

#### Delft University of Technology Faculty of Electrical Engineering, Mathematics and Computer Science Delft Institute of Applied Mathematics

#### Continuous Term Structures for Implied Recovery

A thesis submitted to the Delft Institute of Applied Mathematics in partial fulfillment of the requirements

for the degree

#### MASTER OF SCIENCE in APPLIED MATHEMATICS

by

Nathan Juliano Meibergen Amsterdam, The Netherlands February 2015



This thesis is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.





#### MSc THESIS APPLIED MATHEMATICS

#### "Continuous Term Structures for Implied Recovery"

NATHAN JULIANO MEIBERGEN

Delft University of Technology

Responsible professor

Prof. dr. C. W. Oosterlee

#### Other thesis committee members

dr. ir. J. H. M. Anderluh.

dr. D. Fokkema

February 2015

Amsterdam, The Netherlands

## Abstract

Credit risk pricing models assume recovery to be at its *historical* average (historical recovery assumption). However, the effect of this assumption is not completely understood.

The heard of this thesis lies in constructing a new pricing model for Credit Default Swaps (CDS), in particularly allowing for negative correlation between recovery and default. This model is denoted as partial differential equations for the CDS legs. By means of an additional Monte Carlo approach we are able to extract continuous implied recovery and default intensity term structures.

These structures can then be used to assess the historical recovery assumption. It is in particularly shown that a constant recovery model overestimates the Credit Value Adjustment (CVA) when allowing for perfect negative correlation. While on the other hand it underestimates CVA when it is adjusted for its implied historical average.

## Preface

One year ago I started a journey. A journey on financial mathematics, with a special focus on credit risk. One part of this journey is this thesis, which had to be completed in order to receive my Masters degree in Applied Mathematics at Delft University of Technology. However, this journey contained much more. It was a rich experience of working on this thesis and being at the EY office five/six/seven days a week, surrounded by many competent people. During this year I was given the opportunity to participate in many project, such as on-site project of two months at a major bank in the Netherlands and being a facilitator for an internal training on financial mathematics given in Frankfurt. All together I have learned much more than what is written inside this thesis.

The first few months were merely a process of realizing that I was nothing but a newbie. So often did I encounter myself thinking, 'yes, I got it!', when less than a minute later my world would collapse again. So many books and so many article on subjects I had never heard before. And eager for knowledge as I am, I tried to understand everything, however soon I realized that some things just had to wait for at least a year.

In short, where this thesis started out as a research on some *key questions* about implied recovery, it ended up in the construction of a model for implied recovery (which on turn was meant to answer these key questions). In order to properly understand 'why and how' this model, here is some guidance.

Chapter 1 starts with an introduction to the introduction, describing the key questions. Its next sections establish required understanding of the most important underlying elements for this thesis. Especially with respect to the recovery it looks at statistics (historical) and its modeling in the pricing framework (forward looking), where I finally end with a modeling criteria on what I believe to be important within an implied recovery model. It then seems that there exists a model that does exactly that, which is described in chapter 2. This model however seemed to have a drawback touching upon its reliability. Therefore, in chapter 3, I construct my own model. Eventually showing this model to be a generalization of the model from chapter 2, without the drawback. This section looks furthermore at the interpretation of the parameters and dynamical stability.

The model that I present is in PDE form and I describe how its solution can be approximated by means of finite differences in section 4. Furthermore I look at properties such as numerical stability and the choice of modeling parameters. Then in chapter 5, I show how the model can be fitted to market data, present some examples and reflect on the initially described key questions. The conclusion then summarizes the findings.

There are also some appendices, where I would like to highlight appendix A, which contains an important proof required for setting up boundary conditions within the finite differences framework. It also discusses some extensions to our PDE model. Also appendix I with Matlab code can be interesting for those willing to learn object orientated programming in Matlab.

On a side note, I would like to mention that proofs are only provided if they are established by me, in any other case a reference is given.

## Symbols

A	Alternative notation for CDS premium leg
В	Alternative notation for CDS protection leg
$\overline{C}$	CDS market spread
C	CDS spread
D	discounting term
$\Phi$	standard normal cumulative distribution function
$\Delta x$	step size in spacial direction
$\Delta t$	step size in time direction
${\cal F}$	instance of reference filtration
$\mathbb{F}$	reference filtration
f	risk-free forward interest rate
${\cal G}$	instance of greater filtration
$\mathbb{G}$	greater filtration
Γ	$\mathbb{F}$ -hazard process
${\cal H}$	instance of default filtration
H	default filtration
Lgd	absolute loss given default.
LGD	percentage loss given default.
$\lambda$	default intensity
$\overline{M}$	martingale compensated bankruptcy jump
	process
M	maximum spacial index on finite differences grid

$\overline{N}$	Notional
N	maximum time index on finite differences grid
$\mathbb{P}$	real-world measure
Prem	premium leg of a CDS
Prot	protection leg of a CDS
Q	risk-neutral measure
R	absolute recovery.
$\rho$	percentage recovery.
r	risk-free interest rate
S	Stock price
T	time of maturity
au	time of default
W	standard Brownian motion
I $\xi$	Poisson distribution with intensity $\lambda$ probability of default over some time span
P	value of a zero coupon bond.

# Acronyms

APR	Absolute Priority Rule
CDS	Credit Default Swap
CRR	Cox Ross Rubinstein
CVA	Credit Value Adjustment
GBM	Geometric Brownian Motion
LGD	Loss Given Default
lhs	Left Hand Side
MCMC	Markov Chain Monte Carlo
MCMC PC	Markov Chain Monte Carlo Piecewise Constant
MCMC PC PD	Markov Chain Monte Carlo Piecewise Constant Probability of Default
MCMC PC PD PDE	Markov Chain Monte Carlo Piecewise Constant Probability of Default Partial differential Equation
MCMC PC PD PDE	Markov Chain Monte Carlo Piecewise Constant Probability of Default Partial differential Equation
MCMC PC PD PDE RFV	Markov Chain Monte Carlo Piecewise Constant Probability of Default Partial differential Equation Recovery of Face Value
MCMC PC PD PDE RFV rhs	Markov Chain Monte Carlo Piecewise Constant Probability of Default Partial differential Equation Recovery of Face Value Right Hand Side
MCMC PC PD PDE RFV rhs RMV	Markov Chain Monte Carlo Piecewise Constant Probability of Default Partial differential Equation Recovery of Face Value Right Hand Side Recovery of Market Value
MCMC PC PD PDE RFV rhs RMV RT	Markov Chain Monte Carlo Piecewise Constant Probability of Default Partial differential Equation Recovery of Face Value Right Hand Side Recovery of Market Value Recovery of present value of debt

## Contents

Abstract	II
Preface	IV
Symbols	VI
Acronyms	VIII

1	Intr	oducti	ion	1
	1.1	The co	orrect measure	3
		1.1.1	Risk-neutrality: the $\mathbb{Q}$ -measure $\ldots$ $\ldots$ $\ldots$ $\ldots$ $\ldots$ $\ldots$ $\ldots$	3
		1.1.2	Physical measure: the $\mathbb{P}$ -measure	4
		1.1.3	The difference between the $\mathbb{P}$ - and $\mathbb{Q}$ -measure	4
	1.2	The C	Credit Default Swap	7
		1.2.1	Settlement	8
		1.2.2	Credit events	9
		1.2.3	Modeling	9
			1.2.3.1 Premium leg	10
			1.2.3.2 Protection leg	11
			1.2.3.3 The CDS spread	12
			1.2.3.4 Piecewise linear intensity	13
	1.3	Recov	rery	15
		1.3.1	Physical recovery	16
			1.3.1.1 Determinants	16
			Firm	17
			Industry	17
			Macro	18
			Contract	18
		1.3.2	Implied recovery	19
			1.3.2.1 Differences to the physical recovery	19
			1.3.2.2 Difficulties	21
		1.3.3	Modeling implied recovery	21
			1.3.3.1 Assumptions on the recovery	21
			1.3.3.2 The identification problem	24
			1.3.3.3 A brief overview of current models for implied recovery .	25
			1.3.3.4 Modeling criteria	27

<b>2</b>	Dis	crete t	erm structures: The D&H model							29
	2.1	Descri	ption							. 30
	2.2	Metho	odology							. 31
		2.2.1	The jump-to-default tree							. 31
		2.2.2	Pricing a CDS							. 33
			Assumptions							. 33
			Backward recursions							. 33
		2.2.3	Calibration on CDSs							. 34
	2.3	Retrie	ving recovery and default							. 35
	2.4	Restri	ction of the method							. 36
	2.5	A test	set by Amazon					•		. 39
3	Cor	ntinuoi	is term structures: The PDE model							41
U	3.1	Jump	to default stock prices							42
	3.2	Infinit	esimal increments in both legs of a CDS	•••	•	•••	• •	•	•	· 12 44
	0.2	3 2 1	The premium leg	•••	·	•••	• •	•	•	. 11 //
		0.2.1 3.9.9	The protection log	• •	•	•••	• •	•	•	. 44
	<b>?</b> ?	Docult	ing CDS DDE <sub>2</sub>	• •	•	•••	• •	•	·	. 40 51
	ე.ე ე_∤	De-II		• •	·	•••	• •	•	·	. 01 E4
	3.4 2.5	D&n D-f	Inoder as a special case	• •	·	•••	• •	•	·	. 04 FO
	3.5	Derau	It and recovery choice	• •	·	• •	• •	•	·	. 58 50
		3.5.1	Default	• •	•	•••	• •	•	•	. 58
		3.5.2	Recovery	• •	·	• •	• •	•	·	. 58
	3.6	Retrie	ving term structures by a Monte Carlo simulation	• •	·	•••	• •	•	·	. 61
	3.7	Paran	neter contribution	• •	·	•••	• •	•	·	. 63
		3.7.1	To the CDS spread term structure	• •	·	•••	• •	•	·	. 63
			Scenario A	• •	•	•••	• •	•	•	. 63
			Scenario B	• •	•	•••	• •	•	•	. 67
		3.7.2	To the default and recovery term structures		•			•	•	. 68
			Default intensity		•		• •	•		. 68
			$Recovery \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots$		•			•		. 68
	3.8	Dynar	nic Stability		•	•••	• •	•	•	. 72
<b>4</b>	Fin	ite diff	erences for the PDE model							74
	4.1	Fully 1	Implicit finite differences approach							. 77
		4.1.1	Initial and boundary conditions		•					. 77
			Premium leg							. 77
			Protection leg							. 78
		4.1.2	Discretization							. 79
			Specifying the grid							. 79
			Derivative truncation							. 79
	4.2	Grid s	tretching							. 83
		4.2.1	Grid stretching on our model							. 83
		4.2.2	Methodology							. 84
		4.2.3	Implementing in the PDE							. 86
		4.2.4	Results							. 87
	4.3	Conve	rgence		-					. 89
	1.0	4.3.1	Well-posedness					•		. 89
		T.O.T		· ·	•	• •		•	•	. 00

		4.3.2 Stability	90
	4.4	Parameter choice	95
		4.4.1 A maximum stock value	95
		4.4.2 Temporal and stock step sizes	97
5	Cali	brations and results	100
	5.1	Calibration process	102
		5.1.1 The objective function	102
		5.1.2 Constraints	104
	5.2	Comparing PDE and D&H model results	105
	5.3	Individual firm calibration with single seniority	107
		5.3.1 Individual firm calibration with multiple seniorities	111
	5.4	Determining CVA	115
	5.5	A case comparison to the Piecewise Constant (PC) model	117
	5.6	Implied recovery at its historical average	120
6	Con	clusion	121
A	The	PDE model: Supplements	123
	A.1	The probability of default under the stock limit	124
	A.2	The switching line	126
		Recovery model by D&H	126
		Back to the PDE recovery model	129
	A.3	More recovery models for the PDE model	131
в	Pric	ing basics	134
	B.1	The risk neutral martingale framework	135
	B.2	Ito's lemma	138
С	Deb	t hierarchy and capital structure	140
D	Moo	leling the default time	142
	D.1	Filtrations	143

D	Modeling the default time	142
	D.1 Filtrations	143
	D.2 The hazard process	145
	The canonical hazard process	145
	D.3 The Cox process	148
Е	Convertible bond and PDE model E.1 A pricing PDE	<b>149</b> 150
$\mathbf{F}$	Von Neumann numerical stability	153
$\mathbf{G}$	Pricing a defaultable bond	157
н	Data	159

Data						
	H.1	Firm data	160			
	H.2	Risk-free interest rates	162			

Ι	Mat	Matlab Code					
	I.1	D&H	model of Das&Hanouna	165			
		I.1.1	Main code for Calibration and pricing of CDSs	165			
		I.1.2	Calculation of multiple spreads	166			
		I.1.3	Calculation of spread given input variables and parameters $\ldots$ .	168			
		I.1.4	Reaching probabilities	169			
	I.2	Main	optimization file	170			
		I.2.1	Classes	171			
		I.2.2	Functions	175			
		I.2.3	Dummy functions	180			
		I.2.4	Monte Carlo for retrieving the term structure	180			

#### Bibliography

183

### 1

## Introduction

"A wise man proportions his belief to the evidence."

- David Hume, An Enquiry Concerning Human Understanding

Recent years financial models started providing better estimates on credit risk. In particular after the financial crisis of 2007-2008 the market came to understand its importance. Counterparty credit risk is most importantly identified by the default probability and recovery. Default probability tells us to what extend we might expect a default to occur, whereas recovery is the amount payed back (recovered) by the defaulted firm. A lot of research has been done on the area of default probability, in particular with respect to structural and reduced form models, see appendix D. However, less research is conducted on the aspect of recovery. Therefore, in the scope of pricing models, market convention usually fixes it at its historical average.

With respect to such pricing models it was initially the goal of this thesis to answer the questions:

- 1. Is modeling recovery as a constant sufficient?
- 2. If it is, should we fix it at its historical average?

However, in order to answer such questions we have to have a model for comparison, that is, a model that allows us to inspect results when the recovery is not fixed at its historical average. This is where the journey actually begins. To answer the two questions above, we have to digress from them by providing a model allowing the recovery to be non constant and to satisfy our desired criteria described more thoroughly in section 1.3.3.4. The model which we created is called the *PDE model* and much of this thesis is dedicated to constructing and analyzing this model. However, before we can get to fully understand this model we have to have some basics.

 $\mathbf{2}$ 

First this chapter will describe the historical and pricing measure, these are important to us as market conventions fixes recovery at its historical average, whereas pricing models that include the recovery require the pricing measure variant, possibly implying a discrepancy between both.

Next in section 1.2 the Credit Default Swap (CDS) is described, a financial product functioning as a sort of insurance. This insurance product became an important indicator of a firm's financial health, and therefore an indicator of default probability and recovery. The model that we construct later is based on pricing such CDS to separately extract from it the probability of default and recovery.

In section 1.3 we will look more closely at the recovery, by first looking at the historical recovery and then at the recovery within the pricing framework. Finally the usual models for recovery are discussed and we end with our modeling criteria.

#### 1.1 The correct measure

When calculating certain probabilities or expectations we are always implicitly basing this upon a certain measure. A measure can be understood as the distribution of probabilities related to events. In measure  $M_1$  event A may have probability 0.3, whereas in measure  $M_2$  this same event may have a probability of 0.4.

The understanding of the usage of measures is especially important in the world of finance. Formulas and deductions may be simplified if a convenient measure is applied, whereas results may only be valid under the correct measure.

We won't get into the details of measure theory as this a very large and specific field within mathematics. What we do want, is to give some intuition for the two most frequently used measures within finance. This intuition is often enough to understand which one to use when.

Two measures are of great importance, namely the physical measure, which contains the probabilities of certain events occurring in our real world: if firms go bankrupt with probability 0.01, then on the long term 1% of the firms should have gone bankrupt. The other measure is the risk-neutral measure, this measure sets probabilities in such way that they agree with our model and the market expectations: even though on the long term 1% of all firms go bankrupt, the risk-neutral measure might tell us that the probability of a bankrupt equals 0.015. We will explain both and their differences more thoroughly.

#### 1.1.1 Risk-neutrality: the $\mathbb{Q}$ -measure

The risk-neutral measure describes a complete world in which there are no risks related to any financial product. All contingent attainable claims are assumed to move according to a risk free rate (see (B.7)). Thus suppose a contingent attainable claim  $V_T$  matures at time T, let the interest rate r be a known constant, then

$$V_t = \mathbb{E}_t^{\mathbb{Q}} \left[ e^{-r(T-t)} V_T \right].$$
(1.1)

As the risk-neutral world induced by  $\mathbb{Q}$ -measure assumes there to be no risks, the terminal payoff  $V_T$  is risk-free and by a no-arbitrage argument it should be discounted by the risk free rate. The convenience of this measure is twofold:

- 1. The discounting term is always set to the risk-free interest rate r, as all products are risk-free. In reality this payoff doesn't have to be risk-free such that deciding on the discounting term is a complicated issue.
- 2. Everything in this world increases at a risk-free rate and we thus don't have to decide upon the correct drift term. The most straightforward example of this convenience

is seen in modeling the movement of stock prices, which in this world have drift term r, such that a simple Monte Carlo approach can be used to determine prices.

In appendix B it is shown to some extend how the risk-neutral measure is constructed as the measure which makes contingent attainable claims a martingale under the money market account numeraire.

As (1.1) tells us, under the risk neutral measure, all that is required to price a product is its resulting claim. And as it is very complicated to quantify risk premiums we prefer determining the claims of products such that discounted by the risk-neutral interest rate we retrieve the fair price. Usually however, we cannot directly fill in the Right Hand Side (rhs) of (1.1) as it will depend upon unknown parameters, think for example about a probability of default. Instead we come up with a model for the claim and then set the unknown parameters such that the expectation provides us with the correct price. This process, in which parameters are adjusted to fit the market prices is called *calibration*. The resulting parameters are called their implied counterparts, e.g. the implied probability of default.

As we will show in section 1.1.3, implied parameters do not represent their real world counterparts due to the lack of risk premiums.

#### 1.1.2 Physical measure: the $\mathbb{P}$ -measure

The  $\mathbb{P}$ -measure corresponds to observed realisations of underlying properties. For example: the probability of default in the  $\mathbb{P}$ -measure corresponds to analysing the probability of default from historical realisations, thus by observing for example an empirical distribution of defaults.

The  $\mathbb{P}$ -measure being related to historical observables provides the misleading understanding<sup>1</sup> that: 'anything that is historical is strictly related to the  $\mathbb{P}$ -measure, whereas anything that is forward looking is in the  $\mathbb{Q}$ -measure.' It is true that observables are  $\mathbb{P}$ -measure, thus historical prices are  $\mathbb{P}$ -measure prices. We may however use these to find historical implied parameters. The process of calibration can easily be applied to historical data such that we can get historical implied data, which may sound as a contradiction. An example of such approach can be found in [SUH14]).

#### 1.1.3 The difference between the $\mathbb{P}$ - and $\mathbb{Q}$ -measure

At the beginning of this section we stated that probabilities in both measures might differ. We shall now show you where this difference comes from. In short the answer is:

<sup>&</sup>lt;sup>1</sup>I have personally frequently heard this faulty understanding, even from people with a decent amount of experience.



FIGURE 1.1: The steps in calibrating a pricing model and extracting implied data.

implied/real-world parameters get adjusted to compensate for risk premiums. Lets clarify this next.

In contrast to what is assumed in the risk-neutral framework, the world is not complete, such that not everything can be hedged perfectly. Think for example about liquidity risk, the risk associated to exiting or entering a required position. Another example would be systematic risk, which is a risk inherent to an entire market and it is therefore a risk for which you cannot diversify.

This however does not stop us from using our  $\mathbb{Q}$ -measure model and fit these to market data anyways. What that means is that implied parameters do not coincide with their  $\mathbb{P}$ -measure counterparts, see figure 1.1.

Now we may also understand why calibration is often performed on vanilla options. As these are very liquid we can be sure that there is no risk-aversion nor liquidity risk present in these prices, such that implied parameters approximate reality better.

Now that we have seen implied parameters to be different from their real world counterparts, lets look at an example to make this completely clear. Suppose we are pricing a simplified risky bond, both using the  $\mathbb{P}$ - and  $\mathbb{Q}$ -measure. It is known that parameters that define risk become larger (see for example [HPW05]) in the  $\mathbb{Q}$ -measure, this is what we shall show next with the probability of default.

Suppose today is time t, and that a bond, paying  $\mathbf{1}_{\{\tau > T\}}$ , that is zero if it defaults before maturity T and one if it does not. In the market we observe the price

'risky bond price' = 
$$P - s$$
,

where P is the risk-neutral price, that is, under the assumption of a complete world, and  $s \ge 0$  is the additional premium to make the bond somewhat cheaper to compensate for the real world risks. The claim is clear such that by the risk-neutral pricing framework we can fit it to the market to find

$$P - s = \mathbb{E}^{\mathbb{Q}} \left[ e^{-r(T-t)} \mathbf{1}_{\tau > T} \right]$$
(1.2)

$$= e^{-r(T-t)} \mathbb{E}^{\mathbb{Q}} \left[ \mathbf{1}_{\tau > T} \right] \tag{1.3}$$

$$= e^{-r(T-t)} \mathbb{Q}(\tau > T). \tag{1.4}$$

On the other hand if the real world would be risk-free, then risk-neutral measure and real-world measure would coincide such that

$$P = \mathbb{E}^{\mathbb{Q}} \left[ e^{-r(T-t)} \mathbf{1}_{\tau > T} \right]$$
(1.5)

$$= \mathbb{E}^{\mathbb{P}} \left[ e^{-r(T-t)} \mathbf{1}_{\tau > T} \right]$$
(1.6)

$$= e^{-r(T-t)} \mathbb{E}^{\mathbb{P}}[1_{\tau>T}]$$
(1.7)

$$= e^{-r(T-t)} \mathbb{P}(\tau > T).$$
(1.8)

Now we indeed find that the implied probability of default turns out larger then it is in reality:

$$\mathbb{Q}(\tau > T) = \mathbb{P}(\tau > T) - e^{r(T-t)}s \quad \leftrightarrow \quad \mathbb{Q}(\tau \le T) \ge \mathbb{P}(\tau \le T).$$
(1.9)

To get a little ahead of ourselves, we already note that the authors of [SUH14] show that implied recovery is influenced by incompleteness and systematic risk.

#### 1.2 The Credit Default Swap

A CDS is a credit derivative which was initially introduced by JP Morgan in 1994 as a type of insurance upon debt for creditors. As we might expect from an insurance, if it is very expensive it is probably related to something very risky, whereas if it is fairly cheap, then the related risk is probably very small. For this reason, prices of CDSs provide an indication towards the size of certain risks (in particular creditworthiness of a counterparty), and it is for this reason that we are interested in their behavior.

The CDS works as follow: in exchange for a certain periodic premium, called the CDS spread/premium, a creditor can buy insurance referencing some kind of debt, debtor and notional value. Whenever one of the credit events, specified by the CDS, gets triggered at the side of the debtor, the insured creditor could physically deliver the debt - with a maximum face value as specified by the notional value - to the insurer in trade for its face value.

To be more specific: a CDS specifies

- 1. the type of debt (usually senior unsecured and senior subordinated);
- 2. the reference entity, which is the debtor of such senior unsecured bonds;
- 3. the maturity/tenor (the 5 years usually being the most liquid ones);
- 4. the notional value;
- 5. credit events which trigger the 'insurance' payoff.
- 6. the CDS spread, which is the periodic payment the premium leg completes until either maturity of the CDS contract or until a specified credit event gets triggered, whichever occurs first;

In case of a credit event, the insured party may deliver as many debt contracts, satisfying the type of debt stated within the CDS contract, referencing the specified debtor as long as the sum of face values remains below the agreed notional value.

The CDS thus contains three main parties: the insured party, which is usually called the *premium leg* as it has to pay a periodic premium in order to be insured by the insurer, which is referred to as the *protection leg* as it delivers protection in case of a credit event. Finally it constitutes the *reference entity* which is the party to whom the insured debt references to. This reference entity is however only stated within the contract but doesn't have to approve or be aware of the transaction.

See figure 1.2 for a visualisation of the CDS.



FIGURE 1.2: The CDS scheme: if the reference entity defaults (satisfies at least one of the credit events), then the protection payment, with size of LGD is triggered. As long as no default occurs the premium leg keeps on paying the premium agreed on.

#### 1.2.1 Settlement

Initially the settlement procedure was *physical settlement*: the premium leg would physically deliver the debt to the protection leg. This used to be a simple and quick procedure for both parties. A resulting drawback however was the cheapest to deliver option: the premium leg could deliver their least valuable types of debt satisfying the CDS agreement, as the CDS accepted a variety of slightly different debts. As a consequence the CDSs would be priced as high as the cheapest to deliver option.

At some point in time the CDS became the primary indicator of an entities' creditsworthiness. Resulting in the volume of traded CDSs to outstrip the volume of referencing bonds outstanding: more CDS were traded on a credit than the outstanding bond issuance of that credit[Sap08]. Physical settlement was not appealing, protection buyers would often have to go to the open market in order to buy a certain bond to complete the settlement while protection sellers would be left with cash positions after the credit event. As a solution towards this problem, cash settlement was introduced. However there was no clear mechanism yet for determining the fair value for the protection leg to pay. To this end the credit event auction was introduced: a process in which the value of debt after the credit event is determined. A result of such auction could be for example that the value of some debt is determined on 40 cents on the dollar. For a more comprehensive overview we refer to [Sap08].

Thus, hypothetically speaking, the premium leg would on average be able to sell his debt on the market for the price specified by the credit event auction. As a result, *cash settlement* comes down to a single payment made by the protection leg without an exchange from the premium to the protection leg, where the size of this payment equals

$$(1-\rho)\cdot\overline{N},$$

where  $\rho$  stands for the value determined within the auction ( $\rho = 0.4$  in the above example) and  $\overline{N}$  the agreed notional in the CDS contract.

#### 1.2.2 Credit events

Credit events specified by the CDS describe the events that trigger the protection payment. Such credit events might differ per CDS. As stated by ISDA<sup>2</sup> the possible credit events are[Mar15]

- 1. Failure to pay:
- 2. Bankruptcy:
- 3. Restructuring:
- 4. Repudiation or moratorium:
- 5. Obligation acceleration and obligation default:

Typically, failure to pay, bankruptcy and restructuring are the most common events [Sap08] [Mar15].

First, the most commonly included credit event is failure to pay. Second, bankruptcy is a credit event for corporate reference entities but not for sovereign entities. Third, restructuring, which refers to actions such as coupon reduction or maturity extension. Fourth, repudiation or moratorium provides for compensation after specified actions of a government reference entity and is generally relevant only to emerging market reference entities. And fifth, obligation acceleration and obligation default, which refer to technical defaults such as violation of a bond covenant, are rarely included.

#### 1.2.3 Modeling

When pricing a CDS we refer to determining the fair value of the CDS spread. The CDS consists of two types of payments, namely the periodic payments made by the premium leg and the possible protection payment completed by the protection leg. The CDS spread can thus be found by determining the present value of both the legs, setting these equal at initiation. We will show a common pricing framework, following [BA03].

Consider three firms: A, the premium leg, B, the protection leg, and a reference entity. Firm B will provide firm A with protection over the period  $(T_a, T_b]$  in exchange for a periodic payment C. The payments are transferred at times  $\{T_{a+1}, ..., T_b\}$ , where the time intervals are equally spaced with  $\Delta T$ . If a default of the reference entity occurs at

 $<sup>^2\</sup>mathrm{ISDA}$  (International Swaps and Derivatives Association) provides standardised over the counter contracts.

time  $T_a < \tau \leq T_b$  the protection leg is obliged to complete a protection payment of value  $LGD_{\tau}$ . Let the current time be denoted by t.

#### 1.2.3.1 Premium leg

The premium leg pays the periodic premium, C, to the protection leg as long as no default occurs:

Periodic payments = 
$$\sum_{i=a+1}^{b} D(t, T_i) C \mathbf{1}_{\{\tau \ge T_i\}}.$$
 (1.10)

We wrote  $\tau \geq T_i$ , instead of  $\tau > T_i$ , because besides paying the period payments until a default occurs, the premium leg should also pay an accrual amount whenever a default occurs in between agreed payment times, such that it actually completes a payment for the extra time it was insured. Therefore, also if a default occurs at time  $T_i$ , a last premium should be paid corresponding to being insured within  $(T_{i-1}, T_i]$ . Furthermore, this extra accrual payment in case of a default can be written as

$$Accrual = D(t,\tau) \frac{\tau - T_{\beta(\tau)-1}}{\Delta T} C \mathbf{1}_{\{T_a < \tau < T_b\}},$$
(1.11)

where  $\beta(\tau)$  is the first premium payment time that would follow the default. Equation (1.11) is indeed in agreement with the accrual model presented by Markit<sup>3</sup>, which satisfies the definition provided by ISDA[Mar12].

Thus summing both terms and taking its expectation given *all* knowledge up till time t - as it contains several stochastic components - we get the expression for the premium leg as

$$\operatorname{Prem}\left(t, T_{b}\right) = \mathbb{E}_{t}^{\mathbb{Q}}\left[D(t, \tau) \frac{\tau - T_{\beta(\tau) - 1}}{\Delta T} C \mathbf{1}_{\{T_{a} < \tau < T_{b}\}} + \sum_{i=a+1}^{b} D(t, T_{i}) C \mathbf{1}_{\{\tau \geq T_{i}\}}\right] \quad (1.12)$$

Expression (1.12) is a very general one. However, on the basis of four common assumptions we can come up with an expression which is more often encountered in practice. Assume 1) that  $t = T_a$  and that no default occurred by time t, that 2) discounting is performed on a continuous basis where the short rate is denoted as  $r_t$ , that 3) the default time is the first jump of a Poisson process with intensity  $\lambda_t$  - this  $\lambda_t$  is called the intensity or hazard rate, it may be stochastic in which case the Poisson process becomes a Cox process. More details on the hazard rate and Cox process can be found in Appendix D. Lastly

 $<sup>^3\</sup>mathrm{Markit}$  is a financial company known for providing both independent data and the pricing of credit derivatives.

4) that interest rate and default are independent (i.e. no wrong-way risk)<sup>4</sup>. Using these assumption we proceed:

$$\begin{aligned} \operatorname{Prem}\left(t,T_{b}\right) &= \int_{t}^{T_{b}} \mathbb{E}_{t}^{\mathbb{Q}} \left[ D(t,u) \frac{u - T_{\beta(u)-1}}{\Delta T} C \mathbf{1}_{\{\tau \in (u,u+du]\}} \right] \\ &+ \sum_{i=a+1}^{b} \mathbb{E}_{t}^{\mathbb{Q}} \left[ D(t,T_{i}) C \mathbf{1}_{\{\tau \geq T_{i}\}} \right] \\ &= \int_{t}^{T_{b}} \mathbb{E}_{t}^{\mathbb{Q}} \left[ D(t,u) \frac{u - T_{\beta(u)-1}}{\Delta T} C \right] P_{t}(\tau \in (u, u+du]) \\ &+ \sum_{i=a+1}^{b} \mathbb{E}_{t}^{\mathbb{Q}} \left[ D(t,T_{i}) C \right] P_{t}(\tau \geq T_{i}) \\ &= C \int_{t}^{T_{b}} \mathbb{E}_{t}^{\mathbb{Q}} \left[ e^{-\int_{t}^{u} r_{s} ds} \frac{u - T_{\beta(u)-1}}{\Delta T} \right] \mathbb{E}_{t}^{\mathbb{Q}} \left[ \lambda_{u} e^{-\int_{t}^{u} \lambda_{u} ds} \right] \\ &+ C \sum_{i=a+1}^{b} \mathbb{E}_{t}^{\mathbb{Q}} \left[ e^{-\int_{t}^{T_{i}} r_{s} ds} \right] \mathbb{E}_{t}^{\mathbb{Q}} \left[ e^{-\int_{t}^{T_{i}} r_{s} ds} \right] \\ &= C \int_{t}^{T_{b}} \mathbb{E}_{t}^{\mathbb{Q}} \left[ e^{-\int_{t}^{u} r_{s} ds} \right] \mathbb{E}_{t}^{\mathbb{Q}} \left[ \lambda_{u} e^{-\int_{t}^{u} \lambda_{u} ds} \right] \\ &+ C \sum_{i=a+1}^{b} \mathbb{E}_{t}^{\mathbb{Q}} \left[ e^{-\int_{t}^{T_{i}} r_{s} ds} \right] \mathbb{E}_{t}^{\mathbb{Q}} \left[ e^{-\int_{t}^{T_{i}} \lambda_{s} ds} \right] \\ &+ C \sum_{i=a+1}^{b} \mathbb{E}_{t}^{\mathbb{Q}} \left[ e^{-\int_{t}^{T_{i}} r_{s} ds} \right] \mathbb{E}_{t}^{\mathbb{Q}} \left[ e^{-\int_{t}^{T_{i}} \lambda_{s} ds} \right]. \end{aligned}$$

$$(1.13)$$

The expectation around the terms of  $\lambda_u$  can be removed if  $\lambda$  is deterministic, i.e. a regular Poisson process (not a Cox process).

#### 1.2.3.2 Protection leg

The protection leg only needs to complete a payment of the unknown value  $LGD_{\tau}$  in case a default occurs within  $(T_a, T_b]$ . This unknown value  $LGD_{\tau}$  is called the Loss Given Default (LGD), which is closely related to the recovery, which we will explain in more detail in section 1.3. This leg can be written as

$$\operatorname{Prot}\left(t, T_{b}\right) = \mathbb{E}_{t}^{\mathbb{Q}}\left[D(t, \tau) \operatorname{LGD}_{\tau} \mathbf{1}_{\left\{T_{a} < \tau \leq T_{b}\right\}}\right].$$
(1.14)

Now, again with the assumptions described in the previous section we may proceed as follows

 $<sup>^{4}\</sup>mathrm{To}$  get the expressions allowing for wrong-way risk we should leave in the conditional probabilities given the defaults.

$$\operatorname{Prot}\left(t, T_{b}\right) = \int_{t}^{T_{b}} \mathbb{E}_{t}^{\mathbb{Q}}\left[D(t, u) \operatorname{LGD}_{u} \mathbf{1}_{\{\tau \in (u, u+du]\}}\right]$$
$$= \int_{t}^{T_{b}} \mathbb{E}_{t}^{\mathbb{Q}}\left[D(t, u) \operatorname{LGD}_{u}\right] P_{t}(\tau \in (u, u+du])$$
$$= \int_{t}^{T_{b}} \mathbb{E}_{t}^{\mathbb{Q}}\left[e^{-\int_{t}^{u} r_{s} ds} \operatorname{LGD}_{u}\right] \mathbb{E}_{t}^{\mathbb{Q}}\left[\lambda_{u} e^{-\int_{t}^{u} \lambda_{u} ds}\right] du \qquad (1.15)$$

#### 1.2.3.3 The CDS spread

At time t we wish to agree upon a fair CDS spread. In order to determine a fair spread both legs should be equal at initiation. In the more general case, that is taking equations (1.14) and (1.12), we'd get

$$\mathbb{E}_{t}^{\mathbb{Q}}\left[D(t,\tau)\frac{\tau-T_{\beta(\tau)-1}}{\Delta T}C\mathbf{1}_{\{T_{a}<\tau< T_{b}\}}+\sum_{T_{a+1}}^{b}D(t,T_{i})C\right]=\mathbb{E}_{t}^{\mathbb{Q}}\left[D(t,\tau)\mathrm{LGD}\mathbf{1}_{\{T_{a}<\tau\leq T_{b}\}}\right],$$
(1.16)

As the CDS spread is a constant value it can be taken out of the equation giving

$$C = \frac{\mathbb{E}_t^{\mathbb{Q}} \left[ D(t,\tau) \operatorname{LGD}_{\tau} \mathbf{1}_{\{T_a < \tau \le T_b\}} \right]}{\mathbb{E}_t^{\mathbb{Q}} \left[ D(t,\tau) \frac{\tau - T_{\beta(\tau)-1}}{\Delta T} \mathbf{1}_{\{T_a < \tau < T_b\}} + \sum_{T_{a+1}}^b D(t,T_i) \right]}.$$
(1.17)

Note that the numerator depends implicitly upon the notional of the contract, which is hidden within the LGD LGD.

For the more practical case, we take (1.13) and (1.15) to find a CDS spread of

$$C = \frac{\mathbb{E}_{t}^{\mathbb{Q}} \left[ \int_{0}^{T_{b}} \mathrm{L}\mathrm{GD}_{u} \lambda_{u} e^{-\int_{0}^{u} (r_{s} - \lambda_{s}) ds} du \right]}{\int_{0}^{T_{b}} \mathbb{E}_{t}^{\mathbb{Q}} \left[ e^{-\int_{t}^{u} (r_{s} + \lambda_{s}) ds} \lambda_{u} \right] \frac{u - T_{\beta(u) - 1}}{\Delta T} du + \sum_{T_{a+1}}^{b} \mathbb{E}_{t}^{\mathbb{Q}} \left[ e^{-\int_{t}^{T_{i}} (r_{s} + \lambda_{s}) ds} \right]}.$$
 (1.18)

Here we would already like to point out that the spread is now denoted in terms of the currency it is being sold in, e.g. dollars. However, it is usually depicted in basis points: the percentage of the spread with respect to the notional, multiplied by 100. To do so we evaluate both sides in fractions of the notional, such that the LGD above becomes a fraction of the notional, which we write as LGD (a small capital letter l), which we call the *percentage LGD*. Assuming the LGD to be a fraction of the notional is not necessary, there exist many other assumptions which modify the protection leg. We will look more closely into these assumptions in section 1.3.3.1.

#### 1.2.3.4 Piecewise linear intensity

The piecewise linear intensity model is a frequently used model in both literature and practice (actually Bloomberg uses a similar model to determine the probability of default[FNS09]). We follow the method described by the authors of [BM06]. Suppose spreads are given at maturities  $T_i$ , i = 1, 2, ..., N, then we wish to determine the deterministic intensities  $\lambda(t) = \lambda_i$ , for  $t \in (T_{i-1}, T_i]$ . As is market convention we fix the recovery at 40% of face value, i.e. LGD = 0.6. Then we determine premium and protection legs at time t = 0

$$\operatorname{Prot}\left(0,T_{i}\right) = \operatorname{LGD} \cdot \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_{0}^{\tau} r_{s} ds} \mathbf{1}_{\{\tau \leq T_{i}\}}\right]$$
$$= \operatorname{LGD} \cdot \int_{0}^{T_{i}} \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_{0}^{u} r_{s} ds}\right] P(\tau \in (u, u + du])$$

In a similar fashion we determine the premium leg supposing the spread,  $C_i$ , to be continuously paid up till time  $T_i$ , given no default, by doing so we may neglect the accrual term. We then get

$$\operatorname{Prem}\left(0,T_{i}\right) = \mathbb{E}^{\mathbb{Q}}\left[\int_{0}^{T_{i}} e^{-\int_{0}^{u} r_{s} ds} C_{i} \mathbf{1}_{\{\tau > u\}} du\right]$$
$$= C_{i} \int_{0}^{T_{i}} \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_{0}^{u} r_{s} ds} \mathbf{1}_{\{\tau > u\}}\right] du$$
$$= C_{i} \int_{0}^{T_{i}} \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_{0}^{u} r_{s} ds}\right] P(\tau > u) du$$

Each intensity  $\lambda_i$  can be determined by solving for the CDS spread

$$C_{i} = \frac{\operatorname{LGD} \cdot \int_{0}^{T_{i}} \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_{0}^{u} r_{s} ds} \right] P(\tau \in (u, u + du])}{\int_{0}^{T_{i}} \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_{0}^{u} r_{s} ds} \right] P(\tau > u) du}.$$

We start by solving  $\lambda_1$ , the intensity over the first year, subsequently for  $\lambda_2$ , etc. To do so, we have to discretize both legs such that this may be evaluated by the computer. Let  $\Delta t = \frac{1}{N}$ , such that  $t_k = k\Delta t$  and  $t_N = 1$ . As the intensity process is modeled by a Poisson process we know that

$$P(\tau \in (t_k, t_{k+1}]) \approx \lambda_{\left\lceil \frac{k+1}{N} \right\rceil} e^{-\sum_{n=1}^k \lambda_{\left\lceil \frac{n}{N} \right\rceil} \Delta t} \Delta t_k$$
$$P(\tau > t_k) \approx e^{-\sum_{n=1}^k \lambda_{\left\lceil \frac{n}{N} \right\rceil} \Delta t}.$$

Note that the intensities are piecewise constant functions such that within a year they remain constant, which is the reason why we round the indexes of the intensities upward. Using this discretization and assuming the interest rates to be known values,  $r_k$  the interest rate between  $(t_{k-1}, t_k]$ , we can determine the spread by

$$C_{i} = \frac{\operatorname{LGD} \cdot \sum_{l=1}^{T_{i} \cdot N} e^{-\sum_{n=1}^{l} r_{n} \Delta t} \lambda_{\lceil \frac{l+1}{N} \rceil} e^{-\sum_{n=1}^{l} \lambda_{\lceil \frac{n}{N} \rceil} \Delta t} \Delta t}{\sum_{l=1}^{T_{i} \cdot N} e^{-\sum_{n=1}^{l} r_{n} \Delta t} e^{-\sum_{n=1}^{l} \lambda_{\lceil \frac{n}{N} \rceil} \Delta t}}.$$

Using this formula we can determine the intensities that fit the market spreads. At the end of this thesis we use this model to compare it to ours. We shall refer to this model as the PC model.

#### 1.3 Recovery

Suppose firms<sup>5</sup> A and B to have a contract in which A lends money to B where B is expected to pay back the debt at a later point in time, say T. However, over time the financial conditions of B might decline and even get so bad that B will no longer be able to pay back its debt and enters bankruptcy<sup>6</sup> at time  $\tau \leq T$ . This doesn't mean that B has absolutely zero money, it means that its liabilities have become larger than its assets. In case of a straight bankruptcy<sup>7</sup> of party B, the leftover assets will be liquidated in order to raise money and pay back the creditors - to which A belongs - as good as possible. This is where recovery comes in. In this specific example, recovery is the money paid back to A. Most of the times this is less than the original debt<sup>8</sup>.

More generally, recovery refers to the amount that the creditor eventually receives on its contract in case of a default of the debtor, where default is the failure to meet the legal obligations of the debt repayment.

Note that we don't state the recovery to be paid by party B: there is a large market associated to buying defaulted contracts. The payment the creditor receives on its defaulted contract, might therefore just as well be a consequence of selling the contract within this market. The valuation of a defaulted contract can be understood as the expected recovery that will finally be made by party B.

Just as was the case for the LGD, the recovery is often stated as a percentage, say 40%, such that 60% has been lost, i.e.  $LGD = 1 - \rho$ . Usually this means that 40% of the face value of the contract has been recovered at default. However, special attention should be paid at the definition of the percentage recovery as it might be defined differently: examples of such are percentage of market value or percentage of present value of face (we will get back to this in section 1.3.3.1).

In a frictionless world, we would expect the recovery to be close to 100%. This would be a result of solvency as long as the firm's assets to liabilities remain above 1, and a direct default when it gets insolvent. However, in practice recoveries of 40% are not exceptional. The reasons for this are frictions: The first one simply being a delay between insolvency and the default. Another well-known friction is related to jumps, these reflect sudden changes of investor's views, influencing their investment patterns 'unexpectedly'. As a last friction we state the bankruptcy costs[CH04].

 $<sup>{}^{5}</sup>A$  and B not necessarily have to be firms, in case of a mortgage, party A is a bank and B can be an individual.

<sup>&</sup>lt;sup>6</sup>This not necessarily has to be a bankruptcy, it might just as well be another form of default.

<sup>&</sup>lt;sup>7</sup>The Chapter 7 bankruptcy is also known as a straight bankruptcy which implies basic liquidation.

<sup>&</sup>lt;sup>8</sup>Situations have occurred in which the recovery was more then the original debt.

#### 1.3.1 Physical recovery

By physical recovery - recovery under the  $\mathbb{P}$ -measure - we refer to realised recovery. Realised recoveries can only be determined in case of a default. To be more precise we distinguish two different types of realised recoveries

- 1. Realised recovery based upon the recovery payment made by the defaulted debtor.
- 2. Realised recovery based on the value of the defaulted contract.

The first measure, called by [ARS05] as the ultimate recovery, used to be the standard before the existence of the secondary loan trading market. A drawback of this measure is the fact that it often takes a very long time before recovery rates can be determined: it usually takes the whole legal process of bankruptcy or restructuring to be finalised. At the start of the 21st century the CDS market started to grow at such speed that at some point more CDSs where being sold than underlying bonds existed, which made the physical-settlement of CDSs especially complicated, giving rise to the cash-settled CDS. However, cash settlement could not wait for legal processes to be completed. The market's answer to this problem was the CDS auction[Sap08]. This auction determines the value of the defaulted contract in the secondary market, corresponding to the second measure. This measure is now regarded as the standard measure for the recovery [ARS05].

Furthermore, note that the second measure might be understood as an expectation of the secondary market of the first measure. Hence, applying underlying determinants of the first measure to estimate the second measure seems valid. The models we describe later might for example implement Absolute Priority Rule (APR), which is not something that has a direct relation to the value of the defaulted contract, but it is to the ultimate recovery.

#### 1.3.1.1 Determinants

There are four main determinant classes that influence the recovery: firm-specific characteristics, industry-specific characteristics, macro-economic effects and contract-specific properties. It is important though to emphasise that this is a discussion of recoveries extracted from historical data, both at default<sup>9</sup> and at emergence. As determinants of physical recoveries are expected to influence implied recoveries[SUH14], it is in our interest to look more closely at these.

<sup>&</sup>lt;sup>9</sup>Recovery at default is obtained as the market value of defaulted securities shortly after default.

**Firm** Under firm-specific characteristics one might think about capital structure (see appendix C), leverage ratio, the percentage of tangible assets and liquid assets. All of these directly influence the liquidation process of the firm and, hence, the expected price of debt securities in distressed situations.

Leverage, roughly said, an indicator for the percentage of a firm that is financed by debt, is often measured as the amount of long term debt to assets ratio. It may be associated with greater dispersed ownership requiring greater coordination among bargaining parties [ABS07], resulting in more complex bankruptcy processes, associated with higher costs [CH04]. However, on several occasions, highly leveraged transactions were easily restructured<sup>10</sup> [ABS07].

Another important factor is liquidity, an indicator for how fast a firm can turn assets into cash. There are many measures for liquidity, e.g. the interest coverage ratio, which determines how easily a company can pay of its debt interest. The assets related to covering these interest expenses generate relative high earnings, and are therefore likely to give a higher recovery[SUH14][ABS07] - the authors of [ABS07] scale it under profitability. Another indicator towards liquidity is the quick ratio, an indicator for the ability of a firm to quickly pay of its liabilities without having to stop functioning.

A company holding more tangible assets is likely to produce higher recovery rates in case of a default. The percentage of tangible assets is measured as hard assets (including property, plants, equipment, etc.) to total assets. A higher recovery can be expected as these assets can be sold directly in case of a liquidation. In [CV05] it was found that firms with a high proportion of tangible assets recover roughly 25% more than firms with a low proportion of tangible assets.

**Industry** By the authors of [AK96] it is reasoned that industry affiliation is likely to be important because the business activity of an enterprise dictates the types of assets and the competitive conditions of firms within different industrial sectors. They study the severity of bond defaults stratified by the industrial sector (using SIC<sup>11</sup> codes) and debt seniority. It is shown that, between 1971 and 1995, recoveries from 'public utilities' averaged at \$70.48 and for 'Chemicals, petroleum, rubber and plastic products at \$62.73, per \$100,- of notional. Furthermore they show that the difference between those sectors and the others are statistically significant, even when adjusting for seniorities.

A primary measure for poor industry conditions is industry distress at the time of default. It is shown to be statistically significant [CV05]. However, an important factor to consider is asset specificity: specific assets are related to for example machinery and equipment, whereas cash and non-industrial real estate are considered non-specific assets. A lower asset specificity makes firms less sensitive to industry distress. This can be understood

<sup>&</sup>lt;sup>10</sup>Restructuring is the action of altering conditions of the contract such that they become more manageable for de debtor. This gives them the opportunity to recover from the distress.

<sup>&</sup>lt;sup>11</sup>SIC (Standard Industrial Classification) is a system for classifying industries by four-digit codes. The first two digit represent the major group and the first three their industry.

due to fire-sales<sup>12</sup>. Actually, the authors of [CV05] even show that the recovery for firms with a very low asset specificity, seem to increase in case of industry distress. Furthermore, variables such as industry liquidity and industry leverage might function as indicators for the condition of the issuer's peer firms. These are expected to negatively influence the liquidation process of a defaulted firm. Industry liquidity however seems more significant at emergence than at default[ABS07]. A higher industry leverage could function as a proxy for difficulty of raising new financing: High leveraged firms may be closer to being financially constrained[ABS07].

**Macro** On a macro economic capital level it seems obvious that forces which increase the default rates will decrease the value of assets of distressed companies, resulting in lower recoveries. The authors of [ABRS03] show in a large scale statistical research, that for the defaulted bond market, the correlation coefficient between default rates and recovery calculates -0.72. Similar results are also provided by the authors of [ABS07] and [CV05] Interestingly it is shown by the authors of [ABS07] that macro economic variables have no statistical significance in the presence of a variable accounting for industry distress. However, when removing this variable, both the GDP and the BDA<sup>13</sup> become significant at a level of 10%. Here we thus see how we should be cautious with blindly accepting results of regressions as explanatory variables might be dependent ones.

**Contract** On a contract specific level, as a result of the APR, the seniority plays a very important role[ABS07]. On a sample of 700 defaulted bonds the seniority was found to play an important role on recoveries: senior secured debt averaged at about 58 percent of face value; senior unsecured, 48 percent; senior subordinate, 34 percent; and junior subordinate, about 31 percent [AK96].

Furthermore [JNS14] highlights the importance of liquidity of the security. Which is also a main result in [ABRS03], which shows that recovery rates are a function of supply and demand for the underlying security.

Two remarks are in place: 1) The influences of the determinants as described above do not imply that a model including all of these will be most optimal, as we have seen how correlations play an important role. 2) All of the above results are deduced from econometric regressions that tend to be sensitive to the choice of explanatory variables and the regression methods.

 $<sup>^{12}</sup>$ Fire-sales is the act of selling goods and assets at heavily discounted prices. This is often found to occur just before a default. Thus indeed, due to the fire-sales a firm with low asset specificity will not be affected much by industry distress.

<sup>&</sup>lt;sup>13</sup>BDA is defined as the total face value amount of defaulted bonds in a year measured at mid-year and in trillions of dollars.

#### 1.3.2 Implied recovery

In the previous section we explained physical recovery, so it seems natural to proceed with the recovery under the  $\mathbb{Q}$ -measure, i.e, *implied* recovery. In section 1.1 we discussed in a general sense that differences exist between the  $\mathbb{P}$ - and  $\mathbb{Q}$ -measure. Such differences are indeed noted when comparing the implied and physical recovery. Let's start by understanding the term implied.

Mechanisms such as, for example, recovery and default are understood by their definition and in this sense can be incorporated in a model as such. However, understanding of the functioning of such a mechanism does not yet provide us with the required input parameters. For example, we might understand that at time T a default occurs with some probability p and if it does a recovery of R is related to it, whereas in case of no default we simply get face value F. Therefore the discounted value of this product at time t equals

$$V_t = D(t,T) \left(F + p \cdot R\right), \qquad (1.19)$$

where D(t,T) is the discounting term over (t,T). We thus have a model for the value of the product, but we cannot yet evaluate it as p and R are still unknown.

The process of configuring these parameters in general is called calibration. If this is done by equaling the market prices with the model prices to extract the variables, then the variables are called their implied counterpart, thus as in the example, respectively implied default probability and implied recovery.

On the other hand if we'd use historical values for the default probability and recovery, then we understand from the discussion in section 1.1 that our prices will likely not fit to those found in the market, hence a difference between both measures will be observed.

Summarizing, an implied variable is the configuration of the parameters based upon the state of the market prices and upon a model.

#### **1.3.2.1** Differences to the physical recovery

In section 1.1.3 we discussed the difference between the  $\mathbb{P}$ - and  $\mathbb{Q}$ -measure. This understanding can be applied to see how physical and implied recovery might differ. We have seen in section 1.3.1.1 how both industry specific and macro economical factors can influence the physical recovery. Therefore, the expectations of investors on the industries' health might very well be reflected within the implied recovery. The implied recovery is determined by two components:



FIGURE 1.3: The mean and standard deviation of global historical and implied recovery. It is clear how mean implied recovery is lower, whereas its standard deviation is higher, implying uncertainty. Source: [SUH14].

- 1. Underlying physical determinants, which influence the investors expectation with respect to resulting recovery.
- 2. The expectations on unhedgeable risk components (such liquidity and systematic risk) result in a price adjustment, hence influencing the underlying components, including the implied recovery.

Indeed, the authors of [SUH14] use an approach to isolate the recovery such that they can estimate its mean, finding the results presented in figure 1.3. As expected, it clearly shows the implied recovery to fall below the physical counterpart.

In order to justify systematic risk to be an important factor of this difference, they subsequently show how uncertainty - represented by the standard deviation - increases in times of distress, which is exactly what one may expect to happen in a risky period. Besides such increase in standard deviation also a decrease of the implied recovery is observed, further strengthening the argument of negative correlation between default and recovery. Finally they also argue that there might be another premium related to the markets' lack of understanding of recovery risk.

Summarizing, we understand that a difference between implied and historical variables is common. However, especially with respect to implied recovery it is significantly different from its historical counterpart. This is a consequence of both systematic risk - recoveries and industry distress are closely related - and a lack of understanding of the recovery, creating hedging difficulties.
#### 1.3.2.2 Difficulties

As default and recovery always occur mutually, there is a well known separation problem - this will be discussed more specifically in section 1.3.3.2 - making independent identification of these elements a complicated issue. Due to this difficulty it is market convention to fix the recovery at around 40% for senior contracts, and at about 20% for subordinated contracts and sovereigns [EY13][Iwa13]. This is not a random choice, for decades, average prices at default have been calculated at approximately 40 cents on the dollar. Indeed, the authors of [AK96] show that the arithmetic average recovery rate on a sample of more than 700 defaulted senior corporate bond issues was found to be \$41.70 per \$100 of face value.

In the context of credit management we could argue this market convention to be sufficient, as observed realised recoveries might provide a decent estimate of future realisations. However in the case of pricing we should be looking at implied recovery. Observing the literature on implied recovery we soon understand that assuming such constant implied recovery is not the most logical choice, although the simplest. We state three main arguments:

- 1. Historical recovery has shown to be very dependent upon many determinants, such as contract seniority and firm specific characteristics. Averages may differ significantly[AK96].
- 2. Taking the historical average recovery as implied recovery might be an overestimation due to incompleteness of the market, and the fact that recovery risk is not very well understood. Furthermore, it was shown that systematic risk seems to be an important underlying risk factor of the recovery[SUH14].
- 3. Assuming the implied recovery to be constant neglects the correlation between the default process and the recovery. A strong negative correlation between both has frequently been shown to exist [ABS07], [ABRS03], [CV05].

The importance of a better estimate of the implied recovery is now understood and slowly more research is being performed within this area.

## 1.3.3 Modeling implied recovery

#### 1.3.3.1 Assumptions on the recovery

Usually recovery, R, and loss given default, LGD, are stated as percentages. However stating the recovery to be 40% leaves open what it is referencing to. The percentage recovery,  $\rho$ , might be a percentage of whatever is being referred to. Changing this reference might provide a better suiting economic interpretation or a more convenient pricing framework. This will become more clear in what follows. Suppose  $\tau$  to be the time of default,  $R(\tau)$  the corresponding recovery with  $\rho(\tau)$  the percentage recovery, F the face value of a defaultable debt maturing at T and t indicating today. Following the authors of [BMZ01] we describe the three assumptions typically made about the recovery:

i **Fractional recovery of face value (RFV)**: The recovery is a fraction of the face value of the corresponding contract upon default,

$$R(\tau) = \rho(\tau)F. \tag{1.20}$$

Besides the simplicity of the assumption, [Guh02] shows that it is economically reasonable as bonds of the same issuer and same seniority are valued equally, independent of their time (left) to maturity. This suggests that dependence is only upon the face value. Furthermore, the simplicity persists when determining a pricing framework, especially when the recovery is assumed to be a constant. In Appendix G the general defaultable bond pricing formula is deduced, such that a zero coupon bond with maturity T looks like

$$V_t = \mathbb{E}_t^{\mathbb{Q}} \left[ e^{-\int_t^T (r_s + \lambda_s) ds} F \right] + \mathbb{E}_t^{\mathbb{Q}} \left[ \int_t^T \lambda_u e^{-\int_t^u (r_s + \lambda_s) ds} R(u) du \right].$$
(1.21)

Suppose for simplicity that the face value is 1, F = 1, and that the recovery is a known constant (as in practice is often assumed), then the general defaultable bond price reduces to

$$V_t^{RFV} = \mathbb{E}_t^{\mathbb{Q}} \left[ e^{-\int_t^T (r_s + \lambda_s) ds} \right] + \rho \ \mathbb{E}_t^{\mathbb{Q}} \left[ \int_t^T \lambda_u e^{-\int_t^u (r_s + \lambda_s) ds} du \right].$$
(1.22)

ii **Fractional recovery of market value (RMV)**: One assumes that upon default a contract looses a given fraction of its value just prior to default.

$$R(\tau) = \rho(\tau)V(\tau-), \qquad (1.23)$$

where  $V(\tau -)$  denotes the market value of the debt just before default  $\tau$ .

The elegance of this assumption is mainly found in the resulting simplicity of the framework: let  $\lambda_t$  the default intensity at time t. Then if X is a defaultable claim at maturity T, the authors of [DS99] show that the price of this claim at time t satisfies

$$V_t^{RMV} = \mathbb{E}_t^{\mathbb{Q}} \left[ e^{-\int_t^T (r_s + \lambda_s (1 - \rho_s)) ds} X \right], \qquad (1.24)$$

where  $r_s$  equals the interest rate process and X = 1 for a zero coupon bond. Indeed the simplicity of the framework comes from the fact that we may define  $\phi_t = r_t + \lambda_t (1 - \rho_t)$ and price a defaultable claim only by modelling this function. A difficulty that is directly apparent is however the fact that the default intensity and recovery are by no means separately identifiable given a single term structure. Furthermore, the authors of [DS99] show that while varying the LGD, the result of  $\lambda_t$ , the default intensity, does not seem to change depending upon the choice of either Recovery of Face Value (RFV) or Recovery of Market Value (RMV). Suggesting that the market value just before default is still close to par. This would suggest both assumptions to be fairly reasonable.

iii Fractional recovery of the present value of debt (RT): In this approach one assumes that upon default, debt is valued at a given fraction of the present value with respect to a treasury bond.

$$R(\tau) = \rho(\tau)B(\tau,T)F, \qquad (1.25)$$

where  $B(\tau, T)$  equals the value of a treasury bond at time  $\tau$  maturing at T. The authors of [BMZ01] state this assumption to present the best fit to their data. Suggesting market participants not to anticipate immediate recovery of face.

The assumption does not provide a pricing framework as easy as the RFV, but it has a convenient form which is exploited by for example the authors of [SUH14]. Substituting the assumption in (1.21) with face value of 1 gives

$$V_t^{RT} = \mathbb{E}_t^{\mathbb{Q}} \left[ e^{-\int_t^T (r_s + \lambda_s) ds} \right] + \mathbb{E}_t^{\mathbb{Q}} \left[ \int_t^T \lambda_u e^{-\int_t^u (r_s + \lambda_s) ds} B(u, T) \rho(u) du \right].$$
(1.26)

Recall that  $B(u,T) = \mathbb{E}_u^{\mathbb{Q}}[D(u,T)] = \mathbb{E}_u^{\mathbb{Q}}\left[e^{-\int_u^T r_s ds}\right]$  from which it follows that

$$\mathbb{E}_{t}^{\mathbb{Q}}\left[\int_{t}^{T}\lambda_{u}e^{-\int_{t}^{u}(r_{s}+\lambda_{s})ds}B(u,T)\rho(u)du\right] = \mathbb{E}_{t}^{\mathbb{Q}}\left[\int_{t}^{T}\lambda_{u}e^{-\int_{t}^{u}(r_{s}+\lambda_{s})ds}\mathbb{E}_{u}^{\mathbb{Q}}\left[e^{-\int_{u}^{T}r_{s}ds}\right]\rho(u)du\right]$$

$$(1.27)$$

$$= \mathbb{E}_{t}^{\mathbb{Q}}\left[\int_{t}^{T}\lambda_{u}e^{-\int_{t}^{u}(r_{s}+\lambda_{s})ds}e^{-\int_{u}^{T}r_{s}ds}\rho(u)du\right]$$

$$(1.28)$$

$$= \mathbb{E}_{t}^{\mathbb{Q}}\left[e^{-\int_{t}^{T}r_{s}ds}\int_{t}^{T}\lambda_{u}e^{-\int_{t}^{u}\lambda_{s}ds}\rho(u)du\right]$$

$$(1.29)$$

$$= \mathbb{E}_t^{\mathbb{Q}^T} \left[ \int_t^T \lambda_u e^{-\int_t^u \lambda_s ds} \rho(u) du \right], \qquad (1.30)$$

where the equality in (1.28) follows from the tower property of conditional expectations and the last equality is a consequence of the T-forward measure (see Appendix B). The defaultable bond pricing formula now looks like

$$V_t^{RT} = \mathbb{E}_t^{\mathbb{Q}} \left[ e^{-\int_t^T (r_s + \lambda_s) ds} \right] + \mathbb{E}_t^{\mathbb{Q}^T} \left[ \int_t^T \lambda_u e^{-\int_t^u \lambda_s ds} \rho(u) du \right].$$
(1.31)

The above are just three common assumptions upon the recovery. In the literature one might find more exotic assumptions, such as the assumption that recovery is a percentage of the future payoffs of a product, e.g. for a coupon paying bond this would take into consideration the future coupon payments and the final notional payment.

Our model (which will be presented in chapter 3) will use the RFV assumption due to both its simplicity and the corresponding reasonable economic interpretation.

#### 1.3.3.2 The identification problem

A common misconception<sup>14</sup> among academics and practitioners is that fixing the recovery at a specific value is necessary to achieve econometric identification<sup>15</sup>.

We have seen that under RMV assumption a zero coupon is priced by the formula

$$V_{n,t}^{RMV} = \mathbb{E}_t^{\mathbb{Q}} \left[ e^{-\int_t^T (r_s + \lambda_s \text{LGD}) ds} \right].$$
(1.32)

Most notably is how the LGD and the default intensity are directly multiplicative linked. We can therefore understand that given a single term structure with n prices we can at best only extract a similar number of parameters. Market practice therefore fixes the LGD such that the intensity process can be calibrated to the market data.

The RFV on the other hand assumes  $1 - LGD_t$  to be a percentage of face value, where the same such bond is priced as

$$V_{n,t}^{RFV} = \mathbb{E}_t^{\mathbb{Q}} \left[ e^{-\int_t^T (r_s + \lambda_s) ds} \right] + \int_t^T (1 - \text{LGD}) \mathbb{E}_t^{\mathbb{Q}} \left[ \lambda_s e^{-\int_t^s (r_u + \lambda_u) du} \right] ds.$$
(1.33)

<sup>15</sup>Econometric identification means that for each solution of a model a unique parameter set is available. Suppose for example the function  $y = a \cdot b$ , fixing the value of y, will not yet identify the values a and b.

<sup>&</sup>lt;sup>14</sup>The authors in [DS99] state regarding their RMV model: "[..] under the assumption of exogenous  $(\lambda, \text{LGD}, r)$ , the value of a non-callable corporate bond is simply the sum of the present values of the promised payments. It follows that knowledge of defaultable bond prices (before default alone is not sufficient to separately identify  $\lambda$  and LGD. At most, we can extract information about the risk-neutral mean loss rate  $\lambda_{t}\text{LGD}_{t}$  [..]". We thus understand from this that identification given the bond's term structure is not impossible, however it is within the RMV framework . Indeed, [DS99] is often misused as reference to state that recovery and default are not separately identifiable (see for example [DH09]). To make up for this general misconception, Singleton states in [PS08]: "A common impression among academics and practitioners alike is that fixing LGD<sub>t</sub> at a specific value is necessary to achieve econometric identification. This is certainly true in an economic environment in which contracts are priced under the RMV convention".

In contrast to (1.32)  $\lambda_t$  and LGD do not enter in the same multiplicative manner under the RFV assumption. Now, separate identification does seem possible, as a change in the LGD changes the dynamics of the rest of the model which then has to adjust. Indeed the authors of [PS08] proceed by showing how they extract a constant recovery using CDS spreads. However, the intensity is modeled by a simple Vasicek model and the interest rate is also assumed to be constant. They were bounded to such simplicity given the scarce amount of market data compared to the complexity of the model.

Given the difficulties of extracting a constant recovery we can understand that finding whole term structures for both the intensity and recovery on a stochastic basis is almost impossible. To the best of our understanding no paper presenting stochastic models for *both* the hazard and recovery rate has successfully calibrated its model. The authors of [Chr05] come closest to such calibration by assuming RFV and introducing a three factor model in which interest rate, default and recovery are simultaneously modeled following Vasicek processes. However the model assumes independence between all three and tractability of the model was found to be unstable.

In our model this problem is solved by assuming a particular relationship between intensity and recovery. Our choice towards such model becomes clear from the next section where we describe the different categories of recovery models.

#### 1.3.3.3 A brief overview of current models for implied recovery

Due to the difficulty of identifying recovery, research upon implied recovery is relatively scarce - especially comparing it to the amount of research on the default process. Research into estimating the implied recovery requires assumptions that can generally be categorized in one of the following groups[SUH14]:

- 1. Assuming a constant recovery;
- 2. Assuming a predefined relation between recovery and default;
- 3. Assuming independence between recovery and default.

As an example of the first case we might look at [PS08]. In this article the recovery is modeled as a constant over time and the default intensity is assumed to follow a lognormal mean reverting process,

$$d\ln(\lambda_t) = \kappa(\theta - \ln(\lambda_t))dt + \sigma dB_t.$$
(1.34)

The CDS price is modelled under the RFV framework<sup>16</sup> in which recovery is identifiable. Subsequently they obtain sovereign CDS prices for Mexico, Turkey and Korea over a period spanning 2001-2006 for different maturities. By means of the maximum likelihood

 $<sup>^{16}\</sup>mathrm{See}$  the section 1.3.3.1

method they determine the best fitting parameters for the recovery and the default intensity. Market convention sets sovereign recoveries at around 20%, however estimates of almost 77% where found. This could be due to many reasons, such as turbulent period within the sample, or simply that the default process doesn't seem to capture the distribution of CDS spreads.

In the second category a relationship between probability of default and recovery is considered. The authors of [JM12] present a model where the properties of affine processes are exploited. They start out with the CDS price as in (1.18), where they assume the LGD to be a percentage of the notional, i.e. RFV, and defining

$$LGD_t = \beta_2 + \beta_0 e^{\beta_1 \lambda_t}, \quad LGD_t \in (0, 1).$$

$$(1.35)$$

By defining so and assuming the interest rate to be a constant and therefore fully neglecting it, they find the protection leg of the CDS to be

$$\int_{t}^{T} \mathbb{E}_{t}^{\mathbb{Q}} \left[ e^{-\int_{t}^{u} \lambda_{s} ds} \lambda_{u} \left( 1 - \beta_{2} - \beta_{0} e^{-\int_{t}^{u} \lambda_{s} ds} \right) \right] du.$$
(1.36)

Its convenience lies in the fact that the moment generating functions of both exponential functions have closed formed solutions including Riccati equations under the assumption of a CIR process. On a dataset of over 4.5 years between January 2004 until May 2008 of daily CDS quotes on ConocoPhilips they performed a Markov chain Monte Carlo algorithm in order to estimate the model parameters for both the stochastic recovery and constant recovery case. Notable is the significant negativity of  $\beta_1$  they stochastic model implies, implying the stochastic recovery model to be strongly supported by the data.

Also, the model presented by [DH09] follows a same sort of assumption in which the recovery and default depend both upon the underlying stock. We will look more deeply into this model in chapter 2. It got our preference above the model from [JM12] due to its simplicity and the extra dimension of freedom it gives us as the recovery is not restricted to one specific function in order for it to be solvable.

The third class of recovery models assumes independence between the default and recovery process. The authors of [UMG03] describe a model that finds a function of risky bond prices, referencing different seniorities, resulting in a function of recovery only. More specifically they deduce that

$$\frac{V_{sr}(T-t) - V_{jr}(T-t)}{P(t,T) - V_{jr}(T-t)} = \frac{\mathbb{E}[\rho_{sr}] - \mathbb{E}[\rho_{jr}]}{1 - \mathbb{E}[\rho_{jr}]},$$
(1.37)

where  $V_{sr}(T-t)$  and  $V_{jr}(T-t)$  represent risky bond prices for respectively senior and junior claims, with a time to maturity of T-t. Furthermore the  $\rho_{sr}$  and  $\rho_{jr}$  are the recovery for respectively senior and junior claims. The right hand side still contains two different types of recoveries, however using the capital structure the density function of the underlying firm-wide recovery can be traced back. Results unfortunately show the model not to fit very nicely, which might be related to the model being too flexible as it introduces extra parameters to account for violation of the APR.

Another model, worth mentioning, is the one presented by the authors of [SUH14]. It doesn't make any of the regular assumptions and is able to present nice results on the mean and variance of the recovery. Calibrating to CDS spreads, it uses a similar approach as in [UMG03], but is able to allow for dependence (between default and recovery) by moving into the T-forward measure (see appendix B).

#### 1.3.3.4 Modeling criteria

We may now proceed to state our modeling criteria based upon our current knowledge of the implied recovery and find a model that meets these particular requirements. Let's start by stating the requirements of the model:

- 1. It should be able to cope with different seniorities;
- 2. It has to allow negative correlation between default probability and recovery;
- 3. It has to provide a term structure.

Let's clarify these requirements. The first requirement comes from the observed fact that the seniority of a contract highly influences the physical recovery. Therefore, calibration of a model might be improved if the model can simultaneously calibrate to different term structures related to seniorities.

The second requirement also results from historical observations. Negative correlation has repeatedly been shown to hold, we would therefore like to observe this in our results.

Finally the last requirement is related to the expectations of general risks (as discussed in both section 1.1.3 and 1.3.2.1), and the need to provide decent estimates for time dependent valuations, such as Credit Value Adjustment (CVA). The market's expectations of the recovery might change over time and consequently modify the valuation of CVA. For such valuations, term structures might provide better estimates.

With these requirements at hand we recall the discussion in section 1.3.3.3 on a brief overview of current models for the implied recovery. We have shown that such models usually are based on either a constant recovery, a predefined relation between recovery and default and/or assuming independence between recovery and default. Now, clearly the first and last types do not meet our criteria, such that we are left assuming a predefined relation between both recovery and default.

To this end we choose to study more carefully the model described by the authors of [DH09]. This model satisfies all of our requirements and shall be described in more detail in chapter 2.

**Remark 1.3.1.** One may note we have referred quite some times to the model described in [SUH14], this model does not fall in any of these groups, however it has the following drawbacks:

- 1. It can only provide us with the expected recovery under the T-forward measure;
- 2. The expected recovery is observed only for maturities at which a CDS quote is provided, i.e. providing us an expected recovery depending on the maturity of the contract but neglecting the possible movements of the recovery before maturity. For example, if the expectation is found to be 0.4, it might very well be straight line between 0.6 at initiation and 0.2 at maturity;
- 3. Only the expected recovery would not be enough in the scope of more complex and time dependent calculations required for formulas such as CVA.

# Discrete term structures: The D&H model

"You cant go by common sense, if we could do things by common sense we wouldn't need physicists."

– Richard Dawkins , The Unbelievers

From our discussion on the recovery modelling criteria we decided to take a closer look at the model described by Das & Hanouna in their paper 'Implied Recovery' [DH09]. It is able to provide us with a term structure for both the recovery and the default probability, it includes the possibility of correlation between both and can finally simultaneously be calibrated to multiple seniorities.

To understand the principles of their model we shall first describe it in more detail. Next in section 2.4 an important restriction of the method will be presented to finally present a test set to Amazon data in section 2.5.

The drawback which is presented in section 2.4 will be solved in our PDE model in section 3.

# 2.1 Description

This model falls into the second category as described in section 1.3.3.3, which assumes a relationship between default and recovery. It is a hybrid model in the sense that it has characteristics of both structural models and reduced form models. The first models default risk as being directly dependent upon stock prices or firm prices, whereas reduced form models assume defaults to occur exogenously.

This model determines, simultaneously, the term structures for both the default intensity and the recovery. This is done by letting the default and recovery depend upon the corresponding stock price, which is modeled by a jump-to-default tree (not the usual trinomial model in which prices go up, down or remain on the same level). Assuming a relation between the stock price and default probabilities is not new, see for example [Lin06],[AFV03] and [DS03]. Furthermore, defaults usually go paired with low resale values of the underlying asset, confirming the negative correlation. Finally to determine the form of both default and recovery the method describes a calibration process upon CDS spreads which uses a backward recursion for both premium and protection legs.



FIGURE 2.1: The jump-to-default tree as introduced by Das & Hanouna in [DH09]

# 2.2 Methodology

#### 2.2.1 The jump-to-default tree

First a jump-to-default tree for the stock price is created. Requirements for this are the current stock price  $S_0$  and its corresponding volatility<sup>1</sup>  $\sigma$ . A regular binomial tree can either go up or down, however in this case we introduce a third option corresponding to a default. Let the tree take constant time steps of size  $\Delta t$ . Suppose t to be some time, then we define

$$S_{t+\Delta t} = \begin{cases} S_t u = S_t e^{\sigma \sqrt{\Delta t}}, & \text{with prob. } q_S (1 - \xi_S), \\ S_t d = S_t e^{-\sigma \sqrt{\Delta t}}, & \text{with prob. } (1 - q_S)(1 - \xi_S), \\ 0, & \text{with prob. } \xi_S, \end{cases}$$
(2.1)

where  $q_S$  and  $\xi_S$  are short forms for respectively  $q_{S_t}$  and  $\xi_{S_t}$ . See figure 2.1 for a visualisation of the tree at initiation.

Without the knowledge of the probabilities, the tree can still be created, given u and d. In a regular tree the probabilities of an increase and decrease can be determined by using the risk free interest rate. A similar approach can be used to find an expression for  $q_S$ , leaving  $\xi_S$  open for definition. To this end let  $f_t$  correspond to the risk free forward interest rate over the period  $(t, t + \Delta t)$ . To find  $q_S$ , note that by the no-arbitrage principle it holds that

<sup>&</sup>lt;sup>1</sup>The volatility may be both historical or implied. The fact that a relationship between stock and defaults exists is enough. This is a consequence of the existence of a relationship between historical and implied variables.

$$q_{S}(1-\xi_{S})\cdot S_{t}u + (1-q_{S})(1-\xi_{S})\cdot S_{t}d = S_{t}e^{f_{t}\Delta t} \quad \to \quad q_{S}(1-\xi_{S})\cdot u + (1-q_{S})(1-\xi_{S})\cdot d = e^{f_{t}\Delta t},$$
(2.2)

such that  $q_S$  should satisfy

$$q_S = \frac{\frac{e^{f_t \Delta t}}{(1 - \xi_S) - d}}{u - d}.$$
(2.3)

Let's assume t to be today. The tree takes i possible levels at time  $t_i = t + i\Delta t$ . For convenience we shall write each node on the tree to be a point of the form [i, j] where  $i \in \mathbb{N}_{\geq 0}$ , the time index and  $0 \leq j \leq i$ , the level index. A level increase corresponds to an upward movement of the stock price: the first node thus corresponds to [0, 0] and at time  $t + \Delta t$  it can either move down to [1, 0] or upwards to [1, 1]. With this notational convention we can define the probability of default within a time step as:

$$\xi[i,j] = 1 - e^{-\lambda[i,j]\Delta t}, \quad \text{with} \quad \begin{cases} \lambda[i,j] = \frac{1}{S[i,j]^b}, \quad S[i,j] > 0\\ 0 & S[i,j] = 0 \end{cases}$$
(2.4)

where we require  $b \in \mathbb{R}_{>0}$  such that, naturally, default risk increases with a decreasing stock price<sup>2</sup>. With S[i, j], q[i, j] and  $\xi[i, j]$  defined, the tree can completely be determined. Next, the functional form of the recovery for the corresponding firm is defined by the following relation:

$$\rho[i,j] = \Phi(a_0 + a_1 \xi[i,j]), \tag{2.5}$$

where  $\Phi(x), x \in \mathbb{R}$ , corresponds to the standard normal distribution, assuming the recovery to remain within the interval (0, 1). We point out that any other form, satisfying the properties of the recovery might be chosen<sup>3</sup>. The authors of [DH09] provide results for two other functional forms and show results to be dynamically stable.

The code related to computing  $q, \xi$  and  $\rho$  can be found in appendix I.1.2.

<sup>&</sup>lt;sup>2</sup>The first section of this thesis presents both  $\mathbb{P}$ - and  $\mathbb{Q}$ -measures. We therefore want to emphasize that (2.2) just made sure that the stock path increases on average by the risk free rate. This thus guarantees that our stock is in the risk-free measure, such that the probability of default is defined in consistency with the  $\mathbb{Q}$ -measure.

 $<sup>^{3}</sup>$ As we shall show in appendix A.3, changing this form introduces slight different underlying assumptions upon the recovery.

### 2.2.2 Pricing a CDS

Assumptions In order to price a CDS we need to find the current values for both premium and protection legs. Suppose, without loss of generality, a CDS with maturity T and premium payments every  $\Delta t$  years, thus a total of  $N = \frac{T}{\Delta t}$  premium payments. Assume that premium payments are made at the end of each period and if a default occurs in period  $(k\Delta t, (k+1)\Delta t]$  where k an integer between 0 and N - 1, then the protection payment and last premium payment are completed at  $(k + 1)\Delta t$ . Furthermore for the percentage recovery we assume recovery of face value. And for convenience we take the face value to be 1.

**Backward recursions** Let A[i, j] and B[i, j] correspond to respectively the expected future payments of the premium and protection legs at time  $t_i$  and level j. Let  $C_T$  the corresponding annualized CDS spread<sup>4</sup>. The backwards recursions that follow are given by

$$A[i,j] = e^{-f_i \Delta t} C_T + e^{-f_i \Delta t} \left( q[i,j](1-\xi[i,j]) A[i+1,j] + (1-q[i,j])(1-\xi[i,j]) A[i+1,j+1] \right),$$
  

$$B[i,j] = e^{-f_i \Delta t} \xi[i,j](1-\rho[i,j]) + e^{-f_i \Delta t} \left( q[i,j](1-\xi[i,j]) B[i+1,j] + (1-q[i,j])(1-\xi[i,j]) B[i+1,j+1] \right),$$
  
(2.6)

with final conditions A[N, j] = B[N, j] = 0, for  $0 \le j \le N$ . To be more precise we should define A[i, j] as the expected premium payments at  $t_{i+}$  (the expected premium payments at  $t_i$  would include the premium payment at exactly time  $t_i$  as well). Both equations are nothing but working back on the tree. First we understand that both legs are worthless at maturity, subsequently the tree provides us with probabilities that can be used to find values for previous time steps, noting that the premium leg pays a premium with some probability and the protection leg a protection payment.

Also, note that, just as the recovery is based on a face value of 1, so does  $C_T$ , which is thus stated as a fraction of notional.

If now the parameters  $a_0$ ,  $a_1$  and b are known, we can proceed to price the CDS. Premium and protection leg should equal at time  $t_0$ , that is, A[0,0] = B[0,0]. Note furthermore that  $C_T$  can be taken out of the recursion: let  $A^1[i,j]$  be the value of A[i,j] with  $C_T = 1$ , then for any  $C_T$  it holds that  $A[0,0] = C_T \cdot A^1[0,0] = B[0,0]$ , from which follows that

$$C_T = \frac{B[0,0]}{A^1[0,0]}.$$

See appendix I.1.3 for the code related to the pricing of a CDS.

<sup>&</sup>lt;sup>4</sup>If in the market we read a spread of  $\overline{C}$  and there are two payments, then we set  $C = \frac{1}{2}\overline{C}$ .

#### 2.2.3 Calibration on CDSs

On the other hand, if  $a_0$ ,  $a_1$  and b are unknown parameters, that is, the sensitivity of the probability of default to the stock price and the recovery to the default are unknown, then these should be estimated. This can be done by letting these parameters undefined and calibrating them to CDS data of interest. The simplest procedure would be to perform a least squares minimisation:

$$\min_{a_0,a_1,b} \sum_{i} \left( C_{T_i}(a_0, a_1, b) - \hat{C}_{T_i} \right)^2,$$
(2.7)

where *i* iterates over all maturities,  $\hat{C}_{T_i}$  represents the observed market spread for a CDS with maturity  $T_i$ .  $C_{T_i}(a_0, a_1, b)$  is its model implied counterpart, based on the  $\{a_0, a_1, b\}$ . It is finally assumed that all CDSs have the same coupon frequency. The master code of calibration can be found in appendix I.1.

# 2.3 Retrieving recovery and default

The term structure that we wish to create is built up of the expected recoveries and expected default probability at future time states. Again let i denote a time index and ja level within the tree with  $J_i$  referencing all possible levels at time i excluding the default states. Then for the probability of default we have

$$P(\tau = i) = \sum_{j \in J_i} p(i, j) \cdot \xi[i, j],$$
(2.8)

where  $\xi[i, j]$  corresponds to the probability of default at node [i, j] and p(i, j) equals the probability of reaching node [i, j] (by means of any possible path) - Appendix I.1.4 contains the code to determine this probability. Thus the probability of default at time i is the sum of probabilities of default at this time multiplied by the probability of reaching this particular node. Note how (2.8) represents a weighted average, where the probabilities of reaching a note are the weights.

For the expected recovery we have a similar expression. What we are interested in is  $\mathbb{E}^{\mathbb{Q}}[\rho|\tau=i]$ , that is, the expected recovery at time *i* given a default<sup>5</sup>. Then

$$\mathbb{E}^{\mathbb{Q}}\left[\rho|\tau=t_i\right] = \sum_{j\in J_i} \rho[i,j] P(\rho=\rho[i,j]|\tau=i).$$
(2.9)

Now, given that we arrived at time *i*, the probability that the recovery is  $\rho[i, j]$  corresponds to the probability that we reach state [i, j] and that no default occurred, i.e.

$$P(\rho = \rho[i, j] | \tau = i) = \frac{p[i, j]}{\sum_{j \in J_i} p[i, j]}.$$
(2.10)

Using this we get that

$$\mathbb{E}^{\mathbb{Q}}\left[\rho|\tau=t_i\right] = \frac{1}{\sum_{j\in J_i} p[i,j]} \sum_{j\in J_i} p[i,j] \rho[i,j]$$
(2.11)

<sup>&</sup>lt;sup>5</sup>Das & Hanouna choose to look at  $\mathbb{E}^{\mathbb{Q}}[\rho]$  instead. We however believe that this has no related practical understanding as a recovery always goes together with a default: what is the recovery not given a default?

# 2.4 Restriction of the method

The D&H model described in this section is both simple and elegant. It however has one serious drawback, which we will clarify in this section.

Only when the step size,  $\Delta t$ , converges towards zero, does the tree approach actually establish the same results as when directly modeling a Geometric Brownian Motion. If we therefore want to use this method for real quantification - not only for qualitative results where a large step size might suffice - we'd require smaller time steps.

Recall (2.3) in which we determined the probability of an upward movement of the stock given no default. We should thus require

$$0 \le \frac{e^{f_t \Delta t} / (1 - \xi_s) - d}{u - d} \le 1.$$
(2.12)

Using the definition of u, d, and  $\xi_S$  we get

$$0 \le \frac{e^{f_t \Delta t} / e^{\lambda_S \Delta t} - e^{-\sigma \sqrt{\Delta t}}}{e^{\sigma \sqrt{\Delta t} - e^{-\sigma \sqrt{\Delta t}}}} \le 1 \quad \to \quad -\sigma \le (f_t + \lambda_S) \sqrt{\Delta t} \le \sigma, \tag{2.13}$$

where we specifically write  $\lambda_S$  to emphasise that  $\lambda$  depends upon the value of the stock. Indeed, recall that  $\lambda_S = \frac{1}{S_t^b}$ . Note that  $f, \lambda, \sigma > 0$ , from which trivially follows that  $-\sigma \leq (f_t + \lambda_S) \sqrt{\Delta t}$ . The other requirement is little more involved as it eventually sets a lower boundary upon the parameter b:

$$(f_t + \lambda_S)\sqrt{\Delta t} \le \sigma \leftrightarrow \left(f_t + \frac{1}{S_t^b}\right)\sqrt{\Delta t} \le \sigma$$
(2.14)

$$\Leftrightarrow S_t^b \ge \frac{1}{\sigma/\sqrt{\Delta t} - f_t} \tag{2.15}$$

$$\Leftrightarrow b \ge \frac{-\ln\left(\sigma/\sqrt{\Delta t} - f_t\right)}{\ln\left(S_t\right)}, \quad \forall t \in [0, T],$$
(2.16)

where T stands for the largest maturity evaluated in the model. As we will now show, this constraint upon b is a fairly inconvenient one when  $\Delta t$  is small. Constraint (2.16) should hold for any t and therefore in our jump-to-default tree also for all possible scenarios of the stock, in particular the worst case scenario, which equals

$$S_0 e^{-n\sigma\sqrt{\Delta t}}, \quad n = 0, 1, ..., T \cdot N.$$
 (2.17)

Applying this stock price path to the constraint gives us



FIGURE 2.2: Plot of the rhs of (2.18) for all n =. Model parameter b should be at least the size of maximum value on each line.

$$b \ge \frac{-\ln\left(\sigma/\sqrt{\Delta t} - f_n\right)}{\ln\left(S_0 e^{-n\sigma\sqrt{\Delta t}}\right)} \quad \leftrightarrow \quad b \ge \frac{-\ln\left(\sigma/\sqrt{\Delta t} - f_n\right)}{\ln\left(S_0\right) - n\sigma\sqrt{\Delta t}}.$$
(2.18)

When  $\ln(S_0) - n\sigma\sqrt{\Delta t}$  switches signs the rhs shoots up in absolute value as it gradually becomes smaller - it might even touch upon zero - and after switching signs again becomes larger in absolute value. One may however wonder whether this will occur when we reduce the value of  $\Delta t$  substantially? The answer is yes, it clearly does as we can see as follows: the denominator switches signs around

$$n = \left\lfloor \frac{S_0}{\sigma \sqrt{\Delta t}} \right\rfloor,\tag{2.19}$$

and as n increases by  $\frac{1}{\Delta t}$  with each year it is of order  $\frac{1}{\Delta t}$ . The switching point will thus unavoidably be reached when decreasing the value of  $\Delta t$ . The maximum size that b will attain at such passage point depends on the values of  $\sigma$  and  $f_t$ , but it can be quite substantial. In figure 2.2 we plot the values of the rhs of (2.18) and as can be seen switching points are related to peaks which create an unstable point for model parameter b. Furthermore, it follows indeed that decreasing the value of  $\Delta t$  is associated to an earlier trigger of this switching point. And finally, the authors of [PS08] find that especially long-maturity premia are essential for identification as the impact of changes in the recovery rate on short-maturity premia is relatively low. With this in mind we note that increasing the maturity will decrease the value of  $S_t$  in the worst case scenario, as in (2.16), resulting in 'almost surely' an unstable switching point.

In theorem 4.3.4 we shall show that violation of this constraint upon b results in a numerical unstable model. In order to cope with this flaw we will provide a new approach: the PDE model.

**Remark 2.4.1.** We might of course choose b such that it always satisfies the required constraint, but then the model looses its interpretational capabilities as b is found to be a 'small' value. See for example [Mur99], where it is shown that b lies between 1.2 and 2 for Japanese corporates.

Amazon data				
Stock price				
Stock volatility			0.972	
Step size				
Maturity	Forward rate	Market spreads (bps)	Model spreads (bps)	
1	0.02259796	899.06	899.19	
2	0.02730759	1130.93	1130.41	
3	0.03292666	1258.20	1259.80	
4	0.03826696	1265.46	1264.72	
5	0.04360217	1285.32	1283.32	
Parameters $(a_0, a_1, b)$		(0.19916, -1.2768, 0.80096)		
RMSE			1.22	

TABLE 2.1: CDS spreads on Amazon in September 2001

# 2.5 A test set by Amazon

As to confirm the validity of our results we provide a test set which we may compare with other results. Indeed the authors of [DH09] provide a small set of CDS spreads on Amazon for September 2001. The required data they provide can be found in table 2.1. We have performed several tests in order to find the best fit. As described in section 2.2.3 a least squares approach is used in order to determine a fit.

Matlab has quite a range of optimisation tools of which we used the methods lsqcurvefit and fmincon. These however might converge towards local minima. To this end we also implemented a search for the global minimum by using multistart and globalsearch. Multistart was performed referencing the lsqcurvefit method, whereas globalsearch is limited to fmincon only. For the parameters  $\{a_0, a_1, b\}$  we used a starting point of  $\{0, 0, 1\}$ . This is a purely random point which is at least within required boundary conditions - it can be shown that for this stock and a maturity of 5 years we have b > 0.71118. Table 2.2 states the fitted results. The most important thing to note here is how the global approach is a highly important one. It is expected that recovery and default should be negatively correlated, which in this particular model is represented by a negative value of the parameter  $a_1$ . The local minimum parameter set doesn't reflect this, whereas the global does.

Finally the resulting expected implied recovery and probability of default over 5 years are plotted in figure 2.3, together with a plot of the fit of the spreads.

Method	Fitted parameters	fmin
lsqcurvefit	$\{-0.4343, 0.20924, 0.93558\}$	4.9846
fmincon	$\{-0.43431, 0.20925, 0.93558\}$	4.9846
multistart	$\{0.19899, -1.2765, 0.801\}$	3.6295
globalsearch	$\{0.19894, -1.2764, 0.80101\}$	3.6295

TABLE 2.2: Trinomial model fitted parameters, for several local and global Matlab methods. fmin indicates the optimised value of the least squares approximation in basis points.



FIGURE 2.3: Expected implied recovery at default and probability of default on Amazon (Based on spreads of September 2001)

# Continuous term structures: The PDE model

"Sapere Aude!"

- Immanuel Kant, Answering the Question: What Is Enlightenment?

In section 2 we described the D&H model as introduced by Das & Hanouna in [DH09], a model satisfying our requirements from section 1.3.3.4.

However, as was shown, it has the drawback that time steps have to be fairly large in order to provide numerical stability and a tree with proper probabilities. To solve this issue we note that binomial trees are special cases of some PDE discretization, this can intuitively be seen from figure 3.1. This means that in a similar way we might find a PDE which provides the D&H model as a special case. The advantage of having a PDE is that we have more liberty with regards to the discretization, and in particular choose to use an implicit discretization which has the favorable property of being conditionally stable. This section is dedicated to finding a generalization of the D&H model and provide some further improvements.

This section is constructed as follows. In section 3.1 we provide an understanding of the underlying dynamics of the continuous model. Then the more technical section 3.2 provides the first steps towards the required PDEs for both premium and protection legs, by deducing expressions for the dynamics of an infinitesimal time step in both legs. Then an application of Ito's lemma is used to combine these results with the underlying dynamics in section 3.3, providing us with the skeleton of the PDE model. After the set up we provide a comparison of our model to that of the model by Das & Hanouna. By now no specific choices have been made upon the form of both default intensity and recovery, this is provided in section 3.4. Finally we look into how each model parameter influences both the CDS spread and the recovery and default term structures.



FIGURE 3.1: An intuitive relation between the PDE discretization scheme and that of a binomial tree.

# 3.1 Jump to default stock prices

In the model by Das & Hanouna we have seen that the underlying dynamics where provided by the stock price. This stock price was modeled by a tree which had two regular tree directions, going up and down such as is the case for a regular Cox Ross Rubinstein (CRR) tree [CRR79], and an additional possibility of jumping down to a default state. In the limit, the regular two options converge towards a regular stock price process under the assumption of a Geometric Brownian Motion (GBM). A jump to default was modeled by the probability presented in (2.4), which is nothing but the probability of default of an exponential distribution with intensity  $\lambda$  depending upon the underlying stock. Therefore we understand that the default is the first jumping time of a Poisson process with changing intensity. Such default process is often used in mathematical finance and is referred to as a hazard process. The intensity of such process is frequently modeled (e.g. a Vasicek or CIR process) and referred to as the intensity process. In our case the intensity depends upon the stock such that it is a stochastic process. Such Poisson process is called a Cox process. For more details and information regarding hazard rate and intensity process we refer to appendix D.

Putting these pieces together, both the GBM part and the jump to default part, provides us with a jump-to-default stock price process described by the jump diffusion process

$$dS_t = (r_t + \lambda_t)S_{t-}dt + \sigma S_{t-}dW_t - S_t dI_t, \qquad (3.1)$$

where r the interest rate,  $\lambda$  the intensity of the default process, W a Brownian motion and I the Poisson process with intensity  $\lambda$ . The validity of (3.1) is a general result which can be found in [Mer76]. Note this process to be the pre-default process, that is, this is the process the path follows up till a default<sup>1</sup>. After a default it remains insolvent, at S = 0 (similarly as was defined in the D&H model).

 $<sup>^{1}</sup>$ Using the term pre-default is conventional and not necessary in this case. We will however use this term more frequently and like to remain consistent.

# 3.2 Infinitesimal increments in both legs of a CDS

The underlying dynamics of the premium and protection legs are functions of the stock prices. Therefore, just as is done for the regular derivation of option prices under Black-Scholes dynamics, we can apply Ito's lemma to find the dynamics of both legs (a more general version of Ito's lemma, accounting for jumps, has to be applied). This application shall be provided in section 3.3. However, usually the next step would be to assume the option of interest to be risk-free under the risk-neutral measure, that is, it is a contingent attainable claim in the sense of definition B.1.3. This then provides the convenient result that

$$\mathbb{E}^{\mathbb{Q}}\left[dV_t|\mathcal{G}_t\right] = rV_t dt,$$

where  $V_t$  the option price, see (B.7). Our premium and protection legs are however not contingent attainable claims. One may intuitively understand these not to be, however in the next section we shall prove the martingale condition to be invalid. Therefore we have to really compute

$$\mathbb{E}^{\mathbb{Q}}\left[d\operatorname{Prem}\left(t,T\right)|\mathcal{G}_{t}\right] \quad \text{and} \quad \mathbb{E}^{\mathbb{Q}}\left[d\operatorname{Prot}\left(t,T\right)|\mathcal{G}_{t}\right].$$

instead of assuming them to be risk-free. Due to this computation this section is a little more technical. We shall look at the premium and protection leg separately, proof them not to be martingales under the Q-measure when adjusted for risk-free discounting. Then finally determine expressions for

$$d(\operatorname{Prem}(t,T))$$
 and  $d(\operatorname{Prot}(t,T))$ .

The results provided next are based upon the assumption that the interest rate follows a continuously compounded process r. This section uses many of the definitions and results presented in appendix D on the default process.

#### 3.2.1 The premium leg

**Proposition 3.2.1.** The time t value of a C-continuous paying premium leg of a CDS with a notional of 1 maturing at time  $T \ge t$ , satisfies

$$\operatorname{Prem}\left(t,T\right) = \mathbb{E}^{\mathbb{Q}}\left[\int_{t}^{T} e^{-\int_{t}^{s} r_{u} du} \mathbf{1}_{\{\tau > s\}} C ds \left|\mathcal{G}_{t}\right].$$
(3.2)

**Lemma 3.2.1.** Let the default time follow a hazard process  $\Gamma_t$  with stochastic hazard rate  $\lambda_t$ . The random variable defined by

$$e^{-\int_{0}^{t} r_{s} ds} \cdot \operatorname{Prem}\left(t, T\right) + \int_{0}^{t} \mathbf{1}_{\{\tau > s\}} e^{-\int_{0}^{s} r_{u} du} C ds$$
(3.3)

is a  $\mathbb G\text{-martingale}$  under the  $\mathbb Q\text{-measure}.$ 

We distinguish two different proofs. The first being a simple one and the second a little more complicated, having nice implications for following corollaries.

## Proof. Proof 1

Note first that the discounted expected premium leg satisfies

$$e^{-\int_0^t r_s ds} \cdot \operatorname{Prem}\left(t, T\right) = \mathbb{E}^{\mathbb{Q}}\left[\int_t^T e^{-\int_0^s r_u du} \mathbf{1}_{\{\tau > s\}} C ds \middle| \mathcal{G}_t\right].$$
(3.4)

Then  $V_t = \mathbb{E}^{\mathbb{Q}} \left[ \int_0^T e^{-\int_0^s r_u du} \mathbf{1}_{\{\tau > s\}} C ds \middle| \mathcal{G}_t \right]$  is a martingale. This follows immediately by the tower property of conditional expectations: let  $t_1 \le t_2 \le T$ , then

$$\mathbb{E}^{\mathbb{Q}}\left[V_{t_2}|\mathcal{G}_{t_1}\right] = V_{t_1}.\tag{3.5}$$

The result now follows by the fact that  $\mathcal{G}_t$  contains all information including the default process

$$V_t = \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T e^{-\int_0^s r_u du} \mathbf{1}_{\{\tau > s\}} C ds + \int_0^t e^{-\int_0^s r_u du} \mathbf{1}_{\{\tau > s\}} C ds \cdot \left| \mathcal{G}_t \right| \right]$$
$$= \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T e^{-\int_0^s r_u du} \mathbf{1}_{\{\tau > s\}} C ds \left| \mathcal{G}_t \right| \right] + \int_0^t e^{-\int_0^s r_u du} \mathbf{1}_{\{\tau > s\}} C ds.$$

Proof. Proof 2

$$\mathbb{E}^{\mathbb{Q}}\left[\int_{t}^{T} e^{-\int_{0}^{s} r_{u} du} \mathbf{1}_{\{\tau > s\}} Cds \left| \mathcal{G}_{t} \right] = \mathbf{1}_{\{\tau > t\}} \mathbb{E}^{\mathbb{Q}}\left[\int_{t}^{T} e^{-\int_{0}^{s} r_{u} du} \mathbf{1}_{\{\tau > s\}} Cds \left| \mathcal{G}_{t} \right] \right]$$
$$= \mathbf{1}_{\{\tau > t\}} e^{\Gamma_{t}} \mathbb{E}^{\mathbb{Q}}\left[\mathbf{1}_{\{\tau > t\}} \int_{t}^{T} e^{-\int_{0}^{s} r_{u} du} \mathbf{1}_{\{\tau > s\}} Cds \left| \mathcal{F}_{t} \right] \right]$$
$$= \mathbf{1}_{\{\tau > t\}} e^{\Gamma_{t}} \mathbb{E}^{\mathbb{Q}}\left[\int_{t}^{T} e^{-\int_{0}^{s} r_{u} du} e^{-\Gamma_{s}} Cds \left| \mathcal{F}_{t} \right]$$
(3.6)

Define the continuous function  $\tilde{V}_t := \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T e^{-\int_0^s r_u du} e^{-\Gamma_s} C ds \Big| \mathcal{F}_t \right]$ . Then

$$\mathbb{E}^{\mathbb{Q}}\left[\frac{d\tilde{V}_t}{dt}\middle|\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{Q}}\left[\lim_{\Delta t \to 0} \frac{\tilde{V}_{t+\Delta t} - \tilde{V}_t}{\Delta t}\middle|\mathcal{F}_t\right].$$
(3.7)

Now the dominated convergence theorem can directly be applied as

$$\lim_{\Delta t \to 0} \frac{\tilde{V}_{t+\Delta t} - \tilde{V}_t}{\Delta t} = \frac{d\tilde{V}_t}{dt} < D \in \mathbb{R}_{\geq 0}.$$
(3.8)

Bringing the expectation inside we proceed

$$\begin{split} \mathbb{E}^{\mathbb{Q}} \left[ \frac{d\tilde{V}_{t}}{dt} \middle| \mathcal{F}_{t} \right] &= \lim_{\Delta t \to 0} \frac{\mathbb{E}^{\mathbb{Q}} \left[ \tilde{V}_{t+\Delta t} \middle| \mathcal{F}_{t} \right] - \tilde{V}_{t}}{\Delta t} \\ &= \lim_{\Delta t \to 0} \frac{\mathbb{E}^{\mathbb{Q}} \left[ \int_{t}^{T} e^{-\int_{0}^{s} r_{u} du} e^{-\Gamma_{s}} C ds - \int_{t}^{t+\Delta t} e^{-\int_{0}^{s} r_{u} du} e^{-\Gamma_{s}} C ds \middle| \mathcal{F}_{t} \right] - \tilde{V}_{t}}{\Delta t} \\ &= \lim_{\Delta t \to 0} \frac{\mathbb{E}^{\mathbb{Q}} \left[ -\int_{t}^{t+\Delta t} e^{-\int_{0}^{s} r_{u} du} e^{-\Gamma_{s}} C ds \middle| \mathcal{F}_{t} \right]}{\Delta t} \\ &= \lim_{\Delta t \to 0} \mathbb{E}^{\mathbb{Q}} \left[ \frac{-\int_{t}^{t+\Delta t} e^{-\int_{0}^{s} r_{u} du} e^{-\Gamma_{s}} C ds}{\Delta t} \middle| \mathcal{F}_{t} \right]. \end{split}$$

The term inside the expectation did not change, such that interchanging limit and expectation is analogously allowed, giving us

$$\mathbb{E}^{\mathbb{Q}}\left[\frac{d\tilde{V}_{t}}{dt}\middle|\mathcal{F}_{t}\right] = \mathbb{E}^{\mathbb{Q}}\left[\lim_{\Delta t \to 0} \frac{-\int_{t}^{t+\Delta t} e^{-\int_{0}^{s} r_{u} du} e^{-\Gamma_{s}} C ds}{\Delta t}\middle|\mathcal{F}_{t}\right]$$
$$= \mathbb{E}^{\mathbb{Q}}\left[\lim_{\Delta t \to 0} \frac{-e^{-\int_{0}^{t} r_{u} du} e^{-\Gamma_{t}} C \Delta t}{\Delta t}\middle|\mathcal{F}_{t}\right]$$
$$= -e^{-\int_{0}^{t} r_{u} du} e^{-\Gamma_{t}} C.$$

Due to continuity of  $\frac{d\tilde{V}_t}{dt}$  it is  $\mathcal{F}_t$  measurable such that

$$-e^{-\int_0^t r_u du} e^{-\Gamma_t} C = \mathbb{E}^{\mathbb{Q}} \left[ \frac{d\tilde{V}_t}{dt} \middle| \mathcal{F}_t \right] = \frac{d\tilde{V}_t}{dt}$$

This is the Radon Nikodým derivative and we may write

$$d\tilde{V} = -e^{-\int_0^t r_u du} e^{-\Gamma_t} C dt.$$
(3.9)

Now only regular Ito calculus is required to finish the proof:

$$d\left(e^{\Gamma_{t}}\tilde{V}_{t}\right) = e^{\Gamma_{t}}d\tilde{V}_{t} + \tilde{V}_{t}d\left(e^{\Gamma_{t}}\right)$$
$$= -e^{\Gamma_{t}}e^{-\int_{0}^{t}r_{u}du}e^{-\Gamma_{t}}Cdt + \tilde{V}_{t}e^{\Gamma_{t}}\lambda_{t}dt$$
$$= \left(\tilde{V}_{t}e^{\Gamma_{t}}\lambda_{t} - e^{-\int_{0}^{t}r_{u}du}C\right)dt.$$

And finally

$$d\left(\mathbf{1}_{\{\tau>t\}}e^{\Gamma_{t}}\tilde{V}_{t}\right) = \mathbf{1}_{\{\tau>t\}}d\left(e^{\Gamma_{t}}\tilde{V}_{t}\right) + e^{\Gamma_{t}}\tilde{V}_{t}d\left(\mathbf{1}_{\{\tau>t\}}\right)$$

$$= (1 - H_{t})d\left(e^{\Gamma_{t}}\tilde{V}_{t}\right) - e^{\Gamma_{t}}\tilde{V}_{t}dH_{t}$$

$$= (1 - H_{t})\left(\tilde{V}_{t}e^{\Gamma_{t}}\lambda_{t} - e^{-\int_{0}^{t}r_{u}du}C\right)dt - e^{\Gamma_{t}}\tilde{V}_{t}dH_{t}$$

$$= -e^{\Gamma_{t}}\tilde{V}_{t}\left(dH_{t} - (1 - H_{t})\lambda_{t}dt\right) - (1 - H_{t})e^{-\int_{0}^{t}r_{u}du}Cdt$$

$$= -e^{\Gamma_{t}}\tilde{V}_{t}d\overline{M}_{t} - (1 - H_{t})e^{-\int_{0}^{t}r_{u}du}Cdt.$$
(3.10)

Recalling  $\overline{M}_t$  to be a martingale (see appendix D), the result follows immediately when integrating both sides.

**Corollary 3.2.1.**  $e^{-\int_0^\tau r_s ds} \operatorname{Prem}(t,T)$  is not a  $\mathbb{G}$ -martingale under the  $\mathbb{Q}$ -measure.

Lemma 3.2.2. An infinitesimal increment on the premium leg satisfies

$$d\left(\operatorname{Prem}\left(t,T\right)\right) = -e^{\int_{0}^{t} r_{s} ds} e^{\Gamma_{t}} \tilde{V}_{t} d\overline{M}_{t} - (1 - H_{t-}) C dt + r_{t} \operatorname{Prem}\left(t,T\right) dt. \quad (3.11)$$

**Proof.** Note from (3.4) and (3.6) that

$$\operatorname{Prem}\left(t,T\right) = e^{\int_{0}^{t} r_{s} ds} \mathbf{1}_{\{\tau > t\}} e^{\Gamma_{t}} \tilde{V}_{t}.$$
(3.12)

Using what was found in (3.10) we proceed

$$d\left(\operatorname{Prem}\,(t,T)\right) = d\left(\operatorname{Prem}\,(t-,T)\right)$$

$$= e^{\int_{0}^{t} r_{s} ds} d\left(\mathbf{1}_{\{\tau > t-\}} e^{\Gamma_{t}} \tilde{V}_{t}\right) + \mathbf{1}_{\{\tau > t-\}} e^{\Gamma_{t}} \tilde{V}_{t} d\left(e^{\int_{0}^{t} r_{s} ds}\right)$$

$$= e^{\int_{0}^{t} r_{s} ds} \left(-e^{\Gamma_{t}} \tilde{V}_{t} d\overline{M}_{t} - (1-H_{t-}) e^{-\int_{0}^{t} r_{u} du} C dt\right) + r_{t} \operatorname{Prem}\,(t-,T) dt$$

$$= -e^{\int_{0}^{t} r_{s} ds} e^{\Gamma_{t}} \tilde{V}_{t} d\overline{M}_{t} - (1-H_{t-}) C dt + r_{t} \operatorname{Prem}\,(t-,T) dt.$$

$$(3.13)$$

Now finally we provide the main result of this section as the following corollary.

**Corollary 3.2.2.** The expectation of an infinitesimal increment on a C-continuous paying premium leg under the  $\mathbb{Q}$ -measure conditioning on  $\mathbb{G}$  satisfies

$$\mathbb{E}^{\mathbb{Q}}\left[d\left(\operatorname{Prem}\left(t,T\right)\right)|\mathcal{G}_{t-}\right] = -(1-H_{t-})Cdt + r_t\operatorname{Prem}\left(t-,T\right)dt.$$
(3.15)

### 3.2.2 The protection leg

In section 3.2.1 results were presented for an infinitesimal increment on the premium leg. In this section similar results with analogous proofs for the protection leg are presented.

**Proposition 3.2.2.** The time t value of the protection leg of a CDS maturing at time  $T \ge t$  with a notional of 1, satisfies

$$\operatorname{Prot}\left(t,T\right) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_{t}^{\tau} r_{s} ds} \operatorname{LGD}_{\tau} \mathbf{1}_{\left\{t < \tau \leq T\right\}} \middle| \mathcal{G}_{t}\right].$$
(3.16)

where  $LGD_{\tau}$  is assumed to be the LGD under the RFV assumption.

**Lemma 3.2.3.** Let the default time follow a hazard process  $\Gamma_t$  with stochastic hazard rate  $\lambda_t$ . The random variable defined by

$$e^{-\int_0^t r_s ds} \cdot \operatorname{Prot}\left(t, T\right) + \int_0^t \left(1 - H_s\right) \lambda_s e^{-\int_0^s r_u du} \operatorname{LGD}_s\right) ds \tag{3.17}$$

is a  $\mathbb{G}$ -martingale under the  $\mathbb{Q}$ -measure.

Proof.

$$e^{-\int_{0}^{t} r_{s} ds} \cdot \operatorname{Prot}\left(t, T\right) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_{0}^{\tau} r_{s} ds} \operatorname{LGD}_{\tau} \mathbf{1}_{\{t < \tau \leq T\}} \middle| \mathcal{G}_{t}\right]$$

$$= \mathbf{1}_{\{\tau > t\}} e^{\Gamma_{t}} \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_{0}^{\tau} r_{s} ds} \operatorname{LGD}_{\tau} \mathbf{1}_{\{t < \tau \leq T\}} \middle| \mathcal{F}_{t}\right]$$

$$= \mathbf{1}_{\{\tau > t\}} e^{\Gamma_{t}} \mathbb{E}^{\mathbb{Q}}\left[\int_{t}^{T} e^{-\int_{0}^{u} r_{s} ds} L_{u} dF_{u} \middle| \mathcal{F}_{t}\right]$$

$$= \mathbf{1}_{\{\tau > t\}} e^{\Gamma_{t}} \mathbb{E}^{\mathbb{Q}}\left[\int_{t}^{T} e^{-\int_{0}^{u} r_{s} ds} L_{u} e^{-\Gamma_{u}} \lambda_{u} du \middle| \mathcal{F}_{t}\right].$$
(3.18)

Define  $\tilde{V}_t = \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T e^{-\int_0^u r_s ds} e^{-\Gamma_u} L_u \lambda_u du \Big| \mathcal{F}_t \right]$ . Using the exact same approach as in the proof of lemma 3.2.1 we determine

$$d\tilde{V}_t = -e^{-\int_0^t r_s ds} e^{-\Gamma_t} L_t \lambda_t dt.$$
(3.19)

The rest of the proof is just an application of regular Ito calculus:

$$d\left(e^{\Gamma_{t}}\tilde{V}_{t}\right) = e^{\Gamma_{t}}d\tilde{V}_{t} + \tilde{V}_{t}d\left(e^{\Gamma_{t}}\right)$$
$$= -e^{\Gamma_{t}}e^{-\int_{0}^{t}r_{s}ds}L_{t}e^{-\Gamma_{t}}\lambda_{t}dt + \tilde{V}_{t}e^{\Gamma_{t}}\lambda_{t}dt$$
$$= \lambda_{t}\left(\tilde{V}_{t}e^{\Gamma_{t}} - e^{-\int_{0}^{t}r_{s}ds}\mathrm{LGD}_{t}\right)dt.$$

And finally to determine  $d\left(e^{-\int_0^t r_s ds} \cdot \operatorname{Prot}\left(t, T\right)\right)$  we solve

$$d\left(e^{-\int_{0}^{t}r_{s}ds}\cdot\operatorname{Prot}\left(t,T\right)\right) = d\left(\mathbf{1}_{\{\tau>t\}}e^{\Gamma_{t}}\tilde{V}_{t}\right)$$

$$= (1-H_{t})d\left(e^{\Gamma_{t}}\tilde{V}_{t}\right) - e^{\Gamma_{t}}\tilde{V}_{t}dH_{t}$$

$$= (1-H_{t})\lambda_{t}\left(\tilde{V}_{t}e^{\Gamma_{t}} - e^{-\int_{0}^{t}r_{s}ds}\operatorname{LGD}_{t}\right)dt - e^{\Gamma_{t}}\tilde{V}_{t}dH_{t}$$

$$= -e^{\Gamma_{t}}\tilde{V}_{t}\left(dH_{t} - (1-H_{t})\lambda_{t}dt\right) - (1-H_{t})\lambda_{t}e^{-\int_{0}^{t}r_{s}ds}\operatorname{LGD}_{t}dt$$

$$(3.20)$$

$$= -e^{\Gamma_{t}}\tilde{V}_{t}d\overline{M}_{t} - (1-H_{t})\lambda_{t}e^{-\int_{0}^{t}r_{s}ds}\operatorname{LGD}_{t}dt.$$

Integrating both sides gives us the desired result.

**Corollary 3.2.3.** The discounted expected protection payment at time t, paying  $LGD_{\tau}$  in case of a default, given by

$$\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_{0}^{\tau} r_{s} ds} \mathrm{LGD}_{\tau} \mathbf{1}_{\{t < \tau \leq T\}} \middle| \mathcal{G}_{t}\right]$$
(3.21)

is not a  $\mathbb G\text{-martingale}$  under the  $\mathbb Q\text{-measure}.$ 

Lemma 3.2.4. An infinitesimal increment on the protection leg satisfies

$$d\left(\operatorname{Prot}\left(t,T\right)\right) = -e^{\int_{0}^{t} r_{s} ds} e^{\Gamma_{t}} \tilde{V}_{t} d\overline{M}_{t} - (1 - H_{t-}) \lambda_{t} \operatorname{LGD}_{t} dt + r_{t} \operatorname{Prot}\left(t,T\right) dt. \quad (3.22)$$

**Proof.** Note that

$$Prot(t,T) = e^{\int_0^t r_s ds} e^{-\int_0^t r_s ds} Prot(t,T)$$
(3.23)

and that we may use (3.20) to find

$$d\left(\operatorname{Prot}\left(t,T\right)\right) = d\left(\operatorname{Prot}\left(t,T\right)\right) \tag{3.24}$$

$$= d\left(e^{\int_0^t r_s ds} e^{-\int_0^t r_s ds} \operatorname{Prot}\left(t, T\right)\right)$$
(3.25)

$$= e^{\int_{0}^{t} r_{s} ds} d\left(e^{-\int_{0}^{t} r_{s} ds} \operatorname{Prot}\left(t, T\right)\right) + e^{-\int_{0}^{t} r_{s} ds} \operatorname{Prot}\left(t, T\right) d\left(e^{\int_{0}^{t} r_{s} ds}\right)$$
(3.26)

$$= e^{\int_0^t r_s ds} \left( -e^{\Gamma_t} \tilde{V}_t d\overline{M}_t - (1 - H_{t-}) \lambda_t e^{-\int_0^t r_s ds} \text{LGD}_t dt \right) + r_t \text{Prot} (t-, T) dt$$
(3.27)

$$= -e^{\int_0^t r_s ds} e^{\Gamma_t} \tilde{V}_t d\overline{M}_t - (1 - H_{t-}) \lambda_t \text{LGD}_t dt + r_t \text{Prot} (t, T) dt$$
(3.28)

**Corollary 3.2.4.** The expectation of an infinitesimal increment on a protection leg, paying  $LGD_t$  as protection payment at default, under the  $\mathbb{Q}$ -measure conditioning on  $\mathbb{G}$  satisfies

$$\mathbb{E}^{\mathbb{Q}}\left[d(\operatorname{Prot}\left(t,T\right))|\mathcal{G}_{t-}\right] = -(1-H_{t-})\lambda_t \operatorname{LGD}_t dt + r_t \operatorname{Prot}\left(t-,T\right) dt.$$
(3.29)

# 3.3 Resulting CDS PDEs

In this section we apply Ito's lemma (see (B.14)) to the stock price process together with the results from section 3.2 to find dynamics, in the form of PDEs, for both CDS legs.

We have seen in section 3.1 that the stock price process of the underlying jump to default process satisfies

$$dS_t = (r_t + \lambda_t) S_{t-} dt + \sigma S_{t-} dW_t - S_{t-} dN_t.$$
(3.30)

Let  $f_t \equiv f(t, S_t) : [t_0, \infty) \times \mathbb{R} \to \mathbb{R}$  with  $f \in C^{1,2}$ . Then note how (3.30) satisfies (B.11) such that we may apply Ito's general lemma, (B.14), to find that

$$df(t, S_{t}) = \frac{\partial f_{t-}}{\partial t} dt + (r_{t} + \lambda_{t}) S_{t-} \frac{\partial f_{t-}}{\partial S} dt + \frac{1}{2} \sigma^{2} S_{t-}^{2} \frac{\partial^{2} f_{t-}}{\partial S^{2}} dt + \frac{\partial f_{t-}}{\partial S} dW_{t} + (f(t, S_{t-} + \Delta S_{t}) - f(t, S_{t-})) dN_{t}.$$
(3.31)

Furthermore, note that  $\Delta S_t$  equals exactly the size of  $S_{t-}$  in case of a jump to default, from which follows

$$df(t, S_t) = \frac{\partial f_{t-}}{\partial t} dt + (r_t + \lambda_t) S_{t-} \frac{\partial f}{\partial S} dt + \frac{1}{2} \sigma^2 S_{t-}^2 \frac{\partial^2 f}{\partial S^2} dt + \frac{\partial f}{\partial S} dW_t + (f(t, 0) - f(t, S_{t-})) dN_t.$$
(3.32)

The next step is to take the risk neutral expectation on both sides conditioning on  $\mathcal{G}_{t-}$ . However in order to do so we need to know the risk neutral expectation of the Cox process N(t) on a small interval - note how  $f(t, 0) - f(t, S_{t-})$  is deterministic under  $\mathcal{G}_{t-}$ :

$$\mathbb{E}^{\mathbb{Q}}\left[dN(t)|\mathcal{G}_{t-}\right] = 1 - e^{-\lambda_t dt} = \left(1 - \left(1 - \lambda_t dt\right)\right) = \lambda_t dt.$$
(3.33)

Thus taking expectations in (3.32) gives us

$$\mathbb{E}^{\mathbb{Q}}\left[df(t,S_t)|\mathcal{G}_{t-}\right] = \frac{\partial f_{t-}}{\partial t}dt + (r_t + \lambda_t)S_{t-}\frac{\partial f_{t-}}{\partial S}dt + \frac{1}{2}\sigma^2 S_{t-}^2\frac{\partial^2 f_{t-}}{\partial S^2}dt + \lambda_t\left(f(t,0) - f(t,S_{t-})\right)dt$$
(3.34)

For both legs we understand that after a default, that is when  $S_t = 0$ , no expected further payments will take place, i.e.  $\operatorname{Prem}(t,T) = \operatorname{Prot}(t,T) = 0$ . Then recall from corollary 3.2.2 that

$$\mathbb{E}^{\mathbb{Q}}\left[d\left(\operatorname{Prem}\left(t,T\right)\right)|\mathcal{G}_{t-}\right] = -(1-H_{t-})Cdt + r_t\operatorname{Prem}\left(t-,T\right)dt.$$
(3.35)

and from corollary 3.2.4 that

$$\mathbb{E}^{\mathbb{Q}}\left[d\left(\operatorname{Prot}\left(t,T\right)\right)|\mathcal{G}_{t-}\right] = -(1-H_{t-})\lambda_t \operatorname{LGD}_t dt + r_t \operatorname{Prot}\left(t-,T\right) dt.$$
(3.36)

Now the resulting theorems can be stated:

**Theorem 3.3.1.** (A PDE for the CDS premium leg) Suppose a C-continuous paying CDS with maturity T and assume the reference entity to have a default time governed by a hazard process with (stochastic) intensity  $\lambda_t$ , possibly depending upon the stock  $S_t$ . Let the underlying stock follow the process as defined in (3.30). The expected future premium payments at time t, denoted by  $A_t$ , satisfy

$$\frac{\partial A_{t-}}{\partial t} + (r_t + \lambda_t) S_{t-} \frac{\partial A_{t-}}{\partial S} + \frac{1}{2} \sigma^2 S_{t-}^2 \frac{\partial^2 A_{t-}}{\partial S^2} - (r_t + \lambda_t) A_{t-} + C = 0 \qquad (3.37)$$

defined on  $S \in \mathbb{R}_{>0}$  and  $t \in (0, T)$  with initial and boundary conditions

$$\begin{cases} A(T,S) = 0, \\ A(t,0) = 0, t \in [0,T]. \end{cases}$$
(3.38)

**Theorem 3.3.2.** (A PDE for the CDS protection leg) Suppose a CDS and assume the reference entity to have a default time governed by a hazard process with (stochastic) intensity  $\lambda_t$ , possibly depending upon the stock  $S_t$ . Let the underlying stock follow the process as defined in (3.30). Assume LGD<sub>t</sub> to be the LGD under the RFV assumption. The expected future protection payment at time t, denoted by  $B_t$ , satisfy

$$\frac{\partial B_{t-}}{\partial t} + (r_t + \lambda_t) S_{t-} \frac{\partial B_{t-}}{\partial S} + \frac{1}{2} \sigma^2 S_{t-}^2 \frac{\partial^2 B_{t-}}{\partial S^2} - (r_t + \lambda_t) B_{t-} + \lambda_t \text{LGD}_t = 0, \quad (3.39)$$

defined on  $S \in \mathbb{R}_{>0}$  and  $t \in (0, T)$  with initial and boundary conditions

$$\begin{cases} B(T,S) = 0, \\ B(t,0) = 0, t \in [0,T]. \end{cases}$$
(3.40)

**Proof.** [For both theorem 3.3.1 and 3.3.2]

The PDE's follow immediately after substituting equations (3.35) and (3.36) into (3.34). One should only that when  $S_{t-} > 0$ , that is, no default occurred up till time t-, then  $H_{t-} = 0$ .

Then to show validity of the boundary and initial conditions we note that when t = T, then by definition of both premium and protection legs (see (1.13) and (1.15)), these are set to zero. Finally when  $S_t = 0$  we have that  $\tau \leq t$ , such that the boundary conditions follow from both

$$\operatorname{Prem}\left(t,T\right) = \mathbf{1}_{\{\tau > t\}} e^{\Gamma_{t}} \mathbb{E}^{\mathbb{Q}}\left[\int_{t}^{T} e^{-\int_{0}^{s} r_{u} du} e^{-\Gamma_{s}} C ds \middle| \mathcal{F}_{t}\right]$$
(3.41)

and

$$\operatorname{Prot}\left(t,T\right) = \mathbf{1}_{\{\tau>t\}} e^{\Gamma_{t}} \mathbb{E}^{\mathbb{Q}}\left[\int_{t}^{T} e^{-\int_{t}^{u} r_{s} ds} L_{u} e^{-\Gamma_{u}} \lambda_{u} du \middle| \mathcal{F}_{t}\right].$$
(3.42)

**Remark 3.3.3.** Theorems 3.3.1 and 3.3.2 provide general expressions for premium and protection legs of CDSs. We did not specify the exact forms of both the intensity process nor the recovery. In section 3.5 we will specify our modeling choice.

# **3.4** D&H model as a special case

Theorems 3.3.1 and 3.3.2 provide PDEs which should be related to the backward recursions in the D&H model. Indeed, we argued before that PDEs and tree models are related. In this section we will more carefully describe this relationship by showing that both are indeed equal under certain conditions. In particular, the backward recursions denoted in (2.6) can be created by applying an explicit discretisation for both PDEs, that is, a backwards time discretisation and central discretisation over S. We shall only provide a deduction for the premium leg PDE (3.37) as the protection leg follows in a similar fashion.

Let  $\Delta t$  the time increment and  $\Delta x$  the stock increment, such that  $t_k = k\Delta t$  and  $T = t_N$ ,  $k \leq N$  (no maximum stock value has to be inquired). Then we provide the following discretisations<sup>2</sup>:

$$\frac{\partial A_t}{\partial t} = \frac{1}{\Delta t} \left( A_{n+1}^m - A_n^m \right)$$

$$\frac{\partial A_t}{\partial S} = \frac{1}{2\Delta x} \left( A_{n+1}^{m+1} - A_{n+1}^{m-1} \right)$$

$$\frac{\partial^2 A_t}{\partial S^2} = \frac{1}{\Delta x^2} \left( A_{n+1}^{m+1} - 2A_{n+1}^m + A_{n+1}^{m-1} \right).$$
(3.43)

Note the time discretization to be Euler forwards, and this is thus an explicit discretization scheme. We first transform the PDE by using  $X = \log(S)$ , giving us

$$\frac{\partial A_{t-}}{\partial t} + \left(r_t + \lambda_t - \frac{1}{2}\sigma^2\right)\frac{\partial A_{t-}}{\partial X} + \frac{1}{2}\sigma^2\frac{\partial^2 A_{t-}}{\partial X^2} - (r_t + \lambda_t)A_{t-} + C = 0.$$
(3.44)

Then to simplify even further, the next transformation we'll apply before starting the deduction is setting  $A = e^{\int_0^t (r_s + \lambda_s) ds} \overline{A}$ , such that we may work with the PDE

$$\frac{\partial \bar{A}_{t-}}{\partial t} + \left(r_t + \lambda_t - \frac{1}{2}\sigma^2\right)\frac{\partial \bar{A}_{t-}}{\partial X} + \frac{1}{2}\sigma^2\frac{\partial^2 \bar{A}_{t-}}{\partial X^2} + e^{-\int_0^t (r_s - \lambda_s)ds}C = 0.$$
(3.45)

We now apply the discretization scheme from above. However, just as the particular choice of the above scheme we have a similar liberty in choosing how to evaluate the integrals. We shall see that our particular choice of grid evaluation corresponds exactly to the assumptions made in the D&H model.

We shall evaluate the integral for  $r_s$  by  $\int_0^{(n+1)\Delta t} r_s ds$ , n < N (that is backwards in time<sup>3</sup>), whereas the integral over  $\lambda_s$  is chosen to be evaluated forward in time, that is  $\int_0^{n\Delta t} \lambda_s ds$ . Next we approximate the integrals by the summation

$$\int_0^{n\Delta t} f(t)dt \approx \sum_{k=0}^{n-1} f(t_k)\Delta t.$$

 $<sup>^{2}</sup>$ The discretization in time is of first order whereas the central spacial discretizations are of second order. More on this in section 4.

<sup>&</sup>lt;sup>3</sup>Why we call this backwards in time is explained more clearly in section 4.

Recall that in the model by D&H the intensity process depended on the stock process only. Therefore we provide the index m to emphasize dependence on the spacial dimension. Using this scheme we get the explicit scheme

$$\frac{1}{\Delta t}\bar{A}_{n}^{m} = \left(\frac{1}{\Delta t} - \frac{\sigma^{2}}{\Delta x^{2}}\right)\bar{A}_{n+1}^{m} + \left(\frac{r_{n} + \lambda^{m} - \frac{1}{2}\sigma^{2}}{2\Delta x} + \frac{\sigma^{2}}{2\Delta x^{2}}\right)\bar{A}_{n+1}^{m+1} \\
+ \left(\frac{\sigma^{2}}{2\Delta x^{2}} - \frac{r_{n} + \lambda^{m} - \frac{1}{2}\sigma^{2}}{2\Delta x^{2}}\right)\bar{A}_{n+1}^{m-1} + e^{\sum_{k=0}^{n} r_{k}\Delta t + \sum_{k=0}^{n-1} \lambda^{m_{k}}\Delta t}C.$$
(3.46)

Note how we wrote  $\lambda^{m_k}$  to emphasize that  $\lambda$  may have changed along the way until it reached the level of m at time n. Multiplying both sides by  $\Delta t$  to find

$$\bar{A}_{n}^{m} = \left(1 - \Delta t \frac{\sigma^{2}}{\Delta x^{2}}\right) \bar{A}_{n+1}^{m} + \left(\frac{r_{n} + \lambda^{m} - \frac{1}{2}\sigma^{2}}{2\Delta x} + \frac{\sigma^{2}}{2\Delta x^{2}}\right) \Delta t \bar{A}_{n+1}^{m+1} + \left(\frac{\sigma^{2}}{2\Delta x^{2}} - \frac{r_{n} + \lambda^{m} - \frac{1}{2}\sigma^{2}}{2\Delta x^{2}}\right) \Delta t \bar{A}_{n+1}^{m-1} + e^{\sum_{k=0}^{n} r_{k}\Delta t + \sum_{k=0}^{n-1} \lambda_{k}^{m}\Delta t} \Delta t C.$$

$$(3.47)$$

Recall that the D&H tree model either goes up or down, it never remains on the same spacial level. Our expression however contains  $\bar{A}_{n+1}^m$ . To remove this term we take

$$\frac{\sigma^2 \Delta t}{\Delta x^2} = 1, \tag{3.48}$$

leaving us with

$$\bar{A}_{n}^{m} = \left(\frac{r_{n} + \lambda^{m} - \frac{1}{2}\sigma^{2}}{2\Delta x} + \frac{\sigma^{2}}{2\Delta x^{2}}\right)\Delta t\bar{A}_{n+1}^{m+1} + \left(\frac{\sigma^{2}}{2\Delta x^{2}} - \frac{r_{n} + \lambda^{m} - \frac{1}{2}\sigma^{2}}{2\Delta x^{2}}\right)\Delta t\bar{A}_{n+1}^{m-1} + e^{\sum_{k=0}^{n} r_{k}\Delta t + \sum_{k=0}^{n-1} \lambda_{n}^{m}\Delta t}\Delta tC.$$
(3.49)

Transforming back  $\bar{A}_{n+1} = e^{-\sum_{k=0}^{n} (r_k + \lambda_k) \Delta t} A_{n+1}$  we get the familiar form of the backwards recursion

$$A_n^m = e^{-r_n \Delta t} \underbrace{\left(\frac{r_n + \lambda^m - \frac{1}{2}\sigma^2}{2\Delta x} + \frac{\sigma^2}{2\Delta x^2}\right) \Delta t}_{p} e^{-\lambda^m \Delta t} A_{n+1}^{m+1} + e^{-r_n \Delta t} \underbrace{\left(\frac{\sigma^2}{2\Delta x^2} - \frac{r_n + \lambda^m - \frac{1}{2}\sigma^2}{2\Delta x^2}\right) \Delta t}_{q} e^{-\lambda^m \Delta t} A_{n+1}^{m-1} + e^{-r_n \Delta t} \Delta t C = e^{-r_n \Delta t} \left(pe^{-\lambda^m \Delta t} A_{n+1}^{m+1} + qe^{-\lambda^m \Delta t} A_{n+1}^{m-1} + \Delta t C\right).$$
(3.50)

Indeed, this is the exact same *form* as in the backward recursion. Now recall that in the backward recursion the probability of no default  $(1 - \xi^m)$  entered the recursion, with  $\xi^m$  corresponding to the probability of default within the interval  $\Delta t$ , depending upon the

level of the stock process. We recognize that

$$\lim_{\Delta t \to 0} e^{-\lambda^m \Delta t} = 1 - \xi^m, \tag{3.51}$$

due to the definition of the default intensity  $\lambda$ .

Note that just as in the D&H model, the sum these probabilities equals 1:

$$p+q = \frac{\sigma^2 \Delta t}{\Delta x^2} = 1. \tag{3.52}$$

Thus we have that q = 1 - p. What is left is to show is

- 1. that p and q satisfy the desired *probabilities*<sup>4</sup> of the D&H model, and
- 2. that the gridpoints at which the values of A are evaluated correspond to those of the D&H model.

Both, the discretization just provided and the D&H model, are of first-order in time. However, we saw that smart transformations let to the above particular results. More specifically, it is shown by the authors of [CRR79] that there are multiple ways in which to create a tree with similar results when time steps converge towards zero<sup>5</sup>. Therefore, direct comparison of our PDE discretization with the D&H model might not suffice: Both being first -order approximations, a fair comparison can only be achieved if they match at this particular order. To this end we start by converting the D&H model parameters by dropping orders larger than 1.

Recall from (2.3)

$$p_{tree} = \frac{e^{(r_n + \lambda^m)\Delta t} - e^{-\sigma\sqrt{\Delta t}}}{e^{\sigma\sqrt{\Delta t}} - e^{-\sigma\sqrt{\Delta t}}}$$
$$= \frac{e^{(r_n + \lambda^m)\Delta t} - e^{-\sigma\sqrt{\Delta t}}}{2\sinh(\sigma\sqrt{\Delta t})}$$
$$= \frac{1}{2} \left( e^{(r_n + \lambda^m)\Delta t} - e^{-\sigma\sqrt{\Delta t}} \right) \operatorname{csch}(\sigma\sqrt{\Delta t}), \qquad (3.53)$$

where  $\operatorname{csch}(\cdot)$  corresponds to the hyperbolic cosecant, which has the Taylor expansion

$$\operatorname{csch}(x) = \frac{1}{x} - \frac{1}{6}x + \frac{7}{360}x^3 + \mathcal{O}\left(x^5\right).$$
(3.54)

<sup>&</sup>lt;sup>4</sup>We shall later see that not really probabilities result from this choice of p and q, but in order to show equivalence between both models we only have to show that their limiting values correspond, whether these are interpretable as probabilities or not.

<sup>&</sup>lt;sup>5</sup>Referring to the original paper of the CRR model[CRR79], they show that their choice of parameters are deduced from the limiting scenario of time towards zero, that is,  $u = e^{\sigma\sqrt{\Delta t}}$ ,  $d = \frac{1}{u}$  correspond exactly to Black-Scholes only for  $\Delta t \to 0$ . However, under these conditions they might just as well use a first-order discretisation of these parameters, but for convenience reasons they leave them continuous.
Using (3.54), we proceed with (3.53):

$$p_{tree} = \frac{1}{2} \left( e^{(r_n + \lambda^m)\Delta t} - e^{-\sigma\sqrt{\Delta t}} \right) \underbrace{\left( \frac{1}{\sigma\sqrt{\Delta t}} - \frac{1}{6}\sigma\sqrt{\Delta t} + \mathcal{O}\left(\Delta t^{3/2}\right) \right)}_{f(\Delta t)}$$
(3.55)

$$= \frac{1}{2} \left( (r_n + \lambda^m) \Delta t + \frac{1}{2} (r_n + \lambda^m) \Delta t^2 + \sigma \sqrt{\Delta t} - \frac{1}{2} \sigma^2 \Delta t + \frac{1}{6} \sigma^3 \Delta t^{3/2} + \mathcal{O} \left( \Delta t^2 \right) \right) f(\Delta t)$$
(3.56)

$$=\frac{(r_n+\lambda^m)\Delta t+\frac{1}{2}(r_n+\lambda^m)\Delta t^2+\sigma\sqrt{\Delta t}-\frac{1}{2}\sigma^2\Delta t+\frac{1}{6}\sigma^3\Delta t^{3/2}}{2\sigma\sqrt{\Delta t}}-\frac{1}{12}\sigma^2\Delta t+\mathcal{O}\left(\Delta t^2\right)$$
(3.57)

$$=\frac{1}{2} + \frac{(r_n + \lambda^m - \frac{1}{2}\sigma^2)\sqrt{\Delta t}}{2\sigma} + \mathcal{O}\left(\Delta t^{3/2}\right).$$
(3.58)

Now, we should confirm that  $p_{tree} = p + \mathcal{O}(\Delta t)$  when dropping high-order terms. Recall the required condition (3.48), such that

$$\Delta x = \sigma \sqrt{\Delta t}.\tag{3.59}$$

Next, we show that equality holds for both probabilities by solving for p:

$$p = \left(\frac{r_n + \lambda^m - \frac{1}{2}\sigma^2}{2\Delta x} + \frac{\sigma^2}{2\Delta x^2}\right)\Delta t$$
$$= \frac{1}{2} + \frac{r_n + \lambda^m - \frac{1}{2}\sigma^2}{2\sigma\sqrt{\Delta t}}\Delta t$$
$$= \frac{1}{2} + \frac{(r_n + \lambda^m - \frac{1}{2}\sigma^2)\sqrt{\Delta t}}{2\sigma}$$
(3.60)

Indeed, we see that (3.58) and (3.60) are equal in the limit  $\Delta t \to 0$ .

Lastly we confirm that the underlying gridpoints are the same either. For the PDE case we have a grid spacing as in (3.59), where we note this to be under log transformation of the underlying asset. For the trinomial case we have, with adjustment for log asset prices:

$$\Delta x_{tree} = \frac{1}{2} \left( \ln \left( e^{\sigma \sqrt{\Delta t}} \right) - \ln \left( e^{-\sigma \sqrt{\Delta t}} \right) \right) = \sigma \sqrt{\Delta t}, \tag{3.61}$$

where the term  $\frac{1}{2}$  comes from the fact that this distance is related to a jump from grid point [n+1, m-1] to [n+1, m+1] and thus skipping one grid point.

We have thus shown both methods to be consistent in the sense that they converge towards the same result.

# **3.5** Default and recovery choice

In section 3.3 the PDEs for both the premium and protection leg were determined. These are general expressions and, as was remarked, do not incorporate any specifics on the intensity and recovery processes. Next our modeling choices shall be discussed, in particular our recovery model shall look a little different compared to that of Das & Hanouna [DH09].

#### 3.5.1 Default

Das & Hanouna [DH09] choose to model the default process as a hybrid process, i.e. a combination of both a hazard model and a structural model. They model the hazard rate as being a variable depending upon the stock process,<sup>6</sup>  $\lambda = S^{-b}$ , where parameter b determines its size. As a convenient property we have that for  $b \ge 0$ , an increase of stock corresponds to a decrease in default intensity. This seems reasonable as we might expect an increase of stock to go hand in hand with better economical conditions of the entity. Thus we define in a similar fashion for our PDE model

**Definition 3.5.1** (Default for the PDE model). The implied default intensity at time t,  $\lambda_t$ , within the PDE framework is defined by

$$\lambda_t = \frac{1}{S_t^b},\tag{3.62}$$

where  $S_t$  the stock value at time t and  $b \in \mathbb{R} \cap [0, 2.5]$ .

**Remark 3.5.1.** The restriction of b to the real interval [0, 2.5] will be discussed in section 5.1.2.

#### 3.5.2 Recovery

Recall that in [DH09] the recovery is modeled by the relationship

$$\rho \equiv \rho \left( a_0 + a_1 \xi \right), \tag{3.63}$$

with  $\xi$  the probability of default at the point of evaluation. This seems a decent model, taking into account a possible correlation between the recovery and the default intensity. However it has some drawbacks:

 $<sup>^{6}</sup>$ Note that it is the power of this model to let the intensity depend upon the stock as this makes the intensity dynamic instead of static.

1. The D&H model incorporated fairly large time steps, such that  $\xi$  obtained a 'proper value' larger than zero. In our model we are interested in letting the time steps converge towards zero, such that the probability of a default converges to zero together with it and the recovery becoming nearly constant. Instead we naturally replace the probability of default  $\xi$  by the default intensity  $\lambda$  (which is still related to the probability of default):

$$\rho \equiv \rho \left( a_0 + a_1 \lambda \right). \tag{3.64}$$

We'd like to emphasize another (intuitive) improvement that this particular model attains. Recall that there is a statistical relation between defaults and the size of the recovery, more specifically, when the default rate increases the recovery seems to decrease. This relationship is found from historical realizations of defaults. Thus we observe a correlation between the default probability given survival up till the point of default and the recovery. And this is exactly what we correlated in our recovery model, as from appendix D we know that when  $t > t_0$ , where  $t_0$  is today,

$$P_t(\tau \in (t, t+dt] | \tau > t) = \lambda_t dt.$$
(3.65)

This model however, still has two drawbacks:

- 2. The influences of parameter  $a_0$  and  $a_1$  have some overlapping, resulting in less convenient interpretability. As an example: suppose we want the initial recovery, that is  $\rho_0 = \Phi(a_0 + a_1\lambda)$  set to 0.4. Then there are many combinations of  $a_0$  and  $a_1$  satisfying this requirement. The question 'what does  $a_0$  or  $a_1$  actually tell us?' is thus somewhat involved.
- 3. The calibration process becomes much more involved due to the increased optimalisation domain. This recovery model leaves both parameters  $a_0$  and  $a_1$  as unbounded. A larger domain needs a more comprehensive search in order to find the optimal solution. Furthermore, due to the overlapping in the previous point, more local solutions were found using that particular approach, thus requiring time consuming global searches.

A modeling choice which we prefer and therefore present as alternative is stated in definition 3.5.2

$$\rho \equiv \rho \left( a_0 + a_1 \left( \frac{\lambda - \lambda_0}{\lambda_0} \right) \right), \qquad (3.66)$$

where  $\rho$  is some function that manages the value of the recovery and  $\lambda_0 = \left(\frac{1}{S_0}\right)^b$ . With this model the two drawbacks described before are no longer present. Parameters  $a_0$  and

 $a_1$  have clear interpretations which we will describe in section 3.7.2. Furthermore, the domain of  $a_0$  can be restricted in a rational way.

There are various models/functions that might be chosen to manage the recovery. Das & Hanouna [DH09] present three different types, namely the Probit, Logit and Arctan models. Each of these has a different influence upon the recovery. A more detailed discussion about their behavior and underlying assumptions can be found in appendix A.3. Their functional forms look like:

$$\rho^{(Probit)}(x) = \Phi(x) (\equiv \text{Normal cdf}),$$

$$\rho^{(Logit)}(x) = \frac{1}{1 + e^{-x}},$$

$$\rho^{(Arctan)}(x) = \frac{1}{2} \left( \arctan(x) \cdot \frac{2}{\pi} + 1 \right),$$
(3.67)

such that  $\rho : \mathbb{R} \to (0, 1)$ .

In this thesis we shall only look at the Probit model and we'd like to leave the other models open for future research. On a large data set it would be nice to see whether one of these present a better fit, possibly informing us about the underlying market assumptions.

Thus summarizing as a definition:

**Definition 3.5.2.** The recovery at time t,  $\rho_t$ , within the PDE framework is defined by

$$\rho_t \equiv \Phi\left(a_0 + a_1\left(\frac{\lambda_t - \lambda_0}{\lambda_0}\right)\right),\tag{3.68}$$

where  $a_0 \in \mathbb{R}$ ,  $a_1 \in \mathbb{R}_{\leq 0}$  model parameters,  $\lambda_t$  the default intensity and  $\lambda_0 = S_0^{-b}$  the initial default intensity.

**Remark 3.5.2.** As will be shown in section 3.7.1, when allowing for positive correlation, i.e.  $a_1 > 0$ , unrealistic behavior is seen in the CDS term structure. As positive correlation has no practical foundation we set it to be negative in this thesis. As discussed in appendix A.3 future research might change this by looking at another recovery function, e.g. the Logit function.

# **3.6** Retrieving term structures by a Monte Carlo simulation

Among one of the reasons for our interest into the D&H model was its ability to provide term structures. It created these by the aid of formulas (2.8) and (2.11) which have their foundations in the probabilities of the jump-to-default tree.

In section 4 we shall describe how finite differences can be used to calibrate the PDE model. As was shown the PDE model is a generalization of the D&H model, but it will not provide a tree as it did in the D&H model. As a result, after calibrating the PDE model we have no straightforward method to create the corresponding term structure.

To solve this issue we implement a Monte Carlo simulation providing us with the term structures for both default intensity and recovery, see algorithm 1.

The foundation of this simulation approach lies in the fact that the default intensities and recoveries are governed by the behavior of the stock. Therefore we may use a Monte Carlo approach for the stock to simulate realisations of intensities and recoveries, and from these we may extract results such as the average intensities (providing the default probability density function) and the average value of the recovery at default.

There are two important comments we would like to make regarding the algorithm:

1. There is a good rationale for why we choose to observe exactly these two term structures, that is, the default probability (provided by the probability density function) and the recovery at default. Reasons being the following: for most practical cases these are the values of interest. Suppose for example a simple zero coupon bond with the pricing formula

$$V(t_0) = \mathbb{E}^{\mathbb{Q}} \left[ \mathbf{1}_{\{\tau > T\}} \big| \mathcal{G}_{t_0} \right] + \mathbb{E}^{\mathbb{Q}} \left[ \rho_{\tau} \mathbf{1}_{\{\tau \le T\}} \big| \mathcal{G}_{t_0} \right].$$
(3.69)

The values that we see here are the cumulative distribution function of the default (given all knowledge up till  $t_0$ ) and the recovery given a default. Another interesting note is that with this particular model we are also able to determine the second expression on the rhs of (3.69). The above algorithm can easily be applied/modified to find the mean of this multiplicative function. This can furthermore be extended to more complex formulas, such as CVA.

2. One may wonder why in the case of a defaulted stock we set  $\lambda$  to zero whereas for  $\rho$  we leave it undefined. The reason being that we'd like to retrieve the probability density function, which is given by (D.11). Now, recall that

#### Algorithm 1 Monte Carlo simulation for retrieving term structures

1: N % amount of steps per year 2: T % amount of years to simulate 3: M % number of simulations 4: dt = 1/N; % step size in years 5:for m:=1 to M do 6: 7:  $S(m,1) = S_0$ 8: for t:=1 to N\*T do if S(m,t) == 0 then 9:  $\lambda(m,t) = 0$ 10:  $\rho(m,t) = [] \%$  Empty 11: else 12: $\lambda(m,t) = S(m,t)^{-b}$ 13: $\rho(m,t) = \Phi\left(a_0 + a_1 * \left(\frac{\lambda(m,t) - \lambda(m,0)}{\lambda(m,0)}\right)\right) \% \text{ Normal CDF}$ 14: 15:end if 16: $dI(t) \sim B\left(1 - e^{\lambda dt}\right)$  % Bernoulli distribution 17: $dW(t) \sim N(0, dt)$  % Normal distribution 18: $noDefault = ((r(t) + \lambda(m, t)) * S(m, t)) * dt + \sigma * S(m, t) * dW(t)$ 19:S(m, t+1) = S(m, t) + noDefault - (S(m, t) + noDefault) \* dI(t)20: end for 21:22: end for 23: % Term structures  $\Lambda(t) = mean(\lambda(m, \cdot)) P(t) = mean(\rho(m, \cdot))$ 

$$P_{t_0}(\tau \in (t, t+dt]) = \mathbb{E}^{\mathbb{Q}}\left[\lambda_t e^{-\int_{t_0}^t \lambda_u du} dt \Big| \mathcal{G}_{t_0}\right], \qquad (3.70)$$

Thus, within this simulation framework, the probability density function can be found by determining the probability of default over a very small interval divided by the size of the interval. Thus while simulating paths: in case no default occurred, the default probability is simply the intensity, whereas if a default did already occur, there is no longer any default probability, hence  $\lambda = 0$ .

In the case of recovery we are interested in the recovery *at default*, thus now we should neglect all defaulted stocks as these cannot default again.

# 3.7 Parameter contribution

#### 3.7.1 To the CDS spread term structure

In this section we will look more closely into what the model parameters  $a_0$ ,  $a_1$  and b, add to the resulting term structure of CDS spreads.

To get a proper insight we begin with two a priori notes:

- 1. Recall that parameter b determines the default intensity based on  $\lambda_t = \frac{1}{S_t^b}$ . What we may note immediately is that we should distinguish two different scenarios:
  - A When  $S_t > 1$ , then the default intensity is decreasing for increasing b. For our investigation we take  $S_0 = 100$ .
  - B When  $S_t < 1$ , then the default intensity is increasing together with b. For our investigation we take  $S_0 = 0.75$ .

The first scenario is of more practical importance, as stock values are usually larger than one.

For both cases we fix the volatility at  $\sigma = 0.5$  and the interest rate at r = 0.03. To get a proper understanding of the influences of each parameter we will vary b between 0.25 to 1.5 and look at 4 different sets of  $\{a_0, a_1\}$ , namely:  $\{-0.53, -7.5\}$ ,  $\{-0.53, 7.5\}$ ,  $\{0.53, -7.5\}$  and  $\{0.53, 7.5\}^7$ .

2. As we shall see there is interesting behavior when the correlation is positive, without real practical interpretability. Combined with the fact that positive correlation itself has no practical foundation we will exclude this type of correlation from our simulations in the section with results.

**Remark 3.7.1.** The results presented are found using implicit finite differences with  $\{0 \le S \le 2000, \Delta S = 0.5\}$  and  $\{0 \le t \le 5, \Delta t = 1/52\}$ . More on the finite differences is presented in section 4.

**Scenario A** In figure 3.2 we see the term structures of CDS spreads for several parameter sets. In particular, each color belongs to a specific choice of parameter b, and at the same time, the differences observed from the left image to the right correspond to differences induced by the different values of  $a_0$  and  $a_1$ .

The first three main consistent observations are

Main observation 3.7.1 (Influence  $a_0$  on CDS spreads).  $a_0$  provides initial vertical shifts: increasing  $a_0$  lowers the initial spread.

<sup>&</sup>lt;sup>7</sup>The choice of  $a_0 = \pm 0.53$  is not a random one: It approximately corresponds to an initial recovery of either 70% or 30%.



FIGURE 3.2: The term structure of CDS spreads given by the PDE model. We modify parameter b between 0.25 and 1.25, while looking at 4 different sets of  $\{a_0, a_1\}$  and  $S_0 = 100, \sigma = 0.5, r = 0.03$ .

Main observation 3.7.2 (Influence  $a_1$  on CDS spreads).  $a_1$  influences the terminal value: increasing  $a_1$  lowers the final spread.

Main observation 3.7.3 (Influence b on CDS spreads). b provides both height and form.

These contributions persist throughout all results and are therefore important ones.

Next we will state some of the additional observations and proceed by explaining their causes. This will provide us with extra intuition for the dynamics. 1) Decreasing *b* results in higher spreads, 2) when  $a_1 < 0$ , smaller *b* appears to give downward trends on the CDS term structure, whereas higher *b* turns out to result in upward trends and 3) observation 2 doesn't seem to hold when  $a_1 > 0$ , instead the opposite seems to be true.

1. Lower b results in higher spreads.

Let  $b_2 > b_1$  then initial default intensities corresponding to these parameters for b result in

$$\left(\frac{1}{S_0}\right)^{b_2} < \left(\frac{1}{S_0}\right)^{b_1}.\tag{3.71}$$

This follows from the fact that  $S_0 > 1$ , Telling us that if we decrease b then we expect the initial default intensity,  $\lambda_0$ , to be higher, and as a result - leaving  $a_0$  and  $a_1$  unchanged - we get higher spreads.

2. When  $a_1 < 0$ , lower and higher values for b result in respectively downward and upward trends. This observation is indeed a correct one, except when  $a_1 > 0$ . The trend of the CDS spreads is determined by the value of the spreads over time. Recall the general formula for the spreads, (1.17). The CDS spread at time t with maturity  $T_b$  may be approximated: Let  $t_0 = t$ ,  $t_N = T_b$  and  $t_i - t_{i-1} = \Delta t$ , for i = 0, 1, ..., N. Then supposing continuous premium payments, we discretize the CDS value by

$$C_t(T_b) = \frac{\sum_{i=0}^{N-1} e^{-r_i(t_i-t)} (1-\rho_i) P_t(\tau \in (t_i, t_{i+1}])}{\sum_{i=1}^{N} e^{-r(T_i-t)}}.$$
(3.72)

For  $\Delta t$  small enough it holds that the probability can be approximated by its value on the probability density function of the related Cox process, (D.11), such that

$$P_t(\tau \in (t_i, t_{i+1}]) \approx \mathbb{E}^{\mathbb{Q}} \left[ \lambda_{t_i} e^{-\int_t^{t_i} \lambda_u du} \Delta t \Big| \mathcal{G}_t \right].$$
(3.73)

Now clearly the probability density function sums up to one, i.e.

$$\int_{t}^{\infty} \mathbb{E}^{\mathbb{Q}} \left[ \lambda_{s} e^{-\int_{t}^{s} \lambda_{u} du} \middle| \mathcal{G}_{t} \right] ds = 1.$$
(3.74)

Furthermore, as  $\lambda_0 := \lambda_t$  is deterministic at time t we have

$$P_t(\tau \in (t, t_1]) \approx \lambda_0 \Delta t. \tag{3.75}$$

From expression (3.74) and (3.75) we may understand the behavior in the following two scenarios.

(a)  $\lambda_0$  small.

In this case it seems reasonable for the probability of default to further increase, as recalling (3.74) it should form a probability density function. However, convergence towards zero is required on the long term.

What this reasoning thus tells us is that for some s > t we may expect

$$P_t(\tau \in (t, t + \Delta t) \le P_t(\tau \in (s, s + \Delta t))$$
  

$$\leftrightarrow \quad \lambda_t \le \mathbb{E}^{\mathbb{Q}} \left[ \lambda_s e^{-\int_t^s \lambda_u du} \middle| \mathcal{G}_t \right]$$
  

$$\leftrightarrow \quad \lambda_t \le \mathbb{E}^{\mathbb{Q}} \left[ \lambda_s \middle| \mathcal{G}_t \right].$$

This however might trigger the question: if the stock is always expected to increase within the risk neutral framework, and we model the intensity as  $S^{-b}$ , then how could the expected intensity possibly be increasing knowing that the stock is becoming more 'safe'.

That may seem counter intuitive, but it's only a consequence of the difficulty



FIGURE 3.3: Term structures for default intensity and recovery for two different parameter sets. The first parameter set resulted in a downward CDS term structure whereas the second results in an upward.

of having a proper intuition for Jensen's inequality: note that  $f(x) = x^{-b}$  is a convex function, such that

$$\mathbb{E}^{\mathbb{Q}}\left[S^{-b}\right] \ge \left(\mathbb{E}^{\mathbb{Q}}\left[S\right]\right)^{-b}.$$
(3.76)

Even though the rhs speaks to our intuition it is an underestimation of the true value. This behavior can be confirmed by a simple example: suppose the stock to have initial value  $S_0 = 10$  and at time 1 it either increases with probability p = 0.52 to  $S_1 = 11$  or decreases with probability 1 - p = 0.48 to  $S_1 = 9$ . Then we have  $\mathbb{E}^{\mathbb{Q}}[S_1] = 10.04$  (thus 0.4% interest rate). Now, let b = 1 such that  $\lambda_0 = 0.1$  and subsequently

$$\mathbb{E}^{\mathbb{Q}} \left[ S^{-b} \right] = 0.52 \cdot \left( \frac{1}{11} \right) + 0.48 \cdot \left( \frac{1}{9} \right) \approx 0.1006 
\left( \mathbb{E}^{\mathbb{Q}} \left[ S \right] \right)^{-b} = \frac{1}{10.04} \approx 0.0996.$$
(3.77)

Thus we see that even in this simple example an expected increase of the stock might correspond to a simultaneously increasing default intensity.

(b)  $\lambda_0$  large.

Using the same reasoning as for small  $\lambda_0$  we may now expect the probability of default to decrease instead.

The above two statements are no proofs but mere intuition. A Monte Carlo simulation confirming this intuition can be found in figure 3.3. We provided term structures for the default probability density function and recovery corresponding to a downward and upward trending CDS term structure. Indeed the above reasoning seems to hold.

3. In contrast to the observation in part 2, when  $a_1 > 0$  the exact opposite seems to occur.

To be more specific, the initial trend seems to be upwards for small b and downwards

when b is large. This observation is a little more involved and a more in deep discussion is presented in appendix A.2. We therefore summarize: due to the positive correlation, decreasing intensity goes together with increasing LGD. However, for very small LGD a decrease of the intensity produces a relatively very large increase of LGD, so large that it increases the spread. More specifically we show that when  $a_1 > \frac{\Phi(-a_0)}{\phi(a_0)}$  then spreads will initially increase, whereas when  $a_1 < \frac{\Phi(-a_0)}{\phi(a_0)}$ , then they will decrease. There thus seems to be an analytically tractable *switching line*. One may verify that the parameter set as in the rightmost sub figure of figure 3.2 indeed corresponds to an increase. This behavior is not persistent and when a certain LGD has been exceeded the spread will start moving downward.

We have not carefully studied this behavior within the Logit or Arctan model, but we expect it to be less present. As our results in section 5 are based on the Probit model we shall neglect positive correlation as it seems more like a flaw of the model than a good representation of the market.

**Scenario B** As announced before, the behavior of the default intensity changes together with  $S_0$ , in particular when  $S_0$  is smaller than 1, then increasing *b* corresponds to decreasing intensity. Besides this there are actually no big differences. What we shall see is that this scenario functions more as a magnifier of results in scenario A. In figure 3.4 we show, analogously to scenario A, CDS term structures to several parameter sets. The first and third parameter set show us that indeed larger *b* induces the initial CDS spread to start higher. Because the initial intensity is of such magnitude and the probability of default has to remain a probability density function, it quickly drops. This corresponds to what was noted before: decreasing spreads for higher initial default intensities (when below the switching line). In contrast to scenario A, the initial spreads remain closer to each other, as a consequence of the fact that the recovery remains close the zero in all cases. Therefore the difference is only a result of the change the default intensity.

In the second and fourth we see two interesting things

1. Spreads have an upward term structure, where higher default intensities induce larger initial gradients.

This is a consequence of all spreads lying above the switching line. Furthermore the spread related to the highest default intensity lies the furthest above this line and therefore increases most significantly (see appendix A.2).

2. The term structures drop sooner in the second sub figure than in the fourth sub figure (see figure 3.4).

Indeed, the closer above the switching line the spread lies, the sooner the drop starts. In the second image we have  $a_0 = -1$  whereas in the fourth  $a_0 = 1$ . As



FIGURE 3.4: Term structure of CDS spreads given by the PDE model. We modify parameter b between 0.25 and 1.75, while looking at 4 different sets of  $\{a_0, a_1\}$  and  $S_0 = 0.75, \sigma = 0.5, r = 0.03$ .

we recall from the switching line it lies higher for smaller  $a_0$ . This can intuitively be understood from the fact that under smaller  $a_0$  the model sets the recovery to be smaller. And as we recall the switching line occurs at a particular point on the normal distribution where differences alleviate (around small recoveries).

#### 3.7.2 To the default and recovery term structures

In the previous section we discussed the contribution of the model parameters to the CDS spread term structure. Eventually however we are also interested in the resulting term structures of the implied default and recovery. We shall look into that now, starting with the default intensity. During this section we take  $S_0 = 10, \sigma = 0.5$  and r = 0.03 constant. Furthermore the results presented are determined by the Monte Carlo approach introduced in section 3.6.

**Default intensity** In figure 3.5 we plot the expected future default intensities for several values of b. This gives us an impression of what parameter b exactly does. The most important thing to note is that the term structure is initially downwards sloping for small b and upwards sloping for large b. Eventually all term structures will converge towards zero.

**Recovery** A priori we expect the following behavior of parameters  $a_0$  and  $a_1$ : Parameter  $a_0$  fixes the initial recovery and  $a_1$  determines the direction and speed at which the



FIGURE 3.5: Term structures for the default intensities implied by the jump to default stock price model;  $S_0 = 10, \sigma = 0.5, r = 0.03$ .

term structure moves away from the initial recovery. There are two main possible term structures for the recovery depending on the pre-default intensity:

- 1. The pre-default intensity drifts upwards This implies the recovery to move downwards if  $a_1 < 0$  and upwards if  $a_1 > 0$ .
- 2. The pre-default intensity drifts downwards Now the recovery does the opposite, it moves downwards if  $a_1 < 0$  and upwards if  $a_1 > 0$ .

Any term structure is one of these or a combination of both.

Now let's look at a simulation of several different recoveries, see figure 3.6. The left and right side of the figure show respectively the results of moving  $a_0$  and  $a_1$ . Our expectations seem to be right:

Main observation 3.7.4 (Influence  $a_0$  on the recovery term structure).  $a_0$  controls the initial recovery by setting it higher for larger  $a_0$ .

Main observation 3.7.5 (Influence  $a_1$  on the recovery term structure). The sign of  $a_1$  fixes the direction of movement away from the initial recovery whereas the norm of  $a_1$  increases the absolute value of the derivative with which it moves away.

There are however two points worth noting.



FIGURE 3.6: Term structures for the recovery at default;  $S_0 = 100, \sigma = 0.5, r = 0.03, b = 1.$ 

1. For small  $a_0$  we note upward drifting of the term structure, whereas for large  $a_0$  it drifts downwards.

This is again a consequence of the fact that the expectation of a function is not equal to the function of an expectation. For all scenarios the expected pre-default intensity is the same, and the recovery is a function of this pre-default intensity.

We may verify that the expected pre-default intensity is increasing and thus the reason for the upward movement when  $a_0$  is low is a consequence of the derivative of the normal distribution: the simulations that result in upward moving default intensities influence the recovery to go down, however the downward drift is slow due to the derivative which starts approximating zero. In contrast, the default intensities moving downwards result in upwards moving recoveries with a larger derivative, resulting in larger values more heavily impact on the expectation. The opposite holds for large  $a_0$ .

It is exactly for this reason why we observe the initial upward bending trends in the third image of figure 3.2: the larger initial recovery results in a quicker decent of the recovery bringing with it the upward bend.

2. A similar result seems notable when moving  $a_1$ . At initiation,  $a_1$  with larger absolute values have a larger derivative, which eventually flattens: The larger the initial derivative the quicker it seems to flatten. This is clearly seen for dark blue line b = -10 which has large initial negative derivative but then quickly seems to reach its lowest point to continue drifting upwards. Furthermore, the second derivative of the recovery function at 50% is exactly zero and thus by Jensen's inequality, there is exact equality between the function of the expectation and the expectation of the function. Such that, heuristically, for a very small time step we may expect (let t = 0)

$$\mathbb{E}_{t}^{\mathbb{Q}}\left[\rho_{t+dt}\right] = \mathbb{E}_{t}^{\mathbb{Q}}\left[\Phi\left(a_{0}+a_{1}\left(\frac{\lambda_{t+dt}-\lambda_{t}}{\lambda_{t}}\right)\right)\right] \approx \Phi\left(a_{0}+a_{1}\left(\frac{\mathbb{E}_{t}^{\mathbb{Q}}\left[\lambda_{t+dt}\right]-\lambda_{t}}{\lambda_{t}}\right)\right),\tag{3.78}$$

explaining the initial downward movement. Subsequently the recovery function pushes upwards as the further upward drifting intensities have less influence on the recovery as their derivative becomes smaller.

Summarizing, for  $a_1 < 0$  the quickly moving downwards at initiation is a consequence of the upward moving pre-default intensity together with the concave functional form of the normal distribution around 50%. Next the moving upwards is related to convex functional form for smaller recovery values. Similarly we understand the results when  $a_1 > 0$ .

# 3.8 Dynamic Stability

We shall now examine whether the framework is overly sensitive to the input. In section 3.7.2 we have already seen that changing the parameter set does not result in abrupt changes within the term structure. In the section 5 we provide a description of the calibration process, however for now we will show some calibration results based on shocking input parameters by one percent. These experiments are performed on datasets for Banco do Brasil, Shell and Air France - KLM, see table H.1.

We are given a stock value, volatility and market spreads for which we have found optimal results (a parameter set and term structures) by means of the calibration process. Now, each of these input components will be shocked by one percent and new results are found by calibration. We did this for three firms, for which we shall discuss the calibration in more detail in section 5.3. The results are portrayed in table 3.1. Across all experiments we see that the calibrated parameters and the resulting term structures do not display any dramatic changes. Air France - KLM represents most risky data followed by Banco do Brasil and finally Shell. It seems that more riskier data is more sensitive to the input.

Overall, these results seems to suggest a stable model.

\_

Banco do Brasil	Base case	Stock shock	Volatility shock	Spreads shock
$S_0$	24.35	24.49	24.35	24.35
$\overset{\circ}{\sigma}$	0.437	0.437	0.441	0.437
$a_0$	0.870	0.878	0.816	0.894
$a_1$	-2.707	-2.735	-2.472	-2.833
b	1.016	1.011	1.023	1.011
RMSE	10.16	10.20	9.78	10.46
pd. diff. <sup>8</sup> (%)	-	0.72	2.39	1.52
$\rho \text{ diff.}^9 (\%)$	-	0.89	1.86	1.17
,				
Shell	Base case	Stock shock	Volatility shock	Spreads shock
$S_0$	28.01	28.29	28.01	28.01
σ	0.174	0.174	0.176	0.174
$a_0$	1.916	1.918	1.902	1.923
$a_1$	-5.932	-5.964	-5.752	-6.021
b	1.108	1.104	1.108	1.107
RMSE	3.37	3.38	3.32	3.44
pd. diff (%)	-	0.31	0.74	0.98
$\rho$ diff. (%)	-	0.26	0.18	0.46
Air France - KLM	Base case	Stock shock	Volatility shock	Spreads shock
$S_0$	7.62	7.69	7.62	7.62
$\sigma$	0.461	0.461	0.466	0.461
$a_0$	-0.077	-0.074	-0.039	-0.086
$a_1$	-0.027	-0.028	-0.028	-0.027
b	1.666	1.654	1.651	1.662
RMSE	6.00	6.13	6.27	6.24
pd. diff. (%)	-	1.12	3.04	1.40
$\rho$ diff. (%)	-	0.31	3.29	0.83

TABLE 3.1: Calibration of scenarios when shocking the stock, volatility and market spreads 1% upwards. The lower two rows correspond to respectively the maximum difference of probability density function of default with respect to the base case and the maximum difference of the recovery at default  $\rho$  with respect to the base case.

# 4

# Finite differences for the PDE model

"Simplicity is the final achievement. After one has played a vast quantity of notes and more notes, it is simplicity that emerges as the crowning reward of art."

- Frédéric Chopin , If Not God, Then What?

In section 3.3 we deduced two PDEs that represent our model towards pricing a CDS. PDEs are build up of partial derivatives of underlying components. Finite differences exploit this property by discretizing the dimensions of these components and the corresponding derivatives.

One may understand that depending on the choice of time discretization (e.g. forward or backwards) the approach is slightly modified. In our scenario both time and stock dimensions have to be discretized. As at maturity we have the initial condition, we have to iterate through time, such that iterations may bring with them errors (more on this in section 4.3), therefore, the choice of discretization is an interesting one.

There are two approaches that we'd like to highlight. To this end, let n = 0, 1, ... be a discretization of the time dimension

1. *Explicit* finite differences: the approach resulting from a forward Euler discretization in the time direction, i.e.

$$\frac{\partial y_n}{\partial t} = \frac{y_{n+1} - y_n}{\Delta t} + \mathcal{O}\left(\Delta t\right). \tag{4.1}$$

2. *Fully implicit* finite differences: the approach resulting from a backward Euler discretization in the time direction, i.e.



FIGURE 4.1: Examples of the explicit and implicit time discretization. In the implicit case the space elements are evaluated at a the time step on which elements are not yet determined.

$$\frac{\partial y_n}{\partial t} = \frac{y_n - y_{n-1}}{\Delta t} + \mathcal{O}\left(\Delta t\right). \tag{4.2}$$

The big Oh notation indicates both discretization to be of first order. This can easily be understood from

$$y(t\pm\Delta t) = y(t)\pm y'(t)\Delta t + \mathcal{O}\left(\Delta t^2\right) \quad \rightarrow \quad \begin{cases} y'(t) = \frac{y(t+\Delta t) - y(t)}{\Delta t} + \mathcal{O}\left(\Delta t\right), \\ y'(t) = \frac{y(t) - y(t-\Delta t)}{\Delta t} + \mathcal{O}\left(\Delta t\right), \end{cases}$$
(4.3)

Other types of discretizations may be deduced in a similar fashion, using the Taylor expansion.

We should remark that by *forward* in time we refer to moving away from the initial condition towards the desired point in time at which we seek a solution, this is not necessarily forward in time in the absolute sense, see figure 4.1 for a visualisation.

The differences between explicit and full implicit seem very small but have some important implications, in particular to the numerical stability of the problem: the explicit method is usually conditionally stable, whereas the implicit method is unconditionally stable. Recall that we mentioned numerical stability to be an issue in the D&H model. By using implicit finite differences we can work around this problem. Due to this last convenient result we will describe the fully implicit method more carefully in the next section.

**Remark 4.0.1.** When we take a single step in the time direction we call the finite difference scheme a one-step scheme. There are also multistep finite difference schemes, e.g. the Leapfrog scheme, which uses central discretization. Such schemes however require additional initial conditions, which is why we leave such schemes aside.

**Remark 4.0.2.** Another scheme which also uses central discretization in time is the Crank Nicolson scheme, a second order scheme in time for the parabolic equations. Although

it uses central discretization it remains a one-step scheme after some modifications. This scheme is applicable to our model, however it did not seem to improve convergence.

This section is build up as follows. First we will discuss the fully implicit finite differences approach, which includes derivations of the PDE's boundary conditions. After describing this scheme, in section 4.2 we will look at *grid stretching*, which is a different way of ordering the grid points in the spacial direction with the aim of improving results. Next we investigate upon the convergence and in particular the stability of the explicit scheme, such that we may conclude the D&H model to be unstable if condition (2.14) is violated. Finally we discuss the choice of the discretization parameters in section 4.4.

# 4.1 Fully Implicit finite differences approach

We shall now show how to deduce the fully implicit finite difference scheme and how to implement it. In order to price CDSs by means of our PDE model we need to evaluate both the premium and protection leg PDEs. Recall these to be

$$\frac{\partial A_{t-}}{\partial t} + (r_t + \lambda_t) S_{t-} \frac{\partial A_{t-}}{\partial S} + \frac{1}{2} \sigma^2 S_{t-}^2 \frac{\partial^2 A_{t-}}{\partial S^2} - (r_t + \lambda_t) A_{t-} + C = 0, \qquad (4.4)$$

defined on  $S \in \mathbb{R}_{>0}$  and  $t \in (0, T)$ , with initial and boundary conditions,

$$\begin{cases} A(T,S) = 0, \\ A(t,0) = 0, t \in [0,T]. \end{cases}$$
(4.5)

And the protection leg

$$\frac{\partial B_{t-}}{\partial t} + (r_t + \lambda_t) S_{t-} \frac{\partial B_{t-}}{\partial S} + \frac{1}{2} \sigma^2 S_{t-}^2 \frac{\partial^2 B_{t-}}{\partial S^2} - (r_t - \lambda_t) B_{t-} + \lambda_t \text{LgD}_t = 0, \quad (4.6)$$

defined on  $S \in \mathbb{R}_{>0}$  and  $t \in (0, T)$ , with initial and boundary conditions,

$$\begin{cases} B(T,S) = 0, \\ B(t,0) = 0, t \in [0,T]. \end{cases}$$
(4.7)

Clearly, both have similar form and their schemes are therefore virtually identical. To this end we only show the scheme for the premium leg PDE. The initial and boundary conditions shall be discussed for both cases.

#### 4.1.1 Initial and boundary conditions

In the next section we will discretize the dimensions of both time and stock. For now we only want to note that the stock is an infinite dimension, there is no bound on a maximum stock value in our PDE. However, in the computer we have to implement some upper bound: the upper bound for the stock will be  $S_{\text{max}}$ , some yet to be determined value. In this section we determine the boundary values for both premium and protection leg at S = 0 and  $S = S_{\text{max}}$ , and the initial condition.

**Premium leg** The initial condition and the boundary condition at S = 0 are straightforward as they follow directly from the definition of the model. However the boundary condition at  $S = S_{max}$ , corresponding to the truncation of the stock space requires more attention. What we are looking for is an accurate estimate for  $A(t, S_{max})$ . Recall that A(S,t), the value of the expected premium leg given stock value S at time t, is defined by (3.2). A priori we might expect a good estimate to be payments until maturity, i.e.

$$A(t, S_{max}) \approx \mathbb{E}^{\mathbb{Q}} \left[ \int_{t}^{T} e^{-\int_{t}^{s} r_{u} du} C ds \middle| \mathcal{G}_{t} \right], \quad t < T,$$

$$(4.8)$$

because the probability of a default becomes very small when the stock price increases sufficiently. This expectation is indeed correct and can be formalized by means of the next theorem.

**Theorem 4.1.1.** Let the default intensity be defined by  $\lambda = \frac{1}{S^b}$ , where S the stock value. Then the C-continuous paying expected premium leg with maturity T and interest rate  $r_t$  satisfies

$$\lim_{S_t \to \infty} \operatorname{Prem}\left(t, T\right) = \mathbb{E}^{\mathbb{Q}}\left[\int_t^T e^{-\int_t^s r_u du} C ds \middle| \mathcal{G}_t\right].$$
(4.9)

**Proof.** Recall that the general expression for the expected premium leg at time t satisfies<sup>1</sup>

$$\mathbb{E}^{\mathbb{Q}}\left[\int_{t}^{T} e^{-\int_{t}^{s} r_{u} du} \mathbf{1}_{\{\tau > s\}} C ds \middle| \mathcal{G}_{t}\right] = \mathbb{E}^{\mathbb{Q}}\left[\int_{t}^{T} e^{-\int_{t}^{s} r_{u} du} C ds \middle| \mathcal{G}_{t}\right] P_{t}(\tau > T) + \mathbb{E}^{\mathbb{Q}}\left[\int_{t}^{T} e^{-\int_{t}^{s} r_{u} du} \mathbf{1}_{\{\tau > s\}} C ds \middle| \mathcal{G}_{t}\right] (1 - P_{t}(\tau > T)).$$

$$(4.10)$$

As both expectations are bounded it holds that the desired result follows immediately if

$$\lim_{S_t \to \infty} P_t(\tau > T) = 1. \tag{4.11}$$

The validity of (4.11) is proven in lemma A.1.1.

**Protection leg** Similar as for the premium leg case the initial condition and the boundary condition at S = 0 are straightforward. The condition at  $S = S_{\text{max}}$  follows by a same approach: Intuitively we may expect that if the stock value increases towards infinity, the probability of default becomes negligible such that the expected value of the protection leg becomes zero, providing us with the estimate

$$B(t, S_{\max}) \approx 0, \quad t < T. \tag{4.12}$$

 $<sup>^{1}</sup>$ Recall that we assume there to be no wrong-way risk, i.e. interest rate and default times are independent.

This statement can indeed be formalized by the following theorem

**Theorem 4.1.2.** Let the default intensity be defined by  $\lambda = \frac{1}{S^b}$ , where S the stock value. Then the expected protection leg with maturity T satisfies

$$\lim_{S_t \to \infty} \operatorname{Prot}\left(t, T\right) = 0 \tag{4.13}$$

**Proof.** Recall that the general expression for the expected premium leg at time t satisfies

$$\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_{t}^{\tau}r_{u}du}(1-\rho_{\tau})\mathbf{1}_{\{\tau\leq T\}}\middle|\mathcal{G}_{t}\right] = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_{t}^{\tau}r_{u}du}(1-\rho_{\tau})\middle|\mathcal{G}_{t}\right](1-P_{t}(\tau>T)). \quad (4.14)$$

As the expectation is bounded it holds that the desired results follow immediately from lemma A.1.1.  $\hfill \Box$ 

### 4.1.2 Discretization

**Specifying the grid** Before anything we have to discretize the dimensions that we are working with. In our case these are time and space (stock). We have to truncate the stock by introducing a maximum stock value,  $S_{max}$ . Furthermore we take N steps within one time unit (in this thesis, one year) and M steps in the stock direction, such that the grid consists of points  $n = 0, 1, ..., T \cdot N$ , m = 0, 1, ..., M, see figure 4.2. In the time direction steps have size  $\Delta t = \frac{1}{N}$  and in the stock direction  $\Delta x = \frac{S_{max}}{M}$ , such that  $A_n^m \equiv A(n\Delta t, m\Delta x)$ .

**Derivative truncation** Each partial derivative needs to be discretized, for the time component we use backwards Euler and for the stock component central discretization, i.e. around  $(n\Delta t, m\Delta x)$ :

$$\frac{\partial A(n\Delta t, m\Delta x)}{\partial t} = \frac{1}{\Delta t} \left( A_{n+1}^m - A_n^m \right) + \mathcal{O} \left( \Delta t \right)$$

$$\frac{\partial A(n\Delta t, m\Delta x)}{\partial S} = \frac{1}{2\Delta x} \left( A_n^{m+1} - A_n^{m-1} \right) + \mathcal{O} \left( \Delta x^2 \right)$$

$$\frac{\partial^2 A(n\Delta t, m\Delta x)}{\partial S^2} = \frac{1}{\Delta x^2} \left( A_n^{m+1} - 2A_{n+1}^m + A_n^{m-1} \right) + \mathcal{O} \left( \Delta x^2 \right).$$
(4.15)

Confusingly this might seem a forward Euler discretization. However, during the evaluation of our PDE we start at the final state and work towards the time of interest, which is a backward movement through time. We know the value of  $A_{n+1}^m$  and we are interested in  $A_n^m$ . Now one may understand this to be a valid backward Euler discretization. Note two things with respect to (4.4):



FIGURE 4.2: The finite differences grid when discretized for 1 year in time.

- 1. The PDE states  $A_t$  only, as if it does not depend upon  $S_t$ . However  $S_t$  on its turn depends on t such that the PDE is a correct representation of  $A(t, S_t) \equiv A(t, S)$ . Note also that the intensity  $\lambda$  depends only on the stock value.
- 2. The index specifies t-, this is however a specification required due to the continuous nature of the equation. If at time t there is a default, then differentials become undefined. However, on the finite differences grid a default occurs only within  $(n\Delta t, (n+1)\Delta t)$ , never at the points  $n\Delta t$  or  $(n+1)\Delta t$ , because at these particular points we are either at S = 0 or S > 0, such that we know whether a default occurred or not. In the continuous case these intervals are nonexistent such that we require the notation t- to emphasize that no default occurred yet. We may thus safely neglect the minus sign.

Substituting (4.18) into (4.4) gives us

$$\frac{A_{n+1}^m - A_n^m}{\Delta t} + (r_n^m + \lambda^m) \, m \Delta x \frac{A_n^{m+1} - A_n^{m-1}}{2\Delta x} + \frac{1}{2} \sigma^2 m^2 \Delta x^2 \frac{A_n^{m+1} - 2A_n^m + A_n^{m-1}}{\Delta x^2} + C = 0. \tag{4.16}$$

Isolating  $A_{n+1}^m$  on the left side and ordering, such that

$$A_{n+1}^{m} = \left(1 + \Delta t \sigma^{2} m^{2} + \Delta t \left(r_{n}^{m} + \lambda^{m}\right)\right) A_{n}^{m}$$
$$-\frac{1}{2} \left(\Delta t \sigma^{2} m^{2} - \Delta t \left(r_{n}^{m} + \lambda^{m}\right)\right) A_{n}^{m+1}$$
$$-\frac{1}{2} \left(\Delta t \sigma^{2} m^{2} + \Delta t \left(r_{n}^{m} + \lambda^{m}\right)\right) A_{n}^{m+1} + \Delta t C.$$

$$(4.17)$$

To keep the overview we write

$$\alpha^m = \Delta t \sigma^2 m^2 
\beta^m_n = \Delta t \left( r^m_n + \lambda^m \right),$$
(4.18)

such that

$$A_{n+1}^{m} = \underbrace{\left(1 + \alpha^{m} + \beta_{n}^{m}\right)}_{d_{n}^{m}} A_{n}^{m} + \underbrace{-\frac{1}{2} \left(\alpha^{m} - \beta_{n}^{m}\right)}_{u_{n}^{m}} A_{n}^{m+1} + \underbrace{-\frac{1}{2} \left(\alpha^{m} + \beta_{n}^{m}\right)}_{l_{n}^{m}} A_{n}^{m-1} + \Delta tC, \quad (4.19)$$

for  $0 \le n < T \cdot N$  and 0 < m < M, where for m = 0 and m = M we have boundary conditions. In particular we can write (4.19) in matrix notation as

$$A_{n+1} = Q_n A_n + b_n + \Delta t C, \qquad (4.20)$$

with

Elements  $A_n^0$  and  $A_n^M$  are given by the boundary conditions, more specifically, let r be the deterministic future continuously compounded interest rate, then

$$A_n^0 = 0$$
 and  $A_n^M = \Delta t C \sum_{j=n}^N e^{-\sum_{k=n+1}^j r_k \Delta t}$ , (4.22)

where for  $A_n^M$  we use the discretized version of (4.8). A visualisation of this top boundary condition can be found in figure 4.3. Finally, we'd like to emphasize that all boundary conditions, except this top conditions of the premium leg, are equal to zero and do not require any extra attention. Also note that our implementation of these boundary conditions can be found in the constructor of the PDE Matlab class<sup>2</sup> in appendix (I.2.1)

 $<sup>^{2}</sup>$ Our implementation of the finite difference approach is object orientated programmed.



FIGURE 4.3: The boundary condition at  $S_{\max}$  given an annual spread of C. In this example, the final time node is supposed to be N

Now recall that we started at time T, i.e.  $n = T \cdot N$ , and we are interested in the premium leg at time zero, that is Prem  $(t = 0, T) = A_0$ . We can therefore use (4.20), to determine backwards

$$A_n = Q^{-1} \left( A_{n+1} - b - \Delta t C \right). \tag{4.23}$$



FIGURE 4.4: Going from a regular spaced grid to a grid stretched grid in which points accumulate around 0.5.

# 4.2 Grid stretching

Grid stretching transforms an underlying grid in such a way that it may improve in particular two things:

- 1. An important area may get more attention by accumulating grid points at this area. This is especially desirable at locations where the function has a large derivative or is even discontinuous.
- 2. Less important areas will get less attention by dissipation of grid points, while at the same time we reach areas more further away, possibly improving approximation errors in the boundary conditions.

In figure 4.4 we provided a visualization of what grid stretching will actually do for us: accumulate near a desired point (in the figure around 0.5) and dissipate when we move away from this point.

## 4.2.1 Grid stretching on our model

Such a transformation might come in handy in our model around small values for the stock, because at these values the default intensity experiences larger changes. Recall that



FIGURE 4.5: The premium and protection legs at time zero for several stock values. To this end we took the recovery to be zero (note that for a recovery value between zero and one it would depend upon the initial stock value, such that to circumvent this we took it to be zero). Furthermore we randomly took b = 0.9.

default intensity 
$$\equiv \lambda = \frac{1}{S^b},$$
 (4.24)

such that the derivative of the intensity increases near zero. Therefore we expect the premium and protection legs also to be subject to larger changes near zero: the premium leg will converge towards zero whereas the protection leg to 100% of notional. Indeed this behavior can be noticed in figure 4.5. It therefore makes sense to use grid stretching around small values of the stock.

#### 4.2.2 Methodology

To get this type of behavior without changing the uniform grid as we have in the regular setup, we transform the grid using the function

$$Y = \psi(S) = \frac{\operatorname{arcsinh}(\theta(S - X))) - c_1}{c_2 - c_1}.$$
(4.25)

This does not directly make sense, so we clarify:  $c_1$  and  $c_2$  are chosen such that  $Y \in [0, 1]$ ,  $\theta$  is called the stretch rate such that larger  $\theta$  accumulate and dissipate more strongly. Finally, X is the desired accumulation point and S the regular grid dimension, in our case that would be the stock value.

In this thesis, when we apply grid stretching, we shall set<sup>3</sup>  $X = S_0$  such that more attention is paid around the initial stock value. Also, we take the stretch rate  $\theta = 10$ , a

<sup>&</sup>lt;sup>3</sup>Indeed one may notice that this contradicts the usage of grid stretching around small values of stock if  $S_0$  is large, we will comment on this later in this section.



FIGURE 4.6: An example of a grid stretching function, for  $S_0 = 20$  and  $\theta = 10$ .

mere result of some trial and error. As an example we take  $S_0 = 20$  and  $S_{\text{max}} = 100$ , then the resulting function is seen in figure 4.6. In practice we shall iterate uniformly over the *y*-axis, where each *y* value is related to some stock value. And indeed for *y* between 0.2 and 0.7 the stock value is near 20, confirming the accumulation.

If we thus choose a uniform grid on Y then this is related to some accumulated grid on the stock. So, what we would like to do, with respect to our PDEs, is to transform the partial derivatives with respect to the stock such that they become with respect to Y. This can be achieved as follows: let  $\psi^{-1}(Y) = \eta(Y) = S$ , then

$$\frac{\partial A}{\partial S} = \frac{\partial A}{\partial Y} \frac{\partial Y}{\partial S} = \frac{\partial A}{\partial Y} \left(\frac{\partial Y}{\partial S}\right)^{-1} = \frac{1}{\eta'(Y)} \frac{\partial A}{\partial Y},$$

$$\frac{\partial^2 A}{\partial S^2} = \left(\frac{\partial S}{\partial Y}\right)^{-1} \frac{\partial}{\partial Y} \left(\left(\frac{\partial S}{\partial Y}\right)^{-1} \frac{\partial A}{\partial Y}\right) = \frac{1}{(\eta'(Y))^2} \frac{\partial^2 A}{\partial Y^2} - \frac{\eta''(Y)}{(\eta'(Y))^3} \frac{\partial A}{\partial Y},$$
(4.26)

This thus leaves us with partial derivatives as desired with some complicated looking expressions in front of them. These expressions however are actually fairly easy. It's easy to show that

$$\eta(Y) = \frac{1}{\theta} \sinh(c_2 Y + c_1(1 - Y)) + S_0,$$
  

$$\eta'(Y) = \frac{c_2 - c_1}{\theta} \cosh(c_2 Y + c_1(1 - Y)),$$
  

$$\eta''(Y) = \frac{(c_2 - c_1)^2}{\theta} \sinh(c_2 Y + c_1(1 - Y)).$$
  
(4.27)

#### 4.2.3 Implementing in the PDE

Again we show the implementation of grid stretching only for the premium leg, as the protection leg follows analogously. Recall the PDE from (3.37), substituting into this the partial derivatives of (4.26) gives us

$$\frac{\partial A}{\partial t} + \left[ \left( r_t + \lambda_{\eta(Y)} \right) \frac{\eta(Y)}{\eta'(Y)} - \frac{1}{2} \sigma^2 \frac{(\eta(Y))^2 \eta''(Y)}{(\eta'(Y))^3} \right] \frac{\partial A}{\partial Y} 
+ \frac{1}{2} \sigma^2 \left( \frac{\eta(Y)}{\eta'(Y)} \right)^2 \frac{\partial^2 A}{\partial Y^2} - \left( r_t + \lambda_{\eta(Y)} \right) A + C = 0,$$
(4.28)

where we write  $\lambda_{\eta(Y)}$  to emphasize that the default intensity, which depends upon the stock, now thus depends upon  $\eta(Y) = S$ . Now we discretize using backward Euler on the temporal dimension and central differences on the Y dimension to find

$$\frac{1}{\Delta t} \left( A_{n+1}^m - A_n^m \right) + \left[ (r_n + \lambda_n^m) - \frac{1}{2} \sigma^2 \frac{\eta^m \overline{\eta^m}}{(\overline{\eta^m})^2} \right] \frac{\eta^m}{\overline{\eta^m}} \frac{1}{2\Delta x} \left( A_n^{m+1} - A_n^{m-1} \right) \\
+ \frac{1}{2} \sigma^2 \left( \frac{\eta^m}{\overline{\eta^m}} \right)^2 \frac{1}{\Delta x^2} \left( A_n^{m+1} - 2A_n^m + A_n^{m+1} \right) - (r_n + \lambda_n^m) A_n^m + C = 0.$$
(4.29)

Here we proceed, just as before, using the notation of  $\Delta x$  to denote the step size in the dimension of Y, the accumulated stock and the overlining indicates the derivative. After ordering the terms of A we find

$$\begin{split} A_{n+1}^{m} &= \left(1 + \sigma^{2} \left(\frac{\eta^{m}}{\overline{\eta^{m}}}\right)^{2} \frac{\Delta t}{\Delta x^{2}} + (r_{n} + \lambda_{m}^{n}) \Delta t\right) A_{n}^{m} \\ &+ \left(\frac{\eta^{m}}{\overline{\eta^{m}}} \frac{\Delta t}{2\Delta x} \left[\frac{1}{2} \sigma^{2} \frac{\eta^{m} \overline{\eta^{m}}}{(\overline{\eta^{m}})^{2}} - (r_{n} + \lambda_{n}^{m})\right] - \frac{1}{2} \sigma^{2} \left(\frac{\eta^{m}}{\overline{\eta^{m}}}\right)^{2} \frac{\Delta t}{\Delta x^{2}} \right) A_{n}^{m+1} \\ &- \left(\frac{\eta^{m}}{\overline{\eta^{m}}} \frac{\Delta t}{2\Delta x} \left[\frac{1}{2} \sigma^{2} \frac{\eta^{m} \overline{\overline{\eta^{m}}}}{(\overline{\eta^{m}})^{2}} - (r_{n} + \lambda_{n}^{m})\right] + \frac{1}{2} \sigma^{2} \left(\frac{\eta^{m}}{\overline{\eta^{m}}}\right)^{2} \frac{\Delta t}{\Delta x^{2}} \right) A_{i}^{m-1} - C\Delta t \end{split}$$

To easy the notation we write

$$\delta^{m} = \frac{\eta^{m}}{\overline{\eta^{m}}}, \qquad \alpha^{m} = \sigma^{2} (\delta^{m})^{2} \frac{\Delta t}{\Delta x^{2}},$$
  
$$\beta_{n}^{m} = (r_{n} + \lambda_{n}^{m}) \Delta t, \quad \gamma_{n}^{m} = \frac{\delta^{m}}{2\Delta x} \left[ \frac{1}{2} \sigma^{2} \delta^{m} \frac{\overline{\eta^{m}}}{\overline{\eta^{m}}} \right], \qquad (4.30)$$

such that we get the simpler looking implicit finite difference scheme

$$A_{n+1}^{m} = \left(1 + \alpha^{m} + \beta_{n}^{m}\right)A_{n}^{m} + \left(\gamma_{n}^{m} - \frac{1}{2}\alpha^{m}\right)A_{n}^{m+1} - \left(\gamma_{n}^{m} + \frac{1}{2}\alpha^{m}\right)A_{n}^{m-1} - C\Delta t, \quad (4.31)$$

where we again emphasize that

$$\lambda^m = \frac{1}{(S^m)^b} = \frac{1}{(\eta^m)^b}.$$
(4.32)

#### 4.2.4 Results

This adjustment might be helpful for our discretization as we have seen that both premium and protection legs have larger derivatives for smaller stock values. This however makes the decision of whether to use grid stretching or not, not straightforward: when the initial stock value is small, then surely grid stretching will be beneficial, but when the initial stock value is large then grid stretching would provide less attention to the smaller stock values. This particular behavior can indeed be confirmed by the example in figure 4.7. We truncate the stock values on the grid between 0 and 2000. Next we determine premium leg, protection leg and spread for the step sizes  $dx = 2^k, k = 2, 1, ..., -3$  and observe these results while changing the initial stock value  $S_0$ . In the figure, each row represents error difference plots for different initial stock values.

As we can see the grid stretching method performs best for the lower two rows and worse for the top two rows. This is exactly in line with our expectations. Indeed when  $S_0 = 100$ the grid stretching adjustments provide extra grid points around S = 100, and less around the low stock values. But exactly at these low stock values both premium and protection leg become more sensible. On the contrary, when  $S_0 = 3$  we are modeling extra grid points around the low stock values, such that grid stretching performs better.

One might suggest to always choose the accumulation value of the grid stretcher to be around, say  $S_0 = 10$ , such that we always take the extreme part of the grid into account. However, we find results to become even worse for large stock values, which is likely to be a consequence of dissipation of the grid points around  $S_0$ . A more proper choice could be made if we'd more carefully adjust the stretch rate such that we gain enough points around  $S_0$  and still accumulate properly around the low values of the stock. This we leave for possible future investigation.

With these results in mind, we choose to set the rule to apply grid stretching when stock values are low, we take  $S_0 \leq 15$ , whereas we use the regular grid when  $S_0 > 15$ .

The focus of this study is on determining CDS spreads and to extract from them the desired term structures. The Greeks  $\Delta$  and  $\Gamma$ , important in the scope of hedging, are therefore less of in issue in this thesis. The interested reader is referred to [OLH05] for a



FIGURE 4.7: Comparison of the performance of grid stretching versus a regular grid. Each row models a different initial stock value where the rest of the parameters are randomly taken to be:  $\sigma = 0.5, r = 0.03$  (constant),  $a_0 = 0.2, a_1 = -1.5, b = 0.9, \Delta t = \frac{1}{124}$ .

more comprehensive study on grid stretching on European and American vanilla options, including convergence of  $\Delta$  and  $\Gamma$ .

Finally, speed is not the main focus of this thesis but it is an important subject nonetheless. To this end we only mention briefly that, yes, some additional calculations have to be provided for grid stretching, but on a similar note, less grid points are required to find the same accuracy as when a regular grid is used. Therefore due to its greater accuracy at a smaller grid the method can be faster when applied correctly.

# 4.3 Convergence

Convergence is usually an important requirement for finite differences schemes in general, without convergence of the scheme we will never know how to interpret the results provided. Intuitively, convergence tells us that for smaller  $\Delta t$  and  $\Delta x$  our solution gets closer to the true solution. This simple statement is not at all clear when there are multiple solutions to the PDE. However, we shall soon see that our PDEs have unique solutions and we may therefore stick to this intuitive interpretation.

An important result is the Lax-Richtmyer Equivalence theorem (see theorem F.0.1). We understand from it that our finite difference scheme converges if

- 1. the corresponding initial-boundary value problem is linear and well-posed,
- 2. the finite difference scheme is consistent,
- 3. and numerically stable.

Recall that we mentioned the D&H model to have numerical stability issues. In this section we shall prove this indeed to be the case.

We will start this section by verifying the well-posedness of our PDEs. Then we will determine the stability region of the explicit method. This then tells us whether the particular choice that results in the D&H model is a stable choice. We will omit confirming consistency of the difference schemes as these are generally known results. In a nutshell, consistency corresponds to the property that a finite difference scheme equals the pde when both  $\Delta x$  and  $\Delta t$  approach zero.

# 4.3.1 Well-posedness

In order to apply the Lax-Richtmyer Equivalence theorem we require well-posedness (among consistency and stability) of the PDEs.

**Definition 4.3.1. Hadamard's well-posedness** An initial-boundary value problem for PDEs is well-posed if

- 1. it has a unique solution (existence and uniqueness),
- 2. the solution depends continuously on the given data (dynamic stability).

Following [Hof92] we can easily see that our PDEs are of the parabolic type and that for a parabolic PDE to be well-posed

- 1. the solution domain D(t, S) must be open in the time (or timelike) direction,
- 2. initial data must be specified along the time (or timelike) boundary,
- 3. continuous boundary conditions must be specified along the physical boundaries of the solution domain,
- 4. the boundary conditions can be of the Dirichlet type, the Neuman type, or the mixed type.

All these conditions are easily met for both the premium and protection leg PDEs: indeed the solution domain in the time direction is the open domian (0,T). Initial data is presented at time T, where T functions as the initial time (even though all other times in the domain are before T). Furthermore the boundary conditions at S = 0 are of the Neumann type as they are set to zero for all t.

# 4.3.2 Stability

By stability we refer to numerical stability. When a finite difference scheme is numerically stable then small errors in the solution do not grow when evaluating the next step.

It is a general result that the fully implicit scheme is numerically stable. Therefore using the Lax-Richtmyer Equivalence theorem we can conclude that this scheme converges. For the explicit scheme the story is somewhat different. It is well known that the explicit scheme is conditionally stable, that is, the set of  $\Delta x$  and  $\Delta t$  (the stability region) that produces numerical stable results is restricted.

Recall the D&H model to be an explicit discretization of our PDE's. We are now in a position to obtain stability results for the explicit finite difference scheme, and thus also for the D&H model.

**Theorem 4.3.1.** Given violation of condition (2.14), the forward-time center-spacing finite difference scheme to the PDEs (3.37) and (4.6) is unconditionally unstable.

**Proof.** We follow the approach described in appendix F. It showed us that we may transform a one-step difference scheme to its Fourier counterpart in order to investigate the stability by the Von Neumann approach. To this end we recall that we may look at the homogeneous counterpart of the PDEs, which are for both PDEs the same. As we are interested in the forward-time center-spacing finite difference scheme we may proceed with (3.47), leaving C = 0, and substituting

$$A_n^m \to e^{im\Delta x\eta} A_n(\eta),$$
 (4.33)

gives us

$$e^{im\Delta x\eta}A_{n}(\eta) = \left(1 - \frac{\sigma^{2}\Delta t}{\Delta x^{2}}\right)e^{im\Delta x\eta}A_{n+1}(\eta) + \left(\frac{r_{n} + \lambda^{m} - \frac{1}{2}\sigma^{2}}{2\Delta x} + \frac{\sigma^{2}}{2\Delta x^{2}}\right)\Delta t e^{i(m+1)\Delta x\eta}A_{n+1}(\eta)$$

$$+ \left(\frac{\sigma^{2}}{2\Delta x^{2}} - \frac{r_{n} + \lambda^{m} - \frac{1}{2}\sigma^{2}}{2\Delta x}\right)\Delta t e^{i(m-1)\Delta x\eta}A_{n+1}(\eta).$$

$$(4.34)$$

Let  $\nu = \frac{\sigma^2 \Delta t}{\Delta x^2}$ ,  $\mu = \frac{r_n + \lambda^m - \frac{1}{2}\sigma^2}{2\Delta x} \Delta t$  and  $\theta = \Delta x \eta$  such that

$$A_n(\eta) = \left(1 - \nu + \left(\mu + \frac{1}{2}\nu\right)e^{i\theta} + \left(\frac{1}{2}\nu - \mu\right)e^{-i\theta}\right)A_n(\eta).$$

$$(4.35)$$

Thus we find the amplification factor

$$g(\theta, \Delta t, \Delta x) = 1 - \nu + \left(\mu + \frac{1}{2}\nu\right)e^{i\theta} + \left(\frac{1}{2}\nu - \mu\right)e^{-i\theta}.$$
(4.36)

In order to find stability and its corresponding stability region  $\Lambda$  we have to verify theorem F.0.2. Next we will show under which conditions stability holds. To this end we start by determining  $|g(\theta, \Delta x, \Delta t)|^2$ :

$$\begin{aligned} |g(\theta, \Delta x, \Delta t)|^2 &= |1 - \nu + \nu \cos(\theta) + 2i\mu \sin(\theta)|^2 \\ &= (1 - \nu + \nu \cos(\theta))^2 + 4\mu^2 \sin^2(\theta) \\ &= (1 - 2\nu \sin^2(\theta/2)) + 4\mu^2 \sin^2(\theta) \\ &= 1 - 4\nu \sin^2(\theta/2) + 4\nu \sin^4(\theta/2) + 4\mu^2 \sin^2(\theta) \\ &= 1 - 4\nu \sin^2(\theta/2) + 4\nu \sin^4(\theta/2) + 4 \left(4\sin^2(\theta/2) + 4\sin^4(\theta/2)\right) \\ &= 1 + \left(16\mu^2 - 4\nu\right) \sin^2(\theta/2) + \left(16\mu^2 + 4\nu^2\right) \sin^4(\theta/2) \,. \end{aligned}$$

As we have  $\sin^4(\theta) \leq \sin^2(\theta)$  where equality when  $\theta = 0, 1$ , thus a necessary condition. Proceeding gives

$$|g(\theta, \Delta x, \Delta t)|^{2} \leq 1 + (32\mu^{2} + 4\nu^{2} - 4\nu) \sin^{2}(\theta/2)$$
  
$$\leq 1 + (32\mu^{2} + 4\nu^{2} - 4\nu)$$
(4.37)

Now suppose  $|g(\theta, \Delta x, \Delta t)|^2 \le 1 + y + C\Delta t$  for constants  $y, C \in \mathbb{R}$  then

$$\begin{split} |g(\theta, \Delta x, \Delta t)| &\leq \sqrt{1 + y + C\Delta t} \\ &= 1 + \frac{1}{2}(y + C\Delta t) - \frac{1}{8}(y^2 + C\Delta t)^2 + \dots \quad \text{by Taylor expansion} \\ &= 1 + \frac{1}{2}y - \frac{1}{8}y + \dots + \mathcal{O}\left(\Delta t\right) \\ &= \sqrt{1 + y} + \mathcal{O}\left(\Delta t\right). \end{split}$$

What we can understand from this is that in order to satisfy theorem F.0.2 we require  $y \leq 0$ . We should thus write (4.37) in the form  $1 + y + C\Delta t$ . Substituting back the values for  $\mu$  and  $\nu$  we get

$$32\mu^{2} + 4\nu^{2} - 4\nu = 8\left(r_{n} + \lambda^{m} - \frac{1}{2}\sigma^{2}\right)^{2}\frac{\Delta t^{2}}{\Delta x^{2}} + 4\sigma^{4}\left(\frac{\Delta t}{\Delta x^{2}}\right)^{2} - 4\sigma^{2}\frac{\Delta t}{\Delta x^{2}}.$$
 (4.38)

We require (4.38) to be smaller or equal to  $y + C\Delta t$ . To this end we have to consider the following two possibilities (other possibilities are trivial):

1. 
$$\frac{\Delta t}{\Delta x^2} \equiv \text{Constant}$$
  
2.  $\frac{\Delta t^2}{\Delta x^2} \equiv \text{Constant}$ 

If we'd consider the second option we would get

$$\frac{\Delta t}{\Delta x^2} = \frac{1}{\Delta t} \frac{\Delta t^2}{\Delta x^2},\tag{4.39}$$

which would grow unboundedly when  $\Delta t \to 0$  such that the requirement would never hold. Therefore, assume the first to holds. Recall that n and m exist such that

$$r_n + \lambda^m \ge \frac{\sigma}{\sqrt{\Delta t}}.$$

At this grid point the frozen coefficient problem can be considered [Str89]. It thus follows that

$$8\left(r_n + \lambda^m - \frac{1}{2}\sigma^2\right)^2 \frac{\Delta t^2}{\Delta x^2} \ge 8\left(\frac{\sigma}{\sqrt{\Delta t}} - \frac{1}{2}\sigma^2\right)^2 \frac{\Delta t^2}{\Delta x^2}$$
$$= 8\sigma^2 \frac{\Delta t}{\Delta x^2} + \mathcal{O}\left(\Delta t^{3/2}\right)$$
We find that a necessary requirement for  $y \leq 0$  becomes

$$8\sigma^{2}\frac{\Delta t}{\Delta x^{2}} + 4\sigma^{4}\left(\frac{\Delta t}{\Delta x^{2}}\right)^{2} - 4\sigma^{2}\frac{\Delta t}{\Delta x^{2}} \le 0$$
  
$$\leftrightarrow \frac{\Delta t}{\Delta x^{2}}\left(4\sigma^{4}\left(\frac{\Delta t}{\Delta x^{2}}\right) + 4\sigma^{2}\right) \le 0.$$

As  $\Delta t, \Delta x > 0$  we find

$$4\sigma^4 \left(\frac{\Delta t}{\Delta x^2}\right) + 4\sigma^2 \le 0$$
$$\frac{\sigma^2 \Delta t}{\Delta x^2} \le -1.$$

This condition can never be satisfied given the Left Hand Side (lhs) to be strictly positive.  $\hfill\square$ 

**Remark 4.3.2.** Note that we if we'd remove the condition of violation we would get the usual stability condition seen for the explicit finite differences scheme, i.e.

$$\frac{\sigma^2 \Delta t}{\Delta x^2} \le 1.$$

**Remark 4.3.3.** As is also noted by [Str89], if the frozen coefficient problem is violated in a small region, the instability phenomena that arise will originate in that area and not grow outside that area.

To us this would mean that instability only arises there where the stock is very small. Ideally, we'd like to quantify the amplification of errors within this area and its consequences for the rest of the procedure. However, it is clear that there are several difficulties which can simply be worked around by applying a different discretization scheme, e.g implicit schemes.

**Theorem 4.3.4.** The model by Das & Hanouna is conditionally stable in a stability region

$$\Lambda = \left\{ \Delta t \in \mathbb{R}_{>0} \mid (r+\lambda)\sqrt{\Delta t} \le \sigma, \text{ for all } (r,\lambda) \text{ in the tree} \right\}.$$

**Proof.** Recall that we have shown that the D&H model is a special case of the PDE model, in particular, the PDE model was found to have the discretisation (3.50). We required a forward-time center-space discretisation with

$$\frac{\sigma^2 \Delta t}{\Delta x^2} = 1. \tag{4.40}$$

In the light of the previous theorem this would not satisfy the stability condition when (2.14) is violated. However, the D&H model is not exactly equal to this discretisation, altough it is in the limit. We can see that the only difference between the D&H model and the forward-time center-space discretization of the PDEs lies in slightly different up and down probabilities. This difference was shown to be  $\mathcal{O}(\Delta t)$ . Consequently the difference in the amplification factor of the FD discretisation and the D&H model is of order  $\mathcal{O}(\Delta t)$ . Now by corollary F.0.1 the result is immediate.

**Remark 4.3.5.** Note how the result of theorem 4.3.4 is inconvenient when we want to let  $\Delta t$  approximate 0. This can immediately be seen as the largest  $\lambda$  in the tree is given by

$$\lambda = \frac{1}{\left(S_0 e^{-N\sigma\sqrt{\Delta t}}\right)^b} = \frac{1}{S_0^b e^{-b\sigma\frac{1}{\sqrt{\Delta t}}}} \to \infty \quad \text{when} \quad \Delta t \to 0.$$

Here we recall that  $N = \frac{1}{\Delta t}$ .

## 4.4 Parameter choice

Before we can run the finite difference iterative scheme we have to set some numerical parameters, in particular the ones that determine the grid that we are working on, these are:

$$\Delta t, \quad \Delta x, \quad S_{\max}.$$

Depending on their sizes results might differ significantly. Therefore, with the eye on practice, we wish to establish some rules, that is, depending on certain conditions set these values such that results are acceptable.

First, by means of a Monte Carlo argument we determine  $S_{\text{max}}$ . Thereafter values for  $\Delta t$  and  $\Delta x$  are determined.

#### 4.4.1 A maximum stock value

Within the finite difference approach the stock value has to be capped, requiring the approximation described in section 4.1.1. Making such approximation introduces an error to the resulting parameter set found within the calibration. Understanding this error is not straightforward. Suppose for example two data sets to be known to have respectively  $b_1 = 1.5$  and  $b_2 = 0.6$ , then in the first data set capping the stock value at 100 leaves an error of 0.001 at the upper bound, whereas for a similar error in the second case we'd require capping the stock at  $10^5$ :

$$0.001 = \frac{1}{100^{1.5}} = \frac{1}{100000^{0.6}}$$

Unfortunately we do not know the correct value of b before calibrating such that choosing the maximum value of the stock is not a straightforward exercise. And in particular, if it is 0.6, do we really want to cap the stock at  $10^5$ ? That seems extremely large, slowing the calibration process.

We can make the following reasoning which will help us significantly: Suppose b = 0.6 with current stock value  $S_0 = 40$  and volatility  $\sigma = 0.5$ . Then suppose that hypothetically we'd use a Monte Carlo approach to calibrate the parameter set to fit the CDS spreads. If we'd introduce the extra condition that the probability of default becomes zero when the stock reaches S = 1000, then this extra condition will probably be of very little influence, since it is highly unlikely for the stock to reach such extreme value. Indeed we performed a simple such Monte Carlo run with  $10^5$  paths, from these paths only 15 reached the value

$S_{\max}$	Parameter set $(a_0, a_1, b)$	$\chi$	RMSE
14.376229938415303 10.881154997725107 50	(0.201, -0.157, 0.989)	$2.332\cdot 10^{-4}$	9.57
100	(0.798, -0.771, 0.717)	$2.503\cdot10^{-4}$	9.80
200	$(0.857, \ -0.893, \ 0.699)$	$2.725\cdot10^{-4}$	10.17
400	(0.871, -0.925, 0.694)	$2.782\cdot10^{-4}$	10.26
800	(0.879, -0.941, 0.690)	$2.790\cdot10^{-4}$	10.27
1600	(0.881, -0.947, 0.689)	$2.792 \cdot 10^{-4}$	10.27

TABLE 4.1: Multiple calibrations for the Amazon data  $(S_0 = 7.756, \sigma = 0.972)$  while changing the maximum value of S within the finite difference grid. Further chosen parameters:  $M = 48, N = 2 \cdot S_{\text{max}}$ , interest rates such as in D&H including linear interpolation as described in section 5.2.

of 1000.

Furthermore, a default occurring for a very large value of S does not add any more value than a default for regular values of S, such that an importance sampling argument would not change this reasoning.

Thus, as this is the case for the Monte Carlo simulation, so should it be for the finite difference approach, simply because results should coincide. We may therefore reason that we do not have to cap the stock at ridiculously large values (such as  $10^5$ ) stated above. We can confirm the validity of this reasoning by performing some calibrations to the Amazon data presented by Das & Hanouna. The parameter set that seems to result takes the value  $b \approx 0.69$ . Results of multiple calibrations are shown in table 4.1. Indeed we see that the parameter sets,  $\chi^4$  and the least squares difference converge towards a solution. Furthermore, where initially differences between the parameter sets are relatively large the Euclidean difference (2-norm) between the last two calibrations is less then 0.007.

There are two things we should take into consideration when choosing the maximum value of the stock:

- 1. We expect the stock to have a large probability of default: we shall set the maximum stock value to 2000. A large probability of default might be expected for example when observing very large CDS spreads.
- 2. We expect the stock not to have extreme probability of default: we shall set the maximum value of the stock at 500.

These considerations are open for discussion if for example the current stock value is already very large. Take for example the stock value of Berkshire Hathaway which reached a (closing) stock value of \$228,615 on December 31st 2014. In this case neither 2000 nor 500 would suffice.

<sup>&</sup>lt;sup>4</sup>The objective function is explained and provided in section 5.1.1.

#### 4.4.2 Temporal and stock step sizes

In this section we shall decide on the choice of step sizes  $\Delta t$  and  $\Delta x$ . Up till now we have chosen these values to give decent estimation such that the behavior of the model would become clear. However, when we calibrate the model to market data we want the results not to deviate too much from the infinitesimal step sizes counterpart. To this end we require the results not to deviate more than 1 Bps with the *exact* solution.

Not knowing the exact solution, we cant give specifications of the error therefore we determine a dummy error for 3 different scenarios. By dummy error we refer to the error made with respect to step sizes

$$\Delta t = 2^{-10}, \quad \Delta x = 2^{-8}. \tag{4.41}$$

The spacial step size  $\Delta x$  is not as small compared to  $\Delta t$ , we shall see however that  $\Delta x$ , does not influence the results very much. Furthermore, we are bound by the capabilities of the computer of creating large matrices (choosing  $\Delta x$  smaller would make the computer run out of memory).

In table 4.2 we present the dummy errors of parameter sets found for Banco do Brasil S.A., Shell and Amazon. These three sets are just randomly selected to give us an impression of the results. There are three things that we'd like to note in particular:

- 1. Changing  $\Delta x$  doesn't seem to influence the resulting errors significantly, i.e. from left two right the differences are minimal compared to the differences when changing  $\Delta t$ .
- 2. As expected, when halving  $\Delta t$  the error also seems to halve more or less. This corresponds to the error being of order  $\Delta t$  in time.
- 3. Even for relatively large values of  $\Delta t$  and  $\Delta x$  the error is fairly small. This means that the model doesn't require much information to have a sense of where the solution resides.

From the table we see that larger spreads seem to be related to larger errors, i.e the extremer the scenario, the larger the error. We could therefore argue that for small spreads the step sizes can be chosen larger such that the calibration would perform faster. However, we prefer to remain consistent and choose one pair of  $\Delta t$  and  $\Delta x$  that we can applied to all scenarios. To this it seems a safe choice to take

$$\Delta t = \frac{1}{48} \quad \text{and} \quad \Delta x = \frac{1}{2}.$$
(4.42)

Shell	Exact 5Y spread 56.4882				
		1	$\Delta x$		
$\Delta t$	$2^{0}$	$2^{-1}$	$2^{-2}$	$2^{-3}$	$2^{-4}$
$12^{-1}$	0.1517	0.1704	0.1751	0.1762	0.1765
$24^{-1}$	0.0632	0.0817	0.0864	0.0875	0.0878
$48^{-1}$	0.0186	0.03718	0.0418	0.0430	0.0433
$96^{-1}$	0.0037	0.0148	0.0195	0.0206	0.0209
$192^{-1}$	0.0148	0.0036	0.0083	0.0094	0.0097
Banco	do Brasil S.A.		Exact 5	Y spread	298.8647
			$\Delta x$		
$\Delta t$	$2^{0}$	$2^{-1}$	$2^{-2}$	$2^{-3}$	$2^{-4}$
$12^{-1}$	0.3291	0.3770	0.3857	0.3880	0.3886
$24^{-1}$	0.1360	0.1823	0.1909	0.1931	0.1837
$48^{-1}$	0.0388	0.0842	0.0928	0.0950	0.0955
$96^{-1}$	0.010	0.0351	0.0436	0.0457	0.0463
$192^{-1}$	0.0345	0.010	0.0190	0.021	0.0216
Amazo	on		Exact 5	Y spread	1286.5833
		4	$\Delta x$		
$\Delta t$	$2^{0}$	$2^{-1}$	$2^{-2}$	$2^{-3}$	$2^{-4}$
$12^{-1}$	1.75301	1.7268	1.7188	1.7168	1.7163
$24^{-1}$	0.8876	0.8612	0.8533	0.8513	0.8508
$48^{-1}$	0.4552	0.4289	0.4209	0.41890	0.4184
$96^{-1}$	0.2392	0.2129	0.2049	0.2029	0.2024
$192^{-1}$	0.1313	0.1049	0.0969	0.0949	0.0944

TABLE 4.2: Errors on the 5 year spread on data of Banco do Brasil S.A., Shell and Amazon. Errors are with respect to the 'exact' solution. The values depicted are in Basispoints.

The choice of  $\Delta t = \frac{1}{48}$  also has the advantage that when using forward interest rates we can assume a month to have 4 weeks, such that for monthly provided forward interest rates we can use one such rate for exactly 4 steps.

**Remark 4.4.1.** Note how in the spacial direction, the errors do not seem to reduce by a fourth, as you might expect do to the second order convergence. This might however

exactly be due to the reason that the errors presented in the time dimension are much larger than those in space, not letting the spacial convergence come to its right.

# Calibrations and results

"Is it possible, in the final analysis, for one human being to achieve perfect understanding of another?

We can invest enormous time and energy in serious efforts to know another person, but in the end, how close can we come to that person's essence? We convince ourselves that we know the other person well, but do we really know anything important about anyone?"

– Haruki Murakami , The Wind-Up Bird Chronicle

We shall fit the PDE model to market data corresponding to individual firms. To do so we require the following information of the particular firm

- Stock value at date of evaluation,
- Historical or implied volatility,
- CDS spreads and,
- Interest rates between date of evaluation and longest maturity.

The first three we extract directly from Bloomberg. For the interest rates we use forward implied interest rates. For the exact procedure we used to extract these from market data we refer to appendix H.2.

This section is constructed as follows. Recall that previously we have constructed PDEs and described how to evaluate these by means of finite differences. First we shall therefore look into how to calibrate the model to its parameters. Next we provide a case comparison to the D&H model on the Amazon data. In section 5.3 we provide calibrations on senior spreads whereas in section 5.3.1 we calibrate the model simultaneously to senior and

subordinated spreads. Then in section 5.4 we provide CVA calculations and compare our model's results to those of the PC model described in section 1.2.3.4. Finally we provide a case comparison to the PC model, where we clearly see the differences for this constant recovery model.

# 5.1 Calibration process

Given market spreads of some firm, we wish to determine the parameter set  $\{a_0, a_1, b\}$  such that our model implies these same spreads. These parameters then tell us how our model should look like for this particular firm and subsequently provide us with the desired term structures.

In order to find such parameters we have to perform a process called calibration. Calibration is nothing but finding the parameters that make a financial pricing model fit best to given market data and it usually consists of an optimization process. This can often be complicated due to the complexity of the objective functions. The same holds for our PDE model, which has no straightforward solution, which is why we require the finite differences numerical solution.

Before looking at the optimization process itself we should decide on the objective function and describe some constraints.

#### 5.1.1 The objective function

Let t be the current time and  $\overline{C}^T$  be a market spread with time to maturity T related to some entity. By Prot (t, T; b) and Prem  $(t, T; a_0, a_1, b, C)$  we refer to the protection and premium legs respectively, where C is the model spread.

We would now like to find the correct parameter set such that the market spreads and the model spreads coincide, this would suggest the minimization of an objective function in the following form

$$\min_{a_0,a_1,b} \left\{ \sum_{i=1}^{N} \left( \frac{\operatorname{Prot}\left(t, T_i; b\right)}{\operatorname{Prem}\left(t, T_i; a_0, a_1, b, 1\right)} - \overline{C}^{T_i} \right)^2 \right\},\tag{5.1}$$

for N observed spreads. Recall that model spreads can be written as

$$C^{T} = \frac{\operatorname{Prot}(t, T_{i}; b)}{\operatorname{Prem}(t, T; a_{0}, a_{1}, b, 1)},$$
(5.2)

which comes from the desired relationship

$$\operatorname{Prem}\left(t, T; a_0, a_1, b, C^T\right) = \operatorname{Prot}\left(t, T\right), \qquad (5.3)$$

that premium and protection legs should coincide (recalling from formula (1.13),  $C^T$  can be taken outside). Thus instead of optimizing equation (5.1) we may look at the more complete objective function

$$\sum_{i=1}^{N} \left( \operatorname{Prem}\left(t, T_i; a_0, a_1, b, \overline{C}^{T_i}\right) - \operatorname{Prot}\left(t, T_i\right) \right)^2.$$
(5.4)

It is more complete in the sense that (after calibration) our model is supposed to produce similar spreads and simultaneously satisfy (5.3). However, the new objective function, (5.4), has the slight drawback that the optimizer seems to converge when both, premium and protection legs are getting smaller. This is somewhat misleading, because we want them to be become equal, not necessarily small. A safer objective function would thus look like

$$\sum_{i=1}^{N} \left( \frac{\operatorname{Prem}\left(t, T_i; a_0, a_1, b, \overline{C}^{T_i}\right)}{\operatorname{Prot}\left(t, T_i\right)} - 1 \right)^2.$$
(5.5)

This then brings us to last issue in which we might be fooling the optimizer, e.g. the optimizer might get stuck into bringing  $a_0$  towards minus infinity. This would leave the premium leg zero such that the objective function would be getting smaller and smaller until it reaches a minimum of N. This is easily solved by noting that both premium and protection are always positive such that (5.3) holds if and only if

$$\left(\operatorname{Prem}\left(t,T;a_{0},a_{1},b,C^{T}\right)\right)^{2} = \left(\operatorname{Prot}\left(t,T\right)\right)^{2}$$
$$\leftrightarrow \frac{\operatorname{Prem}\left(t,T_{i};a_{0},a_{1},b,\overline{C}^{T_{i}}\right)}{\operatorname{Prot}\left(t,T_{i}\right)} = \frac{\operatorname{Prot}\left(t,T_{i}\right)}{\operatorname{Prem}\left(t,T_{i};a_{0},a_{1},b,\overline{C}^{T_{i}}\right)}.$$

Such that a new and less ambiguous objective function can be created.

**Definition 5.1.1** (PDE CDS pricing objective function). The calibration objective function to price a CDS under the dynamics of the PDEs from theorem 3.3.1 and 3.3.2 is defined by

$$\chi(a_0, a_1, b) = \sum_{i=1}^{N} \left( \frac{\operatorname{Prem}\left(t, T_i; a_0, a_1, b, \overline{C}^{T_i}\right)}{\operatorname{Prot}\left(t, T_i\right)} - \frac{\operatorname{Prot}\left(t, T_i\right)}{\operatorname{Prem}\left(t, T_i; a_0, a_1, b, \overline{C}^{T_i}\right)} \right)^2.$$
(5.6)

This function guarantees none of the above problems to arise, in particular the last problem is now avoided as when the premium leg converges towards zero the second expression will converge to infinity.

#### 5.1.2 Constraints

In this optimization we deal with three parameters, namely  $\{a_0, a_1, b\}$ . At this point we have only decided upon requiring  $b \ge 0$ , as negative values of b correspond to a decreasing intensity when the stock increases, which has no practical support. We are thus left with optimizing (5.6) on the domain

$$\Omega = \{ (a_0, a_1, b) \in \mathbb{R} | a_0 \in (-\infty, \infty), a_1 \in (-\infty, \infty), b > 0 \}.$$
(5.7)

This seems an unnecessary large domain. Recall how parameter  $a_0$  is connected to the time zero recovery, i.e. the recovery at evaluation. Depending upon the convergence of the recovery function towards zero and one, we can create bounded intervals on which most of the recovery values are represented. We choose the initial recovery values to be between [-0.001, 0.999], such that for zero  $a_1$  we require

Parameter  $a_1$  is related to the correlation between recovery and default intensity. We know from both historical and implied data that it is expected to be negative. To this end, and recalling the discussion in section 3.7.1, it seems reasonable to assume  $a_1 \leq 0$ . For the lower boundary we cannot create an easy constraint. Very small/large default intensity may imply very small/large  $a_1$ . For computational purposes we may fix a  $|a_1| < 5 \cdot 10^3$ . The author of [Mur99] shows that for Japanese companies, b lies between 1.2 and 2. From our simulations so far we have indeed not found any b to exceed 2. We found them smaller then 1.2 quite often though. To be on the safe side and compensate for exceptions we take  $b \in [0, 2.5]$ .

We thus find the final optimization domain to be

**Definition 5.1.2** (Optimization domain). The calibration optimization domain for pricing a CDS under the dynamics of the PDEs from theorem 3.3.1 and 3.3.2 is defined by

$$\Omega = \{ (a_0, a_1, b) \in \mathbb{R} | a_0 \in A_0, a_1 \in (-5 \cdot 10^3, 0), b \in [0, 2.5] \}.$$
(5.9)

# 5.2 Comparing PDE and D&H model results

This section is dedicated to reflecting on the D&H model. Let's fit the Amazon data just as was done by Das & Hanouna[DH09], but now also to our PDE model. In table 5.1 we provides the input data for Amazon in September 2001 (specific dates are not provide by [DH09]) and fitting results for both the D&H and the PDE model. The results are presented with three decimals as we have seen in table 3.1 that sensitivity of even the third decimal can be of influence on the results.

Before going to a discussion of the results we'd like to explain our procedure with respect to the usage of the forward interests rates provided by Das & Hanouna within the PDE model. As can be seen in the table only 5 forward interests rates are provided. The PDE model however, requires many more rates, one for every time step. In their paper [DH09], Das & Hanouna explain that the forward interest rates they provide are averages for that particular year. Therefore we provide an interpolation such that each year's average forward interest rate corresponds to that of Das & Hanouna. In particular we took the initial forward interest rate as the average forward interest rate for the first year minus half times the smallest difference between all average forward interest rates. Subsequently one may find the rates at each year and interpolate to find continuous results<sup>1</sup>, i.e. let  $f'_i$ , i = 0, 1, ...4 the forward rate between year i and i + 1, provided by Das & Hanouna, we then set our rates by

$$f_0 = f'_0 - \frac{1}{2} \min \left\{ \operatorname{abs} \left( f'_i - f'_{i-1} \right) | i = 1, \dots 4 \right\},$$
  

$$f_{N(i+1)} = 2 \cdot f'_i - f_{Ni}, \quad \text{for } i = 0, \dots 4,$$
(5.10)

where N as usual indicates the amount of steps in one year for the PDE discretization. Linear interpolation can now be applied to find the forward interest rates within the year.

What immediately seems to stand out is the relative poor fit of the PDE model: its RMSE is much larger than that of the D&H model. However, when looking at the first sub figure of figure 5.1 we see that the fit does seem very reasonable, it furthermore becomes clear that the high RMSE comes from the relatively large spread at the third year. Recalling the sample term structures we have seen in section 3.7.1 it seems unreasonable for the PDE model to provide a fit just as good as that of the D&H model. The term structure of the market spreads seem to lie on an unrealistic curve with respect to the PDE dynamics.

Let's now look at the term structures of both models. The difference in recovery at default is very large, the PDE model provides almost twice as large recoveries. The probability density functions of both functions are also different but not as striking as the recovery.

<sup>&</sup>lt;sup>1</sup>The reason that we used an interpolation and did not provide continuous data from for example Bloomberg is twofold: Das & Hanouna [DH09] do not specify the exact date of evaluation and their curve choice. Second, for a fair comparison we do not want to provide extra information by looking up real rates.

Now that we have seen the fits of both models on comparable data we are in a position to summarize

Main observation 5.2.1 (PDE model versus D&H model: Credibility). Both, PDE and D&H models, are based on the same dynamics (see section 3.4), such that they *should* provide the same results. We have seen however that the D&H model requires large steps in time to overcome both numerical stability and interpretability of the probabilities on the tree, whereas the PDE model is less than 1 Bps from its exact solution. These large steps thus seem to bring with them a substantial amount of errors.

Furthermore, in the light of dynamic stability the resulting differences in recovery are unacceptable recalling that small deviations in spread should result in small deviations in the underlying term structures (see section 3.8).

Thus the better fit of the D&H model seems a misleading appearance.

Main observation 5.2.2 (PDE model versus D&H model: Usefulness). In contrast to the PDE model it is not able to provide continuous term structures. This might especially provide fairly different results when quantitative calculations have to be provided. Furthermore, in the light of the Monte Carlo simulation the PDE model is capable of providing calculations including correlation between recovery and default. Think of CVA which depends continuously on time and on the correlations between both processes.

**Remark 5.2.1.** Finally, we would like to comment on the choice of modeling the expected recovery as done by Das & Hanouna. Recall that we have chosen to model the expected recovery at default within the framework of Das & Hanouna by (2.11). In their paper they instead provide results for

$$\sum_{j \in \{0, J_i\}} p[i, j] \rho[i, j], \tag{5.11}$$

see section 2.3 for more details.

This however is not the expected recovery at default, but, as they call it, the expected recovery. An important question which then arises is: what exactly is expected recovery if it is not at default?

As a result of their definition/usage of the expected recovery, it will always converge towards zero at infinity. This has the particular consequence that the term structures they present are seemingly continuously declining, as default states always adds a recovery value of zero to the summation.

Amazon fit comparison (Das & Hanouna and the PDE model)						
Sto	ock price			7.756		
Ste	Stock volatility 0.972					
Tri	nomial step siz	e		1		
Т	Fwd interest	Market spread (bps)	Trinomial (bps)	PDE (bps)		
1	0.0226	899.06	899.19	898.10		
2	0.0273	1130.93	1130.41	1140.06		
3	0.0329	1258.20	1259.80	1239.40		
4	0.0383	1265.46	1264.72	1274.85		
5	0.0436	1285.32	1283.32	1286.15		
Parameters D&H $(a_0, a_1, b)$ $(0.199, -1.277, 0.801)$ Parameters PDE $(a_0, a_1, b)$ $(0.924, -1.036, 0.681)$						
$\frac{1.22}{1.22} = 1000, 0.001$				10.26		
$\chi$	$\chi$ - 0.000278					

 TABLE 5.1: Input and fitted data for both the trinomial and the PDE model, corresponding to the Amazon data of September 2001



FIGURE 5.1: Spreads and both the default intensity and recovery term structures for both the D&H and the PDE model. The input data is based upon Amazon data from September 2001.

# 5.3 Individual firm calibration with single seniority

We shall now provide three calibrations on different firms and present their results. The three cases that we will look at are:

#### 1. Banco do Brasil S.A.

This Brazilian state owned bank is the largest in Latin America. It performs a variety of services including those performed by regular commercial banks. But their services are more profound and from [SIC15] we find many more industry affiliation, such as agriculture and transportation. Financial institutions as a whole

list very low on the recovery rank list<sup>2</sup>, however being state owned is expected to promote the recovery as a bailout might be expected in case of bankruptcy.

2. Shell

Shell is known for its oil and gas extraction but its industry reaches much further, including public utilities and manufacturing. These last two list very high on the historical recovery rate [AK96]. Besides being part of an industry which seems to promote high recoveries, it has a lot of tangible assets, think about example crude oil and factories.

3. Air France - KLM

Being one of the largest airliners in Europe it had some complications in 2014. At the end of September 2014 a strike lasted for over a week pressing the financial wealth of the company, indeed this resulted in trimming their forecast on earning before interest, tax, depreciation and amortization [Reu14]. Such happenings change the market's expectations on default probabilities. Furthermore, air transportation estimates a historical recovery of around 40% [AK96].

In table 5.2 we present the initial data, such as date of evaluation and the corresponding stock price and volatility. The volatility is computed by Bloomberg, it corresponds to the historical<sup>3</sup> volatility Bloomberg uses to price a vanilla option with a 5 year maturity. Furthermore, figures 5.2 - 5.4 provide the visualisation of the fit and the corresponding term structures for both probability density function of the default and the expected recovery at default.

It is first of all interesting to note that all expected recoveries start of at their highest value and always decline. The authors of [DH09] reason this to be related to the higher recovery when it should happen suddenly within a very short period from now. As this comes so unexpected the firm was probably not for too long in a bad situation such that they did not spend all their assets yet, resulting in a larger recovery. The level of the recovery seems very reasonable given the industries of the firms. Indeed for Shell one might expect a large recovery due to the large amount of tangible assets. For Banco do Brasil S.A. the expectation is actually somewhat uncertain: financials usually do not have the reputation to provide large recoveries, however, considering this bank to be state owned and providing many more services besides regular commercial banks, the somewhat high recovery seems reasonable. Finally for Air France the recovery drifts around the 46%, somewhat larger than the historical recovery found within the industry of air transportation.

Now looking at the probability density function for the coming five years the market seems

<sup>&</sup>lt;sup>2</sup>Financial institutions usually have a bad name when it comes to recovery. Indeed financial institutions have historical recovery of around 35% of notional [AK96]. This should be interpreted with caution, as when looking at a more specific list we see that for example mortgage banks list second on highest average recoveries. We can therefore not simply state that all banks are expected to have low recoveries.

 $<sup>^{3}</sup>$ For completeness we note again that historical volatility is allowed in this implied framework, reason being that we only require the underlying stochastic process, in this case the stock, to be related to the implied data. As implied and historical data are related, the historical volatility suffices.

	Banco do	Brasil S.A.	S	hell	Air Fra	nce - KLM
Date	20-12	2-2014	24-12	2-2014	24-1	2-2014
$(S_0,\sigma)$	(24.35,	43.68%)	(18.01,	17.38%)	(7.62,	46.16%)
RMSE	11.1	$1 \mathrm{Bps}$	3.44	$4 \mathrm{Bps}$	6.1	$7 \mathrm{Bps}$
$\chi$	5.537	$' \cdot 10^{-3}$	2.146	$5 \cdot 10^{-2}$	1.791	$1 \cdot 10^{-3}$
$(a_0, a_1, b)$	(0.870, -2)	2.707, 1.015)	(1.916, -	$5.93, \ 1.108)$	(-0.077, -	0.0268, 1.664)
Т	Mkt. spr.	Model spr.	Mkt. spr.	Model spr.	Mkt. spr.	Model spr.
1	159.10	156.46	22.80	21.88	229.88	225.34
2	201.01	209.43	31.04	34.21	262.05	267.36
3	240.75	248.18	40.93	43.65	293.38	302.08
4	278.68	277.30	53.41	50.95	335.22	327.81
5	318.48	298.96	62.23	56.53	356.60	345.60





FIGURE 5.2: The calibrated PDE model on data of Banco do Brasil S.A. and the resulting term structures it induces.

to foresee difficulties for Air France where the default probability seems to increase. The density for Shell on the other hand seems fairly stable.

Finally, the fit on the spreads of Air France - KLM seems to be very reasonable, the curvature is somewhat towards the right, however, in the cases of Shell and Banco do Brasil S.A., the spreads seem to lie more on a linear line, which the model has difficulty with fitting, which leads us to the following observation:

Main observation 5.3.1 (A drawback). The PDE model has difficulty in fitting spreads that lie on a straight lie. This can be explained as follows:



FIGURE 5.3: The calibrated PDE model on data of Shell and the resulting term structures it induces.



FIGURE 5.4: The calibrated PDE model on data of Air France - KLM and the resulting term structures it induces.

- 1. The future default probability will decline due to the expected increase of underlying stock.
- 2. Due to the negative assumed correlation the recovery will increase together with it.

These two factors together make the spread decrease on the longer term explaining the tendency for curving to the right. Recalling that the PDE model provides perfect negative correlation, this tendency of bending to the right might be decreased by introducing less correlation, e.g. allowing for positive correlation with for example the Logit recovery model, or introducing a stochastic variable to alternate between positive and negative correlations.

**Remark 5.3.1.** The PDE is a fairly parsimonious model, providing only 3 model parameters and extending it with an additional model parameter allowing for extra flexibility is therefore possible. Furthermore, the CDS term structure contains quotes between 6 months up to 10 years, so additional market data might be of help within the calibration process.

**Remark 5.3.2.** A difficulty of implied data and its reliability is that it cannot be tested for. The best way in which we might establish the reliability of our results is to provide inter-comparison and relate this to historical results. For example, from historical data we know that firms from the sector of public utilities, on average, provide larger recoveries than financials [AK96]. Therefore the same might be expect from implied recovery. For such investigation, the model should be applied to a large data set.

#### 5.3.1 Individual firm calibration with multiple seniorities

A benefit of our model is that we can use it to calibrate the term structures to multiple seniority curves. We can take both senior and subordinated curves quoted on the same firm and calibrate these by letting parameter b be the same for both curves while letting the recovery be different. This is what is expected: when a senior CDS is triggered due to a default, then so is the subordinated, such that both have the same probability of default, i.e. the same parameter b. Thus for two such curves we have to calibrate

Parameter set = 
$$(a_0^0, a_1^0, a_0^1, a_1^1, b),$$
 (5.12)

with the aid of (in our case) 10 quotes. As was discussed before, the seniority is an important underlying of the recovery, therefore, introducing such extra curve could help us find better implied results.

We only found financials having curves for multiple seniorities. This could very well be due to the fact that these companies trade more in such products. We have therefore chosen to look at JPMorgan Chase & Co and ING group, both being very large banks.

With respect to the term structures we have one single expectation, namely that the senior secured recovery should be larger than the senior unsecured.

In table 5.3 and 5.4 we present the initial data and results. A visualization of the fit on the spreads and the term structures of both JPMorgan Chase & Co. and ING Group can be found in respectively figures 5.5 and 5.6.

We expected the senior recovery to be larger than the subordinated recovery, which is clearly the case for the ING Group data. However, for JPMorgan Chase & Co. we see

JPMorgan Chase & Co.					
Date	20-12-2014				
$(S_0,\sigma)$	(61.93, 18.	$(61.93, \ 18.69\%)$			
senior $RMSE$	$4.65 \mathrm{~Bps}$				
subordinated $RMSE$	$8.85 \mathrm{~Bps}$				
$\chi$	$4.057 \cdot 10^{-1}$	2			
$(a_0^0, a_1^0, a_0^1, a_1^1, b)$	(2.292, -9.101, 3.100, -19.940, 0.812)				
	Se	nior	Subor	dinated	
Т	Se Mkt. spr.	nior Model spr.	Subor Mkt. spr.	dinated Model spr.	
T 1	Se Mkt. spr. 22.13	nior Model spr. 21.34	Subor Mkt. spr. 34.50	dinated Model spr. 36.80	
T 1 2	Se Mkt. spr. 22.13 32.39	nior Model spr. 21.34 34.76	Subor Mkt. spr. 34.50 56.80	dinated Model spr. 36.80 56.96	
T 1 2 3	Se Mkt. spr. 22.13 32.39 39.26	nior Model spr. 21.34 34.76 43.39	Subor Mkt. spr. 34.50 56.80 65.40	dinated Model spr. 36.80 56.96 67.59	
T 1 2 3 4	Se Mkt. spr. 22.13 32.39 39.26 48.54	nior Model spr. 21.34 34.76 43.39 48.88	Subory Mkt. spr. 34.50 56.80 65.40 69.77	dinated Model spr. 36.80 56.96 67.59 73.49	

TABLE 5.3: Data and results for calibration on the multiple seniority curves for JPM organ Chase & Co.



FIGURE 5.5: The calibrated PDE model on data of JPMorgan Chase & Co. and the resulting term structures it induces.

ING Group					
Date	24-12-2014				
$(S_0,\sigma)$	$(10.95, \ 28.34\%)$				
senior $RMSE$	$0.82 \mathrm{~Bps}$	$0.82 \mathrm{~Bps}$			
subordinated $RMSE$	$2.10 \mathrm{~Bps}$	2.10 Bps			
$\chi$	$1.934 \cdot 10^{-3}$				
$(a_0^0,a_1^0,a_0^1,a_1^1,b)$	(1.81, -0.77, 0.82, -0.71, 1.51)				
	Senior Subordinated			dinated	
Т	Mkt. spr.	Model spr.	Mkt. spr.	Model spr.	
1					
	15.33	15.50	72.22	70.96	
2	15.33 25.81	$15.50 \\ 24.95$	72.22 85.15	70.96 87.21	
2 3	$     15.33 \\     25.81 \\     34.86 $	15.50 24.95 35.94	$72.22 \\ 85.15 \\ 100.14$	70.96 87.21 102.75	



56.25

130.03

127.86

55.50

5



FIGURE 5.6: The calibrated PDE model on data of ING Group and the resulting term structures it induces.

that at initiation the recoveries cross each another. This can be understood by the fact that the extrapolation of the spreads within our model also cross each other, and as this cannot be explained by the default intensity which is the same for both it has to be due to the recovery. Unfortunately this contradicts the rules of the market as we would expect them to be. But from these spreads it is also clear that liquidity plays a big role within these spreads as the five year CDS spread is unreasonably large compared to the other spreads. Liquidity will unfortunately always be a difficulty as it functions as non uniform noise on the spreads<sup>4</sup>.

The fit on ING Group seems to be near perfect considering the shape that our model accepts. Note also that the recovery turns out fairly high.

<sup>&</sup>lt;sup>4</sup>Liquidity is a known issue with respect to CDSs. It should however be noted that in the light of capital charge, Basel III will recognize a CDS as a hedge, providing incentive to buy CDSs and hopefully decreasing liquidity spreads.

# 5.4 Determining CVA

Up till now the usage of our model was merely to determine term structures for both the default intensity and the recovery. These structures can also be used to calculate credit risk adjustments, such as CVA. This particular adjustment is used to determine the counterparty credit risk on the value of financial products, such as interest rate swaps. As the focus of our model was not on interest rates, and instead we used the forward rate to forecast these values, it seems unreasonable to determine the CVA on an interest rate swap. Furthermore, we want to compare the CVA of our model with that of the PC model, so that focus on the exposure term of the CVA is not of great importance. To this end we determine the CVA on a 5 year plain vanilla call option on the corresponding stock price sold by the same party. Furthermore, we set the call option to be ATM, i.e. strike value is  $S_0$ , and use the average forward value as the interest rate to evaluate the Black Scholes formula<sup>5</sup>.

For a more comprehensive discussion on CVA we refer to [BMP13], a book on counterparty credit risk. We shall use the formula they provide for unilateral CVA evaluated on a product maturing at time T:

$$U_{\text{CVA}}(t,T) := \mathbb{E}_{t}^{\mathbb{Q}} \left[ \text{LGD}_{\tau} \mathbf{1}_{\{t < \tau \leq T\}} D(t,\tau) \text{Ex}(\tau) \right]$$
(5.13)

$$= \int_{t}^{T} \mathbb{E}_{t}^{\mathbb{Q}} \left[ \mathrm{LGD}_{u} D(t, u) \mathrm{Ex}(u) \right] \mathbb{Q} \left( \tau \in (u, u + du) \right)$$
(5.14)

$$= \int_{t}^{T} \mathbb{E}_{t}^{\mathbb{Q}} \left[ \mathrm{LGD}_{u} D(t, u) \mathrm{Ex}(u) \right] \lambda_{u} e^{-\int_{t}^{u} \lambda_{s} ds} du, \qquad (5.15)$$

where Ex(u) corresponds to the exposure at time u and D(t, u) the discount factor. Furthermore, in the second equation we neglect wrong-way risk and in the third we take the default time to be modeled by a hazard process with intensity  $\lambda$ , as is the case for both the PC and PDE model.

For the plain vanilla option, the exposure at time u is simply the value of the option at time u, being the expected value we stand to receive, which is always positive.

Finally, for the PC model we assume the LGD to be constant, that is, based on market convention, 40% for senior debt and 20% for subordinated debt. For the PDE model we can evaluate (5.15) by means of a Monte Carlo simulation. Each run simulates a stock

<sup>&</sup>lt;sup>5</sup>Evaluating the Black Scholes model in this sense provides in no way a fair price of such plain vanilla option. However, it is not our goal to model this price in a perfect way, instead we just want some exposure structure which makes sense to some extend. Note furthermore that plain vanilla call options are usually very liquid such that determing the CVA is not required, as the credit risk is likely to already be mingled within the price of the option.

CVA values				
	1			
	PDE model	PC model		
Banco do Brasil S.A.	1122.9	1460.8		
Royal Dutch Shell	239.3	308.38		
Air France - KLM	1208.5	1521.6		
JPMorgan Chase & Co. (senior)	216.1	310.3		
JPMorgan Chase & Co. (subordinated)	318.4	485.6		
ING Group (senior)	219.0	275.6		
ING Group (subordinated)	518.4	630.4		

TABLE 5.5: CVA of plain vanilla 5Y ATM call options sold by the same company. The value is presented in basispoints with respect to the corresponding currency.

path corresponding to a particular intensity and recovery process, giving us all we require to determine the CVA.

In table 5.5 we present the results of the CVA for both the PDE and the PC model. It provides CVA in terms of basis points with respect to the price of the plain vanilla option. The results lead us to the next immediate observation

Main observation 5.4.1 (CVA estimates). It follows clearly from table 5.5 that the PDE model values CVA strictly lower than the PC model.

In section 5.5 we show this to be a consequence of correlation.

Shifted PDE model for Air France - KLM			
Т	Market spreads	Model spreads	
1	229.88	230.82	
2	262.05	273.75	
3	293.38	309.12	
4	327.81	335.32	
5	356.60	353.44	

TABLE 5.6: Results of the shifted PDE spreads on Air France - KLM, corresponding to the parameter set (-0.110, -0.0268, 1.664).

## 5.5 A case comparison to the PC model

We shall now look at a particular example in which differences between our PDE model and the more simplistic PC model become evident. More specifically we shall show a scenario in which it becomes clear that our PDE model fits similar data but sets the recovery higher and the intensity lower, i.e. less overall expected risky results. This might be important for results such as CVA.

To this end we look at the Air France - KLM data which we calibrated before in section 5.3. We have calibrated the PC model such that it 'perfectly' fits the CDS spreads. However, as we have seen in figure 5.4, the PDE model spreads do not fit perfectly and sometimes are somewhat lower than the data. One might therefore argue that this might not be a fair comparison. To this end we recall from section 3.7.1 that  $a_0$  shifts the CDS spread curve vertically and from section 3.8 we know this to occur gradually. Furthermore it does not influence the probability of default. Trial and error shows us that shifting  $a_0$  from -0.077 to -0.110 provides a term structure which reaches above all spreads. Thus, for fair comparison we use the parameter set

$$(a_0, a_1, b) = (-0.110, -0.0268, 1.664).$$
 (5.16)

The corresponding spreads are found in table 5.6, from which we indeed see that the model spreads are at least as large as their market counterparts. Furthermore the resulting curve is presented in figure 5.7.

As the PDE model's curve now lies above the market spreads, it represents a somewhat more risky scenario. However, when comparing the term structures for the PC and PDE models (see figure 5.8), that is, comparing their expected implied recovery and default intensities, we see that the PDE model presents both an overall larger recovery and smaller default intensity. This is furthermore confirmed by the resulting CVA of 1244.1 Bps in comparison to the 1521.6 Bps in the PC framework.



FIGURE 5.7: The shifted Air France - KLM PDE model fit with respect to the related market spreads.



FIGURE 5.8: The shifted Air France - KLM PDE model fit versus the PC model.

This result can be understood by the fact that, for our PDE model, each time the intensity goes up, the recovery will go down, such that many risky scenarios are covered by their mutual movement. However, one might argue that the same thing occurs for downward movements of the intensity, which is true, however, the upwards movements are unbounded and seem therefore more influential. It is for this reason that modeling the correlation may be important, and hence, so is modeling the recovery. We can understand this reasoning in the following simplified example:

**Example 5.5.1** (Influence of negative correlation). Suppose a spread is simply determined by LGD times intensity and that we have 5 scenarios in which one is extreme, that is we have 4 times the intensity-recovery pair (0.03, 0.4) resulting in the spreads  $C^0 = 0.0180$  and once the riskier scenario (0.1, 0.2) resulting in  $C^1 = 0.08$ . The average intensity and recovery are found to be

$$mI = \frac{4 \cdot 0.03 + 0.1}{5} = 0.044 \qquad mR = \frac{4 \cdot 0.4 + 0.2}{5} = 0.36. \tag{5.17}$$

If we'd use the average recovery to find the same spreads we'd require 4 times  $I_{mR}^0 = 0.0281$  and once  $I_{mR}^1 = 0.1250$ , so that the average intensity for constant recovery equals

$$mI_{mR} = \frac{4 \cdot 0.0281 + 0.1250}{5} = 0.0475.$$
(5.18)

This indeed shows very simplistically that  $mI \leq mI_{mR}$ , i.e. the intensity is lower when an extreme upward case is present and the recovery is negatively related.

Main observation 5.5.1 (Importance of correlation). In table 5.5 we have seen that CVA is strictly lower in the PDE model compared to the PC model. Subsequently we have looked at a comparison favoring the PC model, providing similar results. Then from example 5.5.1 we understand this to be a consequence of negative correlation, enlightening us with its modeling importance.

ρ	Air France - KLM	Banco do Brasil	Shell	JPMorgan (senior)
0.5	1492.51	1450.92	311.15	308.13
0.4	1529.97	1481.27	312.38	309.35
0.3	1557.37	1504.29	313.25	310.18
0.2	1578.49	1521.60	313.91	310.80
0.1	1595.15	1535.19	314.42	311.28
0	1608.63	1546.16	314.84	311.66

TABLE 5.7: CVA values in basispoints when adjusting the recovery.

# 5.6 Implied recovery at its historical average

Finally, this last section will only look at the PC model. It has been discussed in both sections 1.1 and 1.3.2.1 how implied and historical data might differ. In particular it was seen that the implied recovery differs significantly from its historical counterpart due to additional premiums. In this section we provide a simple experiment to access whether setting the recovery at its historical average is justified.

For four data sets (Air France - KLM, Banco do Brasil and Shell) we calibrate the PC model and determine its CVA as explained in section 5.4 for five different recovery values, namely  $\rho = \{0, 0.1, 0.2, 0.3, 0.4, 0.5\}$ . The results are portrayed in table 5.7. We can clearly see that for the safer two firms, that is Shell and JPMorgan, a decrease of 10% in recovery corresponds in almost all cases to an increase of less than 1% in CVA. For the two more riskier sets, Air France - KLM and Banco do Brasil, a similar decrease in recovery results in differences up to almost 40 bps.

Main observation 5.6.1 (Implied or historical recovery?). Recall from section 1.3.2.1 that implied recoveries are found to be much lower than their historical counter parts. Taking into consideration the results of table 5.7, the PC model seems to substantially underestimate CVA.

**Remark 5.6.1.** We have seen that the PC model overestimates CVA in comparison with the PDE model, but that it underestimates the CVA if it is adjusted for its implied average. This once again emphasizes the necessity of conducting more research on the recovery in order to come up with more satisfactory results. 6

# Conclusion

"You stand to win everything. Call it."

- Anton Chigurh , No Country for Old Men

#### The original questions

This study originated from two questions on the constant recovery, of whether modeling the recovery as a constant suffices and if it does, whether it should be fixed at its historical average. In the literature study it became apparent that, due to additional risk premiums, a discrepancy between historical and implied average recovery exists. In particular, the implied recovery is substantially smaller than its historical counterpart. We have then shown that in a constant recovery framework a lower recovery results in larger CVA, especially notable for riskier firms, thus implying underestimation of market convention. However, under the specifications of the PDE model with perfect negative correlation between recovery and default we have seen CVA to be lower in comparison to the PC model, such that the assumption of constant recovery seems conservative in the light of capital requirements.

These results thus do not provide a satisfactory answer to our key questions. Future research might resolve this by adjusting the PDE model, as described next, and using a larger data set.

#### The PDE model

To the best of our knowledge this thesis is first in presenting a continuous not necessarily constant term structure for implied recovery. This is done by setting up a PDE, underlying the stock value, for both legs of the CDS. The recovery can be specified freely as a function of time, stock and/or default intensity. We choose to look at the former, allowing for negative correlation between recovery and default, and provided a similar specification as in the model introduced by Das & Hanouna [DH09], which under the correct assumptions and an explicit discretization is a special case of our PDE model. Their model results in a jump-to-default tree which lacks interpretability and numerical stability when time steps

converge to zero and they consequently required large steps. Under these large time steps we have shown that results lack credibility.

The PDE model provides numerical stability, seems to indicate dynamic stability and its model parameters show clear interpretation. With the stock as underlying we are able to determine term structures for the expected default intensity and recovery at default by means of a Monte Carlo simulation. Within this Monte Carlo framework we can in turn provide more complex calculations, e.g. CVA, and account for negative correlation.

Not all market CDS spreads could be explained completely by means of the PDE model. This could be a consequence of the following three reasons:

- 1. The model is fairly parsimonious, possibly not obtaining the required flexibility.
- 2. To overcome the identification problem a relationship between recovery and intensity was assumed. In particular it was set up to reflect perfect negative correlation. This in turn might be too stringent for a fair representation.
- 3. Prices might be affected by a lack of liquidity possibly introducing a premium for risk-aversion.

For future research, a different formulation for the intensity and/or recovery might allow for extra flexibility, e.g. the intensity depending on the stock with an additional variable. In this thesis we used spreads up to five years, however extra market data from the tenyear term structure might be used. Also, to reduce the amount of correlation one might model the dependence of the recovery on the intensity stochastically, providing a more realistic scenario without perfect negative correlation. If such issues are solved, then we may again reflect upon our key questions. Appendix A

# The PDE model: Supplements

### A.1 The probability of default under the stock limit

**Lemma A.1.1.** Let the default intensity be defined by  $\lambda = \frac{1}{S^b}$ , where S the stock value. Then when the stock increases towards infinity we have the following expression for the default probability

$$\lim_{S_t \to \infty} P_t(\tau > T) = 1, \tag{A.1}$$

for  $t \leq T < \infty$ .

**Proof.** What we need to prove is a statement of convergence in distribution, with the probability as the random variable. However, it is a general result that almost surely convergence implies convergence in distribution, and we therefore proceed to show such convergence to be valid.

Recall by the definition of the hazard process that

$$P(\tau > T) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_{t}^{T} S_{s}^{-b}} ds \middle| \mathcal{G}_{t}\right], \qquad (A.2)$$

where the stock price corresponds to the *pre*-default stock price, which follows a time inhomogeneous Geometric Brownian Motion.

The definition of almost surely convergence states that,  $\forall \epsilon > 0$ 

$$P\left(\lim_{s_t \to \infty} \left\{ \omega \in \Omega : \left| \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_t^T S_s^{-b} ds} \left| \mathcal{G}_t \right] - 1 \right| < \epsilon \right\} \right) = 1$$
(A.3)

should hold. Note that  $\Omega$  scales together with the limit such that it consistently contains all the possible stock paths starting at the correct initial value,  $S_t$ . We shall prove this statement by contradiction.

Suppose (A.3) not to hold, then a subsequence  $(S^{(n)})_{n\in\mathbb{N}}$  exists with  $\lim_{n\to\infty} S^{(n)} = \infty$ on which the path sets  $\Omega_n = \{\omega : S_t = S^{(n)}\}$  are defined, such that  $\forall n \in \mathbb{N}$  we can find an  $\omega_n \in \Omega_n$  for which

$$\left| e^{-\int_t^T S_s^{-b}(\omega_n) ds} - 1 \right| > \epsilon.$$
(A.4)

As the lhs of (A.4) is strictly negative, it implies that

$$1 - e^{-\int_t^T S_s^{-b}(\omega_n)ds} > \epsilon \quad \leftrightarrow \quad -\int_t^T S_s^{-b}(\omega_n)ds < \ln(1-\epsilon)$$
$$\leftrightarrow \quad \int_t^T S_s^{-b}(\omega_n)ds > \ln\left(\frac{1}{1-\epsilon}\right).$$

Let  $\overline{S} \in \mathbb{R}$  be the value for which it holds that

$$\int_{t}^{T} S_{s}^{-b}(\omega_{n}) ds = \frac{T-t}{\overline{S}^{b}},$$
(A.5)

then we know that  $S_{\min}(\omega_n) := \min_{t \leq s \leq T} \{S_s(\omega_n)\} \leq \overline{S}$ . From this we find that

$$S_{\min}^{b}(\omega_{n}) < T - t/\ln\left(\frac{1}{1-\epsilon}\right).$$
(A.6)

As for very small  $\epsilon$  we have

$$\ln\left(\frac{1}{1-\epsilon}\right) \approx \epsilon,\tag{A.7}$$

we may approximate (A.6) by

$$S_{\min}^b(\omega_n) < \frac{T-t}{\epsilon}.$$
 (A.8)

As this holds for all  $\omega_n$  this is a contradiction to the property of bounded variation.



FIGURE A.1: The term structure of CDS spreads given by the PDE model. We modify parameter b between 0.25 and 1.75, while looking at 4 different sets of  $\{1_0, a_1\}$  and  $S_0 = 100, \sigma = 0.5, r = 0.03$ .

# A.2 The switching line

During the discussion of the contribution of the model parameters on the PDE we found there to be an interesting behavior for positive correlation within the Probit framework, that is when  $a_1 > 0$ . Recall the initial upward drifting of the CDS spread from figure 3.2. In this section we show more carefully the dynamics behind this behavior and work out the formula presented in section 3.7.1.

The reason for this behavior is actually a consequence of a characteristic that is even more profoundly present in the recovery model more similar to that of the D&H model. To this end we take a step back and show how a similar scenario as figure 3.2 would look like using the recovery model

$$\rho(a_0 + a_1\lambda). \tag{A.9}$$

#### **Recovery model by D&H** In figure A.1 we plotted a similar result when using (A.9).

In the case of  $a_0 = 1$ ,  $a_1 = 3$  we see that the case b = 0.25 behaves somewhat strange in contrast to the rest. It moves upwards at initiation and it starts lower than when b = 0.5. As we shall see this is analogous to our initial observation.

For each parameter set  $\{1, 3, 0.25\}$ ,  $\{1, 3, 0.5\}$  and  $\{1, 3, 0.75\}$  we shall work out synthetic CDS spreads with maturity  $\Delta t$ , where we take  $\Delta t = 0.001$ , some very small time step into the future. Then we will analyze what causes the spread to start low, then increase

and subsequently decrease.

Note first that for the pricing of a CDS in the very near future we can use the approximation<sup>1</sup>:

$$C = \frac{\text{Protection leg}}{\text{Premium leg}} \approx \frac{\lambda_0 (1 - \rho_0) \Delta t}{1 - \lambda_0 \Delta t}$$
(A.10)

• Parameter set  $\{1, 3, 0.25\}$ :  $\lambda_0 = \left(\frac{1}{100}\right)^{0.25} \approx 0.316$  and  $\rho_0 = \Phi(1+3\lambda_0) \approx 0.9743$ , such that  $1 - \rho_0 \approx 0.0257$ . Then

$$C_{0.25} = \frac{0.316(1 - 0.9743) \cdot 0.001}{1 - 0.316 \cdot 0.001} = 8.1191 \cdot 10^{-6}.$$
 (A.11)

• Parameter set  $\{1, 3, 0.5\}$ :  $\lambda_0 = \left(\frac{1}{100}\right)^{0.5} = 0.1$  and  $\rho_0 = \Phi(1 + 3\lambda_0) \approx 0.9032$ , such that  $1 - \rho_0 \approx 0.0968$ . Then

$$C_{0.5} = \frac{0.1(1 - 0.9032) \cdot 0.001}{1 - 0.1 \cdot 001} = 9.6810 \cdot 10^{-6}.$$
 (A.12)

• Parameter set  $\{1, 3, 0.75\}$ :  $\lambda_0 = \left(\frac{1}{100}\right)^{0.75} \approx 0.0316$  and  $\rho_0 = \Phi(1+3\lambda_0) \approx 0.8632$ , such that  $1 - \rho_0 \approx 0.1368$ . Then

$$C_{0.75} = \frac{0.0316(1 - 0.8632) \cdot 0.001}{1 - 0.0316 \cdot 0.001} = 4.3257 \cdot 10^{-6}.$$
 (A.13)

As expected we get  $C_{0.25} < C_{0.5}$  and  $C_{0.75} < C_{0.5}$ . Just as we observed in figure A.1 the spread first increases and subsequently decreases. Now that we have worked out the above synthetic spreads we can understand where this behavior comes from: it is due to the fact that the LGD makes large relative jumps when it is very small. More specifically, looking back at the example we see that the LGDs are respectively 0.0257, 0.0968 and 0.1368, such that

$$\frac{0.0968}{0.0257} \approx 3.767, \qquad \frac{0.1368}{0.0968} \approx 0.707.$$
 (A.14)

This can better be understood when looking at LGD with respect to the log axis in figure A.2. Indeed, the derivative of log(LGD) gets steeper for larger values of  $a_0 + a_1\lambda$ .

There thus seems to be a switching point, which is an interesting issue, as it tells us something about the dynamics of the model. Suppose for example that we have priced a 5 year CDS today, and we expect the intensity to increase between today and tomorrow,

<sup>&</sup>lt;sup>1</sup>The premium leg stated next corresponds to the premium leg taking the spread to be 1. Furthermore, to get the spread as in figure A.1, we have to divide by the step size,  $\Delta t$ , (the spreads in the figure A.1 are with respect to one year) and multiply by 10<sup>4</sup> due to the basispoints representation.



FIGURE A.2: The recovery and LGD with respect to the log y axis. Increments become larger when they approach zero.

then how would the price of a 5 year CDS look like tomorrow? We know now whether the CDS spread curve will shift up or down, this is determined by the switching point. We shall determine this switching point.

We have thus seen that the synthetic spread for some initial intensity, given by

$$C_{\lambda_0} = \frac{\lambda_0 (1 - \Phi(a_0 + a_1 \lambda_0)) \Delta t}{1 - \lambda_0 \Delta t}, \qquad (A.15)$$

increases for small  $\lambda_0$  and decreases for larger  $\lambda_0$ . Its not too complicated to show that the derivative of the dummy spread is given by

$$\frac{dC_{\lambda_0}}{d\lambda_0} = \frac{\left((1-\Phi)\Delta t - a_1\lambda_0\phi\Delta t\right)\left(1-\lambda_0\Delta t\right) + \lambda_0(1-\Phi)\left(\Delta t\right)^2}{\left(1-\lambda_0\Delta t\right)^2},\tag{A.16}$$

where  $\phi$  corresponds to the probability density function of the normal distribution and we have used the short notation  $\Phi \equiv \Phi(a_0 + a_1\lambda)$  and  $\phi \equiv \phi(a_0 + a_1\lambda)$ . Recalling that the dummy spread is defined for very small values of  $\Delta t$  only, such that we may use the following approximation by neglecting higher orders of  $\Delta t$ :

$$\frac{dC_{\lambda_0}}{d\lambda_0} = \frac{(1-\Phi)\Delta t - a_1\lambda_0\phi\Delta t}{1-2\lambda_0\Delta t}.$$
(A.17)

The switching point is thus where the numerator equals zero. Setting  $x = a_0 + a_1\lambda_0$ , we are interested in the points where

$$1 - \Phi(x) = (x - a_0)\phi(x) \qquad \leftrightarrow \qquad a_0 = x - \frac{\Phi(-x)}{\phi(x)},\tag{A.18}$$


FIGURE A.3: The switching point at which the spread increases or decreases, in terms of  $a_1\lambda_0$  for several values of  $a_0$ .

where we used the fact that  $1 - \Phi(x) = \Phi(-x)$ . We can compute the right hand side of (A.18) for several values of x such that we determine simultaneously the values of x and  $a_0$  that lie exactly on the switching line. Subsequently we can use that  $a_1\lambda_0 = x - a_0$ . In figure A.3 we provided a plot of this line for several values of  $a_0$ , that is, for each  $a_0$  we know  $a_1\lambda$  such that there is neither increase nor decrease in the spread.

In figure A.3 we see that the value of  $a_1\lambda$  which determines the switching point for  $a_0 = 1$  is approximately 0.5. Recall that we were working with an example where  $a_0 = 1$  and  $a_1 = 3$ . We now know that the value for  $a_1\lambda_0$  required to put us on the switching point is approximately 0.5, such that  $\lambda_0 = 0.167$ . And indeed, this result seems to be in line with the previous observation that the switching line was found to lie between  $\left(\frac{1}{100}\right)^{0.5} = 0.1$  and  $\left(\frac{1}{100}\right)^{0.75} \approx 0.0316$ .

**Back to the PDE recovery model** Now that we have a proper notion of such switching line we can understand the analogy to our model. Note that

$$a_0 + a_1 \left(\frac{\lambda - \lambda_0}{\lambda_0}\right) = \underbrace{[a_0 - a_1]}_{a_0^*} + \underbrace{\left[\frac{a_1}{\lambda_0}\right]}_{a_1^*} \lambda.$$
(A.19)

Thus the same result holds for our model, in contrast however we shall not note this difference at initiation but over time. Thus for our case, when  $a_0 = 0.53$  and  $a_1 = 7.5$ , then  $a_0^* = -6.97$  and  $a_1^* = \frac{a_1}{\lambda_0}$ . As at initiation we have  $\lambda = \lambda_0$  we find that  $a_1^*\lambda_0 = a_1$ . From figure A.5 we see that for  $a_0^* = -6.97$  the switching line is at  $a_1 = a_1^*\lambda_0 \approx 5.626$ .



FIGURE A.4: The switching point at which the spread increases or decreases, in terms of  $a_1\lambda_0$  for several values of  $a_0$ .

This implies that the spreads will move in the opposite direction in comparison to the default intensity. Furthermore, we may note that now

$$\frac{dC_{\lambda}}{d\lambda}\Big|_{\lambda=\lambda_0} = \frac{(1-\Phi(a_0^*+a_1^*\lambda))dt - a_1^*\lambda\phi dt}{1-2\lambda dt}\Big|_{\lambda=\lambda_0} \tag{A.20}$$
$$(1-\Phi(a_0))dt - a_1\phi dt$$

$$=\frac{(1-\Phi(a_0))dt - a_1\phi dt}{1-2\lambda_0 dt},$$
 (A.21)

such that

$$\left. \frac{dC_{\lambda}}{d\lambda} \right|_{\lambda=\lambda_0} > 0 \leftrightarrow a_1 > \frac{\Phi(-a_0)}{\phi(a_0)}. \tag{A.22}$$

This means that when  $a_1$  is larger than the rhs of (A.22) the spreads will initially move into the opposite direction of  $\lambda$ , whereas when it is below it moves in the same direction. In particular, when  $a_0 = 0.53$ , we require  $a_1 = 0.86$  to set the time derivative in the initial spreads at zero. This can be confirmed by looking at figure A.5.

We thus conclude that the existence of this switching line is no coincidence and that it could be specified exactly for the Probit model. Finally we'd like to note that depending upon the choice of the recovery model this behavior might differ (see appendix A.3).



FIGURE A.5: Spreads with maturity less than one year for several values of b;  $a_0 = 0.53, a_1 = 0.86$  within the model  $\left(a_0 + a_1 \frac{\lambda - \lambda_0}{\lambda_0}\right)$ 

#### A.3 More recovery models for the PDE model

In this thesis we have only looked at the Probit model for the recovery. We are however not bounded to this particular model, there are many other possibilities, two of which we shall discuss here. Recall our discussion on the switching line in section 3.7.1, where we noted the results of a quicker increasing LGD compared to the intensity. This gives us an insight in the importance of the underlying recovery model.

Suppose at time  $t_0$  the default intensity to be 0.1 and the LGD 0.4. Now at time  $t_1$  the default intensity has dropped to 0.025, that is, it decreases by a factor 4. What may we expect from the LGD, with which rate will it decrease? This has no straightforward answer, but there are two important notes we should make:

- 1. We can control the amount at which it decreases to some extend by changing the value of  $a_1$ : when  $a_1$  is large this means that the jump from 0.1 tot 0.025 influences the recovery even more.
- 2. On the other hand we are bounded by the functional form of the recovery model.

It is exactly the second note that is important to us for understanding why a certain model might fit better. Suppose  $\rho_t$  and  $\rho_{t+dt}$  to be the recoveries at respectively time t and time t + dt where dt is taken to be small. The recovery thus increases by a factor

$$\frac{\rho_{t+dt}}{\rho_t}.$$

This particular relative increment depends on the functional form of the recovery function. And the speed at which it increases can be observed by looking at the derivative of  $\log(\rho_t)$ . However, the speed can be adjusted by simply modifying the value of  $a_1$ , which is the amount of change within the speed bounded by the functional form which we cannot influence by our model choice. Thus what we should really be looking at is the derivative of this speed (the acceleration of the recovery), i.e. the second derivative of  $\log(\rho_t)$ . In figure A.6 we plot these two functions together with the recovery value itself on the left. We can now clearly see some particular underlying assumptions that each model brings with it.

Each model has an underlying practical assumption:

- 1. The Probit model has its highest relative speed at minus infinity and then steadily decelerates towards a recovery value of 1. Practically: for small recovery values, a small movement in the default intensity is expected to result in large relative jumps (for the recovery). Whereas for large recovery, differences in the intensity are not expected to make any changes to the recovery.
- 2. The Logit model is more symmetrical. To both sides of the spectrum it converges to a steady state with a constant relative speed. However in the middle, that is around a 50% of recovery, it reduces its speed. This can practically be understood as assuming the recovery to have a constant sensitivity to the default intensity at the left of the spectrum, and then slowly (and symmetrically) becoming less sensitive until it becomes completely insensitive at infinity.
- 3. At minus infinity, the Arctan model starts with zero acceleration and then around a 50% recovery it changes from acceleration to quick and large deceleration to finally end in a steady state. Practically this would mean that the recovery slowly becomes more sensitive to intensity changes, being most sensitive at around 30 40% of recovery. Then it finally drops in sensitivity to reach a steady state of zero sensitivity at infinity.

Now one may make a better modeling choice upon his believes of the recovery. In this thesis we will calibrate only the Probit model and leave the others for further research. At this point we might expect the Probit and the Logit model to be more natural than the Arctan due to its sudden drops in sensitivity.



FIGURE A.6: The log of the LGD for the three different recovery functions. Each has a different speed at which it converges to zero.

## Appendix B

## **Pricing basics**

This section presents basics on pricing theory, in particular it defines the money market, its resulting risk neutral discounting and the martingale property. The martingale property for pricing is one frequently used when a derivative can be hedged, a property which the CDS legs lack. Finally we recall Ito's lemma, a practical tool towards a pricing PDE.

#### B.1 The risk neutral martingale framework

The definitions and theorems presented can be found in [BM06].

**Definition B.1.1** (Money-market account). Define B(t) the value of a bank account at time  $t \ge 0$ . Assume B(0) = 1 - an initial investment of 1 - and that the bank evolves according to

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

where  $r_t$  a positive function of time. It thus follows that

$$B(t) = e^{\int_0^t r_s ds} \tag{B.2}$$

It is thus stated that an initial investment of B(0) in the money market yields a value of B(t) at time t.

We will often speak of a discount rate, defined by

$$D(t,T) = \frac{B(t)}{B(T)} = e^{-\int_t^T r_s ds}, \quad t \le T,$$
(B.3)

where for convenience we define D(t,T) = 0, t > T. It can intuitively be understood as the value you need to invest in the money market at time t such that at time T you will be left with 1.

Consider a time horizon T > 0 and a probability space  $(\Omega, \mathcal{F}, \mathbb{Q})$ . Here  $\mathcal{F}$  gives rise to a right continuous filtration  $\mathcal{F}_t$ . Let the price process of a non-dividend paying security defined by  $S = \{S_t : 0 \le t \le T\}$ .

**Definition B.1.2.** A 1-dimensional trading strategy is a locally bounded and predictable process  $\phi = \{\phi_t : 0 \leq t \leq T\}$ . The value process associated with such strategy is defined by

$$V_t(\phi) = \phi_t S_t. \tag{B.4}$$

The predictability of  $\phi$  corresponds to the condition that  $\phi_t$  is known just before t, i.e. at t- (this condition is required in case  $S_t$  doesn't follow a continuous path).

Now [BM06] shows that the definition of self-financing can be reduced to

$$dV_t(\phi) = \phi_t dS_t. \tag{B.5}$$

Intuitively, the self financing strategy means that the value only changes due to changes in the corresponding assets. No additional cash in- or outflow is required (after initial time). Note that  $dV_t(\phi) = d(\phi_t S_t)$  would follow trivially from (B.4), but that (B.5) is not at all a trivial statement. **Definition B.1.3** (Contingent attainable claim). A contingent claim is a squareintegrable and positive random variable on  $(\Omega, \mathcal{F}, \mathbb{Q})$ . A contingent claim H is attainable if there exists some self-financing strategy  $\phi$  exists such that  $V_T(\phi) = H$ .

Definition B.1.4. A numeraire is any positive non-divident-paying asset.

The changing of numeraire gives rise to a new probability measure, more specifically an equivalent martingale measure, as defined by [BM06]:

**Definition B.1.5** (Equivalent martingale measure). An equivalent martingale measure  $\mathbb{Q}^N$  is a probability measure on the space  $(\Omega, \mathcal{F})$  such that

- i  $\mathbb{Q}$  and  $\mathbb{Q}^N$  are equivalent measures, i.e.  $\mathbb{Q}(A) = 0 \leftrightarrow \mathbb{Q}^N(A) = 0, \forall A \in \mathcal{F}.$
- ii The Radon-Nikodym derivative  $\frac{d\mathbb{Q}^N}{d\mathbb{Q}}$  belongs to  $L^2(\Omega, \mathcal{F}, \mathbb{Q})$ .
- iii The discounted asset price' process  $D(0, \cdot)S$  is an  $(\mathcal{F}, \mathbb{Q}^N)$ -martingale, i.e.  $\mathbb{E}^{\mathbb{Q}^N} [D(0, t)S_t | \mathcal{F}_u] = D(0, u)S_u$ , for  $0 \le u \le t \le T$ .

**Proposition B.1.1** (Changing of numeraire). Let U be an arbitrary numeraire, then there exists a probability measure  $\mathbb{Q}^U$ , equivalent to  $\mathbb{Q}$ , such that the price of any attainable claim Y, normalized under U is a martingale under  $\mathbb{Q}^U$ :

$$\frac{Y_t}{U_t} = \mathbb{E}^{\mathbb{Q}^U} \left[ \frac{Y_T}{U_T} | \mathcal{F}_t \right]$$
(B.6)

The risk-neutral measure  $\mathbb{Q}$  - on which we provide some intuition in section 1.1 - is a measure that leaves any attainable claim normalised by the money market account as a martingale, that is:

$$\frac{Y_t}{B(t)} = \mathbb{E}^{\mathbb{Q}}\left[\frac{Y_T}{B(T)}|\mathcal{F}_t\right] \quad \to \quad Y(t) = \mathbb{E}^{\mathbb{Q}}\left[D(t,T)Y_T|\mathcal{F}_t\right], \quad 0 \le t \le T.$$
(B.7)

As can be seen, the risk-neutral measure uses the money market as its numeraire. However, as proposition B.1.1 states, any numeraire might be used: the T-Forward measure is defined as an equivalent measure with respect to the zero-coupon numeraire. The authors of [BM06] define the zero-coupon bond with maturity T as a contract that guarantees its holder a payment of one unit at time T, without any intermediate payments, more specifically it holds that <sup>1</sup>

$$P(t,T) = \mathbb{E}_t^{\mathbb{Q}} \left[ D(t,T) \right]. \tag{B.8}$$

<sup>&</sup>lt;sup>1</sup>The zero coupon bond we refer to here is assumed to be risk free.

Where P(t,T) denotes its contract value at time t, clearly P(T,T) = 1. The T-Forward measure, exploits the fact that P(T,T) = 1 by defining its numeraire as  $P(\cdot,T)$ . Denote  $\mathbb{Q}^T$  as the T-Forward measure, then we have that

$$\frac{Y_t}{P(t,T)} = \mathbb{E}^{\mathbb{Q}^T} \left[ \frac{Y_T}{P(T,T)} | \mathcal{F}_t \right] \quad \to \quad Y(t) = P(t,T) \mathbb{E}^{\mathbb{Q}^T} \left[ Y_T | \mathcal{F}_t \right], \quad 0 \le t \le T.$$
(B.9)

Note finally how the rhs of (B.9) would follow from (B.7) when assuming independence between the money market account and the attainable claim.

#### B.2 Ito's lemma

We end this appendix by recalling Ito's lemma. Ito's lemma is not exclusively used in the pricing context, however it is a practical tool towards pricing PDEs, hence, we state it in this particular appendix.

We shall present its more general version, applicable for jump diffusion processes, used within this thesis.

Consider a general jump diffusion process X(t) for  $t \le t_0 = 0$  given by

$$X(t) = X(t_0) + \int_{-\infty}^{\infty} b(u, X_{u-}) du + \int_{-\infty}^{\infty} \sigma(u, X_{u-}) dW_u + \sum_{n=1}^{N(t)} \Delta X_n,$$
(B.10)

or in short form notation

$$X(t) = b(t, X_{t-})dt + \sigma(t, X_{t-})dW_t + \Delta X_n dN(t),$$
(B.11)

where  $b(t, X_{t-})$  and  $\sigma(t, X_{t-})$  are continuous adapted processes (that is measurable and predictable with respect to the canonical filtration  $\mathcal{F}_t$ ), with  $\mathbb{E}_{t_0}^{\mathbb{Q}} \left[ \int_{t_0}^t \sigma(u, X_{u-})^2 du \right] < \infty$ and  $\Delta X_n = X_{T_n} - X_{T_{n-}}$  where  $T_n, n = 1, ..., N(t)$  denote the jumping times of  $X_t$ . Consider a function  $f : [t_0, \infty) \times \mathbb{R} \to \mathbb{R}$  with  $f \in C^{1,2}$ . Note that stating f to be twice continuous differentiable does note mean that the resulting function may not have jumps do to a underlying jumping process.

When  $t \in (T_n, T_{n+1})$ , that is in between two jumping times, Ito's lemma for a regular diffusion process holds, that is in short form:

$$df(t, X_t) = \frac{\partial f}{\partial t}(t, X_{t-})dt + b(t, X_{t-})\frac{\partial f}{\partial X}(t, X_{t-})dt + \frac{1}{2}\sigma^2(t, X_{t-})\frac{\partial^2 f}{\partial X^2}dt + \sigma(t, X_{t-})\frac{\partial f}{\partial X}dW_t$$
(B.12)

If however t corresponds to a jumping time, then the resulting change in  $f(t, X_t)$  constitutes of the extra term

$$f(t, X_{t-} + \Delta X_t) - f(t, X_{t-}).$$
 (B.13)

Ito's lemma for a jump diffusion process can be determined as the sum of the continuous part and the jumping part, that is

$$df(t, X_t) = \frac{\partial f}{\partial t}(t, X_{t-})dt + b(t, X_{t-})\frac{\partial f}{\partial X}(t, X_{t-})dt + \frac{1}{2}\sigma^2(t, X_{t-})\frac{\partial^2 f}{\partial X^2}(t, X_{t-})dt + \sigma(t, X_{t-})\frac{\partial f}{\partial X}(t, X_{t-})dW_t + (f(t, X_{t-} + \Delta X_t) - f(t, X_{t-}))dN_t.$$
(B.14)

This is not a full proof, but it does provide a decent understanding of the result.

## Appendix C

# Debt hierarchy and capital structure

Suppose a firm with two debts: debts A and B. If this firm goes bankrupt it usually doesn't have enough money to completely pay back both creditors. When paying back the corresponding recovery th efirm has to follow the APR: debt with a higher rank should be settled first. Such rank is called the seniority and the related hierarchy is called the debt hierarchy. This debt hierarchy roughly looks as follows:

- Senior secured;
- Senior unsecured/unsubordinated;
- Senior subordinated;
- Subordinated;
- Junior subordinated.

Proceeding the example: if debt A is senior unsubordinated and debt B is subordinated, then following the APR, debt A will be settled *completely* prior to settling B.

The term secured in the hierarchy refers to being backed up by collateral. It should be noted that senior secured has full priority over unsecured debt, but only up to the value of the collateral. Only after all creditors are payed back, the shareholders receive their part of the deal.

The above list is not the most detailed one may find. For example, secured debt has a special further breakdown called lien: 1st lien, 2nd lien, etc., which indicates the order of priority the creditor/debtholder has against the secured assets. The coarse overview however suffices for our understanding.

Besides intuïtively understanding the importance of the capital structure in case of a

default, it is indeed empirically shown to be a highly influential determinant of recovery. For this reason it can be exploited when creating models within credit risk, see for example [SUH14] and [UMG03].

The *capital structure* of a firm is an overview of its outstanding debts based upon the debt hierarchy. It shows us the percentage of each type of debt with respect to its total liabilities. It can be of great importance to know the capital structure of a firm when lending it money. Suppose a firm to have 20% of secured debt and over 70% of unsubordinated debt, then an institution should be cautious with lending it a senior subordinated obligation, whereas this would seem less of a problem if unsubordinated debt would constitute, for example, only 10% of the total liabilities.

## Appendix D

## Modeling the default time

An important issue of working with credit derivatives is modeling the default process. There are two main approaches to model the default process. The first one being the structural approach, which comes down to making the default time dependent on the firm values. An example of such may be seen in [FJOS10], where default is supposed to occur when the firm value hits a certain lower threshold. The second approach is called the reduced-form approach. This approach is based on the assumption that default occurs by 'surprise': it is an exogenous process or more precisely its stopping time is totally inaccessible. The latter meaning that no increasing sequence of stopping times exists such that the default can be announced. The term *reduced-form* comes from the fact that the flow of information is reduced in the sense that there is no information telling us something about when a default might occur.

What follows is a deeper understanding and some required results of the hazard process approach, an often used reduced-form model. Many results, and in particular those in section 3.2, depend on a proper understanding of this approach.

If not specified differently, definitions and results come from [BJR07].

#### D.1 Filtrations

A filtration is an increasing sequence of  $\sigma$ -algebras defined on a measurable space, that is, suppose the probability space  $(\Omega, W, P)$ , then a filtration is a sequence of sub- $\sigma$ -algebras  $\{W_t\}_{t\geq 0}$  where for  $t_1 \leq t_2$  it holds that  $W_{t_1} \subseteq W_{t_2}$ .

Intuïtively, a filtration in the way that we will use it, might be understood as a flow of information: at each time a certain amount of information is observed and we do not forget any information.

We fix a finite horizon date by T, e.g. the maturity of a product of interest.

**Definition D.1.1.** The reference filtration  $\mathbb{F} = (\mathcal{F}_t)_{0 \le t \le T}$  is generated by prices of assets and other common economic factors, e.g. interest rates and volatilities.

**Definition D.1.2.** Let  $\tau$  be the time of default and associate to this a rightcontinuous increasing process defined by

$$H_t := \mathbf{1}_{\{t \ge \tau\}} = \begin{cases} 1, & t \ge \tau \\ 0, & t < \tau \end{cases}$$
(D.1)

With the usage of definition D.1.2 we can introduce a second filtration generated by this particular default indicator function.

**Definition D.1.3.** The *default filtration*  $\mathbb{H} = (\mathcal{H}_t)_{0 \le t \le T}$  is generated by  $H_t$ , more specifically:  $\mathcal{H}_t = \sigma(H_u : u \le t)$ . Intuïtively this sigma algebra contains exclusively information regarding the default process.

By definition of generating  $\sigma$ -algebras,  $\mathbb{H}$  is the smallest filtration making the function  $H_t$  measurable.

Finally we define an enlarged filtration containing both information regarding economic factors and the default process:

**Definition D.1.4.** The greater filtration is defined as  $\mathbb{G} = \mathbb{F} \cup \mathbb{H} = (F_t \cup H_t)_{0 \le t \le T}$ .

**Remark D.1.1** (Alternative notation of the greater filtration). When we write a conditional probability or expectation using the subscript of a time, e.g.  $P_t(\cdot)$  or  $\mathbb{E}_t[\cdot]$ , then this is always with respect to the greater filtration.

Definitions are in place in order to state an important lemma regarding change of filtration. Let  $(\Omega, \mathcal{G}, P)$  be a probability space, where  $\mathbb{G}$  is a filtration on it. The following lemma is proved in several references such as [BM01] and [BJR07]: **Lemma D.1.1** (Change of filtration). For any  $\mathcal{G}$ -measurable random variable Y such that  $\mathbb{E}^{P}[Y] < \infty$ , and any  $t \in \mathbb{R}_{\geq 0}$ , it holds that

$$\mathbb{E}^{P}\left[\mathbf{1}_{\{\tau>t\}}Y\big|\mathcal{G}_{t}\right] = \mathbf{1}_{\{\tau>t\}}\mathbb{E}^{P}\left[Y|\mathcal{G}_{t}\right] = \mathbf{1}_{\{\tau>t\}}\frac{\mathbb{E}^{P}\left[\mathbf{1}_{\{\tau>t\}}Y\big|\mathcal{F}_{t}\right]}{P\left(\tau\leq t|\mathcal{F}_{t}\right)}.$$
 (D.2)

Finally, a random variable X is said to be  $\mathcal{W}$ -measurable if and only if  $\mathbb{E}[X|\mathcal{W}] = X$ . Intuitively this means that the outcome of the random variable is knowable based on the information of  $\mathcal{W}$ .

#### D.2 The hazard process

Even though the default time is not known within the filtration  $\mathbb{F}$  we can associate probabilities to it. This is exactly what is done with the hazard process.

Let  $(\Omega, \mathcal{G}, P)$  be a probability space such that  $\mathbb{G}$  is a filtration on it. For any  $t \in \mathbb{R}_{\geq 0}$ we write  $F_t = P(\tau > t | \mathcal{F}_t)$ . Note how  $F_t$  corresponds to the distribution function of the  $\mathbb{F}$ -default time. It however incorporates a stochastic feature by letting the distribution function depend on  $\mathbb{F}$ . This implies that at some time s < t, we may not know  $F_t$ , a feature especially interesting for a Cox process which will be defined soon. Furthermore define  $G_t = 1 - F_t = P(\tau > t | \mathcal{F}_t)$ , the survival function of the default process.

An  $\mathbb{F}$ -hazard process - if the  $\mathbb{F}$  filtration underlying it is obvious we will omit referring to it - can now be defined as follows:

**Definition D.2.1.** Assume that  $F_t < 1$  for any  $t \in \mathbb{R}_{\geq 0}$ . The  $\mathbb{F}$ -hazard process of  $\tau$  under P is denoted by  $\Gamma$  and is defined by satisfying  $G_t := e^{-\Gamma_t}$ .

An especially interesting result from the hazard process is the following martingale proposition

**Proposition D.2.1** (the martingale compensated bankruptcy jump process[Lin06])). If the process  $\Gamma_t$  is increasing and continuous (which we will see to be valid for the canonical hazard rate process), then the process defined by  $M_t := H_t - \Gamma(\max\{t, \tau\})$  is a  $\mathbb{G}$ -martingale.

For the proof of this proposition we refer to [BJR07].

**The canonical hazard process** The canonical hazard process is the most commonly used construction of a hazard process. A nice feature is that its form corresponds to that of the interest rate. The canonical hazard process is constructed by taking a non-negative,  $\mathbb{F}$ -progressively measurable process  $\lambda = (\lambda_t)_{0 \leq t \leq T}$ , called the intensity process, and subsequently defining

$$\Gamma(t) \equiv \Gamma_t = \int_0^t \lambda_s ds. \tag{D.3}$$

We are not yet ready however, because the definition of the hazard process (definition D.2.1) states that there is a relation between the survival function and (D.3). Indeed, we might simply implement this, but the beauty lies in the fact that we can more specifically define the default time  $\tau$ : let  $\xi \sim \text{Exp}(1)$ , then, following the authors of[BM01] and [BJR07], we define the default time by

$$\tau := \inf\{t \in \mathbb{R}_{>0} : \Gamma_t \le \xi\},\tag{D.4}$$

such that as written in [BM06] we have

$$\Gamma_{\tau} = \xi \quad \leftrightarrow \quad \tau = \Gamma^{-1}(\xi).$$
 (D.5)

The exogenous property of the default time is thus understood by being related to the time value of the random variable  $\xi$ . Our model has no influence upon  $\xi$ . Of course a very large  $\lambda_t$  indicates that the default time should be reached earlier, but more than this cannot be said.

With the definition of the default time as in (D.4) we can indeed make sense of definition D.2.1:

$$P(\tau > t | \mathcal{F}_t) = P(\Gamma_\tau > \Gamma_t | \mathcal{F}_t) = e^{-\Gamma_t}.$$
(D.6)

Finally we note that  $\Gamma_t$  is continuous and increasing by construction, from which follows that the proposition D.2.1 is satisfied. Furthermore the fact that  $\lambda_t$  is progressively measurable tells us that it may depend upon components underlying  $\mathcal{F}_t$ .

At some point in this thesis we will define  $\lambda_t = S_t^{-b}$  and as a consequence the survival function will be a stochastic process: a Cox process, or also called, doubly stochastic Poisson process.

We end this section with two simple propositions which will be used several times within this thesis.

**Proposition D.2.2.** Let  $\Gamma_t$  be a canonical hazard process, with intensity  $\lambda_t$ , then the following two statements are valid

i 
$$dF_t = e^{-\Gamma_t} \lambda_t dt$$
,  
ii  $d(e^{\Gamma_t}) = e^{\Gamma_t} \lambda_t dt$ .

**Proof.** Recall that  $\Gamma_t = \int_0^t \lambda_s ds$ , from which follows that

$$d(\Gamma_t) = d\left(\int_0^t \lambda_s ds\right) = \int_0^{t+dt} \lambda_s ds - \int_0^t \lambda_s ds = \int_t^{t+dt} \lambda_s ds = \lambda_t dt.$$
(D.7)

The propositions now follow from working out only the first one, the second is immediate when this deduction is understood.

$$dF_t = d\left(1 - e^{-\Gamma_t}\right) = -d\left(e^{-\Gamma_t}\right) = e^{-\Gamma_t}d\left(\Gamma_t\right) = e^{-\Gamma_t}\lambda_t dt \tag{D.8}$$

#### D.3 The Cox process

A Cox process is a canonical hazard process where the intensity itself is stochastic,  $\mathcal{F}_{t}$ addapted and right continuous (and thus progressive). As it is a canonical hazard process
it follows the same building blocks for the default time, such that we still have

$$\Gamma_{\tau} = \xi. \tag{D.9}$$

What now changes are the probabilities of default: the rhs of (D.6) is not a valid probability in the Cox process as the intensities in the exponent are unknown within the filtration  $\mathcal{F}_t$ . Therefore we state the following results as a proposition, such that we may use these throughout this thesis:

**Proposition D.3.1.** Let the default time  $\tau$  follow a Cox process with intensity  $\lambda$ , then for  $s \geq t$  the survival probability is given by

$$P_t(\tau > s) = \mathbb{E}^{\mathbb{Q}}\left[ e^{-\int_t^s \lambda_u du} \Big| \mathcal{G}_t \right].$$
(D.10)

Furthermore, the probability density function of the default, given no default up till time t, is given by

$$\mathbb{E}^{\mathbb{Q}}\left[\lambda_{s}e^{\int_{t}^{s}\lambda_{u}du}\Big|\mathcal{G}_{t}\right].$$
(D.11)

They have their proof in [BM06] which is a result of using the tower property and a similar deduction as was used to deduce the survival probability for the regular canonical process. From the probability density function we thus also find [BM06]

$$P(\tau \in (s, s + ds]) = \mathbb{E}^{\mathbb{Q}} \left[ \lambda_s e^{\int_t^s \lambda_u du} ds \middle| \mathcal{G}_t \right].$$
(D.12)

### Appendix E

## Convertible bond and PDE model

A typical convertible bond is a callable bond that also includes an option for the investor to convert the bond into a given number of shares of equity (known as the conversion ratio)[DS03]. More specifically a convertible bond pays the maximum between the agreed face value and the value of the stock (times the conversion ratio). Whenever the issuer decides to call a convertible the investor still has the option to either accept the called upon value or the shares.

We are especially interested in modeling such convertible bonds, as we will see these can be represented by similar PDEs as we have found for both legs of the CDS, with some additional interpretation.

#### E.1 A pricing PDE

The deduction that follows comes from [AFV03]. Denote by V(t, S) the value of a convertible bond, where t the time of valuation and S the corresponding stock. Let's furthermore assume the simplified case in which there are no call features as described above, this means that conversion can only be performed at maturity of the contract. Define by  $S^+$  the value of the stock right after a default and  $S^-$  just before the default such that  $S^+ = (1 - \eta)S^-$  for some  $\eta \in [0, 1]$ . For  $\eta = 1$  we are looking at stock process that jumps to zero in case of default (so is the case for the jump to default model used to determine the PDE model). Finally, denote by  $\gamma$  the conversion rate.

Next we proceed to apply a hedging argument to determine a model for the convertible. To this end we construct a hedging portfolio

$$\Pi = V - \beta S,\tag{E.1}$$

where  $\beta$  indicates the amount of stocks in the portfolio. For convenience we shall first determine the PDE for this portfolio when the stock is risk free (in the sense that it doesn't have the ability of jumping to default), that is

$$dS = \mu S dt + \sigma S dW_t. \tag{E.2}$$

By Ito's lemma we find that

$$d\Pi = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial S^2}dS^2 - \beta dS$$
  
=  $\left(\frac{\partial V}{\partial t} + \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial S^2} - \beta \mu S\right)dt + \sigma S\left(\frac{\partial V}{\partial S} - \beta\right)dW_t$   
=  $\left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial S^2}\right)dt,$ 

where for the last equation we took  $\beta = \frac{\partial V}{\partial S}$ . Let's now look at the *risky* stock, that is the stock that jumps to  $(1 - \eta)S^-$  in case of a default. To this end we take  $\lambda$  to be the hazard rate and  $\rho$  the recovery of the bond. Now, we assume the following to hold at default: the bond holder may choose between

- 1. the recovery of the bond, or
- 2. shares worth  $\gamma S^+ = \gamma S^-(1-\eta)$ .

The change of value in the portfolio during  $t \to t + dt$  for very small dt is

$$d\Pi = (1 - \lambda dt) \left( \frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 \frac{\partial^2 V}{\partial S^2} \right) dt + \lambda dt \left( \max\left\{ \gamma S^-(1 - \eta), \rho \right\} - V + \beta \left( S^+ - S^- \right) \right) \\ = (1 - \lambda dt) \left( \frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 \frac{\partial^2 V}{\partial S^2} \right) dt - \lambda \left( V - \frac{\partial V}{\partial S} S\eta \right) dt + \lambda \max\left\{ \gamma S^-(1 - \eta), \rho \right\} dt \\ = \left( \frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 \frac{\partial^2 V}{\partial S^2} \right) dt - \lambda \left( V - \frac{\partial V}{\partial S} S\eta \right) dt + \lambda \max\left\{ \gamma S^-(1 - \eta), \rho \right\} dt,$$
(E.3)

where we have taken the convention that  $S \equiv S^-$  and removed higher orders of dt. Assuming the expected return on the portfolio to be risk free, that is

$$\mathbb{E}_t^{\mathbb{Q}}\left[d\Pi\right] = r\Pi dt. \tag{E.4}$$

Using this expectation in (E.3), we find

$$r\Pi dt = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2\frac{\partial^2 V}{\partial S^2}\right)dt - \lambda\left(V - \frac{\partial V}{\partial S}S\eta\right)dt + \lambda\max\left\{\gamma S^-(1-\eta),\rho\right\}dt$$
  

$$\leftrightarrow r\left(V - \frac{\partial V}{\partial S}S\right) = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2\frac{\partial^2 V}{\partial S^2}\right) - \lambda\left(V - \frac{\partial V}{\partial S}S\eta\right) + \lambda\max\left\{\gamma S^-(1-\eta),\rho\right\}$$
  

$$\leftrightarrow \frac{\partial V}{\partial t} + (r+\lambda\eta)S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2\frac{\partial^2 V}{\partial S^2} - (\lambda+r)V + \lambda\max\left\{\gamma S^-(1-\eta),\rho\right\} = 0.$$
(E.5)

The similarity between (E.5) and both (3.37) and (4.6) is clear. In particular it looks very similar to the (4.6) where instead of the LGD we now take the recovery. If we furthermore make the assumption that the stock jumps to zero in case of a default, i.e.  $\eta = 1$ , we get the exact same PDE as (4.6):

$$\leftrightarrow \frac{\partial V}{\partial t} + (r+\lambda) S \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 \frac{\partial^2 V}{\partial S^2} - (\lambda+r) V + \lambda \rho = 0.$$
(E.6)

Note that the similarity is only in the form of the PDE, that is, in the behavior of both values. Boundary conditions are different and consequently so are the solutions. But we can understand why both behave the same way: the protection leg corresponds to the protection that the CDS seller has to pay in case of a default, in other words, it pays exactly the amount the CDS buyer may have lost on a reference bond. The buyer of a convertible bond receives exactly the amount that was recovered from such bond. Thus both move the same way, as they are related to the same object, however from the convertible bond it receives the recovery, whereas from the CDS it gets the LGD.

Recall that we have already mentioned that such hedging approach would not work for our CDS legs. Compared to the convertible bond, our PDE lacks two things: 1) the stock jumps to zero such that the stock is worthless and 2) the option holder still has the right on the bond, representing the recovery.

#### Appendix F

## Von Neumann numerical stability

The Von Neumann stability analysis provides a very general method to investigate stability of finite differences schemes. In this section we explain how this method works and how it may be used to evaluate one-step finite differences schemes for linear inhomogenous variable coefficient PDE's, such as (3.3.1) and (3.3.2).

We'd like to emphasize here that stability in this sense refers to numerical stability, that is numerical errors generated during the iterations of discretized equations should not be magnified, in contrast to dynamic stability which refers to a system in which small variations from a reference state will not grow.

Before we get down to the Von Neumann stability analysis we should answer the question, why are we interested in obtaining numerical stability? Obviously a first answer would be: numerical instability would interfere with our result and therefore the result might not be a good approximation of the true solution. That is of course true. But even a deeper answer lies upon the result of the Lax-Richtmyer Equivalence theorem, stated next[GO86].

**Theorem F.0.1. The Lax-Richtmyer Equivalence theorem** A well-posed linear problem with a consistent approximation is convergent if and only if it is numerical stable.

In other words, if we are able to show that a linear well-posed PDE has a consistent and stable finite difference scheme, then we can assert convergence.

The Von Neumann stability analysis works by specifying a condition in the Fourier transformed domain. To this end a Fourier transformation should be applied, therefore this is where we start our explanation.

Let  $u_n^m$  correspond to the solution of a finite difference scheme where n = 0, 1, ..., N corresponds to the time index and m = 0, 1, ..., M to the spacial index. Furthermore we take

stepsize  $\Delta t$  in the time direction and  $\Delta x$  in the spacial direction. The Fourier transform of  $u_n^m$  indexing over the stock index is defined by

**Proposition F.0.1.** A simple Fourier transform Define  $\Delta x = \sum_{n=1}^{M} x_n x^n$ 

$$\hat{u}_n(\eta) = \frac{\Delta x}{\sqrt{2\pi}} \sum_{m=0}^{M} e^{-im\Delta x\eta} u_n^m, \quad \eta \in \left[-\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x}\right],$$
(F.1)

then

$$u_n^m = \frac{1}{\sqrt{2\pi}} \int_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} e^{im\Delta x\eta} \hat{u}_n(\eta) d\eta.$$
(F.2)

Proof.

$$\frac{1}{\sqrt{2\pi}} \int_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} e^{im\Delta x\eta} \hat{u}_n(\eta) d\eta = \frac{1}{\sqrt{2\pi}} \int_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} e^{im\Delta x\eta} \left(\frac{h}{\sqrt{2\pi}} \sum_{l=0}^M e^{-iml\Delta x\eta} u_n^l\right) d\eta$$
$$= \frac{\Delta x}{2\pi} \sum_{l=0}^M u_n^l \int_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} e^{im\Delta x\eta} e^{-il\Delta x\eta} d\eta$$
$$= \frac{\Delta x}{2\pi} \sum_{l=0}^M u_n^l \int_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} e^{i(m-l)\Delta x\eta} d\eta$$
$$= \frac{\Delta x}{2\pi} u_n^m \frac{2\pi}{\Delta x}$$
$$= u_n^m.$$

Next we provide a toy example to introduce the amplification factor. Subsequently we shall state the stability theorem and show how this works out on this particular example.

**Example F.0.1.** Suppose the simple PDE given by  $v_t + av_x = 0$ , x, t > 0. Lets discretize this PDE with a forward-time backward-space scheme, such that we get:

$$\frac{u_{n+1}^m - u_n^m}{\Delta t} + a \frac{u_n^m - u_n^{m-1}}{\Delta x} = 0.$$
 (F.3)

This scheme may be rewritten to

$$u_{n+1}^m = (1 - a\lambda)u_n^m + a\lambda u_n^{m-1}, \tag{F.4}$$

where  $\lambda = \frac{\Delta t}{\Delta x}$ . Next, apply (F.2) such that (F.4) becomes

$$\frac{1}{\sqrt{2\pi}} \int_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} e^{im\Delta x\eta} \hat{u}_{n+1}(\eta) d\eta = \frac{1}{\sqrt{2\pi}} \int_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} e^{im\Delta x\eta} \left( (1-a\lambda) + a\lambda e^{-i\Delta x\eta} \right) \hat{u}_n(\eta) d\eta.$$
(F.5)

Due to the fact that the Fourier transform is unique we understand that the integrands of (F.5) should equal, that is

$$\hat{u}_{n+1}(\eta) = \left( (1 - a\lambda) + a\lambda e^{-i\Delta x\eta} \right) \hat{u}_n(\eta), \quad \eta \in \left[ -\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x} \right].$$
(F.6)

Now, the amplification factor is  $g(\theta)$  where  $\theta = \Delta x \eta$ , such that  $\hat{u}_{n+1}(\eta) = g(\theta)\hat{u}_n(\eta)$ . Thus in this example we have  $g(\theta) = (1 - a\lambda) + a\lambda e^{-i\Delta x\eta}$ .

We are now in a position to provide the theorem for stability that we require. This theorem is stated and proven in [Str89].

**Theorem F.O.2.** (Stability condition)

A one-step finite difference scheme is stable in a stability region  $\Lambda$  if and only if there exists a constant  $C \in \mathbb{R}$  (independent of  $\theta$ ,  $\Delta x$  and  $\Delta t$ ) such that for the amplification factor it holds that

$$|g(\theta, \Delta x, \Delta t)| \le 1 + C\Delta t, \tag{F.7}$$

where  $(\Delta t, \Delta x) \in \Lambda$ .

The restriction placed upon the amplification factor makes sense in the non Fourier transformed world. However as the theorem states stability is granted when such restriction is placed in the Fourier transformed world. This result can however be understood by Parsevals identity which states that

$$\|u_n^m\|_2 = \|\hat{u}_n(\eta)\|_2. \tag{F.8}$$

where the subscripts refer to the  $L^2$ -norm. And thus the restriction persists.

There are two important notes we'd like to make regarding the Stability condition:

- 1. The Fourier transformation presented in proposition F.0.1 and the deduction in the toy example show us that in order to investigate on stability, all that is required is to substitute  $u_n^m$  by  $e^{im\Delta x\eta}u_n(\eta)$  and retrieve the amplification factor.
- 2. The original theorem as in [Str89] requires the PDE to be constant coefficient. This requirement may however be removed as it also states in [Str89] that the **variable coefficient** PDE can be evaluated in the exact same way, where one might consider the so called *frozen* coefficient problem, that is, fixing the problem for all possible coefficients and solving (F.7).
- 3. As may be seen in the toy example, the amplification factor was easily found. However when observing a **inhomogeneous PDE** this is not straightforward as an additional constant would enter (F.6) additively. However, due to Duhamel's principle which states that the solution of an imhomogeneous PDE is a superposition

of solutions to the homogeneous PDE, it follows that a finite difference scheme for an inhomogeneous scheme is stable if the corresponding homogeneous scheme is stable[Str89].

**Remark F.0.3.** Theorem F.0.2 gives an analogous result to stability as in the matrix approach. In a nutshell: if the PDE may be discretized into the form

$$u_{n+1} = Au_n,\tag{F.9}$$

for some discretization matrix A, then a necessary and sufficient condition for stability satisfies

$$\|A\| \le 1 + C\Delta t \tag{F.10}$$

for some constant  $C \in \mathbb{R}$ .

We now state a corollary which will also be used in this thesis. For the proof we refer to [Str89].

**Corollary F.0.1.** If a scheme as in theorem F.0.2 is modified such that the modifications result only in the addition to the amplification factor of terms that are  $\mathcal{O}(\Delta t)$  uniformly in  $\eta$ , then the modified scheme is stable if and only if the original scheme is stable.

## Appendix G

## Pricing a defaultable bond

Suppose a defaultable bond with maturity T, face value F, and paying n coupons of value C at times  $\left\{\frac{T-t}{n}, 2\frac{T-t}{n}, ..., T\right\}$ , where t indicates the current time. Let  $\tau$  be the time of default modeled by a hazard process with time dependent intensity  $\lambda_s$ . The price of the defaultable bond today, t, can be written as

$$V_{t} = \mathbb{E}_{t}^{\mathbb{Q}} \left[ \sum_{j=1}^{n} C e^{-\int_{t}^{t+\frac{T-t}{n}j} r_{s} ds} \mathbf{1}_{\left\{\tau > t+\frac{T-t}{n}j\right\}} \right] + \mathbb{E}_{t}^{\mathbb{Q}} \left[ e^{-\int_{t}^{T} r_{s} ds} F \mathbf{1}_{\left\{\tau > T\right\}} \right] + \mathbb{E}_{t}^{\mathbb{Q}} \left[ \int_{t}^{T} e^{-\int_{t}^{u} r_{s} ds} R(\tau) d\mathbf{1}_{\left\{\tau < u\right\}} \right],$$
(G.1)

where  $r_s$  corresponds to the interest rate process and  $R(\tau)$  to the recovery at default. The first expectation represents the coupon payments which are only paid if no default occurred by the time of the required payment. The second expectation is the discounted value final face value payment. This payment will only be completed in case no default occurred within the lifetime of the contract. Finally, the last expectation corresponds to the recovery payment in case of a default. We can simplify this expression somewhat by first applying the tower property for expectations: let  $\mathcal{F}_T$  be as defined in appendix D.

$$V_{t} = \mathbb{E}_{t}^{\mathbb{Q}} \left[ \mathbb{E}^{\mathbb{Q}} \left[ \sum_{j=1}^{n} Ce^{-\int_{t}^{t+\frac{T-t}{n}j} r_{s}ds} \mathbf{1}_{\{\tau > t+\frac{T-t}{n}j\}} \middle| \mathcal{F}_{T} \right] \right] + \mathbb{E}_{t}^{\mathbb{Q}} \left[ \mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_{t}^{T} r_{s}ds} F \mathbf{1}_{\{\tau > T\}} \middle| \mathcal{F}_{T} \right] \right] + \mathbb{E}_{t}^{\mathbb{Q}} \left[ \mathbb{E}^{\mathbb{Q}} \left[ \int_{t}^{T} e^{-\int_{t}^{u} r_{s}ds} R(\tau) d\mathbf{1}_{\{\tau < u\}} \middle| \mathcal{F}_{T} \right] \right].$$
(G.2)

Under this filtration the interest rate process,  $r_s$ , is known, thus:

$$V_{t} = \mathbb{E}_{t}^{\mathbb{Q}} \left[ \sum_{j=1}^{n} C e^{-\int_{t}^{t+\frac{T-t}{n}j} r_{s} ds} \mathbb{E}^{\mathbb{Q}} \left[ \mathbf{1}_{\{\tau > t+\frac{T-t}{n}j\}} \middle| \mathcal{F}_{T} \right] \right] + \mathbb{E}_{t}^{\mathbb{Q}} \left[ e^{-\int_{t}^{T} r_{s} ds} F \mathbb{E}^{\mathbb{Q}} \left[ \mathbf{1}_{\{\tau > T\}} \middle| \mathcal{F}_{T} \right] \right] + \mathbb{E}_{t}^{\mathbb{Q}} \left[ \int_{t}^{T} e^{-\int_{t}^{u} r_{s} ds} R(\tau) \mathbb{E}^{\mathbb{Q}} \left[ d\mathbf{1}_{\{\tau < u\}} \middle| \mathcal{F}_{T} \right] \right].$$
(G.3)

Now, under the filtration at time T, also the realised intensity process is known. We therefore recall

$$\begin{cases} \mathbb{E}_T^{\mathbb{Q}}\left[\mathbf{1}_{\{\tau>s\}}\right] &= \mathbb{Q}\left(\tau>s\right) = e^{-\int_t^s \lambda_u du} \\ \mathbb{E}_T^{\mathbb{Q}}\left[d\mathbf{1}_{\{\tau$$

where t is supposed to be the initial time of interest. Using this for (G.3) we find the general defaultable bond pricing formula

$$V_{t} = \mathbb{E}_{t}^{\mathbb{Q}} \left[ \sum_{j=1}^{n} C e^{-\int_{t}^{t+\frac{T-t}{n}j} (r_{s}+\lambda_{s})ds} \right] + \mathbb{E}_{t}^{\mathbb{Q}} \left[ e^{-\int_{t}^{T} (r_{s}+\lambda_{s})ds} F \right]$$

$$+ \mathbb{E}_{t}^{\mathbb{Q}} \left[ \int_{t}^{T} \lambda_{T} e^{-\int_{t}^{u} (r_{s}+\lambda_{s})ds} R(u) du \right].$$
(G.5)

For similar deductions in this thesis we assumed no wrong-way risk, where instead of the tower property we use conditional expectations. The above formulas therefore allow wrong-way risk to exist in the model.

# Appendix H

# Data

#### H.1 Firm data

The data used in this thesis is summarized in table H.1.

ING Group (sub.) 24-12-2014 EUR 10.950 28.34 (0.820, -0.715, 1.508) 2.10	Market         PDE           72.22         70.96           85.15         87.21           100.14         102.75           118.60         116.49           130.03         127.86	C Amazon 09-2001 - 7.756 90.72 (0.891, -0.959, 0.693) 10.26	MarketPDE899.06898.101130.931140.061258.201239.401265.461274.851285.321286.15	d simultaneously and comberg with the
<b>G Group</b> 4-12-2014 USD 10.950 28.34 -0.773, 1.508) 0.82	$\begin{array}{c} PDE \\ 15.50 \\ 24.95 \\ 35.94 \\ 46.73 \\ 56.25 \end{array}$	utch Shell PL 4-12-2014 EUR 18.010 17.38 -5.932, 1.108) 3.44	PDE 21.88 34.21 43.65 50.95 56.53	PM are calibrated ta is found in Blo
<b>IN</b> 2 <sup>,</sup> (1.815,	Market 15.33 25.81 34.86 47.63 55.50	(.) Royal D 22 22, (1.916,	Market 22.80 31.04 40.93 53.41 62.23	oth ING and J. interest rate da
<b>co do Brasil S.A.</b> 20-12-2014 USD 24.350 43.68 70, -2.707, 1.015) 11.11	$\begin{array}{c} PDE \\ 156.46 \\ 209.43 \\ 248.18 \\ 277.30 \\ 298.96 \end{array}$	n Chase & Co. (sub 20-12-2014 USD 61.930 18.69 00, -19.940, 0.812 8.85	PDE 38.80 56.96 67.59 73.49 76.76	tor the selected firms. B
Ban (0.8	Market 159.10 201.01 240.75 278.68 318.48	JPMorga 3.1	Market 34.50 56.80 65.40 69.77 95.94	del fitted data
rance - KLM 4-12-2014 EUR 7.615 46.16 46.16 , -0.027, 1.664) 6.17	$\begin{array}{c} PDE \\ 225.34 \\ 267.36 \\ 302.08 \\ 327.81 \\ 345.60 \end{array}$	<b>an Chase &amp; Co.</b> D-12-2014 USD 61.930 18.69 -9.101, 0.812 4.65	PDE 21.34 34.76 43.39 48.88 52.32	it data and PDE mo
Air F 2 (-0.077	Market 229.88 262.05 293.38 335.22 356.60	<b>JPMorg</b> . 2 2.292,	Market 22.13 32.39 39.26 48.54 61.50	E H.1: Inpu
Data Currency $S_0$ $\sigma$ $(\%)$ $(a_0, a_1, b)$ RMSE	Spreads $(T)$ 1 2 3 3 4 5	Data Currency $S_0$ $\sigma$ $(\%)$ $(a_0, a_1, b)$ RMSE	Spreads $(T)$ 1 2 3 4 4 5	TABL.

#### H.2 Risk-free interest rates

One of the model input parameters is the risk free interest rate  $r_t$ . We have modeled the interest rate as an annual interest rate with continuous compounding, i.e. if  $X_0$  is the initial investment and we let it gain interest over the period  $(0, \Delta t)$  years then we end up with

$$X_0 e^{\int_0^{\Delta t} r_t dt}.$$

We'd like to find such future interest rates and provide these to our model as input. This could be done by for example using a model for the interest rate, such as the simple Vasicek or the more complicated CIR model (for more interesting models and their calibration see [BM01][BA03]). However this thesis' focus is not on such model nor its calibration and we therefore provide another approximation. Instead we shall search for forward risk neutral interest rates, similarly was done by Das & Hanouna [DH09].

Forward rates can be extracted from the fixed leg of an interest rate swap and can be understood as expected interest rates, in the Q-measure, at future states over the time of the swap. Thus, for example, the 6-month EURIBOR forward swap rate over 2 years gives us an indication of the expected interest rate corresponding to a 6 month period over 2 years from now. So, ideally we'd need a continuous stream of interest rate swaps referencing a risk free interest rate with infinitesimal maturities. This would then provide us with  $r_t$  at every t. This requirement can be split up into two sub questions:

- 1. Which curve provides us with risk free interest rates? Before the credit crunch in 2007/2008 it was common practice assume the EURIBOR or LIBOR rates as being risk free. They both correspond to average interest rates at which large banks lend each other money over some predefined period, e.g. 1 month or 6 months. However during the credit crunch it became clear that even large banks might default and that therefore the corresponding interest rate have some credit spread on them. To minimize this, it became market practice to look at overnight interest rates, the rates at which money is lend overnight. This has the advantage that it minimizes default risk as it corresponds to a very short time. For the EURO zone such average interest rate is quoted by EONIA (Euro Overnight Index Average). This is thus a better indication of the risk free interest rate and we shall therefore look at swaps on EONIA curves.
- 2. How can we approximate a continuous stream of interest rate swaps with infinitesimal maturities? With the aid of Bloomberg, the best we can do is to use 1 month EONIA swaps with intervals of 1 month. Bloomberg doesn't provide any shorter tenors nor shorter intervals. Furthermore, the interest rates are quotes are as annual rates without compounding. Therefore we have to transform the provided quotes to continuously compounded counterparts.

The quotes are stated as annual quotes under the day count convention actual/360. This convention states that if we want to calculate a fraction of years then we should use

$$\frac{'\text{Actual amount of days in the period'}}{360}.$$

Thus the total interest that we get over 1 month would approximately be

$$1 + \frac{30}{360}r^*$$

where  $r^*$  the interest rate of the 1-month swap. Thus to find the corresponding continuously compounded interest rate r we determine

$$e^{\frac{30}{360}r} = 1 + \frac{30}{360}r^*$$
  
$$\leftrightarrow r = \frac{360}{30}\ln\left(1 + \frac{30}{360}r^*\right).$$

When  $r^*$  is very small both interest rates will be virtually equal. This is the case for the current market where interest rates are very low. However, for completeness and future scenarios in which the interest rate may increase this transformation might be of influence. Appendix I

Matlab Code
# I.1 D&H model of Das&Hanouna

# I.1.1 Main code for Calibration and pricing of CDSs

This code functions as master code. It contains a section for inputting market data and a section for test data. For the market data an extra section is included in order to perform the calibration. Finally a section is presented where results of the recovery and probability of default are plotted.

```
%--- Main: binomial callibration of default and recovery (Das&Hanouna) ---%
% Setup
format short G;
format compact;
global S D f blb; % Defining the global variables
addpath('Functions')
S0=7.756;
sigma=0.972;
data=[1 1 899.06; 2 1 1130.93; 3 1 1258.20; 4 1 1265.46; 5 1 1285.32];
xdata=data(:,1:end-1);
ydata=data(:,end);
h=1;
maxT=max(data(:,1));
N=maxT/h;
f=[0.0226 0.0273 0.0329 0.0283 0.0436]';
constantR=false;
%% Discount and Stock tree
[S, D, blb] = TrinomialTree(maxT, h, S0, sigma, f);
%% Optimization
[xmin,fmin,flag,outpt,allmins] = tri_optimize(ydata,xdata,h,'lsqcurvefit');
%% Creating and plotting the term structures
[C, P, R, Lambda] = calcSpreads(xmin,xdata,h,true);
Default = sum(P.*Lambda,2);
expectedRecovery = (sum(P,2)).(-1).*(sum(R.*P,2));
subplot(1,2,1)
plotsTerm=plotyy(1:N, Default(1:end), 1:N, expectedRecovery(1:end));
subplot(1,2,2)
```

```
plotsQuotes(1)=plot(xdata(:,1), ydata, 'ko');
hold on
plotsQuotes(2)=plot(xdata(:,1), C, 'go');
```

# I.1.2 Calculation of multiple spreads

Here is a master file for calculating spreads given the required input data. The recovery and probability of default at each node are calculated in this code.

```
$ $ $ $ _____
% CalcSpreads.m
% Calculates multiple spreads for given CDSs
% INPUT
% local
% - (3x1)-vector 'params': parameters [a0 a1 b]
8
  - (ix2)-vector 'xdata': Each row corresponds to a spread
  calculation and the columns take [T N]
8
% global
8
  - (maxT x maxT)-matrix 'S': tril-binomial tree of stock prices
% - (maxT x 1)-vector 'D': forward discount values
% OUTPUT
% - (ix1)-vector 'C': with the calculated spreads
function [C, P, R, Lambda, A, B] = calcSpreads(params,xdata,h,...
                          compute_totalProbabilities)
mult_seniorities=false;
if(length(params)==5)
   mult_seniorities=true;
end
% Retreived data
global S D;
a01=params(1);
all=params(2);
if(mult_seniorities)
   a02=params(3);
   al2=params(4);
   b=params(5);
else
   b=params(3);
end
% Creating matrices Lambda, Q and Recovery
Lambda_withtriu=1-exp(-(1./S.^b).*h);
Lambda=tril(Lambda_withtriu);
```

```
% Form by D&H
R1_withtriu=0.5*erfc((-a01-a11*Lambda)/sqrt(2));
% Alternative form based on default intensity
%10=1./S(1,1).^b;
%R1_withtriu=0.5*erfc((-a01-a11*(1./S.^b-l0)./l0)/sqrt(2));
R1=tril(R1_withtriu);
if(length(S)>1)
    u=S(2,2)/S(1,1); d=1/u;
    Q_withtriu = (repmat((1./D), 1, length(S))./(1-Lambda)-d)./(u-d);
    Q=tril(Q_withtriu);
end
\% in order to proceed computing with Q we modify the inf values to 1e13
% (simply very large)
Q(Q==inf)=1e14;
if(mult_seniorities)
    R2\_withtriu=0.5*erfc((-a02-a12*Lambda)/sqrt(2));
    R2=tril(R2_withtriu);
end
if(compute_totalProbabilities)
   P = totalProbabilities(Q,Lambda);
end
% Determine each CDS spread by means of calcSpread
s=size(xdata);
C=zeros(s(1),1);
for idx=1:s(1)
    T=xdata(idx,1);
    N=xdata(idx,2);
    seniority=1;
    if(mult_seniorities)
    seniority=xdata(idx,3);
    end
    if(seniority==1)
        [C(idx), A(idx), B(idx)] =calcSpread(Lambda, Q, R1, T, N, h);
    elseif(seniority==2)
        [C(idx), A(idx), B(idx)] =calcSpread(Lambda, Q, R2, T, N, h);
    end
end
if(mult_seniorities)
    R(:,:,1)=R1;
    R(:,:,2)=R2;
else
    R=R1;
end
```

#### I.1.3 Calculation of spread given input variables and parameters

This is the final layer of calibration, where a single spread is determined based upon all input (market/test) parameters and the previously found default probabilities and recovery values.

```
888-----
% CalcSpread.m
% Calculates multiple spreads for given CDSs
<u>%</u>
% INPUT
% local
% - (maxT x maxT)-matrix 'Lambda': probilities of default
  - (maxT x maxT)-matrix 'Q': up probability given no default
8
% - (maxT x maxT)-matrix 'R': percentage recovery
% - scalar 'T': maturity in years
00
  - scalar 'N': amount of premium payment if no default occurs % - scalar 'h': binomial-tree s
% global
% - (>T x 1)-vector 'D': forward discount values
  - (>T x 1)-vector 'f': forward interest rates
8
% OUTPUT
% - scalar 'C': calculated spread
function [C, A, B] = calcSpread(Lambda, Q, R, T, N, h)
global D;
% Maturity should be a multiple of the stepsize
Steps=round(T/h)+1;
if (mod(Steps, 1) ~= 0); error('Stepsize and maturity do not match'); end
% Backwards recursion for premium and protection leg
% At the end, both legs, have expected future payments of 0.
% The last premium payment is always calculated as discounted payment in
% the time step before (thus also for the last time step)
A=zeros(1, Steps);
B=zeros(1, Steps);
for t=Steps-1:-1:1
   A(1:t) = D(t) * h + D(t) * ...
        (Q(t,1:t).*(1-Lambda(t,1:t)).*A(2:t+1) ...
         + (1-Q(t,1:t)).*(1-Lambda(t,1:t)).*A(1:t));
    B(1:t) = D(t) * Lambda(t, 1:t) . * (1-R(t, 1:t)) + D(t) * ...
        (Q(t,1:t).*(1-Lambda(t,1:t)).*B(2:t+1)...
        + (1-Q(t,1:t)).*(1-Lambda(t,1:t)).*B(1:t));
end
A=A(1,1);
B=B(1,1);
C=B/A*10000;
```

# I.1.4 Reaching probabilities

This code corresponds to determining the probability of reaching nodes in the jump-todefault tree.

```
% totalProbabilities.m
% Determining the probablity of reaching
% nodes in the trinomial.
888-----
                          _____
% INPUT
% local
% - (>T x >T)-matrix 'Lambda': probilities of default
% - (>T x >T)-matrix 'Q': up probability given no default
% OUTPUT
% - (maxT x maxT)-matrix 'P': the probability of reaching the nodes
function P = totalProbabilities(Q,Lambda)
u=Q.*(1-Lambda);
d=(1-Q) . * (1-Lambda);
% Test code to see if it corresponds to our expectations
%syms ull u21 u22 u31 u32 u33 d11 d21 d22 d31 d32 d33;
%u=[u11 0 0 ; u21 u22 0 ; u31 u32 u33];
%d=[d11 0 0 ; d21 d22 0 ; d31 d32 d33];
n=length(Q);
P=zeros(n);
P(1,1)=1;
for i=2:n
   P(i,1:i-1)=P(i-1,1:i-1).*d(i-1,1:i-1);
   P(i,2:i)=P(i,2:i)+P(i-1,1:i-1).*u(i-1,1:i-1);
end
```

# I.2 Main optimization file

```
global S0 f sigma T S;
addpath('functions');
addpath('../../Data');
addpath('../../MC/functions')
recoverytype={'Probit'};
multiseniority=0;
S.min=0;
S.max=500;
S.N=2*S.max;
T.min=0;
T.N=48;
[S0, sigma, data, currency, f, xmin] = getData('airfrance',T.N,multiseniority);
tMax=max(data(:,1));
T.max=tMax;
times=data(:,1);
%% Optimisation
[xmin,fmin,flag,outpt,allmins]=pde_optimize(data,'fmincon patternsearch',...
                                             recoverytype, xmin);
%% Create grid
if(multiseniority)
   paramtemp=zeros(1,3);
  paramtemp(1,:)=[xmin(1:2) xmin(5)];
  paramtemp(2,:)=[xmin(3:4) xmin(5)];
else
  paramtemp=xmin;
end
K=multiseniority+1;
for k=1:K
    params=paramtemp(k,:);
    grid = Grid(S0, params, f, sigma, T, S, recoverytype);
    Spreads.duration=data(5*(k-1)+1:5*k,1);
    Spreads.spread=data(5*(k-1)+1:5*k,2)/1e4;
    Spreads.nr=size(Spreads.duration,1);
    dt=0.1;
    maturities=dt:dt:tMax;
    for i=1:length(maturities)
```

```
pde(2*i-1)=PDE(grid,[],0,...
                        maturities(i), 1, 'premium leg');
        pde(2*i)=PDE(grid,[],0,maturities(i),[],'protection leg');
    end
    pdelist=PDEList(pde);
    [prem, prot] = FiniteDifference(grid,pdelist,'implicit');
    Sinner=grid.Sgrid(2:end-1,1);
    % retrieve CDS spreads
    C=prot(:,:,1)./prem(:,:,1)*10000;
    spreads = interp1(Sinner',C,S0)';
    subplot(1,2,1)
    plot (maturities, spreads)
    hold on
    C=ismember(maturities, data(5*(k-1)+1:5*k,1));
    idx=find(C);
    RMSE = sqrt(sum((spreads(idx)-data(5*(k-1)+1:5*k,2)).^2/5));
end
plot(times(1:5), data(1:5,2))
if(k==2)
   hold on
    plot(times(6:10), data(6:10,2))
end
\% Calling the termstructures by means of the monte carlo approach
subplot (1, 2, 2)
```

[mL,RatD] = mc(S0,sigma,f,T.max,xmin',recoverytype,true,false);

#### I.2.1 Classes

```
classdef PDE
properties
% grid should contain the object Grid
grid
% boundary should contain
% boundary.left
% boundary.bottom
% boundary.right
% boundary.top
% Length of grid.t and the boundaries do not have to be equal.
```

```
% The PDE may be defined on a subgrid.
    % grid.s and boundary.left and boundary.right should be of same
    % size.
    boundary
    % specifies the domain where the pde is defined
    tmin
    tmax
    % type sets whether the PDE corresponds to the 'premium leg' or the
    % 'protection leg'.
    type
    spread
end
methods
    function obj = PDE(Grid, Boundary, Tmin, Tmax, Spread, Type)
        % if some of the components are defined as constant we expand
        % them into being defined all over the grid
        % we fix the sizes of the boundary conditions if needed
        ls=Grid.s.N+1;
        lt=Grid.t.N*Grid.t.max+1;
        % if the Boundary delivered is empty we create it with the
        % default values for the premium and protection leg
        if(isempty(Boundary))
           Boundary.bottom=0;
           Boundary.right=0;
           if(strcmp(Type, 'premium leg'))
               duration=round(Grid.t.N*Tmax);
               R=Grid.r(1:duration);
               R=repmat(R,length(R),1);
               R=sum(triu(cumprod(exp(-R*Grid.t.d),2)),2);
               disc=Grid.t.d.*Spread*R;
               Boundary.top = [disc' 0];
           elseif(strcmp(Type, 'protection leg'))
               Boundary.top=0;
           end
        end
        if(isfield(Boundary, 'left'))
            Boundary.left = fixSize(Boundary.left,1,ls);
        end
        if(isfield(Boundary, 'bottom'))
            Boundary.bottom = fixSize(Boundary.bottom, lt, 1);
        end
        if(isfield(Boundary, 'right'))
            Boundary.right = fixSize(Boundary.right,1,ls);
```

```
end
if(isfield(Boundary,'top'))
Boundary.top = fixSize(Boundary.top,lt,1);
end
obj.boundary = Boundary;
obj.grid = Grid;
obj.tmin = Tmin;
obj.tmax = Tmax;
obj.type = Type;
obj.spread = Spread;
end
```

end

```
classdef PDEList
    properties
        pdes
        tmax
    end
    methods
        function obj = PDEList(pdes)
            obj.pdes = pdes;
            obj.tmax = max([pdes.tmax]);
        end
        function var = getboundary(obj,idx,boundtype, x)
            if(strcmp(boundtype, 'bottom'))
                bound = [obj.pdes(idx).boundary];
                bot = vertcat(bound.bottom);
                var = bot(:,x)';
            elseif(strcmp(boundtype, 'top'))
                bound = [obj.pdes(idx).boundary];
                top = vertcat(bound.top);
                var = top(:, x)';
            elseif(strcmp(boundtype, 'left'))
                bound = [obj.pdes(idx).boundary];
                lef = horzcat(bound.left);
                var = lef(x, :);
            elseif(strcmp(boundtype,'right'))
                bound = [obj.pdes(idx).boundary];
                rig = horzcat(bound.right);
                var = rig(x,:);
            end
        end
```

```
function spreads = getspreads(obj, idx)
    spreads=[obj.pdes(idx).spread];
end
function idx = firstindex(obj,str)
    idx = find(~cellfun(@isempty.strfind({obj.pdes.type},str)),1,'first');
end
function a = amount(obj,str)
    a = sum(cell2mat(strfind({obj.pdes.type},str)));
end
end
end
```

```
% Constructing the required functions for grid stretching
classdef gridstretcher
    properties
       k
       xi
       а
       b
       n
       phi % This contains phi normal and its first two derivatives
    end
    methods
        function obj = gridstretcher(K, Xi, A, B, N)
            obj.xi=Xi;
            obj.k=K;
            obj.a=A;
            obj.b=B;
            obj.n=N;
            cl=asinh(Xi*(A-K));
            c2=asinh(Xi*(B-K));
            y=(0:1/N:1)';
            eval=c2*y+c1*(1-y);
            obj.phi.d0=max(1/Xi*sinh(eval)+K,0);
            obj.phi.d1=(c2-c1)/Xi*cosh(eval);
            obj.phi.d2=(c2-c1)^2/Xi*sinh(eval);
        end
    end
```

end

# I.2.2 Functions

```
function [xmin,fmin,flag,outpt,allmins] = pde_optimize(data, optimtype,...
                                recoverytype, init)
% Constraints
if(strcmp(recoverytype, 'Probit'))
lb=[-3.1 ; -5e3 ; 0];
ub=[3.1 ; 0 ; 2.5];
elseif(strcmp(recoverytype, 'Logit'))
lb=[-7 ; -5e3 ; 0];
ub=[7; 0; 2.5];
elseif(strcmp(recoverytype, 'Arctan'))
lb=[-319 ; -5e3 ; 0];
ub=[319; 0; 2.5];
end
multiseniority=length(init)>3;
if(multiseniority)
    lb(5)=lb(3); lb(3)=lb(1); lb(4)=lb(2);
    ub(5)=ub(3); ub(3)=ub(1); ub(4)=ub(2);
end
x0=init;
% Defaults
flag=0;
outpt=0;
allmins=0;
w=[1 1 1 1 1]';
w=repmat(w,multiseniority+1,1);
w=w./sum(w);
optimfun = @(x) sum( w.*(calc_pdes(x,data, recoverytype, multiseniority)).^2 );
if (~isempty(strfind(optimtype,'lsqcurvefit')))
    optimfun = @(x,xdata)calc_pdes(x,xdata, recoverytype);
    [xmin, fmin] = lsqcurvefit(optimfun, x0, data, ydata, lb, ub);
end
if(~isempty(strfind(optimtype,'fmincon')))
    % Local
    optimoptions(@fmincon,'Display','iter','TolFun',1e-10,'TolX', 1e-10);
    [xmin, fmin] = fmincon(optimfun, x0, [], [], [], [], lb, ub);
    x0=xmin;
end
if(~isempty(strfind(optimtype, 'patternsearch')))
```

```
opts = psoptimset('PlotFcns', {@psplotbestf,@psplotfuncount},...
        'Display','iter','InitialMeshSize',10,'TolX',1e-5, 'MaxIter', 1e4);
        [xmin,fmin] = patternsearch(optimfun,x0,[],[],[],[],lb,ub,[],opts);
end
if(~isempty(strfind(optimtype,'simannealing')))
        [xmin,fmin,~,~] = simulannealbnd(optimfun,x0,lb,ub);
end
```

```
function [diff, spreads] = calc_pdes(params, data, ...
                                         recoverytype, multiseniority)
global S0 f sigma T S;
if(multiseniority)
   paramtemp=zeros(1,3);
  paramtemp(1,:) = [params(1:2) params(5)];
   paramtemp(2,:)=[params(3:4) params(5)];
else
   paramtemp=params;
end
K=multiseniority+1;
for k=1:K
    params=paramtemp(k,:);
    grid = Grid(S0, params, f, sigma, T, S, recoverytype);
    Spreads.duration=data(5*(k-1)+1:5*k,1);
    Spreads.spread=data(5*(k-1)+1:5*k,2)/1e4;
    Spreads.nr=size(Spreads.duration,1);
    for i=1:Spreads.nr
        pde(2*i-1)=PDE(grid,[],0,...
                        Spreads.duration(i), Spreads.spread(i), 'premium leg');
        pde(2*i)=PDE(grid,[],0,Spreads.duration(i),[],'protection leg');
    end
    pdelist=PDEList(pde);
    [prem, prot] = FiniteDifference(grid,pdelist,'implicit');
    Sinner=grid.Sgrid(2:end-1,1);
    % retrieve CDS spreads
    C=prot(:,:,1)./prem(:,:,1)*1e4;
    spreads = interp1(Sinner',C,S0)';
```

```
premS0=interp1(Sinner',prem(:,:,1),S0)';
protS0=interp1(Sinner',prot(:,:,1),S0)';
if(sum(sum([premS0 protS0]<0))>0)
    diff(:,k)=Inf*ones(size(premS0));
else
    diff(:,k)=protS0./premS0-premS0./protS0;
end
```

diff=diff(:);

```
function [prem, prot] = FiniteDifference(grid,pdelist,scheme)
% pdelist contains the list array of pdes (it is created as it contains
% other practical methods for the list).
pdes=pdelist.pdes;
Spreads.times=grid.t.N*[pdes.tmax]+1;
% Setting initial (or actually final) conditions.
% We assume that each of all premium legs have the same inital condition,
% the same is assumed for the protection leg. Thus we pick the first
% boundary condition for each.
iprem=firstindex(pdelist, 'premium leg');
nrprem=amount(pdelist, 'premium leg');
iprot=firstindex(pdelist, 'protection leg');
nrprot=amount(pdelist, 'protection leg');
% the initial conditions are set to the whole grid for each pde
% dimension: (S,t,amount)
if(iprem>0)
    prem=repmat(pdes(iprem).boundary.right(2:end-1),...
        1,nrprem,pdelist.tmax*grid.t.N+1);
end
if(iprot>0)
    prot=repmat(pdes(iprot).boundary.right(2:end-1),...
        1,nrprot,pdelist.tmax*grid.t.N+1);
end
if(strcmp(scheme,'cranknicolson'))
    i=pdelist.tmax*grid.t.N+1;
    if(strcmp(grid.type, 'gridstretching'))
            % The diagonal matrix which has to be inverted.
            % Will be used for evaluation of premium and protection leg.
            l=-(grid.pdevar.del(2:grid.s.N,i)+0.5.*grid.pdevar.alp(2:grid.s.N));
            d=1+grid.pdevar.alp(2:grid.s.N)+grid.pdevar.bet(2:grid.s.N,i);
            u=grid.pdevar.del(2:grid.s.N,i)-0.5.*grid.pdevar.alp(2:grid.s.N);
        elseif (strcmp(grid.type, 'regular'))
```

```
j=(1:1:grid.s.N-1)'; % only consider the inner-grid
            j2=j.^2;
            l=-0.5*(grid.pdevar.alp*j2-grid.pdevar.bet(2:grid.s.N,i).*j);
            d=1+grid.pdevar.alp*j2+grid.pdevar.bet(2:grid.s.N,i);
            u=-0.5*(grid.pdevar.alp*j2+grid.pdevar.bet(2:grid.s.N,i).*j);
    end
    1=1./2;
    d=d./2+1/2;
    u=u./2;
end
% Performing finite difference by looping over time
for i=pdelist.tmax*grid.t.N:-1:1
    premactive=(i<=Spreads.times & strcmp({pdes.type}, 'premium leg'));</pre>
    protactive=(i<=Spreads.times & strcmp({pdes.type}, 'protection leg'));</pre>
    Premactive=sum(premactive);
    Premactiveindex=find(premactive);
    Protactive=sum(protactive);
    Protactiveindex=find(protactive);
    if(strcmp(scheme, 'cranknicolson'))
       le=-l;
       de = -d + 2;
       ue=-u;
    end
    if(strcmp(grid.type, 'gridstretching'))
        % The diagonal matrix which has to be inverted.
        % Will be used for evaluation of premium and protection leq.
        l=-(grid.pdevar.del(2:grid.s.N,i)+0.5.*grid.pdevar.alp(2:grid.s.N));
        d=1+grid.pdevar.alp(2:grid.s.N)+grid.pdevar.bet(2:grid.s.N,i);
        u=grid.pdevar.del(2:grid.s.N,i)-0.5.*grid.pdevar.alp(2:grid.s.N);
    elseif (strcmp(grid.type, 'regular'))
        j=(1:1:grid.s.N-1)'; % only consider the inner-grid
        j2=j.^2;
        l=-0.5*(grid.pdevar.alp*j2-grid.pdevar.bet(2:grid.s.N,i).*j);
        d=1+grid.pdevar.alp*j2+grid.pdevar.bet(2:grid.s.N,i);
        u=-0.5*(grid.pdevar.alp*j2+grid.pdevar.bet(2:grid.s.N,i).*j);
    end
    if(strcmp(scheme,'cranknicolson'))
        1=1./2;
        d=d./2+1/2;
        u=u./2;
        Ae=spdiags(le(2:end),-1,grid.s.N-1,grid.s.N-1)...
            +spdiags(de,0,grid.s.N-1,grid.s.N-1)...
            +spdiags([0 ; ue(1:end-1)],1,grid.s.N-1,grid.s.N-1);
    end
```

```
A=spdiags(l(2:end),-1,grid.s.N-1,grid.s.N-1)...
    +spdiags(d,0,grid.s.N-1,grid.s.N-1)...
    +spdiags([0 ; u(1:end-1)],1,grid.s.N-1,grid.s.N-1);
% setting the boundary contidions for premium and protection leg
if(Premactive>0)
    spread=getspreads(pdelist,Premactiveindex);
   bA=-grid.t.d*repmat(spread,length(2:grid.s.N),1);
   bA(1,:)=bA(1,:)+l(1,:)...
                    .*getboundary(pdelist, Premactiveindex, 'bottom', i);
   bA(grid.s.N-1,:)=bA(grid.s.N-1,:)+u(end)...
                .*getboundary(pdelist,Premactiveindex,'top',i);
    idxPrem=(nrprem-Premactive+1):1:nrprem;
    % Solving the Crank Nicolson scheme
    if(strcmp(scheme, 'cranknicolson'))
        bAe=zeros(length(2:grid.s.N), Premactive);
       bAe(1,:)=le(1,:).*getboundary(pdelist,Premactiveindex,'bottom',i+1);
       bAe(grid.s.N-1,:)=bAe(grid.s.N-1,:)+ue(end)...
                    .*getboundary(pdelist,Premactiveindex,'top',i+1);
        prem(:,idxPrem,i)=A\(Ae*prem(:,idxPrem,i+1)-bA+bAe);
   elseif(strcmp(scheme,'implicit'))
        % Solving the implicit finite difference equation
        prem(:,idxPrem,i)=A\(prem(:,idxPrem,i+1)-bA);
   end
end
if(Protactive>0)
   bB=-grid.t.d*grid.lambda(2:grid.s.N,i).*(1-grid.rho(2:grid.s.N,i));
   bB=repmat(bB,1,Protactive);
   bB(1,:)=bB(1,:)+l(1,:)...
                    .*getboundary(pdelist,Protactiveindex,'bottom',i);
   bB(grid.s.N-1,:)=bB(grid.s.N-1,:)+u(end)...
                .*getboundary(pdelist,Protactiveindex,'top',i);
    idxProt=(nrprot-Protactive+1):1:nrprot;
    % Solving the Crank Nicolson scheme
    if(strcmp(scheme, 'cranknicolson'))
        bBe=zeros(length(2:grid.s.N),Protactive);
       bBe(1,:)=le(1,:).*getboundary(pdelist,Protactiveindex,'bottom',i+1);
        bBe(grid.s.N-1,:)=bBe(grid.s.N-1,:)+ue(end)...
                    .*getboundary(pdelist,Protactiveindex,'top',i+1);
        prot(:,idxProt,i)=A\(Ae*prot(:,idxProt,i+1)-bB+bBe);
    elseif(strcmp(scheme,'implicit'))
        % Solving the implicit finite difference equation
        prot(:,idxProt,i)=A\(prot(:,idxProt,i+1)-bB);
    end
end
```

# I.2.3 Dummy functions

```
function a = amount(celarray,str)
```

```
a = sum(cell2mat(strfind(celarray,str)));
```

end

```
function idx = index(celarray,str)
```

idx = find(~cellfun(@isempty,strfind(celarray,str)),1,'first');

end

```
% Takes any variable and extends its size to that of the
 % specificied dimensions.
function var = fixSize(var, lX, lY)
    if(length(var)==1)
        var=kron(var, ones(lY,lX));
    end
    if(size(var,1)==1)
       var=repmat(var, lY, 1);
    end
    if(size(var,2)==1)
       var=repmat(var,1,lX);
    end
    if(size(var,1)<=lY)</pre>
       var(size(var,1)+1:1Y,:)=0;
    end
    if(size(var,2)<=lX)</pre>
       var(:, size(var, 2) +1:1X) =0;
    end
end
```

#### I.2.4 Monte Carlo for retrieving the term structure

function [mL,RatD,LGDPD] = mc(S0,sigma,f,T,params,modeltype,doplot,CVA)

```
multiseniority=length(params)>3;
```

```
N=length(f)./T;
dt=1/N;
M=1e5;
S=S0*ones(M,1);
if(multiseniority)
   b=params(5);
else
    b=params(3);
end
lambda=zeros(M,N);
for t=1:(N*T)
    lambda(:,t)=S(:,t).^(-b);
    lambda(lambda==Inf)=0;
    dNt=binornd(1,dt*lambda(:,t));
    dWt=sqrt(dt) *randn(M,1);
    incr=((f(t)+lambda(:,t)).*S(:,t))*dt+sigma*S(:,t).*dWt;
    S(:,t+1) = S(:,t)+incr-(S(:,t)+incr).*dNt;
end
% Mean expected stock without default
t=0:dt:T;
mL=mean(lambda);
K=multiseniority+1;
for k=1:K
    a0=params(2*(k-1)+1);
    a1=params(2*k);
    % Recovery at default
    10=(1/S0)^b;
    x=a0+a1*(lambda-l0)./l0;
    if(strcmp(modeltype, 'Probit'))
        R=normcdf(x);
    elseif(strcmp(modeltype, 'Logit'))
        R=1./(1+exp(-x));
    elseif(strcmp(modeltype, 'Atan'))
        R=0.5*(atan(a0+a1*x)*2/pi+1);
    end
    R(lambda==0)=0; % set to zero where a default already occured
    RatD=sum(R,1)./sum(R~=0,1);
    % CVA
    if(CVA)
        PD=dt.*lambda.*exp(-1*cumsum(lambda,2)*dt);
        discount=exp(-1.*cumsum(repmat(f,M,1),2)*dt);
        LGDPD(:,:,k)=discount.*(1-R).*PD;
    end
```

```
% Plot both recovery and default intensity
if(doplot)
 %subplot(1,sets,i)
 %particularly for amazon the probability density function:
 if(~exist('plots'))
      [plots, line_default, line_recovery] = plotyy(t(1:end-1), mL(:),...
                                                t(1:end-1), RatD(:));
 else
      hold(plots(2));
      f=plot(plots(2),t(1:end-1),RatD(:));
      legend(f,'Recovery subordinated');
 end
 ylabel(plots(1), 'Intensity')
 ylabel(plots(2), 'Fraction of notional')
 legend([line_default ; line_recovery], {'Default' 'Recovery'})
end
```

# Bibliography

- [ABRS03] Edward I. Altman, Brooks Brady, Andrea Resti, and Andrea Sironi. The link between default and recovery rates: Theory, empirical evidence and implications. Journal of Business, 78:2203–2228, 2003.
  - [ABS07] Viral V. Acharya, Sreedhar T. Bharath, and Anand Srinivasan. Does industrywide distress affect defaulted firms? - evidence from creditor recoveries. Journal of Financial Economics, 85(3):787–821, September 2007.
  - [AFV03] E. Ayache, Peter. A. Forsyth, and Ken R. Vetzal. Valuation of convertible bonds with credit risk. The Journal of Derivatives, 11(1):9–29, 2003.
  - [AK96] Edward I. Altman and Vellore M. Kishore. Almost everything you wanted to know about recoveries on defaulted bonds. *Financial Analysts Journal*, 52(6):57–64, November/December 1996.
  - [ARS05] Edward Altman, Andrea Resti, and Andrea Sironi. Default recovery rates in credit risk modeling: A review of the literature and empirical evidence. Journal of Finance Literature, 1:21–45, 2005.
  - [BA03] Damiano Brigo and Aurélien Alfonsi. Credit default swaps calibration and option pricing with the ssrd stochastic intensity and interest-rate model. In Proceedings of the 6-th Columbia=JAFEE Conference, pages 15–16, 2003.
  - [BJR07] Tomasz R. Bielecki, Monique Jeanblanc, and Marek Rutkoski. Valuation and hedging of credit derivatives. Lecture Notes, April 2007. URL: https:// cel.archives-ouvertes.fr/cel-00398075/document.
  - [BM01] Damiano Brigo and Fabio Mercurio. *Interest rate models: theory and practice*. Springer finance. Springer, Berlin, Heidelberg, Paris, 2001.
  - [BM06] Damiano Brigo and Fabio Mercurio. Interest rate models theory and practice: With smile, inflation and credit (springer finance). August 2006. URL: http://www.amazon.com/exec/obidos/redirect?tag= citeulike07-20&path=ASIN/B0042JTAP6.
  - [BMP13] Damiano Brigo, Massimo Morini, and Andrea Pallavicini. Counterparty Credit Risk, Collateral and Funding. Wiley, 2013.

- [BMZ01] Gurdip Bakshi, Dilip Madan, and Frank Zhang. Understanding the role of recovery in default risk models: Empirical comparisons and implied recovery rates. Working Paper, 2001.
- [CES06] Richard Cantor, Kenneth Emery, and Pamela Stumpp. Probability of default ratings and loss given default assessments for non-financial speculative-grade corporate obligors in the united states and canada. Technical Report 98771, Rating Methodology: Moody's, August 2006.
- [CH04] Daniel Covitz and Song Han. An empirical analysis of bond recovery rates: Exploring a structural view of default. Working Paper, December 2004. URL: http://www.federalreserve.gov/pubs/feds/2005/200510/ 200510pap.pdf.
- [Chr05] Jens H. E. Christensen. Joint estimation of default and recovery risk: A simulation study. Working Paper, January 2005.
- [CRR79] John C. Cox, Stephen A. Ross, and Mark Rubinstein. Option pricing: A simplified approach. Journal of Financial Economics, (7):229–263, 1979.
  - [CV05] Richard Cantor and Praveen Varma. Determinants of recovery rates on defaulted bonds and loans for north american corporate issuers: 1983-2003. Journal of Fixed Income, 14(4):29–44, March 2005.
  - [DH09] Sanjiv R. Das and Paul Hanouna. Implied recovery. Journal of Economic Dynamics and Control, 33(11):1837–1857, 2009.
  - [DS99] Darrell Duffie and Kenneth J. Singleton. Modeling term structures of defaultable bonds. *The Review of Financial Studies*, 12(4):687–720, 1999.
  - [DS03] Darrell Duffie and Kenneth J. Singleton. Credit Risk: Pricing, Measurement and Management. Princeton University Press, 2003.
  - [EY13] EY. Credit valuation adjustments at non-financial organizations, 2013. accessed 15-01-2015. URL: http://www.ey.com/ Publication/vwLUAssets/Credit\_valuation\_adjustments\_at\_ non-financial\_organizations/\$FILE/2013\_Credit\_valuation\_ adjustments.pdf.
- [FJOS10] Fang Fang, Henrik Jönsson, Cornelis W. Oosterlee, and Wim Schoutens. Fast valuation and calibration of credit default swaps under lévy dynamics. *Journal* of Computational Finance, 14(2):1–30, February 2010.
- [FNS09] Bjorn Flesaker, Madhu Nayakkankuppam, and Igor Shkurko. The bloomberg cds model. Technical report, Bloomberg, 2009. URL: http://img2.timg. co.il/forums/1\_172333558.pdf [cited 30-01-2015].

- [For14] Peter Forsyth. An introduction to computational finance without agonizing pain [online]. December 2014. URL: http://www.cs.uwaterloo.ca/ ~paforsyt/agon.pdf [cited 02-02-2015].
- [GO86] David Gottlieb and A. Steven Orszag. Numerical analysis of spectral methods : theory and applications. Society for Industrial and Applied Mathematics, 1986.
- [Guh02] Rajiv Guha. Recovery of face value at default: Theory and empirical evidence. Working Paper, 2002.
- [Hof92] Joe D Hoffman. Numerical methods for engineers and scientists. McGraw-Hill, New York, NY, 1992.
- [HPW05] John Hull, Mirela Predescu, and Alan White. Bond prices, default probabilities and risk premiums. Journal of Credit Risk, 1(2):53–60, 2005.
  - [Iwa13] Yukinori Iwashita. Conventions for single-name credit default Technical report, OpenGamma, 2013.URL: swaps. http: //developers.opengamma.com/quantitative-research/ Conventions-Single-Name-Credit-Default-Swaps-OpenGamma. pdf.
  - [JM12] Marcin Jaskowski and Michael McAleer. Estimating implied recovery rates from the term structure of cds spreads. KIER Working Papers 836, Kyoto University, Institute of Economic Research, 2012. URL: http://eprints. ucm.es/17581/1/1228.pdf.
  - [JNS14] Rainer Jankowitsch, Florian Nagler, and Marti G. Subrahmanyam. The determinants of recovery rates in the us corporate bond market. Journal of Financial Economics, 114(1):155–177, February 2014.
  - [Lin06] Vadim Linetsky. Pricing equity derivatives subject to bankruptcy. Journal of Mathematical Finance, 16(2):255–282, 2006.
  - [LR12] Alexander Lipton and Andrew Rennie, editors. The Oxford Handbook of Credit Derivatives. Oxford University Press, 2012. URL: http://EconPapers. repec.org/RePEc:oxp:obooks:9780199669486.
  - [Mar12] Markit. Isda model accrual on default fix, December 2012. URL: http://www.cdsmodel.com/assets/cds-model/docs/ISDAModelfix.pdf.
  - [Mar15] ISDA CDS Marketplace. About cds [online]. 2015. URL: http://www. isdacdsmarketplace.com/about\_cds\_market [cited 02-02-2015].
  - [Mer76] Robert C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144, 1976.

- [MU00] Dilip Madan and Haluk Unal. A two-factor hazard-rate model for pricing risky debt and the term structure of credit spreads. Journal of Financial and Quantitative Analysis, 35:43–65, 2000.
- [Mur99] Y. Muromachi. The growing recognition of credit risk in corporate and financial bond markets. Working Paper, 1999. NLI Research Institute.
- [OLH05] Cornelis W. Oosterlee, Coen C.W. Leentvaar, and Xinzheng Huang. Accurate american option pricing by grid stretching and high order finite differences. Working paper, 2005.
  - [PS08] Jun Pan and Kenneth J. Singleton. Default and recovery implicit in the term structure of sovereign cds spreads. The Journal of Finance, 63(5):2345–2384, October 2008.
- [Reu14] Reuters. Air france cuts full-year ebitda goal in third profit warning [online]. December 2014. URL: http://www.reuters.com/article/2014/12/ 18/air-france-klm-warning-idUSL6N0U24UG20141218 [cited 28-01-2015].
- [Sap08] Nishul Saperia. Credit event auction primer. Technical report, Markit, 2008. URL: http://www.creditfixings.com/information/ affiliations/fixings/auctions/2008/credit\_event\_auction\_ primer.pdf.
- [Sch04] Til Schuermann. What do we know about loss given default, chapter 9, pages 249–274. Risk Books, 2nd edition, 2004.
- [SIC15] SIC. Search [online]. 2015. URL: http://siccode.com/en/pages/ what-is-a-sic-code [cited 02-01-2015].
- [Str89] John C. Strikwerda. Finite Difference Schemes and Partial Differential Equations. Wadsworth Publishing, Belmont, CA, USA, 1989.
- [SUH14] Timo Schläfer and Marliese Uhrig-Homburg. Is recovery risk priced? Journal of Banking & Finance, 40(C):257–270, 2014.
- [UMG03] Haluk Unal, Dilip Madan, and Levent Guntay. Pricing the risk of recovery in default with absolute priority rule violation. Journal of Banking & Finance, 27(6):1001–1025, June 2003.