

Shifted Regime Switching CIR Diffusion Tree for Credit Options

Inaugural dissertation in partial fulfillment of the requirements
for the degree of Doctor rerum naturalium (Dr. rer. nat.)

Faculty of Mathematics and Computing Science, Physics, and
Geography of the Justus-Liebig-University Giessen

by

Johannes Weckend

January 2014

Referees

Prof. Dr. Ludger Overbeck

Prof. Dr. Winfried Stute

Contents

List of Figures	vii
List of Tables	ix
List of Symbols	xi
1 Introduction	1
2 Definitions and Basics	7
2.1 Mathematical Basis	7
2.2 Interest Rate and its Products	7
2.3 Price System and Risk-Neutral Measure	10
2.4 Default Time	11
2.5 Sub-Filtration Approach	13
2.6 Credit Products	15
3 Interest Rate and Default Intensity Model	19
3.1 Short Rate Model - CIR Model and Extensions	19
3.1.1 Derivation of the CIR Diffusion Process	19
3.1.2 Properties of the CIR Process	21
3.1.3 Feller's Stability Condition	23
3.1.4 Extension of CIR Model: CIR++	24
3.1.5 Addition of Jumps to the Model: CIR-EJ++	27
3.2 Two-Dimensional CIR(-EJ)++ Model	30
3.3 Calibration to Market Quotes - Term Structure	31
3.4 Volatility Calibration of Interest Rates	33
3.4.1 Market Model	33
3.4.2 CIR++ Model	34
3.4.3 Calibration Results	35
3.5 Volatility Calibration of Default Intensity	37
3.5.1 Market Model	38
3.5.2 CIR++ and CIR-EJ++ Model	40
3.5.3 Calibration Results	45
3.5.4 CDS Option Price in the CIR(-EJ)++ Model - Adaption to 3.5.1 Market Model	48

3.5.5	Analytic Price for the Contingent Credit Line in the CIR(-EJ)++ Model	49
4	Regime Switching	55
4.1	Statistical Analysis of CDS Spreads	55
4.1.1	Summary Statistics	56
4.1.2	Normal Mixture Distribution	56
4.1.3	Regime Switching EM Algorithm	58
4.1.4	Regime Switching Normal Distribution	61
4.1.5	Detecting Regime Switchings	63
4.2	Regime Switching Model	64
4.2.1	General Regime Switching Framework	64
4.2.2	RS CIR(-EJ)++ Model	66
5	Numerical Implementation to a Trinomial Tree	69
5.1	Recombining of CIR Process	70
5.1.1	To Find a Recombining Auxiliary Process \tilde{x}	71
5.1.2	Second Recombining Process \tilde{x}	72
5.1.3	Differences in the Choices of \tilde{x}	73
5.2	CIR Trinomial Tree	73
5.2.1	Building a Tree - Motivation and Convergence	73
5.2.2	The Approach of NAWALKHA & BELIAEVA (2007) (NB)	76
5.2.3	The Approach of BRIGO & MERCURIO (2006) (BM)	80
5.2.4	Differences between the Tree Methods	83
5.3	Tree Building - Handling of Jumps, Regime-Switching and others	85
5.3.1	Tree with Exponential Jumps	85
5.3.2	Regime Switching Tree	87
5.3.3	The Shift Term in the Tree	92
5.3.4	Adding a Default Branch to the Tree	94
5.3.5	Building the Tree - the Steps	95
5.4	Combining the Interest and Default Trees	96
5.4.1	The Independent Case	97
5.4.2	The Case with Correlation	97
5.4.3	The HULL & WHITE (1994b) Approach	97
5.4.4	The Copula Approach	99
5.5	Tree Valuation	101
6	Results	103
6.1	CDS Call Option Results	104
6.1.1	Results from Applying the Calibrated Models	104
6.1.2	Results in Case of Valid Feller Condition	108
6.2	Loan and Contingent Credit Line Results	111
6.2.1	Results from Applying the Calibrated Models	111

6.2.2 Results in Case of Valid Feller Condition	114
7 Conclusion	117
Bibliography	119
Acknowledgments	125

Contents

List of Figures

1.1.1 iTraxx® Main 5 year historic spread changes (6th July 2012, Copyright©2012 Markit Group Limited)	2
3.1.1 Probability density function for CIR and CIR-EJ	28
3.3.1 Zero interest rates (6th July 2012, Copyright©2012 Market Data Distribution Service, Commerzbank AG)	31
3.3.2 Default intensity of Allianz CDS quotes	32
3.4.1 Implied cap volatilities (6th July 2012, Copyright©2012 Bloomberg L.P.)	35
3.4.2 Sensitivities of CIR interest rate parameters	36
3.4.3 Cumulative density distribution of CIR processes	37
3.5.1 Default intensity of market, CIR and CIR-EJ (Allianz)	45
3.5.2 Implied CDS option volatilities (6th July 2012, Copyright©2012 Bloomberg L.P.)	46
3.5.3 Sensitivities of CIR default intensity parameters	47
3.5.4 Sensitivities of CIR-EJ default intensity parameters	48
4.1.1 Conditional probabilities for state one of two (good economy state)	64
4.1.2 Conditional probabilities for iTraxx® Main with three regimes	65
5.1.1 The CIR tree - non recombining	70
5.1.2 Recombining tree transformation	72
5.2.1 Probability density of CIR processes (lh) and tree convergence (rh)	76
5.2.2 The NB tree with multiple node movements	77
5.2.3 The BM tree with different time steps	80
5.2.4 Convergence of different CIR trees to bond prices	84
5.3.1 Exponential jump distribution and its approximation	86
5.3.2 Convergence of CIR-EJ trees to bond price.	87
5.3.3 Regime switching tree	88
5.3.4 Movements on lowest nodes in RS tree	91
5.3.5 Incorporating default branch in the tree	94
5.4.1 Combined tree	96
5.4.2 Probability calculation in the copula model	100
6.1.1 CDS option prices under the RS CIR model	107

List of Figures

6.1.2 CDS option prices under the RS CIR-EJ model	108
6.1.3 CDS option prices under the RS CIR model	110
6.1.4 CDS option prices under the RS CIR-EJ model	110

List of Tables

3.3.1 CDS quotes (in bps) (6th July 2012, Copyright©2012 Markit Group Limited)	32
4.1.1 Summary statistics of indices (Copyright©2012 Markit Group Limited)	57
4.1.2 Augmented Dickey-Fuller test	57
4.1.3 Test results for distribution of spread changes	57
4.1.4 Results for regime switching normal distribution with two regimes (standard errors in parentheses)	61
4.1.5 Daily transition probabilities for two regimes	62
4.1.6 Results for regime switching normal distribution with three regimes (standard errors in parentheses)	62
4.1.7 Daily transition probabilities for three regimes	63
5.4.1 Probabilities of the combined trees without correlation	97
5.4.2 Probabilities of the combined trees with positive correlation	98
5.4.3 Probabilities of the combined trees with negative correlation	98
5.4.4 Probabilities of the combined trees in case of two branches	99
5.4.5 Probabilities of the combined trees with copula	101
6.0.1 Parameters in the calibrated model	103
6.0.2 Parameters in case of valid Feller condition	104
6.1.1 CDS option price in calibrated model convergence	105
6.1.2 CDS option prices in calibrated model under correlation	105
6.1.3 CDS option prices in calibrated model under jumps	106
6.1.4 CDS option prices under correlation	109
6.1.5 CDS option prices under jumps	109
6.2.1 Loan prices in the CIR++ model (in bps)	112
6.2.2 Loan prices under jumps (in bps)	112
6.2.3 Loan prices in the RS CIR++ model (in bps)	113
6.2.4 Loan prices in the RS CIR-EJ++ model (in bps)	113
6.2.5 Loan prices in the CIR++ model (in bps)	114
6.2.6 Loan prices under jumps (in bps)	114
6.2.7 Loan prices in the RS CIR++ model (in bps)	115
6.2.8 Loan prices in the RS CIR-EJ++ model (in bps)	115

List of Tables

List of Symbols

Roman Symbols			
$A(t)$	Generator matrix at time t	K	Strike, swap rate
$A(t, T)$	Factor in bond price formula	$L(t, T)$	Spot interest rate at time t maturing at time T
$B(t)$	Bank account at time t	M	Market value, mean of tree node
$B(t, T)$	Exponent in bond price formula	N	Notional
b	NB tree parameter	$N(t; \lambda)$	Cox process at time t with intensity λ
C	Coupon	$\bar{N}(t; \varsigma)$	Jump arrival rate of default intensity process
$C_{a,b}(t)$	Present value per basis point	$P(t, T)$	Zero-coupon bond at time t for maturity T
$\hat{C}_{a,b}(t)$	Defaultable present value per basis point	$\hat{P}(t, T)$	Defaultable zero-coupon bond at time t for maturity T
$D(t, T)$	Stochastic discount factor at time t for maturity T	p	Default probability in the tree
e_i	Unit vector	p_{ij}	Transition probability from state i to state j
F	Commitment fee	p_u, p_m, p_d	Tree probabilities for up, middle and down movements
$F(t, T, S)$	Forward interest rate at time t for expiry T and maturity S	$Q(t, T)$	Transition probability matrix between t and T
$f(x)$	Probability density function	$q(t)$	Risk-neutral default probability at time t
$f(t, T)$	Instantaneous forward interest rate at time t for time T	$R_{a,b}(t)$	Forward CDS rate at time t
$\hat{H}(t, T)$	Survival probability at time t up to time T	$r(t)$	Instantaneous spot interest rate at time t
h	Auxiliary value in CIR formulas	$S(t)$	Hidden Markov chain process at time t
$h(t)$	Deterministic function in option pricing	$S_{a,b}(t)$	Forward swap rate at time t
I_q	Modified Bessel function of the first kind of the order q	V	Variance of tree node
J	Exponential jump distribution, tree parameter to determine the next nodes	$V(t)$	Price process at time t

$\Pi(t)$	Discounted payoff at time t	Mathematical Symbols	
π_n^{is}	State price of node (n, i, s)	$\mathcal{C}_\rho^{\text{Ga}}$	Gaussian copula with correlation ρ
ρ	Correlation of Brownian motions of interest rate and default intensity	\mathbb{E}	Expected value under risk neutral measure
σ	Volatility parameter, standard deviation of normal distribution	\mathbb{F}	Filtration of market
ς	Jump arrival rate	\mathcal{F}	σ -algebra of market
τ	Default time	\mathbb{G}	Filtration of market and default
Φ	Standard normal CDF	\mathbb{H}	σ -algebra of market and default
φ	Deterministic shift function	\mathbb{I}	Filtration of default
χ^2	Non-central chi-square CDF	\mathcal{H}	σ -algebra of default
Ψ	Integrated deterministic shift function	$\mathcal{N}(\nu, \sigma^2)$	Normal distribution with mean μ and variance σ^2
Ω	Set of all possible states	\mathbb{P}	Real probability measure
Sub/Superscripts		\mathbb{Q}	Risk neutral probability measure
\cdot^{CIR}	Denotes the CIR process	\mathbb{R}	Real numbers
\cdot^{M}	Denotes a market value	\mathcal{S}	Finite state space
\cdot_{T}	Denotes the derivative $\frac{\partial \cdot}{\partial T}$	\mathcal{T}	Set of time points
$\hat{\cdot}$	Estimator (in Chapter 4)		

List of Symbols

1 Introduction

This thesis studies interest rate and single-name credit risk modeling. The focus is on the valuation of credit risk products such as credit default swaps (CDS), CDS options and loans. Loans are the most interesting products to price as they offer the main optionality. Among others, drawdown options, prepayment options and, in case of contingent credit lines, multiple drawing options can be included in a loan. A semi-analytical price is found for the valuation of the contingent credit line on a floating rate loan. The numerical model presented in this thesis can price any non-path dependent options in interest rate and single name default risk.

The modeling of the interest component is done by a short rate model offering a simple handling of the whole term structure curve with one single variable only. There is a large variety of short rate models with different properties. The CIR++ model is applied here because it prevents negative rates by definition, is analytically tractable and matches the term structure exactly. This model is an extension of the COX, INGERSOLL JR & ROSS (1985a) (CIR) model with a deterministic shift introduced by BRIGO & MERCURIO (2001). The CIR++ model provides analytical solutions for many products, and the deterministic shift is used to match the term structure exactly.

The default is modeled by a reduced form model implying that the default time is not predictable and the probability of default is handled by a default intensity. The default intensity is comparable with the short rate in the interest rate setting, especially in relation to the desired properties. Based on the reasons mentioned above the CIR++ model is as well used for the default intensity. In BRIGO & MERCURIO (2006) it is called shifted square-root diffusion model in case of default. Due to the large jumps in the credit market the model is extended by an exponential jump component. Exponential jumps keep the property of positivity, but do not enable the model to reflect large down movements of the default intensity which are also visible in specific market environments. The influence of this component on the valuation and especially on the optionality of the products will be studied.

The calibration to the market volatility is done by the standard procedure for the interest rate process. The relative differences between the implied market cap volatilities and the CIR implied volatilities are minimized. In case of default the CDS option is the product to calibrate the CIR default process. The semi-analytical formula of BRIGO & EL-BACHIR (2010) is used to calculate the prices in the CIR model. The volatilities are obtained through

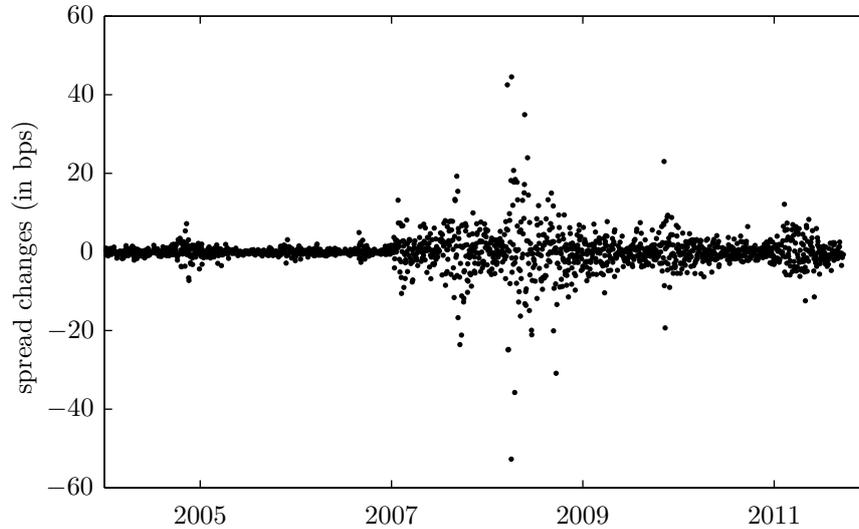


Figure 1.1.1: iTraxx® Main 5 year historic spread changes (6th July 2012, Copyright©2012 Markit Group Limited)

a Black market model formula for CDS options. The market liquidity and availability of market quotes are taken into consideration.

Figure 1.1.1 shows the absolute daily spread changes of the most liquid default product in Europe - the iTraxx® Main 5 year. The rolling series is used here corresponding to the changes of the newest series.

Obviously, in Fig. 1.1.1 the distribution of the spread changes is not the same over the whole period. There is at least one regime switching at the beginning of the credit crisis in the mid of 2007. Over a longer time period there are more regime switchings, and there are economic cycles with irregular economy state changes (e.g., HAMILTON (1989)).

These changes of regime switching are handled in the model by introducing a hidden Markov process to simulate the possibility of different distributions depending on the states. If there are two states only these may be interpreted as good and bad economy, but a finer grid of states is possible as well. This regime switching component affects both, the interest and the default component. Usually, in case of a bad economy, the interest rate decreases and the default intensity increases, and vice versa in a good economy.

Upon introduction of the regime switching into the CIR++ process the analytical tractability is lost. The calibration to the term structure is done by the deterministic shift, the volatility calibration is not possible in an analytical way.

A tree setup is used for the numerical implementation of the model specified above. There are several ways to build a tree for the CIR process. Two tree building procedures are compared

here: The first is the procedure of NELSON & RAMASWAMY (1990), further developed by NAWALKHA & BELIAEVA (2007), the second is the tree procedure presented in BRIGO & MERCURIO (2006). Advantages of each tree are shown and evaluated, and differences are discussed. In the end, the better of both trees is taken for valuation.

For both the interest rate and default intensity process a single tree is built which is combined with a fixed correlation afterwards as shown in SCHÖNBUCHER (2000b) for the Hull-White model. The correlation is incorporated into the processes in two different ways. In the standard procedure of HULL & WHITE (1994b) the tree probabilities get a positive (negative) shift depending on the correlation factor. On the edges of the trees the whole correlation is not applicable, because this would result in negative probabilities. The second incorporation approach is new and uses copulas to model the correlation between the two single trees.

The exponential jumps are simply added to the tree in a standard way. The different parameters resulting from regime switching states are incorporated to one tree using the same nodes. This is done in a new way by allowing that the diffusion process skips nodes in case of very high volatility differences for different regime states. This construction is numerically very efficient since it only requires the separate computation of the different probabilities for each state. Therefore the computation requires only the standard tree memory multiplied by the number of states.

Other important issues like hedging, market completeness, stochastic recovery rates or counterparty risks are not addressed.

Literature Review

This review of literature considers the different topics of this thesis and includes standard books, influential and specialized papers. It is a selection of favorite books and papers only and cannot claim to be complete.

A large number of books and papers has been published in the area of interest and credit risk. Starting generically with the interest rate modeling there are standard books as BRIGO & MERCURIO (2006) or NAWALKHA, SOTO & BELIAEVA (2007). Short rate models are the starting point of interest rate modeling. In these models the short rate is the only stochastic component. References are the fundamental papers of VASICEK (1977) and COX, INGERSOLL JR & ROSS (1985b) with time-homogenous processes and HULL & WHITE (1990) with time-varying parameters. These models have the advantages of easy handling and of analytical solutions for a lot of products. The drawbacks are that the market is more complex, and different market movements of forward curves are not replicable using one variable only. There are more developed versions, e.g., the HJM forward model of HEATH, JARROW & MORTON (1992) or the market model introduced by JAMSHIDIAN (1997) and BRACE, GATAREK & MUSIELA (1997). A more detailed literature review can be found in

GIBSON, LHABITANT & TALAY (2001). Since the interest rate plays a minor role for credit risk products only, it is modeled by a specific short rate model in this thesis.

Credit risk modeling is a more recent field of research. Standard books are BLUHM, OVERBECK & WAGNER (2008), BIELECKI & RUTKOWSKI (2004a) or LANDO (2004). Furthermore, the two papers JEANBLANC & RUTKOWSKI (2000a) and JEANBLANC & RUTKOWSKI (2000b) give a good introduction to the research field. There are at least three types to model the default: Structural models, reduced-form models and rating transition models. The structural models use directly the firm value process and therefore these are also called firm's value models. This approach originates from the famous work of MERTON (1974). Default is defined as the first hitting of a certain barrier which can be deterministic or stochastic. This approach very intuitively refers to economic fundamentals. Nonetheless its problems are the determination of the firm value (not a tradable security), the calibration to market data (fitting of default barrier) and the predictability of default. In the reduced-form model, the next type, the default is treated as an unpredictable event governed by a hazard rate process. In one version the default is defined as the first jump of a Cox process as introduced by LANDO (1998). Other representatives of the reduced form model are JARROW & TURNBULL (1995), DUFFIE & SINGLETON (1999) and SCHÖNBUCHER (1998). Compared to the structural models the calibration is easier for the reduced-form models, but the default intensity is less meaningful than the firm value. The rating transition models use the firm's rating process. This process is Markovian and determines the rating transitions and default according to the market transition probabilities. A standard model is shown in JARROW, LANDO & TURNBULL (1997), a model with stochastic credit spreads was examined by LANDO (1998). In this thesis, the reduced-form modeling is used together with the Cox settings. The advantage of this approach is the same treatment of the short rate and of the default intensity with one stochastic variable only.

There are various ways to model the short rate and the default intensity. Here, the CIR process of COX, INGERSOLL JR & ROSS (1985b) is applied. This process was already examined by FELLER (1951) but without the context of short rates. The advantages of this process are the simple handling and its simple calibration to a given term structure. There are a lot of extensions of the CIR process such as time-inhomogeneous parameters as in MAGHSOODI (1996) or deterministic shift extensions as in BRIGO & MERCURIO (2001). The latter one is used as it preserves the analytical tractability and matches the term structure exactly. Several papers describe the estimation of the parameters in the CIR model such as BROWN & DYBVIG (1986) or OVERBECK & RYDÉN (1997).

One difficulty of the default modeling is to find an (equivalent) measure such that the forward swap rate is a martingale with respect to this measure. The method used in this thesis is the sub-filtration approach of JAMSHIDIAN (2004) introduced by JEANBLANC & RUTKOWSKI (2000b). The filtration is split into two sub-filtrations: on the one hand the default information, and on the other all the remaining information, called market information. The advantage of this method is the unpredictability of default in the pricing formula

depending on the market filtration. This ensures positivity of a defaultable numeraire and therefore the existence of an equivalent measure under a defaultable numeraire. A further option is the pre-intensity approach of DUFFIE & SINGLETON (1999) which is also in line with the Cox process setting. The pre-intensity is the positive extension of the default intensity which is also applicable as numeraire. A third approach is stated in SCHÖNBUCHER (2003) and is called measure of survival. This measure is based on survival and therefore it is absolutely continuous only and not equivalent to the risk-neutral pricing measure.

The default process is calibrated based on CDS options. The market model is comparable to the famous work of BLACK (1976) where Black volatilities are translated into prices. The first papers in this regard came from HULL & WHITE (2003), JAMSHIDIAN (2004) and BRIGO & MORINI (2005). BRIGO & ALFONSI (2005) and BRIGO & EL-BACHIR (2010) derived analytical formulas for the CDS option price in the CIR++ model and its jump-extended version. The authors in BIELECKI, JEANBLANC & RUTKOWSKI (2011) obtain the appropriate replicating strategy in this model and a method to hedge CDS options.

As described in the introduction the default process has at least two different distributions. This property is handled by a regime switching process. Such an approach can be traced back to early work of LINDGREN (1978). It became popular after the seminal work of HAMILTON (1988). He compared a short rate model with constant parameters to one with regime switching parameters by application to the empirical behavior of US interest curves between 1962 and 1987. There are many empirical papers in the interest rate world such as GRAY (1996) or ANG & BEKAERT (2002). The CIR model has been studied in WONG & WONG (2008). In credit risk there also exist empirical studies such as DAVIES (2004) and MAALAOU CHUN, DIONNE & FRANÇOIS (2010a). SIU, ERLWEIN & MAMON (2008) use the regime switching technique in a structural model pricing CDS. ALEXANDER & KAECK (2008) have already examined the regime dependency of the iTraxx® Main in a regression model. ELLIOTT, AGGOUN & MOORE (1995) is a newsworthy book on hidden Markov models.

The numerical valuation procedure is performed in trees. COX, ROSS & RUBINSTEIN (1979) showed the weak convergence for the stock price binomial model to the continuous process. A good starting point is given in the book of SHREVE (2005). The CIR process is not recombining and will end in a highly inefficient and memory intensive tree. This problem was solved first by NELSON & RAMASWAMY (1990) in a binomial model. To extend this approach to a trinomial tree, BRIGO & MERCURIO (2006) and NAWALKHA & BELIAEVA (2007) found two different ways which will be discussed in this thesis. These two approaches have an advantage over the earlier extensions described in LI, RITCHKEN & SANKARASUBRAMANIAN (1995) or ACHARYA & CARPENTER (2002). Therefore, the latter two are not discussed here.

For other short rate models as the Hull-White model we refer to HULL & WHITE (1994a) and for a two-factor model to HULL & WHITE (1994b). The first handling of regime switching in a tree model was done for two regimes in BOLLEN (1998) by modeling a log-normal stock price process. It was enforced for multi regime cases in YUEN & YANG (2010). Until now a

regime switching tree was neither developed for a default intensity process nor for the CIR model.

In this thesis, the interest and the default components are modeled in a two-dimensional tree where each component has one axis. This was first done by SCHÖNBUCHER (2000b) in the Hull-White model. GARCIA, VAN GINDEREN & GARCIA (2003) followed this approach and proposed the non-negative Black-Karazinsky model for the default intensity. Moreover, there exist two papers, XU & LI (2009a) and XU & LI (2009b), which use a two-dimensional CIR tree wherein the second axis models the equity component. The equity is then taken to model the default and to price CDS and credit spread options.

The main focus of this thesis is to price optionality in credit risk and especially in loans. LOUKOIANOVA, NEFTCI & SHARMA (2007) priced a contingent credit line using the structural model. In ENGELMANN (2011) loans with embedded options are priced in a rating transition model. LEUNG & YAMAZAKI (2010) investigated step-up and down options using a Levy process. A subsidiary aspect should also be considered, the modeling of swing options as modeled in CARMONA & TOUZI (2008). Swing options can be seen as contingent claims with multiple exercise rights of Bermudan or American style.

Structure of the Dissertation

This thesis is divided into seven chapters including the introduction. Chapter 2 contains already known results and is an introduction to the subject of this thesis, including the mathematical basis, product definitions, pricing system, default time, and sub-filtration approach. In Chapter 3 an introduction to the CIR processes is given. Additionally the applied single regime processes are presented with the corresponding calibration methods. Chapter 4 contains a statistical analysis and the handling of regime switching, which results in the RS CIR-(EJ)++ model for interest rate and default intensity. In Chapter 5 the tree implementation is presented. It starts with two different procedures of CIR trees. Afterwards the best of both trees is enforced to handle the jump component, the regime-switching, and the shifting component. Finally, the combination and valuation of the two-dimensional tree is performed. Chapter 6 contains the numerical results, and in Chapter 7 the conclusions are drawn.

2 Definitions and Basics

This chapter contains the basics about the instruments, the pricing and the CIR process. It starts with the mathematical basis. Afterwards the definitions of the applied interest and credit instruments are given. Moreover, in this chapter the pricing system, the default handling and the sub-filtration approach are presented in detail.

2.1 Mathematical Basis

The mathematical basis is the probability space. A probability space $(\Omega, \mathcal{G}, \mathbb{P})$ is considered in which Ω is the set of all possible states, \mathcal{G} is the σ -algebra containing all possible events of interest and $\mathbb{P} : \mathcal{G} \rightarrow [0, 1]$ is the real or historical probability measure. The probability space is assumed to be equipped with a filtration $\mathbb{G} = (\mathcal{G}_t)_t$. A filtration is a non-decreasing family of σ -subalgebras of \mathcal{G} such that $\mathcal{G}_s \subseteq \mathcal{G}_t \subseteq \mathcal{G}$ for all $0 \leq s \leq t$. The filtrated probability space $(\Omega, \mathcal{G}, \mathbb{P}, \mathbb{G})$ is assumed to satisfy the usual conditions. The time horizon is a finite interval $[0, T_\infty]$. The current time is set to $t = 0$, while \mathcal{G}_0 is the trivial σ -subalgebra and $\mathcal{G} = \mathcal{G}_{T_\infty}$.

Each \mathcal{G}_t describes the information available at $t \in [0, T_\infty]$. It is assumed that $\mathcal{G}_t = \mathcal{H}_t \vee \mathcal{F}_t$ is the union of two σ -subalgebras \mathcal{H}_t and \mathcal{F}_t for each $t \in [0, T_\infty]$. The first sub-filtration $\mathbb{H} := (\mathcal{H}_t)_t$ carries the information about the default time. The second sub-filtration $\mathbb{F} := (\mathcal{F}_t)_t$ represents all other market information available to traders such as interest rates and default intensities. The formal definition is stated in Section 2.4. As in the case of \mathcal{G} it follows that $\mathcal{H} = \mathcal{H}_{T_\infty}$ and $\mathcal{F} = \mathcal{F}_{T_\infty}$.

For stochastic processes, the state $\omega \in \Omega$ is omitted and $X(t)$ is written instead of $X(t, \omega)$. Whether the process is deterministic or stochastic should be clear from context.

2.2 Interest Rate and its Products

Here, the interest rate and its products such as bonds, forwards, swaps and caps are defined. For further details a good reference is BRIGO & MERCURIO (2006), among others.

Definition 2.2.1. (*Bank account*) The value of a bank account is $B(t)$ at time $t \geq 0$. The bank account follows the process

$$dB(t) = r(t) B(t) dt, \quad B(0) = 1$$

where the interest rate $r(t)$ is a random process of time greater than zero.

Investing one monetary unit at time zero yields $B(t) = \exp\left(\int_0^t r(u) du\right)$ monetary units at time t . The rate r at which the bank account accrues is referred to as the instantaneous spot rate, or as the short rate. The value of the bank account or the short rate defines the discount factor.

Definition 2.2.2. (Stochastic discount factor) The (stochastic) discount factor $D(t, T)$ is the amount at t that is equivalent to one monetary unit payable at time T , and is given by

$$D(t, T) := \frac{B(t)}{B(T)} = \exp\left(-\int_t^T r(u) du\right).$$

In many pricing applications r is assumed to be a deterministic function of time. Here, this is true for the volatility calibration of the default process only. However, in the product pricing r is modeled by a stochastic process in which case the bank account and the discount factors will be stochastic processes as well. This shows the effects of stochastic interest rates to credit products.

Definition 2.2.3. (Zero-coupon bond) A zero-coupon bond (or discount bond) with maturity T is a contract that guarantees its holder the payment of one monetary unit at time T . The contract value at time $t < T$ is denoted as $P(t, T)$. Obviously, $P(T, T) = 1$ for all T .

If the interest rate r is deterministic, D is deterministic as well, and $D(t, T) = P(t, T)$ for each pair (t, T) . However, in case of stochastic rates, $D(t, T)$ is a random quantity at t depending on the future evolution of rates r between t and T . Instead, the zero-coupon bond price $P(t, T)$ is known at t . $P(t, T)$ is the expected value of the random variable $D(t, T)$ under a specific probability measure.

Definition 2.2.4. (Zero-bond curve) The zero-bond curve at time t is given by the graph of the function

$$T \rightarrow P(t, T), \quad T > t.$$

The graph is a function decreasing with T starting from $P(t, t) = 1$, because interest rates are positive quantities. Such a curve is also referred to as the term structure of discounted factors.

Definition 2.2.5. (Spot interest rate) The spot interest rate $L(t, T)$ at time t and maturing at time T is the interest rate for this period:

$$L(t, T) := \frac{1}{\alpha(t, T)} \left(\frac{1}{P(t, T)} - 1 \right) = \frac{1 - P(t, T)}{\alpha(t, T) P(t, T)} \quad (2.2.1)$$

The $\alpha(t, T)$ is the length of the period between t and T in years as given by the day count convention.

Definition 2.2.6. (Simply-compounded forward interest rate) The simply-compounded forward interest rate $F(t, S, T)$ at time t for expiry S and maturity T is the future zero rate implied by today's term structure of the interest rate:

$$F(t, S, T) := \frac{1}{\alpha(S, T)} \left(\frac{P(t, S)}{P(t, T)} - 1 \right) = \frac{P(t, S) - P(t, T)}{\alpha(S, T) P(t, T)} \quad (2.2.2)$$

The forward interest rate and the spot rate are equal for $t = S$.

Definition 2.2.7. (Instantaneous forward interest rate) The instantaneous forward interest rate $f(t, T)$ is the forward rate for the infinitesimal small time period starting at T as seen at t :

$$f(t, T) := \lim_{T \rightarrow S^+} F(t, S, T) = -\frac{\partial \ln P(t, T)}{\partial T}$$

This definition implies that

$$P(t, T) = \exp \left(- \int_t^T f(t, u) du \right).$$

The products defined below are more complex. Because of a missing pricing system and for the sake of simplicity, the discounted payoffs denoted by Π are given. The pricing measure is introduced in Section 2.3.

Definition 2.2.8. (Zero-coupon bond option) A zero coupon bond call/put option gives the holder the right to buy/sell a zero-coupon bond at strike K . The option date and the bond maturity are denoted by $S \geq t$ and $T \geq S$, respectively. The discounted payoffs at time t are equal to

$$\begin{aligned} \Pi_{\text{ZBC}(S, T; K)}(t) &:= D(t, S) \cdot (P(S, T) - K)^+ \quad (\text{call option}) \\ \Pi_{\text{ZBP}(S, T; K)}(t) &:= D(t, S) \cdot (K - P(S, T))^+ \quad (\text{put option}) \end{aligned}$$

The option values depend on the volatility of the interest rate r . The higher the volatility the higher the option value. The bond prices in the option formula are not known at time t , $t < S$.

Definition 2.2.9. (Interest rate swap) An interest rate swap (IRS) is the swap of fixed and floating payments for a set of times $\mathcal{T}_{a,b} = \{T_a, \dots, T_b\}$ and year fractions $\bar{\alpha} := \{\alpha_{a+1}, \dots, \alpha_b\}$; $\alpha_i := \alpha(T_{i-1}, T_i)$. The swap rate is denoted by K and payed on the notional N . The floating payments are the spot interest rates at the settlement dates. The discounted payoff at time t is given by

$$\Pi_{\text{IRS}_{a,b}(K)} := \pm N \sum_{i=a+1}^b \alpha_i D(t, T_i) (L(T_{i-1}, T_i) - K)$$

where \pm depends on the direction of the swap.

The fair swap rate is chosen such that the value of the interest rate swap is zero in the beginning. This is satisfied if fixed and floating legs are equal.

Definition 2.2.10. (*Forward swap rate*) The forward swap rate $S_{a,b}(t)$ at time t for the set of times $\mathcal{T}_{a,b}$ and the year fractions $\bar{\alpha}$ is the rate at which the interest rate swap is fair.

The value of the interest rate swap depends on the price system. Using the risk-neutral price measure defined in Eq. (2.3.1), the forward swap rate is computed as

$$S_{a,b}(t) := \frac{P(t, T_a) - P(t, T_b)}{\sum_{i=a+1}^b \alpha_i P(t, T_i)}. \quad (2.2.3)$$

The denominator of the forward swap rate is a positive portfolio of zero coupon bonds

$$C_{a,b}(t) := \sum_{i=a+1}^b \alpha_i P(t, T_i) \quad (2.2.4)$$

which is called the present value per basis point (PVBP).

Definition 2.2.11. (*Cap*) A cap is the sum of options on the payments made in one payer IRS. The discounted payoff at time t for the time set $\mathcal{T}_{a,b}$, the year fractions $\bar{\alpha}$ and the strike K is given by

$$\Pi_{\text{Cap}_{a,b}(K)}(t) := N \sum_{i=a+1}^b \alpha_i D(t, T_i) (L(T_{i-1}, T_i) - K)^+.$$

The summands are called caplets. The cap gives its holder the right to decide on each settlement date to swap the spot rate against the fixed payment. The market quotes of caps are used to calibrate the volatility components in the interest rate model.

2.3 Price System and Risk-Neutral Measure

Section 2.3 introduces the pricing system and the change of measure technique. Its content is standard and follows in some aspects JAMSHIDIAN (2004).

Assumption 2.3.1. (*State price deflator*) A state price deflator $\xi(t) \geq 0$, $t \in [0, T_\infty]$ exists such that for all $t \leq T \leq T_\infty$

$$t \rightarrow \xi(t) D(t, T)$$

is a \mathbb{P} -martingale.

The state price deflator defines a pricing formula. The price process of a secure payoff V at time T is the unique process $V(t)$, $t \leq T$ such that $V(T) = V$ and $\xi(t) V(t)$ is a \mathbb{P} -martingale. A price process n is called numeraire if this process is a.s. strictly positive. The

associated numeraire measure \mathbb{P}^n is defined as

$$\frac{d\mathbb{P}^n}{d\mathbb{P}} := \frac{\xi_n}{n(0)}.$$

Taking the bank account B as numeraire defines the risk-neutral measure \mathbb{Q} . The expected value \mathbb{E} is the expected value under \mathbb{Q} if no other measure is stated. The price process of each product under the risk-neutral measure is then given by

$$V(t) = \mathbb{E}[D(t, T) \cdot V | \mathcal{G}_t]. \quad (2.3.1)$$

The risk-neutral measure is an equivalent martingale measure as the discounted price process is a martingale under this measure by definition of the state price deflator. The existence of an equivalent martingale measure ensures the non-existence of arbitrage.

The pricing formula (2.3.1) shows that the price is the expected value of the discounted payoff under the risk-neutral measure.

2.4 Default Time

In this thesis the default time is defined by a reduced form model. Following LANDO (1998), the first jump of a Cox process defines the default time. This definition is one special instance of the sub-filtration approach introduced in the beginning. It is discussed in more detail in Section 2.5. The default is always the default of one specific entity just considered.

An inhomogeneous Poisson process N with a non-negative and deterministic intensity function λ satisfies

$$\mathbb{Q}(N(t; \lambda) - N(s; \lambda) = k) = \frac{\left(\int_s^t \lambda(u) du\right)^k}{k!} \exp\left(-\int_s^t \lambda(u) du\right)$$

which for $N(0) = 0$ corresponds to

$$\mathbb{Q}(N(t; \lambda) = 0) = \exp\left(-\int_0^t \lambda(u) du\right).$$

Definition 2.4.1. (*Cox process*) N is called a Cox process, if there is a non-negative and \mathbb{G} -predictable stochastic process $\lambda(t)$ (intensity) with $\int_0^t \lambda(u) du < \infty$ for all $t > 0$, and conditional on the realization $\{\lambda(t)\}_{\{t>0\}}$ of the intensity, $N(t; \lambda)$ is a time-inhomogeneous Poisson process with intensity $\lambda(t)$.

A Cox process is a generalization of the inhomogeneous Poisson process including a stochastic intensity.

Assumption 2.4.2. (Time of default)

- The time of default is the first jump of the Cox process $N(t; \lambda)$ with intensity parameter $\lambda(t)$.
- The time of default will be referred to as τ .

The default intensity λ (sometimes also called the hazard rate) is the probability of default in the next infinitesimal interval provided there was no earlier default.

$$\lim_{h \downarrow 0} \frac{1}{h} \mathbb{Q} \{ \tau \in [t, t+h) \mid \tau \geq t, \mathcal{F}_t \} = \lambda(t)$$

The cumulated intensity or hazard function is the integral over the time interval of the default intensity:

$$\Lambda(t) := \int_0^t \lambda(s) ds$$

The default time defined as the first jump of a Cox process is equivalent to

$$\tau = \inf \{ t \in \mathbb{R}_+ \mid \Lambda(t) \geq \varsigma, \varsigma \text{ random variable with unit exponential distribution} \}. \quad (2.4.1)$$

Note. By comparing the intensity λ with the short rate r one can see the same effect of discounting by the following two equations:

$$P(0, t) = \mathbb{E} \left(e^{-\int_0^t r(u) du} \right) \quad (2.4.2)$$

$$\mathbb{Q}(\tau > t) = \mathbb{E}(\mathbf{1}_{\{\tau > t\}}) = \mathbb{E} \left(e^{-\int_0^t \lambda(u) du} \right) \quad (2.4.3)$$

where the second equation is the survival probability up to time t .

The survival probability \hat{H} at time t up to time T is defined as

$$\hat{H}(t, T) := \mathbb{E}[\mathbf{1}_{\{\tau > T\}} \mid \mathcal{G}_t]. \quad (2.4.4)$$

Remark. This short remark deals with the relation between risk-neutral default probabilities and the default intensity as in HULL & WHITE (2000). Let $q(t)$ be the risk-neutral default probability at time t in the next infinitesimal interval as seen from time zero. Then the survival probability until t using q is given by

$$\mathbb{Q}(\tau > t) = 1 - \int_0^t q(u) du. \quad (2.4.5)$$

The difference between the risk-neutral default probability and the default intensity is the condition of no earlier default. Equating and differentiating (2.4.3) and (2.4.5) yields

$$q(t) = \mathbb{E} \left(\lambda(t) \exp \left(- \int_0^t \lambda(u) du \right) \right).$$

The pricing formula (2.3.1) changes upon incorporation of default. The expectation of a discounted payoff depending on T and additionally on τ is given by:

$$\mathbf{1}_{\{\tau > t\}} V(t) = \mathbb{E} \left(D(t, T) \mathbf{1}_{\{\tau > T\}} V + D(t, \tau) \mathbf{1}_{\{t < \tau \leq T\}} V(\tau) \mid \mathcal{G}_t \right).$$

The proof is based on the stopping time property of the default time. The value is composed of the survival case up to time T and the recovery payment $V(\tau)$ in case of default between t and T .

2.5 Sub-Filtration Approach

Here, the sub-filtration approach gets more formal by using the definition of the default time introduced before. As stated in the beginning the filtration \mathbb{G} is divided into two sub-filtrations as $\mathbb{G} = \mathbb{H} \vee \mathbb{F}$. The sub-filtration \mathbb{H} contains information on the default time only, and its σ -subalgebras are defined as $\mathcal{H}_t := \sigma(\{\tau < u\}, u \leq t)$ for every $t \in \mathbb{R}$. Consequently τ is an \mathbb{H} -stopping time as well as a \mathbb{G} -stopping time. The second sub-filtration \mathbb{F} contains all other market information such as interest rates and default intensities. This means that the processes $r(t)$ and $\lambda(t)$ are \mathcal{F}_t -measurable for every $t \in \mathbb{R}$, and in particular that τ is not an \mathbb{F} -stopping time.

In case of the default time definition in 2.4.2 this means that the Cox process, the default intensity and the default time are all \mathbb{G} -adapted, that the Cox process and default intensity also are \mathbb{F} -adapted, and that the default time also is \mathbb{H} -adapted. In case of Eq. (2.4.1) the random variable ς is independent of the filtration \mathbb{F} .

The importance of the sub-filtration approach becomes evident in the quantity

$$\mathbb{Q}(\tau > t \mid \mathcal{F}_t).$$

As this term will often appear in the denominator it is necessary to have a non-zero term. By not including the default time the sub-filtration \mathbb{F} always ensures the positivity of the term. This was first introduced and discussed in ELLIOTT, JEANBLANC & YOR (2000) and JEANBLANC & RUTKOWSKI (2000b). One important corollary is reproduced from JEANBLANC & RUTKOWSKI (2000b), Lemma 3.2 which shows the connection of the filtrations \mathbb{G} and \mathbb{F} in terms of conditional expectations. Additionally, the corollary is Lemma 5.1.2 from BIELECKI & RUTKOWSKI (2004a). It is applied to $\mathbf{1}_{\{\tau > T\}} Y$.

Corollary 2.5.1. *For any \mathcal{G} -measurable random variable Y the following equation applies for any $0 < t \leq T$:*

$$\mathbb{E} [\mathbf{1}_{\{\tau > T\}} Y \mid \mathcal{G}_t] = \frac{\mathbf{1}_{\{\tau > t\}}}{\mathbb{Q}(\tau > t \mid \mathcal{F}_t)} \mathbb{E} [\mathbf{1}_{\{\tau > T\}} Y \mid \mathcal{F}_t]$$

Proof. It can be shown that any \mathcal{G}_t -measurable random variable \tilde{Y} for the set $\{\tau > t\}$ coincides with some \mathcal{F}_t -measurable random variable. Therefore,

$$\mathbb{E} [\mathbf{1}_{\{\tau > t\}} \tilde{Y} | \mathcal{G}_t] = \mathbf{1}_{\{\tau > t\}} \mathbb{E} [\tilde{Y} | \mathcal{G}_t] = \mathbf{1}_{\{\tau > t\}} X$$

where X is an \mathcal{F}_t -measurable random variable. Taking conditional expectation with respect to \mathcal{F}_t yields

$$\mathbb{E} [\mathbf{1}_{\{\tau > t\}} \tilde{Y} | \mathcal{F}_t] = \mathbb{Q}(\tau > t | \mathcal{F}_t) \cdot X.$$

Since $\tilde{Y} = \mathbf{1}_{\{\tau > T\}} Y$ is a \mathcal{G} -measurable random variable the proof can be completed. ■

The corollary is first applied to the survival probability defined in Eq. (2.4.4) which results in

$$\hat{H}(t, T) = \mathbb{E} [\mathbf{1}_{\{\tau > T\}} | \mathcal{G}_t] = \mathbf{1}_{\{\tau > t\}} H(t, T)$$

where

$$H(t, T) := \frac{\mathbb{E} [\mathbf{1}_{\{\tau > T\}} | \mathcal{F}_t]}{\mathbb{Q}(\tau > t | \mathcal{F}_t)} = \mathbb{E} \left[\exp \left(- \int_t^T \lambda(u) du \right) \middle| \mathcal{F}_t \right]$$

As seen above the computation of this term gives the same result as that of the zero-coupon bond price using the "short rate" λ . The derivative of H with respect to T is the negative expected value of survival until T and default in the next infinitesimal small time interval:

$$-\partial_T H(t, T) = \mathbb{E} \left[\exp \left(- \int_t^T \lambda(u) du \right) \lambda(T) \middle| \mathcal{F}_t \right]$$

In the sub-filtration approach one more assumption is required regarding the behavior of martingales between the filtration \mathbb{G} and its sub-filtration \mathbb{F} .

Assumption 2.5.2. *Every \mathbb{F} (square-integrable) martingale is a \mathbb{G} (square-integrable) martingale.¹*

This property is discussed in JEANBLANC & RUTKOWSKI (2000b) and in JAMSHIDIAN (2004). It is called martingale invariance property or conditionally-independent sub-filtration approach. This assumption implies that the \mathbb{F} -Brownian motion remains a Brownian motion in the enlarged filtration. It can be shown that this assumption is equivalent to:

For any $t \in [0, T_\infty]$ the σ -algebras \mathcal{G}_t and \mathcal{F} are conditionally independent of a given \mathcal{F}_t .

¹The square-integrable property is not necessary as the time horizon $T_\infty < \infty$ is finite.

2.6 Credit Products

The credit risky products are introduced in this section. As in the interest rate case all products are presented here that are used for calibration and pricing. The simplest risky product is treated first, the defaultable bond:

Definition 2.6.1. (Defaultable bond) A defaultable zero-coupon bond with maturity T guarantees its holder the payment of one monetary unit at time T in case of no earlier default. $\hat{P}(t, T)$ is the contract value at time $t \leq T$. Obviously, $\hat{P}(T, T) = 1$ for all $T < \tau$. In case of early default the recovery payment is denoted by RR at time τ , $t \leq \tau \leq T$. The recovery amount is assumed to be fixed and not handled stochastically.

The price of the defaultable bond is calculated by

$$\hat{P}(t, T) := \mathbf{1}_{\{\tau > t\}} \bar{P}_{\text{RR}}(t, T)$$

where

$$\bar{P}_{\text{RR}}(t, T) := \frac{\mathbb{E} \left[D(t, T) \mathbf{1}_{\{\tau > T\}} + D(t, \tau) \text{RR} \mathbf{1}_{\{t < \tau \leq T\}} \mid \mathcal{F}_t \right]}{\mathbb{Q}(\tau > t \mid \mathcal{F}_t)}.$$

In case of zero recovery rate this quantity is the expectation of the stochastic default discount factor

$$\bar{P}_0(t, T) = \mathbb{E} \left[\exp \left(- \int_t^T (r(u) + \lambda(u)) du \right) \mid \mathcal{F}_t \right].$$

which is simplified in case of independence of default intensity and interest rate to $\bar{P}_0(t, T) = P(t, T) H(t, T)$.

The second standard product in the credit world is the insurance against default of the reference entity. Therefore payments of premium (insurance) are necessary.

Definition 2.6.2. (Credit default swap) A credit default swap (CDS) is a swap of premium payments K at times T_{a+1}, \dots, T_b in exchange for a single protection payment $\text{LGD} (= 1 - \text{RR})$ at default time τ , provided that $T_a < \tau \leq T_b$.

To simplify the forthcoming formulas the notional is set to one. The CDS can be split into two parts, such as, the default leg for the default payments and the premium leg for the insurance payments. The discounted payoffs at time t are equal to

$$\mathbf{1}_{\{T_a < \tau \leq T_b\}} D(t, \tau) \text{LGD} \tag{2.6.1}$$

resp.

$$K \left(\sum_{i=a+1}^b D(t, T_i) \alpha_i \mathbf{1}_{\{\tau \geq T_i\}} + D(t, \tau) (\tau - T_{\beta(\tau)-1}) \mathbf{1}_{\{T_a < \tau < T_b\}} \right) \tag{2.6.2}$$

where $\beta(t) = \min \{k \mid T_k > t\}$ is the next date in the tenor structure after t , thus, $t \in [T_{\beta(t)-1}, T_{\beta(t)})$. The premium leg contains two parts. The first one reflects the regular

payments in case of no early default. The second part contains the accrued premium between the last payment and the default date. The discounted payoff of a CDS is obtained by the difference between its two legs (protection seller view):

$$\begin{aligned} \Pi_{\text{CDS}_{a,b}(K)}(t) := & K \left(\sum_{i=a+1}^b D(t, T_i) \alpha_i \mathbf{1}_{\{\tau \geq T_i\}} + D(t, \tau) (\tau - T_{\beta(\tau)-1}) \mathbf{1}_{\{T_a < \tau < T_b\}} \right) \\ & - \mathbf{1}_{\{T_a < \tau \leq T_b\}} D(t, \tau) \text{LGD}. \end{aligned} \quad (2.6.3)$$

Remark. The CDS for $t = T_a$ becomes a forward CDS at $t < T_a$. If $t \leq \tau \leq T_a$ (reference entity defaults prior to the start date), the contract is terminated and has no cash flows.

The price of the CDS is the expectation of the discounted payoff. Corollary 2.5.1 implies that

$$\begin{aligned} \text{CDS}_{a,b}(t; K) &= \mathbb{E} [\Pi_{\text{CDS}_{a,b}}(t; K) \mid \mathcal{G}_t] \\ &= \frac{\mathbf{1}_{\{\tau > t\}}}{\mathbb{Q}(\tau > t \mid \mathcal{F}_t)} \mathbb{E} [\Pi_{\text{CDS}_{a,b}}(t; K) \mid \mathcal{F}_t]. \end{aligned} \quad (2.6.4)$$

Definition 2.6.3. (Forward CDS rate) The forward CDS rate $R_{a,b}(t)$ again is the rate which makes the CDS fair. It is defined for the case of no early default only:

$$\mathbf{1}_{\{\tau \geq t\}} R_{a,b}(t) := \mathbf{1}_{\{\tau \geq t\}} \frac{\text{LGD} \cdot \mathbb{E} [D(t, \tau) \mathbf{1}_{\{T_a < \tau \leq T_b\}} \mid \mathcal{G}_t]}{\mathbb{E} \left[\sum_{i=a+1}^b D(t, T_i) \alpha_i \mathbf{1}_{\{\tau \geq T_i\}} + D(t, \tau) (\tau - T_{\beta(\tau)}) \mathbf{1}_{\{T_a < \tau < T_b\}} \mid \mathcal{G}_t \right]}$$

The forward CDS rate is extended as in BRIGO (2005): The sub-filtration approach in the second Eq. of (2.6.4) implies that the CDS is zero in case of early default by the indicator function, regardless of $R_{a,b}$. Hence, due to the sub-filtration ensuring positivity of the denominator, the swap rate can be extended to any case according to

$$R_{a,b}(t) = \frac{\text{LGD} \cdot \mathbb{E} [D(t, \tau) \mathbf{1}_{\{T_a < \tau \leq T_b\}} \mid \mathcal{F}_t]}{\mathbb{E} \left[\sum_{i=a+1}^b D(t, T_i) \alpha_i \mathbf{1}_{\{\tau \geq T_i\}} + D(t, \tau) (\tau - T_{\beta(\tau)}) \mathbf{1}_{\{T_a < \tau < T_b\}} \mid \mathcal{F}_t \right]} \quad (2.6.5)$$

The product defined next adds the optionality to the credit products and enables calibration of the default intensity process.

Definition 2.6.4. (Credit default swap option) A credit default swap option or credit default swaption is an option on a CDS. It gives the holder the right to enter a CDS at its beginning T_a at a predefined level (strike) K . The notation is the same as in Definition 2.6.2. There are two different types of options. The discounted payoff of a call option (CallCDS) is given by

$$\Pi_{\text{CallCDS}_{a,b}(K)}(t) := D(t, T_a) [\text{CDS}_{a,b}(T_a; R_{a,b}(T_a)) - \text{CDS}_{a,b}(T_a; K)]^+$$

whereas the put option's (PutCDS) discounted payoff is

$$\Pi_{\text{PutCDS}_{a,b}(K)}(t) := D(t, T_a) [\text{CDS}_{a,b}(T_a; K) - \text{CDS}_{a,b}(T_a; R_{a,b}(T_a))]^+.$$

Simplifications of the CallCDSs are shown, PutCDSs can be handled in the same way. The formula in the definition enables simplification in two directions. The first one is based on the fact that the forward CDS rate is fair. Using $\text{CDS}(T_a, T_a, T_b, R_{a,b}(T_a), \text{LGD}) = 0$ implies that

$$\Pi_{\text{CallCDS}_{a,b}(K)}(t) = D(t, T_a) [-\text{CDS}_{a,b}(T_a; K)]^+. \quad (2.6.6)$$

The second simplification avails that the two CDSs in the non-simplified CallCDS definition above differ with respect to their coupons only. Therefore their default payments are equal and cancel each other. The resulting discounted payoff can be rewritten by (2.6.3) and (2.6.4) as

$$\begin{aligned} \Pi_{\text{CallCDS}_{a,b}(K)}(t) &= (R_{a,b}(T_a) - K)^+ \cdot D(t, T_a) \frac{\mathbf{1}_{\{\tau > T_a\}}}{\mathbb{Q}(\tau > T_a | \mathcal{F}_{T_a})} \\ &\cdot \left[\sum_{i=a+1}^b \alpha_i \mathbb{Q}(\tau > T_a | \mathcal{F}_{T_a}) \bar{P}_0(T_a, T_i) + \mathbb{E} \left\{ D(T_a, \tau) (\tau - T_{\beta(\tau)-1}) \mathbf{1}_{\{\tau < T_b\}} | \mathcal{F}_{T_a} \right\} \right] \end{aligned} \quad (2.6.7)$$

These formulas are used in Section 3.5 to develop the market model for default intensities.

The oldest product with credit risk is the loan. The following definition of the loan is applied in this thesis.

Definition 2.6.5. (Loan) A plain vanilla loan (without any options) with start date T_a , end date T_b , notional $N = 1$ and fixed coupon C has a discounted payoff of

$$\begin{aligned} \Pi_{\text{FixedLoan}_{a,b}}(t) &:= D(t, T_b) \mathbf{1}_{\{\tau > T_b\}} - D(t, T_a) \mathbf{1}_{\{\tau > T_a\}} \\ &+ C_{\text{fix}} \left(\sum_{i=a+1}^b D(t, T_i) \alpha_i \mathbf{1}_{\{\tau \geq T_i\}} \right) \\ &+ \mathbf{1}_{\{T_a < \tau \leq T_b\}} D(t, \tau) \text{RR}. \end{aligned} \quad (2.6.8)$$

The terms in the first row are the notional transactions at the beginning and at the end. The second row is the fee payment on payment dates T_{a+1}, \dots, T_b and the third row is the recovery payment in case of default. Loans with floating payments also exist. In each period, the fee is split into the market interest rate and an entity specific spread. The discounted payoff is

$$\begin{aligned} \Pi_{\text{FloatLoan}_{a,b}}(t) &:= D(t, T_b) \mathbf{1}_{\{\tau > T_b\}} - D(t, T_a) \mathbf{1}_{\{\tau > T_a\}} \\ &+ \sum_{i=a+1}^b (L(T_{i-1}, T_i) + C_{\text{float}}) D(t, T_i) \alpha_i \mathbf{1}_{\{\tau \geq T_i\}} \end{aligned}$$

$$+ \mathbf{1}_{\{T_a < \tau \leq T_b\}} D(t, \tau) \text{RR}. \quad (2.6.9)$$

Many options could be embedded in a loan. A list of options considered and priced in this thesis:

- Drawdown option: The drawdown option gives the holder the right to drawdown the loan during the live time of the option. For the undrawn time period the holder of the option has to pay a commitment fee $F < C$.
- Prepayment option: The prepayment option gives the holder the right to prepay the loan at any time after drawdown. If used for a refinancing in the market at cheaper than the current conditions, this option is exercised in a rational manner.
- Multiple options: Some loans work like contingent credit lines (CCL). In this case the credit can be (partially) drawn and undrawn at any time during the life of the loan. The payments are split into a commitment fee F and a coupon C for the undrawn and drawn portion, respectively.

3 Interest Rate and Default Intensity Model

The first section is an introduction into the CIR model and its extensions with shifts or jumps, and to the corresponding analytical pricing formulas. The model which is used throughout the thesis is presented below. The calibration procedure is shown in detail at the end of this chapter.

3.1 Short Rate Model - CIR Model and Extensions

There is a large number of different short rate models such as the Vasicek, the Ho-Lee, or the Cox-Ingersoll-Ross model. Each model has advantages and disadvantages. The following list shows the most important issues for the valuation of a model:

- distribution of the short rate
- analytical tractability
- calibration to market data
- positivity of the rates (necessary for interest rates and default intensities)
- numerical suitability (for Monte Carlo simulation or for building recombining lattices/trees)

The model of Cox-Ingersoll-Ross (CIR) introduced in their paper COX, INGERSOLL JR & ROSS (1985a) is used in this thesis. It prevents negative rates by definition and is analytically tractable. Moreover, when using the shift extension the model matches the term structure exactly. This section contains an introduction into the model and its different extensions and starts with a short derivation of the process. Thereafter the properties and analytical prices are studied.

3.1.1 Derivation of the CIR Diffusion Process

This subsection mostly follows ROGERS (1995) and MAGHSOODI (1996) and introduces the CIR process as the sum of Gaussian squared processes. Let (W_1, \dots, W_d) be a d -dimensional

Brownian motion under \mathbb{Q} and positive constants κ, σ . Then X_j is obtained as the solution of Langevin's equation

$$dX_j(t) = -\frac{1}{2}\kappa X_j(t) dt + \frac{1}{2}\sigma dW_j(t), \quad X_j(0) \in \mathbb{R}$$

for $j = 1, \dots, d$. The short rate under the risk-neutral measure \mathbb{Q} is defined by

$$r(t) := \sum_{j=1}^d X_j^2(t), \quad (3.1.1)$$

Itô calculus reveals

$$\begin{aligned} dr(t) &= \sum_{j=1}^d 2X_j \left(-\frac{1}{2}\kappa X_j(t) dt + \frac{1}{2}\sigma dW_j(t) \right) + \sum_{j=1}^d \frac{1}{4}\sigma^2 \langle dW_j(t), dW_j(t) \rangle \\ &= \left(\frac{d\sigma^2}{4} - \kappa r(t) \right) dt + \sigma \sqrt{r(t)} \sum_{j=1}^d \frac{X_j(t)}{\sqrt{r(t)}} dW_j(t). \end{aligned}$$

Defining $W(t) := \sum_{j=1}^d \int_0^t \frac{X_j(u)}{\sqrt{r(u)}} dW_j(u)$ which is a Brownian motion¹ shows the dynamics

$$dr(t) = \left(\frac{d\sigma^2}{4} - \kappa r(t) \right) dt + \sigma \sqrt{r(t)} dW(t).$$

Define $\theta := \frac{d\sigma^2}{4\kappa}$, then the CIR process follows the equation²

$$dr(t) = \kappa(\theta - r(t)) dt + \sigma \sqrt{r(t)} dW(t). \quad (3.1.2)$$

Each of the constants has got a name and a meaning:

- θ is called the mean-reversion (log-run average) level. On the long run the process is always pushed back to this value.
- κ is the rate (speed) of mean-reversion. The greater κ the faster the return to the mean-reversion level.
- σ is the volatility parameter of the process. The greater σ the bigger the shocks of the process.

The usage of the square-root of the process in the variance has several effects. First the incremental variance is proportional to the current level of the process. And as this term approaches zero when the process approaches zero the CIR process is non-negative always.

¹Shown by $dW(t) = \sum_{j=1}^d \frac{X_j(t)}{\sqrt{r(t)}} dW_j(t)$ and $\langle dW(t), dW(t) \rangle = \sum_{j=1}^d \frac{X_j^2(t)}{r(t)} dt = dt$. A formal proof is given by MAGHSOODI (1996), Lemma 2.1.

²The derivation and the representation (3.1.1) is only true for $d \in \mathbb{N}$. But the CIR process allows all positive constants θ .

BROWN & DYBVIK (1986) conducted an empirical study by applying this model to the U.S. interest rates. Therefore, the parameters of the CIR process are sometimes called Brown-Dybvig parameters.

3.1.2 Properties of the CIR Process

References for this subsection are COX, INGERSOLL JR & ROSS (1985a) at first, besides the two term structure books NAWALKHA, SOTO & BELIAEVA (2007) and BRIGO & MERCURIO (2006).

The CIR process (3.1.2) has a rescaled non-central chi-square distribution. The probability density function at time t conditional on the current information at time s is given by

$$f(r(t) | \mathcal{F}_s) = c \cdot e^{-u-v} \left(\frac{v}{u}\right)^{q/2} I_q \left(2(u \cdot v)^{1/2}\right) \quad (3.1.3)$$

where

$$\begin{aligned} c &= \frac{2\kappa}{\sigma^2 (1 - e^{-\kappa(t-s)})} \\ u &= c \cdot r(s) e^{-\kappa(t-s)} \\ v &= c \cdot r(t) \\ q &= \frac{2\kappa\theta}{\sigma^2} - 1 \end{aligned}$$

and where $I_q(\cdot)$ is the modified Bessel function of the first kind and of the order q . The cumulative distribution function of the CIR process is given by the non-central chi-squared function

$$\chi^2 [2c \cdot r(t); 2q + 2, 2u] \quad (3.1.4)$$

where $2q + 2 = d$ (see Subsection 3.1.1) are the degrees of freedom and $2u$ is the parameter of non-centrality which is proportional to the current spot rate.

The mean and variance of $r(t)$ conditional on information up to time s are given by

$$\begin{aligned} \mathbb{E}[r(t) | \mathcal{F}_s] &= \frac{2q + 2 + 2u}{2c} = r(s) e^{-\kappa(t-s)} + \theta \left(1 - e^{-\kappa(t-s)}\right) \\ \text{Var}[r(t) | \mathcal{F}_s] &= \frac{2(2q + 2 + 4u)}{(2c)^2} = r(s) \frac{\sigma^2}{\kappa} \left(e^{-\kappa(t-s)} - e^{-2\kappa(t-s)}\right) + \theta \frac{\sigma^2}{2\kappa} \left(1 - e^{-\kappa(t-s)}\right)^2. \end{aligned}$$

The CIR model is an affine model, i.e., the price of a zero-coupon bond (Definition 2.2.3) at time t with maturity T can be written in the form

$$P^{\text{CIR}}(t, T) = A(t, T) \exp[-B(t, T) r(t)].$$

In the CIR model the explicit expressions of $A(t, T)$ and $B(t, T)$ are given by

$$A(t, T) = \left[\frac{2h \exp\left(\frac{(\kappa+h)(T-t)}{2}\right)}{2h + (\kappa + h) (\exp[(T-t)h] - 1)} \right]^{\frac{2\kappa\theta}{\sigma^2}} \quad (3.1.5)$$

$$B(t, T) = \frac{2 (\exp[(T-t)h] - 1)}{2h + (\kappa + h) (\exp[(T-t)h] - 1)} \quad (3.1.6)$$

$$h = \sqrt{\kappa^2 + 2\sigma^2} \quad (3.1.7)$$

Moreover, the instantaneous forward rate is analytical and equal to

$$\begin{aligned} f^{\text{CIR}}(t, T) &= -\frac{\partial \ln P^{\text{CIR}}(t, T)}{\partial T} \quad (3.1.8) \\ &= \kappa\theta \cdot B(t, T) + r(t) \cdot B_T(t, T) \\ &= \kappa\theta \cdot B(t, T) + r(t) \left(1 - \kappa \cdot B(t, T) - \frac{1}{2}\sigma^2 B^2(t, T) \right) \\ &= \frac{2\kappa\theta (\exp[(T-t)h] - 1)}{2h + (\kappa + h) (\exp[(T-t)h] - 1)} \\ &\quad + r(t) \frac{4h^2 \exp[(T-t)h]}{[2h + (\kappa + h) (\exp[(T-t)h] - 1)]^2} \end{aligned}$$

The superscript \cdot^{CIR} denotes the model; the subscript \cdot_T is the abbreviation of the derivative $\frac{\partial}{\partial T}$. They are given in Eq. (3.1.9) and (3.1.10). The dynamics of the forward rate curve are obtained by Itô:

$$df^{\text{CIR}}(t, T) = (B_{Tt}(t, T) - \kappa \cdot B_T(t, T)) r(t) + B_T(t, T) \sigma \sqrt{r(t)} dW(t)$$

It can be shown that the CIR model is consistent with the forward rates drift condition of HEATH, JARROW & MORTON (1992).³ The initial forward rate curve is completely determined in the CIR model as

$$f^{\text{CIR}}(0, T) = \kappa\theta \cdot B(0, T) + r(0) \cdot B_T(0, T).$$

The derivatives of (3.1.5) and (3.1.6) above are given by

$$A_T(t, T) = \partial_T A(t, T) = \frac{-2\kappa\theta (\exp[(T-t)h] - 1)}{2h + (\kappa + h) (\exp[(T-t)h] - 1)} \cdot A(t, T) \quad (3.1.9)$$

$$\begin{aligned} B_T(t, T) = \partial_T B(t, T) &= \frac{4h^2 \exp[(T-t)h]}{[2h + (\kappa + h) (\exp[(T-t)h] - 1)]^2} \\ &= 1 - \kappa \cdot B(t, T) - \frac{1}{2}\sigma^2 B^2(t, T) \end{aligned} \quad (3.1.10)$$

Moreover, there are further analytical price formulas for products in the CIR model. All

³For the drift μ_f and the volatility σ_f the equation $\mu_f(t, T) = \sigma_f(t, T) \int_t^T \sigma_f(t, u) du$ must hold.

formulas are presented which are applied in the finally chosen extension of the CIR model.

3.1.3 Feller's Stability Condition

As stated in Subsection 3.1.1 the CIR process is always non-negative. As shown in FELLER (1951) there exist conditions on the corresponding parameters of the process such that the process is strictly positive.

Theorem 3.1.1. (*Feller condition*) *The condition $2\kappa\theta > \sigma^2$ ensures that the process r is strictly positive.*

In his paper Feller studied the parabolic equation

$$u_t = (axu)_{xx} - ((bx + c)u)_x, \quad 0 < x < \infty$$

where $u = u(t, x)$, and a, b, c are constants, $a > 0$. This is the Fokker-Planck equation of the process

$$dx(t) = (bx(t) + c)dt + \sqrt{2a}\sqrt{x(t)}dW(t).$$

This is a CIR process which corresponds to the process in Eq. (3.1.2) upon setting $a = \frac{1}{2}\sigma^2$, $b = -\kappa$ and $c = \kappa\theta$. The equivalence to the condition in Theorem 3.1.1 is $c > a$. Based on this condition, Feller proved the following statements:

- If $0 < c < a \Leftrightarrow 0 < 2\kappa\theta < \sigma^2$ the situation is almost the same as in case of regular equations. But now a 'reflecting barrier' solution exist for the initial value problem. There exist infinity many other positivity preserving norm decreasing solutions, but one and only one satisfies the condition $u(t, 0) < \infty$. Moreover there exists a unique solution with vanishing initial values and prescribed flux at $x = 0$; it is integrable, but generally unbounded.
- If $c > a \Leftrightarrow 2\kappa\theta > \sigma^2$, a positivity and norm preserving solution of the initial value problem exists such that both it and its flux vanish at $x = 0$ which means that $x = 0$ is an absorbing as well as a reflecting barrier, and no homogenous boundary condition can be imposed.

JAMSHIDIAN (1995) suggests a more intuitive way to see the effects of the Feller condition. By defining $X(t) := \ln(r(t))$ and using Itô's lemma the process can be described by

$$dX(t) = \left[\left(\kappa\theta - \frac{\sigma^2}{2} \right) \exp(-X(t)) - \kappa \right] dt + \sigma \cdot \exp\left(-\frac{X(t)}{2}\right) dW(t).$$

The coefficients of $X(t)$ are well behaved for large positive X , but blow up for large negative X . If $2\kappa\theta < \sigma^2$ (Feller condition broken), the drift is negative for large negative X , pulling X toward $-\infty$. But for $2\kappa\theta > \sigma^2$ the drift is positive, pulling X away from $-\infty$.

Nevertheless, there are market situations for which it is impossible to hold both, the Feller condition and a volatility calibration. These market situations have a large volatility in the underlying product and there is an additional restriction to the parameters. The CIR value has to be lower than the market term structure as explained in the next subsection. The market term structure has a certain level and thus, limits the mean-reversion level in the calibration process. Moreover, a higher speed of the mean-reversion allows a higher volatility parameter in the Feller condition, but simultaneously reduces the CIR volatility. Overall the condition limits the volatility of the CIR process for a given specific term structure. Therefore the Feller condition cannot be met in all real applications. The final calibrated model presented in Subsection 3.4.3 is an example for such a market situation where the Feller condition is broken.

3.1.4 Extension of CIR Model: CIR++

The CIR process defined in (3.1.2) has only four parameters as inputs (mean-reversion level, its speed, volatility, and start value) and hence, has four degrees of freedom only. The term structure of discount factors has an infinite number of values (by interpolation of market values). Therefore the CIR model is not sufficient to match the term structure directly. Some applicable extensions are presented in this subsection, together with some analytical solutions for the CIR++ model.

COX, INGERSOLL JR & ROSS (1985a) directly introduced a first idea to match the exact term structure by using a time-dependent mean-reversion level. Besides, HULL & WHITE (1990) enforced this approach by considering all parameters to be time-dependent. In the Vasicek model this approach works well,⁴ but in the CIR model it is not applicable in practice. Even in the case of a time-dependent mean-reversion level, the bond price cannot be calculated in a simple way anymore, but the integral in Eq. (3.1.5) needs to be solved numerically. In conclusion, a root-search algorithm is required to find the time-dependent mean-reversion level in each time step of the calibration. Besides, as stated in HEATH, JARROW & MORTON (1992) the Feller condition is fulfilled if the initial forward rate curve 3.1.1 satisfies

$$f(0, T) \geq r(0) B_T(0, T) + \frac{\sigma^2 B(0, T)}{2}$$

The last equation cannot be fulfilled in any market situation. Thus, the Feller condition is broken in such cases.

There have been several attempts to extend the flexibility of the CIR model into different directions. ROGERS (1995) introduced a simple square-root model. Firstly he described time-changes by a differentiable deterministic function $\varphi(t)$ which results in a different differential equation for the short rate process. The deterministic function is used to exactly

⁴The Vasicek short rate model with time dependent parameter is called Hull-White (HW) model and has been introduced in HULL & WHITE (1990). This model is the same as the CIR without the square root term in the volatility.

3.1. Short Rate Model - CIR Model and Extensions

match the initial term structure by analytical values. A potential second step is achieved by the multiplication of a deterministic function $g(t) > 0$ with the time-changed process resulting in

$$r(t) = g(t) \cdot x_r^\alpha(\varphi_r^{\text{CIR}}(t))$$

where $x_r^\alpha(t)$ is a standard CIR process defined in (3.1.2). The subscript r refers to the related process, the superscript α contains the defining parameters of the CIR process

$$\alpha = (x(0), \kappa, \theta, \sigma). \quad (3.1.11)$$

Rogers shows that this process is equivalent to a CIR process with time-dependent parameters in case they fulfill $\kappa(t)\theta(t)/\sigma^2(t) \equiv \text{constant}$. This is discussed above in the second passage in this subsection.

In this thesis a different extension of the CIR process is used. On the one hand this extension preserves the analytical tractability and on the other it is flexible enough to match the initial term structure. This approach was introduced in BRIGO & MERCURIO (2001). Instead of time-change and multiplication with a deterministic function, a deterministic shift $\varphi^{\text{CIR}}(t) \geq 0$ is added to the CIR process. Hence, the short rate is given by

$$r(t) = x_r^\alpha(t) + \varphi^{\text{CIR}}(t; \alpha) \quad (3.1.12)$$

The exact term structure of discount factors in the continuous time model for a given α is obtained by $\varphi^{\text{CIR}}(t; \alpha)$ as

$$\varphi^{\text{CIR}}(t; \alpha) = f^{\text{M}}(0, t) - f^{\text{CIR}}(0, t; \alpha), \quad (3.1.13)$$

where $f^{\text{M}}(0, t) = -\frac{\partial \ln P^{\text{M}}(0, t)}{\partial t}$ with \cdot^{M} denoting the market value and f^{CIR} is defined according to (3.1.8).

Remark 3.1.2. The parameter vector α has the following restrictions:

- α has to fulfill the Feller-condition to guarantee positivity of the rate. As stated above, in some cases it is broken for an appropriate calibration result.
- α has to be chosen according to

$$\varphi^{\text{CIR}}(t; \alpha) \geq 0 \quad \forall t \geq 0$$

which is equivalent to the condition

$$f^{\text{CIR}}(0, t; \alpha) < f^{\text{M}}(0, t) \quad \forall t \geq 0.$$

This condition ensures the match of the term structure by a positive rate retaining the positivity of the short rate. This is discussed in BRIGO & MERCURIO (2001).

Equation (3.1.13) is equivalent to the formula of the market term structure:

$$\begin{aligned}\ln P^{\text{M}}(0, T) &= \ln A(0, T) - x_r^\alpha(0) B(0, T) - \int_0^T \varphi^{\text{CIR}}(t; \alpha) dt \\ \Psi^{\text{CIR}}(t, T; \alpha) &:= \int_t^T \varphi^{\text{CIR}}(s, \alpha) ds \\ &= \ln A(0, T) - \ln A(0, t) - \ln P^{\text{M}}(0, T) \\ &\quad + \ln P^{\text{M}}(0, t) - x_r^\alpha(0) \cdot (B(0, T) - B(0, t))\end{aligned}$$

As stated above the CIR++ model maintains the analytical tractability. The zero-coupon bond in the CIR++ model still has its affine form:

$$\begin{aligned}P^{\text{CIR}++}(t, T) &= A^{\text{CIR}++}(t, T) \exp[-B(t, T) r(t)] \\ &= A(t, T) \exp\left[-\int_t^T \varphi^{\text{CIR}}(s) ds - B(t, T) x_r^\alpha(t)\right]\end{aligned}\tag{3.1.14}$$

where

$$A^{\text{CIR}++}(t, T) = \frac{P^{\text{M}}(0, T) A(0, t) \exp[-B(0, t) x_r^\alpha(0)]}{P^{\text{M}}(0, t) A(0, T) \exp[-B(0, T) x_r^\alpha(0)]} A(t, T) \exp[B(t, T) \varphi^{\text{CIR}}(t)]$$

The following calibration procedure uses the put option on a zero-coupon bond. Its price at time t is analytical with strike K , maturity S on a bond with maturity T :

$$\begin{aligned}\text{ZBP}(t, S, T; K) &= \\ P(t, T) &\left[\chi^2 \left(2\hat{r}[\rho + \psi + B(S, T)]; \frac{4\kappa\theta}{\sigma^2}, \frac{2\rho^2 [r(t) - \varphi^{\text{CIR}}(t; \alpha)] \exp[h(S - t)]}{\rho + \psi + B(S, T)} \right) - 1 \right] \\ &- K \cdot P(t, S) \left[\chi^2 \left(2\hat{r}[\rho + \psi]; \frac{4\kappa\theta}{\sigma^2}, \frac{2\rho^2 [r(t) - \varphi^{\text{CIR}}(t; \alpha)] \exp[h(S - t)]}{\rho + \psi} \right) - 1 \right]\end{aligned}\tag{3.1.15}$$

where

$$\begin{aligned}\hat{r} &= \frac{1}{B(S, T)} \left[\ln \frac{A(S, T)}{K} - \ln \frac{P^{\text{M}}(t, S) A(t, T) \exp[-B(t, T) x_r^\alpha(0)]}{P^{\text{M}}(t, T) A(t, S) \exp[-B(t, S) x_r^\alpha(0)]} \right] \\ \rho &= \rho(S) := \frac{2h}{\sigma^2 (\exp[h(S - t)] - 1)} \\ \psi &= \frac{\kappa + h}{\sigma^2}\end{aligned}$$

Among others, the derivation of the option price in the CIR model can be found in MAGHSOODI (1996).

In the model designation CIR++ the first + sign refers to a constant shift extension to the CIR process, the second one to a deterministic shift term.

3.1.5 Addition of Jumps to the Model: CIR-EJ++

To further extend the flexibility of the model and to facilitate the adaption to real data, a jump term is introduced into the CIR part of the short rate. Processes with a jump term of this type are called jump diffusion processes. The jump term was first introduced into the CIR process by AHN & THOMPSON (1988). Among others, the importance and influence of a jump component on option valuations is shown in AMIN (1993).

The dynamics of the CIR process extended by a jump are given by

$$dx(t) = \kappa(\theta - x(t))dt + \sigma\sqrt{x(t)}dW(t) + f(x, J)dN(t; \varsigma) \quad (3.1.16)$$

where the last term consists of two components. $dN(t; \varsigma)$ represents a homogeneous Poisson process with constant intensity $\varsigma > 0$ (jump arrival rate), i.e., the jump arrival rate is independent of time t . As has been explained in Subsection 2.4 on default time, the probability of having no jump until time t (in case of $N(0) = 0$) is given by

$$\mathbb{Q}(N(t; \varsigma) = 0) = \exp(-\varsigma \cdot t). \quad (3.1.17)$$

It is assumed that N is independent of the Brownian motion W . The other component $f(x, J)$ describes the distribution in case of a jump. The function can depend on the current level x and on a random variable J . This random variable J again is independent of the other stochastic components. There are several possibilities for the jump function and the distribution of J . BELIAEVA, NAWALKHA & SOTO (2007) used the normal and the exponential distribution. For negative jumps bigger than the rate, a jump component results in negative values with positive probability. Therefore, the jump component must be carefully chosen to obtain the positivity property of the CIR process.

In this thesis it is assumed that $f(x, J) = J$ and J is exponentially distributed with a positive mean γ . Among others, these assumptions have been used by DUFFIE, PAN & SINGLETON (2000) and BRIGO & EL-BACHIR (2006). Therefore the notation $J(\gamma)$ is used in the following, and positive jumps are conceded to the process only. This model is called CIR-EJ (Exponential Jump). To account for the jump the parameter vector in Eq.(3.1.11) is extended to six components:

$$\alpha = (x(0), \kappa, \theta, \sigma, \varsigma, \gamma)$$

The distribution of this jump diffusion process is made up of two parts: A non-central chi-squared distribution and an exponential distributed random variable with arrival rate ς . The calculation of the moments is described in ZHOU (2001). The jump component has no influence on the Feller condition which ensures positivity in the CIR-EJ just as in the CIR model. An example of a probability density function of CIR and CIR-EJ is shown in Fig. 3.1.1 where the parameter vector (0.0165, 0.4, 0.026, 0.14, 0.25, 0.15) is applied with a time horizon of one year. The last two components of the parameter vector are only used in the

CIR-EJ model. As expected, the exponential jump component shifts the density to higher values.

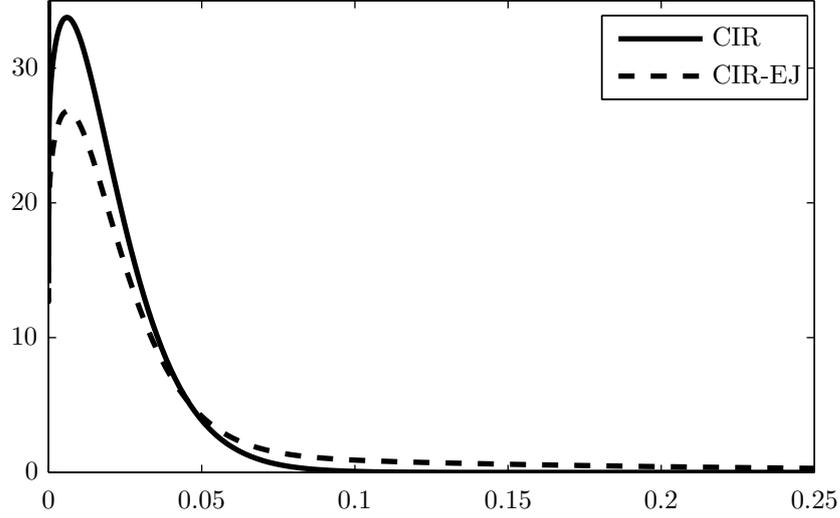


Figure 3.1.1: Probability density function for CIR and CIR-EJ

This model is affine as well, the model class is called affine jump-diffusion (AJD). Without shift extension (assuming $r(t) = x(t)$), the bond prices are equal to

$$P^{\text{CIR-EJ}}(t, T) = A^{\text{CIR-EJ}}(t, T) \exp[-B(t, T) r(t)]$$

$$A^{\text{CIR-EJ}}(t, T) = A(t, T) \cdot \bar{A}(t, T)$$

where

$$\bar{A}(t, T) = \left[\frac{2h \exp\left(\frac{h+\kappa+2\gamma}{2}(T-t)\right)}{2h + (\kappa + h + 2\gamma) (\exp[(T-t)h] - 1)} \right]^{\frac{2\kappa\gamma}{\sigma^2 - 2\kappa\gamma - 2\gamma^2}}$$

$$\text{resp. if } \gamma = \frac{h - \kappa}{2}$$

$$\bar{A}(t, T) = 1$$

The second equation of \bar{A} for the special case $\gamma = \frac{h-\kappa}{2}$ is necessary as in this case the denominator $\sigma^2 - 2\kappa\gamma - 2\gamma^2$ in the exponent becomes zero. But at the same time, namely when $\gamma \rightarrow \frac{h-\kappa}{2}$, the term in the brackets converges to one. The term $\bar{A}(t, T)$ is the bond adjustment for the EJ component of the process. Under the CIR-EJ process the instantaneous forward rate is given by

$$f^{\text{CIR-EJ}}(t, T; \alpha) = \kappa\theta B(t, T) + x(t) \left(1 - \kappa B(t, T) - \frac{1}{2}\sigma^2 B^2(t, T) \right) + \frac{\varsigma\gamma B(t, T)}{1 + \gamma B(t, T)}$$

3.1. Short Rate Model - CIR Model and Extensions

$$\begin{aligned}
&= \kappa \theta B(t, T) + x(t) \cdot B_T(t, T) + \frac{\varsigma \gamma B(t, T)}{1 + \gamma B(t, T)} \\
&= \frac{2\kappa \theta (\exp[(T-t)h] - 1)}{2h + (\kappa + h) (\exp[(T-t)h] - 1)} \\
&+ x(t) \frac{4h^2 \exp[(T-t)h]}{[2h + (\kappa + h) (\exp[(T-t)h] - 1)]^2} \\
&+ \frac{2\varsigma \gamma \exp[(T-t)h - 1]}{2h + (\kappa + h + 2\gamma) \exp[(T-t)h - 1]} \\
&= x(t) B_T(t, T) - \frac{A_T^{\text{CIR-EJ}}(t, T)}{A^{\text{CIR-EJ}}(t, T)}
\end{aligned}$$

The two derivatives of the jump model are equal to

$$\begin{aligned}
A_T^{\text{CIR-EJ}}(t, T) &= \partial_T A^{\text{CIR-EJ}}(t, T) = A(t, T) \partial_T \bar{A}(t, T) + \bar{A}(t, T) \partial_T A(t, T) \\
\bar{A}_T(t, T) &= \partial_T \bar{A}(t, T) = \frac{-2\varsigma \gamma (\exp[(T-t)h] - 1)}{2h + (\kappa + h + 2\gamma) (\exp[(T-t)h] - 1)} \cdot \bar{A}(t, T) \\
&\text{resp. if } \gamma = \frac{h - \kappa}{2} \\
\bar{A}_T(t, T) &= \partial_T \bar{A}(t, T) = 0
\end{aligned}$$

The matching of the CIR-EJ model with the exact term structure is done exactly as in the CIR++ model. The short rate is equal to

$$r(t) = x_r^\alpha(t) + \varphi^{\text{CIR-EJ}}(t; \alpha) \quad (3.1.18)$$

for a deterministic function $\varphi^{\text{CIR-EJ}}$. For a given α the calibration is again obtained by $\varphi^{\text{CIR-EJ}}(t; \alpha)$ as

$$\begin{aligned}
\varphi^{\text{CIR-EJ}}(t; \alpha) &= f^{\text{M}}(0, t) - f^{\text{CIR-EJ}}(0, t; \alpha), \\
&\text{resp.} \\
\ln P^{\text{M}}(0, T) &= \ln A^{\text{CIR-EJ}}(0, T) - x_r^\alpha(0) B(0, T) - \int_0^T \varphi^{\text{CIR-EJ}}(t; \alpha) dt \\
\Psi^{\text{CIR-EJ}}(t, T; \alpha) &= \int_t^T \varphi^{\text{CIR-EJ}}(s, \alpha) ds \\
&= \ln A^{\text{CIR-EJ}}(0, T) - \ln A^{\text{CIR-EJ}}(0, t) - \ln P^{\text{M}}(0, T) \\
&\quad + \ln P^{\text{M}}(0, t) - x_r^\alpha(0) \cdot (B(0, T) - B(0, t))
\end{aligned}$$

The restrictions for the parameter vector α are the same as in the CIR model, see Remark 3.1.2.

The price $P^{\text{CIR-EJ}++}(t, T)$ of the zero-coupon bond in the CIR-EJ++ model is similar to

formula (3.1.14) of the CIR++ model:

$$P^{\text{CIR-EJ}++}(t, T) = A^{\text{CIR-EJ}}(t, T) \exp \left[- \int_t^T \varphi^{\text{CIR-EJ}}(s) ds - B(t, T) r(t) \right].$$

The equation for the bond option price is semi-analytic and can be found in Proposition 3.5.7 where it is directly applied to CDS option pricing.

3.2 Two-Dimensional CIR(-EJ)++ Model

In this section the interest rate and the default intensity process are modeled under consideration of their correlation. The CIR++ process and the CIR-EJ++ process are also called shifted square-root diffusion (SSRD) process and shifted square-root jump diffusion (SSRJD) process, respectively, as described in BRIGO & MERCURIO (2006).

Assumption 3.2.1. *The short rate r is assumed to be a CIR++ process as defined in Eq. (3.1.12). The CIR part is defined as x_r^α , the deterministic shift as $\varphi_r(\cdot; \alpha)$ with a parameter vector $\alpha = (x_r^\alpha(0), \kappa_r, \theta_r, \sigma_r)$. The Brownian motion of the interest rate process is W .*

Assumption 3.2.2. *The default intensity λ is assumed to be a CIR++ or a CIR-EJ++ process as defined in Eq. (3.1.12) and (3.1.18), respectively. The CIR(-EJ) part is defined as x_λ^β , the deterministic shift as $\varphi_\lambda(\cdot; \beta)$ with a parameter vector $\beta = (x_\lambda^\beta(0), \kappa_\lambda, \theta_\lambda, \sigma_\lambda, \varsigma_\lambda, \gamma_\lambda)$. The last two parameters exist only in the CIR-EJ model. The Brownian motion and the jump arrival rate of the default intensity process are \bar{W} and \bar{N} , respectively.*

Assumption 3.2.3. *The Brownian motions W and \bar{W} are correlated according to*

$$d\langle W(t), \bar{W}(t) \rangle = \rho dt$$

The jump component is assumed to be independent of the Brownian motions.

These assumptions mean that both, the interest rate and the default intensity, are stochastic. The interest rate has no jump component because it plays a minor role only in the valuation of credit products. Moreover, in the world of interest rates a jump component with positive shocks only is not reasonable. In the credit world shocks are possible which cause a positive jump of the default intensity. For example the default intensity of a firm jumps in case of bad news. The downward movement after such bad news is mostly smoother which is covered by the diffusion part. Therefore the one-sided jumps are used.

Both Brownian motions W and \bar{W} and the jump component \bar{N} are adapted to the market filtration \mathbb{F} . This is in line with the definition of sub-filtrations in Section 2.5.

3.3 Calibration to Market Quotes - Term Structure

At least two different methods can be used to determine the model parameters. The choice mainly depends on the scope of usage. In the first method the parameters are estimated from historic data (e.g., maximum likelihood as in KLADIVKO (2007)). The resulting calibrated model is under the historic measure \mathbb{P} . The second approach calibrates the parameters according to the actual prices on the market. The model calibrated in this way is able to price non-traded products, providing prices which coincide with the current prices under the risk-neutral measure \mathbb{Q} . In this thesis the second approach is chosen for model calibration since the market provides prices for a large amount of financial products. The quality of the model is assessed by its ability to match these market prices. REBONATO (1998) is a standard reference on this method.

The term structure of interest rates must be extracted from several interest rate products such as cash deposits, forward rate agreements/futures and swap rates. This method is standard and therefore it is not described here. Interested readers are referred to a detailed description of the bootstrapping algorithm in O'KANE (2008), Chapter 2. The starting point is the resulting term structure of zero-coupon bonds which is a requirement for every pricing procedure.

Figure 3.3.1 shows the zero interest rate of the market and the calibrated CIR zero rate. The distance between both lines is filled by the deterministic shift function. The calibration procedure is shown below.

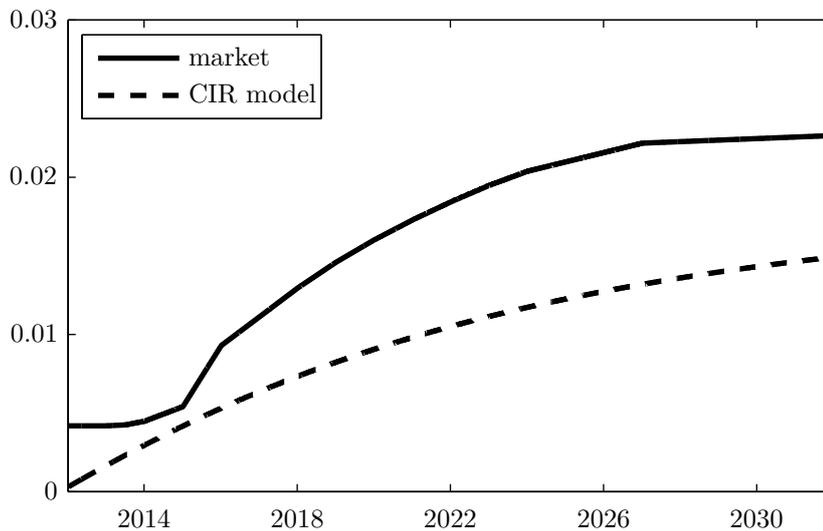


Figure 3.3.1: Zero interest rates (6th July 2012, Copyright©2012 Market Data Distribution Service, Commerzbank AG)

The default term structure is either derived from credit default swaps or from corporate

bonds. As CDS options and loans are the target of the model of this thesis, the CDS is the natural choice for the determination of the term structure.⁵ The market quotes CDS spreads as a curve with different maturities as shown in Table 3.3.1. The standard recovery rate is 40%. In the bootstrapping procedure it is assumed that the interest rates are deterministic.

time (in years)	0.5	1	2	3	4	5	7	10
Allianz	64	73	88	105	120	135	144	154
iTraxx® Main				140		171	182	189

Table 3.3.1: CDS quotes (in bps) (6th July 2012, Copyright©2012 Markit Group Limited)

In case of given market CDS spreads, the default probability is the only free parameter to equalize the protection and the premium leg (CDS Definition 2.6.2). The bootstrapping method starts by matching the default probability to the CDS with the shortest maturity. Based on the CDS with the second shortest maturity the next default probability is determined. This step is repeated up to the last maturity using all earlier default probabilities in every step. The market does not determine the shape of the default intensity between the different maturities. The standard approach is to assume a constant or linear behavior between subsequent maturity buckets as shown in Fig. 3.3.2. The drawback of linear interpolation is that it might become negative under extreme market conditions.⁶

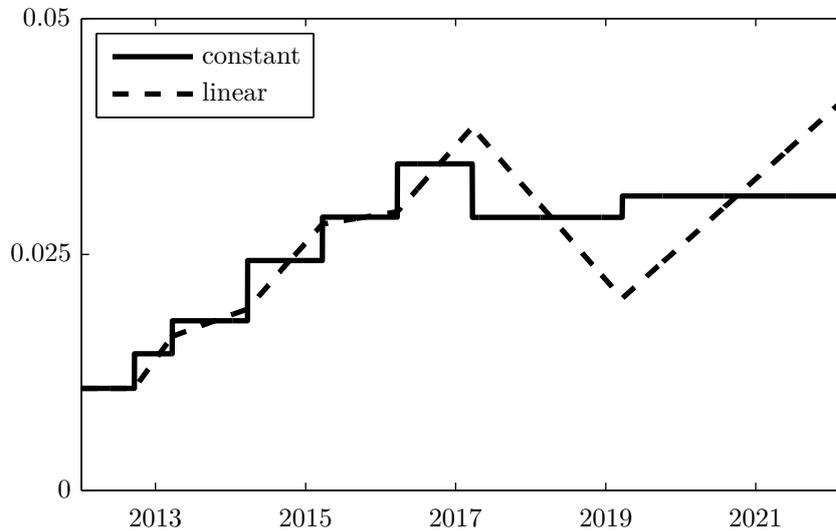


Figure 3.3.2: Default intensity of Allianz CDS quotes

As stated in Subsection 3.1.4 the deterministic shift functions $\varphi_r(\cdot; \alpha)$ and $\varphi_\lambda(\cdot; \beta)$ ensure

⁵An empirical analysis is given in NORDEN & WAGNER (2008).

⁶A discussion can be found in BRIGO & MERCURIO (2006), page 765 et seq.

the exact match of the term structures in case of interest rate and default intensity, respectively. All other parameters of the CIR(-EJ)++ model can be freely chosen to match the model to the volatility structure given by market quotes. The restrictions for α and β are stated in Remark 3.1.2. The volatility calibration is described in Sections 3.4 and 3.5.

3.4 Volatility Calibration of Interest Rates

This section contains the standard procedure to calibrate the CIR++ interest rate model to the market volatility.

The market volatility can be determined using market prices of option products or historical volatilities. In case of a non-liquid option market the historical volatilities could be of interest. But instead traded implied volatilities are used here since - due to the intended pricing issues - market data are preferred in this case as well.

The considered option product is the cap. The market provides prices or implied volatilities where the conversion is done by the standard market model based on the famous model of BLACK (1976). Therefore the implied volatilities are called Black implied volatilities. The conversion is described in Subsection 3.4.1.

The volatility calibration of the CIR++ model is done by using the parameter vector α . α defines the CIR part of the short rate and has a degree of freedom of four: the start value, the mean-reversion level, the speed of mean-reversion, and the volatility parameter. Subsection 3.4.2 outlines the analytical cap pricing formula in the CIR++ model.

The number of maturity buckets can be larger than the number of parameters (i.e., $n > 4$) or generally the market prices are not matched exactly by parameter restrictions. Then, calibration is achieved by minimizing the sum of square percentage differences between CIR++ model and market prices, i.e.,

$$\min \sqrt{\sum_{i=1}^n \left(\frac{Price_i^{CIR++} - Price_i^M}{Price_i^M} \right)^2} \quad (3.4.1)$$

The calibration results for one market situation are shown in Subsection 3.4.3. Additionally this subsection outlines the sensitivities of the cap volatilities in the CIR++ model to the model parameters.

3.4.1 Market Model

The market model formula for pricing caps is presented here. The formula is usable in both directions which is the conversion from volatilities into prices and vice versa. The formula is based on the standard Black formula which is based on the assumption that forward interest

rates follow a drift-less geometric Brownian motion under the pricing measure. Definition 2.6.2 introduces the caps as the sum of options on a payer IRS.

Theorem 3.4.1. *According to the Black formula the cap price at time $t \leq T_a$, strike K and notional N is given by*

$$\begin{aligned} \text{Cap}_{a,b}(t; K) \\ = N \sum_{i=a+1}^b P(t, T_i) \alpha_i \text{Bl} \left(K, F(t, T_{i-1}, T_i), v_{a,b} \sqrt{T_{i-1} - t} \right) \end{aligned}$$

where

$$\text{Bl}(K, F, v) = F\Phi(d_1(K, F, v)) - K\Phi(d_2(K, F, v)) \quad (3.4.2)$$

$$d_{1,2}(K, F, v) = \frac{\ln\left(\frac{F}{K}\right) \pm \frac{v^2}{2}}{v} \quad (3.4.3)$$

Φ denotes the standard normal cumulative distribution function.

The proof is standard and therefore it is omitted here.⁷ The market prices of at-the-money caps are obtained by inserting the forward swap rate $S_{a,b}(t)$ as strike and the implied cap volatilities $v_{a,b}^{\text{M,Cap}}$ at $t = 0$. The forward swap rate can be calculated as in Eq. (2.2.3).

3.4.2 CIR++ Model

This subsection describes shortly the pricing of caps in the CIR++ model.

The single terms of the sum of a cap are called caplets. Caplets are calculated as put options on zero-coupon bonds according to

$$\begin{aligned} \text{Caplet}_{T_{i-1}, T_i}(t; K) &= N\mathbb{E} \left[\alpha_i D(t, T_i) (L(T_{i-1}, T_i) - K)^+ \mid \mathcal{F}_t \right] \\ &= N(1 + \alpha_i K) \mathbb{E} \left[D(t, T_{i-1}) \left(\frac{1}{1 + \alpha_i K} - P(T_{i-1}, T_i) \right)^+ \mid \mathcal{F}_t \right] \\ &= N(1 + \alpha_i K) \text{ZBP} \left(t, T_{i-1}, T_i; \frac{1}{1 + \alpha_i K} \right) \end{aligned}$$

Therefore, the cap prices can be calculated by the sum over put options as

$$\begin{aligned} \text{Cap}_{a,b}(t; K) \\ = N \sum_{i=a+1}^b (1 + \alpha_i K) \text{ZBP} \left(t, T_{i-1}, T_i, \frac{1}{1 + \alpha_i K} \right) \end{aligned}$$

⁷For details see, e.g., BRIGO & MERCURIO (2006), page 198 et seq.

and hence, the price of a cap is analytic in the CIR++ model using the pricing formula 3.1.15 of put options. The model prices of the at-the-money caps are obtained with the forward swap rate $S_{a,b}$ as strike and the CIR model parameters at $t = 0$.

3.4.3 Calibration Results

The calibration to the market implied cap volatility is performed according to (3.4.1). The result of the calibration is the CIR model described by the parameter vector $\alpha = (0.0003, 0.0707, 0.0395, 0.13)$. This parameter vector does not fulfill the Feller condition.

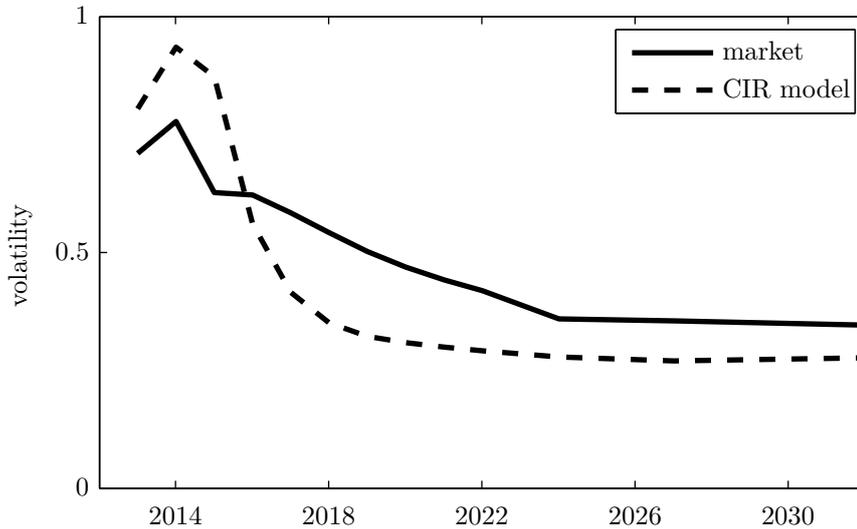


Figure 3.4.1: Implied cap volatilities (6th July 2012, Copyright©2012 Bloomberg L.P.)

Figure 3.4.1 compares the calibrated Black implied volatilities of the CIR model to those of the market. The market volatilities are directly quoted. Black’s formula allows to recalculate the Black volatilities from the CIR model prices. The shapes of the curves are almost the same. The approximation of the CIR model works fairly well taking into account that the CIR model has four parameters only to fit 13 time buckets.

Figures 3.4.2 depict the sensitivities of the Black volatilities in comparison to the CIR model parameters. In each of the four figures one parameter is changed only, to four values higher and four values lower than the calibrated one, while the other three parameters are fixed to their calibrated values shown above. The upper left chart shows the effect of the start value. The higher the value the higher the volatility. The effect weakens over time as the process is pulled to the mean reversion level. The rate of mean reversion, shown in Fig.

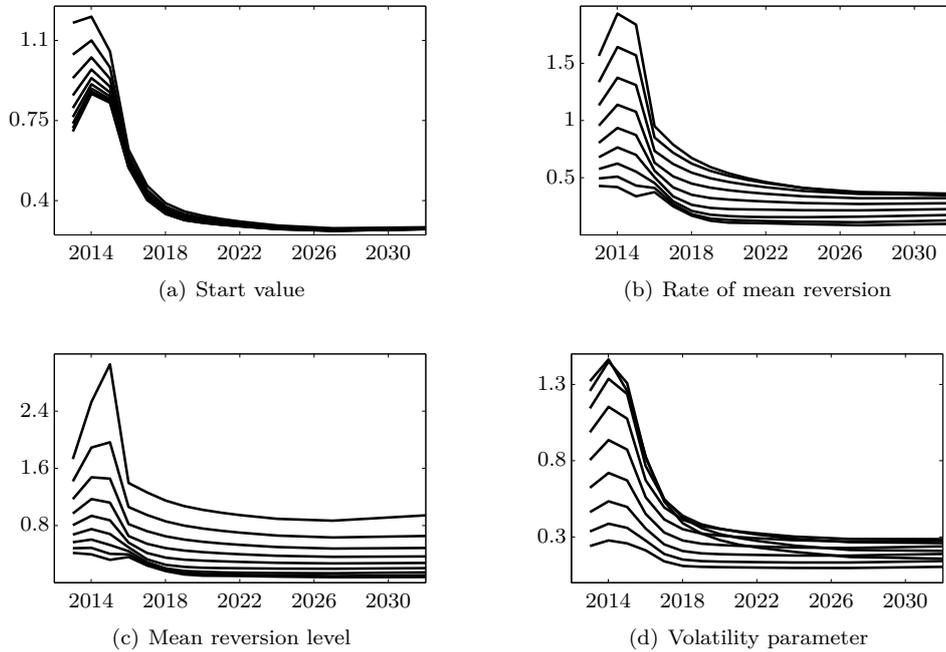


Figure 3.4.2: Sensitivities of CIR interest rate parameters

3.4.2(b), has a similar effect in this parameter constellation. Because of the low start level a higher rate results in a faster move to the higher mean reversion level and therefore to a higher volatility of the process. Note that there are cases with the opposite effect where a faster return to the mean reversion level lowers the volatility. In Fig. 3.4.2(c) the mean reversion level is changed. The higher the mean reversion level the higher is the volatility due to the square-root term in the CIR process. This parameter has the highest impact on the volatilities in this specific parameter vector.

The influence of the volatility parameter is shown in the bottom right chart. The effect is obvious in case of parameters which fulfill the Feller condition - the higher the parameter the higher the volatility. Here, in contrast the volatilities of the CIR model having the two highest volatility parameters cross the lines with lower volatility parameters. The impact of the volatility parameter on the cumulative distribution function of the CIR process is shown in Fig. 3.4.3. The 20 year distributions of Eq. (3.1.4) are plotted for the calibrated volatility parameter as well as for the smallest and the highest value used in the sensitivity plot. The distribution function of the calibrated interest rate process increases smoothly, while the function of the high volatility parameter shows a high density at zero and at extreme values of the CIR process. The cap prices are high in case of these extreme scenarios but with very low probability only. In contrast, because of its high probability to be above the strike, the cap is more often exercised for the calibrated process in that case. The density to be above 0.1 is higher in the calibrated process than in the high volatility case. The statement that

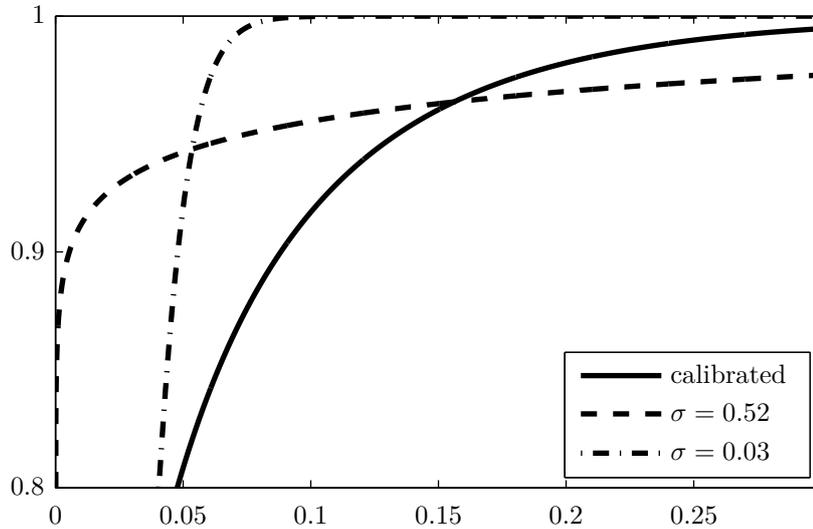


Figure 3.4.3: Cumulative density distribution of CIR processes

a higher volatility parameter results in a higher implied volatility is therefore broken above a specific level.

3.5 Volatility Calibration of Default Intensity

The overall procedure of finding the minimum sum is the same as in the last Section 3.4 for interest rates. The task is to find the minimum sum of squared percentage differences between CIR(-EJ)++ model and market prices of all maturity buckets. Including the exponential jumps to the model component implies that the parameter vector β has two more degrees of freedom than without. The market product for calibration is the CDS call option introduced in Definition 2.6.4. Similar to the procedure in Section 3.4 for the interest rate, the standard market formula which is used to convert implied volatilities to prices is presented first in Subsection 3.5.1. Then a semi-analytical formula for the CIR(-EJ)++ model is introduced in Subsection 3.5.2. The calibration results and a sensitivity analysis are provided in Subsection 3.5.3. A short comparison of the assumptions of the market model and the CIR model in Subsection 3.5.4 follows. The section concludes by a semi-analytical formula for the CCL.

For calibration purposes it is assumed that the correlation between short rate and intensity is zero since otherwise no analytical solution is obtainable for the term structure and the volatility of the default process. The effect of correlation is shown in Chapter 6 using

a numerical valuation technique. Compared to the calibration of the interest rate, more details are given and discussed in this chapter with respect to the calibration of the default intensity, because the stronger focus of this thesis is on credit risk modeling, and because the procedure for the default intensity is not as much standardized as the other in the interest rate world.

Remark. The CDS call option is the market product used to calibrate the volatility process. As the market for this product is not very liquid the related information have to be handled with care. For the situation that the available market data of the CDS options are not sufficient to calibrate the CIR model, BRIGO & ALFONSI (2005) supposed to just minimize $\int_0^{T_\infty} \varphi_\lambda(s; \beta) ds$. But this procedure is not a calibration to market data, and the output volatility obtained is randomly related to the real market volatility. In such situations the historical volatility of the CDS spreads is applicable.

3.5.1 Market Model

SCHÖNBUCHER (2000a) has been the first to apply the idea and methodology of a swap measure to credit risk as among others did JAMSHIDIAN (1997) for interest rates. This subsection is based on the subsequent works of JAMSHIDIAN (2004) and BRIGO & MORINI (2005).

The market model disregards the accrued coupon term $(\tau - T_{\beta(\tau)-1})$ in the default case of the CDS, since its effect is negligible compared to the effects of the remaining terms and the notation $\overline{\text{CDS}}$ is used. Excluding this term Eq. (2.6.7) results in

$$\begin{aligned} & \Pi_{\text{Call}\overline{\text{CDS}}_{a,b}}(t; K) \\ &= D(t, T_a) \frac{\mathbf{1}_{\{\tau > T_a\}}}{\mathbb{Q}(\tau > T_a | \mathcal{F}_{T_a})} \left[\sum_{i=a+1}^b \alpha_i \mathbb{Q}(\tau > T_a | \mathcal{F}_{T_a}) \bar{P}_0(T_a, T_i) \right] (\bar{R}_{a,b}(T_a) - K)^+ \end{aligned} \quad (3.5.1)$$

$$= \mathbf{1}_{\{\tau > T_a\}} D(t, T_a) \left[\sum_{i=a+1}^b \alpha_i \bar{P}_0(T_a, T_i) \right] (\bar{R}_{a,b}(T_a) - K)^+. \quad (3.5.2)$$

The quantity between the square brackets in the first line is called "defaultable present value per basis point (DPVBP)" and is comparable to the non-default portfolio of zero-coupon bonds in Eq. (2.2.4):

$$\hat{C}_{a,b}(t) := \mathbb{Q}(\tau > t | \mathcal{F}_t) \bar{C}_{a,b}(t) \quad \text{with} \quad \bar{C}_{a,b}(t) := \sum_{i=a+1}^b \alpha_i \bar{P}_0(t, T_i)$$

The DPVBP is always positive due to the market sub-filtration \mathbb{F} which does not include the default time. Hence, it can be used as a numeraire for the new equivalent measure called DPVBP measure and denoted by $\hat{\mathbb{Q}}^{a,b}$. SCHÖNBUCHER (2000a) introduced a similar measure

3.5. Volatility Calibration of Default Intensity

called default swap measure. But since he did not apply the sub-filtration, positivity is not guaranteed for his measure. But in case of $t = 0$ his results were the same as those presented below.

Without the accrued term the forward CDS rate $\bar{R}_{a,b}(t)$ of Definition 2.6.3 in combination with the sub-filtration approach of Eq. (2.6.5) is simplified to

$$\bar{R}_{a,b}(t) = \frac{\text{LGD} \cdot \mathbb{E} \left[D(t, \tau) \mathbf{1}_{\{T_a < \tau < T_b\}} \mid \mathcal{F}_t \right]}{\widehat{C}_{a,b}(t)}, \quad t \leq T_a. \quad (3.5.3)$$

Lemma 3.5.1. $\bar{R}_{a,b}(t)$ is an \mathcal{F}_t -martingale (and hence, a \mathcal{G}_t -martingale as stated in assumption 2.5.2) under the DPVBP measure for $t \leq T_a$.

Proof. Applying Baye's rule it follows that

$$\begin{aligned} \widehat{\mathbb{E}}^{a,b} [\bar{R}_{a,b}(t) \mid \mathcal{F}_s] &= \frac{B(s)}{\widehat{C}_{a,b}(s)} \mathbb{E} \left[\frac{\widehat{C}_{a,b}(t) \cdot \text{LGD} \cdot \mathbb{E} [D(t, \tau) \mathbf{1}_{\{T_a < \tau < T_b\}} \mid \mathcal{F}_t]}{B(t) \cdot \widehat{C}_{a,b}(t)} \mid \mathcal{F}_s \right] \\ &= \frac{\mathbb{E} [\text{LGD} \cdot D(s, \tau) \mathbf{1}_{\{T_a < \tau < T_b\}} \mid \mathcal{F}_s]}{\widehat{C}_{a,b}(s)} \\ &= \bar{R}_{a,b}(s). \end{aligned}$$

■

Assumption 3.5.2. The volatility of $\bar{R}_{a,b}$, denoted by $\bar{v}_{a,b}$, is constant.⁸

Using the martingale representation theorem, this assumption in combination with the martingale property yields the drift less geometric Brownian motion

$$d\bar{R}_{a,b}(t) = \bar{v}_{a,b} \bar{R}_{a,b}(t) dW^{a,b}(t)$$

where $W^{a,b}(t)$ is a Brownian motion under $\widehat{\mathbb{Q}}^{a,b}$. Therefore the following Black formula for CDS options can be derived:

Theorem 3.5.3. The market formula of a CallCDS is given by

$$\text{CallCDS}_{a,b}(t; K) = \mathbf{1}_{\{\tau > t\}} \bar{C}_{a,b}(t) \text{Bl} \left(K, \bar{R}_{a,b}(t), \bar{v}_{a,b} \sqrt{T_a - t} \right)$$

with $\text{Bl}(K, F, v)$ as defined in (3.4.2).

Proof. Starting from Eq. (3.5.2) for the discounted payoff and applying the sub-filtration corollary, Baye's rule and the martingale property of $\bar{R}_{a,b}$, yields the price equation

$$\text{CallCDS}_{a,b}(t; K) = \mathbb{E} \left\{ \mathbf{1}_{\{\tau > T_a\}} D(t, T_a) \left[\sum_{i=a+1}^b \alpha_i \bar{P}_0(T_a, T_i) \right] (\bar{R}_{a,b}(T_a) - K)^+ \mid \mathcal{G}_t \right\}$$

⁸How to find a deterministic volatility approximation is discussed in JAMSHIDIAN (2004).

$$\begin{aligned}
&= \frac{\mathbf{1}_{\{\tau>t\}}}{\mathbb{Q}(\tau>t|\mathcal{F}_t)} \mathbb{E} \left\{ D(t, T_a) \cdot \mathbb{E} [\mathbf{1}_{\{\tau>T_a\}} | \mathcal{F}_{T_a}] \bar{C}_{a,b}(T_a) (\bar{R}_{a,b}(T_a) - K)^+ \middle| \mathcal{F}_t \right\} \\
&= \frac{\mathbf{1}_{\{\tau>t\}}}{\mathbb{Q}(\tau>t|\mathcal{F}_t)} \mathbb{E} \left\{ D(t, T_a) \hat{C}_{a,b}(T_a) (\bar{R}_{a,b}(T_a) - K)^+ \middle| \mathcal{F}_t \right\} \\
&= \frac{\mathbf{1}_{\{\tau>t\}}}{\mathbb{Q}(\tau>t|\mathcal{F}_t)} \hat{C}_{a,b}(t) \hat{\mathbb{E}}^{a,b} \left\{ (\bar{R}_{a,b}(T_a) - K)^+ \middle| \mathcal{F}_t \right\} \\
&= \mathbf{1}_{\{\tau>t\}} \bar{C}_{a,b}(t) \hat{\mathbb{E}}^{a,b} \left[(\bar{R}_{a,b}(T_a) - K)^+ \middle| \mathcal{F}_t \right] \\
&= \mathbf{1}_{\{\tau>t\}} \bar{C}_{a,b}(t) \text{Bl} \left(K, \bar{R}_{a,b}(t), \bar{v}_{a,b} \sqrt{T_a - t} \right)
\end{aligned} \tag{3.5.4}$$

■

The market prices of the at-the-money $\overline{\text{CDS}}$ options are obtained by inserting the forward $\overline{\text{CDS}}$ rate $\bar{R}_{a,b}(t)$ as strike and the implied $\overline{\text{CDS}}$ option volatilities $\bar{v}_{a,b}^{\text{M,CallCDS}}$ at $t = 0$. The forward $\overline{\text{CDS}}$ rate at time $t = 0$ is given by the market implied default rates:

$$\bar{R}_{a,b}(0) = \frac{-\text{LGD} \int_{T_a}^{T_b} P(0, u) d \left(e^{-\Lambda^{\text{M}}(u)} \right)}{\sum_{i=a+1}^b \alpha_i P(0, T_i) e^{-\Lambda^{\text{M}}(T_i)}}.$$

3.5.2 CIR++ and CIR-EJ++ Model

This subsection describes the pricing of CDS options in the CIR(-EJ) model as first done by BRIGO & ALFONSI (2005) and BRIGO & EL-BACHIR (2010).

Assumption 3.5.4. *In this subsection the interest rates are assumed to be deterministic.*

This assumption is a requirement to obtain a semi-analytic result for the CDS options in the CIR(-EJ) model. A consequence is that the stochastic discount factor and the zero-coupon bond are equal ($D(t, T) = P(t, T)$). The impact of stochastic interest rates to CDS options can be found in Chapter 6. The CDS pricing formula is simplified by using deterministic interest rates separated from the default component. Taking into account that

$$\mathbb{E} [\mathbf{1}_{\{T_a < \tau < T_b\}} | \mathcal{G}_t] = -\mathbf{1}_{\{\tau > t\}} \int_{T_a}^{T_b} \partial_u H(t, u) du$$

the CDS is given by

$$\begin{aligned}
\text{CDS}_{a,b}(t; K) &= \mathbf{1}_{\{\tau > t\}} \left[K \left(\sum_{i=a+1}^b D(t, T_i) \alpha_i H(t, T_i) \right. \right. \\
&\quad \left. \left. - \int_{T_a}^{T_b} D(t, u) (u - T_{\beta(u)-1}) \partial_u H(t, u) du \right) + \text{LGD} \cdot \int_{T_a}^{T_b} D(t, u) \partial_u H(t, u) du \right].
\end{aligned} \tag{3.5.5}$$

3.5. Volatility Calibration of Default Intensity

For the valuation of the CallCDS the notation of the survival probability $H(t, T, x_\lambda^\beta(0))$ must be extended by a third parameter, the starting value of the intensity driving process at time 0. This third parameter is changed in a second parameter set to simplify Eq. (3.5.6) to (3.5.8) below.

Proposition 3.5.5. *The CallCDS price at time t is given by*

$$\text{CallCDS}_{a,b}(t; K) = \mathbf{1}_{\{\tau > t\}} D(t, T_a) \mathbb{E} \left[e^{-\int_t^{T_a} \lambda(s) ds} \left\{ \text{LGD} - \int_{T_a}^{T_b} h(u) H(T_a, u; x_\lambda^\beta(0)) du \right\}^+ \middle| \mathcal{F}_t \right] \quad (3.5.6)$$

where

$$h(u) := D(T_a, u) \cdot [\text{LGD}(r(u) + \delta_{T_b}(u)) + K(1 - (u - T_{\beta(u)-1}) \cdot r(u))]$$

and $\delta_x(y)$ is the Dirac delta function centered at x .

The term is shortened by applying the Dirac delta function $\int_0^T \delta_x(y) dy = x$ for $x \in [0, T]$.

Proof. Combining formula (3.5.5) with the equation for the price of the CallCDS option at time t yields

$$\begin{aligned} \text{CallCDS}_{a,b}(t; K) &= \mathbb{E} \left\{ \mathbf{1}_{\{\tau > T_a\}} D(t, T_a) \cdot [-\text{CDS}_{a,b}(T_a; K)]^+ \middle| \mathcal{G}_t \right\} \\ &= D(t, T_a) \cdot \mathbb{E} \left[\mathbf{1}_{\{\tau > T_a\}} \left\{ K \int_{T_a}^{T_b} D(T_a, u) (u - T_{\beta(u)-1}) \partial_u H(T_a, u) du \right. \right. \\ &\quad - K \sum_{i=a+1}^b \alpha_i D(T_a, T_i) H(T_a, T_i) \\ &\quad \left. \left. - \text{LGD} \int_{T_a}^{T_b} D(T_a, u) \partial_u H(T_a, u) du \right\}^+ \middle| \mathcal{G}_t \right] \end{aligned}$$

Integration by parts and setting $q(u) := -\partial_u D(T_a, u) = D(T_a, u) \cdot r(u)$ implies that the first integral can be rewritten as follows:

$$\begin{aligned} &\int_{T_a}^{T_b} D(T_a, u) (u - T_{\beta(u)-1}) \partial_u H(T_a, u) du \\ &= \sum_{i=a+1}^b \int_{T_{i-1}}^{T_i} D(T_a, u) (u - T_{i-1}) \partial_u H(T_a, u) du \\ &= \sum_{i=a+1}^b \left\{ D(T_a, u) (u - T_{i-1}) H(T_a, u) \Big|_{T_{i-1}}^{T_i} - \int_{T_{i-1}}^{T_i} (\partial_u [D(T_a, u) (u - T_{i-1})]) H(T_a, u) du \right\} \\ &= \sum_{i=a+1}^b D(T_a, T_i) \alpha_i H(T_a, T_i) - \int_{T_a}^{T_b} (D(T_a, u) - q(u) (u - T_{\beta(u)-1})) H(T_a, u) du \end{aligned}$$

The second integral can be reformulated as

$$\begin{aligned}
 & \int_{T_a}^{T_b} D(T_a, u) \partial_u H(T_a, u) du \\
 &= D(T_a, u) H(T_a, u) \Big|_{T_a}^{T_b} - \int_{T_a}^{T_b} \partial_u D(T_a, u) H(T_a, u) du \\
 &= D(T_a, T_b) H(T_a, T_b) - 1 + \int_{T_a}^{T_b} q(u) H(T_a, u) du \\
 &= \int_{T_a}^{T_b} \delta_{T_b}(u) D(T_a, u) H(T_a, u) du - 1 + \int_{T_a}^{T_b} q(u) H(T_a, u) du \\
 &= -1 + \int_{T_a}^{T_b} (\delta_{T_b}(u) D(T_a, u) + q(u)) H(T_a, u) du
 \end{aligned}$$

The proposition now follows from Corollary 2.5.1 and the fact that

$$\frac{\mathbb{Q}(\tau > T_a | \mathcal{F}_t)}{\mathbb{Q}(\tau > t | \mathcal{F}_t)} = \mathbb{E} \left[e^{-\int_t^{T_a} \lambda(s) ds} | \mathcal{F}_t \right]. \quad \blacksquare$$

Corollary 3.5.6. *Assuming that the short rate is bounded ($0 \leq r(u) \leq 1$) and payment dates occur at least once a year, $(u - T_{\beta(u)-1}) < 1$ (these assumptions are natural and need no further investigations) then, if the integral*

$$\int_{T_a}^{T_b} [\text{LGDD}(T_a, u) \partial_u H(T_a, u, 0) + K \cdot H(T_a, u, 0) D(T_a, u) (1 - (u - T_{\beta(u)-1}) \cdot r(u))] du \quad (3.5.7)$$

is positive, the CallCDS price is given by

$$\begin{aligned}
 \text{CallCDS}_{a,b}(t; K) &= \mathbf{1}_{\{\tau > t\}} D(t, T_a) \int_{T_a}^{T_b} h(u) \quad (3.5.8) \\
 &\cdot \mathbb{E} \left[\exp \left(-\int_t^{T_a} \lambda_s ds \right) \left(H(T_a, u; x_\lambda^{\beta^*}(0)) - H(T_a, u; x_\lambda^\beta(0)) \right)^+ \Big| \mathcal{F}_t \right] du
 \end{aligned}$$

where $\beta^* = (x_\lambda^{\beta^*}(0), \kappa_\lambda, \theta_\lambda, \sigma_\lambda)$ with $x_\lambda^{\beta^*}(0) \geq 0$ satisfies the equation

$$\int_{T_a}^{T_b} h(u) H(T_a, u; x_\lambda^{\beta^*}(0)) du = \text{LGD} \quad (3.5.9)$$

If the integral is not positive, the default swaption price is simply given by the corresponding forward default swap value $\text{CDS}_{a,b}(t; K)$.

Proof. The conditions imply that $h(u) \geq 0$ for all u . Note that $h(u)$ is a deterministic function and the survival probability $H(T_a, u; y)$ is monotonically decreasing with increasing start value y for all T_a and u . Hence,

$$\int_{T_a}^{T_b} h(u) H(T_a, u; y) du$$

3.5. Volatility Calibration of Default Intensity

is a monotonically decreasing function of y which, as y tends to the infinity, is below the LGD:

$$\lim_{y \rightarrow \infty} \int_{T_a}^{T_b} h(u) H(T_a, u; y) du = 0 < \text{LGD}$$

BRIGO & EL-BACHIR (2010) showed in detail that if (3.5.7) is positive the limiting value of y against zero is above LGD

$$\lim_{y \rightarrow 0^+} \int_{T_a}^{T_b} h(u) H(T_a, u; y) du \geq \text{LGD}$$

and that $0 \leq x_\lambda^{\beta^*}(T_a) < \infty$ exists and satisfies Eq. (3.5.9). The subsequent result is obtained by application of the decomposition introduced in JAMSHIDIAN (1989) to the result of Proposition 3.5.5. In case of the same sign for of the integration part for all u in the integration interval, the decomposition allows the permutation of the sequence of taking the positive part and integration. If on the other hand the integral (3.5.7) is negative, the payoff of the option is always positive and exercised. Therefore, the CallCDS has the same value as the forward CDS. ■

The expectation in the CDS call option pricing formula is treated in exactly the same manner as the zero-coupon bond put option, using the default intensity λ as short rate and $H(T_a, u; x_\lambda^{\beta^*}(0))$ as strike. The zero-coupon bond put option price is analytic in the CIR model and is given in formula (3.1.15). The final price of a CallCDS is the discounted integral over a zero-coupon bond put option on the intensity multiplied by a deterministic function:

$$\text{CallCDS}_{a,b}(t; K) = \mathbf{1}_{\{\tau > t\}} D(t, T_a) \int_{T_a}^{T_b} h(u) \text{ZBP} \left(t, T_a, u; H(T_a, u; x_\lambda^{\beta^*}(0)); \beta \right) du \quad (3.5.10)$$

where β in the ZBP part indicates the application of the default intensity CIR process.

The equation resulting from the CIR-EJ model is more complex than the one of the CIR model, but it still is semi-analytic. The results of BRIGO & EL-BACHIR (2010) are repeated here. Using Eq. (3.5.8), the CIR-EJ++ term can be divided into the CIR-EJ part and the deterministic shift resulting in

$$\begin{aligned} \text{CallCDS}_{a,b}(t; K) &= \mathbf{1}_{\{\tau > t\}} D(t, T_a) \cdot e^{-\Psi^{\text{CIR-EJ}}(t, T_a; \beta)} \cdot \int_{T_a}^{T_b} h(u) \cdot A^{\text{CIR-EJ}}(T_a, u) \\ &\quad e^{-\Psi^{\text{CIR-EJ}}(T_a, u; \beta)} \cdot \mathbb{E} \left[e^{-\int_t^{T_a} x_\lambda^\beta(s) ds} \left(e^{-B(T_a, u) x_\lambda^{\beta^*}(0)} - e^{-B(T_a, u) x_\lambda^\beta(0)} \right)^+ \middle| \mathcal{F}_t \right] du. \end{aligned}$$

The default process subscript λ is omitted for readability in Proposition 3.5.7.

Proposition 3.5.7. *The expected value in the term above is given by*

$$\begin{aligned} & \mathbb{E} \left[e^{-\int_t^{T_a} x_\lambda^\beta(s) ds} \left(e^{-B(T_a, u) x_\lambda^{\beta^*}(0)} - e^{-B(T_a, u) x_\lambda^\beta(0)} \right)^+ \middle| \mathcal{F}_t \right] \\ &= e^{-B(T_a, u) x_\lambda^{\beta^*}(0)} \Upsilon \left(T_a - t, x_\lambda^\beta(0), x_\lambda^{\beta^*}(0), 0 \right) - \Upsilon \left(T_a - t, x_\lambda^\beta(0), x_\lambda^{\beta^*}(0), B(T_a, u) \right) \end{aligned}$$

where

$$\Upsilon(T, y(0), \zeta, \varrho) = \frac{1}{2} \alpha(T) e^{-\beta(T)y(0)} - \frac{1}{\pi} \int_0^\infty \frac{e^{Uy(0)} [S \cos(Wy(0) + v\zeta) + R \sin(Wy(0) + v\zeta)]}{v} dv$$

and

$$\begin{aligned} \beta(T) &= \frac{2\varrho h + (2 + \varrho(h - \kappa))(e^{hT} - 1)}{2h + (h + \kappa + \varrho\sigma^2)(e^{hT} - 1)} \\ \alpha(T) &= \left[\frac{2h \exp\left(\frac{\kappa + hT}{2}\right)}{2h + (h + \kappa + \varrho\sigma^2)(e^{hT} - 1)} \right]^{\frac{2\kappa\varrho}{\sigma^2}} \\ &\quad \cdot \left[\frac{2h(1 + \varrho\gamma) \exp\left(\frac{(h^2 - (\kappa + 2\gamma)^2)(1 - \frac{\varrho}{2}(h + \kappa))}{2(h - \kappa - 2\gamma + \varrho(\gamma(h + \kappa) - \sigma_\lambda^2))} T\right)}{2h(1 + \rho\gamma) + [h + \kappa + \varrho\sigma^2 + \gamma(2 + \varrho(h - \kappa))](e^{hT} - 1)} \right]^{\frac{2\kappa\gamma}{\sigma^2 - 2\kappa\gamma - 2\gamma^2}} \\ R &= (J^2 + K^2)^{\frac{D}{2}} e^G \left[E \cos\left(H + D \arctan\left(\frac{K}{J}\right)\right) - F \sin\left(H + D \arctan\left(\frac{K}{J}\right)\right) \right] \\ S &= (J^2 + K^2)^{\frac{D}{2}} e^G \left[F \cos\left(H + D \arctan\left(\frac{K}{J}\right)\right) + E \sin\left(H + D \arctan\left(\frac{K}{J}\right)\right) \right] \\ U &= \frac{\delta + \varepsilon e^{hT} + \phi e^{2hT}}{N} \\ W &= -\frac{4vh^2 e^{hT}}{N} \\ E &= (\tilde{x}^2 + \tilde{y}^2)^{\frac{\kappa\theta}{\sigma^2}} \cos\left(\frac{2\kappa\theta}{\sigma^2} \arctan\left(\frac{\tilde{y}}{\tilde{x}}\right)\right) \\ F &= (\tilde{x}^2 + \tilde{y}^2)^{\frac{\kappa\theta}{\sigma^2}} \sin\left(\frac{2\kappa\theta}{\sigma^2} \arctan\left(\frac{\tilde{y}}{\tilde{x}}\right)\right) \\ \tilde{x} &= \frac{2he^{(h+\kappa)\frac{T}{2}} [2h + (h + \kappa + \varrho\sigma^2)(e^{hT} - 1)]}{N} \\ \tilde{y} &= -\frac{2he^{(h+\kappa)\frac{T}{2}} v\sigma^2 [e^{hT} - 1]}{N} \\ D &= \frac{-2\gamma\varsigma}{\sigma^2 - 2\gamma\kappa - 2\gamma^2} \\ G &= \frac{\varsigma\gamma T [(2 - \varrho(h + \kappa))(h - \kappa - 2\gamma - \varrho[\sigma^2 - \gamma(h + \kappa)]) + v^2(h + \kappa)[\sigma^2 - \gamma(h + \kappa)]]}{(h - \kappa - 2\gamma - \varrho[\sigma^2 - \gamma(h + \kappa)])^2 + v^2[\sigma^2 - \gamma(h + \kappa)]^2} \\ H &= \frac{\varsigma\gamma T v [(2 - \varrho(h + \kappa))[\sigma^2 - \gamma(h + \kappa)] - (h + \kappa)(h - \kappa - 2\gamma - \varrho[\sigma^2 - \gamma(h + \kappa)])]}{(h - \kappa - 2\gamma - \varrho[\sigma^2 - \gamma(h + \kappa)])^2 + v^2[\sigma^2 - \gamma(h + \kappa)]^2} \end{aligned}$$

$$\begin{aligned}
 J &= 1 + \frac{(e^{hT} - 1) [(h + \kappa + 2\gamma)(1 + \varrho\gamma) + (\sigma^2 + \gamma(h - \kappa)) [\varrho(\varrho\gamma + 1) + v^2\gamma]]}{2h(1 + \varrho\gamma)^2 + 2hv^2\gamma^2} \\
 K &= - \frac{(e^{hT} - 1)v [2\gamma\kappa + 2\gamma^2 - \sigma^2]}{2h(1 + \varrho\gamma)^2 + 2hv^2\gamma^2} \\
 N &= (2h + (h + \kappa + \varrho\sigma^2) [e^{hT} - 1])^2 + v^2\sigma^4 [e^{hT} - 1]^2 \\
 \delta &= 2(h - \kappa) - 4\sigma^2\varrho + \varrho^2\sigma^2(h + \kappa) + v^2\sigma^2(h + \kappa) \\
 \varepsilon &= 4\kappa - 4\kappa^2\varrho - 2\kappa\varrho^2\sigma^2 - 2v^2\sigma^2\kappa \\
 \phi &= -2(h + \kappa) - 4\sigma^2\varrho - \varrho^2\sigma^2(h - \kappa) - v^2\sigma^2(h - \kappa)
 \end{aligned}$$

Remark. This was firstly published in the master thesis of CHRISTENSEN (2002). Additionally, it can be found in LANDO (2004), Appendix E and in CHRISTENSEN (2007), in Proposition 6 on page 59.

3.5.3 Calibration Results

The calibration results for the default intensity in the CIR and in CIR-EJ model are the parameter vectors $\beta = (0.0055, 0.0851, 0.0965, 0.446)$ and $\beta = (0.0038, 0.0307, 0.0312, 0.4078, 0.0621, 0.1445)$, respectively. Both vectors do not fulfill the Feller condition. The forward default intensities of the market, the CIR and the CIR-EJ model are shown in Fig. 3.5.1. The distance between the market value and the model value is filled by the deterministic shift function.

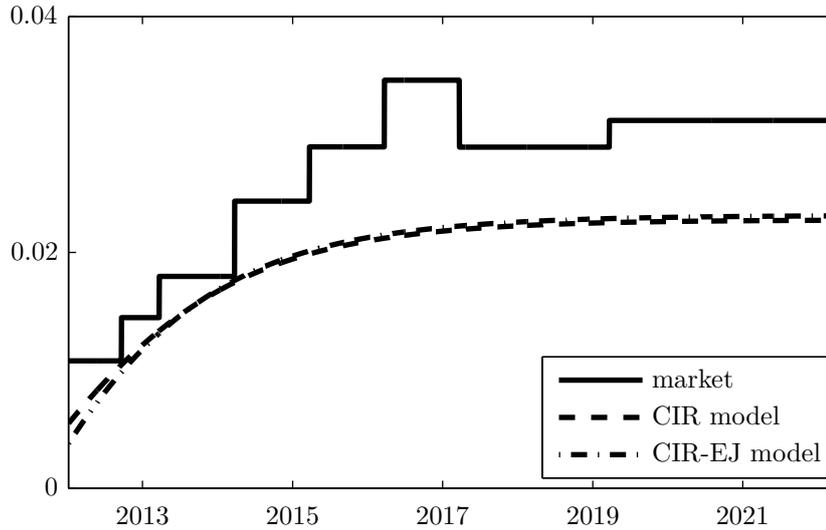


Figure 3.5.1: Default intensity of market, CIR and CIR-EJ (Allianz)

The calibrated CIR(-EJ) processes have to be below the market line to ensure a positive

shift component. The shape of the curve is very similar for the CIR and CIR-EJ model. Both forward rates take the maximum values below the market values at one point in time.

Figure 3.5.2 compares the Black implied CDS option volatilities of the market to those of the CIR and of the CIR-EJ process. In the market there are liquid quotes for the indices only, and not for the enterprises such as Allianz. Therefore, the implied volatility of iTraxx® Main is taken as a proxy. Quotes are given for four different option maturities only, i.e., one, two, three and six months. The underlying CDS or CDS index has a maturity of five years in any case. In addition, Fig. 3.5.2 shows the volatilities of the CDS option maturing in one year for the CIR and CIR-EJ model.

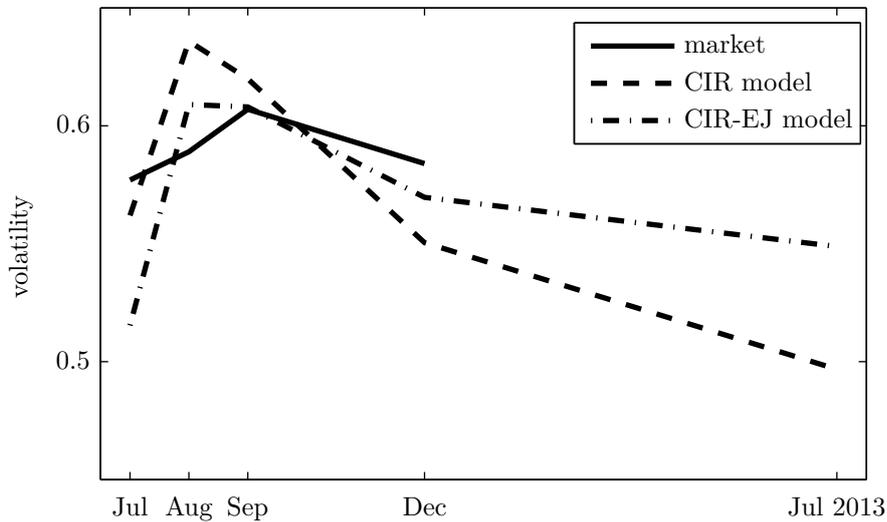


Figure 3.5.2: Implied CDS option volatilities (6th July 2012, Copyright©2012 Bloomberg L.P.)

At a first sight the calibration of the CIR-EJ model looks worse than the CIR model. However, Fig. 3.5.2 shows the volatilities, but the quadratic distance of the prices is minimized to find the CIR(-EJ) processes. Hence, the big distance of CIR-EJ model volatility in the first month has less effect compared to the distance of the CIR model in August and December. The calibration result of the CIR-EJ model has to be better (and is better) compared to the CIR model as there are two more degrees of freedom. Although there are more free parameters to choose than market data points there is not a perfect match. This is due to the dependencies in the parameters and the restriction that the CIR forward curve must be under the market curve.

The sensitivities of the CDS option volatilities to the CIR model parameters are shown in Fig. 3.5.3 in the same manner as those of the cap volatilities in Fig. 3.4.2 to the interest

3.5. Volatility Calibration of Default Intensity

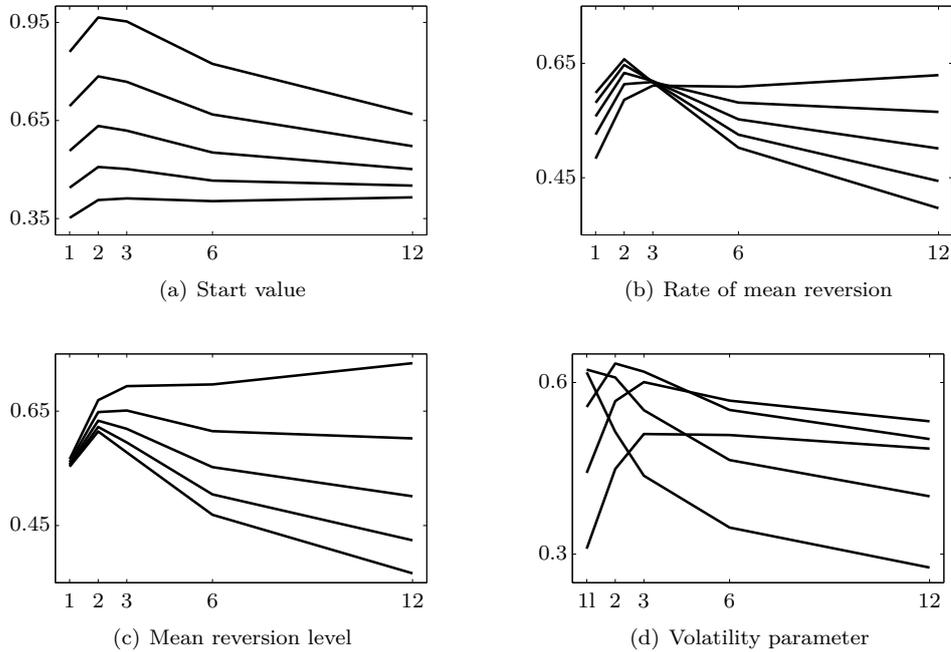


Figure 3.5.3: Sensitivities of CIR default intensity parameters

rates. The x-axis in Fig. 3.5.3 is the maturity in months. In each figure one parameter is changed only, to two values higher and to two values lower than the calibrated one. The other three parameters are fixed to the calibrated values of the CIR process shown above. The upper left chart shows the effect of the start value. The higher the start value the higher the volatility. The effect weakens over time as the process is pulled to the mean reversion level. The rate of mean reversion, shown in Fig. 3.5.3(b), has a contrary effect. Because of the low start level, a higher rate results in a faster movement to the higher mean reversion level and therefore, to a higher volatility on the first two option dates. The third option date shows the same values for all rates of mean reversion. Afterwards a lower rate of mean reversion results in a higher volatility and vice versa. The reason for this effect is that a higher rate of mean reversion lowers the overall volatility as the process returns faster to its mean reversion level. In Fig. 3.5.3(c) the mean reversion level is changed. Due to the square-root term in the CIR process, the higher the mean reversion level the higher is the volatility. This effect increases over time as the process is pushed to this level. The lower right chart exhibits the effect of the volatility parameter. As in the case of the interest rate the effect is not obvious as the distribution is very sensitive to the volatility parameter being far from fulfilling the Feller condition. At the first option date, the lines are sorted by the volatility parameter. The higher the parameter the higher the option volatility. Only the highest and second highest value are interchanged. But both lines with higher volatility than the calibrated value strictly decrease over time. The three other lines are humped. This indifferent picture shows that the effect of the parameters is not stable and that the

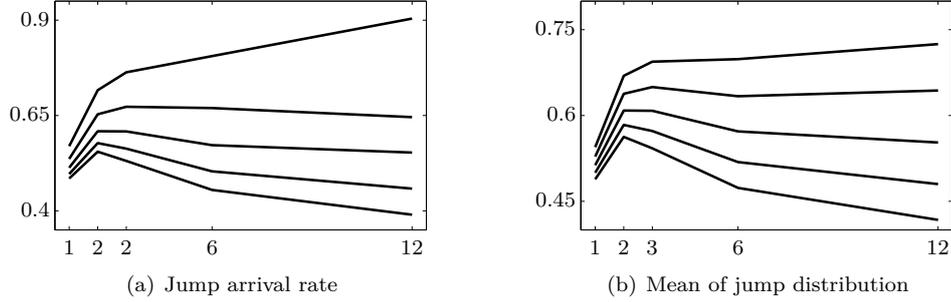


Figure 3.5.4: Sensitivities of CIR-EJ default intensity parameters

calibration of the CIR model to the CDS option volatility is difficult. This effect is due to the fact that the Feller condition is broken for the parameter vector.

Figure 3.5.4 shows the sensitivities to the jump parameters based on the calibrated CIR-EJ model. As the jumps increase the volatility and as they are positive only the effect is obvious. The statement that the higher the parameter the higher are the volatilities for all dates holds for both cases. The jump arrival rate has a higher impact than the mean of jump distribution.

3.5.4 CDS Option Price in the CIR(-EJ)++ Model - Adaption to 3.5.1 Market Model

While the market model omits the accrued coupon term in case of default to enable an analytic result, the option price formula in the CIR(-EJ)++ model shown above takes into account this term. Next, a modified formula in the CIR(-EJ)++ model is presented. This formula is obtained by using the same approach as in the market model. Again, deterministic interest rates are required. Starting point is the third Eq. (3.5.4):

$$\text{Call}\overline{\text{CDS}}_{a,b}(t; K) = \frac{\mathbf{1}_{\{\tau > t\}}}{\mathbb{Q}(\tau > t | \mathcal{F}_t)} \mathbb{E} \left\{ D(t, T_a) \cdot \widehat{C}_{a,b}(T_a) (\bar{R}_{a,b}(T_a) - K)^+ \mid \mathcal{F}_t \right\}$$

If there is no early default, the forward CDS rate of formula (3.5.3) (not an accrued term) in the CIR(-EJ)++ model is given by

$$\mathbf{1}_{\{\tau > t\}} \bar{R}_{a,b}(t) = \mathbf{1}_{\{\tau > t\}} \frac{-\text{LGD} \int_{T_a}^{T_b} D(t, u) \partial_u H(t, u) du}{\bar{C}_{a,b}(t)}.$$

and the Call $\overline{\text{CDS}}$ price under the same condition by

$$\begin{aligned} \text{Call}\overline{\text{CDS}}_{a,b}(t; K) &= \frac{\mathbf{1}_{\{\tau > t\}}}{\mathbb{Q}(\tau > t | \mathcal{F}_t)} \mathbb{E} \left\{ D(t, T_a) \cdot \widehat{C}_{a,b}(T_a) (\bar{R}_{a,b}(T_a) - K)^+ \mid \mathcal{F}_t \right\} \\ &= \mathbf{1}_{\{\tau > t\}} D(t, T_a) \cdot \mathbb{E} \left\{ e^{-\int_t^{T_a} \lambda(s) ds} \right. \\ &\quad \left. \cdot \left(-\text{LGD} \int_{T_a}^{T_b} D(T_a, u) \partial_u H(T_a, u) du - K \cdot \bar{C}_{a,b}(T_a) \right)^+ \mid \mathcal{F}_t \right\} \end{aligned}$$

By rearranging the terms, the same representation as in Proposition 3.5.5 is obtained - with a difference in the auxiliary function $h(u)$ only. Consequently, the second integral can be rewritten in the same way

$$\int_{T_a}^{T_b} D(T_a, u) \partial_u H(T_a, u) du = -1 + D(T_a, T_b) H(T_a, T_b) + \int_{T_a}^{T_b} r(u) D(T_a, u) H(T_a, u) du.$$

By use of the equation

$$K \cdot \bar{C}_{a,b}(T_a) = K \cdot \sum_{i=a+1}^b \alpha_i D(T_a, T_i) H(T_a, T_i)$$

a new auxiliary function $\bar{h}(u)$ is obtained:

$$\bar{h}(u) = D(T_a, u) \cdot \left[\text{LGD}(r(u) + \delta_{T_b}(u)) + K \cdot \sum_{i=a+1}^b \alpha_i \delta_{T_i}(u) \right]$$

Note that $\bar{h}(u) \geq 0$ for all u , and that Corollary 3.5.6 is applicable provided that the new condition $\int_{T_a}^{T_b} H(T_a, u; 0) \bar{h}(u) du \geq \text{LGD}$ is satisfied.

In case of the last quoted CDS option which expires in December the option price difference of the Call $\overline{\text{CDS}}$ and CallCDS is 0.6 bps absolute and below 1% relative, only.

3.5.5 Analytic Price for the Contingent Credit Line in the CIR(-EJ)++ Model

The payoff of the floating rate loan is almost the same as of an $\overline{\text{CDS}}$ without the accrued term in case of default. Due to the definition of the loan (2.6.9) and the spot interest rate (2.2.1) the following equation holds:

$$\begin{aligned} \Pi_{\text{FloatLoan}_{a,b}}(t) &= \Pi_{\overline{\text{CDS}}_{a,b}}(C_{\text{float}})(t) \\ &\quad + \sum_{i=a+1}^b \mathbf{1}_{\{T_{i-1} < \tau \leq T_i\}} (D(t, \tau) - D(t, T_{i-1})) \end{aligned} \quad (3.5.11)$$

The difference between a floating rate loan and a $\overline{\text{CDS}}$ without accrued term is the payment date in case of default only.

The credit contingent line gives its holder the right to draw or pay back the loan at any time. The assumption is that the CCL contract is of the Bermudan type on a predefined set of points in time. These multiple options can be seen as a sum over CDS options where each underlying CDS has only one coupon period, plus the undrawn fee payments. Therefore the option price can be calculated analytically as in the CDS option case. Again, a prerequisite is the assumption that the interest rate is deterministic. As in the CDS option case the contract sign is changed for the option pricing.

Proposition 3.5.8. *The contingent credit line price for a floating rate loan at time t is given by*

$$\begin{aligned} \text{FloatLoan}_{a,b}^{\text{CCL}}(t) &= \mathbf{1}_{\{\tau > t\}} \sum_{i=a+1}^b D(t, T_{i-1}) & (3.5.12) \\ &\mathbb{E} \left[e^{-\int_t^{T_{i-1}} \lambda(s) ds} \left\{ \text{LGD} - \int_{T_{i-1}}^{T_i} h_i^{\text{CCL}}(u) H(T_{i-1}, u; x_\lambda^\beta(0)) du \right\}^+ \middle| \mathcal{F}_t \right] \\ &- \mathbf{1}_{\{\tau > t\}} F_{\text{float}} \sum_{i=a+1}^b \alpha_i D(t, T_i) H(t, T_i) \end{aligned}$$

where

$$\begin{aligned} h_i^{\text{CCL}}(u) &:= \delta_{T_i}(u) [1 + (C_{\text{float}} - F_{\text{float}}) D(T_{i-1}, T_i) \alpha_i - (1 - \text{LGD}) D(T_{i-1}, T_i)] \\ &- (1 - \text{LGD}) r(u) D(T_{i-1}, u). \end{aligned}$$

Proof. The additional term of Eq. (3.5.11) can be written as

$$\begin{aligned} &\mathbb{E} [\mathbf{1}_{\{T_{i-1} < \tau \leq T_i\}} (D(t, \tau) - D(t, T_{i-1})) | \mathcal{G}_t] \\ &= -\mathbf{1}_{\{\tau > t\}} D(t, T_{i-1}) \int_{T_{i-1}}^{T_i} \partial_u H(t, u) (D(T_{i-1}, u) - 1) du \\ &= -\mathbf{1}_{\{\tau > t\}} D(t, T_{i-1}) \left[D(T_{i-1}, T_i) H(T_{i-1}, T_i) \right. \\ &\quad \left. + \int_{T_{i-1}}^{T_i} r(u) D(T_{i-1}, u) H(T_{i-1}, u) du - H(T_{i-1}, T_i) \right] \end{aligned}$$

The functions h_i^{CCL} are determined by \bar{h} from $\overline{\text{CallCDS}}$ as follows:

$$\begin{aligned} h_i^{\text{CCL}}(u) &= \bar{h}_{i-1,i}(u) - D(T_{i-1}, T_i) \delta_{T_i}(u) - r(u) D(T_{i-1}, u) + 1 \\ &= D(T_a, u) \cdot [\text{LGD}(r(u) + \delta_{T_b}(u)) + (C_{\text{float}} - F_{\text{float}}) \alpha_i \delta_{T_i}(u)] \\ &\quad - D(T_{i-1}, T_i) \delta_{T_i}(u) - r(u) D(T_{i-1}, u) + 1 \\ &= \delta_{T_i}(u) [1 + (C_{\text{float}} - F_{\text{float}}) D(T_{i-1}, T_i) \alpha_i - (1 - \text{LGD}) D(T_{i-1}, T_i)] \end{aligned}$$

3.5. Volatility Calibration of Default Intensity

$$- (1 - \text{LGD}) r(u) D(T_{i-1}, u).$$

Alternatively, the same way as for CallCDS and CallCDS is possible. ■

Corollary 3.5.9. *If (3.5.16) is negative, then for all $i = a + 1, \dots, b$ there exists a $\beta_i^* = (x_\lambda^{\beta_i^*}(0), k_\lambda, \theta_\lambda, \sigma_\lambda)$ with $x_\lambda^{\beta_i^*}(0) \geq 0$ which satisfies the equation*

$$\int_{T_{i-1}}^{T_i} h_i^{\text{CCL}}(u) H(T_{i-1}, u; x_\lambda^{\beta_i^*}(0)) du = \text{LGD}. \quad (3.5.13)$$

If $x_\lambda^{\beta_i^*}(0)$ of Eq. (3.5.13) is unique and if for all $i = a + 1, \dots, b$

$$\int_{T_{i-1}}^{T_i} h_i^{\text{CCL}}(u) H(T_{i-1}, u; x_\lambda^{\beta_i^*}(0)) (1 - \exp(-B(T_{i-1}, u)x)) du \quad (3.5.14)$$

is positive for all $x > 0$, then the price of a floating rate CCL is given by

$$\begin{aligned} \text{FloatLoan}_{a,b}^{\text{CCL}}(t) &= \mathbf{1}_{\{\tau > t\}} \sum_{i=a+1}^b D(t, T_{i-1}) X(T_i) \\ &\quad - \mathbf{1}_{\{\tau > t\}} F_{\text{float}} \sum_{i=a+1}^b \alpha_i D(t, T_i) H(t, T_i). \end{aligned} \quad (3.5.15)$$

If the term

$$\text{LGD} - \int_{T_{i-1}}^{T_i} h_i^{\text{CCL}}(u) H(T_{i-1}, u; 0) du \quad (3.5.16)$$

is negative, then

$$\begin{aligned} X(T_i) &= \int_{T_{i-1}}^{T_i} h_i^{\text{CCL}}(u) \\ &\quad \cdot \mathbb{E} \left[\exp \left(- \int_t^{T_{i-1}} \lambda_s ds \right) \left(H(T_{i-1}, u; x_\lambda^{\beta_i^*}(0)) - H(T_{i-1}, u; x_\lambda^\beta(0)) \right)^+ \middle| \mathcal{F}_t \right] du \end{aligned}$$

otherwise $X(T_i)$ is simply the value of the single floating rate loan with coupon $(C_{\text{float}} - F_{\text{float}})$ in period $[T_{i-1}, T_i]$.

Proof. The functions

$$\begin{aligned} h_i^{\text{CCL}}(u) &= \delta_{T_i}(u) [1 + (C_{\text{float}} - F_{\text{float}}) D(T_{i-1}, T_i) \alpha_i - (1 - \text{LGD}) D(T_{i-1}, T_i)] \\ &\quad - (1 - \text{LGD}) r(u) D(T_{i-1}, u). \end{aligned}$$

are negative in the intervals $T_{i-1} < u < T_i$ and positive in T_i . If (3.5.16) is negative it

follows that

$$\lim_{y \rightarrow 0^+} \int_{T_{i-1}}^{T_i} h_i^{\text{CCL}}(u) H(T_{i-1}, u; y) du > \text{LGD}$$

and additionally that

$$\lim_{y \rightarrow \infty} \int_{T_{i-1}}^{T_i} h_i^{\text{CCL}}(u) H(T_{i-1}, u; y) du = 0 \leq \text{LGD}.$$

The functions $H(T_{i-1}, u; x_\lambda^{\beta_i^*}(0))$ provide a solution for $x_\lambda^{\beta_i^*}(0)$ for all i by continuity and monotonicity. The uniqueness of each solution is a prerequisite because of the different signs of h_i^{CCL} .

The terms $H(T_{i-1}, u; x_\lambda^{\beta_i^*}(0)) - H(T_{i-1}, u; x_\lambda^\beta(0))$ have the same sign for all u . The condition that (3.5.14) is positive for all $x > 0$ ensures that

$$\int_{T_{i-1}}^{T_i} h_i^{\text{CCL}}(u) \left(H(T_{i-1}, u; x_\lambda^{\beta_i^*}(0)) - H(T_{i-1}, u; x_\lambda^\beta(0)) \right) du$$

is positive if and only if

$$H(T_{i-1}, u; x_\lambda^{\beta_i^*}(0)) - H(T_{i-1}, u; x_\lambda^\beta(0))$$

is positive for $T_{i-1} \leq u \leq T_i$. Therefore the decomposition of JAMSHIDIAN (1989) is applicable which results in

$$\begin{aligned} & \left(\int_{T_{i-1}}^{T_i} h_i^{\text{CCL}}(u) \left(H(T_{i-1}, u; x_\lambda^{\beta_i^*}(0)) - H(T_{i-1}, u; x_\lambda^\beta(0)) \right) du \right)^+ \\ &= \int_{T_{i-1}}^{T_i} h_i^{\text{CCL}}(u) \left(H(T_{i-1}, u; x_\lambda^{\beta_i^*}(0)) - H(T_{i-1}, u; x_\lambda^\beta(0)) \right)^+ du \end{aligned}$$

for all $i = a + 1, \dots, b$. ■

Remark. The fulfillment of the condition (3.5.14) depends on the CIR parameters as well as on the market environment and the reference entity. For usual conditions

$$\int_{T_{i-1}}^{T_i} h_i^{\text{CCL}}(u) H(T_{i-1}, u; y) du$$

is large positive because of

$$\begin{aligned} & 1 + (C_{\text{float}} - F_{\text{float}}) D(T_{i-1}, T_i) \alpha_i - (1 - \text{LGD}) D(T_{i-1}, T_i) \\ & \gg \int_{T_{i-1}}^{T_i} (1 - \text{LGD}) r(u) D(T_{i-1}, u) du, \end{aligned}$$

and therefore Corollary 3.5.9 is applicable.

3.5. Volatility Calibration of Default Intensity

Again, the final price of a CCL on a floating rate loan is the sum of discounted integrals over a zero-coupon bond put option on the intensity multiplied by a deterministic function:

$$\begin{aligned}
 & \text{FloatLoan}_{a,b}^{\text{CCL}}(t) \\
 &= \mathbf{1}_{\{\tau > t\}} \sum_{i=a+1}^b D(t, T_{i-1}) \int_{T_{i-1}}^{T_i} h_i^{\text{CCL}}(u) \text{ZBP} \left(t, T_{i-1}, u; H \left(T_{i-1}, u; x_{\lambda}^{\beta^*}(0) \right); \beta \right) du. \\
 & \quad - \mathbf{1}_{\{\tau > t\}} F_{\text{float}} \sum_{i=a+1}^b \alpha_i D(t, T_i) H(t, T_i).
 \end{aligned}$$

4 Regime Switching

Figure 1.1.1 in the introduction indicates that the distribution of iTraxx® Main spread changes has significantly changed between the time period before and after mid of 2007. Without the use of any statistical method the distribution shows that there is at least one structural break in this time series. Constant component models would omit this economic behavior by using one distribution only for the whole time period. The regime switching models offer one option to overcome this problem. In these models a state variable is introduced which may be an economic one. The distribution of the process thus, becomes state dependent. This state variable can be modeled as a Markov process.

The regime switching models can be traced back to the early work of LINDGREN (1978) and became popular after the seminal work of HAMILTON (1988). There are many papers on this topic as, among others, GRAY (1996) or ANG & BEKAERT (2002). Most papers on regime switching have no pricing background. In this thesis the effect and the importance of the regime switching will be shown.

This chapter is divided in several sections. The first section contains a statistical analysis on spread data. After summary statistics estimates are given on mixture distributions and regime switching distributions, followed by a short introduction into structural breaks. The second section introduces the mathematical treatment of regime switching in general, and summarizes the finally chosen RS CIR-EJ++ model.

4.1 Statistical Analysis of CDS Spreads

This section substantiates the application of a regime switching model for pricing based on one sufficient requirement, namely that all distribution assumptions include the normal distribution. This provides easy handling of and direct meaning for all parameters. In the final step, a model of the CIR class is used which implies a non-central chi-squared distribution to avoid negative rates. This is a necessary simplification since a more detailed analysis would by far exceed this thesis and its main purpose. A more detailed analysis including testing of CIR parameters on bonds is done in WONG & WONG (2008). It uses the efficient method of moments developed in BANSAL, GALLANT & TAUCHEN (1995) and GALLANT & TAUCHEN (1996). This method requires the semi-nonparametric density estimation which is also described there.

The calibration of the CIR(-EJ) model in Chapter 3 uses the default intensities extracted from CDS spread data. The five year bucket is the most liquid CDS tenor in the market. A simple transformation of the CDS spreads into default intensities is obtained by the credit triangle approximation:¹

$$\lambda_{5\text{yr}} \approx \frac{\text{CDS}_{5\text{yr}}}{1 - \text{RR}}$$

The CDS indices are the most liquid instruments in the market, the indices of interest being the main indices of the European and US market, iTraxx® Main five year and CDX® IG five year, respectively. Both indices contain the 125 most liquid names with investment grade rating regarding some index rules. The special features of the index compared to single name CDS are not discussed here.² The spreads are taken from the on-the-run series as again this is the most liquid data. The time horizon of the data is from mid 2004 until March 2012 containing almost 2,000 business days.

The analysis starts with a summary statistic and a unit root test. Afterwards the normal and the normal-mixture distributions are tested and compared. Finally, the regime switching parameters are estimated using an EM algorithm which alternates between expectation (E) and maximization (M) steps. The E-step computes the log-likelihood using the current estimate of parameters. The M-step computes parameters maximizing the expected log-likelihood found in the E-step.

4.1.1 Summary Statistics

Table 4.1.1 presents the summary statistic of the data. The shown data is in basis points (bps). The iTraxx® Main and the CDX® IG behave in the same manner. The spread range is from 20 to 280 bps, whereas the range of daily spread changes extends from -50 and +50 bps. But upon exclusion of the upper and lower 5% the range of daily spread changes reduces to -6 to +14 bps only indicating very few extreme scenarios.

A formal test for unit roots in the variables was done by an augmented Dickey-Fuller (ADF) test³ based on a simple AR(1) model. The test shows that the spread levels contain a unit root indicating non-stationarity. In contrast, the spread changes do not have a unit root and are stationary. Therefore, the spread changes have been further tested for their distributions. The test results are summarized in Table 4.1.2.

4.1.2 Normal Mixture Distribution

Here, tests for normal and normal mixture distributions were performed. These tests do not take into account different regimes. The parameters are estimated using an EM algorithm.

¹See O'KANE (2008), page 130, for a more detailed discussion. The standard CDS recovery rate in the market is 40%.

²Further information are available on www.markit.com.

³Significance level: 5%, critical value: -1.9416

4.1. Statistical Analysis of CDS Spreads

(in bps)	iTraxx® Main		CDX® IG	
value	spread level	spread changes	spread level	spread changes
mean	80	0.036	90	0.013
st. dev.	49	3.9	49	4.3
skewness	0.64	-0.45	1.1	-0.2
kurtosis	2.4	15	4.2	35
maximum	216	23	280	45
95% quantile	198	11	244	13
5% quantile	24	-5.9	35	-5.4
minimum	20	-39	29	-53

Table 4.1.1: Summary statistics of indices (Copyright©2012 Markit Group Limited)

	iTraxx® Main		CDX® IG	
value	spread level	spread changes	spread level	spread changes
stationary	no	yes	no	yes
p value	0.45	1e-3	0.35	1e-3
ADF t-stat	-0.56	-39	-0.83	-43

Table 4.1.2: Augmented Dickey-Fuller test

Afterwards, a Kolmogorov-Smirnov (KS) test is applied.⁴The results are shown in Table 4.1.3. The '# normal' column gives the numbers of normal distributions in the mixture model where its weight is determined by w . μ and σ are the mean and the standard deviation of the normal distribution.

# normal	iTraxx® Main spread changes				CDX® IG spread changes			
	p value	w	μ	σ^2	p value	w	μ	σ^2
1	7e-38	1	0.036	15	8e-49	1	0.013	19
2	0.0094	0.48	-0.07	0.37	0.0014	0.73	-0.016	1.7
		0.52	0.13	28		0.27	0.092	65
3	0.36	0.35	-0.056	0.16	0.45	0.43	-0.076	0.39
		0.49	0.018	8.9		0.51	0.12	11
		0.16	0.3	67		0.06	-0.3	218

Table 4.1.3: Test results for distribution of spread changes

⁴Again, a 5% significance level is used for the KS test. The test statistics determines the maximum difference between the empirical cumulative distribution function (CDF) $F(x)$ and the tested CDF $G(x)$:

$$\max(|F(x) - G(x)|)$$

The p value for the normal distribution demonstrates that one distribution is not adequate to explain the whole time series. The mixture of two normal distributions on the one hand exhibits a distribution with negative mean and small variance indicating a good economy, on the other a normal distribution with high positive mean and large variance indicating a bad economy. The weights and the levels of the second distribution are different for iTraxx® Main and CDX® IG. The weights of iTraxx® Main are both half whereas the weights of CDX® IG are three quarter and one quarter. The p values of both are much higher but still fail the 1% and 5% level. A mixture distribution comprising of three normal distributions has a p value greater than 35% for both series. The splitting into three distributions reveals another extreme distribution with a small weight. Of the remaining two distributions, one almost has no movement, while the other one is in-between.

4.1.3 Regime Switching EM Algorithm

This subsection is based on the work of ENGEL & HAMILTON (1990) and MAALAOU CHUN, DIONNE & FRANÇOIS (2010a) and presents the regime switching normal model as well as the EM algorithm to find its best parameter estimation.

Regime switching means that a hidden Markov process $S(t) = \{1, \dots, Z\}$ with Z different regimes exists. The spread distribution is assumed to be different in the different states. Let $y(t)$ be the realized spread change at time t and assume that in each state $i = 1, \dots, Z$, the spread dynamics are normally distributed with mean μ_i and variance σ_i^2 , then

$$y(t) |_{S(t)=i} \sim \mathcal{N}(\mu_i, \sigma_i^2), \quad i = 1, \dots, Z.$$

In this subsection, discrete time is used for spreads on a daily basis. Besides let p_{ij} be the transition probability of the Markov process to move from one state to another:

$$\mathbb{P}(S(t) = j | S(t-1) = i) =: p_{ij}, \quad i, j = 1, \dots, Z \quad (4.1.1)$$

where the sum of each row is one $\sum_{j=1}^Z p_{ij} = 1$ for $i = 1, \dots, Z$. The distribution parameters and the transition probabilities completely determine the probability law of $y(t)$. They are summarized in the vector $\theta := (\mu_1, \dots, \mu_Z, \sigma_1, \dots, \sigma_Z, p_{11}, p_{12}, \dots, p_{ZZ})$. The probability density function conditional on the regime and the parameters is

$$f(y(t) | S(t) = i, \theta) = \frac{1}{\sqrt{2\pi\sigma_i}} \exp\left(-\frac{(y(t) - \mu_i)^2}{2\sigma_i^2}\right), \quad i = 1, \dots, Z. \quad (4.1.2)$$

Let ξ_i be the unconditional probability that the first observation is in state i

$$\xi_i := \mathbb{P}(S(1) = i; \theta).$$

The estimator is calculated on data by

$$\hat{\xi}_i = \frac{\sum_{j \neq i} p_{ji}}{\sum_{j \neq i} p_{ij} + \sum_{j \neq i} \sum_{k \neq j} p_{jk}}$$

which in case of $Z = 2$ is reduced to

$$\hat{\xi}_1 = \frac{(1 - p_{22})}{(1 - p_{11}) + (1 - p_{22})} \text{ and } \hat{\xi}_2 = 1 - \hat{\xi}_1.$$

The model specification above is supplemented by the following description of the EM algorithm. Incorporating Bayesian priors $(\zeta_1, \zeta_2, \zeta_3)$ ⁵ the different normal distribution parameters are computed for $j = 1, \dots, Z$ as

$$\hat{\mu}_j = \frac{\sum_{t=1}^T y(t) \cdot \mathbb{P}(S(t) = j | y(1), \dots, y(T); \hat{\theta})}{\zeta_3 + \sum_{t=1}^T \mathbb{P}(S(t) = j | y(1), \dots, y(T); \hat{\theta})} \quad (4.1.3)$$

$$\hat{\sigma}_j^2 = \frac{1}{\zeta_1 + 1/2 \sum_{t=1}^T \mathbb{P}(S(t) = j | y(1), \dots, y(T); \hat{\theta})} \cdot \left(\zeta_2 + \frac{1}{2} \sum_{t=1}^T (y(t) - \hat{\mu}_j)^2 \mathbb{P}(S(t) = j | y(1), \dots, y(T); \hat{\theta}) + \frac{1}{2} \zeta_3 \hat{\mu}_j^2 \right) \quad (4.1.4)$$

In case of not taking into account the Bayesian priors the estimators can be explained as follows. Both, the mean and the variance are simply the weighted sum and the weighted sum of squared differences, respectively. The weighting is over the probabilities to appear in a specific regime.

The estimators for the transitions probabilities are given by

$$\hat{p}_{ii} = \frac{\sum_{t=2}^T \mathbb{P}(S(t) = i, S(t-1) = i | y(1), \dots, y(T); \hat{\theta})}{\sum_{t=2}^T \mathbb{P}(S(t-1) = i | y(1), \dots, y(T); \hat{\theta}) + a_i} \quad i = j \quad (4.1.5)$$

$$\hat{p}_{ij} = \frac{\sum_{t=2}^T \mathbb{P}(S(t) = j, S(t-1) = i | y(1), \dots, y(T); \hat{\theta}) + a_i / (Z - 1)}{\sum_{t=2}^T \mathbb{P}(S(t-1) = i | y(1), \dots, y(T); \hat{\theta}) + a_i} \quad i \neq j \quad (4.1.6)$$

$$a_i = \hat{\xi}_i - \mathbb{P}(S(1) = i | y(1), \dots, y(T); \hat{\theta})$$

for all $i, j = 1, \dots, Z$. The \hat{p}_{ij} are the fractions of transitions from state i to state j adjusted by a_i . These adjustments favor states characterized by high start probabilities, and yet unlikeliness to obey the ratio $(\hat{\xi}_i)$. The size of adjustments is very small normally.

The task remaining is to compute the conditional probabilities in the formulas above. Details of this computation can be found in the appendix of HAMILTON (1990). The probabilities

⁵As described in ENGEL & HAMILTON (1990), these prevent singularities in the likelihood function in specific situations.

are conditional on the full sample $y(1), \dots, y(T)$ which is called smoothed inference. In the first step, the conditional probabilities of $S(t)$ are calculated using all information of y up to time t (called filter inference). In the second step the smoothed probabilities are extracted. The calculation is performed iteratively and starts at $t = 2, \dots, T$ using

$$\begin{aligned} & f\left(y(t) \mid y(t-1), \dots, y(1); \hat{\theta}\right) \\ &= \sum_{j=1}^Z \sum_{i=1}^Z p_{ij} \cdot f\left(y(t) \mid S(t) = j; \hat{\theta}\right) \cdot \mathbb{P}\left(S(t-1) = i \mid y(t-1), \dots, y(1); \hat{\theta}\right) \end{aligned}$$

$$\begin{aligned} & \mathbb{P}\left(S(t) = j \mid y(t), \dots, y(1); \hat{\theta}\right) \\ &= \frac{\sum_{i=1}^Z p_{ij} \cdot f\left(y(t) \mid S(t) = j; \hat{\theta}\right) \cdot \mathbb{P}\left(S(t-1) = i \mid y(t-1), \dots, y(1); \hat{\theta}\right)}{f\left(y(t) \mid y(t-1), \dots, y(1); \hat{\theta}\right)} \end{aligned}$$

for all $j = 1, \dots, Z$. The p_{ij} are given by equations (4.1.5) and (4.1.6), $f\left(y(t) \mid S(t) = j; \hat{\theta}\right)$ by (4.1.2). All other terms are known by the respective previous steps of iteration. The iteration is initialized at $t = 1$ by the equations

$$\begin{aligned} f\left(y(1) \mid \hat{\theta}\right) &= \sum_{j=1}^Z f\left(y(1) \mid S(1) = j; \hat{\theta}\right) \cdot \xi_j \\ \mathbb{P}\left(S(1) = j \mid y(1); \hat{\theta}\right) &= \frac{f\left(y(1) \mid S(1) = j; \hat{\theta}\right) \cdot \xi_j}{f\left(y(1) \mid \hat{\theta}\right)}, \quad \text{for all } j = 1, \dots, Z \end{aligned}$$

The smoothed inference probabilities $\mathbb{P}\left(S(t) = i \mid y(T), \dots, y(1); \hat{\theta}\right)$ are the targets for each $i = 1, \dots, Z$ and all $t = 1, \dots, T$. This is done by a further iterative calculation starting at $\tau = t + 1$ with

$$\begin{aligned} & \mathbb{P}\left(S(\tau) = j, S(t) = i \mid y(\tau), \dots, y(1); \hat{\theta}\right) \\ &= p_{ij} \cdot f\left(y(\tau) \mid S(\tau) = j; \hat{\theta}\right) \cdot \frac{\mathbb{P}\left(S(t) = i \mid y(t), \dots, y(1); \hat{\theta}\right)}{f\left(y(\tau) \mid y(t), \dots, y(1); \hat{\theta}\right)} \end{aligned}$$

for all $j = 1, \dots, Z$. The iteration steps for $\tau = t + 2, \dots, T$ use the equation

$$\begin{aligned} & \mathbb{P}\left(S(\tau) = k, S(t) = i \mid y(\tau), \dots, y(1); \hat{\theta}\right) \\ &= \sum_{j=1}^Z p_{jk} \cdot f\left(y(\tau) \mid S(\tau) = k; \hat{\theta}\right) \cdot \frac{\mathbb{P}\left(S(\tau-1) = j, S(t) = i \mid y(\tau-1), \dots, y(1); \hat{\theta}\right)}{f\left(y(\tau) \mid y(\tau-1), \dots, y(1); \hat{\theta}\right)}. \end{aligned}$$

Using the iteration results, the smoothed inference probabilities are computed as

$$\mathbb{P}\left(S(t) = i | y(T), \dots, y(1); \hat{\theta}\right) = \sum_{j=1}^Z \mathbb{P}\left(S(T) = j, S(t) = i | y(T), \dots, y(1); \hat{\theta}\right).$$

The probabilities $\mathbb{P}(S(t) = j, S(t-1) = i | y(1), \dots, y(T); \theta)$ necessary for the transition probabilities are computed in the same way.

The EM algorithm uses a start parameter vector θ . The maximum likelihood function increases in each step as proven in HAMILTON (1990) by recalculating the equations (4.1.3)-(4.1.6). Using the Bayesian prior the iteration provides the maximum of the generalized objective function

$$z(\theta) = \log f(y(T), \dots, y(1); \theta) - \sum_{i=1}^Z \left(\frac{\zeta_3 \cdot \mu_i^2}{2 \cdot \sigma_i^2} + \zeta_1 \cdot \log \sigma_i^2 + \frac{\zeta_2}{\sigma_i^2} \right)$$

instead of the log-likelihood function which is calculated as

$$\log f(y(T), \dots, y(1); \theta) = \sum_{t=1}^T \log f(y(t) | y(t-1), \dots, y(1); \theta).$$

Thus, the algorithm is fully described. The results of its application are given in the next subsection.

4.1.4 Regime Switching Normal Distribution

Tables 4.1.4 and 4.1.5 show the results for the case of two states ($Z = 2$). The first table contains the distribution parameters of different regimes. The second table presents the daily transition probabilities for the switching between the regimes.

regime	iTraxx® Main spread changes			CDX® IG spread changes		
	p value	μ	σ^2	p value	μ	σ^2
1 of 2	3e-4	-0.062 (0.018)	0.25 (0.017)	0.006	-0.067 (0.025)	0.59 (0.034)
2 of 2	1e-5	0.099 (0.14)	24.40 (2.1)	4e-12	0.086 (0.19)	35 (4.7)

Table 4.1.4: Results for regime switching normal distribution with two regimes (standard errors in parentheses)

Again, the results are pretty similar for the iTraxx® Main and CDX® IG. Both indices have two really different regimes. With a slightly negative mean and small variance the first

regime	iTraxx® Main		CDX® IG	
	to 1 of 2	to 2 of 2	to 1 of 2	to 2 of 2
from 1 of 2	0.979	0.021	0.963	0.037
from 2 of 2	0.014	0.986	0.035	0.965

Table 4.1.5: Daily transition probabilities for two regimes

regime relates to a good state of economy. The second regime has a positive mean and a huge variance indicating a bad economy. The p values are calculated by running a KS test presuming a normal distribution without regime switching conditional on all values where the specific state has the highest probability. The p values again indicate that two states are not sufficient to explain the distribution of the spread changes. The probabilities that the state switches from one day to another are 1.4% to 3.7%. These numbers are fairly high having in average more than three regime switchings per year. These results are in line with ALEXANDER & KAECK (2008) which do a regime dependent regression analysis on the iTraxx® family. But there time horizon is only from 2004 to 2007 where the probability for the good state of economy is higher.

In the next step the application of three different regimes is examined. Tables 4.1.6 and 4.1.7 show the results.

regime	iTraxx® Main spread changes			CDX® IG spread changes		
	p value	μ	σ^2	p value	μ	σ^2
1 of 3	0.004	-0.063 (0.017)	0.2 (0.014)	0.009	-0.065 (0.024)	0.44 (0.028)
2 of 3	0.41	0.07 (0.089)	5.7 (0.33)	0.36	0.024 (0.097)	8.6 (0.47)
3 of 3	0.35	0.11 (0.28)	44 (4.15)	0.017	0.19 (0.62)	103 (17)

Table 4.1.6: Results for regime switching normal distribution with three regimes (standard errors in parentheses)

The first of the three regimes is almost the same as the first of two regimes, while the second of two regimes is split into two different regimes now. The variances of the second and third regimes can be categorized as middle and very high, respectively. Moreover, both have a positive mean. The p values are much higher than in case of two regimes. Besides the test cannot be rejected on a 1% significance level in most regimes.

The daily transition probabilities presented in Table 4.1.7 are similar to the two-regimes

4.1. Statistical Analysis of CDS Spreads

regime	iTraxx® Main			CDX® IG		
	to 1 of 3	to 2 of 3	to 3 of 3	to 1 of 3	to 2 of 3	to 3 of 3
from 1 of 3	0.985	0.015	0.000	0.986	0.014	0.000
from 2 of 3	0.014	0.976	0.01	0.012	0.98	0.008
from 3 of 3	0.000	0.011	0.989	0.000	0.027	0.973

Table 4.1.7: Daily transition probabilities for three regimes

case. The likelihood for remaining in the regime is low. The probabilities to skip the middle regime are almost zero.

The findings with three different regimes are in line with the work of GIESECKE ET AL. (2011). They suggested a high, a middle and a low default risk based on an empirical analysis of the bond market.

As mentioned in the beginning the reason of this subsection is to justify the use of the regime switching models. The target is not the presentation of pricing parameters which besides would not be possible because the estimation is based on normal distributions and not on non-central chi squared distributions. Moreover, the results of this section are done under the historical measure and not under the risk-neutral pricing measure.

4.1.5 Detecting Regime Switchings

Various methods have been developed to detect regime switchings. The Chow test is an obvious choice to find structural breaks or changes in the distribution. MAALAOUI CHUN, DIONNE & FRANÇOIS (2010b) investigated the detection of regime shifts in corporate credit spreads. They used a three-step procedure starting with a pre-whitening step of the process, while steps two and three check on shifts in the mean and variance, respectively.

GIESECKE ET AL. (2011) analyzed the average duration of an NBER (national bureau of economic research) default cycle and recession. Using 150 years of historic data they found that a default cycle has an average duration of 3.2 years where the recession period takes 1.5 years.

In this subsection the state probabilities conditional on the historical spread changes are shown. The conditional probabilities for state one (good economy) are plotted in Fig. 4.1.1 for iTraxx® Main (a) and CDX® IG (b).

These figures show that the probabilities are either close to zero or in the near of one. Less than 3% of the probabilities are in the interval $[1/3, 2/3]$. Again, the iTraxx® Main and CDX® IG behave very similar. Both series indicate the good regime from 2004 to the mid of 2007. Since 2007 the second regime is predominant with some smaller interceptions of the first state, especially in case of the US index CDX® IG.

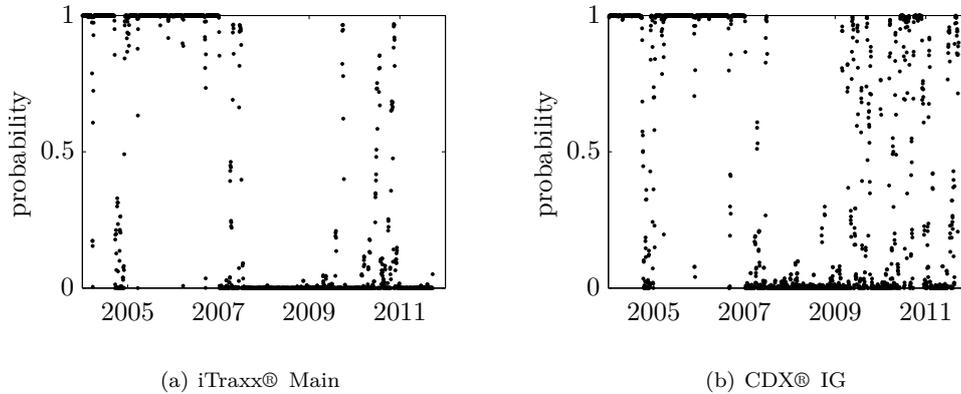


Figure 4.1.1: Conditional probabilities for state one of two (good economy state)

Figures 4.1.2(a)-(c) present the conditional probabilities for the three regimes for iTraxx® Main. Figure 4.1.2(a) shows a very high probability for the first regime until mid of 2007 and is almost the same as 4.1.1(a) for two regimes. In the subsequent period the other two regimes alternate. But even then stable periods are indicated lasting for a half year or longer.

In Chapter 6 the results will be used for pricing. All, one to three regimes will be applied and the effects in pricing are discussed.

4.2 Regime Switching Model

This section describes the formal mathematical definition of regime switching. The first subsection presents the general framework. The final shifted regime switching square root (jump) diffusion (RS CIR(-EJ)++) model is determined in the second part.

4.2.1 General Regime Switching Framework

In the literature on interest rates and default intensities there are different approaches to model regime switching processes. One of them is the hidden Markov model summarized in the book of ELLIOTT, AGGOUN & MOORE (1995) where the drift and diffusion parameters are modeled by an underlying Markov process. A different approach is the conditional Markov chain as discussed in YIN & ZHANG (1997) and, with detail of its application, in BIELECKI & RUTKOWSKI (2004b) where the conditional Markov process models the migration process of credit rating classes. This thesis takes into account the first approach only as the second one is kind of a rating migration model, and therefore, is not an issue of this thesis.

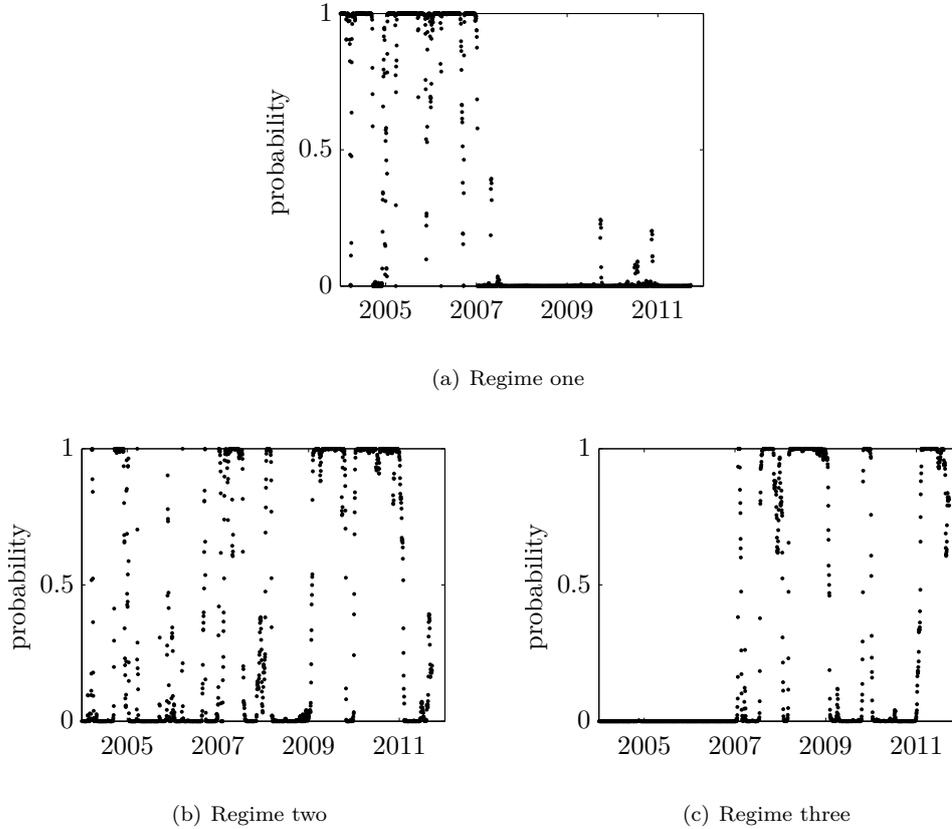


Figure 4.1.2: Conditional probabilities for iTraxx® Main with three regimes

The hidden Markov model was split further resulting, e.g., in the model of LANDÉN (2000) where the Markov process is driven by a market point process, and in the model of ELLIOTT, AGGOUN & MOORE (1995) where a generator matrix determines the state process and its transitions. In this thesis the second approach using generator matrices is applied.

Assumption 4.2.1. *The economic state variable is modeled by an \mathbb{F} -adapted continuous-time hidden Markov chain process $S(t)$ with a finite state space $\mathcal{S} := (s_1, \dots, s_Z)$. As in ELLIOTT, CHAN & SIU (2005) the state space is described by a finite set of unit vectors $\{e_1, \dots, e_Z\}$ where $e_i = (0, \dots, 1, \dots, 0) \in \mathbb{R}^Z$ without loss of its generality.*

$p_{ij}(t, T) := \mathbb{P}(S(T) = e_j | S(t) = e_i)$ denotes the transition probabilities of S from state e_i to state e_j for all times $t \leq T$, $i, j = 1, \dots, Z$, $Q(t, T) := [p_{ij}(t, T)]$. Let $A(t) = [a_{ij}(t)]_{i,j=1,\dots,Z}$ denote the generator matrix of the Markov chain process. The real matrix $A(t)$ satisfies the usual requirements as $a_{ii}(t) = -\sum_{j \neq i} a_{ij}(t)$ and $a_{ij}(t) \geq 0$ for all $i \neq j$. The transition probability matrix and the generator matrix are linked by the forward

Kolmogorov equation

$$\frac{dQ(t, T)}{dT} = A(T) Q(t, T), \quad Q(T, T) = I.$$

In case of a constant generator matrix $A(t) \equiv A$ the solution is given by

$$Q(t, T) = \exp((T - t) A).$$

From page 198 in ELLIOTT, AGGOUN & MOORE (1995), the semi-martingale representation of $S(t)$ is given by

$$S(t) = S(0) + \int_0^t A(u) S(u) du + M(t)$$

where $M(t)$ is an \mathbb{R}^Z -valued \mathbb{P} -martingale process.

4.2.2 RS CIR(-EJ)++ Model

In this subsection the final model is presented which will be used in the pricing tools. Assumptions 3.2.1 and 3.2.2 are extended to the regime-switching case. The state variable S is driven by the generator matrix A as described in Assumption 4.2.1.

Assumption 4.2.2. *Each parameter in the parameter vector is itself a vector of length Z (number of states). The resulting RS CIR(-EJ)++ model for interest rates is given by⁶*

$$r(t) = x_r^\alpha(t) + \varphi^{RS \text{ CIR}}(t; \alpha) \text{ where}$$

$$dx_r^\alpha(t) = \langle \kappa_r, S(t) \rangle (\langle \theta_r, S(t) \rangle - x_r^\alpha(t)) dt + \langle \sigma_r, S(t) \rangle \sqrt{x_r^\alpha(t)} dW(t)$$

and for the default intensity by

$$\lambda(t) = x_\lambda^\beta(t) + \varphi^{RS \text{ CIR-EJ}}(t; \beta) \text{ where}$$

$$\begin{aligned} dx_\lambda^\beta(t) = & \langle \kappa_\lambda, S(t) \rangle (\langle \theta_\lambda, S(t) \rangle - x_\lambda^\beta(t)) dt + \langle \sigma_\lambda, S(t) \rangle \sqrt{x_\lambda^\beta(t)} d\bar{W}(t) \\ & + J(\langle \gamma_\lambda, S(t) \rangle) d\bar{N}(t; \langle \varsigma_\lambda, S(t) \rangle) \end{aligned}$$

In case of a zero jump-diffusion component ($\varsigma_\lambda = 0$) this model is simplified into the RS CIR++ model. The state space process S is independent of the Brownian motions W and \bar{W} and of the jump component \bar{N} .

The deterministic shift is again chosen to match the exact term structure as in the CIR++ or CIR-EJ++ model. Besides the starting point of the process, each parameter is now a

⁶ $\langle \cdot, \cdot \rangle$ is a scalar product in \mathbb{R}^Z that for any $a, b \in \mathbb{R}^Z$ is: $\langle a, b \rangle = \sum_{i=1}^Z a_i b_i$.

4.2. Regime Switching Model

vector of length Z . The other way round each state $i = 1, \dots, Z$ has one parameter vector with constants $(x(0), \kappa_i, \theta_i, \sigma_i, \varsigma_i, \gamma_i)$. The Feller condition is satisfied if it is satisfied for each of the Z one-dimensional parameter vectors.

5 Numerical Implementation to a Trinomial Tree

The models presented in Chapters 3 and 4 provide analytical solutions for some products as presented above. However, there is a large number of products without analytical solutions such as products with early exercise feature or multiple optionality. In case of regime switching most products do not have a closed form solution. Numerical procedures are necessary to price such products. Many different numerical techniques can be used such as Monte-Carlo simulation, tree valuation, finite difference scheme and so on. Each of these techniques has some advantages and disadvantages upon usage. Monte-Carlo simulation, for example, can easily handle a path-dependent payoff, but runs into difficulties in the treatment of American options. In case of the tree it is the other way around: optionality is easy to handle, but the treatment of path-dependency in the payoff needs more effort. Both procedures share the need of a discrete approximation of the continuous process.

Here, a tree procedure is the natural choice, as the pricing of optionality is one of the main issues of this thesis. A requirement of the numerical implementation is that the parameter vectors α (interest rate) and β (default intensity) are calibrated to the market data already. The deterministic shift function is free to match the term structure exactly.

For the general tree implementation for a CIR++ model, in the first instance all direct references to interest rates or default intensities are avoided. The process is

$$a(t) = x^v(t) + \varphi(t; v), \quad t \geq 0,$$

where x^v follows the CIR process

$$dx^v(t) = \kappa(\theta - x^v(t))dt + \sigma\sqrt{x^v(t)}dW(t) \quad (5.0.1)$$

with parameter vector $v = (x^v(0), \kappa, \theta, \sigma)$, $x^v(0)$, κ , θ and σ are positive constants. φ depends on v and is integrable on closed intervals. After building the CIR tree the process is extended by the jump components and the regime-switching process. The index v is omitted in this chapter as there is one process only, and there are many indices in the tree building.

Tree building for a (RS) CIR(-EJ)++ process is characterized by several difficulties:

- The CIR process is naturally not recombining because of its rate depending volatility.

- The jump component has to be reflected in the tree.
- The regime-switching property must be covered in a numerically manageable way.
- The tree has to reflect two dimensions: the interest and the default component.
- The two-dimension tree has to handle the correlation structure between the two processes.

Chapter 5 presents at least one solution for each of these five items. The first section presents a solution to avoid the non-recombining property of the CIR process. In the second section two different ways of one-dimensional tree building procedure are shown and compared. In the third section the handling of jumps, of regime switching and deterministic shift is presented. The fourth section combines the interest rate and default intensity trees and contains two different ways to incorporate the correlation. The last section shows the tree valuation procedure.

5.1 Recombining of CIR Process

The volatility of the CIR model is rate dependent which means that a direct tree is non-recombining: The value of an up-down movement is different from a down-up movement as shown in Fig. 5.1.1. Each distance for an up movement exceeds always the respective distance of a down movement.

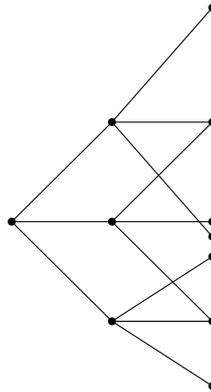


Figure 5.1.1: The CIR tree - non recombining

Direct implementation creates much more branches than there are in a recombining tree and therefore, enhances the simulation time and the memory usage by far. An approach to solve this problem is to find a new process \tilde{x} with constant volatility calculated from x .

There are many possibilities to define a process \tilde{x} with constant volatility. In the next subsections two different approaches are presented. The first one suggested by NELSON

& RAMASWAMY (1990) sets the volatility of \tilde{x} to one. The second approach of BRIGO & MERCURIO (2006) defines $\tilde{x} := \sqrt{x}$.

5.1.1 To Find a Recombining Auxiliary Process \tilde{x}

The volatility of the x -process does not recombine. A new process \tilde{x} is defined by $\tilde{x}(t) := f(x(t), t)$. The target is constant volatility for \tilde{x} which implicates a recombining tree for \tilde{x} . The full calculation by Itô's lemma is as follows:¹

$$d\tilde{x}(t) = \frac{\partial \tilde{x}(t)}{\partial t} dt + \frac{\partial \tilde{x}(t)}{\partial x(t)} dx(t) + \frac{1}{2} \frac{\partial^2 \tilde{x}(t)}{\partial x(t)^2} d\langle x(t) \rangle. \quad (5.1.1)$$

Inserting (5.0.1) into (5.1.1) results in

$$d\tilde{x}(t) = \mu(\tilde{x}, t) dt + \sigma_{\tilde{x}}(\tilde{x}, t) dW(t)$$

where

$$\mu(\tilde{x}, t) = \frac{\partial \tilde{x}(t)}{\partial t} + \frac{\partial \tilde{x}(t)}{\partial x(t)} \kappa(\theta - x(t)) + \frac{1}{2} \frac{\partial^2 \tilde{x}(t)}{\partial x(t)^2} \sigma^2 x(t) \quad (5.1.2)$$

and

$$\sigma_{\tilde{x}}(\tilde{x}, t) = \frac{\partial \tilde{x}(t)}{\partial x(t)} \sigma \sqrt{x(t)} \quad (5.1.3)$$

NELSON & RAMASWAMY (1990) proposed the easiest constant volatility and set $\sigma_{\tilde{x}}(\tilde{x}, t)$ to one:

$$\frac{\partial \tilde{x}(t)}{\partial x(t)} = \frac{1}{\sigma \sqrt{x(t)}}$$

Integration results in

$$f(x(t), t) = \tilde{x}(t) = \int^{x(t)} \frac{1}{\sigma \sqrt{u(t)}} du = \frac{2}{\sigma} \sqrt{x(t)},$$

or respectively

$$x(t) = \frac{\sigma^2}{4} \tilde{x}(t)^2.$$

The dynamics of the new process \tilde{x} are given by

$$d\tilde{x}(t) = \mu(\tilde{x}, t) dt + dW(t) \quad (5.1.4)$$

where the drift is calculated as

$$\begin{aligned} \mu(\tilde{x}, t) &= \frac{\partial \tilde{x}(t)}{\partial t} + \frac{\partial \tilde{x}(t)}{\partial x(t)} \kappa(\theta - x(t)) + \frac{1}{2} \frac{\partial^2 \tilde{x}(t)}{\partial x(t)^2} \sigma^2 x(t) \\ &= \frac{\partial^2 \sqrt{x(t)}}{\partial t} + \frac{\partial^2 \sqrt{x(t)}}{\partial x(t)} \kappa(\theta - x(t)) + \frac{1}{2} \frac{\partial^2 \frac{2\sqrt{x(t)}}{\sigma}}{\partial x(t)^2} \sigma^2 x(t) \end{aligned}$$

¹ $\langle \cdot \rangle$ is the quadratic variation and not the scalar product.

$$\begin{aligned}
 &= \frac{1}{\sqrt{x(t)}\sigma} \kappa(\theta - x(t)) - \frac{1}{4} \frac{\sigma}{\sqrt{x(t)}} \\
 &= \frac{2}{\tilde{x}(t)\sigma^2} \kappa\left(\theta - \frac{\tilde{x}(t)^2\sigma^2}{4}\right) - \frac{1}{2} \frac{1}{\tilde{x}(t)} \\
 &= \frac{1}{\tilde{x}(t)} \left[\frac{1}{2} \kappa\left(\frac{4\theta}{\sigma^2} - \tilde{x}(t)^2\right) - \frac{1}{2} \right]
 \end{aligned}$$

The tree is build for \tilde{x} as shown in Fig. 5.1.2(a) and transformed afterwards into the original process x as shown in (b). The symmetry of the transformed process \tilde{x} is due to the constant volatility. The resulting tree is humped to cover the higher volatility of higher process values.

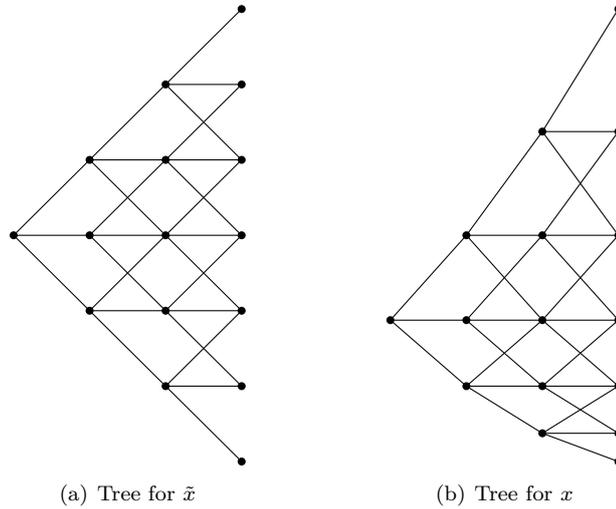


Figure 5.1.2: Recombining tree transformation

5.1.2 Second Recombining Process \tilde{x}

BRIGO & MERCURIO (2006) used a different approach defining the new process \tilde{x} as

$$\tilde{x}(t) := \sqrt{x(t)}.$$

By Itô's lemma the dynamic of \tilde{x} is given by

$$d\tilde{x} = \underbrace{\frac{1}{\tilde{x}(t)} \left[\frac{1}{2} \kappa\left(\theta - \tilde{x}(t)^2\right) - \frac{1}{8} \sigma^2 \right]}_{=: \mu(\tilde{x}, t)} dt + \frac{\sigma}{2} dW(t) \tag{5.1.5}$$

5.1.3 Differences in the Choices of \tilde{x}

The difference between the two methods are marginal and have no impact on the convergence or numerical behavior.

The decision on the method to be applied will be described later. In case of regime-switching the choice of \tilde{x} is important. This is due to the fact that the volatility σ affects the recalculation to the original process in the first recombining process. A more detailed discussion is presented in Subsection 5.3.2.

5.2 CIR Trinomial Tree

There are several approaches for building a tree. Famous papers on term structure models include HULL & WHITE (1994a) and HULL (1996) which present the trees for the Hull-White and Ho-Lee models.

The basic ideas on the trees used in this thesis stem from NELSON & RAMASWAMY (1990) with the first presentation of a binomial tree for the CIR process. Two different approaches following this paper are shown. As discussed later each approach has different advantages and disadvantages. The target tree is trinomial as its convergence is faster and more stable than that of a binomial tree. Trinomial means that each node has three subsequent branches.

In this section, after basic information on building a tree, the two different tree building approaches of NAWALKHA & BELIAEVA (2007) and BRIGO & MERCURIO (2006) are shown. Afterwards they are compared in the last subsection.

5.2.1 Building a Tree - Motivation and Convergence

The basics in tree building include the requirements to obtain a convergence in distribution of the discrete version to the continuous process. Somehow the product prices are conditional expectations of continuous functions. Therefore, the tree convergence is sufficient for the convergence of the prices based on the discrete approximation to the real prices. The proof of the convergence of the Hull-White tree in LESNE, PRIGENT & SCAILLET (2000) forms a basis for this subsection.

Tree approximation requires discretization of the continuous time process. Let t_0, \dots, t_N be the discrete time points for $0 = t_0 < \dots < t_N = T_\infty$. The distances between two points in time are denoted by $\Delta t_n = t_{n+1} - t_n$ for $n = 0, \dots, N - 1$. The discretized process is constant between two time points:

$$x_n = x_n(t) := x(t_n), \text{ defined on } t \in [t_n, t_{n+1}) \text{ for } n = 0, \dots, N.$$

Besides time discretization, the tree uses discrete points in space. These are equi-spaced at each point of time because of the constant volatility of \tilde{x} , the distance is denoted by $\Delta\tilde{x}_n$ for all $n = 0, \dots, N$. \tilde{x}_n^i is the node (n, i) at time $\sum_{k=1}^n \Delta t_{k-1}$ and state $i \cdot \Delta\tilde{x}_n$.

As stated before, each trinomial tree node has three subsequent branches. The three typical movements of \tilde{x}_n^i in one time step are:

- up movement (upper branch) to $\tilde{x}_n^i + \Delta\tilde{x}_{n+1}$ with probability p_u
- remaining constant (middle branch) at \tilde{x}_n^i with probability p_m
- down movement (lower branch) to $\tilde{x}_n^i - \Delta\tilde{x}_{n+1}$ with probability p_d

p_u , p_m and p_d denote the conditional probabilities of the branching processes for an up, middle and down movement, respectively. The state and time dependencies of these probabilities are omitted in the notation. The tree grid of a subsequent time step may differ from that of its preceding one in case of different sizes in time steps, causing the middle node \tilde{x}_n^i to move to \tilde{x}_{n+1}^j for some specified j and according up and down movements of upper and lower branch. This behavior is not treated in this subsection as it does not influence the convergence property.

The probabilities of the nodes become very small or negative for an upward movement at the upper limit and also for a downward movement at the lower limit. To avoid negative probabilities the tree is restricted at these points by a parameter J which inserts multiple node jumps shifting the branches one node down ($J = -1$) or up ($J = 1$) on the upper or lower side, respectively. The nodes within the limits are not shifted ($J = 0$). The exact calculation depends on the branch probabilities and is shown in the next subsections. J depends on state space and on time but it is omitted in the notation.

To achieve consistency of the tree with the time-homogenous dynamics in the first two moments the following probabilities must hold at all nodes

$$\begin{aligned} \mathbb{E} [\tilde{x}(t_{n+1}) | \tilde{x}(t_n) = \tilde{x}_n^i] &= M_n^i \\ \text{Var} [\tilde{x}(t_{n+1}) | \tilde{x}(t_n) = \tilde{x}_n^i] &= (V_n^i)^2 \end{aligned} \tag{5.2.1}$$

where M_n^i and $(V_n^i)^2$ are the mean and the variance of \tilde{x} at time t_{n+1} conditional on $\tilde{x}(t_n) = \tilde{x}_n^i$. Additionally, the probabilities for an up, middle and down movement must sum up to one:

$$p_u + p_m + p_d = 1 \tag{5.2.2}$$

This leads to a system of three equations with three unknowns for the determination of the probabilities for an up, middle and down movement at each node. The calculation of these probabilities is performed in the next two subsections depending on the respective tree building procedure.

The mean M_n^i and variance V_n^i depend on the process \tilde{x} used for the tree. The drift at time t_n and state $i \cdot \Delta\tilde{x}_n$ is denoted as μ_n^i . The three quantities can be computed either for the

dynamics according to Eq. (5.1.4) as

$$\begin{aligned} M_n^i &= \mu_n^i \cdot \Delta t_n + \tilde{x}_n^i \\ V_n &= \sqrt{\Delta t_n} \\ \mu_n^i &= \frac{1}{\tilde{x}_n^i} \left[\frac{1}{2} \kappa \left(\frac{4\theta}{\sigma^2} - (\tilde{x}_n^i)^2 \right) - \frac{1}{2} \right] \end{aligned}$$

or for the dynamics according to the second approach defined in Eq. (5.1.5) as

$$\begin{aligned} M_n^i &= \mu_n^i \cdot \Delta t_n + \tilde{x}_n^i \\ V_n &= \frac{\sigma}{2} \cdot \sqrt{\Delta t_n} \\ \mu_n^i &= \frac{1}{\tilde{x}_n^i} \left[\frac{1}{2} \kappa \left(\theta - (\tilde{x}_n^i)^2 \right) - \frac{1}{8} \sigma^2 \right] \end{aligned}$$

Next the convergence of the discretized tree process to the continuous process must be proved. The Euler approximation of the CIR process according to (5.0.1) is transformed by (5.1.4) to

$$\begin{aligned} \tilde{x}_n - \tilde{x}_{n-1} &= \frac{1}{\tilde{x}_{n-1}} \left[\frac{1}{2} \kappa \left(\frac{4\theta}{\sigma^2} - \tilde{x}_{n-1}^2 \right) - \frac{1}{2} \right] \Delta t_{n-1} + \epsilon_n \\ x_n &= \frac{\sigma}{2} \tilde{x}_n^2 \end{aligned} \quad (5.2.3)$$

The tree determination defined in Eq. (5.2.1) and (5.2.2) and the Euler approximation in Eq. (5.2.3) enable the determination of ϵ_n by the two following equations

$$\begin{aligned} \tilde{x}_n - \tilde{x}_{n-1} &= \frac{1}{\tilde{x}_{n-1}} \left[\frac{1}{2} \kappa \left(\frac{4\theta}{\sigma^2} - \tilde{x}_{n-1}^2 \right) - \frac{1}{2} \right] \Delta t_{n-1} + \epsilon_n \\ \tilde{x}_n - \tilde{x}_{n-1} &= p_u (J+1) \Delta \tilde{x}_n + p_m J \Delta \tilde{x}_n + (J-1) p_d \Delta \tilde{x}_n \end{aligned}$$

Resolving these equations to ϵ_n yields

$$\epsilon_n = p_u (J+1) \Delta \tilde{x}_n + p_m J \Delta \tilde{x}_n + (J-1) p_d \Delta \tilde{x}_n - \frac{1}{\tilde{x}_{n-1}} \left[\frac{1}{2} \kappa \left(\frac{4\theta}{\sigma^2} - \tilde{x}_{n-1}^2 \right) - \frac{1}{2} \right] \Delta t_{n-1}.$$

The choice of mean and variance according to Eq. (5.2.1) leads to

$$\mathbb{E}[\epsilon_n | \mathcal{F}_{n-1}] = 0 \quad \text{and} \quad \mathbb{E}[\epsilon_n^2 | \mathcal{F}_{n-1}] = \Delta t_{n-1}$$

Therefore, $\sum_n \epsilon_n$ will converge to a Brownian motion for $\Delta t_k \rightarrow 0$, $k = 0, \dots, n-1$ as shown in JACOD & SHIRYAEV (1987), Theorem 3.32, page 437. Additionally all requirements of Proposition 1 of LESNE, PRIGENT & SCAILLET (2000) are fulfilled showing the convergence in distribution of the processes x_n and \tilde{x}_n to x and \tilde{x} . On the whole, tree building procedures matching the first two moments converge in distribution to the continuous process.

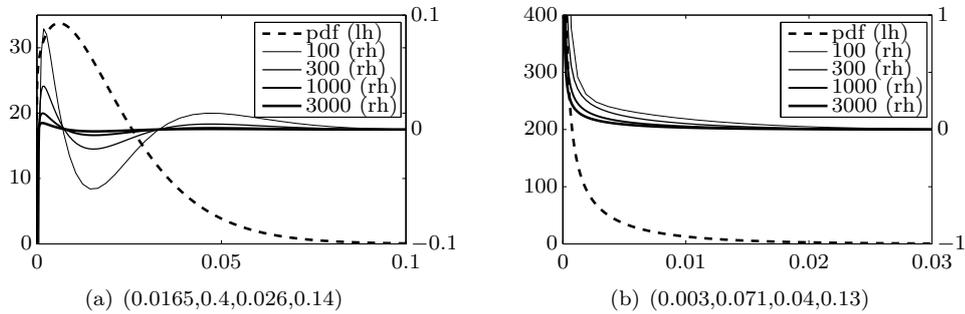


Figure 5.2.1: Probability density of CIR processes (lh) and tree convergence (rh)

Figures 5.2.1 show the densities (pdf) of the CIR processes after one year (a) for the parameter vector (0.0165 0.4 0.026 0.14) and (b) for the calibrated process of the interest rate. The shapes of the densities are different as the left one fulfills the Feller condition and the right one does not. Besides the densities the convergences of the CIR trees are shown by the solid lines. The deviations of the tree value from the analytical densities are presented for different numbers of tree time steps. As shown in the last passage the trees converge to the true values by increasing the number of time steps and thereby decreasing the step size Δt_n . The NB tree introduced in the next subsection is used to demonstrate the convergence in Fig. 5.2.1.

5.2.2 The Approach of Nawalkha & Beliaeva (2007) (NB)

This approach was first introduced by NELSON & RAMASWAMY (1990) and later improved by NAWALKHA & BELIAEVA (2007). The starting point for the process \tilde{x} is calculated by (5.1.4) and (5.1.5), respectively. The forward movements in the tree are characterized by three new branches from each node. In this approach the step size for time Δt_k is fixed to Δt for all k . Therefore, variance V^2 , mean M^i and drift μ^i are all independent of time, but the latter two are state dependent. The tree state space is also fixed for all n and defined by the distance

$$\Delta \tilde{x} := bV$$

for a value b specified later. The parameter b ensures that $\tilde{x}_0 = i \cdot \Delta \tilde{x}$ for one $i \in \mathbb{N}$ and that the tree exactly hits the zero point which is necessary to prevent numerical difficulties at values close to zero. The tree building is shown in Fig. 5.2.2.

The tree building is split into two cases, i.e., case one with tree building for all nodes above zero, and case two for all nodes at zero.

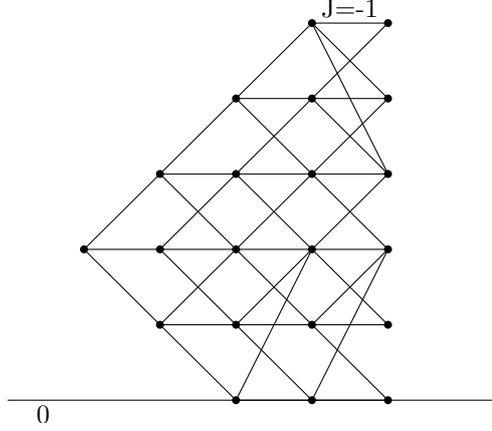


Figure 5.2.2: The NB tree with multiple node movements

Case $\tilde{x}_n^i > 0$:

The NB tree uses the parameter J to insert multiple node jumps at the vertical edges of the tree. Hence, the next up, middle and down movements from a given point \tilde{x}_n^i are given by

$$\begin{aligned}\tilde{x}_{n+1}^{i+J+1} &= \tilde{x}_n^i + (J+1) \cdot \Delta\tilde{x} \\ \tilde{x}_{n+1}^{i+J} &= \tilde{x}_n^i + J \cdot \Delta\tilde{x} \\ \tilde{x}_{n+1}^{i+J-1} &= \tilde{x}_n^i + (J-1) \cdot \Delta\tilde{x}\end{aligned}$$

The value of J will be specified in Eq. (5.2.5) below.

The trinomial probabilities for matching the drift $\mu^i \Delta t$ and the variance V^2 of the process \tilde{x} at time t_n and state i are calculated from Eq. (5.2.1) and (5.2.2)

$$\begin{aligned}\mathbb{E}[\tilde{x}(t_{n+1}) | \tilde{x}(t_n) = \tilde{x}_n^i] &= \tilde{x}_n^i + \mu^i \Delta t \\ \text{Var}[\tilde{x}(t_{n+1}) | \tilde{x}(t_n) = \tilde{x}_n^i] &= V^2 \\ p_u + p_m + p_d &= 1\end{aligned}$$

Inserting the tree probabilities into the expectation value and the variance leads to

$$\begin{aligned}p_u((J+1) \cdot \Delta\tilde{x}) + p_m(J \cdot \Delta\tilde{x}) + p_d((J-1) \cdot \Delta\tilde{x}) &= \mu^i \Delta t \\ p_u((J+1) \cdot \Delta\tilde{x})^2 + p_m(J \cdot \Delta\tilde{x})^2 + p_d((J-1) \cdot \Delta\tilde{x})^2 &= V^2 + (\mu^i \Delta t)^2 \\ p_u + p_m + p_d &= 1\end{aligned}$$

resulting in

$$p_u = \frac{V^2 + (\mu^i \Delta t)^2}{2 \cdot \Delta\tilde{x}^2} - \frac{(2J-1)\mu^i \Delta t}{2 \cdot \Delta\tilde{x}} + \frac{J^2 - J}{2}$$

$$\begin{aligned}
 p_m &= 1 - \frac{V^2 + (\mu^i \Delta t)^2}{\Delta \tilde{x}^2} + \frac{2J\mu^i \Delta t}{\Delta \tilde{x}} - J^2 \\
 p_d &= \frac{V^2 + (\mu^i \Delta t)^2}{2 \cdot \Delta \tilde{x}^2} - \frac{(2J+1)\mu^i \Delta t}{2 \cdot \Delta \tilde{x}} + \frac{J^2 + J}{2}
 \end{aligned} \tag{5.2.4}$$

These probabilities are not easy to handle, and it is difficult to find a direct solution for J and b where all probabilities are between zero and one. Therefore, NAWALKHA, SOTO & BELIAEVA (2007) suggested a simplified approach to get probabilities between zero and one.² Almost all tree steps use the standard case where $J = 0$. Hence, for convergence purposes it is sufficient to match variance and mean in these steps. In case of positive or negative J the correct mean has to be fitted only. In this approach the probabilities for an up, middle and down movement are set to

$$\begin{aligned}
 p_u &= \frac{1}{2b^2} + A \\
 p_m &= 1 - \frac{1}{b^2} \\
 p_d &= \frac{1}{2b^2} - A
 \end{aligned}$$

for each A which ensures that the variance of \tilde{x} is matched for $J = 0$. Using $A = -\frac{J}{2} + \frac{\mu^i \Delta t}{2bV}$ shows that this is the case for all $J \geq 0$ and leads to the following probabilities

$$\begin{aligned}
 p_u &= \frac{1}{2b^2} - \frac{J}{2} + \frac{\mu^i \Delta t}{2bV} \\
 p_m &= 1 - \frac{1}{b^2} \\
 p_d &= \frac{1}{2b^2} + \frac{J}{2} - \frac{\mu^i \Delta t}{2bV}
 \end{aligned}$$

The next step is to find those values for b and J which ensure that the probabilities are well defined. A value below 1 for b induces an impossible negative probability for the middle movement. Besides restrictions for the probabilities for up and down movements are $\frac{1}{b^2} \geq p_u, p_d \geq 0$. Therefore, the following two inequalities are obtained:

$$\begin{aligned}
 \frac{1}{b^2} &\geq \frac{1}{2b^2} - \frac{J}{2} + \frac{\mu^i \Delta t}{2bV} \geq 0 \\
 \frac{1}{b^2} &\geq \frac{1}{2b^2} + \frac{J}{2} - \frac{\mu^i \Delta t}{2bV} \geq 0
 \end{aligned}$$

In summary J has to fulfill

$$\frac{\mu^i \Delta t}{bV} - \frac{1}{b^2} \leq J \leq \frac{\mu^i \Delta t}{bV} + \frac{1}{b^2}$$

J is the number of shifting nodes on the edges of the tree and hence, an integer value. This

²In Subsection 5.2.4 the exact conditions of these probabilities are shown.

restricts b to be less than $\sqrt{2}$. J is set to

$$J = \text{floor} \left(\frac{\mu^i \Delta t}{bV} + \frac{1}{b^2} \right). \quad (5.2.5)$$

As stated above b must be between 1 and $\sqrt{2}$. The exact value for b is chosen as $\sqrt{1.5}$ plus an adjustment term in such a way that the start value is part of the tree and the tree exactly hits the zero point. This avoids calculation problems by closely approaching zero for \tilde{x} . The auxiliary variables b_e and b_c are calculated according to

$$b_e = \frac{\tilde{x}(0)/V}{\text{floor} \left(\frac{\tilde{x}(0)}{\sqrt{1.5}V^2} \right)} \quad \text{and} \quad b_c = \frac{\tilde{x}(0)/V}{\text{floor} \left(\frac{\tilde{x}(0)}{\sqrt{1.5}V^2} + 1 \right)}$$

The value closer to $\sqrt{1.5}$ is chosen for b :

$$b = \begin{cases} b_c & \text{if } |b_c - \sqrt{1.5}| < |b_e - \sqrt{1.5}| \\ b_e & \text{otherwise} \end{cases}$$

In the case of $J = 0$ the normal tree is applied with up-across-down branch. When $J = -1$ the upper limit of the tree is reached, i.e., a further up branch is not possible, and therefore, there are a horizontal branch and two down branches, one of which exhibits a multiple node jump. In case of $J = 1$, the lower limit is reached with respective reverse properties. Both cases are shown in Fig. 5.2.2.

Case $\tilde{x}(t) = 0$:

From $\tilde{x}(t) = 0$ it follows that $x(t) = 0$ and that the variance of the original process is zero. Therefore, NAWALKHA, SOTO & BELIAEVA (2007) proposed to match the mean of the process only and to drop the middle node which results in a binomial tree movement. The probabilities for an up and down movement in that case are given by

$$\begin{aligned} p_u &= \frac{\kappa\theta\Delta t}{x_u} \\ p_d &= 1 - p_u \end{aligned}$$

where x_u is the recalculated value of the original process. The lower branch remains at zero:

$$\tilde{x}_d = 0$$

The up movement is not necessarily one node as this could result in negative probabilities. Using $J \geq 1$ provides, depending on the relationship between x and \tilde{x} ,

$$\tilde{x}_u = bV(J+1) \geq \frac{2}{\sigma} \sqrt{\kappa\theta\Delta t}$$

and

$$\tilde{x}_u = bV(J + 1) \geq \sqrt{\kappa\theta\Delta t},$$

respectively.

This is the NB tree for the approximation of the continuous CIR process. The building procedures as well as the advantages and drawbacks are discussed in Subsection 5.2.4.

5.2.3 The Approach of Brigo & Mercurio (2006) (BM)

This approach is also based on the ideas of NELSON & RAMASWAMY (1990) and has been developed by BRIGO & MERCURIO (2006). Here, the time steps do not need to be equally spaced. At each time point t_n there is a finite number of equ-spaced states with constant vertical step size $\Delta\tilde{x}_n$. This size depends on the time step size and is determined at the end in such a way that positive probabilities are ensured.

Similarly, the variance V_n^2 depends on time when the step size changes. But the state space has no effect as the process \tilde{x} has constant volatility. Mean M_n^i and drift μ_n^i depend on both, time and space.

The BM tree construction uses a different approach to define each subsequent node than the NB tree. Assuming state i at time t_n the process \tilde{x}_n^i can move to \tilde{x}_{n+1}^{j+1} , \tilde{x}_{n+1}^j , \tilde{x}_{n+1}^{j-1} at time t_{n+1} . The value of j defines the absolute tree level and will be defined below. The CIR process is always positive which must hold for the tree. Therefore, the tree nodes are shifted by a predefined level $\varepsilon > 0$ which can be arbitrarily chosen close to zero. But the computational instability in the range very close to zero must be considered.

The BM tree building is shown in Fig. 5.2.3 including details of the shifting by ε , the different horizontal (time) and vertical (space) step sizes, and the fact that the next middle node is not necessarily stable. Always, the subsequent middle node is the node nearest to the expectation value as determined below.

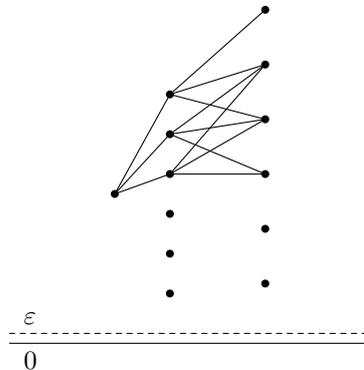


Figure 5.2.3: The BM tree with different time steps

Again, as in the NB tree procedure equations (5.2.1) and (5.2.2) are used to determine the probabilities for the up, middle and down movement at the tree point \tilde{x}_n^i :

$$\begin{aligned} p_u \left(\tilde{x}_{n+1}^j + \Delta \tilde{x}_{n+1} \right) + p_m \tilde{x}_{n+1}^j + p_d \left(\tilde{x}_{n+1}^j - \Delta \tilde{x}_{n+1} \right) &= M_n^i \\ p_u \left(\tilde{x}_{n+1}^j + \Delta \tilde{x}_{n+1} \right)^2 + p_m \left(\tilde{x}_{n+1}^j \right)^2 + p_d \left(\tilde{x}_{n+1}^j - \Delta \tilde{x}_{n+1} \right)^2 &= (V_n)^2 + (M_n^i)^2 \\ p_u + p_m + p_d &= 1 \end{aligned}$$

Contrary to the NB tree, the absolute levels are taken into account. By rearranging the terms the following equations are obtained

$$\begin{aligned} \tilde{x}_{n+1}^j + (p_u - p_d) \Delta \tilde{x}_{n+1} &= M_n^i \\ \left(\tilde{x}_{n+1}^j \right)^2 + 2\tilde{x}_{n+1}^j \Delta \tilde{x}_{n+1} (p_u - p_d) + (\Delta \tilde{x}_{n+1})^2 (p_u + p_d) &= (V_n)^2 + (M_n^i)^2 \\ p_u + p_m + p_d &= 1 \end{aligned}$$

Setting $\eta_n^{i,j} := M_n^i - \tilde{x}_{n+1}^j$ as the difference between the mean and the next middle node and using the first equation to simplify the second one yields

$$\begin{aligned} \Delta \tilde{x}_{n+1} (p_u - p_d) &= \eta_n^{i,j} \\ (p_u + p_d) (\Delta \tilde{x}_{n+1})^2 &= (V_n)^2 + (\eta_n^{i,j})^2 \\ p_u + p_m + p_d &= 1 \end{aligned}$$

Over the whole the probabilities are calculated as

$$\begin{aligned} p_u &= \frac{(V_n)^2}{2(\Delta \tilde{x}_{n+1})^2} + \frac{(\eta_n^{i,j})^2}{2(\Delta \tilde{x}_{n+1})^2} + \frac{\eta_n^{i,j}}{2\Delta \tilde{x}_{n+1}} \\ p_m &= 1 - \frac{(V_n)^2}{(\Delta \tilde{x}_{n+1})^2} - \frac{(\eta_n^{i,j})^2}{(\Delta \tilde{x}_{n+1})^2} \\ p_d &= \frac{(V_n)^2}{2(\Delta \tilde{x}_{n+1})^2} + \frac{(\eta_n^{i,j})^2}{2(\Delta \tilde{x}_{n+1})^2} - \frac{\eta_n^{i,j}}{2\Delta \tilde{x}_{n+1}} \end{aligned} \quad (5.2.6)$$

The level of j is chosen in such a way that the subsequent middle node \tilde{x}_{n+1}^j is as close as possible to the mean M_n^i :

$$j := \text{round} \left(\frac{M_n^i}{\Delta \tilde{x}_{n+1}} \right) \quad (5.2.7)$$

By this the distance between the subsequent middle node value and its expectation value is defined to never be greater than half the state space distance, i.e., $|\eta_n^{i,j}| \leq \frac{\Delta \tilde{x}_{n+1}}{2}$.

In the next step these probabilities are shown to be in the range of zero to one.

$$p_u \geq \frac{(V_n)^2}{2(\Delta \tilde{x}_{n+1})^2} + \frac{\left(\frac{\Delta \tilde{x}_{n+1}}{2} \right)^2}{2(\Delta \tilde{x}_{n+1})^2} - \frac{\frac{\Delta \tilde{x}_{n+1}}{2}}{2\Delta \tilde{x}_{n+1}}$$

$$\begin{aligned}
 &= \frac{(V_n)^2}{2(\Delta\tilde{x}_{n+1})^2} - \frac{1}{8} \stackrel{!}{\geq} 0 \\
 p_u &\leq \frac{(V_n)^2}{2(\Delta\tilde{x}_{n+1})^2} + \frac{\left(\frac{\Delta\tilde{x}_{n+1}}{2}\right)^2}{2(\Delta\tilde{x}_{n+1})^2} + \frac{\frac{\Delta\tilde{x}_{n+1}}{2}}{2\Delta\tilde{x}_{n+1}} \\
 &= \frac{(V_n)^2}{2(\Delta\tilde{x}_{n+1})^2} + \frac{3}{8} \stackrel{!}{\leq} 1 \\
 \Rightarrow \frac{2V_n}{\sqrt{5}} &\leq \Delta\tilde{x}_{n+1} \leq 2V_n
 \end{aligned}$$

$\bar{\eta}_n^{i,j} = -\frac{\Delta\tilde{x}_{n+1}}{2}$ is the minimum of the function $f(\bar{\eta}_n^{i,j}) = \frac{(\bar{\eta}_n^{i,j})^2}{2(\Delta\tilde{x}_{n+1})^2} + \frac{\bar{\eta}_n^{i,j}}{2\Delta\tilde{x}_{n+1}}$ which explains the first inequality. For p_d the same inequalities are obtained as for p_u . p_m never is greater than one and always above zero in case of

$$\begin{aligned}
 \bar{p}_m &\geq 1 - \frac{(V_n)^2}{(\Delta\tilde{x}_{n+1})^2} - \frac{\left(\frac{\Delta\tilde{x}_{n+1}}{2}\right)^2}{(\Delta\tilde{x}_{n+1})^2} \\
 &= \frac{3}{4} - \frac{(V_n)^2}{(\Delta\tilde{x}_{n+1})^2} \stackrel{!}{\geq} 0 \\
 \Rightarrow \Delta\tilde{x}_{n+1} &\geq \frac{2V_n}{\sqrt{3}}
 \end{aligned}$$

On the whole the conditions result in

$$\begin{aligned}
 \frac{2V_n}{\sqrt{3}} &\leq \Delta\tilde{x}_{n+1} \leq 2V_n \\
 \Rightarrow \Delta\tilde{x}_{n+1} &= bV_n \text{ with } \frac{2}{\sqrt{3}} \leq b \leq 2
 \end{aligned} \tag{5.2.8}$$

b can freely be chosen in this range, the standard value is $\sqrt{3}$ for good convergence (see HULL & WHITE (1993) and HULL & WHITE (1994a)). By insertion of the step sizes the probabilities are simplified to

$$\begin{aligned}
 p_u &= \frac{1}{2b^2} + \frac{(\eta_n^{i,j})^2}{2b^2V_n^2} + \frac{\eta_n^{i,j}}{2bV_n} \\
 p_m &= 1 - \frac{1}{b^2} - \frac{(\eta_n^{i,j})^2}{b^2V_n^2} \\
 p_d &= \frac{1}{2b^2} + \frac{(\eta_n^{i,j})^2}{2b^2V_n^2} - \frac{\eta_n^{i,j}}{2bV_n}
 \end{aligned} \tag{5.2.9}$$

where

$$\eta_n^{i,j} = \mu_n^i \Delta t_n + \tilde{x}_n^i - \tilde{x}_{n+1}^j.$$

Like the NB tree, the BM tree is bounded from above and below. This is natural due to

choosing the middle node as the node nearest to the expectation value. When the tree continuously moves up or down, the drift gets increasingly negative or positive, respectively, and at some point stops further movements of the tree in the same direction. It is the other way around for the lower bound.

5.2.4 Differences between the Tree Methods

At the first glance there are several differences between the NB and BM trees presented in the subsections before. The differences listed below will be discussed thereafter:

- The size of time steps is fixed or free for different intervals, respectively. In each time step the same probabilities are used or new computations of probabilities are necessary, respectively.
- The start values $x(0)$ and $\tilde{x}(0)$ are included / not included in the tree, respectively.
- The tree hits the zero point exactly to prevent numerical difficulties / is shifted by ϵ to show the fact that the rate is non-zero.
- Subsequent nodes are found in a different manner.

Fixed time steps in the NB tree differ from variable ones in the BM tree, but upon taking fixed time step sizes in the BM tree the same advantages are archived as in the NB tree. The probabilities for each node in the vertical direction need to be calculated only once. This is one more degree of freedom for the BM tree only.

In the NB tree the start value is exactly matched in the tree grid and not in the second tree. In the BM tree the parameter b in (5.2.8) is usually set to $\sqrt{3}$ for convergence purposes. But it can still be freely chosen in a determined range to match the start value in the tree. As in the NB tree, a sufficient number of time steps is necessary.

The NB tree includes the zero point to avoid computational problems close to zero, while the BM tree ensures its positivity by a small shift parameter ϵ greater than zero. Both tree building procedures are not affected by the choice of the handling of the zero point, i.e., both may include or not include the zero. In case the Feller condition is satisfied, zero is an unreachable point, a condition which is reflected by the small shift parameter. Nevertheless, the convergences of both trees are of the same order as shown in Fig. 5.2.4(a) for the parameter vector (0.0165 0.4 0.026 0.14). In case the Feller condition is broken, the handling of the zero is indeed becoming important. Omitting the zero point by applying a shift parameter can result in tree divergence, while the inclusion of the zero results in a proper convergence of the NB tree even in such parameter situations. Figure 5.2.4(b) illustrates the effect for the calibrated process of the interest rate. The NB tree and the modified BM tree (defined below) converge directly and more rapidly, and therefore both are not visible on the larger scales of the y-axis applied to show the divergence of the BM tree.

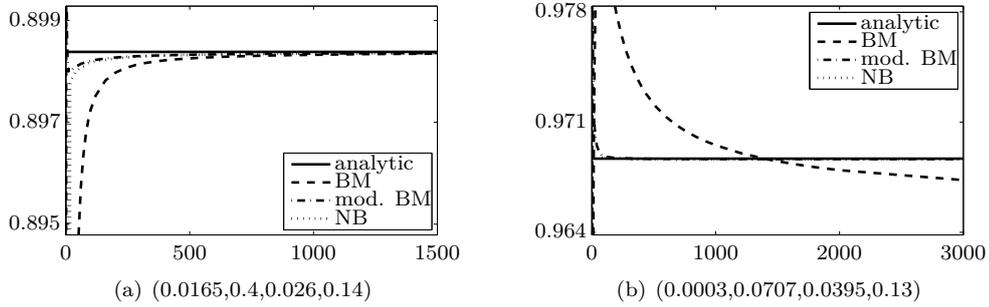


Figure 5.2.4: Convergence of different CIR trees to bond prices

The main difference between the latter two trees in their way to determine the following nodes. In the standard cases in the NB tree, they are found by expectation values using the parameter J in Eq. (5.2.5) in the standard case. The BM tree determines the next nodes by j (Eq. (5.2.7)) choosing the nodes closest to the expectation values in all cases. It is possible to switch from one tree scheme to the other if the time step sizes are equal ($V = V_n = V_{n-1}$):

$$\begin{aligned}
 \tilde{x}_{n+1}^{i+J} - \tilde{x}_n^i &= J \cdot \Delta \tilde{x} \\
 &= JbV \\
 &= (j - i)bV \\
 &= jbV_n - ibV_{n-1} \\
 &= \tilde{x}_{n+1}^j - \tilde{x}_n^i
 \end{aligned}$$

At first the NB tree is used, the final outcome is the same as in the BM tree by setting $J = j - i$. Be aware that J is the relative difference of two successive nodes contrary to i and j which are absolute numbers of successive nodes. Applying the equation

$$\begin{aligned}
 \eta_n^{i,j} &= \mu_n^i \Delta t_n + \tilde{x}_n^i - \tilde{x}_{n+1}^j \\
 &= \mu_n^i \Delta t_n - JbV_n
 \end{aligned}$$

to the probabilities in the BM tree in Eq. (5.2.9) results in the original probabilities of the NB tree in Eq. (5.2.4). The condition that the probabilities must be in the range from zero to one is the same therefore for both trees, the NB tree as well as the BM tree.

On the whole the main tree building procedures are not really different. The comparisons demonstrated that the BM approach is superior to the NB approach at all points listed below except the zero handling:

- The BM tree can handle different time steps to match all product dates exactly. In addition, for an option with a short time horizon on an underlying product with long dated maturity different time steps can be chosen for them. Very small time steps may be chosen in the option live, and large time steps during the product live time. These

choices enforce the convergence and can save a lot of memory space and computation time.

- The BM tree method is much more flexible, a property which is necessary in case of incorporation of the regime switching into the process.
- In the BM tree the next middle node is always the nearest to the expectation value. Due to Eq. (5.2.5) in case of b close to one the NB tree may use unnatural next middle nodes.
- The BM tree in all cases matches the first and second moments in all nodes - the NB tree only in case of $J \geq 0$.

Due to the convergence problems of the BM tree in the cases of a broken Feller condition the tree finally used is a modified BM tree. The modification is that the tree nodes start at zero and therefore - as in the NB tree - the building procedure is divided into nodes above zero and nodes at zero. The nodes above zero are handled by the BM, the nodes at zero by the NB tree procedure. In the following, only this modified tree is used, extended by add-ons described below. Also, all results in Chapter 6 are computed based on this modified BM tree.

5.3 Tree Building - Handling of Jumps, Regime-Switching and others

5.3.1 Tree with Exponential Jumps

This section captures some ideas of NAWALKHA, SOTO & BELIAEVA (2007), Chapter 6. But the exact handling here is slightly different with respect to the partition of the exponential space. Based on the CIR-EJ model

$$dx(t) = \kappa(\theta - x(t))dt + \sigma\sqrt{x(t)}dW(t) + J(\gamma)dN(t; \varsigma),$$

the tree handling of the jump distribution (exponential jumps) and the jump arrival rate (homogeneous Poisson process) is presented in the following. The diffusion part has been discussed in the previous section. At each time step the two parts will be treated successively. The maximum of jumps per time step of the tree is assumed to be one. This is not a constraint since the step size converges to zero ($\Delta t \rightarrow 0$). The probability of a jump in the interval $[t_n, t_{n+1}]$ is given by

$$dN(t; \varsigma) = \begin{cases} 0 & \text{with probability } \exp(-\varsigma \cdot \Delta t_n) \\ 1 & \text{with probability } 1 - \exp(-\varsigma \cdot \Delta t_n) \end{cases}.$$

To keep this analysis general an arbitrary node \tilde{x}_n^i is taken at arbitrary time t_n and state $i \cdot \Delta \tilde{x}$. The distribution of an exponential variable ranges from zero to infinity, its cumulative distribution function is given by

$$F(x) = 1 - \exp\left(-\frac{1}{\gamma} \cdot x\right).$$

To discretize the distribution the space above the node \tilde{x}_n^i is subdivided into M sections. There are at least two versions to subdivide the space of the exponential distribution:

- M equal spaced sections with different probabilities (Fig. 5.3.1 (a))
- M sections with different space sizes but the same probabilities (Fig. 5.3.1 (b))

In both subfigures the dashed lines represent the usual diffusion movements of the process. The solid lines display the M movements of the jump component. The curve on the right hand side of each graph indicates the probability distribution of the jumps.

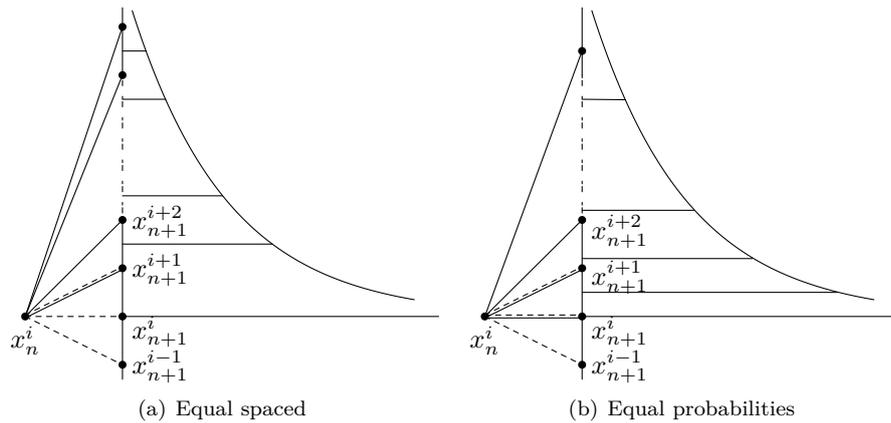


Figure 5.3.1: Exponential jump distribution and its approximation

All probabilities of one section are transferred to the node nearest to the middle of the specific section. In the following the version with equal probabilities but different sizes is applied. The sections are calculated by subdividing $F(x)$ into M intervals. The interval midpoints are the x -values that satisfy the condition

$$F(x) = \frac{2 \cdot k - 1}{2 \cdot M} \quad \text{for } k = 1, \dots, M.$$

Each of these x values has the probability

$$\frac{1 - \exp(-\zeta \cdot \Delta t_n)}{M}.$$

The jumps are not bounded from above as the exponential distribution ranges to infinity. In non-reasonable regions of interest rate and default intensity the probabilities become very

small. Therefore, a limit \tilde{x}_{\max} (resp. x_{\max}) is introduced specifying a maximum value of each tree, and all probabilities beyond this node are transferred to this maximum value.

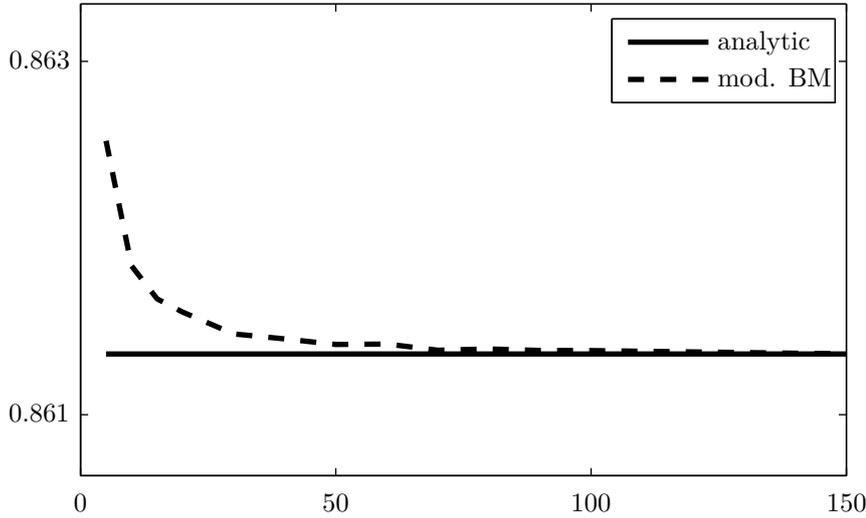


Figure 5.3.2: Convergence of CIR-EJ trees to bond price.

By including the jump component into the tree high state spaces are reached with positive probability. The usual diffusion tree does not reach these states as it is bounded from above. Therefore the treatment in the different trees is discussed. In case of the BM tree and of the modified BM tree, the respective middle node of the diffusion process is determined by the mean. Therefore, all probabilities are above zero, and the procedure does not need additional calculations. The situation is different for the NB tree. In its original form the J values determining the shift between the actual and the next tree node (not the jump component) would be highly negative. In this case the probabilities and the value J have to be recomputed as suggested in NAWALKHA, SOTO & BELIAEVA (2007). This is a procedural issue only and it is not relevant herein since the modified BM tree is applied.

The convergence of the tree for the CIR-EJ process (0.0165, 0.4, 0.026, 0.14, 0.05, 0.15) to its analytic value is shown in Fig. 5.3.2. The analytic value is matched using 150 intervals for the exponential jump component.

5.3.2 Regime Switching Tree

This section was inspired by two papers, BOLLEN (1998) and YUEN & YANG (2010). They provide the handling of regime switching in case of a log-normal process in the context of equity option pricing. BOLLEN (1998) developed the concept for the case of two states, YUEN & YANG (2010) extended this approach to multi-regime states.

The target is to use one tree only for all regimes, i.e., the same nodes. An important advantage of one single set of nodes is memory reduction and easy handling of regime switching. Additionally the whole tree recombines throughout all states. In each node the probabilities for the different regimes are saved separately. The main problem of this approach is to ensure positive moving probabilities in the whole tree and thus to allow different volatilities in different regimes. This problem is solved by allowing to skip nodes between the middle branch and the upper or lower branch, respectively. The detailed tree construction procedure is presented in this subsection.

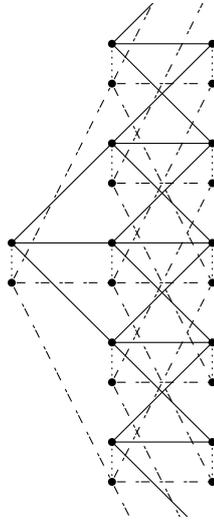


Figure 5.3.3: Regime switching tree

First, Fig. 5.3.3 gives a graphical introduction into the regime switching tree. In this three-dimensional representation the links between the two regime switching levels are shown by spotted lines which demonstrates that the different regimes use the same state space nodes. Additionally the spotted lines indicate that in each node the probabilities of being in a specific regime are changed by the corresponding transition probabilities. The upper tree (solid lines) represents the usual tree movements. However, in the lower tree (dashed lines) one node is skipped to ensure positive probabilities reflecting its higher volatility.

The inclusion of regime switching requires two steps. The first step is to include the trees for the regimes $2, \dots, Z$ into the tree of the first regime, with tree movements and probability determination, using the same state spaces. Moreover, the probability range from zero to one has to be ensured. In the second step the trees must be connected by the hidden Markov process and its transition probabilities.

As described in Chapter 4, $(S(t))_{t \in [0, T]}$ is a continuous-time Markov process. The handling of two states is shown first and then extended to the multi-regime switching case. The two-state space $\mathcal{S} := (s_1, s_2)$ represents the economic condition. Depending on the state, the parameter vector values of the CIR process are:

5.3. Tree Building - Handling of Jumps, Regime-Switching and others

- s_1 (good condition): $v = (x(0), \kappa, \theta, \sigma)$ and
- s_2 (bad condition): $\bar{v} = (x(0), \bar{\kappa}, \bar{\theta}, \bar{\sigma})$

The first parameter is the actual value and therefore the same in both states. The jump parameters are not included as they do not affect the handling of regime switching. They can be built into each regime separately. Without loss of generality it is assumed that $\bar{\sigma} \geq \sigma$. The trinomial tree for the good economy is constructed as explained in Subsection 5.2.3. The state is added to the tree node notation as the second component of the superscript. \tilde{x}_n^{is} denotes the node (n, i, s) in the tree at time t_n , space state $i \cdot \Delta \tilde{x}_n$ and economic state s .

The first point to consider is the auxiliary process \tilde{x} which ensures the recombination property of the tree as discussed in Section 5.1. If the transformation of x to \tilde{x} depends on the volatility, this is not possible for volatilities changing for different states. In these cases, the volatility-independent BM transformation must be used:

$$\tilde{x}(t) := \sqrt{x(t)}$$

As one single tree is used for both regimes there is one single node distance $\Delta \tilde{x}_n$ at one time step only. This would cause negative probabilities in the second tree for large differences between the values of σ and $\bar{\sigma}$. This is prevented by skipping nodes in the second tree which additionally enhances the property of convergence. The variable $1 \leq \nu \in \mathbb{N}$ is introduced for skipping branches. The value of ν depends on the relation between σ and $\bar{\sigma}$ and is determined at the end of this subsection. The second tree probabilities of an up, middle and down movement are calculated in the same way as described in Subsection 5.2.3 on the BM tree. Incorporating the skipping parameter Eq. 5.2.6 are reformulated to

$$\begin{aligned} \bar{p}_u &= \frac{(\bar{V}_n^i)^2}{2(\nu \Delta \tilde{x}_{n+1})^2} + \frac{(\bar{\eta}_n^{i,j})^2}{2(\nu \Delta \tilde{x}_{n+1})^2} + \frac{\bar{\eta}_n^{i,j}}{2\nu \Delta \tilde{x}_{n+1}} \\ \bar{p}_m &= 1 - \frac{(\bar{V}_n^j)^2}{(\nu \Delta \tilde{x}_{n+1})^2} - \frac{(\bar{\eta}_n^{i,j})^2}{(\nu \Delta \tilde{x}_{n+1})^2} \\ \bar{p}_d &= \frac{(\bar{V}_n^j)^2}{2(\nu \Delta \tilde{x}_{n+1})^2} + \frac{(\bar{\eta}_n^{i,j})^2}{2(\nu \Delta \tilde{x}_{n+1})^2} - \frac{\bar{\eta}_n^{i,j}}{2\nu \Delta \tilde{x}_{n+1}} \end{aligned} \quad (5.3.1)$$

\bar{V}^2 (variance) and $\bar{\eta}$ (mean minus next middle node value) are the equivalents for the second parameter vector \bar{v} . The value of j is again chosen to achieve a minimal distance between \tilde{x}_{n+1}^j and \bar{M}_n^i :

$$j := \text{round} \left(\frac{\bar{M}_n^i}{\Delta \tilde{x}_{n+1}} \right)$$

$|\eta_n^{i,j}| \leq \frac{\Delta \tilde{x}_{n+1}}{2}$ due to the definition of j independently of ν . The assumption $\sigma \leq \bar{\sigma}$ implies that $\bar{V}_n = \frac{\bar{\sigma}}{2} \sqrt{\Delta t_n} \geq \frac{\sigma}{2} \sqrt{\Delta t_n} = V_n = \frac{\Delta \tilde{x}_{n+1}}{b}$. In the next step the probabilities of (5.3.1) are shown to be between zero and one always. Calculations similar to those of Subsection

5.2.3 yield

$$\begin{aligned}
 \bar{p}_u &\geq \frac{(\bar{V}_n)^2}{2(\nu\Delta\tilde{x}_{n+1})^2} + \frac{\left(\frac{\Delta\tilde{x}_{n+1}}{2}\right)^2}{2(\nu\Delta\tilde{x}_{n+1})^2} - \frac{\frac{\Delta\tilde{x}_{n+1}}{2}}{2\nu\Delta\tilde{x}_{n+1}} \\
 &= \frac{(\bar{V}_n)^2}{2(\nu\Delta\tilde{x}_{n+1})^2} + \frac{1}{8\nu^2} - \frac{1}{4\nu} \stackrel{!}{\geq} 0 \\
 \bar{p}_u &\leq \frac{(\bar{V}_n)^2}{2(\nu\Delta\tilde{x}_{n+1})^2} + \frac{\left(\frac{\Delta\tilde{x}_{n+1}}{2}\right)^2}{2(\nu\Delta\tilde{x}_{n+1})^2} + \frac{\frac{\Delta\tilde{x}_{n+1}}{2}}{2\nu\Delta\tilde{x}_{n+1}} \\
 &= \frac{(\bar{V}_n)^2}{2(\nu\Delta\tilde{x}_{n+1})^2} + \frac{1}{8\nu^2} + \frac{1}{4\nu} \stackrel{!}{\leq} 1 \\
 &\Rightarrow \sqrt{\frac{(\bar{V}_n)^2}{2(\Delta\tilde{x}_{n+1})^2} + \frac{9}{64} + \frac{1}{8}} \leq \nu \leq \frac{2(\bar{V}_n)^2}{(\Delta\tilde{x}_{n+1})^2} + \frac{1}{2}
 \end{aligned}$$

where the first inequality is determined as in the standard tree. Again, one obtain the same inequalities for p_d as for p_u . p_m is never greater than 1 by definition. To ensure a value greater than zero the following condition must hold:

$$\begin{aligned}
 \bar{p}_m &\geq 1 - \frac{(\bar{V}_n^j)^2}{(\nu\Delta\tilde{x}_{n+1})^2} - \frac{\left(\frac{\Delta\tilde{x}_{n+1}}{2}\right)^2}{(\nu\Delta\tilde{x}_{n+1})^2} \\
 &= 1 - \frac{(\bar{V}_n^j)^2}{(\nu\Delta\tilde{x}_{n+1})^2} - \frac{1}{4\nu} \stackrel{!}{\geq} 0 \\
 &\Rightarrow \nu \geq \sqrt{\frac{(\bar{V}_n^j)^2}{(\Delta\tilde{x}_{n+1})^2} + \frac{1}{4}}
 \end{aligned}$$

The inequality $\sqrt{\frac{(\bar{V}_n)^2}{(\Delta\tilde{x}_{n+1})^2} + \frac{1}{4}} \geq \sqrt{\frac{(\bar{V}_n)^2}{2(\Delta\tilde{x}_{n+1})^2} + \frac{9}{64} + \frac{1}{8}}$ is satisfied for all $\frac{(\bar{V}_n)^2}{(\Delta\tilde{x}_{n+1})^2} = \frac{\bar{\sigma}^2}{\sigma^2 b^2} \geq \frac{1}{b^2} \geq \frac{1}{4}$. Therefore, the branch skipping variable must be in the range of

$$\sqrt{\frac{(\bar{V}_n)^2}{(\Delta\tilde{x}_{n+1})^2} + \frac{1}{4}} \leq \nu \leq \frac{2(\bar{V}_n)^2}{(\Delta\tilde{x}_{n+1})^2} + \frac{1}{2}. \quad (5.3.2)$$

Using the variance and the vertical step size $\Delta\tilde{x}_{n+1} = bV_n$, the condition for ν is

$$\sqrt{\frac{\bar{\sigma}^2}{\sigma^2 b^2} + \frac{1}{4}} \leq \nu \leq 2\frac{\bar{\sigma}^2}{\sigma^2 b^2} + \frac{1}{2}. \quad (5.3.3)$$

The fraction $\frac{\bar{\sigma}^2}{\sigma^2 b^2}$ is always greater or equal to $\frac{1}{4}$. If $\frac{\bar{\sigma}^2}{\sigma^2 b^2} \geq \frac{3}{4}$, then the distance between the lower and the upper bound in Eq. (5.3.3) is greater or equal to one and an integer solution exists. If $\frac{\bar{\sigma}^2}{\sigma^2 b^2} < \frac{3}{4}$, then is $\nu = 1$ the solution in the given range. Again, b is set to the standard value of $\sqrt{3}$. Finally, ν is chosen as the nearest integer to $\frac{\bar{\sigma}}{\sigma}$ which satisfies Eq.

(5.3.3).

A different handling is necessary when the next middle node is one of the ν lowest tree nodes, because the downward movement would result in nodes below zero. Therefore a downward skipping variable $\bar{\nu} \leq \nu$ is introduced. At each node, $\bar{\nu}$ is the largest number below ν such that the downward movement leads to the first tree node. The movement probabilities are calculated by the same approach as in the standard case and are given by

$$\begin{aligned} \bar{p}_u &= \frac{(\bar{V}_n^i)^2}{(\nu^2 + \nu\bar{\nu})(\Delta\tilde{x}_{n+1})^2} + \frac{(\bar{\eta}_n^{i,j})^2}{(\nu^2 + \nu\bar{\nu})(\Delta\tilde{x}_{n+1})^2} + \frac{\bar{\nu} \cdot \bar{\eta}_n^{i,j}}{(\nu^2 + \nu\bar{\nu})\Delta\tilde{x}_{n+1}} \\ \bar{p}_m &= 1 - \frac{(\bar{\nu} + \nu) \left((\bar{V}_n^j)^2 + (\bar{\eta}_n^{i,j})^2 \right) + (\bar{\nu}^2 - \nu^2) \bar{\eta}_n^{i,j} \Delta\tilde{x}_{n+1}}{(\nu^2\bar{\nu} + \nu\bar{\nu}^2)(\Delta\tilde{x}_{n+1})^2} \\ \bar{p}_d &= \frac{(\bar{V}_n^i)^2}{(\bar{\nu}^2 + \nu\bar{\nu})(\Delta\tilde{x}_{n+1})^2} + \frac{(\bar{\eta}_n^{i,j})^2}{(\bar{\nu}^2 + \nu\bar{\nu})(\Delta\tilde{x}_{n+1})^2} - \frac{\nu \cdot \bar{\eta}_n^{i,j}}{(\bar{\nu}^2 + \nu\bar{\nu})\Delta\tilde{x}_{n+1}} \end{aligned} \quad (5.3.4)$$

These probabilities are always in the range between zero and one as long as $\nu \geq \frac{2(\bar{V}_n^j)^2}{(\Delta\tilde{x}_{n+1})^2} - \frac{1}{2}$ even for the most extreme case $\bar{\nu} = 1$. This condition is stronger than the one stated in Eq. (5.3.3). If the probabilities of Eq. (5.3.4) are not in the range between zero and one or if the mean of the actual node is the closest to the zero node, a second auxiliary procedure must be applied. In these cases, the binomial tree is chosen in which the downward movement is set to zero as in the second case of the NB tree in Subsection 5.2.2. The full binomial tree is described in NAWALKHA & BELIAEVA (2007). The upward movement is chosen to be higher

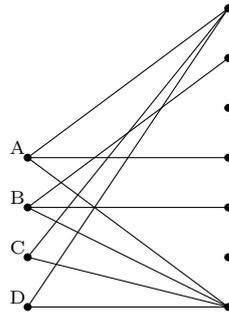


Figure 5.3.4: Movements on lowest nodes in RS tree

than the drift movement of x transformed to \tilde{x} ($x = \tilde{x}^2$). Therefore, the inequality

$$\begin{aligned} \tilde{x}_u &= (J + 1) \nu \Delta\tilde{x}_{n+1} \geq \sqrt{\tilde{x}_n^2 + \kappa(\theta - \tilde{x}_n^2) \Delta t_n} \\ \tilde{x}_d &= 0 \end{aligned}$$

has to be fulfilled for $J \geq 1$. Finally, the probabilities are calculated as

$$\begin{aligned} p_u &= \frac{\kappa(\theta - x_n) \Delta t_n + x_n}{x_u} \\ p_d &= 1 - p_u \end{aligned}$$

Figure 5.3.4 presents all possible movements of a RS CIR tree with $\nu = 3$. The tree movements at node A are the standard movements. At node B the first auxiliary procedure is applied, by which the downward movement becomes smaller than ν and the probabilities of Eq. (5.3.4) are in the range between zero and one. The binomial auxiliary procedure is applied at node C. At node D the standard procedure at zero is performed as described in the NB tree in Subsection 5.2.2.

In the following, this complete determination of the two regime case is extended to the multi-regime case. In the tree construction, σ is defined for Z different regimes to be

$$\sigma = \min_{i=1, \dots, Z} \sigma_i.$$

In each state a skipping parameter can be found by Eq. (5.3.3) as in the case of two regimes.

At this stage there are two or more parallel trees using the same space states with well-defined probabilities. But there is neither a probability transfer nor a connection between these parallel trees so far. The transition and its probabilities are incorporated by applying a constant generator matrix $A(t) \equiv A$ as described in Chapter 4. Due to the assumption that the matrix is constant over time, the defined transition probabilities $p_{ij}(\Delta t_k)$ depend on the length of the tree time steps Δt_k only, not on the start and end time. Therefore, all p_{ij} are given by

$$Q(t, t + \Delta t_k) = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1Z} \\ p_{21} & p_{22} & \dots & p_{2Z} \\ \vdots & \vdots & \ddots & \vdots \\ p_{Z1} & p_{Z2} & \dots & p_{ZZ} \end{pmatrix} = \exp(A \cdot \Delta t_k).$$

As in the normal case the tree construction starts with the actual state having a probability of one. The regime switching step is incorporated either before or after each diffusion step. The sequence is not important as $\Delta t \rightarrow 0$. This regime switching step is a movement of probabilities between the different regimes on each node. The size of each movement is determined by the transition probabilities. Bear in mind that the product value is different in different regimes on the same node.

5.3.3 The Shift Term in the Tree

The shift term φ is the parameter applied to match the term structure exactly. This deterministic function is analytic in the continuous CIR and CIR-EJ models. As shown before

5.3. Tree Building - Handling of Jumps, Regime-Switching and others

the CIR trees converge quickly, but nevertheless, this analytic φ would not exactly match the term structure in the tree - especially for a small number of time steps. For these two models and for the RS CIR tree without analytical solution, it is shown here how to find the exact shift parameter for each tree step.

Incorporation of the shift term requires that the tree \tilde{x} is built with all presented features and recalculated to the original process x at each node. In the tree the nodes x_n^{is} at time t_n for all i and s are shifted by the same amount φ_n . Let P_n denote the bond price $P(0, t_n)$ in case of the default free tree and the survival probability $H(0, t_n)$ in case of the default tree.

The idea is to compute P_n as

$$P_n = P_n^* \exp\left(-\sum_{k=0}^{n-1} \varphi_k \Delta t_k\right)$$

where P_n^* is the pseudo bond price (resp. survival probability) based on the discounted risk-neutral expectation using the state variable x as the pseudo short rate (default intensity). The values for φ at each time step in the tree can iteratively be obtained by

$$\sum_{k=0}^{n-1} \varphi_k \Delta t_k = \ln P_n^* - \ln P_n.$$

The remaining task to find P_n^* for each time t_n requires a more detailed and technical view. π_n^{is} is defined as the state price of node (n, i, s) , i.e.,

$$\pi_n^{is} := \mathbb{E} \left[\mathbf{1}_{\{x_n = x_n^{is}\}} \prod_{k=0}^{n-1} e^{-x_k \Delta t_k} \right].$$

π_n^{is} is the risk-neutral expectation that the discretized process x on the tree hits the node (n, i, s) discounted dependent on all intermediate values x_k , $k = 0, \dots, n-1$. Here all tree paths are included which meet x_n^{is} at time t_n . It is important to take the values of the original process x instead of \tilde{x} . It follows that $P_n^* = \sum_{s \in \mathcal{S}} \sum_{i=1}^I \pi_n^{is}$ in case of I space states at time t_n and a set of \mathcal{S} economic states.

The tree is now fitted by forward induction. Starting from the initial node it is shown how to fit the next time-level $n \rightarrow n+1$ to the given P_{n+1} .

Initialization $n = 0$:

Starting at $n = 0$ the state price is obtained as $\pi_0^i = 1$ for the start node x_0^i . P_1 needs to be matched with the shift φ_0 which is satisfied by

$$\varphi_0 = -x_0^i - \frac{1}{\Delta t_0} \ln P_1$$

Iteration $n \rightarrow n+1$:

Assume that the shift values $\varphi_0, \dots, \varphi_n$ are matched to P_1, \dots, P_{n+1} . The state prices π_k^{is} are known for all i and $k \leq n$. The new state prices for time $n + 1$ are calculated as

$$\pi_{n+1}^{ju} = \sum_{(i,s) \in \text{Pre}(j,u)} p_n^{(is)(ju)} \pi_n^{is} e^{-x_n^i \Delta t_n} \quad (5.3.5)$$

where $p_n^{(is)(ju)}$ is the transition probability from (n, i, s) to $(n + 1, j, u)$. The sum is taken over all predecessor nodes of $(n + 1, j, u)$. Instead this can also be done by using a loop over all nodes at time n and calculating the successor node state price amount for each node by multiplying $\pi_n^{is} e^{-x_n^i \Delta t_n}$ with its probabilities. The latter method is easier as it does not require to find the predecessor nodes.

The new φ_{n+1} is given by

$$P_{n+2} = e^{-\sum_{k=0}^n \varphi_k \Delta t_k} \sum_{s \in \mathcal{S}} \sum_j \pi_{n+1}^{ju} e^{-(x_{n+1}^j + \varphi_{n+1}) \Delta t_{n+1}}$$

and thus,

$$\varphi_{n+1} = \frac{\ln \left(\sum_{s \in \mathcal{S}} \sum_j \pi_{n+1}^{ju} e^{-x_{n+1}^j \Delta t_{n+1}} \right) - \ln P_{n+2} - \sum_{k=0}^n \varphi_k \Delta t_k}{\Delta t_{n+1}}$$

These formulas already include the handling of the jump component and of the regime switching. Both are indirectly included by the predecessor nodes in Eq. (5.3.5). They increase the number of the predecessor nodes which is more complicated. Therefore, again it is easier to use the loop over the successor nodes to determine the state price amount.

5.3.4 Adding a Default Branch to the Tree

The tree of the default intensity λ has to incorporate the default event which is done by an extra default branch as shown in Fig. 5.3.5. Before moving forward with the diffusion process the default branch is reached with the probability of default for the next time step given there was no previous default.

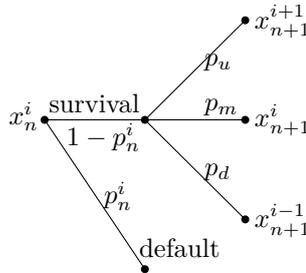


Figure 5.3.5: Incorporating default branch in the tree

The default probability p_n^i at node x_n^i is independent of the regime s and given by

$$1 - p_n^i = e^{-(x_n^i + \varphi_n)\Delta t_n} = e^{-\lambda_n^i \Delta t_n}$$

where x and φ are the components of the default intensity λ . The overall default probability at time t_n is obtained by the state price and the shift component as shown in the last subsection.

Given survival, the 'normal' tree continues with the probabilities multiplied by $(1 - p_n^i)$. As already mentioned the default probability in the reduced-form model is calculated like 'discounting' with the default intensity.

5.3.5 Building the Tree - the Steps

The topic of this subsection is to bring more structure into the tree construction. It is a short summary of the tree building construction for a single tree comprising all features presented above:

1. Calculate the auxiliary process \tilde{x} of the given (RS) CIR(-EJ) process.
2. Find the b and ν values for the step length and the node shift parameter for all states in case of regime-switching.
3. Build the tree for \tilde{x} with the lowest volatility starting at \tilde{x}_0 and recalculate the original value x for all tree nodes.
4. Calculate the successor nodes and the probabilities of diffusion and jumps for each node and all regimes.
5. Calculate the transition probabilities of regime-switching for each node.
6. Starting at \tilde{x}_0 to move forward in the tree. At each time step apply the following steps in the given order:
 - a) In case of default tree: Calculate the default and survival probabilities.
 - b) Calculate the probabilities to reach the tree nodes at the next time step after diffusion.
 - c) Incorporate the jump component which again changes the probabilities.
 - d) Apply the regime switching component which also changes the probabilities between the different states at each node.
 - e) Determine the shift parameter φ for the actual step based on the new state prices of the actual node.
7. Repeat step six until the final time point is reached.

This is the complete tree construction procedure. For tree valuation all subitems of step six are performed in reverse order. The valuation in each tree node is shown in Section 5.5.

5.4 Combining the Interest and Default Trees

The first combined tree was introduced by HULL & WHITE (1994b) who presented the merge of two interest rate trees for different currencies. The first combination of an interest and a default tree was done by SCHÖNBUCHER (2000b) and GARCIA, VAN GINDEREN & GARCIA (2003).

The two trees can only be merged if they have the same time step sizes Δt_k for all k . The vertical number of nodes and the vertical step sizes Δx_r^α , Δx_λ^β can be different. They depend on the volatility and the tree construction procedure described in the last sections. x_n^{ijs} denotes the node (n, i, j, s) at time t_n , space states $i \cdot \Delta x_r^\alpha$ and $j \cdot \Delta x_\lambda^\beta$ and economic state s .

The starting point of the combined tree is determined by the single starting points. Assuming a CIR tree without jump and regime switching, each node has 3^2 successor nodes - all combinations of the up, middle and down movements of each single tree. This is shown in Fig. 5.4.1 where the interest rate tree is vertical and the default intensity is horizontal. The default branch is the tenth branch of the combined tree (not shown in the figure). The probability for each of the nine diffusion movements is presented in the next subsections. They depend on the correlation between the interest rate and the default intensity.

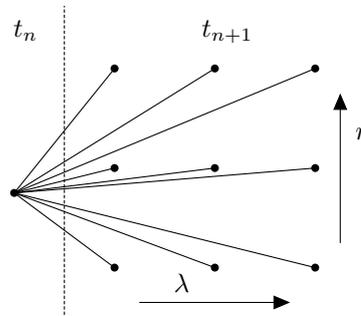


Figure 5.4.1: Combined tree

The jump components of the two trees are assumed to be independent of each other and to the diffusion processes and are incorporated as in the single trees. The regime switching component is also included as in the single tree. For each regime an interest and a default tree is built. After combining the two trees the transition probabilities are applied for each time step.

Altogether the size of the combined tree is

$$\text{'time steps'} \times \text{'rstate size'} \times \text{'lambda state size'} \times \text{'economic states'}$$

5.4.1 The Independent Case

Here, the combined tree building is described for the case of no correlation. The implementation starts by using both trees created as stated in Subsection 5.3.5. The combination of the two trees is achieved by simply multiplying all moving probabilities of the first tree by all of the second one as shown in Table 5.4.1. If there is no correlation, nothing else has to be considered, and the combined tree matches the marginal probabilities.

		default intensity		
		down	middle	up
interest rate	up	$p_u^r \cdot p_d^\lambda$	$p_u^r \cdot p_m^\lambda$	$p_u^r \cdot p_u^\lambda$
	middle	$p_m^r \cdot p_d^\lambda$	$p_m^r \cdot p_m^\lambda$	$p_m^r \cdot p_u^\lambda$
	down	$p_d^r \cdot p_d^\lambda$	$p_d^r \cdot p_m^\lambda$	$p_d^r \cdot p_u^\lambda$
default probability: p				

Table 5.4.1: Probabilities of the combined trees without correlation

5.4.2 The Case with Correlation

The incorporation of a correlation between interest rate and default intensity is more complicated. The calibration procedure of the default intensity in Chapter 3 is performed under the assumption of independence. In case of correlation the calibration needs to be re-adjusted after combining the two trees to exactly reflect the market prices. However, the numerical calibration would be very difficult as there is no analytical solution. Since the correlation has not much impact on the prices of CDS and CDS options, this problem is not treated more precisely in this thesis. Nevertheless the effect and the resulting failure will be shown in the valuation in Chapter 6.

Different methods have been developed to incorporate the correlation into the trees by approximate the adjustment of the moving probabilities. One of them has been presented by HULL & WHITE (1994b) (HW) and is enlarged to the CIR model in Subsection 5.4.3. It uses an auxiliary parameter to incorporate the correlation into the combined tree probabilities. In Subsection 5.4.4 the newly developed model is described. This model uses copulas to incorporate the correlation.

5.4.3 The Hull & White (1994b) Approach

HW adjusts the moving probabilities at each node as shown in the Tables 5.4.2 and 5.4.3 depending on the sign of the correlation. The moving probabilities are affected independently of the underlying process. Therefore the method can be applied to the CIR process.

		default intensity		
		down	middle	up
interest rate	up	$p_u^r \cdot p_d^\lambda - \varepsilon$	$p_u^r \cdot p_m^\lambda - 4\varepsilon$	$p_u^r \cdot p_u^\lambda + 5\varepsilon$
	middle	$p_m^r \cdot p_d^\lambda - 4\varepsilon$	$p_m^r \cdot p_m^\lambda + 8\varepsilon$	$p_m^r \cdot p_u^\lambda - 4\varepsilon$
	down	$p_d^r \cdot p_d^\lambda + 5\varepsilon$	$p_d^r \cdot p_m^\lambda - 4\varepsilon$	$p_d^r \cdot p_u^\lambda - \varepsilon$

default probability: p

Table 5.4.2: Probabilities of the combined trees with positive correlation

		default intensity		
		down	middle	up
interest rate	up	$p_u^r \cdot p_d^\lambda + 5\varepsilon$	$p_u^r \cdot p_m^\lambda - 4\varepsilon$	$p_u^r \cdot p_u^\lambda - \varepsilon$
	middle	$p_m^r \cdot p_d^\lambda - 4\varepsilon$	$p_m^r \cdot p_m^\lambda + 8\varepsilon$	$p_m^r \cdot p_u^\lambda - 4\varepsilon$
	down	$p_d^r \cdot p_d^\lambda - \varepsilon$	$p_d^r \cdot p_m^\lambda - 4\varepsilon$	$p_d^r \cdot p_u^\lambda + 5\varepsilon$

default probability: p

Table 5.4.3: Probabilities of the combined trees with negative correlation

The adjustment of the moving probabilities depends on the auxiliary parameter ε . The absolute sum over all adjustments is 36ε giving the following correlation:

$$\varepsilon := \begin{cases} \frac{1}{36}\rho & \text{for } \rho > 0 \\ -\frac{1}{36}\rho & \text{for } \rho < 0 \end{cases} \quad (5.4.1)$$

This correlation does not affect the marginal probabilities of the single trees as the ε of each column and row sum up to zero. The parameter values are chosen based on the fact that in the limit $\Delta t_k \rightarrow 0$ the probabilities for an up, middle and down movement approach $1/6$, $2/3$ and $1/6$, respectively.³ The values of the matrix are chosen to obtain the correct values in case of correlations -1 , 0 and 1 . In between the correlation is interpolated. A more detailed discussion can be found in HULL & WHITE (1994b).

The adjustment is done by adding to or subtracting from each probability a specific portion. Therefore, the resulting probabilities could become negative in cases of high correlation and at nodes in the edge of the tree where further movements have very low probabilities. To prevent negative probabilities ε is corrected at each node to the highest (smallest) value less (greater) or equal to Eq. (5.4.1) which does not cause to negative probabilities for positive (negative) correlation. Nodes with low probabilities are affected only which have no big impact on the resulting tree correlation. In the limit ($\Delta t_k \rightarrow 0$) the probabilities converge to the values mentioned above, i.e., there is no further adjustment required.

In case of a zero value in the CIR tree or in some cases near zero in the RS CIR tree, there

³This is true for the HW tree for the HW process, but also for the BM tree for the CIR process. The NB tree for the CIR process tend to $1/3$, $1/3$ and $1/3$.

		default intensity	
		down	up
interest rate	up	$p_u^r \cdot p_d^\lambda - \tilde{\varepsilon}$	$p_u^r \cdot p_u^\lambda + \tilde{\varepsilon}$
	down	$p_d^r \cdot p_d^\lambda + \tilde{\varepsilon}$	$p_d^r \cdot p_u^\lambda - \tilde{\varepsilon}$
		default probability: p	

Table 5.4.4: Probabilities of the combined trees in case of two branches

are only two branches instead of three. Therefore the procedure needs to be enlarged to such cases. The probabilities are adjusted accordingly to Table 5.4.4 where $\tilde{\varepsilon} = \frac{1}{4}\rho$. Again the probabilities are restricted by the maximum amount of the table which does not result in a negative probability. These adjustments are applied on nodes where one direction has three branches and the other direction has two branches, too.

5.4.4 The Copula Approach

The copula is a further natural choice to insert a correlation structure which does not effect the marginal distributions. The following is the mathematical definition of the copula (of BLUHM, OVERBECK & WAGNER (2008)):

Definition 5.4.1. (*Copula*) A copula (function) is a multivariate distribution (function) such that its marginal distributions are standard uniform.

In analytic terms this means, \mathcal{C} is a d -dimensional copula if

- $\mathcal{C}(u_1, \dots, u_{i-1}, 0, u_{i+1}, \dots, u_d) = 0$
- $\mathcal{C}(1, \dots, 1, u_i, 1, \dots, 1) = u_i$
- \mathcal{C} is d -increasing

Here, $d = 2$ is taken for the correlation of a two-dimensional tree. The Gaussian copula $\mathcal{C}_\rho^{\text{Ga}}$ with correlation ρ is applied which is defined as

$$\mathcal{C}_\rho^{\text{Ga}}(u_1, u_2) := \Phi_\rho(\Phi^{-1}(u_1), \Phi^{-1}(u_2)) \quad \text{for } u_i \in [0, 1], i = 1, 2.$$

Auxiliary variables \tilde{p} are introduced to achieve a distribution in the interval $[0, 1]$ and applicability of the copula. They are defined as

$$\begin{aligned} \tilde{p}_d &:= p_d \\ \tilde{p}_m &:= p_d + p_m \\ \tilde{p}_u &:= p_d + p_m + p_u = 1 \end{aligned}$$

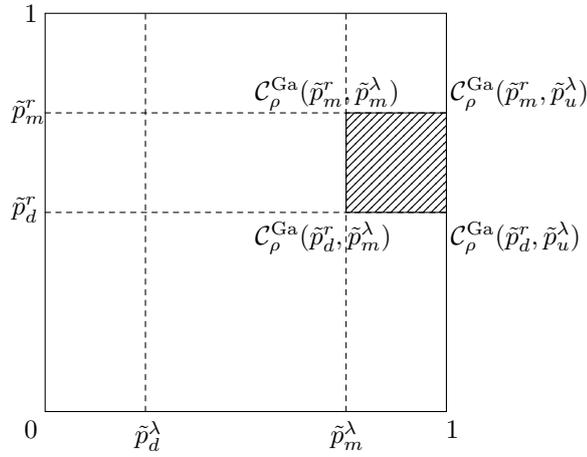


Figure 5.4.2: Probability calculation in the copula model

for both, the interest rate r and the default intensity λ . Now, the probabilities for the joint movements can be extracted based on Fig. 5.4.2 where the intervals are divided by the auxiliary variables.

In Fig. 5.4.2 an exemplary probability calculation is shown for a middle interest rate movement and an up movement of the default intensity. In the figure this probability is indicated by the striped area. The copula entries provide the probabilities for having interest rate and default intensity below this value. Therefore, the searched probability is given by the summation

$$p_{mu} = \mathcal{C}_\rho^{\text{Ga}}(\tilde{p}_m^r, \tilde{p}_u^\lambda) - \mathcal{C}_\rho^{\text{Ga}}(\tilde{p}_m^r, \tilde{p}_m^\lambda) - \mathcal{C}_\rho^{\text{Ga}}(\tilde{p}_d^r, \tilde{p}_u^\lambda) + \mathcal{C}_\rho^{\text{Ga}}(\tilde{p}_d^r, \tilde{p}_m^\lambda).$$

Simplifying the term by using $\tilde{p}_u^\lambda = 1$ and the copula property yields

$$p_{mu} = \tilde{p}_m^r - \mathcal{C}_\rho^{\text{Ga}}(\tilde{p}_m^r, \tilde{p}_m^\lambda) - \tilde{p}_d^r + \mathcal{C}_\rho^{\text{Ga}}(\tilde{p}_d^r, \tilde{p}_m^\lambda).$$

The results for all movements are presented in Table 5.4.5. The copula approach results in the correct values for the correlations -1, 0 and 1 in the limit which is discussed in the preceding subsection.

There is no need for a special handling of the two branches case as the same equations can be applied.

The use of the copula function to incorporate the correlation into the combined tree provides advantages and disadvantages. One advantage is that the best possible correlation can be applied in all nodes without changing any marginal distributions. On the other side the computational effort is much higher as one needs to calculate these probabilities for each node in the tree. This is applicable in case of a limited number of different time steps in the tree, only, as the probabilities and the copulas have to be calculated once only per time

step size.

		default intensity		
		down	middle	up
interest rate	up	$\tilde{p}_d^\lambda - \mathcal{C}_\rho^{\text{Ga}}(\tilde{p}_m^r, \tilde{p}_d^\lambda)$	$\tilde{p}_m^\lambda - \tilde{p}_d^\lambda - \mathcal{C}_\rho^{\text{Ga}}(\tilde{p}_m^r, \tilde{p}_m^\lambda) + \mathcal{C}_\rho^{\text{Ga}}(\tilde{p}_m^r, \tilde{p}_d^\lambda)$	$1 - \tilde{p}_m^\lambda - \tilde{p}_m^r + \mathcal{C}_\rho^{\text{Ga}}(\tilde{p}_m^r, \tilde{p}_m^\lambda)$
	middle	$\mathcal{C}_\rho^{\text{Ga}}(\tilde{p}_m^r, \tilde{p}_d^\lambda) - \mathcal{C}_\rho^{\text{Ga}}(\tilde{p}_d^r, \tilde{p}_d^\lambda)$	$\mathcal{C}_\rho^{\text{Ga}}(\tilde{p}_m^r, \tilde{p}_m^\lambda) - \mathcal{C}_\rho^{\text{Ga}}(\tilde{p}_m^r, \tilde{p}_d^\lambda) - \mathcal{C}_\rho^{\text{Ga}}(\tilde{p}_d^r, \tilde{p}_m^\lambda) + \mathcal{C}_\rho^{\text{Ga}}(\tilde{p}_d^r, \tilde{p}_d^\lambda)$	$\tilde{p}_m^r - \mathcal{C}_\rho^{\text{Ga}}(\tilde{p}_m^r, \tilde{p}_m^\lambda) - \tilde{p}_d^r + \mathcal{C}_\rho^{\text{Ga}}(\tilde{p}_d^r, \tilde{p}_m^\lambda)$
	down	$\mathcal{C}_\rho^{\text{Ga}}(\tilde{p}_d^r, \tilde{p}_d^\lambda)$	$\mathcal{C}_\rho^{\text{Ga}}(\tilde{p}_d^r, \tilde{p}_m^\lambda) - \mathcal{C}_\rho^{\text{Ga}}(\tilde{p}_d^r, \tilde{p}_d^\lambda)$	$\tilde{p}_d^r - \mathcal{C}_\rho^{\text{Ga}}(\tilde{p}_d^r, \tilde{p}_m^\lambda)$

default probability: p

Table 5.4.5: Probabilities of the combined trees with copula

The comparison of the two approaches shows that both meet the basis requirements such as retaining the marginal probabilities and matching the probabilities for correlations -1, 0 and 1 in the limit. The HW approach is computationally faster but has the drawback of adjustments to prevent negative probabilities. The reflection of the in between correlation is not examined and given by interpolation. The copula approach is computationally slower but does not need any adjustment. The distribution of the in between correlation is fully determined by the copula function. Additionally, there is a wide range of different copulas which could be used for this purpose. The effects of the different methods are shown in Chapter 6.

5.5 Tree Valuation

The tree valuation is a standard procedure, repeated here for completeness. It follows in some points the work of SCHÖNBUCHER (2000b) but is extended to cover the regime switching case.

The completely constructed tree can now be used to price claims which only depend on interest rate and default risk. As the target is to price defaultable products at least two cases have to be considered: The default and the survival case in each node. Moreover, an optionality feature such as early exercise or multiple options in case of a contingent credit line can be priced in the tree. Therefore, there exist the following four values or payoffs in each node:

- f_n^{ijs} the payoff of the derivative if a default happens in node (n, i, j, s)
- F_n^{ijs} the payoff of the derivative if no default happens in node (n, i, j, s)
- G_n^{ijs} the optionality payoff/value in node (n, i, j, s)

- V_n^{ijs} the value of the derivative in node (n, i, j, s)

Important is that these payoffs and the value at each node depend on the economic state variable s . The value is retrieved by standard backward induction:

Initialization $n = N$:

The value at the final time is determined by its payoff profile in each final node:

$$V_n^{ijs} = F_n^{ijs}$$

Iteration $n + 1 \rightarrow n$:

Assume that the values V_{n+1}^{klu} are given for all k, l and u at time $n + 1$. At first consider the values without early exercise. For each node (n, i, j, s) the value in case of survival is

$$\check{V}_n^{ijs} = e^{-r_n^i \Delta t_n} \sum_{(k,l,u) \in \text{Succ}(i,j,s)} p_n^{(ijs)(klu)} V_{n+1}^{klu},$$

where $\text{Succ}(i, j, s)$ are the successor nodes of (i, j, s) except default and $p_n^{(ijs)(klu)}$ are the transition probabilities from (n, i, j, s) to $(n + 1, k, l, u)$. Incorporation of the default branch leads to the value

$$\dot{V}_n^{ijs} = e^{-\lambda_n^j \Delta t_n} \check{V}_n^{ijs} + \left(1 - e^{-\lambda_n^j \Delta t_n}\right) f_n^{ijs} + F_n^{ijs}.$$

The optionality is incorporated by taking the maximum (resp. minimum depending on the side of the optionality) of the continuation value without option and the value in case of drawing the option at node (n, i, j, s) :

$$V_n^{ijs} = \max(\dot{V}_n^{ijs}, G_n^{ijs})$$

The value G_n^{ijs} depends on the optionality of the product. The early exercise feature is incorporated in many corporate loans. In this case the optionality payoff could be a (negative) penalty fee for early exercising or simply zero. In case of more optionalities as in contingent credit lines the optionality payoff also comprises the continuation value and only reflects the drawing/undrawing of the actual fee period.

6 Results

In this chapter the previously discussed models and numerical techniques are applied. The programming language MATLAB® is used to obtain these results. At first Section 6.1 shows the effects to a CDS call option. Then, in Section 6.2 loans are priced with drawdown and prepayment options. Additionally the optionality to draw or pay back in each period is shown by pricing contingent credit lines.

For all calculations the 6th July 2012 is the valuation date $t = 0$. The model parameters calibrated to Allianz market data of Section 3.4 are applied. Additionally uncalibrated meaningful parameters are applied for the second and third state in the regime switching models. All parameters are summarized in Table 6.0.1.

model	variable	$x(0)$	κ	θ	σ	ς	γ
CIR++	α	0.0003	0.0707	0.0395	0.13	-	-
2-RS CIR++	α	0.0003	0.026	0.06	0.05	-	-
3-RS CIR++	α	0.0003	0.01	0.02	0.2	-	-
CIR++	β	0.0055	0.0851	0.0965	0.446	-	-
2-RS CIR++	β	0.0055	0.0851	0.06	0.246	-	-
3-RS CIR++	β	0.0055	0.0851	0.15	0.9	-	-
CIR-EJ++	β	0.0038	0.0307	0.0312	0.4078	0.0621	0.1445
2-RS CIR-EJ++	β	0.0038	0.0307	0.0312	0.4078	0.03	0.05
3-RS CIR-EJ++	β	0.0038	0.0307	0.0312	0.4078	0.09	0.2

Table 6.0.1: Parameters in the calibrated model

The regime switching model is presented for two and for three regimes. For the two and the three regimes one parameter vector is given in Table 6.0.1 only. The reason is that the one and two regime parameter vectors are taken additionally for the two and three regime models. That means that the calibrated parameter vectors of the single regime is applied in all models and is used as start regime.

The second and third regime parameter vectors are not calibrated to a specific market situation. The second regime reflects a better economy state than the first one. The third regime represents a very bad economy state with high volatilities and high default intensities.

These states are chosen accordingly to the statistical results in Chapter 4. The applied generator matrices of the transition probabilities are given by

$$A = \begin{pmatrix} -0.014 & 0.014 \\ 0.021 & -0.021 \end{pmatrix}, \quad A = \begin{pmatrix} -0.024 & 0.014 & 0.01 \\ 0.015 & -0.0151 & 0.0001 \\ 0.011 & 0.0001 & -0.0111 \end{pmatrix}$$

in the two and three regime models, respectively. The parameter vectors in the different states does not satisfy the Feller condition. The condition that the model forward curve is below the market forward curve is satisfied for the applied transition probabilities.

Both products are additionally priced on parameter vectors for which the Feller condition in the current regime is satisfied. It is done by adjusting the parameters in the different models as stated in Table 6.0.2.

model	variable	$x(0)$	κ	θ	σ	ς	γ
CIR++	α	0.0003	0.0707	0.032	0.065	-	-
2-RS CIR++	α	0.0003	0.0707	0.05	0.01	-	-
3-RS CIR++	α	0.0003	0.707	0.02	0.12	-	-
CIR++	β	0.0055	0.2	0.035	0.11	-	-
2-RS CIR++	β	0.0055	0.2	0.02	0.05	-	-
3-RS CIR++	β	0.0055	0.2	0.07	0.3	-	-
CIR-EJ++	β	0.0038	0.0307	0.03	0.04	0.02	0.1445
2-RS CIR-EJ++	β	0.0038	0.0307	0.03	0.04	0.01	0.05
3-RS CIR-EJ++	β	0.0038	0.0307	0.03	0.04	0.06	0.25

Table 6.0.2: Parameters in case of valid Feller condition

6.1 CDS Call Option Results

As already announced all different models are applied to the CDS call option. The maturity date of the option is the 20th December 2012, the final CDS date is the 20th June 2017. The strike is chosen to be at the money. The day count convention of the CDS call option is act/360.

6.1.1 Results from Applying the Calibrated Models

The number of steps per year can be chosen freely. For convergence purposes it is useful to use relatively more steps in the period until option maturity than in the period of the underlying CDS. In the tables below 3,000/5,000 indicates that 3,000 time steps are applied up to option maturity and 5,000 time steps during the CDS lifetime. Recalculating this

number per day means that 18 steps per day are used until option maturity and only three steps per day for the CDS lifetime.

(in bps)	mod. BM tree			analytical price
time steps	1,000/1,000	2,000/3,000	3,000/5,000	-
deterministic IR	87.72	87.32	87.23	87.19
correlation -1	88.74	88.33	88.23	-
correlation 0	87.72	87.32	87.23	-
correlation 1	82.86	82.38	82.25	-

Table 6.1.1: CDS option price in calibrated model convergence

In Table 6.1.1 the convergence to the analytical CDS option price is shown for different time step pairs. Having 1,000/1,000 the relative difference to the analytical price is already below 1%. The highest time step pair 3,000/5,000 has a relative difference of 0.1% which is quite close. Additionally, it can be seen that the same price is obtained for deterministic interest rates and stochastic interest rates with zero correlation. This result is in line with the remarks in BRIGO & MERCURIO (2006). A positive correlation between default intensity and interest rates leads to a lower CDS call option price. A correlation of one lowers the price by around 6%. In contrary negative correlation results in higher prices whereby the effect is weaker. Upon application of correlation minus one the price increases by 1% only.

(in bps)	HW correlation	Gaussian copula	t_3 copula	t_1 copula
correlation -1	88.71	88.74	88.74	88.74
correlation -0.75	88.65	88.55	88.56	88.56
correlation -0.5	88.48	88.36	88.35	88.36
correlation -0.25	88.19	88.09	88.08	88.09
correlation 0	87.72	87.72	87.72	87.72
correlation 0.25	87.02	87.22	87.24	87.22
correlation 0.5	86.06	86.58	86.59	86.57
correlation 0.75	84.73	85.68	85.61	85.68
correlation 1	83.25	82.86	82.86	82.86

Table 6.1.2: CDS option prices in calibrated model under correlation

In Table 6.1.2 the results of different correlation approaches are presented. The HW technique described in Subsection 5.4.3 is compared to different copula approaches discussed in Subsection 5.4.4. For correlation one and minus one, the copula application can incorporate more correlation than the HW application. This due to the fact that the HW application is not fully applicable in edge regions with low probabilities. The behavior in between corre-

lation one and zero is vice versa. The prices decreases faster in the HW applications than in the copula applications. The influence of different copulas, Gaussian and t distributed, is not visible. In all cases the time steps 1,000/1,000 are applied.

Table 6.1.3 shows the results of applying the calibrated CIR-EJ process for the default intensity. In the no jumps column the jump component is set to zero to determine its effect. Again, upon assuming deterministic interest rates the prices are very close to the analytical prices. Time steps 3,000/5,000 are used and the number of jumps per tree node is set to 100. The correlation effect is weaker in case of jumps. The reason is that the correlation is applied to the diffusion process only. The jumps are independent of the diffusion movements.

(in bps)	no jumps	jumps
analytical price	53.44	90.27
deterministic IR	53.60	90.43
correlation -1	54.03	91.15
correlation 0	53.60	90.43
correlation 1	51.04	87.57

Table 6.1.3: CDS option prices in calibrated model under jumps

Tables 6.1.1 and 6.1.3 demonstrate that the calibrated CIR and CIR-EJ models do not match the market prices exactly since the analytical CIR and CIR-EJ prices are not exactly the same as the last CDS option price used for calibration.

The effect of regime switching is shown in the two figures below. The time steps are 1,000/1,000 and the number of jumps per node (if required) is 15. Figures 6.1.1 and 6.1.2 show the prices of the CDS option obtained by the RS CIR and RS CIR-EJ model for the default intensity, respectively. In both cases, the interest rate is modeled by the same RS CIR model and the correlation between default intensity and interest rate is set to zero. All parameter vectors and generator matrices are given at the beginning of this chapter. The prices are shown in relation to a factor for the generator matrix. The specified generator matrix is used per year in case of a factor equal to one. The statistics in Chapter 4 are calculated on a daily basis. Therefore the factor goes up to 250 business days which means that the generator matrix is rescaled to a daily basis. A factor of 250 results in very high transition probabilities which are not in line with economic cycles and with durations of several years at least. But its application is useful to show some effects of regime switching and its impact on CDS option prices.

In case of the RS CIR model the second regime is reflected by using a lower mean-reversion level and a lower volatility parameter in the default process and a higher mean-reversion level and lower volatility parameter in the interest rate process. The parameters of the third state are taken reversely as following. The mean-reversion level and the volatility parameter are higher than in the second state for the default intensity. In case of interest

the mean-reversion level is lower and the volatility is higher.

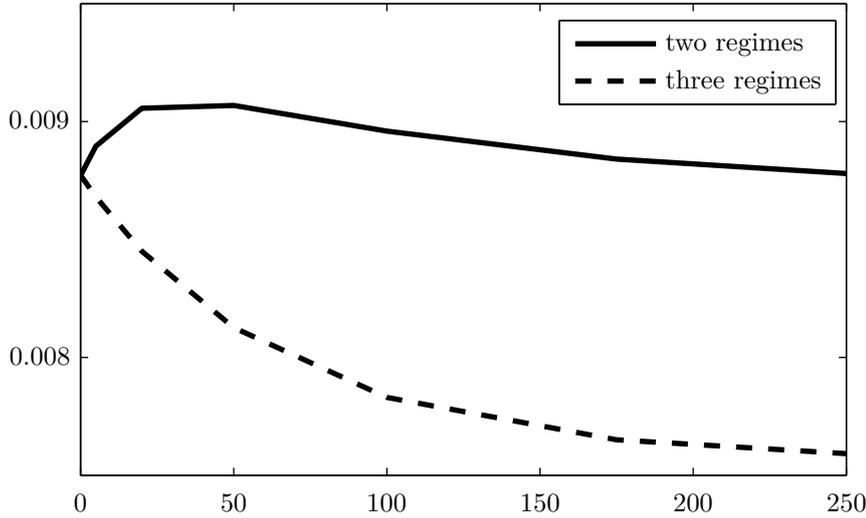


Figure 6.1.1: CDS option prices under the RS CIR model

In most situations, these value changes in the different regimes would result in a CDS option price which is monotonically decreasing in the two regime case by the factor for the generator matrix as the less volatile state is more likely. But upon applying the presented parameters the effect is different as shown in Fig. 6.1.1, and it is comparable to the one in the situation discussed in Subsection 3.4.3.

The density distribution of the default intensity is extreme, since the probability is very high not only for very high default intensities, but also for default intensities at zero. Increasing the volatility parameter strengthens this effect even more which increases the density at zero and at very high values. Altogether this results in less density of the default intensity where the CDS option is exercised. In the non regime switching case the sensitivity behavior of the volatility parameter is corresponding, i.e., a lower volatility parameter results in a higher CDS option price. But there is a factor for the generator matrix the effect goes into reverse and the CDS option prices are decreasing along the x-axis as expected in the two regime case.

The results of the three regimes case are in line with the discussed two states case. The values in state three are more extreme than in the good economy state. The higher volatility results in even higher densities at very high values and at zero resulting in a lower CDS option price.

For the default intensity the RS CIR-EJ model provides a lower jump arrival rate and a lower mean of the jump distribution in the good economy state, and much higher ones in

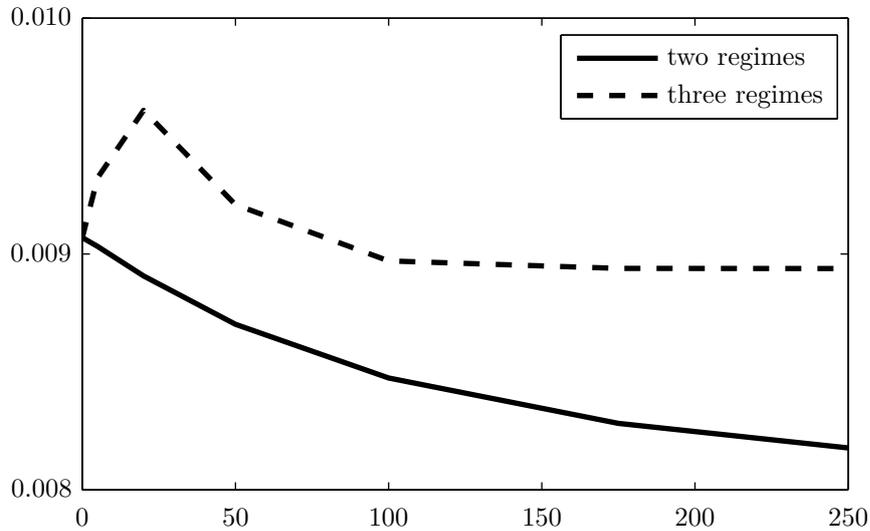


Figure 6.1.2: CDS option prices under the RS CIR-EJ model

the bad economy state. The same diffusion parameters are taken for all regimes to show the effect of regime switching to the jump parameters separately. In case of two states the CDS option price is monotonically decreasing by increasing the factor as the second regime has fewer and lower jumps. In the beginning of the three regimes case the effect is reversed as the third regime again is more extreme with higher jumps which occur more often resulting in a higher CDS option price. But from a peak onwards, the prices are decreasing. This effect can be explained by two reasons. The parameters are extreme resulting in a high probability for the process to stay at zero. For the tree nodes at zero, the CDS option price is zero on the exercise day in case of regime one or two. But the third regime assigns a value greater than zero to the zero node already. However this value decreases by a higher factor as the transition probabilities increases upon changes from regime three to regime one or two. Beyond one point this effect exceeds the effect that all CDS option prices in regimes one and two increases with increasing multiplier. The second reason is that the transition probabilities are above 35% for a factor higher than 50. The regimes are switched often and become more one mixture process. The parameters and especially the jump components are restricted from above by the market forward rate. Therefore, the CDS option price tends to the same value for the factors zero and 250.

6.1.2 Results in Case of Valid Feller Condition

In this subsection the effects are shown on parameter vectors in which the Feller condition is satisfied. The applied time steps are $1,000/1,000$. For the CIR default intensity model and in case of deterministic interest rates the analytical price reduces to 36.49 bps. The

explanation is that the volatility is much smaller compared to the calibrated model. Table 6.1.4 shows the results in the CIR++ model and additionally the effect of correlation for the parameter vector which satisfies the Feller condition. The price for deterministic interest rates is the same as the uncorrelated price. The effects are proportional pretty much the

(in bps)	HW correlation	Gaussian copula
correlation -1	36.94	36.96
correlation -0.5	36.66	36.56
correlation 0	36.21	36.21
correlation 0.5	35.60	35.76
correlation 1	34.88	34.83

Table 6.1.4: CDS option prices under correlation

same as in the calibrated model. The prices are lower for positive correlation and higher for negative correlation. The effects of the different correlation models - HW correlation and Gaussian copula - are as in the calibrated model. The HW correlation has a higher effect for correlation of ± 0.5 , but can not include the full correlation of ± 1 compared to the Gaussian copula model. The difference between the effects of t-copulas and Gaussian copula are very small and therefore not shown here again.

(in bps)	no jumps	jumps
analytical price	15.08	26.60
deterministic IR	14.76	26.39
correlation -1	15.09	26.62
correlation 0	14.76	26.39
correlation 1	14.24	25.99

Table 6.1.5: CDS option prices under jumps

Table 6.1.5 presents the effect of the jump component. The impact of the jump component to the CDS option price is proportional the same as in the calibrated model. By including jumps, the correlation has a lower impact again because of the independence of the jump component to the interest rate process.

In Fig. 6.1.3 the regime switching model is shown without jumps. The difference between the calibrated model and the model satisfying the Feller condition is the biggest as the effect of the two and three regime cases are inversely. Here, no extreme volatility parameter is chosen, and the corresponding lower volatility parameter leads to a lower price as expected. This is shown for the two regime case where the second regime has a lower volatility parameter. The small hump in the beginning can be explained by the volatility through the regime

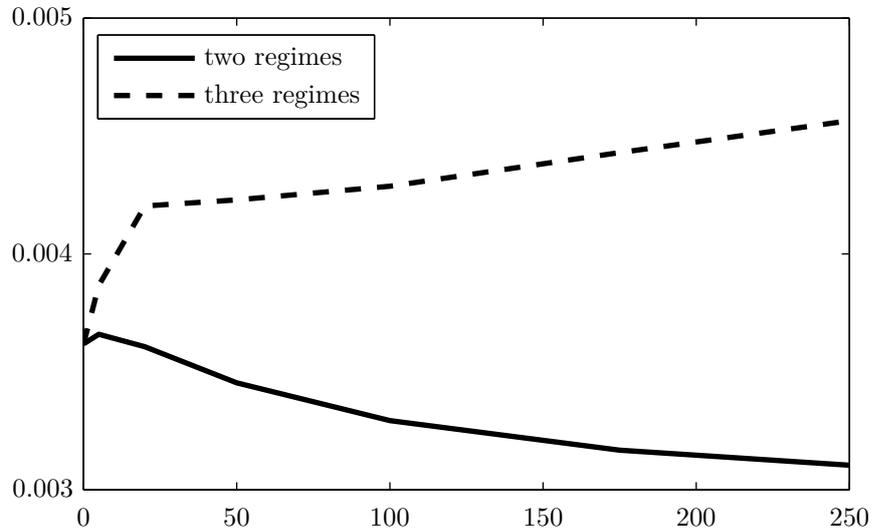


Figure 6.1.3: CDS option prices under the RS CIR model

switching. The three regime case is exactly vice versa. The additional volatile regime leads to increased prices by an increased multiplicative factor. The slope is much higher in the beginning than for factors higher than 20. For higher factors the process is more like a mixture process and the slope becomes smaller.

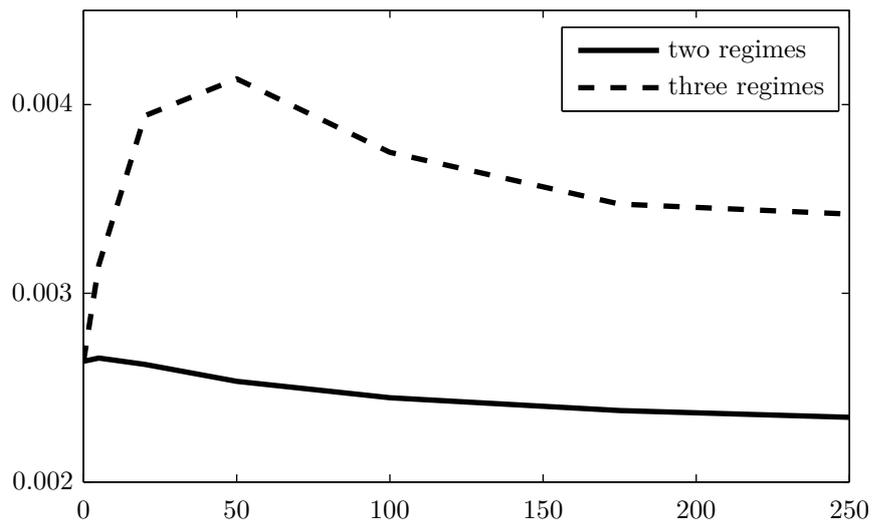


Figure 6.1.4: CDS option prices under the RS CIR-EJ model

The effects in case of the RS CIR-EJ++ model are presented in Fig. 6.1.4. The shapes are

very similar to the calibrated model parameters in Fig. 6.1.2. The only difference is that the prices of the CDS option for the multiplicative factor of zero and 250 do not coincide. The parameters of the one regime case is not as extreme on the forward curve compared to the calibrated model. Therefore the extreme scenario can be chosen more extreme compared to the calibrated model. Therefore the prices are higher for a multiplicative factor of 250.

6.2 Loan and Contingent Credit Line Results

In this section the effects of the different models on loans and contingent credit lines are discussed. The duration of the products is set to ten years in each case which means that the loans mature at the 6th July 2022. The loans are either fixed rate or floating rate loans and may include drawdown options (DD), prepayment options (PP) and contingent credit lines (CCL). It is always assumed that the options are exercised rationally and that the full amount of the loan is drawn or undrawn. The coupon is chosen to be at the money to achieve a loan of zero value without options and without correlation. The day count convention of the loan is act/act. The recovery rate is assumed to be 40%, the payments are quarterly. The resulting coupons of the fixed rate and floating rate loans are 330 bps and 157 bps, respectively.

In case of drawdown option or contingent credit line, a fee is paid for the undrawn amount. This fee is set to half of the coupon if not stated differently. For simplicity a prepayment fee is not included. Neither funding costs, liquidity costs nor commitment fees are considered in the loan pricing. All results are obtained by using 1,000 time steps in the tree. Since option and product times are equal one step size is applied until maturity, only.

6.2.1 Results from Applying the Calibrated Models

Table 6.2.1 presents the results for the CIR++ model including correlations of -1, 0 and 1. Both fixed rate and floating rate loans are priced. Options DD, PP, and CCL are marked by 'x' if included.

Negative or positive correlations result in higher or lower prices, respectively. The reason for the latter is that the loan has a positive value for high default intensities. The correlation of one implies high interest rates which results in lower discount factors and in general in a lower price. The negative correlation has the opposite effect on the prices. They can be more extreme in case of fixed rates because the interest rate risk is included. Therefore, the price of the options on fixed rate loans is higher than those on floating rate loans. In the presented results, the drawdown option prices are higher than the prepayment option prices. The reason is that the coupon is constant over the loan lifetime, but the market implied default intensities are generally increasing apart from some small decreases for time points later than five years. Therefore the drawdown option is more likely used compared to the

correlation			-1	0	1	-1	0	1
DD	PP	CCL	fixed rate loan			floating rate loan		
			51	0	-163	12	0	-34
x			240	197	55	137	126	94
	x		125	92	49	113	105	82
x	x		313	285	168	226	218	191
		x	264	226	89	216	205	174
	x	x	336	311	195	272	262	233

Table 6.2.1: Loan prices in the CIR++ model (in bps)

prepayment option. The price of a loan with embedded drawdown and prepayment options is higher than the price of a contingent credit line in Table 6.2.1. The contingent credit line only allows to choose between coupon and fee payment while the prepayment option cancels the contract completely.

The analytical price of the floating rate CCL is 188 bps. The numerical price with 1,000 time steps and deterministic interest rates is 205 bps. By increasing the number of time steps to 6,000 the price is 199 bps and converges to the analytical price. Taking into account the long option maturities the deviation is of the same size as in the CDS option case. Again, the calibrated parameter vector converges worse than the Feller valid parameter vector below.

A further question is to find the fair coupon in case of a CCL. Unfortunately, this question can only be answered by an iterative procedure in tree valuation. But the convergence is really quick in case of well adjusted coupons. The fair coupons of the fixed and floating loans would be 364 bps and 197 bps, respectively.

jumps			n	y	n	y
DD	PP	CCL	fixed rate loan		floating rate loan	
			0	0	0	0
x			126	225	25	194
	x		36	86	41	101
x	x		133	311	45	259
		x	134	245	26	233
	x	x	141	329	45	283

Table 6.2.2: Loan prices under jumps (in bps)

In Table 6.2.2 the effect of jumps is shown. Again the calibrated β in the CIR-EJ++ model is applied with and without the jump component indicated in the first row. The effect of the exponential jumps has a huge impact on the option prices. Apart from that similar effects as in the CIR++ model can be seen. The analytical prices of the floating rate CCL with

6.2. Loan and Contingent Credit Line Results

and without jumps are 219 bps and 25 bps, respectively. After applying 6,000 time steps the numerical values are 229 bps and 25 bps again showing the convergence to the analytical price.

#states			1	2	3	1	2	3
DD	PP	CCL	fixed rate loan			floating rate loan		
			0	0	0	0	0	0
x			197	189	191	126	110	106
	x		92	96	96	105	118	92
x	x		285	265	252	218	215	190
		x	226	211	213	205	179	165
	x	x	311	286	270	262	250	228

Table 6.2.3: Loan prices in the RS CIR++ model (in bps)

The prices in the RS CIR++ model are presented in Table 6.2.3. The first row shows the number of states used in the calculation. Again, the loans are fair in case of no option and zero correlation since all models match the term structure exactly. The prices of the different options for the different number of states are not obvious. Mostly, the highest prices are obtained for one regime, the lowest for three regimes. Normally, the third regime should have the highest prices because of the highest volatility. But as in the CDS options case, the extreme parameters are responsible for the curious effect. Again, the same calculation is done in Table 6.2.7 for parameters which satisfy the Feller condition. In that case the results are as expected.

#states			1	2	3	1	2	3
DD	PP	CCL	fixed rate loan			floating rate loan		
			0	0	0	0	0	0
x			225	211	226	194	112	174
	x		86	104	113	101	124	180
x	x		311	284	301	259	214	294
		x	245	223	239	233	132	209
	x	x	329	295	313	283	225	313

Table 6.2.4: Loan prices in the RS CIR-EJ++ model (in bps)

The results of the RS CIR-EJ++ model are presented in Table 6.2.4. They are consistent with the volatilities in the different regimes. The two regime model exhibits lower option prices than the one regime model because the volatility of the second regime is lower. The three regime model provides the highest volatility and therefore, as expected, the highest

option prices.

The differences between the option prices of the different models and correlation assumptions are large. But both calibrated models - the one regime CIR++ and CIR-EJ++ model without correlation - provide the same price range for the different products.

6.2.2 Results in Case of Valid Feller Condition

The same loan embedded options are additionally priced in all models under the parameters satisfying the Feller condition. In general, the volatilities of these parameter vectors are smaller and therefore the loan prices are smaller, too. Table 6.2.5 shows the effects in the CIR++ model which are quite similar to the calibrated model.

correlation			-1	0	1	-1	0	1
DD	PP	CCL	fixed rate loan			floating rate loan		
			34	0	-49	10	0	-14
x			117	100	67	37	27	143
	x		35	42	46	106	101	92
x	x		117	132	142	131	125	114
		x	119	105	76	65	57	43
	x	x	119	137	150	136	130	120

Table 6.2.5: Loan prices in the CIR++ model (in bps)

The tree price for the floating rate CCL is 57 bps for 1,000 time steps and deterministic interest rate. This price is very close to the analytical price of 57 bps. The fair coupons of the fixed rate and floating rate loans are reduced to 344 bps and 165 bps, respectively.

jumps			n	y	n	y
DD	PP	CCL	fixed rate loan		floating rate loan	
			0	0	0	0
x			77	103	10	18
	x		9	28	21	69
x	x		80	125	28	85
		x	78	104	10	19
	x	x	80	126	28	86

Table 6.2.6: Loan prices under jumps (in bps)

The effects in the CIR-EJ++ model are shown in Table 6.2.6. In the table the effect of the jump components is separated. The analytical values for the floating rate CCL with and

6.2. Loan and Contingent Credit Line Results

without jumps are 20 bps and 10 bps, respectively.

#states			1	2	3	1	2	3
DD	PP	CCL	fixed rate loan			floating rate loan		
			0	0	0	0	0	0
x			100	98	108	27	26	31
	x		42	37	50	101	85	124
x	x		132	122	144	125	106	150
		x	105	103	114	57	38	72
	x	x	137	126	150	130	110	164

Table 6.2.7: Loan prices in the RS CIR++ model (in bps)

The prices under the RS CIR++ model are presented in Table 6.2.7. Again, the loans are fair in case of no option as all models match the term structure exactly. Here, all models behave as expected for all different numbers of regimes. The single regime model provides more volatility than the two regime model, with the second regime being a good economy regime with low volatility. Therefore the loans with embedded options have a higher price in the single regime model. The third regime included in the three regime model has the highest volatility and therefore the highest prices.

#states			1	2	3	1	2	3
DD	PP	CCL	fixed rate loan			floating rate loan		
			0	0	0	0	0	0
x			103	101	114	18	17	22
	x		28	26	43	69	64	130
x	x		125	119	143	85	79	148
		x	104	102	115	19	18	39
	x	x	126	120	145	86	79	149

Table 6.2.8: Loan prices in the RS CIR-EJ++ model (in bps)

Table 6.2.8 presents the prices of the RS CIR-EJ++ model. The fixed rate loan values are pretty much the same as in case of the non-jump regime switching model.

The relative price differences of the Feller valid parameters are pretty much the same as of the calibrated parameters. Only for the RS CIR++ model the differences are large due to the effects on the calibrated parameters discussed above.

7 Conclusion

In this thesis default intensities and interest rates are modeled using variations of the CIR model. This chapter presents the conclusion including a short summary of the most important results and a short outlook on potential further developments in this research area.

Two different tree models and their convergence properties to analytical prices are evaluated and discussed. In the end, a combination of both tree construction procedures is the best option. It offers the advantages of both procedures. In addition to the model combination, a further jump handling method is introduced and its impact on the CIR model is discussed.

Up to now, this is the first CIR(-EJ) tree for modeling the default intensity process and its application for CDS and loan option pricing. For some examples where the semi-analytical solutions of BRIGO & ALFONSI (2005) and BRIGO & EL-BACHIR (2010) are applicable, the prices coincide very well. A semi-analytical price for the floating rate CCL is proven following their approach.

For the first time the interest rate and the default intensity are handled together within one combined CIR tree. A new two dimensional tree correlation procedure based on copulas is developed here, and its results are compared to those of the HULL & WHITE (1994b) methodology. Both methods are used to approximate the correlation in the discrete tree. The new copula approach is advantageous at extreme points.

Looking at the historic CDS data rendered obvious that the distribution of the time series cannot be handled using one and the same parameter vector for all time periods. Further exploration of the historic data revealed that there are at least three different states to explain the spread returns occurring for example on the iTraxx® Main and CDX® IG historic time series.

To handle the regime switching component the one state CIR tree has been extended. The same nodes are used for all regimes which makes the handling very memory efficient. The required memory size is not more than that of the single state CIR tree multiplied just by the number of states. In case of states with different volatilities, the tree steps can skip nodes in the higher volatility regime if necessary. Near the zero line two further extensions are required, a lower skipping node for the downward movement or a handling as in the binomial tree.

Finally all methods have been applied to real data to show their performance and the effects of their components. Both the jump component and the regime switching have huge impacts on the prices.

For further model development, the calibration of the RS CIR model to market data is a necessary step. This will require a (semi-) analytic solution of the bond option prices in case of regime switching.

The first dimension of the two dimensional CIR tree is used for the interest rate, the second dimension for the default intensity. The valuation is performed for credit products only for which the stochastic interest rate is primarily used for discounting. Interest rate payments are significant in case of floating loans only. The presented two dimensional tree can additionally be used for the counterparty valuation adjustment of interest rate products such as swaps or swaptions. Additionally, the effect of positive and negative correlation can be shown. A further application is to use both dimensions to model two different default intensities including their correlation structure. This potential of the tree can be used for valuation of double default products or counterparty valuation adjustments in credit products such as CDS and CDS options.

Further potential developments include the application of the discussed tree building procedures like regime switching to CEV processes of the form

$$dr(t) = k(\theta - r(t))dt + \sigma r^\beta(t) dW(t)$$

for $\beta \geq 0$. The CIR process is a CEV process for $\beta = \frac{1}{2}$.

Bibliography

- Acharya, V. V. & J. N. Carpenter (2002). Corporate bond valuation and hedging with stochastic interest rates and endogenous bankruptcy. *The Review of Financial Studies*, 15(5), 1355–1383.
- Ahn, C. M. & H. E. Thompson (1988). Jump-diffusion processes and the term structure of interest rates. *The Journal of Finance*, 43(1), 155–174.
- Alexander, C. & A. Kaeck (2008). Regime dependent determinants of credit default swap spreads. *Journal of Banking & Finance*, 32(6), 1008–1021.
- Amin, K. I. (1993). Jump diffusion option valuation in discrete time. *The Journal of Finance*, 48(5), 1833–1863.
- Ang, A. & G. Bekaert (2002). Regime switches in interest rates. *Journal of Business and Economic Statistics*, 20(2), 163–182.
- Bansal, R., A. R. Gallant & G. Tauchen (1995). Nonparametric estimation of structural models for high-frequency currency market data. *Journal of Econometrics*, 66(1-2), 251–287.
- Beliaeva, N. A., S. K. Nawalkha & G. M. Soto (2007). Pricing American interest rate options under the jump-extended CIR and CEV short rate models.
<http://ssrn.com/abstract=985839>.
- Bielecki, T. R., M. Jeanblanc & M. Rutkowski (2011). Hedging of a credit default swaption in the CIR default intensity model. *Finance and Stochastics*, 15(3), 541–572.
- Bielecki, T. R. & M. Rutkowski (2004a). *Credit risk: Modeling, valuation and hedging*. Springer Science+Business Media, 2nd edition.
- Bielecki, T. R. & M. Rutkowski (2004b). Modeling of the defaultable term structure: Conditionally Markov approach. *IEEE Transactions on Automatic Control*, 49(3), 361–373.
- Black, F. (1976). The pricing of commodity contracts. *Journal of Financial Economics*, 3(1-2), 167–179.
- Bluhm, C., L. Overbeck & C. Wagner (2008). *Introduction to credit risk modeling*. Chapman & Hall/CRC Financial Mathematics Series, 2nd edition.
- Bollen, N. (1998). Valuing options in regime-switching models. *The Journal of Derivatives*, 6(1), 38–49.

Bibliography

- Brace, A., D. Gatarek & M. Musiela (1997). The market model of interest rate dynamics. *Mathematical Finance*, 7(2), 127–155.
- Brigo, D. (2005). Market models for CDS options and callable floaters. *Risk*, 18(1), 89–94.
- Brigo, D. & A. Alfonsi (2005). Credit default swap calibration and derivatives pricing with the SSRD stochastic intensity model. *Finance and Stochastics*, 9(1), 29–42.
- Brigo, D. & N. El-Bachir (2006). Credit derivatives pricing with a smile-extended jump stochastic intensity model. http://www.defaultrisk.com/pp_crdrv_96.htm.
- Brigo, D. & N. El-Bachir (2010). An exact formula for default swaptions pricing in the SSRJD stochastic intensity model. *Mathematical Finance*, 20(3), 365–382.
- Brigo, D. & F. Mercurio (2001). A deterministic-shift extension of analytically-tractable and time-homogeneous short-rate models. *Finance and Stochastics*, 5(3), 369–387.
- Brigo, D. & F. Mercurio (2006). *Interest rate models - theory and practice: With smile, inflation and credit*. Springer Science+Business Media, 2nd edition.
- Brigo, D. & M. Morini (2005). CDS market formulas and models. http://www.defaultrisk.com/pp_crdrv171.htm.
- Brown, S. J. & P. H. Dybvig (1986). The empirical implications of the Cox, Ingersoll, Ross theory of the term structure of interest rates. *The Journal of Finance*, 41(3), 617.
- Carmona, R. & N. Touzi (2008). Optimal multiple stopping and valuation of swing options. *Mathematical Finance*, 18(2), 239–268.
- Christensen, J. (2002). *Kreditderivater og deres prisfastsættelse*. Master thesis, University of Copenhagen.
- Christensen, J. (2007). *Default and recovery risk modeling and estimation*. Ph.d thesis, Copenhagen Business School. <http://www.gbv.de/dms/zbw/541699032.pdf>.
- Cox, J. C., J. E. Ingersoll Jr & S. A. Ross (1985a). A theory of the term structure of interest rates. *Econometrica*, 53(2), 385–407.
- Cox, J. C., J. E. Ingersoll Jr & S. A. Ross (1985b). An intertemporal general equilibrium model of asset prices. *Econometrica*, 53(2), 363–384.
- Cox, J. C., S. A. Ross & M. Rubinstein (1979). Option pricing: A simplified approach. *Journal of Financial Economics*, 7, 229–263.
- Davies, A. (2004). Credit spread modeling with regime-switching techniques. *The Journal of Fixed Income*, 14(3), 36–48.
- Duffie, D., J. Pan & K. J. Singleton (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6), 1343–1376.
- Duffie, D. & K. J. Singleton (1999). Modeling term structures of defaultable bonds. *Review of Financial Studies*, 12(4), 687–720.

- Elliott, R. J., L. Aggoun & J. B. Moore (1995). *Hidden markov models: Estimation and control*. Springer Science+Business Media, 1st edition.
- Elliott, R. J., L. Chan & T. K. Siu (2005). Option pricing and Esscher transform under regime switching. *Annals of Finance*, 1(4), 423–432.
- Elliott, R. J., M. Jeanblanc & M. Yor (2000). On models of default risk. *Mathematical Finance*, 10(2), 179–195.
- Engel, C. & J. D. Hamilton (1990). Long swings in the Dollar: Are they in the data and do markets know it? *The American Economic Review*, 80(4), 689–713.
- Engelmann, B. (2011). A framework for pricing and risk management of loans with embedded options. Tech. rep., Working Paper.
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1781991.
- Feller, W. (1951). Two singular diffusion problems. *The Annals of Mathematics*, 54(1), 173–182.
- Gallant, A. R. & G. Tauchen (1996). Which moments to match? *Econometric Theory*, 12, 657–681.
- Garcia, J., H. van Ginderen & R. Garcia (2003). On the pricing of credit spread options: A two factor HW-BK algorithm. *International Journal of Theoretical and Applied Finance*, 6(5), 491–506.
- Gibson, R., F. Lhabitant & D. Talay (2001). Modeling the term structure of interest rates: A review of the literature. <http://ssrn.com/abstract=275076>.
- Giesecke, K., F. A. Longstaff, S. Schaefer & I. Strebulaev (2011). Corporate bond default risk: A 150-year perspective. *Journal of Financial Economics*, 102(2), 233–250.
- Gray, S. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*, 42(1), 27–62.
- Hamilton, J. D. (1988). Rational-expectations econometric analysis of changes in regime. *Journal of Economic Dynamics and Control*, 12, 385–423.
- Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2), 357–384.
- Hamilton, J. D. (1990). Analysis of time series subject to changes in regime. *Journal of Econometrics*, 45, 39–70.
- Heath, D., R. A. Jarrow & A. Morton (1992). Bond pricing and the term structure of interest rates : A new methodology for contingent claims valuation. *Econometrica*, 60(1), 77–105.
- Hull, J. (1996). Using Hull-White interest rate trees. *Journal of Derivatives*, 3(3), 26–36.
- Hull, J. & A. White (1990). Pricing interest-rate-derivative securities. *Review of Financial Studies*, 3(4), 573–592.

Bibliography

- Hull, J. & A. White (1993). One-factor interest-rate models and the valuation of interest-rate derivative securities. *Journal of Financial and Quantitative Analysis*, 28(2), 235–254.
- Hull, J. & A. White (1994a). Numerical procedures for implementing term structure models I: Single-factor models. *The Journal of Derivatives*, 2(1), 7–16.
- Hull, J. & A. White (1994b). Numerical procedures for implementing term structure models II: Two-factor models. *The Journal of Derivatives*, 2(2), 37–48.
- Hull, J. & A. White (2000). Valuing credit default swaps I: No counterparty default risk. *Journal of Derivatives*, 8(1), 29–40.
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1295226.
- Hull, J. & A. White (2003). The valuation of credit default swap options. *Journal of Derivatives*, 10(3), 40–50.
- Jacod, J. & A. N. Shiryaev (1987). *Limit theorems for stochastic processes*. Springer-Verlag.
- Jamshidian, F. (1989). An exact bond option formula. *The Journal of Finance*, 44(1), 205–209.
- Jamshidian, F. (1995). A simple class of square-root interest-rate models. *Applied Mathematical Finance*, 2(1), 61–72.
- Jamshidian, F. (1997). Libor and swap market models and measures. *Finance and Stochastics*, 1(4), 293–330.
- Jamshidian, F. (2004). Valuation of credit default swaps and swaptions. *Finance and Stochastics*, 8(3), 343–371.
- Jarrow, R. A., D. Lando & S. M. Turnbull (1997). A Markov model for the term structure of credit risk spreads. *Review of Financial Studies*, 10(2), 481.
- Jarrow, R. A. & S. M. Turnbull (1995). Pricing derivatives on financial securities subject to credit risk. *Journal of Finance*, 50(1), 53–85.
- Jeanblanc, M. & M. Rutkowski (2000a). Modelling of default risk: An overview. *Modern Mathematical Finance: Theory and Practice*, 171–269.
- Jeanblanc, M. & M. Rutkowski (2000b). Modelling of default risk: Mathematical tools. *preprint*, 1–67. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.139.678&rep=rep1&type=pdf>.
- Kladivko, K. (2007). Maximum likelihood estimation of the Cox-Ingersoll-Ross process: The Matlab implementation. Tech. Rep. 3, working paper.
http://dsp.vscht.cz/konference_matlab/MATLAB07/prispevky/kladivko_k/kladivko_k.pdf.
- Landén, C. (2000). Bond pricing in a hidden Markov model of the short rate. *Finance and Stochastics*, 4(4), 371–389.

- Lando, D. (1998). On Cox processes and credit risky securities. *Review of Derivatives Research*, 2(2), 99–120.
- Lando, D. (2004). *Credit risk modeling: Theory and applications*. Princeton University Press, 1st edition.
- Lesne, J.-P., J.-L. Prigent & O. Scaillet (2000). Convergence of discrete time option pricing models under stochastic interest rates. *Finance and Stochastics*, 4(1), 81–93.
- Leung, T. S. T. & K. Yamazaki (2010). American step-up and step-down credit default swaps under Levy models. <http://adsabs.harvard.edu/abs/2010arXiv1012.3234S>.
- Li, A., P. Ritchken & L. Sankarasubramanian (1995). Lattice models for pricing American interest rate claims. 50(2), 719–737.
- Lindgren, G. (1978). Markov regime models for mixed distributions and switching regressions. *Scandinavian Journal of Statistics*, 5(2), 81–91.
- Loukoianova, E., S. N. Neftci & S. Sharma (2007). Pricing and hedging of contingent credit lines. *The Journal of Derivatives*, 14(3), 61–80.
- Maalaoui Chun, O., G. Dionne & P. François (2010a). Credit spread changes within switching regimes. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1341870.
- Maalaoui Chun, O., G. Dionne & P. François (2010b). Detecting regime shifts in corporate credit spreads. <http://ssrn.com/abstract=1146583>.
- Maghsoodi, Y. (1996). Solution of the extended CIR term structure and bond option valuation. *Mathematical Finance*, 6(1), 89–109.
- Merton, R. (1974). On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance*, 29(2), 449–470.
- Nawalkha, S. K. & N. A. Beliaeva (2007). Efficient trees for CIR and CEV short rate models. *The Journal of Alternative Investments*, 10(1), 71–90.
- Nawalkha, S. K., G. M. Soto & N. A. Beliaeva (2007). *Dynamic term structure modeling: The fixed income valuation course*. John Wiley & Sons, 1st edition.
- Nelson, D. & K. Ramaswamy (1990). Simple binomial processes as diffusion approximations in financial models. *Review of Financial Studies*, 3(3), 393–430.
- Norden, L. & W. Wagner (2008). Credit derivatives and loan pricing. *Journal of Banking and Finance*, 32(12), 2560–2569.
- O’Kane, D. (2008). *Modelling single-name and multi-name credit derivatives*. John Wiley & Sons, 1st edition.
- Overbeck, L. & T. Rydén (1997). Estimation in the Cox-Ingersoll-Ross model. *Econometric Theory*, 13(3), 430–461.

Bibliography

- Rebonato, R. (1998). *Interest-rate option models: Understanding, analysing and using models for exotic interest-rate options*. John Wiley & Sons, 2nd edition.
- Rogers, L. (1995). Which model for term-structure of interest rates should one use? *Mathematical Finance*, 65, 93–116.
- Schönbucher, P. (1998). Term structure modelling of defaultable bonds. *Review of Derivatives Research*, 2(2), 161–192.
- Schönbucher, P. J. (2000a). A Libor market model with default risk. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=261051.
- Schönbucher, P. J. (2000b). A tree implementation of a credit spread model for credit derivatives. <http://www.ssrn.com/abstract=240868>.
- Schönbucher, P. J. (2003). A note on survival measures and the pricing of options on credit default swaps. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.14.1824&rep=rep1&type=pdf>.
- Shreve, S. E. (2005). *Stochastic calculus for finance I: The binomial asset pricing model*. Springer Finance.
- Siu, T. K., C. Erlwein & R. Mamon (2008). The pricing of credit default swaps under a Markov-modulated Merton's structural model. *North American Actuarial Journal*, 12(1), 19–46.
- Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5, 177–188.
- Wong, H. Y. & T. L. Wong (2008). Reduced-form models with regime switching: An empirical analysis for corporate bonds. *Asia-Pacific Financial Markets*, 14(3), 229–253.
- Xu, R. & S. Li (2009a). A tree model for pricing credit default swaps with equity, market and default risk. *Management and Service Science, 2009. MASS'09*, 1(4).
- Xu, R. & S. Li (2009b). A tree model for pricing credit spread options subject to equity, and market risk. *2009 ISECS International Colloquium on Computing, Communication, Control, and Management*, 1(4), 46–50.
- Yin, G. G. & Q. Zhang (1997). *Continuous-time Markov chains and applications: A singular perturbation approach*. Springer Science+Business Media, 1st edition.
- Yuen, F. L. & H. Yang (2010). Option pricing with regime switching by trinomial tree method. *Journal of Computational and Applied Mathematics*, 233(8), 1821–1833.
- Zhou, H. (2001). Jump-diffusion term structure and Ito conditional moment generator. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=278440.

Acknowledgments

First, I want to take the opportunity to express my deep gratitude to my supervisor Prof. Dr. Ludger Overbeck. His trust in my person, his very useful comments and ideas has been a great support for me in writing this thesis. He was always helpful and encouraged me to work hard and steadily. I thank Prof. Dr. Winfried Stute to be a referee of my doctoral thesis.

Also, I want to thank my colleagues at Commerzbank AG for fruitful discussions and the first insights into the financial world. And I want to thank Commerzbank AG which enabled the performance of the work by its financial support.

I would like to thank Markit Group Limited, Bloomberg L.P. and Commerzbank AG for the permission to include copyrighted numbers, tables and graphs.

My final very important thanks go to my family and friends for their patience, constant support and encouragement - especially Denise, Matteo and my parents. Thank you.