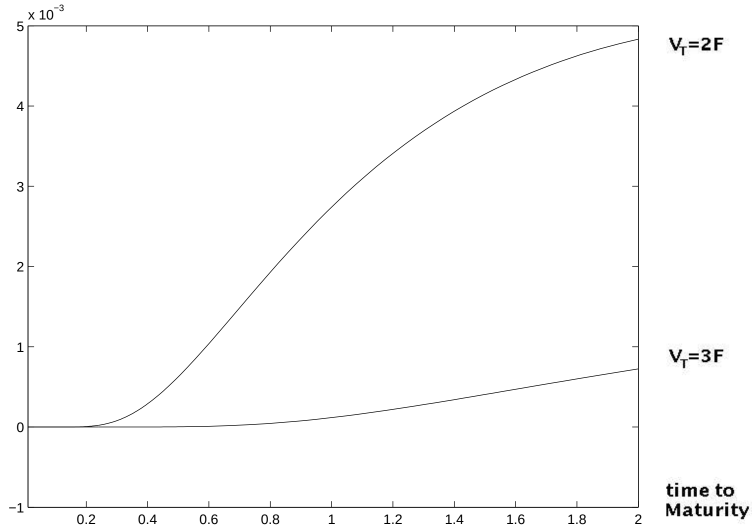

Figure 1.1: This plot shows the credit spread versus time to maturity in the range from zero to two years. The upper line is the price of a bond issued by a company whose firm value equals twice the liabilities while for the second the liabilities are three times as high. Note that if maturity is below 0.3 years the credit spreads approach zero.

is uncertain whether the full amount of money will be paid back or not; cf. Wei and Guo (1991) and Jones, Mason and Rosenfeld (1984).

The reason for this are the assumptions of the model, in particular continuity and log-normality of the firm value process. On the other hand, the intrinsic modeling of the default event may also be questionable. In reality there can be many reasons for a default which are not covered by this model.

- The model is not designed for different bonds with different maturities. Also it can happen that not all bonds default at the same time (*seniority*).
- In practice not all liabilities of a firm have to be paid back at the same time. One distinguishes between short-term and long-term liabilities. To determine the critical level where the company might default Vasiček (1984) introduced the *default point* as a mixture of the level of outstandings. This concept is discussed in Section 1.8.1.
- The interest rates are assumed to be constant. This assumption is relaxed, for example, by Kim, Ramaswamy and Sundaresan (1993), as discussed in Section 1.2.4.
- As there are only few parameters which determine the price of the bond, this model cannot be calibrated to all traded bonds on the market, which reveals arbitrage possibilities.

Geske and Johnson (1984) extended the Merton model to coupon-bearing bonds while Shimko, Tejima and van Deventer (1993) considered stochastic interest rates using the interest rate model proposed in Vasiček (1977). The second extension is essentially equivalent to pricing a European put option with Vasiček interest rates, where closed-form solutions are available. Of course, any other interest rate model can be used in this framework, like Cox, Ingersoll and Ross (1985) or Heath, Jarrow and Morton (1992).

1.2.2 Longstaff and Schwartz (1995)

As already mentioned defaults in the Merton model are restricted to happen only at maturity, if at all. In practice defaults may happen at any time. Also, when a company offers more than one bond with different maturities or seniorities, inconsistencies in the Merton model show up which can be solved by the following approach.

Black and Cox (1976) first used *first passage time models* in the context of credit risk. This means that a default happens at the first time, when the firm value falls below a pre-specified level. They used a time dependent boundary, $F(t) = ke^{-\gamma(T-t)}$, which resulted in a random default time τ. Unfortunately, this framework proves to be unsatisfactory.

Longstaff and Schwartz (1995) extended the Merton, respectively Black and Cox, framework with respect to the following issues:

- Default may happen at the first time, denoted by τ, when the firm value V_t drops below a certain level F.
- Interest rates are stochastic and assumed to follow the Vasiček model.

As a consequence, the firm value at default equals F. In the Merton model the value of the defaulted bond was assumed to be V_T/F which equals 1 in this context. The recovery value of the bond is therefore assumed to be a pre-specified constant $(1-w)$. This is the fraction of the principal the bond holder receives at maturity. Since further defaults are excluded in this model, the bond value at default equals $\bar{B}(\tau,T) = (1-w)B(\tau,T)$, where $B(t,T)$ is the value of a risk-free bond maturing at T. This assumption is often referred to as *recovery of treasury value.*

In the following, we present the model of Longstaff and Schwartz (1995) in greater detail. The firm value is assumed to follow the stochastic differential equation

$$\frac{dV(t)}{V(t)} = \mu(t)\,dt + \sigma\,dW_V(t),$$

and the spot rate is modeled according to the model of Vasiček (1977):

$$dr(t) = \nu(\theta - r(t))\,dt + \eta\,dW_r(t). \tag{1.2}$$

Moreover,

$$\mathbb{E}(W_V(s)\,W_r(t)) = \rho\cdot(s\wedge t) \qquad \text{for all } t \text{ and } s.$$

The last equation reveals a possible correlation between the two Brownian Motions W_V and W_r.

The Vasiček model exhibits a mean-reversion behavior at level θ and easily allows for an explicit representation of r_t. It is a classical model used in interest rate theory and often taken as a starting point for more sophisticated models. A drawback of this model is the fact that it may exhibit negative interest rates with positive probability. See, for example, Brigo and Mercurio (2001) and the discussions therein.

For the price of the defaultable bond they obtain

$$\begin{aligned}\bar{B}_{\mathrm{LS}}(t,T) &= B(t,T)\cdot \mathbb{E}^{Q^T}\left[1_{\{\tau>T\}} + (1-w)1_{\{\tau\le T\}}\,\middle|\,\mathcal{F}_t\right] \\ &= B(t,T)\cdot \left[w\,Q^T(\tau>T|\mathcal{F}_t) + (1-w)\right]. \end{aligned} \tag{1.3}$$

Note that $Q^T(\tau > T|\mathcal{F}_t)$ is the conditional probability (under the T-forward measure[3]) that the default does not happen before T.

To the best of our knowledge, a closed-form solution for this probability is not available[4]. Nevertheless there are certain quasi-explicit results provided by Longstaff and Schwartz (1995). See also Lehrbass (1997) for an implementation of the model.

In the empirical investigation of Wei and Guo (1991), the Longstaff and Schwartz model reveals a performance worse than the Merton model. According to these authors this is mainly due to the exogenous character of the recovery rate.

1.2.3 Jump Models - Zhou (1997)

Another approach to solve the problem of *short maturity spreads* is to extend the firm value process to allow for jumps. Mason and Bhattacharya (1981) extended the Black and Cox (1976) model to a pure jump process for the firm value. The size of the jumps has a binomial distribution. In this model there is some considerable probability for the default to happen even just before maturity.

Alternatively, Zhou (1997) extended the Merton model by assuming the firm value to follow a jump-diffusion process. The immediate consequence is that defaults are not predictable. The model is formulated directly under an equivalent martingale measure Q, and the firm value is assumed to follow

$$dV_t/V_{t-} = (r_t - \lambda\nu)dt + \sigma dW_V(t) + (\Pi_t - 1)dN_t. \tag{1.4}$$

$(N_t)_{t\ge 0}$ is a Poisson process with constant intensity λ. The jumps are $\Pi_t := U_{N_t}$, where $U_1, U_2, \ldots$ are i.i.d. and assumed to be independent of (N_t), (r_t) and $(W_V(t))$. Denote $\nu := \mathbb{E}(U_1) - 1$. Note that the integral of $(\Pi_t - 1)\,dN_t$ is shorthand for

$$Y_s := \int_0^s (\Pi_t - 1)\,dN_t = \sum_{i=1}^{N_s}(U_i - 1),$$

so that $(Y_t)_{t\ge 0}$ is a marked point process. It can be proved[5] that $(Y_t - \lambda\nu t)_{t\ge 0}$ is a martingale so that consequently the discounted firm value is a martingale under the measure Q.

[3]The T-forward measure is the risk neutral measure which has the risk-free bond with maturity T as numeraire. For details see Björk (1997).

[4]See discussions in Bielecki and Rutkowski (2002) and Goldstein (1997).

[5]See, for example, Brémaud (1981).

The interest rate is assumed to be stochastic and follow the Vašiček model; see (1.2). The recovery rate is determined by a deterministic function w, so that the bond holder receives

$$\big(1 - w(V_\tau/F)\big)$$

at default. The function w represents the loss of the bond's value due to the reorganization of the firm. For $w = 1$ we have the zero recovery case.

Zhou considers two models. The first, more general model, assumes that default happens at the first time when the firm value falls below a certain threshold. See the previous chapter for more examples of this class of models. Since in this case no closed-form solutions are available, the author proposes an implementation via Monte-Carlo techniques.

In the second, more restrictive model, the author obtains closed form solutions. For this a constant interest rate and log-normality of the U_i's is assumed and default happens only at maturity T, when $V_T < F$. Furthermore w is assumed to be linear, i.e., $w(x) = 1 - \tilde{w}\,x$. For $\tilde{w} = 1$ we obtain the recovery structure of the Merton model.

Equation (1.4) takes the form of a Doleans-Dade exponential and can be explicitly solved under these assumptions, cf. Protter (1992, p. 77):

$$V_t = V_0 \exp\Big[\sigma_V W_V(t) + (r - \frac{1}{2}\sigma_V^2 - \lambda\nu)t\Big] \prod_{i=1}^{N_t} U_i.$$

We then have the following

Proposition 1.2.1 (Zhou). *Denote* $\sigma_U^2 := \mathrm{Var}(\ln U_1)$ *and* $\tilde{\nu} := 1 + \nu$. *Then the price of a defaultable bond in the above model equals*

$$\begin{aligned}\bar{B}_{ZH}(0,T) &= \frac{\tilde{w}}{F} V_0 e^{-\lambda T\tilde{\nu}} \sum_{j=0}^{\infty} \frac{(\lambda\tilde{\nu}T)^j}{j!} \Phi\Big(\frac{\ln\frac{F}{V_0} - (r + \frac{1}{2}\sigma_V^2 - \lambda\nu)T - j(\ln\tilde{\nu} + \frac{1}{2}\sigma_U^2)}{\sqrt{\sigma_V^2 T + j\sigma_U^2}}\Big) \\ &\quad - e^{-(r+\lambda)T} \sum_{j=0}^{\infty} \frac{(\lambda T)^j}{j!} \Phi\Big(- \frac{\ln\frac{F}{V_0} - (r - \frac{1}{2}\sigma_V^2 - \lambda\nu)T - j(\ln\tilde{\nu} - \frac{1}{2}\sigma_U^2)}{\sqrt{\sigma_V^2 T + j\sigma_U^2}}\Big).\end{aligned}$$

Proof. The payoff of the bond equals

$$\begin{aligned}\bar{B}_{ZH}(t,T) &= 1_{\{\tau > T\}} + 1_{\{\tau \le T\}}\big(1 - w(V_T/F)\big) \\ &= 1_{\{\tau > T\}} + 1_{\{\tau \le T\}}\tilde{w}\frac{V_T}{F} = 1 + 1_{\{\tau \le T\}}\big(\tilde{w}\frac{V_T}{F} - 1\big).\end{aligned}$$

To compute the present value of the bond we consider the expectation of the discounted payoff

$$\begin{aligned}\bar{B}_{ZH}(t,T) &= \mathbb{E}^Q\Big[e^{-r(T-t)} \cdot \Big(1 + 1_{\{\tau \le T\}}\big(\tilde{w}\frac{V_T}{F} - 1\big)\Big)\Big|\, \mathcal{F}_t\Big] \\ &= e^{-r(T-t)}\Big[1 + \mathbb{E}^Q\Big(1_{\{V_T < F\}}\big(\tilde{w}\frac{V_T}{F} - 1\big)\Big)\Big|\, \mathcal{F}_t\Big] \\ &= e^{-r(T-t)}\Big[1 + \frac{\tilde{w}}{F}\mathbb{E}^Q\Big(1_{\{V_T < F\}} V_T\Big|\, \mathcal{F}_t\Big) - \mathbb{E}^Q\Big(1_{\{V_T < F\}}\Big|\, \mathcal{F}_t\Big)\Big].\end{aligned}$$

Note that conditionally on $\{N_T = j\}$ we obtain a log-normal distribution for V_T:

$$\begin{aligned}
\mathbb{P}(V_T < F | N_T = j) &= \mathbb{P}\Big(V_0 \exp\big[(r - \frac{1}{2}\sigma_V^2 - \lambda\nu)T + \sigma_V W_V(T)\big] \prod_{i=1}^{N_T} U_i < F \big| N_T = j\Big) \\
&= \mathbb{P}\Big(\ln V_0 + (r - \frac{1}{2}\sigma_V^2 - \lambda\nu)T + \sigma_V W_V(T) + \sum_{i=1}^{j} \ln U_i < \ln F\Big) =: \mathbb{P}(\xi_j < \ln F),
\end{aligned}$$

where $\sigma_V W(T) + \sum_{i=1}^j \ln U_i$ as a sum of independent normally distributed random variables is again normally distributed. Recall σ_U^2, the variance of $\ln U_1$. As $\mathbb{E}(\ln U_i) = \ln(1+\nu) - \frac{1}{2}\sigma_U^2$, we get

$$\begin{aligned}
\xi_j &\sim \mathcal{N}\Big(\ln V_0 + (r - \frac{1}{2}\sigma_V^2 - \lambda\nu)T + j(\ln\tilde{\nu} - \frac{1}{2}\sigma_U^2), \sigma_V^2 T + j\sigma_U^2\Big) \\
&=: \mathcal{N}(\tilde{\mu}(j), \tilde{\sigma}^2(j)).
\end{aligned}$$

It is an easy exercise to verify that for $\xi \sim \mathcal{N}(\mu, \sigma^2)$

$$\mathbb{E}\big(e^{\xi} 1_{\{e^{\xi} < F\}}\big) = e^{\mu + \frac{1}{2}\sigma^2} \Phi\Big(\frac{\ln F - \mu}{\sigma} - \sigma\Big).$$

Conclude that

$$\begin{aligned}
\mathbb{E}^Q\Big[1_{\{V_T < F\}} V_T\Big] &= \sum_{j=0}^{\infty} Q(N_T = j) \mathbb{E}^Q(1_{\{V_T < F\}} V_T | N_T = j) \\
&= \sum_{j=0}^{\infty} e^{-\lambda T} \frac{(\lambda T)^j}{j!} \exp(\tilde{\mu}(j) + \frac{1}{2}\tilde{\sigma}^2(j)) \Phi\Big(\frac{\ln F - \tilde{\mu}(j)}{\tilde{\sigma}(j)} - \tilde{\sigma}(j)\Big) \\
&= e^{-\lambda T} V_0 e^{(r - \lambda\nu)T} \sum_{j=0}^{\infty} \frac{(\lambda\tilde{\nu}T)^j}{j!} \\
&\quad \cdot \Phi\Big(\frac{\ln\frac{F}{V_0} - (r + \frac{1}{2}\sigma_V^2 - \lambda\nu)T - j(\ln\tilde{\nu} + \frac{1}{2}\sigma_U^2)}{\sqrt{\sigma_V^2 T + j\sigma_U^2}}\Big).
\end{aligned}$$

We therefore obtain

$$\begin{aligned}
\bar{B}_{ZH}(0,T) &= e^{-rT} + \frac{\tilde{w}}{F} V_0 e^{-\lambda T(1+\nu)} \\
&\quad \cdot \sum_{j=0}^{\infty} \frac{(\lambda\tilde{\nu}T)^j}{j!} \Phi\Big(\frac{\ln\frac{F}{V_0} - (r + \frac{1}{2}\sigma_V^2 - \lambda\nu)T - j(\ln\tilde{\nu} + \frac{1}{2}\sigma_U^2)}{\sqrt{\sigma_V^2 T + j\sigma_U^2}}\Big) \\
&\quad - e^{-(r+\lambda)T} \sum_{j=0}^{\infty} \frac{(\lambda T)^j}{j!} \Phi\Big(\frac{\ln\frac{F}{V_0} - (r - \frac{1}{2}\sigma_V^2 - \lambda\nu)T - j(\ln\tilde{\nu} - \frac{1}{2}\sigma_U^2)}{\sqrt{\sigma_V^2 T + j\sigma_U^2}}\Big).
\end{aligned}$$

Noting that

$$e^{-rT} = e^{-(r+\lambda)T} \sum (\lambda T)^j / (j!),$$

the proof is complete.

■

In the case where no jumps are present, i.e., $\lambda = 0$, the sum reduces to the summand with $j = 0$ so that the bond price formula of Merton (1.1) is obtained as a special case.

This model features some properties which are also found in empirical investigations on credit risk:

- The term structure of the credit spreads can be "upward-sloping", flat, humped or "downward-sloping".
- The "short maturity spreads" can be significantly higher than in the Merton model.
- As the firm value at default is random, especially not equal to F as in the Longstaff and Schwartz (1995) model, the recovery is more realistic.
- The recovery rate is correlated with the firm value also just before default.

1.2.4 Further Structural Models

Kim, Ramaswamy and Sundaresan (1993) extended the first passage time models to also incorporate stochastic interest rates following the model of Cox, Ingersoll and Ross (1985). In their model there is an additional possibility for a default to happen at maturity. The payoff they considered equals $\min(F, V)$. Possibly the company is not able to meet its liabilities at maturity but did not face a default up to this time.

Nielsen, Saà-Requejo and Santa-Clara (1993) extended these models to incorporate a stochastic default boundary. For the interest rate they used the model of Hull and White (1990) but were only able to obtain explicit formulas in the special case of the Vasiček model, cf. formula (1.2). Denote $\sigma_U^2 := \mathrm{Var}(\ln U_1)$ and $\tilde{\nu} := 1 + \nu$.

In the work of Ammann (1999) *vulnerable claims* are considered. These are possibly stochastic payoffs which face a counterparty risk. Counterparty risk plays a role if the buyer of a claim considers the default probability of the seller as significant. He therefore will ask for a risk premium which compensates for a possible loss in case of a default. The default is assumed to happen if $V_T < F$, similar to Merton's model. In that case the buyer of the claim X receives the fraction $\frac{V_T}{F} \cdot X$. Explicit prices are derived for the Heath, Jarrow and Morton (1992) forward rate structure and Merton-like firm dynamics.

This section on structural models heavily relies on the assumption that the firm's value is observable or even tradeable. From a practical point of view this seems not justifiable as the firm's value is not tradeable and even difficult to observe. This difficulty is discussed by Buffett (2002) and also solved in the KMV-model; see Section 1.8.1.

1.3 Hazard Rate Models

In comparison to structural models, intensity based models or hazard rate models use a totally different approach for modeling the default. In the structural approach default occurs when the firm value falls below a certain boundary. The hazard rate approach takes the default time as an exogenous random variable and tries to model or fit its probability to default. The main tool for this is a Poisson process with possibly random intensity λ_t, and jumps denoting the default events. As in the first passage time models recovery is not intrinsic to this model and is often assumed to be a somehow determined constant.

The reason for this new approach lies in the very different causes for default. Precise determination as done in structural models seems to be very difficult. Furthermore, in structural models the calibration to market prices often causes difficulties, while intensity based models allow for a better fit to available market data.

In some approaches basic ideas of these model classes are combined, for example by Madan and Unal (1998) and Ammann (1999) where the default intensity explicitly depends on the firm value. These models are called *hybrid models* and will be discussed in Section 1.6. As the firm value approaches a certain boundary, intensity increases sharply and default becomes very likely. So basic features of the structural models are mimicked.

A more involved hybrid model is presented by Duffie and Lando (2001) where a firm value model with incomplete accounting data is considered.

Basically we may distinguish three types of hazard rate models. In the first approach the default process is assumed to be independent of most economic factors, sometimes it is even modeled independently from the underlying.

The *rating based approach* incorporates the firm's rating as this constitutes readily available information on the company's creditworthiness. In principle one tries to model the company's way through different rating classes up to a possible fall to the lowest rating class which determines the default.

A third and very recent class is in the line of the famous market models of Jamshidian (1997) and Brace, Gatarek and Musiela (1995), see Chapter 1.7.

1.3.1 Mathematical Preliminaries

In this section we consider the modeling of the default process in greater detail. The approach is mainly based on Lando (1994) and also discussed in many articles and books like Jeanblanc (2002) and Bielecki and Rutkowski (2002). We first present a brief introduction to *Cox processes*. More details can be found in Appendix A.

As already mentioned different stopping times denoting the default events need to be modeled. The Poisson process is taken as a starting point. Constant intensity seems too restrictive so one uses Cox processes, which can be considered as Poisson processes with random intensities[6]. A special case which suits well for our purposes is the following:

Consider a stochastic process λ_t which is adapted to some filtration $\mathcal{G}_t$. For a Poisson process $(N_t)_{t\geq 0}$ with intensity 1 independent of $\sigma(\lambda_s : 0 \leq s \leq T^*)$ set

$$\tilde{N}_t := N\Big(\int_0^t \lambda_u \, du\Big), \qquad t \leq T^*.$$

$(\tilde{N}_t)_{t\geq 0}$ is a Cox process. Observe that for positive λ_t the process $\int_0^t \lambda_u \, du$ is strictly increasing and so $\tilde{N}$ can be viewed as a Poisson process under a random change of time. This reveals a very powerful concept for the problems considered in credit risk.

If just one default time τ is considered, this will be equal to the first jump τ_1 of $\tilde{N}_t$. If more default events are considered, for example, transition to other rating classes, further jumps τ_i are taken into account. The bigger λ is, the sooner the next jump may be expected to occur. We obtain, for any $t < T^*$,

$$\begin{aligned} \mathbb{P}(\tau > t) &= \mathbb{E}\big[\mathbb{P}\left(\tau > t|(\lambda_s)_{0\leq s\leq t}\right)\big] \\ &= \mathbb{E}\Big[\exp\big(-\int_0^t \lambda_u \, du\big)\Big]. \end{aligned}$$

Conclude that conditionally on $\sigma(\lambda_s : 0 \leq s \leq T^*)$ the jumps are exponentially distributed with parameter $\int_0^t \lambda_u \, du$.

It may be recalled that a fundamental assumption to obtain this is the independence of λ and N.

1.3.2 Jarrow and Turnbull (1995-2000)

In the work of Jarrow and Turnbull (1995) a binomial model is considered. In extension of the classical Cox., Ross and Rubinstein (1979) approach the authors also modeled the non-default and the default state. So for every time period four possible states may be attained: {up,down} $\times$ {non-default,default}. They discovered an analogy to the foreign-exchange markets. As the intensity of the model is assumed to be constant we do not discuss it in greater detail.

In Jarrow and Turnbull (2000) a Vasiček model for the spot rate is used and the hazard rate is explicitly modeled. Correlation of the hazard rate and spot rates are allowed. Denote by Z_t and W_t Brownian motions under the risk neutral measure Q, with constant correlation ρ. Z_t can be some economic factor, like an index or the logarithm of the firm value.

[6]For a full treatment of Cox processes see Brémaud (1981) and Grandell (1997).

Assume the following dynamics

$$\begin{aligned} dr_t &= \kappa(\theta - r_t)\,dt + \sigma dW_t, \\ \lambda_t &= a_0(t) + a_1(t)r_t + a_2(t)Z_t. \end{aligned}$$

Note that λ may take on negative values with positive probability.

Recovery must be modeled exogenously and the authors use the already mentioned *recovery of treasury value*[7]. This means if default happens prior to maturity of the bond, the bond holder receives a fraction $(1-w)$ of the principal at maturity. For the value of the bond we calculate the expectation of the discounted payoff under the risk-neutral measure Q. For ease of notation we consider $t = 0$. By equation (1.3),

$$\bar{B}(0,T) = (1-w)B(0,T) + w\mathbb{E}^Q\Big[\exp\Big(-\int_0^T r_u\,du\Big)1_{\{\tau>T\}}\Big].$$

In the model of Jarrow and Turnbull we obtain

$$\begin{aligned} \bar{B}(0,T) &= (1-w)B(0,T) + w\mathbb{E}^Q\Big[\exp(-\int_0^T r_u\,du)Q(\tau \leq T|\lambda_s : 0 \leq s \leq T)\Big] \\ &= (1-w)B(0,T) + w\mathbb{E}^Q\Big[\exp[-\int_0^T (r_u + \lambda_u)\,du]\Big] \\ &= (1-w)B(0,T) + w\exp(-\mu_T + \frac{1}{2}v_T). \end{aligned}$$

In the last equation μ_T and v_T denote expectation and variance of $\int_0^T (r_u + \lambda_u)\,du$. Under the stated assumptions this integral is normally distributed and μ and v can be easily calculated.

The flexibility of the model leads to a good fit to market data, which is not obtained by most structural models. Also the model incorporates economic factors (Z_t).

1.3.3 Duffie and Singleton (1999)

The paper by Duffie and Singleton (1999) combines two very successful model classes in interest rate modeling to access Credit Risk: exponential affine models and the Heath, Jarrow and Morton (1992) methodology.

For the exponential affine model the authors model a vector of hidden factors which underlie the term structure of interest rates. This vector is assumed to follow a multidimensional Cox-Ingersoll-Ross model:

$$d\mathbf{y}(t) = \mathbf{K}(\mathbf{\Theta} - \mathbf{y}(t))dt + \mathbf{\Sigma}\ \mathrm{diag}(\mathbf{y}(t))^{1/2}d\mathbf{W}(t).$$

[7] See the Longstaff and Schwartz model, Section 1.2.2.

Consequently the components of $\mathbf{y}$ are nonnegative random numbers. Spot and hazard rate are assumed to be linear in $\mathbf{y}(t)$:

$$\begin{aligned} r(t) &= \delta_0 + \delta'\mathbf{y}(t), \\ \lambda(t)(1-\theta(t)) &= \gamma_0 + \gamma'\mathbf{y}(t). \end{aligned}$$

A main feature of the exponential affine models is that the solution of the above SDE can be explicitly expressed in an exponential affine form. Hence we obtain deterministic functions $a(\cdot), b(\cdot)$ such that

$$\mathbb{E}\Big[\exp\Big(i\boldsymbol{\xi}'\int_0^t \mathbf{y}(u)\,du\Big)\Big] = \exp[a(t,\boldsymbol{\xi}) + b(t,\boldsymbol{\xi}'\mathbf{y}(0))].$$

Thus the price of the defaultable bond can be calculated in closed form as the value of the characteristic function at a proper point.

The second approach uses the well known Heath-Jarrow-Morton model of forward rates. Denote by $\bar{f}(t,T)$ the forward rates determined by the term structure of the defaultable bond prior to default[8] and by $\mathbf{W}(t,T)$ a d-dimensional standard Brownian motion. Assume the dynamics of the forward rate to be

$$\bar{f}(t,T) = \bar{f}(0,T) + \int_0^t \mu(u,T)\,du + \int_0^t \boldsymbol{\sigma}(u,T)\,d\mathbf{W}(u).$$

Similar to Heath, Jarrow and Morton (1992) the authors specify the dynamics under the objective measure and consider an equivalent measure Q. For arbitrage-freeness it is sufficient - see the work of Harrison and Pliska (1981) - that all discounted price processes are martingales. Naturally this heavily relies on the recovery assumption.

Duffie and Singleton (1999) introduced the *recovery of market value* which means that immediately at default the bond loses a fraction of its value. This setup is particularly well suited for working with SDEs. The loss rate w_t is assumed to be an adapted process. Hence

$$\bar{B}(\tau,T) = (1-w_t)\bar{B}(\tau-,T).$$

Under these assumptions the authors derived the following *drift condition* for μ and σ:

$$\mu(t,T) = \sigma(t,T)\Big(\int_t^T \sigma(u,T)\,du\Big)'.$$

On the other hand, using the above mentioned *recovery of treasury value* (cf. 1.2.2) and denoting the riskless forward rate by $f(t,T)$, the authors obtained

$$\mu(t,T) = \sigma(t,T)\Big(\int_t^T \sigma(u,T)\,du\Big)' + \theta(t)\lambda(t)\frac{v(t,T)}{p(t,T)}(\bar{f}(t,T) - f(t,T)).$$

[8]The forward rate is by definition $\bar{f}(t,T) = -\frac{\partial}{\partial T}\ln \bar{B}(t,T)$.

1.4 Credit Ratings Based Methods

Simple hazard rate models are often criticized because they do not incorporate available economic fundamental information like firm value or credit ratings. This section reveals some models which incorporate these data. This is also a basic feature of commercial models; see Section 1.8.

Credit ratings constitute a published ranking of the creditor's ability to meet his obligations. Such ratings are provided by independent agencies, for example Standard & Poor's or Moody's and mostly financed by the gauged companies. The firms are rated even if they are not willing to pay, but for a fee they get detailed insight in the results of the examinations and might retain fundamental insights in their internal divisions to identify weaknesses.

Each rating company uses a different system of letters to classify the creditworthiness of the rated agencies. Standard & Poor's, for example, describes the highest rated debt (triple-A=AAA) with the words "Capacity to pay interest and repay principal is extremely strong". An obligation with the lowest rating, 'D', is in state of default or is not believed to make payments in time or even during a grace period. The lower the rating, the higher is the risk that interest or principal payments will not be made.

1.4.1 Jarrow, Lando and Turnbull (1997)

The model proposed by Jarrow, Lando and Turnbull (1997) circumvents some disadvantages of the hitherto introduced models. Especially the use of credit ratings is an attractive feature. The movements between the single rating classes is modeled by a time homogenous Markov chain, the entry into the lowest rating class yielding a default. For example, if a bond is rated AAA, it is a member of the highest rating class (= class 1). If there exist $K-1$ rating classes, denote by K the class of default. Default is assumed to be an absorbing state, restructuring after default is not considered in this model. The generator of the Markov chain is defined as

$$\mathbf{\Lambda} = \begin{pmatrix} -\lambda_1 & \lambda_{12} & \lambda_{13} & \cdots & \lambda_{1K} \\ \lambda_{21} & -\lambda_2 & \lambda_{23} & \cdots & \lambda_{2K} \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ \lambda_{K-1,1} & \lambda_{K-1,2} & \cdots & -\lambda_{K-1} & \lambda_{K-1,K} \\ 0 & 0 & \cdots & \cdots & 0 \end{pmatrix}.$$

The transition rates for the first rating class are in the first row. So $\lambda_1 = \sum_{j\neq 1} \lambda_{1j}$ is the rate for leaving this class, while λ_{12} is the rate for downgrading to class 2 and so on. The rate for a default directly from class one is λ_{1K}.

We denote

$$q_{ij}(0,t) := \mathbb{P}(\text{Rating is in class } i \text{ at } 0 \text{ and in class } j \text{ at } t),$$

and by $\mathbf{Q}(t)$ the matrix of the transition probabilities $q_{ij}(0,t)$.

The transition probabilities can be computed from the intensity matrix via[9]

$$\mathbf{Q}(t) = \exp(t\mathbf{\Lambda}) := \mathrm{id}_n + t\mathbf{\Lambda} + \frac{1}{2!}(t\mathbf{\Lambda})^2 + \frac{1}{3!}(t\mathbf{\Lambda})^3 + \ldots,$$

where id_n is the $n \times n$ identity-Matrix.

Under the recovery of treasury assumption[10] we obtain for the price of a zero coupon bond under default risk

$$\begin{aligned}
\bar{B}(t,T) &= 1_{\{\tau>t\}}\mathbb{E}_t\Big[e^{-\int_t^\tau r_s\,ds}\cdot\delta B(\tau,T)1_{\{\tau\le T\}} + e^{-\int_t^T r_s\,ds}\cdot 1_{\{\tau>T\}}\Big] \\
&= 1_{\{\tau>t\}}\mathbb{E}_t\Big[\delta 1_{\{\tau\le T\}}e^{-\int_t^T r_s\,ds} + 1_{\{\tau>T\}}e^{-\int_t^T r_s\,ds}\Big] \\
&= 1_{\{\tau>t\}}\Big[\delta B(t,T) + \mathbb{E}_t\Big((1-\delta)e^{-\int_t^T r_s\,ds}1_{\{\tau>T\}}\Big)\Big] \\
&= 1_{\{\tau>t\}}B(t,T)\Big[\delta + (1-\delta)Q_t^T(\tau>T)\Big]. \qquad (1.5)
\end{aligned}$$

Q^T is the T-forward measure[11]. It is therefore crucial to have a model which determines the transition probabilities under this measure. While rating agencies estimate the transition probabilities using historical observations, i.e., under the objective measure P, Jarrow, Lando and Turnbull (1997) propose a method which uses the defaultable bond prices and calculates transition probabilities under the the risk-neutral measure Q.

Consider the bond with rating "i" and set $Q_t^{T,i}(\tau > T)$ the probability that the bond will not default until T given it is rated "i" at t. As it makes no sense to talk about bond prices after default, we further on just consider the bond price on $\{\tau > t\}$ and get

$$\bar{B}_i(t,T) = B(t,T)\Big(\delta + (1-\delta)Q_t^{T,i}(\tau > T|\tau > t)\Big). \qquad (1.6)$$

Jarrow, Lando and Turnbull (1997) split the intensity matrices into an empirical part (under P) and a risk adjustment like a market price of risk: They assume that the intensities under Q^T have the form $\mathbf{U\Lambda}$ and $\mathbf{U}$ denotes a diagonal matrix where the entries are the risk adjusting factors μ_i. For the transition probabilities this yields that $q_{ij}(t,T)$ is the ij'th entry of the matrix $\exp(\mathbf{U\Lambda})$. Time homogeneity of μ would entail exact calibration being impossible.

For the discrete time approximation, $[0,T]$ is divided into steps of length 1. Starting with (1.6) one obtains

$$\begin{aligned}
Q_t^{T,i}(\tau \le T|\tau > t) &= 1 - \frac{B_i(t,T) - \delta B(t,T)}{(1-\delta)B(t,T)} \\
&= \frac{B(t,T) - \bar{B}_i(t,T)}{B(t,T)(1-\delta)}. \qquad (1.7)
\end{aligned}$$

[9]See, for example, Israel, Rosenthal and Wei (2001).

[10]The bond holder receives δ equivalent and riskless bonds in case of default. See Section 1.2.2.

[11] The T-forward measure is the risk neutral measure which has the risk-free bond with maturity T as numeraire. For details see Björk (1997).

Denote the empirical probabilities from the rating agency by $p_{ij}(t,T)$. This leads to $Q_0^{T,i}(\tau \leq 1) = \mu_i(0) p_{iK}(0,1)$, and we obtain

$$\mu_i(0) \quad = \quad \frac{Q_0^{T,i}(\tau \leq 1)}{p_{iK}(0,1)} = \frac{B(0,1) - \bar{B}_i(0,1)}{p_{iK}(0,1) \cdot B(0,1)(1-\delta)}.$$

By this one obtains $(\mu_1, \ldots, \mu_{K-1})'$ and consequently $q_{ij}(0,1)$. For the step from t to $t+1$ use

$$Q_0^{T,i}(\tau \leq t+1) = Q_0^{T,i}(\tau \leq t+1 | \tau > t) \cdot Q_0^{T,i}(\tau > t)$$

to get

$$\begin{aligned} Q_0^{T,i}(\tau \leq t+1) &= \mu_i(t) P^i(\tau \leq t+1 | \tau > t) \cdot \sum_{j=1}^{K-1} q_{ij}(0,t) \\ &= \mu_i(t) p_{iK}(t,t+1) \cdot \sum_{j=1}^{K-1} q_{ij}(0,t). \end{aligned}$$

This leads to

$$\begin{aligned} \mu_i(t) &= \frac{Q_0^{T,i}(\tau \leq t+1)}{\sum_{j=1}^{K-1} q_{ij}(0,t) \cdot p_{iK}(t,t+1)} \\ &\overset{(1.7)}{=} \frac{B(0,t+1) - \bar{B}_i(0,t+1)}{B(0,t+1)(1-\delta)\left(\sum_{j=1}^{K-1} q_{ij}(0,t)\right) p_{iK}(t,t+1)}, \end{aligned}$$

and, via $q_{ij}(0,t+1) = \mu_i(t) p_{ij}(0,t+1)$, the required probabilities are obtained.

This model extends Jarrow and Turnbull (1995) using time dependent intensities but still working with constant recovery rates. Das and Tufano (1996) propose a model which also allows for correlation between interest rates and default intensities.

It seems problematic that all bonds with the same rating automatically have the same default probability. In reality this is definitely not the case. Naturally different credit spreads occur for bonds with the same rating.

A further restrictive assumption is the time independence of the intensities. The yield of a bond in this model may only change if the rating changes. Usually the market price precedes the ratings with informations on a possible rating change which is an important insight of the KMV model; see Section 1.8.1.

1.4.2 Lando (1998)

The work of Lando (1998) uses a conditional Markov chain[12] to describe the rating transitions of the bond under consideration. All available market information like interest rates, asset values or other company specific information is modeled as a stochastic process $(X_t)_{t \geq 0}$. This is analogous to the case without ratings, where Lando used $\lambda_t = \lambda(X_t)$.

[12]See also Section 11.3 in Bielecki and Rutkowski (2002).

Assume that a risk-neutral martingale measure Q is already chosen. Then the arbitrage-free price of a contingent claim is the conditional expectation under this measure Q. The author lays out the framework for rating transitions where all probabilities are already under the risk-neutral measure and calibrates them to available market prices. As no historical information is used the probability distribution under the objective measure is not needed. If one wants to consider risk-measures like Value-at-Risk, note that the objective measure is still required.

We denote the generator of the conditional Markov chain $(C_t)_{t\geq 0}$ by

$$\mathbf{\Lambda}(s) = \begin{pmatrix} -\lambda_1(s) & \lambda_{12}(s) & \lambda_{13}(s) & \cdots & \lambda_{1K}(s) \\ \lambda_{21}(s) & -\lambda_2(s) & \lambda_{23}(s) & \cdots & \lambda_{2K}(s) \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ \lambda_{K-1,1}(s) & \lambda_{K-1,2}(s) & \cdots & -\lambda_{K-1}(s) & \lambda_{K-1,K}(s) \\ 0 & 0 & \cdots & \cdots & 0 \end{pmatrix},$$

where for all s

$$\lambda_i(s) = \sum_{j=1,\, j\neq i}^{K} \lambda_{ij}(s), \quad i = 1, \ldots, K-1.$$

We assume $(\lambda_{ij}(t))_{t\geq 0}$ to be adapted and nonnegative processes.

It is important for the intensities to depend on both time and interest rates. Especially for low rated companies the default rates vary considerably over time[13]. It was observed by Duffee (1999), e.g., that default rates significantly depend on the term structure of interest rates. It is certainly bad news for companies with high debt when interest rates increase whereas for other companies it might be good news.

The construction of $(C_t)_{t\geq 0}$ can be done as follows. Consider a series of independent exponential(1)-distributed random variables $E_{11}, \ldots, E_{1K}, E_{21}, \ldots, E_{2K}, \ldots$ which are also independent of $\sigma(\mathbf{\Lambda}(s) : s \geq 0)$ and denote the rating class of the company at the beginning of the observation by η_0. Define

$$\tau_{\eta_0,i} := \inf\{t : \int_0^t \lambda_{\eta_0,i}(s)\, ds \geq E_{1i}\}, \quad i = 1, \ldots, K$$

and

$$\tau_0 := \min_{i\neq\eta_0} \tau_{\eta_0,i}, \qquad \eta_1 := \arg\min_{i\neq\eta_0} \tau_{\eta_0,i}.$$

The $\tau_{\eta_0,i}$ model the possible transitions to other rating classes starting from rating η_0. The first transition to happen determines the transition that really takes place, compare Figure 1.2. The reached rating class is denoted by η_1 while τ_0 denotes the time at which this occurs. Analogously, the next change in rating starting in η_1 is defined, and similarly for η_i and τ_i. Then, for $\tau_{i-1} \leq t < \tau_i$, C_t is defined by $C_t := \eta_i$.

Default is assumed to be an absorbing state of the Markov chain and we denote the overall-time to default by τ. This is the first time when $\eta_i = K$.

[13]Cf. Chapter 15 in Caouette, Altmann and Narayanan (1998).

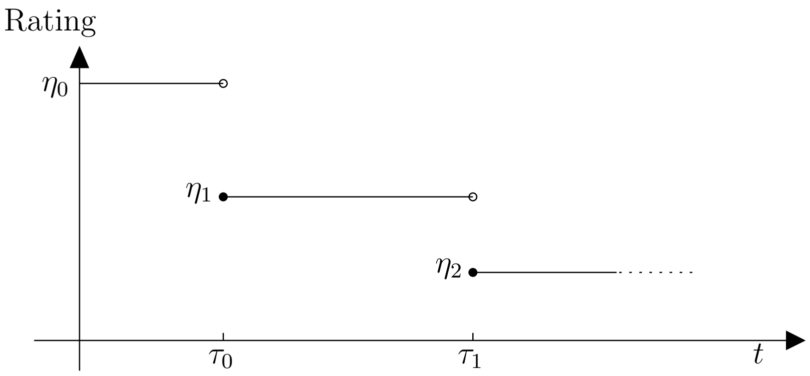

Figure 1.2: A possible realization of rating transitions. The rating starts in η_0 and drops to η_1 at τ_0. The next change is at τ_1, to rating class η_2.

The transition probabilities $P(s,t)$ for the time interval (s,t) satisfy Kolmogorov's backward differential equation[14]

$$\frac{\partial P_X(s,t)}{\partial s} = -\mathbf{\Lambda}(s)\, P_X(s,t).$$

Consider the price of a defaultable zero recovery bond at time t, $\bar{B}^i(t,T)$, which has maturity T and is rated in class i at time t. Then we obtain the following Theorem.

Theorem 1.4.1. *Under the above assumptions the price of the defaultable bond equals*

$$\bar{B}^i(t,T) = 1_{\{C_t=i\}}\mathbb{E}\Big(\exp\big(-\int_t^T r_s\, ds\big)(1 - P_X(t,T)_{i,K})\Big|\, \mathcal{F}_t\Big).$$

Here $P_X(t,T)_{i,K}$ is the (i,K)-th element of the matrix of transition probabilities for the time interval (t,T), $P_X(t,T)$.

Proof. As already mentioned the Markov chain is modeled under Q so that the arbitrage-free price of the bond is the following conditional expectation:

$$\bar{B}^i(t,T) = \mathbb{E}\Big(\exp\big(-\int_t^T r_s\, ds\big)1_{\{\tau>T,C_t=i\}}\Big|\, \mathcal{F}_t\Big).$$

Using conditional expectations and the independence of E_{1K} and $(\mathbf{\Lambda}(s))$ one concludes

$$\begin{aligned}\bar{B}^i(t,T) &= 1_{\{C_t=i\}}\mathbb{E}\Big(\exp\big(-\int_t^T r_s\, ds\big)\mathbb{P}\big(\tau > T\big|\sigma(\mathbf{\Lambda}_s : 0 \le s \le T) \vee \mathcal{F}_t\big)\Big|\, \mathcal{F}_t\Big)\\ &= 1_{\{C_t=i\}}E\Big(\exp\big(-\int_t^T r_s\, ds\big)(1 - P_X(t,T)_{i,K})\Big|\, \mathcal{F}_t\Big).\end{aligned}$$

∎

[14]For non-commutative $\mathbf{\Lambda}$ the solution is in general not of the form $P_X(s,t) = \exp\int_s^t \mathbf{\Lambda}(u)\, du$. See Gill and Johannsen (1990) for solutions using product integrals.

For the calibration to observed credit spreads explicit formulas are needed and therefore further assumptions will be necessary. Lando chooses an Eigenvalue-representation of the generator.

Denote with $\mathbf{A}(s)$ the matrix with entries $\lambda_1(s), \dots, \lambda_{K-1}(s), 0$ on the diagonal and zero otherwise. Assume that $\mathbf{\Lambda}(s)$ admits the representation

$$\mathbf{\Lambda}(s) = \mathbf{B}\,\mathbf{A}(s)\,\mathbf{B}^{-1},$$

where $\mathbf{B}$ is the $K \times K$-matrix of the Eigenvectors of $\mathbf{\Lambda}(s)$.

We conclude $P_X(s,t) = \mathbf{B}\,\mathbf{C}(s,t)\,\mathbf{B}^{-1}$ with

$$\mathbf{C}(s,t) = \begin{pmatrix} \exp\int_s^t \lambda_1(u)du & 0 & \cdots & 0 \\ 0 & \ddots & \cdots & \vdots \\ \vdots & \cdots & \exp\int_s^t \lambda_{K-1}(u)du & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}.$$

It is easy to see that $P_X(s,t)$ satisfies the Kolmogorov-backward differential equation. For uniqueness, see Gill and Johannsen (1990).

Under these additional assumptions the price of the defaultable bond in Theorem 1.4.1 simplifies considerably.

Proposition 1.4.2. *Denoting by $\beta_{ij} := \mathbf{B}_{ij}\mathbf{B}^{-1}_{jK}$, the price of the defaultable bond equals*

$$\bar{B}^i(t,T) = \sum_{j=1}^{K-1} -\beta_{ij}\mathbb{E}\Big[\exp\Big(\int_t^T (\lambda_j(u) - r_u)\,du\Big)\Big|\,\mathcal{F}_t\Big].$$

Proof. In this setup the conditional probability for a default when the bond is in rating class i equals

$$\mathbb{P}_X(t,T)_{i,K} \;=\; 1_{\{\tau>t\}}\sum_{j=1}^{K}\mathbf{B}_{ij}\,\exp(\int_t^T \lambda_j(u)du)\mathbf{B}^{-1}_{jK}.$$

With $\mathbf{B}_{iK}\mathbf{B}^{-1}_{KK} = 1$ we obtain

$$1 - \mathbb{P}_X(t,T)_{i,K} = \sum_{j=1}^{K-1} -\mathbf{B}_{ij}\mathbf{B}^{-1}_{jK}\,\exp(\int_t^T \lambda_j(u)du)$$

and the conclusion follows as in 1.4.1. ∎

Using the readily available tools for hazard rate models it is now easy to consider options which explicitly depend on the credit rating or credit derivatives with a credit trigger.

Calibration

Assuming a Vasiček model[15] for the interest rate we are in the position to use the model laid out above for calibration to observed credit spreads. There are no economic factors considered other than the interest rate and, as a consequence, λ_t must be adapted to $\mathcal{G}_t = \sigma(r_s : 0 \le s \le t)$.

Furthermore, we assume

$$\lambda_j(s) = \gamma_j + \kappa_j r_s, \quad j = 1, \dots, K-1,$$

with constants γ_j, κ_j.

The dynamics of the generator matrix is $\mathbf{\Lambda}(s) = \mathbf{B}\,\mathbf{A}(s)\,\mathbf{B}^{-1}$ and $\mathbf{B}$ has to be estimated from historical data while γ_j, κ_j are calibrated.

The credit spread is the difference of the offered yield to the spot rate. By Theorem 1.4.1 the bond price satisfies

$$\bar{B}^i(t,T) = -\sum_{j=1}^{K-1} -\beta_{ij}\mathbb{E}\Big[\exp\Big(\int_t^T (\gamma_j - (1-\kappa_j)r_u)\,du\Big)\Big|\,\mathcal{F}_t\Big].$$

Therefore, we obtain for the bond's yield

$$\begin{aligned}
-\frac{\partial}{\partial T}\Big|_{T=t} \log \bar{B}^i(t,T) &= -\frac{\partial}{\partial T}\Big|_{T=t} \sum_{j=1}^{K-1} \beta_{ij}\mathbb{E}\Big[\exp\Big(\int_t^T (\gamma_j - (1-\kappa_j)r_u)\,du\Big)\Big|\,\mathcal{F}_t\Big] \\
&= -\sum_{j=1}^{K-1} \beta_{ij} \lim_{T\to t} \mathbb{E}\Big[(\gamma_j + (\kappa_j - 1)r_T)\exp\Big(\int_t^T (\gamma_j + \kappa_j r_u - r_u)\,du\Big)\Big|\,\mathcal{F}_t\Big] \\
&= -\sum_{j=1}^{K-1} \beta_{ij}(\gamma_j + (\kappa_j - 1)r_t).
\end{aligned}$$

Hence the credit spread equals

$$s^i(t) = -\sum_{j=1}^{K-1} \beta_{ij}(\gamma_j + \kappa_j r_t).$$

For calibration a second relation is needed. Lando uses the sensitivity of the credit spreads w.r.t. the spot rate:

$$\frac{\partial}{\partial r_t} s^i(t) = -\sum_{j=1}^{K-1} \beta_{ij}\kappa_j.$$

Denote by $\hat{s}_0, d\hat{s}_0$ the observed credit spreads and their estimated sensitivities. One finally has to solve the following equation to calibrate the model:

$$\begin{aligned}
-\beta(\gamma + \kappa r_0) &= \hat{s}_0 \\
-\beta\kappa &= d\hat{s}_0.
\end{aligned}$$

[15] see equation (1.2).

It turns out to be problematic that observed credit spreads are not always monotone with respect to the ratings. The author argues that in practice this would occur rather seldom.

1.5 Basket Models

Usually there is a whole portfolio under consideration instead of just one single asset. Therefore the so far presented models were extended to models which may handle the behavior of a larger number of individual assets with default risk, a so-called *portfolio* or *basket*.

There are several approaches in the literature and they can be grouped into models which use a conditional independence concept and others which are based on copulas.

From the first class we present the methods of Kijima and Muromachi (2000), which provide a pricing formula for a credit derivative on baskets with a first- or second-to-default feature. An example is the first-to-default put, which covers the loss of the first defaulted asset in the considered portfolio, see also Section 1.9.5. From the second class we discuss an implementation based on the normal copula in Section 1.5.2.

Besides that, Jarrow and Yu (2001) model a kind of direct interaction between default intensities of different companies. In their model the default of a *primary* company has some impact on the hazard rate of a *secondary* company, whose income significantly depends on the primary company.

1.5.1 Kijima and Muromachi (2000)

Consider a portfolio of n defaultable bonds and denote by τ_i the default time of the i-th bond. Let $(\mathcal{G}_t)_{t\geq 0}$ represent the general market information (see Appendix A). Furthermore assume that for any $t_1, \ldots, t_n \leq T$

$$Q(\tau_1 > t_1, \ldots, \tau_n > t_n | \mathcal{G}_T) \quad = \quad Q(\tau_1 > t_1 | \mathcal{G}_T) \cdot \cdots \cdot Q(\tau_n > t_n | \mathcal{G}_T), \tag{1.8}$$

where Q is assumed to be the unique risk neutral measure. Using the representation via Cox processes, this yields

$$(1.8) = \exp(-\sum_{i=1}^{n} \int_0^{t_i} \lambda_i(s)\, ds).$$

In the recovery of treasury model, the loss of bond i upon default equals the pre-specified constant $w_i := (1 - \delta_i)$. So the first-to-default put is the option which pays w_i if the ith asset is the first one to default before T and zero if there is no default. Denote the event that the first defaulted bond is number i by

$$D_i := \{\tau_i \leq T, \tau_j > \tau_i, \forall j \neq i\}.$$

Then, using the risk neutral valuation principle, the price of the bond can be computed as the expectation w.r.t. the risk-neutral measure Q and equals

$$\begin{aligned}\bar{S}_F &= \mathbb{E}\Big[\exp(-\int_0^T r_u\,du)\sum_{i=1}^n w_i 1_{A_i}\Big]\\ &= \sum_{i=1}^n w_i \mathbb{E}\big[\exp(-\int_0^T r_u\,du) Q(A_i|\mathcal{G}_T)\big].\end{aligned}$$

We obtain this probability using the factorization

$$\begin{aligned}&\mathbb{P}(\tau_i \le T, \tau_k > \tau_i,\, \forall k \ne i|\mathcal{G}_T \vee \{\tau_i = x\})\\ &\qquad = 1_{\{x\le T\}}\mathbb{P}(\tau_k > x,\, \forall k \ne i|\mathcal{G}_T \vee \{\tau_i = x\})\\ &\qquad = 1_{\{x\le T\}}\exp(-\sum_{k\ne i}\int_0^x \lambda_k(s)\,ds).\end{aligned}$$

With Theorem A.1.2 we obtain

$$\begin{aligned}&\mathbb{P}(\tau_i \le T, \tau_k > \tau_i,\, \forall k \ne i|\mathcal{G}_T)\\ &\quad = \mathbb{E}\Big[1_{\{\tau_i\le T\}}\exp(-\sum_{k\ne i}\int_0^{\tau_i}\lambda_k(s)\,ds)|\mathcal{G}_T\Big]\\ &\quad = \mathbb{E}\Big[\int_0^T \lambda_i(u)\exp(-\int_0^u \lambda_i(s)\,ds)\exp(-\sum_{k\ne i}\int_0^u \lambda_k(s)\,ds)\,du\Big]\\ &\quad = \int_0^T \mathbb{E}\big[\lambda_i(u)\exp(-\int_0^u\sum_{k=1}^n \lambda_k(s)\,ds)\big]\,du.\end{aligned}$$

We conclude for the price of the first-to-default put:

$$\bar{S}_F = \sum_{i=1}^n w_i \int_0^T \mathbb{E}\Big[\lambda_i(u)\exp(-\int_0^T r_s\,ds - \sum_{k=1}^n\int_0^u \lambda_k(s)\,ds)\Big]\,du.$$

This formula simplifies considerably if $w_i \equiv w$, as in that case

$$\begin{aligned}\bar{S}_F &= w\mathbb{E}\Big[\int_0^T\sum_{i=1}^n \lambda_i(u)\exp(-\int_0^u\sum_{k=1}^n\lambda_k(s)\,ds)\,du\exp(-\int_0^T r_s\,ds)\Big]\\ &= w\mathbb{E}\Big[\big(-\exp(-\sum_{i=1}^n\int_0^T \lambda_i(u)\,du)\big)\Big|_0^T\cdot\exp(-\int_0^T r_s\,ds)\Big]\\ &= (1-\delta)B(0,T)\big[1-\mathbb{E}^T\big(\exp(-\int_0^T\sum_{i=1}^n\lambda_i(u)\,du)\big)\big].\end{aligned}$$

Using similar methods, we determine the swap-price, if w_i is paid immediately at default to the swap-holder. Set

$$\bar{S}_F^* \;=\; \mathbb{E}\Big[\exp(-\int_0^{\tau} r_u\,du)\cdot\sum_{i=1}^{n} w_i 1_{A_i}\Big].$$

Certainly, $\int_0^{\tau} r_u\,du$ is not $\mathcal{G}_T$-measurable, so that a slight modification of the previously used method is necessary. We obtain for the factorization

$$\begin{aligned}
\mathbb{E}\big[\exp(-\int_0^{x} r_u\,du)1_{\{x\le T\}}1_{\{\tau_k>x,\forall k\neq i\}}\big|\mathcal{G}_T\vee\{\tau_i=x\}\big] \\
= 1_{\{x\le T\}}\exp\Big[-\int_0^{x}(r_u+\sum_{k\neq i}\lambda_k(u))\,du\Big]
\end{aligned}$$

and conclude

$$\bar{S}_F^* \;=\; \sum_{i=1}^{n} w_i\int_0^{T}\mathbb{E}\Big[\lambda_i(u)\exp[-\int_0^{u}(r_s+\sum_{k=1}^{n}\lambda_k(s))\,ds]\Big]\,du.$$

Similarly, the authors provide the following price of a (first and) second-to-default swap, which protects the holder against the first two defaults in the portfolio:

$$\begin{aligned}
\bar{S}_S \;=\; & \sum_{i=1}^{n}\delta_i\mathbb{E}\big[\exp(-\int_0^{T}\lambda_i(s)\,ds)\big]-B(0,T)\sum_{i=1}^{n}\delta_i \\
& +\sum_{i\neq j}(\delta_i+\delta_j)\int_0^{T}\mathbb{E}\Big[\lambda_k(u)\exp(-\int_0^{T}r_s\,ds-\sum_{j=1}^{n}\int_0^{u}\lambda_j(s)\,ds)\big]\,du \\
& -(n-2)\sum_{i=1}^{n}\delta_i\int_0^{T}\mathbb{E}\big[\lambda_i(u)\exp(-\int_0^{T}r_s\,ds-\int_0^{u}\sum_{j=1}^{n}\lambda_j(s)\,ds)\big]\,du
\end{aligned}$$

Extended Vasiček implementation

Kijima and Muromachi (2000) discuss a special case of the above implementation. The main idea is to perform a calibration similar to the one of Hull and White (1990) for credit risk models. Assume for the dynamics of the hazard rates

$$d\lambda_i(t) = \big(\phi_i(t)-a_i\lambda_i(t)\big)\,dt+\sigma_i\,dw_i(t), \qquad i=1,\dots,n, \tag{1.9}$$

where w_i are standard Brownian motions with correlation ρ_{ij}, which is sometimes stated as $dw_i dw_j=\rho_{ij}\,dt$. Furthermore, assume for the short rate r_t

$$dr_t = \big(\phi_0(t)-a_0 r_t\big)\,dt+\sigma_0\,dw_0(t).$$

Note that equations of the type (1.9) admit explicit solutions, see Schmidt (1997). From this, we get

$$\lambda_i(t) = \lambda_i(0)e^{-a_i t} + \int_0^t \phi_i(s)e^{-a_i(t-s)}\,ds + \sigma_i \int_0^t e^{-a_i(t-s)}\,dw_i(s).$$

Using the recovery of treasure assumption the bond price equals

$$\bar{B}_i(0,t) = \delta_i B(0,t) + (1-\delta_i)\mathbb{E}\big[\exp(-\int_0^t (r_u + \lambda_i(u))\,du)\big].$$

Note that $\int (r_u + \lambda_i(u))\,du$ is normally distributed and therefore the expectation equals the Laplace transform of a normal random variable with mean

$$\begin{aligned}\mathbb{E}\big[-\int_0^t (r_u + \lambda_i(u))\,du\big] &= -\int_0^t \big(r_0 e^{-a_0 u} + \int_0^u (\phi_0(s)e^{-a_0(u-s)})\,ds\big)\,du \\ &\quad -\int_0^t \big(\lambda_i(0)e^{-a_i u} + \int_0^u (\phi_i(s)e^{-a_i(u-s)})\,ds\big)\,du\end{aligned}$$

and variance

$$\begin{aligned}&\operatorname{Var}\big[\int_0^t (r_u + \lambda_i(u))\,du\big] \\ &\qquad = \operatorname{Var}\Big[\int_0^t \sigma_0 \int_0^u e^{-a_0(u-s)} dz_0(s)\,du + \int_0^t \sigma_i \int_0^u e^{-a_i(u-s)} dw_i(s)\,du\Big].\end{aligned}$$

To compute the variances it is sufficient to calculate the variances of all summands and the covariances. Setting $\rho_{ii} = 1$, we have

$$\begin{aligned}&\mathbb{E}\Big[\int_{u_1=0}^t \int_{u_2=0}^t \sigma_i\sigma_j \int_0^{u_1}\int_0^{u_2} \exp(-a_i(u_1-s_1) - a_j(u_2-s_2))\,dw_j(s_2)\,dw_i(s_1)\,du_2\,du_1\Big] \\ &= \sigma_i\sigma_j\mathbb{E}\Big[\int_0^t\int_0^t\int_0^{s_1}\int_0^{s_2} \exp(-a_i(u_1-s_1) - a_j(u_2-s_2))\,du_2\,du_1\,dw_j(s_2)\,dw_i(s_1)\Big] \\ &= \sigma_i\sigma_j\mathbb{E}\Big[\int_0^t\int_0^t e^{a_i s_1 + a_j s_2}\frac{1}{a_i a_j}(1-e^{-a_i s_1})(1-e^{-a_j s_2})\,dw_j(s_2)\,dw_i(s_1)\Big] \\ &= \sigma_i\sigma_j\rho_{ij}\int_0^t e^{a_i s + a_j s}\frac{1}{a_i a_j}(1-e^{-a_i s})(1-e^{-a_j s})\,ds \\ &= \frac{\sigma_i\sigma_j\rho_{ij}}{a_i a_j}\Big[t + \frac{1}{a_i}(e^{-a_i t}-1) + \frac{1}{a_j}(e^{-a_j t}-1) + \frac{1}{a_i+a_j}(1-e^{-(a_i+a_j)t})\Big] \\ &=: c_{ij}(t)\end{aligned}$$

Therefore,

$$\begin{aligned}\operatorname{Var}\Big[\int_0^t \sigma_i \int_0^u e^{-a_i(u-s)} dw_i(s)\, du\Big] \\ &= \frac{\sigma_i^2}{a_i^2}\Big[t + \frac{2}{a_i}(e^{-a_i t} - 1) + \frac{1}{2a_i}(1 - e^{-2a_i t})\Big] =: v_2(t).\end{aligned}$$

Recall that we want to calibrate the model to the bond prices, which means calculating $\phi_i(s)$. $\phi_0(s)$ is computed as in the risk neutral case, see Hull and White (1990). Consider

$$\frac{1}{B(0,t)}\mathbb{E}\big[\exp(-\int_0^t (r_u + \lambda_i(u))\, du)\big] = \frac{1}{1-\delta_i}\Big[\frac{B_i(0,t)}{B(0,t)} - \delta_i\Big] =: \gamma_i(t),$$

which can be obtained from available prices, since δ_i is assumed to be known. Note that $\gamma_i(t)$ does not involve $\phi_0(s)$ as

$$\begin{aligned}\gamma_i(t) &= \exp\Big[-\int_0^t \big(\lambda_i(0)e^{-a_i u} + \int_0^u \phi_i(s)e^{-a_i(u-s)}\, ds\big)\, du \\ &\quad + \frac{1}{2}\big(c_{0i}(t) + v_2(t)\big)\Big].\end{aligned}$$

As we want to solve this expression for ϕ_i, we consider the following derivatives:

$$\begin{aligned}-\frac{\partial}{\partial t}\ln\gamma_i(t) &= \lambda_i(0)e^{-a_i t} + \int_0^t \phi_i(s)e^{-a_i(t-s)}\, ds - \frac{1}{2}\Big[c_{0i}(t) + v_2(t)\Big]' \\ &=: g_i(t)\end{aligned}$$

With

$$\frac{\partial}{\partial t}g_i(t) = -a_i\lambda_i(0)e^{-a_i t} + \phi_i(t) - a_i e^{-a_i t}\int_0^t \phi_i(s)e^{a_i s}\, ds - \frac{1}{2}\Big[c_{0i}(t) + v_2(t)\Big]''$$

we conclude

$$\phi_i(t) = \frac{\partial}{\partial t}g_i(t) + a_i g_i(t) + a_i\frac{1}{2}\Big[c_{0i}(t) + v_2(t)\Big]' + \frac{1}{2}\Big[c_{0i}(t) + v_2(t)\Big]''.$$

Hence

$$\begin{aligned}a_i c_{0i}(t)' + c_{0i}(t)'' &= \sigma_0\sigma_i\rho_{0i}\Big[\frac{1}{a_0} - \frac{1}{a_0}e^{-a_0 t} - \frac{1}{a_0}e^{-a_i t} + \frac{1}{a_0}e^{-(a_0+a_i)t}\Big] \\ &\quad + \sigma_0\sigma_i\rho_{0i}\Big[\frac{1}{a_i}e^{-a_0 t} + \frac{1}{a_0}e^{-a_i t} - \frac{a_0 + a_i}{a_0 a_i}e^{-(a_0+a_i)t}\Big] \\ &= \sigma_0\sigma_i\rho_{0i}\Big[\frac{1 - e^{-a_0 t}}{a_0} + e^{-a_0 t}\frac{1 - e^{-a_i t}}{a_i}\Big]\end{aligned}$$

and

$$\begin{aligned} a_i v_2(t)' + v_2(t)'' &= \sigma_i^2 \Big[\frac{1}{a_i} - \frac{2}{a_i} e^{-a_i t} - \frac{1}{a_i} e^{-2a_i t} + \frac{2}{a_i} e^{-a_i t} + \frac{2}{a_i} e^{-2a_i t}\Big] \\ &= \frac{\sigma_i^2}{a_i}\Big[1 + e^{-2a_i t}\Big] \end{aligned}$$

which finally leads to

$$\phi_i(t) = \frac{\partial}{\partial t} g_i(t) + a_i g_i(t) + \frac{\sigma_i^2}{2a_i}(1 - e^{-2a_i t}) + \frac{1}{2}\sigma_0 \sigma_i \rho_{0i}\Big[\frac{1 - e^{-a_0 t}}{a_0} + e^{-a_0 t}\frac{1 - e^{-a_i t}}{a_i}\Big].$$

Using similar methods Kijima and Muromachi (2000) obtain an explicit formula for the first-to-default swap. In Kijima (2000) these methods are extended to pricing a credit swap on a basket, which might incorporate a first-to-default feature.

1.5.2 Copula Models

The concept of copulas is well known in statistics and probability theory, and has been applied to finance quite recently. Modeling dependent defaults using copulas can be found, for example, in Li (2000) or Frey and McNeil (2001). We give an outline of Schmidt and Ward (2002), who apply a special copula, the normal copula, to the pricing of basket derivatives.

Fix $t = 0$. The goal of the model is to present a calibration method. Consider the default times $\tau_1, \dots, \tau_n$ and assume for the beginning that $t = 0$. The link between the marginals $Q_i(t) := Q(\tau_i \le t)$ and the joint distribution is the so-called copula $C(t_1, \dots, t_n)$. Assuming continuous marginals, $U_i := Q_i(\tau_i)$ is uniformly distributed. The joint distribution of the transformed random times is the copula

$$C(u_1, \dots, u_n) := Q(U_1 \le u_1, \dots, U_n \le u_n)$$

and defines the joint distribution of the τ_i's via

$$Q(\tau_1 \le t_1, \dots, \tau_n \le t_n) = C\big(Q_1(t_1), \dots, Q_n(t_n)\big).$$

For more detailed information on copulas see Nelsen (1999).

The choice of the copula certainly depends on the application. Schmidt and Ward (2002) choose the normal copula because in a Merton framework with correlated firm value processes such a dependence is obtained, and secondary the normal copula is determined by correlation coefficients which can be estimated from data.

Assume that $(Y_1, \dots, Y_n)$ follows an $n-$dimensional normal distribution with correlation matrix $\mathbf{\Sigma} = (\rho_{ij})$, where $\rho_{ii} = 1$ for all i. Denoting their joint distribution function by $\Phi_n(y_1, \dots, y_n, \mathbf{\Sigma})$ yields the *normal copula*

$$C(u_1, \dots, u_n) = \Phi_n\big(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_n)\big).$$

For modeling purposes it is useful to note that setting

$$\tau_i := Q_i^{-1}(\Phi(Y_i)),$$

results in $\{\tau_1, \dots, \tau_n\}$ having a normal copula with correlation matrix $\mathbf{\Sigma}$.

The above methods enable us to calculate the joint distribution of n default times, and the required correlations can be estimated using historical data. Thus, a value at risk can be determined.

For the pricing of a derivative with first-to-default feature, note that

$$Q(\tau^{1\text{st}} \leq T) = 1 - Q(\tau_1 > T, \dots, \tau_n > T) \tag{1.10}$$

which can be calculated from the copula and the marginals. A more involved, but also explicit formula can be obtained for a kth-to-default option.

For example, consider a *first-to-default swap*, which is also discussed in Section 1.9.5. This is a derivative which offers default protection against the first defaulted asset in a specified portfolio. Under the assumption, that all credits have the same recovery rate $\delta_i \equiv \delta$, the swap pays $(1-\delta)$ at $\tau^{1\text{st}}$ if $\tau^{1\text{st}} \leq T$. In exchange to this, the swap holder pays the premium S at times $T_1, \dots, T_m$, but at most until $\tau^{1\text{st}}$. As explained in Section 1.9.2, calculating expectations of the discounted cash flows yields the first-to-default swap premium. Thus, using equation (1.14), we obtain

$$S_{1\text{st}} = \frac{(1-\delta)\mathbb{E}\big[\exp(-\int_t^{\tau^{1\text{st}}} r_u\, du) 1_{\{\tau^{1\text{st}} \leq T\}}\big]}{\sum_{i=1}^m \mathbb{E}\big[\exp(-\int_0^{T_i} r_u\, du) 1_{\{\tau^{1\text{st}} > T_i\}}\big]}.$$

To calculate the expectations, the distribution of $\tau^{1\text{st}}$ under any forward measure is needed. Assuming, for simplicity, independence of the default intensity and the risk-free interest rate, one obtains

$$\mathbb{E}\big[\exp(-\int_0^{T_i} r_u\, du) 1_{\{\tau^{1\text{st}} > T_i\}}\big] = B(0, T_i) Q(\tau^{1\text{st}} > T_i).$$

The bond prices are readily available and the probability can be calculated via (1.10), once the copula is determined.

For the second expectation, use

$$\begin{aligned}\mathbb{E}\big[\exp(-\int_t^{\tau^{1\text{st}}} r_u\, du) 1_{\{\tau^{1\text{st}} \leq T\}}\big] \\ = \int_0^T B(0,s)\mathbb{E}\big[\exp(-\int_t^s \lambda_u^{1\text{st}}\, du)\lambda^{1\text{st}}(s)\big]\, ds.\end{aligned}$$

Note that this expectation can be obtained via

$$\begin{aligned}\frac{\partial}{\partial s} Q(\tau^{1\text{st}} > s) &= \frac{\partial}{\partial s}\mathbb{E}\big[\exp(-\int_t^s \lambda_u^{1\text{st}}\, du)\big] \\ &= \mathbb{E}\big[\exp(-\int_t^s \lambda_u^{1\text{st}}\, du)\lambda^{1\text{st}}(s)\big].\end{aligned}$$

Further on, Schmidt and Ward (2002) derive interesting results on spread widening, once a default occurred. For example, if one of two strongly related companies defaults, it might be likely that the remaining one gets into difficulties, and therefore credit spreads increase. It seems interesting that traders have a good intuition on this amount of spread widening, which also could be used as an input parameter to the model, which determines the copulas.

1.6 Hybrid models

Hybrid models incorporate both preceding models, for example the firm value is modeled, and a hazard rate framework is derived within this model.

1.6.1 Madan and Unal (1998)

The approach of Madan and Unal (1998) mimics the behavior of the Merton model in a hazard rate framework. They assume the following structure for the default intensity:

$$\lambda(t) = \frac{c}{\left(\ln \frac{V(t)}{F\cdot B(t)}\right)^2}.$$

Here $V(t)$ denotes the firm value which as in Merton's model is assumed to follow a geometric Brownian motion. $B(t)$ is the discounting factor $\exp(-\int_0^t r_u\, du)$ and F is the amount of outstanding liabilities. If the firm value approaches F the default intensity increases sharply and it is very likely that the bond defaults. As defaults can happen at any time this model is much more flexible than the Merton model. Unlike in Longstaff and Schwartz's model, the default can even happen when the firm value is far above F, though with low probability.

The authors also consider parameter estimation in their model. A closed form solution for the bond price is not available and for calculating the prices of derivatives numerical methods need to be used.

Further hybrid models of this type can be found in Ammann (1999) or Bielecki and Rutkowski (2002).

1.6.2 Duffie and Lando (2000)

The model of Duffie and Lando (2001) accounts for the fact that bond holders receive only imperfect information on the issuer's assets. The approach starts with a structural model for the firm value and assumes that the bond holder obtains observations on the firm value disturbed and only at discrete time points, which leads to a hazard rate model. After presenting the framework proposed by the authors we derive the hazard rate explicitly.

Suppose the firm value can be modeled by a geometric Brownian motion, as in the Merton framework, i.e.,

$$V_t = V_0 \exp((\mu - \frac{\sigma^2}{2})t + \sigma W_t) =: V_0 \exp(mt + \sigma W_t).$$

The firm is operated by equity owners, which have complete information on the firm's assets, represented by $\mathcal{F}_t = \sigma(V_s : 0 \leq s \leq t)$. The first step is to determine the optimal liquidation policy.

Assume that the drift of the firm value is smaller than the risk-free interest rate, $\mu < r$, and further on, the firm generates cash flow at the rate δV_t for some constant $\delta > 0$. Then the present value of the firm's future cash flow is finite, respectively

$$\mathbb{E}\Big[\int_t^\infty e^{-r(s-t)} \delta V_s \, ds \big| \mathcal{F}_t\Big] \quad = \quad \delta V_t \int_t^\infty e^{(\mu-r)(s-t)} \, ds = \frac{\delta V_t}{r-\mu}.$$

If $\mu \geq r$ the present value of the firm's future cash flow is infinite. This case poses several problems and an optimal exercise policy like the one determined in equation (1.11) below is not available. Nevertheless, one could assume that equity owners liquidate the firm at the first time when the firm value falls below a certain boundary, thus, assuming directly that (1.11) holds.

If the equity holders choose to liquidate the firm, a fraction $\alpha \in [0, 1]$ of the assets is lost because of liquidation costs. The outstanding debt D has to be paid to the debt-holders, if possible, and the remaining value goes to the equity holders, that is

$$\begin{aligned} &\min(D, (1-\alpha)\frac{\delta V_t}{r-\mu}) && \longrightarrow \text{debtholders} \\ &\max(0, (1-\alpha)\frac{\delta V_t}{r-\mu} - D) && \longrightarrow \text{equity.} \end{aligned}$$

If the debt takes the form of a consol bond, meaning that the coupons are paid continuously at rate $C > 0$ and the tax benefit therefore yields the constant rate θC, we conclude for the initial value of equity, according to a certain liquidation policy represented by a $(\mathcal{F}_t)_{t\geq 0}$-stopping time τ, that

$$F(V_0, C, \tau) \quad = \quad \mathbb{E}\Big[\int_0^\tau e^{-rt}\big(\delta V_t + (\theta - 1)C\big) \, dt + e^{-r\tau} \max(0, \frac{(1-\alpha)\delta V_\tau}{r-\mu} - D)\Big].$$

As the equity owners will choose the liquidation policy maximizing the initial value of equity, this leads to the optimization problem

$$S_0 = \sup_{\tau \in \mathcal{T}} F(V_0, C, \tau),$$

where $\mathcal{T}$ is the set of all $(\mathcal{F}_t)_{t\geq 0}$-stopping times. The optimal strategy, as shown by Leland and Toft (1996), is given by

$$\tau(V_B) = \inf\{t : V_t \leq V_B\}, \tag{1.11}$$

with a certain level V_B which can be determined by solving a Hamilton-Jacobi-Bellman differential equation[16]. For "conventional parameters", the authors are able to show that

$$V_B = V_B(C) = \frac{(1-\theta)C\gamma(r-\mu)}{r(1+\gamma)\delta}, \qquad \gamma = \frac{m + \sqrt{m^2 + 2r\sigma^2}}{\sigma^2}.$$

Turning to the bond holder's perspective, we notice that they receive information on the firm value just at selected times $t_1, t_2, \ldots$. This is modeled by a noisy observation of V_{t_i}, i.e., instead of observing V_{t_i} the market participants observe[17]

$$\tilde{V}_{t_i} := V_{t_i} \cdot \exp(Z_i - \frac{\sigma_Z^2}{2}).$$

The Z_i are assumed to be independent normally distributed random variables with variance σ_Z^2 and being independent of $(W_s)_{s\geq 0}$.

If we assume for simplicity that equity is not traded on the public market, the information available to the bond holder is

$$\mathcal{H}_t = \sigma(\tilde{V}_{t_1}, \ldots, \tilde{V}_{t_n}, 1_{\{\tau \leq s\}} : t_1, \ldots, t_n \leq t \text{ and } 0 \leq s \leq t).$$

In this framework the probability for no default until T equals

$$1_{\{\tau > t\}} \mathbb{P}(\tau > T | \mathcal{H}_t) = 1_{\{\tau > t\}} \mathbb{P}(\inf_{s \in (t,T]} V_s > V_B | \mathcal{H}_t).$$

Fix t and denote by t_k the last t_n which is smaller than or equal to t. Because $(W_t)_{t\geq 0}$ is a Markov process, it is sufficient to condition on a smaller $\sigma-$algebra, and therefore

$$\begin{aligned}
&1_{\{\tau > t\}} \mathbb{P}(\tau < T | \mathcal{H}_t) \\
&\qquad = 1_{\{\tau > t\}} \mathbb{P}\big(V_t \cdot \inf_{s \in (t,T]} \exp(m(s-t) + \sigma(B_s - B_t)) > V_B \big| \tilde{V}_{t_k}, 1_{\{\tau \leq t\}}\big) \\
&\qquad = 1_{\{\tau > t\}} \mathbb{P}\big(\inf_{s \in (t,T]} m(s-t) + \sigma(B_s - B_t) > \ln \frac{V_B}{V_t} \big| \tilde{V}_{t_k}, 1_{\{\tau \leq t\}}\big),
\end{aligned} \tag{1.12}$$

[16]For a detailed treatment of optimization problems in the financial context, see Korn and Korn (1999, Chapter V).

[17]Duffie and Lando (2001) use $\tilde{V}_{t_i} := V_{t_i} \cdot \exp(Z_i)$ instead. This is equivalent in terms of information, but seems counterintuitive as in that case the expectation of $\tilde{V}_{t_i}$ is not V_{t_i}.

where

$$\begin{aligned}\ln\frac{V_B}{V_t} &= \ln V_B - \ln V_{t_k} - m(t-t_k) - \sigma(B_t - B_{t_k})\\ &= \ln V_B - \ln\tilde{V}_{t_k} + Z_k - \frac{\sigma_Z^2}{2} - m(t-t_k) - \sigma(B_t - B_{t_k}).\end{aligned}$$

Applying Lemma B.1.1 with $X_1 := Z_k \sim \mathcal{N}(0,\sigma_Z^2)$ and $X_2 := \ln V_{t_k} - mt_k \sim \mathcal{N}(0,\sigma^2 t_k)$ yields the representation

$$Z_k = \frac{\sigma_Z^2(\ln\tilde{V}_{t_k} - mt_k + \frac{\sigma_Z^2}{2})}{\sigma_Z^2+\sigma^2 t_k} + \frac{\sigma_Z\sigma\sqrt{t_k}}{\sqrt{\sigma_Z^2+\sigma^2 t_k}}\,\xi,$$

where ξ has a standard normal distribution and is independent of $\tilde{V}_{t_k}$. We obtain the decomposition

$$\begin{aligned}\ln\frac{V_B}{V_t} &= \ln\tilde{V}_{t_k}\Big(\frac{\sigma_Z^2}{\sigma_Z^2+\sigma^2 t_k} - 1\Big) + \frac{\sigma_Z\sigma\sqrt{t_k}}{\sqrt{\sigma_Z^2+\sigma^2 t_k}}\,\xi - \sigma(B_t - B_{t_k}) + M(t,t_k)\\ &= \ln\tilde{V}_{t_k}\frac{\sigma^2 t_k}{\sigma_Z^2+\sigma^2 t_k} + \frac{\sigma_Z\sigma\sqrt{t_k}}{\sqrt{\sigma_Z^2+\sigma^2 t_k}}\,\xi - \sigma(B_t - B_{t_k}) + M(t,t_k),\end{aligned}$$

where we set

$$\begin{aligned}M(t,t_k) &= \ln V_B + \frac{\sigma_Z^2(\frac{\sigma_Z^2}{2} - mt_k)}{\sigma_Z^2+\sigma^2 t_k} - \frac{\sigma_Z^2}{2} - m(t-t_k)\\ &= \ln V_B - mt + \frac{\sigma^2 t_k}{\sigma_Z^2+\sigma^2 t_k}\Big(\frac{\sigma_Z^4}{2\sigma^2 t_k} - \frac{\sigma_Z^2}{2} + mt_k\Big).\end{aligned}$$

This decomposition of V_t into independent random variables will enable us to calculate the desired probability. Consider

$$\begin{aligned}(1.12) &= 1_{\{\tau>t\}}\mathbb{P}\Big(\inf_{s\in(t,T]} m(s-t) + \sigma(B_s - B_t) >\\ &\qquad \ln\tilde{V}_{t_k}\frac{\sigma^2 t_k}{\sigma_Z^2+\sigma^2 t_k} + \frac{\sigma_Z\sigma\sqrt{t_k}}{\sqrt{\sigma_Z^2+\sigma^2 t_k}}\,\xi - \sigma(B_t - B_{t_k}) + M(t,t_k)\Big|\,\tilde{V}_{t_k}, 1_{\{\tau\le t\}}\Big)\\ &= 1_{\{\tau>t\}}\mathbb{E}\Big[\mathbb{P}\Big(\inf_{s\in(t,T]} m(s-t) + \sigma(B_s - B_t) > \eta\Big|\tilde{V}_{t_k}, 1_{\{\tau\le t\}}, B_t - B_{t_k}, \xi\Big)\Big|\,\tilde{V}_{t_k}, 1_{\{\tau\le t\}}\Big]\end{aligned}$$

with

$$\eta := \ln\tilde{V}_{t_k}\frac{\sigma^2 t_k}{\sigma_Z^2+\sigma^2 t_k} + \frac{\sigma_Z\sigma\sqrt{t_k}}{\sqrt{\sigma_Z^2+\sigma^2 t_k}}\,\xi - \sigma(B_t - B_{t_k}) + M(t,t_k).$$

Because η is measurable w.r.t. $\mathcal{B} := \sigma(\tilde{V}_{t_k}, 1_{\{\tau\le t\}}, B_t - B_{t_k}, \xi)$ and $B_s - B_t$ is independent of $\mathcal{B}$, we can apply equation (B.2). Recall that equation (B.2) yields for $c<0$

$$\begin{aligned}&\mathbb{P}\Big(\inf_{s\in(t,T]} m(s-t) + \sigma(B_s - B_t) > c\Big)\\ &\qquad = \Phi\Big(\frac{m(T-t) - c}{\sigma\sqrt{T-t}}\Big) - e^{2cm/\sigma^2}\Phi\Big(\frac{m(T-t)+c}{\sigma\sqrt{T-t}}\Big)\end{aligned}$$

and zero otherwise. Therefore we obtain

$$(1.12) \quad = \quad 1_{\{\tau>t\}}\mathbb{E}\Big[1_{\{\eta<0\}}\Big(\Phi\big(\frac{m(T-t)-\eta}{\sigma\sqrt{T-t}}\big) - e^{2\eta m/\sigma^2}\Phi\big(\frac{m(T-t)+\eta}{\sigma\sqrt{T-t}}\big)\Big)\Big|\tilde V_{t_k}\Big]. \qquad (1.13)$$

Furthermore, η is (conditionally on $\tilde V_{t_k}$) normally distributed with mean

$$\ln \tilde V_{t_k}\frac{\sigma^2 t_k}{\sigma_Z^2+\sigma^2 t_k} + M(t,t_k) \quad = \quad \ln V_B - mt + \sigma^2 t_k \frac{\ln \tilde V_{t_k} + mt_k + \frac{\sigma_Z^4}{2\sigma^2 t_k} - \frac{\sigma_Z^2}{2}}{\sigma_Z^2+\sigma^2 t_k} =: \mu_\eta(t_k)$$

and deterministic variance

$$\frac{\sigma_Z^2\sigma^2 t_k}{\sigma_Z^2+\sigma^2 t_k} + \sigma^2(t-t_k) = \sigma^2 t - \frac{\sigma^4 t_k^2}{\sigma_Z^2+\sigma^2 t_k} =: \sigma_\eta^2.$$

It is easy to check that $t \geq t_k$ implies $\sigma_\eta^2 \geq 0$.

One of the main assertions of Duffie and Lando (2001) is that this imperfect information model results in a hazard rate model with a certain hazard rate λ_t. Our objective is to compute this hazard rate explicitly. The calculations are postponed to the appendix and in Lemma B.3.2 we come up with the hazard rate

$$\begin{aligned}\lambda_t &= -1_{\{\tau>t\}}\frac{\partial}{\partial T}\Big|_{T=t}\ln\Big[\mathbb{P}(\tau>T|\mathcal{H}_t)\Big]\\ &= -1_{\{\tau>t\}}\frac{1}{2\sqrt{2\pi}}\big(\mu_\eta\frac{\sigma^2}{\sigma_\eta^2}+m\big)\cdot\exp\big(-\frac{\mu_\eta^2}{2\sigma_\eta^2}\big).\end{aligned}$$

1.7 Market Models with Credit Risk

Schönbucher (2000) discusses the framework for a defaultable *market model*. The difference between the market models and the "continuous maturity" models is that market models rely only on a finite number of bonds, whereas continuous maturity models assume a continuity of bonds traded in the market, that is bonds for *all* maturities in a certain range. As a matter of fact, many important variables are not available in market models as, for example, the short rate or continuously derived forward rates, which form the basis for the setting in Heath, Jarrow and Morton (1992). Introductions to market models without default risk can be found for example in Brace, Gatarek and Musiela (1995), Rebonato (1996) or Brigo and Mercurio (2001).

Assume we are given a collection of settlement dates $T_1 < \cdots < T_K$, the tenor structure, which denotes the maturities of all traded bonds.

Denote by $B_k(t) := B(t,T_k)$ the riskless bonds traded in the market. The discrete forward rate for the interval $[T_k, T_{k+1}]$ is defined as

$$F(t,T_k,T_{k+1}) =: F_k(t) = \frac{1}{T_{k+1}-T_k}\Big(\frac{B_k(t)}{B_{k+1}(t)} - 1\Big).$$

The defaultable zero coupon bond is denoted by $\bar{B}(t, T_k)$. As a starting point for modeling, it is assumed that this is a zero recovery bond, i.e., at default the value of the bond falls to zero. Put $\bar{B}_k(t) = \bar{B}(t, T_k) = 1_{\{\tau > t\}} \bar{B}(t, T_k)$. The *default risk factor* is denoted by

$$D_k(t) := \frac{\bar{B}_k(t)}{B_k(t)}.$$

If there exists an equivalent martingale measure Q we have

$$\begin{aligned} D_k(t) &= \frac{1}{B_k(t)} \mathbb{E}^Q \Big[\exp\big(-\int_0^{T_k} r_u \, du\big) 1_{\{\tau > T_k\}} \Big| \mathcal{F}_t \Big] \\ &= \frac{B_k(t)}{B_k(t)} \mathbb{E}^{T_k} \big[1_{\{\tau > T_k\}} \big| \mathcal{F}_t \big] \\ &= Q^{T_k} \big(\tau > T_k \big| \mathcal{F}_t \big) \end{aligned}$$

where Q^{T_k} denotes the T_k-forward measure[18] and $\mathbb{E}^{T_k}$ the expectation w.r.t. this measure. So $D_k(t)$ denotes the probability that, under the forward measure, the bond survives time T_k.

Remark 1.7.1. In a recovery of treasury model[19] the defaultable bond is modeled as a sum of zero recovery bond $B^0(t, T)$ and a risk-free bond

$$\bar{B}_k(t) = w B_k^0(t) + (1 - w) B_k(t).$$

We immediately conclude that in this case

$$D_k(t) = w Q_t^{T_k} \big(\tau > T_k \big| \mathcal{F}_t \big) + (1 - w).$$

Define

$$H(t, T_k, T_{k+1}) := H_k(t) = \frac{1}{T_{k+1} - T_k} \Big(\frac{D_k(t)}{D_{k+1}(t)} - 1 \Big).$$

To simplify the notation we write B_1 for $B_1(t)$ (similarly for F, D, H) and $T_{j+1} - T_j = \delta_j$.

This leads to the following decomposition

$$\begin{aligned} \bar{B}_k &= \bar{B}_1 \prod_{j=1}^{k-1} \frac{\bar{B}_{j+1}}{\bar{B}_j} \\ &= \bar{B}_1 \prod_{j=1}^{k-1} \frac{\bar{B}_{j+1}}{B_{j+1}} \frac{B_j}{\bar{B}_j} \frac{B_{j+1}}{B_j} \\ &= D_1 \prod_{j=1}^{k-1} \frac{D_{j+1}}{D_j} \, B_1 \prod_{j=1}^{k-1} \frac{B_{j+1}}{B_j} \\ &= D_1 B_1 \prod_{j=1}^{k-1} \big(1 + \delta_j H_j\big)^{-1} \cdot \big(1 + \delta_j F_j\big)^{-1}. \end{aligned}$$

[18] The T_k-forward measure is the risk neutral measure which has the risk-free bond with maturity T_k as numeraire. For details see Björk (1997).

[19] See Section 1.3.2.

The discrete forward rates of the defaultable bond are split into a risk-free part and a risky part which is represented by the "discrete-tenor hazard rate" H.

Defining the *credit spread*

$$S_k(t) = S(t, T_k, T_{k+1}) := \bar{F}_k(t) - F_k(t),$$

we immediately obtain

$$\begin{aligned} S_k(t) &= \frac{1}{\delta_k}\Big(\frac{\bar{B}_k}{\bar{B}_{k+1}} - 1\Big) - \frac{1}{\delta_k}\Big(\frac{B_k}{B_{k+1}} - 1\Big) \\ &= \frac{B_k}{B_{k+1}}\,\frac{1}{\delta_k}\Big(\frac{\bar{B}_k\, B_{k+1}}{\bar{B}_{k+1}\, B_k} - 1\Big) \\ &= (1 + \delta_k F_k)\, H_k. \end{aligned}$$

The main motivation for market models was to reproduce Black-like formulas for prices of caps and swaptions. This was particularly possible in the so-called LIBOR-market models. The basic assumption in these models is that the discrete forward rate has a log-normal distribution. There are also other models, see, for example, Andersen and Andreasen (2000).

Schönbucher (2000) concentrates on LIBOR-like models and assumes

$$\begin{aligned} \frac{dF_k(t)}{F_k(t)} &= \mu_k^F(t)\, dt + \boldsymbol{\sigma}_k^F \cdot d\mathbf{W}(t) \\ \frac{dS_k(t)}{S_k(t)} &= \mu_k^S(t)\, dt + \boldsymbol{\sigma}_k^S \cdot d\mathbf{W}(t). \end{aligned}$$

Here $\mathbf{W}$ denotes a N-dimensional standard Brownian motion, whereas $\boldsymbol{\sigma}_k$ are constant vectors and μ_k are adapted processes.

Alternatively, also the dynamics of H could be specified and the dynamics of S derived.

Since $H_k = S_k/(1 + \delta_k F_k)$, we obtain

$$\begin{aligned} dH_k(t) &= \frac{1}{(1+\delta_k F_k)^2}\Big[(1+\delta_k F_k) S_k(\mu_k^S(t)\, dt + \boldsymbol{\sigma}_k^S \cdot d\mathbf{W}_t) \\ &\qquad - S_k \delta_k F_k(\mu_k^F(t)\, dt + \boldsymbol{\sigma}_k^F \cdot d\mathbf{W}_t) - S_k\delta_k F_k \boldsymbol{\sigma}_k^S \cdot \boldsymbol{\sigma}_k^F\, dt\Big] \\ &\quad + \frac{S_k}{(1+\delta_k F_k)^3}\delta_k^2 F_k^2 \sigma_k^F \cdot \sigma_k^F\, dt \\ &= \ldots\, dt + \frac{S_k}{1+\delta_k F_k}\Big[\boldsymbol{\sigma}_k^S - \frac{\delta_k F_k}{1+\delta_k F_k}\boldsymbol{\sigma}_k^F\Big] \cdot d\mathbf{W}_t \\ &=: H_k(t)\Big[\mu_k^H(t)\, dt + \boldsymbol{\sigma}_k^H(t) \cdot d\mathbf{W}_t]. \end{aligned}$$

Note that $\boldsymbol{\sigma}_k^H$ is not a constant, but an adapted process with

$$\boldsymbol{\sigma}_k^H(t) = \boldsymbol{\sigma}_k^S - \frac{\delta_k F_k(t)}{1+\delta_k F_k(t)}\boldsymbol{\sigma}_k^F.$$

Using Itô's formula we obtain for the dynamics of the defaultable forward rates

$$\begin{aligned} d\bar{F}_k(t) &= dS_k(t) + dF_k(t) + d < S_k, F_k >_t \\ &= \Big[S_k\mu_k^S + F_k\mu_k^F + S_kF_k\boldsymbol{\sigma}_k^S \cdot \boldsymbol{\sigma}_k^F\Big]\,dt + \big(S_k\boldsymbol{\sigma}_k^S + F_k\boldsymbol{\sigma}_k^F\big)\cdot d\mathbf{W}_t \\ &=: \bar{F}_k(t)\Big[\mu_k^{\bar{F}}(t)\,dt + \boldsymbol{\sigma}_k^{\bar{F}}(t)\cdot d\mathbf{W}_t\Big]. \end{aligned}$$

The main reason for the popularity of the market models lies in the agreement between the model and well-established market formulas for basic derivative products. Therefore the model is usually calibrated to actual market data and afterwards used, for example, to price more complicated derivatives. For this reason the dynamics are directly modeled under the risk-neutral measure, or even more conveniently, under the T_k-forward measures. In search of something analogous for market models with credit risk, the T_k-survival measure turns up naturally. It is the measure under which the defaultable bond $\bar{B}_k(t)$ becomes a numeraire.

The T_k-*survival measure* $\bar{Q}_k$ is defined by the density

$$\bar{L}_k := \frac{\exp(-\int_0^{T_k} r_s\,ds)1_{\{\tau>T_k\}}}{\bar{B}_k(0)} = \frac{d\bar{Q}_k}{dQ}.$$

Note that the density has Q-expectation 1 but becomes zero if the default happens before T_k. In view of this, $\bar{Q}_k$ is not equivalent to Q but only absolutely continuous w.r.t. Q.

At this point different changes of measures can be obtained. Changes from the survival to the forward measure and the analogy of the spot LIBOR measure in a credit risk context are also discussed in Schönbucher (2000).

Finally, consider an $\mathcal{F}_T$-measurable claim X_T, which is paid only when $\tau > T$. Assuming zero recovery, then this claim can be valued by the following result, see Bielecki and Rutkowski (2002):

$$S_t = \bar{B}(t,T)\bar{E}_k\big(X_T\big|\mathcal{F}_t\big).$$

Here $\bar{E}_k$ denotes the expectation with respect to $\bar{Q}_k$.

1.8 Commercial Models

The models presented in this section, the so-called commercial models, are quite different from the models presented up to now. These models were developed by several companies and are widely accepted in practice. They all offer an implemented software, but the complete procedure of this implementation is published only for some models.

1.8.1 The KMV Model (1995) - CreditMonitor

The procedure of KMV is based on Merton's approach (see Section 1.2.1) and combines it with historical information via a statistical procedure.

KMV do not publish the exact procedure implemented in their software but the following illustrative example may be considered to be very close to their approach.

In Merton's model the firm value of the company was assumed to be observable. In reality this is unfortunately not the case. Usually shares of a company are traded but the real firm value is even difficult to estimate for internals. Using the traded shares as an estimate of the unknown firm value dates back to Modigliani and Miller, see Caouette, Altmann and Narayanan (1998, p. 142 p.p.) for more information. The share is viewed as a call option on the firm value, where the exercise price is the level of the company's debt.

With the dynamics chosen as in Merton's model and denoting by D the debt level at time T, the value of the shares E corresponds to the Black-Scholes formula

$$E = V\,\Phi(d_1) - De^{-r(T-t)}\Phi(d_2),$$

where the constants d_1, d_2 are

$$\begin{aligned} d_1 &= \frac{\ln \frac{V}{De^{-r(T-t)}} + \frac{1}{2}\sigma^2(T-t)}{\sigma\sqrt{T-t}} \\ d_2 &= d_1 - \sigma\sqrt{T-t}. \end{aligned}$$

Inverting this relation results in the firm value. Also an estimate for the volatility of the share results in an estimate of the firm's value.

KMV found that in general firms do not default when their asset value reaches the book value of their total liabilities. This is due to the long-term nature of some of their liabilities which provides some breathing space. The *default point* therefore lies somewhere in between the total liabilities and the short-term (or current) liabilities. For this reason set

$$\text{default point} := \text{short-term debt} + 50\%\ \text{long-term debt}.$$

In the next step they calculate the *distance-to-default*

$$DD = \frac{\text{firm value} - \text{default point}}{\text{firm value} \times \text{vola of firm value}}.$$

Finally KMV obtains the default probability from data on historical default and bankruptcy frequencies including over 250,000 company-years of data and over 4,700 incidents of bankruptcy[20].

[20]See Crosbie and Bohn (2001) for further information.

1.8.2 Moody's

Besides Merton's approach, which is often stated as contingent claims analysis (CCA), there are statistical approaches, pioneered by Altman (1968), which predict default events using market information and accounting variables via econometric methods. Moody's public firm risk model bridges between these models and is therefore named a 'hybrid' model. The procedure, as described in Sobehart and Klein (2000), uses a variant of Merton's CCA as well as rating information (if available), certain reported accounting information and some macroeconomic variables to represent the state of the economy and of specific industries through logistic regression. On this basis they provide a one-year estimated default probability (EDP).

1.8.3 CreditMetrics

CreditMetrics was originally developed by J.P. Morgan and belongs to RiskMetrics Group since 1998. The procedure is totally published to clarify the model and the used data are provided in the Internet.

The target of CreditMetrics is the valuation of a whole portfolio. This includes different assets and derivatives like loans, bonds, commitments to lend, financial letters-of-credit, receivables and market driven instruments like swaps, forwards and options.

The determination of the actual price of the portfolio proceeds in three steps. First the probability of a default is determined, second the probability of changes in rating (which directly results in a different price) and third the determination of the changes in value which are evoked by either a default or a change in rating.

For the three steps certain inputs are needed. They can be obtained by historical estimation or are observable in the market[21]:

- Transition matrices - transition probabilities for changes in rating,
- Recovery rates in default - ordered by seniority, countries and sectors,
- Risk-free yield curve,
- Credit spreads - for all maturities and ratings.

The transition matrices are also provided by Moody's and Standard & Poor's and therefore have to be listed separately (Moody's rates in eight and Standard & Poor's in 18 classes). In our example we consider the Table 1.1.

Observe that there are some unusual figures in this table. For example, the probability that a company rated CCC is rated AAA after one year equals 0.22 %. This seems to be unusually high in comparison to the other entries. As there are few CCC ratings this

[21] See www.riskmetrics.com/products/data/datasets/creditmetrics.

Table 1.1: The table displays the transition probabilites (in %) for the time horizon of 1 year.

Rating (now)	Rating in 1 year							
	AAA	AA	A	BBB	BB	B	CCC	D
AAA	90.81	8.33	0.68	0.06	0.12	0	0	0
AA	0.7	90.65	7.79	0.64	0.06	0.14	0.02	0
A	0.09	2.27	91.05	5.52	0.74	0.26	0.01	0.06
BBB	0.02	0.33	5.95	86.93	5.3	1.17	0.12	0.18
BB	0.03	0.14	0.67	7.73	80.53	8.84	1	1.06
B	0	0.11	0.24	0.43	6.48	83.46	4.07	5.2
CCC	0.22	0	0.22	1.3	2.38	11.24	64.86	19.79

seems to be a consequence of an exceptional event. Also critical is that the probability to default for a company rated AAA or AA equals zero. For sure there is a small but positive probability that such an event may happen. At this point smoothing algorithms are recommended to obtain a transition-matrix which is well suited for further calculations; see Gupton, Finger and Bhatia (1997, p. 66-67).

For the second set of data, recovery rates are estimated on a historical basis. Usually this information is provided by rating agencies. There are some studies on recovery rates, and we discuss an example of Asarnow and Edwards (1995). CreditMetrics though uses just mean and standard deviation. The use of a beta distribution is discussed but not implemented.

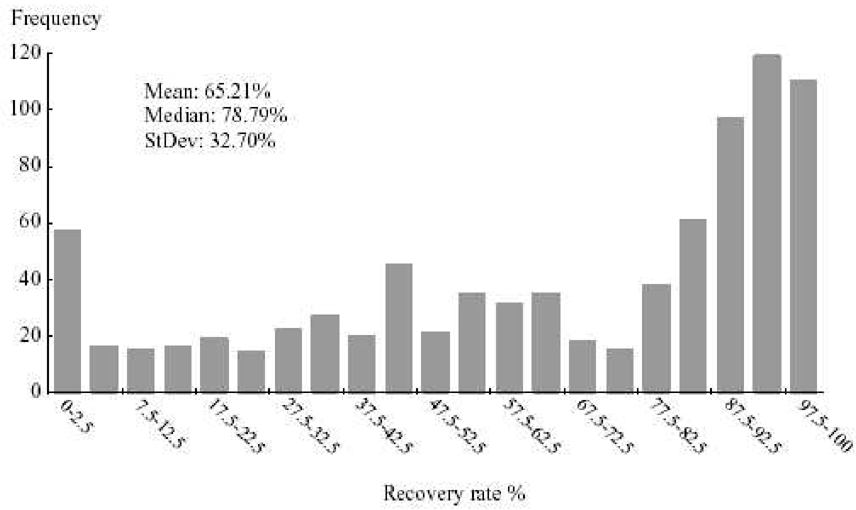

Figure 1.3: Recovery Rates

The seniority of the bond certainly has a significant influence on the recovery rate. Table 1.2 illustrates this.

Table 1.2:

Seniority	mean (%)	SD (%)
Senior Secured	53.80	26.86
Senior Unsecured	51.13	25.45
Senior Subordinated	38.52	23.81
Subordinated	32.74	20.18
Junior Subordinated	17.09	10.90

CreditMetrics also uses the actual term structure of interest rates and observable credit spreads. As the target is the valuation of bonds in a year's horizon not only default information should be used but also price changes due to rating changes. One needs to answer the question "What will be the value of a bond rated XXX in a year?". This is done by calculating stripped forward rates with respect to the rating. Stripping is the procedure to calculate zero coupon prices from a set of bonds offering coupons.

Assume for now that the current credit spreads do not change. The risk-free term structure provides forward rates and the current credit spreads are added to obtain the future (defaultable) forward-rates.

We show the full procedure in the context of an example. We face the problem to price a BBB-rated senior unsecured bond with maturity 5Y and annual coupons of 6%. Face value is 100 USD.

As described above one strips the bond prices to obtain the defaultable forward zero coupon curve. We want to explain this procedure in greater detail using the figures in Table 1.3.

Assume the bond has rating A at the end of the year. The forward value then becomes

$$FV = 6 + \frac{6}{1+3.72\%} + \frac{6}{(1+4.32\%)^2} + \frac{6}{(1+4.93\%)^3} + \frac{106}{(1+5.32\%)^4} = 108.64.$$

The other forward values are

Rating	AAA	AA	A	BBB	BB	B	CCC
Forward Value($)	109.35	109.17	108.64	107.53	102.01	98.09	83.63

The results may be found in Table 1.4.

Table 1.3:

Category	1Y	2Y	3Y	4Y (in %)
AAA	3.60	4.17	4.73	5.12
AA	3.65	4.22	4.78	5.17
A	3.72	4.32	4.93	5.32
BBB	4.10	4.67	5.25	5.63
BB	5.55	6.02	6.78	7.27
B	6.05	7.02	8.03	8.52
CCC	15.05	15.02	14.03	13.52

The value at default is assumed to be the mean of historical recovery values for senior unsecured debt. In the above calculation we followed the CreditMetrics Technical Document. For the standard deviation they do not include the estimated standard deviation of the recovery rates. If this is incorporated (SD for senior unsecured debt = 25.45%, see the table on the previous page) one obtains a standard deviation of 10.11 which is considerably higher.

Table 1.4:

State in 1Y	Prob. (%)	Forward Value	$(FV - \bar{FV})^2$
AAA	0.02	109.35	5.21
AA	0.33	109.17	4.42
A	5.95	108.64	2.48
BBB	86.93	107.53	0.21
BB	5.3	102.01	25.63
B	1.17	98.09	80.70
CCC	0.12	83.63	549.60
Default	0.18	51.13	3129.21
	mean/ SD:	**107.07**	**8.94**

1.9 Credit Derivatives

In this section we introduce several types of derivatives that relate to credit risk. Unless explicitly mentioned, we assume that the protection seller has no default risk. In reality, strong correlations between protection seller and underlying prove to be quite dangerous. The protection seller might default shortly after the underlying and the protection becomes worthless.

Additionally to the derivatives presented in this section, there exist so-called *vulnerable options*. These are derivatives whose writer may default, thus facing a counterparty risk. They are considered, for example, in Ammann (1999) or Bielecki and Rutkowski (2002). We do not consider derivatives on large baskets like *collateralized debt obligations* or others. See Blum, Overbeck and Wagner (2003) for more information.
A *credit default swap* or a *credit default option* is an exchange of a fee for a contingent payment if a credit default event occurs. The fee is usually called *default swap premium*. The difference between swap and option is determined by the way the fee is paid. If the fee is paid up-front, the agreement is called option, while if the fee is paid over time, it is called swap[22].

The "default event" is not a precise notion. Quite contrary, the event, which triggers the payment, is negotiable. It could be a certain level of spread widening, occurrence of publicly available information of failure to pay or an event, that the partners can agree upon. See Das (1998) for examples of credit derivatives and the underlying contracts. Not surprisingly, terms of documentation risk or legal risk arise in the context of credit risk.

If the payoff is some predetermined constant, the derivative is called *digital*, for example default digital put or default digital swap.

There are also options on a basket which have specific features. For example, a first-to-default swap is based on a basket of underlyings, where the protection seller agrees to cover the exposure of the first entity triggering a default event. The first-to-default structure is similar to a collateralized bond or loan obligation. Usually there are bonds or loans with similar credit ratings in the basket, because otherwise the weakest credit would dominate the derivative's behavior.

Like in the interest rate case, there are options with early exercise possibility, called American, credit derivatives with knock-in/out features, options directly on the credit spreads or leveraged credit default structures, see Tavakoli (1998). Also reduced loss credit default options are mentioned therein, which yields a way to reduce the cost of default protection. In this contract the protection buyer still takes a fixed percentage of the loss on a default event, while the further loss is covered by the protection seller.

[22]See, for example, Tavakoli (1998, p.61 p.p.).

1.9.1 Digital Options

In the case of a digital swap or option the payment, which is exchanged if the default event occurs within the lifetime of the option, is fixed. Assume, for simplicity, that the payoff equals 1. There are two possibilities for the time, when the payoff is exchanged, either at maturity T of the option or directly at default τ:

(i) If the payoff takes place at maturity, the price of the option (usually called put) at time t, if there was no default before t, equals[23]

$$\begin{aligned} 1_{\{\tau>t\}} P_d(t,T) &= 1_{\{\tau>t\}} \mathbb{E}_t\Big[\exp(-\int_t^T r_u\,du) 1_{\{\tau\le T\}}\Big] \\ &= 1_{\{\tau>t\}} B(t,T)\, Q_t^T\big[\tau \le T\big]. \end{aligned}$$

Remark 1.9.1. The payoff of the digital default put in this case is similar to the payoff of the zero recovery bond. In fact, if we denote the defaultable bond with zero recovery and maturity T by $B^0(\cdot,T)$, we obtain

$$\begin{aligned} 1_{\{\tau>t\}} P_d(t,T) &= 1_{\{\tau>t\}} \mathbb{E}_t\big[\exp(-\int_t^T r_u\,du)(1 - 1_{\{\tau>T\}})\big] \\ &= 1_{\{\tau>t\}} [B(t,T) - B^0(t,T)]. \end{aligned}$$

So, once the price of the zero recovery bond is known, the price of the default put can be easily calculated. Economically spoken, as a defaultable put and a zero recovery bond with same maturities guarantee the payoff 1, their price must be equal to the price of a risk-free bond, which is $B(t,T)$.

(ii) If the payoff is done at default, Theorem A.1.3 yields for $t \in [0,T]$

$$\begin{aligned} 1_{\{\tau>t\}} P_d(t,T) &= 1_{\{\tau>t\}} \mathbb{E}_t\Big[\exp(-\int_t^\tau r_u\,du) 1_{\{\tau\le T\}}\Big] \\ &= 1_{\{\tau>t\}} \mathbb{E}_t\Big[\int_t^T \exp(-\int_t^s (r_u+\lambda_u)\,du)\lambda_s\,ds\Big] \\ &= 1_{\{\tau>t\}} \int_t^T B(t,s)\, \mathbb{E}_t^s\big[\exp(-\int_t^s \lambda_u\,du)\lambda_s\big]\,ds. \end{aligned}$$

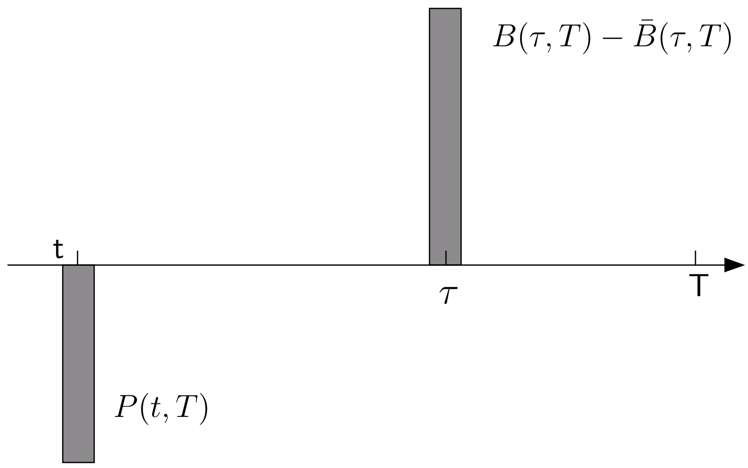

Figure 1.4: Cash flows for a default put. Default occurs at τ before the option expires. The payoff is agreed to be the "difference to an equivalent default-free bond", which is denoted by $B(\tau, T) - \bar{B}(\tau, T)$. The price of the default put is denoted by $P(t, T)$ and is paid initially at t.

1.9.2 Default Options and Credit Default Swap

To clarify the payments taking place for a *default option* or a *credit default swap*, consider figures 1.4 and 1.5. In the case of the default option, the protection buyer pays a fee up-front, which equals the price of the option. For the credit default swap (CDS) the premium $\bar{S}$ is paid at time points $T_1, \ldots, T_n$ until either maturity of the contract or default.

There are two structural possibilities for the default payment[24].

1. *Difference to par.* If a default event occurs, the protection seller has either to pay the par value (which we always assume to be 1) in exchange for the defaulted bond, or pay the par value minus the post-default price of the underlying bond. The payoff is equivalent to
$$1 - \bar{B}(\tau, T), \qquad \text{if } \tau \leq T.$$
2. *Difference to an equivalent bond.* The payoff in the case that a default event occurs is the value of an equivalent, default-free bond minus the market value of the defaulted bond. In this case the payoff equals
$$B(\tau, T) - \bar{B}(\tau, T), \qquad \text{if } \tau \leq T.$$

In the case of a coupon bond, there is usually a protection of the principal, and possibly of the accrued interest.

The first step in pricing the defaultable swap is the pricing of the defaultable option with the same payoff. The price of the option, denoted by $P(t, T)$, yields the discounted value

[23] For convenience we write $\mathbb{E}_t(\cdot)$ for $\mathbb{E}^Q(\cdot|\mathcal{F}_t)$ and $\mathbb{E}_t^T(\cdot)$ for $E^{Q^T}(\cdot|\mathcal{F}_t)$, when Q^T is the T-forward measure.

[24] See, for example, Das (1998, p. 63).

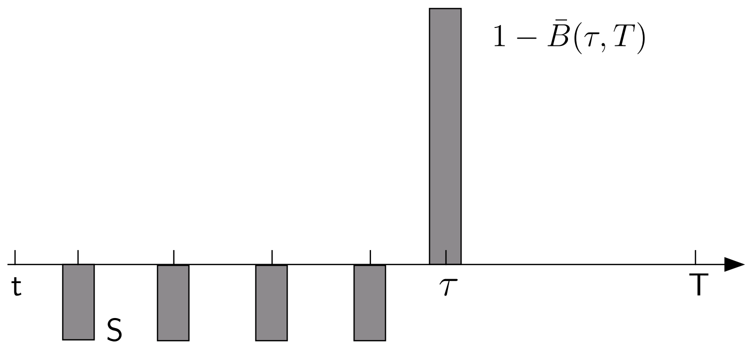

Figure 1.5: Cash flows for a credit default swap. Default occurs at τ before the option expires. The payoff is agreed to be the "difference to par", $1-\bar{B}(\tau,T)$. The default swap spread, $\bar{S}$, is paid regularly at times $T_1,\ldots,T_4$ (until default).

of the payoff at time t. The premium $\bar{S}$ is paid at times $T_1,\ldots,T_n$, but at most until a default event occurs. Denoting the price of a zero recovery bond by $B^0(t,T)$, this yields

$$P(t,T)=\sum_{i=1}^{n}\bar{S}\cdot B^0(t,T_i).$$

Consequently, the swap premium can be obtained, once the price of the defaultable option and the zero recovery bond prices are known, as

$$\bar{S}(t)=\frac{P(t,T)}{\sum_{i=1}^{n}B^0(t,T_i)}. \tag{1.14}$$

For example, if we assume recovery of treasury for the defaultable bond, we have

$$P(t,T)=\mathbb{E}_t\big[\exp(-\int_t^\tau r_u\,du)\,(1-\delta)1_{\{t<\tau\le T\}}\big],$$

which can be expressed using the default digital put as

$$P(t,T)=(1-\delta)P_d(t,T).$$

As already mentioned, this gets slightly more difficult if the underlying is a coupon bond, see Schmid (2002) for details.

1.9.3 Default Swaptions

A *credit default swaption* offers the right, but not the obligation, to buy or sell a credit default swap at a future time point T for a pre-specified swap premium K. The contract is knocked out if a default of the reference entity occurs before T. We refer to a *credit default swap call* (CDS call) if the assigned right is to buy a credit default swap and otherwise to a credit default put (CDS put). Credit default swaptions are not yet standard instruments which are liquidly traded, but, for example, Hull and White (2002) report that a market for such contracts is developing.

Denoting the tenor structure of the underlying swap by $\mathcal{T} = \{T_1, \ldots, T_n\}$ and the price of the CDS call at time t by $C_S(t, T, \mathcal{T})$, we obtain for the payoff of the CDS call at maturity $T \leq T_1$

$$C_S(T, T, \mathcal{T}) = \left[\bar{S}(T) - K\right]^+ \sum_{i=1}^{n} B^0(T, T_i) 1_{\{\tau > T\}}.$$

$\bar{S}(T)$ is the swap premium at time T. For simplicity we set the day-count fraction to one[25].

If the swap offers the replacement of the difference to an equivalent default-free bond in the case of a default, the swap rate equals

$$\bar{S}(T) = \frac{B(T, T_n) - \bar{B}(T, T_n)}{\sum_{i=1}^{n} B^0(T, T_i)}.$$

We conclude for the price of the CDS call

$$C_S(0, T, \mathcal{T}) \;=\; \mathbb{E}\Big[\exp(-\int_0^T r_u\, du)\Big(B(T, T_n) - \bar{B}(T, T_n) - K \sum_{i=1}^{n} B^0(T, T_i)\Big)^+ 1_{\{\tau > T\}}\Big].$$

Otherwise, if difference to par is considered, the swap price depends on the recovery. In a recovery of treasury model, the swap rate, as shown in the previous section, equals

$$\bar{S}(T) = \frac{(1 - \delta)\, P_d(T, T_n)}{\sum_{i=1}^{n} B^0(T, T_i)}.$$

This yields that the price of the CDS call can be computed via

$$C_S(0, T, \mathcal{T}) \;=\; \mathbb{E}\Big[\exp(-\int_0^T r_u\, du)\Big((1 - \delta)\, P_d(T, T_n) - K \sum_{i=1}^{n} B^0(T, T_i)\Big)^+\Big].$$

1.9.4 Credit Spread Options

A *credit spread option* is an option which depends on the credit spread, that is the difference between the yield of the underlying defaultable bond and the yield of a reference bond, which is usually assumed to be default-free. For example, a *credit spread call* with strike (yield) K at maturity T has the payoff

$$\Big(\bar{B}(T, T') - e^{-K(T'-T)} B(T, T')\Big)^+,$$

where $T' > T$ is the maturity of the underlying defaultable bond.

[25]For a discussion on the different day-count fractions, see James and Webber (2000, p. 51 p.p.). With arbitrary day-count fraction Δ_i we would have to consider $\sum_{i=1}^{n} \Delta_i B^0(T, T_i)$.

Thus the call is in the money if the yield of the defaultable bond is higher than the yield of the riskless bond plus the strike (yield) K. We use continuous compounding[26] of the yield rate, and note that this represents an annual yield, if the time scale is denoted in entities of 1 year.

Schmid (2002) discusses credit spread options with a *knock-out* feature. In this case a credit spread call option with maturity T on an underlying defaultable bond with maturity T' and strike K, knocked out at default, has the payoff

$$1_{\{\tau > T\}} \Big(\bar{B}(T, T') - e^{-K(T'-T)} B(T, T') \Big)^+.$$

In contrast to the option-specific payoff, a *credit spread swap* with strike K and maturity T has the payoff

$$\bar{B}(T, T') - e^{-K(T'-T)} B(T, T').$$

To replicate the payoff of the credit spread swap, the seller buys a portfolio at time t, which consists of the defaultable bond with maturity T' and sells $\exp[-K(T'-T)]$ risk free bonds with maturity T'. A replicating argument yields the value at time t of the above payoff to be $\bar{B}(t, T') - B(t, T') \exp[-K(T'-T)]$. Consequently, the credit spread swap premium, which has to be paid at times $T_1, \ldots, T_n$, equals

$$\bar{S}(t) = \frac{\bar{B}(t, T') - e^{-K(T'-T)} B(t, T')}{\sum_{i=1}^n B(t, T_i)}.$$

If the credit spread swap is knocked out at default of the underlying, the premium relates to zero recovery bonds $B^0(\cdot, T')$, which promise the par value, 1, if the reference bond $\bar{B}(\cdot, T')$ did not default until its maturity T' and zero otherwise. Then the premium equals

$$\bar{S}(t) = \frac{\bar{B}(t, T') - e^{-K(T'-T)} B(t, T')}{\sum_{i=1}^n B^0(t, T_i)}.$$

1.9.5 kth-to-default Options

Derivatives with a *kth-to-default* feature are quite common in the market. For example, a first-to-default put covers the loss of the first defaulted asset in a considered portfolio. These types of products offer a cheaper protection against losses, if one considers more than k assets to default in a certain time interval as unlikely, and therefore offer tailor-made credit risk profiles, which may be used to redistribute credit risk or release regulatory capital.

[26] The relation to the discrete time value of money concept is the following. The discounting factor for a time period of T years are

$$\frac{1}{(1+y)^{nT}} = e^{-K \cdot T},$$

if the yield y is paid n times a year. This yields the relation

$$y = (\ln K)^{\frac{1}{n}}.$$

Once a price for a kth-to-default put is obtained, the premium of a kth-to-default swap can be calculated via formula (1.14). See Section 1.5 for applications, where we already obtained the following formula for the premium of a first-to-default swap

$$S_{1\text{st}}(t) \quad = \quad 1_{\{\tau^{1\text{st}}>t\}} \frac{(1-\delta)\mathbb{E}_t\big[\exp(-\int_t^{\tau^{1\text{st}}} r_u\,du)1_{\{\tau^{1\text{st}}\le T\}}\big]}{\sum_{i=1}^m \mathbb{E}_t\big[\exp(-\int_t^{T_i} r_u\,du)1_{\{\tau^{1\text{st}}>T_i\}}\big]}.$$

Chapter 2

SDEs on Hilbert Spaces

This section develops some theory of stochastic processes on Hilbert spaces, in particular, the Itô - calculus for such processes. A detailed treatment may be found in the work of Da Prato and Zabczyk (1992). We adapt their methodology to our framework and provide an introduction to stochastic analysis on Hilbert spaces. The Itô - formula is extended from real-valued functions to functions which have values in a Hilbert space. Note that there is a similar extension in Filipović (2001, sec. 2.3.1) using a different proof. For the tools from analysis and functional analysis we refer to Dieudonné (1969), Yosida (1971) or Werner (2000).

The term structure of interest rates and its evolution may be described by the set of forward rates $\{f(s,t) : s \le t\}$. If we fix the time s, the forward rate curve $x \mapsto f(s,x)$ appears as an element of a functional space. Therefore, a stochastic process $(f(s))_{s\ge0}$ which itself takes values in a functional space may well serve as a model for the forward rates.

To formulate the dynamics of the forward rates, we develop some methodology for Wiener processes in functional spaces. As pointed out by Yor (1974), there are fundamental problems defining the stochastic integral of a Banach space valued process, while Hilbert spaces are more suitable. This leads us to stochastic processes with values in Hilbert spaces.

For technical reasons we always consider a finite time horizon $T^* \in \mathbb{R}$, i.e., investigate the process $(f(s))_{s\in[0,T^*]}$.

2.1 Preliminaries

Consider a separable Hilbert space H with an inner product $<\cdot,\cdot>$. The space of linear, continuous mappings from H into itself is denoted by $L(H)$. Note that $L(H)$ is a Banach space and for $D \in L(H)$ and $h \in H$ we often write $D \cdot h$ instead of $D(h)$. The Borel σ-algebra $\mathcal{B}(H)$ is the σ-algebra induced from the norm of H.

We start by defining normality for probability measures in H.

Definition 2.1.1. A probability measure μ on $(H, \mathcal{B}(H))$ is said to be Gaussian, if and only if for any $h \in H$ there exist $p, q \in \mathrm{I\!R}$ with $q \geq 0$, such that

$$\mu\{x \in H, < h, x > \in A\} = \mathcal{N}(p, q)(A), \qquad \forall A \in \mathcal{B}(\mathrm{I\!R}).$$

Here, $\mathcal{N}(p, q)$ denotes the Gaussian measure on $(\mathrm{I\!R}, \mathcal{B}(\mathrm{I\!R}))$ with mean p and variance q.

Second, we generalize the concept of mean and variance.

Definition 2.1.2. For a Gaussian measure μ on $(H, \mathcal{B}(H))$ the element $m \in H$ such that

$$\int\limits_H < h, x > \ \mu(dx) = < m, h > \qquad \forall h \in H$$

is called the *mean* of μ. The symmetric, nonnegative operator $D \in L(H)$ with

$$\int\limits_H < h_1, x > < h_2, x > \mu(dx) - < m, h_1 > < m, h_2 > = < Dh_1, h_2 > \qquad \forall h_1, h_2 \in H$$

is called the *covariance operator* of μ.

Here D is symmetric in the sense, that $< Dh_1, h_2 > = < Dh_2, h_1 >$. As for $\mathrm{I\!R}^n$, mean and covariance operator uniquely determine μ. For a random variable with distribution μ we write $\xi \sim \mathcal{N}(m, D)$ and $\mathrm{Cov}(\xi) := D$.

The connection with Definition 2.1.1 is the following. For $h \in H$ and $p, q \in \mathrm{I\!R}$ such that

$$\mu\{x \in H, < h, x > \in A\} = \mathcal{N}(p, q)(A),$$

we have $p = < m, h >$ and $q = < Dh, h >$.

Remark 2.1.3. If $H = \mathrm{I\!R}^n$, a measure is Gaussian, iff the characteristic function takes the form

$$\varphi(\boldsymbol{\lambda}) = \exp[\mathrm{i}\boldsymbol{\lambda}^\top \mathbf{m} - \frac{1}{2}\boldsymbol{\lambda}^\top \boldsymbol{\Sigma}\ \boldsymbol{\lambda}], \qquad \boldsymbol{\lambda} \in \mathrm{I\!R}^n,$$

for appropriate $\mathbf{m} \in \mathrm{I\!R}^n$ and $\boldsymbol{\Sigma} \in \mathrm{I\!R}^n \times \mathrm{I\!R}^n$. The appearing terms equal

$$\boldsymbol{\lambda}^\top \mathbf{m} = \int < \boldsymbol{\lambda}, \mathbf{x} > \ \mu(d\mathbf{x})$$

and

$$\begin{aligned}\boldsymbol{\lambda}^\top \boldsymbol{\Sigma}\ \boldsymbol{\lambda} &= \boldsymbol{\lambda}^\top \big(\int (\mathbf{x} - \mathbf{m})\,(\mathbf{x} - \mathbf{m})^\top \mu(d\mathbf{x}) \big)\ \boldsymbol{\lambda} \\ &= \int < \boldsymbol{\lambda}, \mathbf{x} > < \boldsymbol{\lambda}, \mathbf{x} > \ \mu(d\mathbf{x}) - < \boldsymbol{\lambda}, \mathbf{m} >^2 .\end{aligned}$$

Coming back to linear operators on H, we introduce an elementary but important tool, the trace.

Definition 2.1.4. Consider an orthonormal basis $\{e_k : k \in \mathbb{N}\}$ of H. For any linear operator D on H we define the *trace* of D through

$$\operatorname{tr} D := \sum_{j=1}^{\infty} < De_j, e_j >,$$

if the above series converges absolutely, and set the trace equal to infinity otherwise. The trace is independent of the chosen basis. If $\sum | < De_j, e_j > | < \infty$, the operator D is called *trace-class*. We denote the Banach space of trace-class operators by $L_1(H)$ and its norm by $|| \cdot ||_1$, compare Da Prato and Zabczyk (1992, Appendix C).

It can be shown that the covariance operator of a Gaussian probability measure is a trace class operator, see Da Prato and Zabczyk (1992, Proposition 2.15). Note that for positive D trace-class already follows from $\operatorname{tr} D < \infty$.

In our applications H will be a space of functions $h : \mathbb{R} \mapsto \mathbb{R}$. For a stochastic process $(X(s))_{s \geq 0}$ which takes values in H we set $X(s,t) := X(s)(t)$.

That is, if we consider the process of forward rates $(f(s,t))_{s \geq 0}$, $f(s,t)$ represents the forward rate at time s with maturity t, while $f(s)$ represents the whole term structure at time s.

Definition 2.1.5. For a symmetric, nonnegative trace-class operator $D \in L(H)$ the H-valued process $(X(s))_{s \geq 0}$ is called a *D-Wiener process* if

(i) $X(0) = 0$,

(ii) X has continuous trajectories,

(iii) X has independent increments,

(iv) the distribution of $X(s_2) - X(s_1)$ is a Gaussian measure on H with mean 0 and covariance operator $(s_2 - s_1)D$.

If the considered probability space admits a filtration $(\mathcal{F}_s)_{s \geq 0}$ satisfying the usual conditions[1], $X(s)$ is $\mathcal{F}_s$-measurable and $X(s_2) - X(s_1)$ is independent of $\mathcal{F}_{s_1}$ for all $s_2 > s_1 \geq 0$, we say that X is a D-Wiener process with respect to $(\mathcal{F}_s)_{s \geq 0}$.

Property (iv) specifies the covariance structure of $(X_s)_{s \geq 0}$. The covariance operator of a certain increment, say $X_{s_2} - X_{s_1}$, may be decomposed into a factor which depends only on time, namely $s_2 - s_1$, and an operator D. The operator D refers to the covariance in the Hilbert space. For $s_1 = 0$ and $s_2 = s$ one might think of $\operatorname{Cov}\big(X(s,t_1), X(s,t_2)\big)$, so the second factor describes the covariance w.r.t. the maturity (t_1, t_2 respectively).

[1] This means, that $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets and the filtration is right-continuous, which is called the usual augmentation of $(\mathcal{F})$, see Revuz and Yor (1994).

Compare to the case $H = \mathbb{R}^n$. A Brownian motion with covariance function $(s \wedge t)\mathbf{\Sigma} = (s \wedge t)\mathbf{A}\mathbf{A}^\top$ is obtained from a Brownian motion $(\mathbf{B}_s)_{s\geq 0}$ with independent components via $(\mathbf{A}\mathbf{B}_s)_{s\geq 0}$. As this procedure can not be transferred to the infinite dimensional case, the covariance operator always needs to be specified explicitly, and this is why we speak of a D-Wiener process.

We now develop the Eigenvalue expansion of a H-valued random variable ξ. This will be crucial if we consider linear operators on $\xi = X(s)$.

Observe that for any orthonormal basis $\{e_k : k \in \mathbb{N}\}$ of H and $f \in H$ we have

$$f = \sum_k < f, e_k > \, e_k,$$

where the Fourier-coefficients $< f, e_k >$ are real-valued random variables. For $D \in L(H)$ and $f, g \in H$ we obtain

$$\begin{aligned} < Df, g > &= < D \sum_k e_k < f, e_k >, \sum_l e_l < g, e_l >> \\ &= \sum_{k,l} < f, e_k >< De_k, e_l >< g, e_l > . \end{aligned}$$

It is interesting to find a basis which simplifies the above expression. Assume that D is a covariance operator, in particular, D is nonnegative and trace-class. The Eigenvectors of D form a complete orthonormal system $\{e_k : k \in \mathbb{N}\}$ while the Eigenvalues λ_k form a bounded sequence of nonnegative real numbers, such that[2]

$$D\, e_k = \lambda_k\, e_k. \tag{2.1}$$

Because D is a trace-class operator, we have $\sum_k \lambda_k < \infty$ and obtain

$$< Df, g > \;=\; \sum_k \lambda_k < e_k, f >< e_k, g > .$$

If we consider $\xi = X(s)$, where $(X(s))$ is a D-Wiener process, this Eigenvalue expansion gives a useful representation.

Proposition 2.1.6. *Consider a D-Wiener process $(X(s))_{s\geq 0}$ and denote by $\{e_k : k \in \mathbb{N}\}$ the Eigenvectors of D. Define*

$$\beta_k(s) :=< X(s), e_k > .$$

Then, for $\lambda_k > 0$, $\frac{1}{\sqrt{\lambda_k}}\beta_k(s)$ are mutually independent Brownian motions. Moreover, we have the decomposition

$$X(s) = \sum_{k=1}^{\infty} \beta_k(s) e_k, \tag{2.2}$$

and the series in (2.2) converges in $L^2(\Omega, \mathcal{A}, \mathbb{P})$.

[2] See Da Prato and Zabczyk (1992, p.86). Note that the system $\{e_k\}$ certainly depends on D. In the following we always refer to this particular $\{e_k\}$ without stressing the dependence on D.

Proof. Clearly β_k is a continuous, centered Gaussian process. Then

$$\mathbb{E}\Big[\frac{\beta_k(s)\,\beta_l(t)}{\sqrt{\lambda_k\lambda_l}}\Big] = \frac{1}{\sqrt{\lambda_k\lambda_l}}\,\mathbb{E}\Big[< X(s), e_k >< X(t), e_l > \Big],$$

and because $(X(s))_{s\geq 0}$ has independent increments, we obtain

$$\mathbb{E}\Big[< X(s), e_k >< X(t), e_l > \Big] = \mathbb{E}\Big[< X(s\wedge t), e_k >< X(s\wedge t), e_l > \Big].$$

According to item (iv) of Definition 2.1.5, the distribution of $X(s\wedge t)$ is Gaussian with mean 0 and covariance operator $(s\wedge t)D$. Therefore

$$\mathbb{E}\big[< X(s\wedge t), e_k >< X(s\wedge t), e_l > \big] = \int < x, e_k >< x, e_l > \mu_{X_{s\wedge t}}(dx) =< (s\wedge t)D\, e_k, e_l >$$

and

$$\begin{aligned}\mathbb{E}\Big[\frac{\beta_k(s)\,\beta_l(t)}{\sqrt{\lambda_k\lambda_l}}\Big] &= \frac{1}{\sqrt{\lambda_k\lambda_l}}(s\wedge t) < De_k, e_l > \\ &= \frac{1}{\sqrt{\lambda_k\lambda_l}}(s\wedge t)\,\lambda_k\delta_{kl} = (s\wedge t)\,\delta_{kl},\end{aligned}$$

where δ_{kl} equals one if $k=l$ and zero otherwise. We conclude that the β_k's are independent Brownian motions[3].

Furthermore,

$$\mathbb{E}(< X(s), e_k >< X(s), e_l >) = s < De_k, e_l >= s\lambda_k\delta_{kl},$$

yields

$$\mathbb{E}\Big|\Big|\sum_{k=n}^{m}\beta_k(s)e_k\Big|\Big|^2 = s\sum_{k=n}^{m}\lambda_k$$

and, because D is a trace-class operator, $\sum_k \lambda_k < \infty$. So the $\beta_k(s)e_k$ form a Cauchy-sequence and (2.2) converges in L^2.

■

The above result also yields a possibility to construct a Wiener process from a series of independent Brownian motions $(W_k)_{t\geq 0}$: For any orthonormal basis $\{e_k : k\in\mathbb{N}\}$ and positive λ_k such that $\sum_k \lambda_k < \infty$, (2.1) determines a covariance operator, say D. Then a D-Wiener process is obtained by putting

$$X(s) := \sum_{k=1}^{\infty}\sqrt{\lambda_k}W_k(s)e_k.$$

Because of this, X is often called "infinite dimensional Brownian motion".

[3] See, for example, Remark 2.9.2 in Karatzas and Shreve (1988).

2.2 The Stochastic Integral

In this section we aim to define the stochastic integral with respect to a D-Wiener process on a Hilbert Space H. In the case where $H = \mathbb{R}^n$, the integral w.r.t. an n-dimensional Brownian motion $(B(s))_{s\geq 0}$ with covariance Matrix $\mathbf{\Sigma}$ (see page 52) of a $\mathbb{R}^n \times \mathbb{R}^n$- valued stochastic process $(\boldsymbol{\sigma}(s))_{s\geq 0}$ is well known and denoted by

$$\int_0^t \boldsymbol{\sigma}(s)\, dB_s.$$

Note that the matrix $\boldsymbol{\sigma}(s)$ is a linear mapping $\mathbb{R}^n \mapsto \mathbb{R}^n$, operating on $B(s)$.

Imitating this, we consider integrands which take values in the space of linear functions from $H \to H$. As before, for $\Phi \in L(H)$ and $f \in H$ we write $\Phi \cdot f$ for $\Phi(f)$.

Definition 2.2.1. Consider the Hilbert Space H and a process $(\Phi(s))_{s\in[0,T^*]}$, which takes values in $L(H)$. $\Phi(s)$ is called **elementary**, if there exist $0 = t_0 < t_1 < \cdots < t_n = T^*$ and $\Phi_k \in L(H)$, measurable with respect to $\mathcal{F}_{t_k}$, such that $\Phi(0) = 0$ and

$$\Phi(t) = \Phi_k \qquad \text{for } t \in (t_k, t_{k+1}], k = 0, \ldots, n-1.$$

In this case we define the stochastic integral for $t \in [0, T^*]$ by

$$\int_0^t \Phi(s) \cdot dX(s) := \sum_{k=0}^{n-1} \Phi_k \cdot \big(X(t_{k+1} \wedge t) - X(t_k \wedge t)\big).$$

Note that the stochastic integral is itself a stochastic process which has values in H. Stochastic integrals prove to be a powerful concept to describe the behavior of martingales.

Denote the norm[4] on H by $\|\cdot\|$.

A stochastic process $(X(s))_{s\in[0,T^*]}$ with $E||X(s)|| < \infty$ for all $s \in [0, T^*]$ is called a *martingale* w.r.t. the filtration $(\mathcal{F}_s)_{s\geq 0}$, iff

$$\mathbb{E}(\, X_t|\mathcal{F}_s) = X_s, \qquad \text{for all } 0 \leq s < t \leq T^*.$$

Usually one considers filtrations of the type $\mathcal{F}_s = \sigma(X_t : 0 \leq t \leq s)$.

As a consequence of its independent increments, a D-Wiener process $(X(s))_{s\in[0,T^*]}$ w.r.t. $(\mathcal{F}_s)_{s\geq 0}$ is a martingale. This leads to the question, under which circumstances this property is inherited by the stochastic integral. The following proposition considers such a case.

[4]The norm in a Hilbert space is induced by the inner product, such that $||h|| :=< h, h >^{\frac{1}{2}}$ for $h \in H$.

Proposition 2.2.2. *For an elementary stochastic process $(\Phi(s))_{s\in[0,T^*]}$ with values in $L(H)$ we have*[5]

$$|||\Phi|||_{T^*} := \Big[\mathbb{E}\Big(\int_0^{T^*} \mathrm{tr}[(\Phi(s)D^{\frac{1}{2}})(\Phi(s)D^{\frac{1}{2}})^*]\,ds\Big)\Big]^{\frac{1}{2}} = \Big[\mathbb{E}\Big(\parallel \int_0^{T^*} \Phi(s)\cdot dX(s) \parallel^2\Big)\Big]^{\frac{1}{2}}.$$

Furthermore, if $|||\Phi|||_{T^*} < \infty$, then the stochastic integral

$$\int_0^t \Phi(s)\cdot dX(s)$$

is a square-integrable martingale for all $t \leq T^*$.

Proof. Enhance the partition by $t =: t_m$. Setting $\Delta_k X = X(t_{k+1}) - X(t_k)$ we obtain

$$\begin{aligned}
\mathbb{E}||\int_0^t \Phi(s)\,dX(s)||^2 &= \mathbb{E}||\sum_{k=0}^{m-1}\Phi_k\cdot\Delta_k X||^2 \\
&= \mathbb{E}\Big(\sum_{k=0}^{m-1}||\Phi_k\cdot\Delta_k X||^2\Big) + \mathbb{E}\Big(2\sum_{i,j=0,i<j}^{m-1} < \Phi_i\cdot\Delta_i X, \Phi_j\cdot\Delta_j X >\Big) \\
&=: (1) + (2).
\end{aligned}$$

Considering the first term, note that, using expansion (2.2), we may write

$$\Delta_k X = X(t_{k+1}) - X(t_k) = \sum_{i=1}^{\infty}\Delta_k\beta_i e_i,$$

where $\Delta_k\beta_i = \beta_i(t_{k+1}) - \beta_i(t_k)$ and $\frac{1}{\sqrt{\lambda_i}}\beta_i$ are mutually independent Brownian motions. This leads to

$$\begin{aligned}
(1) &= \sum_{k=0}^{m-1}\mathbb{E} < \Phi_k\cdot\Delta_k X, \Phi_k\cdot\Delta_k X > \\
&= \sum_{k=0}^{m-1}\mathbb{E} < \Phi_k\cdot\sum_{i=1}^{\infty}\Delta_k\beta_i e_i, \Phi_k\cdot\sum_{j=1}^{\infty}\Delta_k\beta_j e_j > \\
&= \sum_{k=0}^{m-1}\mathbb{E}\Big(\sum_{i,j=1}^{\infty}\mathbb{E}\Big[\Delta_k\beta_i\,\Delta_k\beta_j\Big|\mathcal{F}_{t_k}\Big] < \Phi_k\cdot e_i, \Phi_k\cdot e_j >\Big).
\end{aligned}$$

As the $\frac{1}{\sqrt{\lambda_k}}\beta_k$ are independent Brownian motions, we have

$$\begin{aligned}
\mathbb{E}\Big(\Delta_k\beta_i\,\Delta_k\beta_j\Big|\mathcal{F}_{t_k}\Big) &= \mathbb{E}\Big[\Delta_k(\frac{\sqrt{\lambda_i}}{\sqrt{\lambda_i}}\beta_i)\,\Delta_k(\frac{\sqrt{\lambda_j}}{\sqrt{\lambda_j}}\beta_j)\Big] \\
&= (t_{k+1}-t_k)\lambda_i\ \delta_{ij}. \qquad (2.3)
\end{aligned}$$

[5] Using positivity and the Eigenvalue expansion of D, we define $D^{\frac{1}{2}}(x) := \sum_k \sqrt{\lambda_k} < x, e_k > e_k$, see Werner (2000, p. 244). Furthermore, for $T \in L(H)$ we denote its Hilbert space adjoint by T^*, see Werner (2000, p. 208). That is, $a, b \in H$ yield $< Ta, b >=< a, T^*b >$.

We conclude

$$\begin{aligned}
(1) &= \sum_{k=0}^{m-1} \mathbb{E}\Big(\sum_{i=1}^{\infty}(t_{k+1}-t_k)\lambda_i < \Phi_k \cdot e_i, \Phi_k \cdot e_i > \Big) \\
&= \sum_{k=0}^{m-1} \mathbb{E}\Big(\sum_{i=1}^{\infty} < \Phi_k \cdot D^{\frac{1}{2}} e_i, \Phi_k \cdot D^{\frac{1}{2}} e_i > \Big)(t_{k+1}-t_k) \\
&= \mathbb{E}\Big(\int_0^t \operatorname{tr}\big[(\Phi(s)D^{\frac{1}{2}})(\Phi(s)D^{\frac{1}{2}})^*\big]\, ds\Big).
\end{aligned}$$

Analogously we obtain

$$(2) = 2 \sum_{i,j=1, i<j}^{m-1} \mathbb{E}\Big(< \Phi_i \cdot \Delta_i X, \Phi_j \cdot \Delta_j X > \Big) = 0.$$

For the martingale property we enhance the partition further by $s =: t_{\tilde{m}}$. Then $s = t_{\tilde{m}} < t = t_m$ so that

$$\begin{aligned}
&\mathbb{E}\Big[\int_0^t \Phi(u)\, dX(u) \Big| \mathcal{F}_s\Big] \\
&\qquad = \int_0^s \Phi(u)\, dX(u) + \mathbb{E}\Big[\sum_{j=\tilde{m}}^{m-1} \Phi_j \cdot \Delta_j X \Big| \mathcal{F}_s\Big] \\
&\qquad = \int_0^s \Phi(u)\, dX(u),
\end{aligned}$$

because of independent increments and zero means of $(X(s))_{s\in[0,T^*]}$.

■

As a next step we want to extend the stochastic integral to more general functions Φ. Therefore we look for a class of processes which can be approximated by elementary functions, such that at the same time the martingale property of the integral is preserved. It turns out that the proper class is formed by certain Hilbert-Schmidt operators.

First, consider the space $H_0 := D^{\frac{1}{2}}(H)$, which, endowed with the inner product[6]

$$< u, v >_0 := \sum_k \frac{1}{\lambda_k} < u, e_k >< v, e_k > = < D^{-\frac{1}{2}} u, D^{-\frac{1}{2}} v >$$

is a Hilbert space. For an orthonormal basis $\{e_k : k \in \mathbb{N}\}$ of H, setting $e_k^0 := D^{\frac{1}{2}} e_k$ yields an orthonormal basis $\{e_k^0 : k \in \mathbb{N}\}$ of H_0.

[6]Similar to $D^{\frac{1}{2}}$, we define $D^{-\frac{1}{2}}(x) := \sum_{k=1}^{\infty} \frac{1}{\sqrt{\lambda_k}} < x, e_k > e_k$.

Then, denote by $L_2(H_0, H)$ the space of all Hilbert-Schmidt operators from H_0 into H, that is, linear operators T, with

$$\sum_{k=1}^{\infty} < Te_k^0, Te_k^0 >^2 \quad < \infty$$

for an orthonormal basis $\{e_k^0 : k \in \mathbb{N}\}$ of H^0. Note that the inner product

$$< S, T >_2 := \sum_{k=1}^{\infty} < Se_k^0, Te_k^0 >$$

induces the norm

$$||T||_2 := \left(\sum_{k=1}^{\infty} |Te_k^0|^2 \right)^{\frac{1}{2}},$$

and $L_2(H_0, H)$ is again a Hilbert space. See, for example, Werner (2000, p. 268 p.p.).

With the above notations we define $|||\Phi|||_T$ also for non-elementary processes with values in $L_2(H_0, H)$ as

$$\begin{aligned} |||\Phi|||_T &:= \left[\mathbb{E} \int_0^T ||\Phi(s)||_2^2 \, ds \right]^{\frac{1}{2}} \\ &= \left[\mathbb{E} \int_0^T \sum_{k=1}^{\infty} < \Phi(s)e_k^0, \Phi(s)e_k^0 > \, ds \right]^{\frac{1}{2}}. \end{aligned}$$

Because $e_k^0 = D^{\frac{1}{2}} e_k$ we obtain

$$\begin{aligned} |||\Phi|||_T &= \left[\mathbb{E} \int_0^T \sum_{k=1}^{\infty} < \Phi(s)D^{\frac{1}{2}} e_k, \Phi(s)D^{\frac{1}{2}} e_k > \, ds \right]^{\frac{1}{2}} \\ &= \left[\mathbb{E} \int_0^T \sum_{k=1}^{\infty} < \Phi(s)D^{\frac{1}{2}} \left(\Phi(s)D^{\frac{1}{2}} \right)^* e_k, e_k > \, ds \right]^{\frac{1}{2}} \\ &= \left[\mathbb{E} \int_0^T \operatorname{tr} \left((\Phi(s)D^{\frac{1}{2}})(\Phi(s)D^{\frac{1}{2}})^* \right) ds \right]^{\frac{1}{2}}. \end{aligned}$$

We then have the following

Lemma 2.2.3. *For a predictable process $(\Phi(s))_{s \in [0,T^*]}$ with values in $L_2(H_0, H)$ and $|||\Phi|||_{T^*} < \infty$, there exists a sequence of elementary processes Φ_n, such that*

$$|||\Phi - \Phi_n|||_{T^*} \to 0 \qquad \textit{as } n \to \infty.$$

Proof. See Da Prato and Zabczyk (1992), Lemma 4.7. ■

Now we define for predictable $(\Phi(s))_{s\in[0,T^*]}$ with values in $L_2(H_0,H)$ and $|||\Phi|||_{T^*}<\infty$

$$\int_0^t \Phi(s)\cdot dX(s) := \lim_{n\to\infty}\int_0^t \Phi_n(s)\cdot dX(s). \tag{2.4}$$

It can be shown that this stochastic integral is well-defined and furthermore a martingale if $|||\Phi|||_{T^*}<\infty$ (see Proposition 2.2.2).

So we finally found the class of suitable integrands which ensure that the stochastic integrals inherit the martingale property, namely predictable processes with values in $L_2(H_0,H)$ which satisfy $|||\Phi|||_{T^*}<\infty$.

Finally the stochastic integral may be, by use of a localization procedure, extended to stochastically integrable processes, i.e., processes, for which

$$\mathbb{P}\Big(\int_0^{T^*} ||\Phi(s)||_2^2\, ds<\infty\Big)=1.$$

Note that, in doing so, the martingale property is lost, but stochastic integrals still remain local martingales. For a full treatment see Da Prato and Zabczyk (1992, p. 94 p.p.).

2.3 Covariances

In this section we consider covariances of the previously defined stochastic integrals. The following definition is in analogy to Definition 2.1.2.

Definition 2.3.1. For two H-valued random variables X_i with mean m_i, $i=1,2$, the symmetric operator $D\in L(H)$, such that

$$\mathbb{E}\big[<X_1,f><X_2,g>\big]-<m_1,f><m_2,g>=<Df,g>$$

is called the covariance of X_1 and X_2 and denoted by $\mathrm{Cov}(X_1,X_2)$.

Proposition 2.3.2. *Assume $(\Phi_1(s))_{s\in[0,T^*]},(\Phi_2(s))_{s\in[0,T^*]}$ are predictable processes with values in $L_2(H_0,H)$, $|||\Phi_1|||_{T^*}<\infty$ and $|||\Phi_2|||_{T^*}<\infty$. Then for all $t\in[0,T^*]$*

$$\mathbb{E}\int_0^t \Phi_i(s)\cdot dX(s)=0,\qquad \mathbb{E}||\int_0^t \Phi_i(s)\cdot dX(s)||^2<\infty,\qquad i=1,2$$

and the covariance operator equals for all $t,s\in[0,T^*]$

$$\mathrm{Cov}\Big(\int_0^t \Phi_1(u)\cdot dX(u),\int_0^s \Phi_2(v)\cdot dX(v)\Big)=$$

$$=\ \mathbb{E}\int_0^{t\wedge s}\big(\Phi_2(u)D^{\frac12}\big)\big(\Phi_1(u)D^{\frac12}\big)^*\,du.$$

Furthermore,

$$\mathbb{E} < \int_0^t \Phi_1(u) \cdot dX(u), \int_0^s \Phi_2(u) \cdot dX(u) > = \mathbb{E} \int_0^{t\wedge s} \operatorname{tr}[(\Phi_2(u)D^{\frac{1}{2}})(\Phi_1(u)D^{\frac{1}{2}})^*]\, du. \tag{2.5}$$

Proof. As $(\Phi_2(u)D^{\frac{1}{2}})$ and $(\Phi_1(u)D^{\frac{1}{2}})^*$ are $L_2(H)$ valued processes, the process $(\Phi_2(u)D^{\frac{1}{2}})(\Phi_1(u)D^{\frac{1}{2}})^*$ takes values[7] in $L_1(H)$. Da Prato and Zabczyk (1992, p. 102) obtain the inequality

$$\mathbb{E} \int_0^{T^*} ||(\Phi_2(u)D^{\frac{1}{2}})(\Phi_1(u)D^{\frac{1}{2}})^*||_1\, du \leq |||\Phi_1|||_{T^*}|||\Phi_2|||_{T^*}, \tag{2.6}$$

which ensures existence of the integral.

Further on, consider elementary processes Φ_1 and Φ_2. We proceed similarly to the proof of Proposition 2.2.2. Assume w.l.o.g. that $s \leq t$ and enhance the partition by t and s at the points t_m and $t_{\tilde{m}}$, say. Then

$$\begin{aligned} &\mathbb{E}\Big(< \int_0^t \Phi_1(u)\, dX(u), a >< \int_0^s \Phi_2(v) dX(v), b > \Big) \qquad (2.7)\\ &= \sum_{i,j=0,i\neq j}^{\tilde{m}} \mathbb{E}\Big(< \Phi_1(t_i)\Delta_i X, a > < \Phi_2(t_j)\Delta_j X, b > \Big) \\ &\quad + \sum_{i=0}^{\tilde{m}} \sum_{j=\tilde{m}+1}^{m} \mathbb{E}\Big(< \Phi_1(t_i)\Delta_i X, a > < \Phi_2(t_j)\Delta_j X, b > \Big) \\ &\quad + \sum_{i=0}^{\tilde{m}} \mathbb{E}\Big(< \Phi_1(t_i)\Delta_i X, a > < \Phi_2(t_i)\Delta_i X, b > \Big). \end{aligned}$$

As for $i \neq j$

$$\mathbb{E}\Big(< \Phi_1(t_i)\Delta_i X, a > < \Phi_2(t_j)\Delta_j X, b > \Big) = 0$$

the first two sums vanish. Furthermore, $\Delta_i X \sim \mathcal{N}(0, (t_{i+1} - t_i)D)$ and we obtain

$$\begin{aligned} &\mathbb{E}\Big(< \Phi_1(t_i)\Delta_i X, a > < \Phi_2(t_i)\Delta_i X, b > \Big) \\ &= \mathbb{E}\Big(< \Phi_1(t_i) \sum_{j=1}^{\infty} \Delta_i\beta_j e_j, a > < \Phi_2(t_i) \sum_{k=1}^{\infty} \Delta_i\beta_k e_k, b > \Big) \\ &= \mathbb{E}\Big[\sum_{j,k=1}^{\infty} < \Phi_1(t_i)e_j, a >< \Phi_2(t_i)e_k, b > \mathbb{E}\big(\Delta_i\beta_j\, \Delta_i\beta_k | \mathcal{F}_{t_i}\big)\Big] \\ &= \mathbb{E}\Big[\sum_{j,k=1}^{\infty} < \Phi_1(t_i)e_j, a >< \Phi_2(t_i)e_k, b > \delta_{jk}\lambda_k(t_{i+1} - t_i)\Big] \\ &= \mathbb{E}\Big[\sum_{k=1}^{\infty} < \Phi_1(t_i)D^{\frac{1}{2}}e_k\, , a >< \Phi_2(t_i)D^{\frac{1}{2}}e_k\, , b > \Big](t_{i+1} - t_i). \qquad (2.8) \end{aligned}$$

[7]Here, $L_1(H)$ is the Banach space of all trace-class operators in $L(H)$, see Page 51.

This yields

$$(2.7) \quad = \quad < \mathbb{E} \int_0^{t\wedge s} (\Phi_2(u) D^{1/2})(\Phi_1(u) D^{1/2})^* \, du \, a, b > .$$

So the conclusion holds for elementary processes. With the bound (2.6) the general conclusion follows from an appropriate approximation through elementary processes.

Observe that equation (2.8) yields

$$(2.7) = \sum_{i=0}^{\tilde{m}} \mathbb{E}\Big[< (\Phi_2(t_i) D^{\frac{1}{2}})(\Phi_1(t_i) D^{\frac{1}{2}})^* a \, , b > \Big] (t_{i+1} - t_i)$$

and (2.5) follows immediately. ■

2.4 Itô's formula

The formula of Itô (1946) yields the chain rule for functions of diffusion processes. In comparison to the fundamental theorem of calculus there appears an unexpected second term. As the formula mainly relies on the Taylor formula this is a result of the non vanishing second-order term and leads to interesting probabilistic interpretations. The reason for its appearance is due to infinite variation of the Brownian motion. Interestingly, there is a close analogue to processes in Hilbert spaces which is derived in this chapter.

We only cite the Taylor formula for Hilbert spaces. A detailed treatment may be found in Dieudonné (1969). Consider a Hilbert space H and an open subset $A \subset H$. It may be recalled that, if the derivative of a continuous mapping $f : A \mapsto H$ denoted by Df exists, it is a continuous and linear mapping form H into H and therefore an element of the Banach space $L(H)$.

Furthermore, if the second derivative $D^2 f$ exists, it is an element of $L(H; L(H))$ and a symmetric[8] mapping. The space $L(H; L(H))$ can be identified[9] with the space of continuous bilinear mappings of $H \times H$ into H, denoted by $L(H, H; H)$.

As a result of the mean value theorem we obtain **Taylor's formula**:

Theorem 2.4.1. *Assume f is a twice continuously differentiable mapping of A into H. If $x + \theta\, t \in A$ for $x, t \in H$ and all $\theta \in [0, 1]$, we have*

$$f(x+t) = f(x) + Df(x) \cdot t + \frac{1}{2} D^2 f(x + \zeta t) \cdot (t, t),$$

where ζ is an element of $[0, 1]$.

[8] $D^2 f$ is symmetric in the sense that $D^2 f \cdot (f, g) = D^2 f \cdot (g, f)$.

[9] By $h \cdot (s, t) \simeq (h \cdot s) \cdot t$.

For the Itô-formula on Hilbert spaces we consider a D-Wiener process $(X(s))_{s\geq 0}$ on H and a predictable process $(\Phi(s))_{s\geq 0}$ with values in $L_2(H_0, H)$, such that $|||\Phi|||_{T^*} < \infty$. Then the stochastic process

$$S(t) = S(0) + \int_0^t \Phi(s) \cdot dX(s) \tag{2.9}$$

is a square-integrable martingale, as already mentioned in the previous section.

Theorem 2.4.2. *For an open subset A of the Hilbert space H, let $f : A \mapsto H$ be a function, whose first and second derivative is uniformly continuous on bounded subsets of A. For $(S(t))_{t\in[0,T^*]}$, as in (2.9) we have for all $t \in [0, T^*]$ $I\!P$-a.s.*

$$f(S(t)) = f(S(0)) + \int_0^t Df(S(u)) \cdot dS(u) + \int_0^t \sum_{k=1}^{\infty} \lambda_k D^2 f(S(u)) \cdot (\Phi(u) \cdot e_k, \Phi(u) \cdot e_k)\, du.$$

Note that the first integral equals

$$\int_0^t Df(S(u)) \cdot \Phi(u) \cdot dX(u).$$

Proof. By a localization procedure we can restrict ourselves to bounded $(X(s))_{s\in[0,T^*]}$ and $(\Phi(s))_{s\in[0,T^*]}$, see Da Prato and Zabczyk (1992, p. 106).

Further on, consider a partition $\Pi = \{t_0, t_1, \ldots, t_n\}$ of $[0, t]$ with $0 = t_0 < t_1 < \cdots < t_n$ and denote its mesh by $||\Pi|| := \max_{1\leq i\leq n}(t_i - t_{i-1})$.

Using the Taylor formula on Banach spaces and writing $(t)^{(2)}$ for (t, t), we obtain

$$\begin{aligned}
f(S_t) - f(S_0) &= \sum_{j=0}^{n-1} f(S(t_{j+1})) - f(S(t_j)) \\
&= \sum_{j=0}^{n-1} Df(S(t_j)) \cdot (S(t_{j+1}) - S(t_j)) + \frac{1}{2} D^2 f(\tilde{S}_j) \cdot (S(t_{j+1}) - S(t_j))^{(2)} \\
&= \sum_{j=0}^{n-1} Df(S(t_j)) \cdot \Delta_j S + \frac{1}{2} D^2 f(S(t_j)) \cdot (\Delta_j S)^{(2)} \\
&\qquad + \frac{1}{2}\left[D^2 f(\tilde{S}_j) - D^2 f(S(t_j))\right] \cdot (\Delta_j S)^{(2)} \\
&= I + II + III,
\end{aligned}$$

where we set $\tilde{S}_j := S(t_j) + \zeta_j(S(t_{j+1}) - S(t_j))$ with $\zeta_j = \zeta_j(\omega) \in [0, 1]$.

Considering I, we intend to approximate $Y_s := Df(S(s))$ by the elementary process

$$Y_s^n := Df(S(0))1_{\{0\}}(s) + \sum_{j=0}^{n-1} Df(S(t_j))1_{(t_j,t_{j+1}]}(s).$$

Indeed, uniform continuity of the derivative and the bounded convergence theorem yield

$$|||Y - Y^n|||_{T^*} = \mathbb{E}\int_0^{T^*} ||Y(u) - Y^n(u)||_2^2 \, du \longrightarrow 0$$

as the mesh of the partition tends to zero.

Then, by Definition (2.4), we have $\mathbb{P}$-a.s. for $||\Pi|| \to 0$,

$$I \longrightarrow \int_0^t Df(S(s)) \cdot dS(s).$$

The third summand, III, converges to 0 $\mathbb{P}$-a.s. for $||\Pi|| \to 0$, because of continuity of the derivative, using a similar argument.

Consider the second term, II. We calculate the conditional expectation of the summands

$$\begin{aligned} &\mathbb{E}\Big(D^2 f(S(t_j)) \cdot (\Delta_j S)^{(2)} \Big| \mathcal{F}_{t_j}\Big) \\ &\qquad = \mathbb{E}\Big(D^2 f(S(t_j)) \cdot (\Phi(t_j) \cdot \Delta_j X)^{(2)} \Big| \mathcal{F}_{t_j}\Big). \end{aligned} \tag{2.10}$$

Using the Eigenvalue expansion of X, we obtain

$$\begin{aligned} (2.10) \quad &= \sum_{k,l=1}^{\infty} \mathbb{E}\Big(\Delta_j\beta_k \, \Delta_j\beta_l \Big| \mathcal{F}_{t_j}\Big) D^2 f(S(t_j)) \cdot (\Phi(t_j) \cdot e_k, \Phi(t_j) \cdot e_l) \\ &\overset{(2.3)}{=} \sum_{k=1}^{\infty} (t_{j+1} - t_j)\lambda_k \, D^2 f(S(t_j)) \cdot (\Phi(t_j) \cdot e_k, \Phi(t_j) \cdot e_k) \\ &=: (t_{j+1} - t_j) \, J(t_j). \end{aligned} \tag{2.11}$$

To show L^2-convergence it suffices to prove that the following expectation converges to zero:

$$\begin{aligned} &\mathbb{E}\Big[\sum_{j=0}^{n-1} \Big(D^2 f(S(t_j))(\Delta_j S)^{(2)} - (t_{j+1} - t_j) \, J(t_j)\Big)\Big]^2 \\ &\qquad = \sum_{j=0}^{n-1} \Big(\mathbb{E}\big[D^2 f(S(t_j))(\Delta_j S)^{(2)}\big]^2 - (t_{j+1} - t_j)^2 \mathbb{E}\big[J(t_j)\big]^2\Big). \end{aligned} \tag{2.12}$$

In the last step we used the fact that the summands are independent from each other and, due to (2.11), have zero mean.

If we expand the first summand via (2.2) and denote $D^2 f(S(t_j))(\Phi(t_j)e_k, \Phi(t_j)e_l) =: \xi_{k,l}^j$

we get

$$\sum_{j=0}^{n-1} \mathbb{E}\Big[D^2 f(S(t_j))(\Delta_j S)^{(2)}\Big]^2$$

$$= \sum_{j=0}^{n-1} \mathbb{E}\Big\{ \sum_{k,l,m,n=1}^{\infty} \mathbb{E}\Big[\Delta_j\beta_k\, \Delta_j\beta_l\, \Delta_j\beta_m\, \Delta_j\beta_n \Big| \mathcal{F}_{t_j}\Big] \xi^j_{k,l}\, \xi^j_{m,n} \Big\}$$

$$= \sum_{j=0}^{n-1} \sum_{k,m=1}^{\infty} \lambda_k\lambda_m (t_{j+1}-t_j)^2 \mathbb{E}(\xi^j_{k,k}\xi^j_{m,m}) + \sum_{j=0}^{n-1}\sum_{k=1}^{\infty} \lambda_k^3 (t_{j+1}-t_j)^2 \mathbb{E}\big((\xi^j_{k,k})^2\big).$$

Second moments of $\xi^j_{k,l}$ are bounded for any j, k, l because $D^2 f$ itself is bounded by assumption. So the last sum converges to zero as $||\Pi|| \to 0$. For the second summand of (2.12) we conclude

$$\sum_{j=0}^{n-1} (t_{j+1}-t_j)^2 \mathbb{E}\Big[\sum_{k=1}^{\infty} \lambda_k\, D^2 f(S(t_j)) \cdot (\Phi(t_j)\cdot e_k, \Phi(t_j)\cdot e_k)\Big]^2$$

$$= \sum_{j=0}^{n-1} (t_{j+1}-t_j)^2 \mathbb{E}\Big[\sum_{k=1}^{\infty} \lambda_k \xi^j_{k,k}\Big]^2$$

$$= \sum_{j=0}^{n-1} (t_{j+1}-t_j)^2 \sum_{k,m=1}^{\infty} \lambda_k\lambda_m \mathbb{E}\big[\xi^j_{k,k}\xi^j_{m,m}\big],$$

which also converges to zero as $\sup_j (t_{j+1}-t_j) \to 0$.

Up to now we obtained convergence in L^2. Considering a subsequence of $\{\Pi^{(n)}\}_{n=1}^{\infty}$ yields the desired $\mathbb{P}$-a.s. convergence, c.f. Karatzas and Shreve (1988, p.152). ■

The Itô-formula can be extended to processes of the type

$$S(t) = S(0) + \int_0^t \mu(s)\, ds + \int_0^t \Phi(s)\, dX(s),$$

where $(\mu(s))_{s\in[0,T^*]}$ is an adapted, H-valued process. Also the function f might be time-dependent. See Da Prato and Zabczyk (1992, p. 105-108).

2.5 The Fubini Theorem

The Fubini theorem is just stated for convenience. For a proof, see Da Prato and Zabczyk (1992, p. 109 p.p.).

Let $(E, \mathcal{E})$ denote a measurable space and μ be a finite, positive measure on $(E, \mathcal{E})$. Furthermore, consider a predictable, measurable mapping[10]

$$\Phi(t, \omega, x) : ([0, T] \times \omega \times E) \to L_2(H_0, H).$$

[10] For details on measurability and predictability in this case, see Da Prato and Zabczyk (1992, p. 109).

Theorem 2.5.1. *Assume that*

$$\int_E |||\Phi(\cdot,\omega,x)|||_T\,\mu(dx) < \infty, \quad \text{for } \mathbb{P}\text{-almost all } \omega.$$

Then it follows that

$$\int_E \Big[\int_0^T \Phi(t,x)\cdot dX(t)\Big]\mu(dx) = \int_0^T \Big[\int_E \Phi(t,x)\,\mu(dx)\Big]\cdot dX(t), \quad \mathbb{P}\text{-a.s.}$$

2.6 Girsanov's Theorem

Recall that we already defined the Hilbert space $H_0 = D^{\frac{1}{2}}(H)$ with inner product $< \cdot, \cdot >_0$ and the induced norm by $|\cdot|_0$ on page 56.

Theorem 2.6.1. *Consider predictable process $(\mu(s))_{s\in[0,T^*]}$ with values in H_0 and set $\Phi(s)(\cdot) := < \mu(s), \cdot >_0$. Assume that*

$$\mathbb{E}\Big(\exp\Big[\int_0^{T^*}\Phi(s)\cdot dX(s) - \frac{1}{2}\int_0^{T^*}|\mu(s)|_0^2\,ds\Big]\Big) = 1,$$

where $(X(s))_{s\in[0,T^*]}$ is a D-Wiener process under the measure P. Then the process

$$\tilde{X}(t) := X(t) - \int_0^t \mu(s)\,ds, \qquad t\in[0,T^*]$$

is a D-Wiener process under the measure $\tilde{P}$, defined by

$$d\tilde{P} := \exp\Big[\int_0^{T^*}\Phi(s)\cdot dX(s) - \frac{1}{2}\int_0^{T^*}|\mu(s)|_0^2\,ds\Big]\,dP.$$

Note that $< \mu(s), \cdot >_0$ is a linear mapping from H into $\mathbb{R}$, thus $\Phi(s) \in L(H,\mathbb{R})$. This requires a slightly more general definition of the stochastic integral as obtained up to now. Nevertheless, this is achieved analogously and the reader is referred to Da Prato and Zabczyk (1992, p. 290), where the theorem is proved.

The process $(\tilde{X}(t))_{t\in[0,T^*]}$ is a so-called D-Wiener process with drift $(\mu_s)_{s\in[0,T^*]}$. The Girsanov theorem shows that one obtains a D-Wiener process with zero drift under the equivalent measure $\tilde{P}$[11].

[11]Under certain circumstances, the Girsanov theorem already describes all equivalent measures, see Bogachev (1991).

Chapter 3

An Infinite Factor Model for Credit Risk

Modeling credit risk may start within the framework of Heath, Jarrow and Morton (1992) (henceforth HJM) and then be extended to credit risk. There are several ways to do this, and in the next two sections we present an approach in a framework due to Duffie and Singleton (1999). We start by formulating the extension of the HJM framework to stochastic differential equations on Hilbert spaces. Our presentation uses the parameterization due to Musiela (1993), see also Bagchi and Kumar (2000) or Filipović (2001).

In Section 3.3 we present an approach based on credit ratings. We use a Markov model in combination with two different recovery structures. For a rating based recovery of market value approach with finitely many factors, see Acharya, Das and Sundaram (2000), and for a rating based recovery of treasury value approach, see Bielecki and Rutkowski (2000). We extend both models using SDEs on Hilbert spaces. Furthermore, recent research in Özkan and Schmidt (2003) extends this to Lévy processes in infinite dimensions.

The arbitrage-free conditions are presented in a fashion which clarifies the connection between the defaultable spot rate, default intensity and the recovery structure.

At this point the question naturally arises, why to consider infinite dimensional models for the term structure of interest rates. Traditionally the infinite number of forward rates in a term structure model are defined via a diffusion driven by a finite number of Brownian motions. This choice enables analytical tractability and is usually justified with a view towards the empirical fact that the first three principal components describe 95% of the observed variance. However, as pointed out in Cont (2001), dealing with interest rate derivatives typically involves expectations of non-linear functions of the forward rate curve. Therefore, a model which might explain the variance of the forward rate quite well may still lack some principal components which have a non-negligible effect on the fluctuations of such derivatives.

Another argument towards infinite dimensional models arises in Chapter 4, namely that a calibration based on such a model may show better numerical results and may help to avoid frequent re-calibrations, while analytical tractability is preserved.

From now on we always consider the objective measure P and a measure Q which is equivalent to P. The following theorems offer conditions, under which all discounted bond prices are martingales under Q. Then Q is called an *equivalent martingale measure* and, as shown by Björk, di Masi, Kabanov and Runggaldier (1997), the market is free of arbitrage.

3.1 An Infinite Factor HJM Extension

To develop our model with credit risk in infinite dimensions, we first discuss the methodology in the case without credit risk. Kennedy (1994) gives an interest rate formulation with Gaussian random fields. This approach was extended to more general models using SDEs on Hilbert spaces by Goldstein (1997), Santa-Clara and Sornette (1997) and Bagchi and Kumar (2000). The framework we present includes the first two and is a special case of the last.

We derive the analogue of the drift condition of Heath, Jarrow and Morton (1992) in an infinite dimensional setting. Starting with a model under some measure Q, we derive a condition under which Q is a martingale measure.

The idea of the HJM approach is to model the dynamics of the forward rates itself rather than to model the dynamics of the instantaneous interest rate and then derive the dynamics of the forward rates. The forward rates have a one-to-one correspondence to bond prices, which in the continuous-time case amounts to

$$B(t,T) = \exp\big[-\int_t^T f(t,u)\,du\big].$$

Usually, the forward rate is modeled via an n-dimensional Brownian motion as

$$df(t,T) = \alpha(t,T)\,dt + \boldsymbol{\sigma}(t,T)\cdot d\mathbf{W}_t, \qquad 0 \le t \le T, \tag{3.1}$$

where $\alpha(t,T) \in \mathbb{R}$ and $\boldsymbol{\sigma}(t,T) \in \mathbb{R}^n$ form predictable processes.

Noticing that the forward-rate curve at time t, denoted by $f(t,\cdot) : [0,T^*] \mapsto \mathbb{R}$, is a function (of T), one could model a stochastic process $f(t)$ which itself takes values in a functional space. So the question arises which functional space to choose. Usually there are forward rates up to a maximum time-to-maturity in the market, say T^{**}. Consequently, on can express the forward rates as $f(t,t+x) : [0,T^*]\times[0,T^{**}] \mapsto \mathbb{R}$. This leads to the so-called Musiela parameterization[1]. One considers

$$r_t(x) := f(t,t+x),$$

where the stochastic process $(r_t)_{t\in[0,T^*]}$ takes values in a functional space $\mathbb{R}^{[0,T^{**}]}$. Sometimes we use r_t to denote the spot rate, $r_t(0)$.

[1]See Musiela (1993).

It turns out that it is appropriate to consider stochastic differential equations on Hilbert spaces. Throughout this chapter H stands for a separable Hilbert space, and our intention is to use a space of real-valued functions on an interval $[0, T^{**}]$. A different approach towards modeling the forward rates uses Gaussian random fields and is presented in Chapter 4.

First we have to restate equation (3.1) in terms of $r_t(x)$. Note that this equation is equivalent to (set $x := T - t$)

$$\begin{aligned} r_t(x) &= f(t,T) \\ &= f(0,T) + \int_0^t \alpha(u,T)\,du + \int_0^t \boldsymbol{\sigma}(u,T)\cdot d\mathbf{W}_u \\ &= r_0(t+x) + \int_0^t \alpha(u,t+x)\,du + \int_0^t \boldsymbol{\sigma}(u,t+x)\cdot d\mathbf{W}_u. \end{aligned}$$

Let $\{S(t) | t \in \mathbb{R}_+\}$ denote the semigroup of right shifts, defined by $S(t)g(x) = g(x+t)$, for any function $g : \mathbb{R}_+ \mapsto \mathbb{R}$. This enables us to obtain a consistent formulation within a functional setting by

$$\begin{aligned} r_t(x) &= S(t)r_0(x) + \int_0^t S(t)\alpha(u,x)\,du + \int_0^t S(t)\boldsymbol{\sigma}(u,x)\cdot d\mathbf{W}_u \\ \Leftrightarrow \qquad r_t &= S(t)r_0 + \int_0^t S(t)\alpha(u)\,du + \int_0^t S(t)\boldsymbol{\sigma}(u)\cdot d\mathbf{W}_u, \end{aligned}$$

where r_0, $\alpha(u)$ and $\boldsymbol{\sigma}(u)$ are itself elements of H. In this formulation the shift operator arises naturally, as forward rates with fixed maturity correspond to forward rates with decreasing time-to-maturity, see also Figure 3.1.

In this section we will generalize the integral with respect to $\mathbf{W}_t$ to Wiener processes on the Hilbert space H.

Consider stochastic processes $\alpha : [0, T^*] \times \Omega \mapsto H$ and $\sigma : [0, T^*] \times \Omega \mapsto L(H; H)$, both predictable w.r.t. $(\mathcal{F}_t)_{t \geq 0}$, satisfying $\mathbb{P}(\int_0^{T^*} \alpha(s)\,ds < \infty) = 1$ and $|||\sigma|||_{T^*} < \infty$. Further on, assume that $(X(t))_{t \geq 0}$ is a D-Wiener process as defined in the preceding chapter. Assume the forward rate dynamics to follow

$$r_t = S(t)r_0 + \int_0^t S(t)\alpha(u)\,du + S(t)\big[\int_0^t \sigma(u)\cdot dX(u)\big]. \tag{3.2}$$

Note that r_t takes values in H, so it represents the whole forward-rate curve, otherwise denoted by $f(t, t+x)$. For α we could explicitly write $\alpha(t,x)$ while this is not possible for σ. Still, even if the index x does not appear directly, it is not obsolete. As the last integral is an element of H for all t we can write it either as

$$\int_0^t \sigma(u)\,dX(u) =: I(t) \in H \tag{3.3}$$

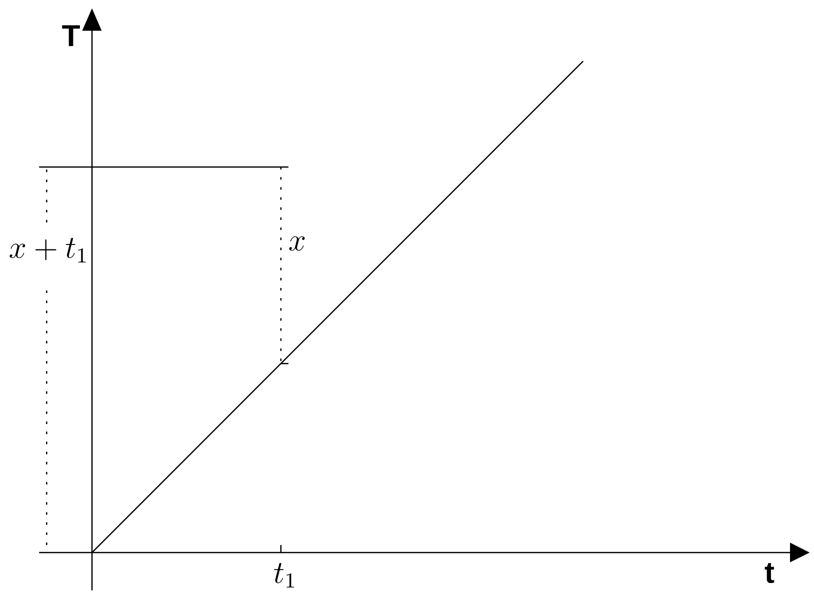

Figure 3.1: The shift occurring in the Musiela parameterization: t denotes current time, while T denotes maturity. The forward rate $f(t_1, x)$ relates naturally to the forward rate $f(0, t_1 + x) = S(t_1)f(0, x)$. This leads to the shift terms in equation (3.2).

or directly as $I(t, x)$. The shift operator therefore yields $S(t)I(t, x) = I(t, t + x)$.

Using the Eigenvalue expansion of X, see equation (2.2), we have the decomposition

$$X(u) = \sum_{k=1}^{\infty} \beta_k(u) e_k,$$

where $\frac{1}{\sqrt{\beta_k(u)}}$ are independent, standard Brownian motions on the real line. Then we denote

$$\sigma_k(u, v) := \big(\sigma(u) \cdot e_k\big)(v),$$

and get, in the above notation, the following

Theorem 3.1.1. *Set $\alpha^*(u, T) := \int_u^T \alpha(u, v)\, dv$ and $\sigma_k^*(u, T) := \int_u^T \sigma_k(u, v)\, dv$. Then all discounted bond prices are martingales iff*

$$\alpha(t, T) = \sum_{k=1}^{\infty} \lambda_k\, \sigma_k^*(t, T) \cdot \sigma_k(t, T) \qquad \forall t \in [0, T^*], T \in [t, t + T^{**}]. \tag{3.4}$$

Equation (3.4) is often referred to as the *drift condition*. Note that the drift condition derived by Heath, Jarrow and Morton (1992) is the special case corresponding to $\lambda_k = 1$ for $k = 1$ and zero otherwise.

Intuitively, the drift condition means that, once the volatility (and dependence) structure is specified, the dynamics under the arbitrage-free measure is fixed. As a change of measure does not change the volatility structure, this could be estimated using historical data. A different approach to obtain the volatility structure uses a calibration to market prices, as discussed in detail in Section 4.5.

Forward rates observed in the market have a time-to-maturity of up to 20 years or more, while the time horizon for credit derivatives is relatively small. This implies for our model that $T^{**} > T^*$, which plays a role, for example, in the drift condition.

Denote the measure under which the above dynamics takes place by Q. If this measure is equivalent to the objective measure P and the drift-condition is satisfied, then the market is free of arbitrage. Completeness follows if the equivalent martingale measure is unique. Conditions under which this holds true in the above setting are to the best of our knowledge not yet available.

Proof of Theorem 3.1.1. In the Musiela parameterization, the bond price equals

$$B(t,T) = \exp(-\int_0^{T-t} r_t(v)\,dv).$$

Setting $y(t,T) := -\int_0^{T-t} r_t(v)\,dv$, we need to derive its dynamics. Using the notation of the stochastic integral via $I(t,v)$, see (3.3), we write

$$y(t,T) = -\int_0^{T-t} r_0(v+t)\,dv - \int_0^{T-t}\int_0^t \alpha(u,v+t)\,du\,dv - \int_0^{T-t} I(t,v+t)\,dv.$$

With $y(0,T) = -\int_0^T r_0(v)\,dv$ we have that

$$\begin{aligned} -\int_0^{T-t} r_0(v+t)\,dv &= y(0,T) + \int_0^T r_0(v)\,dv - \int_0^{T-t} r_0(v+t)\,dv \\ &= y(0,T) + \int_0^t r_0(v)\,dv. \end{aligned}$$

As we would like to apply Itô's formula to prove the martingale property we need to have some dynamics of y w.r.t. dX, which requires interchanging the integration. With the aid of the Eigenvalue expansion we have

$$\begin{aligned} I(t) &= \int_0^t \sigma(u)\,dX(u) \\ &= \int_0^t \sigma(u)\,d\Big(\sum_{k=1}^\infty \beta_k(u)e_k\Big) \\ &= \sum_{k=1}^\infty \int_0^t \sigma(u)\cdot e_k\,d\beta_k(u). \end{aligned}$$

The last equality holds because $|||\sigma|||_{T^*} < \infty$, see Da Prato and Zabczyk (1992, p. 99). Both $I(t)$ and $\sigma(u)\cdot e_k$ are elements of H and may be written as $I(t,v)$ and $(\sigma(u)\cdot e_k)(v)$, respectively. We obtain the representation

$$I(t,v) \;=\; \sum_{k=1}^{\infty} \int_0^t \sigma_k(u,v)\, d\beta_k(u). \tag{3.5}$$

The integrability condition $|||\sigma|||_{T^*} < \infty$ allows us to use the stochastic Fubini Theorem 2.5.1, see Filipović (2001). This yields

$$\begin{aligned}\int_0^{T-t} I(t,v+t)\, dv &= \int_0^{T-t} \sum_{k=1}^{\infty} \int_0^t \sigma_k(u,v+t)\, d\beta_k(u)\, dv \\ &= \sum_{k=1}^{\infty} \int_0^t \int_0^{T-t} \sigma_k(u,v+t)\, dv\, d\beta_k(u).\end{aligned}$$

Applying the obtained representation leads to

$$\begin{aligned}y(t,T) \;=\; & y(0,T) + \int_0^t r_0(v)\, dv - \int_0^{T-t} \int_0^t \alpha(u,v+t)\, du\, dv \\ & - \sum_{k=1}^{\infty} \int_0^t \int_0^{T-t} \sigma_k(u,v+t)\, dv\, d\beta_k(u).\end{aligned} \tag{3.6}$$

Using the standard Fubini theorem we can interchange the order of the α-integral. We want to introduce the spot rate, $r_t(0)$ into the above formula. By its dynamics (3.2) we obtain

$$\begin{aligned}\int_0^t r_v(0)\, dv &= \int_0^t \Big[r_0(v) + \int_0^v \alpha(u,v)\, du + [\int_0^v \sigma(u)\cdot dX(u)](v) \Big]\, dv \\ &= \int_0^t r_0(v)\, dv + \int_0^t \int_0^v \alpha(u,v)\, du\, dv + \int_0^t [\int_0^v \sigma(u)\cdot dX(u)](v)\, dv.\end{aligned}$$

Again using the decomposition (3.5) and Fubini's theorem yields

$$\begin{aligned}\int_0^t r_v(0)\, dv &= \int_0^t r_0(v)\, dv + \int_0^t \int_0^v \alpha(u,v)\, du\, dv + \int_0^t \sum_{k=1}^{\infty} \int_0^v \sigma_k(u,v) d\beta_k(u)\, dv \\ &= \int_0^t r_0(v)\, dv + \int_0^t \int_u^t \alpha(u,v)\, dv\, du + \sum_{k=1}^{\infty} \int_0^t \int_u^t \sigma_k(u,v)\, dv\, d\beta_k(u).\end{aligned}$$

Note that $\int r_0(v)dv$ also appears in (3.6). Thus, we obtain

$$\begin{aligned}
y(t,T) &= y(0,T) + \int_0^t r_v(0)\,dv \\
&\quad - \int_0^t \int_u^t \alpha(u,v)\,dv\,du - \int_0^t \int_t^T \alpha(u,v)\,dv\,du \\
&\quad - \sum_{k=1}^{\infty} \int_0^t \int_t^T \sigma_k(u,v)\,dv\,d\beta_k(u) - \sum_{k=1}^{\infty} \int_0^t \int_u^t \sigma_k(u,v)\,dv\,d\beta_k(u) \\
&= y(0,T) + \int_0^t r_u(0)\,du \\
&\quad - \int_0^t \alpha^*(u,T)\,du - \sum_{k=1}^{\infty} \int_0^t \sigma_k^*(u,T)\,d\beta_k(u),
\end{aligned}$$

where we used $\alpha^*(u,T) = \int_u^T \alpha(u,v)\,dv$ and $\sigma_k^*(u,T) = \int_u^T \sigma_k(u,v)\,dv$.

To apply the Itô - formula 2.4.2 we look for a representation in a more functional analytic way. Define an operator $\Phi : [0,T^*] \times \Omega \mapsto L(H;H)$, by

$$[\Phi(u)\cdot f](\cdot) := \int_u^{\cdot} [\sigma(u)\cdot f](v)\,dv.$$

Then

$$\begin{aligned}
[\Phi(u)\cdot dX(u)](T) &= \sum_{k=1}^{\infty} \big[\Phi(u)\cdot e_k\big](T)\,d\beta_k(u) \\
&= \sum_{k=1}^{\infty} \int_u^T [\sigma(u)\cdot e_k](v)\,dv\,d\beta_k(u) \\
&= \sum_{k=1}^{\infty} \sigma_k^*(u,T)\,d\beta_k(u).
\end{aligned}$$

Setting $\mu(u,\cdot) := r_u(0) - \alpha^*(u,\cdot)$ we obtain

$$y(t) = y(0) + \int_0^t \mu(u)\,du - \int_0^t \Phi(u)\cdot dX(u).$$

This is the representation of y that we were looking for. The second step is to derive the dynamics of the bond price $B(t,T) = \exp(y(t,T))$. To apply Itô's formula, we define

$$F : A \mapsto H, \qquad g(\cdot) \to \exp(g(\cdot)).$$

Here A is chosen in a way, such that $\exp(g(\cdot))$, defined by $x \mapsto \exp(g(x))$, for all $x \in \mathbb{R}$, is again an element of H . Then we have $B(t,\cdot) = [F(y(t))](\cdot)$ or $B(t) = F(y(t))$, respectively.

We compute the first and second derivative of F. First, define for $f, g \in H$ the product of f and g by

$$(f \times g)(\cdot) := f(\cdot)g(\cdot)$$

and write

$$g^k := \underbrace{g \times \cdots \times g}_{k \text{ times}}.$$

Then $F(g(\cdot)) = \exp(g(\cdot)) = \sum_{k=1}^{\infty} \frac{g(\cdot)^k}{k!}$. The derivative of g^2 is

$$D(g^2)(x) = 2x \times \mathrm{id},$$

where id is the identity on H. This is true, because for $x, x_0 \in \mathbb{R}$

$$\begin{aligned} g^2(x) - g^2(x_0) - 2x_0 \times (x - x_0) &= x \cdot x - x_0 \cdot x_0 - 2x_0 \cdot (x - x_0) \\ &= (x - x_0)^2 \end{aligned}$$

and therefore

$$\begin{aligned} \lim_{x \to x_0} \frac{||g^2(x) - g^2(x_0) - 2x_0 \times (x - x_0)||}{||x - x_0||} &= \lim_{x \to x_0} \frac{||(x - x_0)^2||}{||x - x_0||} \\ &\leq \lim_{x \to x_0} \frac{||x - x_0||}{||x - x_0||} \, ||x - x_0|| = 0. \end{aligned}$$

The derivative of g^n is easily obtained by induction and we may conclude

$$DF(g) = F(g) \times \mathrm{id}$$

as well as

$$D^2F(g) = F(g) \times \mathrm{id} \times \mathrm{id}\,.$$

Applying Itô's formula yields

$$\begin{aligned} dB(t) &= DF(B(u)) \cdot [\mu(t)\, dt - \Phi(t) \cdot dX(t)] \\ &\quad + \frac{1}{2} \sum_{k=1}^{\infty} \lambda_k D^2F(B(t))\, (\Phi(t) \cdot e_k, \Phi(t) \cdot e_k)\, dt \\ &= B(t) \times [\mu(t)\, dt - \Phi(t) \cdot dX(t)] \\ &\quad + \frac{1}{2} \sum_{k=1}^{\infty} \lambda_k B(t) \times (\Phi(t) \cdot e_k) \times (\Phi(t) \cdot e_k)\, dt. \end{aligned}$$

Evaluating $B(t,\cdot)$ at maturity T leads to

$$\begin{aligned} dB(t,T) &= B(t,T) \Bigg[(r_t(0) - \alpha^*(t,T))\, dt - \sum_{k=1}^{\infty} \sigma_k^*(t,T)\, d\beta_k(t) \\ &\quad + \frac{1}{2} \sum_{k=1}^{\infty} \lambda_k [\sigma_k^*(t,T)]^2\, dt \Bigg]. \end{aligned} \tag{3.7}$$

Define the discounting process $D_t := \exp(-\int_0^t r_u\,du)$. Note that as D_t is differentiable, it is of finite variation. Applying the common Itô-formula[2] to the discounted bond price therefore yields

$$\begin{aligned} d[D_tB(t,T)] &= (-r_t)\,D_tB(t,T)\,dt + D_t\,dB(t,T) \\ &= D_tB(t,T)\Big[\Big(r_t(0) - r_t - \alpha^*(t,T) + \frac{1}{2}\sum_{k=1}^{\infty}\lambda_k[\sigma_k^*(t,T)]^2\Big)\,dt \\ &\quad -\sum_{k=1}^{\infty}\sigma_k^*(t,T)\,d\beta_k(t)\Big]. \end{aligned} \tag{3.8}$$

Note that we stress the dependence on $(r_t(0) - r_t)$, which is in this case equal to 0. In the case with credit risk we consider $\bar{r}_t(0)$ instead of $r_t(0)$ and this term will not vanish.

Consequently the discounted bond price is a martingale under $|||\sigma|||_{T^*} < \infty$, iff

$$\alpha^*(t,T) = \frac{1}{2}\sum_{k=1}^{\infty}\lambda_k[\sigma_k^*(t,T)]^2\,dt, \qquad \forall T \in [t, t+T^{**}].$$

Using the definitions of α^* and σ^* we obtain

$$\int_t^T \alpha(t,u)\,du = \frac{1}{2}\sum_{k=1}^{\infty}\lambda_k\Big[\int_t^T \sigma_k(t,u)\,du\Big]^2\,du\,dt.$$

Taking the partial derivative w.r.t. T, we get

$$\alpha(t,T) \;=\; \sum_{k=1}^{\infty}\lambda_k\,\sigma_k^*(t,T)\cdot\sigma_k(t,T).$$

■

3.1.1 Change of Measure

Up to now we considered the model under a measure Q and obtained conditions, under which Q is a martingale measure. In fact, the observed dynamics takes place under the objective measure P, and we have to perform a change of measure to obtain the risk-neutral dynamics, which is necessary for pricing and hedging.

The main tools for doing so is the Girsanov Theorem 2.6.1. Once the drift condition is obtained, the procedure for obtaining Q is similar throughout all models.

Observe that the dynamics remains the same under all measures, just the properties of the considered processes change. In particular, if $(X(s))_{s\in[0,T^*]}$ is a D-Wiener process under P,

$$\tilde{X}(s) := X(s) - \int_0^s \mu(u)\,du$$

[2] As D_t is of finite variation, this equals the product rule, compare Revuz and Yor (1994, p. 199 p.p.).

is a D-Wiener process under Q, if

$$dQ := \exp\Big[\int_0^{T^*} \Phi(s) \cdot dX(s) - \frac{1}{2}\int_0^{T^*} |\mu(s)|_0^2 \, ds\Big] \, dP$$

and $\Phi(s)(\cdot) := < \mu(s), \cdot >_0$.

We obtain the following

Proposition 3.1.2. *If there exists a predictable process $(\mu(s))_{s\in[0,T^*]}$ which satisfies the conditions for Theorem 2.6.1 and*

$$\big[\sigma(t)\cdot\mu(t)\big](T) = \alpha(t,T) - \sum_{k=1}^{\infty} \lambda_k \, \sigma_k^*(t,T)\cdot\sigma_k(t,T),$$

for all $t \in [0,T^*]$ and $T \in [t, t+T^{**}]$, then the measure Q as defined above is an equivalent martingale measure.

Proof. The dynamics of the forward rates equal

$$\begin{aligned}
r_t &= S(t)r_0 + \int_0^t S(t)\alpha(u)\,du + S(t)\Big[\int_0^t \sigma(u)\cdot dX(u)\Big] \\
&= S(t)r_0 + \int_0^t S(t)\alpha(u)\,du + S(t)\Big[\int_0^t \sigma(u)\cdot d\big(\tilde{X}(u) - \int_0^u \mu(v)\,dv\big)\Big] \\
&= S(t)r_0 + \int_0^t S(t)\big[\alpha(u) - \sigma(u)\cdot\mu(u)\big]\,du + S(t)\Big[\int_0^t \sigma(u)\cdot d\tilde{X}(u)\Big] \\
&=: S(t)r_0 + \int_0^t S(t)\tilde{\alpha}(u)\,du + S(t)\Big[\int_0^t \sigma(u)\cdot d\tilde{X}(u)\Big].
\end{aligned}$$

Girsanov's theorem yields that $(\tilde{X}(s))_{s\geq 0}$ is a D-Wiener process under Q. Therefore, if the drift condition for $(\tilde{\alpha}(s))$ is satisfied, Q is an equivalent martingale measure. The drift condition reveals

$$\begin{aligned}
\tilde{\alpha}(t,T) &= \alpha(t,T) - \big[\sigma(t)\cdot\mu(t)\big](T) \\
&= \sum_{k=1}^{\infty} \lambda_k \, \sigma_k^*(t,T)\cdot\sigma_k(t,T).
\end{aligned}$$

Thus, the change to the risk-neutral measure Q is possible and the market is free of arbitrage. ■

If credit risk is incorporated in this setting, the change of measure furthermore results in a change of the intensity. This is also true for the ratings model of Section 3.3, cf. Bielecki and Rutkowski (2002, Sections 4.4 and 7.2).

3.2 Models with Credit Risk

At this point we add default risk to our model. In the HJM framework with finite dimension this was first considered by Duffie and Singleton (1999). In the following we extend their results to infinite dimensions.

Consider a hazard-rate model, that is, for a given filtration $(\mathcal{G}_t)_{t\geq 0}$ of general market information, the default time τ admits an intensity $(\lambda_t)_{t\geq 0}$ which is adapted to $(\mathcal{G}_t)_{t\geq 0}$. For details see Appendix A.

As previously, we consider a separable Hilbert space H, whose elements are intended to be functions $f : [0, T^{**}] \mapsto \mathbb{R}$.

The following assumption is basic for the next two sections and summarizes the infinite dimensional setting for the defaultable forward rates.

Assumption (A1): Let $\bar{\alpha} : [0, T^*] \times \Omega \mapsto H$ and $\bar{\sigma} : [0, T^*] \times \Omega \mapsto L(H; H)$ be stochastic processes, which are predictable w.r.t. $(\mathcal{F}_t)_{t\geq 0}$ and satisfy $\mathbb{P}(\int_0^{T^*} \bar{\alpha}(s)\, ds < \infty) = 1$ and $|||\bar{\sigma}|||_{T^*} < \infty$. Furthermore, assume that the defaultable forward rate follows

$$\bar{r}_t = S(t)\bar{r}_0 + \int_0^t S(t)\bar{\alpha}(u)\, du + S(t)\big[\int_0^t \bar{\sigma}(u) \cdot d\bar{X}(u)\big],$$

where $(\bar{X}(s))_{s\in[0,T^*]}$ is a D-Wiener process.

3.2.1 Recovery of Market Value

For methods using SDEs the *recovery of market value* model is particularly well suited. In this model the dynamics before a default occurs is modeled analogously to the risk-free case. If a default occurs, say at τ, the bond loses a random fraction q_τ of its pre-default value, where $(q_s)_{s\in[0,T^*]}$ is a predictable process with values in $[0, 1]$. The remaining value is instantaneously paid to the bond holder, and therefore no more subject to default risk.

The dynamics of the defaultable bond until a default occurs is modeled by specifying the dynamics of the defaultable forward rates, denoted by $\bar{r}_t(x)$. Hence,

$$1_{\{\tau>t\}}\bar{B}(t, T) = 1_{\{\tau>t\}} \exp(-\int_0^{T-t} \bar{r}_t(u)\, du).$$

If the bond defaults within its lifetime its value at default is assumed to become

$$1_{\{\tau\leq T\}}\bar{B}(\tau, T) = 1_{\{\tau\leq T\}}(1 - q_\tau)\bar{B}(\tau-, T).$$

In contrast to other recovery models the value of the bond immediately before default has some influence on the repayment, which seems reasonable.

The value of $(1-q_\tau)\bar{B}(\tau-,T)$ is immediately available to the bond owner at default and no more subject to any risk. Therefore, the value of the defaultable bond can be represented by

$$\bar{B}(t,T) = 1_{\{\tau>t\}} \exp(-\int_0^{T-t} \bar{r}_t(u)\,du) + 1_{\{\tau\le t\}} \exp(\int_\tau^t r_u\,du)(1-q_\tau)\bar{B}(\tau-,T).$$

Now we can state the following

Theorem 3.2.1. *Assume that $\bar{\alpha}(s,x)$ is continuous in s for any $x\in[0,T^{**}]$ and assumption (A1) holds. Under the recovery of market value model, discounted bond prices are martingales, iff the following two conditions are satisfied on $\{\tau>t\}$:*

(i) For any $t\in[0,T^*]$, $t\le T\le t+T^{**}$

$$\bar{\alpha}(t,T) = \sum_{k=1}^{\infty} \bar{\lambda}_k\,\bar{\sigma}_k^*(t,T)\cdot\bar{\sigma}_k(t,T). \tag{3.9}$$

(ii) For any $t\in[0,T^*]$

$$\bar{r}_t(0) = r_t(0) + q_t\lambda_t. \tag{3.10}$$

Proof. If we denote the discounting factor by $D_t = \exp(-\int_0^t r_u\,du)$, the discounted gains process $G(t,T) := D_t\bar{B}(t,T)$ equals

$$\begin{aligned} G(t,T) &= 1_{\{\tau>t\}}D_t\bar{B}(t,T) + 1_{\{\tau\le t\}} \exp\big[-\int_0^t r_u\,du + \int_\tau^t r_u\,du\big](1-q_\tau)\bar{B}(\tau-,T) \\ &= 1_{\{\tau>t\}}D_t\bar{B}(t,T) + 1_{\{\tau\le t\}}D_\tau(1-q_\tau)\bar{B}(\tau-,T) \\ &= 1_{\{\tau>t\}}D_t\bar{B}(t,T) + \int_0^t D_s(1-q_s)\bar{B}(s-,T)\,d\Lambda_s. \end{aligned}$$

For the last representation we set

$$\Lambda_s := 1_{\{\tau\le s\}}.$$

The t-dynamics of $G(t,T)$ becomes

$$dG(t,T) = d[(1-\Lambda_t)D_t\bar{B}(t,T)] + (1-q_t)D_t\bar{B}(t-,T)\,d\Lambda_t =: (1)+(2).$$

Taking into account that Λ_t is of finite variation the first summand equals

$$(1) = -d\Lambda_t\,D_t\bar{B}(t,T) + (1-\Lambda_t)\,d[D_t\bar{B}(t,T)].$$

The computation of the discounted bond's dynamics is analogous to the risk-free case. Using formula (3.8) with $\bar{\lambda}_k, \bar{\beta}_k$, respectively, we obtain

$$\begin{aligned} d[D_t\bar{B}(t,T)] &= D_t\bar{B}(t,T)\bigg\{\Big(\bar{r}_t(0) - r_t - \bar{\alpha}^*(t,T) + \frac{1}{2}\sum_{k=1}^{\infty}\bar{\lambda}_k[\bar{\sigma}_k^*(t,T)]^2\Big)\,dt \\ &\qquad -\sum_{k=1}^{\infty}\bar{\sigma}_k^*(t,T)\,d\bar{\beta}_k(t)\bigg\}. \end{aligned} \tag{3.11}$$

$\bar{r}_s(\cdot)$ is continuous in s, because $\bar{\alpha}(s,\cdot)$ is continuous by assumption and $\bar{X}(s)$ by definition. Therefore, on $\{\tau > t\}$, we have $\bar{B}(t-,T) = \bar{B}(t,T)$.

By definition of $(\lambda_s)_{s\geq 0}$, we have that $\Lambda_s - \int_0^{s\wedge\tau} \lambda_s\, ds$ is a $\mathcal{H}$-martingale, which implies that

$$d\tilde{M}_t := d\Lambda_t - 1_{\{t\leq\tau\}}\lambda_t\, dt = d\Lambda_t - (1-\Lambda_t)\lambda_t\, dt$$

is the differential of a $\mathcal{H}$-martingale. See Bielecki and Rutkowski (2002, Lemma 4.2.1). This leads to

$$\begin{aligned} dG(t,T) &= \Big[- D_t\bar{B}(t,T) + (1-q_t)D_t\bar{B}(t,T)\Big]\, d\Lambda_t \\ &\quad +(1-\Lambda_t)\, D_t\bar{B}(t,T)\bigg\{ \Big(\bar{r}_t(0) - r_t - \bar{\alpha}^*(t,T) + \frac{1}{2}\sum_{k=1}^{\infty}\bar{\lambda}_k[\bar{\sigma}_k^*(t,T)]^2\Big)\, dt \\ &\qquad\qquad - \sum_{k=1}^{\infty}\bar{\sigma}_k^*(t,T)\, d\bar{\beta}_k(t)\bigg\} \\ &= D_t\bar{B}(t,T)\bigg\{ - q_t d\tilde{M}_t - \sum_{k=1}^{\infty}\bar{\sigma}_k^*(t,T)\, d\bar{\beta}_k(t) \\ &\qquad +(1-\Lambda_t)\Big[- q_t\lambda_t + \bar{r}_t(0) - r_t - \bar{\alpha}^*(t,T) + \frac{1}{2}\sum_{k=1}^{\infty}\bar{\lambda}_k[\bar{\sigma}_k^*(t,T)]^2\Big]\, dt\bigg\}. \end{aligned}$$

Hence the dt-term represents the drift. As $(G(t,T))_{t\geq 0}$ is a martingale, iff the drift is zero, it is a martingale, iff

$$1_{\{\tau>t\}}\Big[- q_t\lambda_t + \bar{r}_t(0) - r_t - \bar{\alpha}^*(t,T) + \frac{1}{2}\sum_{k=1}^{\infty}\bar{\lambda}_k[\bar{\sigma}_k^*(t,T)]^2\Big] = 0 \qquad \forall t \leq T. \tag{3.12}$$

Note that this is needed only for $t \leq \tau$. This is due to the assumption that the recovery value is instantaneously paid to the bond holder and therefore there is no risky dynamics after default. Consequently, equation (3.12) is true under (3.9) and (3.10).

For the converse, since this equation must hold for any $t \leq \tau \wedge T$ and the $*$-terms equal zero if $T = t$ we obtain (3.9) and then (3.10). ■

Remark 3.2.2. If one prefers a drift condition which does not depend on a particular realization of τ, the equivalency in Theorem 3.2.1 might be dropped. That is, if conditions (3.9) and (3.10) hold true for any $t \in [0, T^*]$ and $T \in [t, t+T^{**}]$, discounted bond prices are martingales, because equation (3.12) holds. Note that the converse follows only on $\{\tau > t\}$.

The underlying measure is a martingale measure iff conditions (3.9) and (3.10) are satisfied, which implies that the market is free of arbitrage. We are not able to conclude that the market is complete, because there is, to our best knowledge, no uniqueness result available yet. If this would be true, results of Björk, di Masi, Kabanov and Runggaldier (1997) could be used to show approximate completeness.

Some simple Models with infinite Factors

In this section we discuss some simple models in the above presented framework. Assuming that $\bar{\sigma}(s) : [0, T^*] \mapsto L(H; H)$ is deterministic immediately results in a Gaussian model.

In analogy to Vargiolu (2000) a historical estimation of the covariance structure using the Karhunen-Loève[3] decomposition is possible. The procedure requires two steps. First, the covariance operator is estimated using historical data. In the second step the first Eigenvectors /values are obtained, say up to a number $\bar{N}$. This results in a $\bar{N}$-factor HJM model which is used as an approximation of the infinite factor model.

Let us consider the procedure in further detail. With Proposition 2.3.2 the covariance operator of $\bar{r}(t)$ becomes

$$\mathrm{Var}(\bar{r}(t)) = \int_0^t (\bar{\sigma}(s)D^{\frac{1}{2}})\,(\bar{\sigma}(s)D^{\frac{1}{2}})^*\,ds.$$

Assuming we consider a time interval which is small enough so that variations of $\bar{\sigma}(s)$ do not play a significant role, one could use[4]

$$\frac{D_n(t_n)}{t_n - t_1} := \frac{1}{n}\sum_{i=1}^{n} \bar{r}(t_i) \otimes \bar{r}(t_i)$$

as an estimator of

$$(\bar{\sigma}(t)D^{\frac{1}{2}})\,(\bar{\sigma}(t)D^{\frac{1}{2}})^*,$$

where $t = t_n$ (or t_1, respectively $t_{n/2}$).

Similar to Section 4.5.1 focusing on the error of a finite dimensional approximation rather than pre-specifying the dimension naturally involves the Karhunen-Loève decomposition in the following way. The first $\bar{n}$ Eigenvalues and Eigenvectors of $D_n(t_n)$ can be obtained as follows. Fix $k_0 \in H$ and define

$$k^{n+1} := D_n(t_n) \cdot k^n.$$

Then k^{n+1} itself is an element of H. Vargiolu (2000) shows that

$$k^n \to e_1 \qquad \text{and} \qquad \frac{\| k^{n+1} \|}{\| k^n \|} \to \lambda_1, \qquad \text{as } n \to \infty.$$

Using $D_1 := D_n(t_n) - \lambda_1 e_1 \otimes e_1$, and applying the procedure to D_1 yields e_2 and λ_2 and so on.

[3] See, for example, Bogachev (1991, p. 55 p.p.), Da Prato and Zabczyk (1992, p. 99 p.p.) or Adler (1981).

[4] Here $\otimes$ denotes the tensor product of elements of H. The decomposition of a linear operator D into its Eigenvectors e_k and Eigenvalues λ_k then can be written in the form $D = \sum_{k=1}^{\infty} \lambda_k\, e_k \otimes e_k$. See Reed and Simon (1974).

The number of Eigenvectors, $\bar{n}$, will be chosen such that the desired precision is obtained. Finally, we approximate

$$(\bar{\sigma}(t)D^{\frac{1}{2}}) \simeq \sum_{k=1}^{\bar{n}} \lambda_k^{\frac{1}{2}} e_k,$$

and this represents the approximating $\bar{n}$-factor classical HJM model.

Cont (2001) also introduces a quite simple model using stochastic processes in Hilbert spaces, and shows that certain statistical features of the term structure of interest rates, which were observed in empirical studies, can be reproduced. In particular, the model captures imperfect correlation between maturities, mean reversion and the structure of principal components of term structure deformations.

3.2.2 Recovery of Treasury

There are different models of recovery, as already discussed in Chapter 1. An alternative to the recovery of market value is the *recovery of treasury* formulation, see Section 1.2.2. In this model, the default entails a reduction of the face value by a pre-specified constant. The reduced face value, denoted by δ, is assumed to be no more subject to default risk and is paid to the bond holder at maturity T. This is certainly equivalent to paying $\delta B(\tau, T)$ immediately at default.

Therefore the value of the defaultable bond in this model is

$$\bar{B}(t,T) = 1_{\{\tau > t\}} \exp(-\int_0^{T-t} \bar{r}_t(u)\, du) + 1_{\{\tau \leq t\}} \delta B(t,T), \qquad 0 \leq t \leq T.$$

Theorem 3.2.3. *Assume a recovery of treasury model and the riskless bond market to be arbitrage-free. Under assumption (A1), discounted defaultable bond prices are martingales, iff on $\{\tau > t\}$ for any $t \in [0, T^*], T \in [t, t+T^{**}]$*

$$\bar{r}_t(0) = r_t + \lambda_t\big(1 - \delta \frac{B(t,T)}{\bar{B}(t,T)}\big)$$

and condition (3.9) holds.

Default always yields loss of money, so a sensible choice of the model's recovery should imply $\delta B(t,T) < \bar{B}(t,T)$, so that the promised interest of the defaultable bond $\bar{r}_t(0)$ exceeds than the risk-free interest rate, r_t.

Proof. With the notation of the previous proof, the discounted gains process in this model becomes

$$G(t,T) = D_t B(t,T) = (1 - \Lambda_t) D_t \exp(-\int_0^{T-t} \bar{r}_t(u)\, du) + \Lambda_t \delta D_t B(t,T)$$

with dynamics

$$
\begin{aligned}
dG(t,T) &= (1-\Lambda_t)d(D_t \exp(-\int_0^{T-t} \bar{r}_t(u)\,du)) - D_t \exp(-\int_0^{T-t} \bar{r}_t(u)\,du) d\Lambda_t \\
&\quad + \Lambda_t \delta d(D_t B(t,T)) + \delta D_t B(t,T) d\Lambda_t.
\end{aligned}
$$

Taking into account that on $\{\tau > t\}$, $\exp(-\int_0^{T-t} \bar{r}_t(u)\,du) = \bar{B}(t,T)$, the value of $d(D_t\bar{B}(t,T))$ is given in equation (3.11). This yields

$$
\begin{aligned}
dG(t,T) &= (1-\Lambda_t)D_t\bar{B}(t,T)\Big\{\Big(\bar{r}_t(0) - r_t - \bar{\alpha}^*(t,T) + \frac{1}{2}\sum_{k=1}^{\infty}\bar{\lambda}_k[\sigma_k^*(t,T)]^2\Big)\,dt \\
&\qquad - \sum_{k=1}^{\infty}\bar{\sigma}_k^*(t,T)\,d\bar{\beta}_k(t)\Big\} \\
&+ \Big[-D_t\exp(-\int_0^{T-t}\bar{r}_t(u)\,du) + \delta D_t B(t,T)\Big]\Big[d\tilde{M}_t + (1-\Lambda_t)\lambda_t\,dt\Big] \\
&+ \Lambda_t\delta d(D_t B(t,T)) \\
&= (1-\Lambda_t)D_t\bar{B}(t,T)\Big\{\bar{r}_t(0) - r_t - \lambda_t - \bar{\alpha}^*(t,T) + \frac{1}{2}\sum_{k=1}^{\infty}\bar{\lambda}_k[\sigma_k^*(t,T)]^2\Big\}\,dt \\
&+ \delta D_t B(t,T)\,(1-\Lambda_t)\lambda_t\,dt \\
&+ d\bar{M}_t,
\end{aligned}
$$

where we denote the sum of all martingale terms by $\bar{M}_t$. Note that $D_tB(t,T)$ is a martingale, as we assumed the riskless bond market to be free of arbitrage.

Consequently the drift of $(G(t,T))_{t\geq 0}$ is zero, iff on $\{\tau > t\}$ for all $0 \leq t \leq T$

$$
\begin{aligned}
0 &= \bar{B}(t,T)\Big\{\bar{r}_t(0) - r_t - \lambda_t - \bar{\alpha}^*(t,T) + \frac{1}{2}\sum_{k=1}^{\infty}\bar{\lambda}_k[\sigma_k^*(t,T)]^2\Big\} + \delta B(t,T)\lambda_t \\
\Leftrightarrow \qquad 0 &= \bar{r}_t(0) - r_t - \lambda_t + \delta\lambda_t\frac{B(t,T)}{\bar{B}(t,T)} - \bar{\alpha}^*(t,T) + \frac{1}{2}\sum_{k=1}^{\infty}\bar{\lambda}_k[\sigma_k^*(t,T)]^2.
\end{aligned}
$$

Similar arguments as for Theorem 3.2.1 yield the desired result.

■

3.3 Models Using Ratings

As ratings are readily available and a widely used tool in markets subject to credit risk, a model should be capable of using this information. In this section we lay out the framework for a model in infinite dimensions that incorporates different rating classes. We present two alternative recovery structures with recovery levels dependent on the pre-default rating.

The basic assumption of the next two sections describes the behavior of the defaultable forward rates with respect to the current rating.

Assumption (A2). Assume that there are $K-1$ ratings, where 1 denotes the highest rating and $K-1$ the lowest, while K is associated with default. Denoting by $\mathcal{K} = \{1,\ldots,K-1\}$ the set of possible ratings and putting $\bar{\mathcal{K}} = \mathcal{K} \cup \{K\}$, we assume that the rating i forward rate satisfies for $t \in [0,T^*]$

$$r_t^i = S(t)r_0^i + \int_0^t S(t)\alpha^i(u)\,du + S(t)\Big[\int_0^t \sigma^i(u)\cdot dX^i(u)\Big],$$

where $(X^i(t))_{t\in[0,T^*]}$ is a D^i-Wiener process. Furthermore, $\alpha^i : [0,T^*]\times\Omega \mapsto H$ and $\sigma^i : [0,T^*]\times\Omega \mapsto L(H;H)$ are stochastic processes, which are predictable w.r.t. $(\mathcal{F}_t)_{t\geq 0}$ and satisfy $\mathbb{P}(\int_0^{T^*} \bar{\alpha}^i(s)\,ds < \infty) = 1$ and $|||\bar{\sigma}^i|||_{T^*} < \infty$, for all $i \in \mathcal{K}$.

To exclude arbitrage we furthermore assume that

$$r_t^{K-1}(x) > \cdots > r_t^1(x) > r_t(x) \qquad \forall x \in [0,T^{**}].$$

This corresponds to the fact that higher rated bonds are more expensive than lower rated ones. If this would not be the case the rating of the bond would seem to be wrong. This could happen because of speculative behavior or when the rating is delayed by some other effects and is not modeled here.

The above relation could be stated equivalently by the condition that the inter-rating spreads must be positive, see Acharya, Das and Sundaram (2000).

The process which describes the current rating of the bond, $(C^1(t))_{t\geq 0}$, takes values in $\bar{\mathcal{K}}$ and is assumed to be a Markov process at this state. Intuitively, this means that the "history of ratings" for this particular bond does not influence the price nor default risk of the bond, only the current rating does[5]. We denote by $C^2(t)$ the previous rating before $C^1(t)$. If there were no changes in rating up to time t we set $C^2(t) = C^1(t)$. The default τ occurs at the first time, when the state K is reached, $\tau := \inf\{t \geq 0 : C^1(t) = K\}$.

Denote the conditional infinitesimal generator of C^1 given $\mathcal{G}_t$ under the measure Q by

$$\Lambda_t = \begin{pmatrix} \lambda_{11}(t) & \lambda_{12}(t) & \lambda_{13}(t) & \cdots & \lambda_{1K}(t) \\ \lambda_{21}(t) & \lambda_{22}(t) & \lambda_{23}(t) & \cdots & \lambda_{2K}(t) \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ \lambda_{K-1,1}(t) & \lambda_{K-1,2}(t) & \cdots & \lambda_{K-1,K-1}(t) & \lambda_{K-1,K}(t) \\ 0 & 0 & \cdots & \cdots & 0 \end{pmatrix}.$$

Each $(\lambda_{ij}(t))_{t\geq 0}$ is a $(\mathcal{G}_t)_{t\geq 0}$-adapted process satisfying the condition

$$\lambda_{ii}(t) = -\sum_{i,j\in\bar{\mathcal{K}}, j\neq i} \lambda_{ij}(t), \qquad \text{for all } t \geq 0. \tag{3.13}$$

[5] Note that this assumption enables us to obtain easier formulas, but might not be fulfilled in reality. For example, if the bond has been downgraded, it is empirically observed that further downgradings are more likely than upgradings.

We state the following proposition which is proved, for example, in Bielecki and Rutkowski (2002, Prop. 11.3.1).

Proposition 3.3.1. *For any function $f : \bar{\mathcal{K}} \mapsto I\!R$ the following process is a martingale:*

$$\begin{aligned}\tilde{M}(t) &= f(C^1(t)) - \int_0^t \sum_{j=1}^K \lambda_{C^1(u),j} f(j)\, du \\ &=: f(C^1(t)) - \int_0^t (\Lambda f)(C^1(u))\, du. \end{aligned} \tag{3.14}$$

For the rating transition to the default state, using equation (11.51) of Bielecki and Rutkowski (2002), we immediately conclude

Proposition 3.3.2. *The process $(M^i(t))_{t\geq 0}$ is a martingale for any $i \in \mathcal{K}$:*

$$M^i(t) = 1_{\{C^2(t)=i, C^1(t)=K\}} - \int_0^t \lambda_{iK}(u) 1_{\{C^1(u)=i\}}\, du. \tag{3.15}$$

3.3.1 Rating Based Recovery of Market Value

Assume the rating i recovery rate $(q^i(t))_{t\geq 0}$ to be a nonnegative stochastic process which is predictable w.r.t. $(\mathcal{F}_t)_{t\geq 0}$ for all $i \in \mathcal{K}$. In extension to Section 3.2.1 we model the defaultable bond with rating transitions for all $t \in [0, T^*]$ and $T \in [t, t + T^{**}]$ by

$$\begin{aligned}\bar{B}(t,T) = 1_{\{C_1(t)\neq K\}} \exp(-\int_0^{T-t} r_t^{C^1(t)}(u)\, du) \\ + 1_{\{C^1(t)=K\}} q_\tau^{C^2(t)} \bar{B}(\tau-, T) \exp(\int_\tau^t r_u\, du). \end{aligned} \tag{3.16}$$

We call this recovery modeling *rating based recovery of market value.* This may be compared to the case without ratings in Section 3.2.1. The advantages of the recovery of market value model carry through to this model.

At this point we can compute the defaultable forward rate, the forward rate offered by the bond $\bar{B}(t,T)$. Setting $x := T - t$ we obtain

$$\begin{aligned}\bar{r}_t(x) &= -\frac{\partial}{\partial T} \ln \bar{B}(t,T) = \frac{-1}{\bar{B}(t,T)} \cdot \frac{\partial}{\partial T} \bar{B}(t,T) \\ &= \frac{-1}{\bar{B}(t,T)} \frac{\partial}{\partial T} \Big[1_{\{C^1(t)\neq K\}} \exp(-\int_0^{T-t} r_t^{C^1(t)}(u)\, du) \\ &\qquad + 1_{\{C^1(t)=K\}} q_\tau^{C^2(t)} \bar{B}(\tau-, T) \exp(\int_\tau^t r_u\, du) \Big]. \end{aligned}$$

Computing the derivative yields

$$\begin{aligned}
\bar{r}_t(x) &= \frac{1}{\bar{B}(t,T)}\Big[1_{\{C^1(t)\neq K\}}\exp(-\int_0^{T-t} r_t^{C^1(t)}(u)\,du)\cdot r_t^{C^1(t)}(T-t) \\
&\qquad +1_{\{C^1(t)=K\}} q_\tau^{C^2(t)}\exp(\int_\tau^t r_u\,du)\exp(-\int_0^{T-\tau} r_\tau^{C^2(t)}(u)\,du)\cdot r_\tau^{C^2(t)}(T-\tau)\Big] \\
&= 1_{\{C^1(t)\neq K\}} r_t^{C^1(t)}(x) + 1_{\{C^1(t)=K\}} r_\tau^{C^2(t)}(x+t).
\end{aligned}$$

Interestingly, this expression does not depend on the different recovery rates, which is due to the fact that the forward rates describe the behavior of *relative* price changes. So the defaultable forward rate equals the forward rate with respect to the bond's rating. If the bond defaulted, the forward rate curve remains static, as there is no further movement except the risk-free interest.

Denote

$$B^i(t,T) = \exp(-\int_0^{T-t} r_t^i(u)\,du).$$

Theorem 3.3.3. *Assume that (A2) and (3.16) hold under the measure Q. Then discounted defaultable bond prices are martingales under Q iff the following two conditions are satisfied on $\{\tau > t\}$:*

(i) For $t \in [0,T^*]$, $T \in [t, t+T^{**}]$,

$$\begin{aligned}
r^{C^1(t)}(0) &= r_t + (1-q_t^{C^1(t)})\lambda_{C^1(t),K}(t) \\
&+ \sum_{j=1, j\neq C^1(t)}^{K-1}\Big[1-\frac{B^j(t,T)}{B^{C^1(t)}(t,T)}\Big]\lambda_{C^1(t),j}(t).
\end{aligned} \tag{3.17}$$

(ii) For $t \in [0,T^*]$, $T \in [t, t+T^{**}]$,

$$\alpha^{C^1(t)}(t,T) = \sum_{k=1}^{\infty}\lambda_k^{C^1(t)}\,\sigma_k^{C^1(t)*}(t,T)\cdot\sigma_k^{C^1(t)}(t,T). \tag{3.18}$$

Under the conditions of the above theorem and, if Q is equivalent to the objective measure P, Q is an equivalent martingale measure and so the market is free of arbitrage.

Proof. Using equation (3.16), we determine the discounted gains process

$$\begin{aligned}
G(t,T) &= D_t\sum_{i=1}^{K-1} 1_{\{C^1(t)=i\}} B^i(t,T) \\
&+ D_\tau\, 1_{\{C^1(t)=K\}}\sum_{i=1}^{K-1} 1_{\{C^2(t)=i\}} q_\tau^i \bar{B}(\tau-,T).
\end{aligned}$$

Note that the indicators have finite variation, just like $(D_t)_{t\geq 0}$, and therefore Itô's formula yields the dynamics

$$\begin{aligned} dG(t,T) &= \sum_{i=1}^{K-1} 1_{\{C^1(t)=i\}} d(D_t B^i(t,T)) \\ &+ \sum_{i=1}^{K-1} D_t B^i(t,T) d1_{\{C^1(t)=i\}} \\ &+ d\Big(\sum_{i=1}^{K-1} 1_{\{C^1(t)=K, C^2(t)=i\}}\Big) q^i_\tau \bar{B}(\tau-,T) D_\tau. \end{aligned}$$

For the last term,

$$q^i_\tau \bar{B}(\tau-,T) D_\tau d1_{\{C^1(t)=K,C^2(t)=i\}} = q^i_t\, B^i(t,T) D_t d1_{\{C^1(t)=K,C^2(t)=i\}},$$

as the indicator changes only at $t=\tau$. Furthermore, because of continuity of the forward rates, $\bar{B}(\tau-,T) = B^{C^2(\tau)}(\tau,T)$.

Using (3.14) with $f^i(x) = 1_{\{x=i\}}$ for $i \in \mathcal{K}$, we have

$$\begin{aligned} d1_{\{C^1(t)=i\}} &= d(\tilde{M}^i(t) + \int_0^t \sum_{j=1}^{\mathcal{K}} \lambda_{C^1(u),j} f^i(j)\, du) \\ &= d(\tilde{M}^i(t) + \int_0^t \lambda_{C^1(u),i}\, du) \\ &= d\tilde{M}^i(t) + \lambda_{C^1(t),i}\, dt. \end{aligned} \tag{3.19}$$

Analogously to the default-free case (see 3.11) the dynamics of each i-rated bond for $t \in [0,T^*]$ and $T \in [t, t+T^{**}]$ can be expressed as[6]

$$\begin{aligned} d(D_t B^i(t,T)) &= D_t B^i(t,T) \Big[\Big(r^i_t(0) - r_t - \alpha^{i*}(t,T) + \frac{1}{2}\sum_{k=1}^{\infty} \lambda^i_k [\sigma^{i*}_k(t,T)]^2\Big)\, dt \\ &\quad - \sum_{k=1}^{\infty} \sigma^{i*}_k(t,T)\, d\beta^i_k(t)\Big]. \end{aligned} \tag{3.20}$$

[6]As before, we use the abbreviations

$$\alpha^{i*}(t,T) = \int_t^T \alpha^i(t,u)\, du \qquad \sigma^{i*}_k(t,T) = \int_t^T \big[\sigma^{i*}(t) \cdot e_k\big](u)\, du.$$

Use (3.15) to obtain

$$\begin{aligned}
dG(t,T) &= \sum_{i=1}^{K-1} 1_{\{C^1(t)=i\}} D_t B^i(t,T)\Big\{\Big(r_t^i(0) - r_t - \alpha^{i*}(t,T) + \frac{1}{2}\sum_{k=1}^{\infty}\lambda_k^i[\sigma_k^{i*}(t,T)]^2\Big)\,dt \\
&\qquad - \sum_{k=1}^{\infty}\sigma_k^{i*}(t,T)\,d\beta_k^i(t)\Big\} \\
&+ \sum_{i=1}^{K-1} D_t B^i(t,T)[d\tilde{M}^i(t) + \lambda_{C^1(t),i}(t)\,dt] \\
&+ \sum_{i=1}^{K-1} q_t^i\, B^i(t,T) D_t [dM^i(t) + \lambda_{i,K}(t) 1_{\{C^1(t)=i\}}\,dt] \\
&= \sum_{i=1}^{K-1} D_t B^i(t,T)\Big\{1_{\{C^1(t)=i\}}\Big(r_t^i(0) - r_t + q_t^i\,\lambda_{i,K}(t) \\
&\qquad -\alpha^{i*}(t,T) + \frac{1}{2}\sum_{k=1}^{\infty}\lambda_k^i[\sigma_k^{i*}(t,T)]^2\Big) + \lambda_{C^1(t),i}(t)\Big\}\,dt \\
&\quad + d\bar{M}_t,
\end{aligned}$$

where we denoted the sum of the martingale parts by $\bar{M}_t$. The dt-term yields the drift, and $G(t,T)$ is a martingale, iff the drift is zero. We split the drift into two parts. The first part consists of

$$1_{\{C^1(t)=i\}}\Big[-\alpha^{i*}(t,T) + \frac{1}{2}\sum_{k=1}^{\infty}\lambda_k^i[\sigma_k^{i*}(t,T)]^2\Big], \qquad i \in \mathcal{K},$$

which is equal to zero (see equation (3.12)), iff on $\{C^1(t) = i\}$

$$\alpha^i(t,T) = \sum_{k=1}^{\infty}\lambda_k^i\,\sigma_k^{i*}(t,T)\cdot\sigma_k^i(t,T).$$

Hence, condition (3.18) follows.

The second part yields

$$\begin{aligned}
0 &= 1_{\{C^1(t)\neq K\}}\Big\{B^{C^1(t)}(t,T)\big[r_t^{C^1(t)}(0) - r_t + q_t^{C^1(t)}\lambda_{C^1(t),K}(t)\big] \\
&\qquad + \sum_{j=1}^{K-1} B^j(t,T)\lambda_{C^1(t),j}(t)\Big\} \\
\Leftrightarrow \quad 0 &= r_t^{C^1(t)}(0) - r_t + q_t^{C^1(t)}\lambda_{C^1(t),K}(t) \\
&\quad + \sum_{j=1}^{K-1}\frac{B^j(t,T)}{B^{C^1(t)}(t,T)}\lambda_{C^1(t),j}(t), \qquad \text{on } \{C^1(t)\neq K\}. \qquad (3.21)
\end{aligned}$$

Using equation (3.13) leads to

$$
\begin{aligned}
&\sum_{j=1}^{K-1} \frac{B^j(t,T)}{B^{C^1(t)}(t,T)} \lambda_{C^1(t),j}(t) = \\
&\qquad = \sum_{j=1, j\neq C^1(t)}^{K-1} \frac{B^j(t,T)}{B^{C^1(t)}(t,T)} \lambda_{C^1(t),j}(t) + \lambda_{C^1(t),C^1(t)}(t) \\
&\qquad = \sum_{j=1, j\neq C^1(t)}^{K-1} \frac{B^j(t,T)}{B^{C^1(t)}(t,T)} \lambda_{C^1(t),j}(t) - \sum_{j=1, j\neq C^1(t)}^{K} \lambda_{C^1(t),j}(t) \\
&\qquad = \sum_{j=1, j\neq C^1(t)}^{K-1} \Big[\frac{B^j(t,T)}{B^{C^1(t)}(t,T)} - 1\Big] \lambda_{C^1(t),j}(t) - \lambda_{C^1(t),K}(t).
\end{aligned}
$$

Finally we obtain, on $\{C^1(t) \neq K\}$,

$$
\begin{aligned}
(3.21) \quad \Leftrightarrow \quad r_t^{C^1(t)}(0) &= r_t + (1 - q_t^{C^1(t)}) \lambda_{C^1(t),K}(t) \\
&+ \sum_{j=1, j\neq C^1(t)}^{K-1} \Big[1 - \frac{B^j(t,T)}{B^{C^1(t)}(t,T)}\Big] \lambda_{C^1(t),j}(t).
\end{aligned}
$$

■

Remark 3.3.4. Again, if one prefers a drift condition not depending on a particular realization of $(C^1(t))_{t\geq 0}$, equivalency in Theorem 3.3.3 cannot be obtained, see Remark 3.2.2. In this case we require the above equations to be satisfied for any $i \in \mathcal{K}$, which leads to the following conditions:

(i) For $t \in [0, T^*]$, $T \in [t, t+T^{**}]$ and $i \in \mathcal{K}$

$$
r^i(0) = r_t + (1 - q_t^i)\lambda_{i,K}(t) + \sum_{j=1, j\neq i}^{K-1} \Big[1 - \frac{B^j(t,T)}{B^i(t,T)}\Big] \lambda_{i,j}(t). \tag{3.22}
$$

(ii) For $t \in [0, T^*]$, $T \in [t, t+T^{**}]$ and $i \in \mathcal{K}$,

$$
\alpha^i(t,T) = \sum_{k=1}^{\infty} \lambda_k^i \, \sigma_k^{i*}(t,T) \cdot \sigma_k^i(t,T). \tag{3.23}
$$

3.3.2 Rating Based Recovery of Treasury

Another way to model recovery is based on the recovery of treasury model developed by Bielecki and Rutkowski (2000). We adapt their framework but extend their model by considering infinite dimensional Wiener processes.

With the notations of the previous section, the defaultable bond with rating transitions is modeled for $t \in [0, T^*]$ and $T \in [t, t + T^{**}]$ by

$$\bar{B}(t,T) = 1_{\{C^1(t) \neq K\}} \exp\Big(-\int_0^{T-t} r_t^{C^1(t)}(u)\, du\Big) + 1_{\{C^1(t)=K\}} \delta_{C^2(t)} B(t,T). \tag{3.24}$$

The rating i-recovery rate δ_i is assumed to be constant. This recovery modeling is referred to as *rating based recovery of treasury.*

Computing the defaultable forward rate in this model yields

$$\bar{r}_t(x) \;=\; 1_{\{C^1(t) \neq K\}} r_t^{C^1(t)}(x) + 1_{\{C^1(t)=K\}} r_t(x).$$

This is similar to the rating based recovery of market value setting, and, of course, differences appear just for the behavior after default. In this model the defaultable forward rate after default equals the default-free rate. Anyway, some part of the invested money is lost.

Theorem 3.3.5. *Assume that (A2) and (3.24) holds under the measure Q. Then discounted defaultable bond prices are martingales under Q, iff for $t \in [0, T^*]$, $T \in [t, t+T^{**}]$ on $\{\tau > t\}$*

$$\begin{aligned} r_t^{C^1(t)}(0) \;=\; & r_t + \sum_{j=1, j \neq C^1(t)}^{K-1} \lambda_{C^1(t),j}(t)\Big(1 - \frac{B^j(t,T)}{B^{C^1(t)}(t,T)}\Big) \\ & + \lambda_{C^1(t),K}(t)\Big(1 - \delta_{C^1(t)} \frac{B(t,T)}{B^{C^1(t)}(t,T)}\Big) \end{aligned} \tag{3.25}$$

and condition (3.18) holds.

At this point suitable parameters should ensure that the sum in equation (3.25) is positive. The last term is positive for $\delta B(t,T) < \bar{B}(t,T)$, as already noted in the case of recovery of treasury without ratings.

Proof. Using the notation of Theorem 3.3.3, the dynamics of $\bar{B}(t,T)$ becomes

$$\begin{aligned} d\bar{B}(t,T) \;=\; & \sum_{i=1}^{K-1} \Big[1_{\{C^1(t)=i\}} dB^i(t,T) + B^i(t,T)\, d1_{\{C^1(t)=i\}} \Big] \\ & + \sum_{i=1}^{K-1} \Big[1_{\{C^1(t)=K, C^2(t)=i\}} \delta_i\, dB(t,T) + B(t,T) \delta_i\, d1_{\{C^1(t)=K, C^2(t)=i\}} \Big]. \end{aligned}$$

Note that the differentials of the indicators are -1 or 1 when a jump occurs and zero otherwise.

Using (3.19), we have

$$\sum_{i=1}^{K-1} B^i(t,T) d1_{\{C^1(t)=i\}} = \sum_{i=1}^{K-1} B^i(t,T)(d\tilde{M}^i(t) + \lambda_{C^1(t),i}\, dt).$$

Furthermore, use (3.15) and (3.20) to obtain

$$\begin{aligned} d\bar{B}(t,T) &= \sum_{i=1}^{K-1} 1_{\{C^1(t)=i\}} B^i(t,T)\Big[\Big(r_t^i(0) - \alpha^{i*}(t,T) + \frac{1}{2}\sum_{k=1}^{\infty} \lambda_k^i [\sigma_k^{i*}(t,T)]^2\Big)\, dt \\ &\qquad\qquad - \sum_{k=1}^{\infty} \sigma_k^{i*}(t,T)\, d\beta_k^i(t)\Big] \\ &+ \sum_{i=1}^{K-1} B^i(t,T)\Big[\lambda_{C^1(t),i}(t)dt + d\tilde{M}_t^i\Big] \\ &+ \sum_{i=1}^{K-1} 1_{\{C^1(t)=K, C^2(t)=i\}} \delta_i\, dB(t,T) \\ &+ \sum_{i=1}^{K-1} \delta_i B(t,T)\, \Big[\lambda_{i,K}(t) 1_{\{C^1(t)=i\}}\, dt + dM_t^i\Big]. \end{aligned}$$

Separating the drift and martingale parts, this leads to

$$\begin{aligned} d\bar{B}(t,T) &= \sum_{i=1}^{K-1} 1_{\{C^1(t)=i\}} B^i(t,T)\Big[r_t^i(0) - \alpha^{i*}(t,T) + \frac{1}{2}\sum_{k=1}^{\infty} \lambda_k^i [\sigma_k^{i*}(t,T)]^2\Big]\, dt \\ &+ \sum_{i=1}^{K-1} B^i(t,T)\lambda_{C^1(t),i}(t)dt + 1_{\{C^1(t)=K, C^2(t)=i\}} \delta_i\, dB(t,T) \\ &+ \sum_{i=1}^{K-1} \delta_i B(t,T) \lambda_{i,K}(t) 1_{\{C^1(t)=i\}}\, dt \\ &- \sum_{i=1}^{K-1} 1_{\{C^1(t)=i\}} B^i(t,T) \sum_{k=1}^{\infty} \sigma_k^{i*}(t,T)\, d\beta_k^i(t) \\ &+ \sum_{i=1}^{K-1} B^i(t,T) d\tilde{M}^i(t) + \delta_i B(t,T)\, dM^i(t). \end{aligned}$$

If we denote the discounting factor by D_t the discounted bond price equals

$$
\begin{aligned}
d(D_t\bar{B}(t,T)) &= (-r_t)D_t\bar{B}(t,T)\,dt + D_t\,d\bar{B}(t,T) \\
&= -r_tD_t\Big[\sum_{i=1}^{K-1} 1_{\{C^1(t)=i\}}B^i(t,T) + \sum_{i=1}^{K-1} 1_{\{C^1(t)=K,C^2(t)=i\}}\delta_i B(t,T)\Big]\,dt \\
&+ D_t\Big\{\sum_{i=1}^{K-1} 1_{\{C^1(t)=i\}}B^i(t,T)\Big[r_t^i(0) - \alpha^{i*}(t,T) + \frac{1}{2}\sum_{k=1}^{\infty}\lambda_k^i[\sigma_k^{i*}(t,T)]^2\Big]\,dt \\
&+ \sum_{i=1}^{K-1} B^i(t,T)\lambda_{C^1(t),i}(t)dt + 1_{\{C^1(t)=K,C^2(t)=i\}}\delta_i\,dB(t,T) \\
&+ \sum_{i=1}^{K-1}\delta_i B(t,T)\lambda_{i,K}(t)1_{\{C^1(t)=i\}}\,dt\Big\} \\
&+ d\tilde{\tilde{M}}_t,
\end{aligned}
$$

where we added the martingale parts up to $d\tilde{\tilde{M}}_t$. As the discounted risk-free bond is a martingale by assumption, we conclude that

$$d(D_tB(t,T)) = -r_tD_tB(t,T)\,dt + D_tdB(t,T)$$

is a martingale and so the $1_{\{C^1(t)=K,C^2(t)=i\}}$-terms sum up to a martingale. We have

$$
\begin{aligned}
&d(D_t\bar{B}(t,T)) \\
&= D_t\Big\{\sum_{i=1}^{K-1} 1_{\{C^1(t)=i\}}B^i(t,T)\Big[-r_t + r_t^i(0) - \alpha^{i*}(t,T) + \frac{1}{2}\sum_{k=1}^{\infty}\lambda_k^i[\sigma_k^{i*}(t,T)]^2\Big]\,dt \\
&\qquad + \sum_{i=1}^{K-1} B^i(t,T)\lambda_{C^1(t),i}(t)dt + \sum_{i=1}^{K-1}\delta_i B(t,T)\lambda_{i,K}(t)1_{\{C^1(t)=i\}}\,dt\Big\} \\
&+ d\bar{M}_t,
\end{aligned}
$$

denoting the martingale part by $\bar{M}_t$.

Therefore, $D_t\bar{B}(t,T)$ is a martingale, iff on $\{C^1(t) \neq K\}$

$$
\begin{aligned}
0 &= B^{C^1(t)}(t,T)\Big[-r_t + r_t^{C^1(t)}(0) - \alpha^{C^1(t)*}(t,T) + \frac{1}{2}\sum_{k=1}^{\infty}\lambda_k^{C^1(t)}[\sigma_k^{C^1(t)*}(t,T)]^2\Big] \\
&+ \sum_{j=1}^{K-1} B^j(t,T)\lambda_{C^1(t),j}(t) + \delta_{C^1(t)}B(t,T)\lambda_{C^1(t),K}.
\end{aligned}
$$

Further on we consider the drift on $\{\tau > t\}$. This leads to

$$
\begin{aligned}
0 &= B^{C^1(t)}(t,T)\Big[r_t^{C^1(t)}(0) - r_t - \alpha^{C^1(t)*}(t,T) + \frac{1}{2}\sum_{k=1}^{\infty}\lambda_k^{C^1(t)}[\sigma_k^{C^1(t)*}(t,T)]^2\Big] \\
&+ \delta_{C^1(t)}B(t,T)\lambda_{C^1(t),K} + \sum_{j=1}^{K-1} B^j(t,T)\lambda_{C^1(t),j}(t). \qquad (3.26)
\end{aligned}
$$

Again, we split the above condition in two parts. The first part consists of

$$-\alpha^{C^1(t)*}(t,T) + \frac{1}{2}\sum_{k=1}^{\infty} \lambda_k^{C^1(t)} [\sigma_k^{C^1(t)*}(t,T)]^2,$$

which is equal to zero (see equation (3.12)), iff

$$\alpha^{C^1(t)}(t,T) = \sum_{k=1}^{\infty} \lambda_k^{C^1(t)} \, \sigma_k^{C^1(t)*}(t,T) \cdot \sigma_k^{C^1(t)}(t,T).$$

We consider the second part on $\{C^1(t) = i\}$ with $i \in \mathcal{K}$. This yields

$$0 \;=\; r_t^i(0) - r_t + \delta_i \frac{B(t,T)}{B^i(t,T)} \lambda_{i,K} + \sum_{j=1}^{K-1} \frac{B^j(t,T)}{B^i(t,T)} \lambda_{i,j}(t).$$

Using equation (3.13) and denoting

$$\tilde{q}_{i,j}(t,T) = \begin{cases} B^j(t,T)/B^i(t,T) & j \neq K \\ \delta_i B(t,T)/B^i(t,T) & j = K \end{cases}$$

leads to

$$\begin{aligned} r_t^i(0) &= r_t - \sum_{j=1}^{K} \tilde{q}_{i,j}(t,T)\lambda_{i,j}(t) \\ &= r_t - \sum_{j=1, j\neq i}^{K} \tilde{q}_{i,j}(t,T)\lambda_{i,j}(t) - \tilde{q}_{ii}(t,T)\big(- \sum_{j=0, j\neq i}^{K} \lambda_{i,j}(t)\big) \\ &= r_t + \sum_{j=1, j\neq i}^{K} \lambda_{i,j}(t)\big(1 - \tilde{q}_{i,j}(t,T)\big), \end{aligned}$$

and we conclude (3.25). ■

Remark 3.3.6. Similar to Remark 3.2.2, we obtain the following condition, which does not depend on C^1 but also implies an arbitrage-free market together with (3.23):

$$\begin{aligned} r_t^i(0) \;=\; & r_t + \sum_{j=1, j\neq i}^{K-1} \lambda_{i,j}(t)\big(1 - \frac{B^j(t,T)}{B^i(t,T)}\big) \\ & + \lambda_{i,K}(t)\big(1 - \delta_i \frac{B(t,T)}{B^i(t,T)}\big), \qquad i \in \bar{\mathcal{K}}. \end{aligned} \tag{3.27}$$

It seems natural that condition (3.10) extends to the rating model. Equation (3.27) represents the relationship under no-arbitrage between the interest offered by a bond rated i, the likelihood of rating changes with their consequences to the bond's price, as well as with default and recovery.

An equivalent but more concise version of (3.26) is obtained on $\{C^1(t)=i\}$ by setting

$$a^i(t,T) := -r_t + r_t^i(0) - \alpha^{i\,*}(t,T) + \frac{1}{2}\sum_{k=1}^{\infty}\lambda_k^i[\sigma_k^{i\,*}(t,T)]^2.$$

Recall that $i \in \mathcal{K}$. Substituting $\lambda_{ii}(t) = -\sum_{j=1,j\neq i}^{K}\lambda_{ij}(t)$, we obtain

$$\begin{aligned} 0 &= B^i(t,T)a^i(t,T) + \sum_{j=1}^{K-1} B^j(t,T)\lambda_{i,j}(t) + \delta_i B(t,T)\lambda_{i,K}(t) \\ &= B^i(t,T)a^i(t,T) + \delta_i B(t,T)\lambda_{i,K}(t) + \sum_{j=1,j\neq i}^{K-1} B^j(t,T)\lambda_{i,j}(t) - B^i(t,T)\sum_{j=1,j\neq i}^{K}\lambda_{i,j}(t) \end{aligned}$$

and hence obtain the equivalent representation of (3.26),

$$0 = B^i(t,T)a^i(t,T) + \sum_{j=1,j\neq i}^{K-1}\big(B^j(t,T) - B^i(t,T)\big)\lambda_{i,j}(t) + \big(\delta_i B(t,T) - B^i(t,T)\big)\lambda_{i,K}(t).$$

The first part of this expression relates to the drift of the bond itself, while the other parts refer to the possible changes into a different rating class. A change of the rating immediately entails a change of the bond's price. These are multiplied with the rate, that such a change may happen. See also Proposition 3.3.1.

Note that

$$D_T B(T,T) = \exp(-\int_0^T r_u\,du)\Big[1_{\{\tau>T\}} + \delta_{C^2(T)}1_{\{\tau\leq T\}}\Big].$$

As we have shown that the discounted bond price is a martingale this leads to

$$\begin{aligned} \bar{B}(t,T) &= \frac{1}{D_t}\mathbb{E}^Q\Big[D_T\bar{B}(t,T)\Big|\mathcal{F}_t\Big] \\ &= \mathbb{E}^Q\Big[\exp(-\int_t^T r_u\,du)\Big(1_{\{\tau>T\}} + \delta_{C^2(T)}1_{\{\tau\leq T\}}\Big)\Big|\,\mathcal{F}_t\Big]. \end{aligned}$$

This is often stated as "the bond price equals the conditional expectation of the discounted payoff", which proves to be true in our setting as well.

3.4 Pricing

Pricing in credit risky models is usually done via computation of the expectation of the discounted contingent claim, see for example Lando (1994), Duffie and Singleton (1999) or Bielecki and Rutkowski (2002). We present a series of examples where we are able to obtain closed form solutions in Section 4.4.

Starting with a model for the default-free and defaultable forward rate, one uses equation (3.10) with a specific assumption on the recovery structure to obtain a model for the default intensity λ_t. Depending on the model it still might be difficult to evaluate the expectation. There are two main possibilities, either to simplify the model to compute the expectation explicitly or otherwise to use Monte-Carlo methods.

Chapter 4

Credit Risk Modeling with Gaussian Random Fields

In this chapter we use Gaussian random fields to model interest rates and credit risk. After introducing the basic terminology for Gaussian random fields, we present the interest rate model of Kennedy (1994), and at the same time simplify some proofs. Section 4.3 extends this model to also incorporate credit risk in different recovery situations, while Section 4.4 presents several explicit pricing formulas and a hedging scheme of an option on a defaultable bond under zero recovery. In Section 4.5 we discuss two calibration methodologies.

The use of random fields in credit risk modeling seems new, and the approach using Gaussian random fields has the advantage of admitting explicit pricing and hedging formulas. As shown in Pang (1998), the explicit formulation and flexibility of Gaussian random fields proves to be advantageous for calibration issues in the interest rate context.

4.1 Preliminaries

A *random field* is a stochastic process indexed by vectors. As in the univariate case, its distribution is uniquely determined by its finite dimensional distributions (fidis). It is called *Gaussian*, if all fidis are Gaussian. For our purpose we only consider random fields in $[0,T^*] \times [0,T^*]$. More general versions can be found in Adler (1981). A detailed treatment of Gaussian measures on Banach spaces may be found in Bogachev (1991).

Definition 4.1.1. A stochastic process $(X(s,t))_{s,t\in[0,T^*]}$ is called a **Gaussian random field**, if for all $(s_i,t_i), i=1,\dots,n$ and $n \geq 1$ the vector

$$(X(s_1,t_1),\dots,X(s_n,t_n))^\top$$

admits a Gaussian law in $\mathbb{R}^n$.

A Gaussian random field may be fully described by its expectation and covariance func-

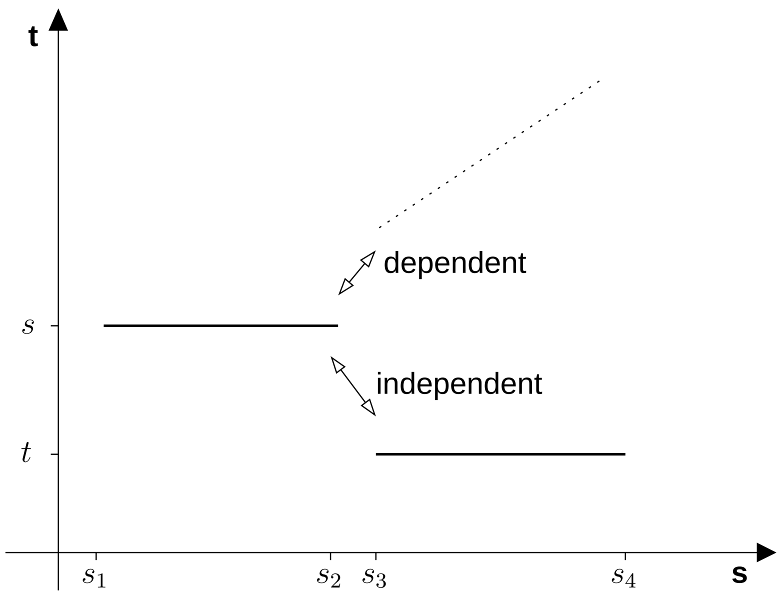

Figure 4.1: The consequences of equation (4.1): Increments like $X(s_4, t) - X(s_3, t)$ and $X(s_2, s) - X(s_1, s)$ are assumed to be independent, while $X(s_4, u) - X(s_3, t)$ and $X(s_2, s) - X(s_1, s)$ may still be dependent (for $u \neq t$).

tion[1]

$$\begin{aligned} \mu(s,t) &:= \mathbb{E}[X(s,t)] \\ c(s,t,u,v) &:= \mathbb{E}\big[(X(s,t)-\mu(s,t))\,(X(u,v)-\mu(u,v))\big]. \end{aligned}$$

Conditions on the covariance function imply a certain smoothness of a random field, for example continuity.

Lemma 4.1.2. *A Gaussian random field $(X(s,t))_{s,t\in[0,1]}$ with zero mean and continuous covariance function has a.s. continuous sample functions, if there exist $0 < C < \infty$ and $\varepsilon > 0$ such that for all $s_1, s_2, t_1, t_2 \in [0,1]$*

$$\mathbb{E}|X(s_1,s_2) - X(t_1,t_2)|^2 \le \frac{C}{|\log \parallel s-t \parallel |^{1+\varepsilon}}.$$

For a proof see Adler (1981, p. 60).

Remark 4.1.3. Assume, for example, that $(X(s,t))_{s,t\in[0,T^*]}$ has zero mean and the covariance function

$$\mathbb{E}(X(s_1,t_1)\,X(s_2,t_2)) = c(s_1 \wedge s_2, t_1, t_2), \tag{4.1}$$

where $c(s,t_1,t_2)$ is a deterministic function. Then $(X(s,t))_{s,t\in[0,T^*]}$ has independent increments in the s-direction, as for $s_4 > s_3 \ge s_2 > s_1$, $s, t \in [0, T^*]$, we have

$$\begin{aligned} &\mathbb{E}\big[(X(s_2,s)-X(s_1,s))(X(s_4,t)-X(s_3,t))\big] \\ &\qquad = c(s_2,s,t) - c(s_2,s,t) + c(s_1,s,t) - c(s_1,s,t) = 0. \end{aligned}$$

[1]Mean and covariance function are sufficient to describe the Gaussian random field, see Bogachev (1991, p. 52 p.p.).

Coming back to the setting in Hilbert spaces, we may compute the covariance operator $D(s_1) := \mathrm{Cov}(X(s_1))$ in many cases. Consider the case where $H = L^2(\lambda)$, the space of square-integrable functions w.r.t. a measure λ. Then Fubini's theorem allows to interchange expectation and the inner product, and the definition of D yields

$$\begin{aligned}
\mathbb{E}(< X(s_1), f >< X(s_1), g >) &= \mathbb{E} \int X(s_1, x) f(x) \lambda(dx) \int X(s_1, y) g(y) \lambda(dy) \\
&= \iint \mathbb{E}(X(s_1, x) X(s_1, y)) f(x) g(y) \lambda(dx) \lambda(dy) \\
&= \int [\int c(s_1, x, y) f(x) \lambda(dx)] g(y) \lambda(dy) \\
&\overset{!}{=} < Df, g > .
\end{aligned}$$

Therefore, $\mathrm{Cov}(X(s))$ is the linear mapping $D(s)$ with

$$D(s) : H \mapsto H, f \to \int c(s, t_1, \cdot) f(t_1) \lambda(dt_1), \tag{4.2}$$

if the integral exists. Note that $x \mapsto \int c(s_1, t_1, x) f(t_1) \lambda(dt_1)$ is a function itself, but $D(s)$ is not necessarily a trace-class operator, cf. Da Prato and Zabczyk (1992, Section 4.3.3).

4.2 A Model without Credit Risk

Before considering bonds with default risk, we present the framework without credit risk: the interest rate case. This section follows the approach of Kennedy (1994), while simplifying some proofs. The forward rates are modeled by a Gaussian random field and we obtain a drift condition under which the model is arbitrage-free. In this section we always consider a finite time horizon T^* and a maximum time-to-maturity T^{**}, so that an overall time horizon $\tilde{T} := T^* + T^{**}$ seems appropriate. The considered market therefore consists of bonds $B(t, T)$ where $t \in [0, T^*]$ and $T \in [t, t + T^{**}]$.

Basic to this section is the following

Assumption (B1): Let $(X(s,t))_{s,t \in [0,\tilde{T}]}$ be a continuous, zero-mean Gaussian random field whose covariance function can be represented by a function $c : \mathbb{R}^3 \mapsto \mathbb{R}$ such that

$$\mathrm{Cov}\left(X_{s_1,t_1}, X_{s_2,t_2}\right) = c(s_1 \wedge s_2, t_1, t_2).$$

Also, $c(0, t_1, t_2) = 0$ (which refers to a deterministic initial term structure). Note that c is symmetric in the sense that $c(\cdot, t_1, t_2) = c(\cdot, t_2, t_1)$.

The information available at time t is described by the σ-algebra

$$\mathcal{F}_t = \sigma\big(X_{u,v} : 0 \leq u \leq t, v \in [u, u + T^{**}]\big).$$

This reveals a basic fact for forward rates, namely, as to $f(t,T)$, the two indices t and T are treated differently. The index t represents the calendar time, while T denotes maturity. For a certain time t, the whole interest rate curve is known, that is, all $\{f(t,T) : T \in [t, t+T^{**}]\}$ are assumed to be observable in the market at time t.

Usually, this information is only available for a discrete tenor structure $T_1, \dots, T_n$, which is a basic motivation to consider market models. On the other hand, one can either interpolate them, using splines or some parametric families, which is discussed in Filipović (2001), or view the discrete observations as a partial information of the whole, but unknown term structure. We take this last viewpoint and, nevertheless, model the whole term structure. Later on, in the calibration process, we account for the discrete observations by an approximation argument.

Take $\mu(s,t) : \mathbb{R}^2 \mapsto \mathbb{R}$ to be a continuous function. The T-forward rate at time t is then modeled through

$$f(t,T) = \mu(t,T) + X(t,T). \tag{4.3}$$

This also specifies the dynamic of the bonds, since

$$B(t,T) = \exp(-\int_t^T f(t,u)\,du).$$

Remark 4.2.1. The Gaussian HJM model is a special case of this. For a deterministic drift $\mu(t,T)$ and volatility $\sigma(t,T)$ and

$$df(t,T) = \mu(t,T)\,dt + \sigma(t,T)\,dW_t,$$

the covariance function of the forward rates becomes

$$\begin{aligned}
\mathrm{Cov}(f(s_1,t_1), f(s_2,t_2)) &= \mathbb{E}\Big[\int_0^{s_1} \sigma(u,t_1)\,dW_u \int_0^{s_2} \sigma(v,t_2)\,dW_v\Big] \\
&= \int_0^{s_1 \wedge s_2} \sigma(u,t_1)\sigma(u,t_2)\,du \\
&\equiv c(s_1 \wedge s_2, t_1, t_2).
\end{aligned}$$

Remark 4.2.2. The model of Hull and White (1990) can be formulated in this framework. The spot rate then satisfies

$$dr_t = [\phi_t - a_t r_t]dt + \sigma_t dW_t.$$

Hull and White (1990) showed that the bond price can be expressed in an exponential-affine form:

$$B(t,T) = \alpha(t,T)\exp(-\beta(t,T)r_t),$$

where α is a deterministic function, and

$$\beta(s,t) = \frac{\beta(0,t) - \beta(0,s)}{\partial\beta(0,s)/\partial s}.$$

It was shown by Schmidt (1997), using time transformations of Brownian motions, that for appropriate deterministic functions f, g, τ the distribution of r has the following form:

$$r(t) \stackrel{\mathcal{L}}{=} f(t) + g(t)W(\tau(t)),$$

where

$$\tau(t) = \int_0^t \Big[\frac{\sigma(u)}{\partial\beta(0,u)/\partial u}\Big]^2 du.$$

As $f(t,T) = -\partial/\partial T \ln B(t,T)$, we obtain

$$\mathrm{Cov}(f(s_1,t_1), f(s_2,t_2)) \;=\; \frac{\partial\beta(s_1,t_1)}{\partial t_1}\frac{\partial\beta(s_2,t_2)}{\partial t_2}\,\mathrm{Cov}(r_{s_1}, r_{s_2}).$$

Hence, as τ is increasing,

$$\mathrm{Cov}(r_{s_1}, r_{s_2}) = g(s_1)g(s_2)\tau(s_1 \wedge s_2).$$

Inserting the definition of β and τ leads to

$$\mathrm{Cov}(f(s_1,t_1), f(s_2,t_2)) = \frac{\partial\beta(s_1,t_1)}{\partial t_1}\frac{\partial\beta(s_2,t_2)}{\partial t_2}\int_0^{s_1\wedge s_2}\Big[\frac{\sigma_u}{\partial\beta(0,u)/\partial u}\Big]^2 du.$$

In this chapter we always consider the objective measure P and a measure Q which is equivalent to P. Assume (B1) to hold under Q. Girsanov's Theorem (see B.4.1) may be used to show that (B1) also holds under P. Furthermore, the following theorems offer conditions, so that all discounted bond prices are martingales under Q. Then Q is called an *equivalent martingale measure* and, as shown by Harrison and Kreps (1979), the market is free of arbitrage.

Theorem 4.2.3 (Kennedy 1994). *Under the assumptions (B1), the measure Q is an equivalent martingale measure iff for all $t \in [0, T^*]$ and $T \in [t, t + T^{**}]$*

$$\mu(t,T) = \mu(0,T) + \int_0^T c(t \wedge v, v, T)\, dv. \tag{4.4}$$

Proof. If all discounted bond prices are martingales, then Q is an equivalent martingale measure. This is equivalent to

$$\mathbb{E}\Big(e^{-\int_0^t r_u\,du} B(t,T)\Big|\mathcal{F}_s\Big) = e^{-\int_0^s r_u\,du} B(s,T) \qquad \forall 0 \le s \le t \le T \le s + T^{**}.$$

We get

$$\begin{aligned}
&\exp\Big(-\int_0^s r_u\,du-\int_s^T f(s,u)\,du\Big)=\mathbb{E}\Big[\exp\Big(-\int_0^t r_u\,du-\int_t^T f(t,u)\,du\Big)\Big|\mathcal{F}_s\Big]\\
\Leftrightarrow\quad &1=\mathbb{E}\Big[\exp\Big(-\int_s^t r_u\,du-\int_t^T f(t,u)\,du+\int_s^T f(s,u)\,du\Big)\Big|\mathcal{F}_s\Big]\\
&=\mathbb{E}\Big[\exp\Big(-\int_s^t f(u,u)-f(s,u)\,du-\int_t^T f(t,u)-f(s,u)\,du\Big)\Big|\mathcal{F}_s\Big] \qquad (4.5)\\
&=:\mathbb{E}\big(e^{-A-B}\big|\mathcal{F}_s\big)=\mathbb{E}\big(e^{-(A+B)}\big).
\end{aligned}$$

The last equation holds because of Remark 4.1.3. In our Gaussian setup the forward rates are normally distributed. Then also the integrals of the forward rates and A and B are normally distributed. The above expectation can therefore be calculated using the Laplace-transform of $-(A+B)$.

By the definition of the forward rates, (4.3), we obtain

$$\mathbb{E}(A+B)=\int_s^t \mu(u,u)-\mu(s,u)\,du+\int_t^T \mu(t,u)-\mu(s,u)\,du.$$

The variances equal

$$\begin{aligned}
\mathrm{Var}(A) &= \mathrm{Var}\Big(\int_s^t X(u,u)-X(s,u)\,du\Big)\\
&= \int_s^t\int_s^t \mathrm{Cov}\Big(X(u,u)-X(s,u),X(v,v)-X(s,v)\Big)\,du\,dv\\
&= \int_s^t\int_s^t c(u\wedge v,u,v)-c(s,u,v)\,du\,dv
\end{aligned}$$

and

$$\mathrm{Var}(B) = \int_t^T\int_t^T c(t,u,v)-c(s,u,v)\,du\,dv,$$

while the covariance of A and B becomes

$$\begin{aligned}
\mathrm{Cov}(A,B) &= \int_s^t\int_t^T c(u,u,v)-c(s,u,v)-c(s,u,v)+c(s,u,v)\,dv\,du\\
&= \int_s^t\int_t^T c(u,u,v)-c(s,u,v)\,dv\,du.
\end{aligned}$$

We therefore obtain for the variance of $A+B$

$$\begin{aligned}
\mathrm{Var}(A+B) &= \int_s^t\int_s^t c(u\wedge v,u,v)-c(s,u,v)\,du\,dv + \int_t^T\int_t^T c(t,u,v)-c(s,u,v)\,du\,dv \\
&\quad +2\int_s^t\int_t^T c(u,u,v)-c(s,u,v)\,dv\,du \\
&= \int_s^T\int_s^T c\big((u\wedge v)\wedge t,u,v\big)-c(s,u,v)\,dv\,du \\
&= 2\int_{u=s}^T\int_{v=s}^u c\big(\underbrace{(u\wedge v)}_{=v}\wedge t,u,v\big)-c(s,u,v)\,dv\,du,
\end{aligned}$$

by the symmetry of c. Equation (4.5) requires $\mathbb{E}[\exp -(A+B)] = \exp\big[-\mathbb{E}(A+B)+\frac{1}{2}\mathrm{Var}(A+B)\big]$ being equal to one. Therefore the exponent needs to be zero, which is equivalent to

$$\begin{aligned}
&\int_{u=s}^T\int_{v=s}^u c(v\wedge t,u,v)-c(s,u,v)\,dv\,du \\
&\qquad = \int_s^t \mu(u,u)-\mu(s,u)\,du - \int_t^T \mu(t,u)-\mu(s,u)\,du \\
&\qquad = \int_s^T \mu(u\wedge t,u)-\mu(s,u)\,du.
\end{aligned}$$

Setting $s=0$ and and taking the partial derivative with respect to T, the following drift-condition is obtained

$$\mu(t,T) = \mu(0,T)+\int_0^T c(t\wedge v,v,T)\,dv.$$

■

4.3 Models with Credit Risk

In this section we consider a market which has two types of bonds. In contrast to the riskless bonds, we denote the price of a bond incorporating a certain default risk by $\bar{B}(t,T)$.

We model the forward rates of the riskless bonds as in the preceding chapter and the forward rates of the defaultable bonds similarly.

Assumption (B2): Assume $(\bar{X}(s,t))_{s,t\in[0,\tilde{T}]}$ is a zero-mean, continuous Gaussian random field with covariance function

$$\mathrm{Cov}\left(\bar{X}_{s_1,t_1}, \bar{X}_{s_2,t_2}\right) = \bar{c}(s_1 \wedge s_2, t_1, t_2),$$

where $\bar{c}(0,t_1,t_2) = 0$. Further on, assume that increments of the type $X(s_2,t) - X(s_1,t)$ and $\bar{X}(s_2,t) - \bar{X}(s_1,t)$ for $0 \le s_1 \le s_2 \le t \le \tilde{T}$ are independent of

$$\mathcal{G}_{s_1} = \sigma\big(X(s,t), \bar{X}(s,t) \quad : \quad 0 \le s \le s_1, t \in [s, s+T^{**}]\big).$$

The defaultable forward rate is modeled with a deterministic function $\bar{\mu}(s,t) : \mathbb{R}^2 \mapsto \mathbb{R}$ by

$$\bar{f}(t,T) \quad := \quad \bar{\mu}(t,T) + \bar{X}(t,T), \tag{4.6}$$

for all $t \in [0,T^*]$ and $T \in [t, t+T^{**}]$.

In the considered hazard rate framework the default intensity $(\lambda_t)_{t\ge 0}$ is assumed to be a nonnegative $(\mathcal{G}_t)_{t\ge 0}$-adapted process, see Appendix A. $\mathcal{G}_t$ can be interpreted as the available information at time t.

4.3.1 Zero Recovery

In a market with credit risk the dynamics of the bond relate to several factors. The risk-free interest rate has certainly a fundamental influence on the behavior of the defaultable bond. Besides that, the creditworthiness of the bond plays an important role. Creditworthiness is represented by the probability of a default, respectively the default intensity. The third component is the price of the bond after default, named recovery. In this first approach we consider the case of zero recovery, that is, the case where the value of the bond after default is zero. Hence only risk-free interest and default intensity remain to be considered.

Theorem 4.3.1. *Assume (B2) and consider a measure Q which is equivalent to the objective measure. If the defaultable forward rates take the form (4.6) under Q and $\int_0^{T^*} \lambda_s\, ds < \infty$ a.s., then Q is an equivalent martingale measure iff the following two conditions hold on $\{\tau > t\}$:*

(i) For all $0 \le t \le T^*$

$$\bar{f}(t,t) = r_t + \lambda_t, \tag{4.7}$$

(ii) For all $t \in [0,T^*]$ and $T \in [t, t+T^{**}]$

$$\bar{\mu}(t,T) = \bar{\mu}(0,T) + \int_0^T \bar{c}(t \wedge v, v, T)\, dv. \tag{4.8}$$

Proof. The case of risk-free bonds was already examined in the previous section. Consider the risky bonds. For $0 \le s \le t \le T^*$ and all $T \in [s, s+T^{**}]$, we show[2]

$$\mathbb{E}\Big(e^{-\int_0^t r_u\,du} \cdot \bar{B}(t,T)\Big|\mathcal{F}_s\Big) \stackrel{!}{=} e^{-\int_0^s r_u\,du}\bar{B}(s,T),$$

which is equivalent to

$$\begin{aligned}
&\mathbb{E}\Big[\exp\big(-\int_0^t r_u\,du\big)\cdot 1_{\{\tau>t\}}\exp\big(-\int_t^T \bar{f}(t,u)\,du\big)\Big|\mathcal{F}_s\Big] \\
&\quad = 1_{\{\tau>s\}}\exp\big(-\int_0^s r_u\,du\big)\mathbb{E}\Big(\exp(-\int_s^t (r_u+\lambda_u)\,du - \int_t^T \bar{f}(t,u)\,du)\Big|\mathcal{F}_s\Big) \\
&\quad \stackrel{!}{=} \exp\big(-\int_0^s r_u\,du\big)1_{\{\tau>s\}}\exp\big(-\int_s^T \bar{f}(s,u)\,du\big).
\end{aligned}$$

The first equality follows using Lemma A.1.2. Condition (4.7) implies, on $\{\tau > s\}$,

$$\begin{aligned}
&\mathbb{E}\Big(\exp(-\int_s^t \bar{r}_u\,du - \int_t^T \bar{f}(t,u)\,du)\Big|\mathcal{F}_s\Big) = \exp\Big(-\int_s^T \bar{f}(s,u)\,du\Big) \\
\Leftrightarrow\quad &\mathbb{E}\Big(\exp(-\int_s^t \bar{f}(u,u)-\bar{f}(s,u)\,du - \int_t^T \bar{f}(t,u)-\bar{f}(s,u)\,du)\Big|\mathcal{F}_s\Big) = 1 \\
\Leftrightarrow\quad &\mathbb{E}\Big(\exp(-\int_s^t \bar{f}(u,u)-\bar{f}(s,u)\,du - \int_t^T \bar{f}(t,u)-\bar{f}(s,u)\,du)\Big) = 1. \qquad (4.9)
\end{aligned}$$

This is exactly (4.5), with defaultable forward rates $\bar{f}(t,T)$ rather than $f(t,T)$. So we obtain analogously the drift condition (4.8) for the default case.

For the converse, if Q is already an equivalent martingale measure, the price of the bond is the expectation of its discounted payoff under Q, i.e.,

$$\begin{aligned}
\bar{B}(t,T) &= \mathbb{E}\big[\exp(-\int_t^T r_u\,du)1_{\{\tau>T\}}\big|\mathcal{F}_t\big] \\
&= 1_{\{\tau>t\}}\mathbb{E}\big[\exp(-\int_t^T r_u+\lambda_u\,du)\big|\mathcal{F}_t\big] \\
&= 1_{\{\tau>t\}}B(t,T)\mathbb{E}^T\big[\exp(-\int_t^T \lambda_u\,du)\big|\mathcal{F}_t\big] \\
&= 1_{\{\tau>t\}}\exp(-\int_t^T f(t,u)\,du)\mathbb{E}^T\big[\exp(-\int_t^T \lambda_u\,du)\big|\mathcal{F}_t\big].
\end{aligned}$$

[2] All expectations are with respect to this equivalent martingale measure, if not stated otherwise. $\mathbb{E}^T$ denotes the expectation w.r.t. the T-forward measure, see page 16.

This yields the defaultable spot rate (on $\{\tau > s\}$)

$$
\begin{aligned}
\bar{f}(t,t) &= -\frac{\partial}{\partial T}\Big|_{T=t} \ln \Big[1_{\{\tau>t\}} \mathbb{E}\Big(\exp(-\int_t^T (r_u + \lambda_u)\, du) \Big| \mathcal{F}_t \Big) \Big] \\
&= \frac{\mathbb{E}\Big(\exp(-\int_t^T (r_u + \lambda_u)\, du)[r_T + \lambda_T] \Big| \mathcal{F}_t \Big)}{\mathbb{E}\Big(\exp(-\int_t^T (r_u + \lambda_u)\, du) \Big| \mathcal{F}_t \Big)}\Bigg|_{T=t} \\
&= r_t + \lambda_t.
\end{aligned}
$$

Using this with the assumption that Q is a martingale measure leads to (4.9) which implies (4.8) as in the risk-free case. ■

We emphasize that no assumption on the dynamics of the risk-free interest rate is needed except (4.7). Nevertheless, this equation strongly connects f, $\bar{f}$ and λ.

The *credit spread* $s(t,T)$ is the difference between defaultable and risk-free rate, and we obtain

$$
\begin{aligned}
s(t,T) &= \bar{f}(t,T) - f(t,T) \\
&= -\frac{\partial}{\partial T} \ln \mathbb{E}\Big(1_{\{\tau>t\}} \exp(-\int_t^T (r_u + \lambda_u)\, du) \Big| \mathcal{F}_t \Big) - f(t,T) \\
&= -\frac{\partial}{\partial T} \ln \Big[1_{\{\tau>t\}} B(t,T) \mathbb{E}^T \Big(\exp(-\int_t^T \lambda_u\, du) \Big| \mathcal{F}_t \Big) \Big] - f(t,T) \\
&= -\frac{\partial}{\partial T} \ln \mathbb{E}^T \big(\exp(-\int_t^T \lambda_u\, du) | \mathcal{F}_t \big).
\end{aligned} \tag{4.10}
$$

An explicit Model for the Intensity

We examine an example in greater detail. Jarrow and Turnbull (2000) use the following model for dynamics of the intensity under Q:

$$\lambda_t = a_0(t) + a_1(t)W(t) + a_2(t)r_t.$$

$(W(t))_{t\geq 0}$ is a Brownian Motion which could represent the log-returns of the asset value of a company or of an index. We assume $(W(t))_{t\geq 0}$ to be independent of $(r_t)_{t\geq 0}$, the calculations for correlated processes being analogous.

This model suggests a specific structure for $(\lambda_t)_{t\geq 0}$, while up to now we considered primarily the forward rates. Nevertheless, starting with a random field model for the risk-free interest rate and the assumptions on $(\lambda_t)_{t\geq 0}$, we will derive $\bar{f}$ and show that it fits well in the above presented defaultable random field model.

Remark 4.3.2. Note that in this model the nonnegativity assumption for $(\lambda_t)_{t\geq 0}$ is violated. Nevertheless, this model can still be seen as an approximation of the model with default intensity $\big((\lambda_t)^+\big)_{t\geq 0}$ if the probability of the default intensity being negative is small, which should be true for a suitable choice of the parameters. This is also a common problem in Gaussian interest models, which admit negative interest rates with positive probability, see Rogers (1995).

To calculate $\bar{f}(t,T)$ via (4.10) we start with[3]

$$\begin{aligned}\mathbb{E}^T\big[\exp(-\int_t^T \lambda_u\,du)\big|\mathcal{F}_t\big] &= \exp(-\int_t^T a_0(u)\,du)\,\mathbb{E}_t\big[\exp(-\int_t^T a_1(u)W(u)\,du)\big]\\ &\quad\cdot\mathbb{E}_t^T\big[\exp(-\int_t^T a_2(u)r_u\,du)\big]. \end{aligned}\tag{4.11}$$

This holds, because W_λ is independent of r. The last factor becomes

$$\begin{aligned}\mathbb{E}_t^T\big[\exp(-\int_t^T a_2(u)r_u\,du)\big] &= B(t,T)^{-1}\mathbb{E}_t\big[\exp(-\int_t^T(1+a_2(u))r_u\,du)\big]\\ &= \exp(\int_t^T \mu(t,u)+X(t,u)\,du)\,\mathbb{E}_t\big[\exp(-\int_t^T(1+a_2(u))r_u\,du)\big].\end{aligned}$$

The exponents are normally distributed. To calculate the Laplace transform we need their expectations and variances, which equal for the first term in (4.11)

$$\mathbb{E}_t\big[\int_t^T a_1(u)W(u)\,du\big] = \int_t^T a_1(u)\mathbb{E}(W(u)|\mathcal{F}_t)\,du = W_t\int_t^T a_1(u)\,du$$

and

$$\begin{aligned}\mathrm{Var}_t\big[\int_t^T a_1(u)W(u)\,du\big] &= \int_t^T\int_t^T a_1(u)a_1(v)\mathbb{E}_t\big((W(u)-W_t)(W(v)-W(t))\big)\,dudv\\ &= \int_t^T\int_t^T a_1(u)a_1(v)(u\wedge v-t)\,dudv.\end{aligned}$$

For the last term in (4.11) we obtain

$$\begin{aligned}\mathbb{E}_t\big[\int_t^T(a_2(u)+1)r_u\,du\big] &= \int_t^T(a_2(u)+1)\mu(u,u)\,du+\int_t^T(a_2(u)+1)\mathbb{E}_t\big(X(u,u)\big)\,du\\ &= \int_t^T(a_2(u)+1)\big(\mu(u,u)+X(t,u)\big)\,du\end{aligned}$$

[3] For convenience we write $\mathbb{E}_t(\cdot)$ for $\mathbb{E}^Q(\cdot|\mathcal{F}_t)$, $\mathbb{E}_t^T(\cdot)$ for $E^{Q^T}(\cdot|\mathcal{F}_t)$ and $\mathrm{Var}_t(\cdot)$ for $\mathrm{Var}^Q(\cdot|\mathcal{F}_t)$.

and

$$
\begin{aligned}
&\mathrm{Var}_t\Big[\int_t^T (a_2(u)+1) r_u \, du\Big] \\
&\quad = 2\int_t^T \int_t^u (a_2(u)+1)(a_2(v)+1)\mathbb{E}_t\big[(X(u,u)-X(t,u))(X(v,v)-X(t,v))\big]\, dv\, du \\
&\quad = 2\int_t^T \int_t^u (a_2(u)+1)(a_2(v)+1)\big(c(v,u,v)-c(t,u,v)\big)\, dv\, du.
\end{aligned}
$$

This yields

$$
\begin{aligned}
(4.11) = \exp\Big[&-\int_t^T a_0(u)\, du - \int_t^T a_1(u) W(t)\, du + \int_t^T \int_t^u a_1(u) a_1(v) v \, du dv - \frac{t}{2}\Big(\int_t^T a_1(u)\, du\Big)^2 \\
&+ \int_t^T \mu(t,u) + X(t,u) - (a_2(u)+1)\big(\mu(u,u)+X(t,u)\big)\, du \\
&+ \int_t^T \int_t^u (a_2(u)+1)(a_2(v)+1)\big(c(v,u,v)-c(t,u,v)\big)\, dv\, du\Big].
\end{aligned}
$$

For the credit spread we therefore obtain

$$
\begin{aligned}
s(t,T) \;=\; & a_0(T) + W(t) a_1(T) - \int_t^T a_1(T) a_1(u) u \, du + t a_1(T)\int_t^T a_1(u)\, du \\
& -\mu(t,T) + (a_2(T)+1)\mu(T,T) + a_2(T) X(t,T) \\
& -(a_2(T)+1)\int_t^T (a_2(u)+1)\big(c(u,u,T)-c(t,u,T)\big)\, du.
\end{aligned}
$$

Directly from the drift condition (4.4) we get

$$
\mu(T,T) - \mu(t,T) = \int_t^T \big(c(u,u,T)-c(t,u,T)\big)\, du, \tag{4.12}
$$

which implies

$$
\begin{aligned}
&-\mu(t,T) + (a_2(T)+1)\Big[\mu(T,T) - \int_t^T (a_2(u)+1)\big(c(u,u,T)-c(t,u,T)\big)\, du\Big] \\
&\quad = -\mu(t,T) + (a_2(T)+1)\Big[\mu(t,T) - \int_t^T a_2(u)\big(c(u,u,T)-c(t,u,T)\big)\, du\Big] \\
&\quad = a_2(T)\mu(t,T) - (a_2(T)+1)\int_t^T a_2(u)\big(c(u,u,T)-c(t,u,T)\big)\, du.
\end{aligned}
$$

We conclude

$$
\begin{aligned}
s(t,T) &= a_0(T) + W(t)a_1(T) - \int_t^T a_1(T)a_1(u)u\,du + ta_1(T)\int_t^T a_1(u)\,du + a_2(T)X(t,T) \\
&\quad + a_2(T)\mu(t,T) - (a_2(T)+1)\int_t^T a_2(u)\big(c(u,u,T) - c(t,u,T)\big)\,du \\
&=: \tilde{\mu}(t,T) + W(t)a_1(T) + a_2(T)f(t,T),
\end{aligned}
$$

$\tilde{\mu}(t,T)$ being a deterministic function. Since

$$\bar{f}(t,T) = f(t,T) + s(t,T),$$

the defaultable forward rate takes the form (4.6) with a Gaussian random field $\tilde{X}$, which has the covariance function

$$\mathrm{Cov}(X_{s_1,t_1}, X_{s_2,t_2}) = a_1(t_1)a_1(t_2)\,(s_1 \wedge s_2) + (1 + a_2(t_1))(1 + a_2(t_2))c(s_1 \wedge s_2, t_1, t_2).$$

So this model fits well into our setup.

4.3.2 Recovery of Treasury Value

In this recovery model the bondholder receives the payoff $1_{\{\tau > T\}} + (1-w)1_{\{\tau \le T\}}$ at maturity. It is useful to split the bond's value into two parts, a zero recovery bond and a riskless bond

$$\bar{B}(t,T)_{TV} = w\bar{B}^0(t,T) + (1-w)B(t,T).$$

For no-arbitrage we need for all $0 \le s \le t \le T^*$ and $T \in [s, s+T^{**}]$

$$
\begin{aligned}
& \mathbb{E}_s\Big[\exp(-\int_s^t r_u\,du)\big(w\bar{B}^0(t,T) + (1-w)B(t,T)\big)\Big] = w\bar{B}^0(s,T) + (1-w)B(s,T) \\
\Leftrightarrow \quad & w\Big[\mathbb{E}_s\Big(\exp(-\int_s^t r_u\,du)\bar{B}^0(t,T)\Big) - \bar{B}^0(s,T)\Big] \\
& \quad + (1-w)\Big[\mathbb{E}_s\Big(\exp(-\int_s^t r_u\,du)B(t,T)\Big) - B(s,T)\Big] = 0.
\end{aligned}
$$

This condition is certainly valid if for both riskless and zero recovery bonds the no-arbitrage condition is valid. If we examine the Gaussian random field setup (cf. 4.2 and 4.3.1) we get that the market is free of arbitrage, if for all $t \in [0, T^*]$ and $T \in [t, t+T^{**}]$

$$\mu(t,T) = \mu(0,T) + \int_0^T c(t \wedge v, v, T)\,dv,$$

and

$$\bar{\mu}(t,T) = \bar{\mu}(0,T) + \int_0^T \bar{c}(t \wedge v, v, T)\, dv.$$

For the converse, use that the default-free bond is traded in the market. Therefore a market free of arbitrage implies

$$\Big[\mathbb{E}_s\Big(\exp(-\int_s^t r_u\, du)B(t,T)\Big) - B(s,T)\Big] = 0.$$

Thus the discounted zero recovery bond has to be a martingale and we conclude that the drift conditions hold.

4.3.3 Fractional Recovery of Treasury Value

The history of a defaultable bond possibly admits several default events. Eventually a coupon can not be paid, while the company is not necessarily forced to default. Also rating migration is sometimes referred to as a default event, and can be incorporated in the following, more general model.

Evidently a model with multiple default events is needed, while this yields that a default event not necessarily leads to default of the bond. Moreover, such a default event could be an upgrading, where the value of the bond increases.

Assumption (B3): In the fractional recovery of treasury value model the bond is modeled for all $t \in [0, T^*], T \in [t, t + T^{**}\ ]$ as

$$\bar{B}_{FR}(t,T) \quad = \quad Q(t) \exp(-\int_t^T \bar{f}(t,u)\, du),$$

with

$$Q(t) := \prod_{\tau_i \le t} (1 - L_{\tau_i}),$$

where the loss process $(L_t)_{t\ge 0}$ takes values in $(0,1)$ and is adapted to $(\mathcal{G}_t)_{t\ge 0}$. The $\tau_1, \tau_2, \ldots$ are the jump times of a Cox Process, cf. Appendix A.

In this setting we define the defaultable forward rates by

$$\bar{f}(t,T) := -\frac{\partial}{\partial T} \ln \frac{\bar{B}_{FR}(t,T)}{Q(t)}.$$

As before[4] we model the defaultable forward rate by a Gaussian random field $\bar{X}$ with covariance function $\bar{c}$ under Q, i.e.,

$$\bar{f}(t,T) = \bar{\mu}(t,T) + \bar{X}(t,T).$$

[4]See Assumption (B2).

Theorem 4.3.3. *Under assumption (B3), discounted bond prices are martingales iff for all $t \in [0, T^*]$*

$$\bar{f}(t,t) = r_t + \lambda_t L_t \tag{4.13}$$

and drift condition (4.8) holds. In this case the market is arbitrage-free.

Proof. All discounted bond prices are martingales iff for all $0 \le s \le t \le T^*$ and $T \in [s, s+T^{**}]$

$$\mathbb{E}_s\Big[\exp(-\int_0^t r_u\,du)\bar{B}(t,T)\Big] = \exp(-\int_0^s r_u\,du)\bar{B}(s,T),$$

which is equivalent to

$$\begin{aligned}
1 &= \mathbb{E}_s\Big[\exp\big[-\int_s^t \big(r_u - \bar{f}(s,u)\big)\,du\big]\frac{Q(t)}{Q(s)}\exp\big[-\int_t^T \big(\bar{f}(t,u) - \bar{f}(s,u)\big)\,du\big]\Big] \\
&= \mathbb{E}_s\Big[\prod_{s<\tau_i\le t}(1-L_{\tau_i})\exp\big[-\int_s^t \big(f(u,u) - \bar{f}(s,u)\big)\,du - \int_t^T \big(\bar{f}(t,u) - \bar{f}(s,u)\big)\,du\big]\Big] \\
&\overset{A.1.4}{=} \mathbb{E}_s\Big[\exp\big[-\int_s^t \lambda_u L_u\,du - \int_s^t \big(f(u,u) - \bar{f}(s,u)\big)\,du - \int_t^T \big(\bar{f}(t,u) - \bar{f}(s,u)\big)\,du\big]\Big].
\end{aligned}$$

Using (4.13) we obtain

$$1 = \mathbb{E}_s\Big[\exp\big[-\int_s^T \big(\bar{f}(u\wedge t,u) - \bar{f}(s,u)\big)\,du\big]\Big],$$

and analogously to (4.5), this is true under (4.8).

For the converse, note that if discounted bond prices are martingales,

$$\begin{aligned}
\bar{B}_{FR}(t,T) &= \mathbb{E}\Big[\exp(-\int_t^T r_u\,du)\cdot Q(T)\Big|\,\mathcal{F}_t\Big] \\
&= Q(t)B(t,T)\mathbb{E}_t^T\Big[\exp(-\int_t^T \lambda_u\,du)\Big],
\end{aligned}$$

which yields for the defaultable spot rate

$$\begin{aligned}
\bar{r}_t &= \bar{f}(t,t) = -\frac{\partial}{\partial T}\Big|_{T=t}\ln\Big[\prod_{\tau_i\le t}(1-L_{\tau_i})\,\mathbb{E}_t\big[\exp\big(-\int_t^T (r_u + \lambda_u L_u)\,du\big)\big]\Big] \\
&= r_t + \lambda_t L_t.
\end{aligned}$$

Similar to the preceding proofs we obtain the conclusion. ∎

Remark 4.3.4. If in the previous theorem $(L_t)_{t\geq 0}$ is admitted to become equal to 1, the value of the bond may drop to zero. In this case the defaultable dynamics vanish and no conditions are needed to ensure an arbitrage-free market. Thus, for equivalency in Theorem 4.3.3, the drift conditions need to be satisfied on $\{Q(t) > 0\}$ only.

We also conclude for the credit spread in this model:

$$\begin{aligned} s(t,T) &= -\frac{\partial}{\partial T}\big[\ln \bar{B}(t,T) - \ln B(t,T)\big] \\ &= -\frac{\partial}{\partial T}\big[\ln Q(t) + \ln B(t,T) + \ln \mathbb{E}_t^T\big(\exp(-\int_t^T \lambda_u L_u \, du)\big) - \ln B(t,T)\big] \\ &= -\frac{\partial}{\partial T}\mathbb{E}_t^T\big[\exp(-\int_t^T \lambda_u L_u \, du)\big]. \end{aligned}$$

4.4 Explicit Pricing Formulas

This section provides explicit pricing formulas for certain credit derivatives introduced already in Section 1.9. They provide the basis for the calibration methods developed later on. Throughout this section, we stay within the notation of Section 4.2 and 4.3.

4.4.1 Default Digitals

A basic derivative based on a credit risky underlying is the default digital put. It promises a fixed payoff, say 1, if a default occurred before maturity, and zero otherwise. We focus on the derivative where the payoff is settled at maturity.

It may be recalled that the default digital put with payoff at maturity is intrinsically related to the zero recovery bond, as

$$P^d(t,T) + B^0(t,T) = B(t,T).$$

Assumption (C1): Assume that both risk-free and defaultable forward rates admit a representation via Gaussian random fields and the drift-conditions (4.4) and (4.8) are satisfied. Further on, assume that the considered defaultable bond admits a fractional recovery of market value[5] with positive, deterministic loss function L_t and (4.13) holds.

If assumption (C1) holds, Theorems 4.2.3 and 4.3.3 yield that the market is free of arbitrage. Furthermore, we deduce from (4.13) that

$$\lambda_t = \frac{\bar{f}(t,t) - f(t,t)}{L_t}. \tag{4.14}$$

Instead of defining the dynamics of $f(t,T)$ and λ_t and then deriving $\bar{f}(t,T)$, we want to propose the dynamics of $f(t,T)$ and $\bar{f}(t,T)$ and investigate the consequences for λ_t. This reflects the fact that λ_t is not observable in the market, while the forward rates are. Therefore, we use equation (4.14) as a starting point for this section.

This immediately has some consequences. By definition, $(\lambda_t)_{t\geq 0}$ is assumed to be a non-negative process. Quite contrary, equation (4.14) suggest λ_t to have a normal distribution, which has a positive probability to be negative. Thus, $(\lambda_t)_{t\geq 0}$ must be deterministic.

We want to relax this rigid assumption and rather take (4.14) as a definition for the stochastic process $(\lambda_t)_{t\geq 0}$. Because L_t is deterministic, λ_t turns out to be a Gaussian random field. Following Remark 4.3.2, for appropriate parameters, $(\lambda_t)_{t\geq 0}$ might be negative with just a small probability and therefore can be used as an approximation of the true default intensity.

[5] See Section 4.3.3.

For ease of notation we write $\bar{f}(u)$ instead of $\bar{f}(u,u)$ and similarly $\mu(u)$, $\bar{\mu}(u)$, $X(u)$ and $\bar{X}(u)$. Furthermore, set $t := 0$.

In the following calculations we will need a measure for interaction between the risk-free and the defaultable forward rate. For this, we set for $s, t_1, t_2 \in [0, \tilde{T}]$

$$\varsigma(s, t_1, t_2) := \mathrm{Cov}\left(\bar{f}(s, t_1), f(s, t_2)\right) = \mathrm{Cov}\left(\bar{X}(s, t_1), X(s, t_2)\right).$$

Note that $\varsigma(s, t_1, t_2)$ is not necessarily symmetric in t_1 and t_2. Furthermore, assumption (A2) immediately yields

$$\mathrm{Cov}\left(\bar{X}(s_1, t_1), X(s_2, t_2)\right) = \varsigma(s_1 \wedge s_2, t_1, t_2).$$

Frequently, we will consider terms similar to

$$r_t + \lambda_t = r_t(1 - \frac{1}{L_t}) + \bar{f}_t \frac{1}{L_t},$$

and therefore set

$$l_t := \left(1 - \frac{1}{L_t}\right).$$

Lemma 4.4.1. *Under (C1), the price of the zero recovery bond is for $T \in [0, T^*]$*

$$\begin{aligned} B^0(0,T) \quad &= \quad \exp\Big[- \int_0^T \left(l_u \mu(0,u) + \frac{\bar{\mu}(0,u)}{L_u}\right) du + 2 \int_0^T \int_0^u \frac{l_u}{L_v} \varsigma(v,v,u)\, dv\, du \\ &\qquad - \int_0^T \int_0^u \left(\frac{l_u}{L_v} c(v,u,v) + \frac{l_v}{L_u} \bar{c}(v,u,v)\right) dv\, du \Big]. \end{aligned}$$

Proof. We obtain for the price of the zero recovery bond

$$\begin{aligned} B^0(0,T) \quad &= \quad \mathbb{E}\Big(\exp(- \int_0^T r_u + \lambda_u \, du) \Big) \\ &= \quad \mathbb{E}\Big(\exp[- \int_0^T (r_u \, l_u + \frac{\bar{f}_u}{L_u}) \, du] \Big). \end{aligned}$$

The exponent's expectation becomes

$$\begin{aligned} \mathbb{E}\Big[- \int_0^T l_u f(u) + \frac{\bar{f}_u}{L_u} \, du \Big] \quad &= \quad - \int_0^T \left(l_u \mu(u) + \frac{\bar{\mu}(u)}{L_u} \right) du \\ &\overset{(4.8)}{=} \quad - \int_0^T \left(l_u \mu(0,u) + \frac{\bar{\mu}(0,u)}{L_u} \right) du \\ &\qquad - \int_0^T \int_0^u \left(l_u c(v,u,v) + \frac{1}{L_u} \bar{c}(v,u,v) \right) dv\, du \end{aligned}$$

and its variance equals

$$
\begin{aligned}
&\operatorname{Var}\Big[\int_0^T l_u X(u) + \frac{\bar{X}_u}{L_u}\,du\Big] \\
&\quad = \int_0^T\int_0^T \Big[l_u l_v c(u\wedge v,u,v) + \frac{1}{L_u L_v}\bar{c}(u\wedge v,u,v) + 2\frac{l_u}{L_v}\,\varsigma(u\wedge v,v,u)\Big]\,dv\,du.
\end{aligned}
$$

The bond price thus equals

$$
\begin{aligned}
B^0(0,T) = \exp\Big\{ &-\int_0^T \Big(l_u\mu(0,u) + \frac{\bar{\mu}(0,u)}{L_u}\Big)\,du \\
&-\int_0^T\int_0^u \Big(l_u c(v,u,v) + \frac{1}{L_u}\bar{c}(v,u,v)\Big)\,dv\,du \\
&+\frac{2}{2}\int_0^T\int_0^u \Big(l_u l_v c(v,u,v) + \frac{1}{L_u L_v}\bar{c}(v,u,v) + 2\frac{l_u}{L_v}\,\varsigma(v,v,u)\Big)\,dv\,du\Big\} \\
= \exp\Big\{ &-\int_0^T \Big(l_u\mu(0,u) + \frac{\bar{\mu}(0,u)}{L_u}\Big)\,du \\
&-\int_0^T\int_0^u \Big(\frac{l_u}{L_v}c(v,u,v) + \frac{l_v}{L_u}\bar{c}(v,u,v)\Big)\,dv\,du \\
&+2\int_0^T\int_0^u \frac{l_u}{L_v}\,\varsigma(v,v,u)\,dv\,du\Big\}. \qquad \blacksquare
\end{aligned}
$$

At a first glance, it seems confusing that the loss rate L_t appears in the price of a zero recovery bond. Note that this follows, because we take equation (4.14) as a definition of the hazard rate.

If the price of the zero recovery bond is available, the following formula allows to calibrate the loss rate L_t. Denoting the forward rate of the zero recovery bond by $f^0(\cdot,T)$, we have

$$
\begin{aligned}
\bar{f}(t) &= r_t + \lambda_t L_t \\
&= r_t + (f^0(t) - f(t))L_t \\
\Leftrightarrow \qquad L_t &= \frac{\bar{f}(t) - f(t)}{f^0(t) - f(t)}.
\end{aligned}
$$

The zero recovery bond is a basis for evaluating more complicated derivatives. It therefore will prove useful to obtain some auxiliary results.

We define the *zero recovery measure* Q^0 by

$$dQ^0 = \frac{\exp(-\int_0^T r_u + \lambda_u \, du)}{B^0(0,T)} \, dQ. \tag{4.15}$$

Then Q^0 is equivalent to Q. Note that we suppress the dependence of T in the notation.

The following lemma gives a representation of $f^0(t,T)$ in terms of $X(t,T)$ and $\bar{X}(t,T)$.

Lemma 4.4.2. *Under (C1), the forward rate offered by the zero recovery bond can be represented for $t \in [0,T^*], T \in [t, t+T^{**}]$ as*

$$\begin{aligned}
f^0(t,T) &= \mu^0(t,T) + l_T X(t,T) + \frac{\bar{X}(t,T)}{L_T}, \qquad (4.16)\\
\mu_0(t,T) &:= l_T \mu(t,T) + \frac{\bar{\mu}(t,T)}{L_T} \\
&+ \int_t^T \Big\{ \frac{L_T - 1}{L_u \, L_T} \big[c(u,u,T) - c(t,u,T) - (\varsigma(u,u,T) - \varsigma(t,u,T)) \big] \\
&\qquad \frac{L_u - 1}{L_u \, L_T} \big[\bar{c}(u,u,T) - \bar{c}(t,u,T) - (\varsigma(u,T,u) - \varsigma(t,T,u)) \big] \Big\} \, du.
\end{aligned}$$

Furthermore, $f^0(t,T)$ has independent increments in the first coordinate and the following covariance function for $s \in [0,T^*]$ and $t_1, t_2 \in [s, s+T^{**}]$:

$$\begin{aligned}
c^0(s,t_1,t_2) &= l_{t_1} l_{t_2} c(s,t_1,t_2) + \frac{\bar{c}(s,t_1,t_2)}{L_{t_1} \, L_{t_2}} \\
&+ \frac{l_{t_1}}{L_{t_2}} \varsigma(s,t_2,t_1) + \frac{l_{t_2}}{L_{t_1}} \varsigma(s,t_1,t_2).
\end{aligned}$$

Proof. By definition,

$$\begin{aligned}
f^0(t,T) &= -\frac{\partial}{\partial T} \ln B^0(t,T) \\
&= \frac{1}{B^0(t,T)} \mathbb{E}_t \big[\exp(-\int_t^T r_u + \lambda_u \, du)(r_T + \lambda_T) \big] \\
&= \mathbb{E}_t^0 (r_T + \lambda_T) = \mathbb{E}_t^0 \Big(l_T r_T + \frac{\bar{f}(T)}{L_T} \Big),
\end{aligned}$$

where $\mathbb{E}^0$ denotes expectation with respect to the zero recovery measure Q^0.

As

$$\mathbb{E}_t^0 (r_T) = \mu(T,T) + X(t,T) + \mathbb{E}_t^0 \big(X(T,T) - X(t,T) \big),$$

we can use Lemma B.4.1 to compute $\mathbb{E}^0\big(X(T,T) - X(t,T)\big)$ and obtain

$$
\begin{aligned}
\mathbb{E}^0_t(r_T) &\overset{4.8}{=} \mu(t,T) + \int_t^T \big(c(u,u,T) - c(t,u,T)\big)\,du + X(t,T) \\
&+ \operatorname{Cov}_t\Big[X(T,T) - X(t,T), -\int_t^T \big(r_u + \lambda_u\big)\,du\Big] \\
&= f(t,T) + \int_t^T \Big([c(u,u,T) - c(t,u,T)](1 - l_u) - \frac{1}{L_u}[\varsigma(u,u,T) - \varsigma(t,u,T)]\Big)\,du.
\end{aligned}
$$

Furthermore,

$$
\begin{aligned}
\mathbb{E}^0_t(\bar{f}(T)) &= \bar{\mu}(t,T) + \bar{X}(t,T) \\
&+ \int_t^T \Big(l_u\,[\bar{c}(u,u,T) - \bar{c}(t,u,T)] - l_u\,[\varsigma(u,T,u) - \varsigma(t,T,u)]\Big)\,du.
\end{aligned}
$$

The assertion about the covariance function remains to be shown. Actually,

$$
\begin{aligned}
&\operatorname{Cov}(f^0(s_1,t_1), f^0(s_2,t_2)) \\
&\quad= \mathbb{E}\Big((l_{t_1}X(s_1,t_1) + \frac{\bar{X}(s_1,t_1)}{L_{t_1}})\,(l_{t_2}X(s_2,t_2) + \frac{\bar{X}(s_2,t_2)}{L_{t_2}})\Big) \\
&\quad= l_{t_1}l_{t_2}c(s_1 \wedge s_2, t_1, t_2) + \frac{\bar{c}(s_1 \wedge s_2, t_1, t_2)}{L_{t_1}\,L_{t_2}} \\
&\quad\quad + \frac{l_{t_1}}{L_{t_2}}\varsigma(s_1 \wedge s_2, t_2, t_1) + \frac{l_{t_2}}{L_{t_1}}\varsigma(s_1 \wedge s_2, t_1, t_2).
\end{aligned}
$$

■

4.4.2 Default Put

In this section we consider a default put with knock-out feature. The put is knocked out if a default occurs before maturity of the contract, which means that the promised payoff is paid only if there was no default until maturity of the contract. So this put protects against market risk but not against the loss in case of a default.

Denoting the price of a (knock out) default put with maturity T on a defaultable bond with maturity T' by $P^k(0,T,T')$, the risk neutral valuation principle yields

$$
P^k(0,T,T') = \mathbb{E}\Big[\exp(-\int_0^T r_u\,du)\big(K - \bar{B}(T,T')\big)^+ 1_{\{\tau > T\}}\Big],
$$

for all $0 \le T < T' \le \tilde{T}$.

Further on, denote by $B^k(0,T,T')$ a knock-out contract on the defaultable bond, which delivers the defaultable bond with maturity T' at time T, if no default happened until T and zero otherwise. This derivative seems a bit synthetic, but if both default put and default call with knock-out are traded, it can be replicated by the following combination of put and call:

$$\begin{aligned}
&C^k(0,T,T') - P^k(0,T,T') \\
&\qquad = \mathbb{E}\Big[\exp(-\int_0^T r_u\,du)\big[(\bar{B}(T,T')-K)^+ - (K-\bar{B}(T,T'))^+\big]1_{\{\tau>T\}}\Big] \\
&\qquad = \mathbb{E}\Big[\exp(-\int_0^T r_u\,du)\bar{B}(T,T')1_{\{\tau>T\}}\Big].
\end{aligned}$$

If the knock-out bond is not available one can use expression (4.17) for a still explicit, but more complicated pricing formula.

Theorem 4.4.3. *The price of a default put with maturity $T \in [0,T^*]$ on a defaultable bond with maturity $T' \in (T,t^*]$, which is knocked out if default occurs before T equals*

$$P^k(0,T,T') \;=\; B^0(0,T)K\Phi(-d_2) - B^k(0,T,T')\Phi(-d_1),$$

with deterministic terms

$$\begin{aligned}
\tilde{\sigma}(T,T') &:= \int_T^{T'}\int_T^{T'} \bar{c}(T,u,v)\,du\,dv, \\
\tilde{\mu}(T,T') &:= -\ln\frac{\bar{B}(0,T')}{\bar{B}(0,T)} + \int_T^{T'}\int_0^{T} l_v[\bar{c}(T,u,v)-\varsigma(T,u,v)]\,dv\,du + \frac{1}{2}\tilde{\sigma}(T,T'), \\
d_2 &:= \frac{-\tilde{\mu}(T,T')-\ln K}{\tilde{\sigma}(T,T')}, \\
d_1 &:= d_2 + \tilde{\sigma}(T,T').
\end{aligned}$$

Proof. The price of the put equals

$$\begin{aligned}
P^k(0,T,T') &= \mathbb{E}\Big[\exp(-\int_0^T r_u+\lambda_u\,du)\big(K-\exp(-\int_T^{T'}\bar{f}(T,u)\,du)\big)^+\Big] \\
&= B^0(0,T)\mathbb{E}^0\Big[\big(K-\exp(-\int_T^{T'}\bar{f}(T,u)\,du)\big)^+\Big].
\end{aligned}$$

From Lemma B.4.1, $\int_T^{T'}\bar{f}(T,u)\,du$ is normally distributed under Q^0, with the same variance as under Q but expectation

$$\begin{aligned}
\mathbb{E}^0\Big(\int_T^{T'} \bar{f}(T,u)\,du\Big) &= \mathbb{E}\Big(\int_T^{T'} \bar{f}(T,u)\,du\Big) + \mathrm{Cov}\Big(\int_T^{T'} \bar{f}(T,u)\,du, -\int_0^T (r_u+\lambda_u)\,du\Big) \\
&= \int_T^{T'} \bar{\mu}(T,u)\,du - \int_T^{T'}\int_0^T \Big(\frac{1}{L_v}\bar{c}(T,u,v) + l_v\varsigma(T,u,v)\Big)\,dv\,du.
\end{aligned}$$

Formula (B.5) reveals the option pricing formula[6]

$$P^k(0,T,T') = B^0(0,T)\Big[K\Phi(-d_2) - e^{\frac{\tilde{\sigma}(T,T')^2}{2} - \tilde{\mu}(T,T')}\Phi(-d_1)\Big],$$

where

$$\tilde{\mu}(T,T') = \int_T^{T'} \bar{\mu}(T,u)\,du - \int_T^{T'}\int_0^T \Big(\frac{1}{L_v}\bar{c}(T,u,v) + l_v\varsigma(T,u,v)\Big)\,dv\,du,$$

the computation of the other parameters being straightforward.

Further on, we have

$$\begin{aligned}
B^0(0,T)e^{\frac{\tilde{\sigma}(T,T')}{2} - \tilde{\mu}(T,T')} &= B^0(0,T)\mathbb{E}\Big(\exp\Big(-\int_0^T (r_u+\lambda_u)\,du - \int_T^{T'} \bar{f}(T,u)\,du\Big)\Big) \\
&= B^0(0,T)\mathbb{E}\Big(\exp\big(-\int_0^T r_u\,du\big)\bar{B}(T,T')1_{\{\tau>T\}}\Big) \\
&= B^k(0,T,T')
\end{aligned}$$

and the proof is finished. ■

If the use of $B^k(0,T,T')$ seems inappropriate, the following is useful:

$$\begin{aligned}
B^k(0,T,T') &= B^0(0,T)\,e^{\frac{\tilde{\sigma}(T,T')}{2} - \tilde{\mu}(T,T')} \\
&= \frac{B^0(0,T)\,\bar{B}(0,T')}{\bar{B}(0,T)}\exp\Big(-\int_T^{T'}\int_0^T l_v[\bar{c}(T,u,v) - \varsigma(T,u,v)]\,dv\,du\Big). \qquad (4.17)
\end{aligned}$$

Using the put-call representation of $B^k(0,T,T')$ the alternative representation

$$P^k(0,T,T') = \frac{B^0(0,T)K\Phi(-d_2) - C^k(0,T,T')\Phi(-d_1)}{\Phi(d_1)},$$

respectively

$$\Phi(d_1)P^k(0,T,T') + \Phi(-d_1)C^k(0,T,T') = B^0(0,T)K\Phi(-d_2)$$

is obtained.

[6] Note that we apply B.5 with respect to Q^0, setting $\xi_2 \equiv 0$.

4.4.3 Credit Spread Options

The pricing of credit spread options can be done in a more or less similar fashion. Consider a put on the credit spread of a defaultable bond which is knocked out at default, i.e., a derivative, which protects against spread widening risk, but not default risk.

Theorem 4.4.4. *Under assumption (C1), the price of the (knock out) credit spread put with maturity $T \in [0, T^*]$ on a defaultable bond with maturity $T' \in (T, t^*]$ equals*

$$P^k_{CS}(0, T, T') = B^k(0, T, T')\Phi\big(d_1\big) - KB^0(0, T)\Phi\big(d_2\big),$$

with the abbreviations

$$\begin{aligned}
\mu_1 &:= -\int_T^{T'} \Big[\bar{\mu}(0, u) - \mu(0, u) + \int_0^u \big(\bar{c}(v \wedge T, v, u) - c(v \wedge T, v, u)\big)\Big]\, dv\, du, \\
\sigma_1 &:= \int_T^{T'}\int_T^{T'} \Big[\bar{c}(u \wedge v, u, v) - \varsigma(T, u, v) - \varsigma(T, v, u) + c(u \wedge v, u, v)\Big]\, dv\, du, \\
\sigma_2 &:= \int_0^{T'}\int_0^{T'} l_1(u, T) l_1(v, T) c(u \wedge v, u, v)\, dv\, du + \int_0^T\int_0^T \frac{\bar{c}(u \wedge v, u, v)}{L_u L_v}\, dv\, du \\
&+ 2\int_0^{T'}\int_0^{T} \frac{l_1(u, T)}{L_v} \varsigma(u \wedge v, v, u)\, dv\, du, \\
\rho &:= \int_0^{T'}\int_T^{T'} l_1(u, T)[\varsigma(u \wedge T, v, u) - c(u \wedge T, v, u)]\, dv\, du \\
&+ \int_0^{T}\int_T^{T'} \frac{1}{L_u}[\bar{c}(u \wedge T, u, v) - \varsigma(u \wedge T), u, v)]\, dv\, du, \\
d_2 &:= \frac{\mu_1 - \ln K}{\sigma_1} + \rho\sigma_2, \\
d_1 &:= d_2 + \sigma_1, \\
l_1(u, T) &:= \begin{cases} l_u & \text{if } u \leq T, \\ 1 & \text{if } u > T. \end{cases}.
\end{aligned}$$

Proof. The price of the credit spread put equals

$$\begin{aligned} P^k_{CS}(0,T,T') &= \mathbb{E}\Big[\exp(-\int_0^T r_u\,du)\Big(\bar{B}(T,T') - KB(T,T')\Big)^+ 1_{\{\tau>T\}}\Big] \\ &= \mathbb{E}\Big[\exp\big[-\int_0^T (r_u+\lambda_u)\,du - \int_T^{T'} f(T,u)\,du\big] \\ &\quad \cdot \Big(\exp\big[-\int_T^{T'} (\bar{f}(T,u) - f(T,u))\,du\big] - K\Big)^+\Big]. \end{aligned}$$

Denote

$$\begin{aligned} \xi_1 &:= -\int_T^{T'} \big(\bar{f}(T,u) - f(T,u)\big)\,du, \\ \xi_2 &:= -\int_0^T (r_u+\lambda_u)\,du - \int_T^{T'} f(T,u)\,du \\ &= \mathbb{E}(\xi_2) - \int_0^{T'} l_1(u,T) X(u\wedge T,u)\,du - \int_0^T \frac{\bar{X}(u)}{L_u}\,du. \end{aligned}$$

Then Lemma B.4.3 applies. We compute

$$\begin{aligned} \mu_1 &= \mathbb{E}(\xi_1) = -\int_T^{T'} \big(\bar{\mu}(T,u) - \mu(T,u)\big)\,du \\ &= -\int_T^{T'} \Big[\bar{\mu}(0,u) - \mu(0,u)\,du + \int_0^u \big(\bar{c}(v\wedge T,v,u) - c(v\wedge T,v,u)\big)\Big]\,dv\,du, \end{aligned}$$

while the other parameters are obtained directly from their definition. Note that

$$\mathbb{E}\Big(\exp\big[-\int_0^T (r_u+\lambda_u)\,du - \int_T^{T'} \bar{f}(T,u)\,du\big]\Big) = B^k(0,T,T')$$

and

$$\mathbb{E}\big(\exp[-\int_0^T (r_u+\lambda_u)\,du]\big) = B^0(0,T).$$

Thus, Lemma B.4.3 yields the desired conclusion. ∎

4.4.4 Credit Default Swap and Swaption

In this section we consider the pricing of a credit default swaption, in particular the price of a so-called CDS call. It may be recalled from Section 1.9.3 that this derivative is a call on the swap premium, which is knocked out if a default of the underlying entity occurs before maturity. It was also shown, that, if the swap offers the replacement of the difference to an equivalent risk-free bond on default, the swap rate is

$$\bar{S}(T) = \frac{B(T,T_n) - \bar{B}(T,T_n)}{\sum_{i=1}^{n} B^0(T,T_i)}.$$

The pricing of the credit default swap therefore mainly relies on the pricing of the zero-recovery bond. Therefore, Lemma 4.4.1 immediately leads to a price of the credit default swap.

Thus, the price of the CDS call equals

$$\begin{aligned} C_S^k(0,T,T') &= \mathbb{E}\Big[\exp(-\int_0^T r_u\,du)\Big(B(T,T_n) - \bar{B}(T,T_n) - K\sum_{i=1}^{n} B^0(T,T_i)\Big)^+ 1_{\{\tau>T\}}\Big] \\ &= \mathbb{E}\Big[\exp(-\int_0^T r_u + \lambda_u\,du)\Big(B(T,T_n) - \exp(-\int_T^{T_n} \bar{f}(T,u)\,du) \\ &\qquad -K\sum_{i=1}^{n}\exp(-\int_T^{T_i} f^0(T,u)\,du)\Big)^+\Big]. \end{aligned} \tag{4.18}$$

Usually the final repayment, represented by $\bar{B}(T,T_n)$, dominates the coupon payments. This justifies the following

Assumption (C2): For the considered maturity $T \in [0,T^*]$ and the tenor structure $T < T_1 < \cdots < T_n \le T^*$ assume that the random variable

$$\exp\big(-\int_T^{T_n} \bar{f}(T,u)\,du\big) + K\sum_{i=1}^{n}\exp\big(-\int_T^{T_i} f^0(T,u)\,du\big) \tag{4.19}$$

can be approximated by a log-normal random variable, denoted by $\tilde{B}(T,T_1,\dots,T_n)$. Furthermore, we assume that discounted zero recovery bonds are martingales, i.e., drift condition (4.8) is satisfied for $\mu^0(\cdot)$ and $c^0(\cdot)$, respectively.

If the drift condition for zero recovery bonds is not satisfied, the following methods can be applied in a similar fashion and it is still possible to obtain explicit pricing formulas.

Under assumption (C2) the pricing of the credit default swaption is very similar to the pricing of a credit spread call, where the underlying is $\tilde{B}(T,T_1\dots,T_n)$.

Denote the mean and variance of $\tilde{B}(T,T_1,\dots,T_n)$ by $\tilde{m}$ and $\tilde{\sigma}^2$. This leads to

Lemma 4.4.5. *Under assumption (C2) the mean and variance of $\tilde{B}(T, T_1, \dots, T_n)$ equal*

$$\begin{aligned}\tilde{m} &= \frac{\bar{B}(0,T_n)}{\bar{B}(0,T)} \exp(-\int_T^{T_n}\int_0^T \bar{c}(v,u,v)\,dv\,du) \\ &\quad + K \sum_{i=1}^n \frac{B^0(0,T_i)}{B^0(0,T)} \exp(-\int_T^{T_i}\int_0^T c^0(v,v,u)\,dv\,du), \\ \tilde{\sigma}^2 &= K^2 \sum_{i,j=1}^n A_{ij} + K \sum_{i=1}^n B_i + C,\end{aligned}$$

where A_{ij}, B_i and C have the explicit expressions (4.20), (4.21) and (4.22), respectively.

Proof. According to assumption (C2),

$$\tilde{B}(T, T_1, \dots, T_n) = e^{\xi},$$

with normally distributed ξ. We want to match the moments and thereby compute mean and variance of ξ. Note that

$$\tilde{m} = \exp[\mathbb{E}\xi + \frac{1}{2}\operatorname{Var}\xi]$$

and therefore $\mathbb{E}\xi = \ln \tilde{m} - \frac{1}{2}\operatorname{Var}\xi$. This leads to

$$\begin{aligned}\tilde{\sigma}^2 &= \exp[2\mathbb{E}\xi + \operatorname{Var}\xi]\big(\exp(\operatorname{Var}\xi) - 1\big) \\ &= \tilde{m}^2\big[\exp(\operatorname{Var}\xi) - 1\big].\end{aligned}$$

Thus,

$$\xi \sim \mathcal{N}\Big(\ln \frac{\tilde{m}}{\sqrt{1 + \frac{\tilde{\sigma}^2}{\tilde{m}}}}, \ln\big(\frac{\tilde{\sigma}^2}{\tilde{m}^2} + 1\big)\Big).$$

Further on, we derive explicit expressions for $\tilde{m}$ and $\tilde{\sigma}^2$. First,

$$\tilde{m} = \mathbb{E}\Big[\exp(-\int_T^{T_n} \bar{f}(T,u)\,du) + K \sum_{i=1}^n \exp(-\int_T^{T_i} f^0(T,u)\,du)\Big].$$

with

$$\begin{aligned}&\mathbb{E}\Big[\exp(-\int_T^{T_n} \bar{f}(T,u)\,du)\Big] \\ &\qquad = \exp\big[-\int_T^{T_n} \bar{\mu}(T,u)\,du + \frac{1}{2}\int_T^{T_n}\int_T^{T_n} \bar{c}(T,u,v)\,dv\,du\big] \\ &\qquad = \frac{\bar{B}(0,T_n)}{\bar{B}(0,T)} \exp(-\int_T^{T_n}\int_0^T \bar{c}(v,u,v)\,dv\,du)\end{aligned}$$

and, if the drift condition for zero recovery bonds is also satisfied,

$$\mathbb{E}\Big[\exp(-\int_T^{T_i} f^0(T,u)\,du)\Big] \;=\; \frac{B^0(0,T_i)}{B^0(0,T)}\,\exp(-\int_T^{T_i}\int_0^T c^0(v,v,u)\,dv\,du).$$

So the assertion on $\tilde{m}$ follows.

The variance of $\tilde{B}(T,T_1,\dots,T_n)$ thus becomes

$$\tilde{\sigma}^2 \;=\; \mathrm{Var}\Big[K\sum_{i=1}^{n}\exp(-\int_T^{T_i} f^0(T,u)\,du)+\exp(-\int_T^{T_n}\bar{f}(T,u)\,du)\Big]$$

which may be simplified to

$$\begin{aligned}\tilde{\sigma}^2 \;=\;& K^2\sum_{i,j=1}^{n}\mathrm{Cov}\,\big(\exp(-\int_T^{T_i} f^0(T,u)\,du),\exp(-\int_T^{T_j} f^0(T,u)\,du)\big)\\ &+2K\sum_{i=1}^{n}\mathrm{Cov}\,\big(\exp(-\int_T^{T_i} f^0(T,u)\,du),\exp(-\int_T^{T_n}\bar{f}(T,u)\,du)\big)\\ &+\mathrm{Var}\,\big(\exp(-\int_T^{T_n}\bar{f}(T,u)\,du)\big).\end{aligned}$$

Therefore $\tilde{\sigma}^2$ has the form $K^2\sum_{i,j=1}^{n}A_{ij}+2K\sum_{i=1}^{n}B_i+C$ and, using equation (B.8), we obtain

$$\begin{aligned}A_{ij} =& \frac{B^0(0,T_i)B^0(0,T_j)}{[B^0(0,T)]^2}\,\exp(-\int_T^{T_i}\int_0^T c^0(v,v,u)\,dv\,du-\int_T^{T_j}\int_0^T c^0(v,v,u)\,dv\,du)\\ &\cdot\Big\{\exp\Big(\int_T^{T_i}\int_T^{T_j}\Big[l_ul_vc(T,u,v)+\frac{l_u}{L_v}\varsigma(T,v,u)\\ &\qquad+\frac{l_v}{L_u}\varsigma(T,u,v)+\frac{1}{L_uL_v}\bar{c}(T,u,v)\Big]\,dv\,du\Big)-1\Big\}.\end{aligned}\tag{4.20}$$

Similarly we obtain

$$B_i=\frac{\bar{B}(0,T_n)B^0(0,T_i)}{\bar{B}(0,T)B^0(0,T)}\,\exp(-\int_T^{T_n}\int_0^T\bar{c}(v,u,v)\,dv\,du-\int_T^{T_i}\int_0^T c^0(v,u,v)\,dv\,du)\tag{4.21}$$

and, finally,

$$C=\Big[\frac{\bar{B}(0,T_n)}{\bar{B}(0,T)}\Big]^2\exp(-2\int_T^{T_n}\int_0^T\bar{c}(v,u,v)\,dv\,du)\cdot\big[\exp(\int_T^{T_n}\int_T^{T_n}\bar{c}(T,u,v)\,du\,dv)-1\big].\tag{4.22}$$

■

We introduce an auxiliary product, which we call the *converting bond*, $B^C(t,T,T')$. It is used as an abbreviation in the pricing formula for the swaption, and an explicit formula for its price is available. The converting bond is a derivative which pays 1 at maturity T' if no default occurred until $T < T'$. Thus, it behaves like a zero recovery bond until T and is converted into a default-free bond at T, if no default occurred so far.

Denoting

$$B^C(t,T,T') := \mathbb{E}_t\Big[\exp(-\int_t^T \lambda_u\,du - \int_t^{T'} r_u\,du)\Big]$$

we obtain the following

Lemma 4.4.6. *The price of the converting bond at time $t=0$ equals for $0 \le T < T' \le T^*$*

$$\begin{aligned}
B^C(0,T,T') = B(0,T')\exp\Big[&-\int_0^T \frac{\bar{f}(0,u) - f(0,u)}{L_u}\,du \\
&+ \frac{1}{2}\int_0^T\!\!\int_0^T \frac{1}{L_u}\Big(\big[l_v\bar{c}(u\wedge v,u,v) + (1+\frac{1}{L_v})c(u\wedge v,u,v)\big] \\
&\qquad - \frac{1}{L_uL_v}\big[\varsigma(u\wedge v,u,v) + \varsigma(u\wedge v,v,u)\big]\Big)\,du\,dv \\
&+ \int_0^T\!\!\int_0^{T'} \frac{\varsigma(u\wedge v,u,v) - c(u\wedge v,u,v)}{L_u}\,dv\,du\Big].
\end{aligned}$$

Proof. Setting $t=0$ we have

$$\begin{aligned}
&\mathbb{E}\Big[\exp(-\int_0^T \lambda_u\,du - \int_0^{T'} r_u\,du)\Big] \\
&= B(0,T')\mathbb{E}\big[\exp(-\int_0^T \lambda_u\,du)\big]\exp\big[\operatorname{Cov}(\int_0^T \lambda_u\,du, \int_0^T r_u\,du)\big] \\
&= B(0,T')\exp\Big[-\int_0^T \frac{\bar{\mu}(u,u) - \mu(u,u)}{L_u}\,du \\
&\qquad + \frac{1}{2}\int_0^T\!\!\int_0^T \frac{1}{L_uL_v}\big[\bar{c}(u\wedge v,u,v) + c(u\wedge v,u,v) \\
&\qquad\qquad - \varsigma(u\wedge v,u,v) - \varsigma(u\wedge v,v,u)\big]\,du\,dv\Big] \\
&\quad \cdot \exp(\int_0^T\!\!\int_0^{T'} \frac{\varsigma(u\wedge v,u,v) - c(u\wedge v,u,v)}{L_u}\,dv\,du).
\end{aligned}$$

Using the drift conditions we may conclude

$$
\begin{aligned}
B^C(0,T,T') = B(0,T')\exp\Big\{ &-\int_0^T \frac{\bar{f}(0,u)-f(0,u)}{L_u}\,du \\
&+\frac{1}{2}\int_0^T\!\!\int_0^T \frac{1}{L_u}\Big[\big[l_v\bar{c}(u\wedge v,u,v)+(1+\frac{1}{L_v})c(u\wedge v,u,v)\big] \\
&\qquad -\frac{1}{L_uL_v}\big[\varsigma(u\wedge v,u,v)+\varsigma(u\wedge v,v,u)\big]\Big]\,du\,dv \\
&+\int_0^T\!\!\int_0^{T'} \frac{\varsigma(u\wedge v,u,v)-c(u\wedge v,u,v)}{L_u}\,dv\,du\Big\}. \qquad\blacksquare
\end{aligned}
$$

Under further assumptions this formula can be simplified considerably. For example, if $L_u \equiv L$

$$
\exp\Big[-\int_0^T \frac{\bar{f}(0,u)-f(0,u)}{L_u}\,du\Big] \;=\; \Big[\frac{\bar{B}(0,T)}{B(0,T)}\Big]^{\frac{1}{L}}.
$$

Recalling the definitions of $\tilde{m}$ and $\tilde{\sigma}^2$ from Lemma 4.4.5, we obtain

Theorem 4.4.7. *Under the assumptions (C1) and (C2) the price of a CDS call equals*

$$
S_C(0,T,T_n) \;=\; B^C(0,T,T_n)\Phi(-d_2)-\big[B^k(0,T,T_n)+K\sum_{i=1}^n B^0(0,T_n)\big]\Phi(-d_1),
$$

with deterministic

$$
\begin{aligned}
\mu_1 &:= \tilde{m}+\frac{B(0,T)}{B(0,T_n)}+\int_T^{T_n}\int_0^u c(v,u,v)\,dv\,du, \\
\sigma_1 &:= \ln\big[\frac{\tilde{\sigma}^2}{\tilde{m}^2}+1\big]+\int_T^{T_n}\!\!\int_T^{T_n} c(T,u,v)\,du\,dv-\big[\tilde{m}+\frac{\tilde{\sigma}^2}{2}\big] \\
&\quad+\ln\Big[\frac{\bar{B}(0,T_n)}{\bar{B}(0,T)}\exp\Big[-\int_T^{T_n}\!\!\int_0^T \bar{c}(v,u,v)\,dv\,du-\int_T^{T_n}\!\!\int_T^{T_n}\varsigma(T,u,v)\,dv\,du\Big] \\
&\quad+K\sum_{i=1}^n\frac{B^0(0,T_i)}{B^0(0,T)}\exp\Big[-\int_T^{T_i}\!\!\int_0^T c^0(v,u,v)\,dv\,du \\
&\qquad\qquad-\int_T^{T_i}\!\!\int_T^{T_n} l_Tc(T,u,v)+\frac{\varsigma(T,u,v)}{L_T}\,dv\,du\Big]\Big],
\end{aligned}
$$

$$\sigma_2 := \int_0^{T_n}\int_0^{T_n} l_2(u,T) l_2(v,T) c(u\wedge v,u,v)\,dv\,du + \int_0^{T}\int_0^{T} \frac{\bar{c}(u\wedge v,u,u)}{L_u L_v}\,dv\,du$$
$$+\,2\int_0^{T_n}\int_0^{T} \frac{l_2(u,T)}{L_v}\varsigma(u\wedge v,v,u)\,dv\,du,$$
$$d_2 := \frac{\mu_1 - \ln K}{\sigma_1} + \rho\sigma_2,$$
$$d_1 := d_2 + \sigma_1,$$
$$l_2(u,T) := \begin{cases} -l_u & \text{for } u \le T,\\ 1 & \text{for } u > T.\end{cases}$$

In equation (B.6) an explicit expression is given for

$$\rho := \mathrm{Cov}\big[\ln\frac{\tilde{B}(T,T_1,\ldots,T_n)}{B(T,T_n)}, -\int_0^T r_u + \lambda_u\,du - \ln B(T,T_n)\big].$$

Proof. Using the abbreviation $B^C(0,T,T_n)$, see Lemma 4.4.6, we obtain

$$\mathrm{I\!E}\big[\exp(-\int_0^T r_u+\lambda_u\,du)\,(\bar{B}(T,T_n) + K\sum_{i=1}^n B^0(T,T_i))\big] = B^C(0,T,T_n) + K\sum_{i=1}^n B^0(0,T_i).$$

The discounted payoffs of the swaption were already derived in equation (4.18). Setting

$$\begin{aligned}\xi_1 &= \ln\frac{\tilde{B}(T,T_1,\ldots,T_n)}{B(T,T_n)}\\ \xi_2 &= -\int_t^T r_u+\lambda_u\,du + \int_T^{T_n} f(T,u)\,du\end{aligned}$$

we can use formula (B.5) to compute the expectation of the discounted payoffs. We need to derive the constants, thus

$$\begin{aligned}\mu_1 &= \mathrm{I\!E}\Big[\ln\frac{\tilde{B}(T,T_1,\ldots,T_n)}{B(T,T_n)}\Big]\\ &= m + \int_T^{T_n}\mu(0,u) + \int_0^u c(v\wedge T,u,v)\,dv\,du\end{aligned}$$

The variances are

$$\begin{aligned}
\sigma_1 &= \mathrm{Var}\Big[\ln\frac{\tilde{B}(T,T_1,\dots,T_n)}{B(T,T_n)}\Big] \\
&= \ln\big[\frac{\sigma^2}{m^2}+1\big] + \int_T^{T_n}\int_T^{T_n} c(T,u,v)\,du\,dv \\
&\quad + \mathrm{Cov}\Big(\ln\tilde{B}(T,T_1,\dots,T_n), -\int_T^{T_n} f(T,u)\,du\Big) \\
&= \ln\big[\frac{\sigma^2}{m^2}+1\big] + \int_T^{T_n}\int_T^{T_n} c(T,u,v)\,du\,dv \\
&\quad + \ln\Big[\frac{\bar{B}(0,T_n)}{\bar{B}(0,T)}\exp\big[-\int_T^{T_n}\int_0^{T}\bar{c}(v,u,v)\,dv\,du - \int_T^{T_n}\int_T^{T_n}\varsigma(T,u,v)\,dv\,du\big] \\
&\quad +K\sum_{i=1}^{n}\frac{B^0(0,T_i)}{B^0(0,T)}\exp\big[-\int_T^{T_i}\int_0^{T} c^0(v,u,v)\,dv\,du \\
&\qquad\qquad - \int_T^{T_i}\int_T^{T_n} l_T c(T,u,v)+\frac{\varsigma(T,u,v)}{L_T}\,dv\,du\big]\Big] \\
&\quad -\big[m+\frac{\sigma^2}{2}\big]
\end{aligned}$$

and

$$\begin{aligned}
\sigma_2 &= \mathrm{Var}\big[-\int_0^{T_n} l_2(u,T)X(u\wedge T,u)\,du + \int_0^{T}\frac{\bar{X}(u)}{L_u}\,du\big] \\
&= \int_0^{T_n}\int_0^{T_n} l_2(u,T)l_2(v,T)c(u\wedge v\wedge T,u,v)\,dv\,du + \int_0^{T}\int_0^{T}\frac{\bar{c}(u\wedge v,u,v)}{L_uL_v}\,dv\,du \\
&\quad -2\int_0^{T_n}\int_0^{T}\frac{l_2(u,T)}{L_v}\varsigma(u\wedge v,v,u)\,dv\,du.
\end{aligned}$$

The assertion now follows using Lemma B.4.4. ■

If one prefers to use the replacement of the "difference to par" instead of the "difference to an equivalent risk-free bond", pricing formulas are obtained proceeding similarly.

4.4.5 Hedging - an Example

Blanchet-Scalliet and Jeanblanc (2001) introduced a hedging methodology for derivatives on underlyings which bear credit risk. Their approach concentrates on derivatives which promise a riskless contingent claim X_T if no default occurred before T. In this framework it is essential that X_T can be replicated on the riskless market.

The above considered derivatives often incorporate a similar knock-out feature, but in most cases the payoff cannot be replicated on the riskless market. For example, consider a call on a zero recovery bond with strike K, offering at T

$$\Big(B^0(T,T') - K\Big)^+.$$

If a default occurred before T, the call is worthless. This is very similar to a knock-out feature, but the call also can become worthless if the value of $B^0(T,T')$ drops below K.

First, we derive an explicit pricing formula for the call option and afterwards suggest a hedging scheme.

Theorem 4.4.8. *The price of a call with maturity $T \in [0,T^*]$ on a zero recovery bond with strike K and maturity $T' \in (T,T^*]$ equals*

$$C(0,T,T) = B^0(0,T')\Phi(d_1) - K\, B^0(0,T)\Phi(d_2), \tag{4.23}$$

with

$$\begin{aligned}
\sigma^2(T,T') &:= \int_T^{T'}\!\!\int_T^{T'} c^0(T,u,v)\,dv\,du,\\
d_2 &:= \frac{\ln \frac{B(0,T')}{KB(0,T)}}{\sigma(T,T')} - \frac{1}{2}\sigma(T,T'),\\
d_1 &:= d_2 + \sigma(T,T').
\end{aligned}$$

Proof. The risk-neutral valuation principle yields

$$\begin{aligned}
C(0,T,T') &= \mathbb{E}\Big(\exp(-\int_0^T r_u\,du)\big[B^0(T,T') - K\big]^+\Big)\\
&= \mathbb{E}\Big(\exp(-\int_0^T r_u\,du)1_{\{B^0(T,T')>K\}}B^0(T,T')\Big)\\
&\quad - K\mathbb{E}\Big(\exp(-\int_0^T r_u\,du)1_{\{B^0(T,T')>K\}}\Big)\\
&=: (1) - (2).
\end{aligned}$$

Observe that $\{B^0(T,T') > K\} \subset \{\tau > T\}$. Hence

$$
\begin{aligned}
(2) &= K\,\mathbb{E}\big(\exp(-\int_0^T r_u\,du)1_{\{B^0(T,T')>K,\tau>T\}}\big)\\
&= K\,\mathbb{E}\big(\exp[-\int_0^T (r_u+\lambda_u)\,du]1_{\{\int_T^{T'} f^0(T,u)\,du>-\ln K\}}\big)\\
&= K\,B^0(0,T)\,Q^0\big(\int_T^{T'} f^0(T,u)\,du > -\ln K\big),
\end{aligned}
$$

where Q^0 denotes the zero recovery measure, see (4.15). According to Girsanov's Theorem B.4.1, the integral $\int_T^{T'} f^0(T,u)\,du$ is normally distributed under Q^0 with the same variance as under Q, namely

$$
\sigma^2(T,T') \quad := \quad \mathrm{Var}\,\big[\int_T^{T'} f^0(T,u)\,du\big] = \int_T^{T'}\int_T^{T'} c^0(T,u,v)\,dv\,du.
$$

Proceeding similarly we obtain for (1):

$$
\begin{aligned}
(1) &= \mathbb{E}\big(\exp[-\int_0^T (r_u+\lambda_u)\,du]\cdot\exp(-\int_T^{T'} f^0(T,u)\,du)1_{\{-\int_T^{T'} f^0(T,u)\,du>\ln K\}}\big)\\
&= B^0(0,T)\mathbb{E}^0\big(\exp(-\int_T^{T'} f^0(T,u)\,du)1_{\{-\int_T^{T'} f^0(T,u)\,du>\ln K\}}\big).
\end{aligned}
$$

Applying Lemma B.4.2 we may conclude

$$
\begin{aligned}
(1) &= B^0(0,T)\mathbb{E}^0\big(\exp(-\int_T^{T'} f^0(T,u)\,du)\big)Q^0\big(-\int_T^{T'} f^0(T,u)\,du > \ln K - \sigma^2(T,T')\big)\\
&= B^0(0,T')Q^0\big(-\int_T^{T'} f^0(T,u)\,du > \ln K - \sigma^2(T,T')\big).
\end{aligned}
$$

Finally, note that

$$
d_2 = -\frac{\ln K + \mu(T,T') + \sigma^2(T,T')}{\sigma(T,T')} = -\frac{\ln\frac{B(0,T')}{KB(0,T)}}{\sigma(T,T')} + \frac{1}{2}\sigma(T,T').
$$

$\blacksquare$

The explicit pricing formula for the default bond option admits an explicit derivation of the hedging strategy.

The price of the option (4.23) depends on two different securities, $B^0(0,T)$ and $B^0(0,T')$. Naturally the hedge consists in trading in these two assets.

Observe that the call price is a continuous function of $B^0(0,T)$ and $B^0(0,T')$ until default. Therefore, the delta-hedging methodology can be applied.

For the first part of the hedge we have

$$\begin{aligned}\Delta_1(s) &:= \frac{\partial C(0,T,T')}{\partial B^0(0,T)} \\ &= \Phi(d_1) + B^0(0,T)\phi(d_1)\frac{\partial d_1}{\partial B^0(0,T')} - KB^0(0,T)\phi(d_2)\frac{\partial d_2}{\partial B^0(0,T')}.\end{aligned}$$

Writing σ for $\sigma(s,t,T)$, we obtain

$$\begin{aligned}\frac{\partial d_{1/2}}{\partial B^0(0,T')} &= \frac{1}{B(0,T')\,\sigma}, \\ \phi(d_1) &= \frac{1}{\sqrt{2\pi}}\exp\Big[-\frac{d_2^2+2d_2\sigma+\sigma^2}{2}\Big] = \phi(d_2)\exp(-\sigma d_2 - \frac{\sigma^2}{2}) \\ &= \phi(d_2)\frac{K\,B^0(0,T)}{B^0(0,T')},\end{aligned}$$

which yields

$$\begin{aligned}\Delta_1(s) &= \Phi(d_1) + \frac{\phi(d_2)}{B^0(0,T')\sigma}\Big[B^0(0,T')\frac{KB^0(0,T)}{B^0(0,T')} - KB^0(0,T)\Big] \\ &= \Phi(d_1). \end{aligned} \tag{4.24}$$

Similarly, we obtain for the hedge w.r.t. $\bar{B}(s,t)$:

$$\begin{aligned}\Delta_2(s) &= B^0(0,T')\phi(d_1)\frac{\partial d_1}{\partial B^0(0,T)} - K\Phi(d_2) - KB^0(0,T)\phi(d2)\frac{\partial d_2}{\partial B^0(0,T)} \\ &= -K\Phi(d_2).\end{aligned}$$

Altogether, the hedge is perfect because up to the jump of the underlying, we have

$$dC(\bar{B}(s,T),\bar{B}(s,t),K) = \Delta_1(s)d\bar{B}(s,T) + \Delta_2(s)d\bar{B}(s,t)$$

and at the discontinuity τ the value of the call and the value of the hedging portfolio both jump to zero.

Note the analogy to the Deltas in the Black-Scholes formula[7].

[7] See, for example Hull (1993).

4.5 Calibration

In this chapter we present two approaches how to calibrate a Gaussian random field model to market data. This is motivated by the results of Pang (1998), who shows that in the interest rate case the calibration of a random field model in comparison to a n-factor HJM model permits more stability over time and frequent re-calibration can be avoided. This is due to the different approaches specifying the number of significant factors.

In n-factor models, n is pre-specified by some reasoning and then the calibration is carried out. In contrast, in random field models, n is specified during the calibration, such that the error of the n-dimensional approximation does not exceed a certain level. Thus, the latter method allows choosing n depending on the data and the required precision.

If we want to avoid assuming a parametric covariance structure as in Kennedy (1997), a relatively large data set needs to be available. We therefore assume that prices of credit default swaps and swaptions are accessible. Nowadays these options are not yet traded liquidly, but as the credit market is increasing rapidly, it is just a question of time until they will be available.

4.5.1 Calibration Using Gaussian Random Fields

As in Pang (1998) we make the following assumptions:

(i) We assume that the riskless model is already calibrated.

(ii) The covariance functions satisfy

$$
\begin{aligned}
\bar{c}(s, t_1, t_2) &= \int_0^s \bar{g}(t_1 - u, t_2 - u)\, du, \\
\varsigma(s, t_1, t_2) &= \int_0^s g(t_1 - u, t_2 - u)\, du.
\end{aligned}
$$

(iii) Furthermore, the surfaces $\bar{g} : \mathbb{R}^2 \mapsto \mathbb{R}$ and $g : \mathbb{R}^2 \mapsto \mathbb{R}$ are piecewise triangular: For nodes $\{u_1, \dots, u_m\}$ any $(u_i, u_i), (u_{i+1}, u_i), (u_{i+1}, u_{i+1})$ or $(u_i, u_i), (u_i, u_{i+1})$, (u_{i+1}, u_{i+1}) define the corners of the surfaces' triangles.

The second assumption yields stationary volatility factors, while the third assumption allows for quick calibration of the covariance function. The $\{u_1, \dots, u_m\}$ do not necessarily coincide with the tenor structure, denoted by $\{T_1, \dots, T_n\}$. For example, in Pang (1998) the t_i are multiples of 0.25 while the tenor structure is $\{1, 2, 3, 5, 7, 10\}$.

If we want to ensure that the term structure of forward rates is continuous or smooth, we would have to assume continuity and boundedness of the covariance function (respectively their second derivative), which is violated by the second assumption. Nevertheless, rounding the edges yields little differences in derivatives' prices and suitable regularity of the forward rates.

For the calibration we would use data of a certain time period, say some weeks or a month, and use standard optimization software to minimize, for example S-Plus with the function `nlmin`, the residual sum of squared differences between the calculated prices and market prices. In this procedure, calculating model prices is done in two steps. First, determine $c(s, t_1, t_2)$ and $\bar{c}(s, t_1, t_2)$ on the basis of $g(u, v)$ and $\bar{g}(u, v)$ for $u, v \in \{u_1, \ldots, u_m\}$, $t_1, t_2 \in \{T_1, \ldots, T_n\}$ and every considered data time $s \in \{s_1, \ldots, s_p\}$. For the second step, the prices of the considered derivatives are computed using the $c(s, t_1, t_2)$ and $\bar{c}(s, t_1, t_2)$ determined in the first step.

4.5.2 Calibration Using the Karhunen-Loève Expansion

An alternative calibration method uses the Karhunen-Loève decomposition[8] inspired by Vargiolu (2000).

The approach presented in this section incorporates a mixture between historical estimation and calibration to actual market data. This has the advantage that on one side the procedure profits from useful historical information, while on the other side the requirements of traders, that a model should calibrate perfectly to market prices, is fulfilled.

The data problem in calibration issues of credit risk models has been addressed in several papers, for example Schönbucher (2002). It therefore seems beneficial that the proposed procedure parsimoniously uses the available data.

Central parameters of our Gaussian model are the covariance functions. With a view towards applications, the flexibility provided by the model encounters the problem that the data for calibration issues is still scarce. In the following we present an intermediary solution which serves both needs.

Consider the covariance function $\bar{c}(s, t_1, t_2)$, where we set $s = 0$. Then $\bar{c}(\cdot)$ can be decomposed into

$$c(0, t_1, t_2) = \sum_k \lambda_k e_k(t_1) e_k(t_2),$$

using an orthonormal basis $\{e_k : k \in \mathbb{N}\}$ of $L^2(\mu)$, the Hilbert space of functions $f : \mathbb{R} \mapsto \mathbb{R}$ which are square integrable w.r.t. a suitable measure μ. For our application a certain period of maturities will be of interest, for example the interval $[0, T^*]$, and we choose for μ the Lebesgue measure.

Note that, to determine the covariance function, one has to specify both the $\{e_k : k \in \mathbb{N}\}$ and the $\{\lambda_k : k \in \mathbb{N}\}$. For the former, we use historical information, while the latter are obtained via calibration.

[8]See, for example, Bogachev (1991, p. 55 p.p.), Da Prato and Zabczyk (1992, p. 99 p.p.) or Adler (1981).

The first step is to estimate the covariance function using a set of historical data. Consider a small time interval, so that stationarity of the considered random fields in this time interval may be assumed. The historical data consists of observations of $\bar{f}(s,t)$ at a set of time points $\mathcal{T} := \{(s_i, t_j) : 1 \le i \le n_1, 1 \le j \le n_2\}$. Following Hall, Fisher and Hoffmann (1994) we suggest an estimator based on *kernel methods*. For the points $\mathbf{a} = (s_1, t_1)$ and $\mathbf{b} = (s_2, t_2)$ we define the covariance estimator by

$$\tilde{c}(\mathbf{a}, \mathbf{b}) := \frac{\sum_{\mathbf{c}_i, \mathbf{d}_j \in \mathcal{T}} K\left(\frac{\mathbf{a}-\mathbf{c}_i}{h}, \frac{\mathbf{b}-\mathbf{d}_j}{h}\right) \cdot \left[X(\mathbf{c}_i) - \bar{X}\right] \left[X(\mathbf{d}_j) - \bar{X}\right]}{\sum_{i,j \in \mathcal{T}} K\left(\frac{\mathbf{a}-\mathbf{c}_i}{h}, \frac{\mathbf{b}-\mathbf{d}_j}{h}\right)},$$

where $K(\mathbf{c}, \mathbf{d})$ is a symmetric kernel Observe that the sum is over all time points in $\mathcal{T}$, labeled $\mathbf{c}_i$ and $\mathbf{d}_j$, respectively. Estimation of the covariance function $\tilde{\rho}(t_1, t_2)$ at a certain time s is thus obtained by considering $s_1 = s_2 = s$.

The following second step is optional, but ensures that the estimator is positive definite, thus a covariance function itself. This yields increased performance for the eigenvector decomposition below. We invert the characteristic function of our estimator,

$$\varphi(\boldsymbol{\lambda}) := \int_{\mathbb{R}^2} \exp(\mathrm{i}\boldsymbol{\lambda}^\top \mathbf{t}) \tilde{\rho}(\mathbf{t})\, d\mathbf{t} \text{ for } \boldsymbol{\lambda} \in \mathbb{R}^2.$$

Because the estimator is symmetric, we have

$$\varphi(\boldsymbol{\lambda}) = \int \cos(\boldsymbol{\lambda}^\top \mathbf{t}) \tilde{\rho}(\mathbf{t})\, d\mathbf{t}.$$

Following Bochner's theorem, we need $\varphi(\boldsymbol{\lambda}) \ge 0$ to ensure that $\tilde{\rho}$ is a covariance function, thus we use the positive part of $\varphi(\boldsymbol{\lambda})$ in the inversion of the Fourier transform and suggest the following estimator of the covariance function

$$\hat{\rho}(\mathbf{t}) = \frac{1}{(2\pi)^2} \int \cos(\boldsymbol{\lambda}^\top \mathbf{t}) [\varphi(\boldsymbol{\lambda})]^+ \, d\boldsymbol{\lambda}.$$

Figure 4.2 shows the result of the covariance estimation on a set of U.S. Treasury data using historical data of 4 weeks. The implementation uses a Gaussian kernel and the covariance estimator is plotted for maturities of 3 months to 3 years.

After obtaining an estimator for the covariance function, we can calculate its Eigenfunctions up to a required precision. Vargiolu (2000) presents a recursive scheme to obtain the Eigenfunctions from the covariance operator[9]. We apply the procedure to our setting within Gaussian random fields.

The eigenvector decomposition is done applying the Mises-Geiringer iteration procedure to our setting, cf. Rutishauser (1976) for the application to $\mathbb{R}^n$. Fix $k_0 \in L^2(\mu)$ and define[10]

$$k^{n+1}(\cdot) := \int \hat{\rho}(\cdot, t) k^n(t)\, dt.$$

[9] As already discussed in Chapter 3, on page 78.

[10] Compare to equation (4.2).

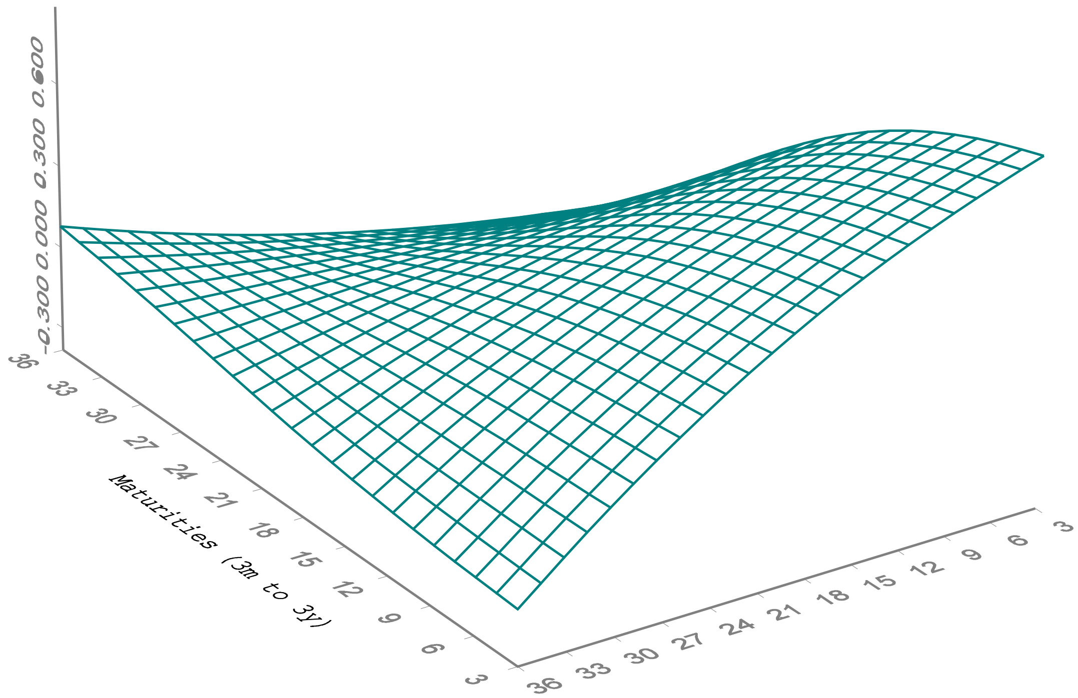

Figure 4.2: Estimated covariance function for U.S. Treasury data (May 2002). The estimation uses a Gaussian kernel and shows maturities of $3, 6, \ldots, 36$ months.

Then k^{n+1} itself is an element of $L^2(\mu)$. The results of Vargiolu (2000) can be used to show that

$$k^n \to e_1 \qquad \text{and} \qquad \frac{\| k^{n+1} \|}{\| k^n \|} \to \lambda_1, \qquad \text{as } n \to \infty.$$

As these Eigenvectors need not be normed, we introduce the normed eigenvectors $\bar{e}_k$. Using $\hat{\rho}_1(t_1, t_2) := \hat{\rho}(t_1, t_2) - \lambda_1 \bar{e}_1(t_1) \bar{e}_1(t_2)$, and applying the procedure to $\hat{\rho}_1$ yields e_2 and λ_2 and so on.

For the application one might want to obtain the covariance function on a certain grid, thus readily available implementations for matrices may be used after a suitable transformation[11].

Figure 4.3 shows the calculated Eigenvectors for the U.S. Treasury data. The first two Eigenvectors show significant Eigenvalues (3.4224 and 0.0569), while the remaining Eigenvalues are of much smaller magnitude. In this example it therefore turns out to be sufficient to use the first two Eigenvectors only.

More generally, assume that we already have determined the first N Eigenfunctions. Then we use the following covariance function for the calibration:

$$\hat{\rho}(\lambda_1, \ldots, \lambda_N, t_1, t_2) := \sum_{k=1}^{N} \lambda_k e_k(t_1) e_k(t_2).$$

[11] Note that the scalar product in $\mathbb{R}^n$ and $\mathrm{L}^2(\mu)$ is different.

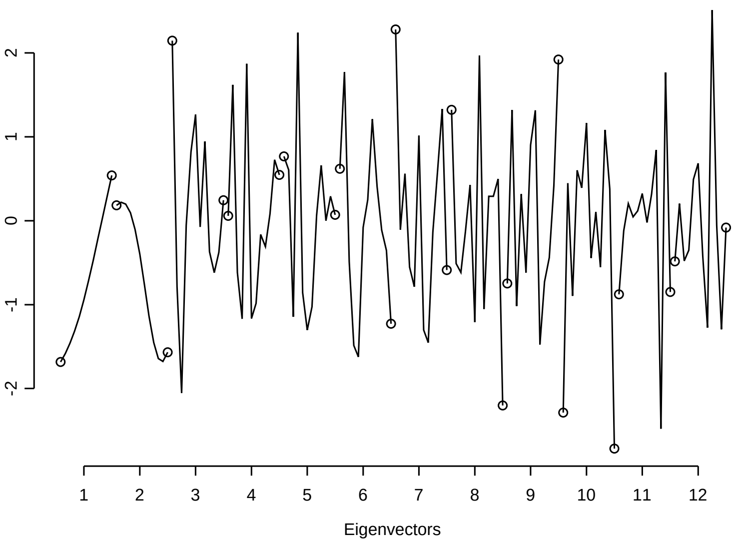

Figure 4.3: Estimated, normed Eigenvectors of the covariance function in Figure 4.2. The first two Eigenvectors correspond to the Eigenvalues 3.4224 and 0.0569, respectively, while the further are of magnitude 10^{-15}.

As before, a standard software package can be used to extract the $\lambda_1, \ldots, \lambda_N$ from observable derivatives prices by a least-squares approach. Note that, in comparison to the previously presented model, a much smaller set of derivatives can be used for the calibration.

The implementation of this last step using credit derivatives data is subject to future research.

Nevertheless, we already analyzed some bond data and estimated the covariance functions and the Eigenvectors / values. Take for example the data from Greece Treasury bonds. The estimation results may be found in Figures 4.5 and 4.6. First, note that the variance for bonds with small maturities is higher than for bonds with large maturities. Second, for the period June to August 2001 negative correlations for bonds with small versus bonds with large maturities were observed. This reflects a movement in opposite direction as to interest rates in this period.

Taking a closer look at the Eigenvectors reveals the components of the covariance function. The first Eigenvector generates more or less the shape of the covariance functions. The already mentioned effect, that larger maturities relate to smaller variances, may be observed here as well. Note that the scale of the z-Axis changes (Max 0.010 to 0.048) with the first Eigenvalue (0.03318 to 0.1371). The second Eigenvector covers the wriggly structure of the covariance function. Note that this is in strong relation with the Eigenvalues.

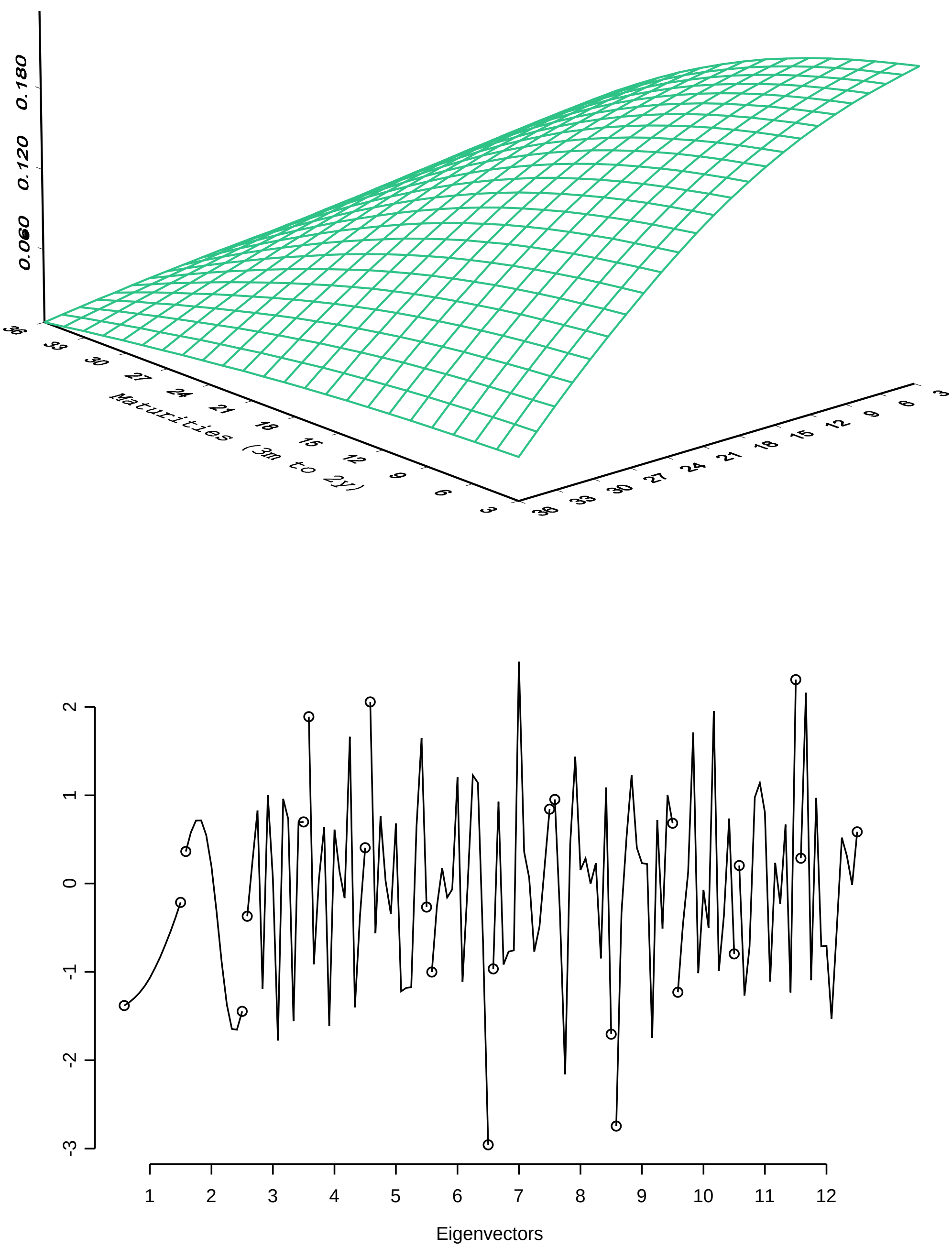

Figure 4.4: Estimated covariance function and normed Eigenvectors for U.S. Treasury data, July - September 2001. The first two Eigenvalues are $1.3658, -0.001186$, the remaining ones being of much smaller magnitude (10^{-9}).

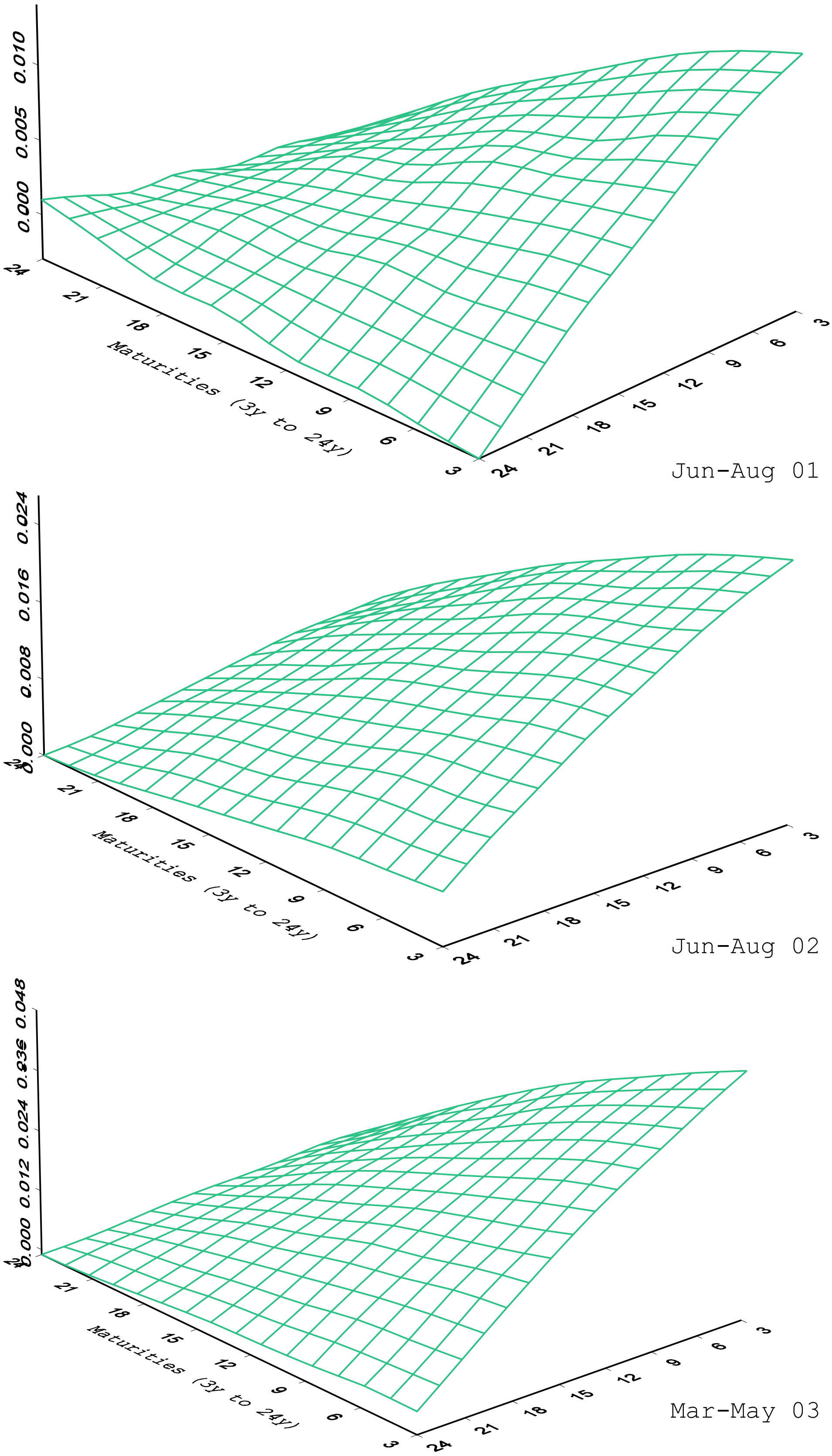

Figure 4.5: Estimated Covariance function for Greece Treasury data. The plots are based on 40 observations of Bonds with maturities $3, 5, 7, 10, 15, 20, 30$ years. The covariance function is plotted for maturities $T = 3, 6, 9, \ldots, 24$ years.

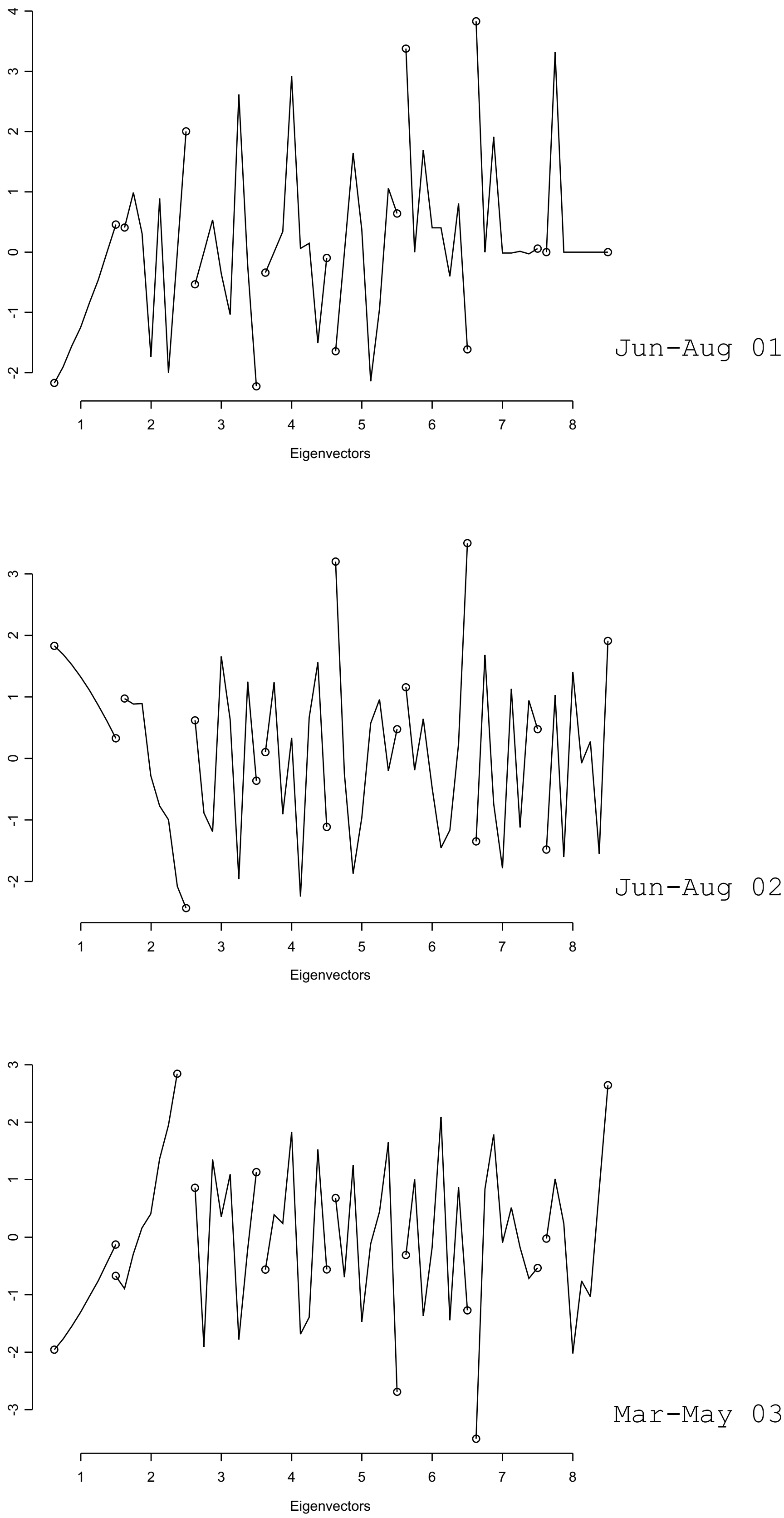

Figure 4.6: Estimated Eigenfunctions for Greece Treasury data. For the first plot Eigenvalues are $0.03318, -0.0009484, -0.0005508, -0.0002444, -0.0001.66$ (the others being of magnitude 10^{-18}). For the second $0.08929, -0.00233, 0.00088127, -0.006616$ (others: 10^{-5}) and for the third $0.1371, -0.00205, 0.0008344, -0.0008202$ (others: 10^{-5}).

Appendix A

Basic Setup for Hazard Rate Models

A detailed treatment of proofs and methods within the hazard rate framework can be found, among others, in Lando (1994), Bielecki and Rutkowski (2002), Jeanblanc and Rutkowski (2000) or Jeanblanc (2002).

Consider a probability space $(\Omega, \mathcal{A}, Q)$, endowed with a filtration $(\mathcal{G}_t)_{t\geq 0}$. The probability measure Q will represent a risk-neutral measure, which is fundamental in pricing contingent claims. The filtration $(\mathcal{G}_t)_{t\geq 0}$ represents the general market information, which could include information on certain indices, interest rates and so on.

Introducing default risk into the model, we consider a default time τ, which is a positive random variable on $(\Omega, \mathcal{A}, Q)$. The associated jump process $1_{\{\tau\leq t\}}$ induces the "default information" represented by $\mathcal{H}_t := \sigma(1_{\{\tau\leq s\}} : 0 \leq s \leq t)$. Therefore, the total information available at time t is $\mathcal{F}_t = \mathcal{H}_t \vee \mathcal{G}_t$.

In hazard-rate models, one uses a specific type of process for $1_{\{\tau\leq t\}}$, namely Cox processes. As some models incorporate more than on jump, we aim at defining a jump process, which jumps at times $\tau_1, \tau_2, \ldots$ and set $\tau := \tau_1$, if just one default event is of interest.

Consider a Poisson process $(\tilde{N}_t)_{t\geq 0}$ with intensity 1, which is independent of $\mathcal{G}_t$ for all t, and a nonnegative, nondecreasing and right-continuous process $(\Lambda_t)_{t\geq 0}$ adapted to $(\mathcal{G}_t)_{t\geq 0}$. We then obtain a *Cox process* through a (random) time change of the process $\tilde{N}$ by setting[1]

$$N_t := \tilde{N}_{\Lambda(t)}.$$

If $(\Lambda(t))_{t\geq 0}$ admits the representation

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

then $(\lambda(t))_{t\geq 0}$ is called the intensity of N.

The default time is represented by the first jump of $(N_t)_{t\geq 0}$, so that

$$\tau := \tau_1 = \inf\{s \geq 0 : N_s = 1\},$$

[1]For a detailed treatment on Cox processes see Grandell (1997).

while the n-th jump is $\tau_n := \inf\{s \geq 0 : N_s = n\}$.

It is easy to deduce the following

Lemma A.1.1. *For a Cox process $(N_t)_{t\geq 0}$ with intensity $(\lambda_t)_{t\geq 0}$ and τ being the first jump of (N_t), we have*

$$Q(\tau > t|\mathcal{G}_t) = \exp(-\int_0^t \lambda_u \, du).$$

For pricing a defaultable bond the following Theorem, first mentioned in Lando (1994), is indispensable:

Theorem A.1.2. *For a $\mathcal{G}_T$-measurable random variable X_T, and τ as well as $(N_t)_{t\geq 0}$ defined as in the preceding Lemma, we have that*

$$\begin{aligned} \mathbb{E}(X_T 1_{\{\tau>T\}}|\mathcal{F}_t) &= 1_{\{\tau>t\}}\mathbb{E}\big(\exp(-\int_t^T \lambda_u \, du)X_T|\mathcal{F}_t\big) \\ &= 1_{\{\tau>t\}}\mathbb{E}\big(\exp(-\int_t^T \lambda_u \, du)X_T|\mathcal{G}_t\big). \end{aligned}$$

Proof. Using the definition of the Cox process as given above yields

$$\begin{aligned} \mathbb{E}(X_T 1_{\{\tau>T\}}|\mathcal{F}_t) &= 1_{\{\tau>t\}}\mathbb{E}\big[X_T\mathbb{E}\big(1_{\{\tau>T\}}|\mathcal{F}_t \vee \mathcal{G}_T\big)|\mathcal{F}_t\big] \\ &= 1_{\{\tau>t\}}\mathbb{E}\big[X_T\mathbb{E}\big(1_{\{\tilde{N}(\Lambda_T)-\tilde{N}(\Lambda_t)=0\}}|\mathcal{F}_t \vee \mathcal{G}_T\big)|\mathcal{F}_t\big]. \end{aligned}$$

By definition of $\mathcal{F}_t$ we have $\sigma(\mathcal{F}_t \vee \mathcal{G}_T) = \sigma(\mathcal{H}_t \vee \mathcal{G}_T)$. Furthermore, since a Poisson process has independent increments, we have, conditionally on $\mathcal{G}_T$ and on $\{\tau > t\}$, that $\tilde{N}(\Lambda_T) - \tilde{N}(\lambda_t)$ is independent of $\mathcal{H}_t$. Therefore

$$\mathbb{E}\big(1_{\{\tilde{N}(\Lambda_T)-\tilde{N}(\Lambda_t)=0\}}|\mathcal{H}_t \vee \mathcal{G}_T\big) = \mathbb{E}\big(\exp[-(\Lambda_T - \Lambda_t)]|\mathcal{G}_T\big)$$

and we may conclude

$$\mathbb{E}(X_T 1_{\{\tau>T\}}|\mathcal{F}_t) = 1_{\{\tau>t\}}\mathbb{E}\big[X_T \exp(-\int_t^T \lambda_u \, du)|\mathcal{G}_t\big]. \qquad \blacksquare$$

Using the arbitrage-free pricing principle, which yields that the fair price of a contingent claim is the expectation of the discounted payoff under an equivalent martingale measure, we obtain the following formula for the defaultable bond $\bar{B}(t,T)$:

$$\begin{aligned} \bar{B}(t,T) &= \mathbb{E}\big(\exp(-\int_t^T r_u \, du)1_{\{\tau>T\}}|\mathcal{F}_t\big) \\ &= 1_{\{\tau>t\}}\mathbb{E}\big(\exp[-\int_t^T (r_u + \lambda_u) \, du]|\mathcal{G}_t\big). \end{aligned}$$

Valuing non-European claims or using different concepts of recovery, one will find the following theorem useful (see Lando (1994)):

Theorem A.1.3. *For a Cox process $(N_t)_{t\geq 0}$ with intensity $(\lambda_t)_{t\geq 0}$ and τ being the first jump of (N_t) and a stochastic process $(Y_s)_{s\geq 0}$, we have*

(i)

$$\mathbb{E}\Big[\int_t^{T\wedge\tau} \exp(-\int_t^s r_u\,du)Y_s\,ds\Big|\mathcal{F}_t\Big] = 1_{\{\tau>t\}}\mathbb{E}\Big[\int_t^T \exp[-\int_t^s (r_u+\lambda_u)\,du]Y_s\,ds\Big|\mathcal{G}_t\Big].$$

(ii)

$$\mathbb{E}\Big[\exp(-\int_t^{\tau} r_u\,du)Y_\tau 1_{\{t\leq\tau\leq T\}}\Big|\mathcal{F}_t\Big] = 1_{\{\tau>t\}}\mathbb{E}\Big[\int_t^T \exp[-\int_t^s (r_u+\lambda_u)\,du]Y_s\lambda_s\,ds\Big|\mathcal{G}_t\Big].$$

The first formula allows for pricing a payoff stream, which is continuously paid until T and stopped at τ. The second formula prices the random payoff Y_τ, which is paid at default.

Proof. For (i), observe that

$$\begin{aligned}
&\mathbb{E}\Big[\int_t^{T\wedge\tau} \exp(-\int_t^s r_u\,du)Y_s\,ds\Big|\mathcal{F}_t\Big] \\
&\quad= \int_t^T \mathbb{E}\Big[1_{\{s\leq\tau\}}\exp(-\int_t^s r_u\,du)Y_s\Big|\mathcal{F}_t\Big]\,ds \\
&\quad= \int_t^T \mathbb{E}\Big[\exp(-\int_t^s r_u\,du)Y_s\mathbb{E}\big(1_{\{s\leq\tau\}}|\mathcal{F}_t\vee\mathcal{G}_T\big)\Big|\mathcal{F}_t\Big].
\end{aligned}$$

As $s>t$ we have $\{s\leq\tau\}=\{t<s\leq\tau\}$. Hence the inner expectation can be represented via $(\tilde{N}(t))_{t\geq 0}$, so that

$$\begin{aligned}
1_{\{\tau>t\}}\mathbb{E}\big(1_{\{s\leq\tau\}}|\mathcal{F}_t\vee\mathcal{G}_T\big) &= 1_{\{\tau>t\}}\mathbb{E}\big(1_{\{s\leq\tau\}}|\mathcal{H}_t\vee\mathcal{G}_T\big) \\
&= \mathbb{E}\big(1_{\{\tau>t\}}\,1_{\{\tilde{N}(\Lambda_s)\leq 1\}}|\mathcal{H}_t\vee\mathcal{G}_T\big).
\end{aligned}$$

On $\{\tau>t\}$ we have that $\tilde{N}(\Lambda_t)=0$, because no jump occurred before t. This yields

$$\begin{aligned}
1_{\{\tau>t\}}\mathbb{E}\big(1_{\{s\leq\tau\}}|\mathcal{F}_t\vee\mathcal{G}_T\big) &= 1_{\{\tau>t\}}\mathbb{E}\big(1_{\{\tilde{N}(\Lambda_s)-\tilde{N}(\Lambda_t)\leq 1\}}|\mathcal{H}_t\vee\mathcal{G}_T\big) \\
&= 1_{\{\tau>t\}}\exp[-(\Lambda_s-\Lambda_t)] = 1_{\{\tau>t\}}\exp(-\int_t^s \lambda_u\,du)
\end{aligned}$$

and (i) follows.

Assertion (ii) of the theorem is covered by Bielecki and Rutkowski (2002, Prop. 8.2.1). ■

If several default events are under consideration, one uses the following

Theorem A.1.4. *If $(L_t)_{t\geq 0}$ is a process which is adapted to $(\mathcal{G})_{t\geq 0}$ we have under the assumptions of A.1.2*

$$\mathbb{E}\Big[\prod_{i=1}^{N_T}(1-L_{\tau_i})\,X_T\Big|\,\mathcal{F}_t\Big] = \prod_{i=1}^{N_t}(1-L_{\tau_i})\mathbb{E}\Big[\exp(-\int_t^T L_u\lambda_u\,du)\,X_T\Big|\,\mathcal{F}_t\Big].$$

Proof. As (L_t) is an adapted process, we have

$$\begin{aligned}
&\mathbb{E}\Big[\mathbb{E}\Big(\prod_{i=1}^{N_T}(1-L_{\tau_i})\,X_T\Big|\mathcal{G}_T\vee\mathcal{H}_t\Big)\Big|\,\mathcal{F}_t\Big] &\text{(A.1)}\\
&\quad= \prod_{i=1}^{N_t}(1-L_{\tau_i})\mathbb{E}\Big[\mathbb{E}\Big(\prod_{i=N_t+1}^{N_T}(1-L_{\tau_i})\,X_T\Big|\mathcal{G}_T\vee\mathcal{H}_t\Big)\Big|\,\mathcal{F}_t\Big].
\end{aligned}$$

Consider the inner expectation

$$\begin{aligned}
&\sum_{k\geq 0}\mathbb{E}\Big(1_{\{N_T-N_t=k\}}\prod_{i=N_t+1}^{N_t+k}(1-L_{\tau_i})\Big|\mathcal{G}_T\vee\mathcal{H}_t\Big)\,X_T &\text{(A.2)}\\
&\quad= \sum_{k\geq 0}\mathbb{E}\Big[1_{\{N_T-N_t=k\}}\mathbb{E}\Big(\prod_{i=N_t+1}^{N_t+k}(1-L_{\tau_i})\Big|\mathcal{G}_T\vee\mathcal{H}_t\vee\sigma(N_T)\Big)\Big|\mathcal{G}_T\vee\mathcal{H}_t\Big]\,X_T.
\end{aligned}$$

The conditional distribution of the τ_i's can be replaced by an unconditional one[2], because

$$\mathcal{L}\big(\tau_{N_t+1},\dots,\tau_{N_T}\big|N_T-N_t=k\big)=\mathcal{L}\big(\eta_{1:n},\dots,\eta_{k:n}\big).$$

Here, the η_i are i.i.d. with density

$$\frac{\lambda_u}{\int_t^T\lambda_u\,du} \qquad \text{on } (t,T].$$

Because the order within the product can be interchanged, it is possible to switch back to the η_i. The inner expectation (A.2) then becomes

$$\begin{aligned}
\mathbb{E}\Big(\prod_{i=1}^{k}(1-L_{\eta_i})\Big|\mathcal{G}_T\vee\mathcal{H}_t\Big) &= \Big(1-\mathbb{E}\big(L_{\eta_1}\big|\mathcal{G}_T\vee\mathcal{H}_t\big)\Big)^k\\
&= \Big(1-\int_t^T L_u\,\frac{\lambda_u}{\int_t^T\lambda_w\,dw}\,du\Big)^k.
\end{aligned}$$

[2]See Rolski, Schmidli, Schmidt and Teugels (1999), p.502. The $\eta_{i:n}$ denote the order statistics of η_i, that is the η_i are ordered, such that $\eta_{1:n}\leq\eta_{2:n}\leq\cdots\leq\eta_{n:n}$.

Therefore we may conclude

$$
\begin{aligned}
(A.2) &= X_T \sum_{k\geq 0} \mathbb{E}\big[1_{\{N_T - N_t = k\}} \big| G_T \vee H_t\big] \Big[1 - \int_t^T L_u \frac{\lambda_u}{\int_t^T \lambda_w \, dw} \, du\Big]^k \\
&= X_T \exp(-\int_t^T \lambda_u \, du) \sum_{k\geq 0} \frac{\big(\int_t^T \lambda_u \, du\big)^k}{k!} \Big[1 - \int_t^T L_u \frac{\lambda_u}{\int_t^T \lambda_w \, dw} \, du\Big]^k \\
&= \exp(-\int_t^T L_u \lambda_u \, du) \, X_T. \qquad \blacksquare
\end{aligned}
$$

Appendix B

Auxiliary Calculations

B.1 Normal Random Variables

Consider two independent normally distributed random variables X_1 and X_2 with zero mean and variances σ_1^2 and σ_2^2. The following lemma may be used to determine the distribution of X_1 conditionally on $X_1 + X_2$.

Lemma B.1.1. *There exists $\xi \sim \mathcal{N}(0,1)$, which is independent of $X_1 + X_2$, and*

$$X_1 = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}(X_1 + X_2) + \frac{\sigma_1 \sigma_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}\, \xi.$$

Proof. We define

$$\xi := \frac{\sigma_2}{\sigma_1 \sqrt{\sigma_1^2 + \sigma_2^2}} X_1 - \frac{\sigma_1}{\sigma_2 \sqrt{\sigma_1^2 + \sigma_2^2}} X_2.$$

Then ξ is normally distributed with expectation zero and variance 1. It remains to show that ξ is independent of $X_1 + X_2$. This follows from

$$\begin{aligned}
\operatorname{Cov}\big(\xi, X_1 + X_2\big) &= \mathbb{E}\Big[\big(\frac{\sigma_2}{\sigma_1 \sqrt{\sigma_1^2 + \sigma_2^2}} X_1 - \frac{\sigma_1}{\sigma_2 \sqrt{\sigma_1^2 + \sigma_2^2}} X_2\big) \cdot \big(X_1 + X_2\big)\Big] \\
&= \frac{\sigma_2}{\sigma_1 \sqrt{\sigma_1^2 + \sigma_2^2}} \sigma_1^2 - \frac{\sigma_1}{\sigma_2 \sqrt{\sigma_1^2 + \sigma_2^2}} \sigma_2^2 \\
&= 0.
\end{aligned}$$

■

B.2 Boundary Crossing Probabilities

We have the following (see, e.g., Pechtl (1996))

Theorem B.2.1. *For a standard Brownian motion $(B_s)_{s \geq 0}$, constants b and m, we have for $b < 0$*

$$\mathbb{P}\big(\inf_{0 < s \leq t} ms + B_s \leq b\big) = \Phi\big(\frac{b - mt}{\sqrt{t}}\big) + e^{2bm} \Phi\big(\frac{b + mt}{\sqrt{t}}\big).$$

Proof. By the reflection principle[1] we conclude for $b < 0$

$$\mathbb{P}(\inf_{0\le s\le t} B_s < b) = 2\mathbb{P}(B_t < b) = 2\Phi\big(\frac{b}{\sqrt{t}}\big). \tag{B.1}$$

Consider a probability measure P^*, defined by

$$dP^* = \exp\big(-mB_t - \frac{m^2t}{2}\big)dP = \exp(-mB_t^* + \frac{m^2t}{2})dP,$$

with $B_t^* := mt + B_t$. P^* is equivalent to P, and the Girsanov theorem yields that B^* is a Brownian motion under P^*. We conclude

$$\begin{aligned} P(\inf_{0\le s\le t} B_s^* < b) &= \int 1_{\{\inf_{0\le s\le t} B_s^* < b\}} \exp(mB_t^* - \frac{m^2t}{2})\, dP^* \\ &= \int \exp(mB_t^* - \frac{m^2t}{2})\, P^*\big[\inf_{0\le s\le t} B_s^* < b \big| B_t^*\big]\, dP^*. \end{aligned}$$

The conditional probability equals one for $B_t^* \le b$. For $B_t^* > b$, a result for conditional expectations yields

$$P^*\big[\inf B_s^* < b \big| B_t^* = x\big] = \lim_{h\downarrow 0} \frac{\frac{1}{h}\int_{B_t^*\in[x,x+h]} 1_{\{\inf_{0\le s\le t} B_s^* < b\}}\, dP^*}{\frac{1}{h}P^*(B_t^*\in[x,x+h])}.$$

The numerator equals

$$\begin{aligned} &\lim_{h\downarrow 0}\frac{1}{h}P^*\big(\inf_{0\le s\le t} B_s^* \le b, B_t^*\in[x,x+h]\big) \\ &\qquad = -\frac{\partial}{\partial x}P^*\big(\inf_{0\le s\le t} B_s^* \le b, B_t^* > x\big) \\ &\qquad = -\frac{\partial}{\partial x}\Phi\big(\frac{2b-x}{\sqrt{t}}\big) \\ &\qquad = \frac{1}{\sqrt{2\pi t}}\exp(-\frac{(x-2b)^2}{2t}), \end{aligned}$$

where we again used the reflection principle[2]. For the denominator we obtain

$$\frac{\partial}{\partial x}\Phi\big(\frac{x}{\sqrt{t}}\big) = \frac{1}{\sqrt{2\pi t}}\exp(-\frac{x^2}{2t}),$$

so that for $x > b$

$$P\big[\inf B_s^* < b \big| B_t^* = x\big] = \exp(-\frac{4b^2 - 4xb}{2t}).$$

[1]See, for example, Karatzas and Shreve (1988).

[2]To conclude, that $(x > b)$

$$P^*\big(\inf B_s^* < b, B_t^* > x\big) = \Phi\big(\frac{2b-x}{\sqrt{t}}\big).$$

Substituting this leads to

$$
\begin{aligned}
&P(\inf_{0\le s\le t} B^*_s < b)\\
&\quad= \frac{1}{\sqrt{2\pi t}}\Big[\int_{-\infty}^{b} \exp(-\frac{(x-mt)^2}{2t})\,dx + \int_{b}^{\infty} \exp\big(mx - \frac{m^2t}{2} - \frac{4b^2-4xb}{2t} - \frac{x^2}{2t}\big)\,dx\Big]\\
&\quad= \frac{1}{\sqrt{2\pi t}}\Big[\int_{-\infty}^{b} \exp(-\frac{(x-mt)^2}{2t})\,dx + \int_{b}^{\infty} \exp\big(-\frac{(x-2b-mt)^2}{2t} + 2bm\big)\,dx\Big]\\
&\quad= \Phi(\frac{b-mt}{\sqrt{t}}) + e^{2bm}\Big[1-\Phi(\frac{-b-mt}{\sqrt{t}})\Big]\\
&\quad= \Phi(\frac{b-mt}{\sqrt{t}}) + e^{2bm}\Phi(\frac{b+mt}{\sqrt{t}}).
\end{aligned}
$$

Note that for $m = 0$ we obtain the special case (B.1).

■

We conclude that, for $c < 0$,

$$
\begin{aligned}
&\mathbb{P}\big(\inf_{s\in(t,T)}\{m(s-t)+\sigma(B_s-B_t)\} > c\big)\\
&\quad= \mathbb{P}\big(\inf_{s\in(0,T-t)}\{ms+\sigma B_s\} > c\big)\\
&\quad= \mathbb{P}\big(\inf_{s\in(0,T-t)}\{\frac{m}{\sigma}s + B_s\} > \frac{c}{\sigma}\big)\\
&\quad= 1-\mathbb{P}\big(\inf_{s\in(0,T-t)}\{\frac{m}{\sigma}s + B_s\} < \frac{c}{\sigma}\big)\\
&\quad= 1-\Big[\Phi(\frac{c-m(T-t)}{\sigma\sqrt{T-t}}) + e^{2cm/\sigma^2}\Phi(\frac{c+m(T-t)}{\sigma\sqrt{T-t}})\Big]\\
&\quad= \Phi(\frac{m(T-t)-c}{\sigma\sqrt{T-t}}) - e^{2cm/\sigma^2}\Phi(\frac{m(T-t)+c}{\sigma\sqrt{T-t}}). \qquad \text{(B.2)}
\end{aligned}
$$

For $c > 0$ this probability equals zero, because $\inf\{\dots\} \le 0$.

B.3 Some Integrals

Lemma B.3.1.

(i)
$$\int_{-\infty}^{0} x\cdot\exp\big(-\frac{(x-a)^2}{2b}\big)\,dx = b\exp\big(-\frac{a^2}{2b}\big) + a\sqrt{2\pi b}\,\Phi\big(-\frac{a}{\sqrt{b}}\big)$$

(ii)
$$
\begin{aligned}
\int_{-\infty}^{0} x\exp\big(-\frac{(x-a)^2}{2b} - \frac{(x-d)^2}{2c}\big)\,dx &= \frac{bc}{b+c}\exp\big(-\frac{bc(a-d)^2+(ac+db)^2}{2bc(b+c)}\big)\\
&\quad+\sqrt{2\pi bc}\frac{ac+db}{(b+c)^{\frac{3}{2}}}\exp\big(-\frac{(a-d)^2}{2(b+c)}\big)\Phi\big(-\frac{ac+db}{\sqrt{bc(b+c)}}\big)
\end{aligned}
$$

Proof. We have

$$\begin{aligned}
\int_{-\infty}^{0} x \exp\Big[-\frac{(x-a)^2}{2b}\Big]\,dx &= -b\int_{-\infty}^{0} \frac{x-a}{b}\exp-\frac{(x-a)^2}{2b}\,dx \\
&\quad +a\int_{-\infty}^{0} \exp-\frac{(x-a)^2}{2b}\,dx \\
&= b\exp(-\frac{a^2}{2b}) + a\sqrt{2\pi b}\frac{1}{\sqrt{2\pi b}}\int_{-\infty}^{0}\exp-\frac{(x-a)^2}{2b}\,dx \\
&= b\exp(-\frac{a^2}{2b}) + a\sqrt{2\pi b}\Phi\big(-\frac{a}{\sqrt{b}}\big).
\end{aligned}$$

For (ii), we have

$$\begin{aligned}
&\int_{-\infty}^{0} x\exp\big(-\frac{(x-a)^2}{2b}-\frac{(x-d)^2}{2c}\big)\,dx \\
&\quad= \int_{-\infty}^{0} x\exp\big(-\frac{x^2(b+c)-2(ac+db)x+a^2c+bd^2}{2bc}\big)\,dx \\
&\quad= \int_{-\infty}^{0} x\exp\big(-\frac{\big(x-\frac{ac+db}{b+c}\big)^2}{2bc/(b+c)}-\frac{(a-d)^2}{2(b+c)}\big)\,dx \\
&\quad\overset{(i)}{=} \exp\big(-\frac{(a-d)^2}{2(b+c)}\big)\Big[\frac{bc}{b+c}\exp\big(-\frac{(ac+db)^2}{2bc(b+c)}\big) \\
&\qquad +\sqrt{2\pi bc}\frac{ac+db}{(b+c)^{\frac{3}{2}}}\Phi\big(-\frac{ac+db}{\sqrt{bc(b+c)}}\big)\Big] \\
&\quad= \frac{bc}{b+c}\exp\big(-\frac{bc(a-d)^2+(ac+db)^2)}{2bc(b+c)}\big) \\
&\qquad +\sqrt{2\pi bc}\frac{ac+db}{(b+c)^{\frac{3}{2}}}\exp\big(-\frac{(a-d)^2}{2(b+c)}\big)\Phi\big(-\frac{ac+db}{\sqrt{bc(b+c)}}\big).
\end{aligned}$$

■

The following lemma is an auxiliary result for Section 1.6.2. With the notation therein, we have

Lemma B.3.2. *The default intensity λ_t equals*

$$\lambda_t = -\frac{m\sqrt{2\pi}}{8\sigma}\exp\big(-\frac{(\ln V_B + (\ln\tilde{V}_{t_k} - mt_k + \frac{\sigma_Z^2}{2})(\sigma_Z^2-1) - mt)^2}{2\,\frac{\sigma_Z^2 t+\sigma^2 t_k(t-t_k)}{\sigma_Z^2+\sigma^2 t_k}}\big).$$

Proof. All conditional expectations are with respect to $\mathcal{H}_t$, so we write $\mathbb{E}_t$ for $\mathbb{E}(\dots|\mathcal{H}_t)$. Note that, by definition of a default intensity,

$$\lambda_t = -1_{\{\tau>t\}} \frac{\partial}{\partial T}\Big|_{T=t} \ln\big[\mathbb{P}(\tau > T|\mathcal{H}_t)\big].$$

Since

$$\frac{\partial}{\partial T} \ln \mathbb{P}_t(\tau > T) = \frac{\frac{\partial}{\partial T}\mathbb{P}_t(\tau > T)}{\mathbb{P}_t(\tau > T)}$$

and $1_{\{\tau>t\}}\mathbb{P}_t(\tau > t) = 1$ we just need to compute the numerator. Using (1.13), the numerator equals

$$\begin{aligned}
&\frac{\partial}{\partial T}\mathbb{E}_t\Big[1_{\{\eta<0\}}\big(\Phi\big(\frac{m(T-t)-\eta}{\sigma\sqrt{T-t}}\big) - e^{2\eta m/\sigma^2}\Phi\big(\frac{m(T-t)+\eta}{\sigma\sqrt{T-t}}\big)\big)\Big] \\
&= \mathbb{E}_t\Big[1_{\{\eta<0\}}\Big(\varphi\big(\frac{m(T-t)-\eta}{\sigma\sqrt{T-t}}\big)\big(\frac{m}{2\sigma\sqrt{T-t}} + \frac{\eta}{2\sigma(T-t)^{\frac{3}{2}}}\big) \\
&\quad -e^{2\eta m/\sigma^2}\varphi\big(\frac{m(T-t)+\eta}{\sigma\sqrt{T-t}}\big)\big(\frac{m}{2\sigma\sqrt{T-t}} - \frac{\eta}{2\sigma(T-t)^{\frac{3}{2}}}\big)\Big)\Big] \\
&= \frac{1}{2\sigma(T-t)^{\frac{3}{2}}}\mathbb{E}_t\Big[\eta 1_{\{\eta<0\}}\big(\varphi\big(\frac{m(T-t)-\eta}{\sigma\sqrt{T-t}}\big) + e^{2\eta m/\sigma^2}\varphi\big(\frac{m(T-t)+\eta}{\sigma\sqrt{T-t}}\big)\big)\Big] \\
&\quad +\frac{m}{2\sigma\sqrt{T-t}}\mathbb{E}_t\Big[1_{\{\eta<0\}}\big(\varphi\big(\frac{m(T-t)-\eta}{\sigma\sqrt{T-t}}\big) - e^{2\eta m/\sigma^2}\varphi\big(\frac{m(T-t)+\eta}{\sigma\sqrt{T-t}}\big)\big)\Big] \\
&= \frac{1}{2\sigma(T-t)^{\frac{3}{2}}}\frac{1}{2\pi\sigma_\eta}\int_{-\infty}^{0} x\exp\big(-\frac{(x-\mu_\eta)^2}{2\sigma_\eta^2}\big) \\
&\qquad \cdot\big[\exp\big(-\frac{(m(T-t)-x)^2}{2\sigma^2(T-t)}\big) + \exp\big(\frac{2mx}{\sigma^2} - \frac{(m(T-t)+x)^2}{2\sigma^2(T-t)}\big)\big]\,dx \quad \text{(B.3)} \\
&\quad +\frac{m}{2\sigma\sqrt{T-t}}\frac{1}{2\pi\sigma_\eta}\int_{-\infty}^{0} \exp\big(-\frac{(x-\mu_\eta)^2}{2\sigma_\eta^2}\big) \\
&\qquad \cdot\big[\exp\big(-\frac{(m(T-t)-x)^2}{2\sigma^2(T-t)}\big) - \exp\big(\frac{2mx}{\sigma^2} - \frac{(m(T-t)+x)^2}{2\sigma^2(T-t)}\big)\big]\,dx. \quad \text{(B.4)}
\end{aligned}$$

Observe that the expression in (B.4) equals zero, so we concentrate on the remaining one. As

$$\begin{aligned}
\exp\big(-\frac{(m(T-t)-x)^2}{2\sigma^2(T-t)}\big) &+ \exp\big(\frac{2mx}{\sigma^2} - \frac{(m(T-t)+x)^2}{2\sigma^2(T-t)}\big) \\
&= 2\exp\big(-\frac{(m(T-t)-x)^2}{2\sigma^2(T-t)}\big)
\end{aligned}$$

we may conclude, using Lemma B.3.1, that

$$\begin{aligned}
(B.3) &= \frac{1}{2\sigma\sigma_\eta\pi(T-t)^{\frac{3}{2}}}\int_{-\infty}^{0} x\exp\Big[-\frac{(x-\mu_\eta)^2}{2\sigma_\eta^2}-\frac{(x-m(T-t))^2}{2\sigma^2(T-t)}\Big)\,dx\\
&= \frac{1}{2\sigma\sigma_\eta\pi(T-t)^{\frac{3}{2}}}\Big\{\frac{\sigma_\eta^2\sigma^2(T-t)}{\sigma_\eta^2+\sigma^2(T-t)}\\
&\qquad\cdot\exp\Big[-\frac{\sigma_\eta^2\sigma^2(\mu_\eta-m(T-t))^2+(\mu_\eta\sigma^2(T-t)+m(T-t)\sigma_\eta^2)^2}{2\sigma_\eta^2\sigma^2(T-t)(\sigma_\eta^2+\sigma^2(T-t))}\Big]\\
&\quad+\sqrt{2\pi\sigma^2\sigma_\eta^2(T-t)}\frac{\mu_\eta\sigma^2(T-t)+m(T-t)\sigma_\eta^2}{(\sigma_\eta^2+\sigma^2(T-t))^{\frac{3}{2}}}\exp\Big[-\frac{(\mu_\eta-m(T-t))^2}{2(\sigma_\eta^2+\sigma^2(T-t))}\Big]\\
&\qquad\cdot\Phi\Big(-\frac{\mu_\eta\sigma^2(T-t)+m(T-t)\sigma_\eta^2}{\sigma_\eta\sigma\sqrt{(T-t)(\sigma_\eta^2+\sigma^2(T-t))}}\Big)\Big\}\\
&=: I+II.
\end{aligned}$$

We first show, that I equals zero when we set $T=t$. We have

$$\begin{aligned}
I &= \frac{c_1}{\sqrt{T-t}(\sigma_\eta^2+\sigma^2(T-t))}\cdot\exp\Big[-\frac{\mu_\eta^2}{(T-t)(\sigma_\eta^2+\sigma^2(T-t))}\\
&\qquad+\frac{\mu_\eta m}{\sigma_\eta^2\sigma^2(\sigma_\eta^2+\sigma^2(T-t))}+(T-t)\frac{c_3}{\sigma_\eta^2+\sigma^2(T-t)}\Big]\\
&\to 0,
\end{aligned}$$

as[3] $T\to t$.

For II we obtain

$$\begin{aligned}
II &= \frac{\sqrt{2\pi\sigma^2\sigma_\eta^2(T-t)}}{2\pi\sigma\sigma_\eta(T-t)^{\frac{3}{2}}}\frac{(T-t)(\mu_\eta\sigma^2+m\sigma_\eta^2)}{(\sigma_\eta^2+\sigma^2(T-t))^{\frac{3}{2}}}\\
&\qquad\cdot\exp\Big[-\frac{(\mu_\eta-m(T-t))^2}{2(\sigma_\eta^2+\sigma^2(T-t))}\Big]\cdot\Phi\Big(-\sqrt{T-t}\frac{\mu_\eta\sigma^2+m\sigma_\eta^2}{\sigma_\eta\sigma\sqrt{\sigma_\eta^2+\sigma^2(T-t)}}\Big)\\
&\to \frac{1}{\sqrt{2\pi}}\big(\mu_\eta\frac{\sigma^2}{\sigma_\eta^2}+m\big)\cdot\exp\big(-\frac{\mu_\eta^2}{2\sigma_\eta^2}\big)\cdot\frac{1}{2}.
\end{aligned}$$

With the definitions of μ_η and σ_η we obtain[4]

$$\lambda_t=-\frac{m\sqrt{2\pi}}{8\sigma}\exp\Big(-\frac{(\ln V_B+(\ln\tilde{V}_{t_k}-mt_k+\frac{\sigma_Z^2}{2})(\sigma_Z^2-1)-mt)^2}{2\,\frac{\sigma_Z^2t+\sigma^2t_k(t-t_k)}{\sigma_Z^2+\sigma^2t_k}}\Big). \qquad \blacksquare$$

[3]See Heuser (1991, p. 289) for $\lim_{\tau\downarrow 0}\frac{1}{\sqrt{\tau}}\exp(c\tau^{-1})=0$, for $c<0$.

[4]Recall, that $m=\mu-\frac{\sigma^2}{2}$.

B.4 Tools for Gaussian Models

The following lemma is a simple version of Girsanov's Theorem 2.6.1.

Lemma B.4.1. *Assume the two random variables ξ and η are jointly normally distributed under a probability measure Q. Then*

$$d\tilde{Q} := \frac{e^{\xi}}{\mathbb{E}(e^{\xi})}\, dQ$$

defines a measure equivalent to Q. η is normally distributed under $\tilde{Q}$ with

$$\begin{aligned} \tilde{E}(\eta) &= \mathbb{E}(\eta) + \mathrm{Cov}(\eta, \xi), \\ \tilde{\mathrm{Var}}(\eta) &= \mathrm{Var}(\eta). \end{aligned}$$

Proof. The definition of $\tilde{Q}$ yields for $\lambda \in \mathbb{R}$

$$\begin{aligned} \tilde{E}\big(\exp(\lambda\eta)\big) &= \mathbb{E}\big(\frac{e^{\xi}}{\mathbb{E}(e^{\xi})}\, e^{\lambda\eta}\big) \\ &= \frac{\exp\Big[\mathbb{E}\xi + \lambda\mathbb{E}\eta + \frac{1}{2}\mathrm{Var}\,\xi + \lambda\,\mathrm{Cov}(\eta,\xi) + \frac{\lambda^2}{2}\mathrm{Var}\,\eta\Big]}{\exp(\mathbb{E}(\xi) + \frac{1}{2}\mathrm{Var}(\xi))}, \end{aligned}$$

which immediately yields the desired result. ■

The following expectation is essential for the derivation of the Black-Scholes formula:

Lemma B.4.2. *For a normally distributed random variable ξ with variance σ^2, we have*

$$\mathbb{E}\big(e^{\xi} 1_{\{\xi > a\}}\big) = \mathbb{E}\big(e^{\xi}\big) \cdot \mathbb{P}\big(\xi > a - \sigma^2\big).$$

Proof. Assume $\mathbb{E}(\xi) = 0$, the statement with nonzero mean being an easy extension. Then

$$\begin{aligned} \mathbb{E}(e^{\xi} 1_{\{\xi > a\}}) &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_a^{\infty} \exp[x - \frac{x^2}{2\sigma^2}]\, dx \\ &= \exp[\frac{\sigma^2}{2}] \int_a^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp[-\frac{(x-\sigma^2)}{2\sigma^2}]\, dx \\ &= \exp[\frac{\sigma^2}{2}]\mathbb{P}(\xi + \sigma^2 > a). \end{aligned}$$

■

Furthermore, we have

Lemma B.4.3. *For $i = 1, 2$ let ξ_i be jointly normal with expectation μ_i, variance σ_i^2 and correlation ρ, respectively. Then*

$$\begin{aligned} \mathbb{E}\big(e^{\xi_2}\big(e^{\xi_1} - K\big)^+\big) &= \mathbb{E}\big[e^{\xi_1+\xi_2}\big]\,\Phi\Big(\frac{\mu_1 - \ln K}{\sigma_1} + \rho\sigma_2 + \sigma_1\Big) \\ &\quad - K\mathbb{E}\big(e^{\xi_2}\big) \cdot \Phi\Big(\frac{\mu_1 - \ln K}{\sigma_1} + \rho\sigma_2\Big). \end{aligned}$$

Proof. First,

$$\begin{aligned}
&\mathbb{E}\Big(e^{\xi_2}\big(e^{\xi_1}-K\big)^+\Big) \\
&\qquad = \mathbb{E}\big(e^{\xi_1+\xi_2}1_{\{\xi_1>\ln K\}}\big) - K\mathbb{E}\big(e^{\xi_2}1_{\{\xi_1>\ln K\}}\big) \\
&\qquad =: (1)+(2).
\end{aligned}$$

We use the decomposition

$$\xi_2 = \mu_2 + \frac{\sigma_2\rho}{\sigma_1}(\xi_1-\mu_1) + \sigma_2\sqrt{1-\rho^2}\xi,$$

where ξ is standard normally distributed and independent of ξ_1. This yields for the first term

$$\begin{aligned}
(1) &= e^{\mu_1+\mu_2}\mathbb{E}\Big[\exp\Big((\xi_1-\mu_1)(1+\frac{\sigma_2\rho}{\sigma_1}) + \sigma_2\sqrt{1-\rho^2}\xi\Big)1_{\{\xi_1>\ln K\}}\Big] \\
&= e^{\mu_1+\mu_2+\frac{\sigma_2^2(1-\rho^2)}{2}}\mathbb{E}\Big[\exp((\xi_1-\mu_1)(1+\frac{\sigma_2\rho}{\sigma_1}))1_{\{(\xi_1-\mu_1)(1+\frac{\sigma_2\rho}{\sigma_1})>(\ln K-\mu_1)(1+\frac{\sigma_2\rho}{\sigma_1})\}}\Big].
\end{aligned}$$

Applying Lemma B.4.2, we obtain

$$\begin{aligned}
(1) &= \exp\big[\mu_1+\mu_2+\frac{\sigma_2^2(1-\rho^2)}{2}+\frac{(\sigma_1+\sigma_2\rho)^2}{2}\big] \\
&\quad \cdot\mathbb{P}\Big((\xi_1-\mu_1)\frac{\sigma_1+\sigma_2\rho}{\sigma_1} > (\ln K-\mu_1)\frac{\sigma_1+\sigma_2\rho}{\sigma_1} - (\sigma_1+\sigma_2\rho)^2\Big) \\
&= \exp\big[\mu_1+\mu_2+\frac{\sigma_1^2+2\rho\sigma_1\sigma_2+\sigma_2^2}{2}\big]\Phi\Big(\frac{\mu_1-\ln K}{\sigma_1}+\sigma_1+\sigma_2\rho\Big).
\end{aligned}$$

For the second term we have

$$\begin{aligned}
&\mathbb{E}(e^{\xi_2}1_{\{\xi_1>\ln K\}}) = \mathbb{E}\Big[\exp\big(\mu_2+\frac{\sigma_2\rho}{\sigma_1}(\xi_1-\mu_1)+\sigma_2\sqrt{1-\rho^2}\xi\big)1_{\{\xi_1>\ln K\}}\Big] \\
&= \exp\big(\mu_2+\frac{\sigma_2^2(1-\rho^2)}{2}\big)\mathbb{E}\Big[e^{\frac{\sigma_2\rho}{\sigma_1}(\xi_1-\mu_1)}1_{\{\frac{\sigma_2\rho}{\sigma_1}(\xi_1-\mu_1)>\frac{\sigma_2\rho}{\sigma_1}\}}(\ln K-\mu_1)\Big] \\
&\overset{B.4.2}{=} \exp\big(\mu_2+\frac{\sigma_2^2}{2}\big)\cdot\mathbb{P}\big(\frac{\sigma_2\rho}{\sigma_1}(\xi_1-\mu_1) > (\ln K-\mu_1)\frac{\sigma_2\rho}{\sigma_1} - \rho^2\sigma_2^2\big) \\
&= \exp\big(\mu_2+\frac{\sigma_2^2}{2}\big)\cdot\Phi\Big(\frac{\mu_1-\ln K}{\sigma_1}+\rho\sigma_2\Big).
\end{aligned}$$

■

Usually the formula in Lemma B.4.3 is abbreviated as

$$\mathbb{E}\big[e^{\xi_1+\xi_2}\big]\Phi(d_1) - K\mathbb{E}\big(e^{\xi_2}\big)\Phi(d_2)$$

and we immediately obtain

$$\mathbb{E}\Big(e^{\xi_2}\big(K-e^{\xi_1}\big)^+\Big) = K\mathbb{E}\big(e^{\xi_2}\big)\Phi(-d_2) - \mathbb{E}\big[e^{\xi_1+\xi_2}\big]\Phi(-d_1). \tag{B.5}$$

We use the notation of Section 4.4.4. It may be recalled that the mean and variance of $\tilde{B}(T, T_1, \ldots, T_n)$ was denoted by m and σ^2 and

$$l_2(u,T) = \begin{cases} -l_u & \text{for } u \leq T, \\ 1 & \text{for } u > T. \end{cases}.$$

The computation of

$$\rho := \operatorname{Cov}\Big[\ln \frac{\tilde{B}(T,T_1,\ldots,T_n)}{B(T,T_n)}, -\int_0^T r_u + \lambda_u \, du - \ln B(T,T_n)\Big]$$

is done in the following

Lemma B.4.4. *Under the assumption (B2) we have*

$$\begin{aligned}
\rho = \ln \Bigg[& \frac{\bar{B}(0,T_n)}{\bar{B}(0,T)} \exp\Big[\int_0^T \bar{c}(v,u,v)\, dv\, du - \int_T^{T_n}\int_0^T l_2(v,T)\varsigma(v \wedge T, u, v)\, dv\, du\Big] \\
& + K \sum_{i=1}^n \frac{B^0(0,T_i)}{B^0(0,T)} \exp\Big[-\int_T^{T_i}\int_0^T c^0(v,u,v)\, dv\, du \\
& \qquad - \int_T^{T_i}\int_0^{T_n} l_2(v,T)\big[l_u c(v\wedge T,u,v) + \frac{\varsigma(v\wedge T,u,v)}{L_u}\big]\, dv\, du\Big]\Bigg] \\
& - \ln\Bigg[\frac{\bar{B}(0,T_n)}{\bar{B}(0,T)} \exp(-\int_T^{T_n}\int_0^T (1+\frac{1}{L_v})\bar{c}(v,u,v)\, dv\, du) \\
& + K \sum_{i=1}^n \frac{B^0(0,T_i)}{B^0(0,T)} \exp\Big[-\int_T^{T_i}\int_0^T c^0(v,u,v)\, dv\, du - \int_T^{T_i}\int_0^T \frac{\bar{c}(v,u,v)}{L_u L_v} + \frac{l_u \varsigma(v,v,u)}{L_v}\, dv\, du\Big]\Bigg] \\
& + \int_T^{T_n}\int_0^{T_n} l_2(v,T)c(v\wedge T,u,v)\, dv\, du - \int_T^{T_n}\int_0^T \frac{\varsigma(u\wedge v,v,u)}{L_v}\, dv\, du. \qquad \text{(B.6)}
\end{aligned}$$

Proof. By the definition of ρ,

$$\begin{aligned}
\rho &= \operatorname{Cov}\Big[\ln \tilde{B}(T,T_1,\ldots,T_n) - \ln B(T,T_n), -\int_0^T r_u + \lambda_u\, du + \int_T^{T_n} f(T,u)\, du\Big] \\
&= \operatorname{Cov}\Big[\ln \tilde{B}(T,T_1,\ldots,T_n), \int_0^{T_n} l_2(v,T)X(v)\, dv - \int_0^T \frac{\bar{X}(v)}{L_v}\, dv\Big] \\
&\quad + \int_T^{T_n}\int_0^{T_n} l_2(v,T)c(v\wedge T,u,v)\, dv\, du - \int_T^{T_n}\int_0^T \frac{\varsigma(u\wedge v,v,u)}{L_v}\, dv\, du. \qquad \text{(B.7)}
\end{aligned}$$

We compute the covariances separately. Observe, that for two jointly normally distributed random variables ξ_1 and ξ_2,

$$\begin{aligned}\mathrm{Cov}(e^{\xi_1}, e^{\xi_2}) &= \mathbb{E}(e^{\xi_1+\xi_2}) - \mathbb{E}(e^{\xi_1})\,\mathbb{E}(e^{\xi_2}) \\ &= \mathbb{E}(e^{\xi_1})\,\mathbb{E}(e^{\xi_2})\big[e^{\mathrm{Cov}(\xi_1,\xi_2)} - 1\big],\end{aligned} \tag{B.8}$$

which is equivalent to

$$\begin{aligned}\mathrm{Cov}(\xi_1, \xi_2) &= \ln\big[1 + \frac{\mathrm{Cov}(e^{\xi_1}, e^{\xi_2})}{\mathbb{E}(e^{\xi_1})\,\mathbb{E}(e^{\xi_2})}\big] \\ &= \ln \mathbb{E}(e^{\xi_1+\xi_2}) - \ln[\mathbb{E}(e^{\xi_1})\,\mathbb{E}(e^{\xi_2})].\end{aligned}$$

First, consider

$$\begin{aligned}&\mathrm{Cov}\,\big[\ln \tilde{B}(T, T_1, \ldots, T_n), \int_0^{T_n} l_2(v, T)X(v)\,dv\big] \\ &= \ln \mathbb{E}\big[\frac{\tilde{B}(T, T_1 \ldots, T_n)}{\exp(-\int_0^{T_n} l_2(v,T)X(v)\,dv)}\big] - \ln \mathbb{E}(\tilde{B}(T, T_1, \ldots, T_n)) \\ &\quad - \ln \mathbb{E}\big(\exp(\int_0^{T_n} l_2(v, T)X(v)\,dv)\big).\end{aligned}$$

Assumption (B2) leads to

$$\begin{aligned}\mathbb{E}\Big[\frac{\tilde{B}(T, T_1, \ldots, T_n)}{\exp(-\int_0^{T_n} l_2(v,T)X(v)\,dv)}\Big] &= \mathbb{E}\Big[\frac{\bar{B}(T, T_n)}{\exp(-\int_0^{T_n} l_2(v,T)X(v)\,dv)}\Big] \\ &\quad + K\sum_{i=1}^{n} \mathbb{E}\Big[\frac{B^0(T, T_i)}{\exp(-\int_0^{T_n} l_2(v,T)X(v)\,dv)}\Big],\end{aligned}$$

where the expectations are

$$\begin{aligned}&\mathbb{E}\Big[\frac{\bar{B}(T, T_n)}{\exp(-\int_0^{T_n} l_2(v,T)X(v)\,dv)}\Big] \\ &\qquad = \mathbb{E}\big[\exp(-\int_T^{T_n} \bar{f}(T, u)\,du + \int_0^{T_n} l_2(v, T)X(v)\,dv)\big] \\ &\qquad = \exp\Big[-\int_T^{T_n} \bar{\mu}(0, u) + \int_0^{T} \bar{c}(v, u, v)\,dv\,du \\ &\qquad\quad + \frac{1}{2}\int_0^{T_n}\int_0^{T_n} l_2(u, T) l_2(v, T) c(u \wedge v, u, v)\,dv\,du \\ &\qquad\quad - \int_T^{T_n}\int_0^{T} l_2(v, t)\varsigma(T \wedge v, u, v)\,dv\,du\Big].\end{aligned}$$

For the computation of the second expectation note that

$$\mathrm{Var}\Big[\int_T^{T_i} l_u X(T,u) + \frac{\bar{X}(T,u)}{L_u}\,du\Big] = \int_T^{T_i}\int_T^{T_i} c^0(T,u,v)\,dv\,du,$$

and, using the drift condition for the zero recovery bond,

$$\begin{aligned} -\int_T^{T_i} \mu^0(T,u)\,du + \frac{1}{2}\int_T^{T_i}\int_T^{T_i} c^0(T,u,v)\,dv\,du & \\ = -\int_T^{T_i} \mu^0(0,u) + \int_0^T c^0(v,u,v)\,dv\,du. & \end{aligned}$$

So, using (4.16), we obtain

$$\begin{aligned} &\mathbb{E}\Big[\frac{B^0(T,T_i)}{\exp(-\int_0^{T_n} l_2(v,T)X(v)\,dv)}\Big] \\ &\quad = \exp\Big[-\int_T^{T_i} \mu^0(0,u) + \int_0^T c^0(v,u,v)\,dv\,du \\ &\qquad + \frac{1}{2}\int_0^{T_n}\int_0^{T_n} l_2(u,T)l_2(v,T)c(u\wedge v,u,v)\,dv\,du \\ &\qquad - \int_T^{T_i}\int_0^{T_n} l_u l_2(v,T)c(T,u,v) + \frac{l_2(v,T)\varsigma(T\wedge v,u,v)}{L_u}\,dv\,du\Big]. \end{aligned}$$

Conclude

$$\begin{aligned} &\mathrm{Cov}\Big[\ln\tilde{B}(T,T_1,\dots,T_n), \int_0^{T_n} l_2(v,T)X(v)\,dv\Big] \\ &\quad = \ln\Big[\exp\Big[-\int_T^{T_n} \bar{f}(0,u) + \int_0^T \bar{c}(v,u,v)\,dv\,du - \int_T^{T_n}\int_0^T l_2(v,T)\varsigma(T\wedge v,u,v)\,dv\,du\Big] \\ &\qquad + K\sum_{i=1}^n \exp\Big[-\int_T^{T_i} f^0(0,u) + \int_0^T c^0(v,u,v)\,dv\,du \\ &\qquad\qquad - \int_T^{T_i}\int_0^{T_n} l_u l_2(v,T)c(T,u,v) + \frac{l_2(v,T)\varsigma(T\wedge v,u,v)}{L_u}\,dv\,du\Big]\Big] \\ &\qquad - \Big[\tilde{m} + \frac{\tilde{\sigma}^2}{2}\Big]. \end{aligned}$$

Consider the second covariance in (B.7),

$$\mathrm{Cov}\Big[\ln\tilde{B}(T,T_1,\dots,T_n), \int_0^T \frac{\bar{X}(u)}{L_u}\,du\Big].$$

We need the following two expectations:

$$\begin{aligned}
&\mathbb{E}\Big[\bar{B}(T,T_n)\exp\Big(\int_0^T \frac{\bar{X}(v)}{L_v}\,dv\Big)\Big] \\
&= \exp\Big[-\int_T^{T_n} \bar{\mu}(0,u)\,du - \int_T^{T_n}\int_0^T (1+\frac{1}{L_v})\bar{c}(v,u,v)\,dv\,du \\
&\quad +\frac{1}{2}\int_0^T\int_0^T \frac{\bar{c}(u\wedge v,u,v)}{L_uL_v}\,dv\,du\Big]
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}\Big[B^0(T,T_i)\exp\Big(\int_0^T \frac{\bar{X}(v)}{L_v}\,dv\Big)\Big] \\
&= \exp\big[-\int_T^{T_i} \mu^0(0,u) + \int_0^T c^0(u\wedge v,u,v)\,dv\,du \\
&\quad + \frac{1}{2}\int_0^T\int_0^T \frac{\bar{c}(u\wedge v,u,v)}{L_uL_v}\,dv\,du \\
&\quad - \int_T^{T_i}\int_0^T \frac{\bar{c}(v,u,v)}{L_uL_v} + \frac{l_u\varsigma(v,v,u)}{L_v}\,dv\,du\Big].
\end{aligned}$$

Thus,

$$\begin{aligned}
&\mathrm{Cov}\,\big[\ln\tilde{B}(T,T_1,\ldots,T_n), \int_0^T \frac{\bar{X}(v,v)}{L_v}\,dv\big] \\
&= \ln\Big[\exp\big(-\int_T^{T_n} \bar{f}(0,u)\,du - \int_T^{T_n}\int_0^T (1+\frac{1}{L_v})\bar{c}(v,u,v)\,dv\,du\big) \\
&\quad +K\sum_{i=1}^{n}\exp\big[-\int_T^{T_i} f^0(0,u) + \int_0^T c^0(v,u,v)\,dv\,du \\
&\quad -\int_T^{T_i}\int_0^T \frac{\bar{c}(v,u,v)}{L_uL_v} + \frac{l_u\varsigma(v,v,u)}{L_v}\,dv\,du\big]\Big] \\
&\quad -\Big[\tilde{m} + \frac{\tilde{\sigma}^2}{2}\Big].
\end{aligned}$$

Finally, putting the above equations together yields the desired result. ■

Bibliography

Acharya, V. V., S. R. Das and R. K. Sundaram (2000). Pricing credit derivatives with rating transitions. *Working paper*.

Adler, R. J. (1981). *The Geometry of Random Fields*. John Wiley & Sons. New York.

Altman, E. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance 23*, 589–609.

Ammann, M. (1999). *Pricing Derivative Credit Risk*, Volume 470 of *Lecture Notes in Economics and Mathematical Systems*. Springer Verlag. Berlin Heidelberg New York.

Andersen, L. and J. Andreasen (2000). Volatility skews and extensions of the LIBOR market model. *Applied Mathematical Finance 7*, 1–32.

Asarnow, E. and D. Edwards (1995). Measuring loss on defaulted bank loans: A 24-year study. *Journal of Commercial Lending 77*, 11–23.

Bagchi, A. and K. S. Kumar (2000). An infinte factor model for the interest rate derivatives. *Working paper*.

Bielecki, T. and M. Rutkowski (2000). Multiple ratings model of defaultable term structure. *Mathematical Finance 10*, 125–139.

Bielecki, T. and M. Rutkowski (2002). *Credit Risk: Modeling, Valuation and Hedging*. Springer Verlag. Berlin Heidelberg New York.

Björk, T. (1997). *Interest Rate Theory*, Volume 1656 of *Lecture Notes in Mathematics*. Springer Verlag. Berlin Heidelberg New York.

Björk, T., G. B. di Masi, Y. Kabanov and W. J. Runggaldier (1997). Towards a general theory of bond markets. *Finance and Stochastics*, 141–174.

Black, F. and J. C. Cox (1976). Valuing corporate securities: Some effects of bond indenture provisions. *Journal of Finance 31*, 351–367.

Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. *Journal of Political Economy 81*, 637–653.

Blanchet-Scalliet, C. and M. Jeanblanc (2001). Hazard rate for credit risk and hedging defaultable contingent claims. *Working paper*.

Blum, C., L. Overbeck and C. Wagner (2003). *Credit Risk Modeling*. Chapman & Hall/CRC.

Bogachev, V. I. (1991). *Gaussian Measures*. American Mathematical Society.

Brace, A., D. Gatarek and M. Musiela (1995). The market model of interest rate dynamics. *Mathematical Finance 7*, 127–155.

Brémaud, P. (1981). *Point Processes and Queues*. Springer Verlag. Berlin Heidelberg New York.

Brigo, D. and F. Mercurio (2001). *Interest Rate Models - Theory and Practice*. Springer Verlag. Berlin Heidelberg New York.

Buffett, E. (2002). Credit risk: The structural approach revisited. *Progress in Probability 52*, 45–53.

Caouette, J. B., E. I. Altmann and P. Narayanan (1998). *Managing Credit Risk*. John Wiley & Sons. New York.

Cont, R. (2001). Modeling term structure dynamics: An infinite dimensional approach. *To appear*.

Cox, J. C., J. W. Ingersoll and S. A. Ross (1985). A theory of the term structure of interest rates. *Econometrica 54*, 385–407.

Cox., J. C., S. A. Ross and M. Rubinstein (1979). Option pricing: A simplified approach. *Journal of Financial Econometrics 7*, 229–265.

Crosbie, P. and J. R. Bohn (1997-2001). *Modeling Default Risk*. KMV Corporation [http://www.kmv.kom/insight/index.html].

Da Prato, G. and J. Zabczyk (1992). *Stochastic Equations in Infinite Dimensions*. Cambridge University Press.

Das, S. (Ed.) (1998). *Credit Derivatives*. John Wiley & Sons. New York.

Das, S. and P. Tufano (1996). Pricing credit-sensitive debt when interest rates, credit ratings and credit spreads are stochastic. *Journal of Financial Engineering 5*, 161–198.

Dieudonné, J. (1969). *Foundations of Modern Analysis*. Academic Press.

Duffee, G. (1999). Estimating the price of default. *Review of Financial Studies 12*, 187–226.

Duffie, D. and D. Lando (2001). Term structures of credit spreads with incomplete accounting information. *Econometrica 69*, 633–664.

Duffie, D. and K. Singleton (1999). Modeling term structures of defaultable bonds. *Review of Financial Studies 12*, 687–720.

Filipović, D. (2001). *Consistency Problems for Heath-Jarrow-Morton Interest Rate Models*, Volume 1760 of *Lecture Notes in Mathematics*. Springer Verlag. Berlin Heidelberg New York.

Frey, R. and A. J. McNeil (2001). Modelling dependent defaults. *Working paper*.

Geske, R. and H. E. Johnson (1984). The valuation of corporate liabilities as compound options: A correction. *Journal of Financial and Quantitative Analysis 19*, 231–232.

Gill, R. and S. Johannsen (1990). A survey of product-integration with a view towards applications in survival analysis. *The Annals of Statistics 18*, 1501–1555.

Goldstein, R. (1997). The term structure of interest rates as a random field. *Working paper*.

Grandell, J. (1997). *Mixed Poisson Processes*, Volume 77 of *Monographs on Statistics and Applied Probability*. Chapman & Hall.

Gupton, G. M., C. Finger and M. Bhatia (1997). CreditMetrics: Technical document. *J.P.Morgan & Incorporated, New York [http://www.riskmetrics.com/ research].*

Hall, P., N. I. Fisher and B. Hoffmann (1994). On the nonparametric estimation of covariance functions. *Annals of Statistics*, 2115–2134.

Harrison, J. M. and D. M. Kreps (1979). Martingales and arbitrage in multiperiod securities market. *Journal of Economic Theory 20*, 381–408.

Harrison, J. M. and S. R. Pliska (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications 11*, 215–260.

Heath, D., R. A. Jarrow and A. J. Morton (1992). Bond pricing and the term structure of interest rates. *Econometrica 60*, 77–105.

Heuser, H. (1991). *Lehrbuch der Analysis, Teil 1 (9. Auflage).* B.G.Teubner Stuttgart.

Hull, J. C. (1993). *Options, Futures, and other Derivative Securities.* Prentice Hall International, Inc.

Hull, J. C. and A. White (1990). Pricing interest-rate derivative securities. *Review of Financial Studies 3*, 573–592.

Hull, J. C. and A. White (2002). The valuation of credit default swap options. *Working paper*.

Israel, R. B., J. S. Rosenthal and J. Z. Wei (2001). Finding generators for markov chains via empirical transition matrices, with applications to credit ratings. *Mathematical Finance 11*, 245–265.

Itô, K. (1946). On a stochastic integral equation. *Proc. Imperial Acad Tokyo 22*, 32–35.

James, J. and N. Webber (2000). *Interest Rate Modelling.* John Wiley & Sons. New York.

Jamshidian, F. (1997). LIBOR and swap market models and measures. *Finance and Stochastics 1*, 293–330.

Jarrow, R., D. Lando and S. Turnbull (1997). A Markov model for the term structure of credit risk spreads. *Review of Financial Studies 10*, 481–523.

Jarrow, R. and S. Turnbull (1995). Pricing options on financial securities subject to default risk. *Journal of Finance 5*, 53–86.

Jarrow, R. and S. Turnbull (2000). The intersection of market and credit risk. *Journal of Banking and Finance 24*, 271–299.

Jarrow, R. and F. Yu (2001). Counterparty risk and the pricing of defaultable securities. *Journal of Finance 56*, 1765–1799.

Jeanblanc, M. (2002). Credit risk. *Munich Spring School 2002*.

Jeanblanc, M. and M. Rutkowski (2000). Modelling of default risk: Mathematical tools. *Working paper*.

Jones, E., S. Mason and E. Rosenfeld (1984). Contingent claim analysis of corporate capital structures: An empirical investigation. *Journal of Finance 39*, 611–625.

Karatzas, I. and S. E. Shreve (1988). *Brownian Motion and Stochastic Calculus.* Springer Verlag. Berlin Heidelberg New York.

Kennedy, D. P. (1994). The term structure of interest rates as a Gaussian random field. *Mathematical Finance 4*, 247–258.

Kennedy, D. P. (1997). Characterizing Gaussian models of the term structure of interest rates. *Mathematical Finance 7*, 107–118.

Kijima, M. (2000). Valuation of a credit swap of the basket type. *Review of Derivatives Research 4*, 81–97.

Kijima, M. and Y. Muromachi (2000). Credit events and the valuation of credit derivatives of basket type. *Review of Derivatives Research 4*, 55–79.

Kim, I. J., K. Ramaswamy and S. Sundaresan (1993). The valuation of corporate fixed income securities. *Working paper, Wharton School, University of Pennsylvania*.

Korn, E. and R. Korn (1999). *Optionsbewertung und Portfolio-Optimierung*. Vieweg.

Lando, D. (1994). *Three Essays on Contingent Claim Pricing*. Ph. D. thesis, Cornell University.

Lando, D. (1998). On Cox processes and credit risky securities. *Review of Derivatives Research 2*, 99–120.

Lehrbass, F. (1997). Defaulters get intense. *Risk - Credit Risk Supplement, July*, 56–59.

Leland, H. E. and K. Toft (1996). Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads. *Journal of Finance 51*, 987–1019.

Li, D. (2000). On default correlation: A copula function approach. *Journal of Fixed Income 9*, 43–54.

Longstaff, F. and E. Schwartz (1995). A simple approach to valuing risky fixed and floating rate debt. *Journal of Finance 50*, 789–819.

Madan, D. B. and H. Unal (1998). Pricing the risks of default. *Review of Derivatives Research 2*, 121–160.

Mason, S. P. and S. Bhattacharya (1981). Risky debt, jump processes, and safety covenants. *Journal of Financial Economics 9*, 281–307.

Merton, R. (1974). On the pricing of corporate debt: the risk structure of interest rates. *Journal of Finance 29*, 449–470.

Musiela, M. (1993). Stochastic PDEs and term structure models. *Journées Internationales de France, IGR-AFFI, La Baule*.

Nelsen, R. B. (1999). *An introduction to copulas*, Volume 139 of *Lecture Notes in Statistics*. Springer Verlag. Berlin Heidelberg New York.

Neugebauer, O. (1969). *Vorlesungen über Geschichte der antiken Mathematischen Wissenschaften, Bd. I.* Springer Verlag. Berlin Heidelberg New York.

Nielsen, T. N., J. Saà-Requejo and P. Santa-Clara (1993). Default risk and interest rate risk: The term structure of default spreads. *Working paper*.

Özkan, F. and T. Schmidt (2003). Credit risk with Lévy random fields. *Working paper*.

Pang, K. (1998). Calibration of Gaussian Heath, Jarrow and Morton and random field interest rate term structure models. *Review of Derivatives Research 4*, 315–346.

Pechtl, A. (1996). A synopsis of certain option pricing formulae. *Risk Publications*.

Protter, P. (1992). *Stochastic Integration and Differential Equations*. Springer Verlag. Berlin Heidelberg New York.

Pye, G. (1974). Gauging the default premium. *Financial Analysts Journal 30*, 49–52.

Rebonato, R. (1996). *Interest-rate Option Models*. John Wiley & Sons. New York.

Reed, M. and B. Simon (1974). *Methods of Modern Mathematical Physics I: Functional Analysis*. Academic Press, New York and London.

Revuz, D. and M. Yor (1994). *Continuous Martingales and Brownian Motion*. Springer Verlag. Berlin Heidelberg New York.

Rogers, L. C. G. (1995). Which model for the term structure of interest rates should one choose ? In D. Duffie and S. E. Shreve (Eds.), *Proceedings of the IMA Workshop on Mathematical Finance*, pp. 93–116. Springer Verlag. Berlin Heidelberg New York.

Rolski, T., H. Schmidli, V. Schmidt and J. Teugels (1999). *Stochastic Processes for Insurance and Finance*. John Wiley & Sons. New York.

Rutishauser, H. (1976). *Vorlesungen über Numerische Mathematik, Band 2*. Birkhäuser Verlag.

Santa-Clara, P. and D. Sornette (1997). The dynamics of the forward interest rate curve with stochastic string shocks. *Review of Financial Studies 14*, 149–185.

Schmid, B. (2002). *Pricing Credit Linked Financial Instruments*, Volume 516 of *Lecture Notes in Economics and Mathematical Systems*. Springer Verlag. Berlin Heidelberg New York.

Schmidt, W. M. (1997). On a general class of one-factor models for the term structure of interest rates. *Finance and Stochastics 1*, 3–24.

Schmidt, W. M. and I. Ward (2002). Pricing default baskets. *Risk*, January 111–114.

Schönbucher, P. (2000). A LIBOR market model with default risk. *Working paper*.

Schönbucher, P. (2002). A tree implementation of a credit spread model for credit derivatives. *Journal of Computational Finance 6*.

Shimko, D., N. Tejima and D. van Deventer (1993). The pricing of risky debt when interest rates are stochastic. *Journal of Fixed Income 3*, 58–66.

Sobehart, J. and R. Klein (2000). Moody's public firm risk model: A hybrid approach to modeling short term default risk. *Moody's investors service*. [http://riskcalc.moodysrms.com/us/research/crm/53853.asp].

Tavakoli, J. (1998). *Credit Derivatives*. John Wiley & Sons. New York.

Vargiolu, T. (2000). Calibration of the Gaussian Musiela model using the Karhunen-Loeve expansion. *Working paper*.

Vasiček, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics 5*, 177–188.

Vasiček, O. (1984). Credit valuation. *Working paper, KMV Corporation*.

Wei, D. G. and D. Guo (1991). Pricing risky debt: An empirical comparison of the Longstaff and Schwartz and Merton models. *Journal of Fixed Income 7*, 8–28.

Werner, D. (2000). *Funktionalanalysis*. Springer Verlag. Berlin Heidelberg New York.

Yor, M. (1974). Sur les intégrales stochastiques à valeurs dans un espace de Banach. *Ann. Inst. H. Poincar'e Sect. B (N.S.)*, 31–36.

Yosida, K. (1971). *Functional Analysis*. Springer Verlag. Berlin Heidelberg New York.

Zhou, C. (1997). A jump-diffusion approach to modeling credit risk and valuing defaultable securities. *Finance and Economics Discussion Paper Series 1997/15, Board of Governors of the Federal Reserve System*.

Index

Erklärung

Hiermit erkläre ich, dass ich die Arbeit selbständig verfasst und nur die angegebenen Hilfsmittel verwendet habe.

Gießen, den 21.07.2003